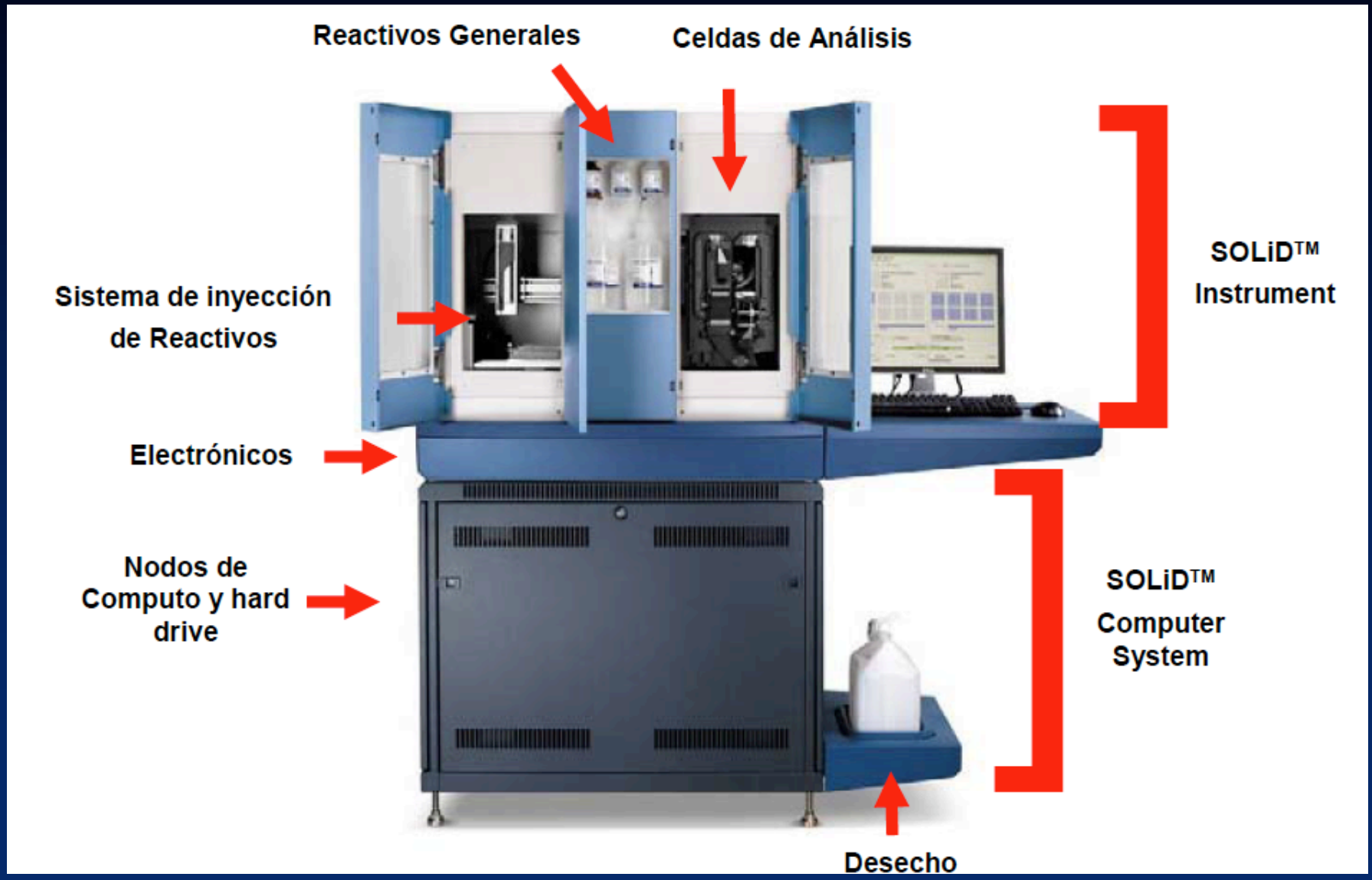
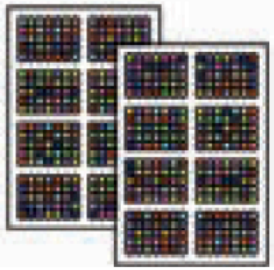
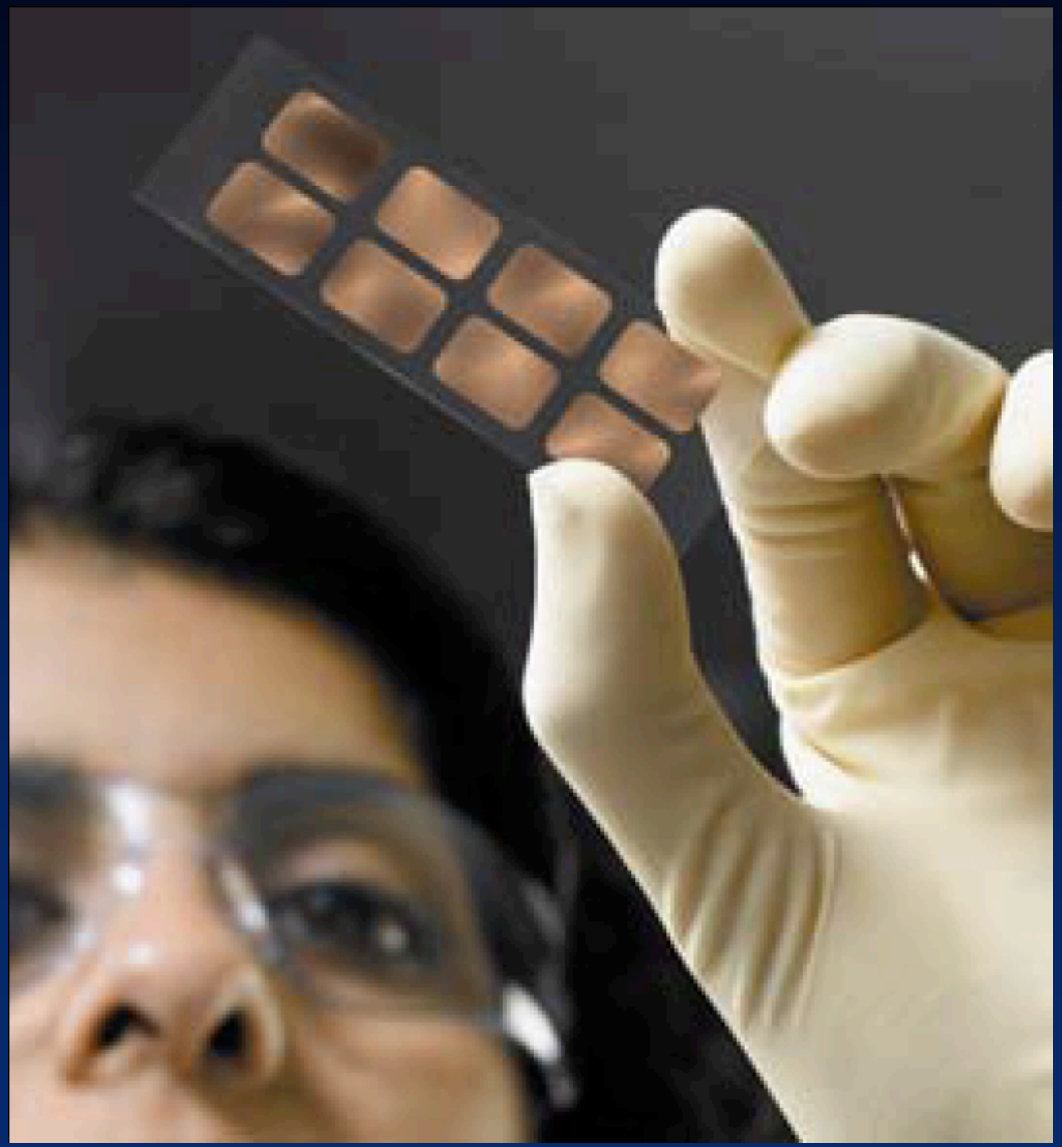
Secuenciación Nueva Generación
SOLiD System

Karol Carrillo Sánchez
kcarrillo@inmegen.gob.mx
21 enero 2010

Sequencing Oligonucleotide by Ligation and Detection

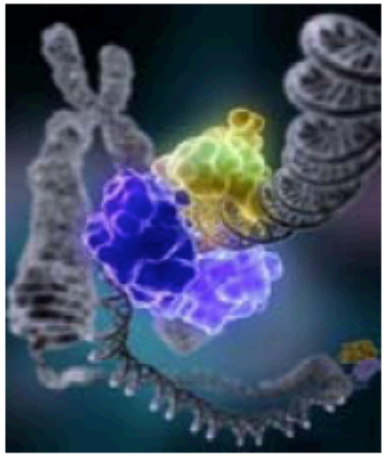

Reactivos Generales
Celdas de Análisis
Sistema de inyección de Reactivos
Electrónicos
Nodos de Computo y hard drive
SOLiD™ Instrument
SOLiD™ Computer System
Desecho

	Unsegmented Full Slide	Quad	Octet
Individual Samples	 2 samples/run	 8 samples/run	 16 samples/run
Multiplexed Samples (16 barcodes)	 ≤16 samples/slide x 2 slides/run ≤ 32 samples/run	 ≤16 samples/segment ≤4 segments/slide x 2 slides/run ≤ 128 samples/run	 ≤16 samples/segment ≤8 segments/slide x 2 slides/run ≤ 256 samples/run

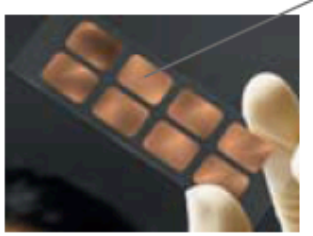


Sequencing Oligonucleotide by Ligation and Detection

Secuenciación por ligación de Oligonucleótidos



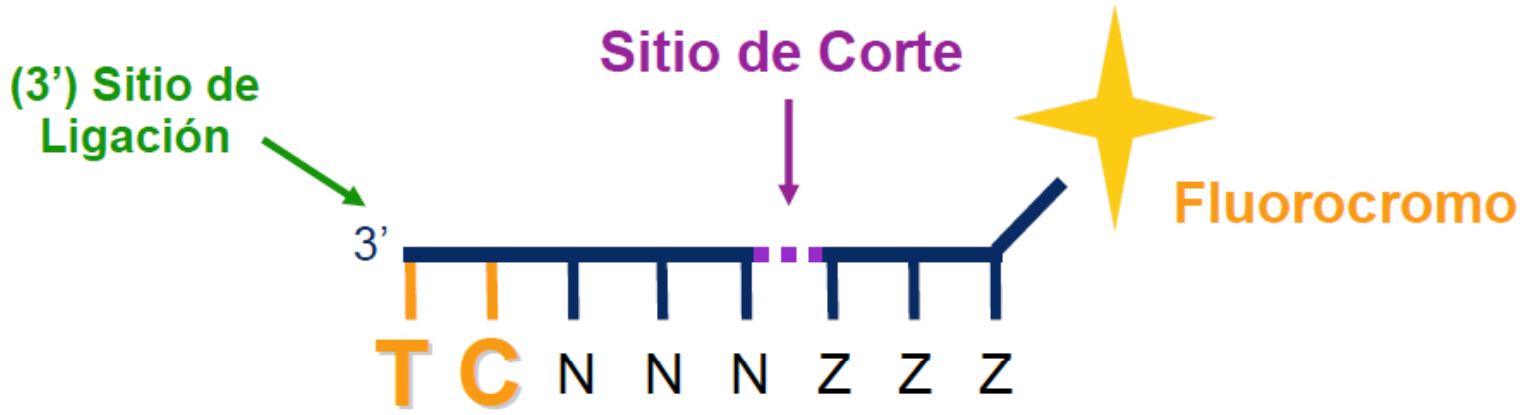
Ligasa



Secuenciación de Alta fidelidad

Ligación ocurre únicamente cuando los oligonucleótidos están perfectamente alineados al DNA complementario.

Características de las Sondas



Separación espacial entre fluorocromo, sitios de ligación y sitio de corte.

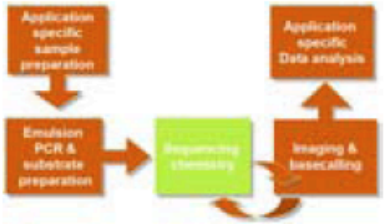
1,024 Sondas Octaméricas (4^5)

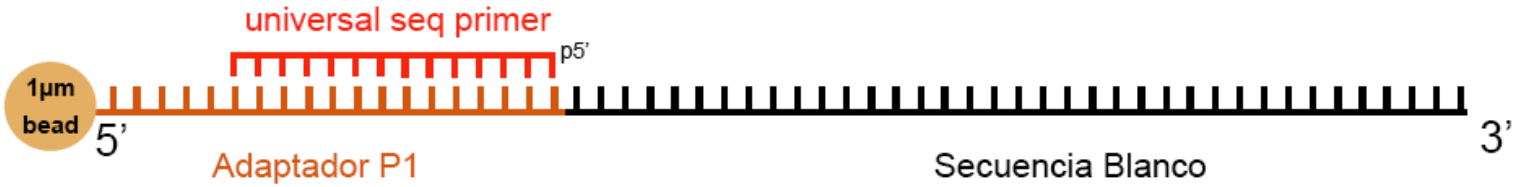
4 Dyes, 4 dinucleótidos

256 sondas por color

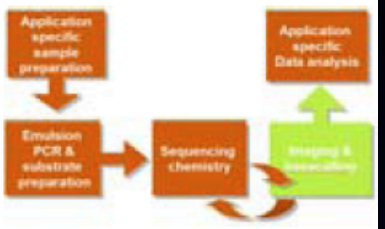
N = bases degeneradas Z = bases universales

