

# Microbiome

2024-06-12

## Installer et charger les packages

```
#install.packages("rmarkdown")
#install.packages("tidyverse")
#install.packages("vegan")
#install.packages("BiocManager")

#if(!requireNamespace("BiocManager")){
#  install.packages("BiocManager")
#}

#BiocManager::install("phyloseq")

library(tidyverse); library(vegan); library(phyloseq)

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.5.1      v purrr  1.0.1
## v tibble  3.2.1      v dplyr  1.1.4
## v tidyr   1.3.1      v stringr 1.5.0
## v readr   2.1.4      v forcats 1.0.0
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## Loading required package: permute
##
## Loading required package: lattice
##
## This is vegan 2.6-4
```

Charger l'objet au format phyloseq et nettoyer les noms

```
# Charger ANT3814_ps.rds dans Posit Cloud en cliquant sur "Upload" depuis l'onglet "Files" dans la fenê

ANT3814_ps <- readRDS("~/Dropbox/Montreal/TEACHING/ANT_3814_ComputationalCourse/GIT/PROBLEM_SETS/ANT3814

# Le code de ce bloc nettoie les noms de taxonomie désordonnés dans le jeu de données pour qu'ils soien

ANT3814_ps_rename <- ANT3814_ps
n_seqs <- seq(ntaxa(ANT3814_ps_rename))
len_n_seqs <- nchar(max(n_seqs))
```

```

taxa_names(ANT3814_ps_rename) <- paste("ASV", formatC(n_seqs,
                                                    width = len_n_seqs,
                                                    flag = "0"), sep = "_")

```

```

ANT3814_ps <- ANT3814_ps_rename
taxa_names(ANT3814_ps)

```

```

## [1] "ASV_001" "ASV_002" "ASV_003" "ASV_004" "ASV_005" "ASV_006" "ASV_007"
## [8] "ASV_008" "ASV_009" "ASV_010" "ASV_011" "ASV_012" "ASV_013" "ASV_014"
## [15] "ASV_015" "ASV_016" "ASV_017" "ASV_018" "ASV_019" "ASV_020" "ASV_021"
## [22] "ASV_022" "ASV_023" "ASV_024" "ASV_025" "ASV_026" "ASV_027" "ASV_028"
## [29] "ASV_029" "ASV_030" "ASV_031" "ASV_032" "ASV_033" "ASV_034" "ASV_035"
## [36] "ASV_036" "ASV_037" "ASV_038" "ASV_039" "ASV_040" "ASV_041" "ASV_042"
## [43] "ASV_043" "ASV_044" "ASV_045" "ASV_046" "ASV_047" "ASV_048" "ASV_049"
## [50] "ASV_050" "ASV_051" "ASV_052" "ASV_053" "ASV_054" "ASV_055" "ASV_056"
## [57] "ASV_057" "ASV_058" "ASV_059" "ASV_060" "ASV_061" "ASV_062" "ASV_063"
## [64] "ASV_064" "ASV_065" "ASV_066" "ASV_067" "ASV_068" "ASV_069" "ASV_070"
## [71] "ASV_071" "ASV_072" "ASV_073" "ASV_074" "ASV_075" "ASV_076" "ASV_077"
## [78] "ASV_078" "ASV_079" "ASV_080" "ASV_081" "ASV_082" "ASV_083" "ASV_084"
## [85] "ASV_085" "ASV_086" "ASV_087" "ASV_088" "ASV_089" "ASV_090" "ASV_091"
## [92] "ASV_092" "ASV_093" "ASV_094" "ASV_095" "ASV_096" "ASV_097" "ASV_098"
## [99] "ASV_099" "ASV_100" "ASV_101" "ASV_102" "ASV_103" "ASV_104" "ASV_105"
## [106] "ASV_106" "ASV_107" "ASV_108" "ASV_109" "ASV_110" "ASV_111" "ASV_112"
## [113] "ASV_113" "ASV_114" "ASV_115" "ASV_116" "ASV_117" "ASV_118" "ASV_119"
## [120] "ASV_120" "ASV_121" "ASV_122" "ASV_123" "ASV_124" "ASV_125" "ASV_126"
## [127] "ASV_127" "ASV_128" "ASV_129" "ASV_130" "ASV_131" "ASV_132" "ASV_133"
## [134] "ASV_134" "ASV_135" "ASV_136" "ASV_137" "ASV_138" "ASV_139" "ASV_140"
## [141] "ASV_141" "ASV_142" "ASV_143" "ASV_144" "ASV_145" "ASV_146" "ASV_147"
## [148] "ASV_148" "ASV_149" "ASV_150" "ASV_151" "ASV_152" "ASV_153" "ASV_154"
## [155] "ASV_155" "ASV_156" "ASV_157" "ASV_158" "ASV_159" "ASV_160" "ASV_161"
## [162] "ASV_162" "ASV_163" "ASV_164" "ASV_165" "ASV_166" "ASV_167" "ASV_168"
## [169] "ASV_169" "ASV_170" "ASV_171" "ASV_172" "ASV_173" "ASV_174" "ASV_175"
## [176] "ASV_176" "ASV_177" "ASV_178" "ASV_179" "ASV_180" "ASV_181" "ASV_182"
## [183] "ASV_183" "ASV_184" "ASV_185" "ASV_186" "ASV_187" "ASV_188" "ASV_189"
## [190] "ASV_190" "ASV_191" "ASV_192" "ASV_193" "ASV_194" "ASV_195" "ASV_196"
## [197] "ASV_197" "ASV_198" "ASV_199" "ASV_200" "ASV_201" "ASV_202" "ASV_203"
## [204] "ASV_204" "ASV_205" "ASV_206" "ASV_207" "ASV_208" "ASV_209" "ASV_210"
## [211] "ASV_211" "ASV_212" "ASV_213" "ASV_214" "ASV_215" "ASV_216" "ASV_217"
## [218] "ASV_218" "ASV_219" "ASV_220" "ASV_221" "ASV_222" "ASV_223" "ASV_224"
## [225] "ASV_225" "ASV_226" "ASV_227" "ASV_228" "ASV_229" "ASV_230" "ASV_231"
## [232] "ASV_232" "ASV_233" "ASV_234" "ASV_235" "ASV_236" "ASV_237" "ASV_238"
## [239] "ASV_239" "ASV_240" "ASV_241" "ASV_242" "ASV_243" "ASV_244" "ASV_245"
## [246] "ASV_246" "ASV_247" "ASV_248" "ASV_249" "ASV_250" "ASV_251" "ASV_252"
## [253] "ASV_253" "ASV_254" "ASV_255" "ASV_256" "ASV_257" "ASV_258" "ASV_259"
## [260] "ASV_260" "ASV_261" "ASV_262" "ASV_263" "ASV_264" "ASV_265" "ASV_266"
## [267] "ASV_267" "ASV_268" "ASV_269" "ASV_270" "ASV_271" "ASV_272" "ASV_273"
## [274] "ASV_274" "ASV_275" "ASV_276" "ASV_277" "ASV_278" "ASV_279" "ASV_280"
## [281] "ASV_281" "ASV_282" "ASV_283" "ASV_284" "ASV_285" "ASV_286" "ASV_287"
## [288] "ASV_288" "ASV_289" "ASV_290" "ASV_291" "ASV_292" "ASV_293" "ASV_294"
## [295] "ASV_295" "ASV_296" "ASV_297" "ASV_298" "ASV_299" "ASV_300" "ASV_301"
## [302] "ASV_302" "ASV_303" "ASV_304" "ASV_305" "ASV_306" "ASV_307" "ASV_308"
## [309] "ASV_309" "ASV_310" "ASV_311" "ASV_312" "ASV_313" "ASV_314" "ASV_315"

