

**CURSOS
PRE-CONGRESO**

PRE-CONGRESS
COURSES

CRITERIOS PARA LA SELECCIÓN DE LA REVISTA ADECUADA PARA PUBLICAR

Camadro E.L. Editora General de *BAG, Journal of Basic and Applied Genetics*, Sociedad Argentina de Genética, Argentina. E- mail: elsacamadro@gmail.com

La publicación es el producto final de toda investigación. Cuando ese producto es validado mediante la revisión de pares, la comunidad científica se enriquece por la exposición a nuevas ideas, enfoques, y procedimientos de calidad. De este modo se construye y evoluciona el conocimiento en todas las disciplinas científicas. Un paso fundamental en el inicio de la redacción de un manuscrito es la selección de la revista adecuada para publicar, ya que de ello depende la difusión de los resultados de la investigación, la calidad de la revisión de pares, y la valoración del artículo publicado en evaluaciones para ingreso y promoción en la carrera del investigador, conformación de jurados docentes, de investigadores, de tesis, premios, y otros. La opinión de los pares sobre la calidad de un artículo o revista es fundamental. No obstante, desde hace algunos años las instituciones han adoptado índices bibliométricos desarrollados por las grandes editoriales para dotar a la evaluación científica de lo que se considera que sería un viso de objetividad. El objetivo de este curso es familiarizar a los participantes con los tipos de obras y editoriales, soportes de publicación, tipos de acceso a bases de datos, revistas y artículos científicos; principales bases bibliográficas, catálogos y directorios; principales métricas para medir el impacto de revistas científicas y limitaciones en el uso, a fin de capacitarlos para desenvolverse con éxito en el sistema científico-tecnológico y de educación superior.

BIOINFORMÁTICA. UNA INTRODUCCIÓN AL ANÁLISIS INFORMÁTICO DE DATOS BIOLÓGICOS MASIVOS

Abelleyro M. Instituto de Medicina Experimental (IMEX), ANM-CONICET, CABA, Argentina. E-mail: martinabelleyro@gmail.com

El gran avance alcanzado por la biología molecular, que permite en la actualidad conocer y comprender los mecanismos asociados a patologías hereditarias y no hereditarias, está íntimamente ligado al análisis bioinformático, para poder diseñar los experimentos con mayor alcance y precisión y para desentrañar la compleja relación que hay entre los resultados obtenidos en el laboratorio con los fenómenos biológico-médicos observados. Las nuevas tecnologías de secuenciación masiva han incrementado la información obtenida de muestras biológicas/clínicas, generando un reto asociado que significa poder entender y tamizar la inmensa cantidad de datos generados, para discriminar, con cierto grado de certeza, aquellos aplicables o transferibles a un posible diagnóstico o dato real de investigación de aquellos posibles artificios propios de la aplicación de métodos experimentales masivos. La presente propuesta tiene como objetivo introducir a los participantes en el análisis de datos de Exomas/Paneles, Genomas y RNAseq. Se propone realizar un recorrido que haga hincapié en conocimientos básicos para los que deseen comenzar a introducirse en el análisis bioinformático, abarcando desde los requerimientos de *hardware* y *software*, tipos y formatos de archivos de partida (e.g., Fastq, BED, etc), archivos de salida (e.g., SAM, BAM, etc.) y programas o algoritmos utilizados en estos tipos de análisis (e.g., STAR, GATK, etc.), cerrando con una discusión de los beneficios y desventajas de los diferentes resultados obtenidos (e.g., llamado de variantes puntuales, CNVs, expresión diferencial, etc.).

MÓDULO I: INTRODUCCIÓN A LAS TECNOLOGÍAS DE SECUENCIACIÓN MASIVA EN PARALELO: DESDE LA MUESTRA HASTA EL LLAMADO DE VARIANTES

Ledesma M., M. Abelleiro Instituto de Medicina Experimental (IMEX), ANM-CONICET, CABA, Argentina. E-mail: mmledesma88@gmail.com

