

## 1. DESCRIPCIÓN DEL SERVICIO

El servicio de Secuenciación Masiva tiene como finalidad el proporcionar asesoramiento y soporte técnico a los grupos de investigación interesados en realizar proyectos de ultrasecuenciación.

Partiendo de las necesidades de cada proyecto, el servicio podrá proporcionar al investigador asesoramiento en el diseño del experimento de ultrasecuenciación y se le proporcionará el presupuesto para su realización. Adicionalmente, el servicio podrá proporcionar soporte técnico o servicios parciales en las siguientes tareas:

- ♣ Preparación de las muestras (librerías) para su ultrasecuenciación.
- ♣ En servicios de control de calidad orientados a la ultrasecuenciación.
- ♣ En la gestión de su secuenciación con la plataforma más conveniente para cada tipo de proyecto.
- ♣ En el análisis bioinformático de los datos generados.

Esta lista no excluye otras actividades que podrán ser planteadas a la dirección técnica del servicio.

Las tarifas establecidas en el anexo I están estimadas de acuerdo a una previsión de un nivel óptimo de ocupación del servicio por lo que podrán ser revisadas a la vista de condiciones particulares. El presupuesto final será elaborado por el servicio tras el trabajo.

## 2. SERVICIOS OFRECIDOS

### Generación de librerías

La preparación de las muestras y la generación de las librerías supone un paso esencial en cualquier proyecto de secuenciación. Además, el protocolo de preparación de librería condiciona de manera determinante el tipo de información que se obtendrá de la secuenciación. El servicio de genómica ofrece inicialmente los protocolos más comunes de generación de librerías. Protocolos no ofrecidos en esta propuesta podrán ser incorporados al portafolio de servicios si existe una demanda por parte de los investigadores y siempre con posterioridad a la puesta a punto de la técnica por parte del responsable del laboratorio. Actualmente sólo se ofrecen protocolos de generación de librerías compatibles con la plataforma de Illumina.

Los protocolos de generación de librerías ofrecidos inicialmente por el servicio son:

a) *Librerías de DNA*: Generación de librerías a partir de muestras de DNA. Este tipo de librerías se utiliza en general para la identificación de sitios de unión de proteínas al DNA (Chip-Seq), el ensamblado de genomas *de novo* o la identificación de variantes de secuencia con respecto a un genoma de referencia (DNA-Seq).

b) *Librerías de mRNA*: Generación de librerías a partir de mRNA purificado. Este tipo de librerías se utiliza generalmente para la identificación de expresión diferencial entre distintos grupos de muestras y puede ser utilizado a su vez para la identificación de desequilibrio alélico, expresión individualizada de diversas isoformas del mismo gen o para la identificación de variantes de splicing en el caso de muestras eucariotas.

c) *Librerías de RNAs pequeños*: Variante del protocolo de preparación de muestras de RNA pero encaminado al análisis de especies de RNA de pequeño tamaño (<50 pb) que normalmente no son capaces de ser identificadas en el protocolo estándar.

d) *Librerías direccionales de RNA*: Con este protocolo, a diferencia del protocolo estándar, se obtiene información sobre qué hebra del DNA original es la codificante de las moléculas de RNA.

#### Control de calidad de muestras/librerías

Además del soporte técnico en la generación de las librerías. Se ofrecerá servicio independiente de control de calidad de las muestras de DNA, RNA o las librerías ya generadas mediante su análisis con el equipamiento Agilent 2100 bioanalyzer.

Complementariamente, también se ofrece un servicio de cuantificación de las librerías mediante PCR cuantitativa utilizando oligonucleótidos específicos de los adaptadores presentes en las librerías preparadas para los equipos de Illumina.

Estos servicios están ya incluidos en el servicio de generación de librerías.

### Secuenciación.

La secuenciación de las muestras podrá ser realizada “in house” o remitida a centros externos de secuenciación en las mejores condiciones a juicio del servicio.

### Análisis bioinformático

La gran cantidad de datos producidos en cualquier proyecto de ultrasecuenciación genera la necesidad de equipamiento y formación específica para extraer la información requerida por parte del investigador. El Servicio de Secuenciación Masiva cuenta con equipamiento y personal especializado para ofrecer una serie de servicios de análisis informático a aquellos investigadores que no dispongan de estas infraestructuras. Inicialmente el servicio ofrece los análisis más demandados, comprometiéndose a actualizar los ofrecidos e incorporar nuevos servicios a medida que exista la demanda de los mismos y se acepte su viabilidad.

Los servicios ofrecidos inicialmente son:

a) *Almacenamiento*: Los datos generados en un proyecto de ultrasecuenciación son de gran volumen. Se ofrece un servicio de almacenamiento tanto accesible en el servidor (por un tiempo prudencial) como en soporte externo como copia de seguridad. Este servicio sin embargo no puede garantizar la seguridad de los datos en casos de fuerza mayor como inundaciones, incendios, etc., al no disponer de la infraestructura necesaria.

b) *Alineamiento con un genoma de referencia*: Alineamiento de cada una de las millones de pequeñas secuencias generadas por la plataforma de ultrasecuenciación con un genoma de referencia permitiendo identificar la localización concreta de la secuencia y cualquier discordancia que presente frente a la referencia. Este análisis requiere una importante capacidad y tiempo de procesamiento y es un paso esencial para cualquier procesamiento posterior.

c) *Cálculo de RPKM (Reads per Kilobase and per million reads)*: En los proyectos de RNA, este análisis permite identificar el número de secuencias/lecturas correspondiente a cada uno de los genes, lo que representa una medida cuantitativa de la expresión relativa de este gen en la muestra.