```

```

## [316] "ASV_316" "ASV_317" "ASV_318" "ASV_319" "ASV_320" "ASV_321" "ASV_322"
## [323] "ASV_323" "ASV_324" "ASV_325" "ASV_326" "ASV_327" "ASV_328" "ASV_329"
## [330] "ASV_330" "ASV_331" "ASV_332" "ASV_333" "ASV_334" "ASV_335" "ASV_336"
## [337] "ASV_337" "ASV_338" "ASV_339" "ASV_340" "ASV_341" "ASV_342" "ASV_343"
## [344] "ASV_344" "ASV_345" "ASV_346" "ASV_347" "ASV_348" "ASV_349" "ASV_350"
## [351] "ASV_351" "ASV_352" "ASV_353" "ASV_354" "ASV_355" "ASV_356" "ASV_357"
## [358] "ASV_358" "ASV_359" "ASV_360" "ASV_361" "ASV_362" "ASV_363" "ASV_364"
## [365] "ASV_365" "ASV_366" "ASV_367" "ASV_368" "ASV_369" "ASV_370" "ASV_371"
## [372] "ASV_372" "ASV_373" "ASV_374" "ASV_375" "ASV_376" "ASV_377" "ASV_378"
## [379] "ASV_379" "ASV_380" "ASV_381" "ASV_382" "ASV_383" "ASV_384" "ASV_385"
## [386] "ASV_386" "ASV_387" "ASV_388" "ASV_389" "ASV_390" "ASV_391" "ASV_392"
## [393] "ASV_393" "ASV_394" "ASV_395" "ASV_396" "ASV_397" "ASV_398" "ASV_399"
## [400] "ASV_400" "ASV_401" "ASV_402" "ASV_403" "ASV_404" "ASV_405" "ASV_406"
## [407] "ASV_407" "ASV_408" "ASV_409" "ASV_410" "ASV_411" "ASV_412" "ASV_413"
## [414] "ASV_414" "ASV_415" "ASV_416" "ASV_417" "ASV_418" "ASV_419" "ASV_420"
## [421] "ASV_421" "ASV_422" "ASV_423" "ASV_424" "ASV_425" "ASV_426" "ASV_427"
## [428] "ASV_428" "ASV_429" "ASV_430" "ASV_431" "ASV_432" "ASV_433" "ASV_434"
## [435] "ASV_435" "ASV_436" "ASV_437" "ASV_438" "ASV_439" "ASV_440" "ASV_441"
## [442] "ASV_442" "ASV_443" "ASV_444" "ASV_445" "ASV_446" "ASV_447" "ASV_448"
## [449] "ASV_449" "ASV_450" "ASV_451" "ASV_452" "ASV_453" "ASV_454" "ASV_455"
## [456] "ASV_456" "ASV_457" "ASV_458" "ASV_459" "ASV_460" "ASV_461" "ASV_462"
## [463] "ASV_463" "ASV_464" "ASV_465" "ASV_466" "ASV_467" "ASV_468" "ASV_469"
## [470] "ASV_470" "ASV_471" "ASV_472" "ASV_473" "ASV_474" "ASV_475" "ASV_476"
## [477] "ASV_477" "ASV_478" "ASV_479" "ASV_480" "ASV_481" "ASV_482" "ASV_483"
## [484] "ASV_484" "ASV_485" "ASV_486" "ASV_487" "ASV_488" "ASV_489" "ASV_490"
## [491] "ASV_491" "ASV_492" "ASV_493" "ASV_494" "ASV_495" "ASV_496" "ASV_497"
## [498] "ASV_498" "ASV_499" "ASV_500" "ASV_501" "ASV_502" "ASV_503" "ASV_504"
## [505] "ASV_505" "ASV_506" "ASV_507" "ASV_508" "ASV_509" "ASV_510" "ASV_511"
## [512] "ASV_512" "ASV_513" "ASV_514" "ASV_515" "ASV_516" "ASV_517" "ASV_518"
## [519] "ASV_519" "ASV_520" "ASV_521" "ASV_522" "ASV_523" "ASV_524" "ASV_525"
## [526] "ASV_526" "ASV_527" "ASV_528" "ASV_529" "ASV_530" "ASV_531" "ASV_532"
## [533] "ASV_533" "ASV_534" "ASV_535" "ASV_536" "ASV_537" "ASV_538" "ASV_539"
## [540] "ASV_540" "ASV_541" "ASV_542" "ASV_543" "ASV_544" "ASV_545" "ASV_546"
## [547] "ASV_547" "ASV_548" "ASV_549" "ASV_550" "ASV_551" "ASV_552" "ASV_553"
## [554] "ASV_554" "ASV_555" "ASV_556" "ASV_557" "ASV_558" "ASV_559" "ASV_560"
## [561] "ASV_561" "ASV_562" "ASV_563" "ASV_564" "ASV_565" "ASV_566" "ASV_567"
## [568] "ASV_568" "ASV_569" "ASV_570" "ASV_571" "ASV_572" "ASV_573" "ASV_574"
## [575] "ASV_575" "ASV_576" "ASV_577" "ASV_578" "ASV_579" "ASV_580" "ASV_581"
## [582] "ASV_582" "ASV_583" "ASV_584" "ASV_585" "ASV_586" "ASV_587" "ASV_588"
## [589] "ASV_589" "ASV_590" "ASV_591" "ASV_592" "ASV_593" "ASV_594" "ASV_595"
## [596] "ASV_596" "ASV_597" "ASV_598" "ASV_599" "ASV_600" "ASV_601" "ASV_602"
## [603] "ASV_603" "ASV_604" "ASV_605" "ASV_606" "ASV_607" "ASV_608" "ASV_609"
## [610] "ASV_610" "ASV_611" "ASV_612" "ASV_613" "ASV_614" "ASV_615" "ASV_616"
## [617] "ASV_617" "ASV_618" "ASV_619" "ASV_620" "ASV_621" "ASV_622" "ASV_623"
## [624] "ASV_624" "ASV_625" "ASV_626" "ASV_627" "ASV_628" "ASV_629" "ASV_630"
## [631] "ASV_631" "ASV_632" "ASV_633" "ASV_634" "ASV_635" "ASV_636" "ASV_637"
## [638] "ASV_638" "ASV_639" "ASV_640" "ASV_641" "ASV_642" "ASV_643" "ASV_644"
## [645] "ASV_645" "ASV_646" "ASV_647" "ASV_648" "ASV_649" "ASV_650" "ASV_651"
## [652] "ASV_652" "ASV_653" "ASV_654" "ASV_655" "ASV_656" "ASV_657" "ASV_658"
## [659] "ASV_659" "ASV_660" "ASV_661" "ASV_662" "ASV_663" "ASV_664" "ASV_665"
## [666] "ASV_666" "ASV_667" "ASV_668" "ASV_669" "ASV_670" "ASV_671" "ASV_672"
## [673] "ASV_673" "ASV_674" "ASV_675" "ASV_676" "ASV_677" "ASV_678" "ASV_679"
## [680] "ASV_680" "ASV_681" "ASV_682" "ASV_683" "ASV_684" "ASV_685" "ASV_686"
## [687] "ASV_687" "ASV_688" "ASV_689" "ASV_690" "ASV_691" "ASV_692" "ASV_693"

