



**UNIVERSIDAD DE SONORA**



**POSGRADO EN BIOCIENCIAS**

**MICROBIOLOGÍA MOLECULAR (2442)**

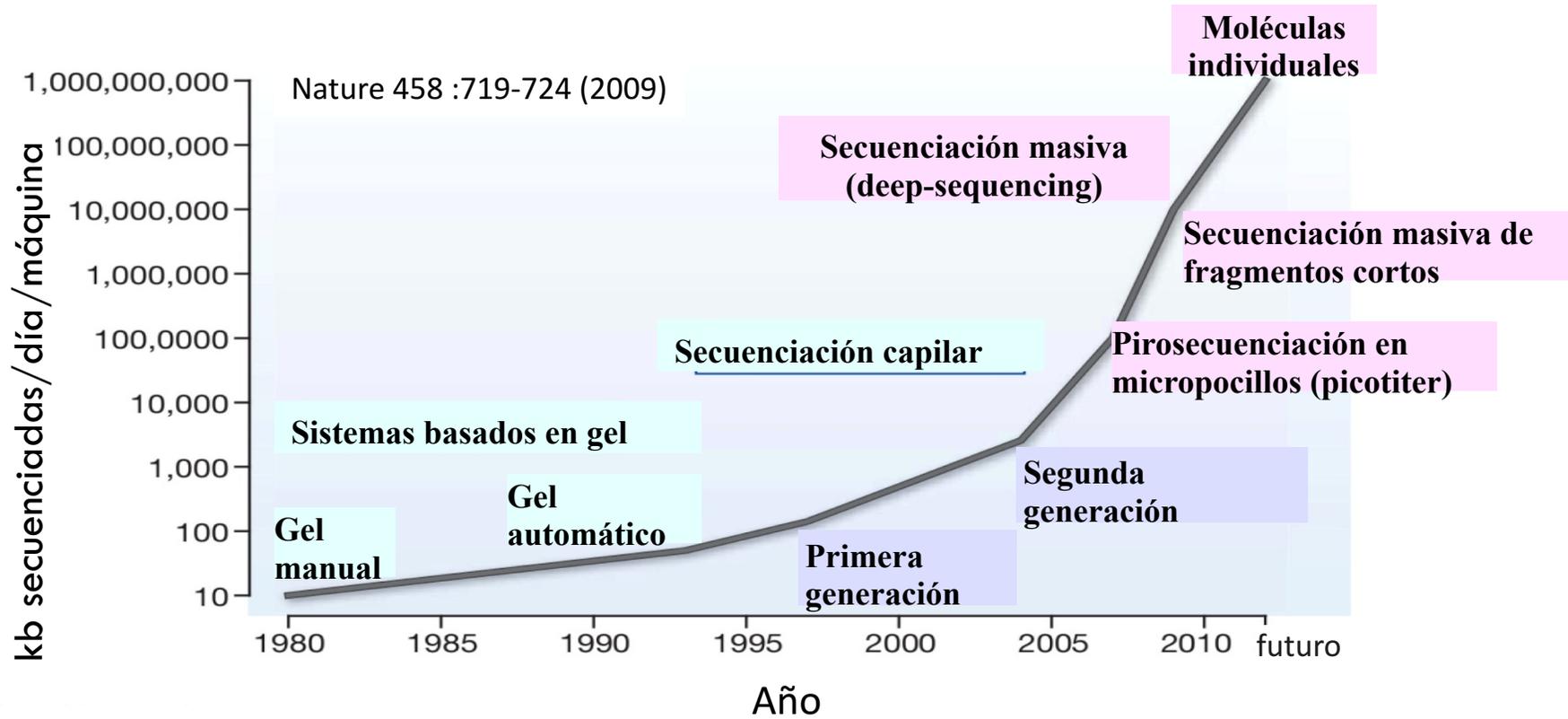
**UNIDAD 7. Secuenciación**

**Elaboró: Dra. Kadiya del C. Calderón Alvarado**

# TÓPICOS

- I. Estrategias de secuenciación
- II. Primera generación (Maxam-Gilbert y Sanger)
- III. Nuevas plataformas (pirosecuenciación, Illumina, Solid, Ion torrent)
- IV. Aplicaciones
- V. Pipelines para el análisis de secuencias

# EVOLUCIÓN DE LOS MÉTODOS DE SECUENCIACIÓN



Maxam-Gilbert

Sanger

Automatizado

Masiva o siguiente generación

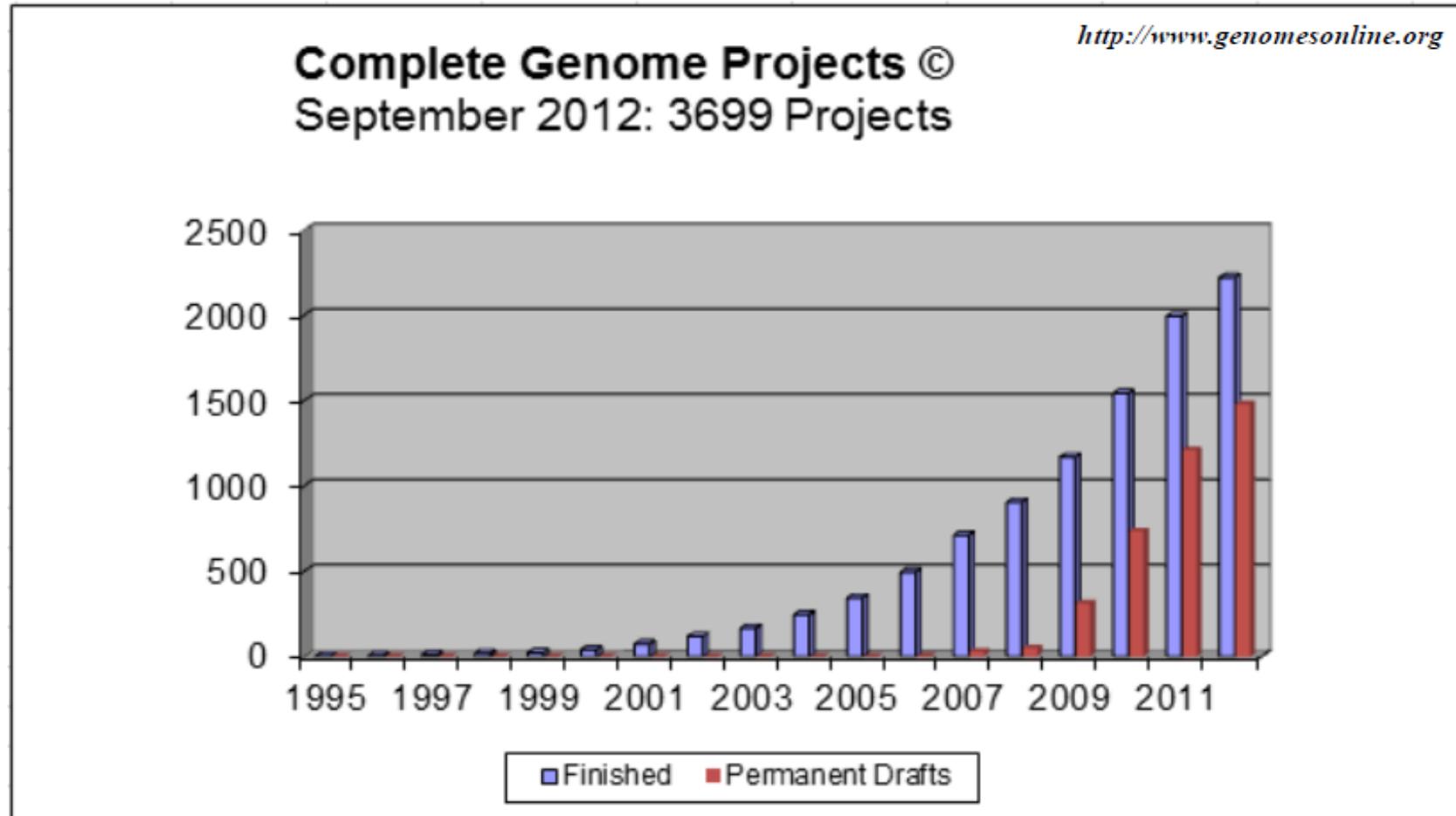
Tercera generación

# Genomas Secuenciados

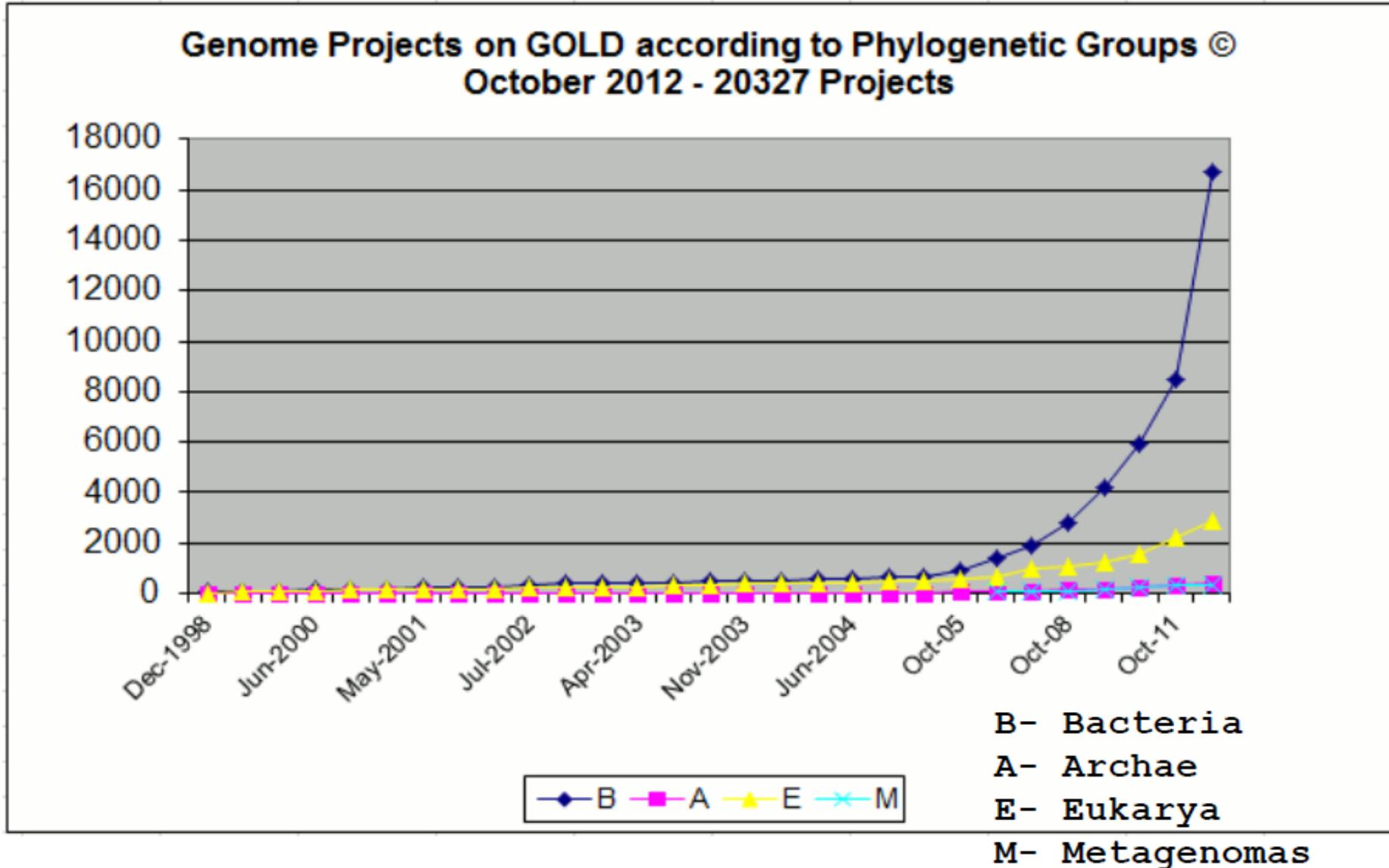
1990 Primer proyecto de genoma

1995 Secuenciación automatizada; aceleración de proyectos genomas

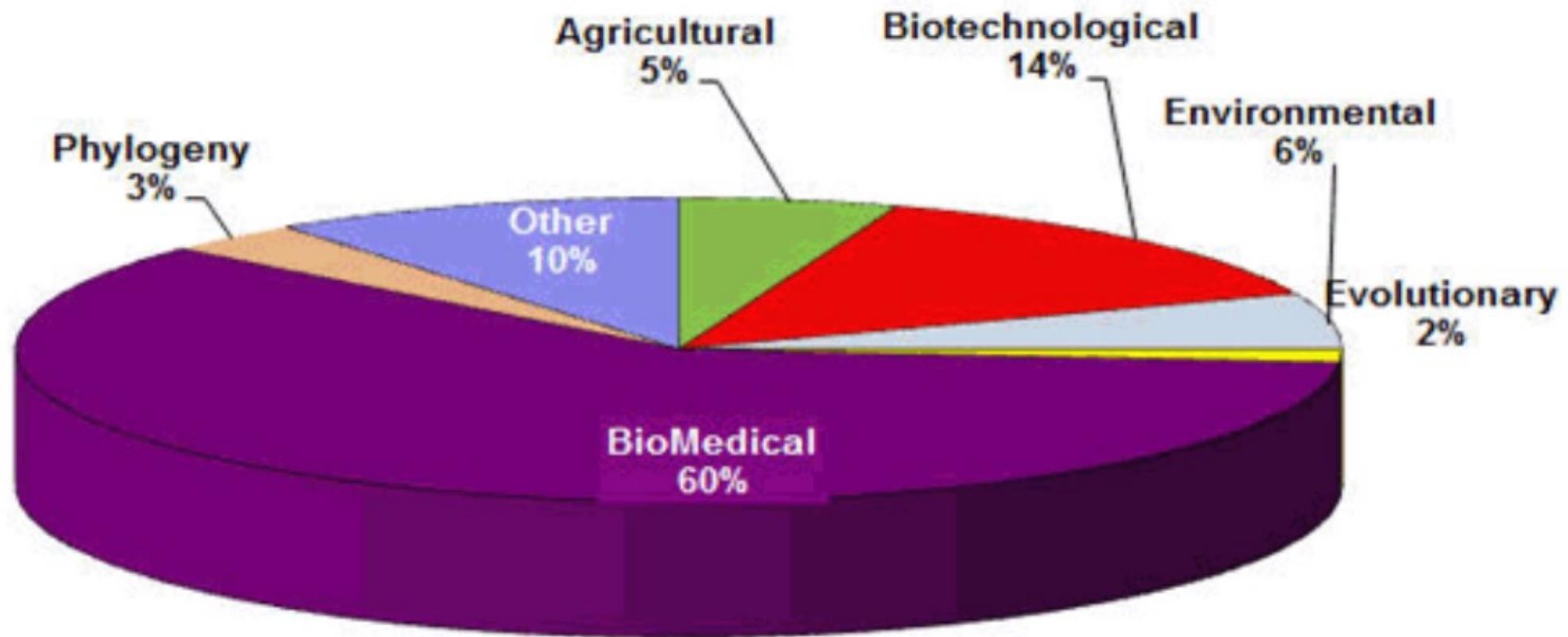
2001 Liberación del genoma humano



# Distribución de Genomas Secuenciados



# Propósitos de financiamiento para Genomas bacterianos



# CONCEPTO DE SECUENCIACIÓN



Es un conjunto de métodos y técnicas bioquímicas, cuya finalidad es determinar el orden de los nucleótidos (A, C, G y T) en una cadena de ADN.

