

Explaining fish species composition in the Amazon basin

AmazonFISH partners

2025-09-24 11:17:33.611773

Contents

1 Charging Biological data	4
1.1 Species composition	4
1.2 Calculating and Decomposing fish dissimilarity	32
1.3 Including PhyloDissimilarity	42
1.3.1 PBsim using Cassemiro et al tree	48
1.3.2 Calculating Dpw	51
1.3.3 Calculating PCD	55
1.4 Functional Beta Diversity	61
1.4.1 Charging and checking morpho data	61
1.4.2 Looking mean morphological position of sub-drainages	74
1.4.3 Quality of the functional	75
2 Charging species richness and attributes of basins	81
3 Preparing environmental data	93
3.1 Calculating <i>Habitat diversity</i>	93
3.2 Soil diversity	95
3.3 Global terrain slope (potential waterfalls)	98
3.4 Current climate	99
3.5 Current Runoff	103
3.6 Net Primary Productivity	104
3.6.1 Combining all variables from current climate in a single <i>data.frame</i>	105
3.7 Past climate	112
3.7.1 CCSM	112
3.7.2 MIROC	114
3.7.3 MPI model	116
3.7.4 Mean variables of LGM models	117
3.7.5 Computing differences between current and past climate	119

3.8	Codding the structure of the drainage basin	123
3.9	Marine Incursions (5Myr and 1Myr)	126
3.10	Hydromorphology type	129
3.11	Habitat variables	133
3.11.1	Setting and transforming variables	133
3.11.2	Computing dissimilarity matrices	137
4	Past-Connectivity among subdrainages as a Palaeo-Connectivity matrix	142
5	Charging <i>Spatial Distance Matrix</i>	146
5.1	Checking the structure of ALL Distance Matrices	148
6	Modelling fish beta diversity	158
6.1	Descriptive statistics	158
6.2	The MRM model	164
6.2.1	Ordinary MRM (multiple regressions using Gaussian Distribution)	165
6.2.2	MRM_glm (multiple regressions using Binomial Distribution)	167
6.2.3	Testing excluding undersampled basins	171
6.2.4	Simple plots	182
6.2.4.1	Taxonomic beta diversity	182
6.2.4.2	Phylogenetic beta diversity (Rabosky genetic + polytomies, Phylo 29k species) .	189
6.2.4.3	Functional Beta Sim 3D na Filled (WITH and WITHOUT OUTLIER)	193
6.3	Spatial autocorrelation	194
7	Plotting maps of biological data	197
8	R + knitr Section Informations	210
Bibliography		212

NOTE_1: *we reduced the number of digits to print when printing numeric values to improve visualization.*

NOTE_2: *this document has been created using Rmarkdown¹.*

¹<http://rmarkdown.rstudio.com/>

1 Charging Biological data

1.1 Species composition

Charging the complete data set and setting θ to all *NA* values.

```
fish_mat=read.csv(paste("StatistiquesSubBasinAmazon_02052016/",
                       "NewData_2019_10/Dataset1_SpeciesList_97bv_311019.csv",
                       sep=""),
# paste("/home/ird/Insync/msdias@unb.br/OneDrive Biz/Amazon",
#       "/AmazonFish_analises/",
#       "StatistiquesSubBasinAmazon_02052016/",
#       "NewData_2017_04/MatrixSpecies_0417.csv",
#       sep="")
sep=";",header=T)

#leaving only valid species names
dim(fish_mat)

## [1] 39462      9

head(fish_mat)

##   Family.Referent.Species Genus.Referent.Species Referent.Species.Name
## 1   Acestrorhynchidae    Acestrorhynchus    Acestrorhynchus.abbreviatus
## 2   Acestrorhynchidae    Acestrorhynchus    Acestrorhynchus.abbreviatus
## 3   Acestrorhynchidae    Acestrorhynchus    Acestrorhynchus.abbreviatus
## 4   Acestrorhynchidae    Acestrorhynchus    Acestrorhynchus.abbreviatus
## 5   Acestrorhynchidae    Acestrorhynchus    Acestrorhynchus.abbreviatus
## 6   Acestrorhynchidae    Acestrorhynchus    Acestrorhynchus.abbreviatus
##   Author.Referent.Species Species.Status Occurrence.Status Major.Tributary.Name
## 1           (Cope, 1878)     native        valid          Amazonas
## 2           (Cope, 1878)     native        valid          Amazonas
## 3           (Cope, 1878)     native        valid      Japur\xe1
## 4           (Cope, 1878)     native        valid      Japur\xe1
## 5           (Cope, 1878)     native        valid        Javari
## 6           (Cope, 1878)     native        valid      Madeira
##   Sub_drainage Endemism_Amazon
## 1      Amazon2            1
## 2      Amazon3            1
## 3      Caqueta            1
## 4      Japura              1
## 5      Javary              1
## 6      Abuna              1

table(fish_mat$Occurrence.Status)

##
##                                marine  questionable      valid
##                                99          155         2535      36673
```

```

fish_mat<-droplevels(fish_mat[fish_mat$Occurrence.Status %in% "valid",])

fish_mat$Referent.Species.Name<-muda_nome(fish_mat$Referent.Species.Name,sep=".[]")
fish_mat[10065,]

##           Family.Referent.Species Genus.Referent.Species
## 10925             Characidae      Hyphessobrycon
##           Referent.Species.Name Author.Referent.Species Species.Status
## 10925 Hyphessobrycon_jackrobertsi          Zarske 2014      native
##           Occurrence.Status Major.Tributary.Name Sub_drainage Endemism_Amazon
## 10925           valid            Mara\xf1on       Pastaza        1

table(grep(fish_mat$Referent.Species.Name,pattern = "[_]"==1:length(fish_mat$Referent.Species.Name))

##  

##  TRUE  

## 36673

# ##excluding additional marine fish families 2022_03_25
# #marine "Atherinopsidae", "Achiridae", "Belonidae", "Clupeidae", "Engraulidae",
# # "Hemiramphidae", "Potamotrygonidae", "Pristigasteridae", "Sciaenidae",
# # "Tetraodontidae"
# #not marine according to Lovejoy et al.2006 "Atherinopsidae"
# fish_mat<-droplevels(fish_mat[!fish_mat$Family.Referent.Species %in%
#                               c("Atherinopsidae", "Achiridae", "Belonidae",
#                                 "Clupeidae", "Engraulidae", "Hemiramphidae",
#                                 "Potamotrygonidae", "Pristigasteridae", "Sciaenidae",
#                                 "Tetraodontidae"),])
# table(fish_mat$Family.Referent.Species)

#####checking if marine species are out
# head(fish_mat)
# marineIncSp<-c("Potamotrygon", "Paratrygon", "Heliotrygon", "Plesiotrygon", "Anchoviella", "Anchovia",
# # "Lyengraulis", "Pterengraulis", "Amazonsprattus", "Jurengraulis", "Pellona", "Ilisha", "Pristigaster",
# # "Rhinosardina", "Potamoraphis", "Belonion", "Pseudotylusurus", "Hyporhamphus",
# # "Plagioscion", "Pachyrops", "Pachyurus", "Petilipinnis", "Colomesus")
# table(fish_mat$Genus.Referent.Species[fish_mat$Genus.Referent.Species %in% marineIncSp])
# fish_mat[fish_mat$Genus.Referent.Species%in% marineIncSp,]
# fish_mat<-fish_mat[!fish_mat$Genus.Referent.Species %in% marineIncSp,]

# ##eliminating Orestias species because they are from Titikaka lake
fish_mat<-droplevels(fish_mat[!fish_mat$Genus.Referent.Species %in% "Orestias",])

#unindo os codigos antigos e novos
fish2=read.csv(paste("StatistiquesSubBasinAmazon_02052016/", ,

```

```

        "NewData_2019_10/stat97bv_311019.csv",
        sep=""),
sep="; ",header=T)[,c("Sub_drainage","Basin2stat")]
head(fish2)

##   Sub_drainage Basin2stat
## 1      Abuna     Abuna
## 2    Amazon1    Amazon1
## 3  Amazon10  Solimoes7
## 4  Amazon11  Solimoes8
## 5  Amazon12  Solimoes9
## 6    Amazon2    Amazon2

fish_mat<-merge(x=fish_mat,y=fish2,by.x="Sub_drainage",by.y="Sub_drainage",all=T)
head(fish_mat)

##   Sub_drainage Family.Referent.Species Genus.Referent.Species
## 1      Abuna           Doradidae       Ossancora
## 2      Abuna           Curimatidae      Potamorhina
## 3      Abuna  Rhamphichthyidae  Rhamphichthys
## 4      Abuna           Characidae      Parecbasis
## 5      Abuna           Anostomidae     Leporinus
## 6      Abuna           Cetopsidae      Cetopsis
##   Referent.Species.Name Author.Referent.Species Species.Status
## 1  Ossancora_punctata          (Kner, 1855)    native
## 2  Potamorhina_latior        (Spix & Agassiz, 1829) native
## 3 Rhamphichthys_marmoratus        Castelnau, 1855 native
## 4  Parecbasis_cyclolepis        Eigenmann, 1914 native
## 5  Leporinus_trifasciatus        Steindachner, 1876 native
## 6  Cetopsis_coecutiens        (Lichtenstein, 1819) native
##   Occurrence.Status Major.Tributary.Name Endemism_Amazon Basin2stat
## 1      valid           Madeira        NA     Abuna
## 2      valid           Madeira        1     Abuna
## 3      valid           Madeira        NA     Abuna
## 4      valid           Madeira        1     Abuna
## 5      valid           Madeira        NA     Abuna
## 6      valid           Madeira        NA     Abuna

fish_mat$basin<-fish_mat$Basin2stat

#fish species lists
head(fish_mat)

##   Sub_drainage Family.Referent.Species Genus.Referent.Species
## 1      Abuna           Doradidae       Ossancora
## 2      Abuna           Curimatidae      Potamorhina
## 3      Abuna  Rhamphichthyidae  Rhamphichthys
## 4      Abuna           Characidae      Parecbasis
## 5      Abuna           Anostomidae     Leporinus
## 6      Abuna           Cetopsidae      Cetopsis

```

```

##   Referent.Species.Name Author.Referent.Species Species.Status
## 1   Ossancora_punctata          (Kner, 1855)      native
## 2   Potamorhina_latior        (Spix & Agassiz, 1829)    native
## 3 Rhamphichthys_marmoratus       Castelnau, 1855      native
## 4   Parecbasis_cyclolepis        Eigenmann, 1914      native
## 5   Leporinus_trifasciatus       Steindachner, 1876      native
## 6   Cetopsis_coecutiens        (Lichtenstein, 1819)      native
## Occurrence.Status Major.Tributary.Name Endemism_Amazon Basin2stat basin
## 1       valid           Madeira          NA     Abuna Abuna
## 2       valid           Madeira          1     Abuna Abuna
## 3       valid           Madeira          NA     Abuna Abuna
## 4       valid           Madeira          1     Abuna Abuna
## 5       valid           Madeira          NA     Abuna Abuna
## 6       valid           Madeira          NA     Abuna Abuna

```

```
sort(unique(fish_mat$Referent.Species.Name))
```

```

## [1] "Abramites_hypselonotus"
## [3] "Acanthobunocephalus_nicoi"
## [5] "Acanthodoras_depressus"
## [7] "Acanthopoma_annectens"
## [9] "Acaronia_nassa"
## [11] "Acestridium_dichromum"
## [13] "Acestridium_gymnogaster"
## [15] "Acestridium_scutatum"
## [17] "Acstrocephalus_acutus"
## [19] "Acstrocephalus_boehlkei"
## [21] "Acstrocephalus_pallidus"
## [23] "Acstrocephalus_stigmatus"
## [25] "Acstrorhynchus_altus"
## [27] "Acstrorhynchus_falcirostris"
## [29] "Acstrorhynchus_heterolepis"
## [31] "Acstrorhynchus_maculipinna"
## [33] "Acstrorhynchus_minimus"
## [35] "Acstrorhynchus_pantaneiro"
## [37] "Acnodon_senai"
## [39] "Acrobrycon_starnesi"
## [41] "Adontosternarchus_clarkae"
## [43] "Adontosternarchus_duartei"
## [45] "Aequidens_diadema"
## [47] "Aequidens_gerciliae"
## [49] "Aequidens_metae"
## [51] "Aequidens_pallidus"
## [53] "Aequidens_plagiozonatus"
## [55] "Aequidens_superomaculatum"
## [57] "Aequidens_tubicen"
## [59] "Agamyxis_albomaculatus"
## [61] "Ageneiosus_akamai"
## [63] "Ageneiosus_inermis"
## [65] "Ageneiosus_lineatus"
## [67] "Ageneiosus_piperatus"
## [1] "Acanthicus_hystrix"
## [3] "Acanthodoras_cataphractus"
## [5] "Acanthodoras_spinosissimus"
## [7] "Acarichthys_heckelii"
## [9] "Acaronia_vultuosa"
## [11] "Acestridium_discus"
## [13] "Acestridium_martini"
## [15] "Acestridium_triplax"
## [17] "Acstrocephalus_anomalus"
## [19] "Acstrocephalus_nigrifasciatus"
## [21] "Acstrocephalus_sardina"
## [23] "Acstrorhynchus_abbreviatus"
## [25] "Acstrorhynchus_falcatus"
## [27] "Acstrorhynchus_grandoculis"
## [29] "Acstrorhynchus_isalineae"
## [31] "Acstrorhynchus_microlepis"
## [33] "Acstrorhynchus_nasutus"
## [35] "Acnodon_normani"
## [37] "Acrobrycon_ipanquianus"
## [39] "Adontosternarchus_balaenops"
## [41] "Adontosternarchus_devenanzi"
## [43] "Adontosternarchus_nebulosus"
## [45] "Aequidens_epae"
## [47] "Aequidens_mauesanus"
## [49] "Aequidens_michaeli"
## [51] "Aequidens_patricki"
## [53] "Aequidens_rondoni"
## [55] "Aequidens_tetramerus"
## [57] "Aequidens_viridis"
## [59] "Agamyxis_pectinifrons"
## [61] "Ageneiosus_atronasus"
## [63] "Ageneiosus_intrusus"
## [65] "Ageneiosus_militaris"
## [67] "Ageneiosus_polystictus"

```

```

## [69] "Ageneiosus_ucayalensis"
## [71] "Ageneiosus_vittatus"
## [73] "Agoniates_halecinus"
## [75] "Aguarunichthys_torosus"
## [77] "Amazonspinther_dalmata"
## [79] "Amblydoras_affinis"
## [81] "Amblydoras_monitor"
## [83] "Ammocryptocharax_elegans"
## [85] "Ammoglanis_amapaensis"
## [87] "Anablepsoides_atratus"
## [89] "Anablepsoides_cajariensis"
## [91] "Anablepsoides_christinae"
## [93] "Anablepsoides_elongatus"
## [95] "Anablepsoides_gamae"
## [97] "Anablepsoides_hoetmeri"
## [99] "Anablepsoides_intermittens"
## [101] "Anablepsoides_jari"
## [103] "Anablepsoides_limoncochae"
## [105] "Anablepsoides_micropus"
## [107] "Anablepsoides_ophiomimus"
## [109] "Anablepsoides_ottonii"
## [111] "Anablepsoides_peruanus"
## [113] "Anablepsoides_rubrolineatus"
## [115] "Anablepsoides_taeniatus"
## [117] "Anablepsoides_urubuiensis"
## [119] "Anablepsoides_xinguensis"
## [121] "Anadoras_regani"
## [123] "Anchovia_surinamensis"
## [125] "Anchoviella_carrikeri"
## [127] "Anchoviella_hernanni"
## [129] "Anchoviella_juruasanga"
## [131] "Ancistrus_bolivianus"
## [133] "Ancistrus_dolichopterus"
## [135] "Ancistrus_heterorhynchus"
## [137] "Ancistrus_jelskii"
## [139] "Ancistrus_latifrons"
## [141] "Ancistrus_malacops"
## [143] "Ancistrus_megalostomus"
## [145] "Ancistrus_nudiceps"
## [147] "Ancistrus_occloii"
## [149] "Ancistrus_ranunculus"
## [151] "Ancistrus_shuar"
## [153] "Ancistrus_temminckii"
## [155] "Ancistrus_triradiatus"
## [157] "Ancistrus_verecundus"
## [159] "Andinoacara_pulcher"
## [161] "Anodus_elongatus"
## [163] "Anostomoides_laticeps"
## [165] "Anostomus_anostomus"
## [167] "Anostomus_ternetzi"
## [169] "Aphanotorulus_rubrocauda"
## [171] "Aphyocharacidium_bolivianum"

```

"Ageneiosus_uranocephalus"
 "Agoniates_anchoviae"
 "Aguarunichthys_inpai"
 "Amaralia_hypsiura"
 "Amazonsprattus_scintilla"
 "Amblydoras_gonzalezi"
 "Amblydoras_nauticus"
 "Ammocryptocharax_minutus"
 "Anablepsoides_amana"
 "Anablepsoides_beniensis"
 "Anablepsoides_chapare"
 "Anablepsoides_derhami"
 "Anablepsoides_erberi"
 "Anablepsoides_henschelae"
 "Anablepsoides_holmiae"
 "Anablepsoides_iridescent"
 "Anablepsoides_jucundus"
 "Anablepsoides_lineasopilatae"
 "Anablepsoides_monticola"
 "Anablepsoides_ornatus"
 "Anablepsoides_parlettei"
 "Anablepsoides_roraima"
 "Anablepsoides_speciosus"
 "Anablepsoides_urophthalmus"
 "Anablepsoides_waimacui"
 "Anadoras_grypus"
 "Anadoras_weddellii"
 "Anchoviella_alleni"
 "Anchoviella_guianensis"
 "Anchoviella_jamesi"
 "Anchoviella_lepidostole"
 "Ancistrus_bufonius"
 "Ancistrus_dubius"
 "Ancistrus_hoplogenys"
 "Ancistrus_krenakarore"
 "Ancistrus_lineolatus"
 "Ancistrus_maximus"
 "Ancistrus_montanus"
 "Ancistrus_occidentalis"
 "Ancistrus_parecis"
 "Ancistrus_sericeus"
 "Ancistrus_tamboensis"
 "Ancistrus_tombador"
 "Ancistrus_variolus"
 "Andeanancistrus_eschwartzae"
 "Anduzeedoras_oxyrhynchus"
 "Anodus_orinocensis"
 "Anostomoides_passionis"
 "Anostomus_longus"
 "Aphanotorulus_horridus"
 "Aphanotorulus_unicolor"
 "Aphyocharax_agassizii"

```

## [173] "Aphyocharax_alburnus"
## [175] "Aphyocharax_colifax"
## [177] "Aphyocharax_nattereri"
## [179] "Aphyodite_apiaka"
## [181] "Aphyodite_tupebas"
## [183] "Aphyolebias_claudiae"
## [185] "Aphyolebias_obliquus"
## [187] "Aphyolebias_rubrocaudatus"
## [189] "Apionichthys_dumerili"
## [191] "Apionichthys_menezesi"
## [193] "Apionichthys_rosai"
## [195] "Aピストグラマ_acrensis"
## [197] "Aピストグラマ_aguarico"
## [199] "Aピストグラマ_allpahuayo"
## [201] "Aピストグラマ_angayuara"
## [203] "Aピストグラマ_atahualpa"
## [205] "Aピストグラマ_barlowi"
## [207] "Aピストグラマ_brevis"
## [209] "Aピストグラマ_cinilabra"
## [211] "Aピストグラマ_diplotaenia"
## [213] "Aピストグラマ_elizabethae"
## [215] "Aピストグラマ_erythrura"
## [217] "Aピストグラマ_feconat"
## [219] "Aピストグラマ_flavipedunculata"
## [221] "Aピストグラマ_gephyra"
## [223] "Aピストグラマ_hippolytae"
## [225] "Aピストグラマ_inconspicua"
## [227] "Aピストグラマ_juruensis"
## [229] "Aピストグラマ_linkei"
## [231] "Aピストグラマ_martini"
## [233] "Aピストグラマ_meinkeni"
## [235] "Aピストグラマ_moae"
## [237] "Aピストグラマ_norberti"
## [239] "Aピストグラマ_panduro"
## [241] "Aピストグラマ_paucisquamis"
## [243] "Aピストグラマ_payaminonis"
## [245] "Aピストグラマ_pertensis"
## [247] "Aピストグラマ_pulchra"
## [249] "Aピストグラマ_resticulosa"
## [251] "Aピストグラマ_rubrolineata"
## [253] "Aピストグラマ_salpinction"
## [255] "Aピストグラマ_sororcula"
## [257] "Aピストグラマ_steindachneri"
## [259] "Aピストグラマ_trifasciata"
## [261] "Aピストグラマ_urteagai"
## [263] "Aピストグラマ_wolli"
## [265] "Aピストオリカリア_condei"
## [267] "Aポマトオセロス_alleni"
## [269] "Aピテロノツス_albifrons"
## [271] "Aピテロノツス_lindalvae"
## [273] "Aライチスス_thoro"
## [275] "Aラパイマ_leptosoma"
## [173] "Aphyocharax_anisitsi"
## [175] "Aphyocharax_erythrurus"
## [177] "Aphyocharax_pusillus"
## [179] "Aphyodite_grammica"
## [181] "Aphyolebias_boticarioi"
## [183] "Aphyolebias_manuensis"
## [185] "Aphyolebias_peruensis"
## [187] "Aphyolebias_schleseri"
## [189] "Apionichthys_finis"
## [191] "Apionichthys_nattereri"
## [193] "Apionichthys_seripierriae"
## [195] "Aピストグラマ_agassizii"
## [197] "Aピストグラマ_alacrina"
## [199] "Aピストグラマ_amoena"
## [201] "Aピストグラマ_arua"
## [203] "Aピストグラマ_baenschi"
## [205] "Aピストグラマ_bitaeiata"
## [207] "Aピストグラマ_cacatuoides"
## [209] "Aピストグラマ_cruzi"
## [211] "Aピストグラマ_eleutheria"
## [213] "Aピストグラマ_eremnopyge"
## [215] "Aピストグラマ_eunotus"
## [217] "Aピストグラマ_flabellicauda"
## [219] "Aピストグラマ_geisleri"
## [221] "Aピストグラマ_gibbiceps"
## [223] "Aピストグラマ_huascar"
## [225] "Aピストグラマ_iniridae"
## [227] "Aピストグラマ_kullanderi"
## [229] "Aピストグラマ_luelingi"
## [231] "Aピストグラマ_megastoma"
## [233] "Aピストグラマ_mendezi"
## [235] "Aピストグラマ_nijsseni"
## [237] "Aピストグラマ_ortegai"
## [239] "Aピストグラマ_pantalone"
## [241] "Aピストグラマ_paulmuelleri"
## [243] "Aピストグラマ_personata"
## [245] "Aピストグラマ_playayacu"
## [247] "Aピストグラマ_regani"
## [249] "Aピストグラマ_rositae"
## [251] "Aピストグラマ_rupununi"
## [253] "Aピストグラマ_similis"
## [255] "Aピストグラマ_staecki"
## [257] "Aピストグラマ_taeniata"
## [259] "Aピストグラマ_uaupesi"
## [261] "Aピストグラマ_wapisana"
## [263] "Aピストグラモイデス_pucallpaensis"
## [265] "Aピストオリカリア_ommation"
## [267] "Aポスチリソマ_myriodon"
## [269] "Aピテロノツス_bonapartii"
## [271] "Aピテロノツス_macrolepis"
## [273] "Aラパイマ_gigas"
## [275] "Aラホラエムス_ferreirai"

```



```

## [381] "Biotoecus_opercularis"
## [383] "Bivibranchia_notata"
## [385] "Boehlke_a_fredcochui"
## [387] "Boulengerella_cuvieri"
## [389] "Boulengerella_lucius"
## [391] "Boulengerella_xyrekes"
## [393] "Brachychalcinus_nummus"
## [395] "Brachychalcinus_reisi"
## [397] "Brachyglanis_frenatus"
## [399] "Brachyglanis_micropthalmus"
## [401] "Brachyhypopomus_alberti"
## [403] "Brachyhypopomus_batesi"
## [405] "Brachyhypopomus_belindae"
## [407] "Brachyhypopomus_bennetti"
## [409] "Brachyhypopomus_bullocki"
## [411] "Brachyhypopomus_hamiltoni"
## [413] "Brachyhypopomus_pinnicaudatus"
## [415] "Brachyhypopomus_regani"
## [417] "Brachyhypopomus_verdii"
## [419] "Brachyplatystoma_capapretum"
## [421] "Brachyplatystoma_juruense"
## [423] "Brachyplatystoma_rousseauxii"
## [425] "Brachyplatystoma_vaillantii"
## [427] "Brachyrhamdia_marthae"
## [429] "Brachyrhamdia_rambarrani"
## [431] "Brittanichthys_axelrodi"
## [433] "Brochis_multiradiatus"
## [435] "Brycon_bicolor"
## [437] "Brycon_coxeyi"
## [439] "Brycon_henni"
## [441] "Brycon_melanopterus"
## [443] "Brycon_stolzmanni"
## [445] "Bryconacidnus_hemigrammus"
## [447] "Bryconadenos_tanaothoros"
## [449] "Bryconamericus_alfredae"
## [451] "Bryconamericus_andresoi"
## [453] "Bryconamericus_brevirostris"
## [455] "Bryconamericus_cinarucoense"
## [457] "Bryconamericus_diaphanus"
## [459] "Bryconamericus_grosvenori"
## [461] "Bryconamericus_macrophthalmus"
## [463] "Bryconamericus osgoodi"
## [465] "Bryconamericus_pectinatus"
## [467] "Bryconamericus_pinnavittatus"
## [469] "Bryconamericus_zamorensis"
## [471] "Bryconexodon_juruenae"
## [473] "Bryconops_affinis"
## [475] "Bryconops_caudomaculatus"
## [477] "Bryconops_collettei"
## [479] "Bryconops_durbini"
## [481] "Bryconops_gracilis"
## [483] "Bryconops_inpai"
## [485] "Bivibranchia_fowleri"
## [487] "Bivibranchia_velox"
## [489] "Boehlke_a_weitzmani"
## [491] "Boulengerella_lateristriga"
## [493] "Boulengerella_maculata"
## [495] "Brachychalcinus_copei"
## [497] "Brachychalcinus_orbicularis"
## [499] "Brachychalcinus_retrospina"
## [501] "Brachyglanis_melas"
## [503] "Brachyglanis_nocturnus"
## [505] "Brachyhypopomus_arrayae"
## [507] "Brachyhypopomus_beebei"
## [509] "Brachyhypopomus_benjamini"
## [511] "Brachyhypopomus_brevirostris"
## [513] "Brachyhypopomus_cunia"
## [515] "Brachyhypopomus_hendersoni"
## [517] "Brachyhypopomus_provenzanoi"
## [519] "Brachyhypopomus_sullivani"
## [521] "Brachyhypopomus_walteri"
## [523] "Brachyplatystoma_filamentosum"
## [525] "Brachyplatystoma_platynemum"
## [527] "Brachyplatystoma_tigrinum"
## [529] "Brachyrhamdia_heteropleura"
## [531] "Brachyrhamdia_meesi"
## [533] "Brachyrhamdia_thayeria"
## [535] "Brittanichthys_myersi"
## [537] "Brycon_amazonicus"
## [539] "Brycon_cephalus"
## [541] "Brycon_falcatus"
## [543] "Brycon_hilarii"
## [545] "Brycon_pesu"
## [547] "Bryconacidnus_ellisi"
## [549] "Bryconacidnus_paipayensis"
## [551] "Bryconadenos_weitzmani"
## [553] "Bryconamericus_alpha"
## [555] "Bryconamericus_boliviensis"
## [557] "Bryconamericus_carlosi"
## [559] "Bryconamericus_deuterodonoides"
## [561] "Bryconamericus_galvisi"
## [563] "Bryconamericus_hypopterus"
## [565] "Bryconamericus_orinocoense"
## [567] "Bryconamericus_pachacuti"
## [569] "Bryconamericus_phoenicopterus"
## [571] "Bryconamericus_subtilisform"
## [573] "Bryconella_pallidifrons"
## [575] "Bryconexodon_trombetasi"
## [577] "Bryconops_alburnoides"
## [579] "Bryconops_chernoffi"
## [581] "Bryconops_disruptus"
## [583] "Bryconops_giacopinii"
## [585] "Bryconops_humeralis"
## [587] "Bryconops_melanurus"

```

```

## [485] "Bryconops_munduruku"
## [487] "Bryconops_sapezal"
## [489] "Bujurquina_apoparuana"
## [491] "Bujurquina_eurhinus"
## [493] "Bujurquina_huallagae"
## [495] "Bujurquina_mariae"
## [497] "Bujurquina_moriorum"
## [499] "Bujurquina_pardus"
## [501] "Bujurquina_robusta"
## [503] "Bujurquina_tambopatae"
## [505] "Bujurquina_zamorensis"
## [507] "Bunocephalus_coracoideus"
## [509] "Bunocephalus_verrucosus"
## [511] "Caenotropus_maculosus"
## [513] "Caenotropus_schizodon"
## [515] "Callichthys_serralabium"
## [517] "Caquetaia_kraussii"
## [519] "Caquetaia_spectabilis"
## [521] "Carnegiella_myersi"
## [523] "Carnegiella_strigata"
## [525] "Centrodoras_brachiatus"
## [527] "Centromochlus_altae"
## [529] "Centromochlus_heckelii"
## [531] "Centromochlus_meridionalis"
## [533] "Centromochlus_perugiae"
## [535] "Centromochlus_schultzi"
## [537] "Ceratobranchia_binghami"
## [539] "Ceratobranchia_elatior"
## [541] "Ceratobranchia_obtusirostris"
## [543] "Cetopsisium_pemon"
## [545] "Cetopsis_candiru"
## [547] "Cetopsis_gobiooides"
## [549] "Cetopsis_oliveirai"
## [551] "Cetopsis_pearsoni"
## [553] "Cetopsis_sandrae"
## [555] "Cetopsorhamdia_filamentosa"
## [557] "Cetopsorhamdia_orinoco"
## [559] "Chaetobranchopsis_australis"
## [561] "Chaetobranchus_flavescens"
## [563] "Chaetostoma_alternifasciatum"
## [565] "Chaetostoma_branickii"
## [567] "Chaetostoma_changae"
## [569] "Chaetostoma_dermorhynchum"
## [571] "Chaetostoma_leucomelas"
## [573] "Chaetostoma_loborhynchos"
## [575] "Chaetostoma_marmorescens"
## [577] "Chaetostoma_mollinasum"
## [579] "Chaetostoma_spondylus"
## [581] "Chaetostoma_taczanowskii"
## [583] "Chaetostoma_vagum"
## [585] "Chalceus_erythrurus"
## [587] "Chalceus_maclepidotus"
## [485] "Bryconops_piracolina"
## [487] "Bryconops_transitoria"
## [489] "Bujurquina_cordemadi"
## [491] "Bujurquina_hophrys"
## [493] "Bujurquina_labiosa"
## [495] "Bujurquina_megalospilus"
## [497] "Bujurquina_ortegai"
## [499] "Bujurquina_peregrinabunda"
## [501] "Bujurquina_sympilus"
## [503] "Bujurquina_vittata"
## [505] "Bunocephalus_aleuropsis"
## [507] "Bunocephalus_knerii"
## [509] "Caenotropus_labyrinthicus"
## [511] "Caenotropus_mestomorgmatus"
## [513] "Callichthys_callichthys"
## [515] "Calophysus_macropterus"
## [517] "Caquetaia_myersi"
## [519] "Carnegiella_marthae"
## [521] "Carnegiella_schereri"
## [523] "Catoprion_mento"
## [525] "Centrodoras_hasemani"
## [527] "Centromochlus_existimatus"
## [529] "Centromochlus_macracanthus"
## [531] "Centromochlus_orca"
## [533] "Centromochlus_reticulatus"
## [535] "Cephalosilurus_apurensis"
## [537] "Ceratobranchia_delotaenia"
## [539] "Ceratobranchia_joanae"
## [541] "Cetopsisium_ferreirai"
## [543] "Cetopsisium_soniae"
## [545] "Cetopsis_coecutiens"
## [547] "Cetopsis_montana"
## [549] "Cetopsis_parma"
## [551] "Cetopsis_plumbea"
## [553] "Cetopsis_starnesi"
## [555] "Cetopsorhamdia_insidiosa"
## [557] "Cetopsorhamdia_phantasia"
## [559] "Chaetobranchopsis_orbicularis"
## [561] "Chaetobranchus_semidasciatus"
## [563] "Chaetostoma_anale"
## [565] "Chaetostoma_breve"
## [567] "Chaetostoma_daidalmatos"
## [569] "Chaetostoma_jegui"
## [571] "Chaetostoma_lineopunctatum"
## [573] "Chaetostoma_marcapatae"
## [575] "Chaetostoma_microps"
## [577] "Chaetostoma_sericeum"
## [579] "Chaetostoma_stroumpoulos"
## [581] "Chaetostoma_trimaculineum"
## [583] "Chalceus_epakros"
## [585] "Chalceus_guaporensis"
## [587] "Chalceus_spilogyrus"

```

```

## [589] "Characidium_boavistae"
## [591] "Characidium_bolivianum"
## [593] "Characidium_declivirostre"
## [595] "Characidium_fasciatum"
## [597] "Characidium_heinianum"
## [599] "Characidium_nana"
## [601] "Characidium_pellucidum"
## [603] "Characidium_purpuratum"
## [605] "Characidium_schindleri"
## [607] "Characidium_summus"
## [609] "Characidium_zebra"
## [611] "Charax_condei"
## [613] "Charax_gibbosus"
## [615] "Charax_leticiae"
## [617] "Charax_michaeli"
## [619] "Charax_rupununi"
## [621] "Chasmocranus_longior"
## [623] "Chasmocranus_quadrimaculatus"
## [625] "Cheirocerus_goeldii"
## [627] "Cheirodon_ortegai"
## [629] "Chilodus_gracilis"
## [631] "Chrysobrycon_eliasi"
## [633] "Chrysobrycon_mojicai"
## [635] "Chrysobrycon_yoliae"
## [637] "Cichla_jariina"
## [639] "Cichla_miriana"
## [641] "Cichla_nigromaculata"
## [643] "Cichla_orinocensis"
## [645] "Cichla_pleiozona"
## [647] "Cichla_thyrorus"
## [649] "Cichlasoma_amazonarum"
## [651] "Cichlasoma_bimaculatum"
## [653] "Clupeocharax_anchoveoides"
## [655] "Colossoma_macropomum"
## [657] "Compsaraia_iara"
## [659] "Copeina_guttata"
## [661] "Copella_arnoldi"
## [663] "Copella_meinkeni"
## [665] "Copella_nigrofasciata"
## [667] "Cordylancistrus_platycephalus"
## [669] "Cordylancistrus_santarosensis"
## [671] "Corydoras_acutus"
## [673] "Corydoras_aeneus"
## [675] "Corydoras_albolineatus"
## [677] "Corydoras_ambiacus"
## [679] "Corydoras_apiaka"
## [681] "Corydoras_armatus"
## [683] "Corydoras_benattii"
## [685] "Corydoras_bilineatus"
## [687] "Corydoras_bondi"
## [689] "Corydoras_burgessi"
## [691] "Corydoras_cervinus"
"Characidium_boehlkei"
"Characidium_crandellii"
"Characidium_etheostoma"
"Characidium_hasemani"
"Characidium_longum"
"Characidium_papachibe"
"Characidium_pteroides"
"Characidium_roesseli"
"Characidium_steindachneri"
"Characidium_xavante"
"Charax_caudimaculatus"
"Charax_delimai"
"Charax_hemigrammus"
"Charax_macleayi"
"Charax_pauciradiatus"
"Charax_tectifer"
"Chasmocranus_peruanus"
"Cheirocerus_eqques"
"Cheirodon_luelingi"
"Chilodus_fritillus"
"Chilodus_punctatus"
"Chrysobrycon_hesperus"
"Chrysobrycon_myersi"
"Cichla_intermedia"
"Cichla_melaniae"
"Cichla_monoculus"
"Cichla_ocellaris"
"Cichla_pinima"
"Cichla_temensis"
"Cichla_vazzoleri"
"Cichlasoma_araguaiense"
"Cichlasoma_boliviense"
"Colomesus_asellus"
"Compsaraia_compsus"
"Compsaraia_samueli"
"Copeina osgoodi"
"Copella_compta"
"Copella_nattereri"
"Copella_vilmae"
"Cordylancistrus_platyrhynchus"
"Corydoras_acrensis"
"Corydoras_adolfoi"
"Corydoras_agassizii"
"Corydoras_amandajanea"
"Corydoras_amphibelus"
"Corydoras_arcuatus"
"Corydoras_atropersonatus"
"Corydoras_bifasciatus"
"Corydoras_blochi"
"Corydoras_brittoi"
"Corydoras_caudimaculatus"
"Corydoras_copei"

```

```

## [693] "Corydoras_coriatae"
## [695] "Corydoras_cruziensis"
## [697] "Corydoras_davidsandsi"
## [699] "Corydoras_desana"
## [701] "Corydoras_elegans"
## [703] "Corydoras_evelynae"
## [705] "Corydoras_geryi"
## [707] "Corydoras_gossei"
## [709] "Corydoras_granti"
## [711] "Corydoras_haraldschultzi"
## [713] "Corydoras_hephaestus"
## [715] "Corydoras_incolicana"
## [717] "Corydoras_julii"
## [719] "Corydoras_knaacki"
## [721] "Corydoras_latus"
## [723] "Corydoras_leucomelas"
## [725] "Corydoras_loxozonus"
## [727] "Corydoras_melanistius"
## [729] "Corydoras_melini"
## [731] "Corydoras_napoensis"
## [733] "Corydoras_negro"
## [735] "Corydoras_noelkempffi"
## [737] "Corydoras_orphnopterus"
## [739] "Corydoras_ourastigma"
## [741] "Corydoras_paragua"
## [743] "Corydoras_pastazensis"
## [745] "Corydoras_pavanelliae"
## [747] "Corydoras_pulcher"
## [749] "Corydoras_rabauti"
## [751] "Corydoras_reynoldsi"
## [753] "Corydoras_robustus"
## [755] "Corydoras_schwartzii"
## [757] "Corydoras_serratus"
## [759] "Corydoras_similis"
## [761] "Corydoras_spectabilis"
## [763] "Corydoras_stenocephalus"
## [765] "Corydoras_sychri"
## [767] "Corydoras_tukano"
## [769] "Corydoras_virginiae"
## [771] "Corydoras_weitzmani"
## [773] "Corydoras_zawadzkii"
## [775] "Creagrutus_amoenus"
## [777] "Creagrutus_barrigai"
## [779] "Creagrutus_cacique"
## [781] "Creagrutus_changae"
## [783] "Creagrutus_cracentis"
## [785] "Creagrutus_flavescens"
## [787] "Creagrutus_gracilis"
## [789] "Creagrutus_ignotus"
## [791] "Creagrutus_manu"
## [793] "Creagrutus_muelleri"
## [795] "Creagrutus_occidaneus"
"Corydoras_crimmeni"
"Corydoras_crypticus"
"Corydoras_delphax"
"Corydoras_duplicareus"
"Corydoras_eques"
"Corydoras_fowleri"
"Corydoras_gomezi"
"Corydoras_gracilis"
"Corydoras_guapore"
"Corydoras_hastatus"
"Corydoras_imitator"
"Corydoras_isbrueckeri"
"Corydoras_kanei"
"Corydoras_lamberti"
"Corydoras_leopardus"
"Corydoras_loretoensis"
"Corydoras_mamore"
"Corydoras_melanotaenia"
"Corydoras_multiradiatus"
"Corydoras_narcissus"
"Corydoras_nijsseni"
"Corydoras_ornatus"
"Corydoras_ortegai"
"Corydoras_panda"
"Corydoras_parallelus"
"Corydoras_paucerna"
"Corydoras_pinheiroi"
"Corydoras_pygmaeus"
"Corydoras_reticulatus"
"Corydoras_robineae"
"Corydoras_sarareensis"
"Corydoras_semiaquilus"
"Corydoras_seussi"
"Corydoras_sodalis"
"Corydoras_splendens"
"Corydoras_sterbai"
"Corydoras_trilineatus"
"Corydoras_urucu"
"Corydoras_vittatus"
"Corydoras_xinguensis"
"Corydoras_zygatus"
"Creagrutus_anary"
"Creagrutus_beni"
"Creagrutus_cerritulus"
"Creagrutus_cochui"
"Creagrutus_ephippiatus"
"Creagrutus_gephyrus"
"Creagrutus_holmi"
"Creagrutus_kunturus"
"Creagrutus_maxillaris"
"Creagrutus_nigrotaeniatus"
"Creagrutus_ortegai"

```

```

## [797] "Creagrutus_ouranonastes"
## [799] "Creagrutus_peruanus"
## [801] "Creagrutus_phasma"
## [803] "Creagrutus_runa"
## [805] "Creagrutus_ungulus"
## [807] "Creagrutus_yanatili"
## [809] "Creagrutus_zephyrus"
## [811] "Crenicara_punctulatum"
## [813] "Crenicichla_adspersa"
## [815] "Crenicichla_anamiri"
## [817] "Crenicichla_chicha"
## [819] "Crenicichla_cyanonotus"
## [821] "Crenicichla_heckeli"
## [823] "Crenicichla_hummelincki"
## [825] "Crenicichla_isbrueckeri"
## [827] "Crenicichla_lenticulata"
## [829] "Crenicichla_lucius"
## [831] "Crenicichla_macrophthalma"
## [833] "Crenicichla_monicae"
## [835] "Crenicichla_pellegrini"
## [837] "Crenicichla_phaiospilus"
## [839] "Crenicichla_proteus"
## [841] "Crenicichla_regani"
## [843] "Crenicichla_rosemariae"
## [845] "Crenicichla_saxatilis"
## [847] "Crenicichla_semicincta"
## [849] "Crenicichla_tigrina"
## [851] "Crenicichla_virgatula"
## [853] "Crenuchus_spilurus"
## [855] "Crossoloricaria_rhami"
## [857] "Ctenobrycon_spilurus"
## [859] "Curculionichthys_jumaorum"
## [861] "Curculionichthys_sabaji"
## [863] "Curimata_aspera"
## [865] "Curimata_cyprinoides"
## [867] "Curimata_inornata"
## [869] "Curimata_ocellata"
## [871] "Curimata_vittata"
## [873] "Curimatella_dorsalis"
## [875] "Curimatella_meyeri"
## [877] "Curimatopsis_evelynae"
## [879] "Curimatopsis_jaci"
## [881] "Curimatopsis_maculosa"
## [883] "Curimatopsis_microlepis"
## [885] "Cyanogaster_noctivaga"
## [887] "Cynodon_septenarius"
## [889] "Cynopotamus_bipunctatus"
## [891] "Cynopotamus_juruенae"
## [893] "Cyphocharax_abramoides"
## [895] "Cyphocharax_festivus"
## [897] "Cyphocharax_gouldingi"
## [899] "Cyphocharax_leucostictus"
"Creagrutus_pearsoni"
"Creagrutus_petilus"
"Creagrutus_pila"
"Creagrutus_tuyuka"
"Creagrutus_vexillapinnus"
"Creagrutus_yudja"
"Crenicara_latruncularium"
"Crenicichla_acutirostris"
"Crenicichla_alta"
"Crenicichla_anthurus"
"Crenicichla_cincta"
"Crenicichla_dandara"
"Crenicichla_hemera"
"Crenicichla_inpa"
"Crenicichla_johanna"
"Crenicichla_lepidota"
"Crenicichla_lugubris"
"Crenicichla_marmorata"
"Crenicichla_notophthalmus"
"Crenicichla_percna"
"Crenicichla_ploegi"
"Crenicichla_pydanielae"
"Crenicichla_reticulata"
"Crenicichla_santosi"
"Crenicichla_sedentaria"
"Crenicichla_strigata"
"Crenicichla_urosema"
"Crenicichla_wallacii"
"Crossoloricaria_bahuaja"
"Ctenobrycon_hauxwellianus"
"Curculionichthys_itaim"
"Curculionichthys_karipuna"
"Curculionichthys_scaius"
"Curimata_cisandina"
"Curimata_incompta"
"Curimata_knerii"
"Curimata_roseni"
"Curimatella_alburna"
"Curimatella_immaculata"
"Curimatopsis_crypticus"
"Curimatopsis_guaporensis"
"Curimatopsis_maclepis"
"Curimatopsis_melanura"
"Curimatopsis_pallida"
"Cynodon_gibbus"
"Cynopotamus_amazonum"
"Cynopotamus_gouldingi"
"Cynopotamus_xinguano"
"Cyphocharax_derhami"
"Cyphocharax_gangamon"
"Cyphocharax_laticlavius"
"Cyphocharax_mestomyllon"

```

```

## [901] "Cyphocharax_multilineatus"
## [903] "Cyphocharax_nigripinnis"
## [905] "Cyphocharax_pantostictos"
## [907] "Cyphocharax_sanctigabrielis"
## [909] "Cyphocharax_spilurus"
## [911] "Cyphocharax_vexillapinnus"
## [913] "Dekeyseria_brachyura"
## [915] "Dekeyseria_pulchra"
## [917] "Denticetopsis_praecox"
## [919] "Denticetopsis_sauli"
## [921] "Dianema_longibarbis"
## [923] "Dicrossus_filamentosus"
## [925] "Dicrossus_maculatus"
## [927] "Distocyclus_conirostris"
## [929] "Doras_higuchii"
## [931] "Duopalatinus_peruanus"
## [933] "Eigenmannia_loretana"
## [935] "Eigenmannia_matintapereira"
## [937] "Eigenmannia_nigra"
## [939] "Eigenmannia_virescens"
## [941] "Elachoharax_geryi"
## [943] "Elachoharax_mitopterus"
## [945] "Electrophorus_electricus"
## [947] "Entomocorus_benjamini"
## [949] "Epapterus_dispilurus"
## [951] "Ernstichthys_megistus"
## [953] "Erythrocharax_altipinnis"
## [955] "Exallodontus_aguanai"
## [957] "Exastilithoxus_hoedemani"
## [959] "Farlowella.altocorpus"
## [961] "Farlowella_gianetii"
## [963] "Farlowella_knerii"
## [965] "Farlowella_odontotumulus"
## [967] "Farlowella_platonymous"
## [969] "Farlowella_smithi"
## [971] "Fluviphylax_pygmaeus"
## [973] "Fluviphylax_zonatus"
## [975] "Furcodontichthys_novaesi"
## [977] "Galeocharax_gulo"
## [979] "Gasteropelecus_sternicla"
## [981] "Gelanoglanis_pan"
## [983] "Geophagus_abalios"
## [985] "Geophagus_argyrostictus"
## [987] "Geophagus_camopiensis"
## [989] "Geophagus_megasema"
## [991] "Geophagus_proximus"
## [993] "Geophagus_taeniopareius"
## [995] "Gephyrocharax_chaparae"
## [997] "Geryichthys_sterbai"
## [999] "Gladioglanis_conquistador"
## [1001] "Glanapteryx_anguilla"
## [1003] "Glanidium_ribeiroi"
## [901] "Cyphocharax_muyrakytan"
## [903] "Cyphocharax_notatus"
## [905] "Cyphocharax_plumbeus"
## [907] "Cyphocharax_spiluropsis"
## [909] "Cyphocharax_stilbolepis"
## [911] "Dekeyseria_amazonica"
## [913] "Dekeyseria_picta"
## [915] "Dekeyseria_scaphirhyncha"
## [917] "Denticetopsis_royeroi"
## [919] "Denticetopsis_seducta"
## [921] "Dianema_urostriatum"
## [923] "Dicrossus_foorni"
## [925] "Dicrossus_warzeli"
## [927] "Doras_carinatus"
## [929] "Doras_phlyzakion"
## [931] "Eigenmannia_limbata"
## [933] "Eigenmannia_macrops"
## [935] "Eigenmannia_muirapinima"
## [937] "Eigenmannia_trilineata"
## [939] "Eigenmannia_waiwai"
## [941] "Elachoharax_junki"
## [943] "Elachoharax_pulcher"
## [945] "Engraulisoma_taeniatum"
## [947] "Entomocorus_melaphareus"
## [949] "Ernstichthys_intonsus"
## [951] "Erythrinus_erythrinus"
## [953] "Etsaputu_relictum"
## [955] "Exastilithoxus_fimbriatus"
## [957] "Exodon_paradoxus"
## [959] "Farlowella.amazonum"
## [961] "Farlowella_gracilis"
## [963] "Farlowella_nattereri"
## [965] "Farlowella_oxyrryncha"
## [967] "Farlowella_schreitmulleri"
## [969] "Fluviphylax_obscurus"
## [971] "Fluviphylax_simplex"
## [973] "Fonchiiloricaria_nanodon"
## [975] "Galeocharax_goeldii"
## [977] "Gasteropelecus_levis"
## [979] "Gelanoglanis_nanonocticolus"
## [981] "Gelanoglanis_travieso"
## [983] "Geophagus_altifrons"
## [985] "Geophagus_brasiliensis"
## [987] "Geophagus_dicrozoster"
## [989] "Geophagus_mirabilis"
## [991] "Geophagus_surinamensis"
## [993] "Geophagus_winemilleri"
## [995] "Gephyrocharax_major"
## [997] "Gladioglanis_anacanthus"
## [999] "Gladioglanis_machadoi"
## [1001] "Glanapteryx_niobium"
## [1003] "Gnathocharax_steindachneri"

```

```

## [1005] "Gnathodolus_bidens"
## [1007] "Grundulus_bogotensis"
## [1009] "Guianacara_dacrya"
## [1011] "Gymnocorymbus_ternetzi"
## [1013] "Gymnorhamphichthys_hypostomus"
## [1015] "Gymnorhamphichthys_rondoni"
## [1017] "Gymnotus_anguillaris"
## [1019] "Gymnotus_carapo"
## [1021] "Gymnotus_chaviro"
## [1023] "Gymnotus_coropinae"
## [1025] "Gymnotus_diamantinensis"
## [1027] "Gymnotus_javari"
## [1029] "Gymnotus_mamiraua"
## [1031] "Gymnotus_obscurus"
## [1033] "Gymnotus_pantanal"
## [1035] "Gymnotus_riberalta"
## [1037] "Gymnotus_tigre"
## [1039] "Gymnotus_ucamara"
## [1041] "Haemomaster_venezuelae"
## [1043] "Harttia_dissidens"
## [1045] "Harttia_rondoni"
## [1047] "Harttia_uatumensis"
## [1049] "Hassar_gabiru"
## [1051] "Hassar_shewellkeimi"
## [1053] "Heliotrygon_rosai"
## [1055] "Helogenes_marmoratus"
## [1057] "Hemiancistrus_landoni"
## [1059] "Hemiancistrus_sabaji"
## [1061] "Hemibrycon_beni"
## [1063] "Hemibrycon_helleri"
## [1065] "Hemibrycon_inambari"
## [1067] "Hemibrycon_metae"
## [1069] "Hemibrycon_orcesi"
## [1071] "Hemibrycon_surinamensis"
## [1073] "Hemidoras_morrisi"
## [1075] "Hemigrammus_aguaruna"
## [1077] "Hemigrammus_arua"
## [1079] "Hemigrammus_bleheri"
## [1081] "Hemigrammus_coeruleus"
## [1083] "Hemigrammus_cylindricus"
## [1085] "Hemigrammus_durbinae"
## [1087] "Hemigrammus_erythrozonous"
## [1089] "Hemigrammus_gracilis"
## [1091] "Hemigrammus_hyanuary"
## [1093] "Hemigrammus_levis"
## [1095] "Hemigrammus_lunatus"
## [1097] "Hemigrammus_marginatus"
## [1099] "Hemigrammus_melanochrous"
## [1101] "Hemigrammus_microstomus"
## [1103] "Hemigrammus_neptunus"
## [1105] "Hemigrammus_ocellifer"
## [1107] "Hemigrammus_pretoensis"
## [1109] "Goeldiella_eques"
## [1111] "Grundulus_cochae"
## [1113] "Gymnocorymbus_flaviolimai"
## [1115] "Gymnocorymbus_thayeri"
## [1117] "Gymnorhamphichthys_petiti"
## [1119] "Gymnorhamphichthys_rosamariae"
## [1121] "Gymnotus_arapaima"
## [1123] "Gymnotus_cataniapo"
## [1125] "Gymnotus_coatesi"
## [1127] "Gymnotus_curupira"
## [1129] "Gymnotus_eyra"
## [1131] "Gymnotus_jonasi"
## [1133] "Gymnotus_melanopleura"
## [1135] "Gymnotus_onca"
## [1137] "Gymnotus_pedanopterus"
## [1139] "Gymnotus_stenoleucus"
## [1141] "Gymnotus_tiquie"
## [1143] "Gymnotus_varzea"
## [1145] "Harttia_depressa"
## [1147] "Harttia_panara"
## [1149] "Harttia_trombetensis"
## [1151] "Hasemania_nambiquara"
## [1153] "Hassar_orestis"
## [1155] "Heliotrygon_gomesi"
## [1157] "Helogenes_gouldingi"
## [1159] "Hemiancistrus_guahiborum"
## [1161] "Hemiancistrus_pankimpuju"
## [1163] "Hemiancistrus_snethlageae"
## [1165] "Hemibrycon_divisorensis"
## [1167] "Hemibrycon_huambonicus"
## [1169] "Hemibrycon_jelskii"
## [1171] "Hemibrycon_mikrostiktos"
## [1173] "Hemibrycon_polyodon"
## [1175] "Hemibrycon_tridens"
## [1177] "Hemidoras_stenopeltis"
## [1179] "Hemigrammus_analis"
## [1181] "Hemigrammus_bellottii"
## [1183] "Hemigrammus_changae"
## [1185] "Hemigrammus_cupreus"
## [1187] "Hemigrammus_diagonicus"
## [1189] "Hemigrammus_elegans"
## [1191] "Hemigrammus_geisleri"
## [1193] "Hemigrammus_haraldi"
## [1195] "Hemigrammus_kuroobi"
## [1197] "Hemigrammus_luelingi"
## [1199] "Hemigrammus_machadoi"
## [1201] "Hemigrammus_megaceps"
## [1203] "Hemigrammus_micropterus"
## [1205] "Hemigrammus_mimus"
## [1207] "Hemigrammus_newboldi"
## [1209] "Hemigrammus_orthus"
## [1211] "Hemigrammus_pulcher"

```



```

## [1213] "Hypessobrycon_nigricinctus"
## [1215] "Hypessobrycon_oritoensis"
## [1217] "Hypessobrycon_pando"
## [1219] "Hypessobrycon_petricolus"
## [1221] "Hypessobrycon_pinnistriatus"
## [1223] "Hypessobrycon_platyodus"
## [1225] "Hypessobrycon_psittacus"
## [1227] "Hypessobrycon_pyrrhonotus"
## [1229] "Hypessobrycon_rosaceus"
## [1231] "Hypessobrycon_scutulatus"
## [1233] "Hypessobrycon_stramineus"
## [1235] "Hypessobrycon_taguae"
## [1237] "Hypessobrycon_taphorni"
## [1239] "Hypessobrycon_tropis"
## [1241] "Hypessobrycon_vanzolinii"
## [1243] "Hypessobrycon_wadai"
## [1245] "Hypoclinemus_mentalis"
## [1247] "Hypomasticus_julii"
## [1249] "Hypomasticus_megalepis"
## [1251] "Hypophthalmus_edentatus"
## [1253] "Hypophthalmus_marginatus"
## [1255] "Hypopomus_artedi"
## [1257] "Hypoptopoma_bianale"
## [1259] "Hypoptopoma_elongatum"
## [1261] "Hypoptopoma_incognitum"
## [1263] "Hypoptopoma_spectabile"
## [1265] "Hypoptopoma_sternoptychum"
## [1267] "Hypopygus_benoneae"
## [1269] "Hypopygus_hoedemani"
## [1271] "Hypopygus_minissimus"
## [1273] "Hypopygus_nijsseni"
## [1275] "Hypopygus_varii"
## [1277] "Hypostomus_bolivianus"
## [1279] "Hypostomus_dardanelos"
## [1281] "Hypostomus_faveolus"
## [1283] "Hypostomus_hemicochliodon"
## [1285] "Hypostomus_kopeyaka"
## [1287] "Hypostomus_latirostris"
## [1289] "Hypostomus_macushi"
## [1291] "Hypostomus_niceforoi"
## [1293] "Hypostomus_pantherinus"
## [1295] "Hypostomus_pyrineusi"
## [1297] "Hypostomus_sculpodon"
## [1299] "Hypostomus_taphorni"
## [1301] "Hypostomus_weberi"
## [1303] "Hypselecaria_temporalis"
## [1305] "Iguanodectes_geisleri"
## [1307] "Iguanodectes_polylepis"
## [1309] "Iguanodectes_rachovii"
## [1311] "Iguanodectes_variatus"
## [1313] "Imparfinis_cochabambae"
## [1315] "Imparfinis_hasemani"

## [1213] "Hypessobrycon_notidanos"
## [1215] "Hypessobrycon_paepkei"
## [1217] "Hypessobrycon_peruvianus"
## [1219] "Hypessobrycon_peugeotii"
## [1221] "Hypessobrycon_piranga"
## [1223] "Hypessobrycon_procyon"
## [1225] "Hypessobrycon_pulchripinnis"
## [1227] "Hypessobrycon_robustulus"
## [1229] "Hypessobrycon_scholzei"
## [1231] "Hypessobrycon_socolofii"
## [1233] "Hypessobrycon_swieglesi"
## [1235] "Hypessobrycon_takasei"
## [1237] "Hypessobrycon_tenuis"
## [1239] "Hypessobrycon_tukunai"
## [1241] "Hypessobrycon_vilmae"
## [1243] "Hypessobrycon_wosiackii"
## [1245] "Hypodoras_forficulatus"
## [1247] "Hypomasticus_lineomaculatus"
## [1249] "Hypomasticus_pachycheilus"
## [1251] "Hypophthalmus_fimbriatus"
## [1253] "Hypophthalmus_oremaculatus"
## [1255] "Hypoptopoma_baileyi"
## [1257] "Hypoptopoma_brevirostratum"
## [1259] "Hypoptopoma_gulare"
## [1261] "Hypoptopoma_psilogaster"
## [1263] "Hypoptopoma_steindachneri"
## [1265] "Hypoptopoma_thoracatum"
## [1267] "Hypopygus_cryptogenes"
## [1269] "Hypopygus_lepturus"
## [1271] "Hypopygus_neblinae"
## [1273] "Hypopygus_ortegai"
## [1275] "Hyporhamphus_brederi"
## [1277] "Hypostomus_carinatus"
## [1279] "Hypostomus_ericius"
## [1281] "Hypostomus_fonchii"
## [1283] "Hypostomus_hoplonites"
## [1285] "Hypostomus_kuarup"
## [1287] "Hypostomus_levis"
## [1289] "Hypostomus_melanephelis"
## [1291] "Hypostomus_oculeus"
## [1293] "Hypostomus_plecostomus"
## [1295] "Hypostomus_rondoni"
## [1297] "Hypostomus_soniae"
## [1299] "Hypostomus_varimaculosus"
## [1301] "Hypselecaria_coryphaenoides"
## [1303] "Iguanodectes_adujai"
## [1305] "Iguanodectes_gracilis"
## [1307] "Iguanodectes_purusii"
## [1309] "Iguanodectes_spilurus"
## [1311] "Ilisha_amazonica"
## [1313] "Imparfinis_guttatus"
## [1315] "Imparfinis_longicaudus"

```



```

## [1525] "Megalodoras_uranoscopus"
## [1527] "Megalonema_platycephalum"
## [1529] "Melanocharacidium_dispilomma"
## [1531] "Melanocharacidium_pectorale"
## [1533] "Melanorivulus_britzkei"
## [1535] "Melanorivulus_kayabi"
## [1537] "Melanorivulus_modestus"
## [1539] "Melanorivulus_zygonectes"
## [1541] "Mesonauta_egregius"
## [1543] "Mesonauta_guyanae"
## [1545] "Mesonauta_mirificus"
## [1547] "Metynnис_fasciatus"
## [1549] "Metynnис_hypsauchen"
## [1551] "Metynnис_longipinnis"
## [1553] "Metynnис_maculatus"
## [1555] "Metynnис_polystictus"
## [1557] "Microcharacidium_geryi"
## [1559] "Microcharacidium_weitzmani"
## [1561] "Microgenys_weyrauchi"
## [1563] "Microglanis_lundbergi"
## [1565] "Microglanis_poecilus"
## [1567] "Micromischodus_sugillatus"
## [1569] "Microphilypnus_acangaquara"
## [1571] "Microphilypnus_hypolyrasimeion"
## [1573] "Microphilypnus_tapajosensis"
## [1575] "Micropoecilia_branneri"
## [1577] "Microschemobrycon_casiquiare"
## [1579] "Microschemobrycon_geisleri"
## [1581] "Microschemobrycon_melanotus"
## [1583] "Microsternarchus_bilineatus"
## [1585] "Mikrogeophagus_altispinosus"
## [1587] "Moema_apurinan"
## [1589] "Moema_hellneri"
## [1591] "Moema_manuensis"
## [1593] "Moema_pepotei"
## [1595] "Moema_portugali"
## [1597] "Moema_rubrocaudatus"
## [1599] "Moema_wischmanni"
## [1601] "Moenkhausia_affinis"
## [1603] "Moenkhausia_atahualpiana"
## [1605] "Moenkhausia_barbouri"
## [1607] "Moenkhausia_beninei"
## [1609] "Moenkhausia_britskii"
## [1611] "Moenkhausia_celibela"
## [1613] "Moenkhausia_chlorophthalma"
## [1615] "Moenkhausia_collettii"
## [1617] "Moenkhausia_conspicua"
## [1619] "Moenkhausia_cosmops"
## [1621] "Moenkhausia_crisnejas"
## [1623] "Moenkhausia_diktyota"
## [1625] "Moenkhausia_eurystaenia"
## [1627] "Moenkhausia_grandisquamis"

## [1525] "Megalonema_amaxanthum"
## [1527] "Melanocharacidium_depressum"
## [1529] "Melanocharacidium_nigrum"
## [1531] "Melanocharacidium_rex"
## [1533] "Melanorivulus_canesi"
## [1535] "Melanorivulus_megaroni"
## [1537] "Melanorivulus_rubroreticulatus"
## [1539] "Melosternarchus_amaru"
## [1541] "Mesonauta_festivus"
## [1543] "Mesonauta_insignis"
## [1545] "Metynnис_argenteus"
## [1547] "Metynnис_guaporensis"
## [1549] "Metynnис_lippincottianus"
## [1551] "Metynnис_luna"
## [1553] "Metynnис_melanogrammus"
## [1555] "Microcharacidium_eleotrioides"
## [1557] "Microcharacidium_gnomus"
## [1559] "Microgenys_lativirgata"
## [1561] "Microglanis_iheringi"
## [1563] "Microglanis_pellopterygius"
## [1565] "Microglanis_zonatus"
## [1567] "Micromyzon_akamai"
## [1569] "Microphilypnus_amazonicus"
## [1571] "Microphilypnus_macrostoma"
## [1573] "Microphilypnus_ternetzi"
## [1575] "Microschemobrycon_callops"
## [1577] "Microschemobrycon_elongatus"
## [1579] "Microschemobrycon_guaporensis"
## [1581] "Microschemobrycon_meyburgi"
## [1583] "Microsternarchus_brevis"
## [1585] "Miuroglanis_platycephalus"
## [1587] "Moema_beucheyi"
## [1589] "Moema_kenwoodi"
## [1591] "Moema_nudifrontata"
## [1593] "Moema_peruensis"
## [1595] "Moema_quiii"
## [1597] "Moema_staecki"
## [1599] "Moenkhausia_abyss"
## [1601] "Moenkhausia_agnesae"
## [1603] "Moenkhausia_aurantia"
## [1605] "Moenkhausia_bellasomniosa"
## [1607] "Moenkhausia_bonita"
## [1609] "Moenkhausia_browni"
## [1611] "Moenkhausia_ceros"
## [1613] "Moenkhausia_chrysargyrea"
## [1615] "Moenkhausia_comma"
## [1617] "Moenkhausia_copei"
## [1619] "Moenkhausia_cotinho"
## [1621] "Moenkhausia_dichroura"
## [1623] "Moenkhausia_dorsinuda"
## [1625] "Moenkhausia_gracilima"
## [1627] "Moenkhausia_hasemani"

```



```

## [1837] "Phenacogaster_beni"
## [1839] "Phenacogaster_megalostictus"
## [1841] "Phenacogaster_napoatilis"
## [1843] "Phenacogaster_pectinatus"
## [1845] "Phenacogaster_retropinnus"
## [1847] "Phenacorhamdia_nigrolineata"
## [1849] "Phreatobius_cisternarum"
## [1851] "Phreatobius_sanguijuela"
## [1853] "Physopyxis_ananas"
## [1855] "Physopyxis_lyra"
## [1857] "Piabucina_elongata"
## [1859] "Piabucus_dentatus"
## [1861] "Piaractus_brachypomus"
## [1863] "Pimelodella_bockmanni"
## [1865] "Pimelodella_breviceps"
## [1867] "Pimelodella_chaparae"
## [1869] "Pimelodella_cristata"
## [1871] "Pimelodella_gracilis"
## [1873] "Pimelodella_hartwelli"
## [1875] "Pimelodella_howesi"
## [1877] "Pimelodella_montana"
## [1879] "Pimelodella_peruana"
## [1881] "Pimelodella_serrata"
## [1883] "Pimelodina_flavipinnis"
## [1885] "Pimelodus_altissimus"
## [1887] "Pimelodus_jivaro"
## [1889] "Pimelodus_ornatus"
## [1891] "Pimelodus_tetramerus"
## [1893] "Pituna_xinguensis"
## [1895] "Plagioscion_monte"
## [1897] "Plagioscion_surinamensis"
## [1899] "Platydoras_birindellii"
## [1901] "Platydoras_hancockii"
## [1903] "Platysilurus_mucosus"
## [1905] "Platystacus_cotylephorus"
## [1907] "Platyurosternarchus_crypticus"
## [1909] "Plectrochilus_diabolicus"
## [1911] "Plectrochilus_wieneri"
## [1913] "Plesiotrygon_iwamae"
## [1915] "Poecilia_waiapi"
## [1917] "Polycentrus_jundia"
## [1919] "Poptella_compressa"
## [1921] "Porotergus_duende"
## [1923] "Porotergus_gymnotus"
## [1925] "Potamorhina_altamazonica"
## [1927] "Potamorhina_pristigaster"
## [1929] "Potamorrhaphis_guianensis"
## [1931] "Potamorrhaphis_petersi"
## [1933] "Potamotrygon_albimaculata"
## [1935] "Potamotrygon_constellata"
## [1937] "Potamotrygon_humerosa"
## [1939] "Potamotrygon_leopoldi"

## [1837] "Phenacogaster_capitulatus"
## [1839] "Phenacogaster_microstictus"
## [1841] "Phenacogaster_ojitatus"
## [1843] "Phenacogaster_prolatus"
## [1845] "Phenacorhamdia_boliviiana"
## [1847] "Phractocephalus_hemiopterus"
## [1849] "Phreatobius_dracunculus"
## [1851] "Phycocharax_rasbora"
## [1853] "Physopyxis_cristata"
## [1855] "Piabarchus_analis"
## [1857] "Piabucus_caudomaculatus"
## [1859] "Piabucus_melanostoma"
## [1861] "Pimelodella_altipinnis"
## [1863] "Pimelodella_boliviiana"
## [1865] "Pimelodella_buckleyi"
## [1867] "Pimelodella_conquetaensis"
## [1869] "Pimelodella_cyanostigma"
## [1871] "Pimelodella_grisea"
## [1873] "Pimelodella_hasemani"
## [1875] "Pimelodella_humeralis"
## [1877] "Pimelodella_nigrofasciata"
## [1879] "Pimelodella_roccae"
## [1881] "Pimelodella_steindachneri"
## [1883] "Pimelodus_albofasciatus"
## [1885] "Pimelodus_blochii"
## [1887] "Pimelodus_maculatus"
## [1889] "Pimelodus_pictus"
## [1891] "Pinirampus_pirinampu"
## [1893] "Plagioscion_auratus"
## [1895] "Plagioscion_squamossissimus"
## [1897] "Planiloricaria_cryptodon"
## [1899] "Platydoras_costatus"
## [1901] "Platynematicichthys_notatus"
## [1903] "Platysilurus.olallae"
## [1905] "Platystomatichthys_sturio"
## [1907] "Platyurosternarchus_macrostomus"
## [1909] "Plectrochilus_machadoi"
## [1911] "Plesiolebias_altamira"
## [1913] "Plesiotrygon_nana"
## [1915] "Poecilocharax_weitzmani"
## [1917] "Poptella_brevispina"
## [1919] "Poptella_longipinnis"
## [1921] "Porotergus_gimbeli"
## [1923] "Potamoglanis_wapixana"
## [1925] "Potamorhina_latior"
## [1927] "Potamorrhaphis_eigenmanni"
## [1929] "Potamorrhaphis_labiatu"
## [1931] "Potamotrygon_adamastor"
## [1933] "Potamotrygon_amazona"
## [1935] "Potamotrygon_falkneri"
## [1937] "Potamotrygon_jabuti"
## [1939] "Potamotrygon_limai"