```

```
## [694] "ASV_694" "ASV_695" "ASV_696" "ASV_697" "ASV_698" "ASV_699" "ASV_700"
## [701] "ASV_701" "ASV_702" "ASV_703" "ASV_704" "ASV_705" "ASV_706" "ASV_707"
## [708] "ASV_708" "ASV_709" "ASV_710" "ASV_711" "ASV_712" "ASV_713" "ASV_714"
## [715] "ASV_715" "ASV_716" "ASV_717" "ASV_718" "ASV_719" "ASV_720" "ASV_721"
## [722] "ASV_722" "ASV_723" "ASV_724" "ASV_725" "ASV_726" "ASV_727" "ASV_728"
## [729] "ASV_729" "ASV_730" "ASV_731" "ASV_732" "ASV_733" "ASV_734" "ASV_735"
## [736] "ASV_736" "ASV_737" "ASV_738" "ASV_739" "ASV_740" "ASV_741" "ASV_742"
## [743] "ASV_743" "ASV_744" "ASV_745" "ASV_746" "ASV_747" "ASV_748" "ASV_749"
## [750] "ASV_750" "ASV_751" "ASV_752" "ASV_753" "ASV_754" "ASV_755" "ASV_756"
## [757] "ASV_757" "ASV_758" "ASV_759" "ASV_760" "ASV_761" "ASV_762" "ASV_763"
## [764] "ASV_764" "ASV_765" "ASV_766" "ASV_767" "ASV_768" "ASV_769" "ASV_770"
```

Regardons les données

```
# Voir les métadonnées
view(sample_data(ANT3814_ps))

# Voir les différents types de microbes dans le jeu de données
TAXONOMY_TABLE <-tax_table(ANT3814_ps)
view(TAXONOMY_TABLE)

# Voir le nombre de chaque type de microbe dans chaque échantillon
OTU_TABLE <-otu_table(ANT3814_ps)
view(OTU_TABLE)

# Voir le nombre de lectures de séquençage dans chaque échantillon
sort(sample_sums(ANT3814_ps))
```

```
## L3S242 L3S341 L3S360 L3S313 L3S294 L3S378 L2S309 L5S240 L5S155 L5S174 L5S203
##      897      962      969      1103      1225      1249      1419      1792      1800      1953      2112
## L5S104 L2S357 L5S222 L2S204 L2S222 L2S155 L2S382 L2S175 L2S240 L6S68 L1S257
##      2227      2373      2521      3158      3187      3932      4089      4386      5061      5980      6388
## L1S281 L6S20  L6S93  L1S8  L1S140 L1S105  L1S76 L1S208 L4S112 L1S57  L4S63
##      6615      6855      6953      7033      7163      7780      7867      8152      8340      8702      9744
## L4S137
##      9820
```

**QUESTION 1 :** Qu'est-ce que l'ASV\_228 ? Dans combien d'échantillons est-il présent ? Pouvons-nous identifier avec confiance son espèce ?

