OMICSPACE

Ethical and secure -omics data space

BDIAB

IIS La Fe

Plataforma de Big Data, IA, Bioestadística & Bioinformática

SERVICE CATALOG



Participants:

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

Ethical and secure -omics data space

SERVICE CATALOG





BIOINFORMATICS

BDIAB

IIS La Fe

Plataforma de Big Data, IA, Bioestadística & Bioinformática

Participants:

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

Table of contents

Bioinformatics Services **3**

Genomic, epigenomic and epitranscriptomic analyses from long-read sequencing with Oxford Nanopore and PacBio **3**

PTBG-IA-ES15 Whole-genome long-read genomic analysis 3

PTBG-IA-ES22 Comprehensive epigenome and epitranscriptome analysis from long reads 4

PTBG-IA-ES23 Epigenome and epitranscriptome analysis of specific regions from long reads (one quarter of a flow cell, basecalling excluded) 5

Oxford Nanopore long-read basecalling services **5**

PTBG-IA-ES24 High-accuracy basecalling with detection of modified bases 6

PTBG-IA-ES25 High-accuracy basecalling for genomic analyses 6

PTBG-IA-ES26 Super-accuracy basecalling with detection of modified bases 7

PTBG-IA-ES27 Super-accuracy basecalling for genomic analyses 7

Genomic analyses with short reads and array-based technologies **8**

PTBG-IA-ES16 Genomic analysis of targeted panels or arrays 8

PTBG-IA-ES17 Genomic Analysis of Exomes (WES) 9

PTBG-IA-ES18 Whole-genome sequencing analysis with short reads 9

Array-based epigenetic analyses **10**

PTBG-IA-ES14 Epigenetic analysis with arrays or small panels 10

Transcriptomic analyses **10**

PTBG-IA-ES19 Transcriptomic analysis from bulk RNA-seq 10

PTBG-IA-ES20 Transcriptomic analysis from expression arrays 11

PTBG-IA-ES21 Transcriptomic analysis from single-cell RNA-seq 12

Metagenomic and metatranscriptomic analyses **13**

PTBG-IA-ES28 Large-scale metagenomic and metatranscriptomic análisis 13

PTBG-IA-ES29 Mid-scale metagenomic and metatranscriptomic análisis 13

Bioinformatics Services

We offer a wide range of advanced bioinformatics services tailored to the needs of researchers and organizations. An expert team combines advanced analytical capabilities and high-performance infrastructure with years of experience to deliver rigorous analyses and high-quality consulting.

Below is a description of the services provided by the platform. Each service includes its scope. Anything not explicitly stated is considered out of scope unless a specific agreement is made. For additional needs or customizations, the platform offers the PTBG-ES32 Custom Solutions option, which allows incorporating service customizations with defined planning, controls, and deliverables.

Genomic, epigenomic and epitranscriptomic analyses from long-read sequencing with Oxford Nanopore and PacBio

Analysis of genomic, epigenomic, and epitranscriptomic data generated using long-read sequencing technologies from Oxford Nanopore and PacBio. These technologies enable direct access to structural and regulatory information in the genome and transcriptome that cannot be detected with short-read sequencing, allowing the capture of complex structural variants, highly repetitive regions, and base modifications detected from the signal.

The services support comprehensive characterization of whole genomes, analysis of specific genomic regions or transcripts, and direct detection of epigenetic and epitranscriptomic modifications.

Services in this category do not include basecalling from POD5 or FAST5 files, as this activity is covered under a separate service category.

PTBG-IA-ES15 Whole-genome long-read genomic analysis

Service Description

Whole-genome analysis from long-read FASTQ files. This approach leverages the ability of long reads to resolve structural variants and complex or repetitive regions of the genome. Execution is carried out using specialized bioinformatics resources with version control, ensuring traceability and reproducibility. Particularly suited for studies in structural genomics, genome assembly, and advanced genomic characterization.

Technical Scope

- Alignment to the reference genome.
- Calling of structural and small-scale variants, including SNPs and indels.

- Specific strategies for hard-to-map regions and repetitive areas.

Results and Delivery Formats

- Quality control report with long-read and alignment metrics.
- BAM file with reads aligned to the reference genome.
- VCF files with detected and annotated variants.
- Visualizations of structural variants, available upon request.

PTBG-IA-ES22 Comprehensive epigenome and epitranscriptome analysis from long reads

Service Description

Global characterization of the epigenome and epitranscriptome from raw data generated on a full flow cell. Long-read technologies such as Oxford Nanopore and PacBio enable direct detection of modifications in native DNA and RNA molecules without prior amplification.