```

```

## [1941] "Potamotrygon_motoro"
## [1943] "Potamotrygon_schroederi"
## [1945] "Potamotrygon_tatianae"
## [1947] "Potamotrygon_wallacei"
## [1949] "Priocharax_nanus"
## [1951] "Prionobrama_filigera"
## [1953] "Pristigaster_cayana"
## [1955] "Pristobrycon_aureus"
## [1957] "Pristobrycon_maculipinnis"
## [1959] "Procerusternarchus_pixuna"
## [1961] "Prochilodus_nigricans"
## [1963] "Prodontocharax_alleni"
## [1965] "Prodontocharax_melanotus"
## [1967] "Propimelodus_eigenmanni"
## [1969] "Psectrogaster_ciliata"
## [1971] "Psectrogaster_essequibensis"
## [1973] "Psectrogaster_rhombooides"
## [1975] "Pseudacanthicus_histrix"
## [1977] "Pseudacanthicus_spinosus"
## [1979] "Pseudancistrus_kayabi"
## [1981] "Pseudancistrus_pectegenitor"
## [1983] "Pseudancistrus_zawadzkii"
## [1985] "Pseudanos_trimaculatus"
## [1987] "Pseudanos_winterbottomi"
## [1989] "Pseudopapterus_cucuhyensis"
## [1991] "Pseudobunocephalus_amazonicus"
## [1993] "Pseudobunocephalus_quadriradiatus"
## [1995] "Pseudohemiodon_apithanos"
## [1997] "Pseudohemiodon_thorectes"
## [1999] "Pseudolithoxus_nicoi"
## [2001] "Pseudopimelodus_bufonius"
## [2003] "Pseudopimelodus_pulcher"
## [2005] "Pseudoplatystoma_punctifer"
## [2007] "Pseudoplatystoma_tigrinum"
## [2009] "Pseudostegophilus_haemomyzon"
## [2011] "Pseudotylosurus_angusticeps"
## [2013] "Pterobunocephalus_depressus"
## [2015] "Pterodoras_granulosus"
## [2017] "Pterophyllum_altum"
## [2019] "Pterophyllum_scalare"
## [2021] "Pterygoplichthys_disjunctivus"
## [2023] "Pterygoplichthys_lituratus"
## [2025] "Pterygoplichthys_pardalis"
## [2027] "Pterygoplichthys_scrophus"
## [2029] "Pterygoplichthys_xinguensis"
## [2031] "Pygidianops_amphioxus"
## [2033] "Pygocentrus_nattereri"
## [2035] "Pyrrhulina_australis"
## [2037] "Pyrrhulina_brevis"
## [2039] "Pyrrhulina_elongata"
## [2041] "Pyrrhulina_lugubris"
## [2043] "Pyrrhulina_maxima"
## [2045] "Potamotrygon_orbignyi"
## [2047] "Potamotrygon_scobina"
## [2049] "Potamotrygon_tigrina"
## [2051] "Priocharax_ariel"
## [2053] "Priocharax_pygmaeus"
## [2055] "Pristella_maxillaris"
## [2057] "Pristigaster_whiteheadi"
## [2059] "Pristobrycon_calmoni"
## [2061] "Pristobrycon_striolatus"
## [2063] "Prochilodus_britskii"
## [2065] "Prochilodus_rubrotaeniatus"
## [2067] "Prodontocharax_howesi"
## [2069] "Propimelodus_caesius"
## [2071] "Psectrogaster_amazonica"
## [2073] "Psectrogaster_curviventris"
## [2075] "Psectrogaster_falcata"
## [2077] "Psectrogaster_rutiloides"
## [2079] "Pseudacanthicus_pirarara"
## [2081] "Pseudancistrus_asurini"
## [2083] "Pseudancistrus_orinoco"
## [2085] "Pseudancistrus_sidereus"
## [2087] "Pseudanos_gracilis"
## [2089] "Pseudanos_varii"
## [2091] "Pseudauchenipterus_nodosus"
## [2093] "Pseudopapterus_hasemani"
## [2095] "Pseudobunocephalus_bifidus"
## [2097] "Pseudohemiodon_amazonum"
## [2099] "Pseudohemiodon_lamina"
## [2101] "Pseudolithoxus_kinja"
## [2103] "Pseudoloricaria_laeviuscula"
## [2105] "Pseudopimelodus_mangurus"
## [2107] "Pseudoplatystoma_fasciatum"
## [2109] "Pseudoplatystoma_reticulatum"
## [2111] "Pseudorinelepis_genibarbis"
## [2113] "Pseudostegophilus_nemurus"
## [2115] "Pseudotylosurus_microps"
## [2117] "Pterobunocephalus_dolichurus"
## [2119] "Pterolebias_longipinnis"
## [2121] "Pterophyllum_leopoldi"
## [2123] "Pterosturisoma_microps"
## [2125] "Pterygoplichthys_gibbiceps"
## [2127] "Pterygoplichthys_multiradiatus"
## [2129] "Pterygoplichthys_punctatus"
## [2131] "Pterygoplichthys_weberi"
## [2133] "Ptychocharax_rhyacophila"
## [2135] "Pygidianops_eigenmanni"
## [2137] "Pygopristis_denticulata"
## [2139] "Pyrrhulina_beni"
## [2141] "Pyrrhulina_eleanorae"
## [2143] "Pyrrhulina_laeta"
## [2145] "Pyrrhulina_marilynae"
## [2147] "Pyrrhulina_melanostoma"

```

```

## [2045] "Pyrrhulina_obermulleri"
## [2047] "Pyrrhulina_spilota"
## [2049] "Pyrrhulina_vittata"
## [2051] "Racenisia_fimbriipinna"
## [2053] "Retroculus_xinguensis"
## [2055] "Rhabdolichops_eastwardi"
## [2057] "Rhabdolichops_lundbergi"
## [2059] "Rhabdolichops_nigrimans"
## [2061] "Rhabdolichops_troscheli"
## [2063] "Rhamdella_montana"
## [2065] "Rhamdia_foina"
## [2067] "Rhamdia_muelleri"
## [2069] "Rhamdia_quelen"
## [2071] "Rhamphichthys_heleios"
## [2073] "Rhamphichthys_longior"
## [2075] "Rhamphichthys_rostratus"
## [2077] "Rhinobrycon_negrensis"
## [2079] "Rhinodoras_boehlkei"
## [2081] "Rhinosardinia_amazonica"
## [2083] "Rhyacoglanis_seminiger"
## [2085] "Rhynchodoras_xingui"
## [2087] "Rhytiodus_elongatus"
## [2089] "Rhytiodus_microlepis"
## [2091] "Rineloricaria_castroi"
## [2093] "Rineloricaria_fallax"
## [2095] "Rineloricaria_hasemani"
## [2097] "Rineloricaria_jurupari"
## [2099] "Rineloricaria_lanceolata"
## [2101] "Rineloricaria_morrowi"
## [2103] "Rineloricaria_phoxocephala"
## [2105] "Rineloricaria_wolfei"
## [2107] "Roeboides_affinis"
## [2109] "Roeboides_descalvadensis"
## [2111] "Roeboides_myersii"
## [2113] "Roestes_molossus"
## [2115] "Salminus_affinis"
## [2117] "Salminus_hilarii"
## [2119] "Sartor_elongatus"
## [2121] "Satanoperca_acuticeps"
## [2123] "Satanoperca_daemon"
## [2125] "Satanoperca_leucosticta"
## [2127] "Satanoperca_pappaterra"
## [2129] "Schizodon_fasciatus"
## [2131] "Scobinancistrus_aureatus"
## [2133] "Scoloplax_baileyi"
## [2135] "Scoloplax_dicra"
## [2137] "Scoloplax_dolicholophia"
## [2139] "Scopaeocharax_atopodus"
## [2141] "Scorpiodoras_calderonensis"
## [2143] "Scorpiodoras_liophysus"
## [2145] "Semaprochilodus_insignis"
## [2147] "Semaprochilodus_taeniurus"

## [2045] "Pyrrhulina_semifasciata"
## [2047] "Pyrrhulina_stoli"
## [2049] "Pyrrhulina_zigzag"
## [2051] "Reganella_depressa"
## [2053] "Rhabdolichops_caviceps"
## [2055] "Rhabdolichops_electrogrammus"
## [2057] "Rhabdolichops_naavalha"
## [2059] "Rhabdolichops_stewarti"
## [2061] "Rhadinoloricaria_macromystax"
## [2063] "Rhamdella_rusbyi"
## [2065] "Rhamdia_laukidi"
## [2067] "Rhamdia_poeyi"
## [2069] "Rhamphichthys_drepanium"
## [2071] "Rhamphichthys_lineatus"
## [2073] "Rhamphichthys_marmoratus"
## [2075] "Rhaphiodon_vulpinus"
## [2077] "Rhinodoras_armbrusteri"
## [2079] "Rhinopetitia_potamorhachia"
## [2081] "Rhyacoglanis_epiblepsis"
## [2083] "Rhynchodoras_woodsi"
## [2085] "Rhytiodus_argenteofuscus"
## [2087] "Rhytiodus_lauzannei"
## [2089] "Rineloricaria_beni"
## [2091] "Rineloricaria_daraha"
## [2093] "Rineloricaria_formosa"
## [2095] "Rineloricaria_heteroptera"
## [2097] "Rineloricaria_konopickyi"
## [2099] "Rineloricaria_melini"
## [2101] "Rineloricaria_nigricauda"
## [2103] "Rineloricaria_teffeana"
## [2105] "Roeboexodon_guyanensis"
## [2107] "Roeboides_biserialis"
## [2109] "Roeboides_dispar"
## [2111] "Roeboides_oligistos"
## [2113] "Roestes_ogilviei"
## [2115] "Salminus_brasiliensis"
## [2117] "Sarcoglanis_simplex"
## [2119] "Sartor_respectus"
## [2121] "Satanoperca_curupira"
## [2123] "Satanoperca_jurupari"
## [2125] "Satanoperca_lilith"
## [2127] "Schizodon_dissimilis"
## [2129] "Schultzichthys_bondi"
## [2131] "Scobinancistrus_pariolispos"
## [2133] "Scoloplax_baskini"
## [2135] "Scoloplax_distolothrix"
## [2137] "Scoloplax_empousa"
## [2139] "Scopaeocharax_rhinodus"
## [2141] "Scorpiodoras_heckelii"
## [2143] "Semaprochilodus_brama"
## [2145] "Semaprochilodus_kneri"
## [2147] "Serrabrycon_magoi"

```

```

## [2149] "Serrapinnus_heterodon"
## [2151] "Serrapinnus_micropterus"
## [2153] "Serrasalmus_altuvei"
## [2155] "Serrasalmus_eigenmanni"
## [2157] "Serrasalmus_gouldingi"
## [2159] "Serrasalmus_hollandi"
## [2161] "Serrasalmus_maculatus"
## [2163] "Serrasalmus_nigricans"
## [2165] "Serrasalmus_rhombeus"
## [2167] "Serrasalmus_serrulatus"
## [2169] "Simpsonichthys_filamentosus"
## [2171] "Sorubim_lima"
## [2173] "Sorubim_trigocephalus"
## [2175] "Spatuloricaria_caquetae"
## [2177] "Spatuloricaria_evansii"
## [2179] "Spatuloricaria_tuira"
## [2181] "Spectracanthicus_murinus"
## [2183] "Spectrolebias_bellidoi"
## [2185] "Spectrolebias_filamentosus"
## [2187] "Spectrolebias_reticulatus"
## [2189] "Squaliforma_emarginata"
## [2191] "Squaliforma_phrixosoma"
## [2193] "Squaliforma_squalina"
## [2195] "Stauroglanis_gouldingi"
## [2197] "Steatogenys_elegans"
## [2199] "Stegophilus_panzeri"
## [2201] "Steindachnerina_bimaculata"
## [2203] "Steindachnerina_dobula"
## [2205] "Steindachnerina_guentheri"
## [2207] "Steindachnerina_leucisca"
## [2209] "Steindachnerina_pupula"
## [2211] "Steindachnerina_serriata"
## [2213] "Stenolicmus_sarmientoi"
## [2215] "Sternarchella_patriciae"
## [2217] "Sternarchella_schotti"
## [2219] "Sternarchella_terminalis"
## [2221] "Sternarchogiton_nattereri"
## [2223] "Sternarchogiton_preto"
## [2225] "Sternarchorhamphus_muelleri"
## [2227] "Sternarchorhynchus_caboclo"
## [2229] "Sternarchorhynchus_cramptoni"
## [2231] "Sternarchorhynchus_curvirostris"
## [2233] "Sternarchorhynchus_hagedornae"
## [2235] "Sternarchorhynchus_inpai"
## [2237] "Sternarchorhynchus_kokraimoro"
## [2239] "Sternarchorhynchus_montanus"
## [2241] "Sternarchorhynchus_retzeri"
## [2243] "Sternarchorhynchus_starksii"
## [2245] "Sternarchorhynchus_taphorni"
## [2247] "Sternopygus_aequilabiatus"
## [2249] "Sternopygus_branco"
## [2251] "Sternopygus_obtusirostris"

## [2149] "Serrapinnus_microdon"
## [2151] "Serrasalmus_altispinis"
## [2153] "Serrasalmus_compressus"
## [2155] "Serrasalmus_elongatus"
## [2157] "Serrasalmus_hastatus"
## [2159] "Serrasalmus_humeralis"
## [2161] "Serrasalmus_manueli"
## [2163] "Serrasalmus_odyssei"
## [2165] "Serrasalmus_sanchezi"
## [2167] "Serrasalmus_spilopleura"
## [2169] "Sorubim_elongatus"
## [2171] "Sorubim_maniradii"
## [2173] "Sorubimichthys_planiceps"
## [2175] "Spatuloricaria_euacanthagenys"
## [2177] "Spatuloricaria_puganensis"
## [2179] "Spectracanthicus_immaculatus"
## [2181] "Spectracanthicus_zuanoni"
## [2183] "Spectrolebias_brousseauti"
## [2185] "Spectrolebias_pilletti"
## [2187] "Spinipterus_acsi"
## [2189] "Squaliforma_horrifica"
## [2191] "Squaliforma_scopularia"
## [2193] "Squaliforma_virescens"
## [2195] "Steatogenys_duidae"
## [2197] "Steatogenys_ocellatus"
## [2199] "Steindachnerina_argentea"
## [2201] "Steindachnerina_binotata"
## [2203] "Steindachnerina_fasciata"
## [2205] "Steindachnerina_hypostoma"
## [2207] "Steindachnerina_planiventris"
## [2209] "Steindachnerina_quasimodoi"
## [2211] "Stenolicmus_ix"
## [2213] "Sternarchella_calhamazon"
## [2215] "Sternarchella_rex"
## [2217] "Sternarchella_sima"
## [2219] "Sternarchogiton_labiatus"
## [2221] "Sternarchogiton_porcinum"
## [2223] "Sternarchogiton_zuanoni"
## [2225] "Sternarchorhynchus_axelrodi"
## [2227] "Sternarchorhynchus_chaoi"
## [2229] "Sternarchorhynchus_curumim"
## [2231] "Sternarchorhynchus_goeldii"
## [2233] "Sternarchorhynchus_higuchii"
## [2235] "Sternarchorhynchus_jaimei"
## [2237] "Sternarchorhynchus_mareikeae"
## [2239] "Sternarchorhynchus_mormyrus"
## [2241] "Sternarchorhynchus_severii"
## [2243] "Sternarchorhynchus_stewarti"
## [2245] "Sternarchorhynchus_villasboasi"
## [2247] "Sternopygus_astrabes"
## [2249] "Sternopygus_macrurus"
## [2251] "Sternopygus_xingu"

```

```

## [2253] "Stethaprion_crenatum"
## [2255] "Stichonodon_insignis"
## [2257] "Sturisoma_brevirostre"
## [2259] "Sturisoma_guentheri"
## [2261] "Sturisoma_nigrirostrum"
## [2263] "Sturisoma_rostratum"
## [2265] "Symphysodon_aequifasciatus"
## [2267] "Symphysodon_tarzoo"
## [2269] "Synbranchus_lampreia"
## [2271] "Synbranchus_marmoratus"
## [2273] "Tahuantinsuyo_chipi"
## [2275] "Tarumania_walkerae"
## [2277] "Tatia_brunnea"
## [2279] "Tatia_dunni"
## [2281] "Tatia_intermedia"
## [2283] "Tatia_musaica"
## [2285] "Tatia_strigata"
## [2287] "Teleocichla_centrarchus"
## [2289] "Teleocichla_monogramma"
## [2291] "Teleocichla_prionogenys"
## [2293] "Teleocichla_wajapi"
## [2295] "Tetragonopterus_carvalhoi"
## [2297] "Tetragonopterus_juruena"
## [2299] "Tetragonopterus_manaos"
## [2301] "Tetragonopterus_paucidens"
## [2303] "Tetranemichthys_quadrifilis"
## [2305] "Thalassophryne_amazonica"
## [2307] "Thayeria_obliqua"
## [2309] "Thoracocharax_securis"
## [2311] "Thrissobrycon_pectinifer"
## [2313] "Tometes_ancylorhynchus"
## [2315] "Tometes_kranponhah"
## [2317] "Trachelyichthys_decaradiatus"
## [2319] "Trachelyopterichthys_anduzei"
## [2321] "Trachelyopterus_ceratophy whole
## [2323] "Trachelyopterus_galeatus"
## [2325] "Trachycorystes_trachycorystes"
## [2327] "Trachydoras_gepharti"
## [2329] "Trachydoras_nattereri"
## [2331] "Trichomycterus_aguarague"
## [2333] "Trichomycterus_barbouri"
## [2335] "Trichomycterus_chaberti"
## [2337] "Trichomycterus_dispar"
## [2339] "Trichomycterus_gabrieli"
## [2341] "Trichomycterus_johnsoni"
## [2343] "Trichomycterus_megantoni"
## [2345] "Trichomycterus_taczanowskii"
## [2347] "Trichomycterus_taeniopterus"
## [2349] "Trichomycterus_weyrauchi"
## [2351] "Tridensimilis_brevis"
## [2353] "Trigonectes_macrophthalmus"
## [2355] "Triportheus_albus"
## [2253] "Stethaprion_erythrops"
## [2255] "Strongylura_marina"
## [2257] "Sturisoma_graffini"
## [2259] "Sturisoma_lyra"
## [2261] "Sturisoma_robustum"
## [2263] "Sturisomatichthys_caquetae"
## [2265] "Symphysodon_discus"
## [2267] "Synaptolaemus_latofasciatus"
## [2269] "Synbranchus_madeirae"
## [2271] "Taeniacara_candidi"
## [2273] "Tahuantinsuyo_macantzatza"
## [2275] "Tatia_auropygia"
## [2277] "Tatia_caxiuaneensis"
## [2279] "Tatia_gyrina"
## [2281] "Tatia_melanoleuca"
## [2283] "Tatia_nigra"
## [2285] "Teleocichla_centisquama"
## [2287] "Teleocichla_gephyrogramma"
## [2289] "Teleocichla_preta"
## [2291] "Teleocichla_proselytus"
## [2293] "Tetragonopterus_argenteus"
## [2295] "Tetragonopterus_chalceus"
## [2297] "Tetragonopterus_kuluene"
## [2299] "Tetragonopterus_ommatus"
## [2301] "Tetranemichthys_barthemi"
## [2303] "Tetranemichthys_wallacei"
## [2305] "Thayeria_boehlkei"
## [2307] "Thayeria_tapajonica"
## [2309] "Thoracocharax_stellatus"
## [2311] "Tocantinsia_piresi"
## [2313] "Tometes_camunani"
## [2315] "Tometes_makue"
## [2317] "Trachelyichthys_exilis"
## [2319] "Trachelyopterichthys_taeniatus"
## [2321] "Trachelyopterus_coriaceus"
## [2323] "Trachycorystes_menezesi"
## [2325] "Trachydoras_brevis"
## [2327] "Trachydoras_microstomus"
## [2329] "Trachydoras_steindachneri"
## [2331] "Trichomycterus_anhangae"
## [2333] "Trichomycterus_bomboizanus"
## [2335] "Trichomycterus_corduvensis"
## [2337] "Trichomycterus_fassli"
## [2339] "Trichomycterus_hasemani"
## [2341] "Trichomycterus_knerii"
## [2343] "Trichomycterus_rivulatus"
## [2345] "Trichomycterus_taenia"
## [2347] "Trichomycterus_tiraquae"
## [2349] "Tridens_melanops"
## [2351] "Tridentopsis_pearsoni"
## [2353] "Trigonectes_rogoguae"
## [2355] "Triportheus_angulatus"

```

```

## [2357] "Triportheus_auritus"
## [2359] "Triportheus_culter"
## [2361] "Triportheus_pictus"
## [2363] "Tucanoichthys_tucano"
## [2365] "Tympanopleura_brevis"
## [2367] "Tympanopleura_longipinna"
## [2369] "Typhlobelus_auriculatus"
## [2371] "Tyttocharax_dorsimaculatus"
## [2373] "Tyttocharax_spinosus"
## [2375] "Tyttocharax_cochui"
## [2377] "Tyttocharax_tambopatensis"
## [2379] "Utiaritichthys_esguiceroi"
## [2381] "Utiaritichthys_sennaebragai"
## [2383] "Vandellia_sanguinea"
## [2385] "Xenurobrycon_polyancistrus"
## [2387] "Xenurobrycon_varii"
## [2389] "Xyliophius_melanopterus"
## [2391] "Zungaro_zungaro"
"Triportheus_brachipomus"
"Triportheus_curtus"
"Triportheus_rotundatus"
"Tympanopleura_atronasus"
"Tympanopleura_cryptica"
"Tympanopleura_rondoni"
"Typhlobelus_ternetzi"
"Tyttobrycon_hamatus"
"Tyttobrycon_xeruini"
"Tyuttocharax_madeirae"
"Uaru_amphiacanthoides"
"Utiaritichthys_longidorsalis"
"Vandellia_cirrhosa"
"Xenurobrycon_heterodon"
"Xenurobrycon_pteropus"
"Xyliophius_lepturus"
"Xyliophius_sofiae"
"Zungaropsis_multimaculatus"

```

```

fish_mat$occ<-1
library(reshape2)
fish_mat<-dcast(fish_mat,basin~Referent.Species.Name,
                 fun.aggregate = sum, value.var = "occ")
dim(fish_mat)

```

```

## [1] 97 2393

```

```

fish_mat[1:10,1:5]

```

	basin	Abramites_hypselonotus	Acanthicus_hystrix	
## 1	Abuna	0	0	
## 2	Amazon1	0	0	
## 3	Amazon2	1	0	
## 4	Amazon5	1	0	
## 5	Amazon8	0	0	
## 6	Amazon9	1	0	
## 7	Andira1	0	0	
## 8	Apaporis	0	0	
## 9	Apurimac1	0	0	
## 10	Arinos	0	0	
	Acanthobunocephalus_nicoi	Acanthodoras_cataphractus		
## 1		0	1	
## 2		0	0	
## 3		0	1	
## 4		0	0	
## 5		0	1	
## 6		0	0	
## 7		0	1	
## 8		0	0	
## 9		0	0	
## 10		0	0	

```

rownames(fish_mat) <- fish_mat$basin
fish_mat <- fish_mat[,-c(1)]
fish_mat[fish_mat>0] <- 1

#setting NA to 0
fish_mat[is.na(fish_mat)] <- 0

#####
# biz<-(fish_mat[c("Pachitea", "Ucayali2", "Apurimac1", "Urubamba",
#                  "Huallaga", "Maranon5", "Maranon3", "Santiago", "Pastaza", "
#                  Curaray", "Napo2"),-1])
# biz2<-fish_mat[c("Pachitea", "Ucayali2", "Apurimac1", "Urubamba",
#                  "Huallaga", "Maranon5", "Maranon3", "Santiago", "Pastaza", "
#                  Curaray", "Napo2"),apply(fish_mat,2,sum,na.rm=T)==1]
# biz2 <- biz2[,apply(biz2,2,sum,na.rm=T)==1]
# dim(biz2)
# rowSums(biz2)
# id1<-grep(pattern = c("Maranon"),rownames(fish_mat))
# id2<-grep(pattern = c("Ucayali"),rownames(fish_mat))
# id3<-grep(pattern = c("Tapaj"),rownames(fish_mat))
# id4<-grep(pattern = c("Xingu"),rownames(fish_mat))
# id<-c(id1,id2,id3,id4);rm(id1,id2,id3,id4)
# biz3<-fish_mat[id,apply(fish_mat,2,sum,na.rm=T)==1]
# biz3<-biz3[,!apply(biz3,2,sum,na.rm=T)==0]
# dim(biz3)
# biz3$basin<-rownames(biz3)
#
# #especies endemicas nos andes
# biz=reshape2::melt(biz3)
# biz[is.na(biz)]<-0
#
# #especies totais nos andes
# biz<-droplevels(biz[biz$value>0,])
# dim(biz)
# head(biz)
# table(biz$variable)

#especies e generos das bacias selecionadas
# biz<-droplevels(biz[biz$value==1,])
# head(biz)
# dim(biz)
# biz<-biz[order(biz$basin),]
# write.table(biz,"especies_generos_OccUnica_baciasSelecionadas.csv",sep=";")

#####

```

```

library(vegan)

## Loading required package: permute

#####
# ## 2024_04_18 eliminating undersampled drainages
# undersampled<-c("Jamanxin", "Jari", "Paru_Este", "Demini", "Tapaua", "Grande", "Apurimac", "Ucayali2", "C
# fish_mat<-fish_mat[!fish_mat$Sub_drainage %in% undersampled,]
#####

```

1.2 Calculating and Decomposing fish dissimilarity

Variation in species composition among pair of sites can be decomposed in two phenomenons: species **turnover** and *nestedness* (Baselga 2009, 2012). The former represents the proportion of species replaced from one site to another and is independent of differences in species richness among sites. The latter quantifies differences in species composition that are related to difference in species richness among sites. This partitioning is useful to separate confounding processes acting on variation on species composition (Baselga 2009, 2012). Such partitioning can be done using two distinct dissimilarity indices (i.e., *Sorensen* and *Jaccard*), but the rationale behind this procedure is the same whatever the chosen index. We opted here to use the *Sorensen* index because it is one of the most used index in ecological studies. The **betapart** package is used to calculate both partitions.

```

library(ape)
library(RColorBrewer)
library(recluster)

D_fish<- betapart::beta.pair(fish_mat,index.family = "sorensen")
str(D_fish)

## List of 3
## $ beta.sim: 'dist' num [1:4656] 0.705 0.337 0.292 0.342 0.346 ...
##   ..- attr(*, "Labels")= chr [1:97] "Abuna" "Amazon1" "Amazon2" "Amazon5" ...
##   ..- attr(*, "Size")= int 97
##   ..- attr(*, "call")= language as.dist.default(m = beta.sim)
##   ..- attr(*, "Diag")= logi FALSE
##   ..- attr(*, "Upper")= logi FALSE
## $ beta.sne: 'dist' num [1:4656] 0.0472 0.2746 0.3213 0.277 0.2548 ...
##   ..- attr(*, "Labels")= chr [1:97] "Abuna" "Amazon1" "Amazon2" "Amazon5" ...
##   ..- attr(*, "Size")= int 97
##   ..- attr(*, "call")= language as.dist.default(m = beta.sne)
##   ..- attr(*, "Diag")= logi FALSE
##   ..- attr(*, "Upper")= logi FALSE
## $ beta.sor: 'dist' num [1:4656] 0.752 0.612 0.613 0.619 0.601 ...
##   ..- attr(*, "Labels")= chr [1:97] "Abuna" "Amazon1" "Amazon2" "Amazon5" ...
##   ..- attr(*, "Size")= int 97
##   ..- attr(*, "call")= language as.dist.default(m = beta.sor)
##   ..- attr(*, "Diag")= logi FALSE
##   ..- attr(*, "Upper")= logi FALSE

```

```

# #excluding species per site matrix
#rm(fish_mat)
fish_tree <- (hclust((D_fish$beta.sim),method = 'average'))
# cores<-cutree(fish_tree,h=0.5)
# cores<-rainbow(length(unique(cores)))[cores]
# names(cores)==as.phylo(fish_tree)$tip.label
# cores<-brewer.pal(n = 11, name = 'Paired')[cores]

#Definindo numero de cluster Kreft & Jetz 2010
d2 <- cophenetic(fish_tree)
#transform hclust to phylogenetic tree (ape)
tree=as.phylo(fish_tree)
#inspect explained diversity for different cuts of a tree
expl_div<-recluster.expl.diss(tree,D_fish$beta.sim,maxcl=30)
#expl_div
R2=expl_div$expl.div
clusters=expl_div$nclust#[ -length(expl_div$nclust)]
library(segmented)

## Loading required package: MASS

## Loading required package: nlme

out.lm<-lm(R2~clusters)
o<-segmented(out.lm,seg.Z=~clusters,psi=list(clusters=c(15)),
control=seg.control(display=FALSE))
slope(o)

## $clusters
##           Est. St.Err. t value CI(95%).l CI(95%).u
## slope1 0.0454  0.0044    10.3    0.0363    0.054
## slope2 0.0091  0.0012     7.4    0.0066    0.012

o

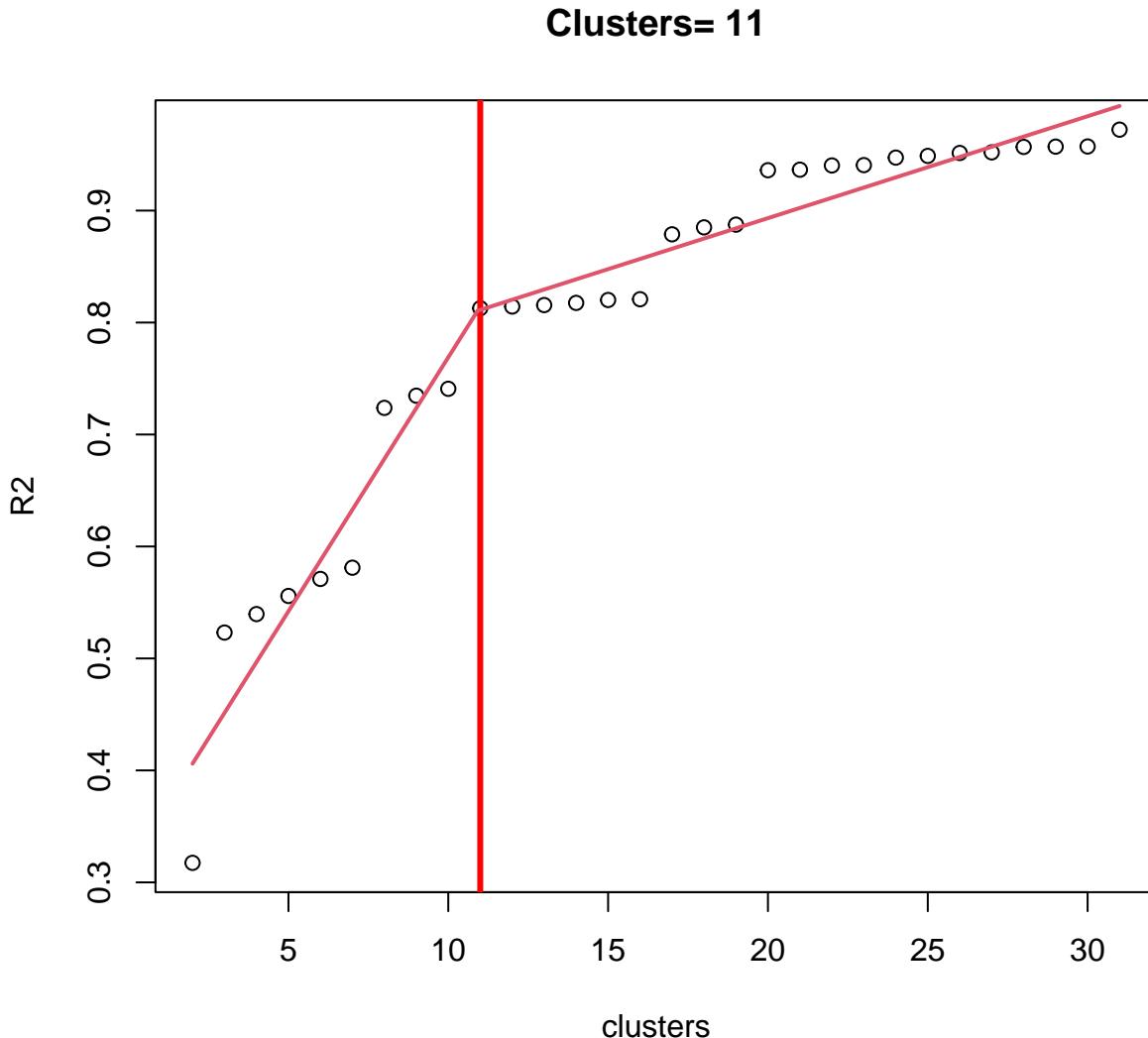
## Call: segmented.lm(obj = out.lm, seg.Z = ~clusters, psi = list(clusters = c(15)),
## control = seg.control(display = FALSE))
##
## Coefficients of the linear terms:
## (Intercept)   clusters  U1.clusters
##      0.3153      0.0454     -0.0363
##
## Estimated Break-Point(s):
## psi1.clusters
##            10.9

k=round(o$psi[,2],0);k

## [1] 11

```

```
plot(R2~clusters,main=paste("Clusters= ",k,sep=""));abline(v=k,col="red",lwd=3);
plot(o,add=T)
```



```
cores <- cutree(fish_tree, k=k)
coresDF<-data.frame(subdrainage=names(cores),cores=cores)

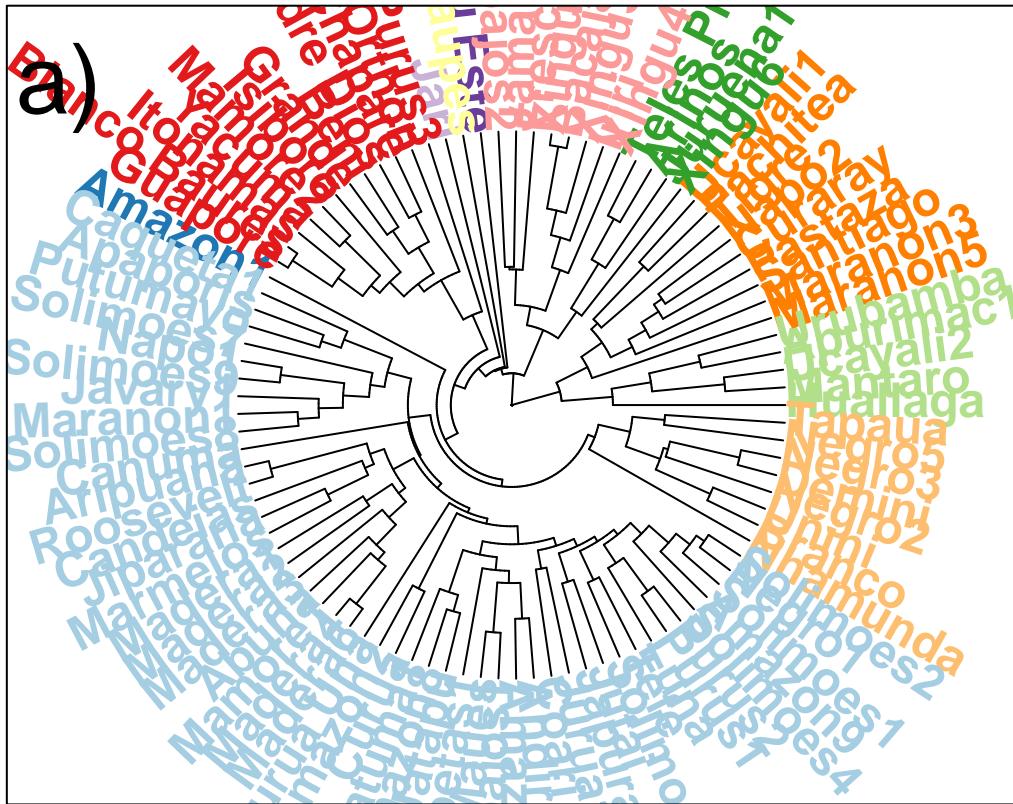
cores<-c(brewer.pal(n=k,name = "Paired"))[cores]
coresDF$cores<-cores

op<-par()
#jpeg("a_upgma_species_compositionBsim.jpeg",width = 3800,height = 1100,res = 200)
# layout(mat = c(1,2))
#par(mar=c(1,1,1,1),mfrow=c(1,3))
#plot(fish_tree,h=-1,las=1,col.label=cores,ylab="Fish dissimilarity (Bsim)");box();grid()
#rect.hclust(fish_tree,h=0.5,border = rainbow(11))
#abline(h=0.6,lwd=5,col="red")
#title("a) Species composition")
plot(as.phylo(fish_tree),type="f",las=1,cex=1.5,
```

```

tip.col=cores,font=c(2));box();#grid()
mtext("a)", adj=0.01, line=-3, cex=3)

```



```

# library(GISTools)
library(RColorBrewer)
library(sf)

```

```

## Linking to GEOS 3.12.1, GDAL 3.8.4, PROJ 9.3.1; sf_use_s2() is TRUE

```

```

amazon=st_read(paste("StatistiquesSubBasinAmazon_02052016/ShapefileBasin/",
                     sep=""),
                layer = "BasinNivel2_v052016")

```

```

## Reading layer 'BasinNivel2_v052016' from data source
##   '/home/ird/Insync/msdias@unb.br/OneDrive Biz/Amazon/AmazonFish_analises_github/AmazonBetaDivers'
##   using driver 'ESRI Shapefile'
## Simple feature collection with 97 features and 19 fields

```

```

## Geometry type: MULTIPOLYGON
## Dimension: XY
## Bounding box: xmin: -80 ymin: -20 xmax: -50 ymax: 5.3
## Geodetic CRS: WGS 84

```

```

#g60<-cutree(tree = fish_tree,h = 0.5)+1
g60 <- data.frame(basin = as.phylo(fish_tree)$tip.label, Cores=cores)
rownames(g60)<-g60$basin
g60<-g60[amazon$BvNiv2,]
table(g60$basin==amazon$BvNiv2)

```

```

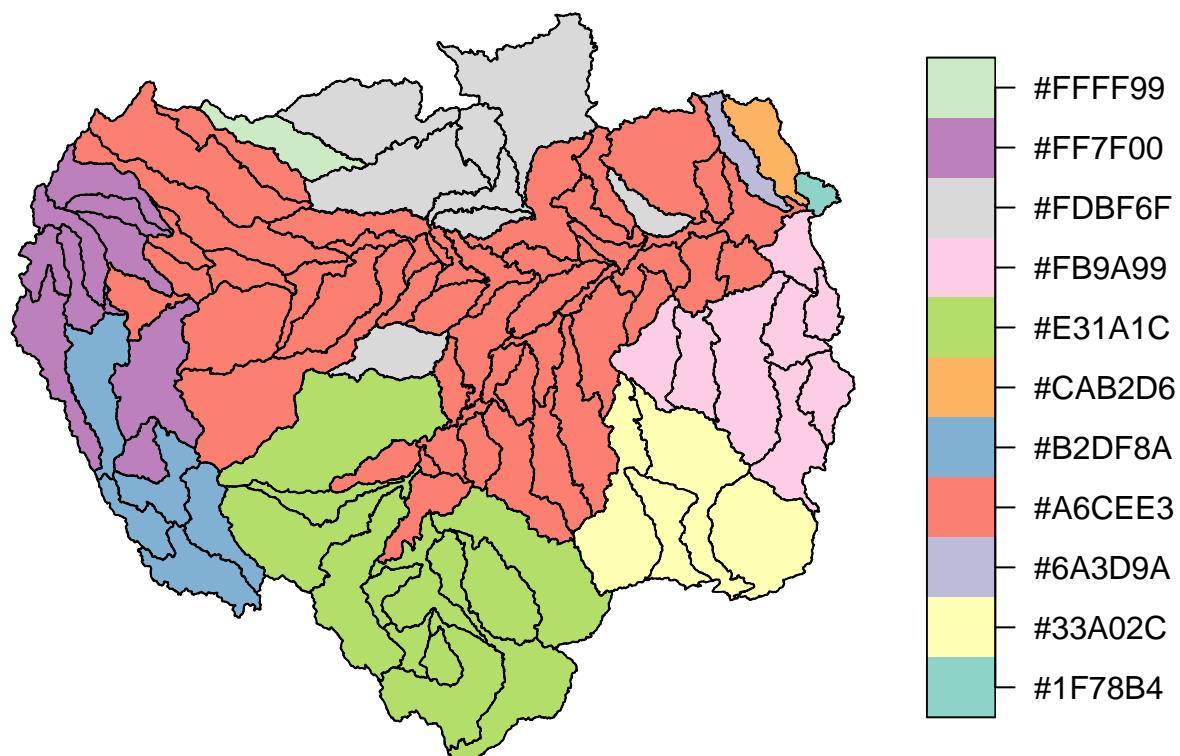
## 
## TRUE
## 97

```

```
amazon$Cores<-g60$Cores
```

```
plot(amazon["Cores"])
```

Cores



```

summary(as.vector(D_fish$beta.sim))

##      Min. 1st Qu. Median    Mean 3rd Qu.    Max.
## 0.0266  0.3786  0.5083  0.5144  0.6408  1.0000

sd(as.vector(D_fish$beta.sim))

## [1] 0.18

#
library(vegan)
set.seed(55444)
NMDS <- metaMDS((D_fish$beta.sim), k = 2, #try = c(1,10),
                  #trymax = 200,
                  autotransform = F, noshare = F,
                  previous.best = T, parallel=4)

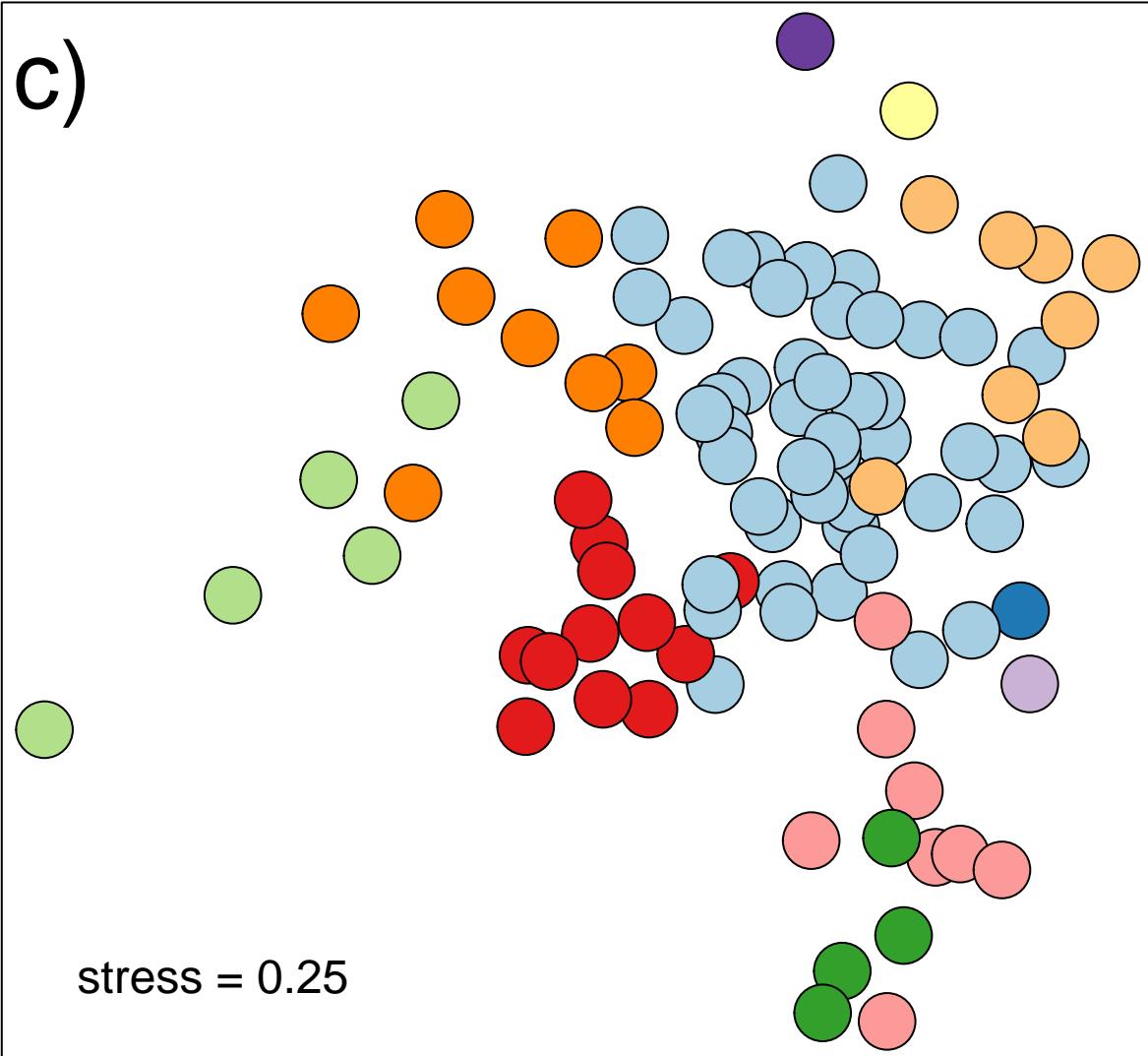
## Run 0 stress 0.88
## Run 1 stress 0.25
## ... New best solution
## ... Procrustes: rmse 0.1 max resid 0.58
## Run 2 stress 0.25
## ... New best solution
## ... Procrustes: rmse 0.041 max resid 0.19
## Run 3 stress 0.25
## ... Procrustes: rmse 0.035 max resid 0.17
## Run 4 stress 0.26
## Run 5 stress 0.27
## Run 6 stress 0.27
## Run 7 stress 0.26
## Run 8 stress 0.26
## Run 9 stress 0.25
## ... New best solution
## ... Procrustes: rmse 0.035 max resid 0.16
## Run 10 stress 0.25
## ... New best solution
## ... Procrustes: rmse 0.04 max resid 0.21
## Run 11 stress 0.26
## Run 12 stress 0.26
## Run 13 stress 0.25
## Run 14 stress 0.27
## Run 15 stress 0.26
## Run 16 stress 0.26
## Run 17 stress 0.25
## Run 18 stress 0.27
## Run 19 stress 0.25
## Run 20 stress 0.27
## *** Best solution was not repeated -- monoMDS stopping criteria:
##      3: no. of iterations >= maxit
##      16: stress ratio > sratmax
##      1: scale factor of the gradient < sfgrmin

```

```

par(mar=c(1,1,1,1))
plot(scores(NMDS,display="sites"),las=1,col="black",
      pch=21,cex=4,bg=coresDF$cores,axes=F);box()
text(x = -0.5,y = -0.3,paste0("stress = ",round(NMDS$stress,2)),cex=1.5)
mtext("c)", adj=0.01, line=-3, cex=3)

```



```
#dev.off()
```

```
par(op)
```

```

## Warning in par(op): graphical parameter "cin" cannot be set

## Warning in par(op): graphical parameter "cra" cannot be set

## Warning in par(op): graphical parameter "csi" cannot be set

## Warning in par(op): graphical parameter "cxy" cannot be set

```

```

## Warning in par(op): graphical parameter "din" cannot be set

## Warning in par(op): graphical parameter "page" cannot be set

##cutting tree
library(cluster)
groups <- cutree(fish_tree,h = 0.4)
groups <- data.frame(basin = names(groups), groups)
prin_canal<-groups[groups$basin=="Amazon5",]$groups
prin_canal <- as.vector(groups[groups$groups%in%prin_canal,]$basin)
# plot(scores(NMDS), las=1, type="n")
# text(scores(NMDS), rownames(scores(NMDS)), cex=0.8, col=groups$groups)

#Mean distance from each group to the main channel
library(plyr)
groups <- ddply(.data = groups,.variables = 'groups',.fun = function(x){
  # bas<-c("Paru_Este", "Tapajos2")
  # bas<-c("Jari")
  bas<-as.vector(x$basin)
  # bas<-as.vector(groups[groups$groups==6,]$basin)
  if( any(prin_canal%in%bas) ) {
    x$Bsim_m <- ( rep(0,length(bas)) )
    return(x)
  } else {
    Bsim<-as.matrix(betapart::beta.pair(fish_mat[c(prin_canal,bas),],
                                           index.family = "sorensen")$beta.sim)
    resu<-mean(Bsim[prin_canal,bas])
    x$Bsim_m <- ( rep(resu,length(bas)) )
    return(x)
  }
})
}

#merging fish data base
amazon
```

```

## Simple feature collection with 97 features and 20 fields
## Geometry type: MULTIPOLYGON
## Dimension:      XY
## Bounding box:  xmin: -80 ymin: -20 xmax: -50 ymax: 5.3
## Geodetic CRS:  WGS 84
## First 10 features:
##      Basin   BvNiv2 SupBv1 SupBv2 WaterColor Area_km Chut_hydro Chut_grand
## 1     Jari     Jari  58206  58206     Clear   58206        12         0
## 2     Curua   Curua  25290  25290     Clear   25290        0         0
## 3 Paru_Este Paru_Este  39288  39288     Clear   39288        2         0
## 4     Curua   Una_Curua  31115  31115    White   31115        0         1
## 5      Tefe     Tefe  24269  24269    Black   24269        0         0
## 6     Jutai   Jutai  78109  78109    Black   78109        0         0
## 7 Putumayo Putumayo 120579 120579    White  120579        1         0
```

```

## 8     Coari     Coari 35587 35587    Black   35587      0      0
## 9     Nhamunda  Nhamunda 28609 28609    Clear   28609      0      0
## 10    Uatumá    Jatapu 67919 34493    Clear   34493      2      0
## NetworkDen Elevatio_2 NbEspeceSi NbEspeceEn NbSitesTot Endemism DistanceEm
## 1       0.15      259       68      145      34      3     154
## 2       0.13      309       43      543      16      0     696
## 3       0.14      314       2      244      1       0     229
## 4       0.12      171       67      655      45      0     505
## 5       0.19      87       433      966      80      3     2291
## 6       0.18     123       171      976      46      0     2688
## 7       0.12     237       517      973     108      2     2908
## 8       0.17      75       310      933      35      0     2055
## 9       0.17     106       203      813      66      0     859
## 10      0.10     156       132      840      20      0     1265
## X23ma_cd      X23Ma X10Ma_cd      X10Ma
## 1       H AffluentsNord      K2 AffluentsNord
## 2       H AffluentsNord      K2 AffluentsNord
## 3       H AffluentsNord      K2 AffluentsNord
## 4       B GroupAtlantik      C2 GroupAtlantik4
## 5       F PebasSudEst3      B2 AcreSystem
## 6       F PebasSudEst3      B2 AcreSystem
## 7       G PebasSystem      D2 AtlantikNord3
## 8       F PebasSudEst3      B2 AcreSystem
## 9       H AffluentsNord      K2 AffluentsNord
## 10      H AffluentsNord      K2 AffluentsNord
##                               geometry   Cores
## 1 MULTIPOLYGON (((-52 -1.1, -... #CAB2D6
## 2 MULTIPOLYGON (((-55 -1.9, -... #A6CEE3
## 3 MULTIPOLYGON (((-53 -1.5, -... #6A3D9A
## 4 MULTIPOLYGON (((-55 -2.6, -... #A6CEE3
## 5 MULTIPOLYGON (((-66 -4.6, -... #A6CEE3
## 6 MULTIPOLYGON (((-70 -6.6, -... #A6CEE3
## 7 MULTIPOLYGON (((-77 0.92, -... #A6CEE3
## 8 MULTIPOLYGON (((-63 -3.9, -... #A6CEE3
## 9 MULTIPOLYGON (((-59 -0.013,... #FDBF6F
## 10 MULTIPOLYGON (((-59 1.5, -5... #A6CEE3

```

```

rownames(groups) <- groups$basin
groups <- groups[amazon$BvNiv2,]
table(groups$basin == amazon$BvNiv2)

```

```

##
## TRUE
## 97

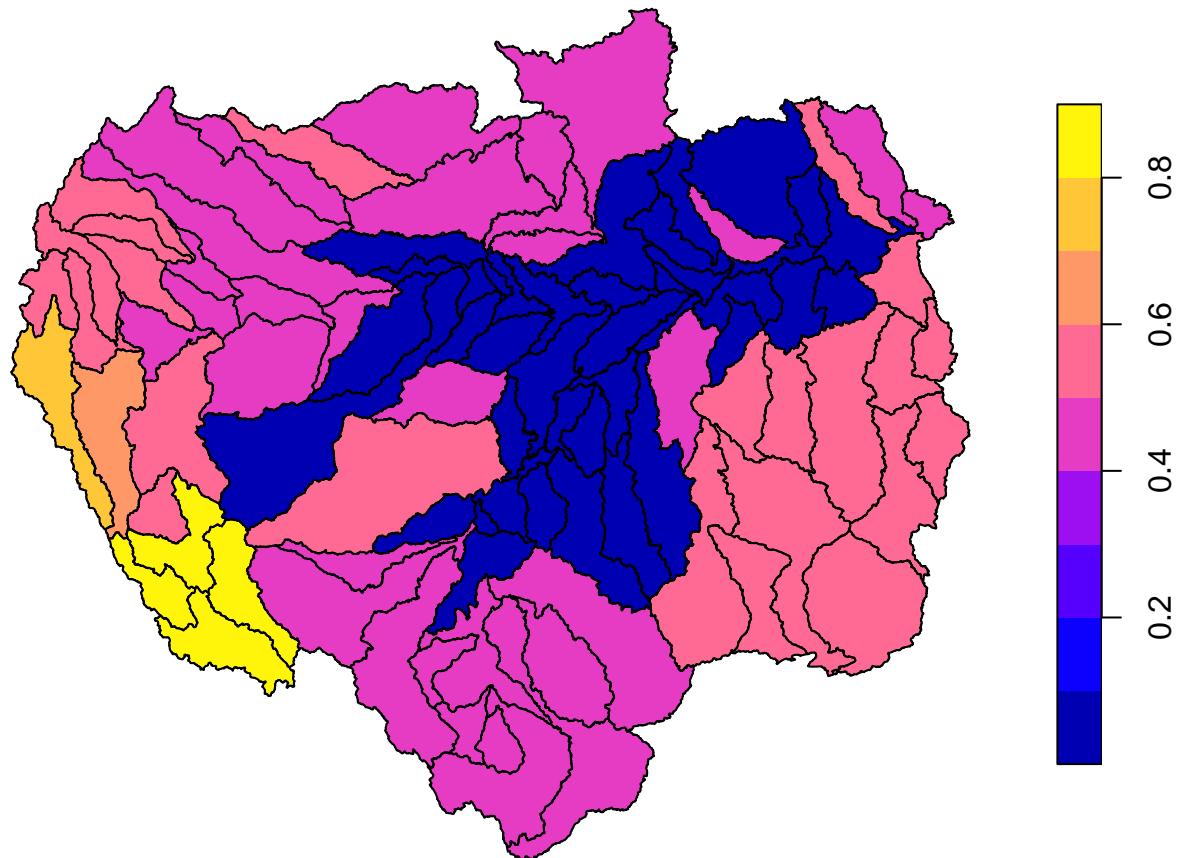
```

```

amazon$Bsim_m <- groups$Bsim_m
plot(amazon["Bsim_m"])

```

Bsim_m



```
par(op)
```

```
## Warning in par(op): graphical parameter "cin" cannot be set
```

```
## Warning in par(op): graphical parameter "cra" cannot be set
```

```
## Warning in par(op): graphical parameter "csi" cannot be set
```

```
## Warning in par(op): graphical parameter "cxy" cannot be set
```

```
## Warning in par(op): graphical parameter "din" cannot be set
```

```
## Warning in par(op): graphical parameter "page" cannot be set
```

```
rm(op,k,groups,fish_tree,g60)
```

1.3 Including PhyloDissimilarity

I also calculated the Bsim component based on phylogenetic beta diversity from *betapart* package. Rabosky et al. published a fish supertree containing genetical of ~8000 sp and included all the other freshwater and marine species (~23000) to this supertree based on the taxonomy. They provided a supertree file (see <https://fishtreeoflife.org/downloads/>) containing 100 posterior draws from the final super tree. I select each of these 100 trees, calculated the Bsim with the AmazonFish data with each tree and then computed mean estimates of the 100 draws. Here is the mean Bsim phylobeta from all the 100 draws

```
PBsim <- read.csv(paste0(getwd(), "/BetaPhylo_meanvalues_Genet_Taxo/PSIMmean_over100posteriorSampling"))
sep = ",", header = TRUE, row.names = 1)
PBsim[1:5, 1:5]

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8
## Abuna     0.00   0.409   0.148   0.115   0.14
## Amazon1   0.41   0.000   0.068   0.065   0.10
## Amazon2   0.15   0.068   0.000   0.107   0.18
## Amazon5   0.11   0.065   0.107   0.000   0.13
## Amazon8   0.14   0.101   0.180   0.134   0.00

PBsim <- as.dist(PBsim)

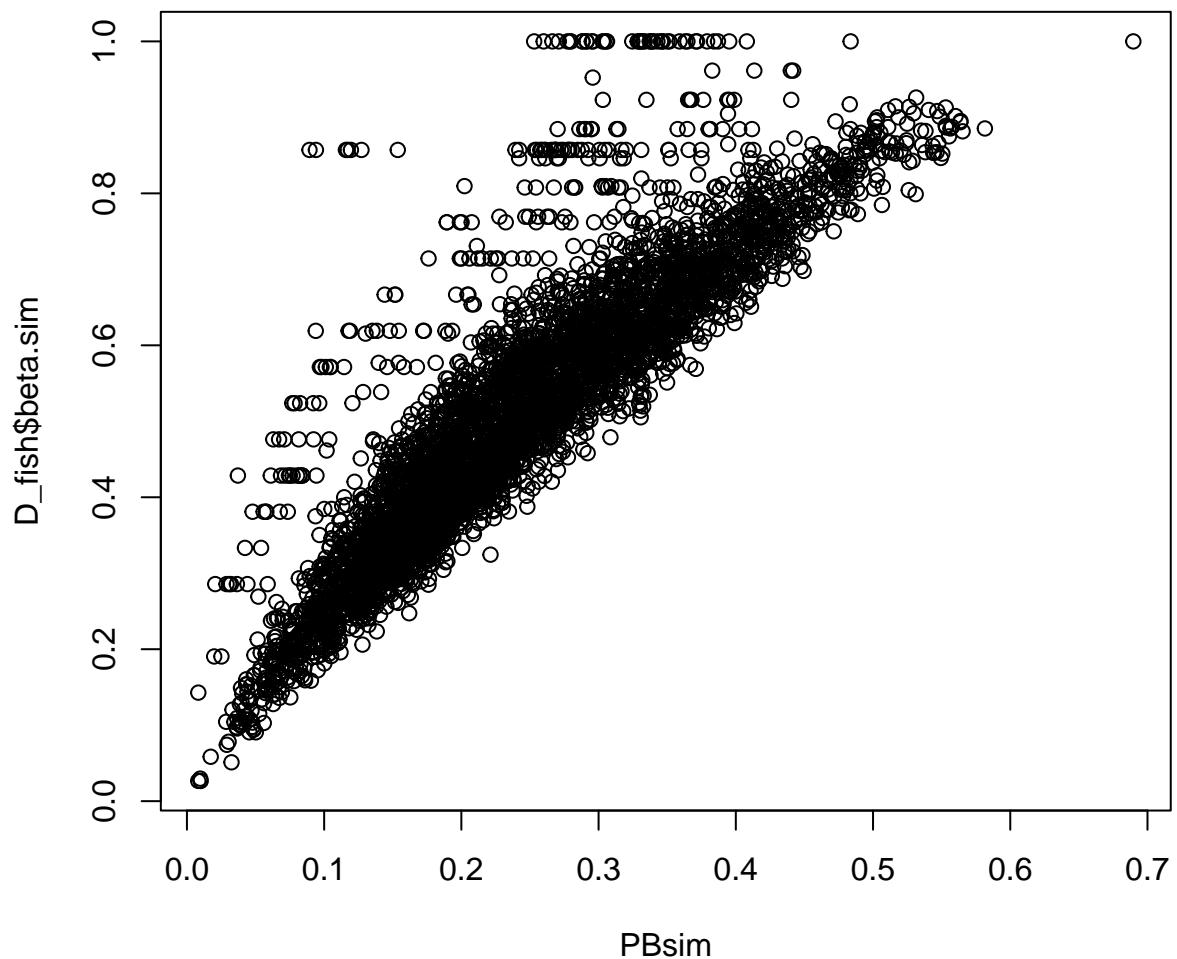
# plot(metaMDS(PBsim, k=2, autotransform = FALSE))
summary(as.vector(PBsim))

##      Min. 1st Qu. Median      Mean 3rd Qu.      Max.
## 0.00824 0.16441 0.23345 0.24394 0.31561 0.68970

sd(as.vector(PBsim))

## [1] 0.11

plot(D_fish$beta.sim ~ PBsim)
```



```
cor.test(D_fish$beta.sim, PBsim, method = "pearson")
```

```
##
## Pearson's product-moment correlation
##
## data: D_fish$beta.sim and PBsim
## t = 132, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.88 0.89
## sample estimates:
## cor
## 0.89

PBsne <- read.csv(paste0(getwd(), "/BetaPhylo_meanvalues_Genet_Taxo/PSNEmean_over100posteriorSampling"))
sep = ",", header = TRUE, row.names = 1)
PBsne[1:5, 1:5]

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8
```

```

## Abuna 0.000 0.047 0.2347 0.266 0.2405
## Amazon1 0.047 0.000 0.3237 0.347 0.3145
## Amazon2 0.235 0.324 0.0000 0.024 0.0038
## Amazon5 0.266 0.347 0.0244 0.000 0.0211
## Amazon8 0.241 0.314 0.0038 0.021 0.0000

PBsne <- as.dist(PBsne)

# plot(metaMDS(PBsne, k=2, autotransform = FALSE))
summary(as.vector(PBsne))

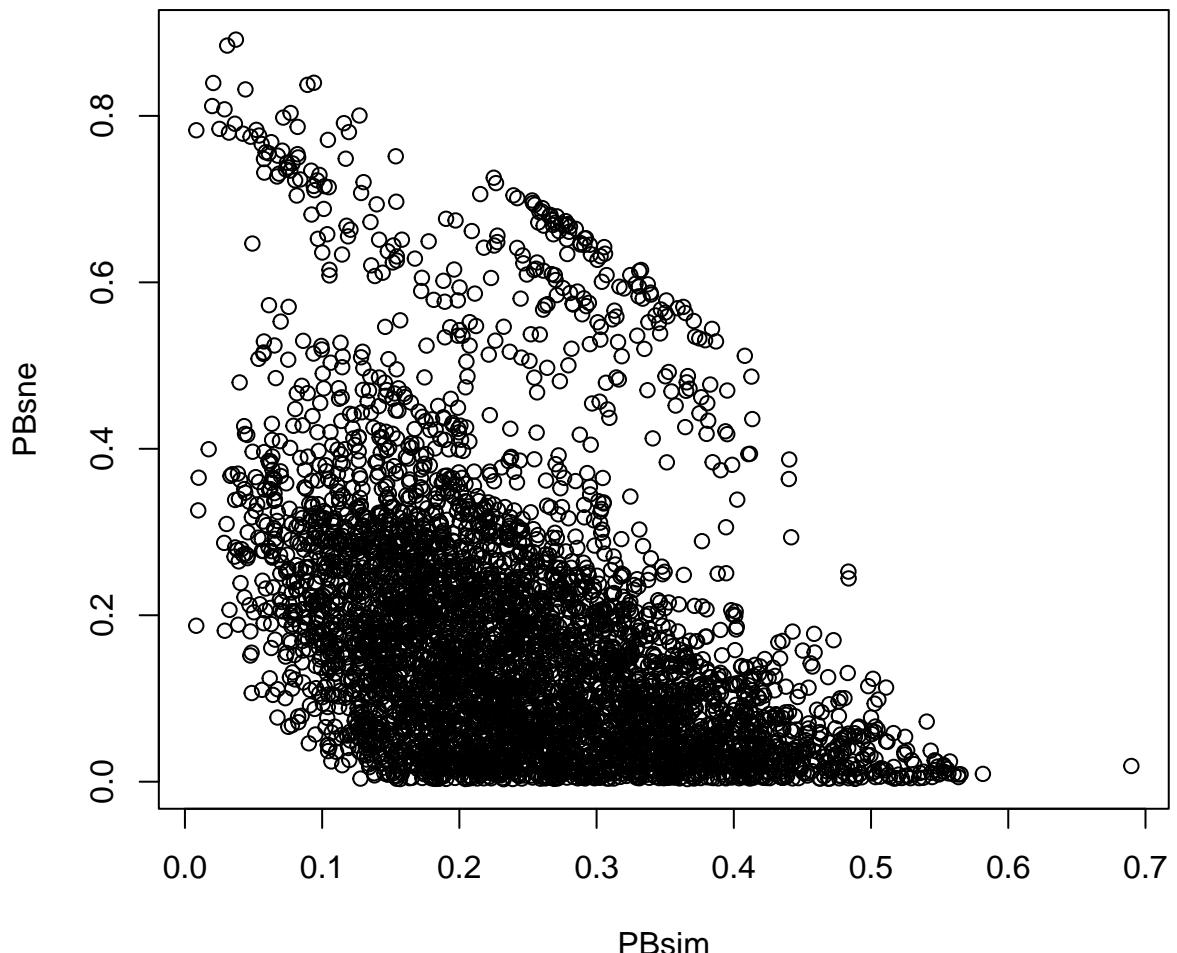
##      Min. 1st Qu. Median    Mean 3rd Qu.    Max.
## 0.00303 0.05613 0.13239 0.17246 0.24032 0.89164

sd(as.vector(PBsne))

## [1] 0.16

plot(PBsne ~ PBsim)

```



```

cor.test(PBsne, PBsim, method = "pearson")

##
## Pearson's product-moment correlation
##
## data: PBsne and PBsim
## t = -29, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.42 -0.37
## sample estimates:
## cor
## -0.39

```

I also calculated the Psim by using the ~635sp with genetic data to compare

```

#####
##### #phylogenetic diversity
#####
##### #phylogeny
filo <- read.tree(file = "actinopt_12k_treePL.tre") #fish

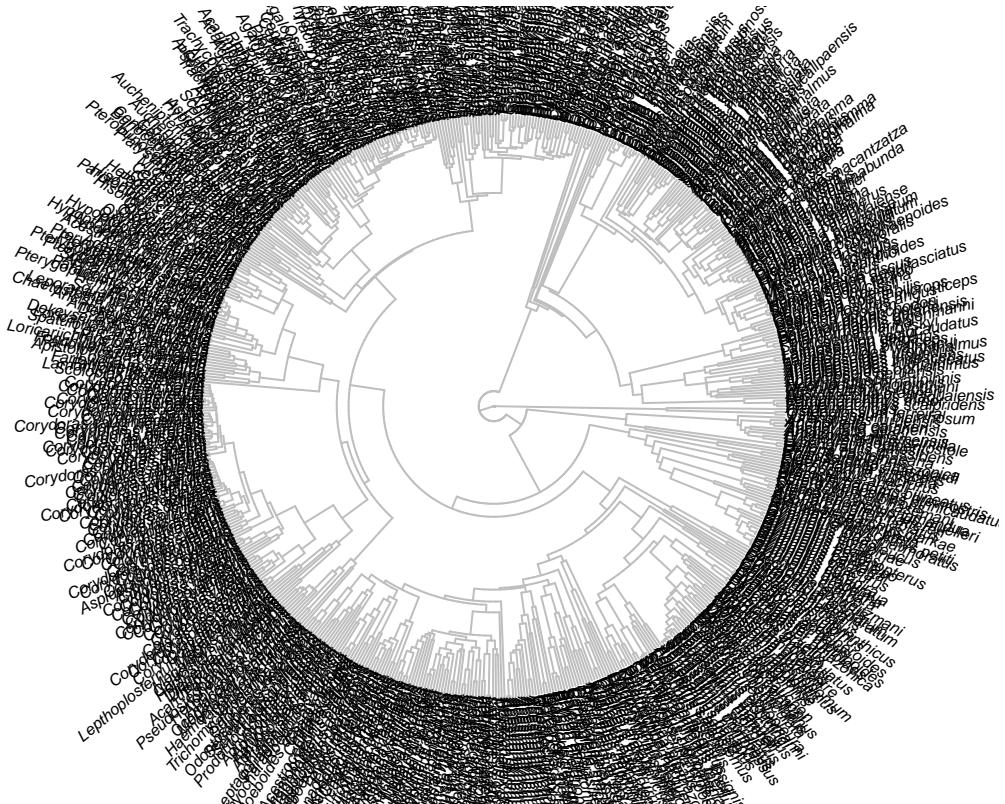
# especies presentes tanto na filogenia quanto no arquivo
# mami
sp <- intersect(filo$tip.label, colnames(fish_mat))
length(sp)

## [1] 635

# excluindo as especies do mapa para reter somente as
# especies presentes na filogenia
fish_mat_philo <- droplevels(fish_mat[, colnames(fish_mat) %in%
  sp])

# deixar na filo as sp da matrix
library(picante)
filo_gamb <- prune.sample(samp = fish_mat_philo, phylo = filo)
plot.phylo(filo_gamb, type = "fan", edge.color = "gray", cex = 0.5)

```

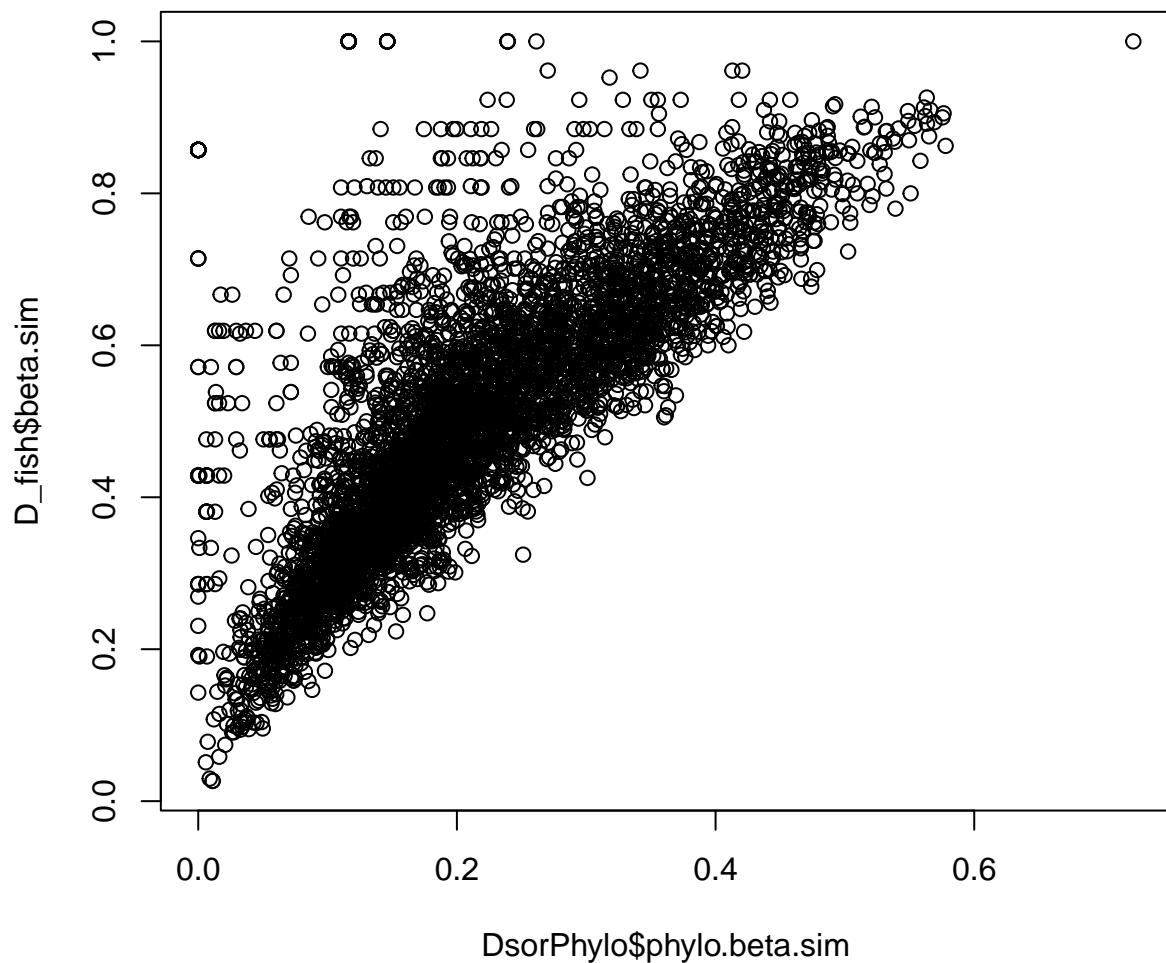


```

library(betapart)
DsorPhylo <- phylol.betta.pair(x = fish_mat_philo, tree = filo_gamb,
  index.family = "sorensen")

plot(DsorPhylo$phylo.betta.sim, D_fish$beta.sim)

