OMICSPACE

Espacio de datos –ómicos ético y seguro

BDIAB

IIS La Fe

Plataforma de Big Data, IA, Bioestadística & Bioinformática

CATÁLOGO DE SERVICIOS



Participantes:

IIS La Fe · Fundación Pública Andaluza Progreso y Salud · Hospital del Mar Barcelona · Hospital Universitario La Paz · Hospital Universitario Ramón y Cajal · iacs · Medicina Xenómica · HARMONY Alliance Foundation

OMICSPACE

Espacio de datos –ómicos ético y seguro

CATÁLOGO DE SERVICIOS





BIOINFORMÁTICA



IIS La Fe

Plataforma de Big Data, IA, Bioestadística & Bioinformática

Participantes:

IIS La Fe · Hospital del Mar Barcelona · Hospital Universitario La Paz · Hospital Universitario Ramón y Cajal · iacs · Medicina Xenómica · HARMONY

Contenido

Servicios de bioinformática 3

Análisis genómicos, epigenómicos y epitranscriptómicos a partir lecturas largas (Nanopore / PacBio) 3

PTBG-IA-ES15 Análisis genómico de genoma completo a partir de lecturas largas 3

PTBG-IA-ES22 Análisis de epigenoma y epitranscriptómica completo a partir de lecturas largas 4

PTBG-IA-ES23 Análisis de epigenoma y epitranscriptoma de regiones específicas a partir de lecturas largas (1/4 flow cell, sin basecalling) 5

Servicios de basecalling de lecturas largas de Oxford Nanopore 6

PTBG-IA-ES24 Basecalling de alta calidad con detección de bases modificadas 6

PTBG-IA-ES25 Basecalling de alta calidad para análisis genómica 7

PTBG-IA-ES26 Basecalling de calidad muy alta con detección de bases modificadas 7

PTBG-IA-ES27 Basecalling de calidad muy alta para análisis genómicos 8

Análisis genómicos con lecturas cortas y tecnologías basadas en arrays 9

PTBG-IA-ES16 Análisis genómico de paneles dirigidos o arrays 9

PTBG-IA-ES17 Análisis genómico de exomas (WES) 9

PTBG-IA-ES18 Análisis genómico de genoma completo con lecturas cortas (Illumina) 10

Análisis epigenéticos basados en arrays 11

PTBG-IA-ES14 Análisis epigenético (arrays o pequeños paneles) 11

Análisis transcriptómicos 11

PTBG-IA-ES19 Análisis transcriptómico a partir de RNA-seq (bulk) 12

PTBG-IA-ES20 Análisis transcriptómico a partir de arrays de expresión 12

PTBG-IA-ES21 Análisis transcriptómico a partir de scRNA-seq 13

Análisis metagenómicos y metatranscriptómicos 14

PTBG-IA-ES28 Análisis metagenómico y metatranscriptómico de tamaño alto 14

PTBG-IA-ES29 Análisis metagenómico y metatranscriptómico de tamaño medio 15

Servicios de bioinformática

Ofrecemos una amplia gama de servicios avanzados de bioinformática, adaptados a las necesidades de investigadores y organizaciones. Un equipo experto combina capacidades analíticas avanzadas e infraestructura de alto rendimiento con años de experiencia para proporcionar análisis rigurosos y asesoramiento de alta calidad.

A continuación, se describen los servicios ofertados por la plataforma. En cada servicio se describe su alcance. Todo aquello que no figure de manera explícita queda fuera del alcance, salvo acuerdo específico. Para necesidades adicionales o personalizaciones, la plataforma dispone de la modalidad PTBG-ES32 Soluciones a medida, a través de la cual pueden incorporarse personalizaciones en los servicios con planificación, controles y entregables definidos.

Análisis genómicos, epigenómicos y epitranscriptómicos a partir lecturas largas (Nanopore / PacBio)

Análisis de datos genómicos, epigenómicos y epitranscriptómicos generados mediante tecnologías de secuenciación de lecturas largas de Oxford Nanopore y PacBio. Estas tecnologías permiten acceder de forma directa a información estructural y reguladora del genoma y del transcriptoma que no resulta detectable mediante secuenciación de lecturas cortas, lo que habilita la captura de variantes estructurales complejas, regiones altamente repetitivas y modificaciones de bases detectadas a partir de la señal.

