**mNGS: patógenos transmitidos por vectores que circulan en una comunidad**

🧪**Actividad:** Análisis de múltiples muestras

Se recolectaron 33 mosquitos, de los cuales 17 estaban alimentados con sangre. En este montaje se tiene 2 controles de agua que se procesaron desde la extracción de ARN hasta la secuenciación. ¡Veamos que encontramos implementando un heatmap!

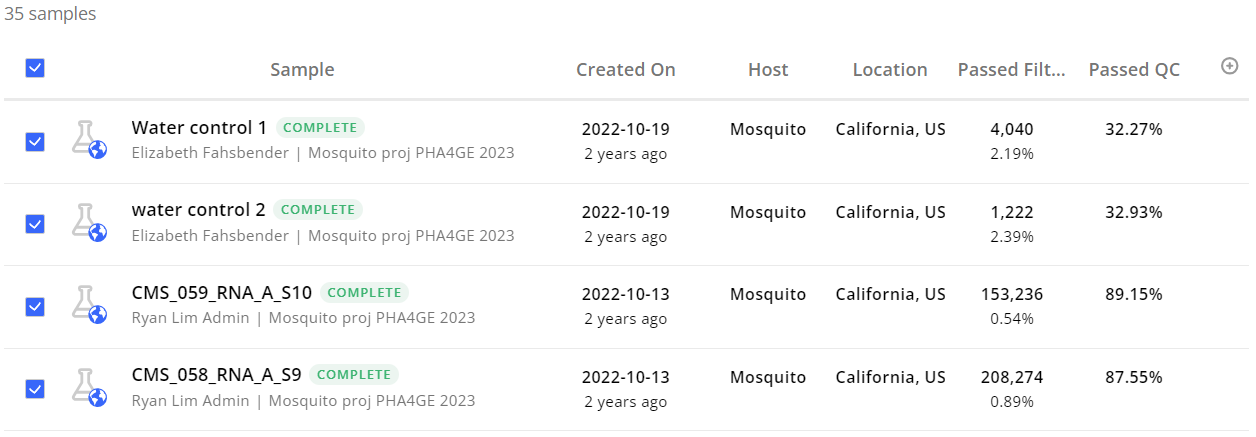
⏳Tiempo: ~ 30 minutos para la actividad y 10 minutos para la discusión

1. Inicien sesión en [**CZ ID**](http://czid.org)

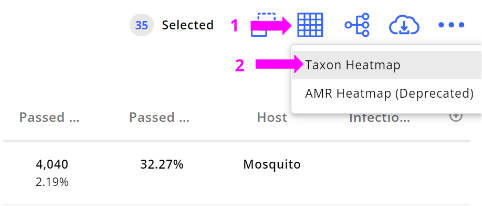
**Username:** correo

**Password:** contraseña

1. En la base de datos públicos busque y acceda al proyecto***Mosquito proj PHA4GE 2023***.
2. Seleccione todas las muestras desde la casilla superior.



1. Cree un heatmap haciendo clic en el ícono de la cuadrícula y seleccione **Taxon Heatmap**.



**Prepare el heatmap para el análisis.** Hay varias opciones para filtrar el heatmap y reducir su complejidad. Aquí te sugerimos algunos filtros iniciales para ayudarte a evaluar los datos a un nivel general.

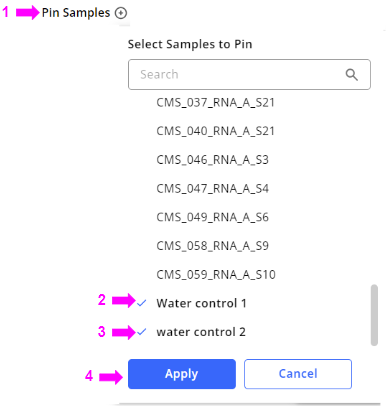
1. Añada en la sección de filtros *background model:*  **PHA4GE 2023** al heatmap.
2. Aplique los siguientes filtros para eliminar resultados no confiables:

**NT rPM ≥ 10** [Quiero ver solo los taxones presentes con al menos 10 lecturas por millón alineadas a la base de datos de nucleótidos].