```



```
cor.test(DsorPhylo$phylo.beta.sim, D_fish$beta.sim)
```

```
##
## Pearson's product-moment correlation
##
## data: DsorPhylo$phylo.beta.sim and D_fish$beta.sim
## t = 72, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.71 0.74
## sample estimates:
## cor
## 0.73
```

```
#####
#####
```

1.3.1 PBsim using Cassemiro et al tree

I also tested whether the results are consistent using the phylogeny provided by Cassemiro et al. (2023, PNAS), the link of data can be found here <https://zenodo.org/record/6672927>.

```

library(ape)
tr = read.nexus(paste0("BetaPhylo_meanvalues_Genet_Taxo/", "Cassemiro et al 2023PNAS/",
"3167_Neotrop_Freshwater_Fish_spp_time_tree.tre"))

# plot(tr,cex=0.1)
tr$tip.label[1:30]

## [1] "Carcharhinidae_Carcharhinus_leucas"
## [2] "Pristidae_Pristis_pristis"
## [3] "Potamotrygonidae_Paratrygon_aiereba"
## [4] "Potamotrygonidae_Heliotrygon_gomesi"
## [5] "Potamotrygonidae_Potamotrygon_tatianae"
## [6] "Potamotrygonidae_Potamotrygon_motoro"
## [7] "Potamotrygonidae_Potamotrygon_tigrina"
## [8] "Potamotrygonidae_Potamotrygon_falkneri"
## [9] "Potamotrygonidae_Potamotrygon_constellata"
## [10] "Potamotrygonidae_Potamotrygon_orbignyi"
## [11] "Potamotrygonidae_Plesiotrygon_nana"
## [12] "Potamotrygonidae_Plesiotrygon_iwamae"
## [13] "Potamotrygonidae_Potamotrygon_yepezi"
## [14] "Potamotrygonidae_Potamotrygon_magdalena"
## [15] "Potamotrygonidae_Potamotrygon_histrix"
## [16] "Lepidosirenidae_Lepidosiren_paradoxa"
## [17] "Lepisosteidae_Atractosteus_tristoechus"
## [18] "Anguillidae_Anguilla_rostrata"
## [19] "Megalopidae_Megalops_atlanticus"
## [20] "Polycentridae_Polycentrus_schomburgkii"
## [21] "Polycentridae_Monocirrhus_polyacanthus"
## [22] "Mugilidae_Joturus_pichardi"
## [23] "Mugilidae_Mugil_curema"
## [24] "Mugilidae_Mugil_liza"
## [25] "Cichlidae_Geophagus_crassilabris"
## [26] "Cichlidae_Geophagus_pellegrini"
## [27] "Cichlidae_Geophagus_steindachneri"
## [28] "Cichlidae_Gymnogeophagus_taroba"
## [29] "Cichlidae_Gymnogeophagus_setequetas"
## [30] "Cichlidae_Gymnogeophagus_rhabdotus"

tr$tip.label <- unlist(lapply(strsplit(tr$tip.label, split = "_"),
  function(x) {
    paste0(x[2:3], collapse = "_")
  }))
tr$tip.label[1:30]

## [1] "Carcharhinus_leucas"      "Pristis_pristis"

```

```

## [3] "Paratrygon_aiereba"      "Heliotrygon_gomesi"
## [5] "Potamotrygon_tatianae"   "Potamotrygon_motoro"
## [7] "Potamotrygon_tigrina"    "Potamotrygon_falkneri"
## [9] "Potamotrygon_constellata" "Potamotrygon_orbignyi"
## [11] "Plesiotrygon_nana"       "Plesiotrygon_iwamae"
## [13] "Potamotrygon_yepezi"     "Potamotrygon_magdalena"
## [15] "Potamotrygon_histrix"    "Lepidosiren_paradoxa"
## [17] "Atractosteus_tristoechus" "Anguilla_rostrata"
## [19] "Megalops_atlanticus"     "Polycentrus_schomburgkii"
## [21] "Monocirrhus_polyacanthus" "Joturus_pichardi"
## [23] "Mugil_curema"           "Mugil_liza"
## [25] "Geophagus_crassilabris"  "Geophagus_pellegrini"
## [27] "Geophagus_steindachneri" "Gymnogeophagus_taroba"
## [29] "Gymnogeophagus_setequedas" "Gymnogeophagus_rhabdotus"

```

```

filo <- tr

# especies presentes tanto na filogenia quanto no arquivo
# mami
sp <- intersect(filo$tip.label, colnames(fish_mat))
length(sp)

```

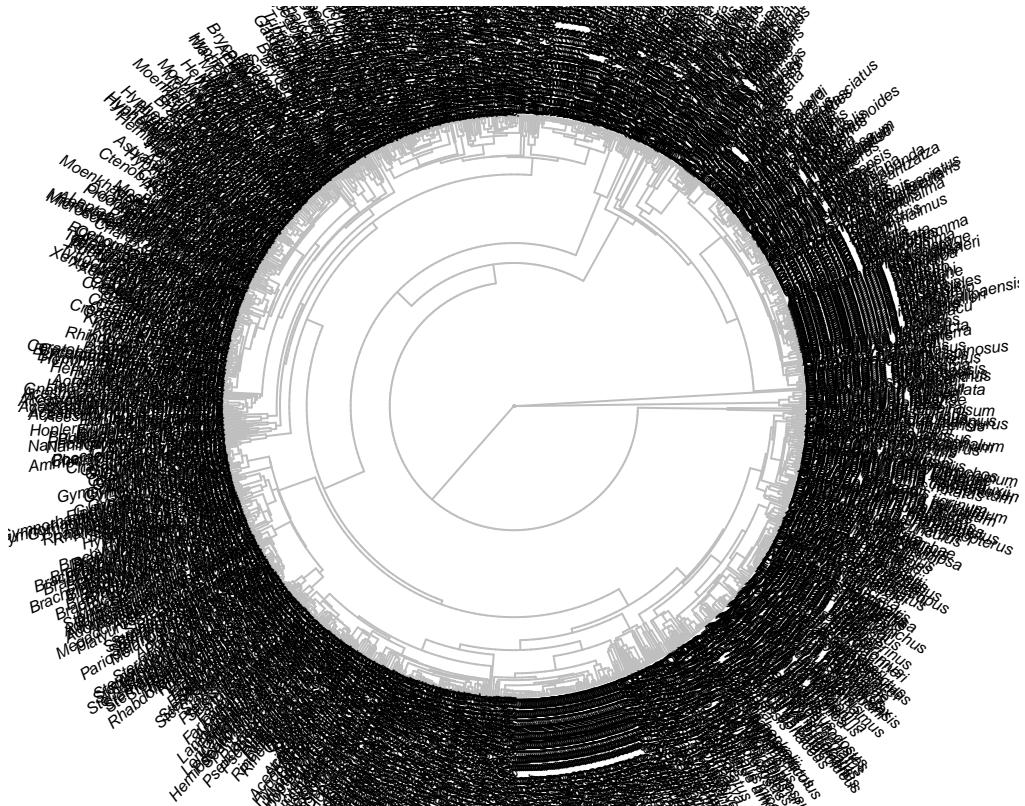
```

## [1] 1111

# excluindo as especies do mapa para reter somente as
# especies presentes na filogenia
fish_mat_philo <- droplevels(fish_mat[, colnames(fish_mat) %in%
                                         sp])

# deixar na filo as sp da matrix
library(picante)
filo_gamb <- prune.sample(samp = fish_mat_philo, phylo = filo)
plot.phylo(filo_gamb, type = "fan", edge.color = "gray", cex = 0.5)

```



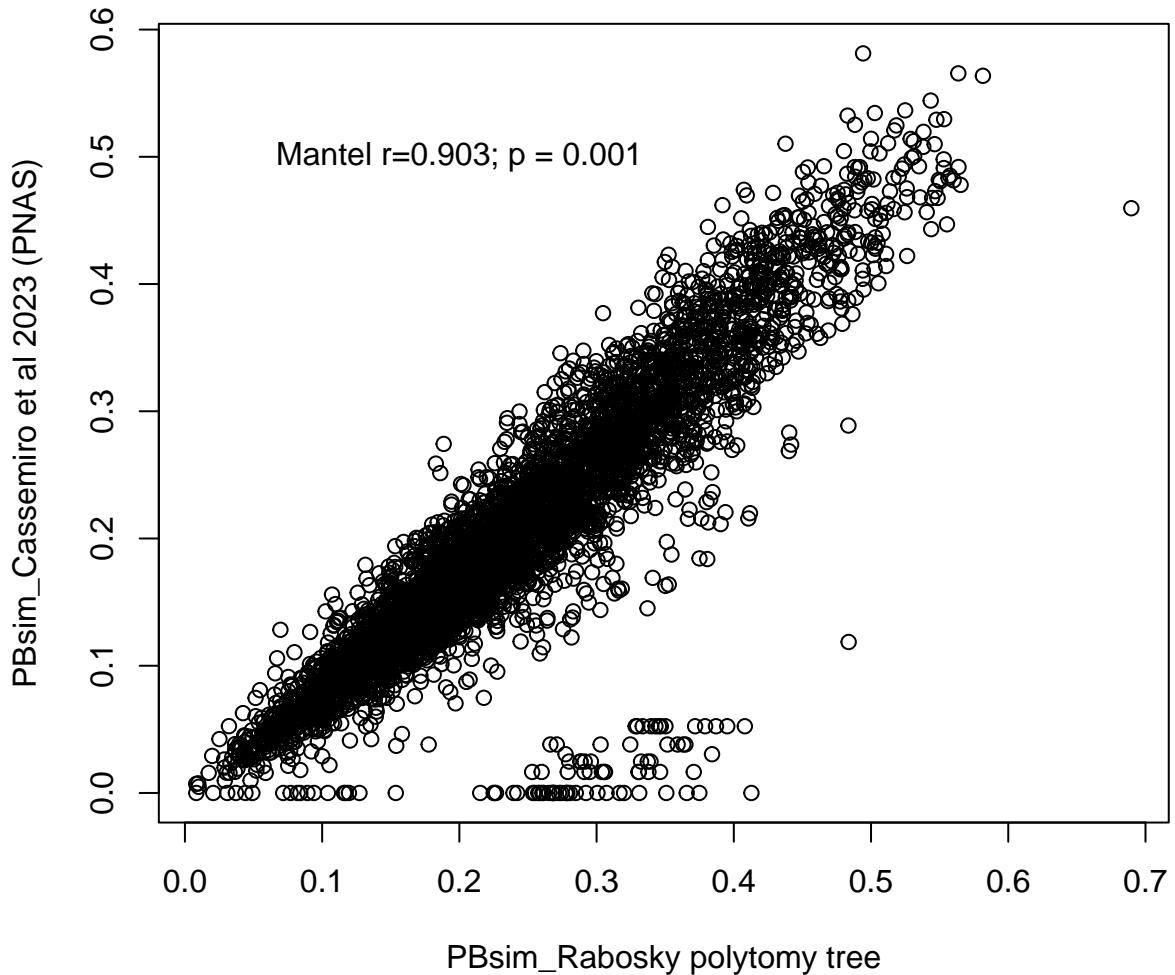
```

library(betapart)
DsorPhylo <- phylo.betta.pair(x = fish_mat_philo, tree = filo_gamb,
  index.family = "sorensen")

PBsimCasse <- DsorPhylo$phylo.betta.sim

#
library(vegan)
plot(PBsimCasse ~ PBsim, ylab = "PBsim_Cassemiro et al 2023 (PNAS)",
  xlab = "PBsim_Rabosky polytomy tree")
res <- vegan::mantel(PBsimCasse, PBsim)
text(0.2, 0.5, paste0("Mantel r=", round(res$statistic, 3), "; ",
  "p = ", round(res$signif, 3)))

```



1.3.2 Calculating Dpw

This is a phylogenetic index between communities that is much more related to the basal origins of species. The index is equivalent to MPD but it is calculated as a distance between communities. We are applying this index to the full fish tree (genetic data + polytomies) as it is rather robust to such peculiarities.

```
nullPhyloMetric<-function(mat,phylo,nreps){
  library(mcreplicate)
  cophe_phylo<-cophenetic(phylo)
  obs<-comdist(mat, cophe_phylo, abundance.weighted=FALSE)
  simu<-mc_replicate(nreps,{
    mat_aleat<-randomizeMatrix(mat, null.model="richness")
    resu<-comdist(mat_aleat, cophe_phylo, abundance.weighted=FALSE)
    resu
  },mc.cores=4)
  stand<-function(x){(x-mean(x))/sd(x)}

  obs_simu<-cbind(obs,simu)
```

```

standardized_Metric<-t(apply(obs_simu,MARGIN = 1,FUN = stand))[, "obs"]

obs[]<-standardized_Metric
return(obs)
}

#TREE GENETICS + POLYTOMIES
#library(PhyloMeasures)
library(picante)

Dpw_mean<-file.exists("BetaPhylo_meanvalues_Genet_Taxo/DpwMean_30nullModelControlRichness_over100trees")
# Dpw_mean<-file.exists("BetaPhylo_meanvalues_Genet_Taxo/DpwMean_30nullModelControlRichness_over100trees")
#Dpw_mean<-FALSE

if(Dpw_mean==FALSE){
  filos<-read.tree(file="BetaPhylo_meanvalues_Genet_Taxo/actinopt_full.trees")#Rabosky et al. 2020 .
  Dpw_poly<-list()

  #loop over all posterior samplings

  for(i in 1:100){
    #i=1
    #posterior samplings
    filo<-filos[[i]] #fazendo com uma arvore so

    #especies presentes tanto na filogenia quanto no arquivo mami
    sp<-intersect(filo$tip.label,colnames(fish_mat));length(sp)

    #fazendo a matriz
    mat<-fish_mat[,sp]
    #mat[1:10,1:6]
    #dim(mat)

    #deixar na filo as sp da matrix
    #library(picante)
    filo_gamb<-prune.sample(samp = mat, phylo = filo)
    #plot.phylo(filo_gamb,type="fan",edge.color="gray",cex=0.5)

    # #Dpw
    # library(PhyloMeasures)
    # Dpw_poly[[i]]<-cd.query(tree=filo_gamb,mat,standardize=FALSE)
    # rownames(Dpw_poly[[i]])<-colnames(Dpw_poly[[i]])<-rownames(mat)

    #Dpw with null model to control for richness 2022_08_24
    # system.time({
      file<-paste0("BetaPhylo_meanvalues_Genet_Taxo/temp_Dpw_mean_withoutMarine/",
                  "DpwMean_30nullModelControlRichness_over100trees_WITHOUT_MARINE","_tree_",i,".csv") #
      if(!file.exists(file)) {
        Dpw_poly[[i]]<-nullPhyloMetric(mat = mat, phylo = filo_gamb,nreps = 30)
      }
    })
  }
}

```

```

    write.table(x = as.matrix(Dpw_poly[[i]]),
                file = paste0("BetaPhylo_meanvalues_Genet_Taxo/temp_Dpw_mean_withoutMarine/", #exclude
                             "DpwMean_30nullModelControlRichness_over100trees_WITHOUT_MARINE", "_tree",
                             sep = ";")

} else {
  Dpw_poly[[i]] <- as.dist(read.table( #exclude WITHOUT_MARINE for all species
    file = paste0("BetaPhylo_meanvalues_Genet_Taxo/temp_Dpw_mean_withoutMarine/",
                  "DpwMean_30nullModelControlRichness_over100trees_WITHOUT_MARINE", "_tree",
                  sep = ";"))
}

}

#      })

print(i)
rm(mat2)
}

Dpw_poly

#mean over all Dpw values
retorno<-function(x,val){
  return(as.matrix(x)[,val])
}

Dpw_poly_mean<-sapply(rownames(as.matrix(Dpw_poly[[1]])),FUN = function(x){
  rowMeans(sapply(Dpw_poly,retorno,val=x))
})

write.table(x = Dpw_poly_mean,
            file = "BetaPhylo_meanvalues_Genet_Taxo/DpwMean_30nullModelControlRichness_over100trees",
            sep = ";")

# write.table(x = Dpw_poly_mean,
#             file = "BetaPhylo_meanvalues_Genet_Taxo/DpwMean_30nullModelControlRichness_over100trees",
#             sep = ";")

} else {
  Dpw_mean<-as.dist(read.table("BetaPhylo_meanvalues_Genet_Taxo/DpwMean_30nullModelControlRichness",
                                sep=";",header=TRUE))
  # Dpw_mean<-as.dist(read.table("BetaPhylo_meanvalues_Genet_Taxo/DpwMean_30nullModelControlRichness",
  #                               sep=";",header=TRUE))
  # Dpw_mean<-as.dist(read.table("BetaPhylo_meanvalues_Genet_Taxo/DpwMean_over100trees_withoutMarine",
  #                               sep=";",header=TRUE))
}

}

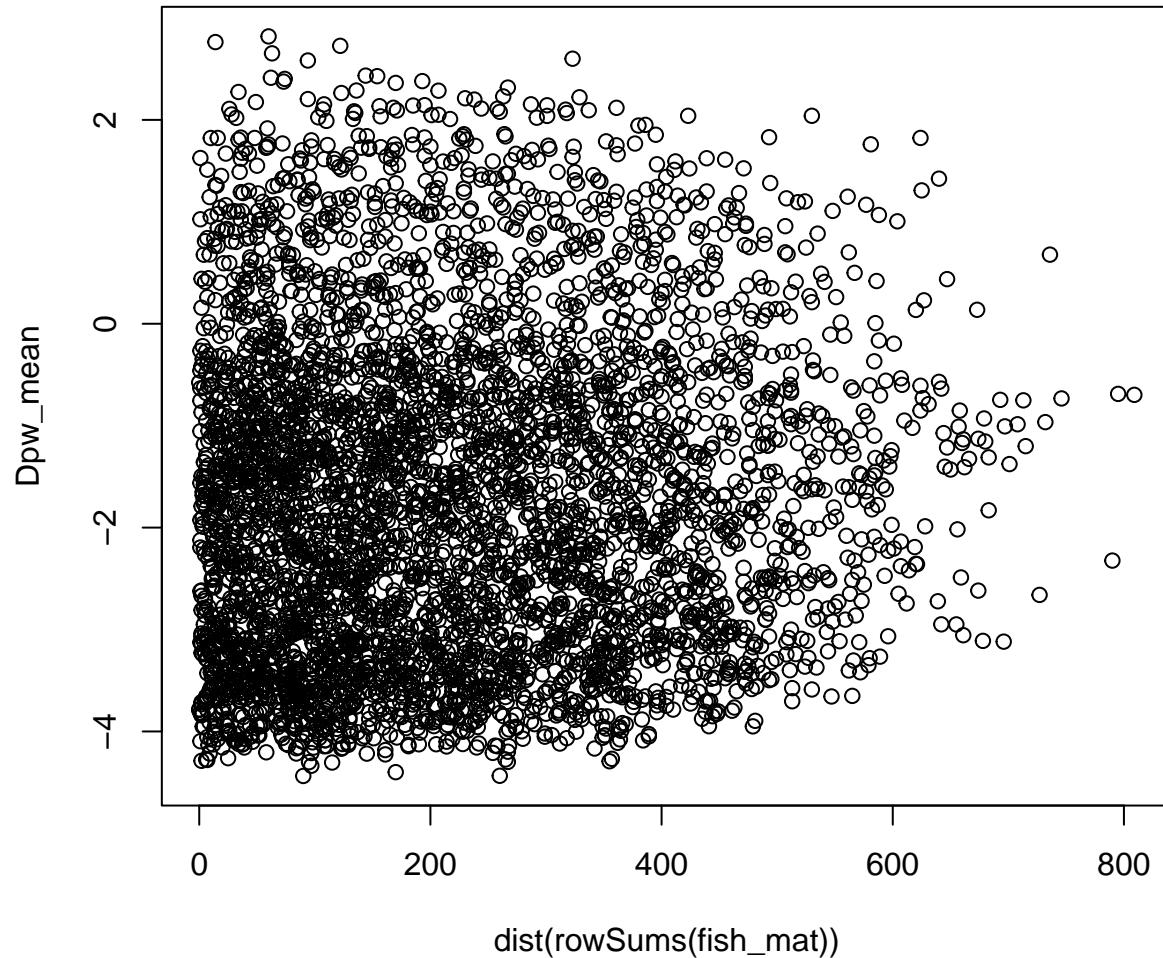
#TREE: GENETICS
dim(fish_mat_philo)

## [1] 97 1111

```

```
filo_gamb
```

```
##  
## Phylogenetic tree with 1111 tips and 1110 internal nodes.  
##  
## Tip labels:  
## Paratrygon_aiereba, Heliotrygon_gomesi, Potamotrygon_tatianae, Potamotrygon_motoro, Potamotrygo  
##  
## Rooted; includes branch length(s).  
  
# Dpw_gen<-(cd.query(tree=filo_gamb,  
#                         matrix.a = fish_mat_philo,  
#                         standardize=FALSE))  
# Dpw_gen<-nullPhyloMetric(mat = fish_mat_philo, phylo = filo_gamb,nreps = 1)  
#  
# rownames(Dpw_gen)<-rownames(fish_mat_philo)  
# colnames(Dpw_gen)<-rownames(fish_mat_philo)  
# Dpw_gen<-as.dist(Dpw_gen)  
#  
# #Dpw genetic tree vs full tree  
# plot(Dpw_mean~Dpw_gen,asp=1);abline(a = 0,b = 1)  
# cor(Dpw_mean,Dpw_gen,method = "spearman")  
plot(Dpw_mean~dist(rowSums(fish_mat)))
```

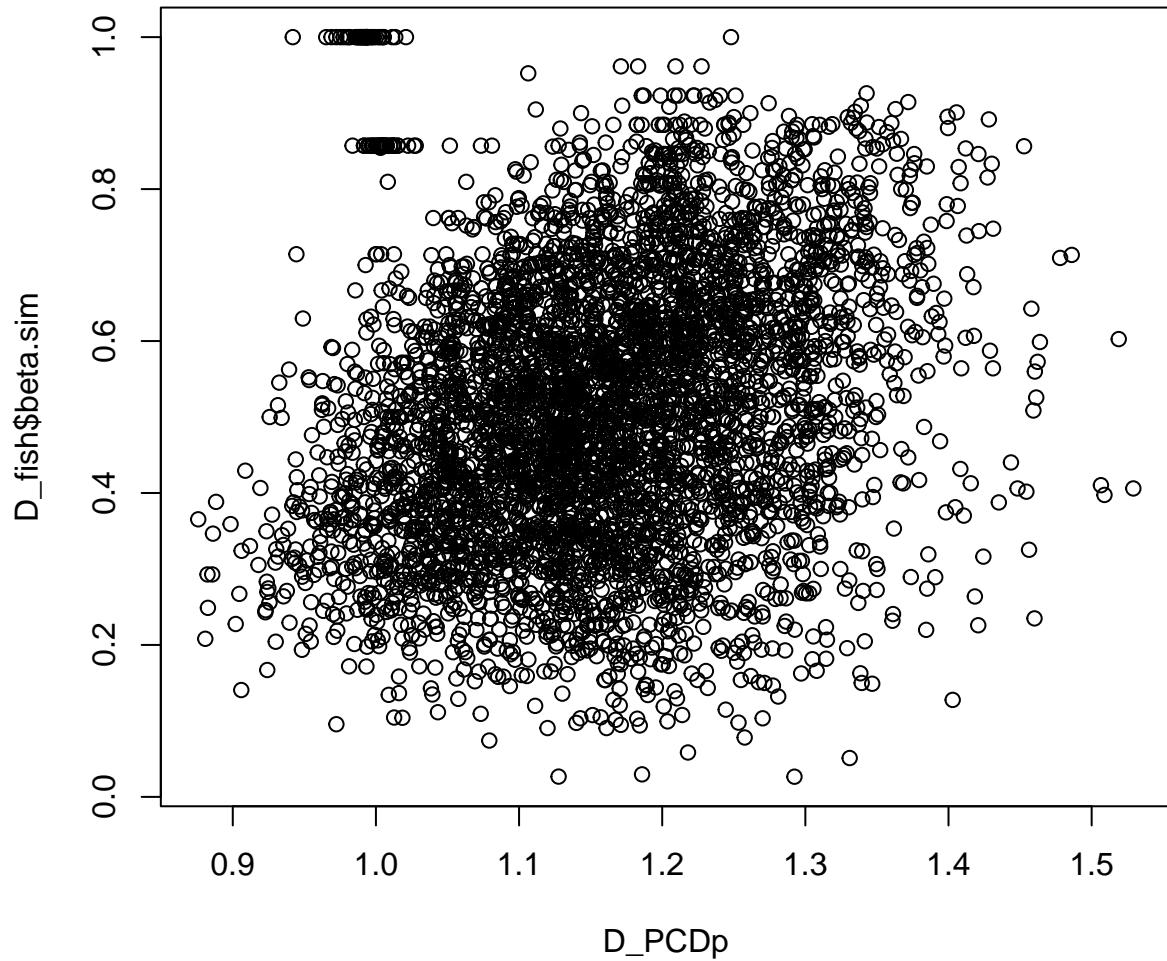


1.3.3 Calculating PCD

PCDp over 635 species

```
library(picante)
# system.time(resu<-pcd(comm = fish_mat,tree =
# filo_gamb,reps=500)) resu$PCD resu$PCDc
# plot(resu$PCDp~resu$PCDc) write.table(x =
# as.matrix(resu$PCDp),
# file='PCD_IvesHelmus/PCDp_500reps.csv', sep=';')
# write.table(x = as.matrix(resu$PCDc),
# file='PCD_IvesHelmus/PCDc_500reps.csv', sep=';')
# write.table(x = as.matrix(resu$PCD),
# file='PCD_IvesHelmus/PCD_500reps.csv', sep=';')
D_PCDp <- as.dist(read.csv("PCD_IvesHelmus/PCDp_500reps.csv",
header = T, row.names = 1, sep = ";"))

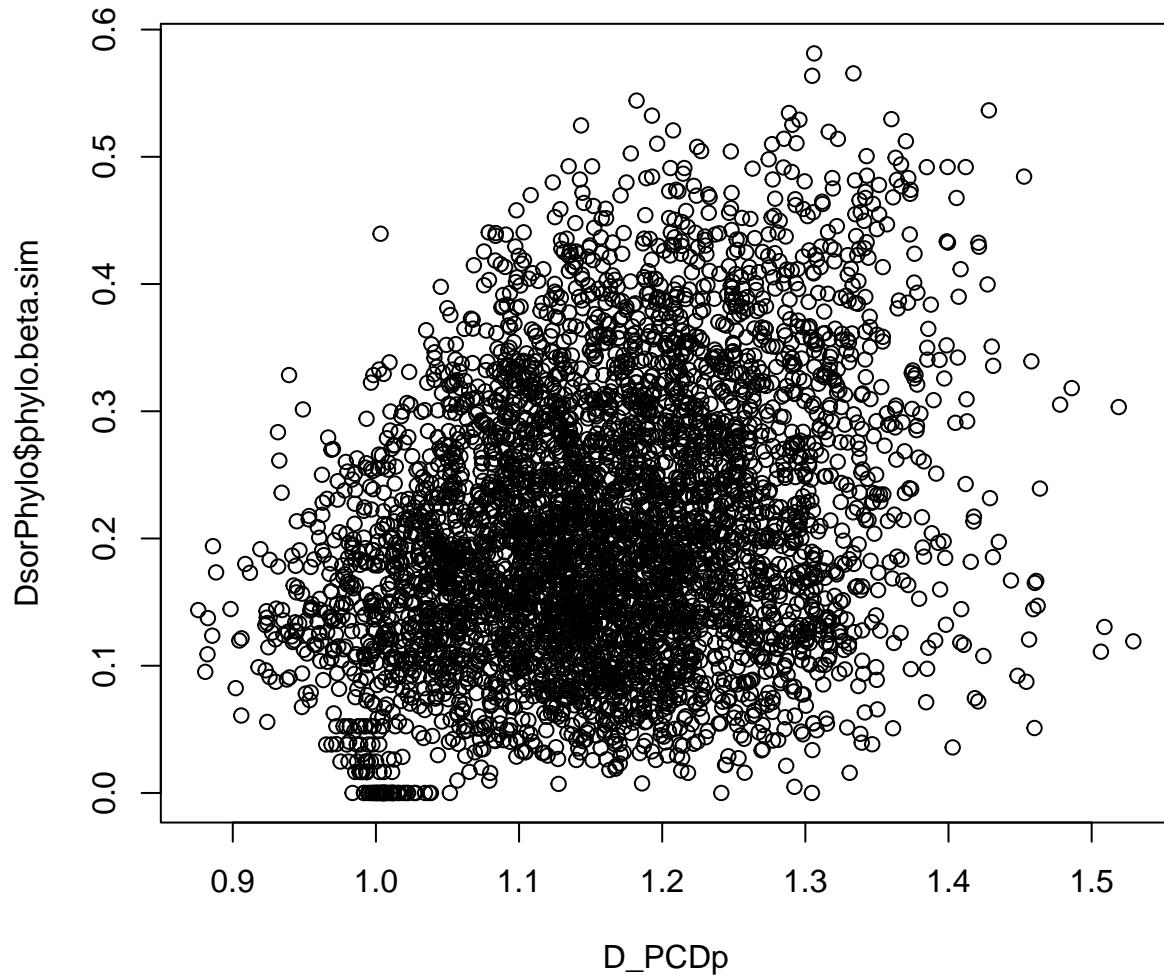
plot(D_fish$beta.sim ~ D_PCDp)
```



```
cor.test(D_fish$beta.sim, D_PCDp)
```

```
##
## Pearson's product-moment correlation
##
## data: D_fish$beta.sim and D_PCDp
## t = 16, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.20 0.26
## sample estimates:
## cor
## 0.23
```

```
plot(DsorPhylo$phylo.beta.sim ~ D_PCDp)
```



```
cor.test(DsorPhylo$phylo.beta.sim, D_PCDp)
```

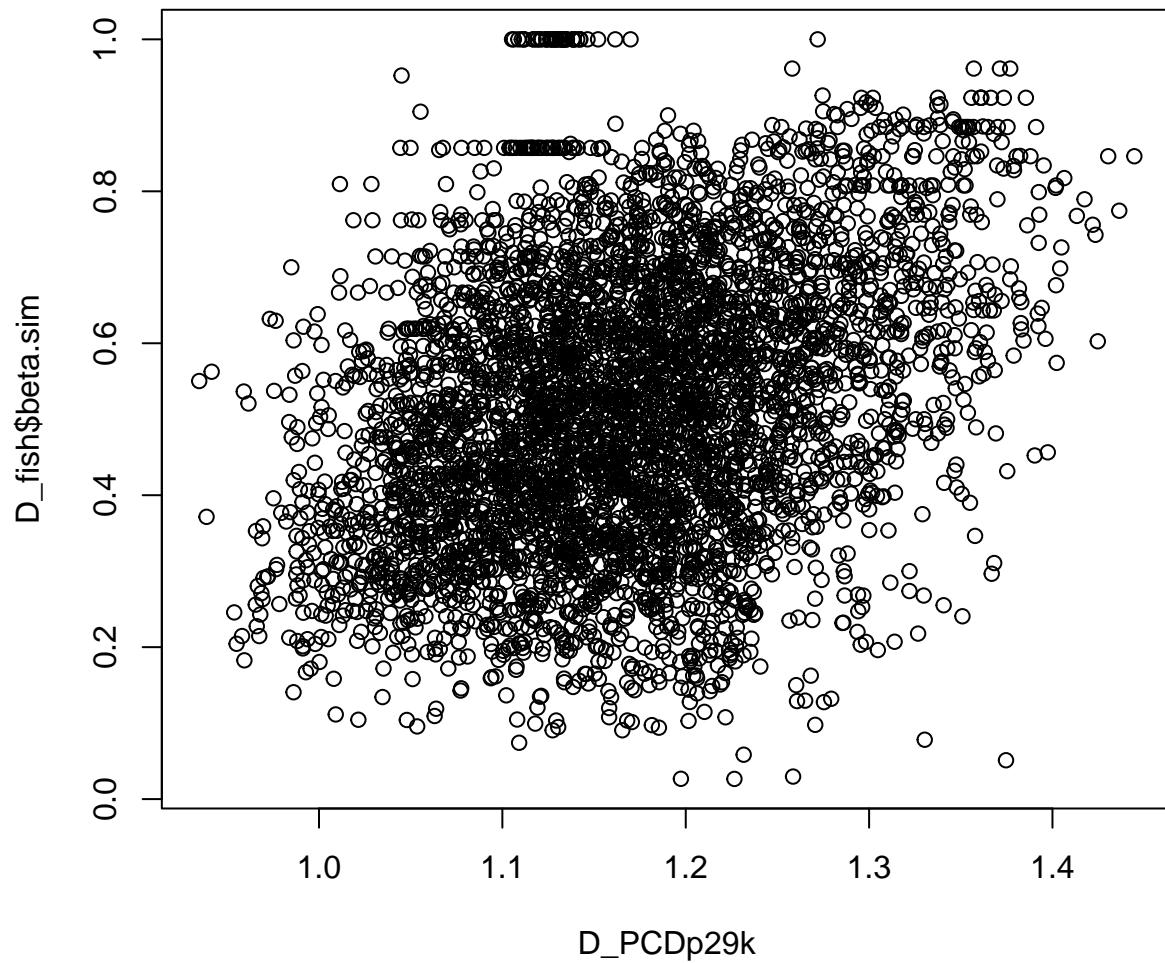
```
##
## Pearson's product-moment correlation
##
## data: DsorPhylo$phylo.beta.sim and D_PCDp
## t = 21, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.27 0.32
## sample estimates:
## cor
## 0.29
```

PCDp over ~2k species (genetics + taxonomy based polytomies)

```
library(picante)
D_PCDp29k <- as.dist(read.csv("BetaPhylo_meanvalues_Genet_Taxo/PCDpMean_over54posteriorSampling_20200120.csv"))
```

```
sep = ";"))

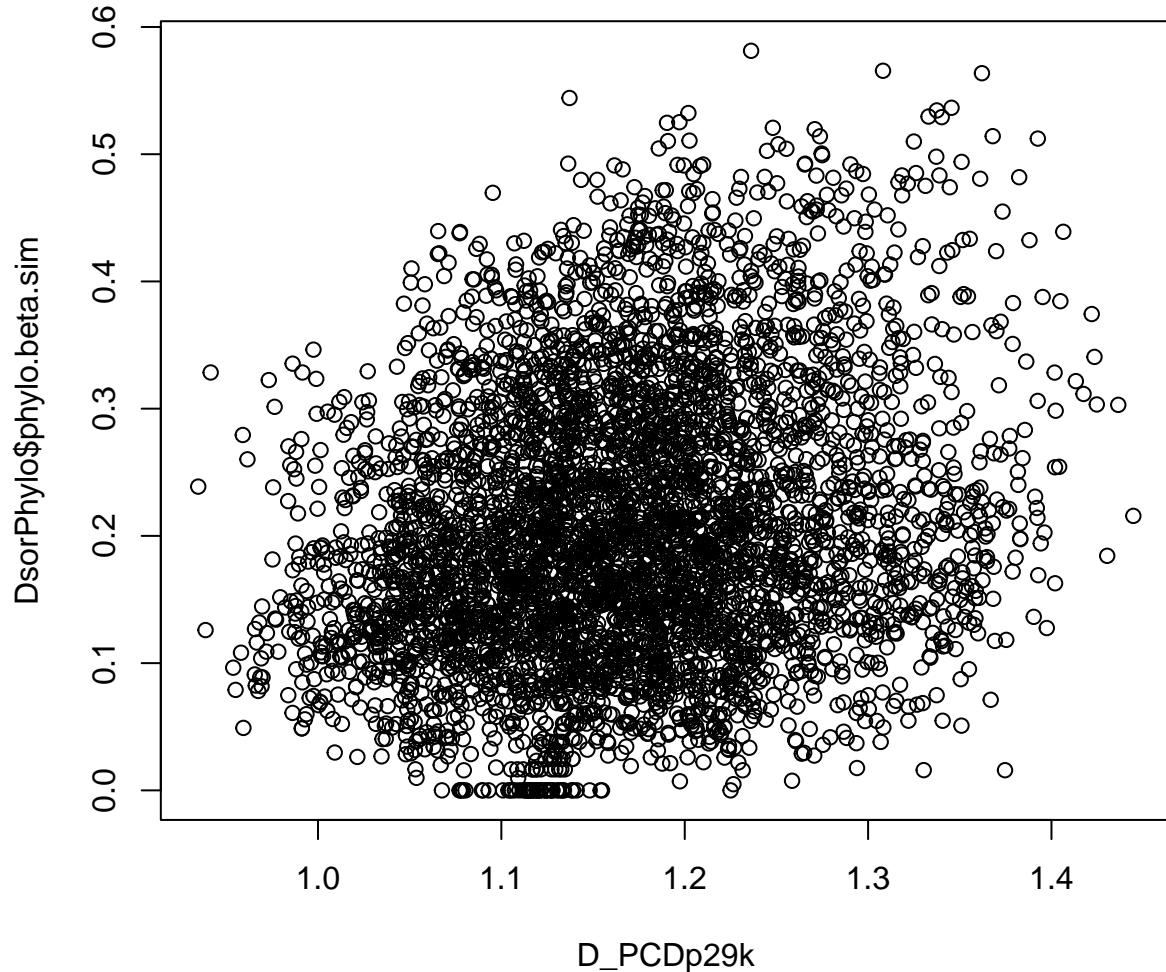
plot(D_fish$beta.sim ~ D_PCDp29k)
```



```
cor.test(D_fish$beta.sim, D_PCDp29k)

##
## Pearson's product-moment correlation
##
## data: D_fish$beta.sim and D_PCDp29k
## t = 24, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.31 0.36
## sample estimates:
## cor
## 0.33
```

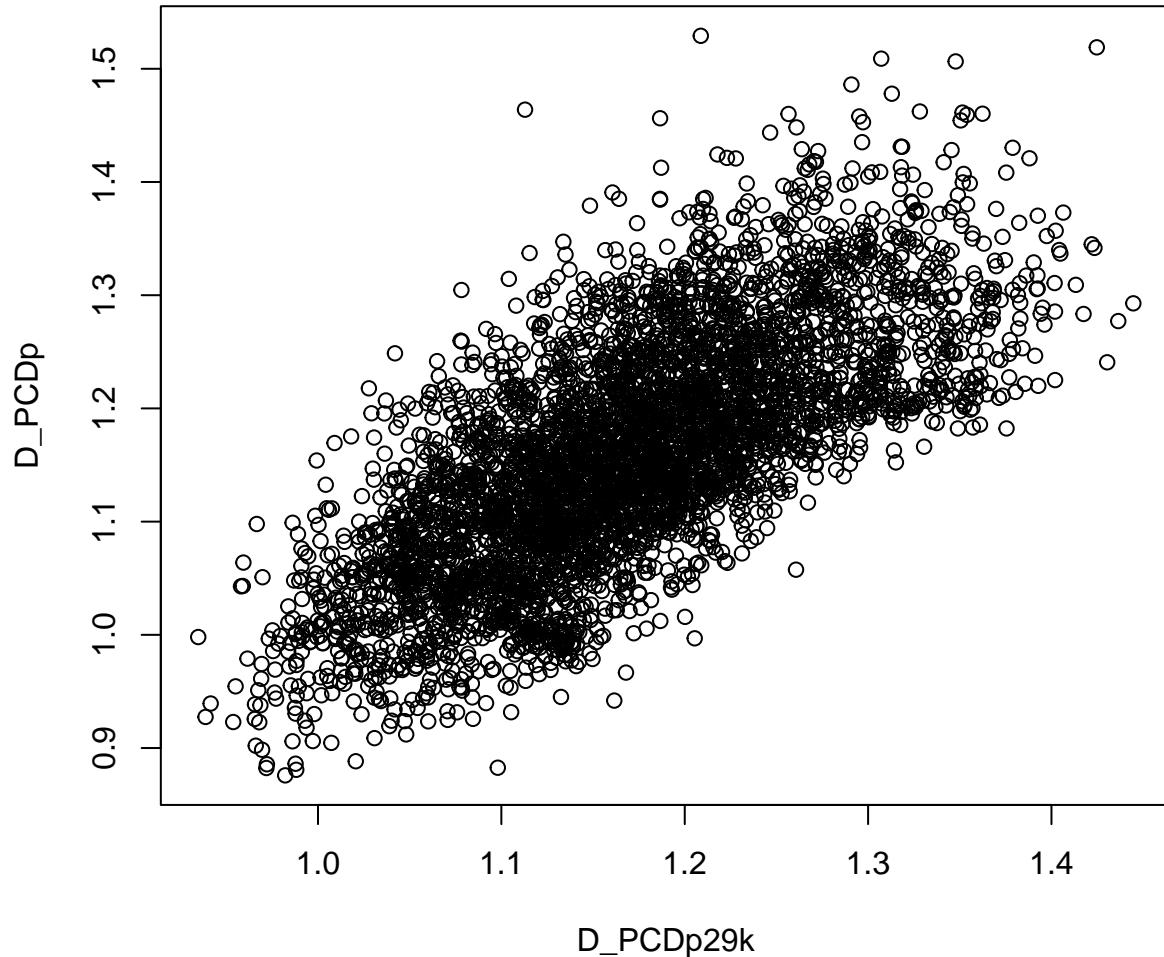
```
plot(DsorPhylo$phylo.beta.sim ~ D_PCDp29k)
```



```
cor.test(DsorPhylo$phylo.beta.sim, D_PCDp29k)
```

```
##  
## Pearson's product-moment correlation  
##  
## data: DsorPhylo$phylo.beta.sim and D_PCDp29k  
## t = 15, df = 4654, p-value <2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.19 0.25  
## sample estimates:  
## cor  
## 0.22
```

```
plot(D_PCDp ~ D_PCDp29k)
```



```
cor.test(D_PCDp, D_PCDp29k)
```

```
##  
## Pearson's product-moment correlation  
##  
## data: D_PCDp and D_PCDp29k  
## t = 70, df = 4654, p-value <2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.70 0.73  
## sample estimates:  
## cor  
## 0.71
```

1.4 Functional Beta Diversity

1.4.1 Charging and checking morpho data

```
# trait<-read.csv('Base morpho
# Amazon_2020_04_15.csv',sep=';',header=T,dec='.')
trait <- read.csv("Base morpho Amazon_NA filled_Sbrosse_2020_05_18.csv",
  sep = ";", header = T, dec = ",")  
  
summary(trait)
```

```
##  Genusspecies.Amazon.fishbase Genusspecies.base.morpho      EdHd
##  Length:2273                  Length:2273                  Min.   :0.000
##  Class  :character          Class  :character          1st Qu.:0.303
##  Mode   :character          Mode   :character          Median  :0.400
##                                         Mean    :0.401
##                                         3rd Qu.:0.510
##                                         Max.   :0.867
##                                         NA's   :355
##  MoBd            JlHd            EhBd            BlBd
##  Min.   :0.000   Min.   :0.000   Min.   :0.157   Min.   : 1.12
##  1st Qu.:0.244   1st Qu.:0.214   1st Qu.:0.489   1st Qu.: 2.96
##  Median  :0.426   Median  :0.322   Median  :0.534   Median  : 3.90
##  Mean    :0.374   Mean    :0.350   Mean    :0.541   Mean    : 4.49
##  3rd Qu.:0.516   3rd Qu.:0.411   3rd Qu.:0.589   3rd Qu.: 4.91
##  Max.   :0.887   Max.   :7.060   Max.   :0.971   Max.   :40.34
##  NA's   :355     NA's   :355     NA's   :355     NA's   :355
##  HdBd            PFiBd           PF1B1           CFdCPd
##  Min.   :0.285   Min.   :0.000   Min.   :0.000   Min.   : 0.00
##  1st Qu.:0.498   1st Qu.:0.157   1st Qu.:0.161   1st Qu.: 2.08
##  Median  :0.560   Median  :0.217   Median  :0.186   Median  : 2.60
##  Mean    :0.576   Mean    :0.228   Mean    :0.192   Mean    : 2.78
##  3rd Qu.:0.647   3rd Qu.:0.325   3rd Qu.:0.219   3rd Qu.: 3.19
##  Max.   :0.988   Max.   :0.844   Max.   :0.511   Max.   :15.45
##  NA's   :355     NA's   :355     NA's   :355     NA's   :355
##  Length
##  Min.   : 1.00
##  1st Qu.: 5.82
##  Median :10.00
##  Mean   :16.55
##  3rd Qu.:19.00
##  Max.   :450.00
##  NA's   :355
```

```
# removing trophic level
trait <- trait[, !colnames(trait) %in% c("Niveau_Trophique")]
summary(trait)
```

```
##  Genusspecies.Amazon.fishbase Genusspecies.base.morpho      EdHd
##  Length:2273                  Length:2273                  Min.   :0.000
```

```

##   Class :character          Class :character          1st Qu.:0.303
##   Mode  :character          Mode  :character          Median :0.400
##                                         Mean   :0.401
##                                         3rd Qu.:0.510
##                                         Max.   :0.867
##                                         NA's    :355
##   MoBd      JlHd      EhBd      BlBd
##   Min.    :0.000  Min.    :0.000  Min.    :0.157  Min.    : 1.12
##   1st Qu.:0.244  1st Qu.:0.214  1st Qu.:0.489  1st Qu.: 2.96
##   Median :0.426  Median :0.322  Median :0.534  Median : 3.90
##   Mean   :0.374  Mean   :0.350  Mean   :0.541  Mean   : 4.49
##   3rd Qu.:0.516 3rd Qu.:0.411  3rd Qu.:0.589  3rd Qu.: 4.91
##   Max.   :0.887  Max.   :0.060  Max.   :0.971  Max.   :40.34
##   NA's    :355   NA's    :355   NA's    :355   NA's    :355
##   HdBd      PFiBd      PF1B1      CFdCPd
##   Min.    :0.285  Min.    :0.000  Min.    :0.000  Min.    : 0.00
##   1st Qu.:0.498  1st Qu.:0.157  1st Qu.:0.161  1st Qu.: 2.08
##   Median :0.560  Median :0.217  Median :0.186  Median : 2.60
##   Mean   :0.576  Mean   :0.228  Mean   :0.192  Mean   : 2.78
##   3rd Qu.:0.647 3rd Qu.:0.325  3rd Qu.:0.219  3rd Qu.: 3.19
##   Max.   :0.988  Max.   :0.844  Max.   :0.511  Max.   :15.45
##   NA's    :355   NA's    :355   NA's    :355   NA's    :355
##   Length
##   Min.   : 1.00
##   1st Qu.: 5.82
##   Median :10.00
##   Mean   :16.55
##   3rd Qu.:19.00
##   Max.   :450.00
##   NA's    :355

trait[1:15, ]

##   Genusspecies.Amazon.fishbase      Genusspecies.base.morpho
## 1           Abramites.hypselonotus  Abramites.hypselonotus 0.47
## 2           Acanthicus.hystrix     Acanthicus.hystrix  0.14
## 3           Acanthobunocephalus.nicoi Acanthobunocephalus.nicoi 0.28
## 4           Acanthodoras.cataphractus Acanthodoras.cataphractus 0.21
## 5           Acanthodoras.depressus   Acanthodoras.depressus 0.31
## 6           Acanthodoras.spinosissimus Acanthodoras.spinosissimus 0.17
## 7           Acanthopoma.annectens    Acanthopoma.annectens 0.38
## 8           Acarichthys.heckelii    Acarichthys.heckelii  0.27
## 9           Acaronia.nassa        Acaronia.nassa  0.51
## 10          Acaronia.vultuosa      Acaronia.vultuosa  0.42
## 11          Acestridium.dichromum   Acestridium.dichromum 0.38
## 12          Acestridium.discus      Acestridium.discus  0.43
## 13          Acestridium.gymnogaster Acestridium.gymnogaster NA
## 14          Acestridium.martini     Acestridium.martini 0.45
## 15          Acestridium.scutatum     Acestridium.scutatum NA
##   EdHd MoBd JlHd EhBd
## 1   0.47 0.48 0.14 0.40
## 2   0.14 0.00 0.00 0.60
## 3   0.28 0.23 0.29 0.53
## 4   0.21 0.33 0.30 0.56
## 5   0.31 0.24 0.32 0.53
## 6   0.17 0.28 0.49 0.49
## 7   0.38 0.00 0.30 0.37
## 8   0.27 0.36 0.25 0.58
## 9   0.51 0.39 0.60 0.53
## 10          Acaronia.vultuosa  0.42
## 11          Acestridium.dichromum 0.38
## 12          Acestridium.discus  0.43
## 13          Acestridium.gymnogaster NA
## 14          Acestridium.martini 0.45
## 15          Acestridium.scutatum NA
##   BlBd HdBd PFiBd PF1B1 CFdCPd Length
## 1   3.0  0.40  0.24 0.168   3.1    14
## 2   4.5  0.79  0.00 0.404   7.7    53

```

```

## 3 5.0 0.58 0.25 0.192 2.1 2
## 4 4.8 0.52 0.21 0.267 2.0 12
## 5 4.7 0.56 0.22 0.197 2.4 8
## 6 3.8 0.58 0.26 0.278 2.3 14
## 7 6.5 0.50 0.25 0.144 1.7 12
## 8 2.9 0.67 0.34 0.236 3.4 13
## 9 2.2 0.60 0.45 0.333 1.8 15
## 10 2.1 0.62 0.39 0.339 1.2 12
## 11 11.3 0.80 0.00 0.092 8.0 6
## 12 17.0 0.83 0.00 0.079 7.2 7
## 13 NA NA NA NA NA NA
## 14 17.2 0.89 0.00 0.076 4.4 7
## 15 NA NA NA NA NA NA

# NA values
# na<-as.vector(trait$MorphoAmazon_baseMorpho[15])
na <- as.vector(trait$Genusspecies.base.morpho[15])
trait[trait == na] <- NA

summary(trait)

##   Genusspecies.Amazon.fishbase Genusspecies.base.morpho      EdHd
## Length:2273                  Length:2273      Min.   :0.000
## Class :character             Class :character  1st Qu.:0.303
## Mode  :character             Mode  :character Median  :0.400
##                                     Mean   :0.401
##                                     3rd Qu.:0.510
##                                     Max.   :0.867
##                                     NA's   :355
##   MoBd          JlHd          EhBd          BlBd
## Min.   :0.000   Min.   :0.000   Min.   :0.157   Min.   : 1.12
## 1st Qu.:0.244   1st Qu.:0.214   1st Qu.:0.489   1st Qu.: 2.96
## Median :0.426   Median :0.322   Median :0.534   Median : 3.90
## Mean   :0.374   Mean   :0.350   Mean   :0.541   Mean   : 4.49
## 3rd Qu.:0.516   3rd Qu.:0.411   3rd Qu.:0.589   3rd Qu.: 4.91
## Max.   :0.887   Max.   :7.060   Max.   :0.971   Max.   :40.34
## NA's   :355     NA's   :355     NA's   :355     NA's   :355
##   HdBd          PFiBd         PF1Bl        CFdCPd
## Min.   :0.285   Min.   :0.000   Min.   :0.000   Min.   : 0.00
## 1st Qu.:0.498   1st Qu.:0.157   1st Qu.:0.161   1st Qu.: 2.08
## Median :0.560   Median :0.217   Median :0.186   Median : 2.60
## Mean   :0.576   Mean   :0.228   Mean   :0.192   Mean   : 2.78
## 3rd Qu.:0.647   3rd Qu.:0.325   3rd Qu.:0.219   3rd Qu.: 3.19
## Max.   :0.988   Max.   :0.844   Max.   :0.511   Max.   :15.45
## NA's   :355     NA's   :355     NA's   :355     NA's   :355
##   Length
## Min.   : 1.00
## 1st Qu.: 5.82
## Median :10.00
## Mean   :16.55
## 3rd Qu.:19.00
## Max.   :450.00

```

```
## NA's :355
```

```
dim(trait)
```

```
## [1] 2273 12
```

```
traitFULL <- trait
```

```
head(traitFULL[, 1:7], 30)
```

```
##      Genusspecies.Amazon.fishbase
## 1          Abramites.hypselonotus
## 2          Acanthicus.hystrix
## 3          Acanthobunocephalus.nicoi
## 4          Acanthodoras.cataphractus
## 5          Acanthodoras.depressus
## 6          Acanthodoras.spinosissimus
## 7          Acanthopoma.annectens
## 8          Acarichthys.heckelii
## 9          Acaronia.nassa
## 10         Acaronia.vultuosa
## 11         Acestridium.dichromum
## 12         Acestridium.discus
## 13         Acestridium.gymnogaster
## 14         Acestridium.martini
## 15         Acestridium.scutatum
## 16         Acestridium.triplax
## 17         Acestrocephalus.acutus
## 18         Acestrocephalus.boehlkei
## 19 Acestrocephalus.nigrifasciatus
## 20         Acestrocephalus.pallidus
## 21         Acestrocephalus.sardina
## 22         Acestrocephalus.stigmatus
## 23         Acestrorhynchus.abbreviatus
## 24         Acestrorhynchus.altus
## 25         Acestrorhynchus.falcatus
## 26         Acestrorhynchus.falcirostris
## 27         Acestrorhynchus.grandoculis
## 28         Acestrorhynchus.heterolepis
## 29         Acestrorhynchus.isalineae
## 30         Acestrorhynchus.lacustris
##      EhBd BlBd
## 1  0.40  3.0
## 2  0.60  4.5
## 3  0.53  5.0
## 4  0.56  4.8
## 5  0.53  4.7
## 6  0.49  3.8
## 7  0.37  6.5
## 8  0.58  2.9
## 9  0.53  2.2
## 10 0.56  2.1
```

```
##      Genusspecies.base.morpho EdHd MoBd JlHd
## 1          Abramites.hypselonotus 0.47 0.48 0.14
## 2          Acanthicus.hystrix 0.14 0.00 0.00
## 3          Acanthobunocephalus.nicoi 0.28 0.23 0.29
## 4          Acanthodoras.cataphractus 0.21 0.33 0.30
## 5          Acanthodoras.depressus 0.31 0.24 0.32
## 6          Acanthodoras.spinosissimus 0.17 0.28 0.49
## 7          Acanthopoma.annectens 0.38 0.00 0.30
## 8          Acarichthys.heckelii 0.27 0.36 0.25
## 9          Acaronia.nassa 0.51 0.39 0.60
## 10         Acaronia.vultuosa 0.42 0.48 0.56
## 11         Acestridium.dichromum 0.38 0.00 0.00
## 12         Acestridium.discus 0.43 0.00 0.00
## 13         <NA>    NA   NA   NA
## 14         Acestridium.martini 0.45 0.00 0.00
## 15         <NA>    NA   NA   NA
## 16         Acestridium.triplax 0.27 0.00 0.00
## 17         <NA>    NA   NA   NA
## 18         Acestrocephalus.boehlkei 0.52 0.48 1.08
## 19         <NA>    NA   NA   NA
## 20         Acestrocephalus.pallidus 0.52 0.50 1.27
## 21         Acestrocephalus.sardina 0.49 0.62 1.04
## 22         <NA>    NA   NA   NA
## 23         Acestrorhynchus.abbreviatus 0.44 0.52 0.37
## 24         Acestrorhynchus.altus 0.39 0.55 1.32
## 25         Acestrorhynchus.falcatus 0.39 0.52 1.23
## 26         Acestrorhynchus.falcirostris 0.46 0.64 1.55
## 27         Acestrorhynchus.grandoculis 0.44 0.54 1.45
## 28         Acestrorhynchus.heterolepis 0.46 0.50 1.56
## 29         Acestrorhynchus.isalineae 0.53 0.51 1.55
## 30         Acestrorhynchus.lacustris 0.38 0.54 1.35
```

```

## 11 0.47 11.3
## 12 0.48 17.0
## 13 NA NA
## 14 0.65 17.2
## 15 NA NA
## 16 0.74 13.9
## 17 NA NA
## 18 0.51 3.6
## 19 NA NA
## 20 0.54 3.6
## 21 0.65 3.4
## 22 NA NA
## 23 0.54 3.0
## 24 0.60 3.6
## 25 0.63 4.6
## 26 0.68 5.9
## 27 0.64 6.2
## 28 0.56 5.1
## 29 0.56 6.4
## 30 0.59 3.5

# # #how many species without trait information?
# cols<-colnames(trait)[-c(1:2)] traitFULL$nNAvalues
# <-apply(traitFULL[,cols],1,function(x){ sum(is.na(x)) })
# table(traitFULL$nNAvalues==0)/nrow(traitFULL)#49% with
# complete information
# table(traitFULL$nNAvalues<=4)/nrow(traitFULL)#73% with 4
# or less NA values
# table(traitFULL$nNAvalues<=8)/nrow(traitFULL)#74% with 8
# or less NA values

# calculating FD based on filled traits (Su et al. 2020
# GEB) 2020_05_15
traitFULL <- droplevels(traitFULL[!is.na(trait$Genusspecies.base.morpho),
])
dim(traitFULL)

## [1] 1918 12

traitFULL$GenusspeciesAmazonfishbase2 <- muda_nome(traitFULL$Genusspecies.Amazon.fishbase,
sep = "[.]")
rownames(traitFULL) <- traitFULL$GenusspeciesAmazonfishbase2
summary(traitFULL)

## Genusspecies.Amazon.fishbase Genusspecies.base.morpho      EdHd
## Length:1918                  Length:1918                  Min.   :0.000
## Class :character             Class :character             1st Qu.:0.303
## Mode  :character             Mode  :character            Median :0.400
##                                         Mean   :0.401

```

```

##                                     3rd Qu.:0.510
##                                     Max.   :0.867
##      MoBd          JlHd          EhBd          BlBd
##  Min.   :0.000  Min.   :0.000  Min.   :0.157  Min.   : 1.12
##  1st Qu.:0.244  1st Qu.:0.214  1st Qu.:0.489  1st Qu.: 2.96
##  Median :0.426  Median :0.322  Median :0.534  Median : 3.90
##  Mean   :0.374  Mean   :0.350  Mean   :0.541  Mean   : 4.49
##  3rd Qu.:0.516  3rd Qu.:0.411  3rd Qu.:0.589  3rd Qu.: 4.91
##  Max.   :0.887  Max.   :7.060  Max.   :0.971  Max.   :40.34
##      HdBd          PFiBd          PF1B1          CFdCPd
##  Min.   :0.285  Min.   :0.000  Min.   :0.000  Min.   : 0.00
##  1st Qu.:0.498  1st Qu.:0.157  1st Qu.:0.161  1st Qu.: 2.08
##  Median :0.561  Median :0.217  Median :0.186  Median : 2.60
##  Mean   :0.576  Mean   :0.228  Mean   :0.192  Mean   : 2.78
##  3rd Qu.:0.648  3rd Qu.:0.325  3rd Qu.:0.219  3rd Qu.: 3.19
##  Max.   :0.988  Max.   :0.844  Max.   :0.511  Max.   :15.45
##      Length          GenusSpeciesAmazonfishbase2
##  Min.   : 1.00  Length:1918
##  1st Qu.: 5.82  Class :character
##  Median :10.00  Mode  :character
##  Mean   :16.55
##  3rd Qu.:19.00
##  Max.   :450.00

```

```

# performing gower on trait values with the same species in
# fish_mat remove species not in the matrix
toto <- unique(traitFULL$GenusSpeciesAmazonfishbase2)
sp <- colnames(fish_mat)
length(toto)

```

```
## [1] 1918
```

```
length(sp)
```

```
## [1] 2392
```

```
head(setdiff(toto, sp), 30) #####Normal: Seb has sent names according Tedesco et al. 2017
```

```

## [1] "Acestrorhynchus_lacustris"      "Achirus_achirus"
## [3] "Acrobrycon_tarijae"             "Adontosternarchus_sachsi"
## [5] "Ageneiosus_brevis"              "Ageneiosus_pardalis"
## [7] "Aguarunichthys_tocantinsensis" "Ammocryptocharax_lateralis"
## [9] "Anableps_anableps"              "Anableps_microlepis"
## [11] "Anablepsoides_immaculatus"     "Anablepsoides_xanthonotus"
## [13] "Anchovia_clupeoides"            "Ancistrus_cirrhosus"
## [15] "Ancistrus_fulvus"               "Ancistrus_leucostictus"
## [17] "Ancistrus_macrophthalmus"      "Ancistrus_maculatus"
## [19] "Andinoacara_coeruleopunctatus" "Apareiodon_piracicabae"
## [21] "Aphyocharax_dentatus"           "Aphyocharax_paraguayensis"
## [23] "Aphyocharax_rathbuni"           "Aphyocheirodon_hemigrammus"

```

```

## [25] "Aphyolebias_wischmanni"      "Apteronotus_macrostomus"
## [27] "Archolaemus_blax"           "Aspidoras_belenos"
## [29] "Aspidoras_pauciradiatus"    "Astroblepus_cyclopus"

head(setdiff(sp, toto), 30)

## [1] "Acestridium_gymnogaster"      "Acestridium_scutatum"
## [3] "Acestrocephalus_acutus"        "Acestrocephalus_anomalus"
## [5] "Acestrocephalus_nigrifasciatus" "Acestrocephalus_stigmatus"
## [7] "Acrobrycon_starnesi"           "Adontosternarchus_duartei"
## [9] "Adontosternarchus_nebulosus"    "Aequidens_superomaculatum"
## [11] "Ageneiosus_akamai"             "Ageneiosus_intrusus"
## [13] "Ageneiosus_lineatus"           "Amazonspinther_dalmata"
## [15] "Amblydoras_gonzalezi"          "Ammoglanis_amapaensis"
## [17] "Anablepsoides_amanan"          "Anablepsoides_cajariensis"
## [19] "Anablepsoides_chapare"         "Anablepsoides_gamae"
## [21] "Anablepsoides_henschelae"       "Anablepsoides_hoetmeri"
## [23] "Anablepsoides_jari"            "Anablepsoides_lineasoppilatae"
## [25] "Anablepsoides_ottonii"          "Anablepsoides_parlettei"
## [27] "Anablepsoides_roraima"          "Anablepsoides_urubuiensis"
## [29] "Anablepsoides_xinguensis"        "Anchoviella_hernanni"

length(intersect(sp, toto))/length(sp) #1661 species remained with trait information (69%)

## [1] 0.69

# AGAIN excluding species from fish_mat according to trait
dim(fish_mat)

## [1] 97 2392

fish_mat_trait <- fish_mat[, intersect(sp, toto)]
dim(fish_mat_trait)

## [1] 97 1661

head(traitFULL)

##                                         Genusspecies.Amazon.fishbase
## Abramites_hypselonotus                 Abramites.hypselonotus
## Acanthicus_hystrix                     Acanthicus.hystrix
## Acanthobunocephalus_nicoi              Acanthobunocephalus.nicoi
## Acanthodoras_cataphractus              Acanthodoras.cataphractus
## Acanthodoras_depressus                 Acanthodoras.depressus
## Acanthodoras_spinosissimus             Acanthodoras.spinosissimus
##                                         Genusspecies.base.morpho EdHd MoBd JlHd EhBd BlBd
## Abramites_hypselonotus 0.47 0.48 0.14 0.40 3.0
## Acanthicus.hystrix 0.14 0.00 0.00 0.60 4.5

```

```

## Acanthobunocephalus_nicoi   Acanthobunocephalus.nicoi 0.28 0.23 0.29 0.53 5.0
## Acanthodoras_cataphractus Acanthodoras.cataphractus 0.21 0.33 0.30 0.56 4.8
## Acanthodoras_depressus     Acanthodoras.depressus 0.31 0.24 0.32 0.53 4.7
## Acanthodoras_spinosissimus Acanthodoras.spinosissimus 0.17 0.28 0.49 0.49 3.8
##
##                                     HdBd PFiBd PF1B1 CFdCPd Length
## Abramites_hypselonotus        0.40 0.24 0.17 3.1    14
## Acanthicus_hystrix           0.79 0.00 0.40 7.7    53
## Acanthobunocephalus_nicoi    0.58 0.25 0.19 2.1    2
## Acanthodoras_cataphractus   0.52 0.21 0.27 2.0    12
## Acanthodoras_depressus      0.56 0.22 0.20 2.4    8
## Acanthodoras_spinosissimus 0.58 0.26 0.28 2.3    14
##
##                                     GenusspeciesAmazonfishbase2
## Abramites_hypselonotus        Abramites_hypselonotus
## Acanthicus_hystrix            Acanthicus.hystrix
## Acanthobunocephalus_nicoi    Acanthobunocephalus.nicoi
## Acanthodoras_cataphractus   Acanthodoras.cataphractus
## Acanthodoras_depressus       Acanthodoras.depressus
## Acanthodoras_spinosissimus  Acanthodoras.spinosissimus

traitPCA <- droplevels(traitFULL[intersect(sp, toto), ])

traitPCA[1:25, ] #c('EdHd', 'MoBd', 'JlHd', 'EhBd', 'BlBd', 'HdBd', 'PFiBd', 'PFlB1', 'CFdCPd', 'Length')

##                                     Genusspecies.Amazon.fishbase
## Abramites_hypselonotus        Abramites.hypselonotus
## Acanthicus_hystrix            Acanthicus.hystrix
## Acanthobunocephalus_nicoi    Acanthobunocephalus.nicoi
## Acanthodoras_cataphractus   Acanthodoras.cataphractus
## Acanthodoras_depressus       Acanthodoras.depressus
## Acanthodoras_spinosissimus  Acanthodoras.spinosissimus
## Acanthopoma_annectens         Acanthopoma.annectens
## Acarichthys_heckelii          Acarichthys.heckelii
## Acaronia_nassa                Acaronia.nassa
## Acaronia_vultuosa             Acaronia.vultuosa
## Acestridium_dichromum         Acestridium.dichromum
## Acestridium_discus            Acestridium.discus
## Acestridium_martini           Acestridium.martini
## Acestridium_triplax           Acestridium.triplax
## Acistrocephalus_boehlkei     Acistrocephalus.boehlkei
## Acistrocephalus_pallidus     Acistrocephalus.pallidus
## Acistrocephalus_sardina      Acistrocephalus.sardina
## Acistrorhynchus_abbreviatus   Acistrorhynchus.abbreviatus
## Acistrorhynchus_altus          Acistrorhynchus.altus
## Acistrorhynchus_falcatus      Acistrorhynchus.falcatus
## Acistrorhynchus_falcirostris  Acistrorhynchus.falcirostris
## Acistrorhynchus_grandoculis   Acistrorhynchus.grandoculis
## Acistrorhynchus_heterolepis   Acistrorhynchus.heterolepis
## Acistrorhynchus_isalineae     Acistrorhynchus.isalineae
## Acistrorhynchus_maculipinna   Acistrorhynchus.maculipinna
##                                     Genusspecies.base.morpho EdHd MoBd JlHd EhBd
## Abramites_hypselonotus        Abramites.hypselonotus 0.47 0.48 0.14 0.40

```

## Acanthicus_hystrix	Acanthicus.hystrix	0.14	0.00	0.00	0.60
## Acanthobunocephalus_nicoi	Acanthobunocephalus.nicoi	0.28	0.23	0.29	0.53
## Acanthodoras_cataphractus	Acanthodoras.cataphractus	0.21	0.33	0.30	0.56
## Acanthodoras_depressus	Acanthodoras.depressus	0.31	0.24	0.32	0.53
## Acanthodoras_spinosissimus	Acanthodoras.spinosissimus	0.17	0.28	0.49	0.49
## Acanthopoma_annectens	Acanthopoma.annectens	0.38	0.00	0.30	0.37
## Acarichthys_heckelii	Acarichthys.heckelii	0.27	0.36	0.25	0.58
## Acaronia_nassa	Acaronia.nassa	0.51	0.39	0.60	0.53
## Acaronia_vultuosa	Acaronia.vultuosa	0.42	0.48	0.56	0.56
## Acestridium_dichromum	Acestridium.dichromum	0.38	0.00	0.00	0.47
## Acestridium_discus	Acestridium.discus	0.43	0.00	0.00	0.48
## Acestridium_martini	Acestridium.martini	0.45	0.00	0.00	0.65
## Acestridium_triplax	Acestridium.triplax	0.27	0.00	0.00	0.74
## Acistrocephalus_boehlkei	Acistrocephalus.boehlkei	0.52	0.48	1.08	0.51
## Acistrocephalus_pallidus	Acistrocephalus.pallidus	0.52	0.50	1.27	0.54
## Acistrocephalus_sardina	Acistrocephalus.sardina	0.49	0.62	1.04	0.65
## Acistrorhynchus_abbreviatus	Acistrorhynchus.abbreviatus	0.44	0.52	0.37	0.54
## Acistrorhynchus_altus	Acistrorhynchus.altus	0.39	0.55	1.32	0.60
## Acistrorhynchus_falcatus	Acistrorhynchus.falcatus	0.39	0.52	1.23	0.63
## Acistrorhynchus_falcirostris	Acistrorhynchus.falcirostris	0.46	0.64	1.55	0.68
## Acistrorhynchus_grandoculis	Acistrorhynchus.grandoculis	0.44	0.54	1.45	0.64
## Acistrorhynchus_heterolepis	Acistrorhynchus.heterolepis	0.46	0.50	1.56	0.56
## Acistrorhynchus_isalineae	Acistrorhynchus.isalineae	0.53	0.51	1.55	0.56
## Acistrorhynchus_maculipinna	Acistrorhynchus.maculipinna	0.52	0.62	1.76	0.62
##					
## Abramites_hypselonotus	3.0	0.40	0.24	0.168	3.1
## Acanthicus_hystrix	4.5	0.79	0.00	0.404	7.7
## Acanthobunocephalus_nicoi	5.0	0.58	0.25	0.192	2.1
## Acanthodoras_cataphractus	4.8	0.52	0.21	0.267	2.0
## Acanthodoras_depressus	4.7	0.56	0.22	0.197	2.4
## Acanthodoras_spinosissimus	3.8	0.58	0.26	0.278	2.3
## Acanthopoma_annectens	6.5	0.50	0.25	0.144	1.7
## Acarichthys_heckelii	2.9	0.67	0.34	0.236	3.4
## Acaronia_nassa	2.2	0.60	0.45	0.333	1.8
## Acaronia_vultuosa	2.1	0.62	0.39	0.339	1.2
## Acestridium_dichromum	11.3	0.80	0.00	0.092	8.0
## Acestridium_discus	17.0	0.83	0.00	0.079	7.2
## Acestridium_martini	17.2	0.89	0.00	0.076	4.4
## Acestridium_triplax	13.9	0.92	0.40	0.087	2.7
## Acistrocephalus_boehlkei	3.6	0.47	0.16	0.143	3.4
## Acistrocephalus_pallidus	3.6	0.54	0.18	0.174	3.8
## Acistrocephalus_sardina	3.4	0.58	0.19	0.145	3.8
## Acistrorhynchus_abbreviatus	3.0	0.46	0.23	0.184	2.9
## Acistrorhynchus_altus	3.6	0.57	0.20	0.171	3.9
## Acistrorhynchus_falcatus	4.6	0.65	0.21	0.166	4.0
## Acistrorhynchus_falcirostris	5.9	0.72	0.17	0.134	3.7
## Acistrorhynchus_grandoculis	6.2	0.77	0.17	0.145	4.6
## Acistrorhynchus_heterolepis	5.1	0.51	0.21	0.080	4.5
## Acistrorhynchus_isalineae	6.4	0.66	0.16	0.132	4.3
## Acistrorhynchus_maculipinna	6.6	0.72	0.22	0.135	3.8
##	GenusspeciesAmazonfishbase2				
## Abramites_hypselonotus	Abramites_hypselonotus				

```

## Acanthicus_hystrix          Acanthicus_hystrix
## Acanthobunocephalus_nicoi   Acanthobunocephalus_nicoi
## Acanthodoras_cataphractus  Acanthodoras_cataphractus
## Acanthodoras_depressus     Acanthodoras_depressus
## Acanthodoras_spinosissimus Acanthodoras_spinosissimus
## Acanthopoma_annectens      Acanthopoma_annectens
## Acarichthys_heckelii       Acarichthys_heckelii
## Acaronia_nassa             Acaronia_nassa
## Acaronia_vultuosa          Acaronia_vultuosa
## Acestridium_dichromum      Acestridium_dichromum
## Acestridium_discus         Acestridium_discus
## Acestridium_martini        Acestridium_martini
## Acestridium_triplax        Acestridium_triplax
## Acistrocephalus_boehlkei   Acistrocephalus_boehlkei
## Acistrocephalus_pallidus   Acistrocephalus_pallidus
## Acistrocephalus_sardina    Acistrocephalus_sardina
## Acestrorhynchus_abbreviatus Acestrorhynchus_abbreviatus
## Acestrorhynchus_altus      Acestrorhynchus_altus
## Acestrorhynchus_falcatus   Acestrorhynchus_falcatus
## Acestrorhynchus_falcirostris Acestrorhynchus_falcirostris
## Acestrorhynchus_grandoculis Acestrorhynchus_grandoculis
## Acestrorhynchus_heterolepis Acestrorhynchus_heterolepis
## Acestrorhynchus_isalineae  Acestrorhynchus_isalineae
## Acestrorhynchus_maculipinna Acestrorhynchus_maculipinna