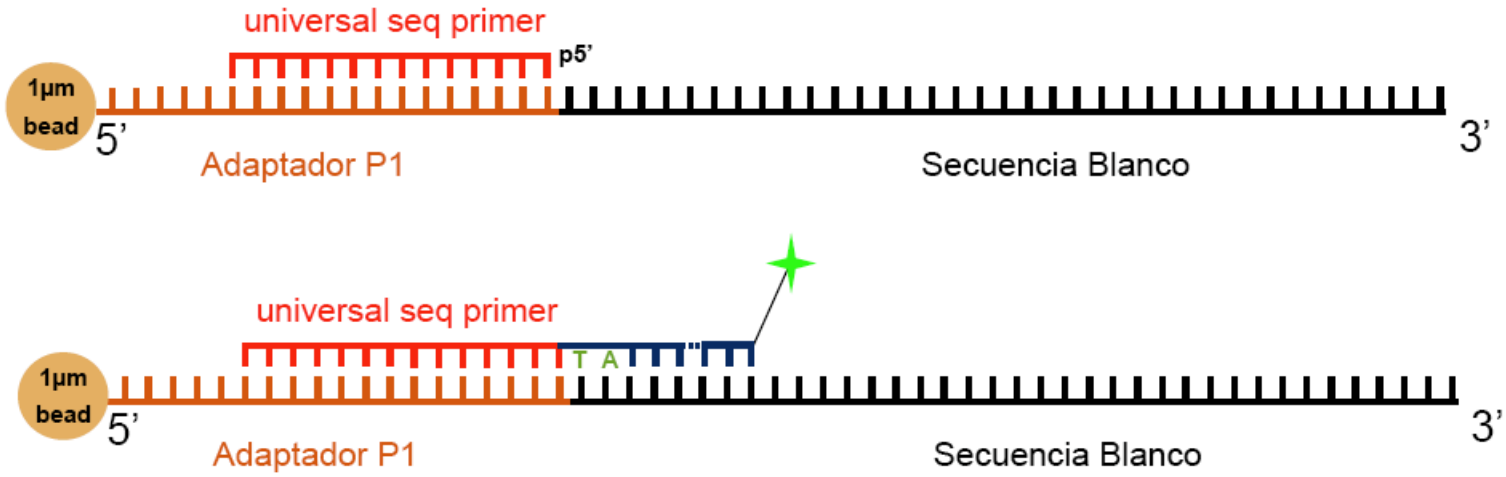
Reacción de Ligación



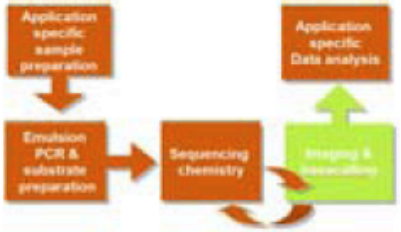


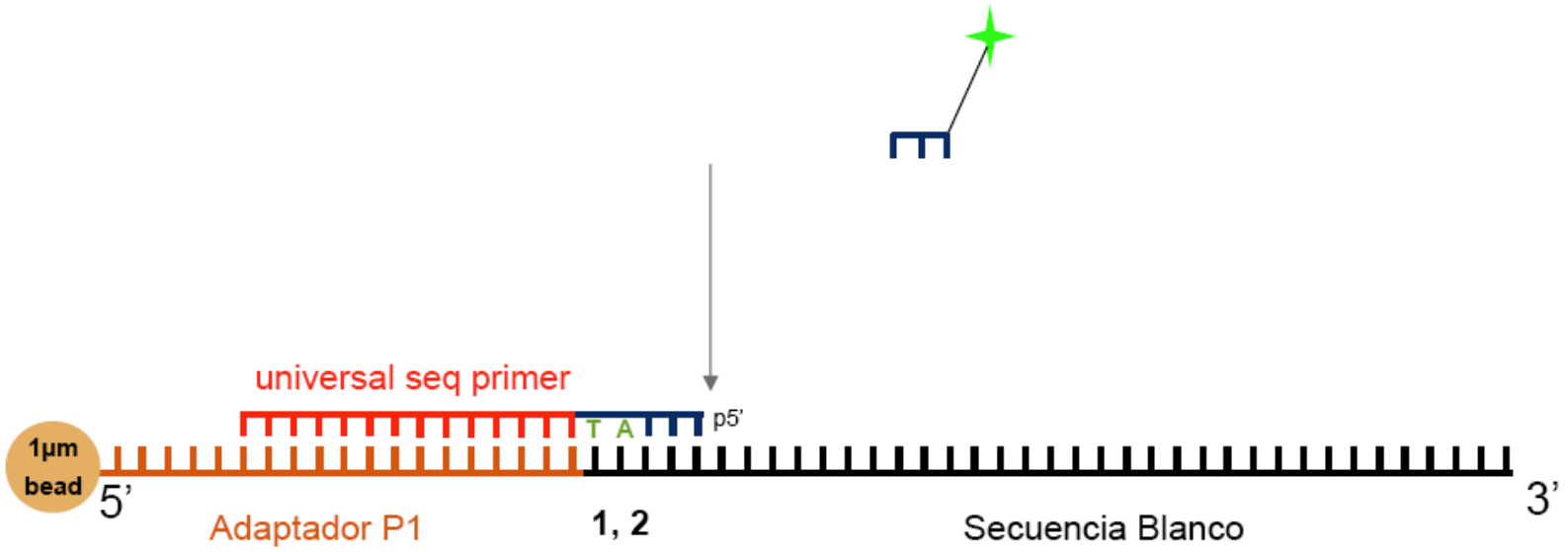
CAP - Defosforilación



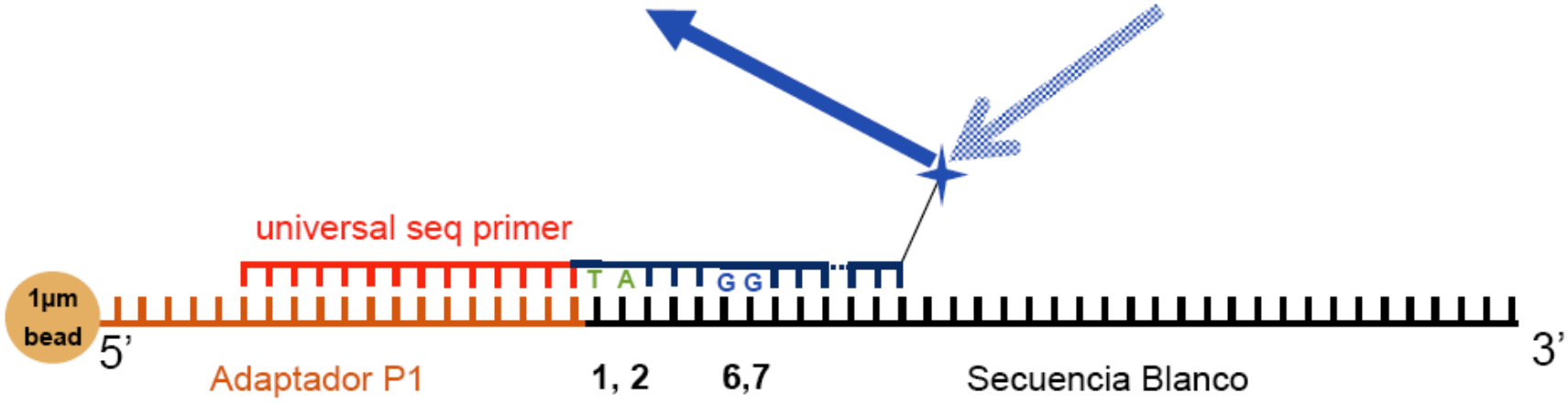


Corte

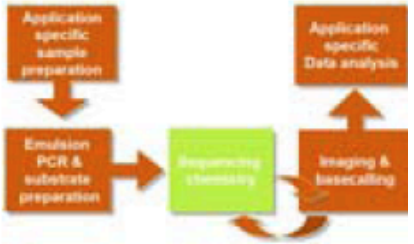

Application specific sample preparation
Emulsion PCR & substrate preparation
Sequencing chemistry
Application specific Data analysis

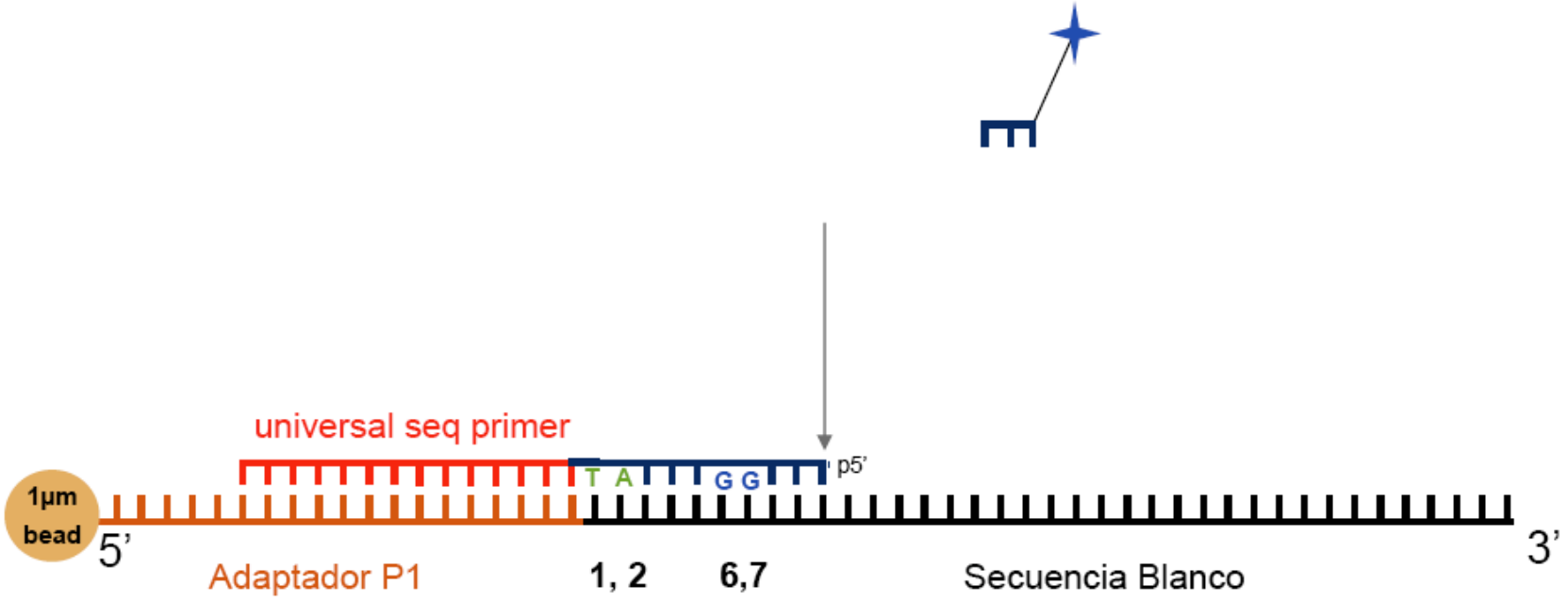

universal seq primer
T A
p5'
1µm bead
5'
Adaptador P1
1, 2
Secuencia Blanco
3'

