Bioinformatics Server Connexion

In a Terminal:

`ssh  -X  etudiant1@172.20.9.115`

Etudiant1 must be replaced by your own identifier (etudiant2, etudiant4 etc), the one that you have chosen !!!

- Do not forget the -X … it's necessary for graphic interface (Rstudio)
- 172.20.9.115 : is the IP address of our bioinformatics server

R package, library and Functions

I want to use Phyloseq for analysing my data!!

1- Download the package Phyloseq:
Install.packages()/Bioconductor

2- load the package in your R session to use it
 library(Phyloseq)

3- Now I can **use the functions of Phyloseq in your R** session
Taxa_sums()

functions of R packages

Function name	Argument	Possible options

Seqs = getSequences(seqtab.nochim, size=…, label=..)

<-

Variable containing the results

= same as <-

Complex command (optimization): function imbrication : from inside to outside

Result = function4(function3(function2(function1())))

Result = function4(function3(function2(function1())))

Res= prune_taxa(names(sort(taxa_sums(Final.TSS.log),TRUE)[1:50]), Final.TSS.log)

Can be Decomposed as:

A = taxa_sums(Final.TSS.log)

B = sort(A,TRUE)

C = names(B[1:50])

D = prune_taxa(C,Final.TSS.log)

Common OTU/ASV table

```
> otu_table(Final2_rar)
OTU Table:               [229 taxa and 18 samples]
                         taxa are columns
        ASV1 ASV2 ASV3 ASV4 ASV5 ASV6 ASV7 ASV8 ASV9 ASV10 ASV11 ASV12 ASV13 ASV14 ASV15 ASV16
S11B     109   24   79   78   89   24   66   31    0     0     0    32    14     0     0    46
S1B       83    0   28    0   59   58    0    0   69    51    67    18     0     0     0     0
S2B       51    0   26   15   44   65    0    0   68    31    47     0     2     0    85     0
S2S       93   93   74    9  127   40   23   68    0     0     0    43    57     0     0    26
        ASV17 ASV18 ASV19 ASV20 ASV21 ASV22 ASV23 ASV24 ASV25 ASV26 ASV27 ASV28 ASV29 ASV30 ASV31
S11B       0     0    69     0    24    31    22    25     0     0     0     0     0     3     0
S1B        0    29     0    47     0     0     0     0     0     0     0     0     0     0     0
S2B        0    46     0    60     0     0    24    20     0    32    42     0     6     0     0
S2S        0     0    34     0     0    10    24    28     0     0     0     0     0     2    19
```

DADA2 by default gives you this

	A	B	C	D	E	F	G	H	I	J	K	L	M
		GGAATATTGCACAAT(	GGAAT CTTG	CGAAAACTTGACA	GGAATCTTGCACA	GGAATCTTGCACAA	GGAATATTG	GGAATATTG	GGAATATTG	GGAATTTTC(	GGAATCTTG	GGAATCTTG	GGAATATTG	AT.
S11B		131	32	97	85	98	41	63	35	0	0	0	45
S1B		73	0	25	0	54	54	0	0	66	59	65	29
S2B		54	0	40	17	45	66	0	0	66	34	45	0
S2S		120	101	83	14	124	47	41	82	0	0	0	50
S3B		65	0	33	24	54	62	0	0	74	64	61	0
S3S		89	10	0	0	44	32	0	0	57	115	54	0
S4B		13	7	44	38	45	47	0	10	45	70	40	0
S4S		79	8	42	0	65	0	0	11	165	140	51	34
S5B		199	18	80	128	0	39	67	15	0	0	48	0
S5S		213	0	39	114	53	0	44	0	37	0	44	30
S6B		187	113	62	31	84	37	78	69	0	0	0	45
S6S		84	378	38	26	0	22	50	84	0	0	0	29
S7B		60	128	83	94	0	34	27	101	0	0	0	0
S7S		91	50	65	119	0	49	84	43	0	0	0	0
S8B		66	81	64	85	0	48	69	73	0	0	0	31
S8S		77	13	99	105	56	48	64	25	0	0	18	45
S9B		86	90	53	68	85	43	66	61	0	0	0	36
S9S		106	86	70	31	38	75	66	77	0	0	47	58