Los servicios permiten la caracterización integral de genomas completos, el análisis de regiones genómicas o transcritos específicos, y la detección directa de modificaciones epigenéticas y epitranscriptómicas.

Los servicios de esta categoría no incluyen el basecalling a partir de ficheros POD5 o FAST5, dado que esta actividad se contempla en una categoría específica de servicios.

PTBG-IA-ES15 Análisis genómico de genoma completo a partir de lecturas largas

Descripción del servicio

Análisis de genoma completo a partir de ficheros FASTQ de lectura larga. El enfoque aprovecha la capacidad de las lecturas largas para resolver variantes estructurales y regiones complejas o repetitivas del genoma. La ejecución se realiza con recursos

bioinformáticos especializados y con control de versiones, garantizando trazabilidad y reproducibilidad. Especialmente indicado para estudios de genómica estructural, ensamblado y caracterización genómica avanzada.

Alcance técnico

- Alineamiento al genoma de referencia.
- Llamada de variantes estructurales y de pequeña escala, incluidas SNP e indels.
- Estrategias específicas para regiones de difícil mapeo y áreas repetitivas.

Resultados y formatos de entrega

- Informe de control de calidad con métricas de lecturas largas y de alineamiento.
- Archivo BAM con las lecturas alineadas al genoma de referencia.
- Archivos VCF con las variantes detectadas y anotadas.
- Visualizaciones de variantes estructurales, disponibles bajo solicitud.

PTBG-IA-ES22 Análisis de epigenoma y epitranscriptómica completo a partir de lecturas largas

Descripción del servicio

Caracterización global del epigenoma y del epitranscriptoma a partir de datos crudos generados en un flow cell completo. Las tecnologías de lectura larga, como Oxford Nanopore y PacBio, permiten la detección directa de modificaciones en moléculas nativas de ADN y ARN sin amplificación previa.

Alcance técnico

- Alineamiento de lecturas largas a referencias de ADN y ARN, según aplique.
- Epigenómica: cuantificación de niveles de metilación por CpG y por región; llamada de regiones diferencialmente metiladas; metilación alelo-específica; integración con variación estructural; anotación funcional y enriquecimiento en elementos reguladores; asociación con expresión, fenotipo o rasgos mediante modelos estadísticos.
- Epitranscriptómica: mapeo y cuantificación de modificaciones por transcrito y por posición; identificación de sitios diferencialmente modificados entre condiciones; vinculación con cambios de expresión o con la estabilidad del ARN.

Resultados y formatos de entrega

- Informes de control de calidad.
- Archivos BAM con las lecturas alineadas.
- Archivos VCF con las variantes detectadas y anotadas en epigenómica, cuando proceda.
- Tablas de abundancia de genes o isoformas en epitranscriptómica.
- Perfiles de metilación por muestra.
- Análisis de regiones diferencialmente metiladas, disponibles bajo solicitud.
- Perfiles de haplotipo, disponibles bajo solicitud.
- Anotaciones de eventos de splicing en epitranscriptómica, disponibles bajo solicitud.
- Representaciones visuales de variantes estructurales, disponibles bajo solicitud.

PTBG-IA-ES23 Análisis de epigenoma y epitranscriptoma de regiones específicas a partir de lecturas largas (1/4 flow cell, sin basecalling)

Descripción del servicio

Servicio equivalente al PTBG-IA-ES22, orientado a regiones específicas y ejecutado sobre una fracción de flow cell. Queda especialmente indicado para estudios focalizados, validación de hipótesis concretas y proyectos piloto, manteniendo la detección directa de múltiples modificaciones de bases a partir de la señal.

Alcance técnico

- Alineamiento de lecturas largas a referencias de ADN y ARN, según aplique.
- Epigenómica: cuantificación de metilación por CpG y por región en áreas de interés; llamada de regiones diferencialmente metiladas y metilación alelo-específica; integración con variación estructural; anotación funcional y enriquecimiento en elementos reguladores; asociación con expresión, fenotipo o rasgos mediante modelos estadísticos.
- Epitranscriptómica: mapeo y cuantificación de modificaciones por transcrito y por posición en regiones seleccionadas; identificación de sitios diferencialmente modificados; vinculación con cambios de expresión o con la estabilidad del ARN.