d) *Expresión diferencial*: El servicio ofrece la identificación de genes que presenten expresión diferencial entre dos grupos de muestras en los que se haya realizado experimentos de RNA seq. Un servicio estándar incluye la realización de hasta 5 análisis diferentes dividiendo el conjunto de muestras cada vez en dos grupos de acuerdo a distintas variables. Por cada variable extra requerida por el investigador se añadirá un coste adicional.

e) *Identificación de variantes de secuencia*: Este análisis permite identificar las variantes de secuencia (sustituciones y variantes estructurales en su caso) que presenta la muestra frente al genoma de referencia. Junto con un listado de estas variantes se ofrecerá información sobre la posible repercusión funcional de este cambio (si se dispone de un modelo de genes para ese organismo) y sobre la existencia de esta variante en las bases de datos públicas en el caso de que provengan de una muestra humana.

f) *Identificación de sitios de unión de proteínas*: En los proyectos de Chip-Seq se ofrecerá un listado de los sitios de enriquecimiento, obtenidos a través de las variaciones de cobertura de las regiones del DNA enriquecidas.

Estos servicios son los más comunes y los que se recogen en las tarifas. Somos conscientes de que cada trabajo tendrá sus particularidades y además queremos flexibilizar el servicio al máximo de forma que un usuario pueda requerir cualquier parte del trabajo de forma independiente. Por ello, esta información debe considerarse orientativa y sólo tras una entrevista usuario-servicio seremos capaces de proporcionar un presupuesto ajustado.

### **3. TIEMPO DE PROCESAMIENTO Y PRIORIDADES**

Los distintos proyectos dirigidos al servicio se procesarán, a no ser que existan circunstancias especiales, según el orden de solicitud. Se estima un procesamiento máximo de 30-40 librerías al mes. El procesamiento informático de un proyecto estándar de 10 muestras requiere por término medio una semana, sin embargo este tiempo es estimativo y podrá variar de acuerdo al número de muestras a procesar y los análisis requeridos.

#### **4. REQUERIMIENTOS DE LAS MUESTRAS**

A) RNA. En el caso de RNA total se requerirán de 1 a 10 microgramos (recomendado 10 microgramos). La muestra deberá presentar un valor de integridad o “RIN” > 8 (medido en el Bionalyzer). En su defecto, en el caso de RNA eucariota se estimará la calidad de la muestra mediante gel de agarosa en el que se muestre una intensidad del rRNA 28S aproximadamente dos veces la intensidad mostrada por el rRNA 18S. Alternativamente, se puede utilizar mRNA aislado y purificado (mínimo de 100 nanogramos).

B) DNA. En el caso de DNA total se requerirá de 1 a 5 microgramos (recomendado 5 microgramos) de DNA no degradado con un ratio de absorbancia 260/280 entre 1.8 y 2. En el caso de DNA enriquecido (Chip-Seq) se requerirán al menos 10 nanogramos con una concentración al menos de 250 µg/L.



## ANEXO 1. TARIFAS

### Control de calidad y generación de librerías

| <i>Servicio</i>                         | <i>&lt; 5 librerías</i> | <i>5 ó más librerías</i> |
|---|-------------------------|--------------------------|
| Análisis Bioanalyzer*                   | 50 €/chip               | 50 €/chip                |
| Librería de DNA Illumina                | 160 €/muestra           | 120 €/muestra            |
| Librería de mRNA Illumina               | 200 €/muestra           | 150 €/muestra            |
| Librería de RNAs de pequeño tamaño      | 280 €/muestra           | 220 €/muestra            |
| Librerías de Chip-seq                   | 250 €/muestra           | 190 €/muestra            |
| Librerías direccionales de RNA Illumina | 390 €/muestra           | 340 €/muestra            |
| Librería shotgun Roche                  | 250 €/muestra           |                          |
| Cuantificación de librerías*            | 35 €/muestra            | 20 €/muestra             |

### Secuenciación

| <i>Plataforma Illumina</i>     | <i>carril externo (18 x 10<sup>7</sup> lecturas)</i> |
|--------------------------------|--|
| 35 bases single-end            | No disponible  |
| 35 bases paired-end            | No disponible  |
| 50 bases single-end            | 1400 €/ 2000 USD                                     |
| 50 bases paired-end            | 1900 €/ 3000 USD                                     |
| 70 bases single-end            | No disponible  |
| 70 bases paired-end            | No disponible  |
| 100 bases single-end           | 2100 €/ 2500 USD                                     |
| 100 bases paired-end           | 30000 €/ 4000 USD                                    |
| <i>Plataforma Junior Roche</i> |  |
| shotgun                        | 1.500 €  |
| amplicones                     | 1.500 €  |

### Análisis informáticos

| <i>Servicio</i>                   | <i>Precio interno</i> | <i>Precio Externo</i> |
|-----------------------------------|-----------------------|-----------------------|
| Almacenamiento                    | 10 €/100 Gb           | 30 €/100Gb            |
| Alineamiento                      | 25 €/muestra          | 75 €/muestra          |
| Cálculo RPKM                      | 25 €/muestra          | 75 €/muestra          |
| Expresión diferencial             | 20 €/análisis         | 60 €/análisis         |
| Identificación de variantes       | 25 €/muestra          | 75 €/muestra          |
| Identificación de sitios de unión | 25 €/muestra          | 75 €/muestra          |

La tarifa interna se aplicará a los usuarios de centros públicos de investigación y la externa a los usuarios comerciales