→ Identifier is the sequence of 250 bp!!!
→ We will apply transformation to obtain common OTU_table with ASV names (more easy to manipulate)

Preprocessing : Dada2

- Avant de filtrer, correction des Erreurs, Derepliquer, assembler (Fonction dada2) …

→ il y a une petite mise en place : gestions des fichiers, noms des samples

Preprocessing : Dada2

> **fnFs <- sort(list.files(path, pattern="_R1.fastq", full.names=TRUE))**

lister les noms de fichiers (**list.files**) situés dans la variable **path** (à savoir répertoire ./data) et qui contiennent dans leurs noms la signature "**_R1.fastq**".

full.names donne le chemin complet (".data/S11B_R1.fastq"), sinon ça n'afficherait que "S11B_R1.fastq". Enfin les classer (**sort**).

- **Pourquoi faire ca?**

On a besoin du chemin complet pour que les fonctions puissent savoir ou sont ces fichiers et les utiliser

→**fnFs** stock simplement l'ensemble **des chemins d'accès** des fichiers fastq **R1**

→**fnFr** stock simplement l'ensemble **des chemins d'accès** des fichiers fastq **R2**

sample.names <- sapply(strsplit(basename(fnFs), "_"), `[`, 1)

- FnFs stocke le chemin complet
- **Basename(fnFs)** permet de recupérer le nom après le dernier séparateur (/)

Dans notre cas ./data/S11B_R1.fastq devient alors S11B_R1.fastq

- Ensuite **strsplit(basename(fnFs), "_")** permet de séparer le nom sur la base du "_"

 ca donne : S11B   R1.fastq   (equivalent à Excel séparateur de champ)

- **sapply(strsplit(basename(fnFs), "_"), `[`, 1)** permet de recupérer le premier élement de la ligne "S11B   R1.fastq" à savoir S11B

- **Pourquoi faire ca?**

Pour récupérer les noms de vos samples : S11B, S1B, S2S,S2B, etc

filtFs <- file.path(path, "Filtered", basename(fnFs))

Construire un chemin (file.path) ici composé de "path" puis de "Filtered" et enfin du nom basename(fnFs)

→ path = ./data

→ ./data/Filtered/S11B_R1.fastq

Etc pour chacun des fichiers

On a construit un Dossier Filtered dans le dossier ./data et qui contient des fichiers vides avec le nom S11B_R1.fastq etc

Pourquoi ?

C'est le dossier qui va accueillir les fichiers filtrés sur la qualité de séquence (Préparation)

To avoid problems …

1- **Set your working directory** to **Project** directory

2- put in your **Project** directory
--> data directory with your sequences
 --> and files

Project directory

- **data**
- MapfileFA.txt
- ggrare.R
- Fonctions_dada2.R
- silva_nr_v132_train_set.fa
- silva_species_assignment_v132.fa.gz

Because In the script

path="./data"

MAP="mapfileFA.txt"

source("./ggrare.R")
source("./Fonctions_dada2.R")

taxa <- assignTaxonomy(seqtab.nochim, "silva_nr_v132_train_set.fa", …)
taxa <- addSpecies(taxa, "silva_species_assignment_v132.fa.gz", …)

Be careful : pattern="_R1.fastq"

Understanding the Phyloseq Object & class level : All is interconnected

Final2

phyloseq-class experiment-level object
otu_table()   OTU Table: [229 taxa and 18 samples]
sample_data() Sample Data: [18 samples by 23 sample variables]
tax_table()   Taxonomy Table: [229 taxa by 7 taxonomic ranks]
phy_tree()   Phylogenetic Tree: [229 tips and 227 internal nodes]
refseq()   DNAStringSet: [229 reference sequences]