Segundo Ciclo - Visualización 2

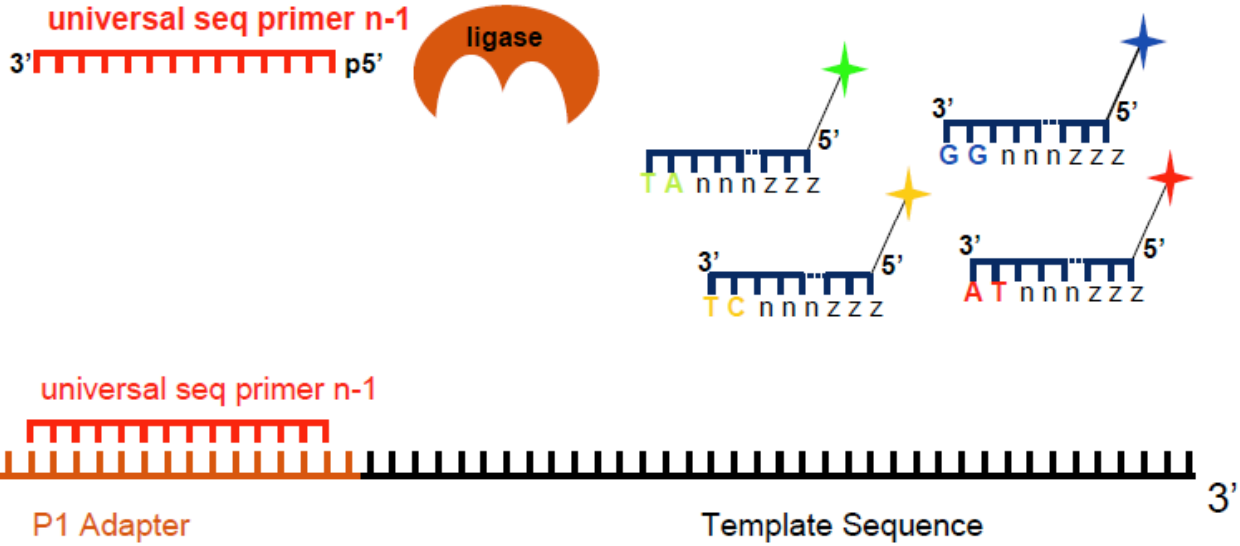


Segundo Ciclo - Corte 2

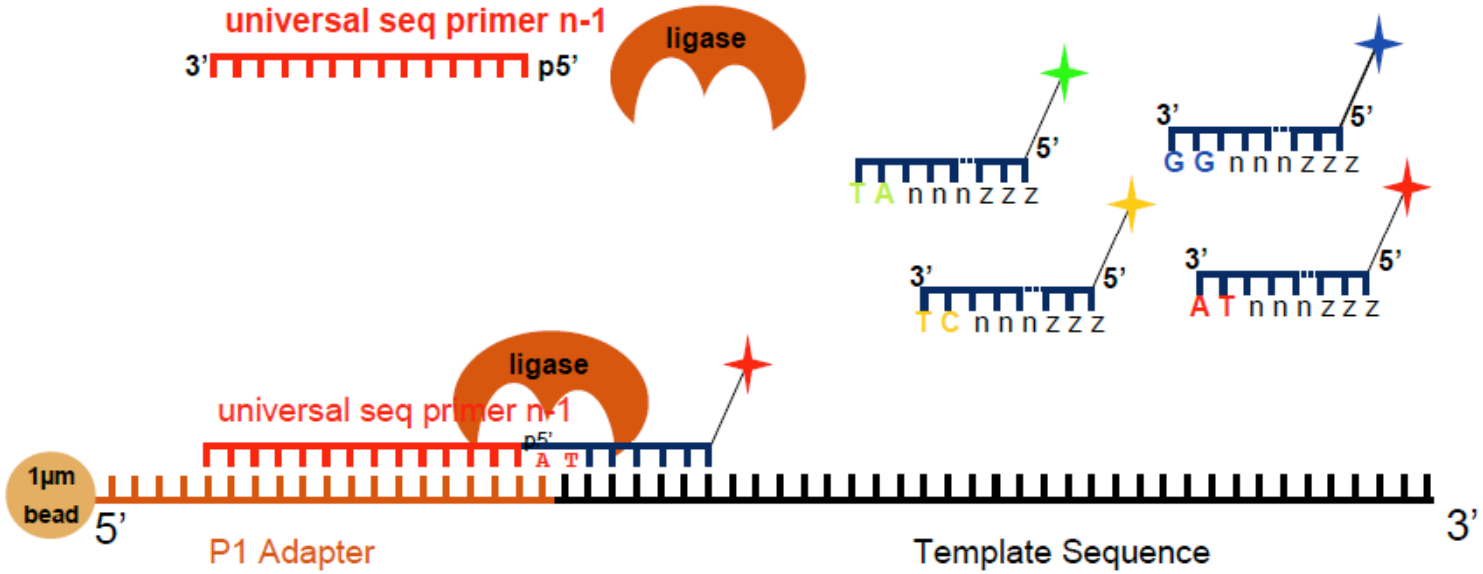




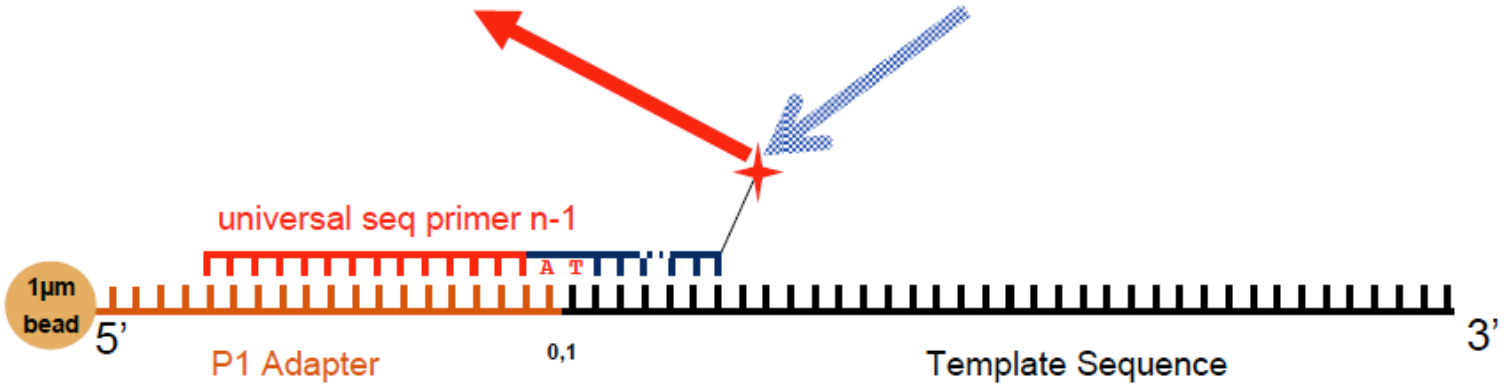
Primer 2
1er ciclo post-reset



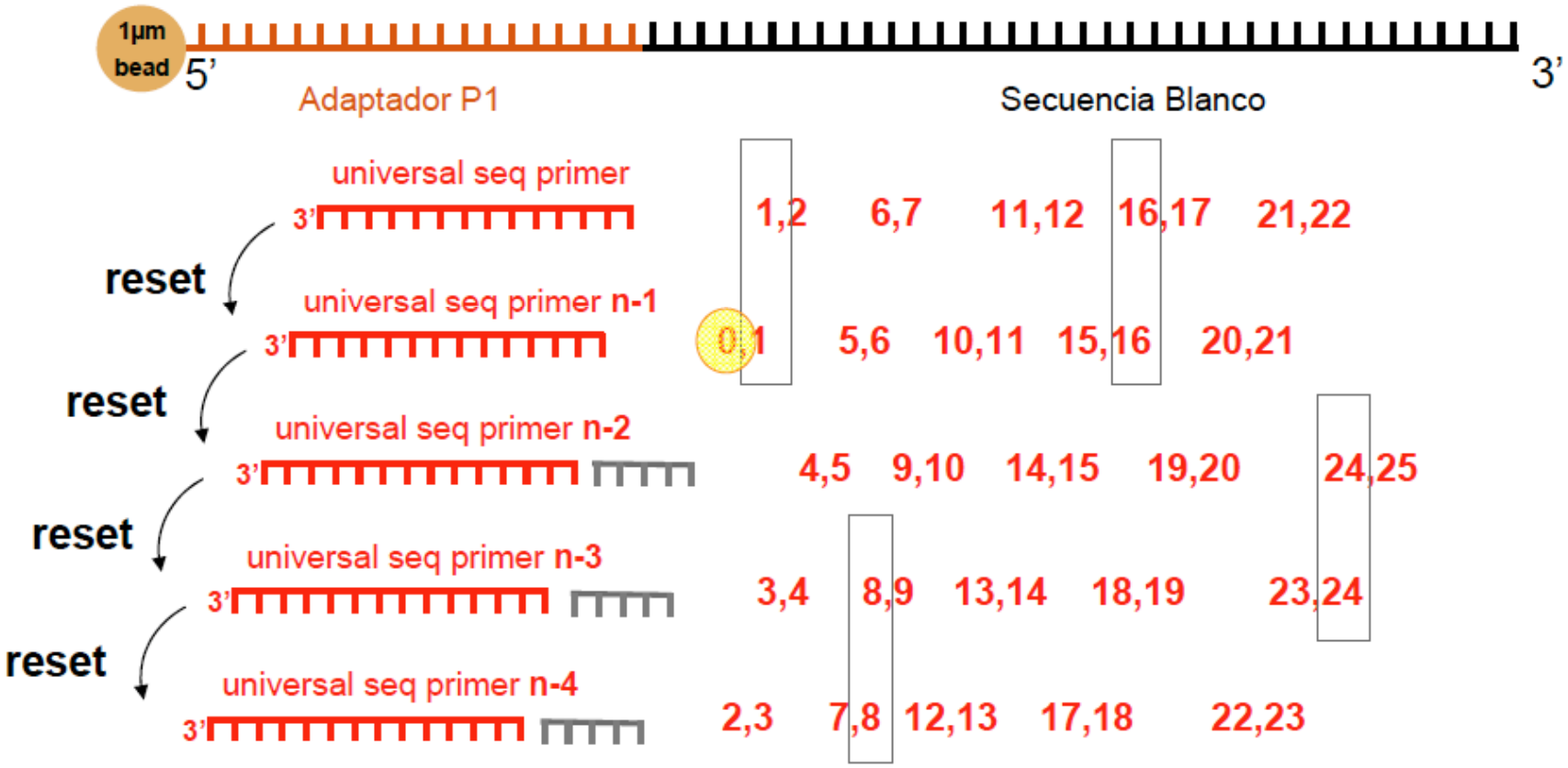
Primer 2
1er ciclo post-reset



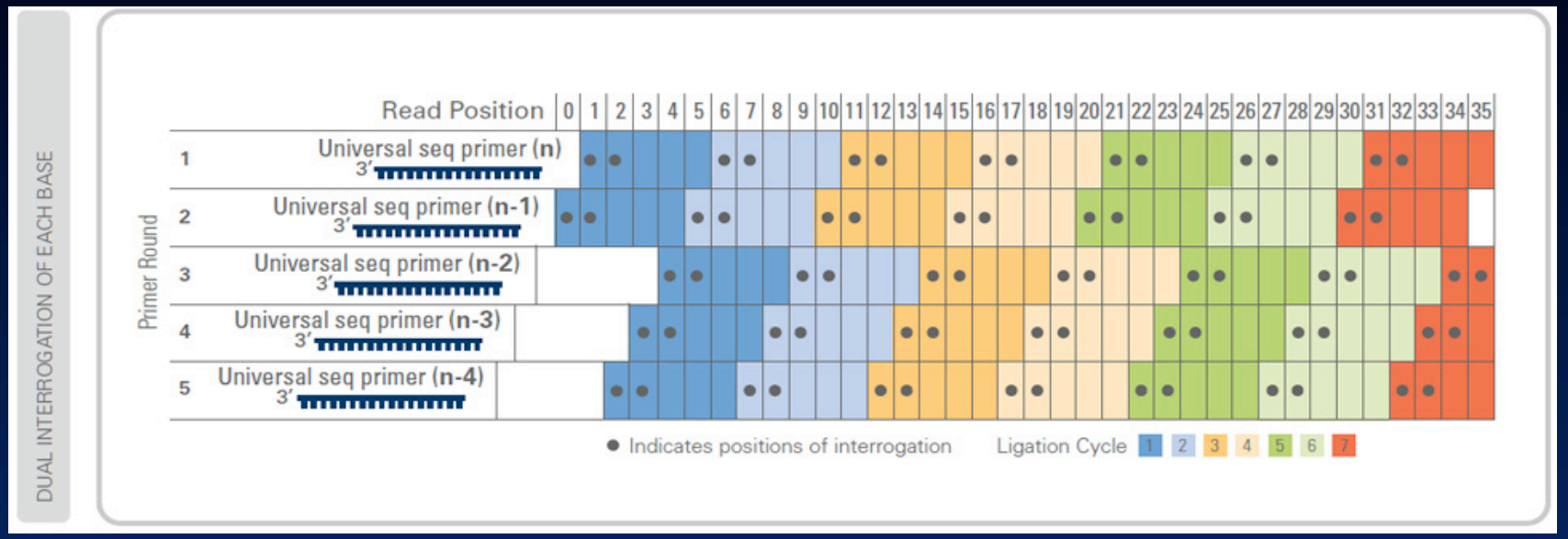
Primer 2
1er ciclo post-reset



Corridas Secuenciales
Múltiples ciclos por corrida

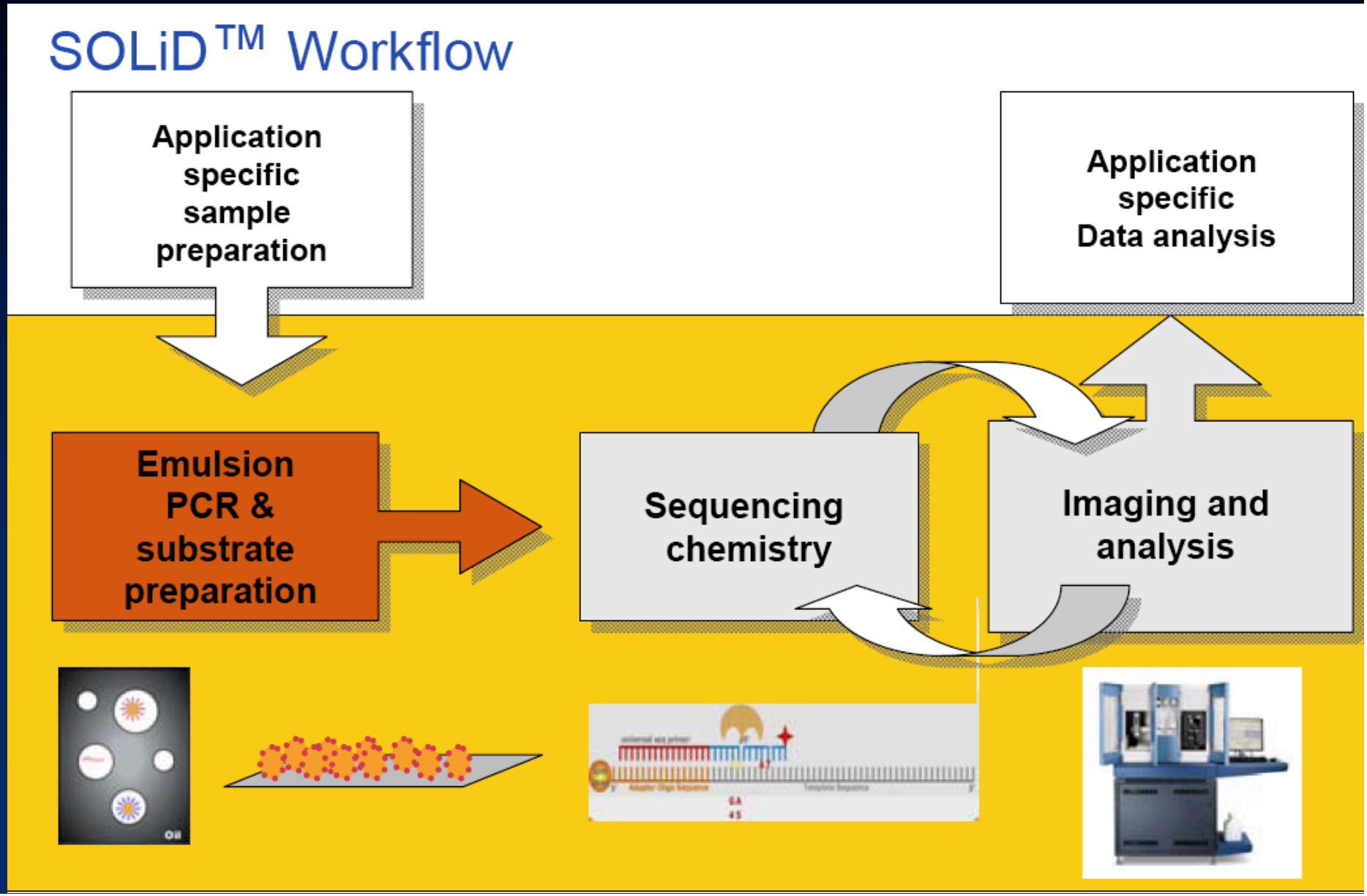


¿Cúantas bases se interrogan por sonda?



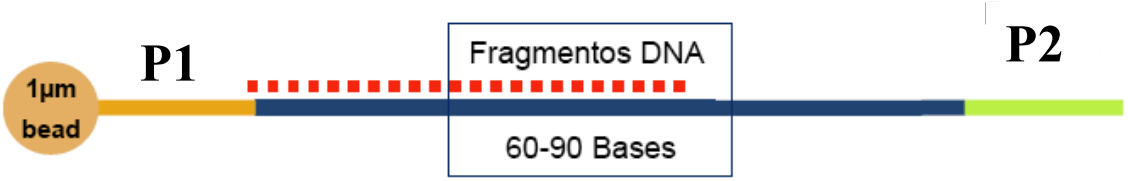
- ✓ Se interrogan 2 bases por sonda
- ✓ Una misma base es interrogada 2 veces

Diagrama de Flujo



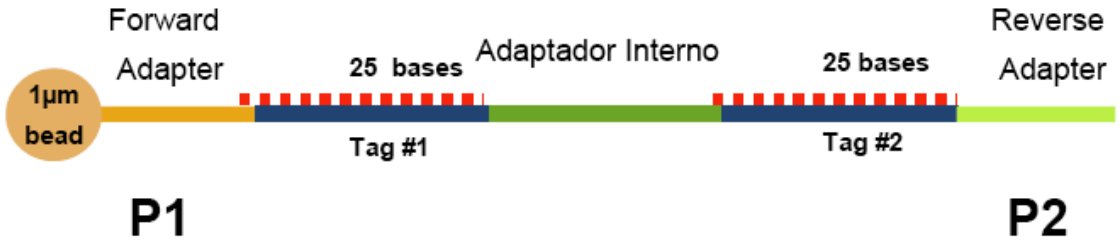
Dos tipos de bibliotecas

Fragmentación



- Análisis de transcriptoma Completo (WTA)
- Resecuenciación dirigida (Productos PCR)
- Descubrimiento de MicroRNAs
- Análisis cuantitativo de RNA
- 3' SAGE o 5' SAGE
- ChIP-Seq
- Descubrimiento SNPs

Mate Pair



- Secuenciación de genomas complejos
- Perfiles de Metilación
- Rearreglos estructurales
- Variación de número de copias
- Descubrimiento de SNPs

Fragmentación de DNA

Covaris

HydroShear

Bibliotecas de Fragmentación

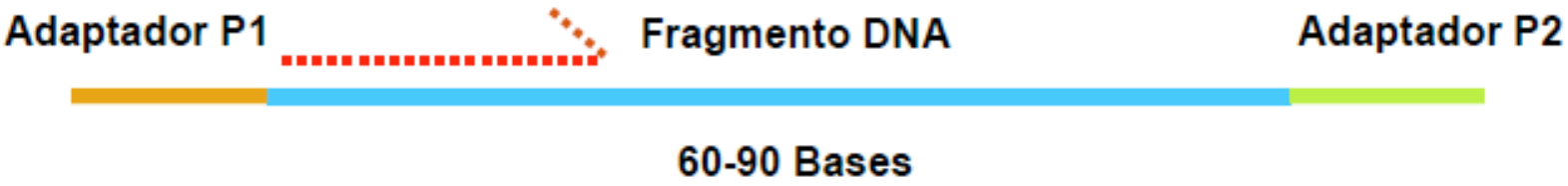
Muestra Compleja

ej. DNA Genómico, Productos de PCR concatenados, TAG library.

