

ANNEXE Fiche 12 BIS : Ecologie microbienne des sols : TP et TD associés

TP Extraction d'ADN pour quantification de la biomasse moléculaire microbienne (1/2 journée)

Protocole

- Extraction par lyse mécanique et chimique

- 1- Peser 1g de SOL LYOPHILISE sur une feuille de papier aluminium puis les transvaser dans le tube falcon 15 mL contenant les billes.
- 2- Ajouter 5 mL de tampon d'extraction. Composition : Tris-HCl (100 mM pH 8), EDTA (100 mM pH 8), NaCl (100mM), SDS (2 %), H₂O UP
- 3- A l'aide d'une seringue et d'un filtre adapté (0,2µm) OU d'une unité de filtration+filtre nitro-cellulose 0,2µm, filtrer la totalité du tampon
- 4- Agiter vigoureusement par retournement et tapotement pour décoller le sol et les billes du fond du tube.
- 5- Réaliser 3 runs de lyse mécanique dans l'agitateur FastPrep à une vitesse de 4 m/sec pendant 30 secondes, avec un temps d'attente de 5 minutes entre chaque run.
- 6- Après avoir vortexé les tubes, réaliser la lyse chimique en incubant les échantillons au bain-marie à 70°C pendant 30 min, avec agitation au vortex au bout de 15 et 30 minutes.
- 7- Transférer la totalité de votre tube dans un nouveau tube de 50 mL
- 8- Centrifuger à 7000 g pendant 5 min à 20°C minimum pour éviter la cristallisation du SDS.
- 9- Récupérer **TOUT** le surnageant à la pipette dans un nouveau tube de **5mL**, homogénéiser, puis prélever et placer 1 mL du lysat dans un tube de 1,5ml.

- Phase de purification de l'ADN du sol : Déprotéinisation, Précipitation et lavage et resuspension du culot d'ADN

Déprotéinisation

- 10- Ajouter 1/10ème du volume (100µl) d'acétate de potassium 3M pH5,5. Homogénéiser par retournement les tubes et incuber-les 10 min dans la glace.
Cette étape permet la précipitation des protéines, le froid favorise la formation de cristaux qui vont piéger les protéines.
- 11- Centrifuger à 14000g pendant 5 min à 4°C, récupérer le surnageant dans un tube de 2 mL.
Cette étape de centrifugation à haute vitesse permet d'obtenir un surnageant clair contenant l'ADN et débarrassé des particules en suspension.

Précipitation

- 12- Ajouter 900µl d'isopropanol à -20°C et agiter doucement les tubes par retournement.

L'ADN précipite

- 13- Placer les tubes à -20°C pendant 30 min minimum.

Lavage du culot d'ADN

- 14- Centrifuger à 13000 rpm pendant 30 min à 4°C.
- 15- Eliminer le surnageant avec précaution dans une poubelle prévue à cet effet à la pipette

- 16- Laver le culot d'ADN (sur la paroi du tube): ajouter 400 μL d'éthanol 70° à -20°C et centrifuger à 13000 rpm pendant 5 min à 4°C .
- 17- Eliminer l'alcool de la même façon qu'à l'étape 12. Les restants d'alcool sur les parois des tubes sont éliminés en plaçant les culots d'ADN (tubes ouverts) à l'étuve à 60°C pendant 10-15 min ou plus si nécessaire (jusqu'à ce qu'il n'y ait plus d'alcool).

Reprise du culot d'ADN

- 18- Resuspendre le culot d'ADN dans 200 μL d' H_2O UP : laisser les tubes à 4°C pour mieux dissoudre le culot pendant au moins une heure, puis homogénéiser la solution d'ADN à l'aide d'une micropipette en aspirant et refoulant la solution très délicatement (il faut éviter de faire des bulles). Un ADN dit « brut » est obtenu.

Quantification de la quantité d'ADN extraite

- 19- La concentration en ADN de chaque échantillon (10 μL ADN + 5 μL de bleu de dépôt/5 μL de smart ladder) est estimée sur gel d'agarose 1% en comparaison avec une gamme d'ADN de thymus de veau. La quantification de la fluorescence est faite après coloration du gel d'électrophorèse au Bromure d'Ethydium. Pour cela, une photo est prise et traitée dans un logiciel d'analyse d'image.