1. Descubrir todas las variaciones en una secuencia
2. Visión general de todos los factores de virulencia de un organismo
3. Estudios taxonómicos y de identificación

# MAXAM — GILBERT - QUÍMICO

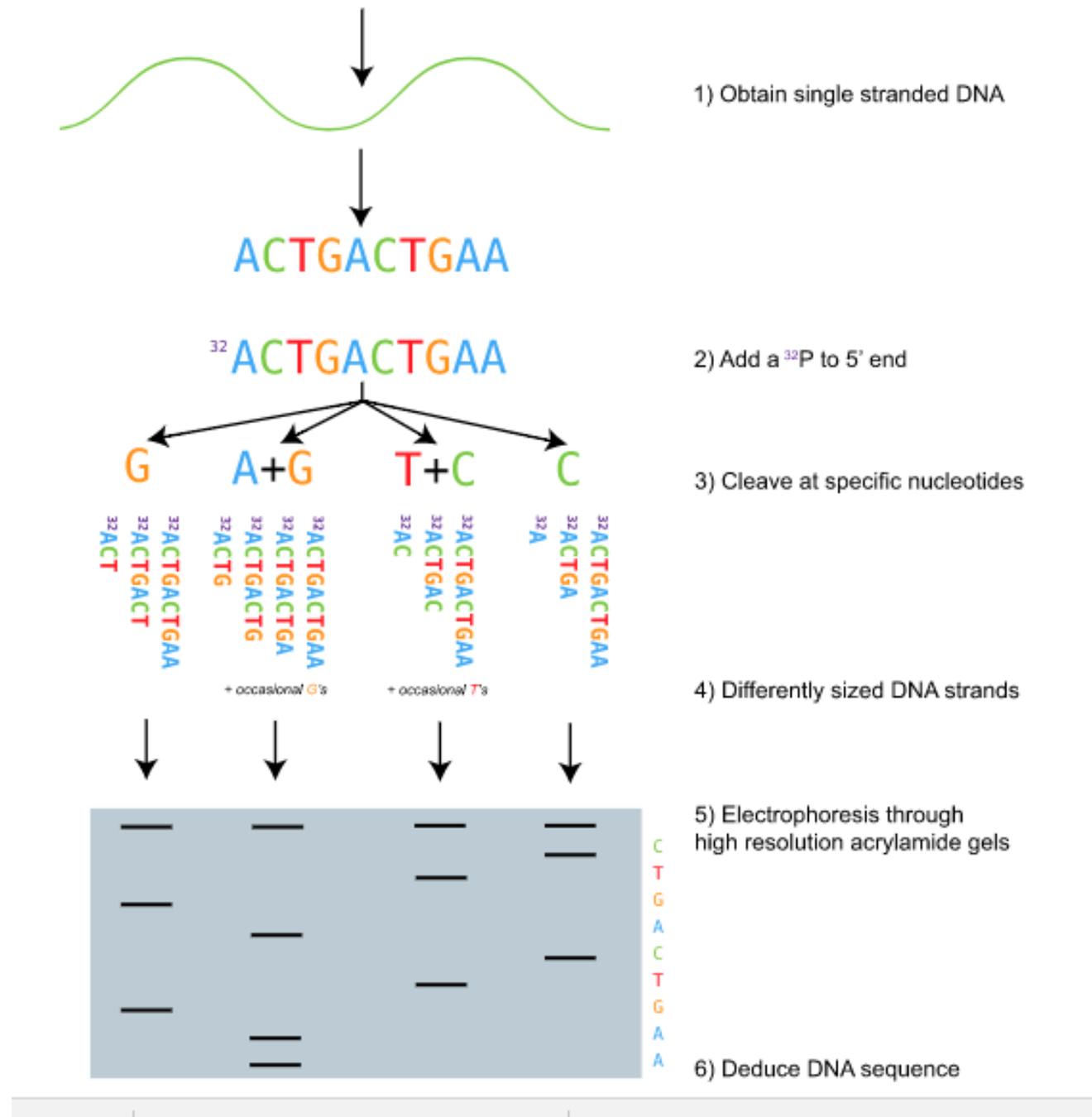
Se realiza un marcaje radioactivo en el extremo 5' del fragmento de ADN con P<sup>32</sup>

El método utiliza 4 reacciones químicas

- G
- A+G
- C
- C+T

# MAXAM – GILBERT

- Fragmentos separados por electroforesis en gel de **poliacrilamida**
- Detecta diferencias de longitud de tan solo un nucleótido.

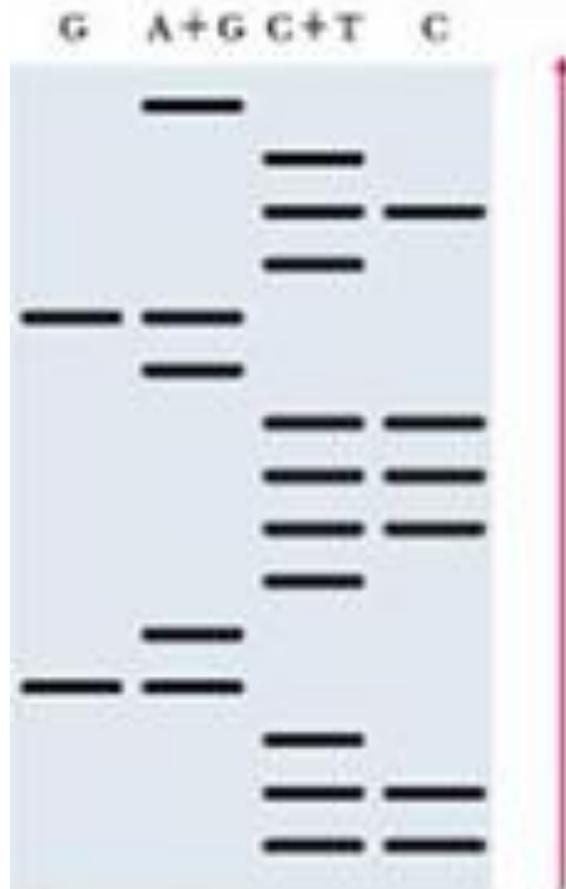


# FUNCIONAMIENTO

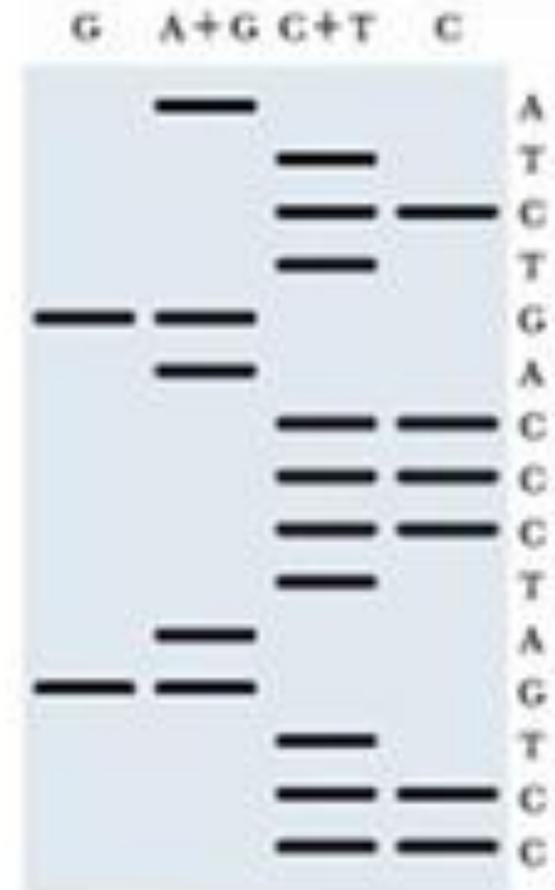
La muestra se divide en cuatro alícuotas y se somete a cuatro reacciones químicas distintas, generando rupturas en el ADN después de la base indicada en cada una de las cuatro reacciones.

Reactivo	Corte en
Hidracina, NaCl	Citocinas
Dimetilsulfato	Guaninas
	Adeninas (Guaninas + Adenina)
Hidracina	Citocinas + Timinas

# EJERCICIO



## Respuesta



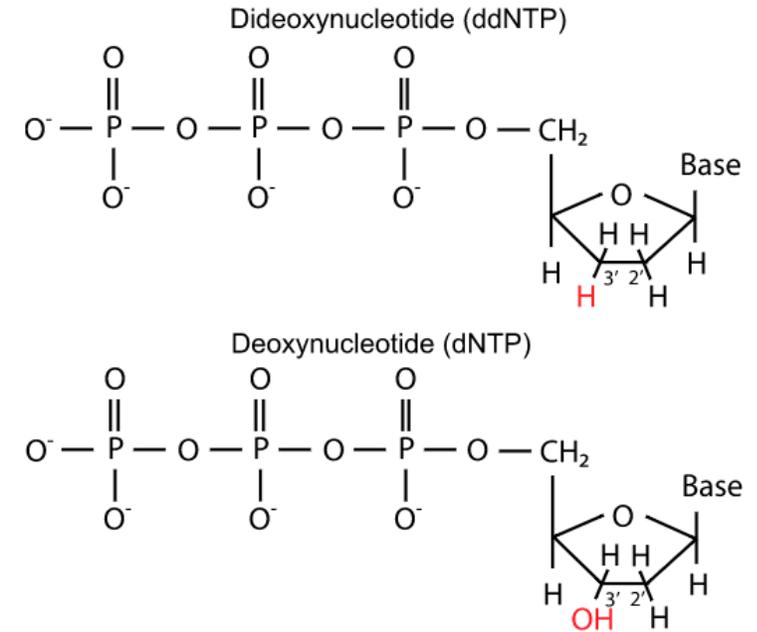
# SANGER - ENZIMÁTICO

1977

Método enzimático de terminación de cadena

Componentes de la reacción

- ADN molde
- Enzima polimerasa que añade nucleótidos por el extremo 3'
- Primer (20 bases y marcado radioactivamente)
- Nucleótidos trifosfato
- Nucleótidos didesoxi marcados (baja concentración)



# MÉTODO SANGER

dATP + dCTP + dGTP + dTTP

DNA polimerasa

ADN molde

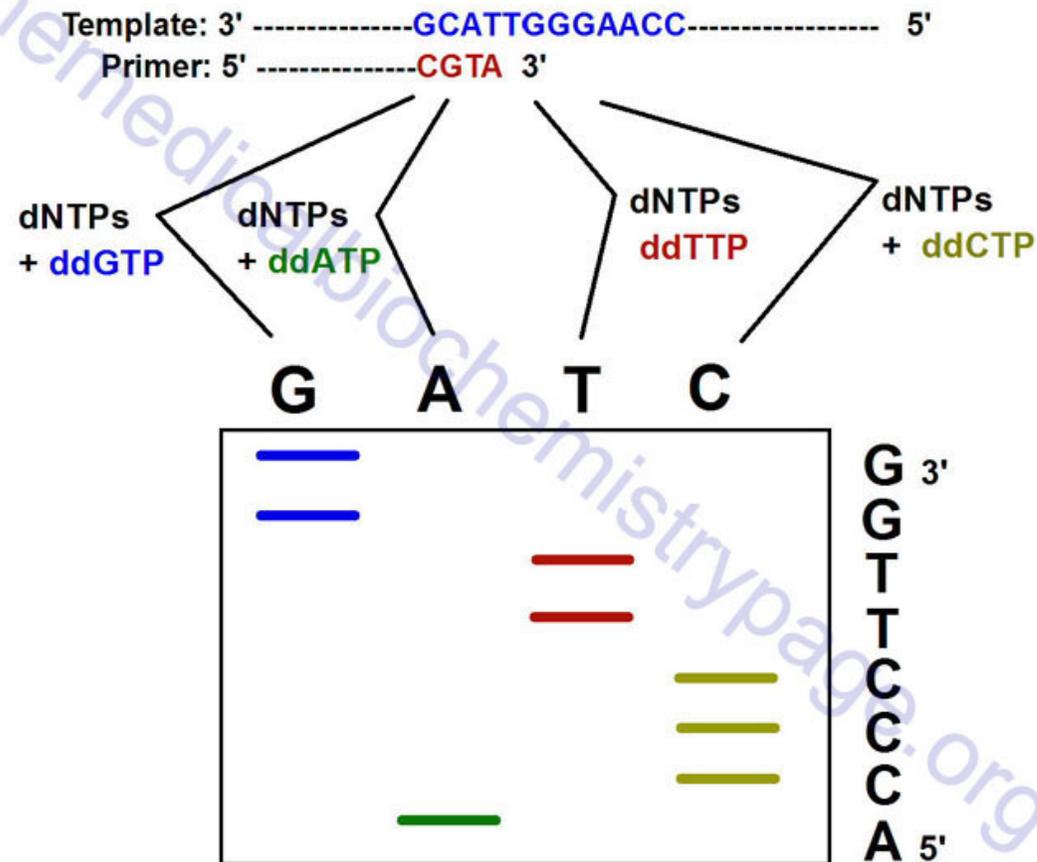
Primer



# MÉTODO SANGER MANUAL

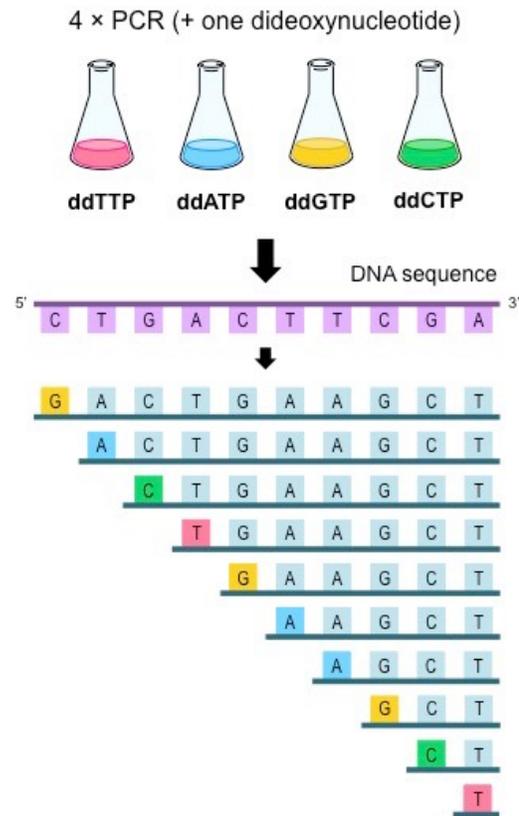
Usa ddNTPs marcados con fluorescencia  
Precisa de una sola reacción, en lugar de 4  
No emplea radioactividad  
La separación por electroforesis capilar mejora el rendimiento