Technical Scope

- Alignment of long reads to DNA and RNA references, as applicable.
- Epigenomics: quantification of methylation levels by CpG and by region; calling of differentially methylated regions; allele-specific methylation; integration with structural variation; functional annotation and enrichment in regulatory elements; association with expression, phenotype, or traits using statistical models.
- Epitranscriptomics: mapping and quantification of modifications by transcript and position; identification of differentially modified sites between conditions; linking to expression changes or RNA stability.

Results and Delivery Formats

- Quality control reports.
- BAM files with aligned reads.
- VCF files with detected and annotated variants in epigenomics, where applicable.
- Gene or isoform abundance tables for epitranscriptomics.
- Methylation profiles per sample.
- Analysis of differentially methylated regions (available upon request).
- Haplotype profiles (available upon request).
- Annotations of splicing events in epitranscriptomics (available upon request).
- Visual representations of structural variants (available upon request).

PTBG-IA-ES23 Epigenome and epitranscriptome analysis of specific regions from long reads (one quarter of a flow cell, basecalling excluded)

Service Description

Equivalent to PTBG-IA-ES22, but focused on specific regions and executed on a fraction of a flow cell. Particularly suited for targeted studies, hypothesis validation, and pilot projects, while maintaining direct detection of multiple base modifications from the signal.

Technical Scope

- Alignment of long reads to DNA and RNA references, as applicable.
- Epigenomics: quantification of methylation by CpG and by region in areas of interest; calling of differentially methylated regions and allele-specific methylation; integration with structural variation; functional annotation and enrichment in regulatory elements; association with expression, phenotype, or traits using statistical models.
- Epitranscriptomics: mapping and quantification of modifications by transcript and position in selected regions; identification of differentially modified sites; linking to expression changes or RNA stability.

Results and Delivery Formats

- Quality control reports.
- BAM files with aligned reads.
- VCF files with detected and annotated variants in epigenomics, where applicable.
- Gene or isoform abundance tables for epitranscriptomics.
- Methylation profiles per sample.
- Analysis of differentially methylated regions (available upon request).
- Haplotype profiles (available upon request).
- Annotations of splicing events in epitranscriptomics (available upon request).
- Visual representations of structural variants (available upon request).

Oxford Nanopore long-read basecalling services

Primary processing of long-read sequencing data generated with Oxford Nanopore technologies, starting from raw signal in FAST5 and POD5 formats. The objective is to transform the electrical signal into nucleotide sequences, producing high-quality unaligned FASTQ and BAM files, ready for downstream genomics, epigenomics and epitranscriptomics analyses.

The services include precision configurations in High Accuracy mode and Super Accuracy mode, with optional detection of base modifications, allowing the processing to be tailored to scientific goals ranging from standard genomic analyses to epigenomic and epitranscriptomic studies with maximum fidelity. The choice between High Accuracy and Super Accuracy makes it possible to balance accuracy, compute time and computational cost according to project needs.

PTBG-IA-ES24 High-accuracy basecalling with detection of modified bases

Service Description

Basecalling in High Accuracy mode, with an approximate precision of 99.4%, including detection of modified bases. Suitable for epigenomic and epitranscriptomic studies, offering a balance between accuracy and computational efficiency.

Technical Scope

- Detection of modified bases according to specified catalogs for DNA and RNA.
    - DNA: 4-methylcytosine, 5-methylcytosine, and 5-hydroxymethylcytosine in CG context, as well as 6-methyladenine.
    - RNA: m5C, m6A in DRACH context, inosine, and pseudouridine.
- Parameter tuning to optimize accuracy and efficiency based on study requirements.

Results and Delivery Formats

- FASTQ or unaligned BAM files with called bases and requested modified bases.
- Basecalling quality report.

PTBG-IA-ES25 High-accuracy basecalling for genomic analyses

Service Description

Basecalling of DNA in High Accuracy mode, without detection of modified bases. Suitable for standard genomic analyses, including whole-genome sequencing, structural variant detection, and genome assembly, optimizing computation time and resource usage.

Technical Scope

- Configuration tailored for WGS, structural variants, and assembly workflows.
- Optimization of computational performance for basecalling.

Results and Delivery Formats

- FASTQ or uBAM (Unmapped BAM) files with called bases.
- Basecalling quality report.

PTBG-IA-ES26 Super-accuracy basecalling with detection of modified bases

Service Description

Basecalling in Super Accuracy mode, with an approximate precision of 99.7%, including detection of modified bases. Recommended for projects requiring maximum accuracy in identifying epigenetic and epitranscriptomic modifications and for studies with high analytical demands.