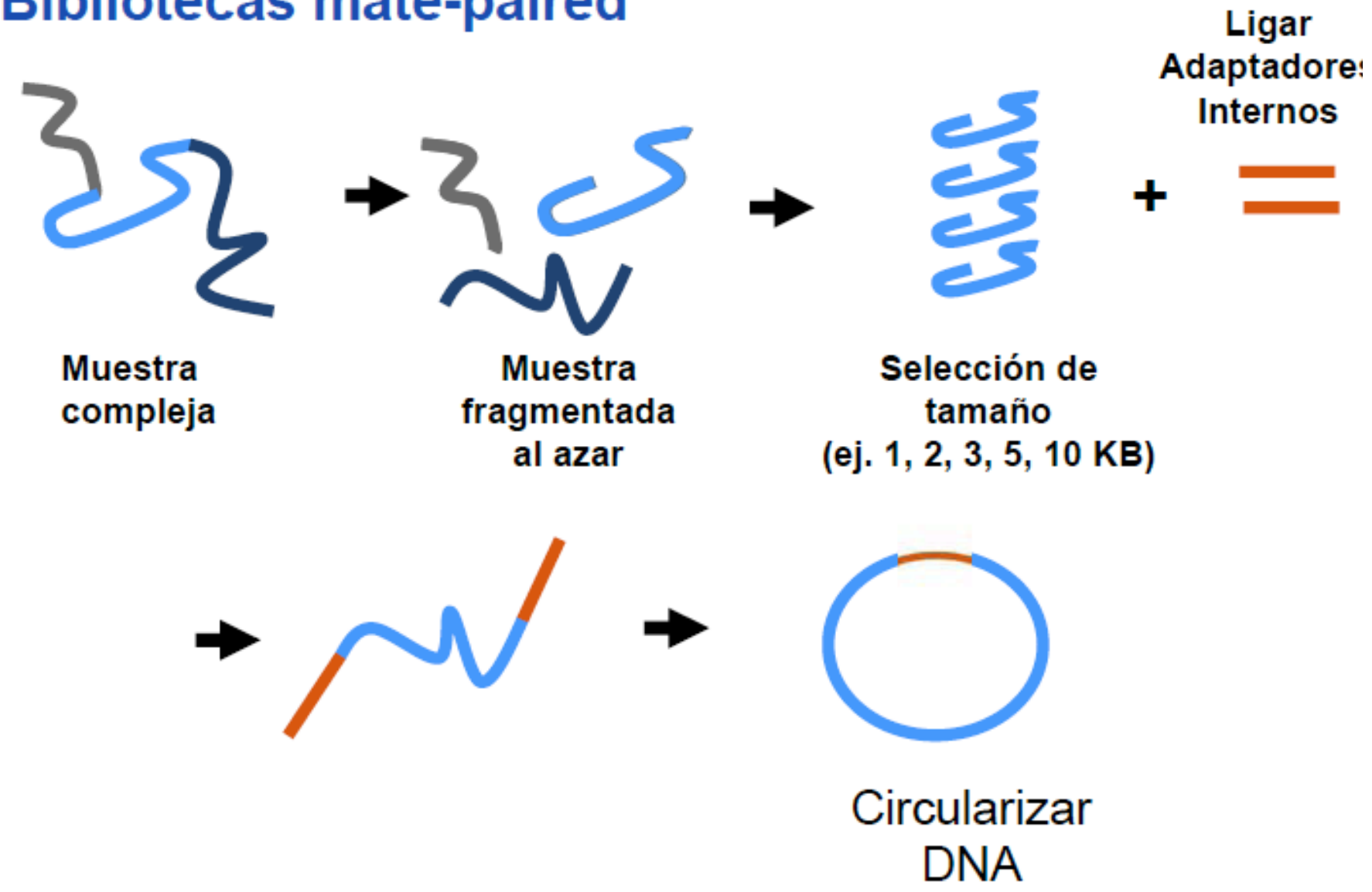
Muestra Fragmentada,

Randomly or Targeted
ej. Sonicación, fragmentación, digestión enzimática

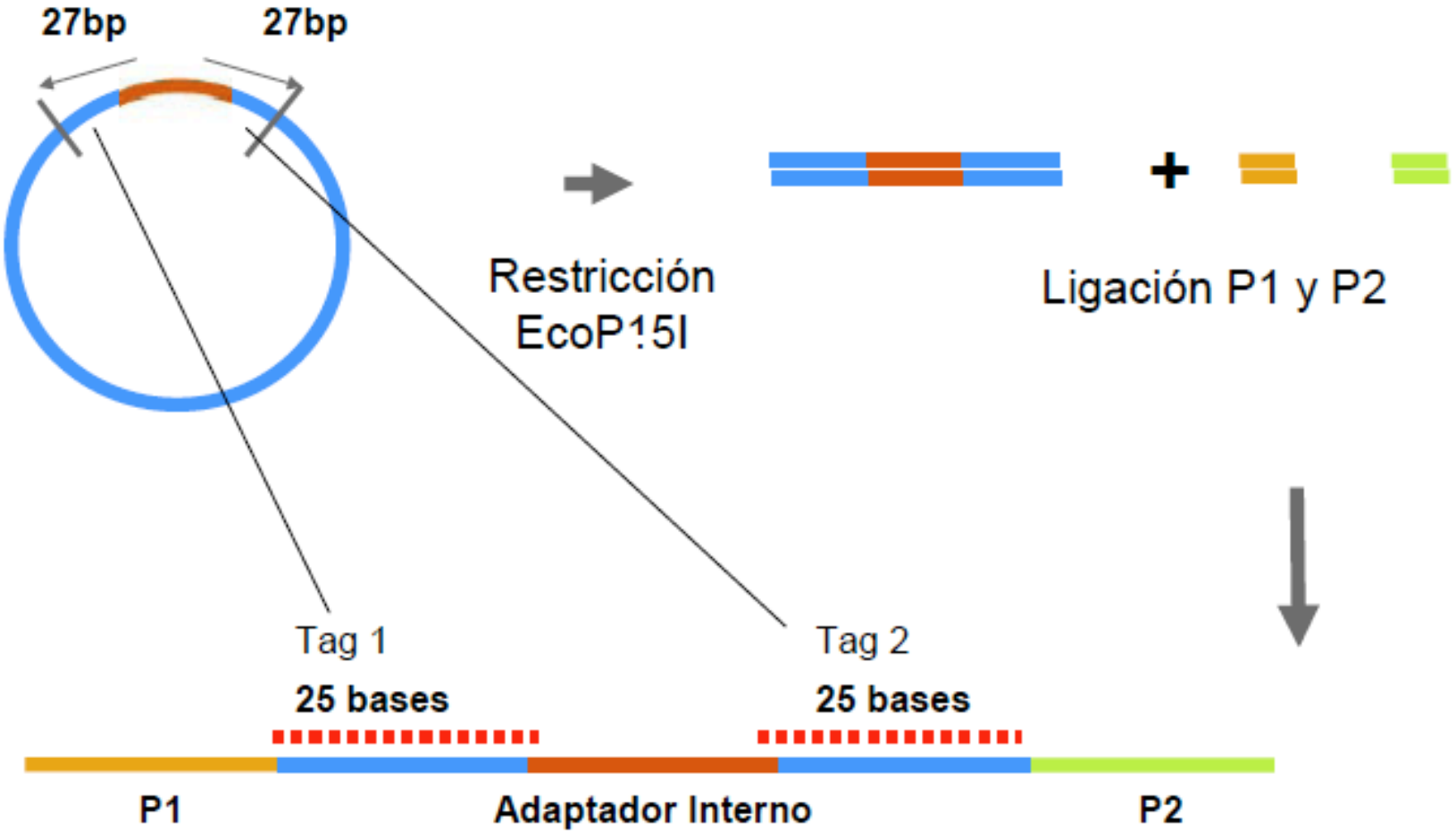
Ligación P1 y P2



Bibliotecas mate-paired

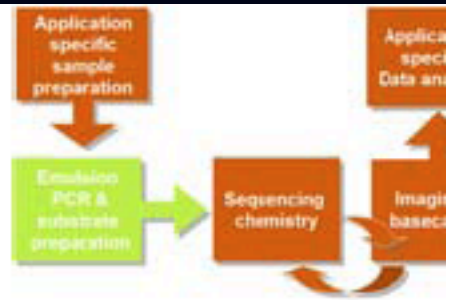


Creación de biblioteca mate-paired (II)



Amplificación Clonal de la Biblioteca

PCR en emulsión



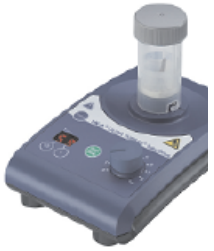
Biblioteca DNA

+

Primers P1<<P2

Perlas acopladas a P1

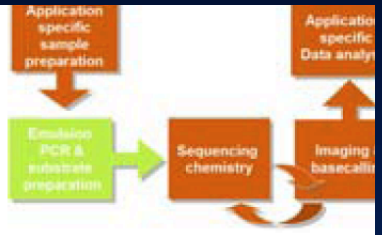
Enzima

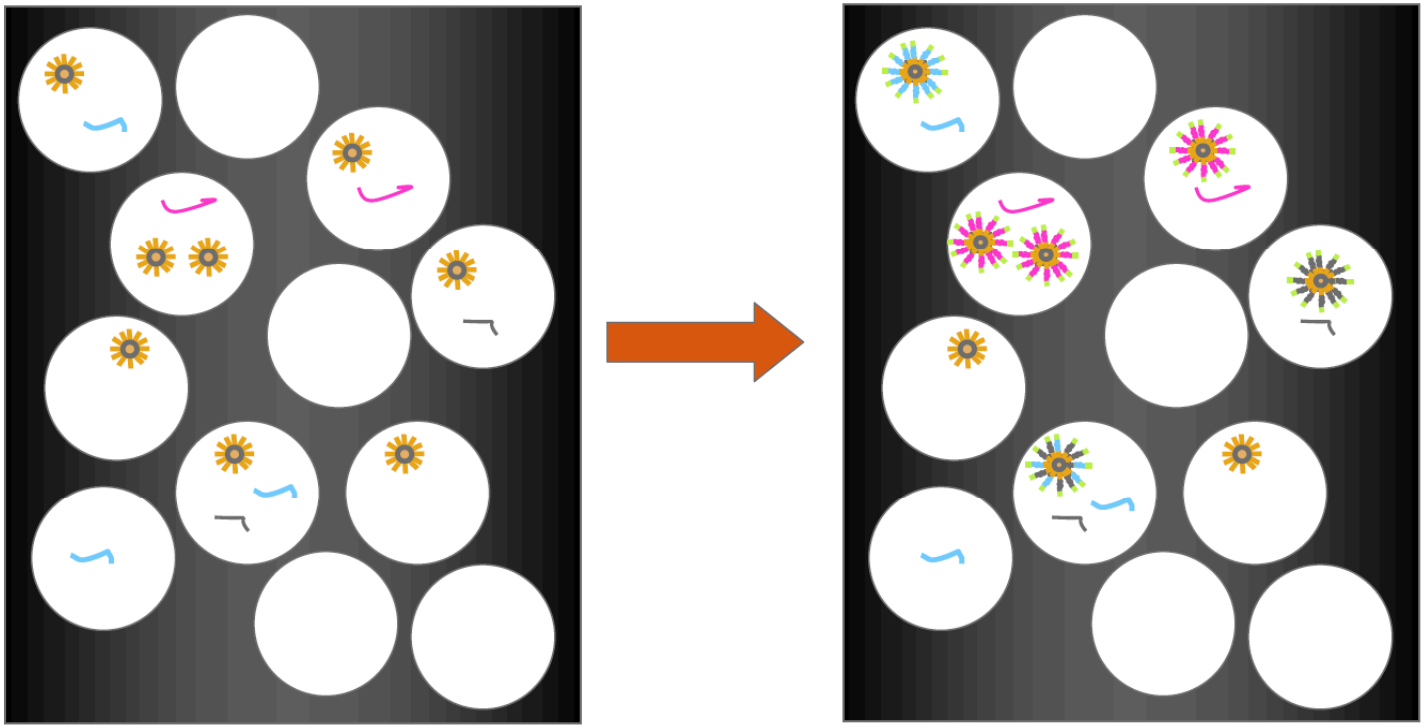


IKA® ULTRA-TURRAX® Tube

PCR en Emulsión

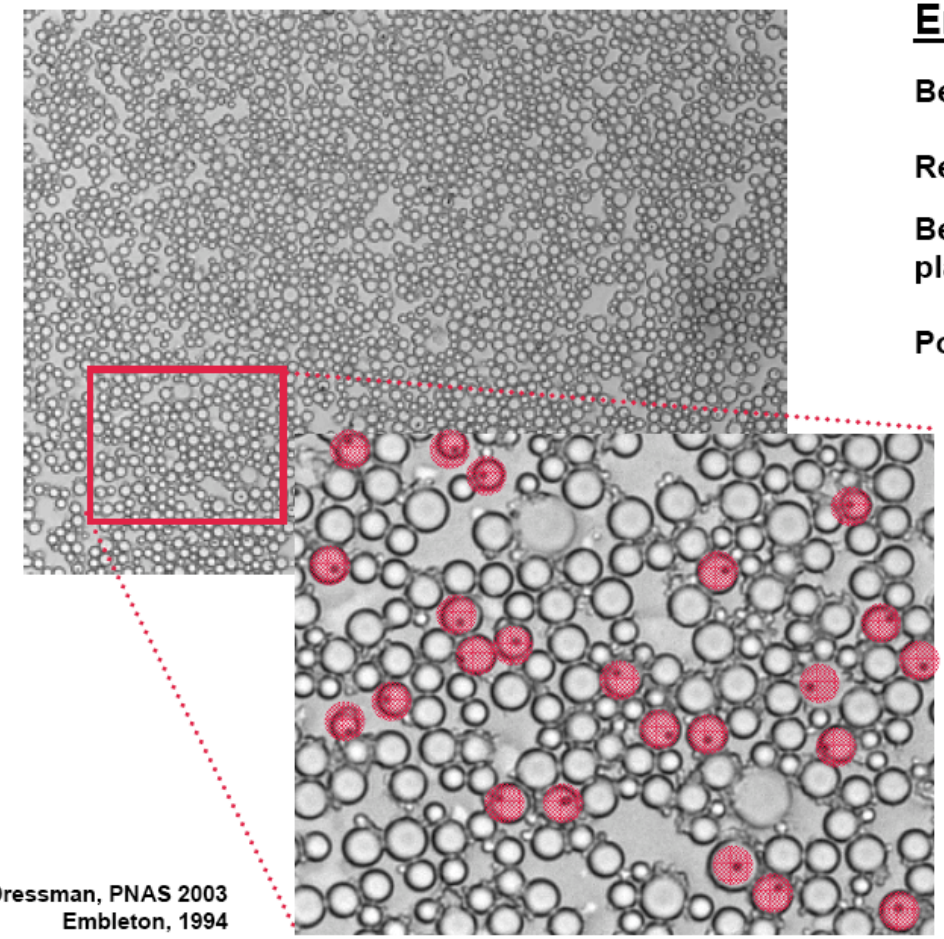
La reacción de amplificación se desarrolla dentro de las pequeñas micelas generadas.