- **OTU/ASV Abundance Table = otu_table()**
- **Metadata (groups, treatments, localisation, sample names, env data) = sample_data()**
- **Taxonomy classification = tax_table()**
- **Phylogenetic tree (i.e Unifract distance) = phy_tree()**
- **DNA sequences of ASV/OTU = refseq()**

otu_table()

```
> otu_table(Final2_rar)
OTU Table:              [229 taxa and 18 samples]
                        taxa are columns
      ASV1 ASV2 ASV3 ASV4 ASV5 ASV6 ASV7 ASV8 ASV9 ASV10 ASV11 ASV12 ASV13 ASV14 ASV15 ASV16
S11B   109   24   79   78   89   24   66   31    0     0     0    32    14     0     0    46
S1B     83    0   28    0   59   58    0    0   69    51    67    18     0     0     0     0
S2B     51    0   26   15   44   65    0    0   68    31    47     0     2     0    85     0
S2S     93   93   74    9  127   40   23   68    0     0     0    43    57     0     0    26
      ASV17 ASV18 ASV19 ASV20 ASV21 ASV22 ASV23 ASV24 ASV25 ASV26 ASV27 ASV28 ASV29 ASV30 ASV31
S11B      0     0    69     0    24    31    22    25     0     0     0     0     0     3     0
S1B       0    29     0    47     0     0     0     0     0     0     0     0     0     0     0
S2B       0    46     0    60     0     0    24    20     0    32    42     0     6     0     0
S2S       0     0    34     0     0    10    24    28     0     0     0     0     0     2    19
```

tax_table()

```
>   tax_table(Final2_rar)
Taxonomy Table:     [229 taxa by 7 taxonomic ranks]:
      Kingdom     Phylum                     Class                     Order                Family
ASV1   "Bacteria" "Cyanobacteria"           "Oxyphotobacteria"        "Synechococcales"    "Cyanobiaceae"
ASV2   "Bacteria" "Proteobacteria"          "Gammaproteobacteria"     "Alteromonadales"    "Pseudoalteromonadaceae"
ASV3   "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "SAR11_clade"        "Clade_I"
ASV4   "Archaea"  "Euryarchaeota"           "Thermoplasmata"          "Marine_Group_II"    NA
ASV5   "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "SAR11_clade"        "Clade_I"
ASV6   "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "SAR11_clade"        "Clade_II"
ASV7   "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "Rhodospirillales"   "AEGEAN-169_marine_group"
ASV8   "Bacteria" "Proteobacteria"          "Gammaproteobacteria"     "Alteromonadales"    "Pseudoalteromonadaceae"
ASV9   "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "Rhodospirillales"   "AEGEAN-169_marine_group"
ASV10  "Bacteria" "Cyanobacteria"           "Oxyphotobacteria"        "Synechococcales"    "Cyanobiaceae"
ASV11  "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "SAR11_clade"        "Clade_I"
ASV12  "Bacteria" "Proteobacteria"          "Alphaproteobacteria"     "SAR11_clade"        "Clade_I"
ASV13  "Bacteria" "Proteobacteria"          "Gammaproteobacteria"     "Alteromonadales"    "Alteromonadaceae"
```