```

```
dim(traitPCA)
```

```
## [1] 1661 13
```

```
summary(traitPCA)
```

```

## Genusspecies.Amazon.fishbase Genusspecies.base.morpho      EdHd
## Length:1661                  Length:1661                  Min.   :0.000
## Class  :character            Class  :character            1st Qu.:0.303
## Mode   :character            Mode   :character            Median :0.405
##                                         Mean   :0.403
##                                         3rd Qu.:0.514
##                                         Max.   :0.867
## MoBd                JlHd           EhBd           BlBd
## Min.   :0.000   Min.   :0.000   Min.   :0.157   Min.   : 1.12
## 1st Qu.:0.242   1st Qu.:0.212   1st Qu.:0.489   1st Qu.: 2.95
## Median :0.422   Median :0.322   Median :0.534   Median : 3.91
## Mean   :0.372   Mean   :0.352   Mean   :0.540   Mean   : 4.50
## 3rd Qu.:0.515   3rd Qu.:0.413   3rd Qu.:0.591   3rd Qu.: 4.91
## Max.   :0.887   Max.   :7.060   Max.   :0.971   Max.   :24.05
## HdBd                PFiBd          PF1Bl          CFdCPd
## Min.   :0.285   Min.   :0.000   Min.   :0.000   Min.   : 0.00
## 1st Qu.:0.503   1st Qu.:0.156   1st Qu.:0.161   1st Qu.: 2.09
## Median :0.562   Median :0.215   Median :0.186   Median : 2.61
## Mean   :0.580   Mean   :0.227   Mean   :0.192   Mean   : 2.79
## 3rd Qu.:0.651   3rd Qu.:0.325   3rd Qu.:0.220   3rd Qu.: 3.20

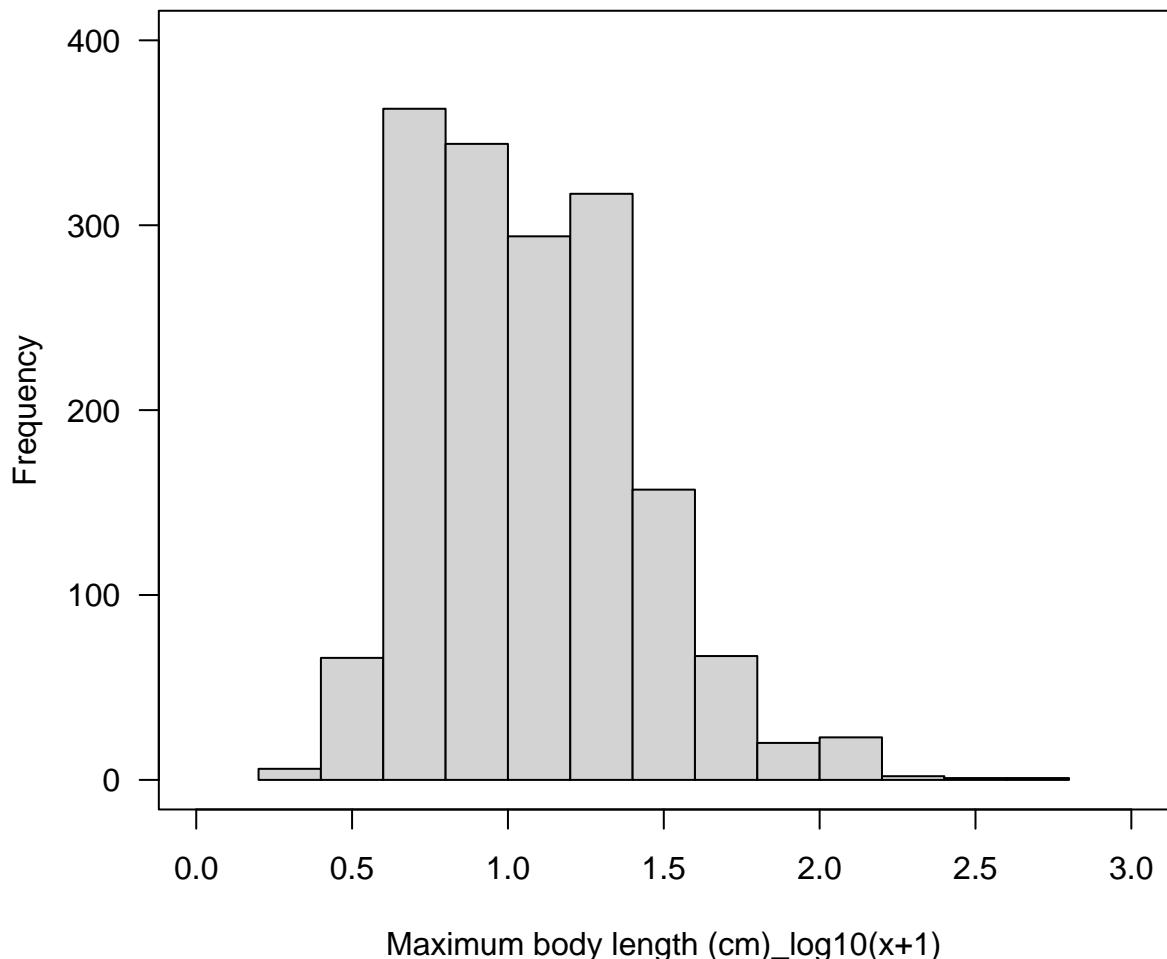
```

```

## Max.    :0.988   Max.    :0.844   Max.    :0.511   Max.    :15.45
## Length   GenusSpeciesAmazonfishbase2
## Min.    : 1.0   Length:1661
## 1st Qu.: 5.0   Class :character
## Median  :10.0   Mode  :character
## Mean    :16.2
## 3rd Qu.:18.5
## Max.    :450.0

# histogram Body size
# jpeg('Body_size_allSpecies.jpeg',width = 1000,height =
# 1000,res = 200)
hist(log10(traitPCA$Length + 1), xlab = "Maximum body length (cm)_log10(x+1)",
     main = "", las = 1, ylim = c(0, 400), xlim = c(0, 3))
box()

```



```

# dev.off()

traitPCA <- apply(traitPCA[, c("EdHd", "MoBd", "JlHd", "EhBd",

```

```

"BlBd", "HdBd", "PFIbD", "PFlBl", "CFdCPd", "Length")], 2,
  function(x) {
    (x - mean(x))/(sd(x))
  }) #standardize traits
head(traitPCA)

##                                     EdHd MoBd JlHd EhBd BlBd HdBd PFIbD
## Abramites_hypselonotus      0.47 0.52 -0.525 -1.533 -0.591 -1.5883 0.056
## Acanthicus_hystrix       -1.83 -1.76 -0.856  0.592  0.015  1.8054 -1.568
## Acanthobunocephalus_nicoi -0.85 -0.68 -0.148 -0.133  0.194  0.0033 0.152
## Acanthodoras_cataphractus -1.34 -0.22 -0.123  0.240  0.123 -0.4761 -0.092
## Acanthodoras_depressus     -0.63 -0.61 -0.088 -0.065  0.081 -0.1573 -0.071
## Acanthodoras_spinosissimus -1.65 -0.44  0.326 -0.506 -0.251  0.0467 0.218
##                                     PFlBl CFdCPd Length
## Abramites_hypselonotus     -0.385   0.20 -0.093
## Acanthicus_hystrix        3.374   3.44  1.558
## Acanthobunocephalus_nicoi -0.005  -0.46 -0.601
## Acanthodoras_cataphractus 1.193  -0.53 -0.178
## Acanthodoras_depressus    0.077  -0.29 -0.347
## Acanthodoras_spinosissimus 1.363  -0.34 -0.093

trait_gow <- dist(traitPCA) #not using Gower distance, but Euclid

#####
# performing PCA
pca_trait <- prcomp(traitPCA, center = FALSE, scale. = FALSE)
# summary(pca_trait)
knitr::kable(as.data.frame(summary(pca_trait)$importance))

```

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	1.58	1.36	1.23	1.13	0.95	0.85	0.69	0.59	0.51	0.40
Proportion of Variance	0.25	0.19	0.15	0.13	0.09	0.07	0.05	0.04	0.03	0.02
Cumulative Proportion	0.25	0.43	0.59	0.71	0.80	0.88	0.92	0.96	0.98	1.00

```
knitr::kable(as.data.frame(summary(pca_trait)$rotation))
```

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
EdHd	-0.22	0.42	0.09	0.53	0.00	0.11	0.49	-0.30	-0.26	-0.28
MoBd	-0.53	0.19	-0.20	0.12	0.17	-0.20	0.15	0.15	0.18	0.70
JlHd	-0.38	-0.20	0.08	0.26	0.21	0.70	-0.30	0.34	0.00	-0.11
EhBd	-0.03	-0.34	-0.59	0.19	0.17	-0.34	0.16	0.38	-0.30	-0.31
BlBd	-0.04	-0.59	0.25	0.28	-0.23	-0.07	-0.05	-0.29	-0.47	0.38
HdBd	0.35	-0.30	-0.33	0.34	-0.08	0.29	0.29	-0.21	0.57	0.13
PFIbD	-0.48	-0.11	-0.34	-0.11	-0.01	-0.10	-0.36	-0.63	0.17	-0.27
PFlBl	0.34	0.26	-0.42	-0.09	0.37	0.32	-0.19	-0.26	-0.44	0.30

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
CFdCPd	0.22	0.02	0.29	0.45	0.59	-0.37	-0.35	-0.08	0.20	-0.05
Length	-0.11	-0.33	0.22	-0.43	0.59	0.09	0.50	-0.18	0.01	-0.05

```

write.table(x = round(as.data.frame(summary(pca_trait)$importance),
  2), file = "MorphoTraits_importance.csv", sep = ";")

write.table(x = round(as.data.frame(summary(pca_trait)$rotation),
  2), file = "MorphoTraits_rotation.csv", sep = ";")

pc_plot <- pca_trait$x[, 1:2]

jpeg("Rplot_PCAaxes.jpeg", width = 12000, height = 6000, res = 400)
par(mfrow = c(1, 2))
biplot(pca_trait, cex = 0.3, pc.biplot = FALSE, center = FALSE,
  scale = FALSE, ylim = c(-9, 4), xlim = c(-9, 4))

## Warning in plot.window(...): "center" is not a graphical parameter

## Warning in plot.xy(xy, type, ...): "center" is not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "center" is not a
## graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "center" is not a
## graphical parameter

## Warning in box(...): "center" is not a graphical parameter

## Warning in title(...): "center" is not a graphical parameter

## Warning in text.default(x, xlabs, cex = cex[1L], col = col[1L], ...): "center"
## is not a graphical parameter

## Warning in plot.window(...): "center" is not a graphical parameter

## Warning in plot.xy(xy, type, ...): "center" is not a graphical parameter

## Warning in title(...): "center" is not a graphical parameter

## Warning in axis(3, col = col[2L], ...): "center" is not a graphical parameter

## Warning in axis(4, col = col[2L], ...): "center" is not a graphical parameter

## Warning in text.default(y, labels = ylabs, cex = cex[2L], col = col[2L], :
## "center" is not a graphical parameter

```

```

plot(pc_plot, cex = 0.6, pch = 21, bg = "gray", ylim = c(-9,
  4), xlim = c(-9, 4))
dev.off()

## pdf
## 2

pdf("Rplot_PCAaxes.pdf", width = 16, height = 8)
par(mfrow = c(1, 2))
biplot(pca_trait, cex = 0.2, pc.biplot = FALSE, center = FALSE,
  scale = FALSE, ylim = c(-9, 4), xlim = c(-9, 4))

## Warning in plot.window(...): "center" is not a graphical parameter

## Warning in plot.xy(xy, type, ...): "center" is not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "center" is not a
## graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "center" is not a
## graphical parameter

## Warning in box(...): "center" is not a graphical parameter

## Warning in title(...): "center" is not a graphical parameter

## Warning in text.default(x, xlabs, cex = cex[1L], col = col[1L], ...): "center"
## is not a graphical parameter

## Warning in plot.window(...): "center" is not a graphical parameter

## Warning in plot.xy(xy, type, ...): "center" is not a graphical parameter

## Warning in title(...): "center" is not a graphical parameter

## Warning in axis(3, col = col[2L], ...): "center" is not a graphical parameter

## Warning in axis(4, col = col[2L], ...): "center" is not a graphical parameter

## Warning in text.default(y, labels = ylabs, cex = cex[2L], col = col[2L], :
## "center" is not a graphical parameter

plot(pc_plot, cex = 0.6, pch = 21, bg = "gray", ylim = c(-9,
  4), xlim = c(-9, 4))
dev.off()

## pdf
## 2

```

1.4.2 Looking mean morphological position of sub-drainages

```

# head(pca_trait$x) fish_mat_trait[1:5,1:3] for(i in
# 1:nrow(fish_mat_trait)){ #i=2 x<-fish_mat_trait[i,]
# x<-x[which(x>0)] X<-pca_trait$x[names(x),]
# media<-colMeans(X) devpad<-apply(X,2,sd)
# jpeg(paste0('mean_morpho/',rownames(x), '.jpeg'))
# plot(pca_trait$x,bg='lightgray',pch=21,col='white',main=rownames(x))
# points(X) hpts <- chull(X) hpts <- c(hpts, hpts[1])
# lines(X[hpts, ]) points(x =
# media[1],y=media[2],cex=3,col='red',pch=19) dev.off() }
# #species from the second right block (in morpho space)
# rownames(pca_trait$x[pca_trait$x[,1]>1,])

```

1.4.3 Quality of the functional

This is really time consuming. Be prepared!

```

# # Uncomment to calculate again! qualidade do espaço
# funcional Marié et al. 2015 GEB
# source('quality_funct_space_fromdist2.R') library(clue)
# quality <-quality_funct_space_fromdist (trait_gow,
# nbdim=7, plot='quality_funct_space_I') quality$meanSD

# resu2<-dudi.pco(d = trait_gow,scannf = F,nf = 10)
# summary(resu2) porc<-resu2$eig/sum(resu2$eig)
# cumsum(porc[1:5])*100 plot(resu2$li[,1:5]) dim(resu2$li)
# pcoa_eix<-as.matrix(resu2$li[,1:5])

# #parallelizing dissimilarity calculation: 14/07/2020
# BsimParallel<-function(mat=fish_mat_trait,
# trait=pca_trait$x[,1:2]){
# library(betapart);
# ID<-rownames(mat); Bsim <- as.matrix(outer(ID, ID,paste))
# rownames(Bsim)<-ID; colnames(Bsim)<-ID
# pegar<-as.data.frame(Bsim); pegar[]<-FALSE
# pegar[lower.tri(pegar)]<-TRUE; diag(pegar)<-TRUE id <-
# Bsim[as.matrix(pegar)] id <- data.frame(Bsim=NA,id=id)
# id2<-do.call('rbind',strsplit(id$id,split = ' '))
# colnames(id2)<-c('id1','id2') id<-data.frame(id,id2);
# rm(id2) id$Bsor<-NA;id$Bsne<-NA; head(id)
# library(foreach);library(doParallel) nc<-7
# registerDoParallel(cores = nc)
# resu<-foreach(a=id$id1,b=id$id2,.combine = rbind,
# .packages = c('betapart'), .export = c('mat', 'trait'))
# %dopar% { beta_funct<-functional.beta.pair(x =
# mat[c(a,b),], traits = trait, index.family = 'sorensen')
# fin<-data.frame(Bsor=as.vector(beta_funct$funct.beta.sor),
# Bsim=as.vector(beta_funct$funct.beta.sim),
# Bsne=as.vector(beta_funct$funct.beta.sne)) (fin) }
# #head(id) Bsim<-Bsne<-Bsor<-pegawai
# Bsor[Bsor==TRUE]<-resu$Bsor Bsim[Bsim==TRUE]<-resu$Bsim
# Bsne[Bsne==TRUE]<-resu$Bsne

```

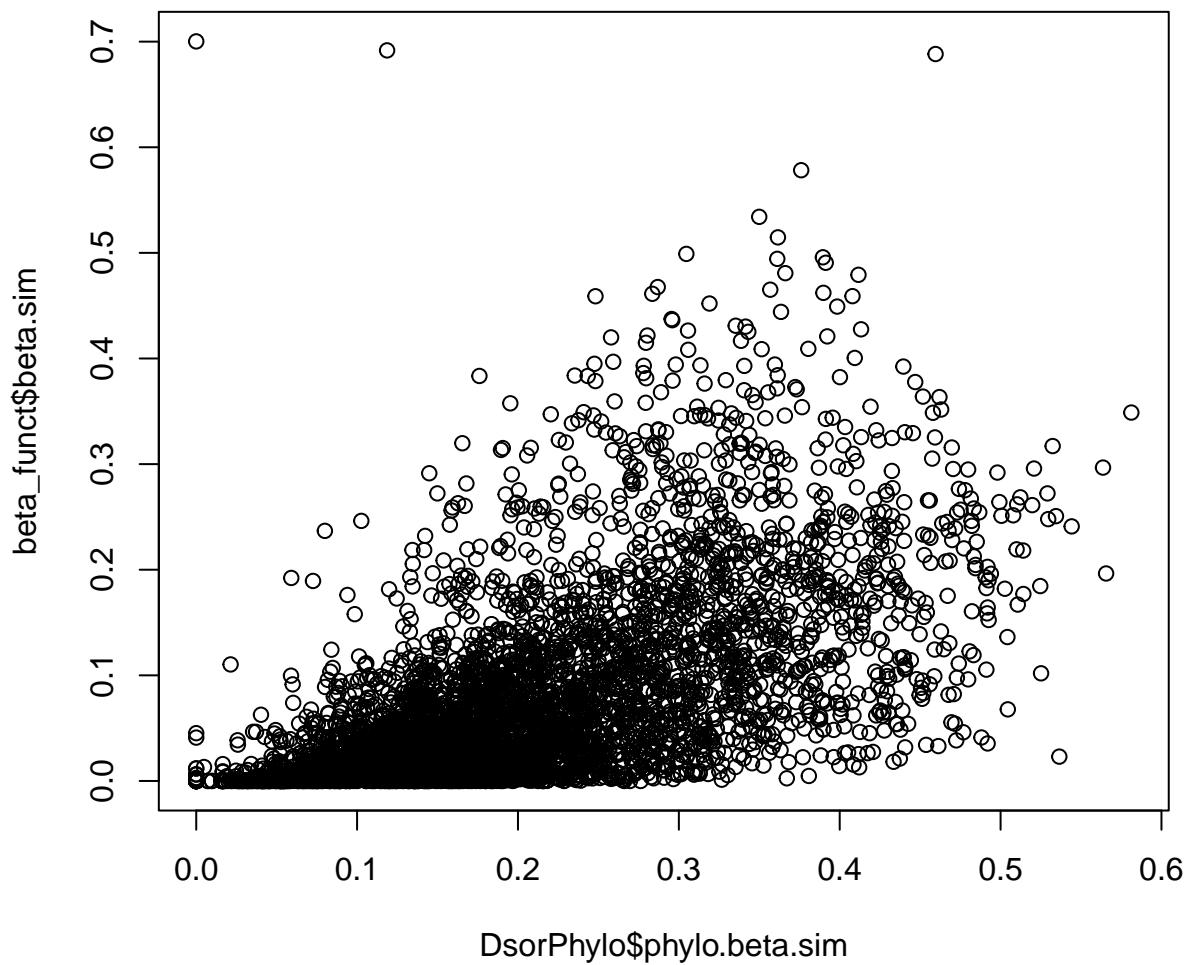
```

# return(list(Bsor=as.dist(Bsor), Bsim=as.dist(Bsim),
# Bsne=as.dist(Bsne))) }
# system.time(beta_funct<-functional.beta.pair(x =
# fish_mat_trait[1:5,], traits = pca_trait$x[,1:3],
# index.family = 'sorensen'))
# system.time(resu<-BsimParallel(mat=fish_mat_trait, trait=pca_trait$x[,1:4]))
# library(betapart)
# system.time(beta_funct<-functional.beta.pair(x =
# fish_mat_trait, traits = pca_trait$x[,1:2], index.family
# = 'sorensen')) write.table(x =
# as.matrix(beta_funct$funct.bet.sim),
# file='BetaFunct/BetaFuncSim_3D_filledNAtrait_SBrosse_2020_05_15.csv',
# sep=';') write.table(x =
# as.matrix(beta_funct$funct.bet.sor),
# file='BetaFunct/BetaFuncSor_3D_filledNAtrait_SBrosse_2020_05_15.csv',
# sep=';') write.table(x =
# as.matrix(beta_funct$funct.bet.sne),
# file='BetaFunct/BetaFuncSne_3D_filledNAtrait_SBrosse_2020_05_15.csv',
# sep=';')

beta_funct <- list(beta.sim = NA, beta.sor = NA, beta.sne = NA)
beta_funct$beta.sim = as.dist(read.table("BetaFunct/4D_BetaFuncSim_filledNAtrait_SBrosse_2021_01_26
sep = ";"))
beta_funct$beta.sor = as.dist(read.table("BetaFunct/4D_BetaFuncSor_filledNAtrait_SBrosse_2021_01_26
sep = ";"))
beta_funct$beta.sne = as.dist(read.table("BetaFunct/4D_BetaFuncSne_filledNAtrait_SBrosse_2021_01_26
sep = ";"))

plot(DsorPhylo$phylo.beta.sim, beta_funct$beta.sim)

```



```
cor.test(DsorPhylo$phylo.beta.sim, beta_funct$beta.sim)
```

```
##
## Pearson's product-moment correlation
##
## data: DsorPhylo$phylo.beta.sim and beta_funct$beta.sim
## t = 51, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.58 0.61
## sample estimates:
## cor
## 0.6
```

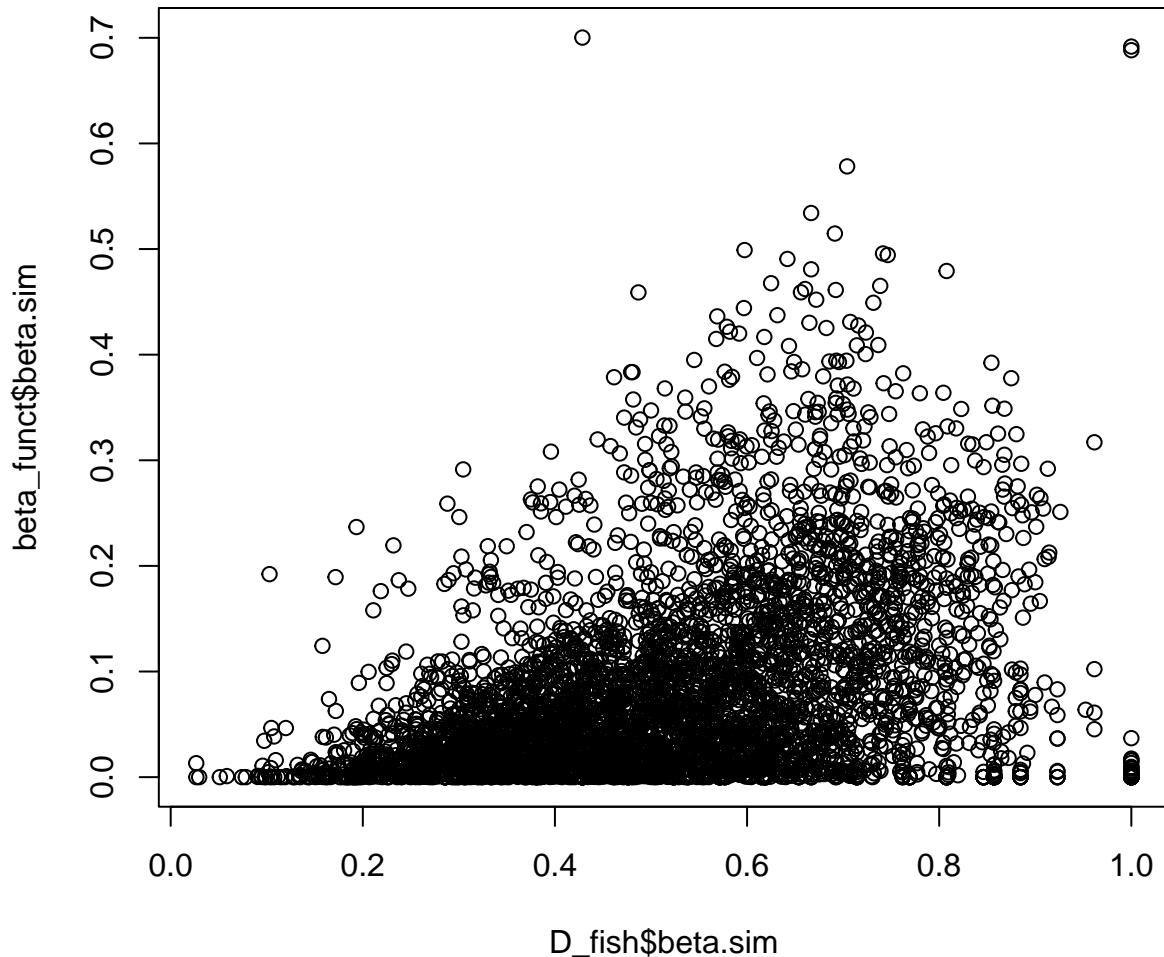
```
# taxo vs funct
summary(as.vector(beta_funct$beta.sim))
```

```
##      Min. 1st Qu. Median      Mean 3rd Qu.      Max.
## 0.0000  0.0140  0.0479  0.0792  0.1117  0.7002
```

```
sd(as.vector(beta_funct$beta.sim))
```

```
## [1] 0.089
```

```
plot(D_fish$beta.sim, beta_funct$beta.sim)
```



```
cor.test(D_fish$beta.sim, beta_funct$beta.sim)
```

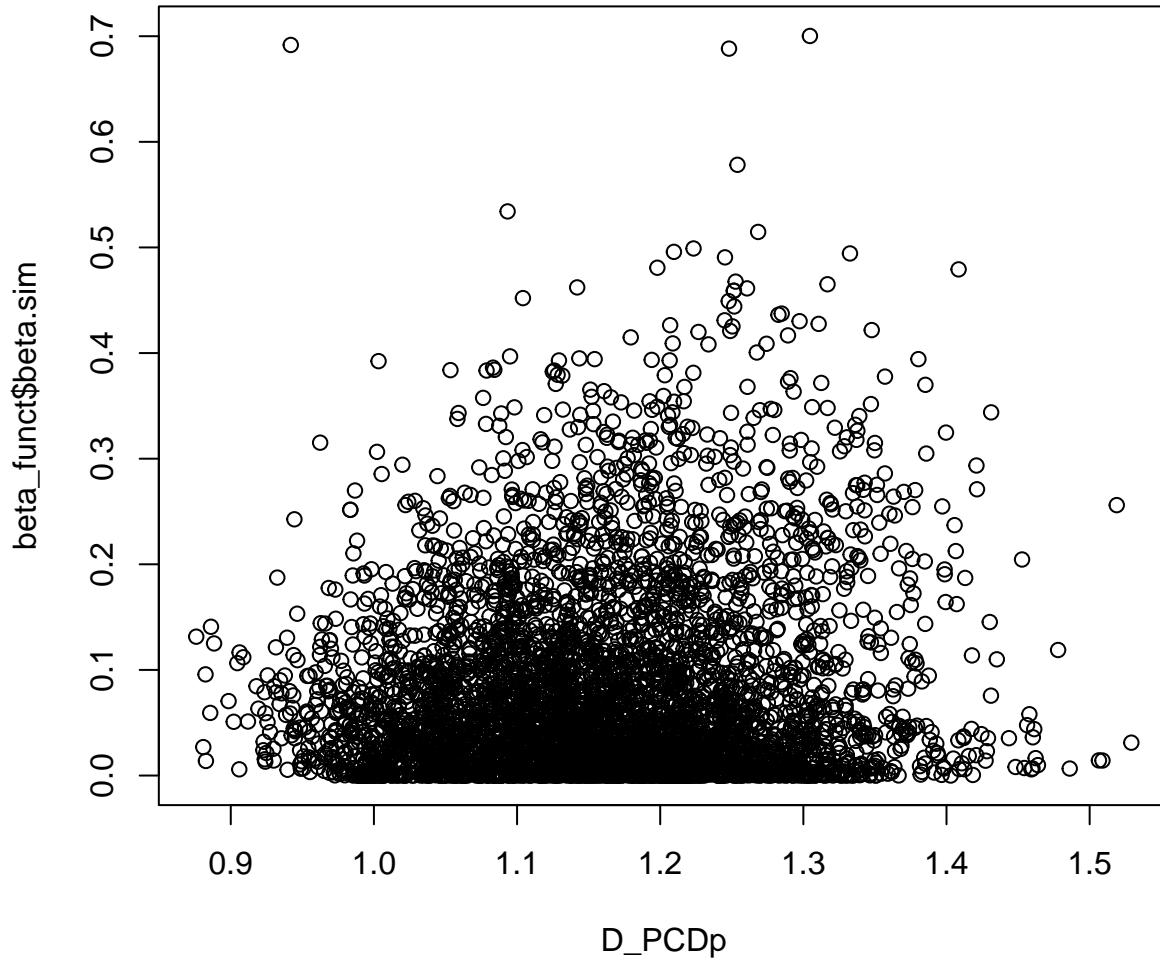
```
##  
## Pearson's product-moment correlation  
##  
## data: D_fish$beta.sim and beta_funct$beta.sim  
## t = 27, df = 4654, p-value <2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.34 0.39  
## sample estimates:
```

```

## cor
## 0.36

# PCDp vs funct
plot(D_PCDp, beta_funct$beta.sim)

```



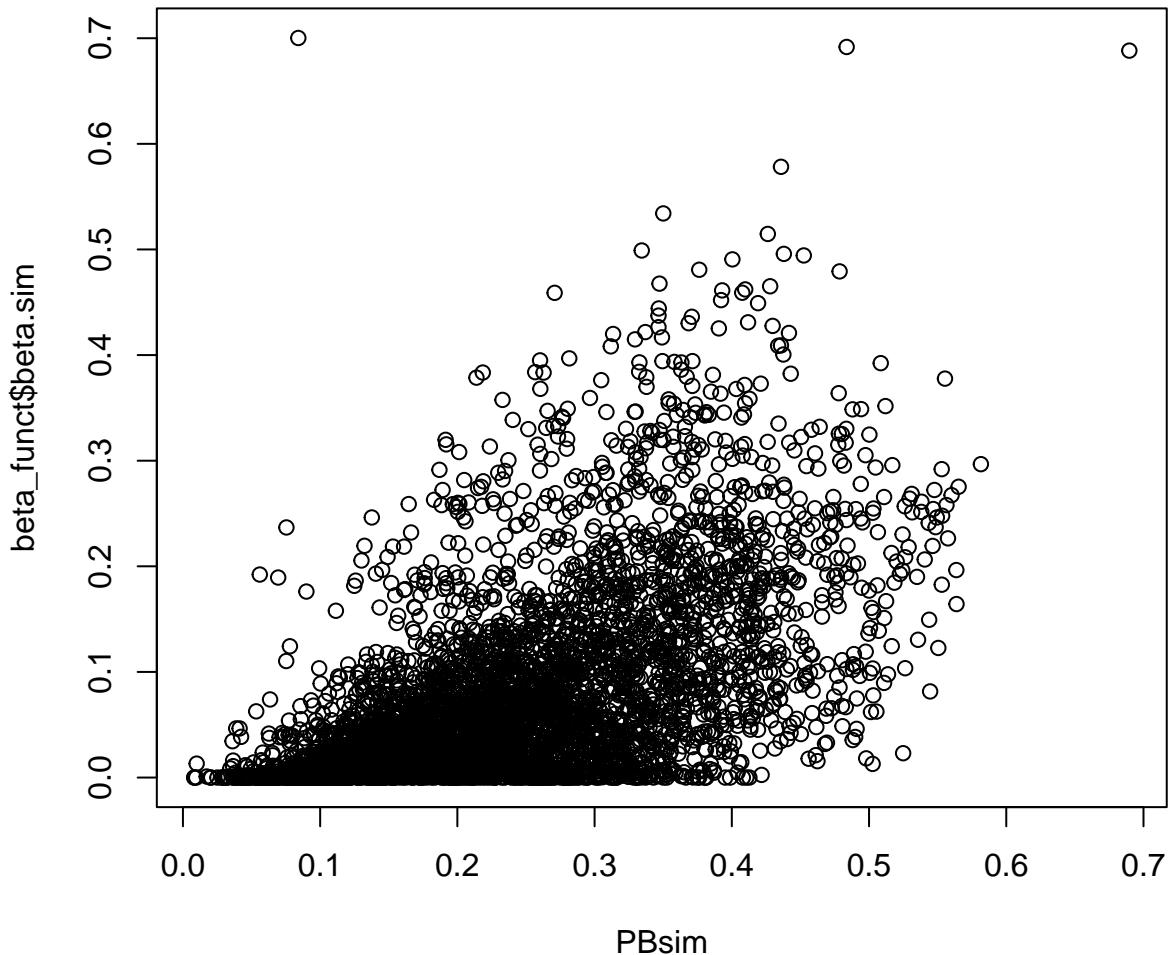
```

cor.test(D_PCDp, beta_funct$beta.sim)

##
## Pearson's product-moment correlation
##
## data: D_PCDp and beta_funct$beta.sim
## t = 7, df = 4654, p-value = 5e-12
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.072 0.129
## sample estimates:
## cor
## 0.1

```

```
# funct vs Phylo30k
plot(PBsim, beta_funct$beta.sim)
```



```
cor.test(PBsim, beta_funct$beta.sim)
```

```
##
## Pearson's product-moment correlation
##
## data: PBsim and beta_funct$beta.sim
## t = 46, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.54 0.58
## sample estimates:
## cor
## 0.56
```

2 Charging species richness and attributes of basins

```
## Richness and endemic species Charging the complete data
## set.
fish = read.csv(paste("StatistiquesSubBasinAmazon_02052016/NewData_2017_04/PhysicalData_0417.csv",
sep = ""), sep = ";", header = T)

summary(fish)
```

	BvNiv2	WaterColor	Area_km	Chut_hydrofall
##	Length:97	Length:97	Min. : 5261	Min. : 0.00
##	Class :character	Class :character	1st Qu.: 28609	1st Qu.: 0.00
##	Mode :character	Mode :character	Median : 42905	Median : 0.00
##			Mean : 60790	Mean : 2.41
##			3rd Qu.: 78109	3rd Qu.: 2.00
##			Max. :226666	Max. :26.00
##	Chut_grand	NetworkDensity	Elevation_min	Elevation_max
##	Min. :0.0000	Min. :0.0596	Min. : 0.0	Min. : 57
##	1st Qu.:0.0000	1st Qu.:0.1289	1st Qu.: 12.0	1st Qu.: 239
##	Median :0.0000	Median :0.1448	Median : 47.0	Median : 645
##	Mean :0.0825	Mean :0.1441	Mean : 71.8	Mean :1457
##	3rd Qu.:0.0000	3rd Qu.:0.1613	3rd Qu.:110.0	3rd Qu.:1759
##	Max. :2.0000	Max. :0.2265	Max. :470.0	Max. :6247
##	Elevation_range	Elevation_mean	Elevation_std	Elevation_basinSup1000m
##	Min. : 53	Min. : 26.2	Min. : 10.5	Min. : 0.00
##	1st Qu.: 195	1st Qu.: 106.4	1st Qu.: 34.2	1st Qu.: 0.00
##	Median : 570	Median : 182.5	Median : 72.7	Median : 0.00
##	Mean :1385	Mean : 399.0	Mean : 247.4	Mean : 8.27
##	3rd Qu.:1619	3rd Qu.: 313.6	3rd Qu.: 166.0	3rd Qu.: 0.58
##	Max. :6082	Max. :3882.3	Max. :1584.0	Max. :99.23
##	Elevation_basinSup1500m	NbEspeceEnveloppe	DistanceEmbouchure_km	
##	Min. : 0.0000	Min. : 18	Min. : 0	
##	1st Qu.: 0.0000	1st Qu.: 543	1st Qu.:1427	
##	Median : 0.0000	Median : 761	Median :2264	
##	Mean : 6.6002	Mean : 686	Mean :2443	
##	3rd Qu.: 0.0505	3rd Qu.: 915	3rd Qu.:3301	
##	Max. :97.9752	Max. :1099	Max. :5793	
##	X23ma_cd	X23Ma	X10Ma_cd	X10Ma
##	Length:97	Length:97	Length:97	Length:97
##	Class :character	Class :character	Class :character	Class :character
##	Mode :character	Mode :character	Mode :character	Mode :character
##				
##				
##				
##	NbSites1km_0417	NbSites_0417	Diversity0417	Endem0417
##	Min. : 4	Min. : 4	Min. : 7	Min. : 0.00
##	1st Qu.: 46	1st Qu.: 65	1st Qu.:195	1st Qu.: 0.00
##	Median :103	Median :154	Median :302	Median : 1.00
##	Mean :123	Mean :202	Mean :332	Mean : 3.08
##	3rd Qu.:162	3rd Qu.:302	3rd Qu.:486	3rd Qu.: 4.00
##	Max. :406	Max. :900	Max. :786	Max. :16.00

```

colnames(fish)[1] <- "basin"

# nested according to Pebas source information from Hoorn
# et al. 2010
fish$PebasLake <- 0
fish$PebasLake[grep("Pebas", fish$X23Ma)] <- 1
head(fish[, c("X23Ma", "PebasLake")], 20)

##          X23Ma PebasLake
## 1      PebasSudEst     1
## 2    GroupAtlantik     0
## 3    GroupAtlantik     0
## 4    GroupAtlantik     0
## 5    GroupAtlantik     0
## 6    GroupAtlantik     0
## 7    GroupAtlantik     0
## 8      PebasNord     1
## 9      PebasSud     1
## 10   GroupAtlantik     0
## 11   GroupAtlantik     0
## 12   GroupAtlantik     0
## 13   PebasSudEst     1
## 14   PebasSudEst     1
## 15   PebasNordEst     1
## 16   PebasSudEst3     1
## 17   GroupAtlantik     0
## 18   PebasSystem     1
## 19   PebasSudEst     1
## 20   PebasSudEst3     1

head(fish)

##      basin WaterColor Area_km Chut_hydrofall Chut_grand NetworkDensity
## 1    Abuna     White   31970            0            0        0.165
## 2 Amazon1     White   15948            0            0        0.135
## 3 Amazon2     White   45757            0            0        0.082
## 4 Amazon5     White   14602            0            0        0.087
## 5 Amazon8     White   17154            2            0        0.151
## 6 Amazon9     White   6190             0            0        0.103
##      Elevation_min Elevation_max Elevation_range Elevation_mean Elevation_std
## 1            95         301           206          172            31
## 2              0         377           379          41            57
## 3              0         818           820          154           166
## 4              0         183           185          26            34
## 5              3         214           211          64            46
## 6              3         152           149          54            37
##      Elevation_basinSup1000m Elevation_basinSup1500m NbEspeceEnveloppe
## 1                  0                   0                777
## 2                  0                   0                126
## 3                  0                   0                668

```

```

## 4          0          0        847
## 5          0          0        964
## 6          0          0        954
##   DistanceEmbouchure_km X23ma_cd      X23Ma X10Ma_cd      X10Ma
## 1           2884       A  PebasSudEst     A2  AcreSud
## 2             0       B GroupAtlantik    B2 AcreSystem
## 3            154       B GroupAtlantik    B2 AcreSystem
## 4            728       B GroupAtlantik    B2 AcreSystem
## 5           1153       B GroupAtlantik    B2 AcreSystem
## 6           1427       B GroupAtlantik    B2 AcreSystem
##   NbSites1km_0417 NbSites_0417 Diversity0417 Endem0417 PebasLake
## 1            31         50        229        0        1
## 2            46        109        154        2        0
## 3           207        465        486        1        0
## 4           136        248        425        1        0
## 5           141        299        566        2        0
## 6           167        433        532        0        0

```

```
fish[, c("basin", "PebasLake", "X23Ma", "X23ma_cd")]
```

	basin	PebasLake	X23Ma	X23ma_cd
## 1	Abuna	1	PebasSudEst	A
## 2	Amazon1	0	GroupAtlantik	B
## 3	Amazon2	0	GroupAtlantik	B
## 4	Amazon5	0	GroupAtlantik	B
## 5	Amazon8	0	GroupAtlantik	B
## 6	Amazon9	0	GroupAtlantik	B
## 7	Andira1	0	GroupAtlantik	B
## 8	Apaporis	1	PebasNord	C
## 9	Apurimac1	1	PebasSud	D
## 10	Arinos	0	GroupAtlantik	B
## 11	Aripuana	0	GroupAtlantik	B
## 12	Bacaja	0	GroupAtlantik	B
## 13	Beni	1	PebasSudEst	A
## 14	Blanco_Baures	1	PebasSudEst	A
## 15	Branco	1	PebasNordEst	E
## 16	Candeias	1	PebasSudEst3	F
## 17	Canuma	0	GroupAtlantik	B
## 18	Caqueta1	1	PebasSystem	G
## 19	Chapare	1	PebasSudEst	A
## 20	Coari	1	PebasSudEst3	F
## 21	Curaray	1	PebasSystem	G
## 22	Curua	0	AffluentsNord	H
## 23	Curua_Una	0	GroupAtlantik	B
## 24	Demini	1	PebasNordEst	E
## 25	Fresco	0	GroupAtlantik2	I
## 26	Grande	1	PebasSudEst	A
## 27	Guapore	1	PebasSudEst	A
## 28	Huallaga	1	PebasOuest3	J
## 29	Iriri	0	GroupAtlantik	B
## 30	Isiboro	1	PebasSudEst	A
## 31	Itonamas	1	PebasSudEst	A

## 32	Jamanxim	0	GroupAtlantik	B
## 33	Japura1	1	PebasSystem	G
## 34	Jari	0	AffluentsNord	H
## 35	Jatapu	0	AffluentsNord	H
## 36	Javary1	1	PebasSystem	G
## 37	Jiparana	0	GroupAtlantik	B
## 38	Jurua1	1	PebasSudEst3	F
## 39	Juruena1	0	GroupAtlantik	B
## 40	Jutai	1	PebasSudEst3	F
## 41	Luna	0	GroupAtlantik	B
## 42	Madera1	0	GroupAtlantik	B
## 43	Madera2	0	GroupAtlantik	B
## 44	Madera3	0	GroupAtlantik	B
## 45	Madera4	1	PebasSudEst3	F
## 46	Madera6	1	PebasSudEst	A
## 47	Madre_Dios	1	PebasSudEst	A
## 48	Mamore1	1	PebasSudEst	A
## 49	Mamore2	1	PebasSudEst	A
## 50	Mantaro	1	PebasSud	D
## 51	Maranon1	1	PebasSystem	G
## 52	Maranon3	1	PebasSystem	G
## 53	Maranon5	1	PebasQuest	K
## 54	Marmelos	0	GroupAtlantik	B
## 55	Maues	0	GroupAtlantik	B
## 56	Napo1	1	PebasSystem	G
## 57	Napo2	1	PebasSystem	G
## 58	Negro1	0	GroupAtlantik	B
## 59	Negro2	1	PebasNordEst	E
## 60	Negro3	1	PebasNordEst2	L
## 61	Negro5	1	PebasNordEst2	L
## 62	Nhamunda	0	AffluentsNord	H
## 63	Orthon	1	PebasSudEst	A
## 64	Pachitea	1	PebasSystem	G
## 65	Paru_Este	0	AffluentsNord	H
## 66	Pastaza	1	PebasSystem	G
## 67	Purus1	1	PebasSudEst3	F
## 68	Purus2	1	PebasSudEst3	F
## 69	Purus3	1	PebasSudEst2	M
## 70	Putumayo	1	PebasSystem	G
## 71	Roosevelt	0	GroupAtlantik	B
## 72	Santiago	1	PebasQuest2	N
## 73	Solimoes1	0	GroupAtlantik	B
## 74	Solimoes2	0	GroupAtlantik	B
## 75	Solimoes3	1	PebasNordEst	E
## 76	Solimoes4	1	PebasNordEst	E
## 77	Solimoes7	1	PebasSystem	G
## 78	Solimoes8	1	PebasSystem	G
## 79	Solimoes9	1	PebasSystem	G
## 80	Tapajos1	0	GroupAtlantik	B
## 81	Tapajos2	0	GroupAtlantik	B
## 82	Tapaua	1	PebasSudEst3	F
## 83	Tefe	1	PebasSudEst3	F

```

## 84    Teles_Pires      0  GroupAtlantik      B
## 85      Tigre          1  PebasSystem       G
## 86    Trombetas1      0  AffluentsNord     H
## 87    Uatumai1        0  GroupAtlantik      B
## 88      Uaupes         1  PebasNord        C
## 89    Ucayali1         1  PebasSystem       G
## 90    Ucayali2         1  PebasSystem       G
## 91      Unini          1  PebasNordEst     E
## 92    Urubamba         1  PebasSud        D
## 93    Xingu1           0  GroupAtlantik      B
## 94    Xingu3           0  GroupAtlantik      B
## 95    Xingu4           0  GroupAtlantik2     I
## 96    Xingu6           0  GroupAtlantik2     I
## 97    Yacuma           1  PebasSudEst      A

fish$PebasConec <- "Other"
fish$PebasConec[grep("Pebas", fish$X23Ma)] <- "Pebas"

# creating new variable based on past connection
coneclinks <- data.frame(basin = fish$basin, ma23 = as.numeric(factor(fish$X23Ma)),
                           ma10 = as.numeric(factor(fish$X10Ma)), Pebas = fish$PebasConec)

# code expressing Basin Links to Pebas basin
matPebas23Mya <- matrix(0, ncol = 97, nrow = 97, dimnames = list(fish$basin,
                                                               fish$basin))
matPebas23Mya[1:10, 1:10]

```

	Abuna	Amazon1	Amazon2	Amazon5	Amazon8	Amazon9	Andira1	Apaporis
## Abuna	0	0	0	0	0	0	0	0
## Amazon1	0	0	0	0	0	0	0	0
## Amazon2	0	0	0	0	0	0	0	0
## Amazon5	0	0	0	0	0	0	0	0
## Amazon8	0	0	0	0	0	0	0	0
## Amazon9	0	0	0	0	0	0	0	0
## Andira1	0	0	0	0	0	0	0	0
## Apaporis	0	0	0	0	0	0	0	0
## Apurimac1	0	0	0	0	0	0	0	0
## Arinos	0	0	0	0	0	0	0	0
##	Apurimac1	Arinos						
## Abuna	0	0						
## Amazon1	0	0						
## Amazon2	0	0						
## Amazon5	0	0						
## Amazon8	0	0						
## Amazon9	0	0						
## Andira1	0	0						
## Apaporis	0	0						
## Apurimac1	0	0						
## Arinos	0	0						

```

# matPebas23Mya['Abuna', 'Abuna']
for (i in unique(fish$basin)) {
  for (k in unique(fish$basin)) {
    # i='Abuna';k='Apurimac1' #Pebas-Pebas
    # i='Abuna';k='Teles_Pires' #Pebas-NonPebas
    # i='Xingu3';k='Teles_Pires' ##NonPebas-NonPebas
    II <- as.vector(conec[conec$basin == i, "Pebas"])
    KK <- as.vector(conec[conec$basin == k, "Pebas"])
    if ((II == KK) & (II == "Pebas")) {
      matPebas23Mya[i, k] <- 1 #Pebas-Pebas
    } else {
      if ((II == KK) & (II != "Pebas")) {
        matPebas23Mya[i, k] <- 0 #NonPebas-NonPebas
      } else {
        matPebas23Mya[i, k] <- 0 #Pebas-NonPebas
      }
    }
  }
}

# connectivity at 23ma
mat23 <- matrix(0, ncol = 97, nrow = 97, dimnames = list(fish$basin,
  fish$basin))
mat23[1:10, 1:10]

```

	Abuna	Amazon1	Amazon2	Amazon5	Amazon8	Amazon9	Andira1	Apaporis
## Abuna	0	0	0	0	0	0	0	0
## Amazon1	0	0	0	0	0	0	0	0
## Amazon2	0	0	0	0	0	0	0	0
## Amazon5	0	0	0	0	0	0	0	0
## Amazon8	0	0	0	0	0	0	0	0
## Amazon9	0	0	0	0	0	0	0	0
## Andira1	0	0	0	0	0	0	0	0
## Apaporis	0	0	0	0	0	0	0	0
## Apurimac1	0	0	0	0	0	0	0	0
## Arinos	0	0	0	0	0	0	0	0
##								
	Apurimac1	Arinos						
## Abuna	0	0						
## Amazon1	0	0						
## Amazon2	0	0						
## Amazon5	0	0						
## Amazon8	0	0						
## Amazon9	0	0						
## Andira1	0	0						
## Apaporis	0	0						
## Apurimac1	0	0						
## Arinos	0	0						

```

lapply(1:length(unique(conec$ma23)), function(x) {
  b <- conec[conec$ma23 == x, "basin"]
  mat23[b, b] <- 1
})

## [[1]]
## [1] 1
##
## [[2]]
## [1] 1
##
## [[3]]
## [1] 1
##
## [[4]]
## [1] 1
##
## [[5]]
## [1] 1
##
## [[6]]
## [1] 1
##
## [[7]]
## [1] 1
##
## [[8]]
## [1] 1
##
## [[9]]
## [1] 1
##
## [[10]]
## [1] 1
##
## [[11]]
## [1] 1
##
## [[12]]
## [1] 1
##
## [[13]]
## [1] 1
##
## [[14]]
## [1] 1

# connectivity at 10ma
mat10 <- matrix(0, ncol = 97, nrow = 97, dimnames = list(fish$basin,
  fish$basin))
mat10[1:10, 1:10]

```

```

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8 Amazon9 Andira1 Apaporis
## Abuna      0      0      0      0      0      0      0      0
## Amazon1    0      0      0      0      0      0      0      0
## Amazon2    0      0      0      0      0      0      0      0
## Amazon5    0      0      0      0      0      0      0      0
## Amazon8    0      0      0      0      0      0      0      0
## Amazon9    0      0      0      0      0      0      0      0
## Andira1    0      0      0      0      0      0      0      0
## Apaporis   0      0      0      0      0      0      0      0
## Apurimac1  0      0      0      0      0      0      0      0
## Arinos     0      0      0      0      0      0      0      0

##          Apurimac1 Arinos
## Abuna      0      0
## Amazon1    0      0
## Amazon2    0      0
## Amazon5    0      0
## Amazon8    0      0
## Amazon9    0      0
## Andira1    0      0
## Apaporis   0      0
## Apurimac1  0      0
## Arinos     0      0

```

```

lapply(1:length(unique(conec$ma10)), function(x) {
  b <- conecl[conecl$ma10 == x, "basin"]
  mat10[b, b] <- 1
})

```

```

## [[1]]
## [1] 1
##
## [[2]]
## [1] 1
##
## [[3]]
## [1] 1
##
## [[4]]
## [1] 1
##
## [[5]]
## [1] 1
##
## [[6]]
## [1] 1
##
## [[7]]
## [1] 1
##
## [[8]]
## [1] 1
##

```

```

## [[9]]
## [1] 1
##
## [[10]]
## [1] 1
##
## [[11]]
## [1] 1
##
## [[12]]
## [1] 1
##
## [[13]]
## [1] 1
##
## [[14]]
## [1] 1
##
## [[15]]
## [1] 1

# connectivity at Oma
library(igraph)

## 
## Attaching package: 'igraph'

## The following objects are masked from 'package:ape':
## 
##     degree, edges, mst, ring

## The following object is masked from 'package:vegan':
## 
##     diversity

## The following object is masked from 'package:permute':
## 
##     permute

## The following objects are masked from 'package:stats':
## 
##     decompose, spectrum

## The following object is masked from 'package:base':
## 
##     union

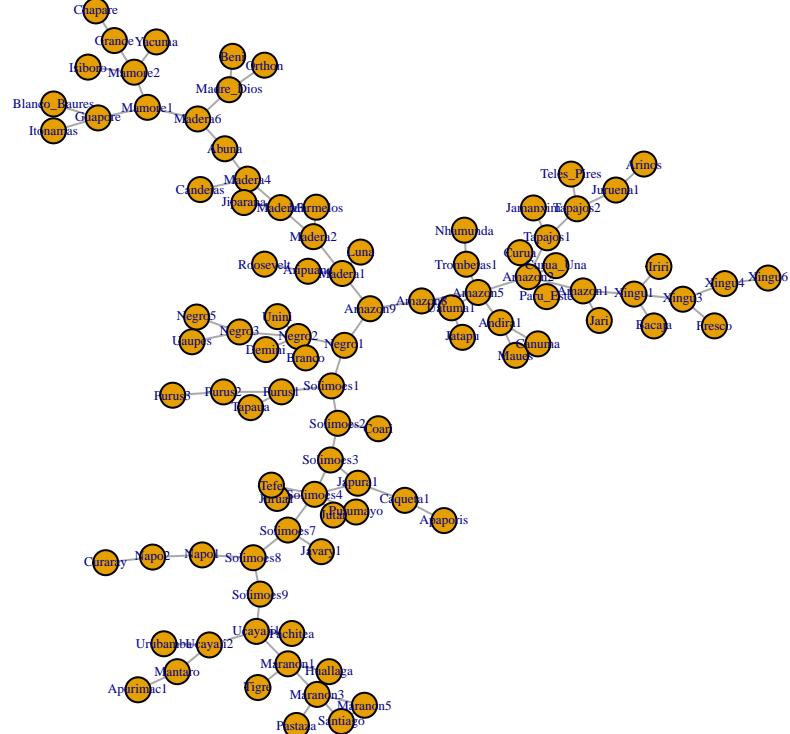
```

```

g <- graph.formula(Amazon1 -> Jari, Amazon1 -> Xingu1, Amazon1 ->
  Amazon2, Bacaja -> Xingu1, Xingu3 -> Xingu1, Iriri -> Xingu1,
  Fresco -> Xingu3, Xingu4 -> Xingu3, Xingu4 -> Xingu6, Amazon2 ->
  Paru_Este, Amazon2 -> Curua_Una, Amazon2 -> Tapajos1, Amazon2 ->
  Curua, Amazon2 -> Amazon5, Trombetas1 -> Amazon5, Trombetas1 ->
  Nhamunda, Chapare -> Grande, Mamore2 -> Grande, Mamore2 ->
  Isiboro, Mamore2 -> Yacuma, Beni -> Madre_Dios, Orthon ->
  Madre_Dios, Madera6 -> Madre_Dios, Madera6 -> Mamore1,
  Mamore2 -> Mamore1, Guapore -> Mamore1, Guapore -> Itonamas,
  Guapore -> Blanco_Baures, Tapajos1 -> Jamanxim, Tapajos1 ->
  Tapajos2, Teles_Pires -> Tapajos2, Juruena1 -> Tapajos2,
  Juruena1 -> Arinos, Negro3 -> Negro5, Negro3 -> Uaupes, Negro3 ->
  Negro2, Demini -> Negro2, Branco -> Negro2, Unini -> Negro2,
  Negro1 -> Negro2, Jatapu -> Uatumal1, Amazon5 -> Uatumal1, Amazon8 ->
  Uatumal1, Curaray -> Napo2, Napo1 -> Napo2, Apaporis -> Caqueta1,
  Japura1 -> Caqueta1, Abuna -> Madera6, Abuna -> Madera4, Purus2 ->
  Purus3, Purus2 -> Purus1, Candeias -> Madera4, Jiparana ->
  Madera4, Madera3 -> Madera4, Jiparana -> Madera3, Marmelos ->
  Madera3, Madera2 -> Madera3, Madera2 -> Marmelos, Maranon5 ->
  Santiago, Maranon5 -> Maranon3, Maranon3 -> Santiago, Maranon3 ->
  Pastaza, Maranon3 -> Huallaga, Maranon3 -> Maranon1, Huallaga ->
  Maranon1, Apurimac1 -> Mantaro -> Ucayali2, Urubamba ->
  Ucayali2, Pachitea -> Ucayali1 -> Ucayali2, Maranon1 ->
  Ucayali1 -> Solimoes9, Maranon1 -> Tigre, Napo1 -> Solimoes8 ->
  Solimoes9, Javary1 -> Solimoes7 -> Solimoes8, Putumayo ->
  Solimoes4 -> Solimoes7, Jutai -> Solimoes4, Jurua1 -> Solimoes4,
  Tefe -> Solimoes4, Solimoes3 -> Solimoes4 -> Japura1, Aripuana ->
  Roosevelt, Amazon5 -> Andira1, Amazon5 -> Amazon8, Andira1 ->
  Maues, Andira1 -> Canuma, Amazon8 -> Amazon9, Madera1 ->
  Amazon9, Madera1 -> Luna, Aripuana -> Madera1 -> Madera2,
  Amazon9 -> Negro1 -> Solimoes1, Purus1 -> Solimoes1 -> Solimoes2,
  Coari -> Solimoes2 -> Solimoes3, Japura1 -> Solimoes3, Tapaua ->
  Purus1 -> Purus2)

V(g)$label.cex = 0.4
plot(g, layout = layout_with_kk, vertex.size = 7)

```



```
m <- as_adj(g)
rm(g)
mat0 <- as.matrix(m)
rm(m)
mat0 <- mat0[sort(rownames(mat0)), sort(colnames(mat0))]
```

```
## [1] TRUE
```

```
# do they have the same structure?  
all.equal(dim(mat0), dim(mat10))
```

```
## [1] TRUE
```

```

all.equal(dim(mat0), dim(matPebas23Mya))

## [1] TRUE

# row names
colnames(mat0)[!colnames(mat0) %in% colnames(mat10)]


## character(0)

colnames(mat10)[!colnames(mat10) %in% colnames(mat0)]


## character(0)

all.equal(rownames(mat0), rownames(mat10))

## [1] TRUE

all.equal(rownames(mat0), rownames(mat23))

## [1] TRUE

all.equal(rownames(mat0), rownames(matPebas23Mya))

## [1] TRUE

# col names
all.equal(colnames(mat0), colnames(mat10))

## [1] TRUE

all.equal(colnames(mat0), colnames(mat23))

## [1] TRUE

all.equal(colnames(mat0), colnames(matPebas23Mya))

## [1] TRUE

# This index is different from the one at the Richness
# analysis
D_conec <- as.dist(mat0 + mat10 + mat23)

rm(mat10, mat23)

```

3 Preparing environmental data

We have two measures of habitat diversity. The first is the surface of the sub-drainage basins, and the second is a direct measure of habitat diversity derived from the relative cover of 14 habitat types found in each sub-drainage. It is widely accepted that, compared to small ones, larger habitats/islands can support larger populations, have higher habitat diversity, and have more barriers to dispersal, thus increasing speciation probability, decreasing extinction probability and supporting more species richness (Kisel & Barraclough 2010; Brown 2013).

3.1 Calculating *Habitat diversity*

We extract vegetation maps from GIS layers² and calculate the relative cover of each vegetation type in each sub-drainage. With the matrix of relative cover per sub-drainages, we applied the *Shannon diversity index* to calculate a measure of habitat diversity based on the vegetation cover. This index is valuable because it generates continuous positive values that can be directly incorporated in statistical analysis³.

```
habitat = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/GlobLandCover.csv",
  sep = ""), sep = ";", header = T)

# setting rownames
rownames(habitat) <- habitat$basin
dim(habitat)

## [1] 97 23

head(habitat)

##          basin X10 X11 X12 X30 X40 X50 X60 X61 X62 X100 X110 X120 X130 X150
## Abuna      Abuna 1.5  0  0 0.4 2.8  95  0  0  0  0.2  0.1  0.1  0.0  0
## Amazon1   Amazon1 0.2  0  0 0.0 0.5  63  0  0  0  1.1  0.8  0.4  0.5  0
## Amazon2   Amazon2 0.8  0  0 0.3 4.3  73  0  0  0  1.7  0.3  0.5  0.1  0
## Amazon5   Amazon5 0.9  0  0 0.5 3.7  41  0  0  0  1.6  0.0  0.3  0.0  0
## Amazon8   Amazon8 0.4  0  0 0.4 1.5  77  0  0  0  0.1  0.0  0.0  0.0  0
## Amazon9   Amazon9 1.2  0  0 0.9 2.6  68  0  0  0  0.5  0.0  0.2  0.0  0
##          X153 X160 X170 X180 X190 X200 X210 X220
## Abuna      0  0.1  0.0  0.0  0.1  0.0  0.0  0
## Amazon1   0 23.0  0.2  0.5  0.0  0.0  9.7  0
## Amazon2   0  3.9  0.0  3.4  0.1  0.0 11.5  0
## Amazon5   0  8.2  0.0 12.6  0.3  0.6 29.8  0
## Amazon8   0  7.1  0.0  2.1  0.2  0.0 11.1  0
## Amazon9   0  6.4  0.0  2.3  0.3  0.0 17.1  0

# does the sum of categories sum up 1?
summary(rowSums(habitat[, -1])) #ok

##    Min. 1st Qu. Median    Mean 3rd Qu.    Max.
## 99.7    99.9  100.0  100.0  100.0  100.2
```

²see http://worldgrids.org/doku.php?id=wiki:layers#land_cover_and_land_use for a list of available layers

³The higher the diversity values, the higher the habitat diversity in the sub-drainage

```

# creating diversity of habitat
head(habitat[, -1])

##          X10 X11 X12 X30 X40 X50 X60 X61 X62 X100 X110 X120 X130 X150 X153 X160
## Abuna    1.5  0   0  0.4 2.8 95   0   0   0   0.2  0.1  0.1  0.0   0   0   0.1
## Amazon1  0.2  0   0  0.0 0.5 63   0   0   0   1.1  0.8  0.4  0.5   0   0 23.0
## Amazon2  0.8  0   0  0.3 4.3 73   0   0   0   1.7  0.3  0.5  0.1   0   0  3.9
## Amazon5  0.9  0   0  0.5 3.7 41   0   0   0   1.6  0.0  0.3  0.0   0   0  8.2
## Amazon8  0.4  0   0  0.4 1.5 77   0   0   0   0.1  0.0  0.0  0.0   0   0  7.1
## Amazon9  1.2  0   0  0.9 2.6 68   0   0   0   0.5  0.0  0.2  0.0   0   0  6.4
##          X170 X180 X190 X200 X210 X220
## Abuna    0.0  0.0  0.1  0.0  0.0   0
## Amazon1  0.2  0.5  0.0  0.0  9.7   0
## Amazon2  0.0  3.4  0.1  0.0 11.5   0
## Amazon5  0.0 12.6  0.3  0.6 29.8   0
## Amazon8  0.0  2.1  0.2  0.0 11.1   0
## Amazon9  0.0  2.3  0.3  0.0 17.1   0

habitat2 <- habitat
habitat$ShannonDiv <- vegan::diversity(habitat[, -1], index = "shannon")
summary(habitat$ShannonDiv)

##      Min. 1st Qu. Median     Mean 3rd Qu.      Max.
## 0.0603  0.2697  0.5204  0.6135  0.9333  2.0395

# diversity of habitat
habitat <- habitat[, c("basin", "ShannonDiv")]
head(habitat)

##          basin ShannonDiv
## Abuna      Abuna      0.28
## Amazon1   Amazon1    1.07
## Amazon2   Amazon2    1.04
## Amazon5   Amazon5    1.52
## Amazon8   Amazon8    0.84
## Amazon9   Amazon9    1.07

# merging to fish data set
fish <- merge(fish, habitat, by = "basin")

# #data control comp <- fish

write.csv(habitat, file = "comparingResults/habitat.csv")
rm(habitat)

```

3.2 Soil diversity

We used the soil classification (from Soil and Terrain Database for Latin America and the Caribbean⁴) of the whole drainage basin to calculate a index of soil (i.e., geological) diversity. For such, we estimated the proportional cover of each soil type in each subdrainage and used the proportions to calculate the Shannon diversity index (exactly as Habitat diversity).

```
soil = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Soil_2017_04/",
  "SoilLithoData/SoterlacLithology.csv", sep = ""), sep = ";",
  header = T)
soil[is.na(soil)] <- 0
rownames(soil) <- soil$basin
head(soil)
```

	basin	AcidIgnRock	AcidMetRock	Andesite	Basalt	BasicMetamRock	
## Abuna	Abuna	0	24.01	0	0	0	
## Amazon1	Amazon1	0	7.82	0	0	0	
## Amazon2	Amazon2	0	26.64	0	0	0	
## Amazon5	Amazon5	0	0.00	0	0	0	
## Amazon8	Amazon8	0	0.22	0	0	0	
## Amazon9	Amazon9	0	0.00	0	0	0	
	ClasticSedimRock	Colluvial	Conglomerate	Eolean	Fluvial	GneissFerr	
## Abuna		45.5	0	0	1.2	0	
## Amazon1		24.8	0	0	65.0	0	
## Amazon2		32.5	0	0	26.9	0	
## Amazon5		3.6	0	0	65.9	0	
## Amazon8		4.6	0	0	23.4	0	
## Amazon9		0.0	0	0	41.9	0	
	GneissMigm	Granite	GranoDiorite	Marl	Pyroclastic	Quartzite	Sandstone
## Abuna		13	0	0	0	0	16.2
## Amazon1		0	0	0	0	0	2.4
## Amazon2		0	0	0	0	0	14.0
## Amazon5		0	0	0	0	0	30.5
## Amazon8		0	0	0	0	0	71.8
## Amazon9		0	0	0	0	0	58.1
	Schist	Shale	Siltstone	Slate			
## Abuna		0	0	0			
## Amazon1		0	0	0			
## Amazon2		0	0	0			
## Amazon5		0	0	0			
## Amazon8		0	0	0			
## Amazon9		0	0	0			

```
# checking whether the categories sum up 1
sort(rowSums(soil[, -1]), decreasing = T) #perfect
```

	Jari	Amazon1	Amazon2	Paru_Este	Xingu1
##	100	100	100	100	100
##	Curua	Bacaja	Trombetas1	Curua_Una	Xingu3

⁴http://geonode.isric.org/layers/geonode:soter_lac_map_unit

```

##      100      100      100      100      100
## Amazon5 Fresco Nhamunda Tapajos1 Branco
##      100      100      100      100      100
## Jatapu Iriri Andira1 Uatumal Xingu4
##      100      100      100      100      100
## Maues Amazon8 Jamanxim Negro1 Demini
##      100      100      100      100      100
## Madera1 Tapajos2 Amazon9 Negro2 Solimoes1
##      100      100      100      100      100
## Canuma Luna Unini Xingu6 Teles_Pires
##      100      100      100      100      100
## Solimoes2 Negro5 Negro3 Madera2 Solimoes3
##      100      100      100      100      100
## Aripuana Arinos Purus1 Marmelos Madera3
##      100      100      100      100      100
## Coari Juruenai Solimoes4 Tefe Roosevelt
##      100      100      100      100      100
## Japurá1 Uaupes Purus2 Jiparana Tapaua
##      100      100      100      100      100
## Candeias Jutai Solimoes7 Madera4 Caqueta1
##      100      100      100      100      100
## Guapore Apaporis Madera6 Putumayo Solimoes8
##      100      100      100      100      100
## Mamore1 Blanco_Baures Abuna Purus3 Napo1
##      100      100      100      100      100
## Javary1 Jurua1 Solimoes9 Itonamas Napo2
##      100      100      100      100      100
## Mamore2 Curaray Tigre Orthon Yacuma
##      100      100      100      100      100
## Chapare Maranon1 Beni Grande Pastaza
##      100      100      100      100      100
## Madre_Dios Ucayali1 Maranon3 Isiboro Santiago
##      100      100      100      100      100
## Urubamba Huallaga Ucayali2 Pachitea Maranon5
##      100      100      100      100      100
## Apurimac1 Mantaro
##      100      100

```

```

Soil2 <- soil
soil$SoilDiv <- vegan::diversity(soil[, -1], index = "shannon")
summary(soil$SoilDiv)

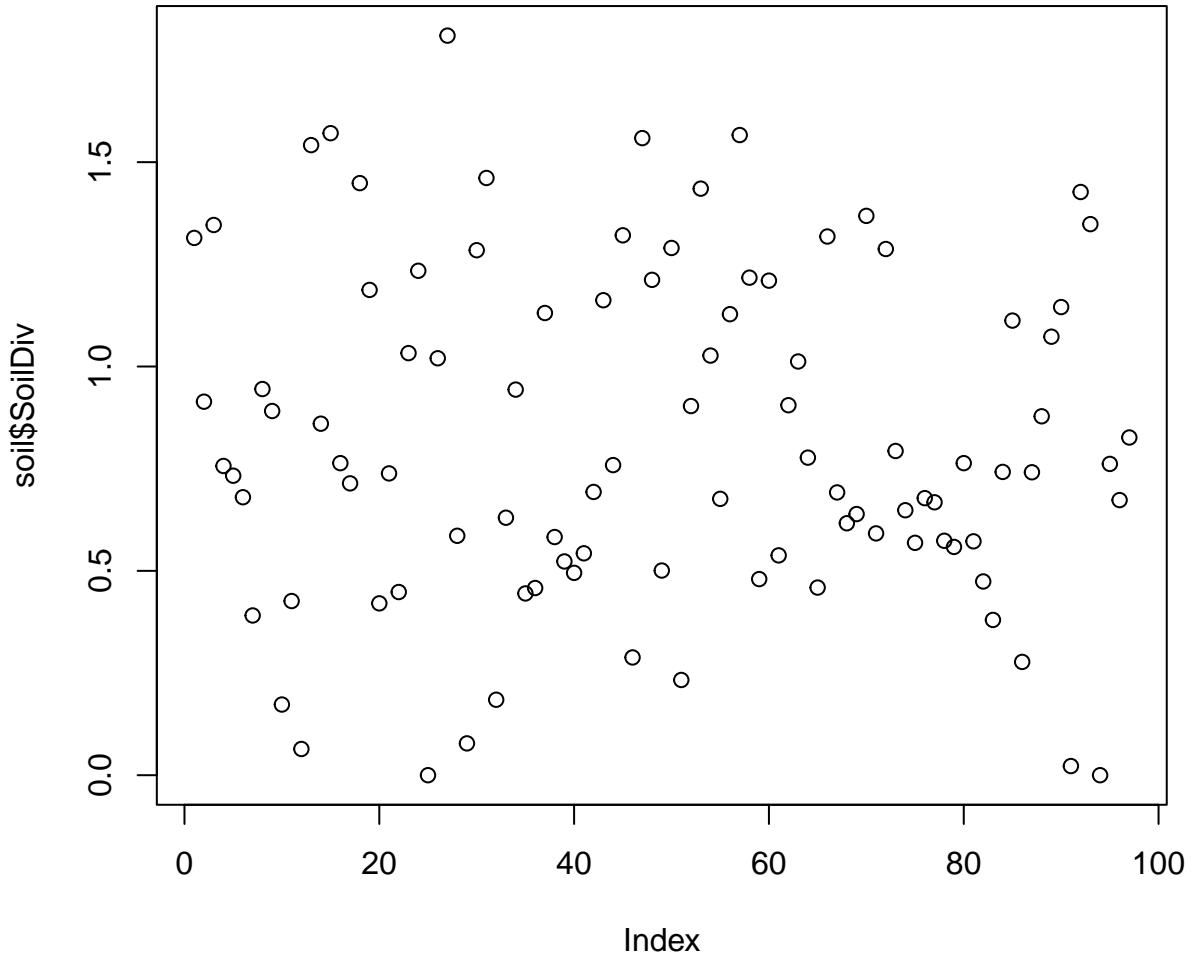
```

```

##      Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.000 0.543 0.757 0.818 1.145 1.810

