

Temario

Metabarcoding de comunidades de eucariontes

Fecha: 20-24 de enero del 2020

Lugar/Sede: La Paz, Baja California Sur

Horario: 9:00 am – 18:00 pm

Visión: Las técnicas de metabarcoding son un conjunto de nuevas herramientas genéticas para evaluar la biodiversidad presente en las comunidades naturales, desde un punto de vista cualitativo y/o cuantitativo. Sus aplicaciones potenciales incluyen (entre otras) análisis de la calidad del agua, evaluación de la biodiversidad de suelos, análisis tróficos o de contenidos digestivos, evaluación de recursos pesqueros, diagnóstico del estado de salud de instalaciones de acuicultura, detección temprana de la presencia de especies no autóctonas, estudios de patrones ecológicos globales, evaluación de impactos antropogénicos (biomonitoreo genético) o reconstrucción de comunidades paleoecológicas con el fin de estudiar cambios ecológicos ocurridos en el pasado.

Misión: Este curso/taller representa el primer esfuerzo en México de realizar un curso en este tema por lo que se pretende abordar tanto aspectos teóricos, como aspectos relacionados con el análisis bioinformático de datos derivados de los estudios de metabarcoding. Se pretende también divulgar las aplicaciones potenciales de la técnica, así como generar ideas sobre posibles proyectos a desarrollar.

Objetivo: En este curso/taller ofreceremos una descripción general de los procedimientos de metabarcoding de eucariontes, haciendo énfasis en la resolución práctica de problemas. Trabajaremos con ejemplos basados en conjuntos de datos reales. Al completar el taller, los participantes deberían estar en condiciones de (1) entender el potencial y las capacidades del metabarcoding, (2) ejecutar análisis completos mediante el uso de pipelines bioinformáticas para obtener inventarios de diversidad y datos ecológicamente interpretables a partir de un conjunto de datos de secuencias en bruto y (3) diseñar sus propios proyectos de metabarcoding, incluyendo el análisis y planificación del muestreo, tareas de laboratorio y la selección de los métodos de software más adecuados. Todos los materiales del curso (incluyendo copias de las presentaciones, ejercicios prácticos, archivos de datos y guiones de ejemplo) será proporcionado electrónicamente a los participantes.

Coordinación general: Dr. Fausto Valenzuela Quiñones, M.C. Tania Valdivia Carrillo

Instructores participantes/Institución: Dr. Owen S. Wangensteen (Universidad de Tromsø, Noruega) y Dr. Carlo Pecoraro (Physalia Courses)

Temario

Lunes 20

Expositor: Dr. Owen S. Wangensteen (Universidad de Tromsø, Noruega) y Dr. Carlo Pecoraro (Physalia Courses)

Tema:

Sesión 1. Introducción al metabarcoding. La pipeline de metabarcoding.

Subtema: En esta sesión explicaremos el formato del curso e introduciremos una serie de conceptos clave en las técnicas de metabarcoding, así como las diferentes plataformas de secuenciación de última generación (NGS) disponibles en la actualidad. Se expondrán algunos ejemplos de resultados que se pueden obtener a partir de proyectos de metabarcoding. Se repasarán las distintas etapas que forman parte de una pipeline de metabarcoding típica y se introducirán algunos conceptos bioinformáticos clave. En la sesión práctica, comprobaremos que la infraestructura informática para el resto del curso está funcionando y que todos los participantes tienen instalado el software necesario.

Conceptos básicos introducidos: Secuenciación de alto rendimiento (NGS/HTS), plataformas de secuenciación, multiplexado, librería para secuenciación, pipeline bioinformática, marcadores genéticos para metabarcoding, algoritmos de clustering y de denoising (reducción de ruido), Unidad Taxonómica Operativa Molecular (MOTU), asignación taxonómica.

Sesión 2. Técnicas de muestreo para ADN ambiental. Metabarcoding en el laboratorio. Buenas prácticas en el laboratorio de ADN ambiental y ADN antiguo. Extracción de ADN. PCR y preparación de librerías.

En esta sesión aprenderemos los fundamentos de los procedimientos de laboratorio necesarios para llevar a cabo las técnicas de metabarcoding siguiendo buenas prácticas de laboratorio. Se discutirán aspectos prácticos para todas las etapas clave en el laboratorio. Explicaremos las técnicas de recolección de muestras para distintos tipos de sustrato, el pretratamiento (filtrado o tamizado) de las muestras y los métodos de extracción de ADN. Se discutirán las medidas de seguridad adicionales para evitar contaminaciones a la hora de trabajar en laboratorios de ADN ambiental o ADN antiguo. Se introducirá la necesidad de utilizar controles negativos y blancos. Se estudiarán los protocolos de PCR para amplificar los marcadores de metabarcoding elegidos y la necesidad de realizar réplicas de PCR para métodos de baja reproducibilidad. Finalmente, se discutirán diversos procedimientos de preparación de librerías para secuenciación.