Sample_data()

```
> sample_data(Final2_rar)
Sample Data:        [18 samples by 22 sample variables]:
     X.SampleID Description Treatment Treatment1 groupe Pres PicoEuk Synec Prochloro NanoEuk Crypto SiOH4   NO2   NO3   NH4   PO4
S11B       S11B        S11B     South     South5B    GGF   35    5370 46830       580    6010   1690 3.324 0.083 0.756 0.467 0.115
S1B         S1B         S1B     North     North1B    NBF   52     660 32195     10675     955    115 1.813 0.256 0.889 0.324 0.132
S2B         S2B         S2B     North     North2B    NBF   59     890 25480     16595     670    395 2.592 0.105 1.125 0.328 0.067
S2S         S2S         S2S     North     North2S    NBS    0     890 25480     16595     670    395 3.381 0.231 0.706 0.450 0.109
S3B         S3B         S3B     North     North3B    NBF   74     835 13340     25115    1115    165 1.438 0.057 1.159 0.369 0.174
S3S         S3S         S3S     North     North3S    NBS    0     715 26725     16860     890    200 1.656 0.098 0.794 0.367 0.095
S4B         S4B         S4B     North     North4B    NBF   78    2220  3130     29835    2120    235 2.457 0.099 1.087 0.349 0.137
S4S         S4S         S4S     North     North4S    NBS   78    2220  3130     29835    2120    235 2.457 0.099 1.087 0.349 0.137
S5B         S5B         S5B     North     North5B    NBF   42    1620 55780     23795    2555   1355 2.028 0.103 1.135 0.216 0.128
S5S         S5S         S5S     North     North5S    NBS    0    1620 56555     22835    2560    945 2.669 0.136 0.785 0.267 0.114
S6B         S6B         S6B     South     South1B    GGF   13    2520 39050       705    3630   1295 2.206 0.249 0.768 0.629 0.236
S6S         S6S         S6S     South     South1S    GGS    0    2435 35890       915    3735   1300 3.004 0.251 0.927 0.653 0.266
S7B         S7B         S7B     South     South2B    GGF   26       0     0         0    4005   1600 3.016 0.157 0.895 0.491 0.176
S7S         S7S         S7S     South     South2S    GGS    0    4535 26545      1340    6585   1355 1.198 0.165 1.099 0.432 0.180
S8B         S8B         S8B     South     South3B    GGF   33       0     0         0    5910   1590 3.868 0.253 0.567 0.533 0.169
```

Phy_tree()

```
> phy_tree(Final2_rar)

Phylogenetic tree with 229 tips and 227 internal nodes.

Tip labels:
        ASV1, ASV2, ASV3, ASV4, ASV5, ASV6, ...

Unrooted; includes branch lengths.
```