```

```
plot(soil$SoilDiv)
```



```
# diversity of soil geology
soil <- soil[, c("basin", "SoilDiv")]

# merging to fish data set
fish <- merge(fish, soil, by = "basin")

write.csv(soil, file = "comparingResults/soil.csv")
rm(soil)
```

3.3 Global terrain slope (potential waterfalls)

We used the an estimation of slope between pixel (from Global Terrain Slope and Aspect Data⁵) of the whole drainage basin to calculate the percentage of area within each cathegory of slope (0–0.5%, 0.5–2%, 2–5%, 5–8%, 8–16%, 16–30%, 30–45%, and > 45%). As the data set is presente as a matrix of surface cover (summing up 1 between cathegories), we opted to sum up classes 16–30%, 30–45%, and > 45% to compose a new variable.

```
slope = read.csv(paste("StatistiquesSubBasinAmazon_02052016/",
  "Slope_Waterfall/Slope.csv", sep = ""), sep = ";", header = T)
slope[is.na(slope)] <- 0
rownames(slope) <- slope$basin
head(slope)

##          basin C0_05 C05_2 C2_5 C5_10 C10_15 C15_30 C30_45 CSup45
## Abuna      Abuna  0.00   19    79    2.4   0.00   0.00     0     0
## Amazon1   Amazon1  0.36   57    25   16.5   0.00   0.67     0     0
## Amazon2   Amazon2  1.12   30    39    9.5   0.55  20.23     0     0
## Amazon5   Amazon5  1.60   58    30   10.4   0.00   0.00     0     0
## Amazon8   Amazon8  0.00   17    39   44.0   0.00   0.00     0     0
## Amazon9   Amazon9  1.38   24    21   53.6   0.00   0.00     0     0

# checking whether the categories sum up 1
table(round(rowSums(slope[, -1]), 1)) #perfect

## 
## 98.8 98.9 99.2 99.4 99.5 99.6 99.7 99.8 99.9 100
##    2     1     1     1     2     5     3     9    13    60

# defyning classes
slope$CSup15 <- rowSums(slope[, c("C15_30", "C30_45", "CSup45")])
slope$CBelow15 <- rowSums(slope[, c("C0_05", "C05_2", "C2_5",
  "C5_10", "C10_15")])

# diversity of soil geology
slope <- slope[, c("basin", "CSup15", "CBelow15")]

# merging to fish data set
fish <- merge(fish, slope, by = "basin")

write.csv(slope, file = "comparingResults/slope.csv")

rm(slope)
```

⁵<http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/global-terrain-slope.html>

3.4 Current climate

Charging all data set.

```
Cclimate = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/climatology_present.csv",
  sep = ""), sep = ", ", header = T)

dim(Cclimate)

## [1] 97 332

# names(Cclimate)
Cclimate[1:10, 1:7]

##      X    BvNiv2 aet_mn1 aet_std1 aet_cv1 aet_mn2 aet_std2
## 1 22   Abuna     152      2.0     1.3     135     1.33
## 2 45 Amazon1    115     10.4     9.0     122     3.32
## 3 46 Amazon2     77      7.9    10.3     100     9.77
## 4 47 Amazon5    114      8.2     7.1     123     2.62
## 5 81 Amazon8    129      5.5     4.2     121     1.70
## 6 82 Amazon9    131      4.2     3.2     122     0.94
## 7 51 Andira1    131      4.9     3.7     122     1.83
## 8 29 Apaporis   129     12.3     9.5     117    19.78
## 9 97 Apurimac1   69     14.5    21.1      77    12.80
## 10 91 Arinos     165      4.0     2.4     147     3.15

# setting row names
rownames(Cclimate) <- Cclimate$BvNiv2

# is everything ok? NAs, NaN values, Extreme values?
# summary(Cclimate) OK
```

Now, let's select the most important variables based on (Leprieur *et al.* 2011). They showed that the following variables are important.

```
# separating variables Temperature: I must divide by 10 to
# get the right temperature Mean temperature of the coldest
# month
tempMin <- Cclimate$bio_mn6/10
# Mean temperature of the warmest month
tempMax <- Cclimate$bio_mn5/10
# Annual mean temperature
tempMean <- Cclimate$bio_mn1/10
# Coefficient of variation of mean monthly temperature
tempCV <- Cclimate$bio_mn4/1000

temp2 <- data.frame(tempMin, tempMax, tempMean, tempCV)
rm(tempMin, tempMax, tempMean, tempCV)
temp2$basin <- Cclimate$BvNiv2
head(temp2)
```

```

##   tempMin tempMax tempMean tempCV   basin
## 1     18     33      26  0.72 Abuna
## 2     22     33      27  0.52 Amazon1
## 3     21     32      26  0.61 Amazon2
## 4     22     33      27  0.58 Amazon5
## 5     22     33      27  0.44 Amazon8
## 6     22     33      27  0.44 Amazon9

summary(temp2)

##      tempMin      tempMax      tempMean      tempCV
## Min.   :-4.0   Min.   :15.9   Min.   : 7.66   Min.   :0.294
## 1st Qu.:16.3   1st Qu.:31.4   1st Qu.:25.07   1st Qu.:0.422
## Median :19.6   Median :32.3   Median :25.76   Median :0.487
## Mean   :18.0   Mean   :31.7   Mean   :24.87   Mean   :0.615
## 3rd Qu.:20.8   3rd Qu.:33.0   3rd Qu.:26.40   3rd Qu.:0.619
## Max.   :22.4   Max.   :34.8   Max.   :27.13   Max.   :2.120
##      basin
## Length:97
## Class :character
## Mode  :character
## 
## 
##
```

```

# precipitation Precipitation of the driest month
precMin <- Cclimate$bio_mn14
# Precipitation of the wettest month
precMax <- Cclimate$bio_mn13
# Anual precipitation
precAnn <- Cclimate$bio_mn12
# Coefficient of variation of mean monthly precipitation
precCV <- Cclimate$bio_mn15
prec2 <- data.frame(precMin, precMax, precAnn, precCV)
rm(precMin, precMax, precAnn, precCV)  #
prec2$basin <- Cclimate$BvNiv2
head(prec2)
```

```

##      precMin precMax precAnn precCV   basin
## 1     20     270    1785     61 Abuna
## 2     59     340    2453     47 Amazon1
## 3     36     305    1836     61 Amazon2
## 4     45     334    2098     58 Amazon5
## 5     86     332    2311     44 Amazon8
## 6     86     296    2250     40 Amazon9
```

```

summary(prec2)

##      precMin      precMax      precAnn      precCV
## Min.   : 2.95   Min.   :147   Min.   : 837   Min.   :13.3
```

```

## 1st Qu.: 28.65   1st Qu.:285   1st Qu.:1963   1st Qu.:33.8
## Median : 62.03   Median :321   Median :2241   Median :50.8
## Mean    : 71.29   Mean    :308   Mean    :2204   Mean    :47.7
## 3rd Qu.:102.43   3rd Qu.:340   3rd Qu.:2514   3rd Qu.:61.2
## Max.    :205.73   Max.    :436   Max.    :3248   Max.    :86.0
##      basin
## Length:97
## Class :character
## Mode  :character
##
##
```

```

# solar radiation Mean Anual solar radiation
SolRad_mn <- Cclimate$solrad_mn13
# Coefficient of variation of anual solar radiation
SolRad_cv <- Cclimate$solrad_cv13
solRad <- data.frame(SolRad_mn, SolRad_cv)
rm(SolRad_mn, SolRad_cv)
solRad$basin <- Cclimate$BvNiv2
head(solRad)
```

```

##   SolRad_mn SolRad_cv   basin
## 1      15   0.1133 Abuna
## 2      15   0.0024 Amazon1
## 3      15   0.0138 Amazon2
## 4      15   0.0118 Amazon5
## 5      15   0.0216 Amazon8
## 6      15   0.0142 Amazon9
```

```
summary(solRad)
```

```

##   SolRad_mn     SolRad_cv       basin
## Min.   :14.2   Min.   :0.00237   Length:97
## 1st Qu.:14.6   1st Qu.:0.03243   Class  :character
## Median :14.7   Median :0.07995   Mode   :character
## Mean   :14.7   Mean   :0.14902
## 3rd Qu.:14.8   3rd Qu.:0.22932
## Max.   :14.8   Max.   :0.91393
```

```

# Actual evapotranspiration
vari <- c("aet_mn1", "aet_mn2", "aet_mn3", "aet_mn4", "aet_mn5",
         "aet_mn6", "aet_mn7", "aet_mn8", "aet_mn9", "aet_mn10", "aet_mn11",
         "aet_mn12")
aet <- Cclimate[, vari]
# Aet minimum
aetMin <- apply(aet, 1, function(x) min(x, na.rm = TRUE))
# Aet maximum
aetMax <- apply(aet, 1, function(x) max(x, na.rm = TRUE))
# Annual Aet
```

```

aetAnn <- rowSums(aet)
# Coefficient of variation of monthly aet
aetCV <- apply(aet, 1, function(x) sd(x, na.rm = TRUE)/mean(x,
  na.rm = TRUE))
aet2 <- data.frame(aetMin, aetMax, aetAnn, aetCV)
rm(aetMin, aetMax, aetAnn, aetCV)
aet2$basin <- Cclimate$BvNiv2
head(aet2)

```

```

##          aetMin aetMax aetAnn aetCV   basin
## Abuna      65     152    1309 0.291  Abuna
## Amazon1    85     138    1433 0.147 Amazon1
## Amazon2    61     134    1228 0.254 Amazon2
## Amazon5    71     136    1298 0.195 Amazon5
## Amazon8   100     134    1429 0.087 Amazon8
## Amazon9   105     135    1436 0.080 Amazon9

```

```
summary(aet2)
```

```

##      aetMin       aetMax       aetAnn       aetCV
## Min.   : 39.0   Min.   : 84.8   Min.   : 696   Min.   :0.0492
## 1st Qu.: 65.8   1st Qu.:136.9   1st Qu.:1310   1st Qu.:0.0858
## Median : 85.5   Median :146.0   Median :1404   Median :0.1627
## Mean   : 86.6   Mean   :143.1   Mean   :1396   Mean   :0.1803
## 3rd Qu.:111.1   3rd Qu.:152.1   3rd Qu.:1523   3rd Qu.:0.2559
## Max.   :126.9   Max.   :165.4   Max.   :1746   Max.   :0.4171
##      basin
## Length:97
## Class :character
## Mode  :character
##
## 
## 
## 
```

```
rm(aet, vari)
```

```

# Potential evapotranspiration
vari <- c("pet_mn1", "pet_mn2", "pet_mn3", "pet_mn4", "pet_mn5",
  "pet_mn6", "pet_mn7", "pet_mn8", "pet_mn9", "pet_mn10", "pet_mn11",
  "pet_mn12")
pet <- Cclimate[, vari]
# Pet minimum
petMin <- apply(pet, 1, function(x) min(x, na.rm = TRUE))
# Pet maximum
petMax <- apply(pet, 1, function(x) max(x, na.rm = TRUE))
# Annual Pet
petAnn <- rowSums(pet)
# Coefficient of variation of monthly pet
petCV <- apply(pet, 1, function(x) sd(x, na.rm = TRUE)/mean(x,

```

```

    na.rm = TRUE))
pet2 <- data.frame(petMin, petMax, petAnn, petCV)
rm(petMin, petMax, petAnn, petCV)
pet2$basin <- Cclimate$BvNiv2
head(pet2)

```

```

##          petMin petMax petAnn petCV basin
## Abuna      122     166   1751  0.101 Abuna
## Amazon1    122     159   1654  0.084 Amazon1
## Amazon2    119     157   1628  0.089 Amazon2
## Amazon5    120     161   1650  0.094 Amazon5
## Amazon8    121     153   1623  0.079 Amazon8
## Amazon9    120     151   1614  0.076 Amazon9

```

```
summary(pet2)
```

```

##          petMin        petMax       petAnn       petCV
##  Min.   : 79.3   Min.   :114   Min.   :1175   Min.   :0.0588
##  1st Qu.:115.5  1st Qu.:153  1st Qu.:1641  1st Qu.:0.0794
##  Median :119.5  Median :159  Median :1672  Median :0.0879
##  Mean   :118.6  Mean   :160  Mean   :1684  Mean   :0.0955
##  3rd Qu.:124.0  3rd Qu.:166  3rd Qu.:1737  3rd Qu.:0.0999
##  Max.   :139.8  Max.   :185  Max.   :1920  Max.   :0.2030

##      basin
##  Length:97
##  Class :character
##  Mode   :character

##
```

```
rm(pet, vari)
```

3.5 Current Runoff

We will calculate runoff as a mean value from two different models

```

# Runoff UNH
unh = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/ro_unh_cmp.csv",
  sep = ""), sep = ",", header = T)

# Mean Annual runoff
ro_unh <- unh[, c("BvNiv2", "rounhcmp_mn13")]
rm(unh)

# Runoff Quest
quest = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/ro_quest.csv",
  sep = ""), sep = ",", header = T)

```

```

# Mean Annual runoff
ro_quest <- data.frame(BvNiv2 = quest$BvNiv2, ro_quest = rowSums(quest[, 
  paste("roqst_mn", 1:12, sep = "")]))
rm(quest)

# merging and computing mean model values
ro <- merge(ro_unh, ro_quest, by = "BvNiv2")
rm(ro_unh, ro_quest)
colnames(ro)[1] <- "basin"

ro$ro_mean <- rowMeans(ro[, -1])
ro <- ro[, c("basin", "ro_mean")]

head(ro)

```

```

##      basin ro_mean
## 1    Abuna   689
## 2 Amazon1  1153
## 3 Amazon2   885
## 4 Amazon5  1345
## 5 Amazon8  1274
## 6 Amazon9  1339

```

3.6 Net Primary Productivity

Here we charge Net Primary Productivity. We will use the **mean**, **standard deviation**, and **coefficient of variation** of NPP in each basin between 2000-2006.

```

npp = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/npp_present.csv",
  sep = ""), sep = ",", header = T)

head(npp)

```

```

##      X     BvNiv2 MEAN STD  cv MEAN00 STD00 cv00 MEAN01 STD01 cv01 MEAN02 STD02
## 1  1      Jari  2684 191 7.1  2832   178  6.3  2628   131  5.0  2593   189
## 2  2     Curua  2386 126 5.3  2663   143  5.4  2448   113  4.6  2009   159
## 3  3 Paru_Este  2572 210 8.2  2737   229  8.4  2578   123  4.8  2443   183
## 4  4 Curua_Una  2534 116 4.6  2802   158  5.6  2577   138  5.4  2353   187
## 5  5      Tefe  2760 123 4.5  2770   220  7.9  3360   221  6.6  2089   304
## 6  6      Jutai  2619 131 5.0  3246   299  9.2  3149   250  8.0  2123   257
##      cv02 MEAN03 STD03 cv03 MEAN04 STD04 cv04 MEAN05 STD05 cv05 MEAN06 STD06 cv06
## 1   7.3   2645   193  7.3   2765    81  2.9   2588   435 16.8   2737   226  8.3
## 2   7.9   2570   182  7.1   2525    92  3.6   2166   133  6.1   2320   123  5.3
## 3   7.5   2647   250  9.4   2677   104  3.9   2326   454 19.5   2593   231  8.9
## 4   8.0   2562   132  5.2   2343   149  6.4   2524   180  7.1   2575   153  6.0
## 5  14.6   2647   282 10.6   3067   110  3.6   2551   138  5.4   2837   171  6.0
## 6  12.1   3018   109  3.6   2399   340 14.2   1847   276 14.9   2549   348 13.7

```

```
# setting variables
npp <- npp[, c("BvNiv2", "MEAN", "STD", "cv")]
colnames(npp) <- c("basin", "npp_mn", "npp_sd", "npp_cv")
rownames(npp) <- npp$basin
head(npp)
```

	basin	npp_mn	npp_sd	npp_cv
## Jari	Jari	2684	191	7.1
## Curua	Curua	2386	126	5.3
## Paru_Este	Paru_Este	2572	210	8.2
## Curua_Una	Curua_Una	2534	116	4.6
## Tefe	Tefe	2760	123	4.5
## Jutai	Jutai	2619	131	5.0

3.6.1 Combining all variables from current climate in a single *data.frame*

Then, we have to combine to whole data set. Column binding process has been done based on the subdrainage labels.

First, we standardize all variables so that they all have zero mean and unity standard deviation. The next step is to sepparate different variables related to energy, availability of water and temperature hypotheses. For each of these groups, we calculated **Distances Matrices** based on **Euclidean distance**.

```
# merging all data.frames
current_climate <- merge(temp2, prec2, by = "basin")
current_climate <- merge(current_climate, aet2, by = "basin")
current_climate <- merge(current_climate, pet2, by = "basin")
current_climate <- merge(current_climate, ro, by = "basin")
current_climate <- merge(current_climate, npp, by = "basin")
current_climate <- merge(current_climate, solRad, by = "basin")
head(current_climate)
```

	basin	tempMin	tempMax	tempMean	tempCV	precMin	precMax	precAnn	precCV	aetMin	
## 1	Abuna	18	33	26	0.72	20	270	1785	61	65	
## 2	Amazon1	22	33	27	0.52	59	340	2453	47	85	
## 3	Amazon2	21	32	26	0.61	36	305	1836	61	61	
## 4	Amazon5	22	33	27	0.58	45	334	2098	58	71	
## 5	Amazon8	22	33	27	0.44	86	332	2311	44	100	
## 6	Amazon9	22	33	27	0.44	86	296	2250	40	105	
	aetMax	aetAnn	aetCV	petMin	petMax	petAnn	petCV	ro_mean	npp_mn	npp_sd	npp_cv
## 1	152	1309	0.291	122	166	1751	0.101	689	1689	96	5.7
## 2	138	1433	0.147	122	159	1654	0.084	1153	2679	261	9.7
## 3	134	1228	0.254	119	157	1628	0.089	885	2568	211	8.2
## 4	136	1298	0.195	120	161	1650	0.094	1345	2256	181	8.0
## 5	134	1429	0.087	121	153	1623	0.079	1274	2512	124	4.9
## 6	135	1436	0.080	120	151	1614	0.076	1339	2556	88	3.4
	SolRad_mn	SolRad_cv									
## 1	15	0.1133									
## 2	15	0.0024									
## 3	15	0.0138									
## 4	15	0.0118									

```

## 5      15    0.0216
## 6      15    0.0142

# naming rows and excluding dummy variable from data.frame
rownames(current_climate) <- current_climate$basin
current_climate <- current_climate[, !colnames(current_climate) %in%
  c("basin")]
head(current_climate)

##          tempMin tempMax tempMean tempCV precMin precMax precAnn precCV aetMin
## Abuna       18     33     26   0.72      20     270    1785    61     65
## Amazon1     22     33     27   0.52      59     340    2453    47     85
## Amazon2     21     32     26   0.61      36     305    1836    61     61
## Amazon5     22     33     27   0.58      45     334    2098    58     71
## Amazon8     22     33     27   0.44      86     332    2311    44    100
## Amazon9     22     33     27   0.44      86     296    2250    40    105
##          aetMax aetAnn aetCV petMin petMax petAnn petCV ro_mean npp_mn npp_sd
## Abuna      152    1309  0.291     122    166    1751  0.101     689    1689    96
## Amazon1    138    1433  0.147     122    159    1654  0.084    1153    2679   261
## Amazon2    134    1228  0.254     119    157    1628  0.089     885    2568   211
## Amazon5    136    1298  0.195     120    161    1650  0.094    1345    2256   181
## Amazon8    134    1429  0.087     121    153    1623  0.079    1274    2512   124
## Amazon9    135    1436  0.080     120    151    1614  0.076    1339    2556   88
##          npp_cv SolRad_mn SolRad_cv
## Abuna      5.7      15    0.1133
## Amazon1    9.7      15    0.0024
## Amazon2    8.2      15    0.0138
## Amazon5    8.0      15    0.0118
## Amazon8    4.9      15    0.0216
## Amazon9    3.4      15    0.0142

write.csv(current_climate, file = "comparingResults/current_climate.csv")

# Standard function (centering and reducing; see first
# chunk)
current_climate_s <- data.frame(apply(current_climate, 2, Standard))
rownames(current_climate_s) <- rownames(current_climate)
colMeans(current_climate_s) #OK

##          tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
## -1.5e-16  1.8e-16 -1.4e-16 -2.6e-17 -4.0e-17  2.1e-16 -1.1e-16 -3.5e-17
##          aetMin aetMax aetAnn aetCV petMin petMax petAnn petCV
## 4.1e-17   2.9e-16 -1.8e-16  5.8e-17  2.6e-16  4.4e-16  1.4e-16 -6.2e-17
##          ro_mean npp_mn npp_sd npp_cv SolRad_mn SolRad_cv
## 4.8e-17  -2.1e-16  4.6e-18  3.5e-17  6.2e-17 -2.4e-18

apply(current_climate_s, 2, sd) #OK

##          tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
##          0.5       0.5       0.5       0.5       0.5       0.5       0.5       0.5
```

```

##    aetMin    aetMax    aetAnn    aetCV    petMin    petMax    petAnn    petCV
##    0.5      0.5      0.5      0.5      0.5      0.5      0.5      0.5
##    ro_mean   npp_mn   npp_sd   npp_cv  SolRad_mn SolRad_cv
##    0.5      0.5      0.5      0.5      0.5      0.5

```

```

# setting predictor' names
colnames(current_climate_s) <- paste(colnames(current_climate_s),
  ".s", sep = "")
current_climate_s$basin <- rownames(current_climate_s)

# naming rows and checking the final dataset
head(current_climate_s)

```

```

##      tempMin.s tempMax.s tempMean.s tempCV.s precMin.s precMax.s precAnn.s
## Abuna     0.05     0.30     0.21   0.1347    -0.49    -0.361   -0.419
## Amazon1   0.45     0.22     0.31   -0.1302    -0.12     0.292    0.249
## Amazon2   0.38     0.13     0.23   -0.0087    -0.34    -0.033   -0.367
## Amazon5   0.47     0.29     0.35   -0.0444    -0.24     0.235   -0.106
## Amazon8   0.44     0.17     0.33   -0.2384     0.14     0.218    0.106
## Amazon9   0.48     0.19     0.35   -0.2352     0.14    -0.117    0.045
##      precCV.s aetMin.s aetMax.s aetAnn.s aetCV.s petMin.s petMax.s petAnn.s
## Abuna     0.365    -0.462    0.33   -0.222    0.556    0.1758   0.249    0.30
## Amazon1  -0.010    -0.029   -0.18    0.092   -0.170    0.1485   -0.050   -0.14
## Amazon2   0.374    -0.539   -0.33   -0.426    0.373    0.0035   -0.128   -0.26
## Amazon5   0.304    -0.338   -0.24   -0.248    0.073    0.0813   0.057   -0.16
## Amazon8  -0.094    0.275   -0.30    0.082   -0.469    0.1141   -0.265   -0.28
## Amazon9  -0.224    0.391   -0.30    0.101   -0.505    0.0624   -0.385   -0.32
##      petCV.s ro_mean.s npp_mn.s npp_sd.s npp_cv.s SolRad_mn.s SolRad_cv.s
## Abuna     0.098    -0.403   -0.721   -0.39    -0.33     -0.27    -0.11
## Amazon1 -0.207     0.039   0.244    -0.12    -0.18     0.38    -0.45
## Amazon2 -0.118    -0.217   0.136    -0.21    -0.24     0.37    -0.41
## Amazon5 -0.018     0.222   -0.168   -0.25    -0.25     0.37    -0.42
## Amazon8 -0.280     0.155   0.081    -0.35    -0.36     0.35    -0.39
## Amazon9 -0.339     0.217   0.124    -0.41    -0.42     0.35    -0.41
##      basin
## Abuna    Abuna
## Amazon1 Amazon1
## Amazon2 Amazon2
## Amazon5 Amazon5
## Amazon8 Amazon8
## Amazon9 Amazon9

```

```

summary(current_climate_s)

```

```

##      tempMin.s      tempMax.s      tempMean.s      tempCV.s
##  Min.  :-2.402    Min.  :-2.766    Min.  :-2.6842    Min.  :-0.42939
##  1st Qu.:-0.183   1st Qu.:-0.045   1st Qu.: 0.0318   1st Qu.:-0.25789
##  Median : 0.178    Median : 0.121    Median : 0.1393   Median : -0.17174
##  Mean   : 0.000    Mean   : 0.000    Mean   : 0.0000   Mean   : 0.00000
##  3rd Qu.: 0.300    3rd Qu.: 0.242   3rd Qu.: 0.2393   3rd Qu.: 0.00464
##  Max.   : 0.479    Max.   : 0.545    Max.   : 0.3536   Max.   : 2.01340

```

```

##   precMin.s      precMax.s      precAnn.s      precCV.s
## Min. :-0.6453    Min. :-1.506    Min. :-1.365    Min. :-0.9745
## 1st Qu.:-0.4026  1st Qu.:-0.221  1st Qu.:-0.241  1st Qu.:-0.3946
## Median :-0.0874   Median : 0.114   Median : 0.037   Median : 0.0885
## Mean   : 0.0000   Mean   : 0.000   Mean   : 0.000   Mean   : 0.0000
## 3rd Qu.: 0.2940   3rd Qu.: 0.295   3rd Qu.: 0.310   3rd Qu.: 0.3827
## Max.   : 1.2694   Max.   : 1.192   Max.   : 1.042   Max.   : 1.0832
##   aetMin.s      aetMax.s      aetAnn.s      aetCV.s
## Min. :-1.0030    Min. :-2.035    Min. :-1.768    Min. :-0.6608
## 1st Qu.:-0.4392  1st Qu.:-0.215  1st Qu.:-0.217  1st Qu.:-0.4766
## Median :-0.0238   Median : 0.103   Median : 0.019   Median :-0.0885
## Mean   : 0.0000   Mean   : 0.000   Mean   : 0.000   Mean   : 0.0000
## 3rd Qu.: 0.5178   3rd Qu.: 0.316   3rd Qu.: 0.319   3rd Qu.: 0.3810
## Max.   : 0.8506   Max.   : 0.781   Max.   : 0.883   Max.   : 1.1939
##   petMin.s      petMax.s      petAnn.s      petCV.s
## Min. :-1.8654    Min. :-1.971    Min. :-2.3313   Min. :-0.6417
## 1st Qu.:-0.1488  1st Qu.:-0.265  1st Qu.:-0.1972  1st Qu.:-0.2808
## Median : 0.0403   Median :-0.050   Median :-0.0546  Median :-0.1327
## Mean   : 0.0000   Mean   : 0.000   Mean   : 0.0000  Mean   : 0.0000
## 3rd Qu.: 0.2537   3rd Qu.: 0.249  3rd Qu.: 0.2403  3rd Qu.: 0.0777
## Max.   : 1.0028   Max.   : 1.100   Max.   : 1.0790  Max.   : 1.8796
##   ro_mean.s      npp_mn.s      npp_sd.s      npp_cv.s
## Min. :-0.9254    Min. :-1.223    Min. :-0.454    Min. :-0.444
## 1st Qu.:-0.3443  1st Qu.:-0.268  1st Qu.:-0.344    1st Qu.:-0.357
## Median :-0.0834   Median : 0.113   Median :-0.238  Median :-0.227
## Mean   : 0.0000   Mean   : 0.000   Mean   : 0.0000  Mean   : 0.0000
## 3rd Qu.: 0.3383   3rd Qu.: 0.334  3rd Qu.: 0.205  3rd Qu.: 0.247
## Max.   : 1.2748   Max.   : 1.006   Max.   : 1.542    Max.   : 1.866
##   SolRad_mn.s    SolRad_cv.s    basin
## Min. :-1.716     Min. :-0.446    Length:97
## 1st Qu.:-0.222   1st Qu.:-0.355  Class :character
## Median : 0.261   Median :-0.210  Mode  :character
## Mean   : 0.000   Mean   : 0.000
## 3rd Qu.: 0.352   3rd Qu.: 0.244
## Max.   : 0.382   Max.   : 2.327

```

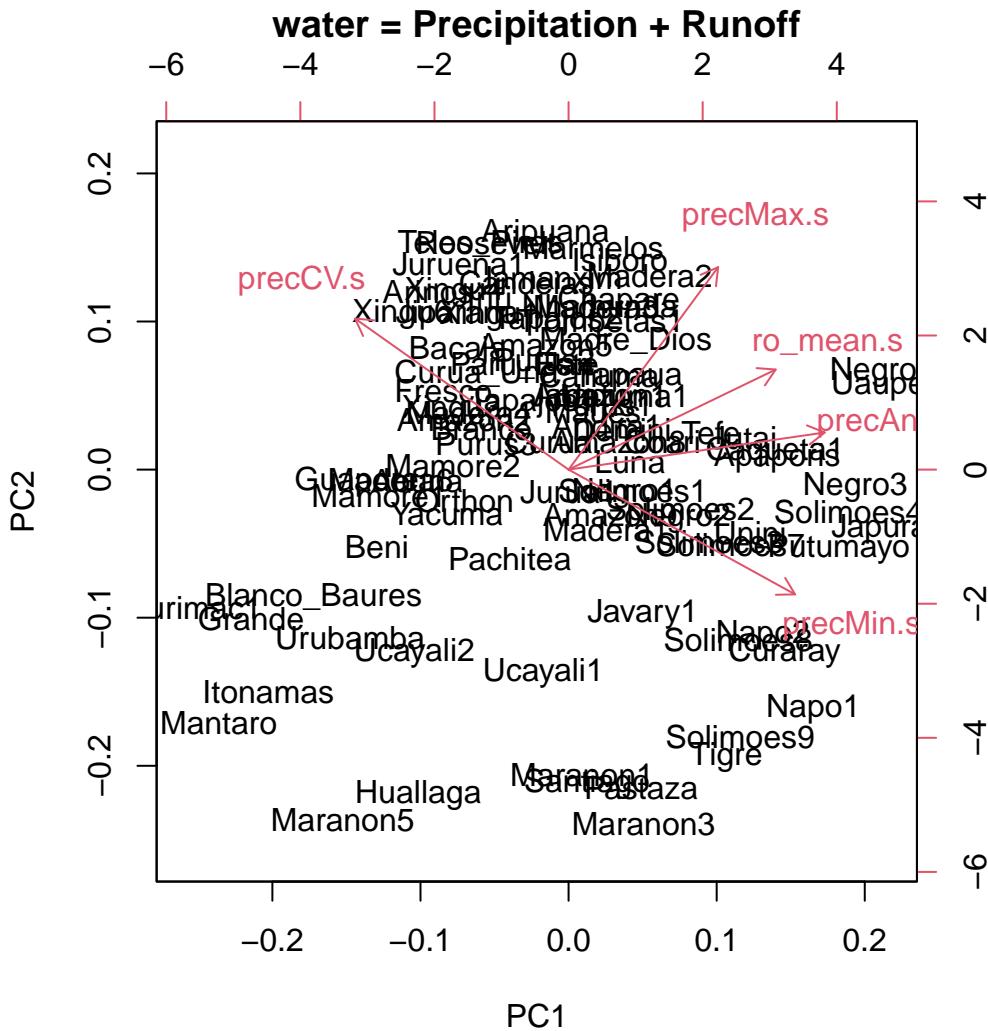
```
write.csv(current_climate_s, file = "comparingResults/current_climate_s.csv")
```

```

#####
##### separating
#####
##### variables
#####
##### according
#####
##### to
#####
##### ecological
#####
##### hypotheses
#####
##### water
#####
##### =
#####
##### Precipitation
#####
##### +
#####
##### Runoff
vari <- c("precMin.s", "precMax.s", "precAnn.s", "precCV.s",
       "ro_mean.s")

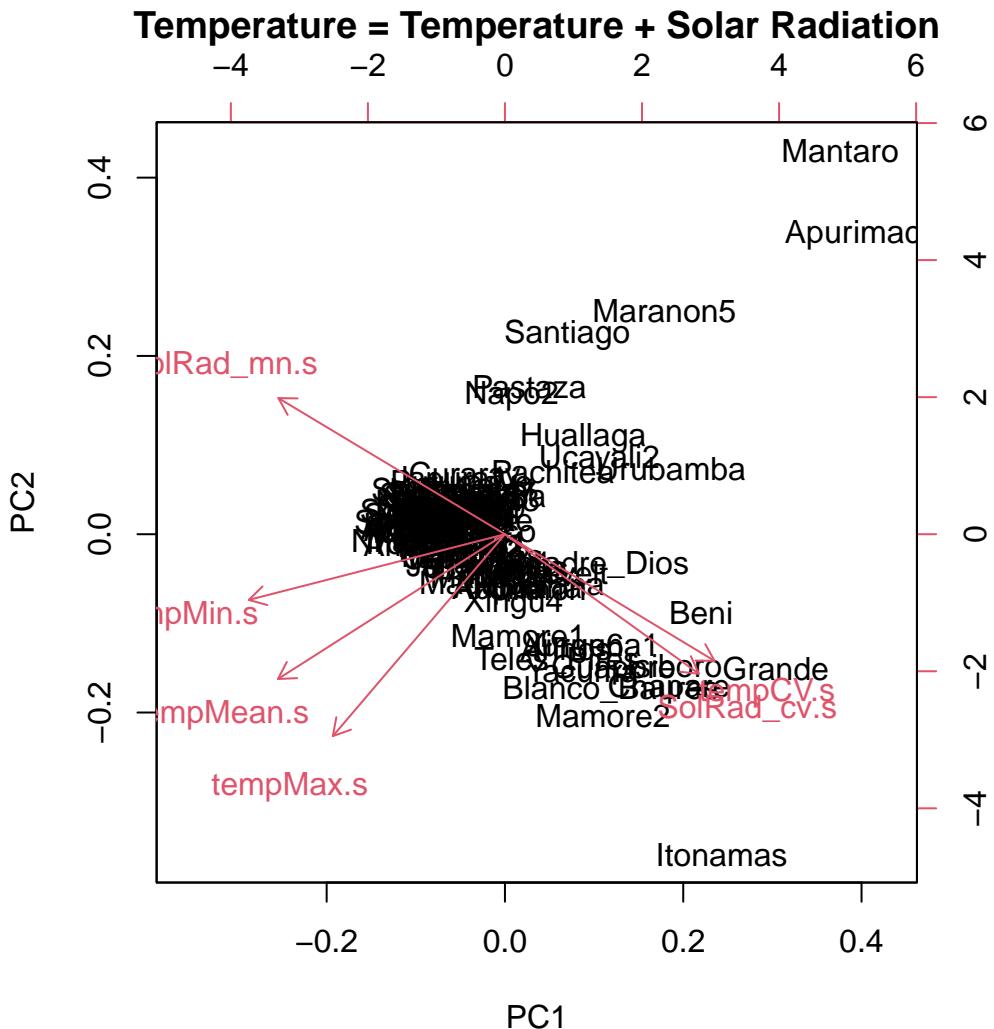
```

```
# summarizing current climate in a Distance matrix
D_water <- dist(current_climate_s[, vari])
res <- prcomp(current_climate_s[, vari], center = F, scale. = F)
biplot(res, main = "water = Precipitation + Runoff")
```



```
rm(res)
rm(vari)

# Temperature = Temperature + Solar Radiation
vari <- c("tempMin.s", "tempMax.s", "tempMean.s", "tempCV.s",
         "SolRad_cv.s", "SolRad_mn.s")
# summarizing in a few PCA axes
D_temp <- dist(current_climate_s[, vari])
res <- prcomp(current_climate_s[, vari], center = F, scale. = F)
biplot(res, main = "Temperature = Temperature + Solar Radiation")
```

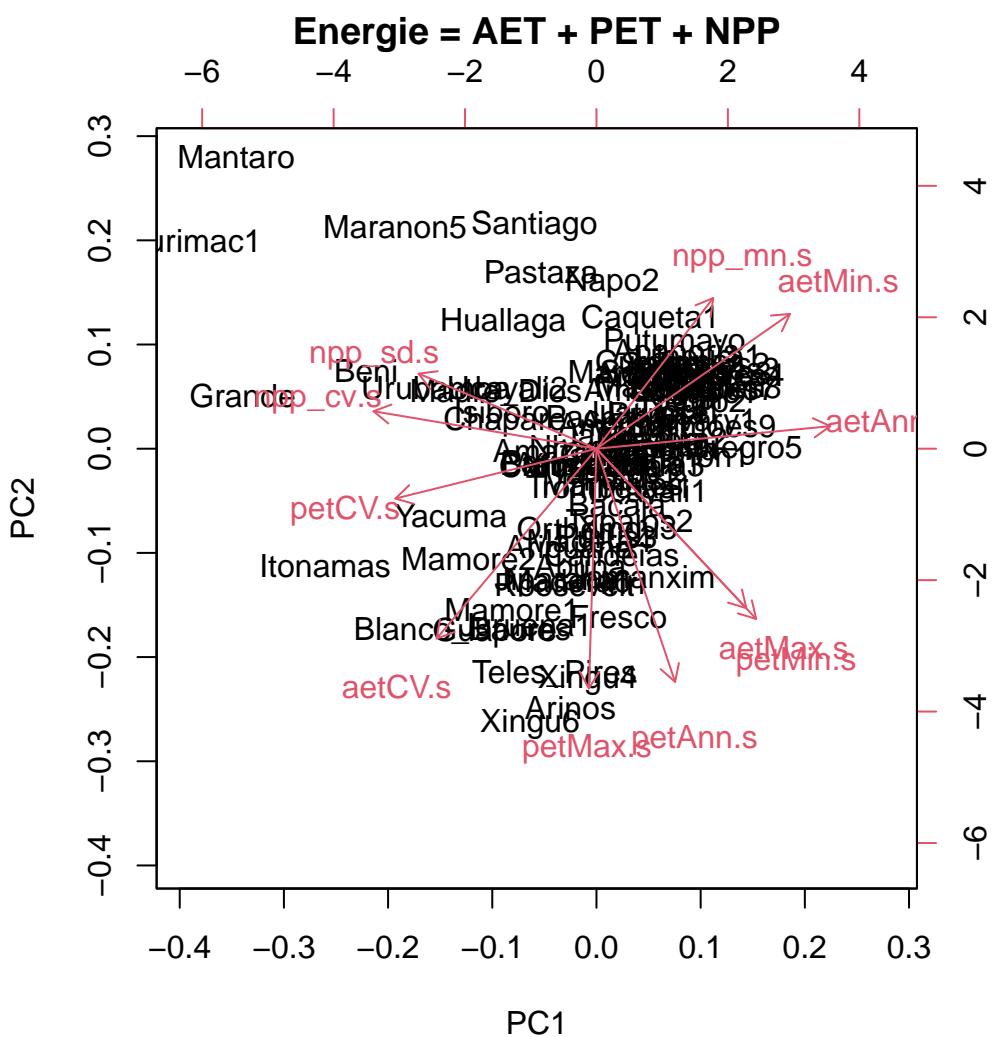


```

rm(res)
rm(vari)

# Energie = AET + PET + NPP
vari <- c("aetMin.s", "aetMax.s", "aetAnn.s", "aetCV.s", "petMin.s",
         "petMax.s", "petAnn.s", "petCV.s", "npp_mn.s", "npp_sd.s",
         "npp_cv.s")
# summarizing in a few PCA axes
D_energ <- dist(current_climate_s[, vari])
res <- prcomp(current_climate_s[, vari], center = F, scale. = F)
biplot(res, main = "Energie = AET + PET + NPP")

```



```

rm(res)
rm(vari)

rm(temp2, prec2, aet2, pet2, Cclimate, ro, npp, solRad)

```

3.7 Past climate

Now, it's time to deal with past climate variables. We first charge data from all past climatic models, separate all mean variables related to temperature and precipitation, aggregate them, and finally calculate the **Euclidean dissimilarity matrix**.

3.7.1 CCSM

```
# Climat LGM : CCSM model
ccsm = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/climatology_lgm_ccsm.csv",
sep = ""), sep = ",", header = T)

# Setting CCSM row names
rownames(ccsm) <- ccsm$BvNiv2

# let's consider only annual variables temperature: I must
# divide by 10 to get the right temperature Mean
# temperature of the coldest month
tempMin <- ccsm[, "bio_mn6"]/10
# Mean temperature of the warmest month
tempMax <- ccsm[, "bio_mn5"]/10
# Annual mean temperature
tempMean <- ccsm[, "bio_mn1"]/10
# Temperature sazonality
tempCV <- ccsm[, "bio_mn4"]/1000 #I must divide by 1000 to get the right CV

temp_ccsm <- data.frame(tempMin, tempMax, tempMean, tempCV)
rownames(temp_ccsm) <- ccsm$BvNiv2
rm(tempMin, tempMax, tempMean, tempCV)
head(temp_ccsm)

##          tempMin tempMax tempMean tempCV
## Abuna      15     29     23    0.67
## Amazon1    18     30     23    0.63
## Amazon2    17     29     22    0.60
## Amazon5    19     29     23    0.45
## Amazon8    18     29     23    0.33
## Amazon9    19     29     23    0.34

summary(temp_ccsm)

##          tempMin        tempMax        tempMean        tempCV
##  Min.   :-7.28   Min.   :11.2   Min.   : 3.41   Min.   :0.245
##  1st Qu.:12.44  1st Qu.:28.1   1st Qu.:21.37  1st Qu.:0.388
##  Median :15.84  Median :28.7   Median :22.37  Median :0.506
##  Mean   :14.14  Mean   :27.9   Mean   :21.26  Mean   :0.625
##  3rd Qu.:16.88  3rd Qu.:29.1   3rd Qu.:22.88  3rd Qu.:0.630
##  Max.   :18.67  Max.   :30.8   Max.   :23.52  Max.   :2.101
```

```

# precipitation Precipitation of the driest month
precMin <- ccsm[, "bio_mn14"]
# Precipitation of the wettest month
precMax <- ccsm[, "bio_mn13"]
# Annual precipitation
precAnn <- ccsm[, "bio_mn12"]
# Coefficient of variation of mean monthly precipitation
precCV <- ccsm[, "bio_mn15"]
prec_ccsm <- data.frame(precMin, precMax, precAnn, precCV)
rownames(prec_ccsm) <- ccsm$BvNiv2
rm(precMin, precMax, precAnn, precCV)
head(prec_ccsm)

```

```

##          precMin precMax precAnn precCV
## Abuna      24     258    1677     58
## Amazon1    124     297    2325     29
## Amazon2     63     319    2015     46
## Amazon5     88     387    2501     44
## Amazon8    123     370    2647     33
## Amazon9    113     382    2569     33

```

```
summary(prec_ccsm)
```

```

##      precMin      precMax      precAnn      precCV
## Min.   : 2.94   Min.   :111   Min.   : 680   Min.   :17.4
## 1st Qu.:31.49  1st Qu.:263   1st Qu.:1901  1st Qu.:30.9
## Median :72.47  Median :319   Median :2304   Median :44.2
## Mean   :74.29  Mean   :310   Mean   :2227   Mean   :44.4
## 3rd Qu.:111.34 3rd Qu.:350   3rd Qu.:2555  3rd Qu.:55.0
## Max.   :181.97  Max.   :627   Max.   :3516   Max.   :78.9

```

```

# do they have the same row sequence?
temp_ccsm$basin <- rownames(temp_ccsm)
prec_ccsm$basin <- rownames(prec_ccsm)
all.equal(temp_ccsm$basin, prec_ccsm$basin) #ok

```

```
## [1] TRUE
```

```

lgm_ccsm <- merge(temp_ccsm, prec_ccsm, by = "basin")
rm(temp_ccsm, prec_ccsm, ccsm)
summary(lgm_ccsm)

```

```

##      basin      tempMin      tempMax      tempMean
## Length:97      Min.   :-7.28   Min.   :11.2   Min.   : 3.41
## Class :character 1st Qu.:12.44  1st Qu.:28.1   1st Qu.:21.37
## Mode  :character Median :15.84  Median :28.7   Median :22.37
##                  Mean   :14.14  Mean   :27.9   Mean   :21.26
##                  3rd Qu.:16.88  3rd Qu.:29.1   3rd Qu.:22.88
##                  Max.   :18.67  Max.   :30.8   Max.   :23.52

```

```

##      tempCV      precMin      precMax      precAnn      precCV
## Min.   :0.245   Min.   : 2.94   Min.   :111   Min.   : 680   Min.   :17.4
## 1st Qu.:0.388   1st Qu.: 31.49   1st Qu.:263   1st Qu.:1901   1st Qu.:30.9
## Median :0.506   Median : 72.47   Median :319   Median :2304   Median :44.2
## Mean    :0.625   Mean    : 74.29   Mean    :310   Mean    :2227   Mean    :44.4
## 3rd Qu.:0.630   3rd Qu.:111.34   3rd Qu.:350   3rd Qu.:2555   3rd Qu.:55.0
## Max.    :2.101   Max.    :181.97   Max.    :627   Max.    :3516   Max.    :78.9

```

3.7.2 MIROC

```

# Clima LGM : MIROC model
miroc = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/climatology_lgm_miroc.csv",
sep = ""), sep = ",", header = T)

# Setting row names
rownames(miroc) <- miroc$BvNiv2

# let's consider only annual variables temperature: I must
# divided by 10 to get the right temperature Mean
# temperature of the coldest month
tempMin <- miroc[, "bio_mn6"]/10
# Mean temperature of the warmest month
tempMax <- miroc[, "bio_mn5"]/10
# Annual mean temperature
tempMean <- miroc[, "bio_mn1"]/10
# Coefficient of variation of mean monthly temperature
tempCV <- miroc[, "bio_mn4"]/1000
temp_miroc <- data.frame(tempMin, tempMax, tempMean, tempCV)
rownames(temp_miroc) <- miroc$BvNiv2
rm(tempMin, tempMax, tempMean, tempCV)
head(temp_miroc, 10)

```

```

##      tempMin tempMax tempMean tempCV
## Abuna     13.6     28      22  1.33
## Amazon1   18.3     31      24  0.96
## Amazon2   16.4     28      22  1.11
## Amazon5   17.6     30      24  0.87
## Amazon8   18.7     30      24  0.61
## Amazon9   19.0     30      25  0.53
## Andira1   18.0     30      24  0.58
## Apaporis  17.3     29      23  0.79
## Apurimac1 -6.8     15       5  1.35
## Arinos    10.3     28      20  1.49

```

```

# precipitation Precipitation of the driest month
precMin <- miroc[, "bio_mn14"]
# Precipitation of the wettest month
precMax <- miroc[, "bio_mn13"]
# Anual precipitation
precAnn <- miroc[, "bio_mn12"]

```

```

# Coefficient of variation of mean monthly precipitation
precCV <- miroc[, "bio_mn15"]
prec_miroc <- data.frame(precMin, precMax, precAnn, precCV)
rownames(prec_miroc) <- miroc$BvNiv2
rm(precMin, precMax, precAnn, precCV)
head(prec_miroc, 5)

##          precMin precMax precAnn precCV
## Abuna      21     256    1916     52
## Amazon1    129     420    2985     38
## Amazon2     56     384    2183     58
## Amazon5     66     265    1706     49
## Amazon8     81     238    1743     33

# do they have the same row sequence?
temp_miroc$basin <- rownames(temp_miroc)
prec_miroc$basin <- rownames(prec_miroc)
all.equal(temp_miroc$basin, prec_miroc$basin) #ok

## [1] TRUE

lgm_miroc <- merge(temp_miroc, prec_miroc, by = "basin")
rm(temp_miroc, prec_miroc)
head(lgm_miroc)

##          basin tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
## 1   Abuna      14      28      22  1.33      21     256    1916     52
## 2 Amazon1     18      31      24  0.96     129     420    2985     38
## 3 Amazon2     16      28      22  1.11      56     384    2183     58
## 4 Amazon5     18      30      24  0.87      66     265    1706     49
## 5 Amazon8     19      30      24  0.61      81     238    1743     33
## 6 Amazon9     19      30      25  0.53      86     224    1770     30

summary(lgm_miroc)

##          basin              tempMin            tempMax            tempMean
## Length:97           Min.   :-6.77           Min.   :13.0           Min.   : 4.62
## Class :character    1st Qu.:12.09        1st Qu.:27.5        1st Qu.:20.30
## Mode  :character    Median :15.67        Median :28.4        Median :22.22
##                   Mean   :14.08        Mean   :28.0        Mean   :21.35
##                   3rd Qu.:17.18        3rd Qu.:29.4        3rd Qu.:23.09
##                   Max.  :19.04        Max.  :33.1        Max.  :24.57
##          tempCV            precMin            precMax            precAnn            precCV
## Min.   :0.453       Min.   : 2.94       Min.   :145       Min.   : 867       Min.   :19.0
## 1st Qu.:0.721      1st Qu.: 28.16     1st Qu.:274     1st Qu.:1915     1st Qu.:34.7
## Median :0.893      Median : 57.63     Median :312     Median :2325      Median :46.9
## Mean   :0.992      Mean   : 61.72     Mean   :314     Mean   :2230      Mean   :45.8
## 3rd Qu.:1.128      3rd Qu.: 85.25     3rd Qu.:352     3rd Qu.:2452     3rd Qu.:56.1
## Max.   :2.316      Max.   :153.72     Max.   :488     Max.   :3562      Max.   :89.8

```

```
rm(miroc)
```

3.7.3 MPI model

```
# Clima LGM : MPI model
mpi = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Statistiques/climatology_lgm_mpi.csv",
  sep = ""), sep = ",", header = T)

# head(mpi)
rownames(mpi) <- mpi$BvNiv2

# temperature: I must divide by 10 to get the right
# temperature Mean temperature of the coldest month
tempMin <- mpi[, "bio_mn6"]/10
# Mean temperature of the warmest month
tempMax <- mpi[, "bio_mn5"]/10
# Annual mean temperature
tempMean <- mpi[, "bio_mn1"]/10
# Coefficient of variation of mean monthly temperature
tempCV <- mpi[, "bio_mn4"]/1000
temp_mpi <- data.frame(tempMin, tempMax, tempMean, tempCV)
rownames(temp_mpi) <- mpi$BvNiv2
rm(tempMin, tempMax, tempMean, tempCV)

# precipitation Precipitation of the driest month
precMin <- mpi[, "bio_mn14"]
# Precipitation of the wettest month
precMax <- mpi[, "bio_mn13"]
# Anual precipitation
precAnn <- mpi[, "bio_mn12"]
# Coefficient of variation of mean monthly precipitation
precCV <- mpi[, "bio_mn15"]
prec_mpi <- data.frame(precMin, precMax, precAnn, precCV)
rownames(prec_mpi) <- mpi$BvNiv2
rm(precMin, precMax, precAnn, precCV)

# do they have the same row sequence?
temp_mpi$basin <- rownames(temp_mpi)
prec_mpi$basin <- rownames(prec_mpi)
all.equal(temp_mpi$basin, prec_mpi$basin) #ok

## [1] TRUE

lgm_mpi <- merge(temp_mpi, prec_mpi, by = "basin")
rm(temp_mpi, prec_mpi)
head(lgm_mpi)

##      basin tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
```

```

## 1 Abuna      14     28     22   1.16     87    249    2048     29
## 2 Amazon1    17     30     23   0.93    112   1741    6117     84
## 3 Amazon2    16     28     22   0.62     58    640    2942     64
## 4 Amazon5    16     28     22   0.92    127   519    3249     40
## 5 Amazon8    16     28     22   1.37    131   549    3384     44
## 6 Amazon9    16     29     23   1.35    136   479    3145     38

```

```
summary(lgm_mpi)
```

	basin	tempMin	tempMax	tempMean	
## Length:97		Min. :-8.05	Min. :11.7	Min. : 3.4	
## Class :character		1st Qu.:12.20	1st Qu.:27.7	1st Qu.:20.9	
## Mode :character		Median :14.57	Median :28.2	Median :21.6	
##		Mean :13.19	Mean :27.6	Mean :20.7	
##		3rd Qu.:15.59	3rd Qu.:28.9	3rd Qu.:22.2	
##		Max. :16.92	Max. :31.3	Max. :23.1	
	tempCV	precMin	precMax	precAnn	precCV
## Min. :0.536		Min. : 3.1	Min. :143	Min. : 849	Min. :18.0
## 1st Qu.:0.905		1st Qu.: 53.2	1st Qu.: 311	1st Qu.:2360	1st Qu.:34.2
## Median :1.138		Median : 94.0	Median : 458	Median :2920	Median :43.2
## Mean :1.105		Mean : 95.5	Mean : 455	Mean :2739	Mean :45.2
## 3rd Qu.:1.254		3rd Qu.:140.5	3rd Qu.: 551	3rd Qu.:3217	3rd Qu.:54.5
## Max. :1.810		Max. :194.6	Max. :1741	Max. :6117	Max. :84.0

```
rm(mpi)
```

3.7.4 Mean variables of LGM models

The first thing is to verify that all data frames from different climatic models have the same sequence of rows and columns. Let's check whether or not this is the case.

```

# all pairwise comparisons of ROW names
all.equal(lgm_miroc$basin, lgm_ccsm$basin) #ok

## [1] TRUE

all.equal(lgm_miroc$basin, lgm_mpi$basin) #ok

## [1] TRUE

all.equal(lgm_ccsm$basin, lgm_mpi$basin) #ok

## [1] TRUE

# all pairwise comparisons of COLUMN names
all.equal(colnames(lgm_miroc), colnames(lgm_ccsm)) #ok

## [1] TRUE

```

```
all.equal(colnames(lgm_miroc), colnames(lgm_mpi)) #ok
```

```
## [1] TRUE
```

```
all.equal(colnames(lgm_ccsm), colnames(lgm_mpi)) #ok
```

```
## [1] TRUE
```

Great, everything is **OK!** As all selected variables from the three LGM models have the same structure, we can produce a mean value for each cell variable by summing and dividing all matrix cells among different data sets⁶.

```
head(lgm_miroc)
```

```
##   basin tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
## 1 Abuna    14     28      22    1.33     21     256    1916     52
## 2 Amazon1   18     31      24    0.96    129     420    2985     38
## 3 Amazon2   16     28      22    1.11     56     384    2183     58
## 4 Amazon5   18     30      24    0.87     66     265    1706     49
## 5 Amazon8   19     30      24    0.61     81     238    1743     33
## 6 Amazon9   19     30      25    0.53     86     224    1770     30
```

```
head(lgm_ccsm)
```

```
##   basin tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
## 1 Abuna    15     29      23    0.67     24     258    1677     58
## 2 Amazon1   18     30      23    0.63    124     297    2325     29
## 3 Amazon2   17     29      22    0.60     63     319    2015     46
## 4 Amazon5   19     29      23    0.45     88     387    2501     44
## 5 Amazon8   18     29      23    0.33    123     370    2647     33
## 6 Amazon9   19     29      23    0.34    113     382    2569     33
```

```
head(lgm_mpi)
```

```
##   basin tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
## 1 Abuna    14     28      22    1.16     87     249    2048     29
## 2 Amazon1   17     30      23    0.93    112    1741    6117     84
## 3 Amazon2   16     28      22    0.62     58     640    2942     64
## 4 Amazon5   16     28      22    0.92    127     519    3249     40
## 5 Amazon8   16     28      22    1.37    131     549    3384     44
## 6 Amazon9   16     29      23    1.35    136     479    3145     38
```

```
# excluding the column that refers to the basin name and
# computing mean values of past climate projections among
# different models
```

```
lgm_meanModels <- (lgm_miroc[, -1] + lgm_ccsm[, -1] + lgm_mpi[,
```

⁶but excluding the first column that refers to the row names

```

-1]) /3

rownames(lgm_meanModels) <- lgm_miroc$basin
head(lgm_meanModels)

##          tempMin tempMax tempMean tempCV precMin precMax precAnn precCV
## Abuna      14       28      22   1.05      44     254    1880      46
## Amazon1    18       30      24   0.84     122     819    3809      50
## Amazon2    17       28      22   0.78      59     448    2380      56
## Amazon5    17       29      23   0.74      94     391    2485      44
## Amazon8    18       29      23   0.77     112     386    2591      37
## Amazon9    18       29      24   0.74     112     362    2495      34

lgm_meanModels <- lgm_meanModels[, !colnames(lgm_meanModels) %in%
  c("tempCV", "precCV")]

write.csv(lgm_meanModels, file = "comparingResults/lgm_meanModels.csv")
rm(lgm_miroc, lgm_ccsm, lgm_mpi)

```

3.7.5 Computing differences between current and past climate

Now, we are ready to subtract the past projection variables from the current climate to have a table corresponding to changes in climate between the Last Glacial Maximum to the current days.

```

head(current_climate)

##          tempMin tempMax tempMean tempCV precMin precMax precAnn precCV aetMin
## Abuna      18       33      26   0.72      20     270    1785      61      65
## Amazon1    22       33      27   0.52      59     340    2453      47      85
## Amazon2    21       32      26   0.61      36     305    1836      61      61
## Amazon5    22       33      27   0.58      45     334    2098      58      71
## Amazon8    22       33      27   0.44      86     332    2311      44     100
## Amazon9    22       33      27   0.44      86     296    2250      40     105
##          aetMax aetAnn aetCV petMin petMax petAnn petCV ro_mean npp_mn npp_sd
## Abuna      152    1309  0.291     122     166    1751  0.101      689    1689      96
## Amazon1    138    1433  0.147     122     159    1654  0.084     1153    2679     261
## Amazon2    134    1228  0.254     119     157    1628  0.089      885    2568     211
## Amazon5    136    1298  0.195     120     161    1650  0.094     1345    2256     181
## Amazon8    134    1429  0.087     121     153    1623  0.079     1274    2512     124
## Amazon9    135    1436  0.080     120     151    1614  0.076     1339    2556      88
##          npp_cv SolRad_mn SolRad_cv
## Abuna      5.7        15    0.1133
## Amazon1    9.7        15    0.0024
## Amazon2    8.2        15    0.0138
## Amazon5    8.0        15    0.0118
## Amazon8    4.9        15    0.0216
## Amazon9    3.4        15    0.0142

```

```
head(lgm_meanModels)
```

```
##      tempMin tempMax tempMean precMin precMax precAnn
## Abuna     14      28      22      44     254    1880
## Amazon1   18      30      24     122     819    3809
## Amazon2   17      28      22      59     448    2380
## Amazon5   17      29      23      94     391    2485
## Amazon8   18      29      23     112     386    2591
## Amazon9   18      29      24     112     362    2495
```

do they have the same structure for lines?

```
all.equal(rownames(current_climate), rownames(lgm_meanModels))
```

```
## [1] TRUE
```

Differences between Current and the LGM climate

```
head(current_climate[, colnames(lgm_meanModels)])
```

```
##      tempMin tempMax tempMean precMin precMax precAnn
## Abuna     18      33      26      20     270    1785
## Amazon1   22      33      27      59     340    2453
## Amazon2   21      32      26      36     305    1836
## Amazon5   22      33      27      45     334    2098
## Amazon8   22      33      27      86     332    2311
## Amazon9   22      33      27      86     296    2250
```

```
head(lgm_meanModels)
```

```
##      tempMin tempMax tempMean precMin precMax precAnn
## Abuna     14      28      22      44     254    1880
## Amazon1   18      30      24     122     819    3809
## Amazon2   17      28      22      59     448    2380
## Amazon5   17      29      23      94     391    2485
## Amazon8   18      29      23     112     386    2591
## Amazon9   18      29      24     112     362    2495
```

```
diff_CurrentLGM <- (current_climate[, colnames(lgm_meanModels)] -
  lgm_meanModels)
summary(diff_CurrentLGM)
```

```
##      tempMin      tempMax      tempMean      precMin
## Min.   :3.31   Min.   :0.836   Min.   :3.18   Min.   :-63.19
## 1st Qu.:3.95  1st Qu.:3.311  1st Qu.:3.55  1st Qu.:-25.10
## Median :4.26   Median :3.965   Median :3.74   Median :-11.38
## Mean   :4.20   Mean   :3.829   Mean   :3.78   Mean   : -5.88
## 3rd Qu.:4.43  3rd Qu.:4.496  3rd Qu.:4.01  3rd Qu.: 1.49
## Max.   :5.18   Max.   :5.644   Max.   :4.51   Max.   : 91.81
##      precMax      precAnn
```

```

## Min.   :-479.786   Min.   :-1355.3
## 1st Qu.: -90.138   1st Qu.: -280.8
## Median : -34.806   Median : -172.3
## Mean   : -51.483   Mean   : -194.5
## 3rd Qu.: -0.514    3rd Qu.: -62.9
## Max.   : 24.699    Max.   : 168.0

write.csv(diff_CurrentLGM, file = "comparingResults/diff_CurrentLGM.csv")

# Computing Distance Matrix on scaled data
diff_CurrentLGM_s <- (apply(diff_CurrentLGM, 2, Standard))
rownames(diff_CurrentLGM_s) <- rownames(diff_CurrentLGM)
colMeans(diff_CurrentLGM_s) #OK

## tempMin tempMax tempMean precMin precMax precAnn
## 4.8e-16 -7.6e-17 2.6e-16 1.1e-17 -1.8e-17 -3.5e-17

apply(diff_CurrentLGM_s, 2, sd) #OK

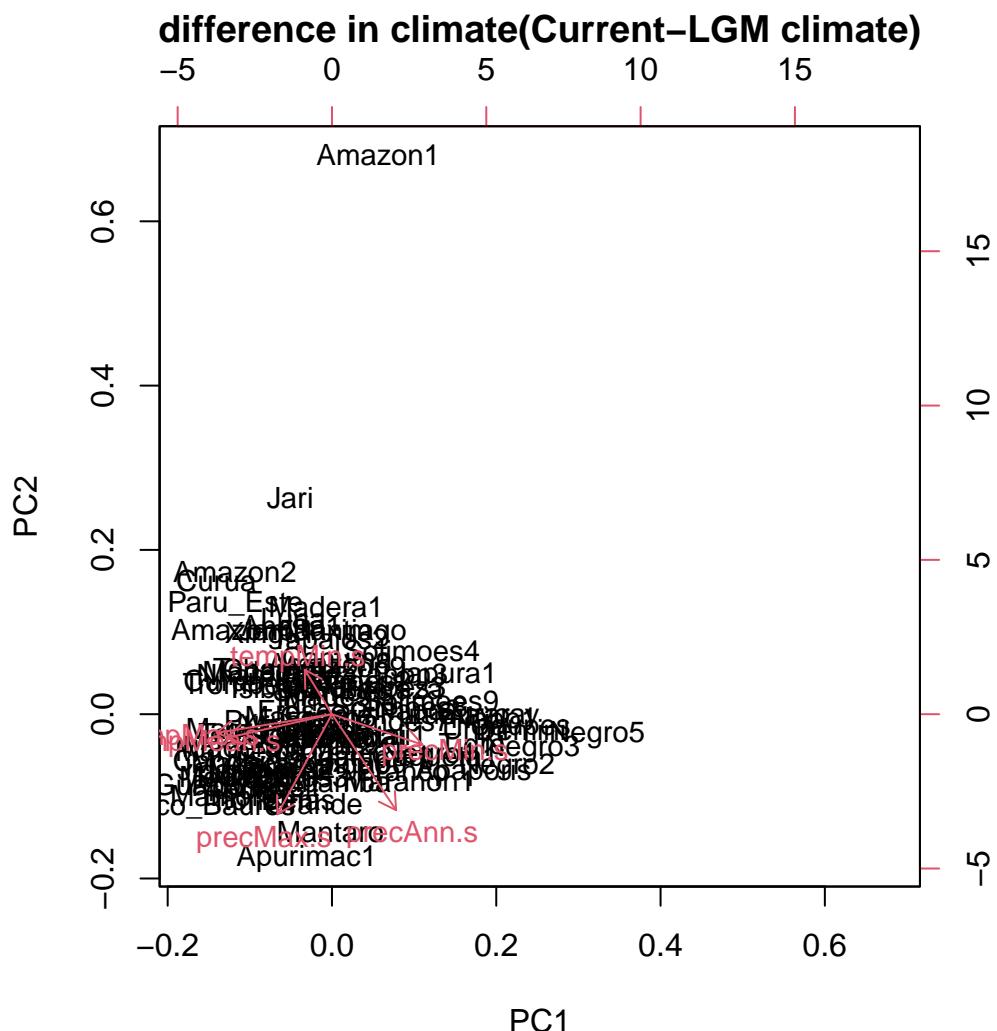
## tempMin tempMax tempMean precMin precMax precAnn
## 0.5      0.5      0.5      0.5      0.5      0.5

# setting predictor' names
colnames(diff_CurrentLGM_s) <- paste(colnames(diff_CurrentLGM_s),
                                         ".s", sep = "")
head(diff_CurrentLGM_s)

##          tempMin.s tempMax.s tempMean.s precMin.s precMax.s precAnn.s
## Abuna     0.023     0.61     0.58     -0.31     0.484     0.22
## Amazon1   0.137    -0.73    -0.94     -0.96    -3.115    -2.64
## Amazon2   0.844     0.30     0.90     -0.29    -0.664    -0.79
## Amazon5   0.977     0.37     0.36     -0.71    -0.040    -0.44
## Amazon8   0.382    -0.15    -0.27     -0.32    -0.018    -0.20
## Amazon9   0.228    -0.19    -0.37     -0.33    -0.106    -0.11

# PCA (centered and reduced) with differences in variables
pca_diff <- prcomp(diff_CurrentLGM_s, center = F, scale. = F)
biplot(pca_diff, main = "difference in climate(Current-LGM climate)",
       cex = 0.9)

```



```

# Distance
D_CurrentLGM <- dist(diff_CurrentLGM_s)

# control
diff_CurrentLGM_s <- data.frame(diff_CurrentLGM_s)
diff_CurrentLGM_s$basin <- rownames(diff_CurrentLGM_s)
# comp <- merge(comp, diff_CurrentLGM_s, by='basin')

rm(lgm_meanModels, pca_diff, diff_CurrentLGM, current_climate_s,
  current_climate)

```

3.8 Codding the structure of the drainage basin

We computed an index to evaluate the effect of tributary position on fish biodiversity. This index is defyned as...

```
pos = read.csv(paste("StatistiquesSubBasinAmazon_02052016/codage_bv2_tributaryPosition_Ago2017.csv",
  sep = ""), sep = ";", header = T)
pos
```

##	Basin	BvNiv2	codage
## 1	Jari	Jari	2
## 2	Curua	Curua	6
## 3	Paru_Este	Paru_Este	4
## 4	Curua_Una	Curua_Una	5
## 5	Tefe	Tefe	17
## 6	Jutai	Jutai	19
## 7	Putumayo	Putumayo	20
## 8	Coari	Coari	15
## 9	Nhamunda	Nhamunda	9
## 10	Uatumá	Jatapu	11
## 11	Negro	Branco	13
## 12	Negro	Demini	13
## 13	Xingu	Iriri	3
## 14	Xingu	Fresco	3
## 15	Xingu	Bacaja	3
## 16	Tapajos	Jamanxim	7
## 17	Madera	Jiparana	12
## 18	Andira	Maues	10
## 19	Madera	Luna	12
## 20	Madera	Marmelos	12
## 21	Madera	Candeias	12
## 22	Madera	Abuna	12
## 23	Purus	Tapaua	14
## 24	Ucayali	Pachitea	24
## 25	Maranon	Tigre	23
## 26	Maranon	Huallaga	23
## 27	Maranon	Pastaza	23
## 28	Maranon	Santiago	23
## 29	Japura	Apaporis	16
## 30	Napo	Curaray	22
## 31	Madera	Orthon	12
## 32	Madera	Madre_Dios	12
## 33	Madera	Beni	12
## 34	Madera	Itonamas	12
## 35	Madera	Yacuma	12
## 36	Madera	Isiboro	12
## 37	Madera	Blanco_Baures	12
## 38	Madera	Guapore	12
## 39	Ucayali	Urubamba	24
## 40	Ucayali	Mantaro	24
## 41	Xingu	Xingu1	3
## 42	Xingu	Xingu3	3

## 43	Xingu	Xingu4	3
## 44	Xingu	Xingu6	3
## 45	Amazon	Amazon1	1
## 46	Amazon	Amazon2	1
## 47	Amazon	Amazon5	1
## 48	Tapajos	Tapajos1	7
## 49	Tapajos	Tapajos2	7
## 50	Tapajos	Teles_Pires	7
## 51	Andira	Andira1	10
## 52	Andira	Canuma	10
## 53	Madera	Madera1	12
## 54	Madera	Madera2	12
## 55	Madera	Madera3	12
## 56	Madera	Madera4	12
## 57	Madera	Madera6	12
## 58	Madera	Mamore1	12
## 59	Madera	Mamore2	12
## 60	Purus	Purus1	14
## 61	Purus	Purus2	14
## 62	Purus	Purus3	14
## 63	Jurua1	Jurua1	18
## 64	Javary1	Javary1	21
## 65	Ucayali	Ucayali1	24
## 66	Ucayali	Ucayali2	24
## 67	Maranon	Maranon1	23
## 68	Maranon	Maranon3	23
## 69	Maranon	Maranon5	23
## 70	Napo	Napo1	22
## 71	Napo	Napo2	22
## 72	Japura	Japura1	16
## 73	Japura	Caqueta1	16
## 74	Negro	Negro1	13
## 75	Negro	Negro2	13
## 76	Negro	Uaupes	13
## 77	Negro	Negro3	13
## 78	Negro	Negro5	13
## 79	Uatuma	Uatumai1	11
## 80	Trombetas	Trombetas1	8
## 81	Amazon	Amazon8	1
## 82	Amazon	Amazon9	1
## 83	Amaz_inter	Solimoes1	1
## 84	Amaz_inter	Solimoes2	1
## 85	Amaz_inter	Solimoes3	1
## 86	Solimoes	Solimoes4	1
## 87	Solimoes	Solimoes7	1
## 88	Solimoes	Solimoes8	1
## 89	Solimoes	Solimoes9	1
## 90	Tapajos	Juruena1	7
## 91	Tapajos	Arinos	7
## 92	Madera	Aripuana	12
## 93	Madera	Roosevelt	12
## 94	Madera	Grande	12

```
## 95      Madera      Chapare     12
## 96      Negro       Unini       13
## 97      Ucayali     Apurimac1   24

# merging the dataset
fish <- merge(x = fish, y = pos, by.x = "basin", by.y = "BvNiv2",
    all.x = T)
fish$Codigo <- fish$codage

rm(pos)
```