TD Compréhension des filtres environnementaux influençant l'abondance des communautés microbiennes

La séance se déroule en deux temps. Un premier temps d'analyse et d'interprétation des résultats acquis en travaux pratiques pour caractériser l'abondance totale (biomasse moléculaire microbienne) des communautés microbiennes des sols d'échantillons collectés sous des occupations de sols différentes. Dans un second temps, un jeu de données complémentaire intégrant des données de biomasse moléculaire microbienne, des données de diversité bactérienne, de propriétés physico-chimiques des sols et d'occupation des sols est proposé (données issues de la recherche dans ce cas). Ce second temps permet d'aller plus loin dans la compréhension des filtres environnementaux de l'abondance des communautés microbiennes et de la diversité des communautés bactériennes des sols.

NB : Dans les annexes, le logiciel R est mobilisé pour l'analyse des données et les lignes de code correspondant aux différentes étapes sont proposées. Ici, trois packages sont mobilisés : ggplot2 (représentation graphique), corr et car (statistiques, inflation de variance sur modèles linéaires). Le reste des fonctions sont des fonctions de base de R.

- **Partie 1 : Analyse des résultats acquis en travaux pratiques : Comparaison de la biomasse moléculaire microbienne des sols entre occupations du sol**

- **Estimation de la biomasse moléculaire microbienne**

Création du fichier de données

Le fichier construit par les étudiants ou fourni aux étudiants contient les informations suivantes :

- ↘ Code_ech : Code de l'échantillon (votre code)
- ↘ Occupation : Occupation du sol
- ↘ Vol_resuspension_ADN_μl : Volume d'eau dans lequel l'ADN a été resuspendu après précipitation à l'isopropanol et lavage à l'éthanol en μl.
- ↘ Vol_aliquot_μl : Volume de lysat prélevé pour précipiter l'ADN en μl.
- ↘ Vol_lysat_μl : Volume total de lysat produit à partir de la quantité de sol extraite en μl.
- ↘ Masse_sol_g : Masse de sol utilisée pour l'extraction en g.
- ↘ No_groupe: Numéro du binôme
- ↘ No_gel: Numéro du gel pour la quantification de l'ADN extrait
- ↘ Dilution: Facteur de dilution pour que la quantité d'ADN déposée sur gel soit dans les bornes de la gamme étalon.
- ↘ Vol_depot_μl : Volume d'ADN resuspendu déposé sur le gel d'électrophorèse en μl.
- ↘ Quantite_ADN : Quantité d'ADN dans la gamme étalon en ng d'ADN.
- ↘ Fluorescence : Mesure de fluorescence sur gel, sans unité.

Etape 1 : Construction de la courbe de calibration à partir de la gamme de thymus de veau :

Les étudiants réalisent la régression entre la concentration en ADN dans les échantillons de la gamme de thymus de veau et leur fluorescence. Elle est de la forme :

$$C_{ADN} = a \times fluorescence + b$$

C_{ADN} en $ng \cdot \mu l^{-1}$. On considère également son coefficient de détermination (r^2).

Dans R, on sélectionnera uniquement les points de la gamme de thymus et on écrira la ligne de code suivante :

```
Calibration <- lm(Quantite_ADN ~ Fluorescence )
```

Pour visualiser le r^2 :

```
r.square(Calibration)
```

Etape 2 : Estimation de la concentration en ADN pour chaque échantillon

Les étudiants appliquent l'équation de calibration pour chaque échantillon et on divise par le volume d'ADN déposé dans chaque puits (10µl). La concentration est alors estimée en ng/µl. Dans R, on aura la ligne de code suivante :

```
C_ADN= (Calibration$coefficients[2]* Fluorescence +Calibration$coefficients[1])/V_depot
```

#3 Etape 3 : Mise en œuvre de l'équation sous le logiciel R pour estimer la biomasse moléculaire microbienne de chaque échantillon en $\mu\text{g}_{\text{ADN}} \cdot \text{g}_{\text{sol}}^{-1}$

Ici, on écrit l'équation issue de l'analyse dimensionnelle. Dans R, elle s'écrit sous la forme :

```
DATA$BMM=DATA$C_ADN*Dilution*DATA$Vol_resuspension_ADN_µl*DATA$Vol_lysat_µl/DATA$Vol_aliquot_µl/1000/DATA$Masse_sol_g
```

- Evaluation de l'effet de l'occupation du sol sur la biomasse moléculaire microbienne

Etape 1. Comparaison graphique des résultats et interprétation

Le site <https://r-graph-gallery.com> s'appuie sur le package ggplot2 pour proposer une large suite de représentations graphiques. Ici, l'exemple est celui d'un boxplot (boîte à moustaches) dans lequel la médiane, le 1^{er} et 3^e quartiles ainsi que les écart-types sont représentés par occupation du sol.