## Sanger ddNTP Chain Termination Sequencing

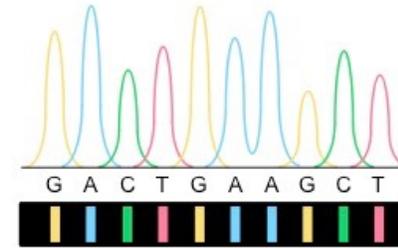


# MÉTODO SANGER AUTOMÁTICO

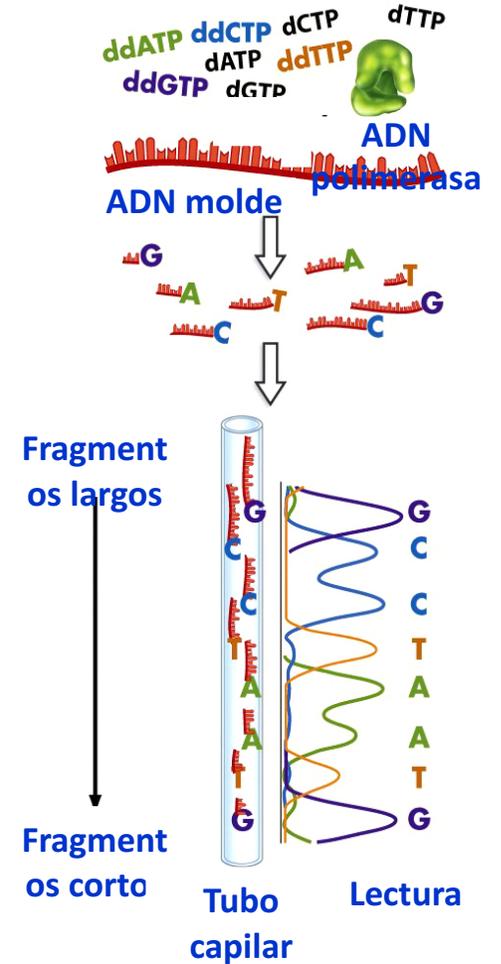
Usa ddNTPs marcados con fluorescencia  
 Precisa de una sola reacción, en lugar de 4  
 No emplea radioactividad  
 La separación por electroforesis capilar mejora el rendimiento



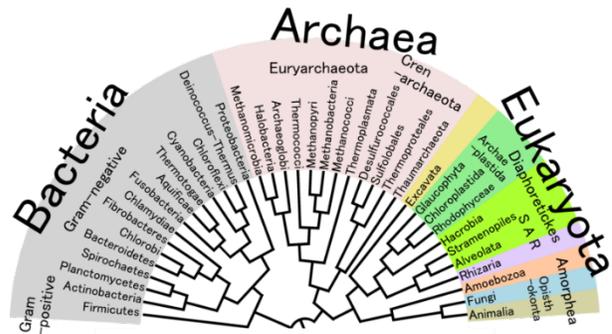
Use a sequencing machine



Separate with a gel



# APORTACIONES CON SANGER



Los tres dominios de la vida



Biblioteca metagenómica



La secuencia completa del genoma humano

# SANGER

Estándares de calidad (Best Practice Guidelines)

## Ventajas

- Baja tasa de error (<0.01%)
- Longitud de la lectura (~1000 pb)

## Desventajas

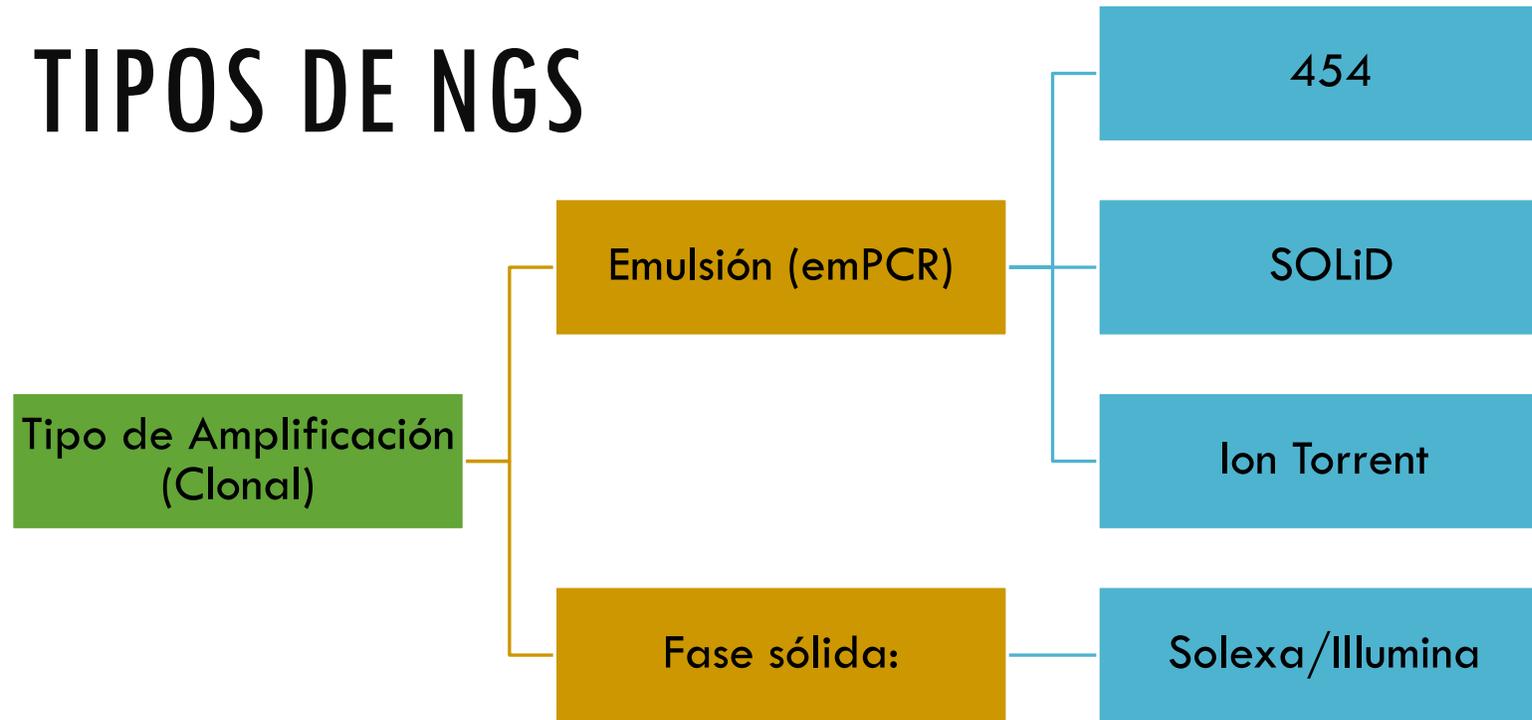
- Alto costo por par de base
- Tiempo largos por corrida
- La necesidad de clonar los fragmento
- Cepas aisladas



**SECUENCIACIÓN DE NUEVA GENERACIÓN**  
**SECUENCIACIÓN MASIVA EN PARALELO**

**NGS**

# TIPOS DE NGS

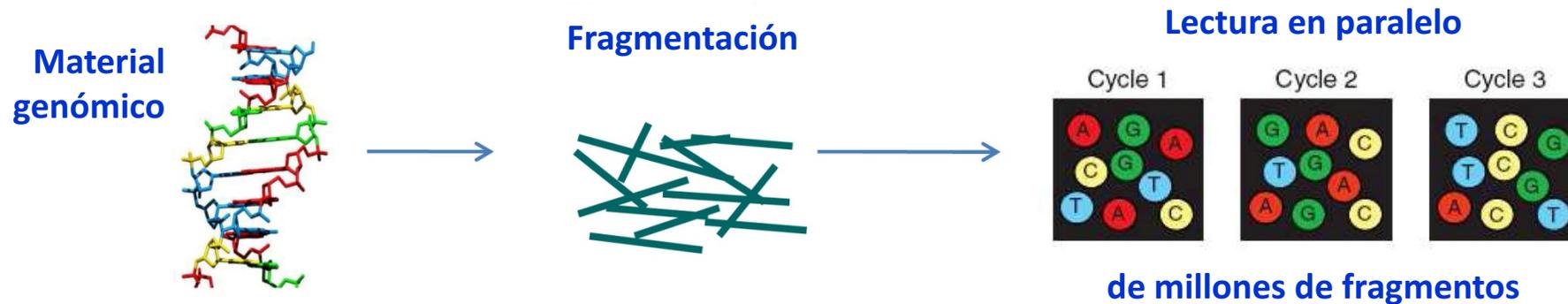


El material genómico se fragmenta y los distintos fragmentos se secuencian simultáneamente, generando millones de lecturas en una sola reacción

# Secuenciación masiva en paralelo

## Secuenciación masiva (2ª y 3ª generación):

El material genómico se fragmenta y los distintos fragmentos se secuencian simultáneamente, generando millones de lecturas en una sola reacción