Resultados y formatos de entrega

- Informes de control de calidad.
- Archivos BAM con las lecturas alineadas.

- Archivos VCF con las variantes detectadas y anotadas en epigenómica, cuando proceda.
- Tablas de abundancia de genes o isoformas en epitranscriptómica.
- Perfiles de metilación por muestra.
- Análisis de regiones diferencialmente metiladas, disponibles bajo solicitud.
- Perfiles de haplotipo, disponibles bajo solicitud.
- Anotaciones de eventos de splicing en epitranscriptómica, disponibles bajo solicitud.
- Representaciones visuales de variantes estructurales, disponibles bajo solicitud.

Servicios de basecalling de lecturas largas de Oxford Nanopore

Procesamiento primario de datos de secuenciación de lecturas largas generados con tecnologías de Oxford Nanopore, a partir de señal cruda en formatos FAST5 y POD5. El objetivo consiste en transformar la señal eléctrica en secuencias de nucleótidos, generando archivos FASTQ y BAM sin alinear de alta calidad, listos para análisis posteriores de genómica, epigenómica y epitranscriptómica.

Los servicios incluyen configuraciones de precisión en modo *High Accuracy* y *Super Accuracy*, con detección opcional de modificaciones de base, lo que permite adaptar el procesado a objetivos científicos que van desde análisis genómicos estándar hasta estudios epigenómicos y epitranscriptómicos con máxima fidelidad.

La elección entre High Accuracy y Super Accuracy permite equilibrar precisión, tiempo de cálculo y coste computacional en función de las necesidades del proyecto.

PTBG-IA-ES24 Basecalling de alta calidad con detección de bases modificadas

Descripción del servicio

Llamada de bases en modo High Accuracy, con precisión aproximada del 99,4 por ciento, y detección de bases modificadas. Adecuado para epigenómica y epitranscriptómica, con equilibrio entre precisión y eficiencia computacional.

Alcance técnico

- Detección de bases modificadas según los catálogos indicados para ADN y ARN. En ADN se contemplan 4-metilcitosina, 5-metilcitosina y 5-hidroximetilcitosina en

contexto CG, así como 6-metiladenina. En ARN se contemplan m5C, m6A en contexto DRACH, inosina y pseudouridina.

- Ajuste de parámetros para optimizar precisión y eficiencia según requisitos del estudio.

Resultados y formatos de entrega

- Archivos FASTQ o BAM sin alinear con las bases llamadas y las bases modificadas solicitadas.
- Informe de calidad de la llamada de bases.

PTBG-IA-ES25 Basecalling de alta calidad para análisis genómica

Descripción del servicio

Llamada de bases de ADN en modo High Accuracy, sin detección de bases modificadas. Adecuado para análisis genómicos estándar, incluyendo secuenciación de genoma completo, detección de variantes estructurales y ensamblado de genomas, optimizando tiempos de cálculo y uso de recursos.

Alcance técnico

- Configuración orientada a WGS, variantes estructurales y ensamblado.
- Optimización del rendimiento computacional del flujo de basecalling.

Resultados y formatos de entrega

- Archivos FASTQ o uBAM (*Unmapped BAM*) sin alinear con las bases llamadas.
- Informe de calidad de la llamada de bases.

PTBG-IA-ES26 Basecalling de calidad muy alta con detección de bases modificadas

Descripción del servicio

Llamada de bases en modo Super Accuracy, con precisión aproximada del 99,7% y detección de bases modificadas. Recomendado para proyectos que requieren máxima precisión en la identificación de modificaciones epigenéticas y epitranscriptómicas y para estudios de alta exigencia analítica.

