

## **Guía de usuarios de SECUENCIACIÓN MASIVA (NGS)** **de ILLUMINA**



Illumina NextSeq500

La Unidad de Genómica de CABIMER ofrece **Servicio de Secuenciación Masiva** o (NGS) utilizando las siguientes plataformas de Illumina:

- [NextSeq500](#)
- [NovaSeq6000](#).

Esta tecnología está basada en secuenciación por síntesis (SBS [SBS video](#)) con nucleótidos marcados con fluorocromos y terminadores reversibles, permitiendo así una secuenciación masiva y paralela en millones de fragmentos de ADN.

Específicamente, estos sistemas resuelven con eficiencia los homopolímeros. Son aptos para diferentes aplicaciones, pudiendo secuenciar desde paneles pequeños de genes hasta exomas, transcriptomas y genomas de humanos y de mayor complejidad en un solo experimento, así como diferentes proyectos de forma paralela para aumentar la rentabilidad.

Dichas plataformas ofrecen una **Flexibilidad** y **Escalabilidad** muy útil debido a:

- **Distintos tipos de matrices** (flowcells) con **diferente rendimiento: desde 130 Millones hasta 20 Billones de lecturas.**
- La secuenciación se puede realizar desde un único extremo de las moléculas de ADN (**single-end**) o desde ambos (**paired-end**).
- Longitud de lectura: **35bp, 50bp, 75bp, 100bp, 150bp, 300bp o 500bp**, según el sistema y la flowcell.
- Todas las muestras están marcadas con un “código de barras o *index*”, pudiendo introducir distintas muestras en la misma carrera de secuenciación, amortizando así la reacción y el experimento en general (**multiplexar**).



Illumina NovaSeq6000

Formatos de secuenciación	Millones Reads/flow <sup>2</sup>	Gigabases	Secuenciador
Mid Output -150 nucleótidos <sup>1</sup>	130M	16-19	NextSeq500
Mid Output -300 nucleótidos <sup>1</sup>		32-39	
High Output - 75 nucleótidos <sup>1</sup>	400M	25-30	
High Output -150 nucleótidos <sup>1</sup>		50-60	
High Output - 300 nucleótidos <sup>1</sup>		100-120	
SP Reagent - 100 nucleótidos <sup>1</sup>	800M	65-80	NovaSeq6000
SP Reagent - 200 nucleótidos <sup>1</sup>		134-167	
SP Reagent - 300 nucleótidos <sup>1</sup>		200-250	
SP Reagent - 500 nucleótidos <sup>1</sup>		325-400	
S1 Reagent - 100 nucleótidos <sup>1</sup>	1.600M	134-167	
S1 Reagent - 200 nucleótidos <sup>1</sup>		266-333	
S1 Reagent - 300 nucleótidos <sup>1</sup>		400-500	
S2 Reagent - 100 nucleótidos <sup>1</sup>	4.100M	333-417	
S2 Reagent - 200 nucleótidos <sup>1</sup>		667-833	
S2 Reagent - 300 nucleótidos <sup>1</sup>		1.000-1.250	
S4 Reagent - 35nucleótidos <sup>1</sup>	10.000M	280-350	
S4 Reagent - 200 nucleótidos <sup>1</sup>		1.600-2.000	
S4 Reagent - 300 nucleótidos <sup>1</sup>		2.400-3.000	

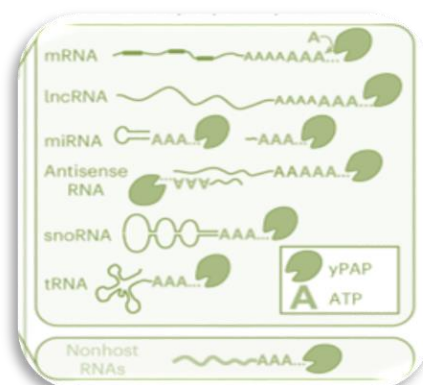
>**Nucleótidos<sup>1</sup>** o pares de bases: Los nucleótidos se pueden distribuir en *single- ó paired-end* (Por ejemplo, 200 nucleótidos se pueden leer como 2 X 100pb ó 1 X 200pb).