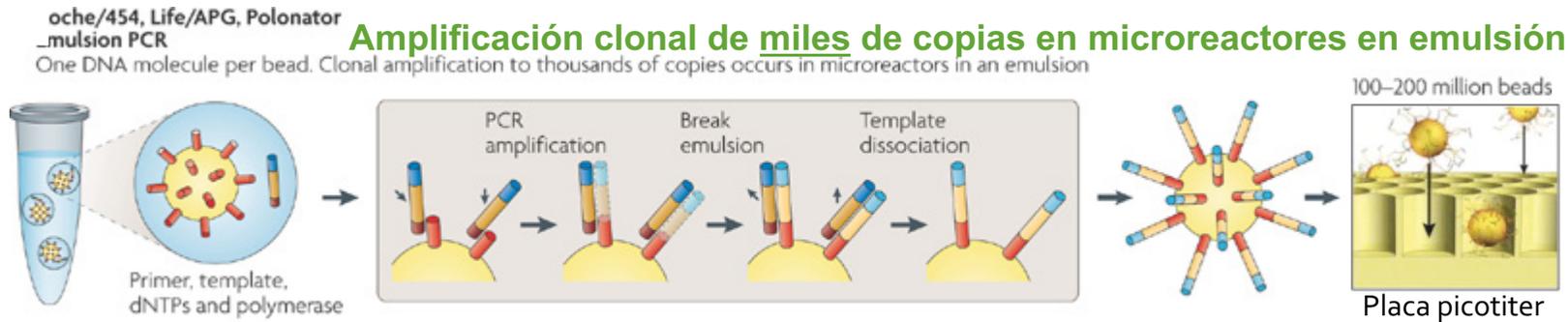


# Plataformas de secuenciación masiva

## Métodos de preparación de las muestras:

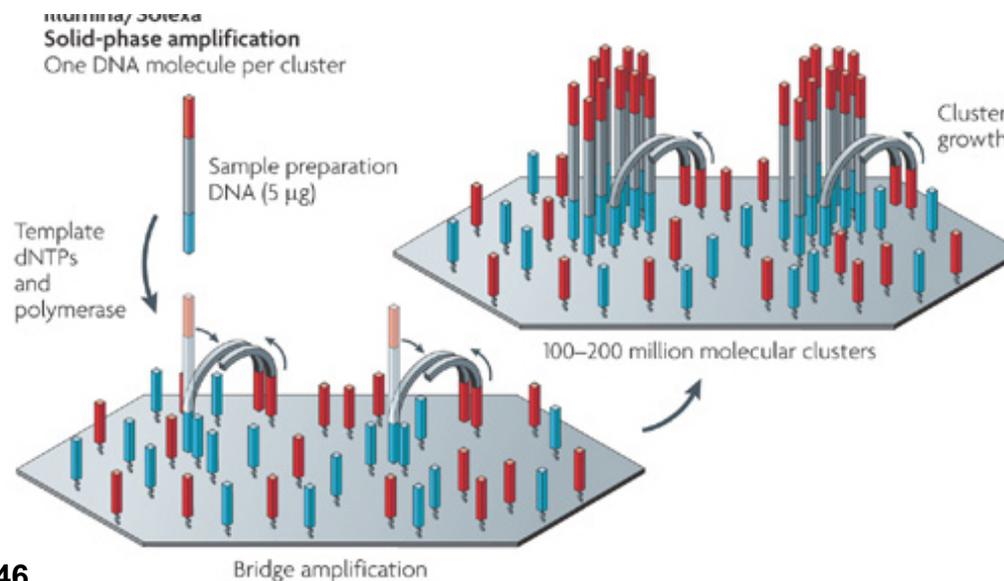
### Clonales → requieren PCR

- emPCR: Roche GS-FLX, Applied Biosystems SOLiD 5500, Lifetech Ion-Torrent

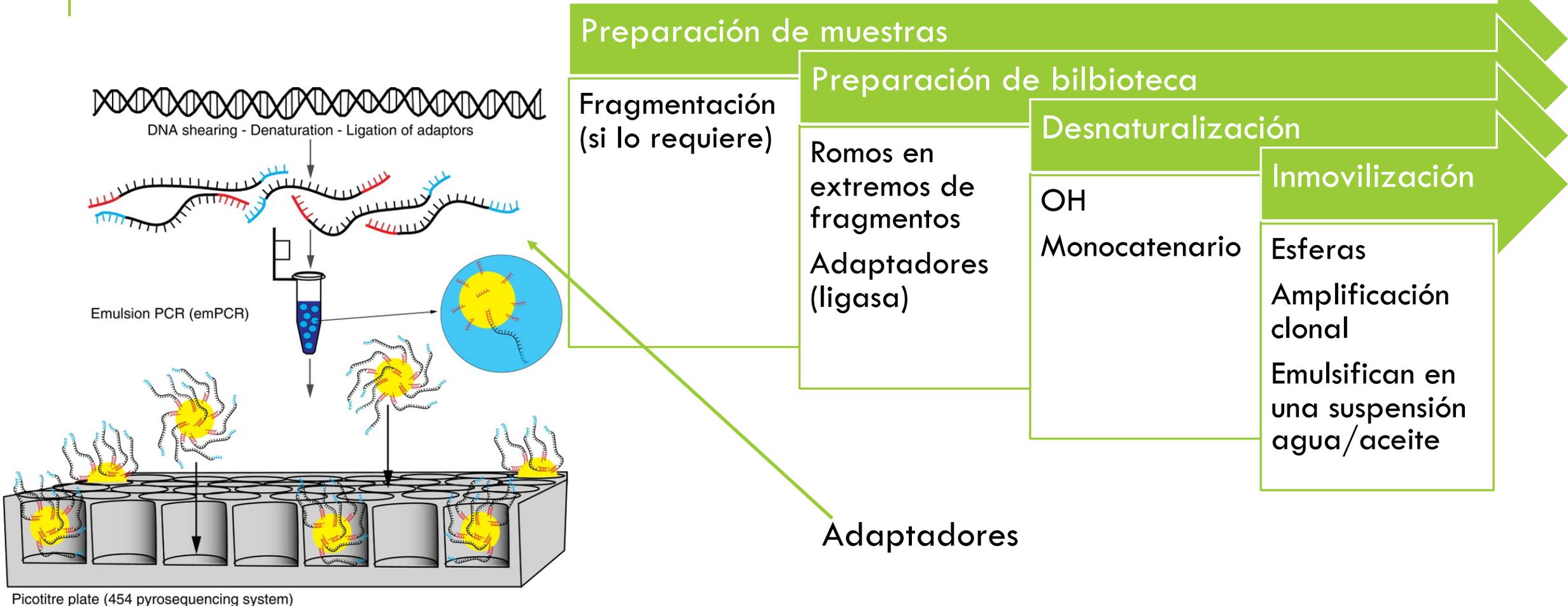


- Amplificación-puente en fase sólida: **Illumina**

**Una molécula de ADN por cluster, 100-200 millones de clusters moleculares**



# PCR EN EMULSIÓN



Metzker, ML (2010) Sequencing technologies- the next generation. Nature Rev Genetics 11:31-46

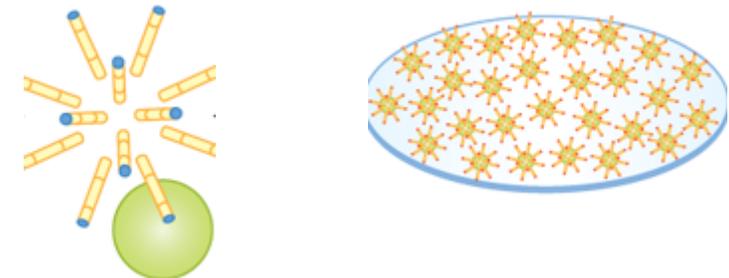
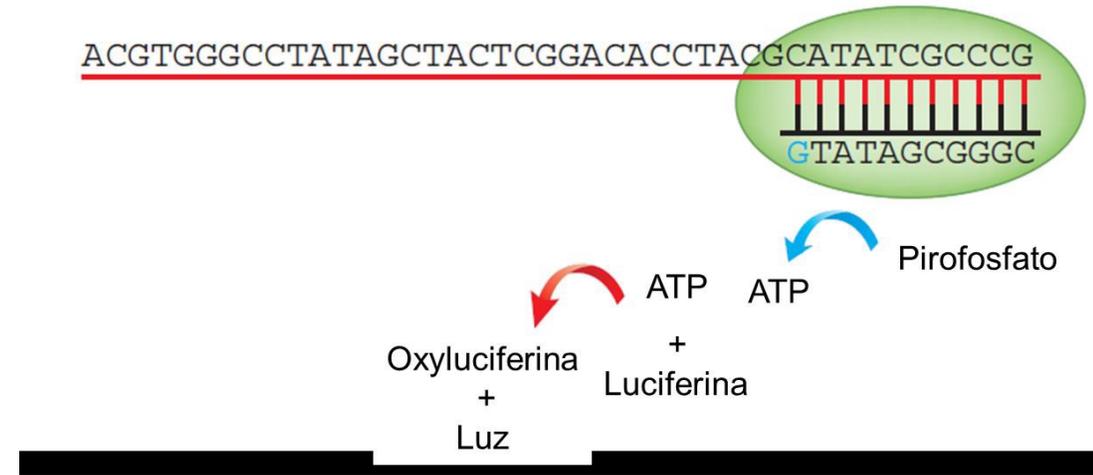
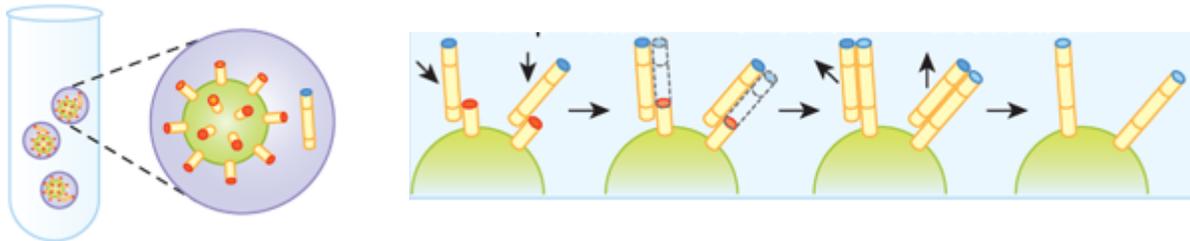
# emPCR - Pirosecuenciación

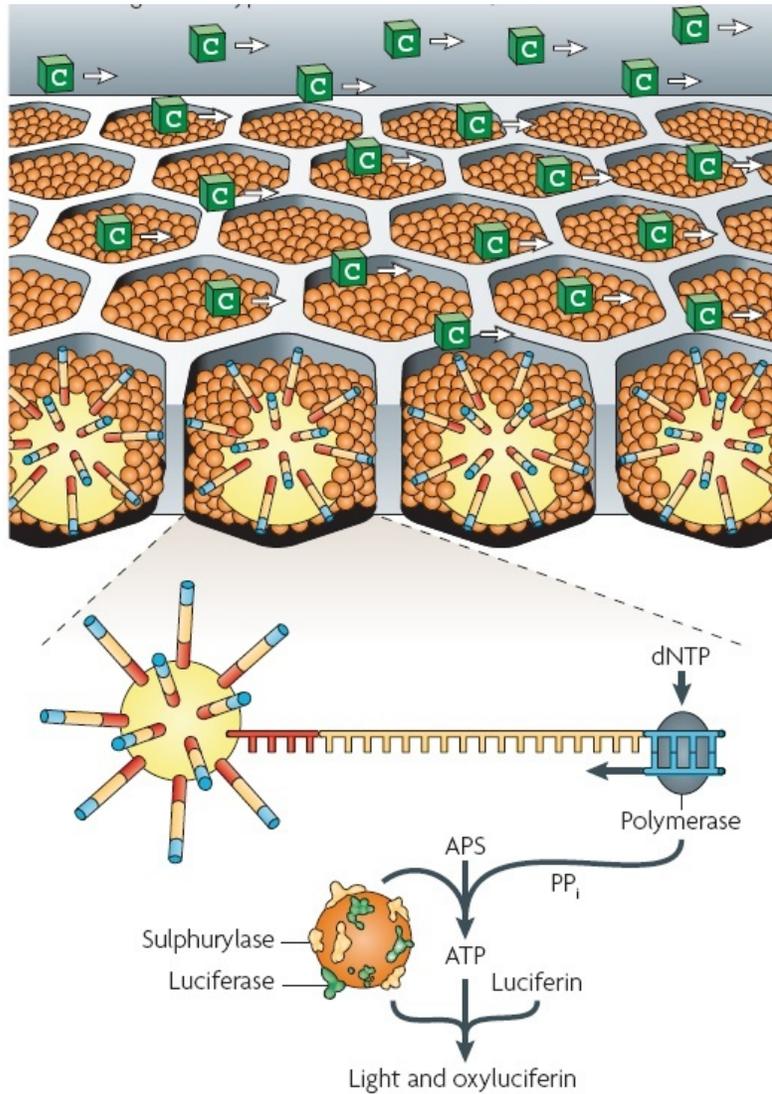
Plataforma 454 de Roche

## Secuenciación por síntesis

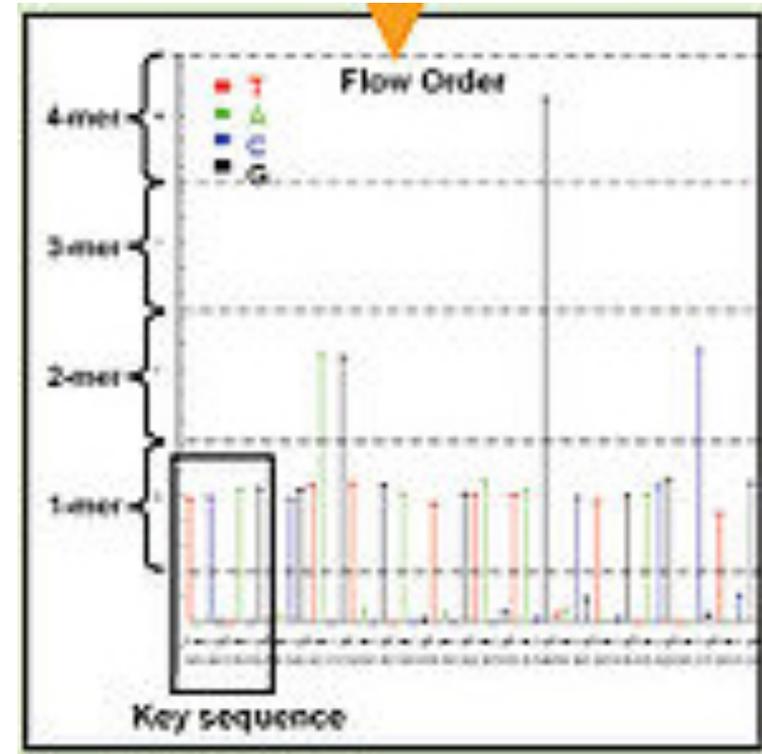
Usa una reacción luminosa para la detección del nucleótido incorporado en la síntesis de la cadena complementaria.

Tecnología no utilizada





Millones de amplificaciones

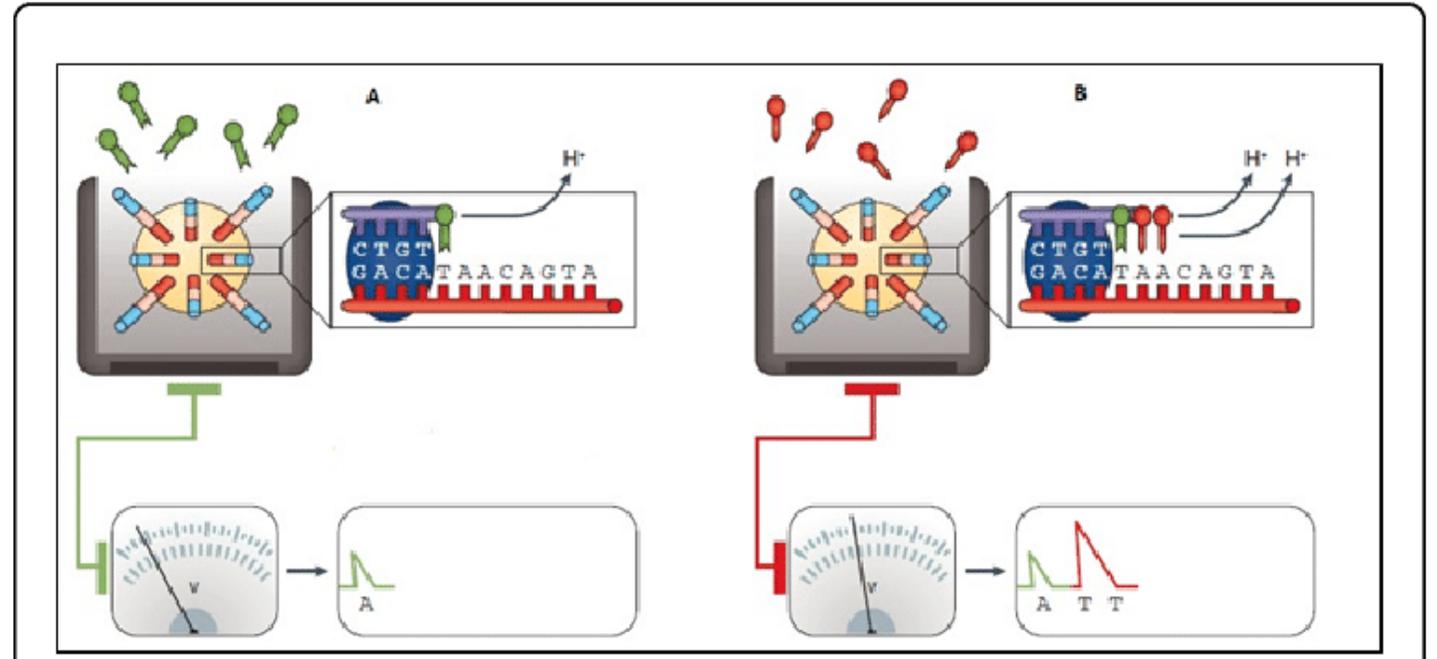


Se traduce a un programa

Metzker, ML (2010) Sequencing technologies- the next generation. Nature Rev Genetics 11:31-46

# ION TORRENT (SEMICONDUCCIÓN)

- Basado en la detección de protones liberados durante el proceso de polimerización del ADN.
- No usan nucleótidos modificados químicamente y
- Detección no se realiza por métodos ópticos, sino por detección de pH.