Conceptos básicos introducidos: Técnicas de muestreo, buenas prácticas de laboratorio, tipos de muestra para metabarcoding, ADN ambiental, ADN de comunidades, ADN extra-celular, laboratorio limpio, conservación del ADN, extracción del ADN, PCR, fuentes de contaminación, indexación de muestras para multiplexado, réplicas de PCR, etiquetas de muestra y etiquetas de librería, secuencias adaptadoras.

Temario

Martes 21

Expositor: Dr. Owen S. Wangensteen (Universidad de Tromsø, Noruega) y Dr. Carlo Pecoraro (Physalia Courses)

Tema:

Sesión 3. La pipeline de OBITools I. Control de calidad y filtrado de secuencias erróneas.

En esta sesión comenzaremos a trabajar con una pipeline de metabarcoding usando el paquete de utilidades OBITools, con un conjunto real de datos en crudo de ejemplo. En la sesión teórica, repasaremos los pasos necesarios para analizar los datos en crudo procedentes del secuenciador y los formatos en que éstos se encuentran. Se estudiarán los procedimientos básicos para el filtrado inicial de calidad de los datos de secuencia y algunos conceptos básicos como alineamiento de hebras apareadas, demultiplexado y de-replicación de secuencias únicas. En la sesión práctica realizaremos las primeras etapas de la pipeline (control de calidad, filtrado de secuencias con errores y eliminación de secuencias quiméricas) con los datos de ejemplo, y aprenderemos como se cambia el formato de los ficheros de datos para adaptarlo a los requerimientos de los distintos paquetes informáticos usados en metabarcoding.

Conceptos básicos introducidos: formatos fasta y fastq, formato OBITools-fasta, escala de calidad Phred, alineamiento de hebras apareadas, demultiplexado, filtrado de secuencias, quimeras, de-replicación, secuencias únicas, número de reads, singletons (secuencias aisladas).

Sesión 4. La pipeline de OBITools II. Métodos de denoising y de clustering,

En esta sesión continuaremos trabajando con la pipeline de OBITools. En la parte teórica se profundizará en los distintos métodos de clustering (agrupamiento de secuencias) y de denoising (reducción de ruido) apropiados para los distintos tipos de marcadores genéticos, así como los paquetes de software que implementan estos métodos. En la sesión práctica llevaremos a cabo el clustering de las secuencias del ejemplo usando la técnica de agregación paso a paso implementada en el software SWARM, así como el recuento de abundancias tras la etapa de clustering.

Conceptos básicos introducidos: denoising vs clustering, clustering mediante secuencias de referencia vs clustering de novo, clustering Bayesiano, clustering por agregación iterativa, umbral de identidad constante vs umbral de identidad variable, recuento de abundancias.

Temario

Miércoles 22

Expositor: Dr. Owen S. Wangensteen (Universidad de Tromsø, Noruega) y Dr. Carlo Pecoraro (Physalia Courses)

Tema:

Sesión 5. La pipeline de OBITools III. Métodos de asignación taxonómica.

En esta sesión repasaremos los distintos métodos para añadir información taxonómica a las secuencias representativas de cada MOTU, obtenidas durante la etapa previa de la pipeline. Estudiaremos las diferencias entre asignación por similaridad y asignación filogenética y aprenderemos cómo las bases de datos de secuencia interactúan con las bases de datos de taxonomía

para llevar a cabo la asignación filogenética. Estudiaremos la importancia de una buena base de datos de referencia para una correcta asignación taxonómica. Se repasarán las distintas bases de datos públicas disponibles para secuencias de referencia con información taxonómica, como Genbank, Silva o BOLD. En la sesión práctica, completaremos la pipeline de OBITools usando el software para asignación taxonómica Ecotag y una base de datos de referencia local.

Conceptos básicos introducidos: base de datos de referencia local, asignación taxonómica filogenética, best match (secuencia con mejor coincidencia), métodos para asignar rangos taxonómicos superiores, bases de datos de taxonomía, identificador taxonómico (taxid).