3.9 Marine Incursions (5Myr and 1Myr)

We calculated the percentage of each sub-drainage basin under 25m of altitude to represent a potential case of marine incursion. The higher the percentage of basin under this threshold, the higher the chance of being under the sea level during the last 1Myr ago.

```
mar = read.csv(paste("StatistiquesSubBasinAmazon_02052016/Marine_Incursion_2018_07/",
  "PercElevSbv2_260618.csv", sep = ""), sep = ";", header = T)
summary(mar)
```

```
##      BvNiv1          BvNiv2          Inf25m        Ent25_100m
## Length:61      Length:61      Min.   : 2.0      Min.   : 1.4
## Class :character  Class :character  1st Qu.: 2.5      1st Qu.: 11.0
## Mode  :character  Mode  :character  Median  : 8.0      Median  : 44.7
##                               Mean   :14.8      Mean   : 47.4
##                               3rd Qu.:22.5      3rd Qu.: 88.1
##                               Max.  :68.0      Max.  :100.0
##                               NA's   :34
```

```
dim(mar)
```

```
## [1] 61 4
```

```
# new and old basin Names
fish2 = read.csv(paste("StatistiquesSubBasinAmazon_02052016/",
  "NewData_2019_10/stat97bv_311019.csv", sep = ""), sep = ";",
  header = T)[, c("Sub_drainage", "Basin2stat")]
dim(fish2)
```

```
## [1] 97 2
```

```
head(fish2)
```

```
## Sub_drainage Basin2stat
## 1       Abuna     Abuna
## 2    Amazon1    Amazon1
## 3  Amazon10  Solimoes7
## 4  Amazon11  Solimoes8
## 5  Amazon12  Solimoes9
## 6    Amazon2     Amazon2
```

```
head(mar)
```

```
##      BvNiv1      BvNiv2 Inf25m Ent25_100m
## 1 Amaz_inter Solimoes1     27      100
## 2 Amaz_inter Solimoes2      8      100
## 3 Amaz_inter Solimoes3      3      100
## 4    Amazon    Amazon1     64      89
## 5    Amazon    Amazon2     31      53
## 6    Amazon    Amazon5     68      95
```

```

intersect(sort(mar$BvNiv2), sort(fish2$Basin2stat))

## [1] "Amazon1"      "Amazon2"       "Amazon5"       "Amazon8"       "Amazon9"
## [6] "Andira1"       "Apaporis"      "Aripuana"      "Branco"        "Candeias"
## [11] "Canuma"        "Caqueta1"     "Coari"         "Curua"         "Curua_Una"
## [16] "Demini"        "Jamanxim"     "Japura1"      "Jari"          "Jatapu"
## [21] "Javary1"       "Jiparana"      "Jurua1"        "Jutai"         "Luna"
## [26] "Madera1"       "Madera2"       "Madera3"      "Madera4"      "Madera6"
## [31] "Maranon1"      "Marmelos"      "Maues"         "Napo1"        "Negro1"
## [36] "Negro2"         "Negro3"        "Negro5"        "Nhamunda"     "Paru_Este"
## [41] "Purus1"         "Purus2"        "Purus3"        "Putumayo"     "Roosevelt"
## [46] "Solimoes1"      "Solimoes2"     "Solimoes3"    "Solimoes4"    "Solimoes7"
## [51] "Solimoes8"      "Solimoes9"     "Tapajos1"      "Tapajos2"     "Tapaua"
## [56] "Tefe"           "Trombetas1"   "Uatumal"       "Uaupes"        "Unini"
## [61] "Xingu1"

```

```

mar <- merge(x = fish2, y = mar, by.x = "Basin2stat", by.y = "BvNiv2",
             all.x = TRUE, all.y = TRUE)

mar$Inf25m[is.na(mar$Inf25m)] <- 0
mar$Ent25_100m[is.na(mar$Ent25_100m)] <- 0
mar <- mar[, colnames(mar) != c("BvNiv1")]
cor(log(mar$Inf25m + 1), log(mar$Ent25_100m + 1))

```

```

## [1] 0.62

```

```

subset(mar, Inf25m != 0)

```

	Basin2stat	Sub_drainage	Inf25m	Ent25_100m
## 2	Amazon1	Amazon1	64	89.1
## 3	Amazon2	Amazon2	31	53.0
## 4	Amazon5	Amazon3	68	94.7
## 5	Amazon8	Amazon4	30	76.3
## 6	Amazon9	Amazon5	34	87.1
## 7	Andira1	Andira	18	88.1
## 17	Canuma	Canuma	2	54.8
## 20	Coari	Coari	2	98.4
## 22	Curua	Curua	3	15.9
## 34	Jari	Jari	2	6.5
## 41	Luna	Luna	11	100.0
## 42	Madera1	Madera1	29	100.0
## 43	Madera2	Madera2	2	90.4
## 55	Maues	Maues	3	46.4
## 58	Negro1	Negro1	4	84.2
## 59	Negro2	Negro2	8	100.0
## 62	Nhamunda	Nhamunda	12	53.2
## 65	Paru_Este	Paru_Este	2	6.9
## 67	Purus1	Purus1	5	100.0
## 73	Solimoes1	Amazon6	27	100.0
## 74	Solimoes2	Amazon7	8	100.0

```

## 75 Solimoes3      Amazon8       3      100.0
## 80 Tapajos1      Tapajos1     10      54.1
## 86 Trombetas1    Trombetas     2      11.6
## 87 Uatumal1      Uatumal      5      61.1
## 91 Unini          Unini        2      100.0
## 93 Xingu1         Xingu1       12      44.7

write.csv(mar, file = "comparingResults/marineIncursions.csv")

# merging marine and fish data set
intersect(sort(mar$Basin2stat), sort(fish$basin))

## [1] "Abuna"           "Amazon1"        "Amazon2"        "Amazon5"
## [5] "Amazon8"         "Amazon9"         "Andira1"        "Apaporis"
## [9] "Apurimac1"       "Arinos"          "Aripuana"       "Bacaja"
## [13] "Beni"            "Blanco_Baures"  "Branco"         "Candeias"
## [17] "Canuma"          "Caqueta1"       "Chapare"        "Coari"
## [21] "Curaray"         "Curua"          "Curua_Una"     "Demini"
## [25] "Fresco"          "Grande"         "Guapore"        "Huallaga"
## [29] "Iriri"           "Isiboro"        "Itonamas"       "Jamanxim"
## [33] "Japura1"         "Jari"            "Jatapu"         "Javary1"
## [37] "Jiparana"        "Jurua1"         "Juruena1"       "Jutai"
## [41] "Luna"             "Madera1"        "Madera2"        "Madera3"
## [45] "Madera4"         "Madera6"        "Madre_Dios"    "Mamore1"
## [49] "Mamore2"         "Mantaro"        "Maranon1"       "Maranon3"
## [53] "Maranon5"        "Marmelos"       "Maues"          "Napo1"
## [57] "Napo2"            "Negro1"          "Negro2"         "Negro3"
## [61] "Negro5"           "Nhamunda"       "Orthon"         "Pachitea"
## [65] "Paru_Este"        "Pastaza"        "Purus1"         "Purus2"
## [69] "Purus3"           "Putumayo"       "Roosevelt"      "Santiago"
## [73] "Solimoes1"        "Solimoes2"      "Solimoes3"      "Solimoes4"
## [77] "Solimoes7"        "Solimoes8"      "Solimoes9"      "Tapajos1"
## [81] "Tapajos2"         "Tapaua"          "Tefe"           "Teles_Pires"
## [85] "Tigre"            "Trombetas1"    "Uatumal1"       "Uaupes"
## [89] "Ucayali1"         "Ucayali2"       "Unini"          "Urubamba"
## [93] "Xingu1"           "Xingu3"          "Xingu4"         "Xingu6"
## [97] "Yacuma"

```

```

fish <- merge(x = fish, y = mar, by.x = "basin", by.y = "Basin2stat",
              all.x = TRUE, all.y = TRUE)

rm(fish2)

```

3.10 Hydromorphology type

Proportion of Hydrogeomorphology type according to Jezequel et al. “Geomorphological diversity of rivers in the Amazon Basin”

```
hydro = read.csv(paste("StatistiquesSubBasinAmazon_02052016/2023_Hydromorphology_Jezequel/",
  "2023_05_Types_Hydromorphology_97bv.csv", sep = ""), sep = ";",
  header = T)
summary(hydro)
```

```
##      BvNiv2          bv2stand       MajType      length_km
## Length:178      Length:178      Length:178      Min.   : 3.17
## Class :character  Class :character  Class :character 1st Qu.:111.93
## Mode  :character  Mode  :character  Mode  :character Median  :264.92
##                                     Mean   :360.83
##                                     3rd Qu.:462.16
##                                     Max.   :3135.74
```

```
dim(hydro)
```

```
## [1] 178   4
```

```
hydro <- acast(hydro, BvNiv2 ~ MajType, fun.aggregate = sum,
  value.var = "length_km")
hydro
```

```
##           A     B     C
## Abuna    13.4 235.0 0.0
## Amazon1 209.1 0.0  0.0
## Amazon2 552.3 0.0  0.0
## Amazon5 357.5 0.0  0.0
## Amazon8 191.9 0.0  0.0
## Amazon9 154.6 0.0  0.0
## Andira1 316.5 81.4 0.0
## Apaporis 22.7 276.6 717.4
## Apurimac1 0.0  0.0 225.4
## Arinos   240.0 263.9 0.0
## Aripuana 614.5 148.5 0.0
## Bacaja   176.3 91.1 0.0
## Beni     203.2 1006.9 415.3
## Blanco_Baures 28.9 181.3 0.0
## Branco   1073.9 685.8 0.0
## Candeias 11.0  60.5 0.0
## Canuma   317.9 367.8 0.0
## Caquetai 915.9 486.3 871.3
## Chapare  0.0  26.2 0.0
## Coari    133.3 72.7 0.0
## Curaray  71.7 428.4 9.9
## Curua    33.4 123.5 0.0
## Curua_Una 75.8 47.5 0.0
```

## Demini	113.5	545.5	0.0
## Fresco	181.7	141.9	0.0
## Grande	22.7	266.1	0.0
## Guapore	323.8	655.8	0.0
## Huallaga	272.9	43.2	655.3
## Iriri	1169.1	234.3	0.0
## Isiboro	3.2	239.6	0.0
## Itonamas	0.0	21.6	0.0
## Jamanxim	417.3	217.1	0.0
## Japura1	791.0	0.0	0.0
## Jari	184.8	9.8	0.0
## Jatapu	108.0	302.9	0.0
## Javary1	159.6	1645.0	0.0
## Jiparana	145.5	605.1	0.0
## Jurua1	353.2	2994.9	0.0
## Juruenal1	717.3	264.2	0.0
## Jutai	463.1	703.1	0.0
## Luna	467.5	387.2	0.0
## Madera1	331.7	0.0	0.0
## Madera2	232.2	163.1	0.0
## Madera3	342.4	0.0	0.0
## Madera4	480.0	0.0	0.0
## Madera6	0.0	111.4	0.0
## Madre_Dios	711.0	477.8	387.2
## Mamore1	34.6	497.9	0.0
## Mamore2	227.3	509.3	0.0
## Mantaro	0.0	0.0	25.3
## Maranon1	360.9	228.0	28.4
## Maranon3	295.4	231.5	248.4
## Maranon5	0.0	116.5	265.9
## Marmelos	0.0	140.8	0.0
## Maues	148.5	0.0	0.0
## Napo1	345.5	0.0	0.0
## Napo2	12.5	0.0	806.6
## Negro1	401.8	224.8	0.0
## Negro2	364.5	265.6	0.0
## Negro3	730.6	953.0	0.0
## Negro5	427.5	467.4	247.8
## Nhamunda	212.4	147.9	0.0
## Orthon	34.1	371.4	0.0
## Pachitea	19.4	0.0	478.1
## Paru_Este	177.1	90.3	0.0
## Pastaza	103.9	68.6	441.5
## Purus1	80.2	858.2	0.0
## Purus2	8.1	396.6	0.0
## Purus3	0.0	3135.7	0.0
## Putumayo	1424.6	53.2	318.3
## Roosevelt	174.2	307.2	0.0
## Santiago	0.0	9.6	388.4
## Solimoes1	396.9	0.0	0.0
## Solimoes2	239.5	105.1	0.0
## Solimoes3	161.6	0.0	0.0

```

## Solimoes4      575.9    0.0    0.0
## Solimoes7      356.6    0.0    0.0
## Solimoes8      442.8    0.0    0.0
## Solimoes9      192.5    0.0    0.0
## Tapajos1       509.9   10.1    0.0
## Tapajos2       471.5    0.0    0.0
## Tapaua          28.8   951.8    0.0
## Tefe            48.2    0.0    0.0
## Teles_Pires    838.9   205.5    0.0
## Tigre           78.5   568.2   72.3
## Trombetas1    1049.5   447.7    0.0
## Uatumal1       301.3   37.5    0.0
## Uaupes          353.5   41.1    0.0
## Ucayali1       1235.7  191.4  396.5
## Ucayali2       327.6    0.0  459.4
## Unini           168.0   318.4    0.0
## Urubamba        0.0    0.0  370.9
## Xingu1          578.7   53.6    0.0
## Xingu3          437.2    0.0    0.0
## Xingu4          737.9   108.8    0.0
## Xingu6          267.1   576.3    0.0
## Yacuma          0.0    9.5    0.0

```

```
hydro <- hydro/rowSums(hydro)
```

```
all.equal(fish$basin, rownames(hydro)) # all corresponding names are correct
```

```
## [1] TRUE
```

```
fish$basin == rownames(hydro) #they are all at the same order
```

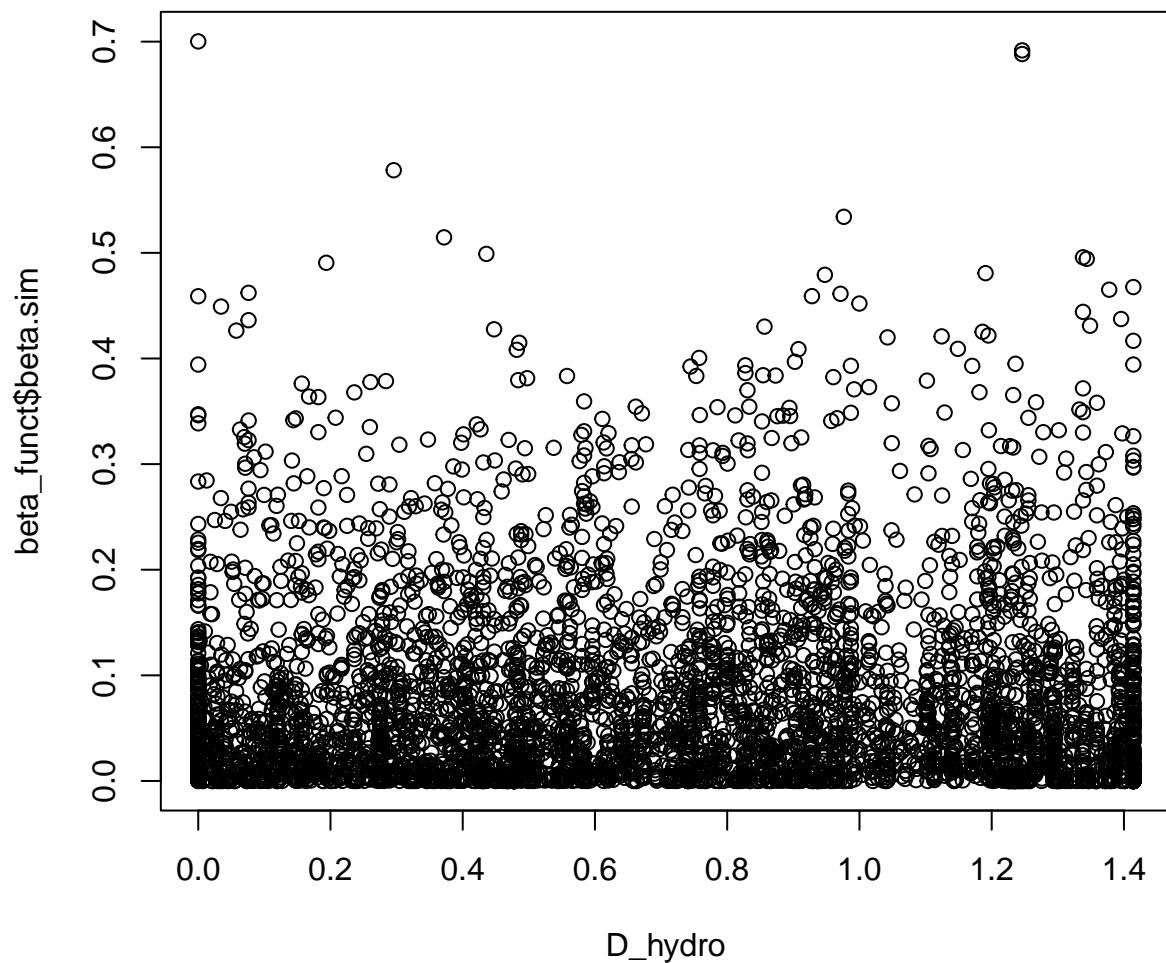
```

## [1] TRUE TRUE
## [16] TRUE TRUE
## [31] TRUE TRUE
## [46] TRUE TRUE
## [61] TRUE TRUE
## [76] TRUE TRUE
## [91] TRUE TRUE TRUE TRUE TRUE TRUE

```

```
D_hydro <- dist(hydro)
```

```
plot(beta_funct$beta.sim ~ D_hydro)
```



```
(beta_func$beta.sim ~ D_hydro)
```

```
## beta_func$beta.sim ~ D_hydro
```

3.11 Habitat variables

3.11.1 Setting and transforming variables

We will compose current habitat based on a series of variables, namely *River network Density*, *Surface of subdrainage basins (in km²)*, *Number of waterfall*, *Mean elevation along subdrainages*, *Elevation range*, and *Percentage of drainage basin above 1000m of altitude*. All continuos variables will be log-transformed (river network, surface, number of waterfalls, mean elevation, and elevation range) and scaled previously (i.e., mean = 0, and sd = 1) so that each variable has the same weight while computing distance matrix.

Habitat diversity (i.e., the proportion of vegetation type cover within subdrainages) will be calculated using the Euclidean difference between proportions in sites. *Water type* was coded differently.

We also computed a new variable related to the sampling effort of each subdrainage basin. For such, we divided the number of sampling points by the surface area of subdrainage so that we have a continuous sampling effort measure independent of the size of subdrainages. Let's start by separating and transforming variables.

```
names(fish)

## [1] "basin"                  "WaterColor"
## [3] "Area_km"                "Chut_hydrofall"
## [5] "Chut_grand"              "NetworkDensity"
## [7] "Elevation_min"            "Elevation_max"
## [9] "Elevation_range"          "Elevation_mean"
## [11] "Elevation_std"            "Elevation_basinSup1000m"
## [13] "Elevation_basinSup1500m"  "NbEspeceEnveloppe"
## [15] "DistanceEmbouchure_km"   "X23ma_cd"
## [17] "X23Ma"                  "X10Ma_cd"
## [19] "X10Ma"                  "NbSites1km_0417"
## [21] "NbSites_0417"            "Diversity0417"
## [23] "Endem0417"               "PebasLake"
## [25] "PebasConec"              "ShannonDiv"
## [27] "SoilDiv"                 "CSup15"
## [29] "CBelow15"                "Basin"
## [31] "codage"                   "Codigo"
## [33] "Sub_drainage"             "Inf25m"
## [35] "Ent25_100m"

# transforming some variables
fish$Area_log <- log(fish$Area_km + 1)
# there are only a few waterfalls on 'Chut_grand', so we
# opt for 'Chut_hydrofall'
fish$Chut_hydrofall_log <- log(fish$Chut_hydrofall + 1)
fish$ElevRge_log <- log(fish$Elevation_range + 1)
fish$ElevMean_log <- log(fish$Elevation_mean + 1)
fish$ShanDiv_log <- log(fish$ShannonDiv + 1)
fish$CSup15_log <- log(fish$CSup15 + 1)
fish$Inf25m_log <- log(fish$Inf25m + 1)

# computing sampling effort variable
fish$SamplingEff <- log(fish$NbSites_0417/fish$Area_km)
```

The variable *Percentage of each basin that is above 1000 meters* (i.e., `Elevation_.basinSup1000m`) is a proportion variable. First, we need to transform it in values between 0-1 by dividing by 100. So, in order to improve linearity, we applied the `arcsin(sqrt(x+1))` transformation. Such transformation has been suggested to deal with this kind of data (see Gotelli 2011; Legendre & Legendre 2012).

```
fish$ElevPro1000_asin <- arcsin(sqrt(fish$Elevation_basinSup1000m/100))
```

We excluded all variables not used in the analyses, and, as a further step, we standardized all variables used in statistical models. **Centering** and **reducing** predictors is important because

- variables are on different scales (km^2 , mm, $^{\circ}\text{C}$, etc);
- their range are very different (0-3000, 19-34, 0.1-0.9, etc);
- thus, coefficients from regression models could not be directly compared as a direct measure of strength.

After centering and reducing predictors, regression coefficients varies on the same scale and can be directly compared as a measure of effect size (Gelman 2007).

```
# Centering and reducing predictor so that mean=0 and sd=0.5
vari <- c(
  #tributary characteristics
  "NetworkDensity", "SamplingEff", "DistanceEmbouchure_km",
  "Area_log", "Chut_hydrofall_log", "codage", ##"Conec_index",
  "ElevRge_log", "ElevMean_log", "ElevPro1000_asin", "Inf25m_log",
  "CSup15_log"#, "ShanDiv_log", "SoilDiv"
)

# Standard function (centering and reducing; see first chunk)
pred <- data.frame(apply(fish[,vari], 2, Standard))
rownames(pred) <- fish$basin
colMeans(pred) #OK
```

##	NetworkDensity	SamplingEff	DistanceEmbouchure_km
##	1.2e-16	-1.6e-16	-5.0e-17
##	Area_log	Chut_hydrofall_log	codage
##	-4.5e-16	-3.5e-17	5.8e-17
##	ElevRge_log	ElevMean_log	ElevPro1000_asin
##	1.6e-17	8.1e-17	3.3e-17
##	Inf25m_log	CSup15_log	
##	-9.7e-18	-2.6e-17	

```
apply(pred, 2, sd) #OK
```

##	NetworkDensity	SamplingEff	DistanceEmbouchure_km
##	0.5	0.5	0.5
##	Area_log	Chut_hydrofall_log	codage
##	0.5	0.5	0.5
##	ElevRge_log	ElevMean_log	ElevPro1000_asin
##	0.5	0.5	0.5
##	Inf25m_log	CSup15_log	
##	0.5	0.5	

```

#setting predictor' names
colnames(pred) <- paste(colnames(pred), ".s", sep="")
pred$basin <- rownames(pred)

# fish <- fish[, !colnames(fish)%in%vari]

#combining fish biological data with transformed predictors
fish <- merge(fish,pred,by="basin")
summary(fish)

##      basin          WaterColor        Area_km     Chut_hydrofall
## Length:97       Length:97       Min.   : 5261    Min.   : 0.00
## Class :character  Class :character  1st Qu.: 28609   1st Qu.: 0.00
## Mode  :character  Mode  :character  Median  : 42905   Median  : 0.00
##                               Mean   : 60790   Mean   : 2.41
##                               3rd Qu.: 78109   3rd Qu.: 2.00
##                               Max.  :226666   Max.  :26.00
##      Chut_grand    NetworkDensity  Elevation_min  Elevation_max
## Min.   :0.0000    Min.   :0.0596   Min.   : 0.0    Min.   : 57
## 1st Qu.:0.0000   1st Qu.:0.1289   1st Qu.: 12.0   1st Qu.: 239
## Median :0.0000   Median :0.1448   Median : 47.0   Median : 645
## Mean   :0.0825   Mean   :0.1441   Mean   : 71.8   Mean   :1457
## 3rd Qu.:0.0000   3rd Qu.:0.1613   3rd Qu.:110.0   3rd Qu.:1759
## Max.   :2.0000   Max.   :0.2265   Max.   :470.0   Max.   :6247
##      Elevation_range Elevation_mean Elevation_std Elevation_basinSup1000m
## Min.   : 53       Min.   : 26.2    Min.   : 10.5   Min.   : 0.00
## 1st Qu.: 195     1st Qu.: 106.4   1st Qu.: 34.2   1st Qu.: 0.00
## Median : 570     Median : 182.5   Median : 72.7   Median : 0.00
## Mean   :1385     Mean   : 399.0   Mean   :247.4   Mean   : 8.27
## 3rd Qu.:1619     3rd Qu.: 313.6   3rd Qu.:166.0   3rd Qu.: 0.58
## Max.   :6082     Max.   :3882.3   Max.   :1584.0   Max.   :99.23
##      Elevation_basinSup1500m NbEspeceEnveloppe DistanceEmbouchure_km
## Min.   : 0.0000    Min.   : 18       Min.   : 0
## 1st Qu.: 0.0000    1st Qu.: 543     1st Qu.:1427
## Median : 0.0000    Median : 761     Median :2264
## Mean   : 6.6002    Mean   : 686     Mean   :2443
## 3rd Qu.: 0.0505    3rd Qu.: 915     3rd Qu.:3301
## Max.   :97.9752    Max.   :1099     Max.   :5793
##      X23ma_cd        X23Ma        X10Ma_cd        X10Ma
## Length:97       Length:97       Length:97       Length:97
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
## 
## 
## 
##      NbSites1km_0417 NbSites_0417 Diversity0417 Endem0417      PebasLake
## Min.   : 4       Min.   : 4       Min.   : 7       Min.   : 0.00    Min.   :0.000
## 1st Qu.: 46      1st Qu.: 65      1st Qu.:195     1st Qu.: 0.00    1st Qu.:0.000
## Median :103      Median :154      Median :302      Median : 1.00    Median :1.000
## Mean   :123      Mean   :202      Mean   :332      Mean   : 3.08    Mean   :0.588
## 3rd Qu.:162      3rd Qu.:302     3rd Qu.:486     3rd Qu.: 4.00    3rd Qu.:1.000

```

```

## Max. :406      Max. :900      Max. :786      Max. :16.00     Max. :1.000
## PebasConec      ShannonDiv      SoilDiv      CSup15
## Length:97       Min. :0.0603     Min. :0.000     Min. : 0.00
## Class :character 1st Qu.:0.2697    1st Qu.:0.543    1st Qu.: 0.00
## Mode :character   Median :0.5204    Median :0.757    Median : 1.45
##                           Mean :0.6135    Mean :0.818    Mean :14.02
##                           3rd Qu.:0.9333   3rd Qu.:1.145   3rd Qu.:16.18
##                           Max. :2.0395    Max. :1.810    Max. :96.02
## CBelow15        Basin          codage        Código
## Min. : 2.88     Length:97       Min. : 1.0      Min. : 1.0
## 1st Qu.: 83.82    Class :character 1st Qu.: 7.0    1st Qu.: 7.0
## Median : 98.47    Mode :character  Median :12.0    Median :12.0
## Mean : 85.86           Median :12.0    Mean :11.7    Mean :11.7
## 3rd Qu.:100.00           Mean :11.7    3rd Qu.:15.0    3rd Qu.:15.0
## Max. :100.00           Max. :24.0    Max. :24.0    Max. :24.0
## Sub_drainage      Inf25m        Ent25_100m    Area_log
## Length:97         Min. : 0.00     Min. : 0.0      Min. : 8.57
## Class :character  1st Qu.: 0.00     1st Qu.: 0.0    1st Qu.:10.26
## Mode :character   Median : 0.00     Median : 6.9    Median :10.67
##                           Mean : 4.11     Mean : 29.8    Mean :10.76
##                           3rd Qu.: 2.00    3rd Qu.: 55.3   3rd Qu.:11.27
##                           Max. :68.00     Max. :100.0   Max. :12.33
## Chut_hydrofall_log ElevRge_log    ElevMean_log  ShanDiv_log
## Min. :0.000       Min. :3.99      Min. :3.30      Min. :0.0586
## 1st Qu.:0.000     1st Qu.:5.28      1st Qu.:4.68    1st Qu.:0.2388
## Median :0.000     Median :6.35      Median :5.21    Median :0.4189
## Mean :0.652       Mean :6.39      Mean :5.32    Mean :0.4428
## 3rd Qu.:1.099     3rd Qu.:7.39      3rd Qu.:5.75    3rd Qu.:0.6592
## Max. :3.296       Max. :8.71      Max. :8.26    Max. :1.1117
## CSup15_log        Inf25m_log    SamplingEff    ElevPro1000_asin
## Min. :0.000       Min. :0.000     Min. :-9.19     Min. :0.0000
## 1st Qu.:0.000     1st Qu.:0.000     1st Qu.:-6.48   1st Qu.:0.0000
## Median :0.896     Median :0.000     Median :-5.98   Median :0.0000
## Mean :1.502       Mean :0.618      Mean :-5.84   Mean :0.1454
## 3rd Qu.:2.844     3rd Qu.:1.099     3rd Qu.:-5.22   3rd Qu.:0.0762
## Max. :4.575       Max. :4.234      Max. :-2.66   Max. :1.4832
## NetworkDensity.s SamplingEff.s    DistanceEmbouchure_km.s Area_log.s
## Min. : -1.3066    Min. : -1.5668    Min. : -0.8897    Min. : -1.4765
## 1st Qu.: -0.2348   1st Qu.: -0.2964   1st Qu.: -0.3700   1st Qu.: -0.3337
## Median : 0.0108    Median : -0.0615    Median : -0.0651   Median : -0.0601
## Mean : 0.0000      Mean : 0.0000      Mean : 0.0000    Mean : 0.0000
## 3rd Qu.: 0.2665    3rd Qu.: 0.2935    3rd Qu.: 0.3126   3rd Qu.: 0.3442
## Max. : 1.2752      Max. : 1.4906      Max. : 1.2201   Max. : 1.0633
## Chut_hydrofall_log.s codage.s      ElevRge_log.s  ElevMean_log.s
## Min. : -0.348     Min. : -0.759     Min. : -0.8917   Min. : -0.9572
## 1st Qu.: -0.348    1st Qu.: -0.335    1st Qu.: -0.4127   1st Qu.: -0.3043
## Median : -0.348    Median :  0.019    Median : -0.0154   Median : -0.0498
## Mean : 0.0000      Mean : 0.0000      Mean : 0.0000    Mean : 0.0000
## 3rd Qu.: 0.238     3rd Qu.: 0.231     3rd Qu.: 0.3721   3rd Qu.: 0.2066
## Max. : 1.410       Max. : 0.868      Max. : 0.8637   Max. : 1.4007
## ElevPro1000_asin.s Inf25m_log.s  CSup15_log.s
## Min. : -0.227     Min. : -0.272     Min. : -0.476

```

```

## 1st Qu.:-0.227      1st Qu.:-0.272      1st Qu.:-0.476
## Median :-0.227      Median :-0.272      Median :-0.192
## Mean   : 0.000      Mean   : 0.000      Mean   : 0.000
## 3rd Qu.:-0.108      3rd Qu.: 0.211      3rd Qu.: 0.426
## Max.   : 2.088      Max.   : 1.591      Max.   : 0.975

```

```

#setting rownames
rownames(fish) <- fish$basin

rm(pred,vari)

```

3.11.2 Computing dissimilarity matrices

Now it's time to compute dissimilarity matrix for Current habitat variables. For most of habitat variables, we will use the *Euclidean* distance to calculate dissimilarity among sampling sites. As water color is categorical (three levels: white, black, and clear), we used our own function to calculate the dissimilarity matrix. The rationale for WaterColor distance matrix is if two sites have exactly the same water type, we coded them 1; if their water type differ, them the code is 0. Therefore, we expected low fish dissimilarity levels among subdrainages with the same water types (conversely, high similarity among subdrainages with the same water type).

```

vari <- c(
  "NetworkDensity.s", "Area_log.s"##"DistanceEmbouchure_km.s",##"codage.s",
  ##"ShanDiv_log.s", "SoilDiv.s"
  )##"ElevRge_log.s"
#Euclidean distance
D_Habitat_size <- dist(fish[,vari])
rm(vari)

#positioning
vari <- c("DistanceEmbouchure_km.s", "codage.s")
#Euclidean distance
D_Habitat_position <- dist(fish[,vari])

#Habitat fragmentation = NumberWaterfalls
vari <- c("Chut_hydrofall_log.s")
D_Habitat_fragm <- dist(fish[,vari,drop=FALSE])
rm(vari)

#Habitat harshness = Elevation + CSum_15_log.s
vari <- c("CSup15_log.s",
         "ElevMean_log.s", "ElevRge_log.s", "ElevPro1000_asin.s")
D_Habitat_harsh <- dist(fish[,vari])
rm(vari)
# plot(D_Habitat_harsh~D_Habitat_fragm)
# cor(as.vector(D_Habitat_harsh),as.vector(D_Habitat_fragm))

```

##Habitat Diversity and Soil diversity

habitat2\$basin

```

## [1] "Abuna"          "Amazon1"        "Amazon2"        "Amazon5"
## [5] "Amazon8"        "Amazon9"         "Andira1"        "Apaporis"
## [9] "Apurimac1"      "Arinos"          "Aripuana"       "Bacaja"
## [13] "Beni"           "Blanco_Baures"  "Branco"         "Candeias"
## [17] "Canuma"          "Caqueta1"       "Chapare"        "Coari"
## [21] "Curaray"         "Curua"          "Curua_Una"     "Demini"
## [25] "Fresco"          "Grande"          "Guapore"        "Huallaga"
## [29] "Iriri"           "Isiboro"         "Itonamas"       "Jamanxim"
## [33] "Japura1"         "Jari"            "Jatapu"         "Javary1"
## [37] "Jiparana"        "Jurua1"          "Juruena1"       "Jutai"
## [41] "Luna"            "Madera1"        "Madera2"        "Madera3"
## [45] "Madera4"         "Madera6"        "Madre_Dios"    "Mamore1"
## [49] "Mamore2"         "Mantaro"         "Maranon1"       "Maranon3"
## [53] "Maranon5"        "Marmelos"        "Maues"          "Napo1"
## [57] "Napo2"           "Negro1"          "Negro2"         "Negro3"
## [61] "Negro5"          "Nhamunda"        "Orthon"         "Pachitea"
## [65] "Paru_Este"        "Pastaza"         "Purus1"         "Purus2"
## [69] "Purus3"          "Putumayo"        "Roosevelt"      "Santiago"
## [73] "Solimoes1"        "Solimoes2"       "Solimoes3"      "Solimoes4"
## [77] "Solimoes7"        "Solimoes8"       "Solimoes9"      "Tapajos1"
## [81] "Tapajos2"         "Tapaua"          "Tefe"           "Teles_Pires"
## [85] "Tigre"            "Trombetas1"      "Uatumal"        "Uaupes"
## [89] "Ucayali1"         "Ucayali2"        "Unini"          "Urubamba"
## [93] "Xingu1"           "Xingu3"          "Xingu4"         "Xingu6"
## [97] "Yacuma"

```

Soil2\$basin

```

## [1] "Abuna"          "Amazon1"        "Amazon2"        "Amazon5"
## [5] "Amazon8"        "Amazon9"         "Andira1"        "Apaporis"
## [9] "Apurimac1"      "Arinos"          "Aripuana"       "Bacaja"
## [13] "Beni"           "Blanco_Baures"  "Branco"         "Candeias"
## [17] "Canuma"          "Caqueta1"       "Chapare"        "Coari"
## [21] "Curaray"         "Curua"          "Curua_Una"     "Demini"
## [25] "Fresco"          "Grande"          "Guapore"        "Huallaga"
## [29] "Iriri"           "Isiboro"         "Itonamas"       "Jamanxim"
## [33] "Japura1"         "Jari"            "Jatapu"         "Javary1"
## [37] "Jiparana"        "Jurua1"          "Juruena1"       "Jutai"
## [41] "Luna"            "Madera1"        "Madera2"        "Madera3"
## [45] "Madera4"         "Madera6"        "Madre_Dios"    "Mamore1"
## [49] "Mamore2"         "Mantaro"         "Maranon1"       "Maranon3"
## [53] "Maranon5"        "Marmelos"        "Maues"          "Napo1"
## [57] "Napo2"           "Negro1"          "Negro2"         "Negro3"
## [61] "Negro5"          "Nhamunda"        "Orthon"         "Pachitea"
## [65] "Paru_Este"        "Pastaza"         "Purus1"         "Purus2"
## [69] "Purus3"          "Putumayo"        "Roosevelt"      "Santiago"
## [73] "Solimoes1"        "Solimoes2"       "Solimoes3"      "Solimoes4"
## [77] "Solimoes7"        "Solimoes8"       "Solimoes9"      "Tapajos1"

```

```

## [81] "Tapajos2"      "Tapaua"        "Tefe"          "Teles_Pires"
## [85] "Tigre"         "Trombetas1"    "Uatuma1"       "Uaupes"
## [89] "Ucayali1"      "Ucayali2"      "Unini"         "Urubamba"
## [93] "Xingu1"        "Xingu3"        "Xingu4"        "Xingu6"
## [97] "Yacuma"

```

```
table(rownames(habitat2)==fish$basin)
```

```

## 
## TRUE
## 97

```

```
table(rownames(Soil2)==fish$basin)
```

```

## 
## TRUE
## 97

```

```
D_Habitat_Div<-dist(cbind(habitat2[,-1],Soil2[,-1]))
```

#Water color: if both sites have the same water color = 0, otherwise = 1

```

vari <- c("WaterColor")
mat_hab <- data.frame(WC=fish[,vari])
rownames(mat_hab) <- rownames(fish)
head(mat_hab,10)

```

```

##           WC
## Abuna     White
## Amazon1   White
## Amazon2   White
## Amazon5   White
## Amazon8   White
## Amazon9   White
## Andira1   White
## Apaporis  White
## Apurimac1 White
## Arinos    Clear

```

#code expressing Basin Water Color similar to to Pebas basin

```

WaterColor <- matrix(0,ncol=97,nrow=97,dimnames = list(fish$basin,fish$basin))
WaterColor[1:10,1:10]

```

	Abuna	Amazon1	Amazon2	Amazon5	Amazon8	Amazon9	Andira1	Apaporis
## Abuna	0	0	0	0	0	0	0	0
## Amazon1	0	0	0	0	0	0	0	0
## Amazon2	0	0	0	0	0	0	0	0
## Amazon5	0	0	0	0	0	0	0	0
## Amazon8	0	0	0	0	0	0	0	0

```

## Amazon9      0      0      0      0      0      0      0      0
## Andira1     0      0      0      0      0      0      0      0
## Apaporis    0      0      0      0      0      0      0      0
## Apurimac1   0      0      0      0      0      0      0      0
## Arinos      0      0      0      0      0      0      0      0
##          Apurimac1 Arinos
## Abuna        0      0
## Amazon1     0      0
## Amazon2     0      0
## Amazon5     0      0
## Amazon8     0      0
## Amazon9     0      0
## Andira1     0      0
## Apaporis    0      0
## Apurimac1   0      0
## Arinos      0      0

for(i in unique(fish$basin)){
  for(k in unique(fish$basin)){
    #i="Abuna";k="Apurimac1"      #White-White
    #i="Abuna";k="Teles_Pires"    #White-Clear
    #i="Xingu3";k="Negro1"       ##Clear-Black
    II<-as.vector(mat_hab[i,"WC"])
    KK<-as.vector(mat_hab[k,"WC"])
    if((II == KK) ) { #same Water Color
      WaterColor[i,k]<-1
    } else { #any distinct Water Color combination
      WaterColor[i,k]<-0
    }
  }
}

D_Habitat_water <- as.dist(WaterColor)
rm(vari, mat_hab,II,KK,WaterColor)

#Sampling effort = log(Density of sampling sites per subdrainage)
vari <- c("SamplingEff.s")
mat_hab <- as.data.frame(fish[,vari])
rownames(mat_hab) <- rownames(fish)
D_Habitat_SampEff <- dist(mat_hab)
rm(vari, mat_hab)

#marine incursion
vari <- c("Inf25m")
mat_hab <- as.data.frame(fish[,vari])
rownames(mat_hab) <- rownames(fish)
colnames(mat_hab) <- "Inf25m"
D_Marine_Incurs <- dist(mat_hab)

```

```
rm(vari, mat_hab)
```

4 Past-Connectivity among subdrainages as a Palaeo-Connectivity matrix

The geological history of Amazon basin has been depicted in a recent research paper(Hoorn *et al.* 2010). We extracted information on past connectivity among subdrainages in two time slices (10million years ago, and 23mya). Then, we used this past connectivity information to construct a binary connectivity matrix depicting history. For such, we need to first create an empty matrix with row and column names corresponding to the current subdrainages, and then we use information from palaeo-connection to fill matrix cells for those subdrainages that were connected 10mya.

```
# information from Hoorn et al. (2010) creating new
# variable based on past connection subbasins with the same
# number were connected
conec <- data.frame(basin = fish$basin, ma23 = as.numeric(as.factor(fish$X23Ma)),
  ma10 = as.numeric(as.factor(fish$X10Ma)))
head(conec, 40)
```

	basin	ma23	ma10
## 1	Abuna	11	3
## 2	Amazon1	2	4
## 3	Amazon2	2	4
## 4	Amazon5	2	4
## 5	Amazon8	2	4
## 6	Amazon9	2	4
## 7	Andira1	2	12
## 8	Apaporis	4	8
## 9	Apurimac1	10	2
## 10	Arinos	2	14
## 11	Aripuana	2	10
## 12	Bacaja	2	12
## 13	Beni	11	3
## 14	Blanco_Baures	11	3
## 15	Branco	5	6
## 16	Candeias	13	10
## 17	Canuma	2	14
## 18	Caqueta1	14	8
## 19	Chapare	11	15
## 20	Coari	13	4
## 21	Curaray	14	9
## 22	Curua	1	5
## 23	Curua_Una	2	12
## 24	Demini	5	6
## 25	Fresco	3	11
## 26	Grande	11	15
## 27	Guapore	11	3
## 28	Huallaga	9	1
## 29	Iriri	2	12
## 30	Isiboro	11	15
## 31	Itonamas	11	15
## 32	Jamanxim	2	12
## 33	Japura1	14	7
## 34	Jari	1	5

```

## 35      Jatapu    1    5
## 36      Javary1   14   4
## 37      Jiparana   2   10
## 38      Jurua1    13   4
## 39      Juruena1  2   14
## 40      Jutai     13   4

# connectivity at 23ma
mat23 <- matrix(0, ncol = 97, nrow = 97, dimnames = list(fish$basin,
  fish$basin))
mat23[1:10, 1:10]

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8 Amazon9 Andira1 Apaporis
## Abuna      0      0      0      0      0      0      0      0
## Amazon1    0      0      0      0      0      0      0      0
## Amazon2    0      0      0      0      0      0      0      0
## Amazon5    0      0      0      0      0      0      0      0
## Amazon8    0      0      0      0      0      0      0      0
## Amazon9    0      0      0      0      0      0      0      0
## Andira1    0      0      0      0      0      0      0      0
## Apaporis   0      0      0      0      0      0      0      0
## Apurimac1  0      0      0      0      0      0      0      0
## Arinos     0      0      0      0      0      0      0      0
##          Apurimac1 Arinos
## Abuna      0      0
## Amazon1   0      0
## Amazon2   0      0
## Amazon5   0      0
## Amazon8   0      0
## Amazon9   0      0
## Andira1   0      0
## Apaporis  0      0
## Apurimac1 0      0
## Arinos    0      0

unlist(lapply(seq_along(unique(conec$ma23)), function(x) {
  b <- conec[conec$ma23 == x, "basin"]
  mat23[b, b] <- 1
  rm(b)
  return()
}))

## NULL

# the result
mat23[1:10, 1:10]

```

```

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8 Amazon9 Andira1 Apaporis
## Abuna      1      0      0      0      0      0      0      0
## Amazon1    0      1      1      1      1      1      1      0

```

```

## Amazon2      0      1      1      1      1      1      1      0
## Amazon5      0      1      1      1      1      1      1      0
## Amazon8      0      1      1      1      1      1      1      0
## Amazon9      0      1      1      1      1      1      1      0
## Andira1      0      1      1      1      1      1      1      0
## Apaporis     0      0      0      0      0      0      0      1
## Apurimac1    0      0      0      0      0      0      0      0
## Arinos       0      1      1      1      1      1      1      0
##          Apurimac1 Arinos
## Abuna        0      0
## Amazon1      0      1
## Amazon2      0      1
## Amazon5      0      1
## Amazon8      0      1
## Amazon9      0      1
## Andira1      0      1
## Apaporis     0      0
## Apurimac1    1      0
## Arinos       0      1

```

```

# connectivity at 10ma
mat10 <- matrix(0, ncol = 97, nrow = 97, dimnames = list(fish$basin,
  fish$basin))
mat10[1:10, 1:10]

```

```

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8 Amazon9 Andira1 Apaporis
## Abuna      0      0      0      0      0      0      0      0
## Amazon1    0      0      0      0      0      0      0      0
## Amazon2    0      0      0      0      0      0      0      0
## Amazon5    0      0      0      0      0      0      0      0
## Amazon8    0      0      0      0      0      0      0      0
## Amazon9    0      0      0      0      0      0      0      0
## Andira1    0      0      0      0      0      0      0      0
## Apaporis   0      0      0      0      0      0      0      0
## Apurimac1  0      0      0      0      0      0      0      0
## Arinos     0      0      0      0      0      0      0      0
##          Apurimac1 Arinos
## Abuna      0      0
## Amazon1    0      0
## Amazon2    0      0
## Amazon5    0      0
## Amazon8    0      0
## Amazon9    0      0
## Andira1    0      0
## Apaporis   0      0
## Apurimac1  0      0
## Arinos     0      0

```

```

unlist(lapply(seq_along(unique(conec$ma10)), function(x) {
  b <- conec[conec$ma10 == x, "basin"]
  mat10[b, b] <- 1
})

```

```

    rm(b)
    return()
}))

## NULL

# the result
mat10[1:10, 1:10]

##          Abuna Amazon1 Amazon2 Amazon5 Amazon8 Amazon9 Andira1 Apaporis
## Abuna      1      0      0      0      0      0      0      0
## Amazon1    0      1      1      1      1      1      0      0
## Amazon2    0      1      1      1      1      1      0      0
## Amazon5    0      1      1      1      1      1      0      0
## Amazon8    0      1      1      1      1      1      0      0
## Amazon9    0      1      1      1      1      1      0      0
## Andira1    0      0      0      0      0      0      1      0
## Apaporis   0      0      0      0      0      0      0      1
## Apurimac1  0      0      0      0      0      0      0      0
## Arinos     0      0      0      0      0      0      0      0
##          Apurimac1 Arinos
## Abuna      0      0
## Amazon1   0      0
## Amazon2   0      0
## Amazon5   0      0
## Amazon8   0      0
## Amazon9   0      0
## Andira1   0      0
## Apaporis  0      0
## Apurimac1 1      0
## Arinos    0      1

```

Now we transform these square matrices in **dist** objects so that we can proceed with dissimilarity analysis.

```

D_pastConnect <- list(C0mya = as.dist(mat0), C10mya = as.dist(mat10),
                      C23mya = as.dist(mat23), Pebas23mya = as.dist(matPebas23Mya))

rm(mat10, mat23, cone)

```

5 Charging *Spatial Distance Matrix*

We computed the distance among pairs of subdrainages following the riverine connections among them. We use such spatial matrix connectivity to create *distance-decay plots*. Charging the spatial matrix:

```
spa_mat = read.csv(paste("StatistiquesSubBasinAmazon_02052016/DistanceMatrixRiver.csv",
  sep = ""), sep = ";", header = T)

# setting matrix row names and colum
rownames(spa_mat) <- spa_mat$MatriceDistance_km
spa_mat <- spa_mat[, -c(1)]/1000 #to give distance in Km

# checking rownames and colnames order
all.equal(rownames(spa_mat), colnames(spa_mat)) #ok!

## [1] TRUE

# checking lower and upper triangles
table(as.dist(spa_mat) == as.dist(t(spa_mat))) #ok!

## 
## TRUE
## 4656
```

Now, we separate only those river subdrainages used in the modelling:

```
spa_mat <- spa_mat[rownames(spa_mat) %in% fish$basin, colnames(spa_mat) %in%
  fish$basin]

# checking again matrix structure: rownames and colnames
# order
all.equal(rownames(spa_mat), colnames(spa_mat)) #ok!

## [1] TRUE

# Are fish and spa_mat on the same order?
table(rownames(spa_mat) == fish$basin)

## 
## TRUE
## 97

table(colnames(spa_mat) == fish$basin)

## 
## TRUE
## 97
```

```
# checking lower and upper triangles
table(as.dist(spa_mat) == as.dist(t(spa_mat))) #GREAT!

##
## TRUE
## 4656

# transforming in a distance object
D_spat <- as.dist(spa_mat)

rm(spa_mat)
```

5.1 Checking the structure of ALL Distance Matrices

Before modelling, it is important to check if **all distance matrices have the same data structure**. *Data structure* means that all distance matrices have identical row and column names, and that rows and columns are exactly on the same order. This is important because as we advance in the modelling procedure, we must be sure that fish dissimilarity between Amazon1-Jari subdrainages, for instance, will be compared with habitat, current and past climate, and geographical distances for the same pair of subdrainages. Therefore, time has come to compare the structure of all dissimilarity matrices.

```
#matrices: D_CurrentLGM, D_energ, D_conec
#           D_Habitat_size, D_Habitat_Div, D_Habitat_fragm, D_Habitat_harsh,
#           D_Habitat_water,
#           D_Habitat_SampEff, D_Marine_Incurs,
#           D_temp, D_water; and F_Bsor, F_Bsim, and F_Bsne, PBsne
F_Bsor <- D_fish$beta.sor
F_Bsim <- D_fish$beta.sim
F_Bsne <- D_fish$beta.sne
C_10mya <- D_pastConnect$C10mya
C_23mya <- D_pastConnect$C23mya
C_Pebas23mya <- D_pastConnect$Pebas23mya
Fun_Bsor<-beta_funct$beta.sor#functional BDiv
Fun_Bsim<-beta_funct$beta.sim
Fun_Bsne<-beta_funct$beta.sne
as.matrix(PBsimCasse)[1:20,1:20]
```

	Abuna	Amazon1	Amazon2	Amazon5	Amazon8	Amazon9	Andira1	Apaporis	
## Abuna	0.000	0.361	0.093	0.092	0.110	0.132	0.289	0.350	
## Amazon1	0.361	0.000	0.053	0.055	0.080	0.096	0.256	0.441	
## Amazon2	0.093	0.053	0.000	0.108	0.132	0.124	0.077	0.096	
## Amazon5	0.092	0.055	0.108	0.000	0.114	0.124	0.049	0.076	
## Amazon8	0.110	0.080	0.132	0.114	0.000	0.113	0.061	0.106	
## Amazon9	0.132	0.096	0.124	0.124	0.113	0.000	0.076	0.120	
## Andira1	0.289	0.256	0.077	0.049	0.061	0.076	0.000	0.195	
## Apaporis	0.350	0.441	0.096	0.076	0.106	0.120	0.195	0.000	
## Apurimac1	0.145	0.216	0.157	0.171	0.191	0.180	0.318	0.220	
## Arinos	0.326	0.318	0.108	0.115	0.166	0.198	0.290	0.310	
## Aripuana	0.091	0.086	0.160	0.185	0.161	0.165	0.095	0.076	
## Bacaja	0.341	0.393	0.119	0.145	0.167	0.206	0.318	0.367	
## Beni	0.102	0.218	0.182	0.211	0.247	0.254	0.270	0.239	
## Blanco_Baures	0.167	0.263	0.137	0.127	0.164	0.187	0.283	0.250	
## Branco	0.116	0.078	0.140	0.138	0.122	0.112	0.068	0.038	
## Candeias	0.103	0.146	0.148	0.181	0.184	0.215	0.164	0.127	
## Canuma	0.344	0.325	0.083	0.090	0.072	0.102	0.253	0.235	
## Caquetá1	0.165	0.166	0.182	0.200	0.225	0.248	0.177	0.072	
## Chapare	0.272	0.358	0.139	0.147	0.168	0.174	0.355	0.394	
## Coari	0.226	0.199	0.074	0.072	0.063	0.089	0.167	0.182	
##		Apurimac1	Arinos	Aripuana	Bacaja	Beni	Blanco_Baures	Branco	
## Abuna		0.145	0.326	0.091	0.34	0.102		0.17	0.116
## Amazon1		0.216	0.318	0.086	0.39	0.218		0.26	0.078
## Amazon2		0.157	0.108	0.160	0.12	0.182		0.14	0.140
## Amazon5		0.171	0.115	0.185	0.14	0.211		0.13	0.138
## Amazon8		0.191	0.166	0.161	0.17	0.247		0.16	0.122

## Amazon9	0.180	0.198	0.165	0.21	0.254	0.19	0.112
## Andira1	0.318	0.290	0.095	0.32	0.270	0.28	0.068
## Apaporis	0.220	0.310	0.076	0.37	0.239	0.25	0.038
## Apurimac1	0.000	0.211	0.143	0.24	0.049	0.11	0.136
## Arinos	0.211	0.000	0.094	0.24	0.244	0.19	0.089
## Aripuana	0.143	0.094	0.000	0.13	0.198	0.17	0.115
## Bacaja	0.239	0.239	0.133	0.00	0.241	0.25	0.118
## Beni	0.049	0.244	0.198	0.24	0.000	0.12	0.200
## Blanco_Baures	0.115	0.192	0.168	0.25	0.121	0.00	0.148
## Branco	0.136	0.089	0.115	0.12	0.200	0.15	0.000
## Candeias	0.143	0.186	0.125	0.21	0.213	0.17	0.136
## Canuma	0.216	0.274	0.086	0.31	0.293	0.33	0.068
## Caqueta1	0.064	0.158	0.168	0.18	0.223	0.20	0.136
## Chapare	0.093	0.335	0.150	0.35	0.040	0.14	0.185
## Coari	0.216	0.265	0.089	0.26	0.248	0.28	0.100
##	Candeias	Canuma	Caqueta1	Chapare	Coari		
## Abuna	0.10	0.344	0.165	0.272	0.226		
## Amazon1	0.15	0.325	0.166	0.358	0.199		
## Amazon2	0.15	0.083	0.182	0.139	0.074		
## Amazon5	0.18	0.090	0.200	0.147	0.072		
## Amazon8	0.18	0.072	0.225	0.168	0.063		
## Amazon9	0.21	0.102	0.248	0.174	0.089		
## Andira1	0.16	0.253	0.177	0.355	0.167		
## Apaporis	0.13	0.235	0.072	0.394	0.182		
## Apurimac1	0.14	0.216	0.064	0.093	0.216		
## Arinos	0.19	0.274	0.158	0.335	0.265		
## Aripuana	0.12	0.086	0.168	0.150	0.089		
## Bacaja	0.21	0.314	0.184	0.354	0.257		
## Beni	0.21	0.293	0.223	0.040	0.248		
## Blanco_Baures	0.17	0.331	0.201	0.139	0.279		
## Branco	0.14	0.068	0.136	0.185	0.100		
## Candeias	0.00	0.177	0.191	0.140	0.133		
## Canuma	0.18	0.000	0.206	0.399	0.219		
## Caqueta1	0.19	0.206	0.000	0.148	0.151		
## Chapare	0.14	0.399	0.148	0.000	0.248		
## Coari	0.13	0.219	0.151	0.248	0.000		

```

# creating a function to compare dimensions,
# colnames, and rownames among pairs of dissimilarity matrices
check_matrix_pair <- function(x,y){
  # mats
  x <- as.matrix(x)
  y <- as.matrix(y)

  #matrix dimention
  res <- all(c(all(dim(x) == dim(y)) ,
  #matrix colnames
  all(colnames(x) == colnames(y)) ,
  #matrix rownames
  all(rownames(x) == rownames(y)) ) )
  return(res)
}

```

```

#####
### testing the function
#1. different column order
fake_mat <- matrix(NA,3,7,dimnames = list(1:3,1:7));fake_mat

##    1 2 3 4 5 6 7
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA

fake_mat2 <- fake_mat[,sample(ncol(fake_mat))];fake_mat2

##    2 6 5 7 3 1 4
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA

check_matrix_pair(fake_mat,fake_mat2)

## [1] FALSE

#2. identical matrices
fake_mat <- matrix(NA,3,7,dimnames = list(1:3,1:7));fake_mat

##    1 2 3 4 5 6 7
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA

fake_mat2 <- fake_mat;fake_mat2

##    1 2 3 4 5 6 7
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA

check_matrix_pair(fake_mat,fake_mat2)

## [1] TRUE

#3. different row order
fake_mat <- matrix(NA,3,7,dimnames = list(1:3,1:7));fake_mat

##    1 2 3 4 5 6 7
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA

```

```

fake_mat2 <- fake_mat[sample(nrow(fake_mat)),];fake_mat2

##    1 2 3 4 5 6 7
## 3 NA NA NA NA NA NA
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA

check_matrix_pair(fake_mat,fake_mat2)

## [1] FALSE

#4. different matrix dimension
fake_mat <- matrix(NA,3,7,dimnames = list(1:3,1:7));fake_mat

##    1 2 3 4 5 6 7
## 1 NA NA NA NA NA NA
## 2 NA NA NA NA NA NA
## 3 NA NA NA NA NA NA

fake_mat2 <- fake_mat[,1:5];fake_mat2

##    1 2 3 4 5
## 1 NA NA NA NA NA
## 2 NA NA NA NA NA
## 3 NA NA NA NA NA

check_matrix_pair(fake_mat,fake_mat2)

## Warning in colnames(x) == colnames(y): longer object length is not a multiple
## of shorter object length

## [1] FALSE

#great, it works perfectly!
rm(fake_mat2)
#####
# Comparing the structure of all distance matrices
# list of all distance matrices
#matrices: D_CurrentLGM, D_energ, D_conec
#          D_Habitat_size, D_Habitat_Div, D_Habitat_fragm, D_Habitat_harsh,
#          D_Habitat_water,
#          D_Habitat_SampEff, D_Marine_Incurs,
#          D_temp, D_water; and F_Bsor, F_Bsim, and F_Bsne

mats <- list(D_CurrentLGM = D_CurrentLGM , D_energ = D_energ ,

```

```

D_Habitat_size = D_Habitat_size ,
D_Habitat_Div = D_Habitat_Div ,
D_Habitat_harsh = D_Habitat_harsh,
D_Marine_Incurs = D_Marine_Incurs,
D_Habitat_fragm = D_Habitat_fragm,
D_Habitat_water = D_Habitat_water ,
D_Habitat_SampEff = D_Habitat_SampEff ,
D_temp = D_temp,
D_Habitat_position = D_Habitat_position,
D_water = D_water , F_Bsor = F_Bsor ,
F_Bsim = F_Bsim , F_Bsne = F_Bsne ,
PBsim = PBsim,PBsne = PBsne,
C_10mya = C_10mya , C_23mya = C_23mya ,
C_Pebas23mya = C_Pebas23mya , D_conec = D_conec,
Fun_Bsor = Fun_Bsor , Fun_Bsim = Fun_Bsim ,
Fun_Bsne = Fun_Bsne ,
D_PCDp = D_PCDp ,
D_PCDp29k = D_PCDp29k, Dpw_mean = Dpw_mean,#Dpw_gen,
D_hydro = D_hydro,
PBsimCasse = PBsimCasse,
fake_mat = fake_mat)

#creating a matrix to receive results from all pairwise matrix combinations
res <- matrix(NA,nrow = length(mats), ncol = length(mats),
              dimnames = list(names(mats),
                             names(mats)))

#looping over all pairwise matrix combinations
for(i in seq_along(mats)){
  for(k in seq_along(mats)){
    res[i,k] <- suppressWarnings(check_matrix_pair(mats[[i]],mats[[k]]))
  }
}

#perfect! Only the fake matrix (fake_mat) has a different structure
res

##                                     D_CurrentLGM D_energ D_Habitat_size D_Habitat_Div
## D_CurrentLGM                      TRUE    TRUE      TRUE      TRUE
## D_energ                         TRUE    TRUE      TRUE      TRUE
## D_Habitat_size                   TRUE    TRUE      TRUE      TRUE
## D_Habitat_Div                    TRUE    TRUE      TRUE      TRUE
## D_Habitat_harsh                  TRUE    TRUE      TRUE      TRUE
## D_Marine_Incurs                  TRUE    TRUE      TRUE      TRUE
## D_Habitat_fragm                  TRUE    TRUE      TRUE      TRUE
## D_Habitat_water                  TRUE    TRUE      TRUE      TRUE
## D_Habitat_SampEff                TRUE    TRUE      TRUE      TRUE
## D_temp                           TRUE    TRUE      TRUE      TRUE
## D_Habitat_position               TRUE    TRUE      TRUE      TRUE
## D_water                          TRUE    TRUE      TRUE      TRUE
## F_Bsor                           TRUE    TRUE      TRUE      TRUE
## F_Bsim                           TRUE    TRUE      TRUE      TRUE
## F_Bsne                           TRUE    TRUE      TRUE      TRUE

```

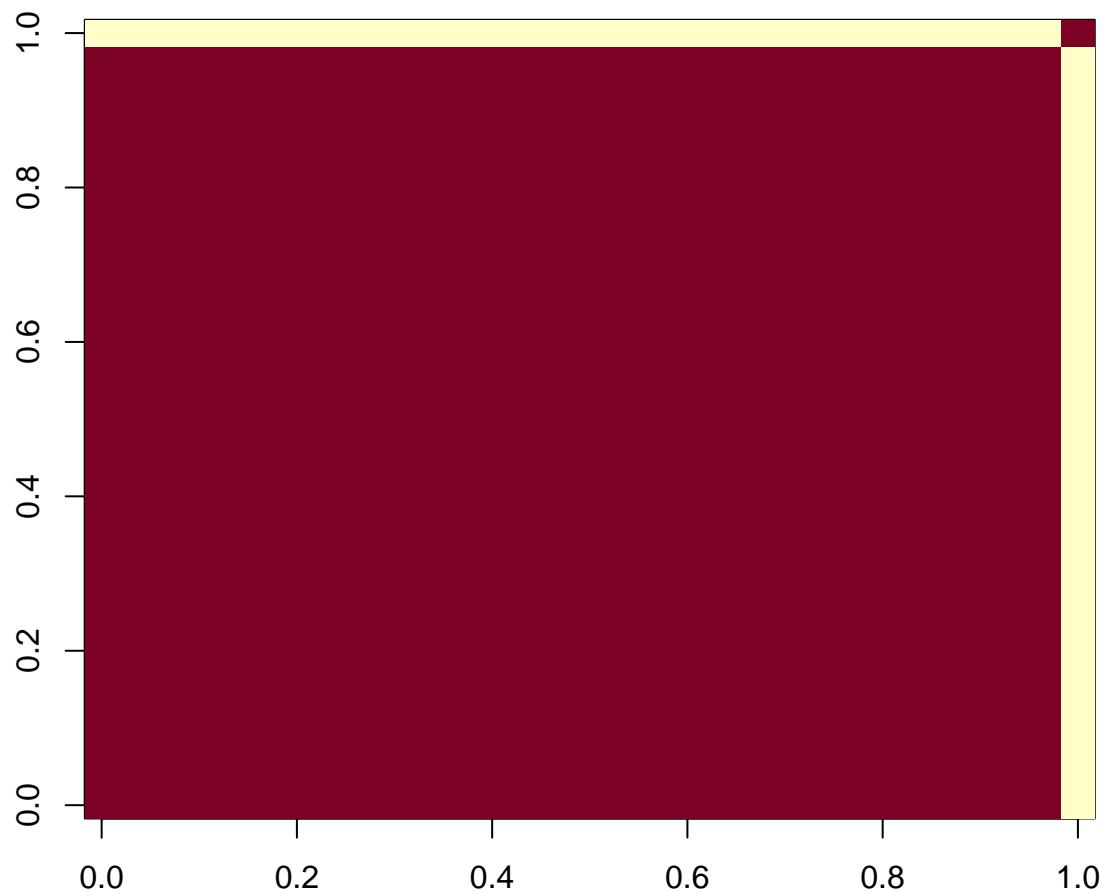
	TRUE	TRUE	TRUE	TRUE
## PBsim	TRUE	TRUE	TRUE	TRUE
## PBsne	TRUE	TRUE	TRUE	TRUE
## C_10mya	TRUE	TRUE	TRUE	TRUE
## C_23mya	TRUE	TRUE	TRUE	TRUE
## C_Pebas23mya	TRUE	TRUE	TRUE	TRUE
## D_conec	TRUE	TRUE	TRUE	TRUE
## Fun_Bsor	TRUE	TRUE	TRUE	TRUE
## Fun_Bsim	TRUE	TRUE	TRUE	TRUE
## Fun Bsne	TRUE	TRUE	TRUE	TRUE
## D_PCDp	TRUE	TRUE	TRUE	TRUE
## D_PCDp29k	TRUE	TRUE	TRUE	TRUE
## Dpw_mean	TRUE	TRUE	TRUE	TRUE
## D_hydro	TRUE	TRUE	TRUE	TRUE
## PBsimCasse	TRUE	TRUE	TRUE	TRUE
## fake_mat	FALSE	FALSE	FALSE	FALSE
##	D_Habitat_harsh	D_Marine_Incurs	D_Habitat_fragm	
## D_CurrentLGM	TRUE	TRUE	TRUE	
## D_energ	TRUE	TRUE	TRUE	
## D_Habitat_size	TRUE	TRUE	TRUE	
## D_Habitat_Div	TRUE	TRUE	TRUE	
## D_Habitat_harsh	TRUE	TRUE	TRUE	
## D_Marine_Incurs	TRUE	TRUE	TRUE	
## D_Habitat_fragm	TRUE	TRUE	TRUE	
## D_Habitat_water	TRUE	TRUE	TRUE	
## D_Habitat_SampEff	TRUE	TRUE	TRUE	
## D_temp	TRUE	TRUE	TRUE	
## D_Habitat_position	TRUE	TRUE	TRUE	
## D_water	TRUE	TRUE	TRUE	
## F_Bsor	TRUE	TRUE	TRUE	
## F_Bsim	TRUE	TRUE	TRUE	
## F_Bsne	TRUE	TRUE	TRUE	
## PBsim	TRUE	TRUE	TRUE	
## PBsne	TRUE	TRUE	TRUE	
## C_10mya	TRUE	TRUE	TRUE	
## C_23mya	TRUE	TRUE	TRUE	
## C_Pebas23mya	TRUE	TRUE	TRUE	
## D_conec	TRUE	TRUE	TRUE	
## Fun_Bsor	TRUE	TRUE	TRUE	
## Fun_Bsim	TRUE	TRUE	TRUE	
## Fun Bsne	TRUE	TRUE	TRUE	
## D_PCDp	TRUE	TRUE	TRUE	
## D_PCDp29k	TRUE	TRUE	TRUE	
## Dpw_mean	TRUE	TRUE	TRUE	
## D_hydro	TRUE	TRUE	TRUE	
## PBsimCasse	TRUE	TRUE	TRUE	
## fake_mat	FALSE	FALSE	FALSE	
##	D_Habitat_water	D_Habitat_SampEff	D_temp	D_Habitat_position
## D_CurrentLGM	TRUE	TRUE	TRUE	TRUE
## D_energ	TRUE	TRUE	TRUE	TRUE
## D_Habitat_size	TRUE	TRUE	TRUE	TRUE
## D_Habitat_Div	TRUE	TRUE	TRUE	TRUE
## D_Habitat_harsh	TRUE	TRUE	TRUE	TRUE

## D_Marine_Incurs	TRUE		TRUE	TRUE		TRUE		
## D_Habitat_fragm	TRUE		TRUE	TRUE		TRUE		
## D_Habitat_water	TRUE		TRUE	TRUE		TRUE		
## D_Habitat_SampEff	TRUE		TRUE	TRUE		TRUE		
## D_temp	TRUE		TRUE	TRUE		TRUE		
## D_Habitat_position	TRUE		TRUE	TRUE		TRUE		
## D_water	TRUE		TRUE	TRUE		TRUE		
## F_Bsor	TRUE		TRUE	TRUE		TRUE		
## F_Bsim	TRUE		TRUE	TRUE		TRUE		
## F_Bsne	TRUE		TRUE	TRUE		TRUE		
## PBsim	TRUE		TRUE	TRUE		TRUE		
## PBsne	TRUE		TRUE	TRUE		TRUE		
## C_10mya	TRUE		TRUE	TRUE		TRUE		
## C_23mya	TRUE		TRUE	TRUE		TRUE		
## C_Pebas23mya	TRUE		TRUE	TRUE		TRUE		
## D_conec	TRUE		TRUE	TRUE		TRUE		
## Fun_Bsor	TRUE		TRUE	TRUE		TRUE		
## Fun_Bsim	TRUE		TRUE	TRUE		TRUE		
## Fun_Bsne	TRUE		TRUE	TRUE		TRUE		
## D_PCDp	TRUE		TRUE	TRUE		TRUE		
## D_PCDp29k	TRUE		TRUE	TRUE		TRUE		
## Dpw_mean	TRUE		TRUE	TRUE		TRUE		
## D_hydro	TRUE		TRUE	TRUE		TRUE		
## PBsimCasse	TRUE		TRUE	TRUE		TRUE		
## fake_mat	FALSE		FALSE	FALSE		FALSE		
##	D_water	F_Bsor	F_Bsim	F_Bsne	PBsim	PBsne	C_10mya	C_23mya
## D_CurrentLGM	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_energ	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_size	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_Div	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_harsh	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Marine_Incurs	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_fragm	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_water	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_SampEff	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_temp	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_position	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_water	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## F_Bsor	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## F_Bsim	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## F_Bsne	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## PBsim	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## PBsne	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## C_10mya	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## C_23mya	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## C_Pebas23mya	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_conec	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## Fun_Bsor	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## Fun_Bsim	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## Fun_Bsne	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_PCDp	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_PCDp29k	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