>**2:** Millones de lecturas en *single-end* (en *paired-end* serían el doble)

Así, con las plataformas de Illumina existentes en la Unidad, el usuario puede trabajar con distinto nº de muestras y una amplia gama de **aplicaciones** para todo tipo de organismos procariotas y eucariotas. Entre las más solicitadas están:

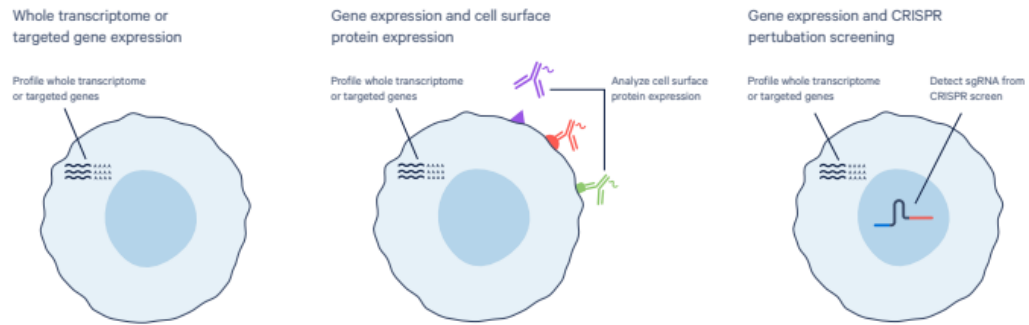
➤ **Transcriptómica (RNA SEQUENCING):**

- [Total RNA-Seq](#)
- [mRNAseq](#)
- [Small RNA-Seq](#)
- [Ribo-Seq](#), [RNA-Exome-Seq](#), [targeted RNA-Seq...](#)



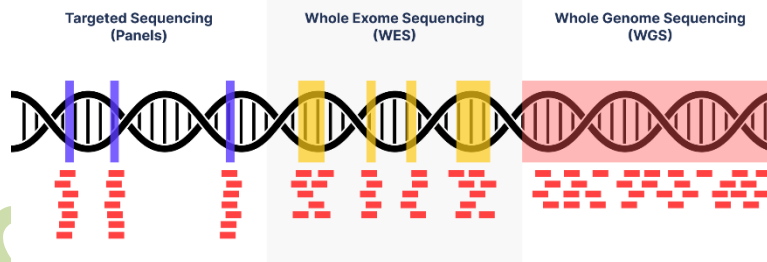
➤ **Secuenciación de Célula única (SINGLE-CELL)**

- **Single-Cell gene expression** (scRNA-Seq y scRNA-Seq PLEX)
- **Single-Cell ATAC** (scATAC-Seq)
- **Single-Cell Multiome** (simultaneous scRNA-seq + ATAC-seq)
- **Single-Cell CRISPR Gene expression** (CRISPR screening)



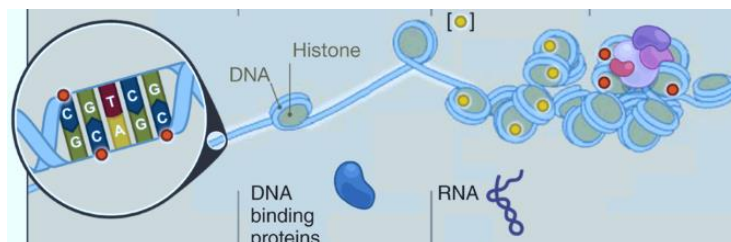
➤ **Secuenciación de Genomas (GENOME SEQUENCING)**

- **Whole-genome sequencing** (WGS)
- **De novo sequencing**
- **Whole-exome sequencing (WES)**
- **Target Capture** (paneles personalizados, de enfermedades,...)