Code R :

```
ggplot(DATA, aes(x=Occupation, y=BMM, fill=Occupation)) + geom_boxplot() +  
theme(legend.position="none") + xlab("Occupation") + ylab("Biomasse moléculaire microbienne  
(µg.gsol-1)")
```

Etape 2. Evaluation de l'effet de l'occupation du sol

```
kruskal.test(x = DATA$BMM, g = DATA$Occupation, p.adj=c("bonferroni"))  
TEST<-kruskal(y= DATA$BMM, trt = DATA$Occupation, p.adj=c("bonferroni"))  
dunn.test(DATA$BMM, DATA$Occupation, method="bonferroni")
```

- Partie 2 : Analyse d'un jeu de données recherche : Évaluation de l'effet de l'occupation du sol et de la physico-chimie du sol sur la biomasse moléculaire microbienne du sol et la diversité bactérienne des sols

Ici, l'objectif est d'analyser un jeu de données permettant d'évaluer l'importance relative des propriétés physico-chimiques et de l'occupation des sols sur la biomasse moléculaire microbienne et la diversité bactérienne des sols. La description de la diversité bactérienne s'appuie ici sur des méthodes basées sur l'extraction directe de l'ADN du sol et sa caractérisation par PCR et séquençage à haut débit. Après traitement bioinformatique, les informations de séquences relatives aux différents échantillons sont classées dans différentes unités taxonomiques opérationnelles (OTU) qui contiennent des séquences différant de moins de 5%. Le nombre d'OTU différentes dans un même échantillon constitue alors une mesure de sa diversité sous forme de richesse taxonomique.

Les étudiants ont à leur disposition un jeu de données contenant les informations relatives :

no_site : numéro de site

Carbone : Taux de carbone organique (g/kg)

C_N : rapport C sur N du sol

PH : ph du sol dans l'eau

Argile : teneur en argiles (g/kg)

Limon : teneur en limons (g/kg)

Sable : Teneur en sables (g/kg)

Occupation : occupation du sol (GC : Grandes cultures, PP : Prairies Permanentes, Fo : Forêts)

Biomasse_moléculaire : Biomasse moléculaire mesurée (µg/g)

Richesse_bacterienne : nombre de taxons bactériens

Etape 1 : Etude des corrélations entre variables explicatives

```
TEMP<-DATA[,colnames(DATA)%in%c("Carbone", "C_N", "PH", "Argile", "Limon", "Sable",  
"Biomasse_moleculaire", "Richesse_bacterienne")  
CORRELATIONS<-cor(TEMP)  
corrplot(CORRELATIONS,type = "upper", order = "hclust", tl.col = "black", tl.srt = 45)
```

Etape 2. Représentation graphique de la variable à expliquer en fonction d'une variable explicative

Variable quantitative (physico-chimie)

Variable_explicative_choisie : la variable en x (pH, Sable...)

Variable_à_expliquer_choisie : la variable en y (Biomasse moléculaire microbienne, Richesse taxonomique bactérienne)

```
ggplot(DATA, aes(x=Variable_explicative_choisie, y= Variable_à_expliquer_choisie)) +  
geom_point() + xlab("Nom de la Variable_explicative_choisie") +  
theme(legend.position="none") + ylab("Nom de la Variable_à_expliquer_choisie")
```

Variable qualitative (occupation du sol) : boxplots

Cette étape remobilise les mêmes outils que l'étape 1 de la partie 1. Il est possible de traiter simultanément les données de biomasse moléculaire microbienne et de richesse taxonomique bactérienne.

Etape 3 : Analyse statistique : Analyse de variance :

Ici, une approche de sélection de variables (stepwise) est mise en œuvre pour construire le modèle le plus parcimonieux ne tenant compte que des variables ayant un effet significatif sur la biomasse moléculaire microbienne ou diversité bactérienne. Les deux analyses doivent être réalisées séparément. L'analyse de la biomasse moléculaire microbienne ne doit pas inclure les données de richesse taxonomique bactérienne et inversement. Il serait faux de le faire au regard de la littérature.

- Biomasse moléculaire microbienne

```
TEMP<-DATA[,-c("Richesse_bacterienne")]  
model=step(lm(TEMP$Biomasse_moleculaire ~.,data= TEMP), direction="both")  
anova(model)  
vif(model)
```

- Richesse taxonomique bactérienne

```
TEMP<-DATA[,-c("Biomasse_moleculaire")]  
model=step(lm(TEMP$Richesse_bacterienne ~.,data= TEMP), direction="both")  
anova(model)  
vif(model)
```