**QUESTION 2 :** Qu'est-ce que l'ASV\_221 ? Dans combien d'échantillons est-il présent ? Pouvons-nous identifier avec confiance son espèce ?

**QUESTION 3 :** Combien d'échantillons ont plus de 1000 lectures de séquençage ?

Créer une courbe de raréfaction du nombre d'ASV pour chaque échantillon. Cela montre combien de nouveaux ASV sont trouvés avec plus de séquençage de l'ADN.

```
sort(sample_sums(ANT3814_ps))
```

```
## L3S242 L3S341 L3S360 L3S313 L3S294 L3S378 L2S309 L5S240 L5S155 L5S174 L5S203
##      897      962      969      1103      1225      1249      1419      1792      1800      1953      2112
## L5S104 L2S357 L5S222 L2S204 L2S222 L2S155 L2S382 L2S175 L2S240 L6S68 L1S257
##      2227      2373      2521      3158      3187      3932      4089      4386      5061      5980      6388
## L1S281 L6S20 L6S93 L1S8 L1S140 L1S105 L1S76 L1S208 L4S112 L1S57 L4S63
##      6615      6855      6953      7033      7163      7780      7867      8152      8340      8702      9744
## L4S137
##      9820
```

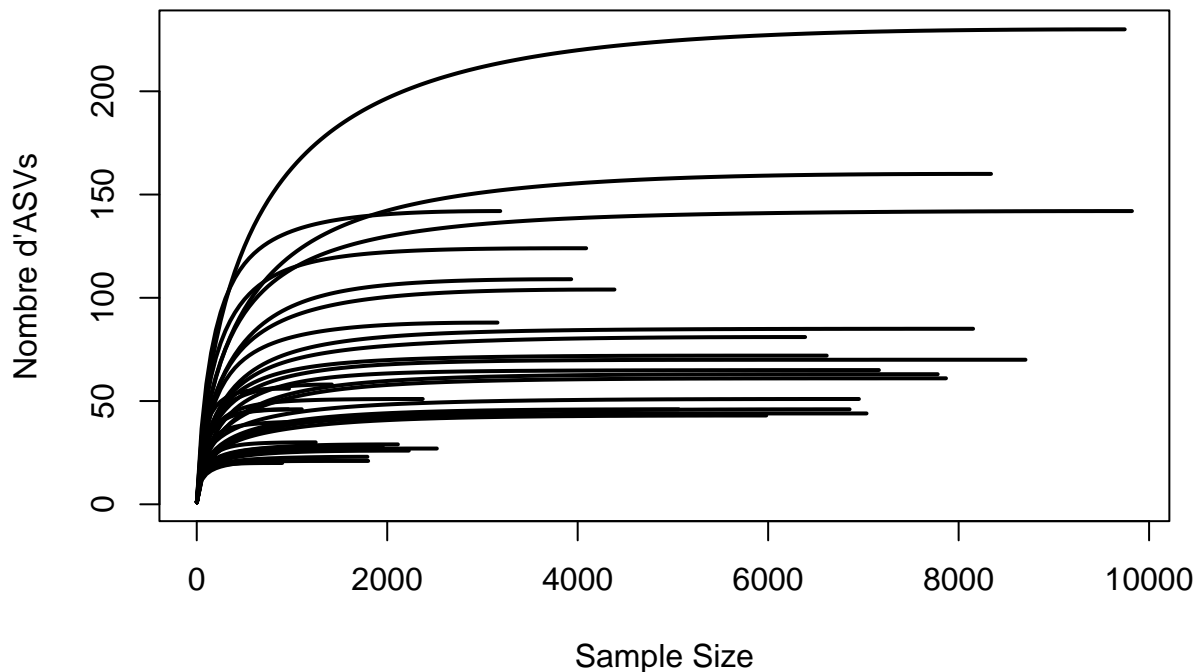
```
tab <- otu_table(ANT3814_ps)
class(tab) <- "matrix" # as.matrix() ne fera rien
```

```
## Warning in class(tab) <- "matrix": Setting class(x) to "matrix" sets attribute
## to NULL; result will no longer be an S4 object
```

```
## vous obtenez un avertissement ici, mais c'est ce dont nous avons besoin
tab <- t(tab) # transposer les observations en lignes
```

```
rarecurve(tab, step=50, lwd=2, ylab="Nombre d'ASVs", label=F, cex=0.5)
```

```
## Warning in rarecurve(tab, step = 50, lwd = 2, ylab = "Nombre d'ASVs", label =
## F, : most observed count data have counts 1, but smallest count is 2
```



**QUESTION 4 :** En général, autour de quel nombre de lectures de séquençage (la valeur sur l'axe des X) le nombre d'ASV se stabilise-t-il ?

Raréfions le jeu de données à 1103 lectures de séquençage (en excluant 3 échantillons). Cela sélectionnera aléatoirement 1103 lectures de séquençage de chaque échantillon pour permettre une comparaison uniforme de la diversité microbienne dans tous les échantillons.

```
set.seed(111) # garder le résultat reproductible
ps.rarefied = rarefy_even_depth(ANT3814_ps, rngseed=1, sample.size=1103, replace=F)
```

```
## 'set.seed(1)' was used to initialize repeatable random subsampling.
```

```
## Please record this for your records so others can reproduce.
```

```
## Try 'set.seed(1); .Random.seed' for the full vector
```

```
## ...
```

```
## 3 samples removed because they contained fewer reads than 'sample.size'.
```

```
## Up to first five removed samples are:
```

```
## L3S242L3S341L3S360
```

```
## ...
```

```
## 1040TUs were removed because they are no longer  
## present in any sample after random subsampling
```

```
## ...
```

```
ps.rarefied #La profondeur de séquençage des 31 échantillons restants est de 1103.
```

```
## phyloseq-class experiment-level object  
## otu_table() OTU Table: [ 666 taxa and 31 samples ]  
## sample_data() Sample Data: [ 31 samples by 8 sample variables ]  
## tax_table() Taxonomy Table: [ 666 taxa by 7 taxonomic ranks ]  
## phy_tree() Phylogenetic Tree: [ 666 tips and 664 internal nodes ]
```