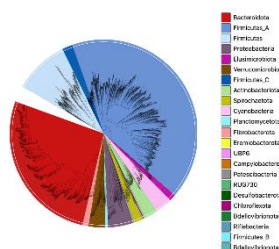
➤ **Epigenómica (EPIGENOMICS)**

- **Chip-Seq**, **ATAC-Seq**, Mnasa-Seq, **3C**, **4C**...
- **RIP-Seq**, **DRIP-Seq**, **DRIPc-Seq**



➤ **Metagenómica (METAGENOMICS)**

- **16S sequencing**
- **Shotgunsequencing**



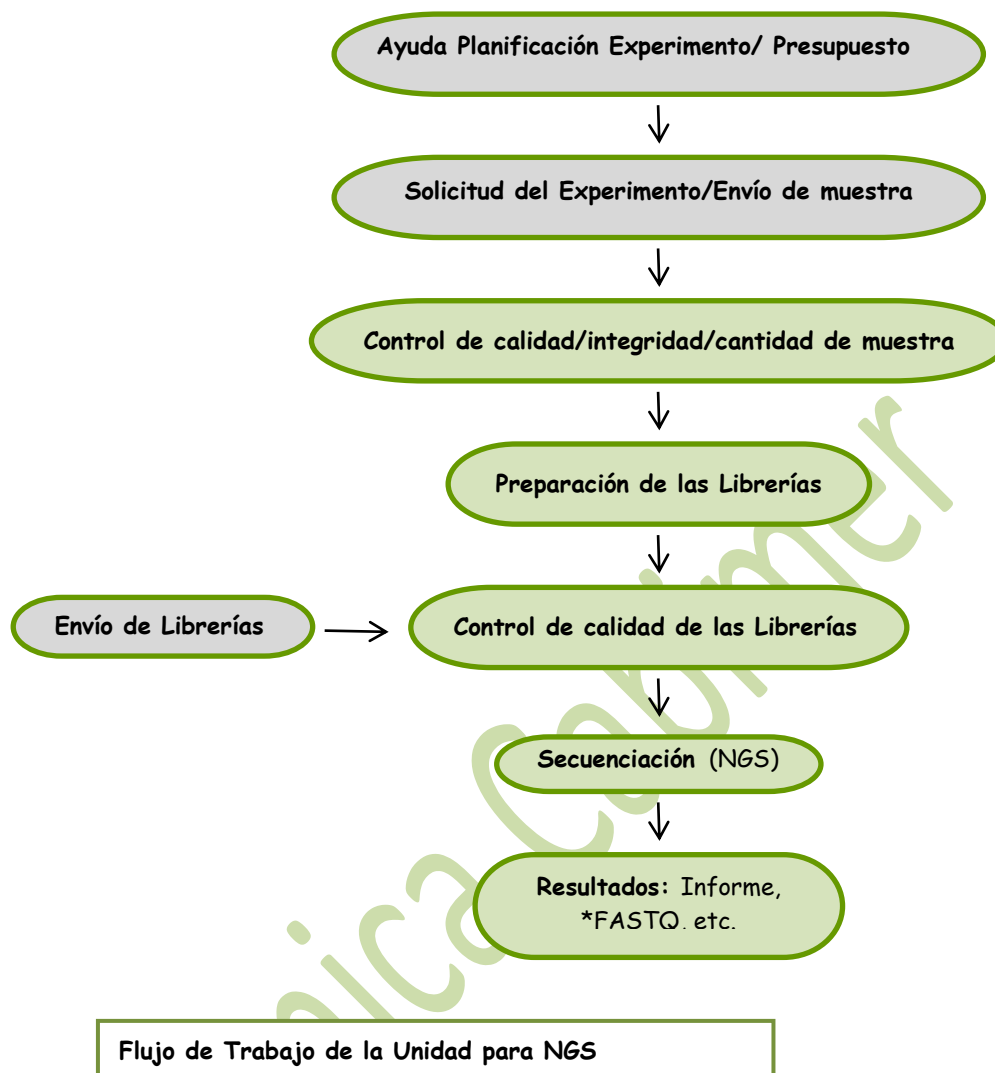
Los datos de los experimentos de Secuenciación se pueden analizar en una amplia gama de programas de código abierto, o analizar y almacenar de manera segura en la nube de Illumina [BaseSpace®](#). Ésta última es un entorno informático de Illumina que incluye herramientas para que la gestión y análisis de datos resulte lo suficientemente sencilla para laboratorios que se estén iniciando en este campo así como lo suficientemente potente para ascender con rapidez en análisis de NGS. Permite el análisis de Exomas, transcriptomas, genomas completos, llamadas de variantes somáticas y metagenómica entre otros.