refseq()

```
> refseq(Final2_rar)
  A DNAStringSet instance of length 229
      width seq                                                             names
  [1]   402 GGAATTTTCCGCAATGGGCGAAAGCCTGACGGAGCAACGCCGCGTGAGGG...TGCTGGGCCATCACTGACGCTCATGGACGAAAGCCAGGGGAGCGAAAGG ASV1
  [2]   425 GGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTGTG...ACCTGGGTCAACACTGACGCTCATGTACGAAAGCGTGGGGAGCAAACAG ASV2
  [3]   400 GGAATCTTGCACAATGGAGGAAACTCTGATGCAGCGATGCCGCGTGAGTG...TTCTGGATCATTACTGACACTGAGGAACGAAAGCATGGGTAGCGAAGAG ASV3
  [4]   383 CGAAAACTTGACAATGCGAGCAATCGTGATCAGGGAACTCCAAGTGGCTG...CACTAGAACGGATCCGACGGTCAGGGACGAAGCCTAGGGGCACGAACCG ASV4
  [5]   400 GGAATCTTGCACAATGGAGGAAACTCTGATGCAGCGATGCCGCGTGAGTG...TTCTGGATCATTACTGACACTGAGGAACGAAAGCATGGGTAGCGAAGAG ASV5
  ...   ... ...