# SECUENCIACIÓN EN FASE SÓLIDA

# ILLUMINA

1. Concentración y tamaño del fragmento
2. Purificación
3. Index PCR
4. Purificación
5. Agrupación
6. Ajuste de concentración de bibliotecas
7. Desnaturalización de bibliotecas.

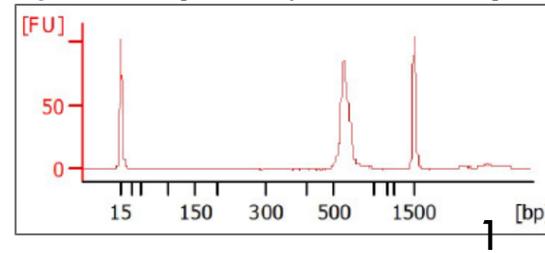


Figure 5 Example Bioanalyzer Trace of Final Library

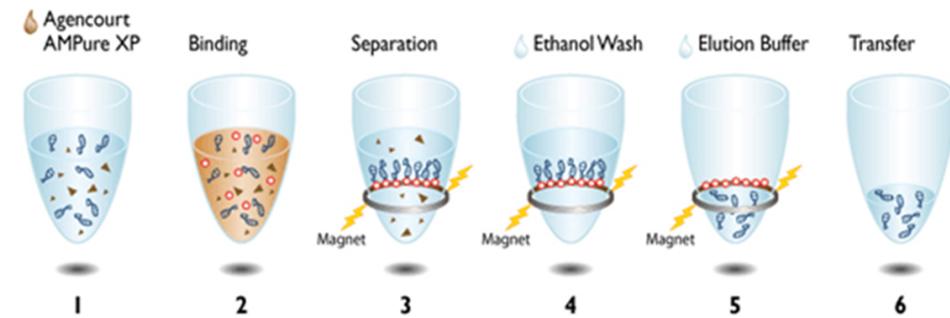
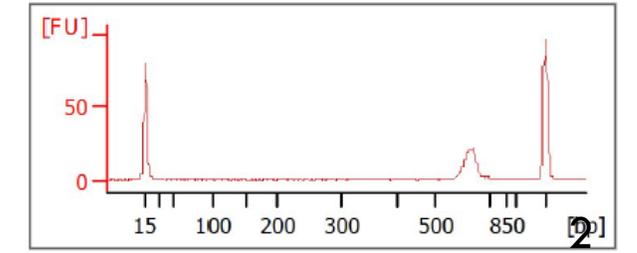
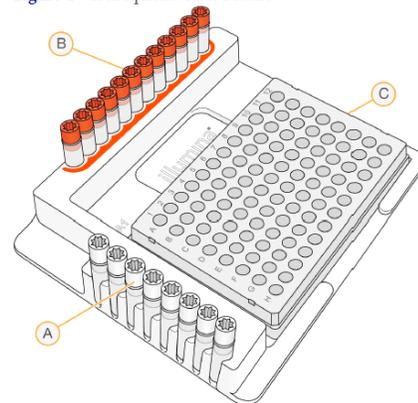
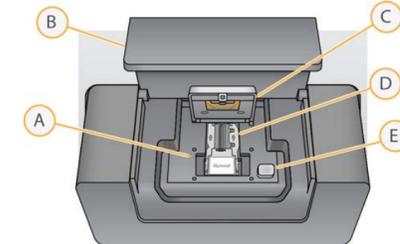


Figure 4 TruSeq Index Plate Fixture



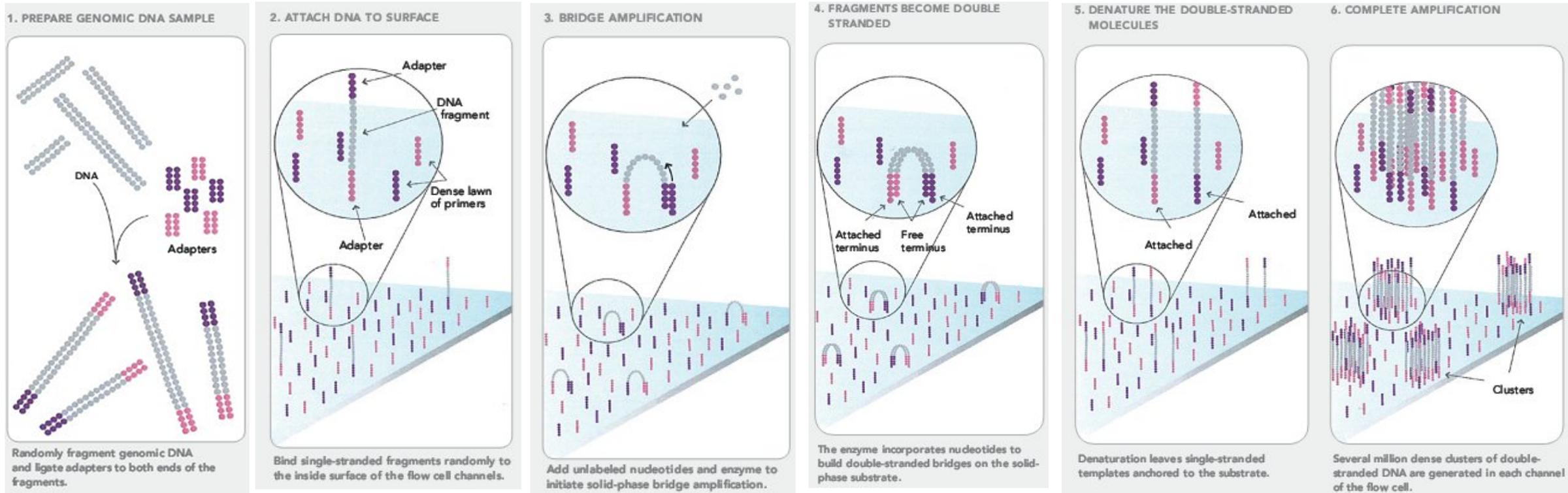
- A Index 2 primers (white caps)
- B Index 1 primers (orange caps)
- C 96-well plate

Compartimento de la celda de flujo



- A Platina de la celda de flujo
- B Puerta del compartimento de la celda de flujo
- C Cierre de la celda de flujo
- D Celda de flujo
- E Botón de apertura del cierre de la celda de flujo

# ILLUMINA



Se incorporan los 4 nucleótidos, marcados con diferentes fluorocromos

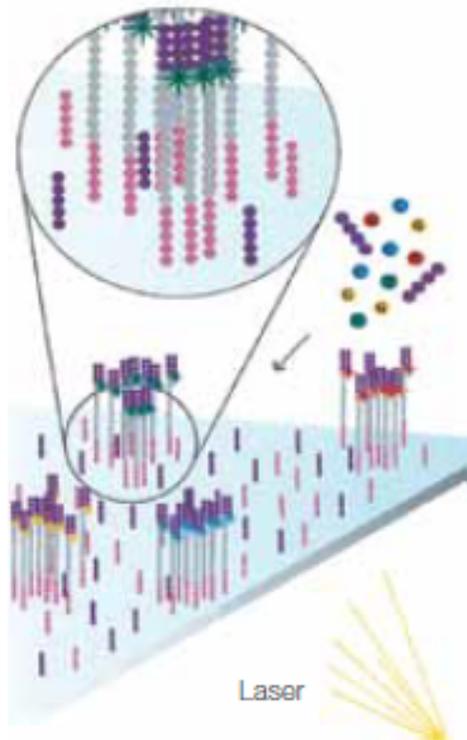
Lavado de los nucleótidos no unidos

Toma de imágenes

Se eliminan los fluorocromos con TCEP

Repetición de ciclos

Figure 8: Determine First Base

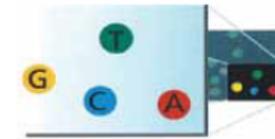


The first sequencing cycle begins by adding four labeled reversible terminators, primers, and DNA polymerase.

Se agregan los dNTPs, cada uno posee un fluoróforo de color particular:

- G - Amarillo.
- C – Azul
- A – Rojo
- T - Verde

Figure 9: Image First Base

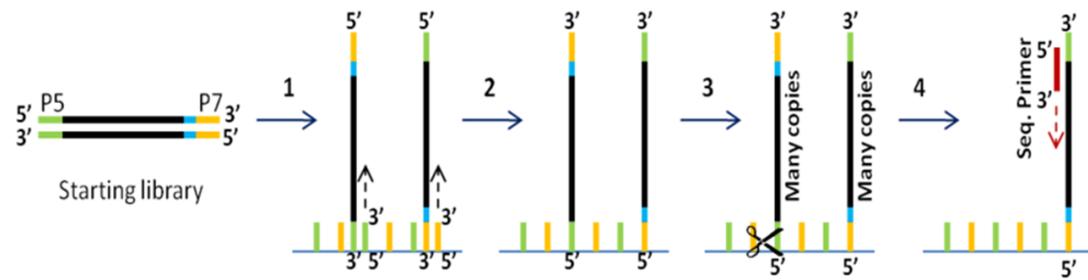


After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.

Se excita con un laser y se captura una imagen.  
La imagen capturada en una posición específica corresponde al primer nucleótido.

# I. READ 1

For Single- and Paired-End Sequencing



# III. READ 2

For Paired-End Sequencing Only

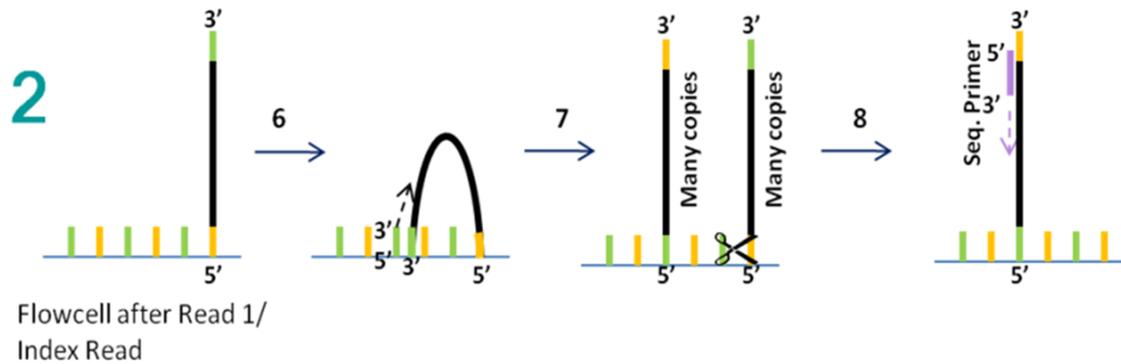
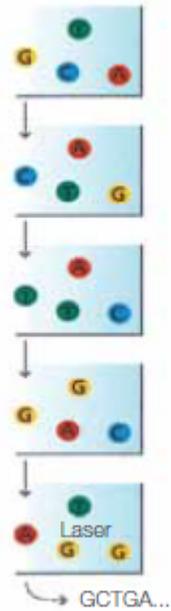


Figure 12: Sequencing Over Multiple Chemistry Cycles



The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time.

Los ciclos de secuenciación se repiten hasta determinar la secuencia del ADN

# APLICACIONES



**Metagenómica  
16S**

**Metaproteómica**



**¿Qué microorganismos conforman una muestra ambiental?**

**Metatranscriptómica**

**¿Qué hacen?**

**¿Cómo lo hacen?**

**¿Quién lo hace?**

**Metagenómica**



# SECUENCIACIÓN SMRT (TERCERA GENERACIÓN)

- Single Molecule Real Time
- Pacific Biosciences (PacBio)
- No requiere de amplificación
- Alto costo
- No hay en México
- Secuencia 1 000 a 1 200 bases.
- Secuenciación genómica de novo
- Diagnóstico in vitro