PCR en emulsión

SOLiD system: Clonal Amplification

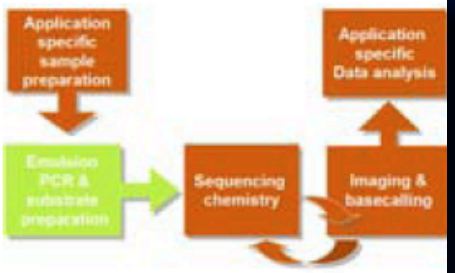


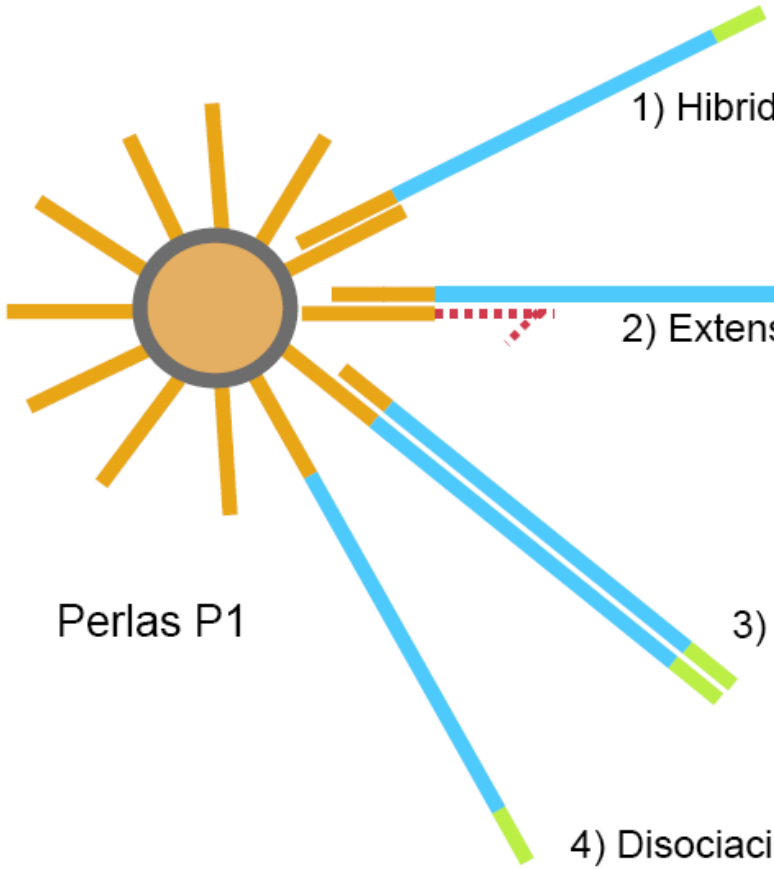
Dressman, PNAS 2003
Embleton, 1994

Emulsion Metrics

Bead size:	1 μm
Reactor size:	8-10 μm
Beads / emulsion plate (96-well):	1600 M beads
Post Enrichment:	~200M / plate

PCR en Emulsión, Amplificación Clonal





1) Hibridación P1 - DNA blanco

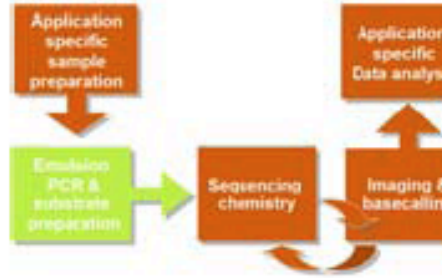
2) Extensión de cadena desde P1

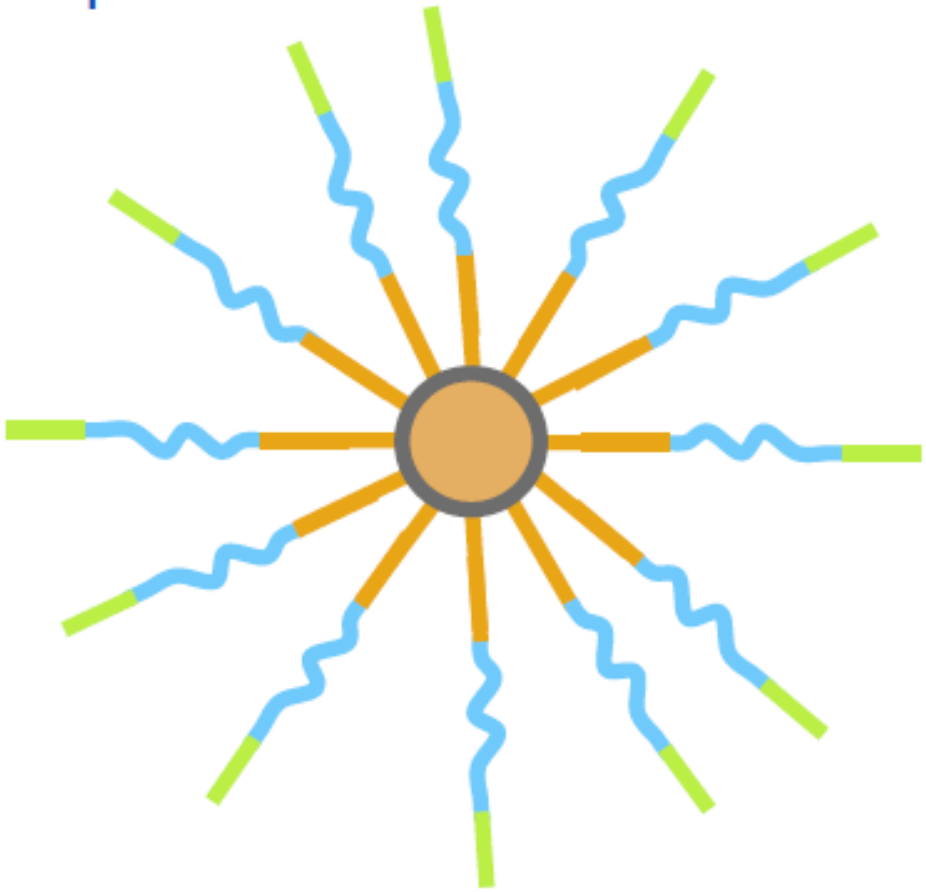
Perlas P1

3) Extensión completa (secuencia complementaria sobre la superficie de las perlas

4) Disociación del DNA blanco

Amplificación Clonal

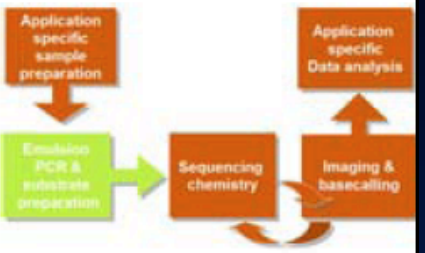


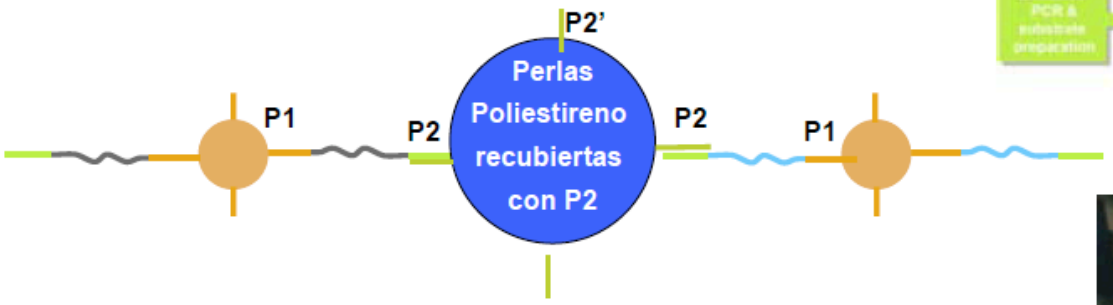


Perlas amplificadas contienen ~30K amplicones generados a partir del material original (ssDNA)

Perlas sin producto

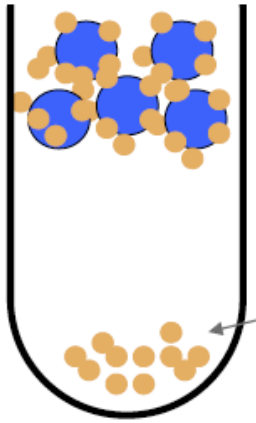
Enriquecimiento





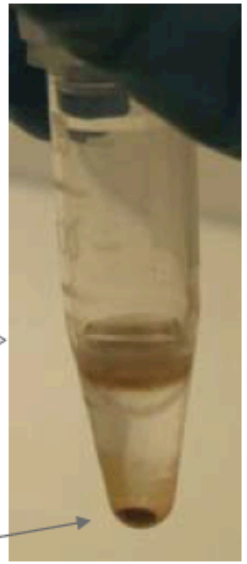
Centrifugar

en gradiente de glicerol.

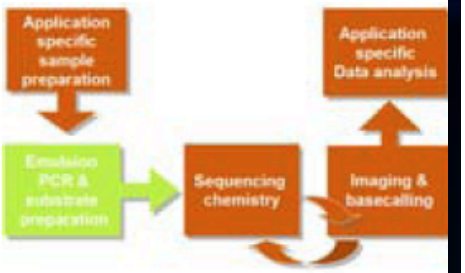


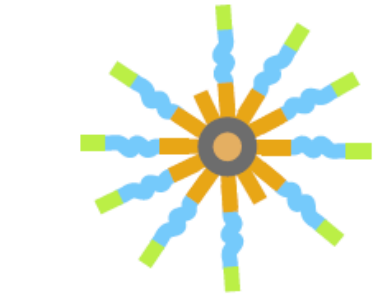
Sobrenadante
Perlas Capturadas (Amplificación +)

Pastilla
Perlas sin amplicon

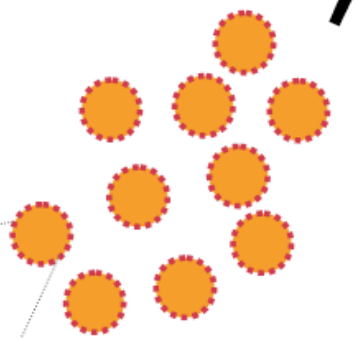


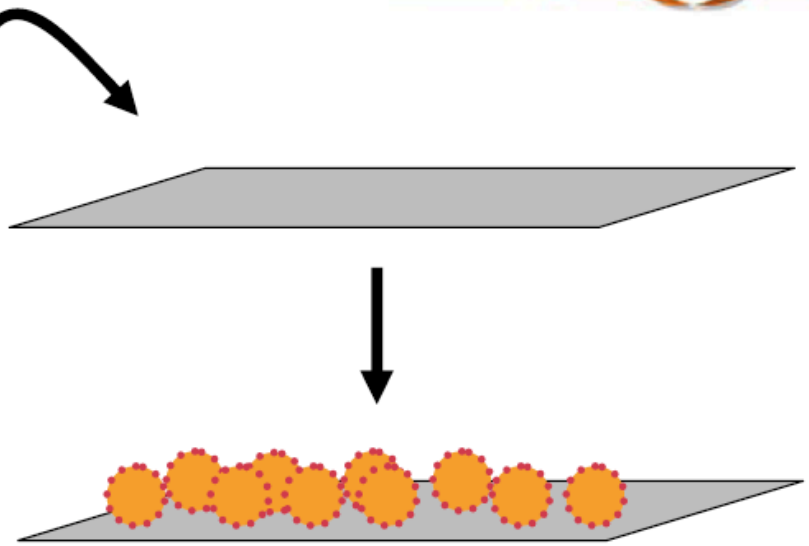
Deposición sobre el soporte





Modificación extremos 3’