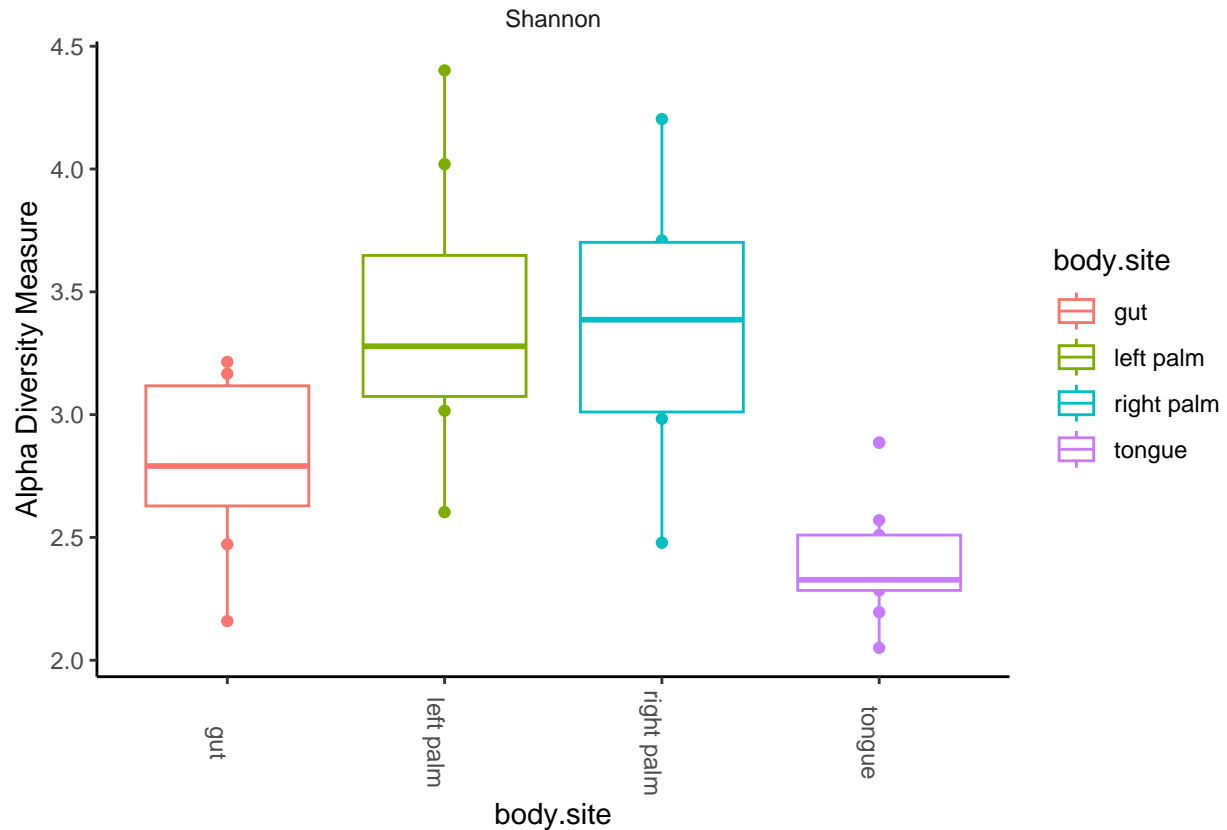
```
sort(sample_sums(ps.rarefied))
```

```
## L1S105 L1S140 L1S208 L1S257 L1S281 L1S57 L1S76 L1S8 L2S155 L2S175 L2S204  
## 1103 1103 1103 1103 1103 1103 1103 1103 1103 1103 1103  
## L2S222 L2S240 L2S309 L2S357 L2S382 L3S294 L3S313 L3S378 L4S112 L4S137 L4S63  
## 1103 1103 1103 1103 1103 1103 1103 1103 1103 1103 1103  
## L5S104 L5S155 L5S174 L5S203 L5S222 L5S240 L6S20 L6S68 L6S93  
## 1103 1103 1103 1103 1103 1103 1103 1103 1103
```

## QUESTION 5 : Combien d'échantillons ont MOINS de 1000 lectures de séquençage dans le jeu de données raréfié ?

La diversité alpha de Shannon tient compte à la fois de la richesse et de l'abondance des ASV dans chaque échantillon. Les boxplots ci-dessous montrent la valeur de la diversité alpha pour chaque échantillon regroupé par le site corporel où il a été collecté.

```
plot_richness(ps.rarefied, x="body.site", measures= "Shannon", color = 'body.site') +  
  geom_boxplot() +  
  theme_classic() +  
  theme(strip.background = element_blank(), axis.text.x.bottom = element_text(angle = -90))
```



## QUESTION 6 : Quel type de site corporel semble avoir la plus faible diversité alpha ?

Testons la différence significative de la diversité alpha de chaque site corporel en utilisant des tests de somme des rangs de Wilcoxon par paires. Comme nous effectuons plusieurs tests, nous devons corriger la valeur p en utilisant la méthode du taux de découverte faux (FDR).

```
rich = estimate_richness(ps.rarefied, measures = "Shannon")
wilcox.shannon <- pairwise.wilcox.test(rich$Shannon,
                                       sample_data(ps.rarefied)$body.site,
                                       p.adjust.method = "BH")

tab.shannon <- wilcox.shannon$p.value %>%
  as.data.frame() %>%
  tibble::rownames_to_column(var = "group1") %>%
  gather(key="group2", value="p.adj", -group1) %>%
  na.omit()
tab.shannon
```

```
##      group1      group2      p.adj
## 1 left palm      gut 0.074825175
## 2 right palm     gut 0.170229770
## 3 tongue         gut 0.071904566
```



```
## 5 right palm left palm 0.851814852
## 6 tongue left palm 0.000987248
## 9 tongue right palm 0.008391608
```

**QUESTION 7 :** En utilisant un seuil de valeur p ajustée de 0.05, quelles paires de sites d'échantillonnage ont des niveaux de diversité alpha significativement différents ?