Alcance técnico

- Llamada de bases en modo Super Accuracy.
- Detección de un catálogo ampliado de bases modificadas en ADN y ARN. El modo Super Accuracy contempla un conjunto ampliado de modificaciones respecto a
- High Accuracy, incluyendo m5C y 2-O-metilcitosina; m6A en contexto DRACH; combinaciones con inosina y 2-O-metiladenosina; pseudouridina; 2-O-metiluridina y 2-O-metilguanosina
- Ajuste fino de parámetros para maximizar la precisión en la detección de modificaciones.

Resultados y formatos de entrega

- Archivos FASTQ o uBAM sin alinear con las bases llamadas, incluyendo las bases modificadas solicitadas.
- Informe de calidad de la llamada de bases.

PTBG-IA-ES27 Basecalling de calidad muy alta para análisis genómicos

Descripción del servicio

Llamada de bases de ADN en modo Super Accuracy, sin detección de bases modificadas. Recomendado para proyectos que requieren máxima fidelidad en la llamada de variantes, validación de resultados o análisis sensibles a errores de secuenciación, permitiendo trabajar con coberturas menores.

Alcance técnico

- Llamada de bases en modo Super Accuracy.
- Configuración orientada a máxima fidelidad en análisis genómicos.
- Ajuste del flujo para entornos de validación y estudios sensibles al error.

Resultados y formatos de entrega

- Archivos FASTQ o uBAM sin alinear con las bases llamadas.
- Informe de calidad de la llamada de bases.

Análisis genómicos con lecturas cortas y tecnologías basadas en arrays

Análisis de datos genómicos obtenidos mediante secuenciación de lecturas cortas (Illumina) o tecnologías basadas en arrays. Los servicios cubren desde estudios de genoma completo hasta exomas y paneles dirigidos, con pipelines estandarizados y resultados interpretables.

PTBG-IA-ES16 Análisis genómico de paneles dirigidos o arrays

Descripción del servicio

Análisis de paneles génicos y de datos genómicos obtenidos mediante arrays.

Alcance técnico

- Normalización y control de calidad de datos de arrays.
- Análisis comparativo y generación de indicadores alineados con el diseño experimental.
- En paneles dirigidos: alineamiento al genoma de referencia y llamada de variantes de pequeña escala, incluidas SNVs e indels, con anotación funcional.
- En arrays: genotipado y llamada de segmentos con alteraciones del número de copias y pérdida de heterocigosidad, cuando proceda.

Resultados y formatos de entrega

- Informes de control de calidad.
- Tablas de genotipos y de variantes puntuales en arrays.
- Listado de segmentos con alteraciones del número de copias y pérdida de heterocigosidad, disponible bajo solicitud.
- Archivos BAM con las lecturas alineadas en paneles dirigidos.
- Archivos VCF con las variantes detectadas y anotadas en paneles dirigidos.
- Diseño y elaboración de visualizaciones científicas, disponible bajo solicitud.

PTBG-IA-ES17 Análisis genómico de exomas (WES)

Descripción del servicio

Análisis de las regiones codificantes del genoma a partir de ficheros FASTQ.

Alcance técnico

- Control de calidad y alineamiento al genoma de referencia.
- Llamada de variantes de pequeña escala, incluidas SNVs e indels.
- Llamada de alteraciones del número de copias en exoma, cuando proceda.
- Anotación funcional con herramientas validadas como VEP, ANNOVAR y SnpEff.

Resultados y formatos de entrega

- Informes de control de calidad.
- Archivos BAM con las lecturas alineadas.
- Archivos VCF con las variantes detectadas y anotadas.
- Listado de regiones con alteraciones del número de copias, disponible bajo solicitud.
- Diseño y elaboración de visualizaciones científicas, disponible bajo solicitud.

PTBG-IA-ES18 Análisis genómico de genoma completo con lecturas cortas (Illumina)

Descripción del servicio

Análisis de genoma completo con lecturas cortas Illumina.

Alcance técnico

- Control de calidad y alineamiento al genoma de referencia.
- Llamada de variantes puntuales: SNVs e indels.
- Llamada de variantes estructurales y de alteraciones del número de copias: SV y CNV.
- Anotación funcional con herramientas como VEP, ANNOVAR y SnpEff.