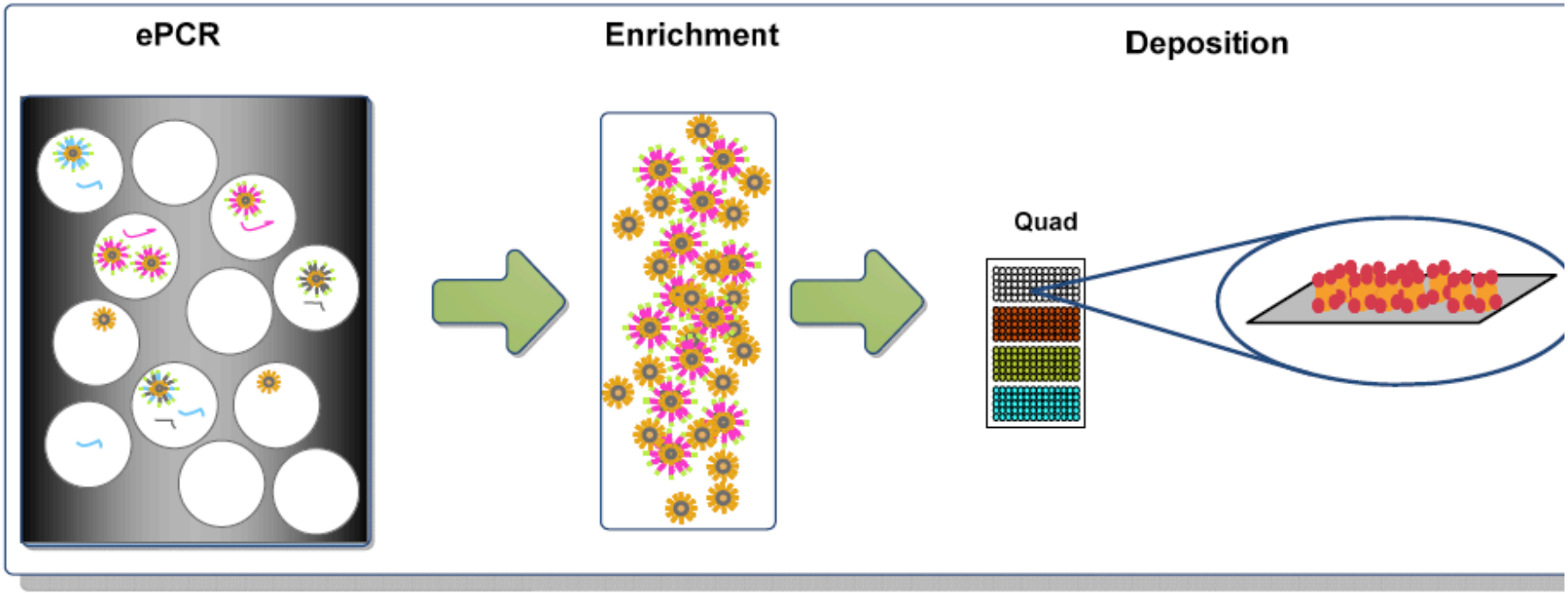


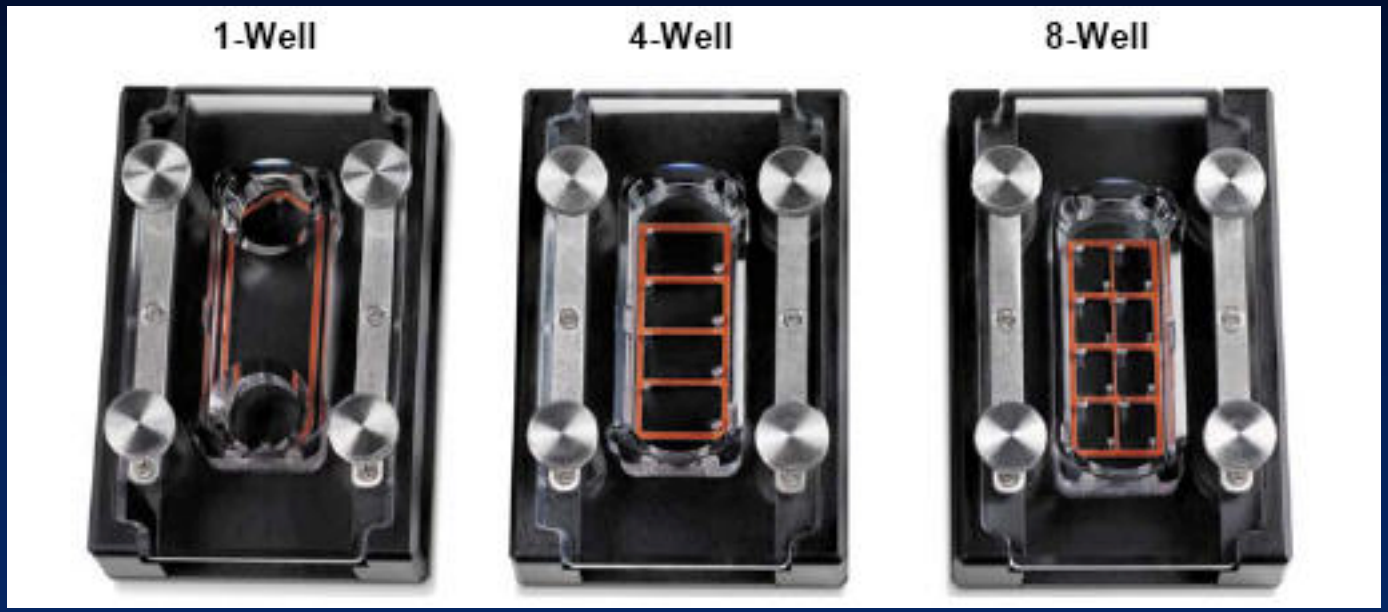


Las perlas son adheridas a la superficie del portaobjetos en un arreglo **aleatorio**.

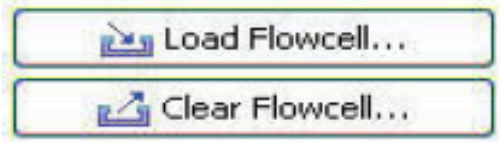
Preparación de perlas y Deposición

Sample Preparation → Library Construction → Emulsion PCR → SOLiD™ Sequencing → Data Analysis



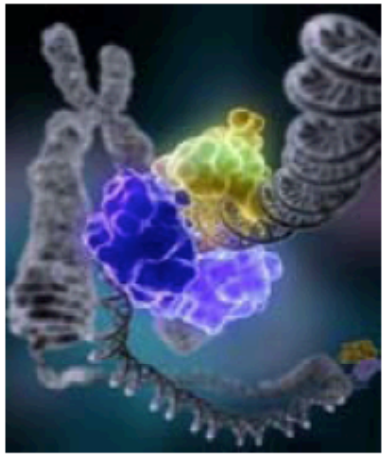

1-Well
4-Well
8-Well

Cargado de portaobjetos en el Instrumento

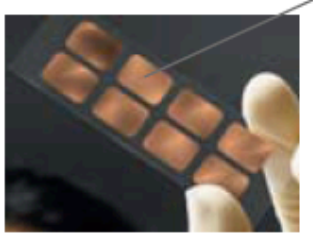


Sequencing Oligonucleotide by Ligation and Detection

Secuenciación por ligación de Oligonucleótidos



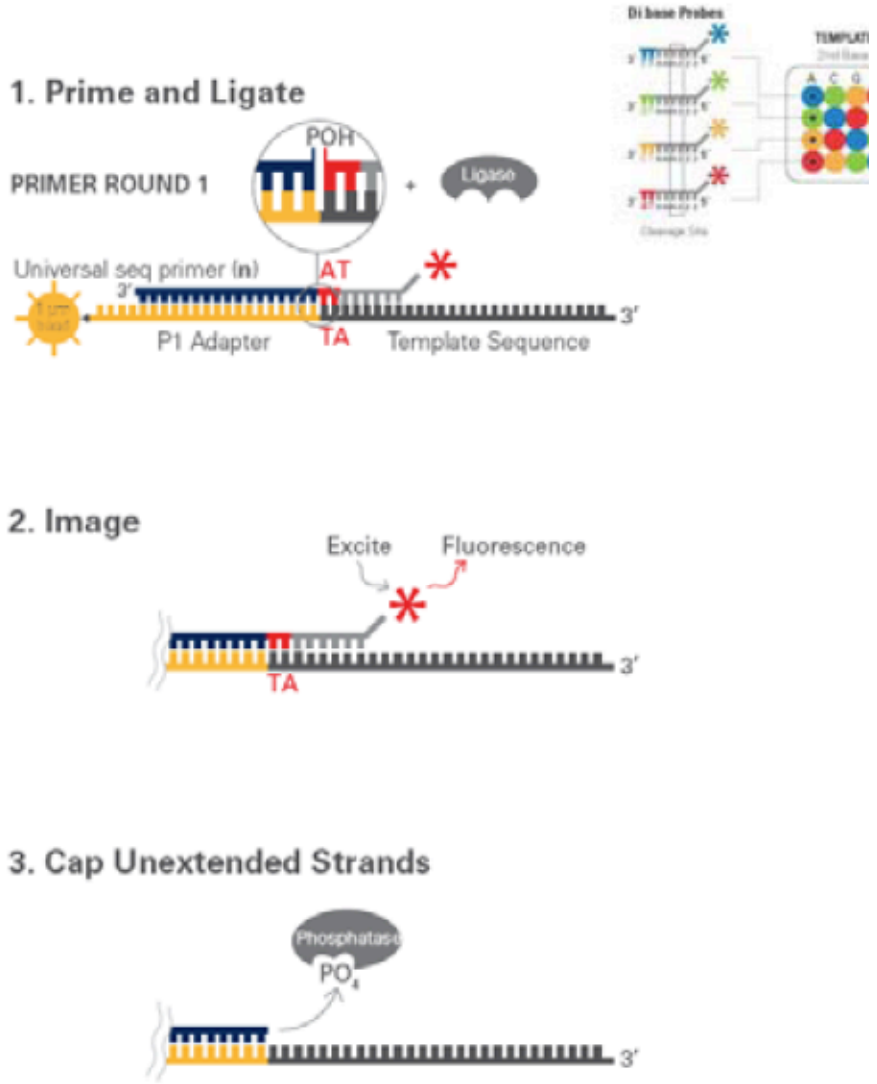
Ligasa



Secuenciación de Alta fidelidad

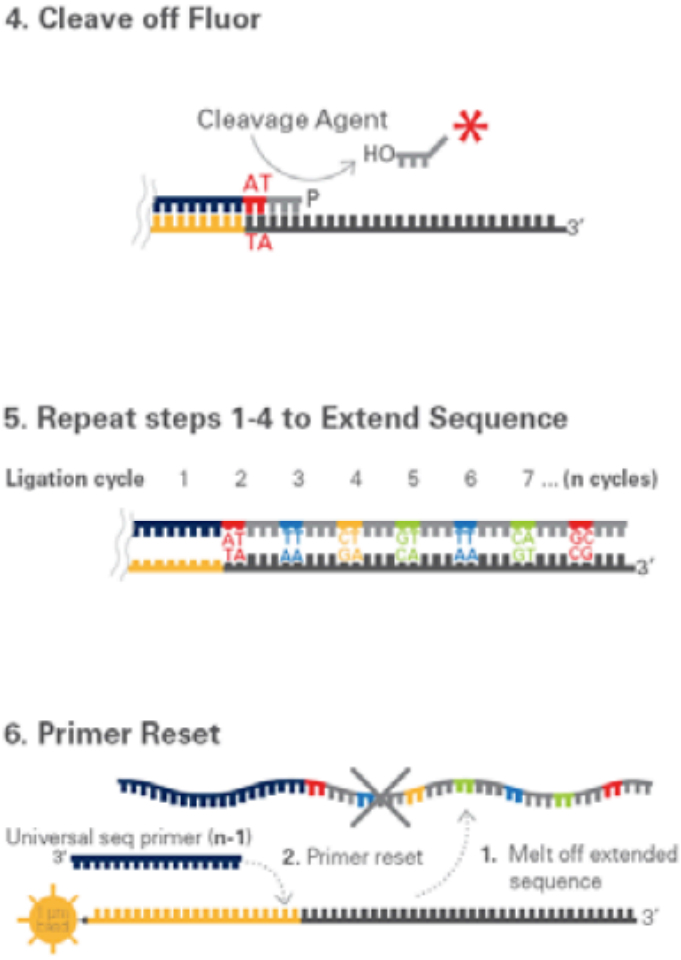
Ligación ocurre únicamente cuando los oligonucleótidos están perfectamente alineados al DNA complementario.

1. Ligación
2. Captura de Imagen
3. Cap
4. Corte
5. Repetir n ciclos de ligación
6. Primer Reset
7. Cambio de Primer



Repetir 1-6 para cada uno de 5 primers.

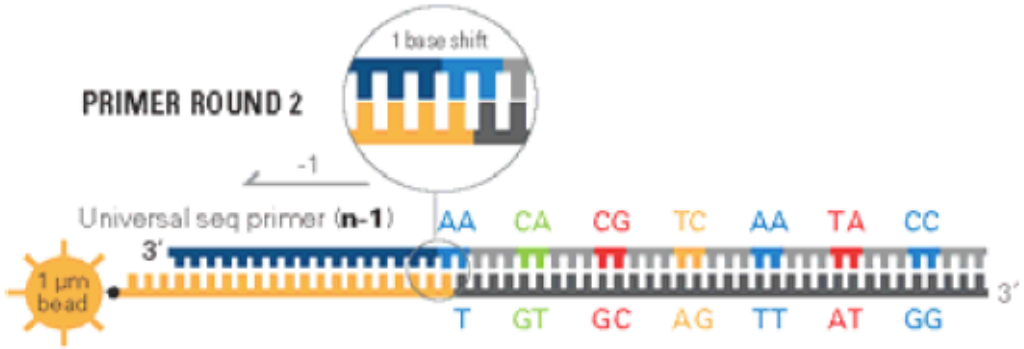
1. Ligación
2. Captura de Imagen
3. Cap
4. Corte
5. Repetir n ciclos de ligación
6. Primer Reset
7. Cambio de Primer

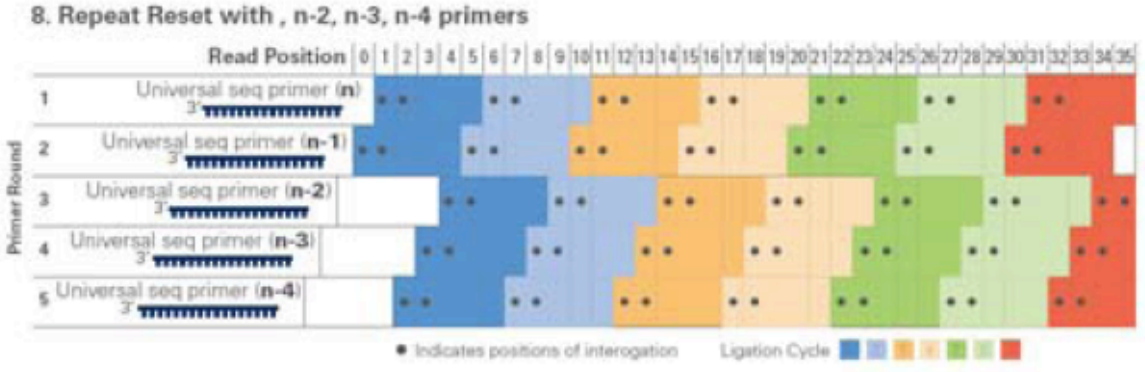


Repetir 1-6 para cada uno de 5 primers.