Diversité bêta

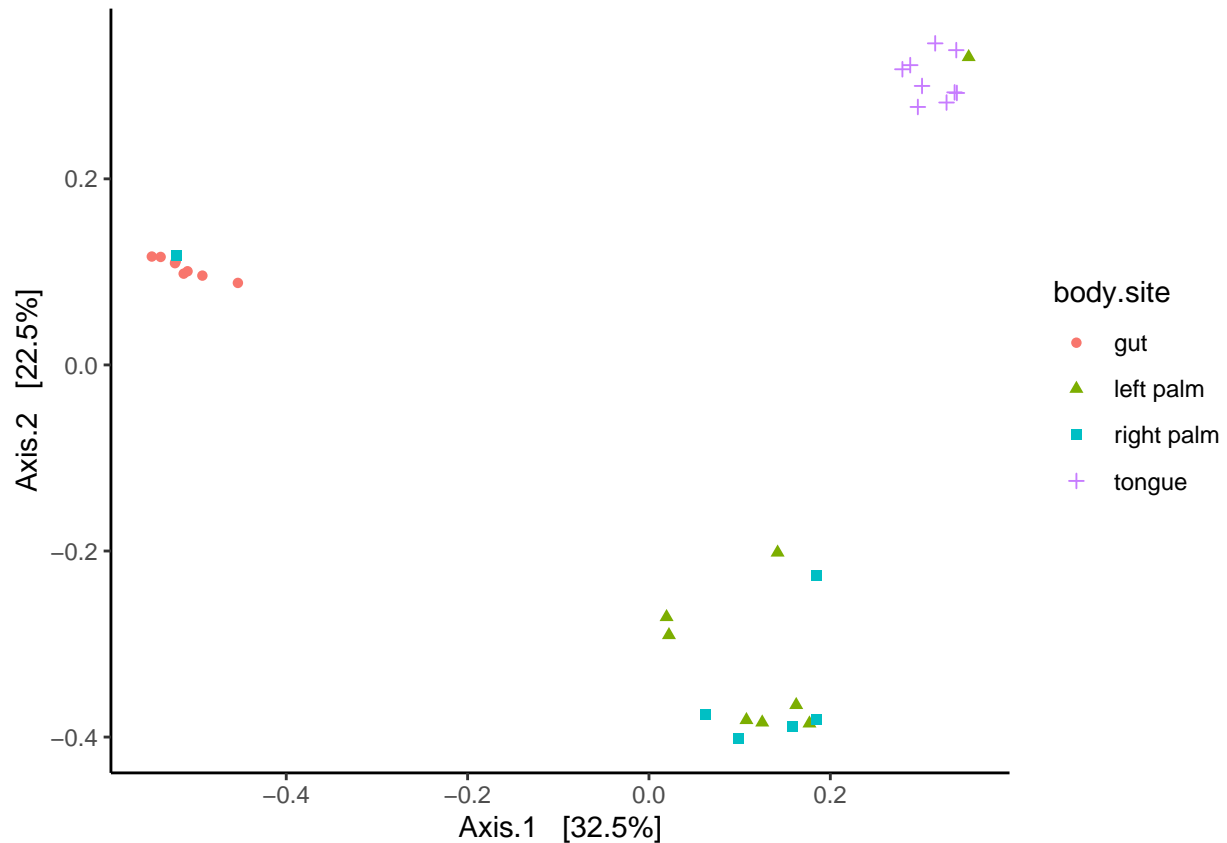
Les métriques de diversité bêta évaluent la dissimilarité entre les écosystèmes, vous indiquant dans quelle mesure une communauté est différente d'une autre. Voici quelques exemples de code utilisant les métriques de distance Bray Curtis et Jaccard binaire.

## PCoA et PERMANOVA/ADONIS

```
#La distance Bray Curtis est en usage commun, donc nous spécifions le package par phyloseq::distance

dist = phyloseq::distance(ps.rarefied, method="bray")
ordination = ordinate(ps.rarefied, method="PCoA", distance=dist)

plot_ordination(ps.rarefied, ordination, color="body.site", shape='body.site') +
  theme_classic() +
  theme(strip.background = element_blank())
```



**QUESTION 8 :** En regardant le graphique de diversité bêta, combien de clusters principaux voyez-vous ? En général, quels sont les quatre types d'échantillons les plus similaires ?

**QUESTION 9 :** Remarquez le point carré bleu dans le cluster rouge et le point triangulaire vert dans le cluster violet. Que pensez-vous qu'il s'est passé ici ?

PERMANOVA/ADONIS

```
# ADONIS

metadata <- data.frame(sample_data(ps.rarefied))
test.adonis <- adonis2(dist ~ body.site, data = metadata)
test.adonis

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
```

```
## adonis2(formula = dist ~ body.site, data = metadata)
##           Df SumOfSqs      R2      F Pr(>F)
## body.site  3   5.2139 0.46984 7.976  0.001 ***
## Residual  27   5.8833 0.53016
## Total     30  11.0972 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#### #PAIRWISE PERMANOVA

```
cbn <- combn(x=unique(metadata$body.site), m = 2)
p <- c()

for(i in 1:ncol(cbn)){
  ps.subs <- subset_samples(ps.rarefied, body.site %in% cbn[,i])
  metadata_sub <- data.frame(sample_data(ps.subs))
  permanova_pairwise <- adonis2(phyloseq::distance(ps.subs, method = "bray") ~ body.site,
                                data = metadata_sub)
  p <- c(p, permanova_pairwise$`Pr(>F)`[1])
}

p.adj <- p.adjust(p, method = "BH")
p.table <- cbind.data.frame(t(cbn), p=p, p.adj=p.adj)
p.table
```