Technical Scope

- Basecalling in Super Accuracy mode.
- Detection of an extended catalog of modified bases in DNA and RNA. Compared to High Accuracy, Super Accuracy includes additional modifications such as:
    - DNA/RNA: m5C and 2′-O-methylcytosine; m6A in DRACH context; combinations with inosine and 2′-O-methyladenosine; pseudouridine; 2′-O-methyluridine and 2′-O-methylguanosine.
- Fine-tuning of parameters to maximize precision in modification detection.

Results and Delivery Formats

- FASTQ or uBAM files with called bases, including requested modified bases.
- Basecalling quality report.

PTBG-IA-ES27 Super-accuracy basecalling for genomic analyses

Service Description

Basecalling of DNA in **Super Accuracy** mode, without detection of modified bases. Recommended for projects requiring maximum fidelity in variant calling, result validation, or analyses sensitive to sequencing errors, enabling work with lower coverage.

Technical Scope

- Basecalling in Super Accuracy mode.
- Configuration optimized for maximum accuracy in genomic analyses.
- Workflow adjustments for validation environments and error-sensitive studies.

Results and Delivery Formats

- FASTQ or uBAM files with called bases.
- Basecalling quality report.

Genomic analyses with short reads and array-based technologies

Analyses of genomic data obtained through short-read sequencing with Illumina or array-based technologies. The services cover whole-genome studies as well as exomes and targeted panels, with standardized workflows and interpretable results.

PTBG-IA-ES16 Genomic analysis of targeted panels or arrays

Service Description

Analysis of gene panels and genomic data obtained through arrays.

Technical Scope

- Normalization and quality control of array data.
- Comparative analysis and generation of indicators aligned with the experimental design.
- For targeted panels: alignment to the reference genome and calling of small-scale variants, including SNVs and indels, with functional annotation.
- For arrays: genotyping and calling of segments with copy number alterations and loss of heterozygosity, where applicable.

Results and Delivery Formats

- Quality control reports.
- Tables of genotypes and point variants for arrays.
- Lists of segments with copy number alterations and loss of heterozygosity (available upon request).
- BAM files with aligned reads for targeted panels.
- VCF files with detected and annotated variants for targeted panels.
- Design and development of scientific visualizations (available upon request).

PTBG-IA-ES17 Genomic Analysis of Exomes (WES)

Service Description

Analysis of coding regions of the genome from FASTQ files.

Technical Scope

- Quality control and alignment to the reference genome.
- Calling of small-scale variants, including SNVs and indels.
- Calling of copy number alterations in exomes, where applicable.
- Functional annotation using validated tools such as VEP, ANNOVAR, and SnpEff.

Results and Delivery Formats

- Quality control reports.
- BAM files with aligned reads.
- VCF files with detected and annotated variants.
- Lists of regions with copy number alterations (available upon request).
- Design and development of scientific visualizations (available upon request).

PTBG-IA-ES18 Whole-genome sequencing analysis with short reads

Service Description

Whole-genome analysis using Illumina short-read sequencing.

Technical Scope

- Quality control and alignment to the reference genome.
- Calling of point variants: SNVs and indels.
- Calling of structural variants and copy number alterations: SV and CNV.
- Functional annotation using tools such as VEP, ANNOVAR, and SnpEff.

Results and Delivery Formats

- Quality control reports.
- BAM files with aligned reads.
- VCF files with detected and annotated variants.
- Lists of regions with copy number alterations (available upon request).
- Design and development of scientific visualizations (available upon request).

Array-based epigenetic analyses

PTBG-IA-ES14 Epigenetic analysis with arrays or small panels

Service Description

Analysis of the epigenetic state of hundreds of thousands of CpG sites in DNA using data generated with methylation chips such as Illumina 450K or Illumina EPIC. Analysis of small panels can also be performed, decoding methylation patterns and other regulatory marks that modulate genome expression profiles.

Technical Scope

- Quality control and normalization of methylation array data.
- Comparative analysis and identification of differentially methylated regions using reproducible criteria.
- For targeted panels: alignment to the reference genome and quantification of methylation at CpG positions with aggregation by functional regions.
- Functional integration and biological interpretation of findings at gene and pathway level.

Results and Delivery Formats

- Quality control reports.
- Tables of normalized intensities for arrays (available upon request).
- BAM files with aligned reads for targeted panels.
- Matrix of methylated CpG positions per sample.
- Lists of differentially methylated regions (available upon request depending on study design).
- Design and development of scientific visualizations (available upon request).

Transcriptomic analyses

Services for gene expression analysis based on bulk transcriptomics, single-cell technologies, and expression arrays, including processing, quantification, comparative analysis, and functional interpretation support.