1. Ligación
2. Captura de Imagen
3. Cap
4. Corte
5. Repetir n ciclos de ligación
6. Primer Reset
7. Cambio de Primer

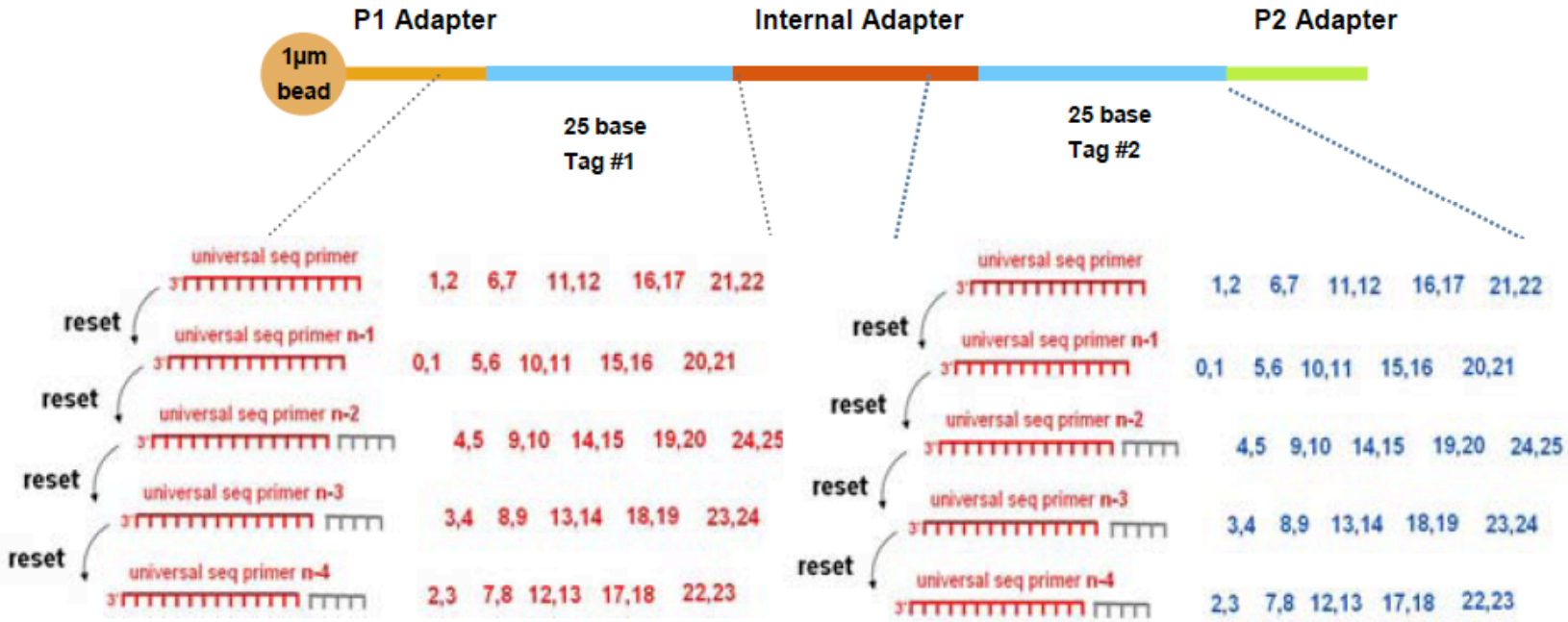
→ Repetir 1-6 para cada uno de los 5 primers



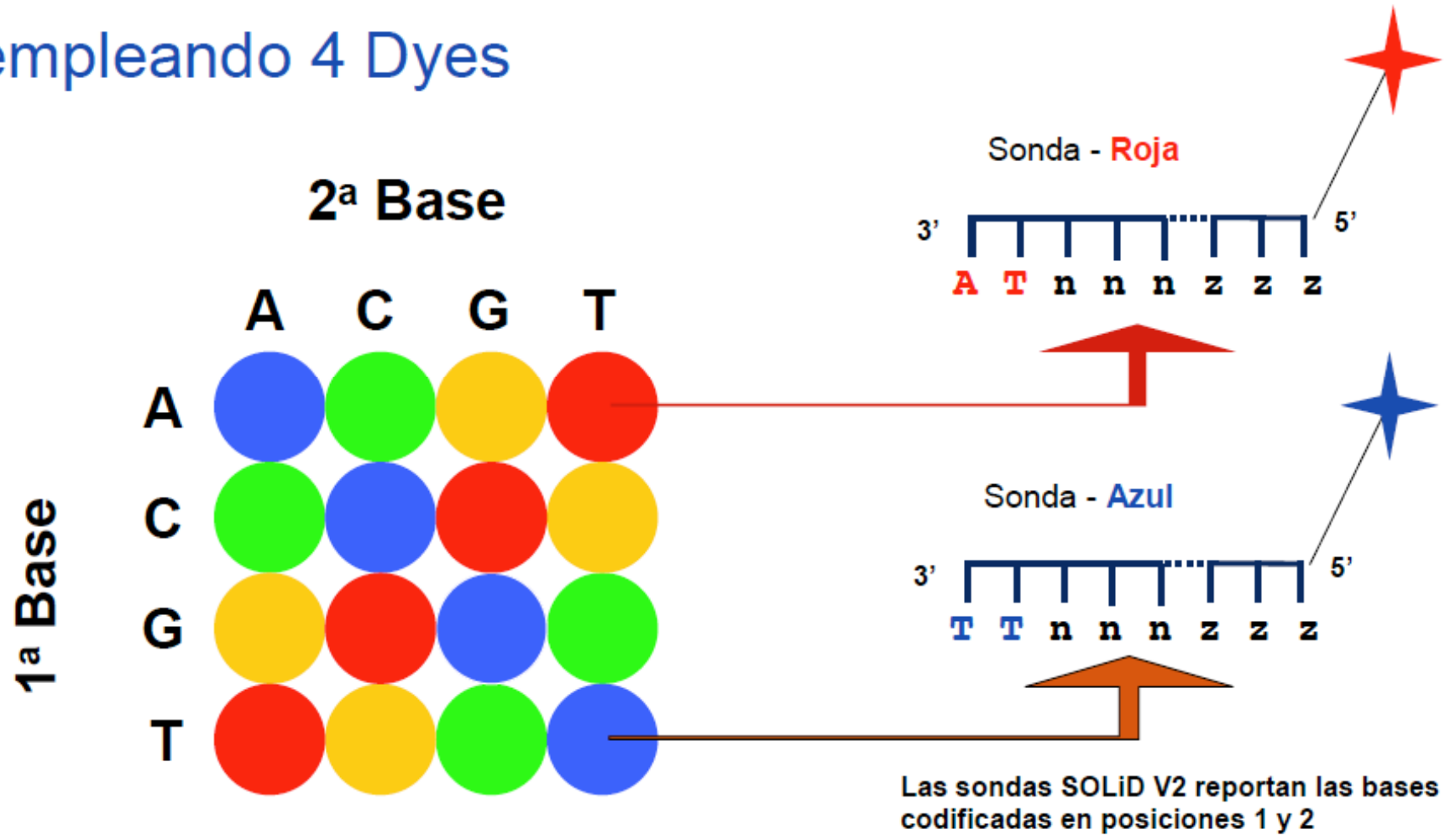


Corridas Mate Pair

generan dos grupos de secuencias

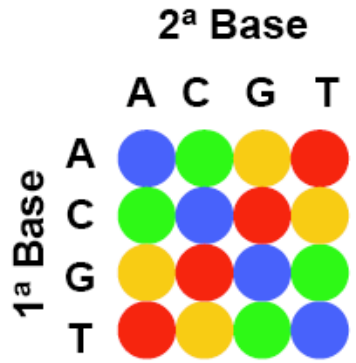


Codificación binaria empleando 4 Dyes



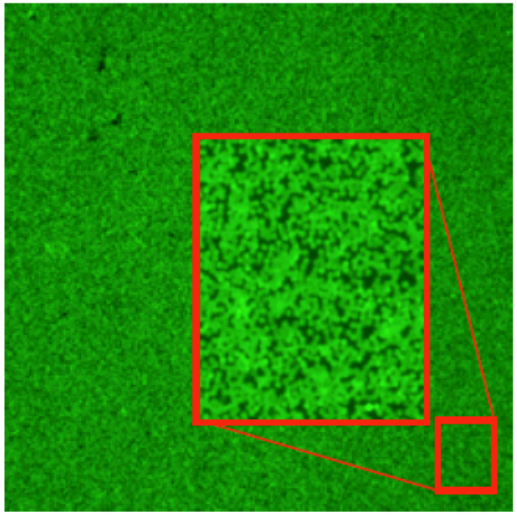
Color Space

La interrogación de dos bases facilita la discriminación entre **errores** sistemáticos y polimorfismos **verdaderos**.