El análisis y procesamiento de datos provenientes de los métodos de análisis masivos, tanto la secuenciación masiva paralela como las tecnologías de *Microarrays*, resultan hoy un desafío clave para el estudio de muestras en la biología molecular. Este módulo hará un recorrido que abarcará los diferentes métodos de análisis masivo de datos genómicos, tipo de muestras posibles para realizar el estudio, acercamientos actuales basados en secuenciación masiva paralela (MPS) y *Microarrays*, como así también los algoritmos de análisis que permite obtener los resultados utilizados en las áreas de investigación y clínica. En el recorrido de este módulo se presentarán los controles de calidad para el procesamiento correcto de las muestras, los requerimientos de *hardware* y *software* para poder realizar el procesamiento de datos de MPS y se detallará la aplicación de los algoritmos bioinformáticos comúnmente utilizados. Se realizará un recorrido por los programas (e.g., BWA, Bowtie, STAR, etc) y los archivos (e.g., Fastq, BED, etc) necesarios para el procesamiento resultados crudos provenientes del secuenciador, se describirán los archivos intermedios obtenidos (e.g., SAM, BAM) y se presentara el pipeline de GATK (*gold estándar*) para el llamado de variantes. Finalmente se presentarán alguno de los visualizadores genómicos conocidos en bioinformática y se discutirán las dificultades que presentan los llamados de variantes estructurales según el origen de los datos genómicos procesados.

MÓDULO II: RNA-SEQ: ANÁLISIS DE EXPRESIÓN DIFERENCIAL Y BÚSQUEDA DE PATRONES TRANSCRIPCIONALES

Lincango Yupanki M. Academia Nacional de Medicina, CABA, Argentina. E-mail: lincango_marco@hotmail.com

El estudio del RNA es esencial para el entendimiento funcional del genoma y de la desregulación de procesos moleculares involucrados en oncopatologías y otras enfermedades. La técnica de RNA-seq permite examinar la cantidad y las secuencias de RNA mediante tecnologías de secuenciación de próxima generación (NGS, por sus siglas en inglés). Aunque esta tecnología puede ser usada con diferentes enfoques metodológicos y/o estudios biológicos, las aplicaciones más comunes son el análisis de expresión génica diferencial (DGE, por sus siglas en inglés) y la evaluación de perfiles transcripcionales. Las estrategias de diseño experimental de RNA-seq y análisis de datos pueden diferir dependiendo de la interrogante biológica planteada. Sin embargo, el objetivo principal de la estrategia elegida es maximizar el valor de los datos obtenidos. Los experimentos de RNA-seq generan un gran volumen de datos y su análisis incluye distintos pasos y una variedad de herramientas a elección. En este módulo se mostrarán aspectos básicos del flujo de trabajo bioinformático involucrado en el análisis de datos de RNA-seq, tal como: adquisición de datos públicamente disponibles, control de calidad y preprocesamiento de datos crudos, análisis de expresión diferencial, identificación de patrones transcripcionales y visualización de datos. Se espera que este módulo sirva como una guía introductoria para el diseño y análisis de datos RNA-seq.

MÓDULO III: ANÁLISIS DE PANELES DE MICROARRAYS. PRESENTACIÓN DEL SOFTWARE GALAXY

Ledesma M., M. Abelleiro. Instituto de Medicina Experimental (IMEX), ANM-CONICET, CABA, Argentina. E-mail: mmledesma88@gmail.com

El análisis de variantes pequeñas puede ser abordado eficientemente utilizando paneles de *Microarrays*; este tipo de herramientas son utilizadas para analizar la presencia de variantes genéticas en una muestra que puede estar o no asociada a una patología específica. Para el análisis de estos resultados se utilizan diversas herramientas

bioinformáticas de acceso libre. En este módulo se profundizará en los aspectos del procesamiento bioinformático de los resultados crudos de *Microarrays*, en particular en el llamado y anotado de variantes (GATK) en *bash*. A modo de ejemplo se expondrá un *pipeline* de análisis estadístico utilizando el *software* R, que tiene como objetivo la búsqueda de biomarcadores diferenciales, es decir, genes o variantes que muestren diferencias significativas en su expresión o frecuencia entre diferentes grupos de muestras. Se aplicarán estrategias de aprendizaje automatizado (*machine learning*) en el análisis de los datos del *Microarray*. Esto permitirá identificar patrones y características que distingan los grupos de muestras y potencialmente descubrir biomarcadores relevantes en el contexto del estudio. Finalmente se presentará la herramienta Galaxy (URL: <https://usegalaxy.org/>), una plataforma bioinformática que proporciona una interfaz web intuitiva y fácil de usar para el análisis de datos genómicos. Galaxy puede ser utilizado para llevar a cabo diversas etapas de los procesos presentados en el curso, desde el pre-procesamiento de datos crudos hasta el análisis de los datos procesados.