PTBG-IA-ES19 Transcriptomic analysis from bulk RNA-seq

Service Description

Comprehensive processing of RNA-seq data from raw reads or pre-quantified expression matrices. Includes quality control, preprocessing, and expression quantification at gene and

transcript level via alignment or pseudo-alignment, with mapping metrics and read distribution. Downstream analysis includes normalization, data exploration, differential expression analysis, and functional enrichment using ontologies such as GO, KEGG, and Reactome. Graphical representation is adapted to project needs.

Technical Scope

- Quality control, filtering, and sequencing metric evaluation.
- Alignment to the reference genome or pseudo-alignment with mapping and distribution metrics.
- Expression quantification at gene and transcript level.
- Normalization, multivariate exploration, and differential expression between conditions.
- Functional enrichment with GO, KEGG, and Reactome, and synthesis of results for interpretation.

Results and Delivery Formats

- Quality control reports.
- BAM files with aligned reads (available upon request).
- Raw and normalized count matrices.
- Tables with differential expression results.
- Tables with functional enrichment results.
- Design and development of scientific visualizations (volcano plots, heatmaps, etc.).

PTBG-IA-ES20 Transcriptomic analysis from expression arrays

Service Description

Analysis of transcriptomic data obtained through microarrays, starting from raw files or pre-processed matrices.

Technical Scope

- Platform-specific quality control and background correction with normalization.
- Probe filtering by quality, variability, and specificity.
- Mapping and updating of probe-to-gene annotations.
- Multivariate exploration using PCA and hierarchical clustering with batch adjustment if applicable.
- Differential expression based on linear models.
- Functional analysis across multiple ontologies (GO, KEGG, Reactome).

Results and Delivery Formats

- Quality control reports.
- Raw or normalized count matrices.
- Tables with differential expression results.
- Tables with functional enrichment results.
- Design and development of scientific visualizations (volcano plots, heatmaps, etc.).

PTBG-IA-ES21 Transcriptomic analysis from single-cell RNA-seq

Service Description

Analysis of single-cell transcriptomic data from raw reads or count matrices.

Technical Scope

- Quality control, filtering of low-quality cells, detection and handling of doublets.
- Dimensionality reduction using PCA, UMAP, and t-SNE; clustering with resolution adjustment.
- Identification of cell types by markers and reference datasets.
- Integration of multiple samples with batch correction and analysis of cell proportions.
- Differential expression adapted to single-cell data (by cluster and pseudo-bulk) with functional interpretation support.

Results and Delivery Formats

- Quality control reports.
- BAM files with aligned reads.
- Raw and normalized count matrices.
- Cell annotation tables.
- Marker lists per cluster or population: CSV/TSV tables with genes, logFC, p-value/FDR, and differential expression statistics between clusters.
- Tables with functional enrichment results (available upon request).
- Design and development of scientific visualizations (PCA, UMAP/t-SNE, clustering, volcano plots, heatmaps, etc.).

Metagenomic and metatranscriptomic analyses

Services for the study of microbial communities from DNA or RNA data in environmental contexts and in human samples.

PTBG-IA-ES28 Large-scale metagenomic and metatranscriptomic análisis

Service Description

Analysis of large-scale projects, including taxonomic identification, relative abundance analysis, metabolic pathway reconstruction, and functional characterization of the analyzed microbiome.

Technical Scope

- Quality control and removal of host reads, where applicable.
- Taxonomic assignment with abundance estimation at family, genus, and species level, as per agreed design; functional quantification at gene, orthology, and metabolic pathway level.
- Differential abundance analysis between groups.
- Reconstruction of metabolic pathways and integrated biological interpretation of the microbiome.

Results and Delivery Formats

- Quality control reports.
- Filtered reads (available upon request).
- Count and abundance matrices at required taxonomic levels.
- Differential abundance analysis between groups (available upon request).
- Functional characterization report of samples (available upon request).
- Design and development of agreed scientific visualizations (abundance, α or β diversity, etc.).

PTBG-IA-ES29 Mid-scale metagenomic and metatranscriptomic análisis

Service Description

Analysis of medium-scale projects using pipelines equivalent to PTBG-IA-ES28, maintaining capabilities for taxonomic identification, abundance quantification, and functional characterization.

Technical Scope

- Quality control and removal of host reads, where applicable.
- Taxonomic assignment with abundance estimation at family, genus, and species level, as per agreed design; functional quantification at gene, orthology, and metabolic pathway level.
- Differential abundance analysis between groups.
- Reconstruction of metabolic pathways and integrated biological interpretation of the microbiome.

Results and Delivery Formats

- Quality control reports.
- Filtered reads (available upon request).
- Count and abundance matrices at required taxonomic levels.
- Differential abundance analysis between groups (available upon request).
- Functional characterization report of samples (available upon request).
- Design and development of agreed scientific visualizations (abundance, α or β diversity, etc.).