Para ampliar la información, consulte en la web de [Illumina](#).

**EI SERVICIO INCLUYE**

Los Experimentos de Secuenciación Masiva (NGS) solicitados en la Unidad **incluyen los siguientes puntos y servicios:**

1. **Asesoramiento** respecto al diseño experimental (Aplicaciones, Cobertura, etc...). Una vez decidido el nº de muestras, matriz de secuenciación, etc.. se da un presupuesto del proyecto. En caso de decidir realizar el Experimento con la Unidad, la **Solicitud** debe formalizarse cumplimentando y enviando el Formulario de Solicitud correspondiente.
2. **Envío de muestras**. Para una mayor flexibilidad, el Servicio de NGS de la Unidad, ofrece **dos puntos de entrada en el flujo de trabajo:**
  - a) **Solicitud de Elaboración de Librerías y Secuenciación**. En este caso, el usuario entrega la muestra con los requisitos que se detallan más adelante y la Unidad realizará el **Control de calidad** de las mismas mediante BioAnalyzer o TapeStation, Fluorimetría (Qubit), Espectrofotometría, qPCR, etc, según determine el servicio por aplicación. Posteriormente, si todo está correcto, informará al usuario y se procederá con la **Preparación, Control de Calidad y Normalización** de las **Librerías**. Este Servicio se lleva a cabo dependiendo de la aplicación.



b) **Solicitud de Secuenciación**: El usuario entrega en la Unidad la librería ya preparada, cuantificada (preferiblemente por Fluorimetría), con rango de tamaño adecuado (*Bioanalyzer* o similar si es posible) y que sea compatible con el Sistema de Secuenciación. Una vez entregada, la Unidad realizará los controles pertinentes para comprobar la eficiencia y calidad de la misma (qPCR, Qubit, etc..) previos a la secuenciación en sí.

3. **Matrices y reactivos** para la **Secuenciación** en la plataforma Illumina.

4. **Análisis primario de los Resultados** que consistirá en los siguientes archivos:

- Secuencias de la/s muestras o Librería/s (archivos FASTQ) en las que ya se ha realizado la transformación de datos brutos (.bcl), el Filtrado y *Trimming* de adaptadores, para hacer entrega de secuencias que superan ya uno parámetros de Calidad

- o Si se usan códigos de Barras o índices, los resultados se entregarán de forma individual para cada uno de ellos.
- o En los experimentos de *single-cell*, se realiza adicionalmente el análisis secundario de los datos que permite obtener los archivos necesarios para visualizar y manipular los *clusters* que se hayan identificado en las muestras analizadas, con las herramientas de 10xGenomics. Para el análisis secundario de otras aplicaciones, consultar con la Unidad
- o Informe \*.pdf con los parámetros del Control de Calidad del proceso y de la Carrera de secuenciación, así como los Datos/Secuencias resultantes del experimento.

### **REQUISITOS DE LAS MUESTRAS DE ILLUMINA**

Si el usuario va a entregar la Librería ya preparada y compatible con Illumina, contactar con la Unidad.