[225]   425 GGAATATTGCGCAATGGGCGAAAGCCTGACGCAGCCATGCCGCGTGTGTG...ATCTGGGATACAACTGACGCTGAGGTACGAAAGCGTGGGGAGCAAACAG ASV226
[226]   418 CGAATCATTCACAATGGGCGAAAGCCTGATGGTGCAATGCCGCGTGGGGG...TACTGGACAATTACTGACGCTCAGGCACGAAAGCATGGGGAGCGAAAGG ASV227
[227]   404 GGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTGTG...ATACCGATGGCGAAGGCAGCCACCTGGGTCAACACTGACGCTCATGTAC ASV229
[228]   360 CGAAAACTTCACACTGCAGGAAACTGTGATGAGGGAACTCCAAGTGCGCG...ACCACCAGTTGCGAAGGCGCCACACTGGAACGGATCCGACGGTCAGGGA ASV230
[229]   400 GGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGAGTG...TTCTGGGCCATATCTGACGCTGAGGTTCGAAAGCGTGGGTAGCAAACAG ASV231
> ’|
```
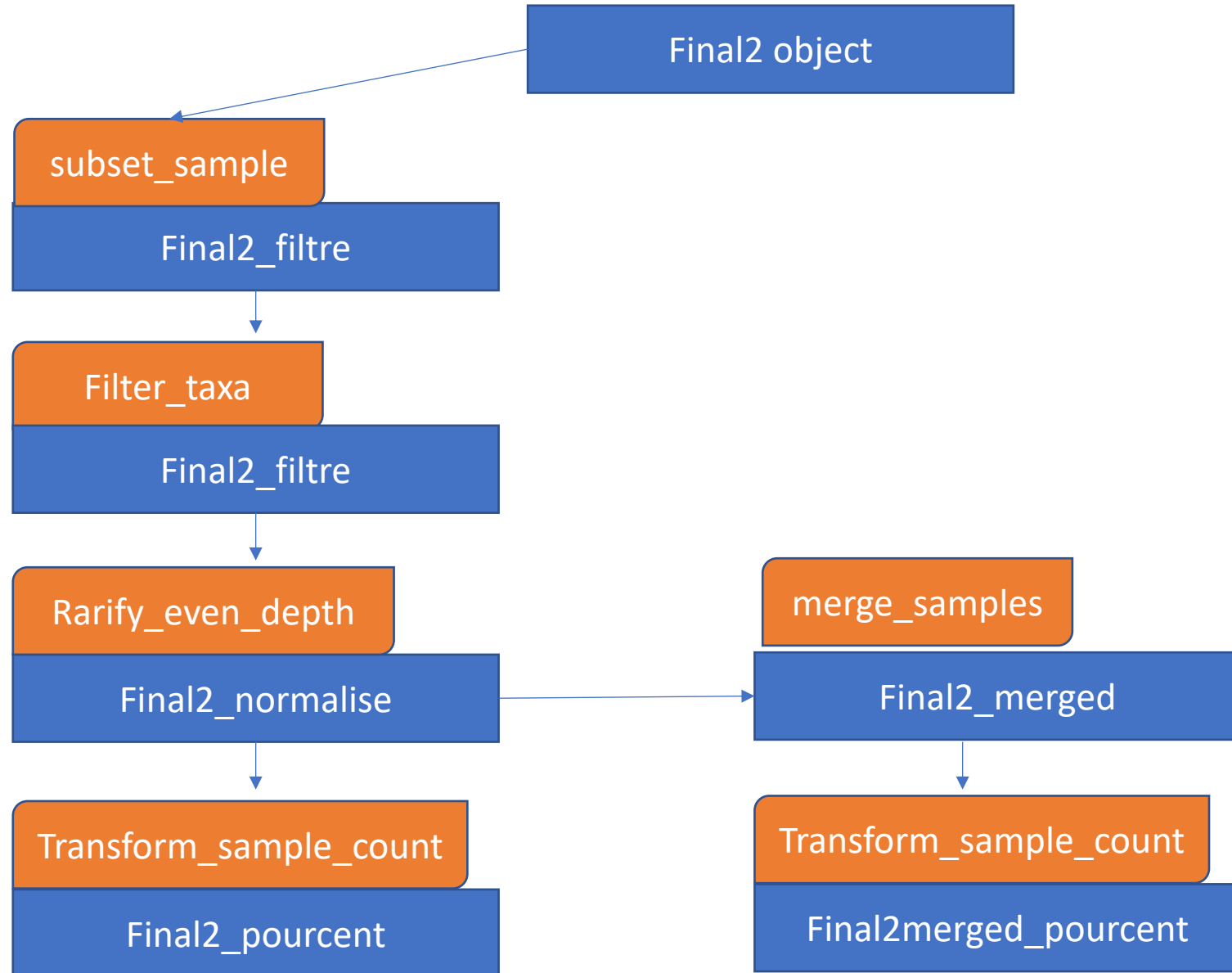
Applied functions onto an Object creates a new object (with same type of classes)
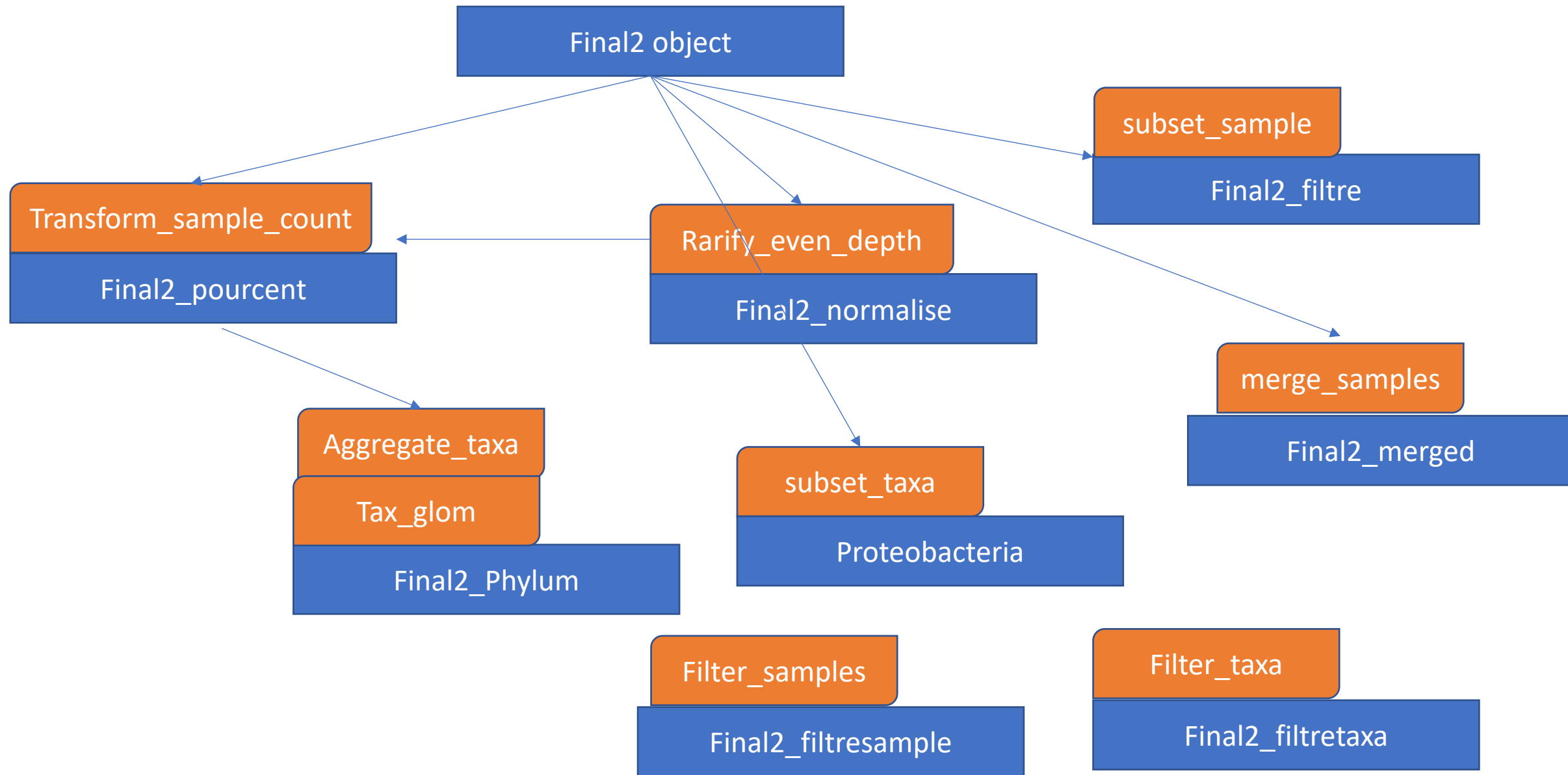
Applied functions of an Object creates a new object (with same features)