**NR rPM ≥ 1** [Quiero ver solo los taxones presentes con al menos 1 lectura por millón alineada a la base de datos de proteínas].

**NT L ≥ 50** [Quiero ver solo los taxones presentes con al menos 50 pares de bases de homología alineadas a la base de datos de nucleótidos].

1. Cambie *taxon level* de Species a **Genus**.
2. Fije al lado del heatmap las controles de agua. Para esto seleccione en la parte superior del heatmap *Pin Samples* y marque los controles de agua.



Con esto, tendrá una visión general del panorama de patógenos en todas las muestras. Por defecto, las muestras (columnas) y los taxones (filas) se agrupan mediante agrupamiento jerárquico. Puedes desplazarte por heatmap para observar patrones como:

* Organismos que aparecen de manera constante en todas las muestras.
* Organismos atípicos que están presentes en un subconjunto de muestras.
* Grupos de organismos que a menudo aparecen juntos pero solo están presentes en un subconjunto de muestras.

1. **Observe sus controles de agua. ¿Hay algún taxón presente tanto en tu control de agua como en tu muestra?. A que se atribuye esto?.**

**Elimine la contaminación.** Asegúrese de haber aplicado *background model* correcto: **PHA4GE 2023**

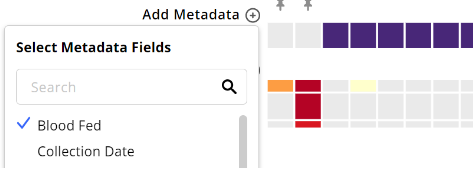
Ahora agregue un filtro de puntuación Z con la opción *Z-score filter*.

**NT Z score ≥ 1** [Solo quiero ver taxones presentes con al menos una desviación estándar mayor que el promedio de rPM alineado con la base de datos de nucleótidos en mi *background model* seleccionado]

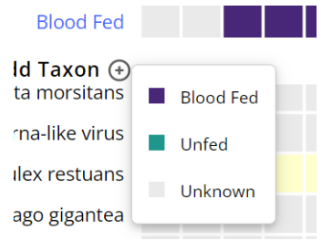
1. **Observe qué muestras están agrupadas. Elija 3 muestras que se agrupen. ¿Qué nota sobre ellas?**

Para observar más específicamente la pregunta de interés: *¿Qué patógenos transmitidos por vectores circulan en el ecosistema?* Es posible que desee agregar metadatos al heatmap para diferenciar qué taxones se encuentran en los animales alimentados con sangre y cuáles en los animales no alimentados.

Entre los nombres de las muestras y los nombres de los taxones en el lado izquierdo del heatmap, se encuentra la opción *add Metadata*. Seleccione el signo más y agregue metadatos adicionales seleccionando el encabezado de metadatos.



1. Seleccione **Blood fed y** desmarque **collection location**.



1. Luego, al pasar el cursor sobre el nombre de categoría, puede ver la leyenda.

Ordene las columnas de las muestras por la metadata haciendo clic en el encabezado de la metadata. Esto separará los dos grupos a lo largo del eje x del mapa de calor y le permitirá visualizar más fácilmente las diferencias en la abundancia de patógenos entre los dos grupos de interés.

1. Ordena por el campo de metadata haciendo clic en **blood-fed**.
2. **¿Hay algún patógeno presente con mayor abundancia/prevalencia en las muestras alimentadas con sangre en comparación con las que no lo están?**

A veces es útil separar el análisis de bacterias, eucariotas y virus. Podemos usar el filtro **Taxon Categories** para seleccionar solo los **eukaryotes**.

1. Puede ser útil eliminar los taxones que no contribuyen al análisis. **¿Qué taxón(es) en el heatmap considera podría (n) eliminarse?.** Elimine los géneros específicos de insectos de su visualización.
2. **¿Qué observa en el mapa de calor?**

Haga lo mismo para los virus y las bacterias. Seleccione una categoría a la vez y responda las siguientes preguntas:

1. **¿Cuáles son las diferencias que ve en el heatmap al observar cada categoría?**

**Heatmap de virus:**

**Heatmap de Bacteria:**

1. **¿Cómo se comparan con el heatmap de eucariotas?**