# CARACTERÍSTICAS DE LAS PLATAFORMAS DE SECUENCIACIÓN

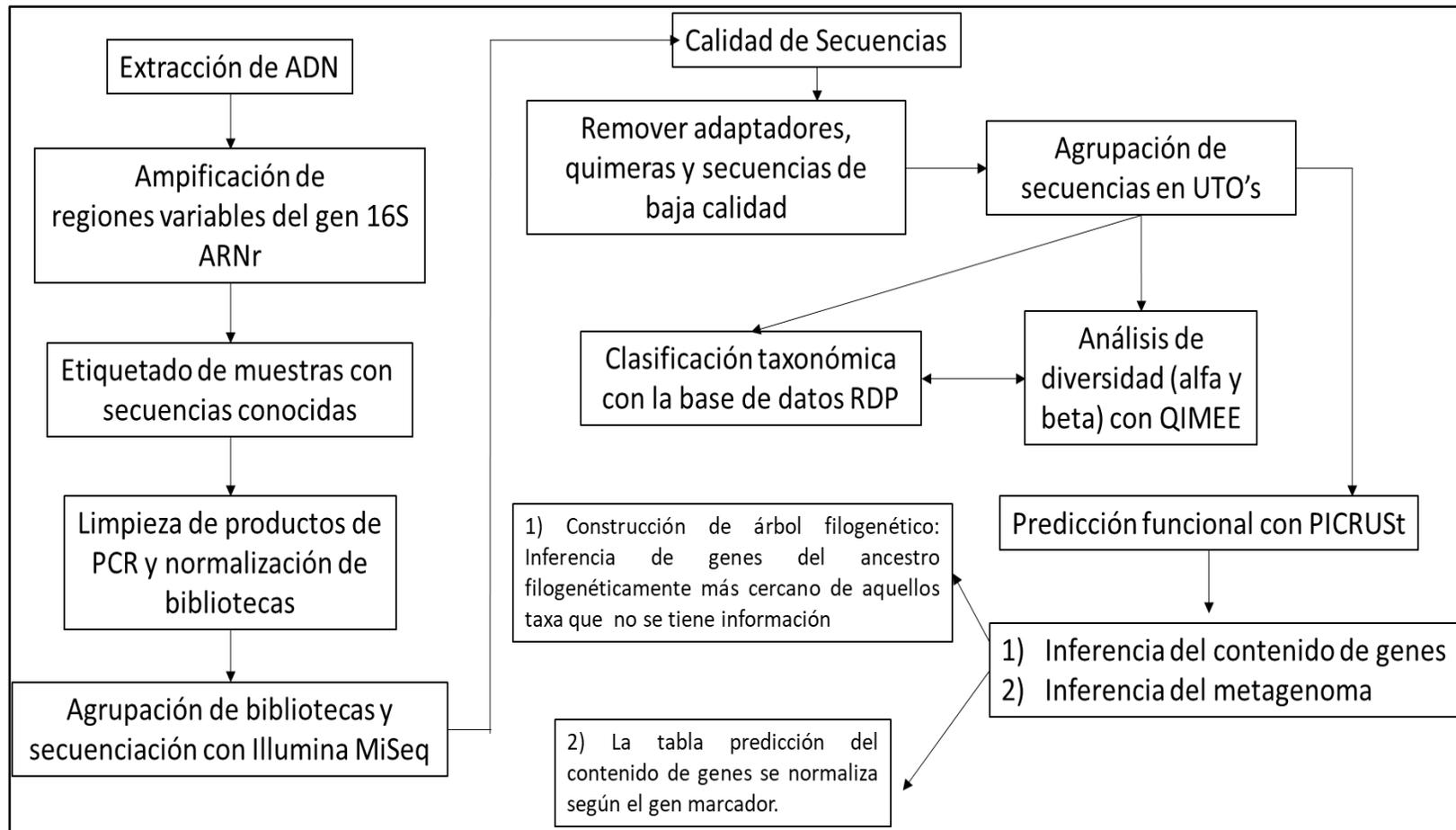
Segunda generación

Tercera generación

Método	Amplificación de las muestras	Tecnología de secuenciación	Longitud de lectura	Rendimiento
Roche GS-FLX Titanium XL	emPCR	Por síntesis	<b>1000 bp</b>	<b>700 Mb/día</b>
Lifetech Ion-Torrent ION PGM	emPCR	Por síntesis	400 bp	Hasta 8 Gb/día
Lifetech Ion-Torrent ION Proton	emPCR	Por síntesis	200 bp	Hasta 60 Gb/día
Illumina Hiseq 2500	En fase sólida	Por síntesis	2 x 50-150 bp	> 80 Gb/día
Illumina Miseq	En fase sólida	Por síntesis	<b>2 x 50-250 bp</b>	<b>&gt; 5 Gb/día</b>
Pacific Biosciences	Moléculas individuales	Por síntesis	1000 bp	-
Helicos HeliScope	Moléculas individuales	Por síntesis	35 bp	> 8 Gb/día
Oxford Nanopore	Moléculas individuales	Por degradación	Sin límite teórico	10 <sup>7</sup> bp/poro × segundo



# COMO ANALIZAR LAS SECUENCIAS?



RDP  
MG RAST

# Composition, diversity and function of intestinal microbiota in pacific white shrimp (*Litopenaeus vannamei*) at different culture stages

Shenzheng Zeng<sup>1,\*</sup>, Zhijian Huang<sup>1,\*</sup>, Dongwei Hou<sup>1</sup>, Jian Liu<sup>1</sup>,  
Shaoping Weng<sup>2</sup> and Jianguo He<sup>1,2</sup>

<sup>1</sup> State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Marine Resources and Coastal Engineering, School of Marine Sciences, Sun Yat-sen University, Guangzhou, China

<sup>2</sup> School of Life Sciences, Sun Yat-sen University, Guangzhou, China

\* These authors contributed equally to this work.



# RESUMEN

## ABSTRACT

Intestinal microbiota is an integral component of the host and plays important roles in host health. The pacific white shrimp is one of the most profitable aquaculture species commercialized in the world market with the largest production in shrimp consumption. Many studies revealed that the intestinal microbiota shifted significantly during host development in other aquaculture animals. In the present study, 22 shrimp samples were collected every 15 days from larval stage (15 day post-hatching, dph) to adult stage (75 dph) to investigate the intestinal microbiota at different culture stages by targeting the V4 region of 16S rRNA gene, and the microbial function prediction was conducted by PICRUSt. The operational taxonomic unit (OTU) was assigned at 97% sequence identity. A total of 2,496 OTUs were obtained, ranging from 585 to 1,239 in each sample. Forty-three phyla were identified due to the classifiable sequence. The most abundant phyla were Proteobacteria, Cyanobacteria, Tenericutes, Fusobacteria, Firmicutes, Verrucomicrobia, Bacteroidetes, Planctomycetes, Actinobacteria and Chloroflexi. OTUs belonged to 289 genera and the most abundant genera were *Candidatus\_Xiphinematobacter*, *Propionigenium*, *Synechococcus*, *Shewanella* and *Cetobacterium*. Fifty-nine OTUs were detected in all samples, which were considered as the major microbes in intestine of shrimp. The intestinal microbiota was enriched with functional potentials that were related to transporters, ABC transporters, DNA repair and recombination proteins, two component system, secretion system, bacterial motility proteins, purine metabolism and ribosome. All the results showed that the intestinal microbial composition, diversity and functions varied significantly at different culture stages, which indicated that shrimp intestinal microbiota depended on culture stages. These findings provided new evidence on intestinal microorganism microecology and greatly enhanced our understanding of stage-specific community in the shrimp intestinal ecosystem.

# INTRODUCCIÓN

## Microbiota intestinal



Organismos acuáticos de importancia comercial



## Camarón blanco



Pérdidas económicas por enfermedades



Enfermedades asociadas a Disbiosis (alteración en la microbiota intestinal)

Funciones que incluyen en la salud del hospedero



Relación Microbiota:Hospedero?



¿Cómo se comporta la microbiota a medida que el hospedero se desarrolla?

Secuenciación del gel 16S ARNr : Diversidad y funcionalidad

# OBJETIVO

Evaluar la diferencias en la microbiota intestinal a diferentes estadios de cultivo.

Comparar la composición, diversidad y funciones de la microbiota intestinal en camarón blanco en sus diferentes etapas de desarrollo.



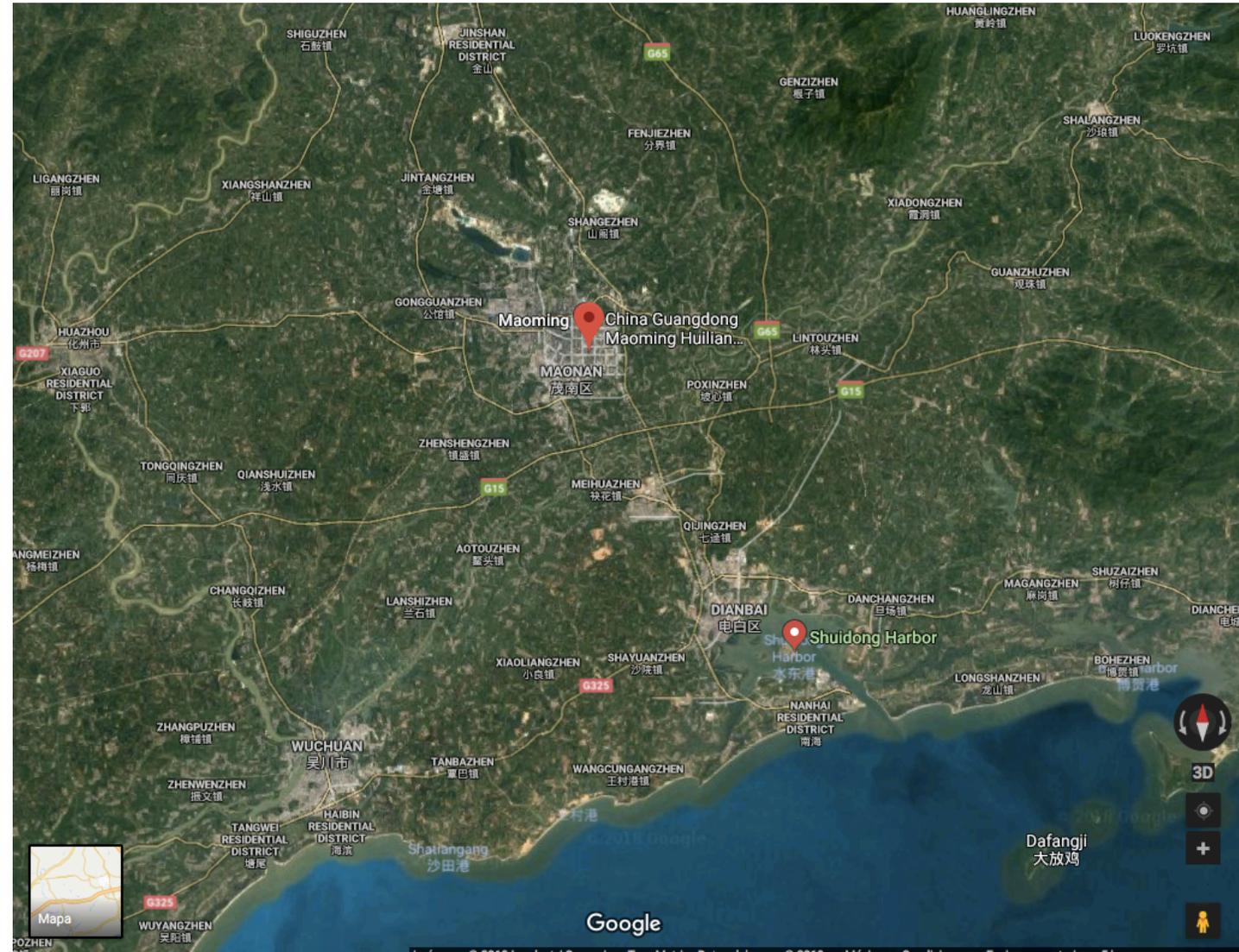
# MATERIALES Y MÉTODOS

Colecta de muestra

Análisis de datos

Predicción funcional de la microbiota basado en el gen ARNr 16s

# 22 MUESTRAS DE INTESTINO DE CAMARÓN DE 5 TANQUES DESDE JULIO A OCTUBRE DE 2015 - CADA 15 DIAS



# TRATAMIENTO DE LA MUESTRA

Disección del intestino



Esterilización de la superficie del camarón con etanol 70%



- ✓ Perlas de vidrio esteriles
- ✓ 1.5 mL de buffer PBS



Vortex por 3 minutos



10,000 g x 1 min



-80°C

# EXTRACCIÓN Y SECUENCIACIÓN DE ADN

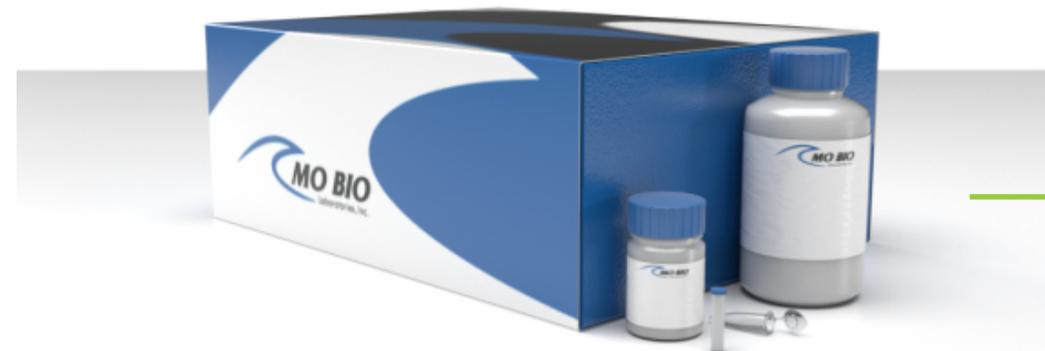
Preparación de librerías de secuencias con kit TruSeq DNA PCR-Free