```
##           1           2      p p.adj
## 1      gut left palm 0.001 0.0020
## 2      gut right palm 0.003 0.0036
## 3      gut  tongue 0.001 0.0020
## 4 left palm right palm 0.942 0.9420
## 5 left palm  tongue 0.001 0.0020
## 6 right palm  tongue 0.002 0.0030
```

**QUESTION 10 :** Un test “adonis” est une analyse de variance utilisant des matrices de distance. Il teste une différence significative dans les positions des clusters parmi les sites corporels. Les clusters des sites corporels sont-ils significativement différents les uns des autres ? Comment le savez-vous ?

Diagramme en barres de l'abondance : Voici des exemples utilisant une table non raréfiée pour visualiser toutes les étiquettes détectées.

**QUESTION 11 :** Dans le test PERMANOVA par paires, nous cherchons à voir quelles paires de sites corporels ont des communautés microbiennes significativement différentes. En regardant la colonne de la valeur p ajustée (p.adj), quelles communautés microbiomes de sites corporels sont significativement différentes les unes des autres ?

Abondances relatives des Phyla

```
# Normaliser le jeu de données en transformant les abondances des ASV en abondances relatives

ps.rel = transform_sample_counts(ANT3814_ps, function(x) x/sum(x)*100)

# Regrouper les ASV au niveau du phylum

glom <- tax_glom(ps.rel, taxrank = 'Phylum', NArm = FALSE)
ps.melt <- psmelt(glom)

ps.melt <- psmelt(glom)

# changer en caractère pour un niveau ajusté facile

ps.melt$Phylum <- as.character(ps.melt$Phylum)

ps.melt <- ps.melt %>%
  group_by(body.site, Phylum) %>%
  mutate(median=median(Abundance))

# sélectionner un groupe médian > 1

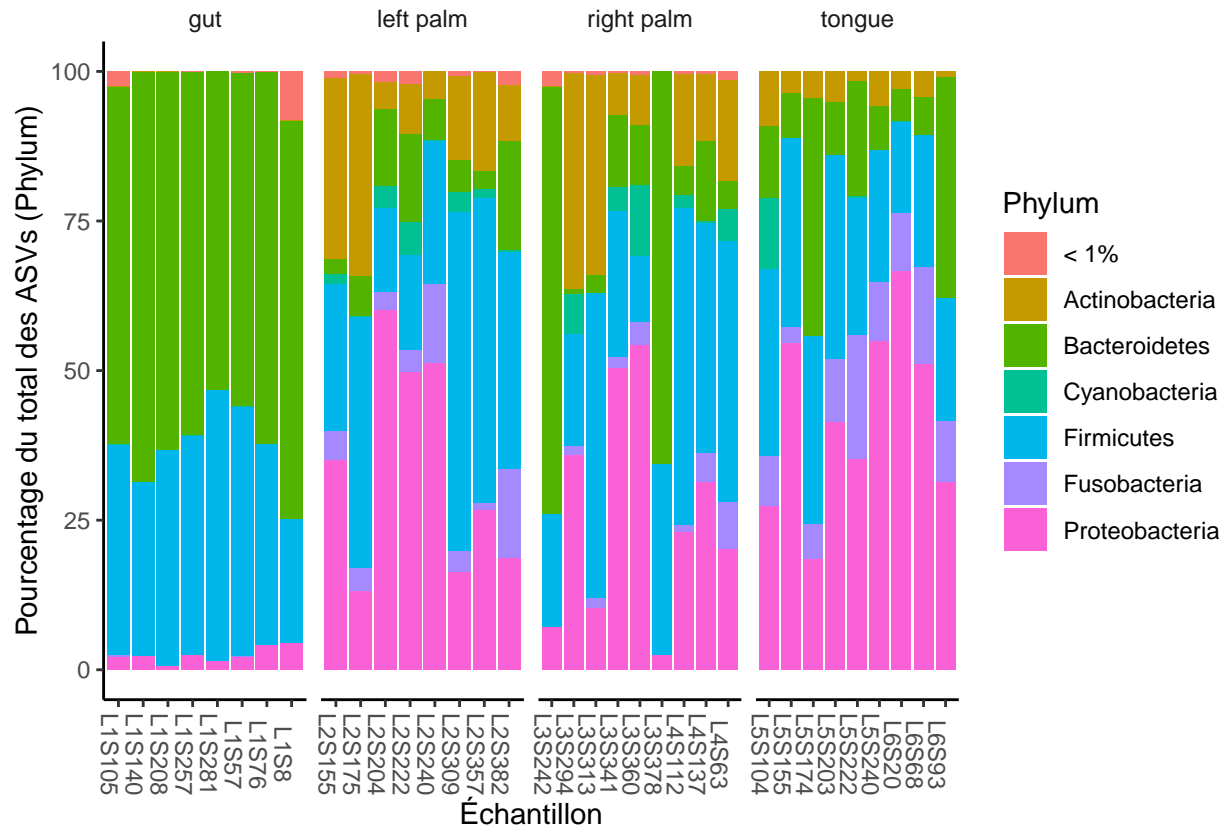
keep <- unique(ps.melt$Phylum[ps.melt$median > 1])
ps.melt$Phylum[!(ps.melt$Phylum %in% keep)] <- "< 1%"

# pour obtenir les mêmes lignes ensemble

ps.melt_sum <- ps.melt %>%
  group_by(Sample, body.site, Phylum) %>%
  summarise(Abundance=sum(Abundance))