Resultados y formatos de entrega

- Informes de control de calidad.
- Archivos BAM con las lecturas alineadas.
- Archivos VCF con las variantes detectadas y anotadas.
- Listado de regiones con alteraciones del número de copias, disponible bajo solicitud.
- Diseño y elaboración de visualizaciones científicas, disponible bajo solicitud.

Análisis epigenéticos basados en arrays

PTBG-IA-ES14 Análisis epigenético (arrays o pequeños paneles)

Descripción:

Análisis del estado epigenético de cientos de miles de sitios CpG en el ADN utilizando datos generados con chips de metilación, tales como Illumina 450K o Illumina EPIC. También se puede realizar el análisis de paneles pequeños, descifrando patrones de metilación y otras marcas reguladoras que modulan los patrones de expresión del genoma.

Alcance técnico

- Control de calidad y normalización de datos de arrays de metilación.
- Análisis comparativo e identificación de regiones diferencialmente metiladas con criterios reproducibles.
- En paneles dirigidos: alineamiento al genoma de referencia y cuantificación de metilación en posiciones CpG con agregación por regiones funcionales.
- Integración funcional y interpretación biológica de los hallazgos a nivel de genes y rutas.

Resultados y formatos de entrega

- Informes de control de calidad.
- Tablas de intensidades normalizadas en arrays, disponibles bajo solicitud.
- Archivos BAM con las lecturas alineadas en paneles dirigidos.
- Matriz de posiciones CpG metiladas por muestra.
- Listado de regiones diferencialmente metiladas, disponible bajo solicitud en función del diseño muestral.
- Diseño y elaboración de visualizaciones científicas, disponible bajo solicitud.

Análisis transcriptómicos

Servicios de análisis de expresión génica basados en tecnologías de transcriptómica bulk, célula única y arrays de expresión, con procesamiento, cuantificación, análisis comparativo y soporte de interpretación funcional.

PTBG-IA-ES19 Análisis transcriptómico a partir de RNA-seq (bulk)

Descripción del servicio

Procesamiento integral de datos de RNA-seq a partir de lecturas crudas o matrices de expresión previamente cuantificadas. Se incluye control de calidad, preprocesamiento y cuantificación de expresión a nivel de gen y de transcrito mediante alineamiento o pseudoalineamiento, con métricas de mapeo y distribución de lecturas. El análisis descendente contempla normalización, exploración de datos, análisis de expresión diferencial y enriquecimiento funcional con ontologías como GO, KEGG y Reactome. La representación gráfica se adapta a las necesidades del proyecto.

Alcance técnico

- Control de calidad, filtrado y evaluación de métricas de secuenciación.
- Alineamiento al genoma de referencia o pseudoalineamiento con métricas de mapeo y distribución de lecturas.
- Cuantificación de expresión a nivel de gen y de transcrito.
- Normalización, exploración multivariante y expresión diferencial entre condiciones.
- Enriquecimiento funcional con GO, KEGG y Reactome, y síntesis de resultados para interpretación.

Resultados y formatos de entrega

- Informes de control de calidad.
- Archivos BAM con las lecturas alineadas, disponibles bajo solicitud.
- Matrices de conteo en versión cruda y normalizada.
- Tablas con los resultados de la expresión diferencial.
- Tablas con los resultados del enriquecimiento funcional.
- Diseño y elaboración de visualizaciones científicas (volcano plots, heatmaps, etc).

PTBG-IA-ES20 Análisis transcriptómico a partir de arrays de expresión

Descripción del servicio

Análisis de datos transcriptómicos obtenidos mediante microarrays, partiendo de ficheros crudos o matrices ya procesadas.

Alcance técnico

- Control de calidad específico de plataforma y corrección de fondo con normalización.
- Filtrado de sondas por calidad, variabilidad y especificidad.
- Mapeo y actualización de anotaciones sonda-gen.
- Exploración multivariante mediante PCA y clustering jerárquico con ajuste por lotes si procede.
- Expresión diferencial basada en modelos lineales
- Análisis funcional a nivel de distintas ontologías (GO, KEGG, Reactome).