Cuantificación de la librería con Qubit 2.0 Fluorometer



Secuenciación en Illumina HiSeq 2500 Platform



PowerFecal DNA Isolation Kit

Amplificación region V4 del gen ARNr 16s



# ANÁLISIS DE DATOS

## Anotación taxonómica



Indices de  
diversidad

**UCHIME**

Remover  
secuencias  
quimera

## Información taxonómica



Abundancia de  
OTUs

## Perfil funcional de la comunidad

**PICRUSt**



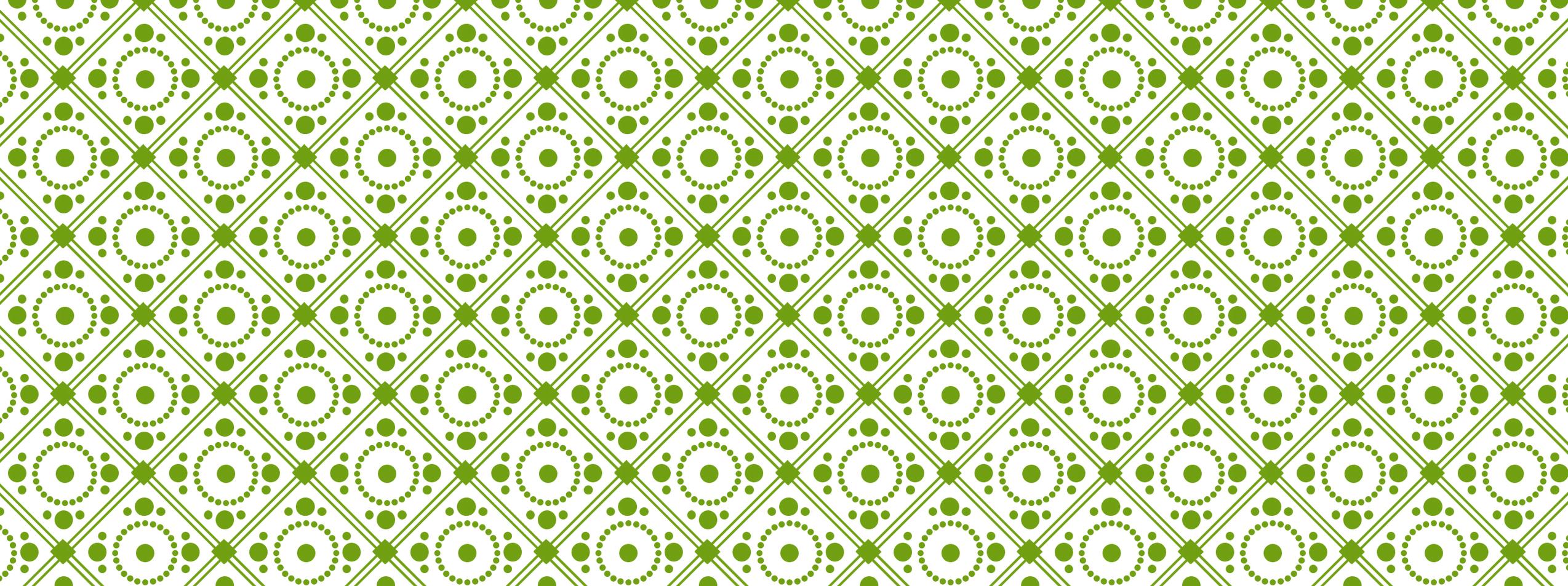
## Análisis estadístico



ANOVA de 1 vía



- ✓ Procedimiento de permutación de respuesta múltiple
- ✓ PerMANOVA



# RESULTADOS Y DISCUSIÓN



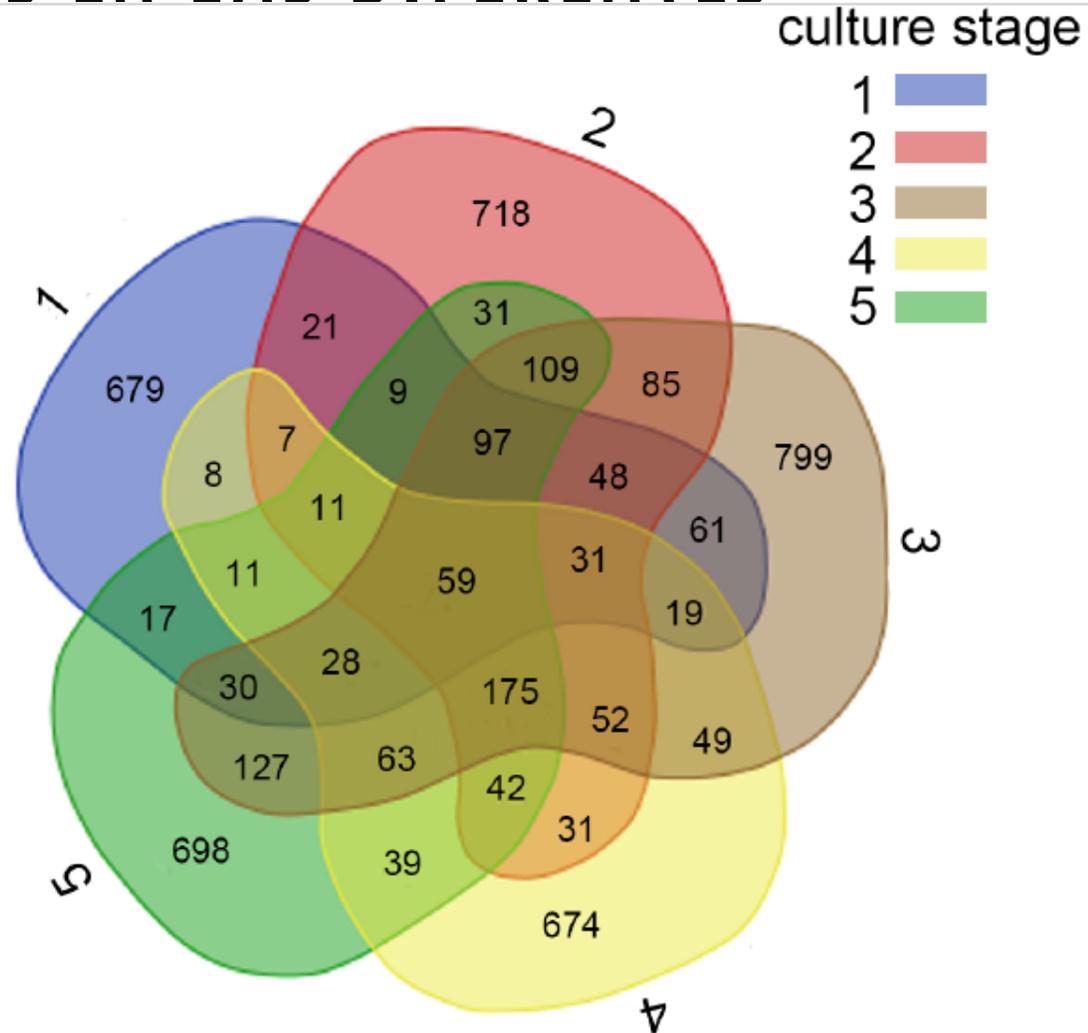
INFORMACIÓN DE SECUENCIACIÓN Y OTU'S. RESUMEN DEL ANÁLISIS DE LECTURAS DE SECUENCIAS, NÚMERO DE OTU'S Y NÚMERO DE OTU'S QUE PUEDEN CLASIFICARSE EN DIFERENTES NIVELES TAXONÓMICOS. A, B, C, D Y E MUESTRA LAS TINAS Y LOS NÚMEROS SON EL ESTADIO.

**Table 1 Sequencing and OTU classification information.** Summary of sequencing read analysis, numbers of OTUs, and numbers of OTUs that can be classified into different levels (phylum, class, order, family and genus). A, B, C, D and E stand for the ponds. 1, 2, 3, 4 and 5 stand for the culture stages.

Sample	Trimmed tags	OTUs	Phylum	Class	Order	Family	Genus
A1	73,996	645	18	46	78	103	73
B1	64,306	591	20	48	84	109	83
C1	64,282	817	21	54	100	131	115
D1	66,164	622	14	41	69	95	69
E1	70,365	737	17	44	70	97	77
A2	61,460	1,169	32	74	121	161	147
D2	70,166	1,170	26	65	105	140	132
E2	63,705	910	18	51	94	105	84
A3	61,660	976	28	64	112	134	108
B3	62,061	971	25	59	101	132	117
C3	38,103	993	31	72	120	153	127
D3	69,588	1,150	27	63	113	154	147
E3	60,883	947	21	50	91	112	91
A4	45,440	978	29	66	101	134	132
B4	38,348	585	22	48	85	109	84
C4	40,220	643	19	46	77	101	83
D4	62,236	722	18	47	81	105	98
E4	61,853	708	18	46	77	100	86
A5	57,143	1,075	28	68	111	143	122
C5	66,736	1,239	27	66	111	156	143
D5	68,803	757	17	40	74	104	87
E5	67,102	972	25	53	90	114	90



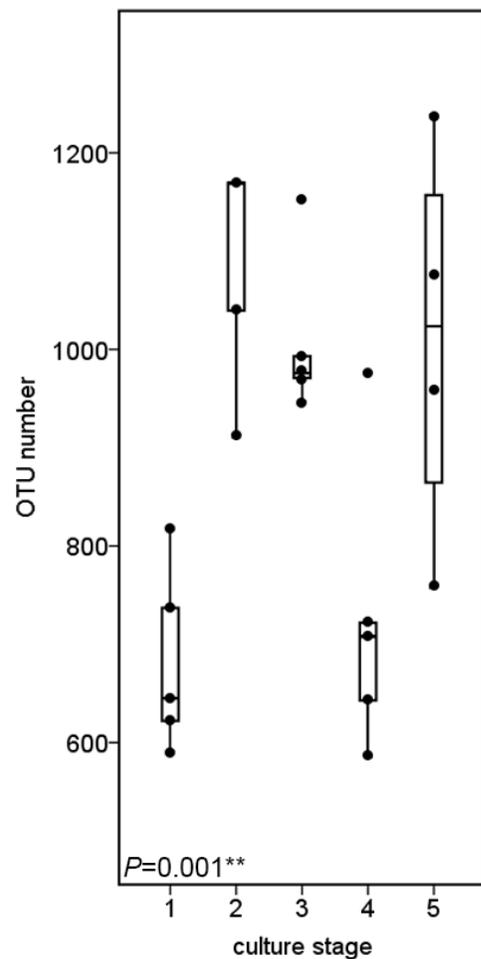
# ANÁLISIS DE OTU'S COMPARTIDAS DE DIFERENTES BIBLIOTECAS. EL DIGRAMA VENN MUESTRA OTU'S ÚNICAS COMPARTIDAS EN LAS DIFERENTES BIBLIOTECAS.





