

Supercomputación Castilla y León

SCAYLE

Curso

Análisis genómico en plantas silvestres utilizando RADSeq: de los datos en bruto al archivo VCF



SCAYLE, enmarcado en el Proyecto DIGIS3 (DIGitalización Inteligente, Sostenible y coheSiva digital concebido como un a Centro de Innovación Digital - DIGIS3), organiza el Curso Análisis genómico en plantas silvestres utilizando RADSeq: de los datos en bruto al archivo VCF.

Dirección y coordinación académica

Supercomputación Castilla y León (SCAYLE), Área de Formación.

Objetivos

Este curso tiene como objetivo principal capacitar a los participantes en el uso de herramientas bioinformáticas para el procesamiento, análisis y visualización de datos obtenidos mediante la técnica RADSeq (Restriction site Associated DNA Sequencing). A lo largo de cinco jornadas, se abordará desde la introducción al entorno de trabajo (Linux y R) y el uso del centro de supercomputación SCAYLE, hasta el preprocesamiento de datos, la ejecución del workflow completo en Stacks 2, y la interpretación de resultados poblacionales a partir de archivos VCF. El curso combinará sesiones teóricas y prácticas intensivas que permitirán a los asistentes adquirir habilidades para aplicar flujos de trabajo reproducibles en el análisis de datos genómicos.

Destinatarios

El curso está dirigido a investigadores interesados en estudios genómicos, a profesionales del sector de las Ciencias Computacionales, Biología y/o Biotecnología relacionados con el diagnóstico genético y a Alumnos Universitarios (titulaciones técnicas del ámbito experimental y/o económico) de posgrado y técnicos de laboratorio interesados en el análisis genómico de poblaciones naturales, con énfasis en plantas silvestres, y en general, cualquier persona afín a la temática tanto en la dimensión de la investigación, como de la innovación y el desarrollo.

Número de Plazas 20

Fecha

Del **29 de septiembre al 3 de octubre de 2025.**

Duración 36 horas

Lugar

Edificio CRAI-TIC, Campus de Vegazana, Universidad de León.

Horario

- Lunes a jueves de 09:00 a 14:00 y de 15:00 a 17:30 horas.
- Viernes de 08:30 a 14:30 horas.

Idioma Español.

Importe matrícula

Matrícula de **225€**. El coste completo del curso es de 450€, al que se le ha aplicado un descuento del 50% gracias a la Cofinanciación al 50% de la Unión Europea y el Ministerio de Industria, Comercio y Turismo y la Fundación EOI del Gobierno de España, en el marco del Mecanismo de Recuperación y Resiliencia financiado por los fondos Next Generación de la Unión Europea. No obstante, los puntos de vista y las opiniones expresadas son únicamente los del autor o autores y no reflejan necesariamente las de la Unión Europea, el Ministerio de Industria, Comercio y Turismo o la Fundación EOI. Ni la Unión Europea ni la autoridad que concede la subvención pueden ser considerados responsables de los mismos.

Nº de cuenta: ES82 2103 4292 8600 3351 0978.

Tras la finalización del curso, SCAYLE ofrece acceso gratuito a recursos HPC a los inscritos durante el mes siguiente a la finalización del mismo, gracias a la financiación del Proyecto DIGIS3, vinculado a la cumplimentación y firma de la documentación completa asociada a los test DMA (Digital Maturity Assessment).

Inscripción

www.scayle.es/formacion

El plazo de Inscripción finalizará una semana antes del comienzo del curso.

Una vez realizada la inscripción, el alumno dispone de un plazo de 7 días para realizar el ingreso de la cuota del curso y formalizar la matrícula, en caso contrario la reserva será anulada.

La adjudicación de las plazas será por riguroso orden de formalización de la matrícula.

Profesorado

Cristina Esteban Blanco.

Postdoctoral Genomic and Bioinformatics Researcher, SCAYLE.

Juan Manuel Gorospe.

Dpto de Botánica de la Facultad de de Ciencias en Charles University (Praga, República Checa).

CONTENIDOS

29 de septiembre de 2025 - Introducción al entorno y preparación de datos. Cristina Esteban Blanco.

Recepción de Alumnos y Entrega de Documentación.

Inauguración del Curso.

09:00 Recepción de participantes y entrega de documentación.

09:00 – 09:30 Inauguración del curso y presentación de objetivos.

09:30 – 10:30 Introducción al entorno Linux (comandos básicos, estructura de ficheros).

11:00 – 11:30 Pausa.

11:30 – 13:30 Uso del Centro de Procesamiento de Datos de SCAYLE: acceso remoto, gestión de cuentas y colas de trabajo .

13:30 – 14:00 Introducción a R y primeros pasos en el análisis de datos (sintaxis básica, importación de tablas).

14:00 – 15:00 Pausa comida.

15:00 – 18:00 Prácticas en Caléndula: combinación de comandos Linux con scripts en R para visualizar datos de ejemplo .

30 de septiembre de 2025 - Preprocesamiento de datos RADseq. Juan Manuel Gorospe.

09:00 – 10:30 Fundamentos de datos RADseq: diseño experimental y formatos de archivo.

10:30 – 11:00 Pausa.

11:00 – 12:30 Preprocesamiento: demultiplexado, recorte de adaptadores y control de calidad con herramientas como FastQC y Trimmomatic.

12:30 – 14:00 Submuestreo de lecturas y estrategias para optimizar tiempo de cómputo.

14:00 – 15:00 Pausa comida.

15:00 – 18:00 Introducción al flujo de trabajo principal en Stacks2: ustacks, cstacks, sstacks y gstacks.

1 de octubre de 2025 - Workflow completo en Stacks2. Juan Manuel Gorospe.

09:00 – 10:30 Flujos de trabajo en Stacks2 con y sin referencia: ref_map y denovo_map.

10:30 – 11:00 Pausa.

10:50 – 12:30 Ejecución práctica del flujo de trabajo en Stacks2.

12:30 – 14:00 Optimización de parámetros en Stacks2: pruebas de m, M y n para maximizar loci informativos.

14:00 – 15:00 Pausa comida.

15:00 – 18:00 Uso del módulo "populations" de Stacks2: cálculo de estadísticas estándar en genética de poblaciones y formatos de datos de salida.

2 de octubre de 2025 - Del archivo VCF a los análisis preliminares. Juan Manuel Gorospe.

09:00 - 10:30 Introducción al formato VCF: estructura de registros y campos principales.

10:30 - 11:00 Pausa.

11:00 - 12:30 Cálculo de estadísticas de calidad de datos (depth, MAF, missingness), visualización en R y filtrado de variantes.

12:30 - 14:00 Generación del archivo VCF final y otros formatos de datos (fasta, structure, PLINK, Genepop).

14:00 - 15:00 Pausa comida.

15:00 - 18:00 Análisis preliminares: redes de NeighborNet (SplitsTree), SVDquartets y Structure.

3 de octubre de 2025 - Evaluación de resultados y conclusiones. Juan Manuel Gorospe.

09:00 - 10:30 Cálculo de estadísticas finales, PCA y clustering.

10:30 - 10:50 Pausa.

10:50 - 12:30 Interpretación de resultados de Structure y SVDquartets: gráficos y extracción de conclusiones.

12:30 - 13:00 Mesa redonda y discusión de casos de estudio en plantas silvestres.

13:00 - 13:15 Clausura del curso y entrega de certificados.

Organiza



Colaboran