	C_Pebas23mya	D_conec	Fun_Bsor	Fun_Bsim	Fun_Bsne	D_PCDp
## D_pw_mean	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_hydro	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## PBsimCasse	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## fake_mat	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##						
## D_CurrentLGM	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_energ	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_size	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_Div	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_harsh	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Marine_Incurs	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_fragm	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_water	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_SampEff	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_temp	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_Habitat_position	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_water	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## F_Bsor	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## F_Bsim	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## F_Bsne	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## PBsim	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## PBsne	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## C_10mya	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## C_23mya	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## C_Pebas23mya	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_conec	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## Fun_Bsor	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## Fun_Bsim	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## Fun_Bsne	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_PCDp	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_PCDp29k	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_pw_mean	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## D_hydro	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## PBsimCasse	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
## fake_mat	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##						
	D_PCDp29k	Dpw_mean	D_hydro	PBsimCasse	fake_mat	
## D_CurrentLGM	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_energ	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_size	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_Div	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_harsh	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Marine_Incurs	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_fragm	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_water	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_SampEff	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_temp	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_Habitat_position	TRUE	TRUE	TRUE	TRUE	FALSE	
## D_water	TRUE	TRUE	TRUE	TRUE	FALSE	
## F_Bsor	TRUE	TRUE	TRUE	TRUE	FALSE	
## F_Bsim	TRUE	TRUE	TRUE	TRUE	FALSE	
## F_Bsne	TRUE	TRUE	TRUE	TRUE	FALSE	
## PBsim	TRUE	TRUE	TRUE	TRUE	FALSE	

```
## PBsne      TRUE  TRUE  TRUE  TRUE  FALSE
## C_10mya    TRUE  TRUE  TRUE  TRUE  FALSE
## C_23mya    TRUE  TRUE  TRUE  TRUE  FALSE
## C_Pebas23mya  TRUE  TRUE  TRUE  TRUE  FALSE
## D_conec    TRUE  TRUE  TRUE  TRUE  FALSE
## Fun_Bsor   TRUE  TRUE  TRUE  TRUE  FALSE
## Fun_Bsim   TRUE  TRUE  TRUE  TRUE  FALSE
## Fun_Bsne   TRUE  TRUE  TRUE  TRUE  FALSE
## D_PCDp     TRUE  TRUE  TRUE  TRUE  FALSE
## D_PCDp29k  TRUE  TRUE  TRUE  TRUE  FALSE
## Dpw_mean   TRUE  TRUE  TRUE  TRUE  FALSE
## D_hydro    TRUE  TRUE  TRUE  TRUE  FALSE
## PBsimCasse TRUE  TRUE  TRUE  TRUE  FALSE
## fake_mat   FALSE FALSE FALSE FALSE  TRUE
```

```
image(res)
```



```
rm(mats,res,fake_mat,i,k,F_Bsor,F_Bsim,F_Bsne,C_10mya,C_23mya,C_Pebas23mya)
```

6 Modelling fish beta diversity

6.1 Descriptive statistics

```
##Table for Mapping
med_simi<-function(x,nome="Bsim"){
  x=as.matrix(x)
  diag(x)<-NA
  x<-rowMeans(x,na.rm = TRUE)
  x<-data.frame(basin=names(x),x)
  colnames(x)[2]<-nome
  return(x)
}

Median_simi<-function(x,nome="Bsim"){
  x=as.matrix(x)
  diag(x)<-NA
  x<-apply(x,1,median,na.rm = TRUE)
  x<-data.frame(basin=names(x),x)
  colnames(x)[2]<-nome
  return(x)
}

one=med_simi(D_fish$beta.sim,nome="TaxoBsim")
two=med_simi(PBsim,nome="PhyloBsim30k")
three=med_simi(D_PCDp29k,nome="PCDp30k")
four=med_simi(Fun_Bsim,nome="FunctBsim")
five=med_simi(Dpw_mean,nome="Dpw_mean3k")
one<-merge(one,two,by="basin")
one<-merge(one,three,by="basin")
one<-merge(one,four,by="basin")
one<-merge(one,five,by="basin")
rm(two, three, four,five)

one$Median_TaxoBsim<-Median_simi(D_fish$beta.sim,nome="Median_TaxoBsim")$Median_TaxoBsim
one$Median_PBsim<-Median_simi(PBsim,nome="Median_PBsim")$Median_PBsim

#turnover vs nestedness comparison
#Taxo
summary(D_fish$beta.sim/D_fish$beta.sor)# mean=0.77, range=0.05-1.00

##      Min. 1st Qu. Median    Mean 3rd Qu.    Max.
## 0.0466  0.6511  0.8132  0.7707  0.9230  1.0000

sd(D_fish$beta.sim/D_fish$beta.sor)# sd= 0.18

## [1] 0.18
```

```

summary(D_fish$beta.sor)

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.217   0.552   0.672   0.664   0.772   1.000

summary(D_fish$beta.sim)

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0266  0.3786  0.5083  0.5144  0.6408  1.0000

summary(D_fish$beta.sne)

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.000   0.048   0.121   0.150   0.223   0.771

sd(D_fish$beta.sor)

## [1] 0.16

sd(D_fish$beta.sim)

## [1] 0.18

sd(D_fish$beta.sne)

## [1] 0.13

#Phylo
summary(PBsim/(PBsim+PBsne)) # mean=0.62, range=0.01-0.99

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0105  0.4249  0.6342  0.6201  0.8345  0.9931

sd(PBsim/(PBsim+PBsne)) # sd= 0.24

## [1] 0.24

summary(PBsim+PBsne)

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.131   0.315   0.401   0.416   0.481   0.951

```

```

summary(PBsim)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.00824 0.16441 0.23345 0.24394 0.31561 0.68970

summary(PBsne)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.00303 0.05613 0.13239 0.17246 0.24032 0.89164

sd(PBsim+PBsne)

## [1] 0.15

sd(PBsim)

## [1] 0.11

sd(PBsne)

## [1] 0.16

#Funct
summary(beta_funct$beta.sim/beta_funct$beta.sor) # mean=0.31, range=0.00-1.00

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.0000 0.0357 0.1859 0.3075 0.5411 1.0000

sd(beta_funct$beta.sim/beta_funct$beta.sor) # sd= 0.31

## [1] 0.31

summary(beta_funct$beta.sor)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.0192 0.2251 0.3217 0.3763 0.4667 0.9997

summary(beta_funct$beta.sim)

##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.0000 0.0140 0.0479 0.0792 0.1117 0.7002

```

```

summary(beta_funct$beta.sne)

##      Min.    1st Qu.     Median      Mean    3rd Qu.      Max.
## 5.40e-06 9.56e-02 2.25e-01 2.97e-01 4.21e-01 1.00e+00

sd(beta_funct$beta.sor)

## [1] 0.23

sd(beta_funct$beta.sim)

## [1] 0.089

sd(beta_funct$beta.sne)

## [1] 0.26

#pairwise correlations
#cor(one[,-1])
cor.test(x = D_fish$beta.sim,y = PBsim)#taxo vs PhyloBsim30k

## 
## Pearson's product-moment correlation
##
## data: D_fish$beta.sim and PBsim
## t = 132, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.88 0.89
## sample estimates:
## cor
## 0.89

cor.test(x = D_fish$beta.sim,y = Fun_Bsim)#taxo vs FuncBsim

## 
## Pearson's product-moment correlation
##
## data: D_fish$beta.sim and Fun_Bsim
## t = 27, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.34 0.39
## sample estimates:
## cor
## 0.36

```

```
cor.test(x = PBsim, y = Fun_Bsim) #PhyloBsim30k vs FuncBsim

##
## Pearson's product-moment correlation
##
## data: PBsim and Fun_Bsim
## t = 46, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.54 0.58
## sample estimates:
## cor
## 0.56
```

```
cor.test(x = PBsim, y = Dpw_mean) #PhyloBsim30k vs Dpw (phylo ancient)
```

```
##
## Pearson's product-moment correlation
##
## data: PBsim and Dpw_mean
## t = -14, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.23 -0.18
## sample estimates:
## cor
## -0.2
```

```
cor.test(x = D_fish$beta.sim, y = Dpw_mean) #taxo vs Dpw (phylo ancient)
```

```
##
## Pearson's product-moment correlation
##
## data: D_fish$beta.sim and Dpw_mean
## t = -16, df = 4654, p-value <2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.26 -0.21
## sample estimates:
## cor
## -0.23
```

```
cor.test(x = Fun_Bsim, y = Dpw_mean) #taxo vs Dpw (phylo ancient)
```

```
##
## Pearson's product-moment correlation
##
## data: Fun_Bsim and Dpw_mean
## t = -5, df = 4654, p-value = 3e-06
```

```

## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.097 -0.040
## sample estimates:
## cor
## -0.068

#descriptive statistic
summary(one)

##      basin          TaxoBsim      PhyloBsim30k      PCDp30k
## Length:97      Min.   :0.308      Min.   :0.127      Min.   :1.07
## Class :character 1st Qu.:0.442      1st Qu.:0.197      1st Qu.:1.15
## Mode  :character Median :0.512      Median :0.254      Median :1.16
##                  Mean   :0.514      Mean   :0.244      Mean   :1.17
##                  3rd Qu.:0.572      3rd Qu.:0.283      3rd Qu.:1.18
##                  Max.   :0.866      Max.   :0.358      Max.   :1.32
##      FunctBsim      Dpw_mean3k      Median_TaxoBsim      Median_PBsim
## Min.   :0.0133     Min.   :-3.273      Min.   :0.297      Min.   :0.127
## 1st Qu.:0.0455    1st Qu.:-2.654      1st Qu.:0.431      1st Qu.:0.193
## Median :0.0781    Median :-2.029      Median :0.524      Median :0.263
## Mean   :0.0792    Mean   :-1.727      Mean   :0.519      Mean   :0.247
## 3rd Qu.:0.1057    3rd Qu.:-0.762      3rd Qu.:0.584      3rd Qu.:0.284
## Max.   :0.2178    Max.   : 0.590      Max.   :0.857      Max.   :0.368

apply(one[,-1], 2, sd)

##           TaxoBsim      PhyloBsim30k      PCDp30k      FunctBsim      Dpw_mean3k
## 0.097       0.052        0.042        0.040       1.075
## Median_TaxoBsim      Median_PBsim
## 0.108       0.055

range((1-Fun_Bsim))

## [1] 0.3 1.0

sumario<-function(x){
  list(Summary=summary(as.vector(x)), SD=sd(as.vector(x)))
}

#standardizing predictors
D_energ_s<-Standard(D_energ) #current energy
D_water_s<-Standard(D_water) #current water hypotheses
D_temp_s<-Standard(D_temp) #current temperature: fast speciation hypotheses
D_Habitat_size_s<-Standard(D_Habitat_size) #Habitat size (area + riverine density)
D_Habitat_Div_s<-Standard(D_Habitat_Div) #Habitat Diversity (forest type + soil types)
D_Habitat_harsh_s<-Standard(D_Habitat_harsh) #Habitat elevation gradient (elev mean, range, CSup15,
D_Habitat_position_s<-Standard(D_Habitat_position) #Habitat position

```

```

D_Habitat_fragm_s<-Standard(D_Habitat_fragm) #Habitat fragmentation
D_Habitat_water_s<-Standard(D_Habitat_water) #Habitat water type
D_Habitat_SampEff_s<-Standard(D_Habitat_SampEff) #Sampling effort
D_Marine_Incurs_s<-Standard(D_Marine_Incurs) #Historical Marine Incursions
D_CurrentLGM_s<-Standard(D_CurrentLGM) #past climate: past climate instability hypotheses
D_Pebas23mya_s<-Standard(D_pastConnect$Pebas23mya) #Connected or not to Pebas Lake
D_hydro_s<-Standard(D_hydro)
D_spat_s<-Standard(D_spat)

##VIF
resu <- lm(as.vector(D_fish$beta.sim) ~
             D_energ_s + D_water_s +#current energy/water hypotheses
             D_temp_s + #current temperature: fast speciation hypotheses
             D_Habitat_size_s + #Habitat size (area + riverine density)
             D_Habitat_Div_s + #Habitat Diversity (forest type + soil types)
             D_Habitat_harsh_s + #Habitat elevation gradient (elev mean, range, CSup15, Pro1000)
             D_Habitat_position_s + #Habitat position
             D_Habitat_fragm_s + #Habitat fragmentation
             D_Habitat_water_s + #Habitat water type
             D_Habitat_SampEff_s + #Sampling effort
             D_Marine_Incurs_s + #Historical Marine Incursions
             D_CurrentLGM_s + #past climate: past climate instability hypotheses
             D_Pebas23mya_s + #Connected or not to Pebas Lake
             D_hydro_s +
             D_spat_s)
car::vif(resu)

##          D_energ_s          D_water_s          D_temp_s
##            8.4              2.4              6.2
##      D_Habitat_size_s      D_Habitat_Div_s      D_Habitat_harsh_s
##            1.1              1.3              2.2
## D_Habitat_position_s      D_Habitat_fragm_s      D_Habitat_water_s
##            2.6              1.1              1.2
##  D_Habitat_SampEff_s      D_Marine_Incurs_s      D_CurrentLGM_s
##            1.2              1.8              1.5
##      D_Pebas23mya_s          D_hydro_s          D_spat_s
##            1.5              1.4              2.8

rm(resu)

```

6.2 The MRM model

One way to model fish similarity⁷ is to use **Multiple Regression on Distance Matrices**. This technique has been revised elsewhere and is a good way to explain variation in species composition based on a series of environmental (i.e., climate, habitat diversity, connectivity, etc) predictors (Lichstein 2006; Dias *et al.* 2014;

⁷We consider fish Similarity as 1-Dissimilarity. We consider fish similarity because it is an intuitive way to think on differences among species composition (i.e., two sub-drainages with *exactly* the same set of species would have similarity of 1, whereas two sub-drainages with *completely distinct* sets of fish species would have similarity of 0).

Murray *et al.* 2015). In this case, significance of variables is tested against a null model based on the permutation of rows and columns on the response matrix (Lichstein 2006; Dias *et al.* 2014).

As *dissimilarity matrices of predictors* vary between 0 and $+\infty$ but have different mean and sd values, **coefficients** of the MRM models **absorb** such variation and therefore prevent comparing the effect size among predictors (**Gelman_2008?**). In order to avoid such differences and to directly estimate coefficients reflecting the strength of the relationship, we rescaled all predictors so that they have **mean = 0** and **sd = 0.5** (**Gelman_2008?**). This has been implemented by applying the **Standard**⁸ function to each predictor directly on the MRM function.

6.2.1 Ordinary MRM (multiple regressions using Gaussian Distribution)

```
#generic function for performing the same model over metrics

modellingMRM_glm<-function(object,nameResponse,
                               family){
  require(ecodist)
  response<-as.vector(object)
  MRMformula<-formula(response ~
    D_energ_s + D_water_s +#current energy/water hypotheses
    D_temp_s + #current temperature: fast speciation hypotheses
    D_Habitat_size_s + #Habitat size (area + riverine density)
    D_Habitat_Div_s + #Habitat Diversity (forest type + soil types)
    D_Habitat_harsh_s + #Habitat elevation gradient (elev mean,
                           # range, CSup15, Pro1000)
    D_Habitat_position_s + #Habitat position
    D_Habitat_fragm_s + #Habitat fragmentation
    D_Habitat_water_s + #Habitat water type
    D_Habitat_SampEff_s + #Sampling effort
    D_Marine_Incurs_s + #Historical Marine Incursions
    D_CurrentLGM_s + #past climate: past climate instability hypotheses
    D_Pebas23mya_s + #Connected or not to Pebas Lake
    D_hydro_s +
    D_spat_s)

  resu <- suppressWarnings(MRM(MRMformula,
                                 nperm = 5,mrank = F,method=family), classes = "warning")
  (resu<-suppressWarnings(do.call("rbind",resu), classes = "warning"))
  colnames(resu)<-c(nameResponse,paste0("pvalue_",nameResponse))
  round(resu,2)
}

library(ecodist)

## Registered S3 method overwritten by 'ecodist':
##   method   from
##   dim.dist proxy
```

⁸Defined as $X_s = (x - \text{mean}(x))/(2 * \text{sd}(x))$

```

##  

## Attaching package: 'ecodist'  

##  

## The following objects are masked from 'package:vegan':  

##  

##     mantel, pco  

##  

#TaxoBeta simpson = species turnover  

sumario(D_fish$beta.sim)  

## $Summary  

##      Min. 1st Qu. Median   Mean 3rd Qu.    Max.  

## 0.0266  0.3786  0.5083  0.5144  0.6408  1.0000  

##  

## $SD  

## [1] 0.18  

##  

resu<-modellingMRM_glm(object = D_fish$beta.sim,  

                         nameResponse = "TaxoBsim",  

                         family = "linear")  

##  

## Filo Beta: 3000sp (Rabosky 2018, 2020)  

resuFilo3k<-modellingMRM_glm(object = PBsim,  

                                 nameResponse = "PBsim3k",  

                                 family = "linear")  

##  

## Filo Beta: 600sp (Rabosky 2018, 2020)  

resuFilo<-modellingMRM_glm(object = DsorPhylo$phylo.beta.sim,  

                            nameResponse = "PBsim_gen",  

                            family = "linear")  

##  

## Filo Beta: 1111sp (Cassemiro et al 2023 PNAS)  

resuFiloCasse<-modellingMRM_glm(object = PBsimCasse,  

                                    nameResponse = "PBsimCasse",  

                                    family = "linear")  

##  

## Funct Beta 3k species  

resuFunc<-modellingMRM_glm(object = Fun_Bsim,  

                           nameResponse = "Fun_Bsim",  

                           family = "linear")

```

```

write.table(round(resu,2),file = "MRM_a_taxo.csv",sep=";")
write.table(round(resuFilo3k,2),file = "MRM_b_filo3k.csv",sep=";") #Rabosky 2018, 2020
write.table(round(resuFilo,2),file = "MRM_c_filo600sp.csv",sep=";") #Rabosky 2018, 2020
write.table(round(resuFiloCasse,2),file = "MRM_b_filoCasse.csv",sep=";") #Cassemiro et al 2023PNAS
write.table(round(resuFunc,2),file = "MRM_d_func.csv",sep=";")



MRM_allMetrics<-data.frame(round(resu,2),#taxo
                           round(resuFilo3k,2),round(resuFilo,2),#PBsim
                           round(resuFiloCasse,2), #PBsim Cassemiro et al 2023
                           round(resuFunc,2)#Funct

                         ))
MRM_allMetrics<-data.frame(Variables=rownames(MRM_allMetrics),MRM_allMetrics)

write.table(MRM_allMetrics,
            file = paste0("MRM_allMetrics_",
                          format(Sys.Date(), "%Y_%m"),
                          ".csv"),
            sep=";",row.names = FALSE)

rm(resu,resuFilo3k,resuFilo,resuFiloCasse,resuFunc)

```

Here are the variables used to compose each predictor:

- Energie = AET (min, max, annual, cv) + PET (min, max, annual, cv) + NPP (mean, sd, cv)
 - Water availability = Precipitation (min, max, annual, cv) + Runoff (mean)
 - Temperature = Temperature (min, max, mean, cv) + Solar Radiation (mean, cv)
 - Habitat + Habitat size = NetworkDensity + Area_log + Habitat diversity = ForestCover (%cover) + SoilCover + Habitat harshness = Csup15 + ElevMean_log + ElevPropBasinAbove1000m + Habitat fragmentation = Chut_hydrofall_log + Habitat WaterColor = WaterColor (categories black, clear, white; Same water type = 1, Distinct water type = 0) + Habitat SamplingEffort = SamplingEffort
 - Diff_MarineInc = Marine Incursions, Proport Inf25m ($\log(x+1)$)
 - Diff_CurrentLGM = temp (min, max, mean) + precip (min, max, annual)
 - Pebas23mya = both connected \rightarrow 1, one connected or no connections to Pebas lake \rightarrow 0
 - Basin position = DistanceEmbouchure + Codage
 - SpatialMat = WaterCourse distance among subdrainages
-

6.2.2 MRM_glm (multiple regressions using Binomial Distribution)

One problem with the previous model is that it is defined to deal with **linear** models (i.e., it is similar to using a simple linear model). However, dissimilarity matrices based on biological data are usually bounded between 0 (minimal similarity) and 1 (maximal similarity), and many observations may assume such extreme values. Linear models based on Gaussian error distribution (or models based on Pearson correlation) are not appropriate

for such a kind of data because they are not theoretically bounded between 0-1 and as such they may produce unrealistic predicted values; this is specially true when a dissimilarity matrix have many observations of value 0 or 1 (Dias *et al.* 2014).

In order to deal with theses problems and better set an appropriate model to deal with dissimilarities, we used Generalized Linear Models (**GLM**) with Binomial error distribution (and the *logit* link function). Binomial GLMs are appropriate to deal with bounded values (between 0 and 1) and can correctly estimate predicted values even under high frequency of *0s* and *1s* on the response variable. For further justification of such model, please check the procedure adopted by (Dias *et al.* 2014). The GLM approach adopted here is equivalent to the MRM models fitted above, but has more advantages as those discussed above. As in the MRM, we calculated significance of predictors based on the simulation procedure described in (Lichstein 2006; Dias *et al.* 2014).

```
library(ecodist)

#TaxoBeta simpson = species turnover
resu<-modellingMRM_glm(object = D_fish$beta.sim,
                         nameResponse = "TaxoBsim",
                         family = "logistic")

# TaxoBeta nestedness
sumario(D_fish$beta.sne)

## $Summary
##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.000   0.048   0.121   0.150   0.223   0.771
##
## $SD
## [1] 0.13

resuTaxoBsne<-modellingMRM_glm(object = D_fish$beta.sne,
                                    nameResponse = "TaxoBsne",
                                    family = "logistic")

#Filo Beta: 3000sp (Rabosky 2018, 2020)
resuFilo3k <- modellingMRM_glm(object = PBsim,
                                 nameResponse = "PBsim3k",
                                 family = "logistic")

# Filo Beta nestedness
resuFilo3kBsns <- modellingMRM_glm(object = PBSne,
                                       nameResponse = "PBSne3k",
                                       family = "logistic")

#Filo Beta: 600sp (Rabosky 2018, 2020)
resuFilo <- modellingMRM_glm(object = DsorPhylo$phylo.beta.sim,
                               nameResponse = "PBsim_gen",
                               family = "logistic")

resuFiloBsne <- modellingMRM_glm(object = DsorPhylo$phylo.beta.sne,
                                    nameResponse = "PBsim_gen",
```

```

family = "logistic")

#Filo Beta: 1111sp (Cassemiro et al 2023PNAS)
resuFiloCasse <- modellingMRM_glm(object = PBsimCasse,
                                    nameResponse = "PBsimCasse",
                                    family = "logistic")
# resuFiloCasseBsne <- modellingMRM_glm(object = PBsneCasse,
#                                         nameResponse = "PBsimCasse",
#                                         family = "logistic")

#Funct Beta 3k species
#rounding extremely low negative values (-1.1e-15) to 0
min(Fun_Bsim)

## [1] -1.1e-15

Fun_Bsim<-as.matrix(Fun_Bsim)
table(Fun_Bsim[Fun_Bsim<0])

##
## -1.12056368432741e-15 -4.17665363053448e-16 -3.6595197815909e-16
##          4                  2                  2
## -3.2427492362413e-16 -2.62101103874906e-16 -2.49628903551016e-16
##          2                  2                 128
## -1.60387674286745e-16 -1.36294603610457e-16 -1.27321299111335e-16
##          2                  2                  2

Fun_Bsim[Fun_Bsim<0]<-0
Fun_Bsim<-as.dist(Fun_Bsim)

resuFunc <- modellingMRM_glm(object = Fun_Bsim,
                               nameResponse = "Fun_Bsim",
                               family = "logistic")

#Funct Beta nestedness
resuFuncBsne <- modellingMRM_glm(object = Fun Bsne,
                                    nameResponse = "Fun_Bsne",
                                    family = "logistic")

write.table(round(resu,2),file = "MRM_a_taxoBsim_glm.csv",sep=";")
write.table(round(resuTaxoBsne,2),file = "MRM_a2_taxoBsne_glm.csv",sep=";")
write.table(round(resuFilo3k,2),file = "MRM_b_filo3k_glm.csv",sep=";") #Rabosky 2018, 2020
write.table(round(resuFilo3kBsne,2),file = "MRM_b2_filo3kBsne_glm.csv",sep=";") #Rabosky 2018, 2020
write.table(round(resuFilo,2),file = "MRM_c_filo600sp_glm.csv",sep=";") #Rabosky 2018, 2020

```

```

write.table(round(resuFiloCasse,2),file = "MRM_b_filoCasse_glm.csv",sep=";") #Cassemiro et al 2023PNAS
write.table(round(resuFunc,2),file = "MRM_d_funcBsim_glm.csv",sep=";") 
write.table(round(resuFuncBsne,2),file = "MRM_d2_funcBsne_glm.csv",sep=";")



MRM_allMetrics<-data.frame(rownames(resu),
                            round(resu,2),
                            round(resuTaxoBsne,2),#TaxoBsne
                            round(resuFilo3k,2), #Rabosky 2018,2020
                            round(resuFilo3kBsns,2),#PBsns
                            round(resuFilo,2), #Rabosky 2018,2020
                            round(resuFiloCasse,2), #Cassemiro et al 2023PNAS
                            round(resuFunc,2),
                            round(resuFuncBsne,2))#Fun_Bsne

write.table(MRM_allMetrics,
            file = paste0("MRM_allMetrics_glm_",
                          format(Sys.time(), "%Y_%m"),
                          ".csv"),
            sep=";",row.names = FALSE)

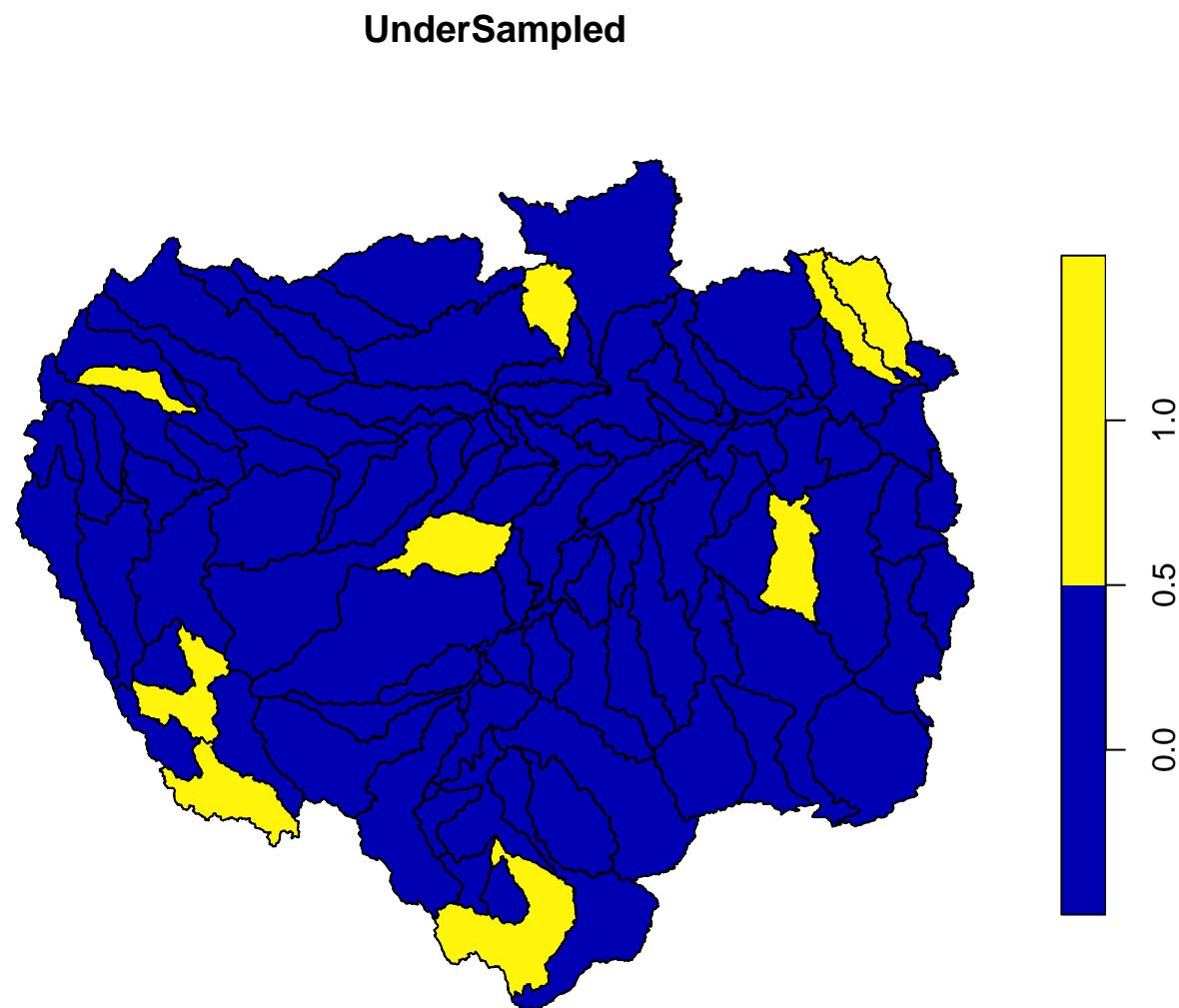
rm(resu,resuFilo3k,resuFilo,resuFiloCasse,resuFunc)

```

6.2.3 Testing excluding undersampled basins

First, I identify the basins undersampled according to (Jézéquel *et al.* 2020).

```
# following Jézéquel et al. 2020 Conservation biology,  
# Figure 1a  
undersample_basins <- c("Demini", "Paru_Este", "Jari", "Jamanxim",  
  "Grande", "Tapaua", "Apurimac1", "Ucayali2", "Curaray")  
amazon$UnderSampled <- 0  
amazon$UnderSampled[amazon$BvNiv2 %in% undersample_basins] <- 1  
  
plot(amazon["UnderSampled"])
```



Then, I constructed a function to remove those pairwise values from distance matrices

```
# x<-D_fish$beta.sim N<-undersample_basins  
  
Exclude_Undersampled <- function(DistanceMatrix = x, NamesToExclude = N) {  
  M <- as.matrix(DistanceMatrix)
```

```

# dim(M)
N <- NamesToExclude
M <- as.dist(M[!rownames(M) %in% N, !colnames(M) %in% N])
return(M)
}

test <- Exclude_Undersampled(DistanceMatrix = D_fish$beta.sim,
                               NamesToExclude = undersample_basins)

names(test)

## [1] "Abuna"          "Amazon1"        "Amazon2"        "Amazon5"
## [5] "Amazon8"        "Amazon9"         "Andira1"        "Apaporis"
## [9] "Arinos"          "Aripuana"        "Bacaja"         "Beni"
## [13] "Blanco_Baures" "Branco"          "Candeias"       "Canuma"
## [17] "Caqueta1"       "Chapare"         "Coari"          "Curua"
## [21] "Curua_Una"      "Fresco"          "Guapore"        "Huallaga"
## [25] "Iriri"          "Isiboro"         "Itonamas"      "Japurai1"
## [29] "Jatapu"          "Javary1"         "Jiparana"       "Jurua1"
## [33] "Juruena1"        "Jutai"           "Luna"           "Madera1"
## [37] "Madera2"         "Madera3"         "Madera4"        "Madera6"
## [41] "Madre_Dios"      "Mamore1"         "Mamore2"        "Mantaro"
## [45] "Maranon1"        "Maranon3"        "Maranon5"       "Marmelos"
## [49] "Maues"           "Napo1"           "Napo2"          "Negro1"
## [53] "Negro2"          "Negro3"          "Negro5"         "Nhamunda"
## [57] "Orthon"           "Pachitea"        "Pastaza"        "Purus1"
## [61] "Purus2"           "Purus3"          "Putumayo"       "Roosevelt"
## [65] "Santiago"         "Solimoes1"        "Solimoes2"       "Solimoes3"
## [69] "Solimoes4"        "Solimoes7"        "Solimoes8"       "Solimoes9"
## [73] "Tapajos1"         "Tapajos2"         "Tefe"           "Teles_Pires"
## [77] "Tigre"            "Trombetas1"       "Uatumal"        "Uaupes"
## [81] "Ucayali1"         "Unini"           "Urubamba"       "Xingu1"
## [85] "Xingu3"           "Xingu4"          "Xingu6"         "Yacuma"

table(names(test) %in% undersample_basins) #corrected excluded

## 
## FALSE
##     88

rm(test)

```

Now I refit the same models without those problematic basins

```

library(ecodist)

#biological data
TaxoBsim_US<-Exclude_Undersampled(DistanceMatrix = D_fish$beta.sim,
                                      NamesToExclude = undersample_basins)

```

```

PBsim_US<-Exclude_Undersampled(DistanceMatrix = PBsim,
                                  NamesToExclude = undersample_basins)

PBsim600_US<-Exclude_Undersampled(DistanceMatrix = DsorPhylo$phylo.beta.sim,
                                     NamesToExclude = undersample_basins)

PBsimCasse_US<-Exclude_Undersampled(DistanceMatrix = PBsimCasse,
                                       NamesToExclude = undersample_basins)

Dpw_mean_US<-Exclude_Undersampled(DistanceMatrix = Dpw_mean,
                                      NamesToExclude = undersample_basins)

Fun_Bsim_US<-Exclude_Undersampled(DistanceMatrix = Fun_Bsim,
                                     NamesToExclude = undersample_basins)

#predictors
#removing & standardizing predictors
energ_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_energ,
                                           NamesToExclude = undersample_basins))

water_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_water,
                                           NamesToExclude = undersample_basins))

temp_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_temp,
                                           NamesToExclude = undersample_basins))

size_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_size,
                                           NamesToExclude = undersample_basins))

Div_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_Div,
                                         NamesToExclude = undersample_basins))

harsh_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_harsh,
                                           NamesToExclude = undersample_basins))

position_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_position,
                                             NamesToExclude = undersample_basins))

fragm_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_fragm,
                                           NamesToExclude = undersample_basins))

water_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_water,
                                           NamesToExclude = undersample_basins))

SampEff_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Habitat_SampEff,
                                            NamesToExclude = undersample_basins))

Marine_Incurs_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_Marine_Incurs,
                                                 NamesToExclude = undersample_basins))

```

```

CurrentLGM_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_CurrentLGM,
                                               NamesToExclude = undersample_basins))

Pebas23mya_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_pastConnect$Pebas23mya,
                                                NamesToExclude = undersample_basins))

hydro_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_hydro,
                                           NamesToExclude = undersample_basins))

spat_US<-Standard(Exclude_Undersampled(DistanceMatrix = D_spat,
                                          NamesToExclude = undersample_basins))

modelTaxoBsim<-formula("as.vector(TaxoBsim_US) ~
                           energ_US + water_US +
                           temp_US +
                           size_US +
                           Div_US +
                           harsh_US +
                           position_US +
                           fragm_US +
                           water_US +
                           SampEff_US +
                           Marine_Incurs_US +
                           CurrentLGM_US +
                           Pebas23mya_US +
                           hydro_US +
                           spat_US")

#TaxoBeta simpson = species turnover

resu <- suppressWarnings(MRM(modelTaxoBsim,
                               nperm = 3, mrank = F, method = "logistic"), classes = "warning")

(resu<-do.call("rbind",resu))

## Warning in rbind(coef = structure(c(-0.0074499572874183, -0.0853753784986295, :
## number of columns of result is not a multiple of vector length (arg 2)

##          b      pval
## (Intercept) -0.0074   1.00
## energ_US     -0.0854   0.67
## water_US     -0.1990   0.33
## temp_US       0.3179   0.33
## size_US      -0.0710   0.67
## Div_US        0.1840   0.33
## harsh_US      0.3513   0.33
## position_US    0.1308   0.67
## fragm_US      0.0202   1.00

```

```

## SampEff_US      -0.0882    0.67
## Marine_Incurs_US -0.2009    0.33
## CurrentLGM_US      0.1283    0.33
## Pebas23mya_US     -0.2038    0.33
## hydro_US          -0.0338    0.67
## spat_US           0.5328    0.33
## dev                249.6326  3813.00

```

```

colnames(resu) <- c("TaxoBsim", "pvalue_TaxoBsim")
round(resu, 2)

```

	TaxoBsim	pvalue_TaxoBsim
## (Intercept)	-0.01	1.00
## energ_US	-0.09	0.67
## water_US	-0.20	0.33
## temp_US	0.32	0.33
## size_US	-0.07	0.67
## Div_US	0.18	0.33
## harsh_US	0.35	0.33
## position_US	0.13	0.67
## fragm_US	0.02	1.00
## SampEff_US	-0.09	0.67
## Marine_Incurs_US	-0.20	0.33
## CurrentLGM_US	0.13	0.33
## Pebas23mya_US	-0.20	0.33
## hydro_US	-0.03	0.67
## spat_US	0.53	0.33
## dev	249.63	3813.00

```

#Filo Beta: 3000sp (Rabosky 2018, 2020)
modelPBsim <- formula("as.vector(PBsim_US) ~
                        energ_US + water_US +
                        temp_US +
                        size_US +
                        Div_US +
                        harsh_US +
                        position_US +
                        fragm_US +
                        water_US +
                        SampEff_US +
                        Marine_Incurs_US +
                        CurrentLGM_US +
                        Pebas23mya_US +
                        hydro_US +
                        spat_US")

```

```

resuFilo3k <- suppressWarnings(MRM(modelPBsim,
                                      nperm = 3, mrank = F, method = "logistic"), classes = "warning")

(resuFilo3k <- do.call("rbind", resuFilo3k))

```

```
## Warning in rbind(coef = structure(c(-1.18211073478133, -0.159854502501044, :  
## number of columns of result is not a multiple of vector length (arg 2)
```

```
## b pval  
## (Intercept) -1.182 0.33  
## energ_US -0.160 0.33  
## water_US -0.115 0.33  
## temp_US 0.140 0.33  
## size_US 0.028 0.67  
## Div_US 0.115 0.33  
## harsh_US 0.226 0.33  
## position_US 0.045 1.00  
## fragm_US -0.021 0.67  
## SampEff_US -0.160 0.33  
## Marine_Incurs_US -0.153 0.33  
## CurrentLGM_US 0.129 0.33  
## Pebas23mya_US -0.125 0.67  
## hydro_US 0.013 0.67  
## spat_US 0.387 0.33  
## dev 146.011 3813.00
```

```
colnames(resuFilo3k) <- c("PBsim3k", "pvalue_PBsim3k")  
round(resuFilo3k, 2)
```

```
## PBsim3k pvalue_PBsim3k  
## (Intercept) -1.18 0.33  
## energ_US -0.16 0.33  
## water_US -0.11 0.33  
## temp_US 0.14 0.33  
## size_US 0.03 0.67  
## Div_US 0.12 0.33  
## harsh_US 0.23 0.33  
## position_US 0.05 1.00  
## fragm_US -0.02 0.67  
## SampEff_US -0.16 0.33  
## Marine_Incurs_US -0.15 0.33  
## CurrentLGM_US 0.13 0.33  
## Pebas23mya_US -0.12 0.67  
## hydro_US 0.01 0.67  
## spat_US 0.39 0.33  
## dev 146.01 3813.00
```

```
#Filo Beta: 600sp (Rabosky 2018, 2020)  
modelPBsim600 <- formula("as.vector(PBsim600_US) ~  
    energ_US + water_US +  
    temp_US +  
    size_US +  
    Div_US +  
    harsh_US +  
    position_US +  
    fragm_US +
```

```

water_US +
SampEff_US +
Marine_Incurs_US +
CurrentLGM_US +
Pebas23mya_US +
hydro_US +
spat_US")

resuFilo <- suppressWarnings(MRM(modelPBsim600,
nperm = 3, mrank = F, method = "logistic"), classes = "warning")

(resuFilo<-do.call("rbind",resuFilo))

```

Warning in rbind(coef = structure(c(-1.38027771130023, -0.205060986969225, :
number of columns of result is not a multiple of vector length (arg 2)

	b	pval
## (Intercept)	-1.380	0.33
## energ_US	-0.205	0.33
## water_US	-0.159	0.33
## temp_US	-0.084	0.67
## size_US	0.063	0.33
## Div_US	0.077	0.33
## harsh_US	0.196	0.33
## position_US	-0.063	0.67
## fragm_US	-0.044	0.67
## SampEff_US	-0.210	0.33
## Marine_Incurs_US	-0.172	0.33
## CurrentLGM_US	0.169	0.33
## Pebas23mya_US	-0.196	0.33
## hydro_US	0.019	0.67
## spat_US	0.493	0.33
## dev	207.401	3813.00

```

colnames(resuFilo)<-c("PBsim_gen","pvalue_PBsim_gen")
round(resuFilo,2)

```

	PBsim_gen	pvalue_PBsim_gen
## (Intercept)	-1.38	0.33
## energ_US	-0.21	0.33
## water_US	-0.16	0.33
## temp_US	-0.08	0.67
## size_US	0.06	0.33
## Div_US	0.08	0.33
## harsh_US	0.20	0.33
## position_US	-0.06	0.67
## fragm_US	-0.04	0.67
## SampEff_US	-0.21	0.33
## Marine_Incurs_US	-0.17	0.33

```

## CurrentLGM_US      0.17      0.33
## Pebas23mya_US     -0.20      0.33
## hydro_US          0.02      0.67
## spat_US           0.49      0.33
## dev                207.40    3813.00

#Filo Beta: 1111sp (Cassemiro et al 2023 PNAS)
modelPBsimCasse<-formula("as.vector(PBsimCasse_US) ~
                           energ_US + water_US +
                           temp_US +
                           size_US +
                           Div_US +
                           harsh_US +
                           position_US +
                           fragm_US +
                           water_US +
                           SampEff_US +
                           Marine_Incurs_US +
                           CurrentLGM_US +
                           Pebas23mya_US +
                           hydro_US +
                           spat_US")

resuFiloCasse <- suppressWarnings(MRM(modelPBsimCasse,
                                         nperm = 3, mrank = F, method = "logistic"), classes = "warning")

(resuFiloCasse<-do.call("rbind",resuFiloCasse))

## Warning in rbind(coef = structure(c(-1.38027771130023, -0.205060986969225, :
## number of columns of result is not a multiple of vector length (arg 2)

##          b      pval
## (Intercept) -1.380  0.33
## energ_US     -0.205  0.33
## water_US     -0.159  0.33
## temp_US      -0.084  0.67
## size_US       0.063  0.67
## Div_US        0.077  0.33
## harsh_US      0.196  0.33
## position_US   -0.063  0.67
## fragm_US      -0.044  0.67
## SampEff_US    -0.210  0.33
## Marine_Incurs_US -0.172  0.67
## CurrentLGM_US   0.169  0.33
## Pebas23mya_US   -0.196  0.33
## hydro_US       0.019  1.00
## spat_US        0.493  0.33
## dev             207.401 3813.00

```

```
colnames(resuFiloCasse) <- c("PBsimCasse", "pvalue_PBsimCasse")
round(resuFiloCasse, 2)
```

	PBsimCasse	pvalue_PBsimCasse
## (Intercept)	-1.38	0.33
## energ_US	-0.21	0.33
## water_US	-0.16	0.33
## temp_US	-0.08	0.67
## size_US	0.06	0.67
## Div_US	0.08	0.33
## harsh_US	0.20	0.33
## position_US	-0.06	0.67
## fragm_US	-0.04	0.67
## SampEff_US	-0.21	0.33
## Marine_Incurs_US	-0.17	0.67
## CurrentLGM_US	0.17	0.33
## Pebas23mya_US	-0.20	0.33
## hydro_US	0.02	1.00
## spat_US	0.49	0.33
## dev	207.40	3813.00

```
#Funct Beta 3k species
# Fun_Bsim_US
summary(Fun_Bsim_US)
```

```
##      Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.0000  0.0179  0.0526  0.0826  0.1168  0.5783
```

```
modelFunBsimUS <- formula("as.vector(Fun_Bsim_US) ~
                           energ_US + water_US +
                           temp_US +
                           size_US +
                           Div_US +
                           harsh_US +
                           position_US +
                           fragm_US +
                           water_US +
                           SampEff_US +
                           Marine_Incurs_US +
                           CurrentLGM_US +
                           Pebas23mya_US +
                           hydro_US +
                           spat_US")
```

```
resuFunc <- suppressWarnings(MRM(modelFunBsimUS,
                                   nperm = 3, mrank = F, method = "logistic"), classes = "warning")

(resuFunc <- do.call("rbind", resuFunc))
```

```

## Warning in rbind(coef = structure(c(-2.45225057436953, -0.284271373661007, :
## number of columns of result is not a multiple of vector length (arg 2)

```

	b	pval
## (Intercept)	-2.452	0.33
## energ_US	-0.284	0.33
## water_US	-0.031	0.67
## temp_US	-0.126	0.33
## size_US	0.039	0.67
## Div_US	0.067	0.33
## harsh_US	0.160	0.33
## position_US	0.087	0.67
## fragm_US	-0.136	0.67
## SampEff_US	-0.396	0.33
## Marine_Incurs_US	-0.361	0.33
## CurrentLGM_US	0.287	0.33
## Pebas23mya_US	-0.011	1.00
## hydro_US	0.117	0.33
## spat_US	0.203	0.33
## dev	323.760	3813.00

```

colnames(resuFunc) <- c("Fun_Bsim", "pvalue_Fun_Bsim")
round(resuFunc, 2)

```

	Fun_Bsim	pvalue_Fun_Bsim
## (Intercept)	-2.45	0.33
## energ_US	-0.28	0.33
## water_US	-0.03	0.67
## temp_US	-0.13	0.33
## size_US	0.04	0.67
## Div_US	0.07	0.33
## harsh_US	0.16	0.33
## position_US	0.09	0.67
## fragm_US	-0.14	0.67
## SampEff_US	-0.40	0.33
## Marine_Incurs_US	-0.36	0.33
## CurrentLGM_US	0.29	0.33
## Pebas23mya_US	-0.01	1.00
## hydro_US	0.12	0.33
## spat_US	0.20	0.33
## dev	323.76	3813.00

```

write.table(round(resu, 2),
            file = "UnderSampled_MRM_glm_a_taxo.csv", sep = ";")

write.table(round(resuFilo3k, 2),
            file = "UnderSampled_MRM_glm_b_filo3k.csv", sep = ";") #Rabosky 2018, 2020

write.table(round(resuFilo, 2),
            file = "UnderSampled_MRM_glm_c_filo600sp.csv", sep = ";") #Rabosky 2018, 2020

```

```

write.table(round(resuFiloCasse,2),
            file = "UnderSampled_MRM_glm_b_filoCasse.csv",sep=";") #Cassemiro et al 2023PNAS

write.table(round(resuFunc,2),
            file = "UnderSampled_MRM_glm_d_func.csv",sep=";")



MRM_allMetrics<-data.frame(round(resu,2),#taxo
                             round(resuFilo3k,2),round(resuFilo,2),#PBsim
                             round(resuFiloCasse,2), #PBsim Cassemiro et al 2023
                             round(resuFunc,2)#Funct
                             #,round(resuFilo_ancient_genetic,2)
                             ))
MRM_allMetrics<-data.frame(Variables=rownames(MRM_allMetrics),MRM_allMetrics)

write.table(MRM_allMetrics,
            file = paste0("UnderSampled_MRM_glm_allMetrics_",
                          format(Sys.time(), "%Y_%m"),
                          ".csv"),
            sep=";",row.names = FALSE)

rm(resu,resuFilo3k,resuFilo,resuFiloCasse,resuFunc)

```

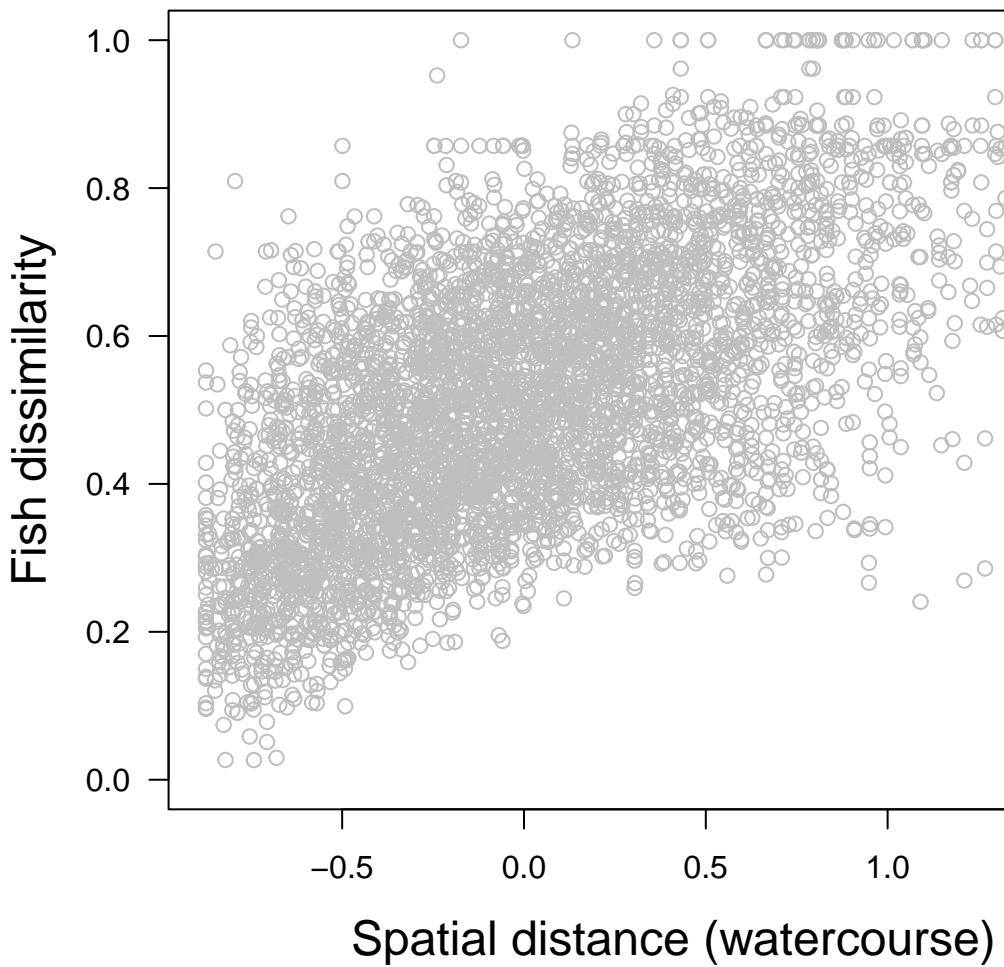
Here are the variables used to compose each predictor:

- Energie = AET (min, max, annual, cv) + PET (min, max, annual, cv) + NPP (mean, sd, cv)
- Water availability = Precipitation (min, max, annual, cv) + Runoff (mean)
- Temperature = Temperature (min, max, mean, cv) + Solar Radiation (mean, cv)
- Habitat + Habitat size = NetworkDensity + Area_log + Habitat diversity = ForestCover (%cover) + SoilCover + Habitat harshness = Csup15 + ElevMean_log + ElevPropBasinAbove1000m + Habitat fragmentation = Chut_hydrofall_log + Habitat WaterColor = WaterColor (categories black, clear, white; Same water type = 1, Distinct water type = 0) + Habitat SamplingEffort = SamplingEffort
- Diff_MarineInc = Marine Incursions, Proport Inf25m ($\log(x+1)$)
- Diff_CurrentLGM = temp (min, max, mean) + precip (min, max, annual)
- Pebas23mya = both connected -> 1, one connected or no connections to Pebas lake -> 0
- Basin position = DistanceEmbouchure + Codage
- SpatialMat = WaterCourse distance among subdrainages ***

6.2.4 Simple plots

Let's see some plots among fish similarity and the predictors used in the modelling procedure.

```
par(mfrow = c(1, 1), cex.lab = 1.5)
# jpeg('CompositionBsim_WaterCourseDistance.jpeg', width =
# 8000, height = 2000, res = 400)
plot(D_fish$beta.sim ~ Standard(D_spat), ylab = "Fish dissimilarity",
     xlab = "Spatial distance (watercourse)", las = 1, cex = 1,
     col = "gray", ylim = c(0, 1))
```

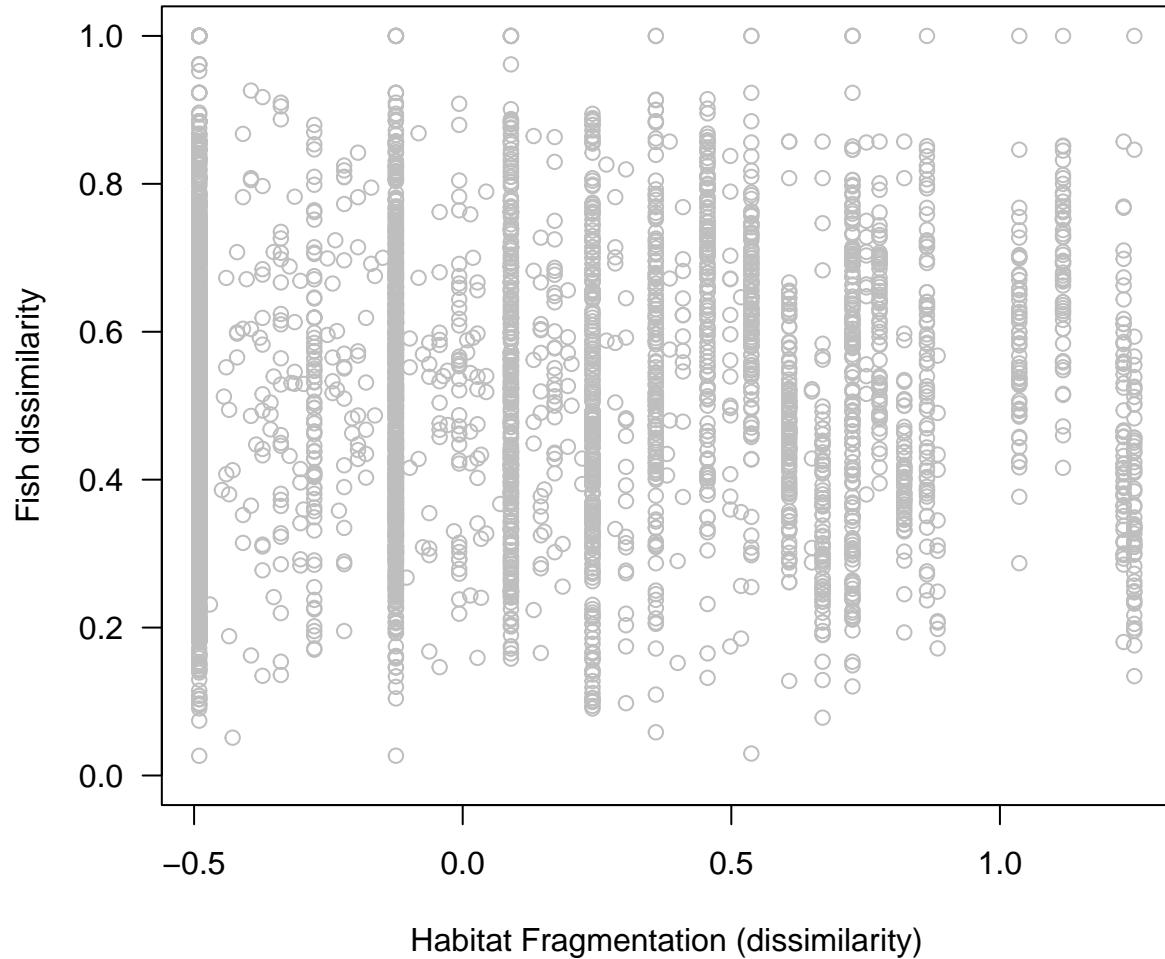


6.2.4.1 Taxonomic beta diversity

```
# curve(plogis(-0.19-0.68*x), add=T) dev.off()

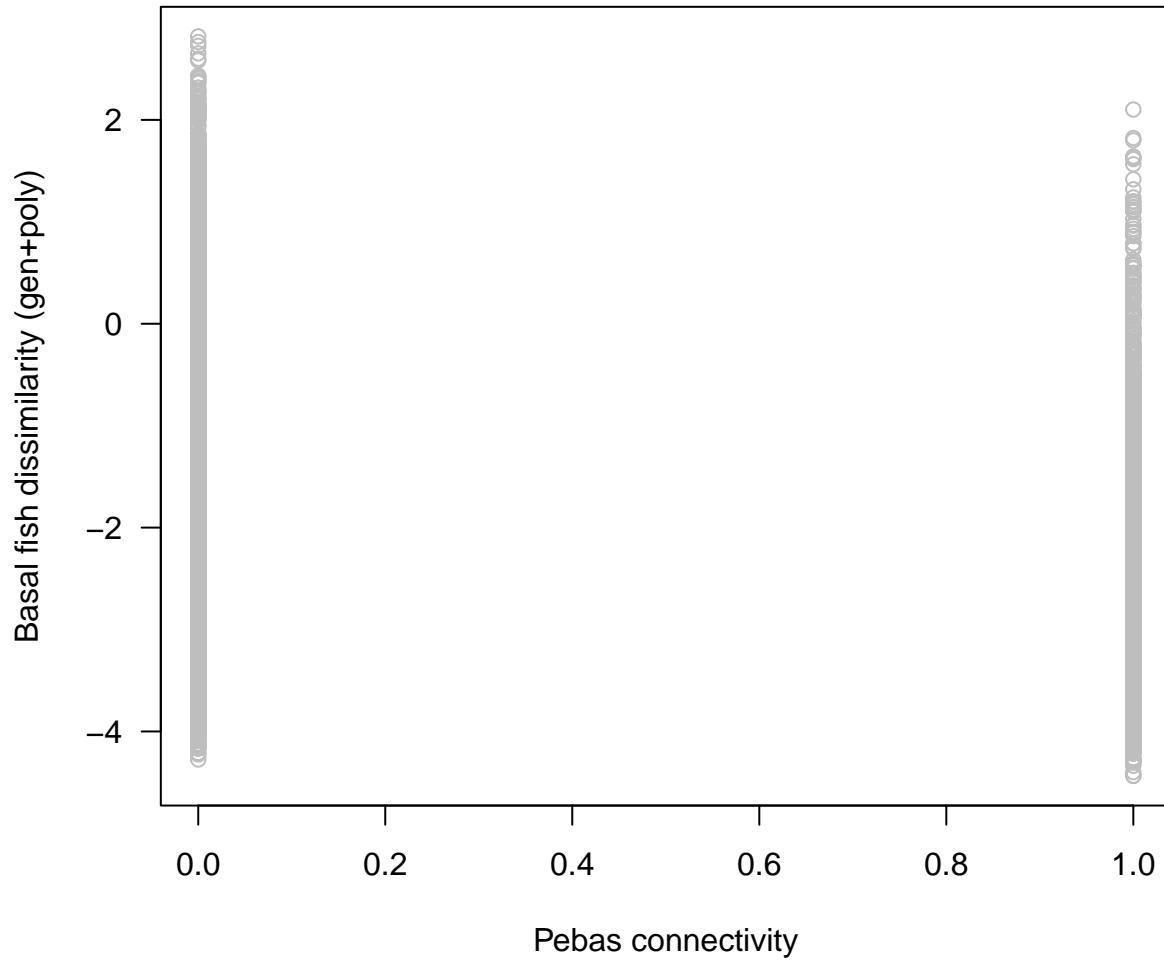
# jpeg('CompositionBsim_Fragmentation.jpeg', width =
# 8000, height = 2000, res = 400)
plot(D_fish$beta.sim ~ Standard(D_Habitat_fragm), ylab = "Fish dissimilarity",
```

```
xlab = "Habitat Fragmentation (dissimilarity)", las = 1,  
cex = 1, col = "gray", ylim = c(0, 1))
```



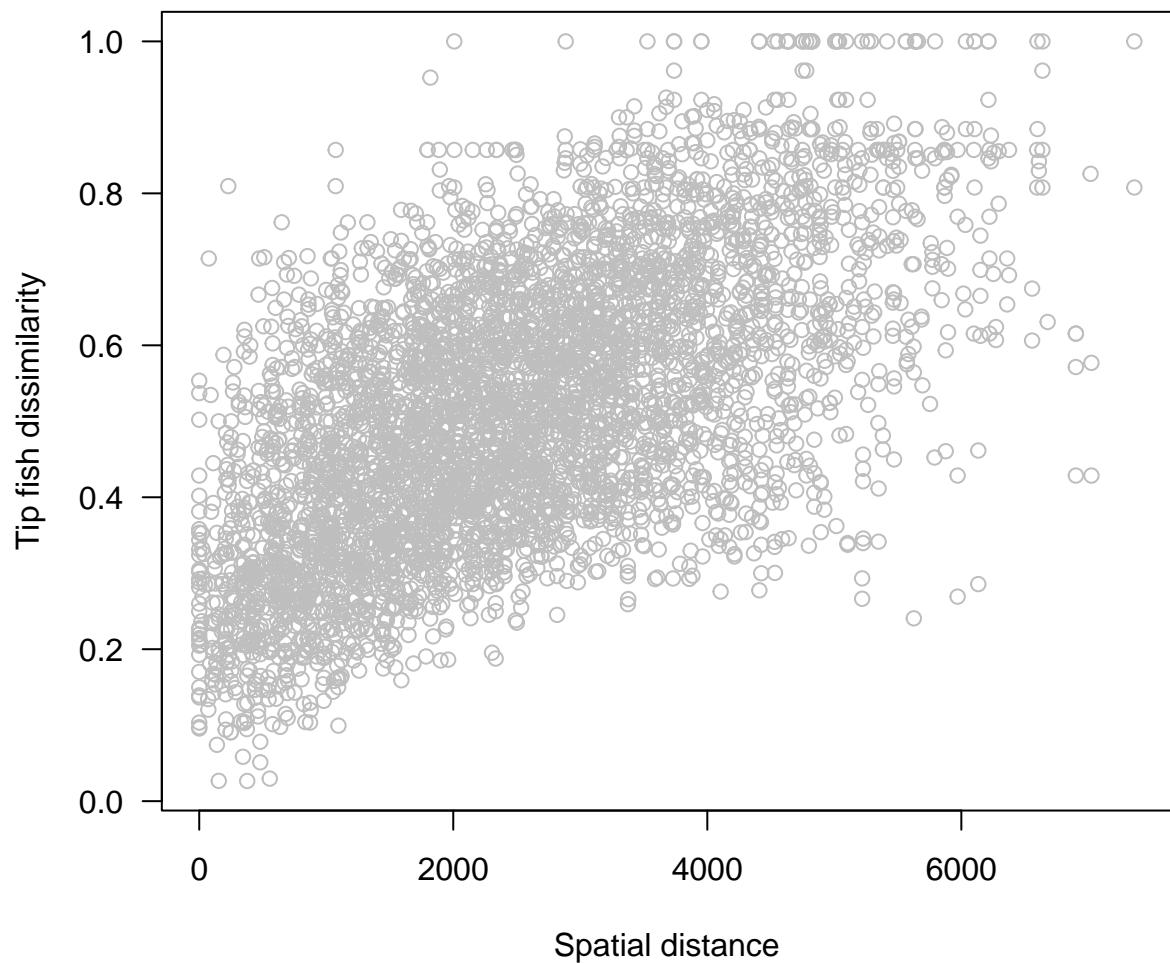
```
# curve(plogis(-0.19-0.41*x),add=T) dev.off()
```

```
# jpeg('.jpeg',width = 2000,height = 2000,res = 400)  
plot(Dpw_mean ~ as.vector(D_pastConnect$Pebas23mya), ylab = "Basal fish dissimilarity (gen+poly)",  
xlab = "Pebas connectivity", las = 1, cex = 1, col = "gray")
```

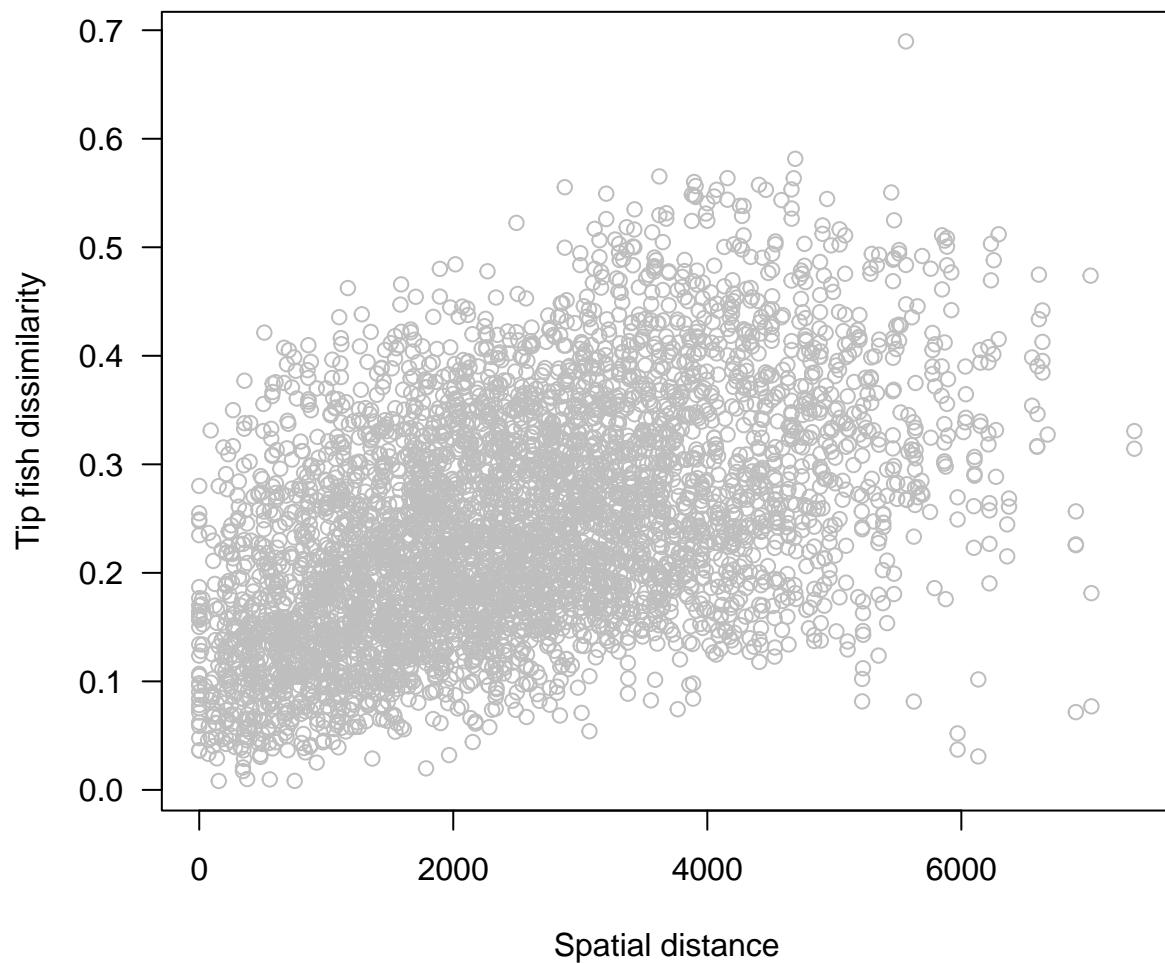


```
# plot(Dpw_gen ~ as.factor(D_pastConnect$Pebas23mya),
# ylab='Basal fish dissimilarity (gen+poly)', xlab='Pebas
# connectivity', las=1, cex = 1, col='gray') dev.off()
```

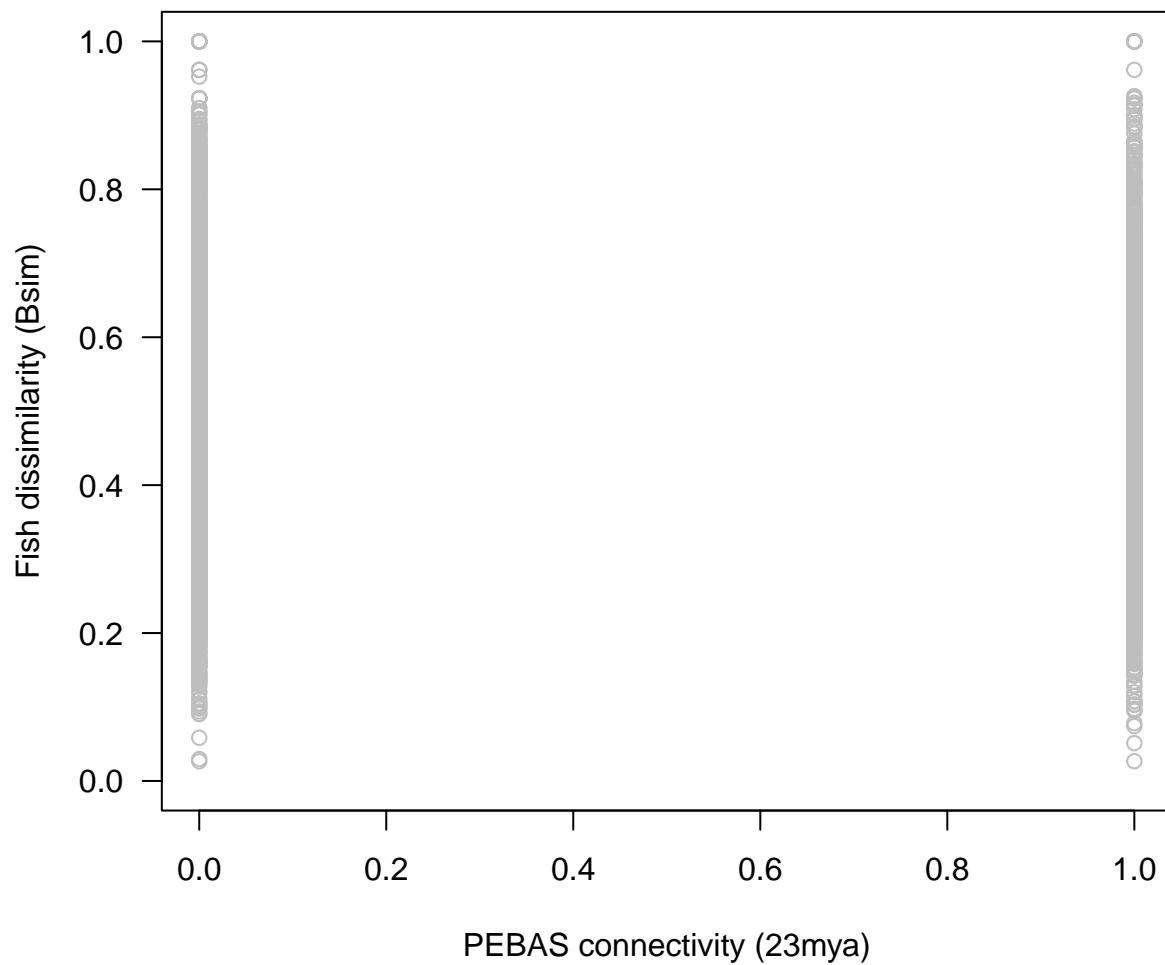
```
plot(D_fish$beta.sim ~ D_spat, ylab = "Tip fish dissimilarity",
xlab = "Spatial distance", las = 1, cex = 1, col = "gray")
```



```
plot(PBsim ~ D_spat, ylab = "Tip fish dissimilarity", xlab = "Spatial distance",
     las = 1, cex = 1, col = "gray")
```