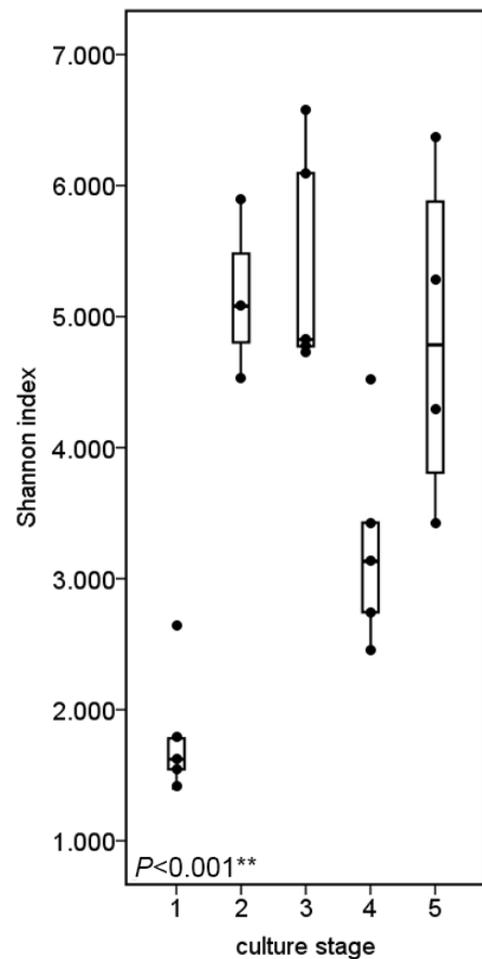
A

Número de OTU



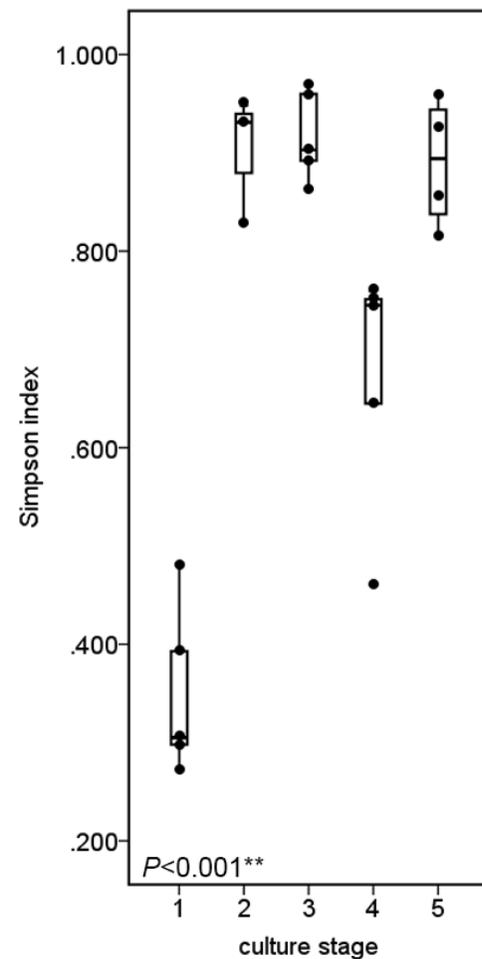
B

Índice de Shannon



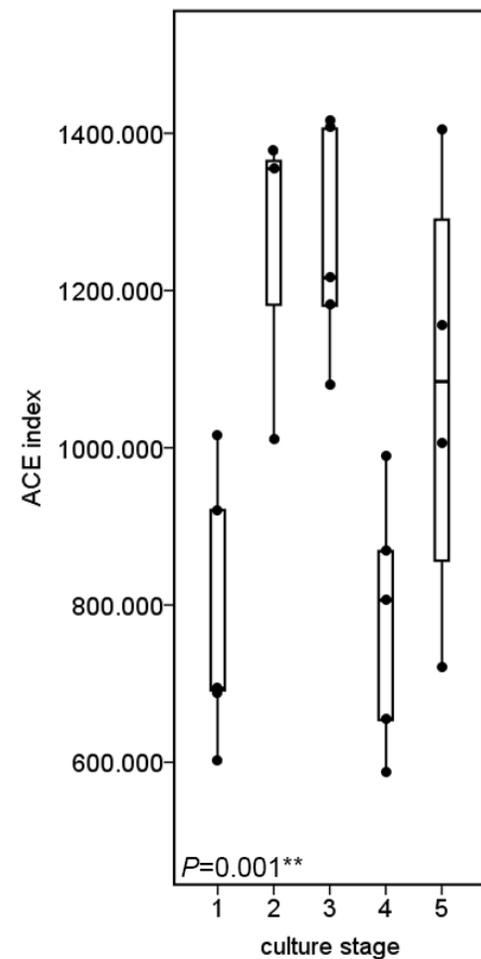
C

Índice de Simpson



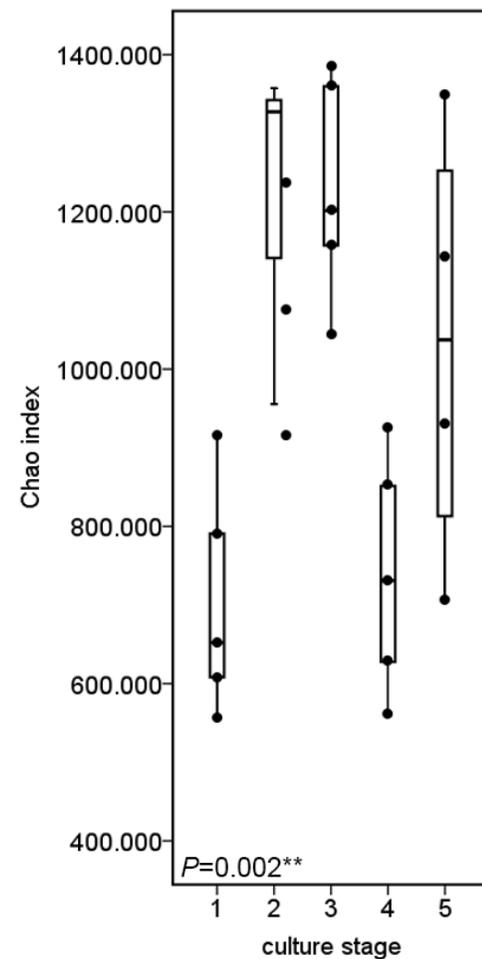
D

Índice de ACE

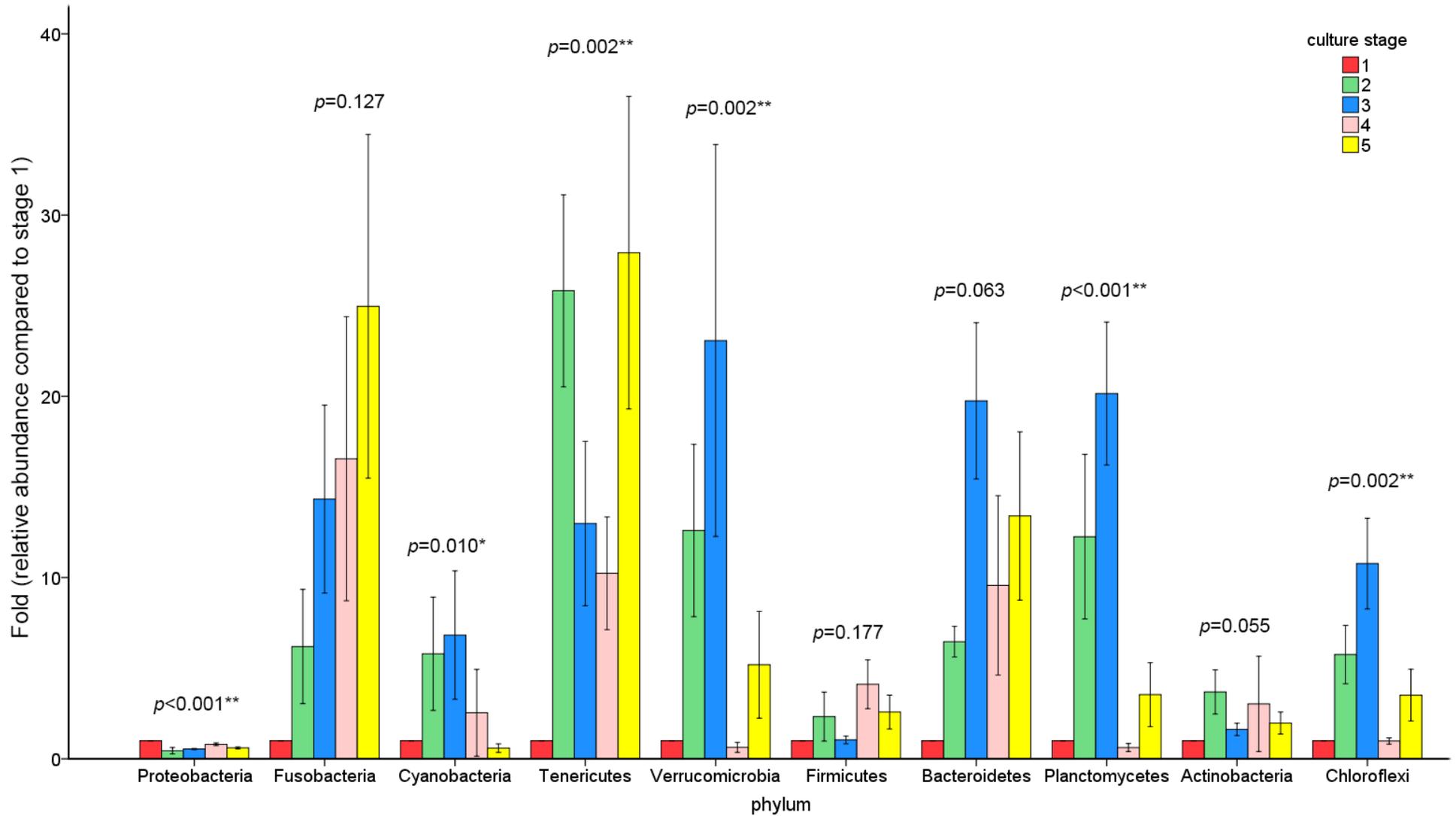


E

Índice de Chao



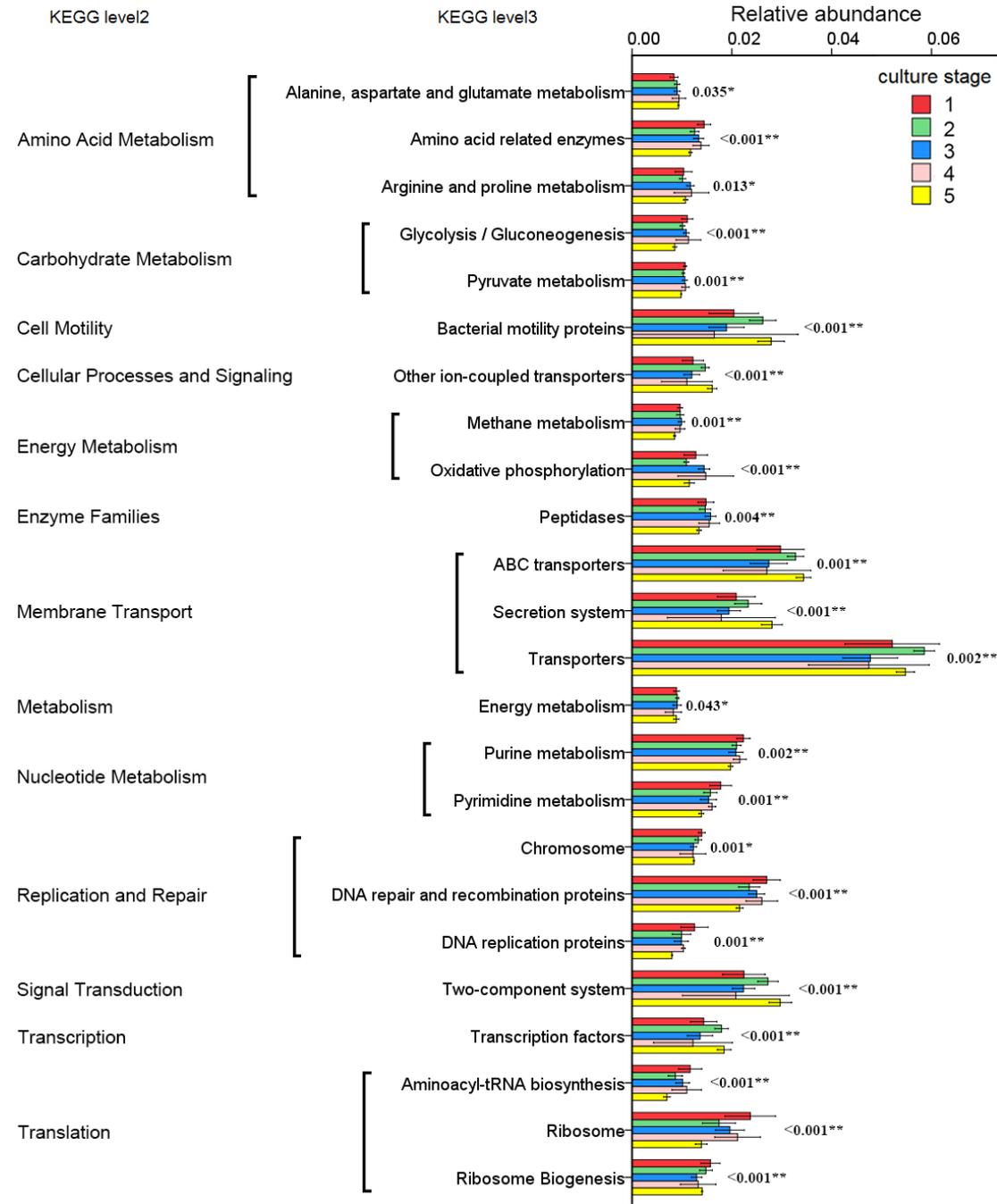
# ABUNDANCIA DE LOS TOP 10 FILOS A DIFERENTES ETAPAS DE CULTIVO



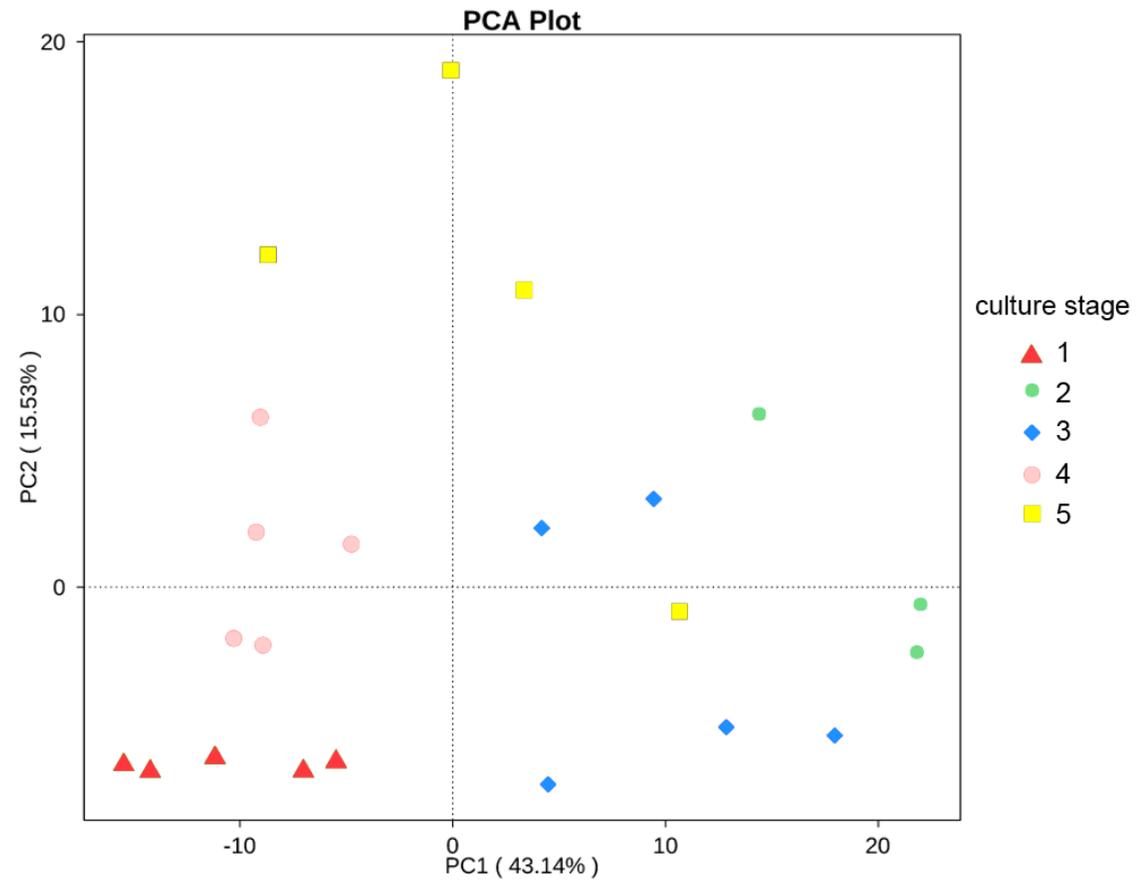
**Table 4 Relative abundance of predicted functions.** KOs in KEGG level3 are listed following the relative average abundance. The KEGG level 2 is also listed.

KOs	KEGG level 2	Relative abundance
Transporters	Membrane transport	4.25 to 6.06
ABC transporters	Membrane transport	2.46 to 3.58
DNA repair and recombination proteins	Replication and repair	2.09 to 2.89
Two component system	Signal transduction	1.79 to 3.16
Secretion system	Membrane transport	1.50 to 2.99
Bacterial motility proteins	Cell motility	1.23 to 3.02
Purine metabolism	Nucleotide metabolism	1.89 to 2.32
Function unknown	Poorly characterized	1.64 to 2.58
Ribosome	Translation	1.28 to 2.72
Pyrimidine metabolism	Nucleotide metabolism	1.32 to 1.94
Peptidases	Enzyme families	1.30 to 1.73
Transcription factors	Transcription	1.03 to 1.97
Ribosome biogenesis	Translation	1.21 to 1.7
Oxidative phosphorylation	Energy metabolism	1.04 to 1.66
Amino acid related enzymes	Amino acid metabolism	1.14 to 1.53
Other ion-coupled transporters	Cellular processes and signaling	0.97 to 1.69
Chromosome	Replication and repair	1.14 to 1.45
Arginine and proline metabolism	Amino acid metabolism	0.94 to 1.28
Chaperones and folding catalysts	Folding, sorting and degradation	1.01 to 1.10
Glycolysis/Gluconeogenesis	Carbohydrate metabolism	0.82 to 1.21
Pyruvate metabolism	Carbohydrate metabolism	0.98 to 1.11
Amino sugar and nucleotide sugar metabolism	Carbohydrate metabolism	0.88 to 1.10
DNA replication proteins	Replication and repair	0.79 to 1.45
Aminoacyl-tRNA biosynthesis	Translation	0.64 to 1.32
Carbon fixation pathways in prokaryotes	Energy metabolism	0.87 to 1.08
Methane metabolism	Energy metabolism	0.84 to 1.06

# PREDICCIÓN FUNCIONAL



# PCA MUESTRA LA SIMILITUD DE LAS MUESTRAS



# CONCLUSIÓN

- Se aporta gran cantidad de información sobre la microbiota intestinal de camarón blanco, sobre su composición, así como los microorganismos que dominaron en todos los estadios. De igual forma, aumenta el conocimiento sobre la microbiota estadio específica .
- Se requieren más estudios para conocer la relación entre los cambios en la microbiota a diferentes estadios de cultivo y la salud del camarón.