Si va entregar la muestra (ADN o ARN) para que la Unidad prepare la Librería debe seguir las siguientes recomendaciones. En general, **La calidad del Material de partida es MUY IMPORTANTE para lograr Datos Significativos, fiables y robustos**. La **cantidad** de partida **determina el protocolo** y el kit que en ese momento se haya optimizado en el laboratorio. **El siguiente listado es orientativo**; PREGUNTE ANTES PARA SU APLICACIÓN EN CONCRETO.

#### **a) Muestras de ADN:**

- Proporcionar la cantidad requerida **resuspendida en agua libre de nucleasas o lowTE**.
- **ADN Bicatenario, no degradado y Sin contaminantes** (RNA, EDTA, Mg2+, etc.).
- Realizar preferiblemente la **purificación utilizando columnas** (Qiagen o similares). Si se escoge otro método, eliminar completamente las trazas de cualquier contaminante (Etanol, fenol, etc...) ya que interfieren en los pasos enzimáticos de preparación de las librerías (**muy sensibles al EDTA**).
- **Cuantificar las muestras de manera precisa**. Es mejor utilizar un **método fluorimétrico** por ejemplo Quant-IT (Qubit, Invitrogen) para medir con precisión la cantidad de ADN de doble cadena. En el caso de DNA-Seq, se puede utilizar la cuantificación por espectrofotometría (ej. NanoDrop); los ratios 260/280 deben estar entre 1.8-2.0, y 260/230 entre 2.0-2.2.
- Siempre que sea posible, realizar una electroforesis en gel de agarosa de las muestras para **confirmar la integridad de las mismas** (Proporcionar la imagen). El DNA genómico e íntegro debe observarse como una única banda >10-12 kb.

- Si la muestra es cDNA, indicar el método utilizado para su síntesis.

**Tabla Orientativa. IMPORTANTE CONFIRMAR CON LA UNIDAD**

Tipo de Muestra	Cantidad Total	Volumen mínimo	Concentración mínima	Cuantificación
<b>DNAg</b>	1-500 ng Rango óptimo: 100-500 ng	<b>20 µl</b>	40 ng/µl	NanoDrop. Por debajo de 100ng, Qubit
<b>ChiP enriched DNA (sonicado)<sup>1</sup></b>	50pg-50ng	<b>15 µl</b>	0,2 ng/µl	Qubit
<b>MNase-Seq, DRIP-Seq</b>	50pg-50ng	<b>15 µl</b>	0,2 ng/µl	Qubit

1. ChiP-enriched, fragmented DNA (200-800bp). Library molecules with shorter inserts (**200-300 bp**) tend to cluster and amplify more efficiently on the Illumina flow cell.

- Para muestras de Chip-Seq además:
  - Comprobaciones previas pertinentes para ver que las muestras han inmunoprecipitado bien.
  - Pruebas de sonicación hasta obtener muestras por debajo de 1kb (recomendable <800pb y no menos de 200bp).
  - Es preferible **no** precipitar las muestras **con glucógeno o glico-blue** ya que altera el análisis de las muestras en bioanalyzer y a veces afecta a las enzimas que se usan en los protocolos de librerías. Si no hay otra opción consultar con la unidad.
  - ADN descroslinkado y purificado (con columna preferiblemente) para quitar todos los restos de proteínas etc.... Un mínimo de 100pg totales (preguntar si es menos) a una concentración superior a 5pg/uL.

Una vez recibidas las muestras y según el tipo, la Unidad evaluará la Calidad/Cantidad de las mismas con distinta metodología. Si **éstas no cumplen los requisitos de control de calidad iniciales**, se le **informará al usuario** por correo electrónico.