## 'summarise()' has grouped output by 'Sample', 'body.site'. You can override
## using the '.groups' argument.

ggplot(ps.melt_sum, aes(x = Sample, y = Abundance, fill = Phylum)) +
  geom_bar(stat = "identity", aes(fill=Phylum)) +
  labs(x="Échantillon", y="Pourcentage du total des ASVs (Phylum)") +
  facet_wrap(~body.site, scales = "free_x", nrow=1) +
  theme_classic() +
  theme(strip.background = element_blank(),
        axis.text.x.bottom = element_text(angle = -90))
```



**QUESTION 12 :** Ici, nous pouvons voir l'abondance relative de chaque PHYLUM microbien. Chaque colonne est un échantillon différent, et l'abondance relative d'un phylum dans un échantillon est codée par couleur par la barre. En regardant les abondances des phyla, quel type d'échantillon est le plus différent ?

Abondances relatives des Genres

```
# Normaliser le jeu de données en transformant les abondances des ASV en abondances relatives

ps.rel = transform_sample_counts(ANT3814_ps, function(x) x/sum(x)*100)

# Regrouper les ASV au niveau du genre
glom <- tax_glom(ps.rel, taxrank = 'Genus', NArm = FALSE)
ps.melt <- psmelt(glom)

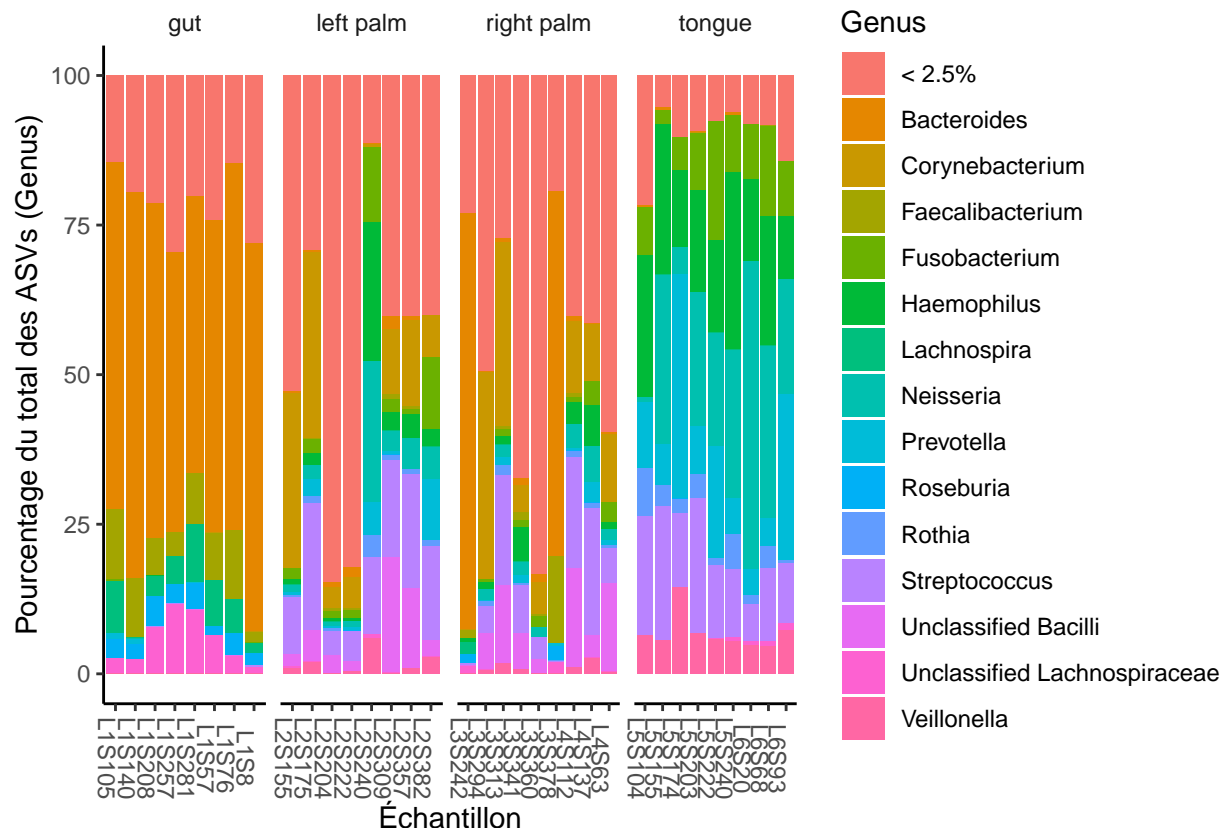
# changer en caractère pour un niveau ajusté facile
ps.melt$Genus <- as.character(ps.melt$Genus)

ps.melt <- ps.melt %>%
  group_by(body.site, Genus) %>%
  mutate(median=median(Abundance))
```

```
# sélectionner un groupe médian > 2.5
keep <- unique(ps.melt$Genus[ps.melt$median > 2.5])
ps.melt$Genus[!(ps.melt$Genus %in% keep)] <- "< 2.5%"
# pour obtenir les mêmes lignes ensemble
ps.melt_sum <- ps.melt %>%
  group_by(Sample,body.site,Genus) %>%
  summarise(Abundance=sum(Abundance))
```

## 'summarise()' has grouped output by 'Sample', 'body.site'. You can override  
## using the '.groups' argument.

```
ggplot(ps.melt_sum, aes(x = Sample, y = Abundance, fill = Genus)) +
  geom_bar(stat = "identity", aes(fill=Genus)) +
  labs(x="Échantillon", y="Pourcentage du total des ASVs (Genus)") +
  facet_wrap(~body.site, scales = "free_x", nrow=1) +
  theme_classic() +
  theme(legend.position = "right",
        strip.background = element_blank(),
        axis.text.x.bottom = element_text(angle = -90))
```



**QUESTION 13 :** Ici, nous pouvons voir l'abondance relative de chaque GENRE microbien. En comparaison avec ce que nous avons vu pour les phyla microbiens, comment la variation des communautés microbiennes des sites corporels diffère-t-elle au niveau des genres ? En général, quels sites corporels sont similaires et lesquels sont différents ?