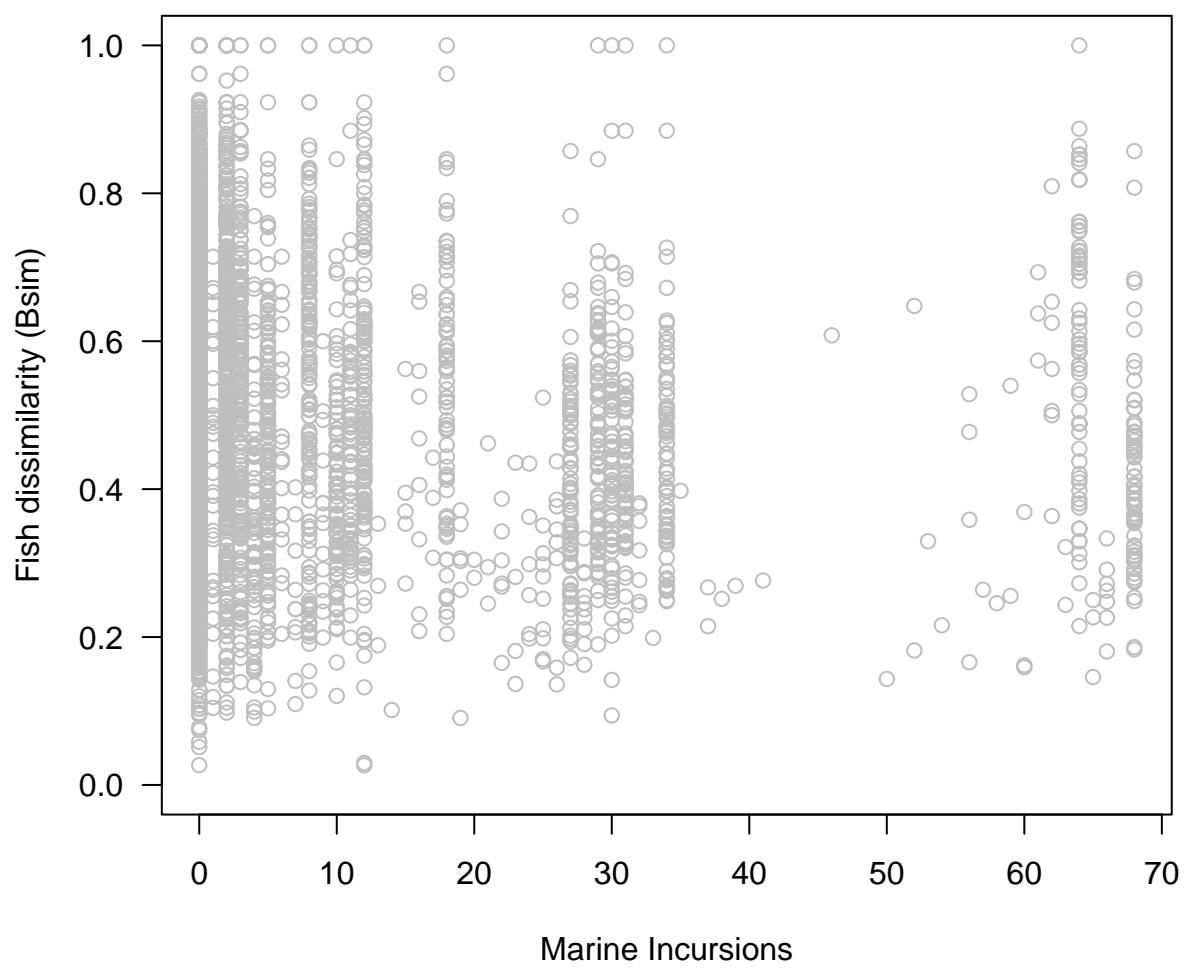


```
# jpeg("CompositionBsim_PebasConnect.jpeg",width = 2000,height = 2000,res = 400)
plot(D_fish$beta.sim ~ (as.vector(D_pastConnect$Pebas23mya)),
      ylab="Fish dissimilarity (Bsim)",
      xlab="PEBAS connectivity (23mya)",#vertical=T,method="jitter",
      las=1, cex = 1, col="gray",ylim=c(0,1))
```



```
# dev.off()
```

```
plot(D_fish$beta.sim ~ D_Marine_Incurs,
      ylab="Fish dissimilarity (Bsim)",
      xlab="Marine Incursions", #vertical=T, m   method="jitter",
      las=1, cex = 1, col="gray", ylim=c(0,1))
```

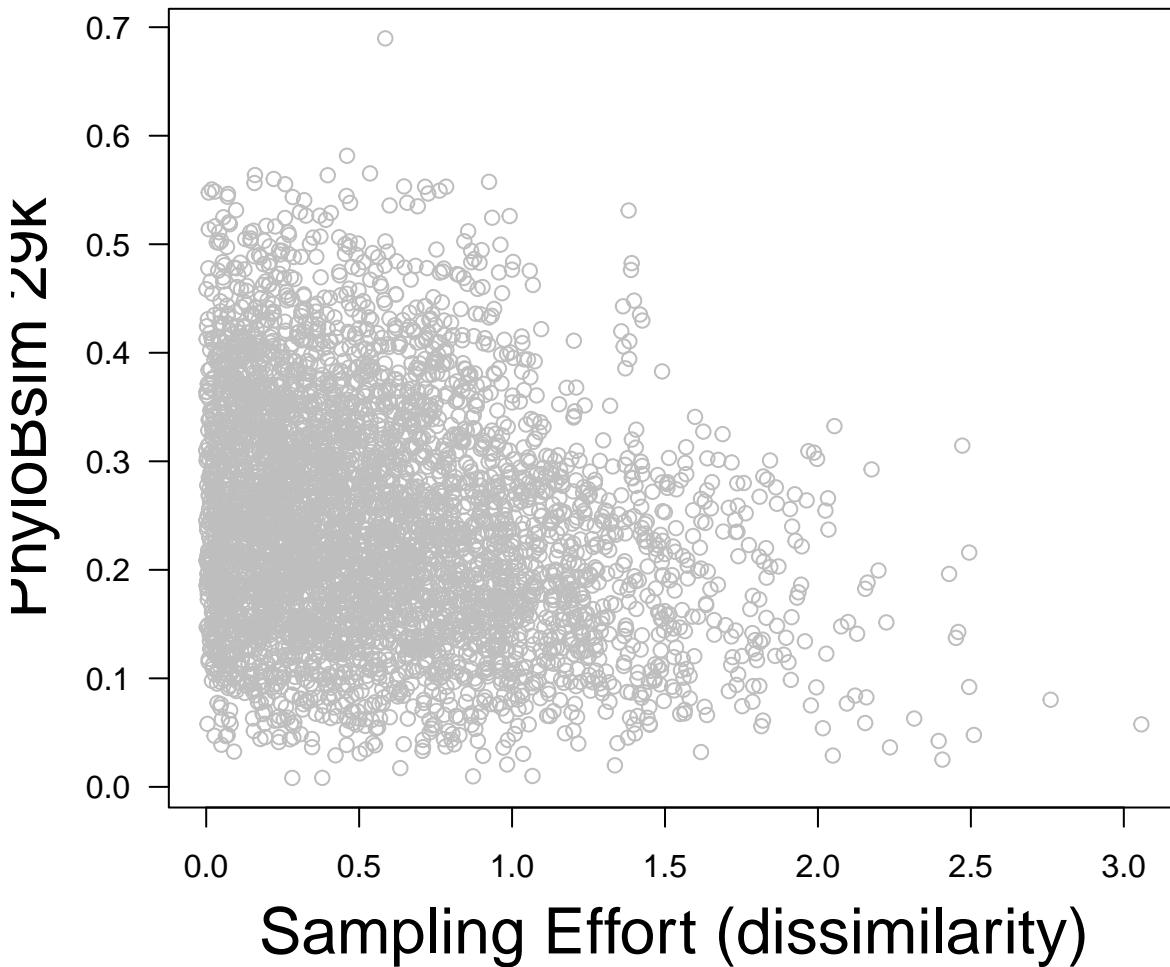


```

par(mfrow=c(1,1),cex.lab=2)
plot(PBsim ~ D_Habitat_SampEff, ylab="PhyloBsim 29k",
      xlab="Sampling Effort (dissimilarity)",
      las=1, cex = 1, col="gray")

```

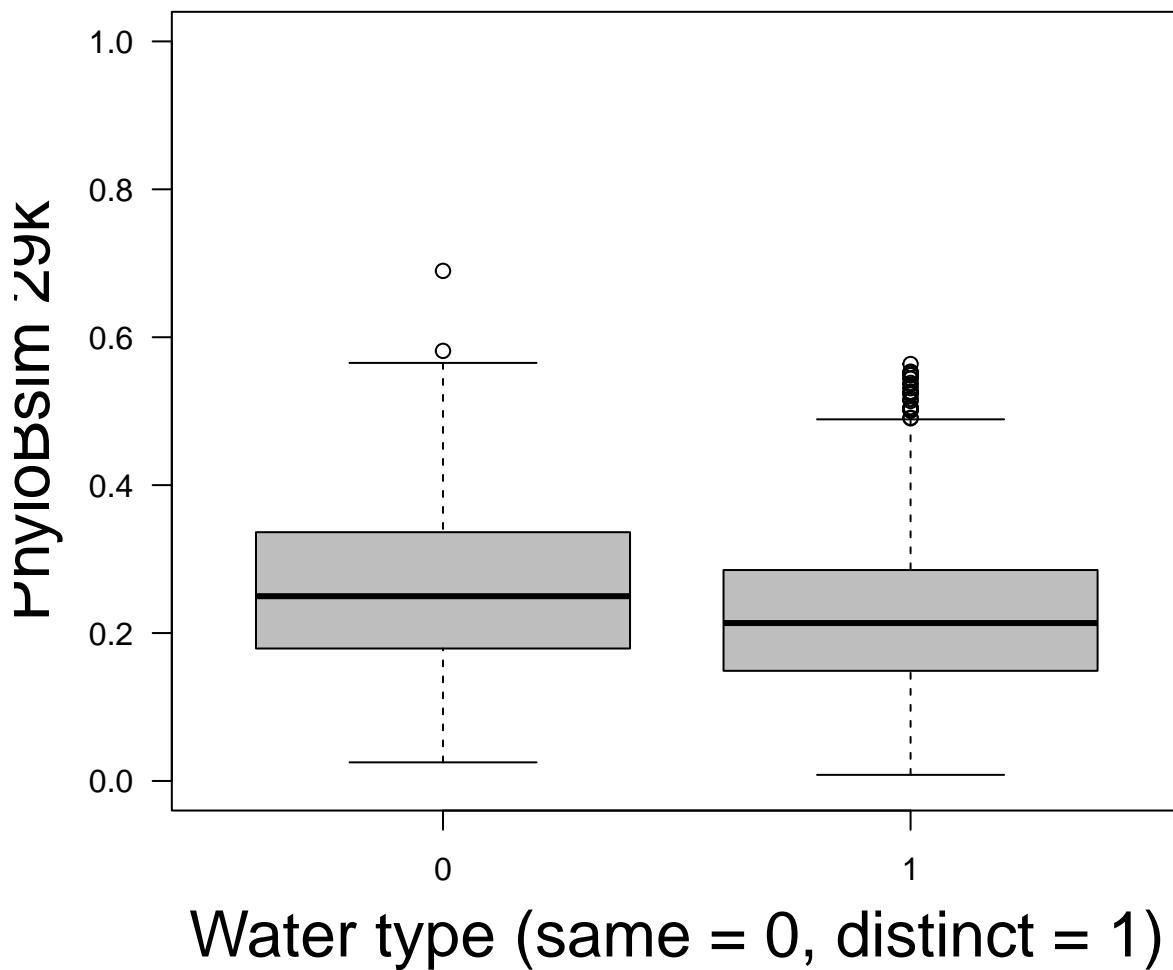
6.2.4.2 Phylogenetic beta diversity (Rabosky genetic + polytomies, Phylo 29k species)



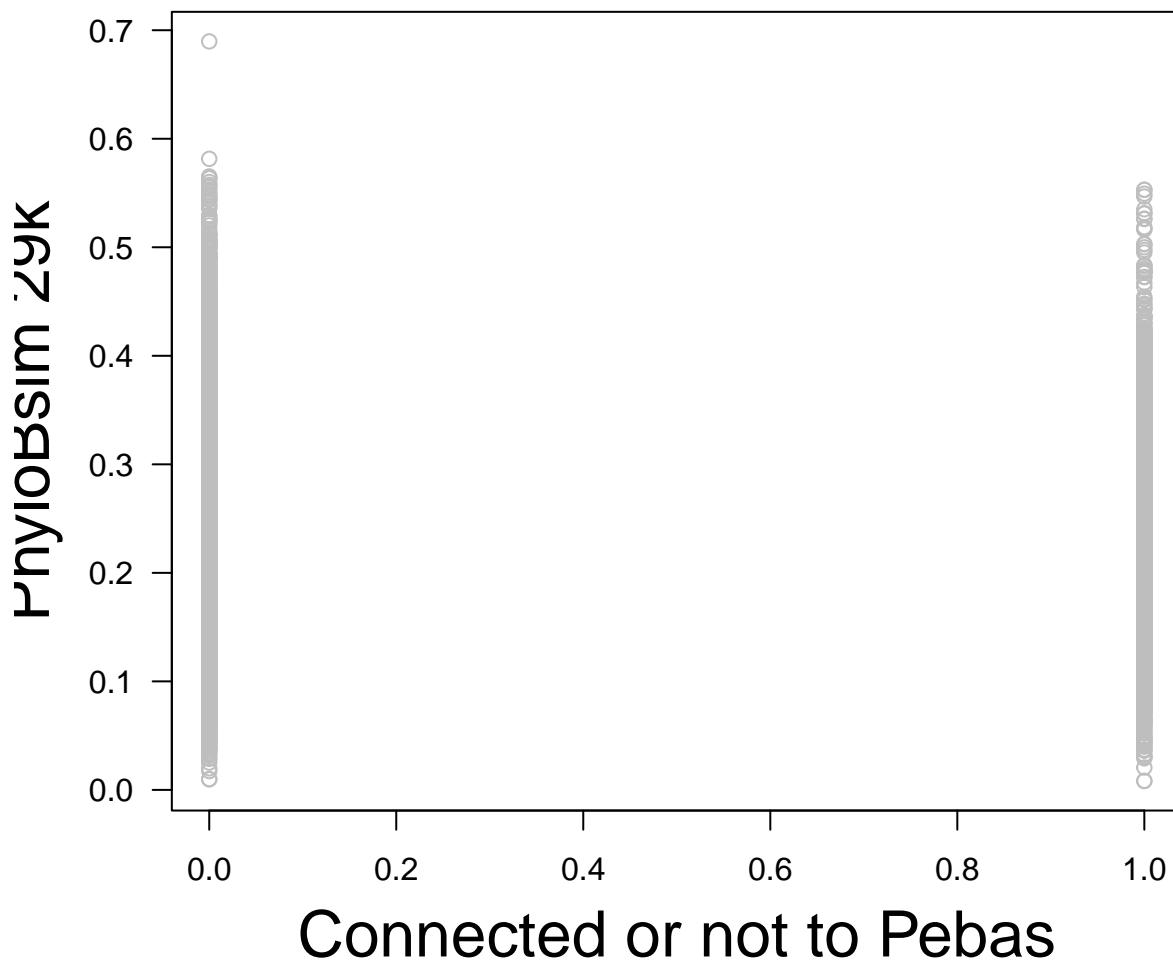
```

# jpeg("CompositionBsim_WaterType.jpeg",width = 2000,height = 2000,res = 400)
plot(as.vector(PBsim) ~ as.factor(as.vector(D_Habitat_water)),
      ylab="PhyloBsim 29k",
      xlab="Water type (same = 0, distinct = 1)",
      #vertical=T,method="jitter",
      las=1, cex = 1, col="gray",ylim=c(0,1))

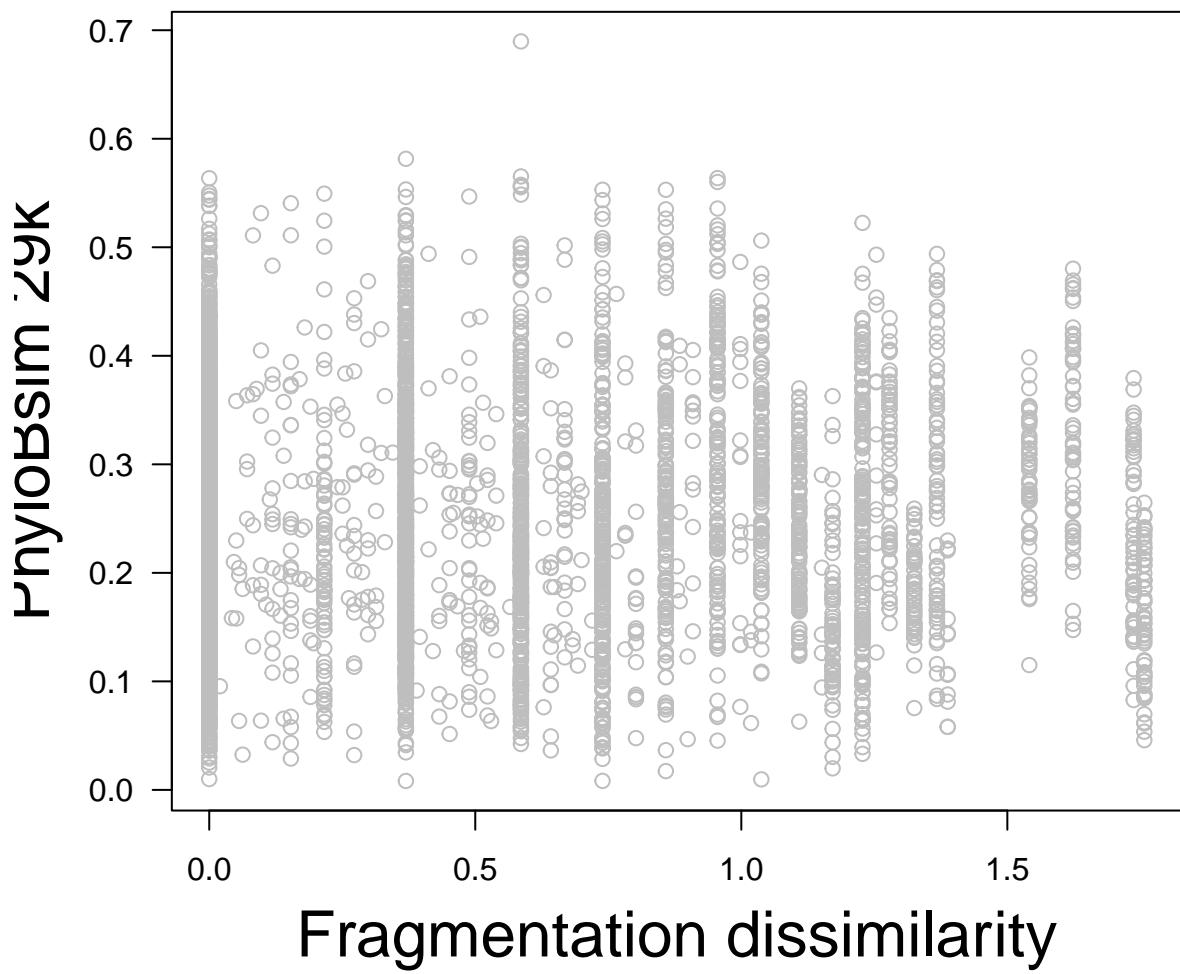
```



```
# dev.off()
plot(PBsim ~ D_pastConnect$Pebas23mya, ylab="PhyloBsim 29k",
      xlab="Connected or not to Pebas",
      las=1, cex = 1, col="gray")
```



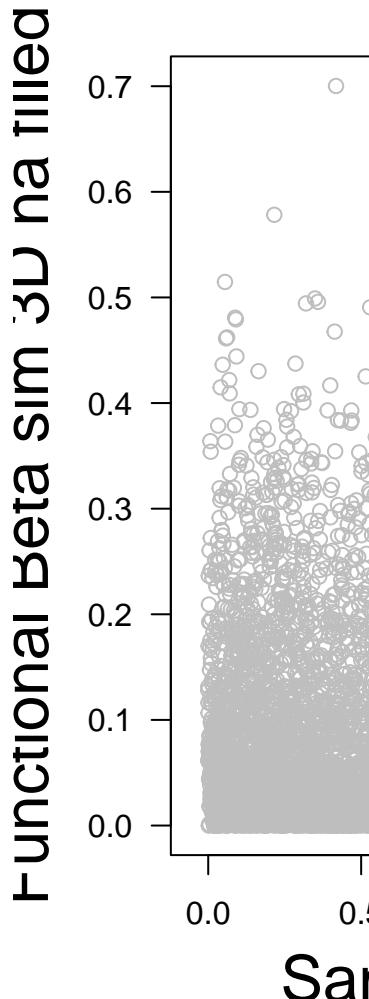
```
plot(PBsim ~ D_Habitat_fragm, ylab="PhyloBsim 29k",
      xlab="Fragmentation dissimilarity",
      las=1, cex = 1, col="gray")
```



```

par(mfrow = c(1, 1), cex.lab = 2)
plot(Fun_Bsim ~ D_Habitat_SampEff, ylab = "Functional Beta sim 3D na filled",
     xlab = "Sampling effort Dissimilarity", las = 1, cex = 1,
     col = "gray")

```



6.2.4.3 Functional Beta Sim 3D na Filled (WITH and WITHOUT OUTLIER)

```

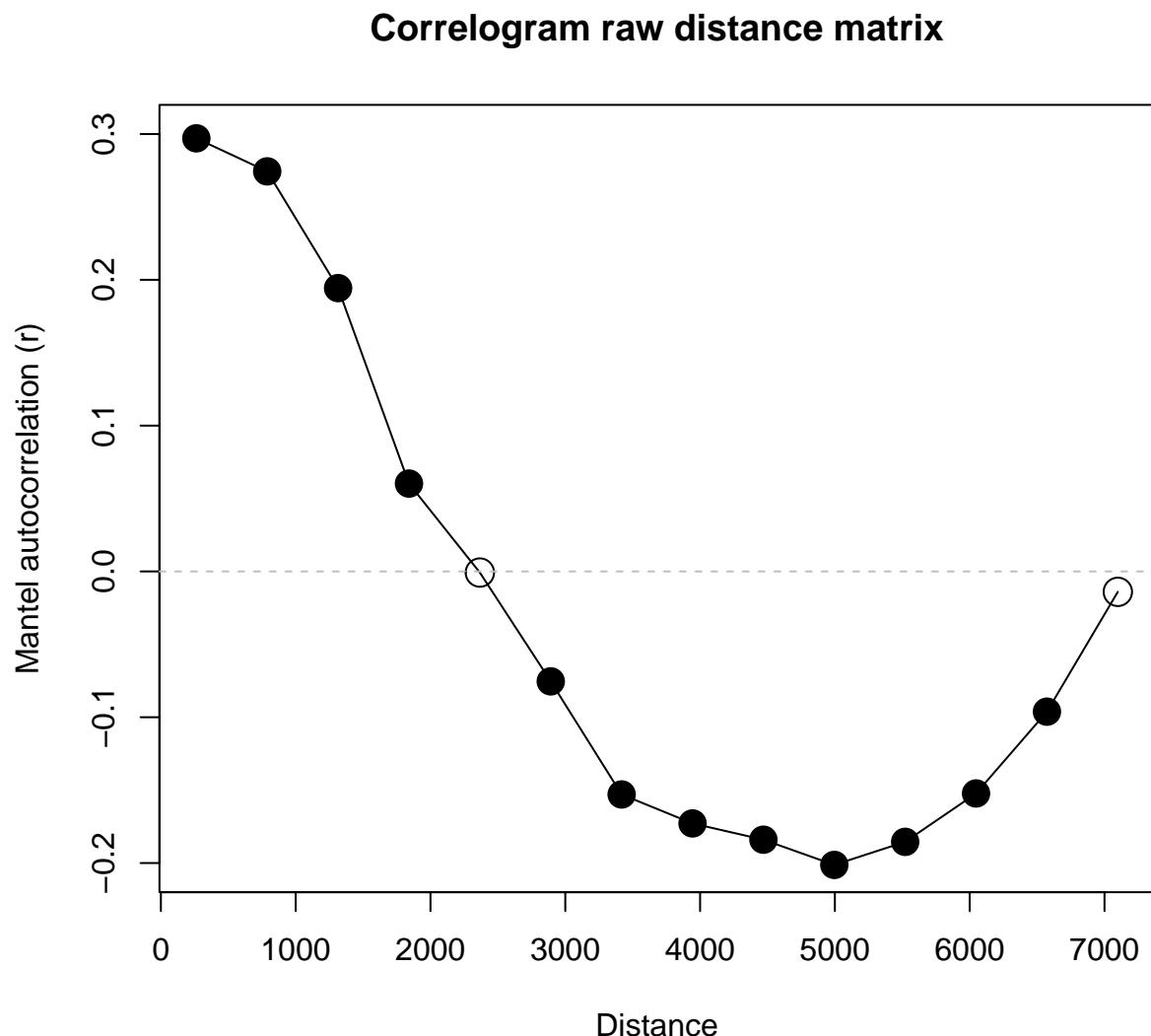
# plot(Fun_Bsim_withoutOutlier ~ D_Habitat_SampEff,
# ylab='Functional Beta sim 3D na filled without OUTLIER',
# xlab='Sampling effort Dissimilarity', las=1, cex = 1,
# col='gray')

```

6.3 Spatial autocorrelation

We are not following (Lichstein 2006; Murray *et al.* 2015) on their strategy for teasing apart the spatial structure from MRM models. Rather, we compute correlograms (note both are on the same y scale) based on the raw fish distance matrix and than on the residuals of our modelling procedure (MRMglm).

```
#Spatial struture in raw fish dissimilarity
library(ecodist)
correlog <- mgram(species.d = D_fish$beta.sim, space.d = D_spat,
                    nperm = 1000);#correlog
#par(mfrow=c(1,2))
plot(correlog,ylim=c(-0.2,0.3),
     main='Correlogram raw distance matrix')
```



```
#spatial structure on MRMglm model residuals
resu <- glm(as.vector((D_fish$beta.sim)) ~
              Standard(D_energ) + Standard(D_water) + #current energy/water hypotheses
              Standard(D_temp) + #current temperature: fast speciation hypotheses
```

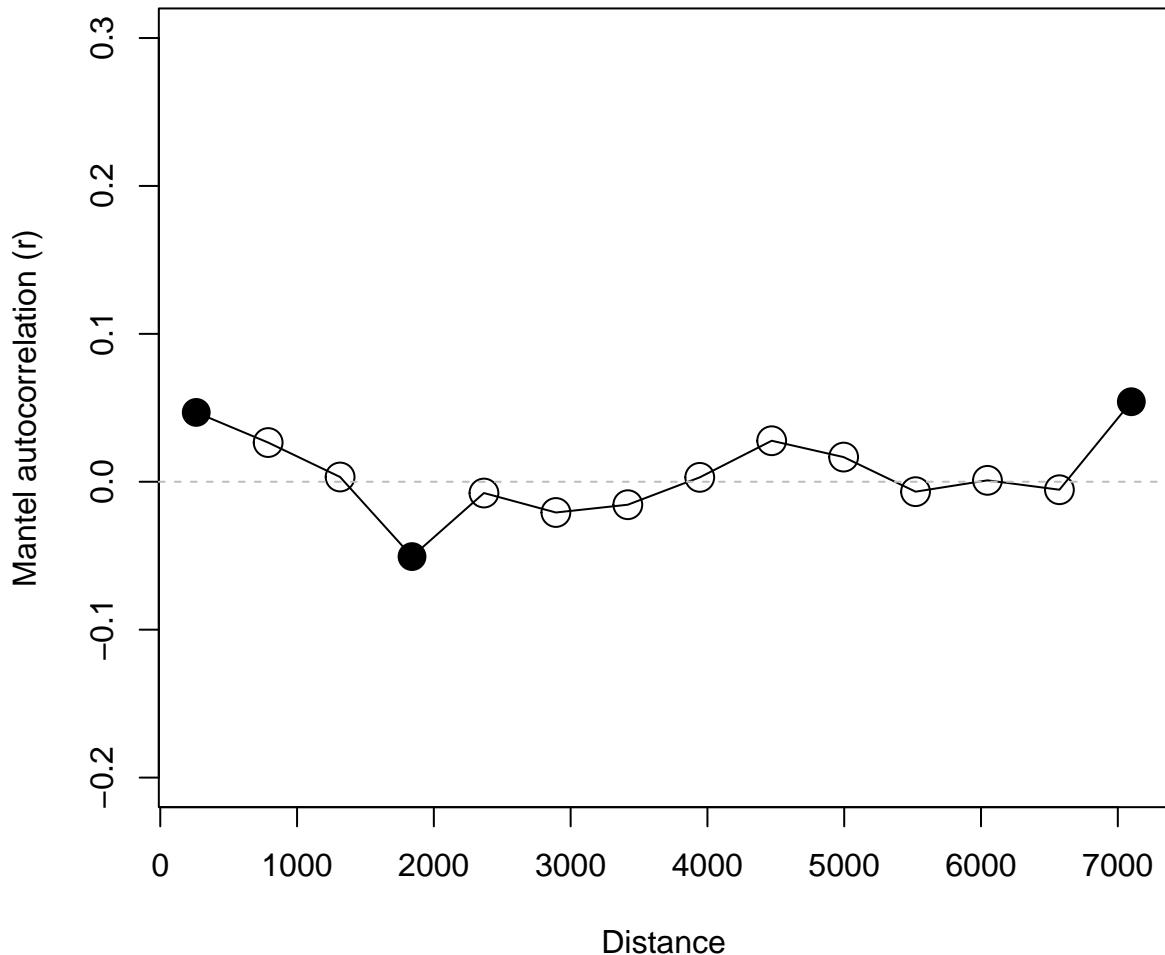
```

Standard(D_Habitat_size) + #Habitat size (area + riverine density)
Standard(D_Habitat_Div) + #Habitat Diversity (forest type + soil types)
Standard(D_Habitat_harsh) + #Habitat elevation gradient (elev mean, range, CSup15, Pr
Standard(D_Habitat_position) + #Habitat position
Standard(D_Habitat_fragm) + #Habitat fragmentation
Standard(D_Habitat_water) + #Habitat water type
Standard(D_Habitat_SampEff) + #Sampling effort
Standard(D_Marine_Incurs) + #Historical Marine Incursions
Standard(D_CurrentLGM) + #past climate: past climate instability hypotheses
Standard(D_pastConnect$Pebas23mya) + #Connected or not to Pebas Lake
Standard(D_spat),
family=binomial(link='logit'))
## Warning in eval(family$initialize): non-integer #successes in a binomial glm!

correlog_resid <- mgram(species.d = resid(resu,type = 'pearson'), space.d = D_spat,
                         nperm = 1000);#correlog_resid
plot(correlog_resid,ylim=c(-0.2,0.3),
     main='Correlogram on MRMglm residuals (on the response scale)')

```

Correlogram on MRMglm residuals (on the response scale)



```
rm(resu,correlog, correlog_resid)
```

7 Plotting maps of biological data

```
fish<-merge(fish,one,by="basin")

#merging fish data base
head(fish)

##      basin WaterColor Area_km Chut_hydrofall Chut_grand NetworkDensity
## 1    Abuna     White   31970          0          0       0.165
## 2 Amazon1     White   15948          0          0       0.135
## 3 Amazon2     White   45757          0          0       0.082
## 4 Amazon5     White   14602          0          0       0.087
## 5 Amazon8     White   17154          2          0       0.151
## 6 Amazon9     White   6190           0          0       0.103
##   Elevation_min Elevation_max Elevation_range Elevation_mean Elevation_std
## 1            95         301          206        172          31
## 2             0         377          379         41          57
## 3             0         818          820        154         166
## 4             0         183          185         26          34
## 5             3         214          211         64          46
## 6             3         152          149         54          37
##   Elevation_basinSup1000m Elevation_basinSup1500m NbEspeceEnveloppe
## 1                  0                      0          777
## 2                  0                      0          126
## 3                  0                      0          668
## 4                  0                      0          847
## 5                  0                      0          964
## 6                  0                      0          954
##   DistanceEmbouchure_km X23ma_cd          X23Ma X10Ma_cd      X10Ma
## 1            2884          A PebasSudEst          A2  AcreSud
## 2              0          B GroupAtlantik          B2 AcreSystem
## 3            154          B GroupAtlantik          B2 AcreSystem
## 4            728          B GroupAtlantik          B2 AcreSystem
## 5           1153          B GroupAtlantik          B2 AcreSystem
## 6           1427          B GroupAtlantik          B2 AcreSystem
##   NbSites1km_0417 NbSites_0417 Diversity0417 Endem0417 PebasLake PebasConec
## 1            31          50          229          0          1      Pebas
## 2            46          109          154          2          0     Other
## 3           207          465          486          1          0     Other
## 4           136          248          425          1          0     Other
## 5           141          299          566          2          0     Other
## 6           167          433          532          0          0     Other
##   ShannonDiv SoilDiv CSup15 CBelow15 Basin codage Código Sub_drainage Inf25m
## 1      0.28    1.31   0.00     100 Madera      12     12      Abuna      0
## 2      1.07    0.91   0.67      98 Amazon       1      1 Amazon1     64
## 3      1.04    1.35  20.23      80 Amazon       1      1 Amazon2     31
## 4      1.52    0.76   0.00     100 Amazon       1      1 Amazon3     68
## 5      0.84    0.73   0.00     100 Amazon       1      1 Amazon4     30
## 6      1.07    0.68   0.00     100 Amazon       1      1 Amazon5     34
##   Ent25_100m Area_log Chut_hydrofall_log ElevRge_log ElevMean_log ShanDiv_log
```

```

## 1      0    10.4          0.0    5.3    5.2    0.24
## 2     89     9.7          0.0    5.9    3.7    0.73
## 3     53    10.7          0.0    6.7    5.0    0.71
## 4     95     9.6          0.0    5.2    3.3    0.92
## 5     76     9.8          1.1    5.4    4.2    0.61
## 6     87     8.7          0.0    5.0    4.0    0.73
##   CSup15_log Inf25m_log SamplingEff ElevPro1000_asin NetworkDensity.s
## 1     0.00     0.0     -6.5       0     0.33
## 2     0.51     4.2     -5.0       0    -0.14
## 3     3.06     3.5     -4.6       0    -0.95
## 4     0.00     4.2     -4.1       0    -0.88
## 5     0.00     3.4     -4.0       0     0.11
## 6     0.00     3.6     -2.7       0    -0.63
##   SamplingEff.s DistanceEmbouchure_km.s Area_log.s Chut_hydrofall_log.s
## 1     -0.29           0.16    -0.259    -0.35
## 2      0.40           -0.89   -0.728    -0.35
## 3      0.59           -0.83   -0.017    -0.35
## 4      0.83           -0.62   -0.788    -0.35
## 5      0.84           -0.47   -0.679     0.24
## 6      1.49           -0.37  -1.367    -0.35
##   codage.s ElevRge_log.s ElevMean_log.s ElevPro1000_asins.s Inf25m_log.s
## 1     0.019      -0.39    -0.077    -0.23    -0.27
## 2    -0.759      -0.17    -0.754    -0.23    1.56
## 3    -0.759      0.12    -0.131    -0.23    1.25
## 4    -0.759      -0.43    -0.957    -0.23    1.59
## 5    -0.759      -0.38    -0.545    -0.23    1.24
## 6    -0.759      -0.51    -0.619    -0.23    1.29
##   CSup15_log.s TaxoBsim PhyloBsim30k PCDp30k FunctBsim Dpw_mean3k
## 1     -0.48      0.51     0.26     1.2    0.218   -2.45
## 2     -0.31      0.56     0.28     1.1    0.089     0.16
## 3      0.49      0.41     0.19     1.2    0.041   -0.29
## 4     -0.48      0.37     0.16     1.1    0.035   -0.43
## 5     -0.48      0.39     0.18     1.2    0.042   -0.82
## 6     -0.48      0.42     0.20     1.2    0.050     0.20
##   Median_TaxoBsim Median_PBsim
## 1      0.57      0.28
## 2      0.58      0.28
## 3      0.40      0.18
## 4      0.36      0.16
## 5      0.38      0.18
## 6      0.39      0.19

```

```

rownames(fish) <- fish$basin
fish <- fish[amazon$BvNiv2,]
fish$BvNiv2 <- fish$basin
names(fish)

```

```

## [1] "basin"                  "WaterColor"
## [3] "Area_km"                "Chut_hydrofall"
## [5] "Chut_grand"              "NetworkDensity"
## [7] "Elevation_min"            "Elevation_max"
## [9] "Elevation_range"          "Elevation_mean"

```

```

## [11] "Elevation_std"           "Elevation_basinSup1000m"
## [13] "Elevation_basinSup1500m" "NbEspeceEnveloppe"
## [15] "DistanceEmbouchure_km"   "X23ma_cd"
## [17] "X23Ma"                  "X10Ma_cd"
## [19] "X10Ma"                  "NbSites1km_0417"
## [21] "NbSites_0417"           "Diversity0417"
## [23] "Endem0417"             "PebasLake"
## [25] "PebasConec"            "ShannonDiv"
## [27] "SoilDiv"                "CSup15"
## [29] "CBelow15"               "Basin"
## [31] "codage"                 "Codigo"
## [33] "Sub_drainage"           "Inf25m"
## [35] "Ent25_100m"              "Area_log"
## [37] "Chut_hydrofall_log"      "ElevRge_log"
## [39] "ElevMean_log"            "ShanDiv_log"
## [41] "CSup15_log"              "Inf25m_log"
## [43] "SamplingEff"             "ElevPro1000_asin"
## [45] "NetworkDensity.s"        "SamplingEff.s"
## [47] "DistanceEmbouchure_km.s" "Area_log.s"
## [49] "Chut_hydrofall_log.s"    "codage.s"
## [51] "ElevRge_log.s"           "ElevMean_log.s"
## [53] "ElevPro1000_asin.s"      "Inf25m_log.s"
## [55] "CSup15_log.s"            "TaxoBsim"
## [57] "PhyloBsim30k"            "PCDp30k"
## [59] "FunctBsim"               "Dpw_mean3k"
## [61] "Median_TaxoBsim"         "Median_PBsim"
## [63] "BvNiv2"

```

names(amazon)

```

## [1] "Basin"          "BvNiv2"          "SupBv1"          "SupBv2"          "WaterColor"
## [6] "Area_km"        "Chut_hydro"       "Chut_grand"     "NetworkDen"     "Elevatio_2"
## [11] "NbEspeceSi"    "NbEspeceEn"      "NbSitesTot"     "Endemism"      "DistanceEm"
## [16] "X23ma_cd"      "X23Ma"           "X10Ma_cd"       "X10Ma"          "geometry"
## [21] "Cores"          "Bsim_m"           "UnderSampled"

```

library(dplyr)

```

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:igraph':
##
##     as_data_frame, groups, union

## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarise

```

```
## The following object is masked from 'package:nlme':
##
##     collapse

## The following object is masked from 'package:MASS':
##
##     select

## The following object is masked from 'package:ape':
##
##     where

## The following objects are masked from 'package:stats':
##
##     filter, lag

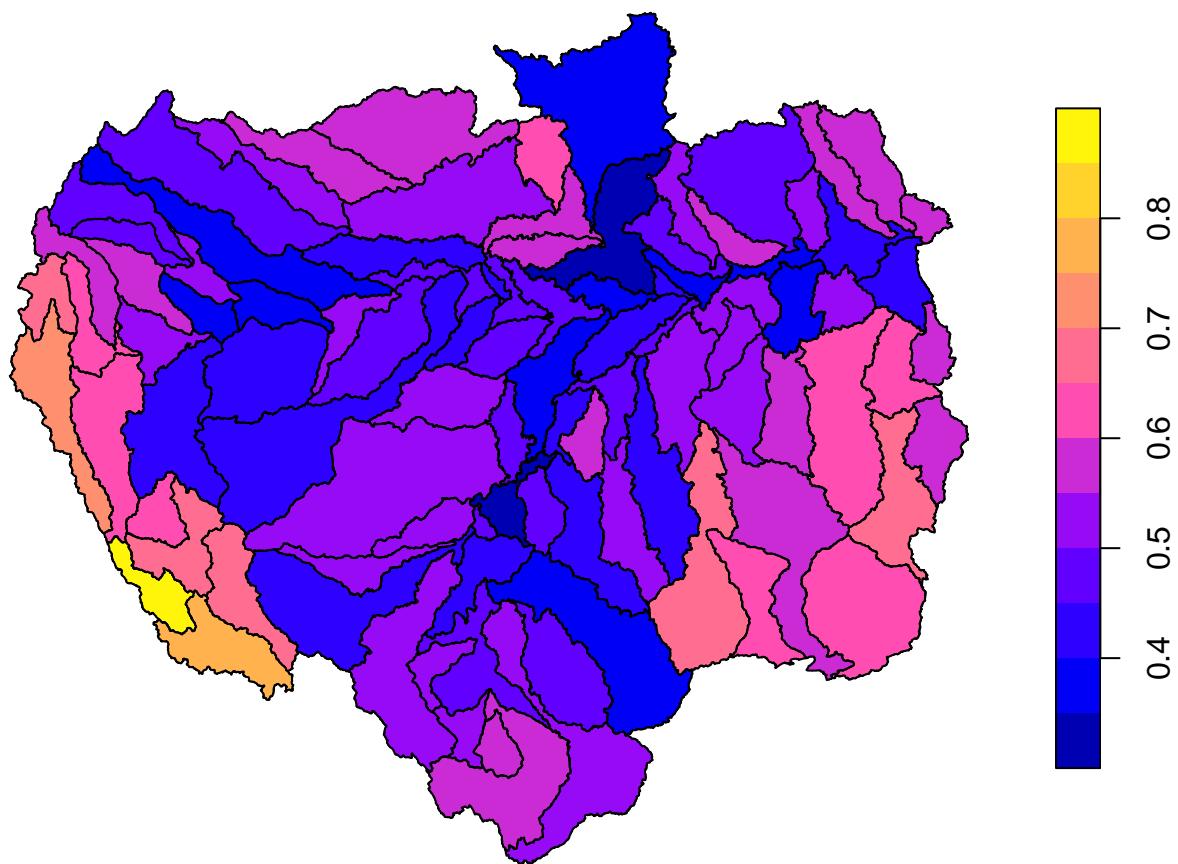
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

amazon = left_join(amazon,fish,by="BvNiv2")
is(amazon)

## [1] "sf"      "oldClass"

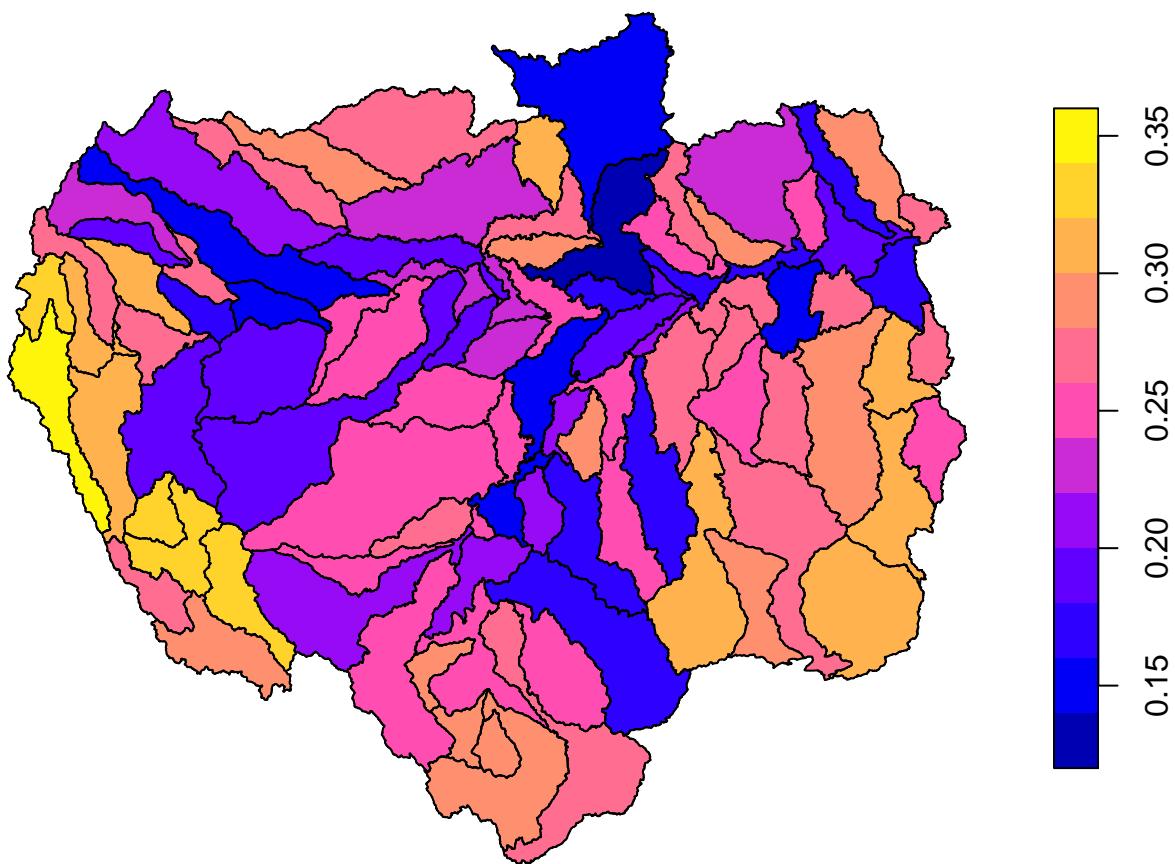
plot(amazon["TaxoBsim"])
```

TaxoBsim



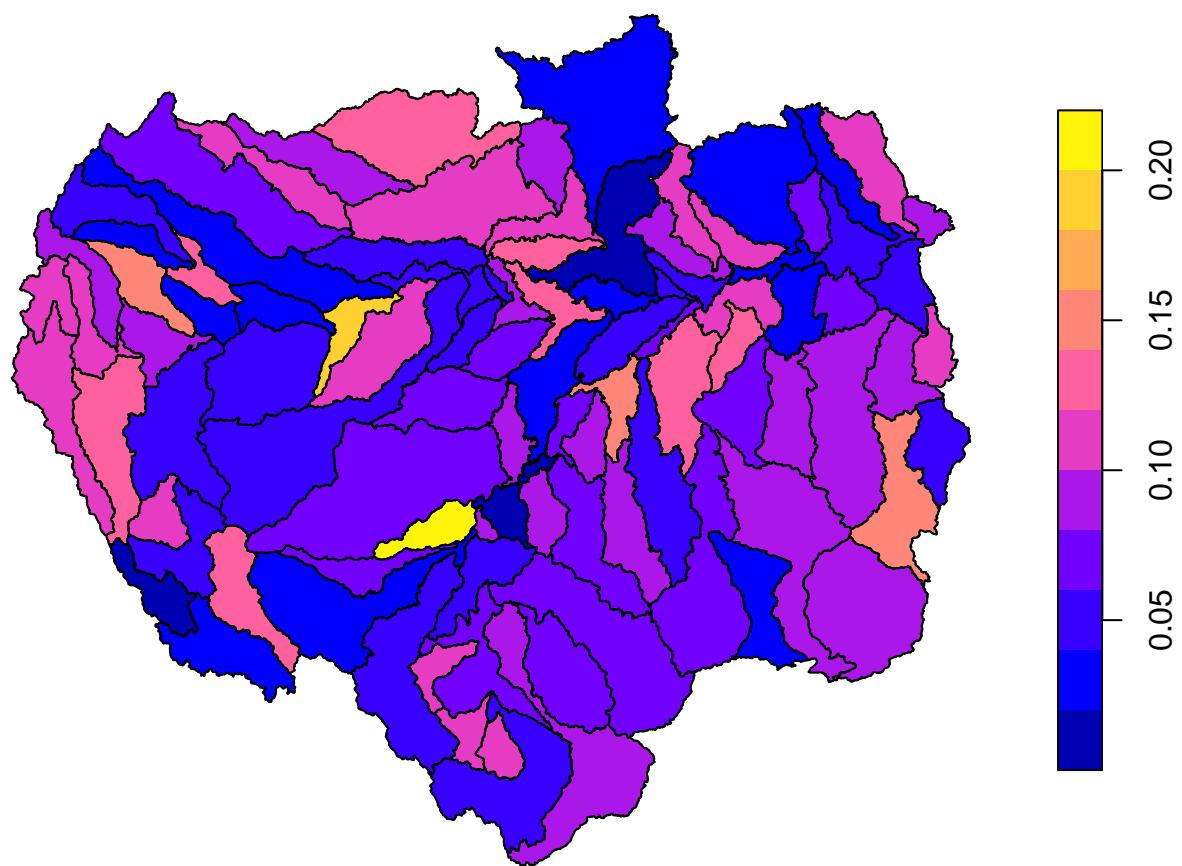
```
plot(amazon["PhyloBsim30k"])
```

PhyloBsim30k



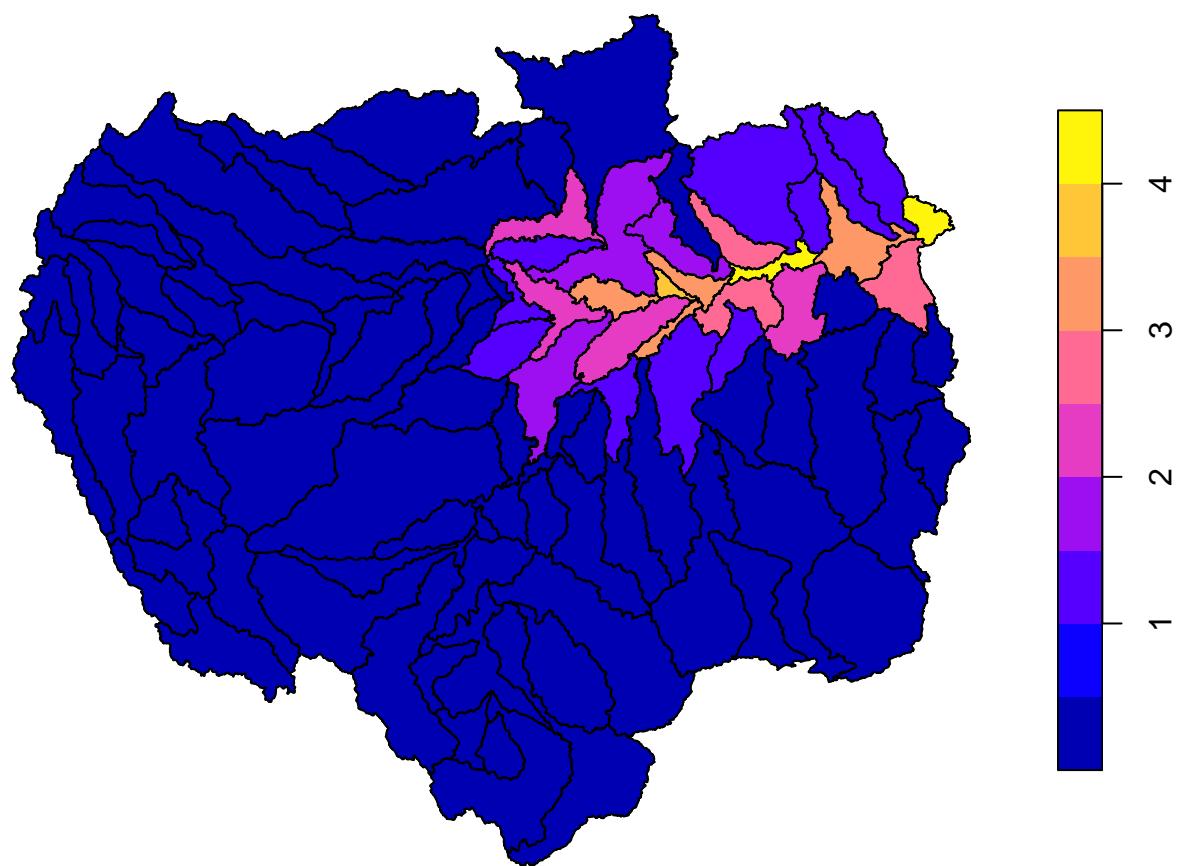
```
plot(amazon["FunctBsim"])
```

FunctBsim



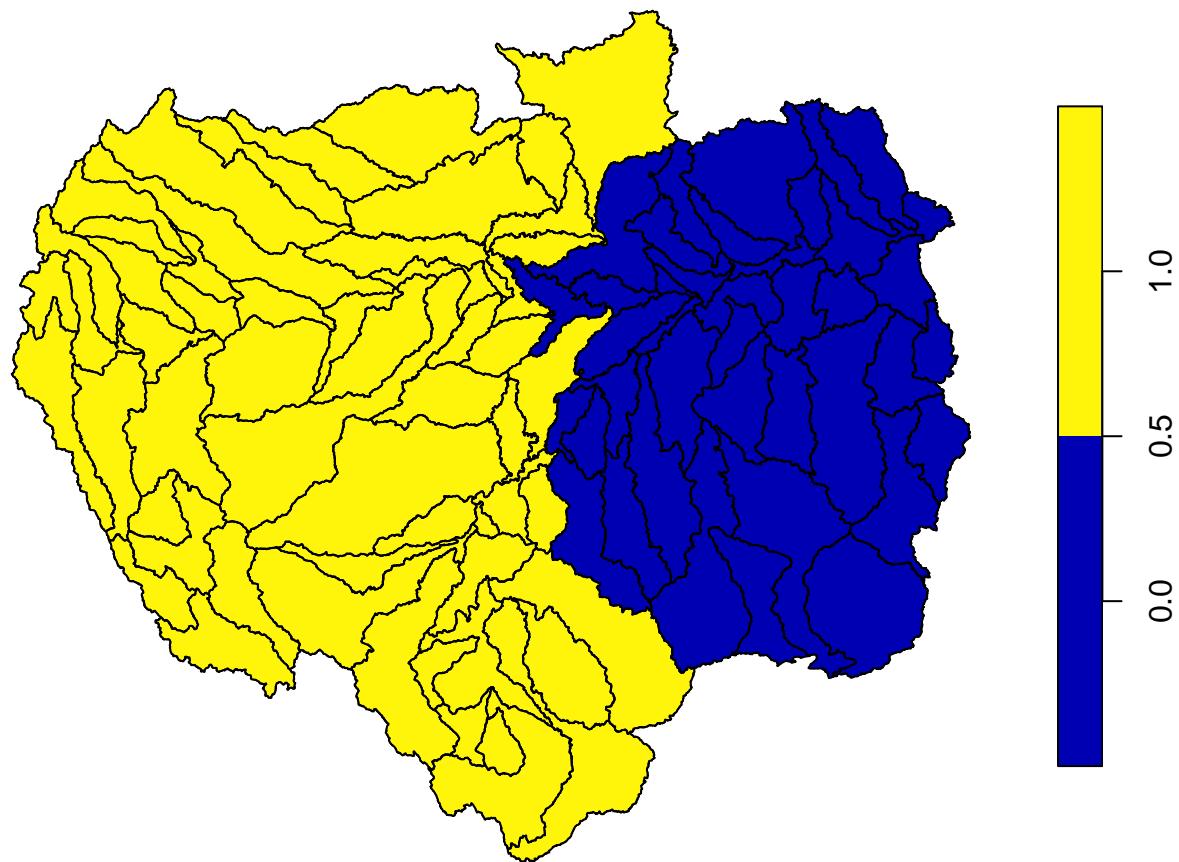
```
plot(amazon["Inf25m_log"],main="marine incursion")
```

marine incursion



```
plot(amazon["PebasLake"])
```

PebasLake



```
#Fig1 with ggplot
```

```
library(ggplot2)
library(dplyr)
library(viridis)
```

```
## Loading required package: viridisLite
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
##
## The following object is masked from 'package:dplyr':
##      combine
```

```

library(maps)

## 
## Attaching package: 'maps'

## The following object is masked from 'package:viridis':
## 
##     unemp

## The following object is masked from 'package:plyr':
## 
##     ozone

## The following object is masked from 'package:cluster':
## 
##     votes.repub

# library(raster)
# library(RStoolbox)

WorldData <- map_data('world')
#WorldData<-fortify(WorldData)
wld<-ggplot() + theme_bw()
wld<-wld+geom_map(data=WorldData, map=WorldData,
                    aes(x=long, y=lat, map_id=region),
                    color="white", fill="gray50", linewidth=0.2) + #fill="gray92", size=0.05, alpha=1,
                    xlim(-81,-50)+ylim(-21,6)

## Warning in geom_map(data = WorldData, map = WorldData, aes(x = long, y = lat, :
## Ignoring unknown aesthetics: x and y

# wld+geom_point(data=pts, aes(x=CoordX, y=CoordY), size=0.7, alpha=0.5)
# wld+geom_point(data=gridA, aes(x=long, y=lat, group=group), size=0.7, alpha=0.5)

# head(amazon)
# dim((amazon))
amazon_d<-amazon
# # head(amazon_d)
# # head(fish)
# amazon_d<-merge(x=amazon_d, y=fish, by.x="BvNiv2", by.y="BvNiv2", all.x=T)
# # head(amazon_d)

ditch_the_axes <- theme(
  axis.text = element_blank(),
  #axis.line = element_blank(),
  axis.ticks = element_blank(),
  panel.border = element_rect(fill = FALSE),
  #panel.grid = element_blank(),

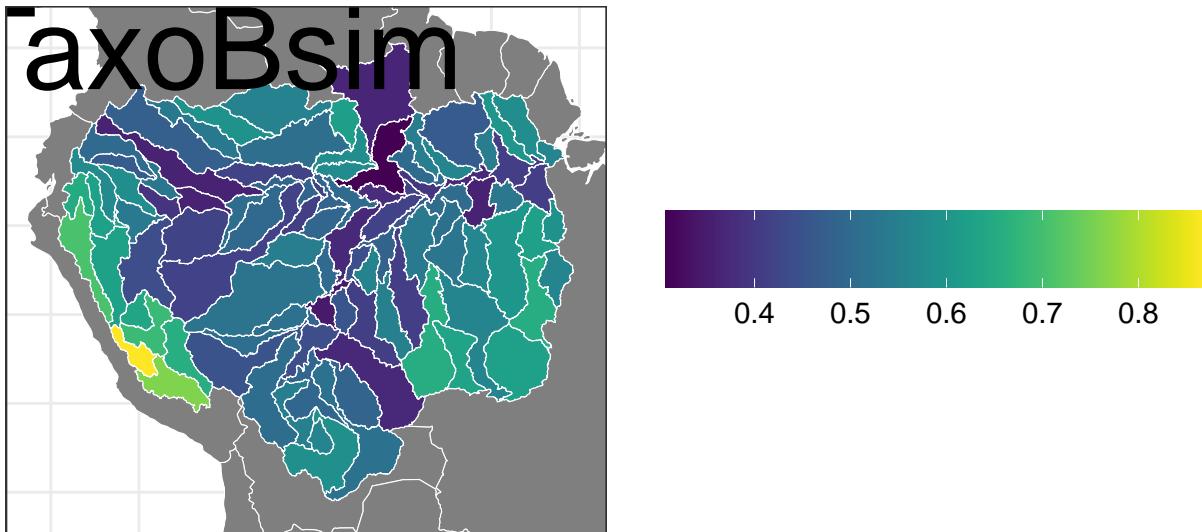
```

```

axis.title = element_blank()
)

fig1a<-wld+
  geom_sf(data=amazon_d,aes(fill=TaxoBsim),#catego_cores(TaxoBsim)
           color="white",size=0.2)+
  scale_fill_viridis(name = "",begin = 0, end = 1,
                     guide = guide_colorbar(barwidth = 14, barheight = 2))+#
  #scale_fill_gradient(na.value = "white",low="blue",high="red")+
  ditch_the_axes+
  theme(legend.position.inside = c(0.23, 0.09),legend.direction = "horizontal",legend.text=element_text(size=15))
  annotate("text", x=-70, y=5, label= "a) TaxoBsim      ",size=15)
fig1a

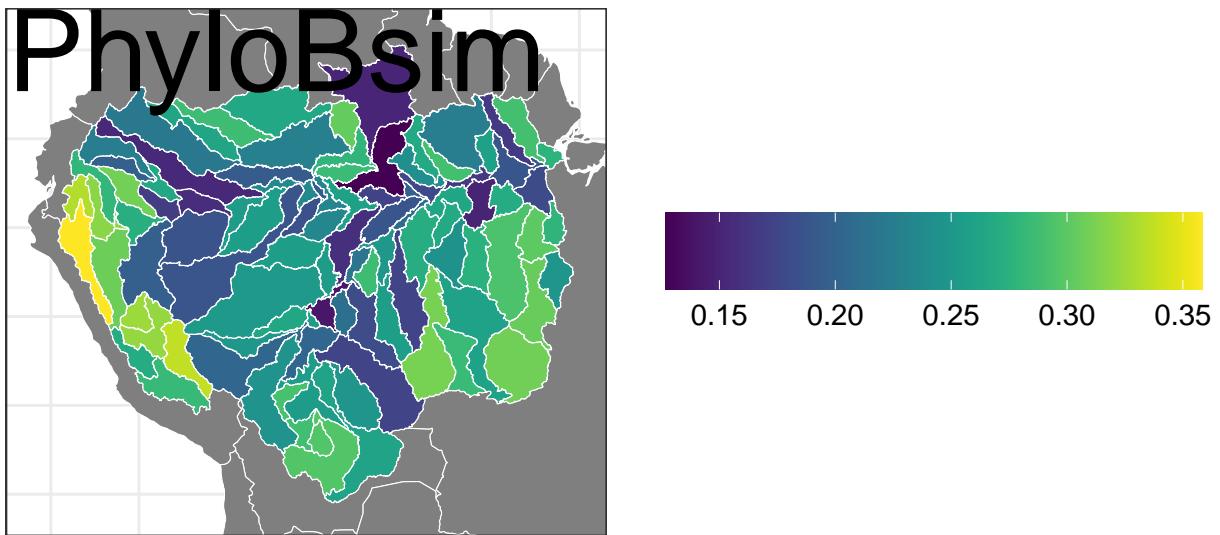
```



```

fig1b<-wld+
  geom_sf(data=amazon_d,aes(fill=PhyloBsim30k),#catego_cores(PhyloBsim30k)
           color="white",size=0.2)+ 
  scale_fill_viridis(name = "",begin = 0, end = 1,guide = guide_colorbar(barwidth = 14, barheight =
#scale_fill_gradient(na.value = "white",low="blue",high="red")+
ditch_the_axes+
  theme(legend.position.inside = c(0.23, 0.09),legend.direction = "horizontal",legend.text=element_text
  annotate("text", x=-71, y=5, label= "b) PhyloBsim",size=15)
fig1b

```



```

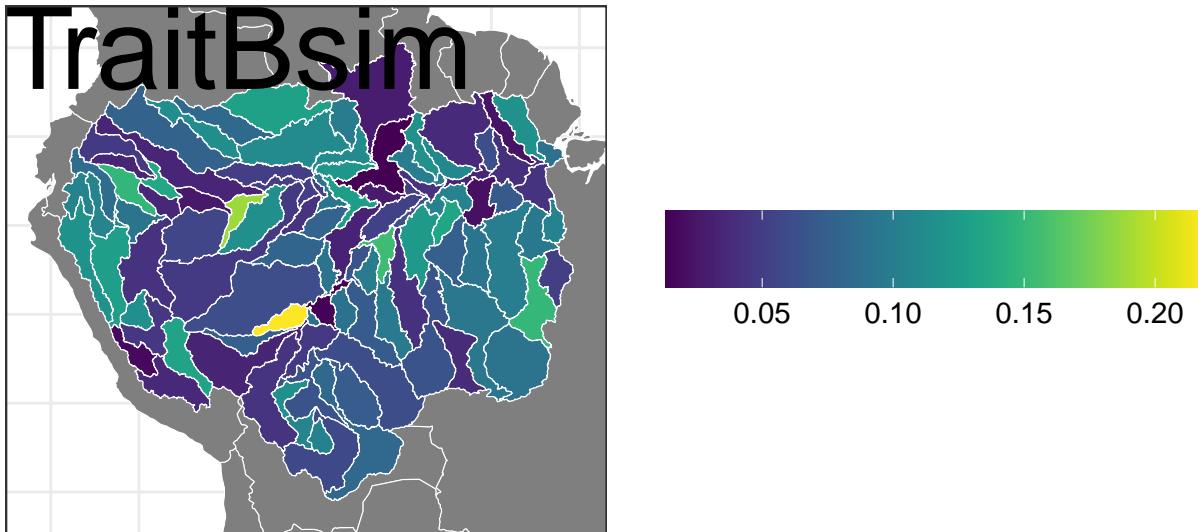
# fig1c<-wld+
#   geom_polygon(data=amazon_d,aes(long,lat,group=group,fill=PCDp30k),color="white",size=0.2)+ 
#   scale_fill_viridis(name = "",begin = 0.1, end = 1,guide = guide_colorbar(barwidth = 14, barheight =
# #scale_fill_gradient(na.value = "white",low="blue",high="red")+
#   ditch_the_axes+
#   theme(legend.position = c(0.22, 0.09),legend.direction = "horizontal",legend.text=element_text
#   annotate("text", x=-75, y=5, label= "C) PCDp      ",size=15)
# fig1c

```

```

fig1d<-wld+
  geom_sf(data=amazon_d,aes(fill=FunctBsim),#catego_cores(FunctBsim))
    color="white",size=0.2)+ 
  scale_fill_viridis(name = "",begin = 0, end = 1,guide = guide_colorbar(barwidth = 14, barheight =
#scale_fill_gradient(na.value = "white",low="blue",high="red")+
ditch_the_axes+
  theme(legend.position.inside = c(0.23, 0.09),legend.direction = "horizontal",legend.text=element_text(size=15))
  annotate("text", x=-73, y=5, label= "c) TraitBsim",size=15)
fig1d

```



```

#save
g <- arrangeGrob(fig1a,fig1b, fig1d, nrow=1) #generates g
#g
ggsave(file="Fig1.jpeg", g,units = "cm",width = 55,height = 15,dpi = 300)

#

```

8 R + knitr Section Informations

```
## R version 4.5.1 (2025-06-13)
## Platform: x86_64-pc-linux-gnu
## Running under: Ubuntu 22.04.5 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblas-p0.3.20.so;  LAPACK version 3.10
##
## locale:
## [1] LC_CTYPE=pt_BR.UTF-8      LC_NUMERIC=C
## [3] LC_TIME=pt_BR.UTF-8      LC_COLLATE=en_US.UTF-8
## [5] LC_MONETARY=pt_BR.UTF-8    LC_MESSAGES=en_US.UTF-8
## [7] LC_PAPER=pt_BR.UTF-8      LC_NAME=C
## [9] LC_ADDRESS=C              LC_TELEPHONE=C
## [11] LC_MEASUREMENT=pt_BR.UTF-8 LC_IDENTIFICATION=C
##
## time zone: America/Sao_Paulo
## tzcode source: system (glibc)
##
## attached base packages:
## [1] stats      graphics   grDevices utils      datasets   methods    base
##
## other attached packages:
## [1] maps_3.4.2          gridExtra_2.3       viridis_0.6.5
## [4] viridisLite_0.4.2    ggplot2_3.5.2       dplyr_1.1.4
## [7] ecodist_2.1.3        igraph_2.0.1.1     betapart_1.6
## [10] picante_1.8.2       plyr_1.8.9         cluster_2.1.8.1
## [13] sf_1.0-20            segmented_2.1-4    nlme_3.1-168
## [16] MASS_7.3-65          recluster_2.9      RColorBrewer_1.1-3
## [19] ape_5.8-1            vegan_2.7-1       permute_0.9-7
## [22] reshape2_1.4.4       formatR_1.14      knitcitations_1.0.12
## [25] knitr_1.50
##
## loaded via a namespace (and not attached):
## [1] DBI_1.2.3           mnormt_2.1.1      phangorn_2.12.1
## [4] rlang_1.1.6          magrittr_2.0.3     e1071_1.7-16
## [7] compiler_4.5.1       mgcv_1.9-1        systemfonts_1.0.5
## [10] vctrs_0.6.5          combinat_0.0-8     quadprog_1.5-8
## [13] stringr_1.5.1       pkgconfig_2.0.3    fastmap_1.2.0
## [16] backports_1.4.1     magic_1.6-1       labeling_0.4.3
## [19] rmarkdown_2.29       intertools_0.1-3   ragg_1.2.5
## [22] xfun_0.52           clusterGeneration_1.3.8 jsonlite_2.0.0
## [25] RefManageR_1.4.0     parallel_4.5.1    R6_2.6.1
## [28] stringi_1.8.7       car_3.1-2         lubridate_1.9.3
## [31] rcdd_1.6             numDeriv_2016.8-1.1 Rcpp_1.0.14
## [34] iterators_1.0.14     optimParallel_1.0-2 snow_0.4-4
## [37] Matrix_1.7-4         splines_4.5.1      timechange_0.3.0
## [40] tidyselect_1.2.1     rstudioapi_0.17.1  abind_1.4-5
## [43] yaml_2.3.10          doParallel_1.0.17  codetools_0.2-19
## [46] minpack.lm_1.2-4     curl_6.2.2        lattice_0.22-5
```

```
## [49] tibble_3.2.1           withr_3.0.2          coda_0.19-4
## [52] evaluate_1.0.3          phytools_2.4-4      units_0.8-4
## [55] proxy_0.4-27           xml2_1.3.8          pillar_1.10.2
## [58] carData_3.0-5          KernSmooth_2.23-26 foreach_1.5.2
## [61] geometry_0.4.7         generics_0.1.3       scales_1.4.0
## [64] class_7.3-23           glue_1.8.0          scatterplot3d_0.3-44
## [67] tools_4.5.1            fastmatch_1.1-4     grid_4.5.1
## [70] plotrix_3.8-4          bibtex_0.5.1         cli_3.6.5
## [73] DEoptim_2.2-8           textshaping_0.3.6   expm_1.0-0
## [76] doSNOW_1.0.20          gtable_0.3.6        digest_0.6.37
## [79] classInt_0.4-11         farver_2.1.2        htmltools_0.5.8.1
## [82] lifecycle_1.0.4          httr_1.4.7
```

Bibliography

1. A. (2009). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19, 134–143. Baselga,
2. A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, 21, 1223–1232. Baselga,
3. J.H. (2013). Why are there so many species in the tropics? *Journal of Biogeography*, 41, 8–22. Brown,
4. M.S., Oberdorff, T., Hugueny, B., Leprieur, F., J'ez'equel, C., Cornu, J., et al. (2014). Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, 17, 1130–1140. Dias,
5. A. (2007). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27, 2865–2873. Gelman,
6. C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., et al. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, 330, 927–931. Hoorn,
7. C., Tedesco, P.A., Darwall, W., Dias, M.S., Frederico, R.G., Hidalgo, M., et al. (2020). Freshwater fish diversity hotspots for conservation priorities in the amazon basin. *Conservation Biology*, 34, 956–965. J'ez'equel
8. Y. & Barraclough, TimothyG. (2010). Speciation has a spatial scale that depends on levels of gene flow. *The American Naturalist*, 175, 316–334. Kisel,
9. F., Tedesco, P.A., Hugueny, B., Beauchard, O., Dürr, H.H., Brosse, S., et al. (2011). Partitioning global patterns of freshwater fish beta diversity reveals contrasting signatures of past climate changes: Partitioning global patterns of fish beta diversity. *Ecology Letters*, 14, 325–334. Leprieur,
10. J.W. (2006). Multiple regression on distance matrices: A multivariate spatial analysis tool. *Plant Ecology*, 188, 117–131. Lichstein,
11. K.A., Preston, N., Allen, T., Zambrana-Torrelío, C., Hosseini, P.R. & Daszak, P. (2015). Global biogeography of human infectious diseases. *Proceedings of the National Academy of Sciences*, 112, 12746–12751. Murray,