#### **b) Muestras de RNA:**

- **Eluir** o resuspender el RNA en **agua libre de RNasas** (previamente **tratado con DNasa**)
- Las muestras deben estar **libres** de **cualquier** tipo de **contaminante**. Realizar la purificación utilizando columnas (Qiagen RNeasy, Invitrogen Ribominus o similares), o precipitación con etanol. En este último caso eliminar completamente las trazas de etanol para que no interfiera en los pasos enzimáticos de preparación de las librerías. **ES NECESARIO ELIMINAR LAS TRAZAS DE gDNA. No se recomienda la**



**inactivación de la DNasa por calor**, afecta a la integridad de la muestra. No sólo basta con inactivarla, es necesario eliminarla (columna, precipitación, etc..).

- **Cuantificar** las muestras de **manera precisa**: Espectrofotometro (NanoDrop®), Fulorímetro (Qubit), según cantidades a determinar.
- Se recomienda utilizar muestras de RNA total con un valor de **RIN>7** y ratio 28S-rRNA:18-rRNA de 2:1. **Existen alternativas para muestras de baja cantidad y/o calidad** (FFPE, etc...), consultad con la Unidad.

**Tabla Orientativa. IMPORTANTE CONFIRMAR CON LA UNIDAD**

Tipo de Muestra	Cantidad Total <sup>1</sup>	Volumen mínimo	Concentración mínima
RNA total "mRNA-Seq"	25-1000 ng	>15 µl	10 ng/µl
RNA total "Total RNA-Seq"	1-1000 ng para RIN≥9 10-100 ng para RIN≥2 o muestras FFPE (DV200>55%)	>15µl	10 ng/µl
Small RNA	500-1000ng total RNA	>15 µl	50 ng/µl
<sup>2</sup> RIP-Seq, CLIP-Seq, Ribo-Seq	10-100 ng	>10 µl	-----

1. Si no puede obtener dicha cantidad, consultar con la Unidad alternativas posibles.
2. El ARN de interés debe tener 5'-monofosfato y 3'-OH. Dichas modificaciones pueden ser añadidas a la mayoría de las moléculas de ARN con T4 PNK (Consultar con la Unidad).

**c) Muestras para Single-cell (células o núcleos):**

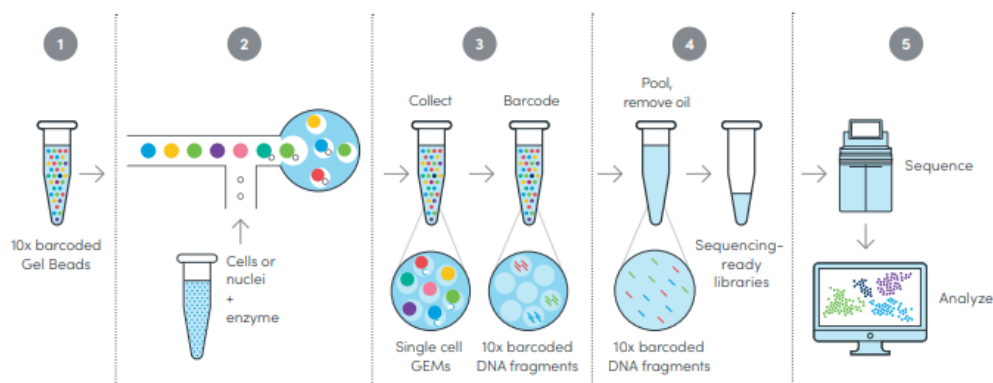
Para diferentes aplicaciones de célula única o single-cell, la Unidad se completa con el equipo [Chromium Controller de 10X Genomics](#). Esta tecnología está basada en la encapsulación de las células de forma individual en esferas, para así marcar el material genético de cada una de manera distintiva y única, poder para hacer un análisis masivo de secuenciación.



Para el catálogo de aplicaciones disponibles, consultar la página de [10x Genomics](#).