Save table with abundance & taxonomic assignment : combine two tables but ...

What you want :

```
         S9B                   S9S                 Kingdom     Phylum            Class
ASV2    "8.68686868686869"   "6.86868686868687"  "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV8    "6.26262626262626"   "8.08080808080808"  "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV32   "2.92929292929293"   "0"                 "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV58   "0.707070707070707"  "0"                 "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV77   "0.808080808080808"  "0"                 "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV120  "0.404040404040404"  "0"                 "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV175  "0.101010101010101"  "0"                 "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
ASV229  "0.101010101010101"  "0"                 "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"
```

What you have :

```
> otu_table(Myselection2)
OTU Table:          [8 taxa and 18 samples]
                    taxa are columns
        ASV2      ASV8      ASV32     ASV58     ASV77     ASV120    ASV175    ASV229
S11B  2.4242424 3.1313131 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S1B   0.0000000 0.0000000 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S2B   0.0000000 0.0000000 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S2S   9.3939394 6.8686869 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S3B   0.0000000 0.0000000 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S3S   0.9090909 0.0000000 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S4B   0.5050505 1.1111111 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S4S   0.4040404 0.8080808 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S5B   1.9191919 0.9090909 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
S5S   0.0000000 0.0000000 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
```

```
> tax_table(Myselection2)
Taxonomy Table:     [8 taxa by 7 taxonomic ranks]:
        Kingdom     Phylum            Class                  Order
ASV2    "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV8    "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV32   "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV58   "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV77   "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV120  "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV175  "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
ASV229  "Bacteria"  "Proteobacteria"  "Gammaproteobacteria"  "Alteromonadales"
```

Need to transpose otu_table :

```
> t(otu_table(Myselection2))
OTU Table:              [8 taxa and 18 samples]
                  taxa are rows
          S11B S1B S2B      S2S S3B       S3S       S4B       S4S       S5B
ASV2    2.424242   0   0 9.393939    0 0.9090909 0.5050505 0.4040404 1.9191919
ASV8    3.131313   0   0 6.868687    0 0.0000000 1.1111111 0.8080808 0.9090909
ASV32   0.000000   0   0 0.000000    0 0.0000000 0.0000000 0.0000000 0.0000000
ASV58   0.000000   0   0 0.000000    0 0.0000000 0.0000000 0.0000000 0.0000000
ASV77   0.000000   0   0 0.000000    0 0.0000000 0.0000000 0.0000000 0.0000000
ASV120  0.000000   0   0 0.000000    0 0.0000000 0.0000000 0.0000000 0.0000000
ASV175  0.000000   0   0 0.000000    0 0.0000000 0.0000000 0.0000000 0.0000000
ASV229  0.000000   0   0 0.000000    0 0.0000000 0.0000000 0.0000000 0.0000000
                   S7B         S7S       S8B       S8S       S9B       S9S
```

So :
tableau=cbind(t(otu_table(Myselection2)),tax_table(Myselection2))
→ is transpose otu_table & join the two tables
-> Always doing this for getting your abundance-taxonomic assignment at any level
-> next, write.table function (to save as file)

One exception

- When you use aggregate_taxa (from microbiome)

→Genus_glom=aggregate_taxa(pourcentS,"Genus")

You can directly save as file because:

```
> otu_table(Genus_glom)
OTU Table:              [36 taxa and 18 samples]
                        taxa are rows
                              S11B        S1B        S2B        S2S        S3B        S3S
Candidatus_Actinomarina    2.424242   8.6868687   8.5858586   2.828283   4.7474747   2.4242424
Sva0996_marine_group       0.000000   0.6060606   1.2121212   0.000000   0.8080808   1.1111111
Marinoscillum              0.000000   0.0000000   0.5050505   0.000000   0.0000000   0.4040404
NS2b_marine_group          0.000000   0.3030303   0.0000000   0.000000   0.2020202   0.0000000
NS4_marine_group           2.525253   0.0000000   2.0202020   2.828283   0.0000000   0.0000000
NS5_marine_group           0.000000   1.7171717   0.0000000   0.000000   1.7171717   0.7070707
Cyanobium_PCC-6307         0.000000   0.0000000   0.0000000   0.000000   0.0000000   0.0000000
Prochlorococcus_MIT9313    0.000000   5.1515152   3.1313131   0.000000   5.4545455   9.4949495
```