Sesión 6. Diseño de primers para metabarcoding

En esta sesión se expondrán los métodos para obtener una base de datos de referencia local específica para el marcador de metabarcoding concreto que se desea utilizar. Se usarán programas para técnicas de PCR in silico (ecoPCR), que permiten obtener automáticamente bases de datos de referencia para los fragmentos de ADN de interés a partir de una base de datos de ADN y de las secuencias de los primers utilizados. Así mismo, se estudiarán algunos procedimientos de diseño de primers para metabarcoding (ecoprimer) y algunos parámetros que permiten evaluar in silico la utilidad de los primers para detectar los organismos objetivo.

Conceptos básicos introducidos: PCR in silico, formato ecoPCR y formato ecoPCR-database, ecoprimer, rango taxonómico (selectividad) de los primers, primers específicos vs primers universales, sesgo introducido por los primers (primer bias), resolución taxonómica de los primers, especies sinónimas, detectabilidad de la especie.

Temario

Jueves 23

Expositor: Dr. Owen S. Wangensteen (Universidad de Tromsø, Noruega) y Dr. Carlo Pecoraro (Physalia Courses)

Tema:

Sesión 7. Refinado final de los resultados. Comparación de resultados de diferentes pipelines.

En esta sesión se expondrán métodos para el refinado y curado de los conjuntos de datos finales para obtener un conjunto definitivo refinado. Se aprenderán técnicas como el colapso taxonómico de MOTUs, la eliminación de errores de lectura del etiquetado (tag-switching) y la corrección de blancos, para eliminar posibles falsos positivos. También se introducirán los filtros de abundancia mínima para eliminar el ruido producido por la presencia de secuencias poco abundantes. Además, compararemos los resultados obtenidos a partir de diferentes pipelines y de diferentes variantes de las pipelines.

Conceptos básicos introducidos: tag-switching, colapso taxonómico, ruido producido por secuencias de escasa abundancia, filtros de abundancia mínima, corrección de blancos.

Sesión 8. Representación gráfica de resultados. Patrones de α - y β - diversidad.

En esta sesión introduciremos los métodos para representación de resultados a partir de los conjuntos finales refinados de datos de metabarcoding y discutiremos cómo interpretar estos resultados desde el punto de vista ecológico. Utilizaremos scripts de R para representar patrones de α - y β - diversidad, así como el uso de gráficos específicos para resultados de metabarcoding que integran la información taxonómica y la abundancia de reads (gráficos krona).

Conceptos básicos introducidos: α -diversidad, riqueza de MOTUs, rarefacción, abundancia de reads absoluta vs relativa, gráficos de violín, β -diversidad, PCA (análisis de componentes principales), RDA (análisis de redundancia), nMDS (escalado multidimensional no lineal), índices de presencia/ausencia vs índices de abundancia, distancias de Jaccard y de Bray-Curtis

Temario

Viernes 24

Expositor: Dr. Owen S. Wangensteen (Universidad de Tromsø, Noruega) y Dr. Carlo Pecoraro (Physalia Courses)

Tema:

Sesión 9. Diseño experimental. Personalización. Diseño de proyectos de metabarcoding.

En esta sesión discutiremos cómo diseñar adecuadamente un proyecto de metabarcoding para alcanzar con éxito el objetivo que se desea estudiar, en función de los recursos disponibles. Discutiremos cuáles son las mejores estrategias para obtener buenos resultados publicables, optimizando el tiempo, el coste, la cantidad de trabajo necesaria y los recursos informáticos disponibles. Se discutirá sobre la selección de la plataforma de secuenciación adecuada, el número de corridas de secuenciación necesarias y el número de muestras que se pueden multiplexar en cada corrida, en función de la profundidad de secuenciación deseada. También se discutirá sobre el número de réplicas ecológicas necesarias para responder a las cuestiones ecológicas planteadas y sobre la necesidad de réplicas de PCR en función del marcador utilizado y de los taxones objetivo. **Conceptos básicos introducidos:** nivel de multiplexado óptimo, réplicas ecológicas, réplicas técnicas (de PCR), profundidad de secuenciación, precio por muestra.

Sesión 10. Brainstorming y presentación de proyectos.

En esta última sesión, los participantes se dividirán en varios grupos y aplicarán los conocimientos obtenidos a lo largo del curso para diseñar un proyecto concreto de metabarcoding, que cada grupo elegirá libremente. Finalmente, cada grupo hará una presentación de 5 minutos exponiendo su proyecto, y todos los participantes discutirán en un debate abierto sobre los pros y los contras del plan experimental presentado. La idea es hacer esta sesión tan abierta y participativa como sea posible, para que todos los participantes expresen sus opiniones y puedan aplicar los conocimientos obtenidos a lo largo del curso de una manera práctica.

Conceptos básicos introducidos: Plan experimental, solicitudes de proyecto, evaluación y mejora mediante revisión por pares.