Para los requisitos de las muestras para las aplicaciones, consultar con la unidad. La información se actualiza constantemente, así como las aplicaciones. El flujo de trabajo, en cualquier caso, es el que refleja el siguiente esquema:



### **ENVÍO DE MUESTRAS**

Las muestras de **RNA/DNA**, una vez purificadas y en las concentraciones indicadas en los apartados anteriores se enviarán congeladas en nieve carbónica **en tubos Eppendorf® de 1,5 mL en Agua libre de RNAsas** ó Low TE, según proceda. Debe acompañarlas una copia original impresa del **Formulario de Solicitud** cumplimentado correctamente.

Para aplicaciones de **Single cell**, consultad con la Unidad.

Si por algún motivo ajeno a la Unidad, una vez realizada la solicitud, el usuario cancelara la misma, el material fungible correspondiente le será facturado.

La dirección a la cual pueden ser remitidas es:

Unidad de Genómica  
Centro Andaluz de Biología Molecular y  
Medicina Regenerativa (Cabimer)  
Avda. Américo Vespucio nº 24.  
Parque Tecnológico Cartuja '93  
41.092 SEVILLA

Horario de recepción de muestras: De Lunes a Jueves: 9:00 h a 16:00 h, los Viernes de 9:00 a 14:00 h (9:00h a 14:00h Julio y Agosto)

El envío de las muestras **debe ser comunicado previamente** por teléfono (954-467828) o por e-mail (eloisa.andujar@cabimer.es/monica.perez@cabimer.es).

Por su parte, la Unidad de Genómica comunicará vía e-mail la recepción de las mismas, así como cualquier problema que pueda surgir tras los controles de calidad realizados sobre éstas.

Para las aplicaciones de single-cell, contactar con el personal de la unidad.

### **ENTREGA DE LOS RESULTADOS**

Una vez terminado el experimento y obtenidos los datos de su análisis, los resultados se le enviarán al usuario vía ftp junto con un informe final (\*.pdf) explicando el procesamiento de muestras y datos.

Como se indicaba anteriormente, se entregan los Datos de secuencia en Formato FASTQ (*raw sequence+quality scores*), que contiene las lecturas que ya han pasado un Filtrado de Calidad.

### **TARIFAS**

Las tarifas que se pueden aplicar dependen mucho de la cantidad de material de partida, número de muestras que se usen y se puedan secuenciar en el mismo experimento de secuenciación, longitud de lecturas, etc... Consultar con la Unidad cada proyecto.

### **INFORMACIÓN DE INTERÉS**

- **NextSeq500** Illumina <https://emea.illumina.com/systems/sequencing-platforms/nextseq.html>
- **NovaSeq6000** Illumina <https://emea.illumina.com/systems/sequencing-platforms/novaseq.html>
- **BaseSpace Sequence Hub** Illumina. *BaseSpace Sequence Hub is a cloud-based genomics analysis and storage platform that directly integrates with all Illumina sequencers.*
  - <https://emea.illumina.com/products/by-type/informatics-products/basespace-sequence-hub.html>
  - <https://emea.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps.html>
  - <https://help.basespace.illumina.com/>
- **Library Prep Selector:** <https://emea.illumina.com/library-prep-array-kit-selector.html>
- **Sequencing Coverage Calculator**  
[https://emea.support.illumina.com/downloads/sequencing\\_coverage\\_calculator.html](https://emea.support.illumina.com/downloads/sequencing_coverage_calculator.html)
- **Paired versus Single end sequencing** <https://emea.illumina.com/science/technology/next-generation-sequencing/plan-experiments/paired-end-vs-single-read.html>

- **Whole Genome (WGS) versus Exome Sequencing**  
<https://emea.illumina.com/techniques/sequencing/dna-sequencing/whole-genome-sequencing/whole-genome-vs-exome.html>
- **Frequently Asked questions** <https://support-docs.illumina.com/TS/FAQ/Content/FAQ.htm>
- **10X Genomics** <https://www.10xgenomics.com/>

Genómica Cabimer