Projet de Visualisation de Données Métagénomiques

Consignes

- Un projet Git par binôme
 - Soit un fichier de script Rmd
 - Soit un fichier ipynb
- Le script doit être portable (VM)
- Un répertoire contenant les figures

Poster en binôme

MADT – TP Métagénomique

OPEN @ ACCESS Freely available online



Microbial Biogeography of Public Restroom Surfaces

Gilberto E. Flores¹, Scott T. Bates¹, Dan Knights², Christian L. Lauber¹, Jesse Stombaugh³, Rob Knight^{3,4}, Noah Fierer^{1,5}*

1 Cooperative Institute for Research in Environmental Science, University of Colorado, Boulder, Colorado, United States of America, 2 Department of Computer Science, University of Colorado, Boulder, Colorado, Boulder, Colorado, United States of America, 3 Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado, United States of America, 4 Howard Hughes Medical Institute, University of Colorado, Boulder, Colorado, United States of America, 5 Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado, United States of America

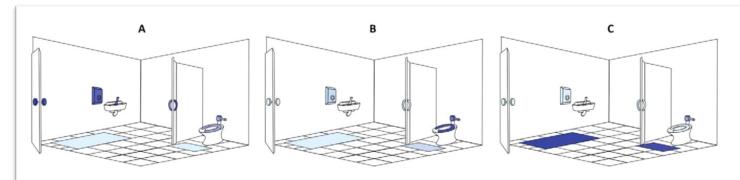


Figure 3. Cartoon illustrations of the relative abundance of discriminating taxa on public restroom surfaces. Light blue indicates low abundance while dark blue indicates high abundance of taxa. (A) Although skin-associated taxa (*Propionibacteriaceae*, *Corynebacteriaceae*, *Staphylococcaceae* and *Streptococcaceae*) were abundant on all surfaces, they were relatively more abundant on surfaces routinely touched with hands. (B) Gut-associated taxa (*Clostridiales*, *Clostridiales*, *Group XI*, *Ruminococcaceae*, *Lachnospiraceae*, *Prevotellaceae* and *Bacteroidaceae*) were most abundant on toilet surfaces. (C) Although soil-associated taxa (*Rhodobacteraceae*, *Rhizobiales*, *Microbacteriaceae* and *Nocardioidaceae*) were in low abundance on all restroom surfaces, they were relatively more abundant on the floor of the restrooms we surveyed. Figure not drawn to scale. doi:10.1371/journal.pone.0028132.g003

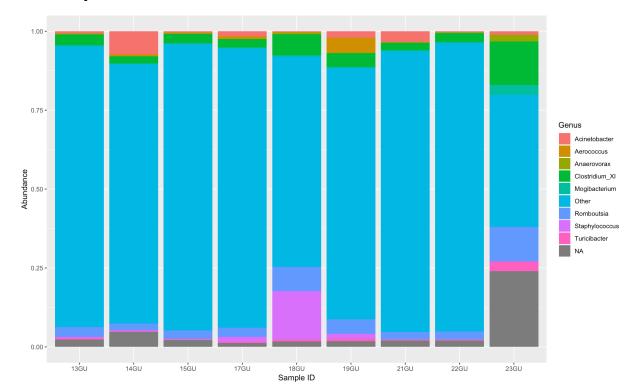
Données formatées

 Une table de comptage (abondance) de chaque OTU dans chaque échantillon avec son affiliation taxonomique

ОТИ	Affiliation taxonomique	Echantillon 1	Echantillon 2	Echantillon 3	•••
OTU_1	Domain1;Phylum1;Class1; Order1;Family1; Genus1	8278	2378	3892	
OTU_2	Domain1;Phylum1;Class1; Order1;Family1; Genus2	7287	1878	3029	
OTU_3	Domain1;Phylum1;Class1; Order1;Family2; Genus1	3231	892	1092	
OTU_4	Domain1;Phylum2;Class1; Order1;Family2; Genus1	2454	672	982	
OTU_5	Domain1;Phylum2;Class2; Order1;Family2; Genus4	1230	450	950	
OTU_6	Domain1;Phylum4;Class2; Order3;Family2; Genus2	472	230	765	
OTU_7	Domain1;Phylum2;Class2; Order2;Family1; Genus3	231	123	109	_

Etape 1 : Construire des visualisations simples

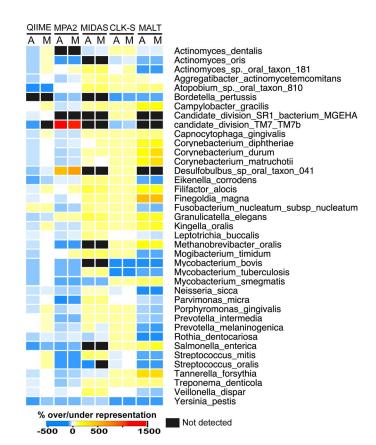
- On occulte l'information hiérarchique et on fixe un niveau taxonomique (e.g. Genre)
- Stacked barplots



Etape 1 : Construire des visualisations simples

 On occulte l'information hiérarchique et on fixe un niveau taxonomique (e.g. Genre)

Heatmap



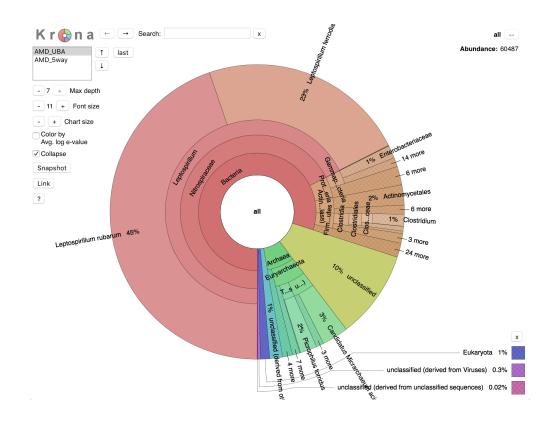
Etape 2 : Construire des visualisations plus complexes

- On exploite l'information hiérarchique
- Treemap



Etape 2 : Construire des visualisations plus complexes

- On exploite l'information hiérarchique
- Sunburst diagram (circular treemap)



Etape 3 : Interactivité

- Permettre à l'utilisateur de choisir l'abondance minimale (slider) pour représenter un OTU et afficher le stacked barplot et la heatmap
- Permettre à l'utilisateur de choisir le niveau taxonomique (select box) pour afficher le stacked barplot et la heatmap
- Permettre à l'utilisateur de rechercher un taxon (e.g. Escherichia) et afficher son abondance dans les différents échantillons sous forme de barplot
- Afficher des informations lors du survol de la souris
- Générer des tableaux de données téléchargeables

Etape 4 : Extraire de l'information

- Chercher une corrélation entre les variables
 - Type d'échantillon (sol, toilettes, mains)
 - Profil taxonomique global
 - Abondance d'un taxa

Analyses multivariées (e.g. ACP)

• Etc.