**Miércoles, 25 de agosto de 202112**

**MockCommunity**

* **La función para la MOCK community.** La creo en un archivo por separado. He revisado cómo hacer la documentación con reoxygen. Es sencillo. Pero, aparte, también se pueden incluir comentarios en cada línea de la función que sean explicativos. Esto lo tengo hecho, tocaría revisar si existen cosas a mejorar en esta función (por ejemplo, la condición del if es muy larga y poco leíble, me da la impresión de que doy muchos pasos en vano, etc…). **A día 26/08.** He mejorado la función con el if statement y hecho los comentarios en github. **Ahora toca hacer los comentarios de reoxygen y dejarla lista.**

**A día 30/08/2021.** He añadido los comentarios de reoxygen. **Habría que revisarlos, añadir comentarios en negrita,** etc. **Buscar la referencia del tipo de corte que hacemos con la MOCK (bibliografía) para añadirlo aquí**.

**A día 31/08.** He añdido las negritas, etc. Me queda ver lo de la referencia, pero no sé bien cual poner (preguntar a Hurry). Para incluir ejemplos de las funciones, etc, voy a utilizar data/, pero ATENCIÓN. Lo voy a hacer de manera preliminar, pero quizá tenga **que hacerlo de manera más adecuada cuando elija el dataset que voy a usar de ejemplo (yo creo que podría usar directamente el de DADA2 y así me ahorro problemas). Cuando lo haga en condiciones, añadiría las secuencias (para la vignette con DADA2 y todo) y también tablas ya formateadas, etc para los ejemplos. Cuando haga esto bien, también tengo que incluir de alguna manera de donde he obtenido esas tablas (mirar el capitulo 14, apartado external data)**

**También me he puesto con el testing de la función:**

* **Voy a testar que:**
* **~~Espero que el resultado sea un dataframe~~**
* **~~Espero que no contenga columnas con la MOCK community~~**
* **Debería crear errores dentro de la función para que:**

1. **Si no se encuentra columnas con la MOCK.**
2. **Si los números no son abundancias absolutas, sino relativas. (todo esto se debe chequear también en los tests)**
3. **Alomejor podría incluir un argumetno que diga: seleccionar el primero, si o no? Entonceds dependiendo de este argumento (normalmente FALSE), se muestran las preguntas. Esto habría que dcoumentarlo, pero sirve para aligerar los tests y para aligerar el workflow cuando se necesite.**

**Voy a dejar así hecha la parte de tests. Lo interesante sería dejar procesándose la base de datos de dada2 (u otra que busque de suelo con mock community)-esto dentro del publicworkflow- y, mientras, escribir los nuevos errores de la función MOCK y comenzar con otras funciones.**

**DATASET de ejemplo. El problema con el dataset de ejemplo de DADA2 es que no tiene para quitar los primers con cutadapt. Yo creo que lo mejor sería con datos nuestrosdatos publicados (por ejemplo el de plátano, lo que pasa que puedeque no se reproduzcan losresultados y meda mal rollo, además, no sé cuándo estará publicado, así que bueno, hablaré con hury para usar otro dataset). Me gustaría comparar DNA de rizosfera de Frantoio frente a DNA de su endosfera de frantoio, cogiendo los datos de linking belowground.**

**A día 06/09/2021, hemos decidio incluir datos de Frantoio (endosfera y rizosfera de CONTROL) del paper de Linkin beloewground. Es un coñazo autentico descargar las seucencias de SAR y me va a llevar más tiempo que directamente hacerlo en el laboratorio, además, no están las MOCKs. Por eso, he preferido continuar colocando las funciones en el paquete. HE introducido balanced y unbalanced anova con modificaciones y sus comentarios. También he incluido los paquetes delos que dependen. Como cosa por hacer de esto, quedaría poner bien los ejemplos (esto en general, cuando incluya los datos para el tutorial) e incluir referencias a los paquetes que utilizo. También, añadir tests a balanced y unbalanced (si es que es necesario) y errores, y los errores que comenté en MOCK (de hecho, hay que hacer algo porque check() no funciona por el read\_line)**

**EN TUKEY. HE arreglñado lo de la formula (tengo que terminar de ponerlo bine , haydos, la primera función). PEROOO . me he dado cuenta de otra forma de incluir la formula y quiero quizá poder incluir la formula con varios factores. Esto habría que hacerlo, pero lo dejo apuntado como segunda opción (primero centrarme en que funcione y en tnerlo listo y luego ya en incluir estas cosas)**

**A día 07/09.** He introducido la función de Tukey, pero me queda la duda si debería cambiar el AOV para unbalanced…. CREO QUE SÍIII