Resultados y formatos de entrega

- Informes de control de calidad.
- Matrices de conteo en versión cruda o normalizada.
- Tablas con los resultados de la expresión diferencial.
- Tablas con los resultados del enriquecimiento funcional.
- Diseño y elaboración de visualizaciones científicas (volcano plots, heatmaps, etc).

PTBG-IA-ES21 Análisis transcriptómico a partir de scRNA-seq

Descripción del servicio

Análisis de datos de transcriptómica a célula única a partir de lecturas crudas o matrices de conteos.

Alcance técnico

- Control de calidad, filtrado de celulas de baja calidad, detección y tratamiento de dobletes.
- Reducción de dimensionalidad mediante PCA, UMAP y t-SNE; clustering con ajuste de resolución.
- Identificación de tipos celulares por marcadores y por referencias.
- Integración de múltiples muestras con corrección de lotes y análisis de proporciones celulares.
- Expresión diferencial adaptada a datos single-cell (por clúster y pseudobulk) con soporte de interpretación funcional.

Resultados y formatos de entrega

- Informes de control de calidad.
- Archivos BAM con las lecturas alineadas.
- Matrices de conteo en versión cruda y normalizada.
- Tablas de anotación de células.

- Listados de marcadores por clúster o población: tablas CSV/TSV con genes, logFC, p-valor/FDR, y estadísticas de expresión diferencial entre clústeres.
- Tablas con los resultados del enriquecimiento funcional, disponibles bajo solicitud.
- Diseño y elaboración de visualizaciones científicas (PCA, UMAP/t-SNE, clustering, volcano plots, heatmaps, etc).

Análisis metagenómicos y metatranscriptómicos

Servicios orientados al estudio de comunidades microbianas a partir de datos de ADN o ARN, tanto en entornos ambientales como en muestras humanas.

PTBG-IA-ES28 Análisis metagenómico y metatranscriptómico de tamaño alto

Descripción del servicio

Análisis de proyectos de gran volumen de datos, incluyendo identificación taxonómica, análisis de abundancia relativa, reconstrucción de rutas metabólicas y caracterización funcional del microbioma analizado.

Alcance técnico

- Control de calidad y eliminación de lecturas del hospedador, cuando proceda.
- Asignación taxonómica con estimación de abundancias a nivel de familia, género y especie, según el diseño acordado; cuantificación funcional a nivel de genes, ortología y rutas metabólicas.
- Análisis de abundancia diferencial entre grupos.
- Reconstrucción de rutas metabólicas e interpretación biológica integrada del microbioma.

Resultados y formatos de entrega

- Informes de control de calidad.
- Lecturas filtradas, disponibles bajo solicitud.
- Matrices de conteos y abundancias a los niveles taxonómicos requeridos.
- Análisis de abundancia diferencial entre grupos, disponibles bajo solicitud.

- Informe de caracterización funcional de las muestras, disponible bajo solicitud.
- Diseño y elaboración de visualizaciones científicas acordadas ((abundancia, diversidad α o β, etc).

PTBG-IA-ES29 Análisis metagenómico y metatranscriptómico de tamaño medio

Descripción del servicio

Análisis de proyectos de tamaño medio, con pipelines equivalentes al servicio PTBG-IA-ES28 manteniendo las capacidades de identificación taxonómica, cuantificación de abundancias y caracterización funcional.

Alcance técnico

- Control de calidad y eliminación de lecturas del hospedador, cuando proceda.
- Asignación taxonómica con estimación de abundancias a nivel de familia, género y especie, según el diseño acordado; cuantificación funcional a nivel de genes, ortología y rutas metabólicas.
- Análisis de abundancia diferencial entre grupos.
- Reconstrucción de rutas metabólicas e interpretación biológica integrada del microbioma.

Resultados y formatos de entrega

- Informes de control de calidad.
- Lecturas filtradas, disponibles bajo solicitud.
- Matrices de conteos y abundancias a los niveles taxonómicos requeridos.
- Análisis de abundancia diferencial entre grupos, disponibles bajo solicitud.
- Informe de caracterización funcional de las muestras, disponible bajo solicitud.
- Diseño y elaboración de visualizaciones científicas acordadas (abundancia, diversidad α o β, etc).