→ **write.table(otu_table(Genus_glom),"./myfile.txt")**

Understand complex command: Select the top 100 ASV
*my*taxa <- names(sort(taxa_sums(Final2_rar), TRUE)[1:100])

taxa_sums(Final2_rar) = Sum of abundance for each ASV

```
> taxa_sums(Final2_rar)
```

ASV1	ASV2	ASV3	ASV4	ASV5	ASV6	ASV7	ASV8	ASV9	ASV10	ASV11	ASV12	ASV13	ASV14
1614	949	905	914	770	667	663	611	461	409	455	374	292	289

ASV15	ASV16	ASV17	ASV18	ASV19	ASV20	ASV21	ASV22	ASV23	ASV24	ASV25	ASV26	ASV27	ASV28
245	218	233	201	241	209	195	221	191	183	192	186	181	160

ASV29 ASV30 ASV31 ASV32 ASV33 ASV34 ASV35 ASV36 ASV37 ASV38 ASV39 ASV40 ASV41 ASV42

sort (taxa_sums(Final2_rar)) : sort ASV abundance → increasing

```
> sort(taxa_sums(Final2_rar))
```

ASV213	ASV226	ASV227	ASV229	ASV231	ASV200	ASV209	ASV211	ASV219	ASV224	ASV225	ASV170	ASV179
1	1	1	1	1	2	2	2	2	2	2	3	3

ASV185	ASV198	ASV205	ASV210	ASV217	ASV218	ASV220	ASV230	ASV168	ASV173	ASV184	ASV187	ASV194
3	3	3	3	3	3	3	3	4	4	4	4	4

sort (taxa_sums(Final2_rar), TRUE): sort ASV abundance → DEcreasing

```
> sort(taxa_sums(Final2_rar),TRUE)
```

ASV1	ASV2	ASV4	ASV3	ASV5	ASV6	ASV7	ASV8	ASV9	ASV11	ASV10
1614	949	914	905	770	667	663	611	461	455	409

ASV15	ASV19	ASV17	ASV22	ASV16	ASV20	ASV18	ASV21	ASV25	ASV23	ASV26
245	241	233	221	218	209	201	195	192	191	186

names(sort (taxa_sums(Final2_rar), TRUE)): get the names

```
> names(sort(taxa_sums(Final2_rar),TRUE))
 [1] "ASV1"    "ASV2"    "ASV4"    "ASV3"    "ASV5"    "ASV6"    "ASV7"    "ASV8"
[11] "ASV10"   "ASV12"   "ASV13"   "ASV14"   "ASV15"   "ASV19"   "ASV17"   "ASV22"
[21] "ASV18"   "ASV21"   "ASV25"   "ASV23"   "ASV26"   "ASV24"   "ASV30"   "ASV27"
[31] "ASV31"   "ASV32"   "ASV35"   "ASV33"   "ASV37"   "ASV34"   "ASV39"   "ASV41"
[41] "ASV28"   "ASV43"   "ASV44"   "ASV43"   "ASV50"   "ASV46"   "ASV47"   "ASV52"
```

Mytop100= names(sort(taxa_sums(Final2_rar),TRUE)[1:100]): get only the top first 100 names

```
> names(sort(taxa_sums(Final2_rar),TRUE)[1:3])
[1] "ASV1" "ASV2" "ASV4"
```

Selection100= prune_taxa(mytop100, Final2_rar): KEEP the "mytop100" from Final2_rar and put it in the object selection100

```
> selection100
phyloseq-class experiment-level object
otu_table()   OTU Table:          [ 100 taxa and 18 samples ]
sample_data() Sample Data:        [ 18 samples by 22 sample variables ]
tax_table()   Taxonomy Table:     [ 100 taxa by 7 taxonomic ranks ]
phy_tree()    Phylogenetic Tree:  [ 100 tips and 98 internal nodes ]
refseq()      DNAStringSet:       [ 100 reference sequences ]
```