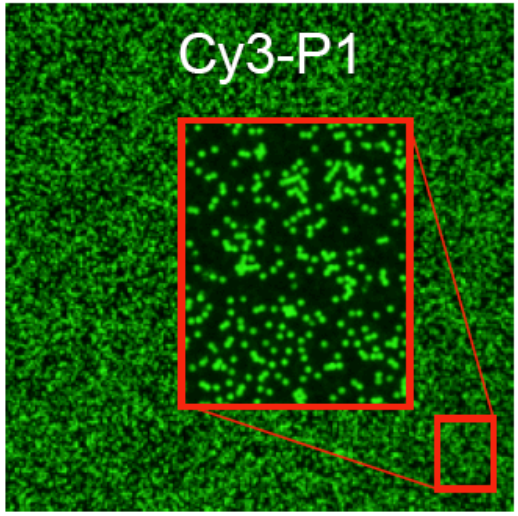


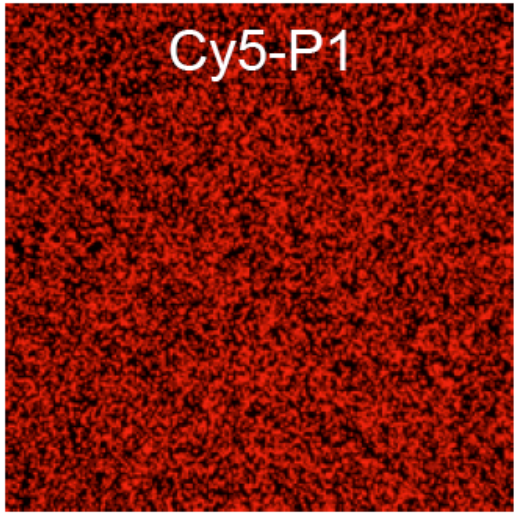
SOLiD 3.0

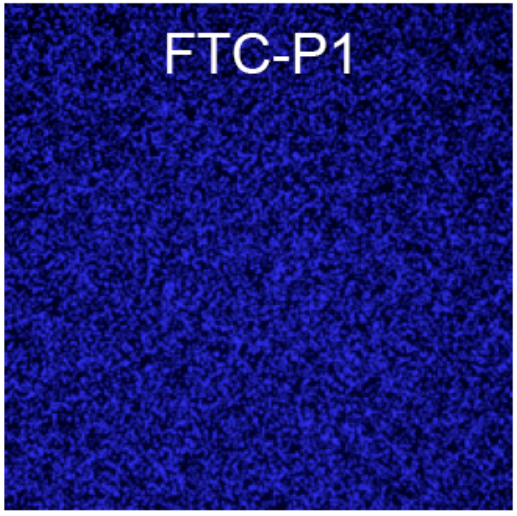
300K/panel

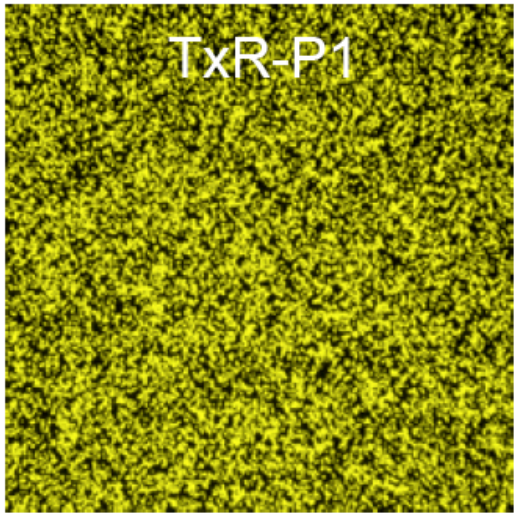


Cy3 Focal Map

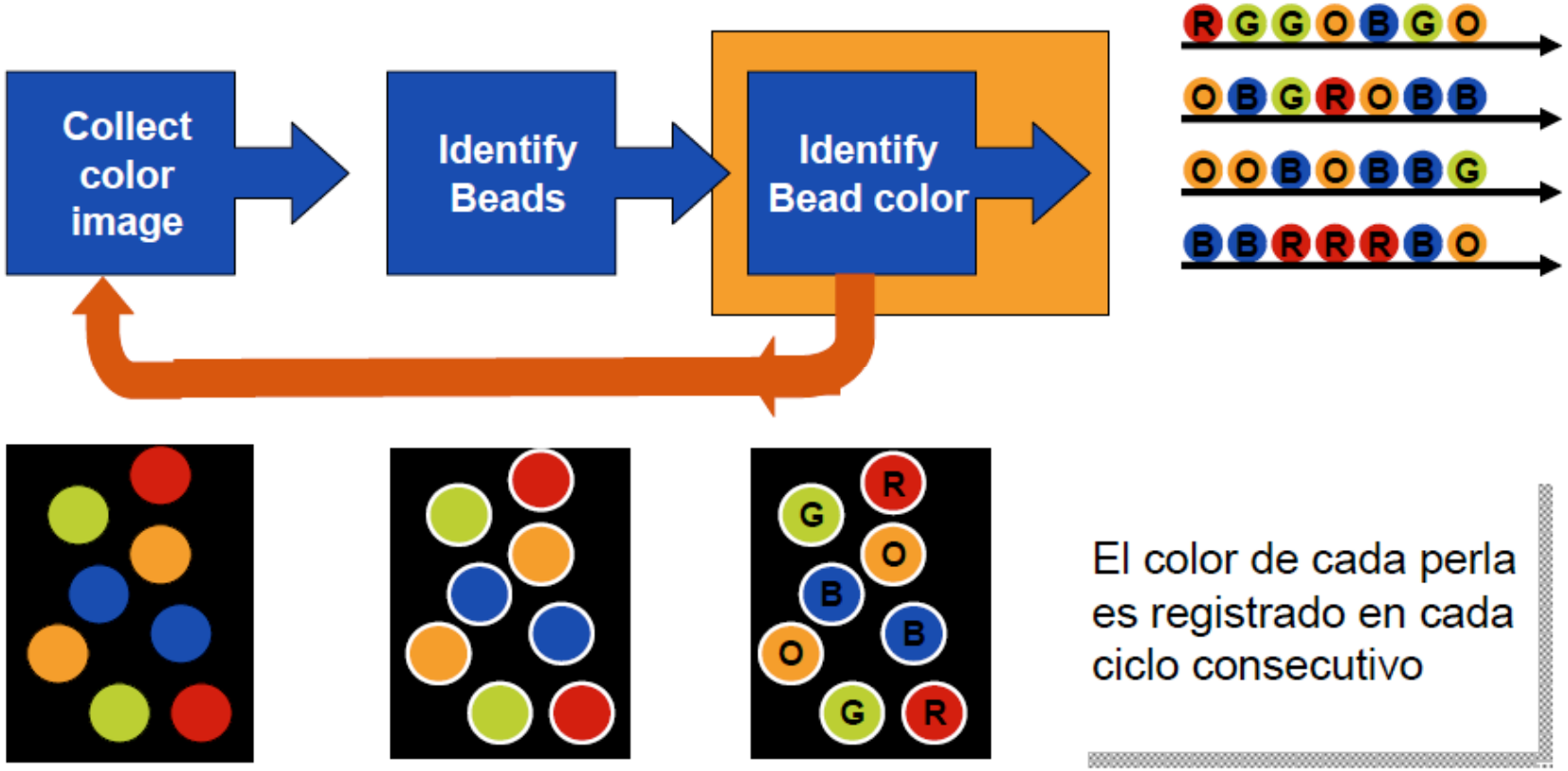






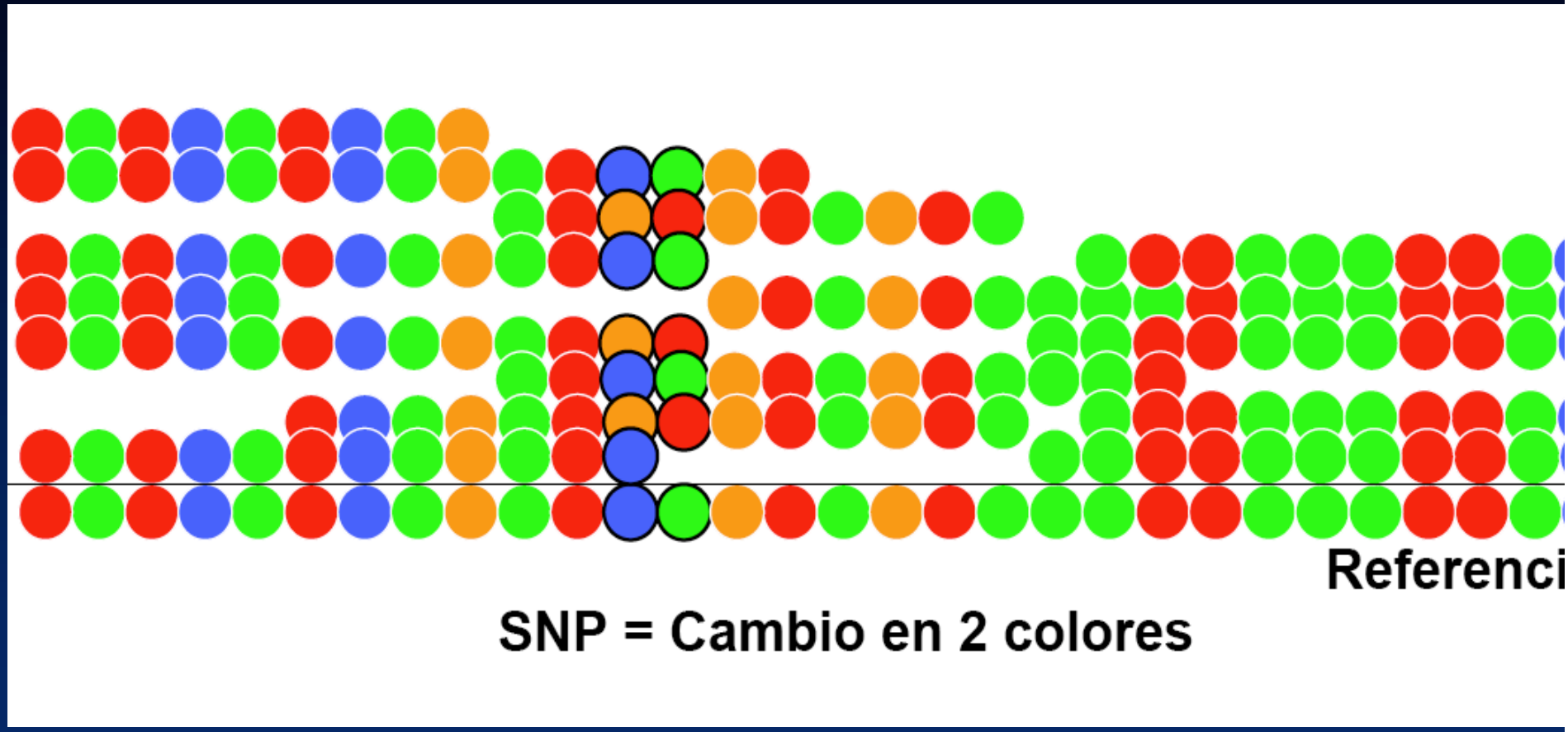


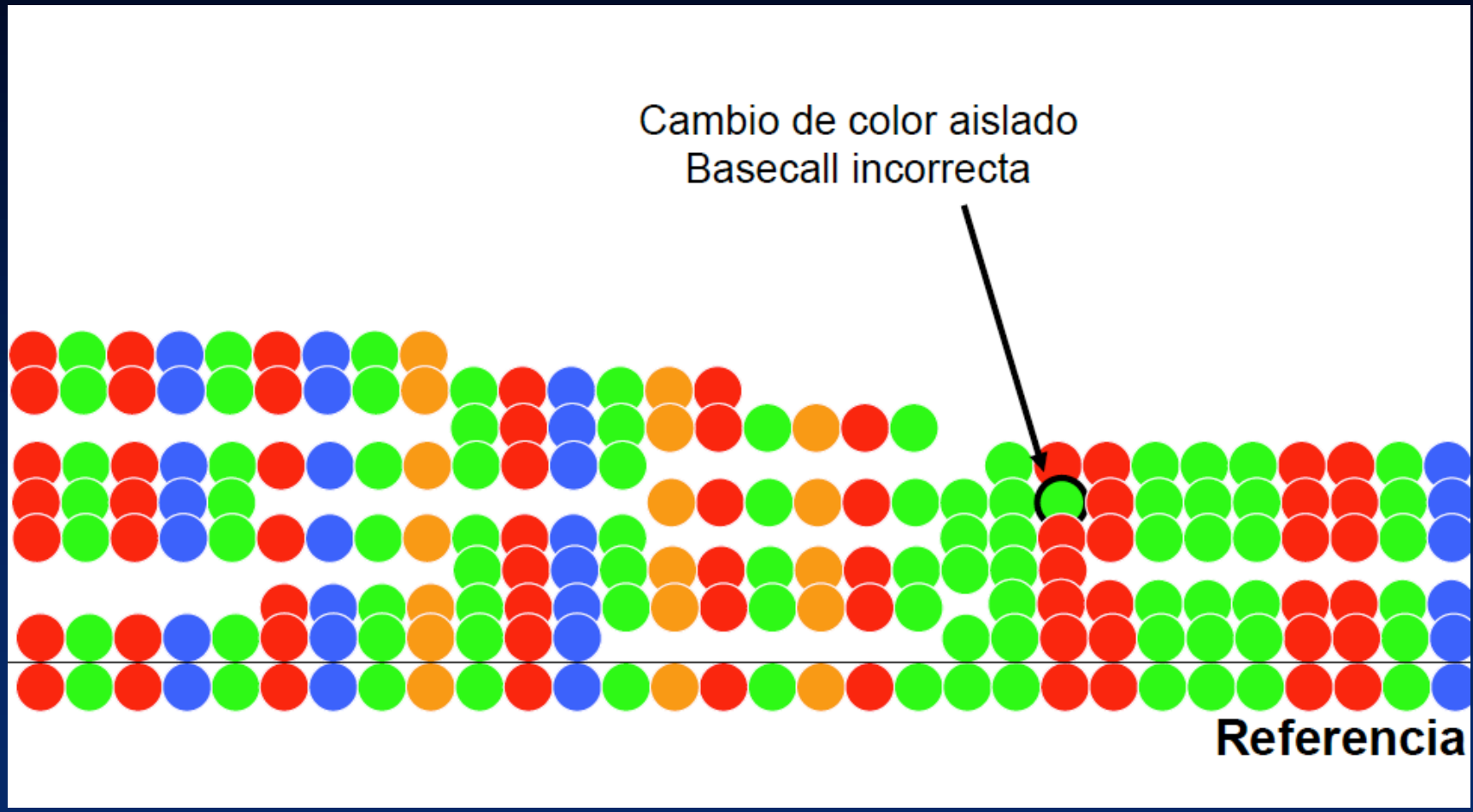
Color Space – Dual Base encoding



El color de cada perla es registrado en cada ciclo consecutivo

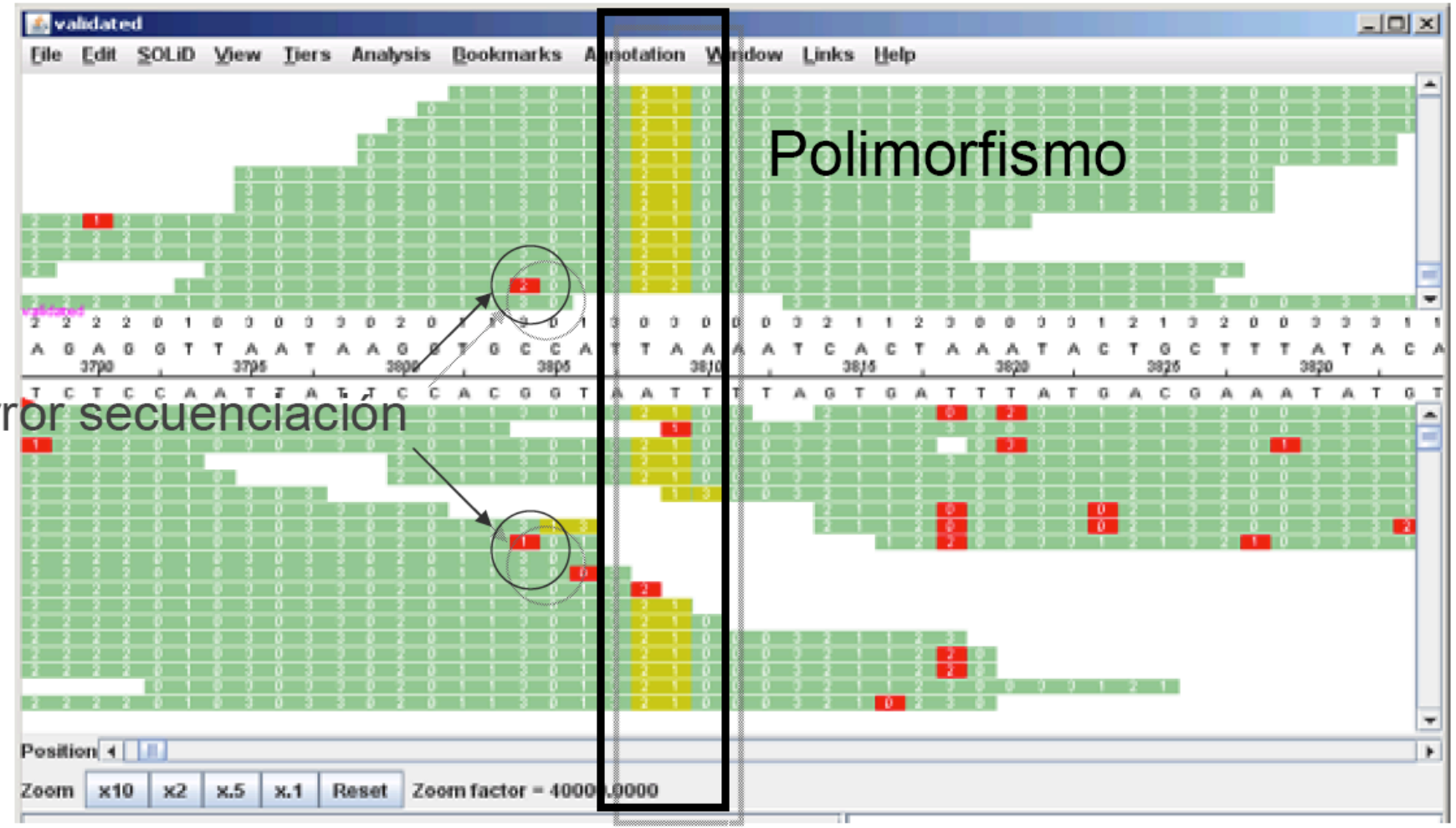
Resultados

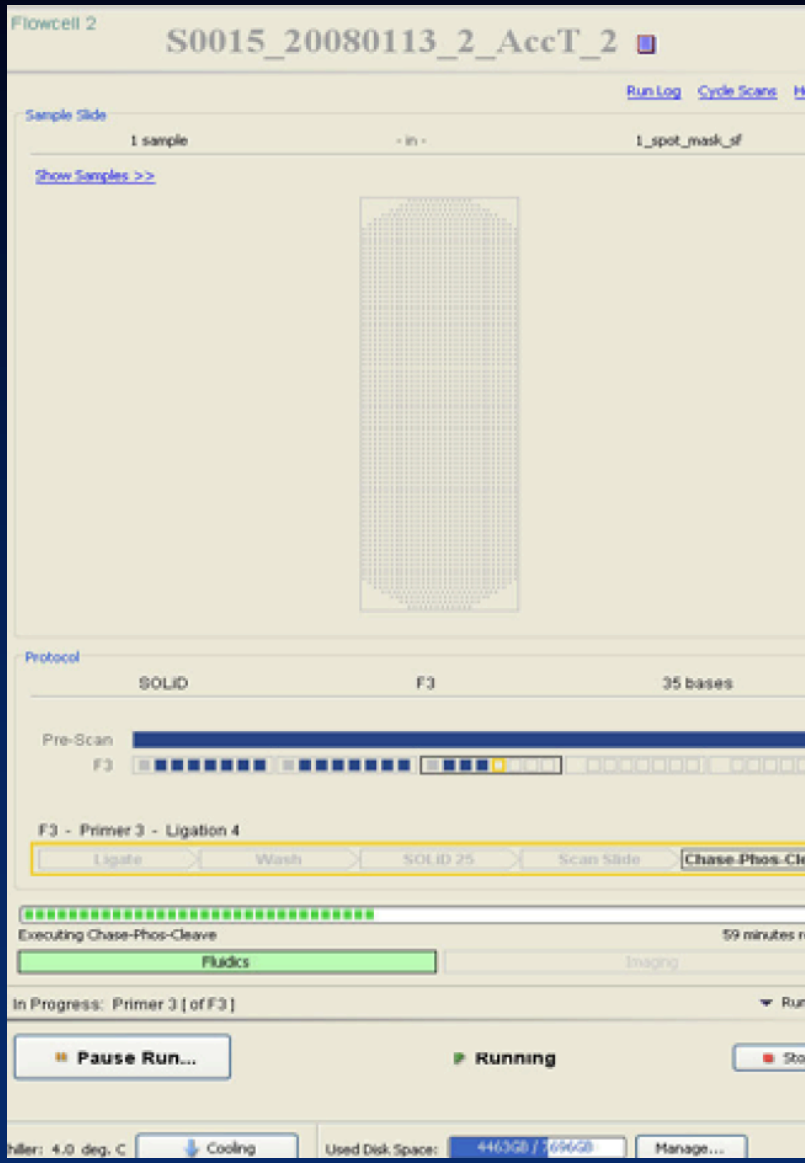


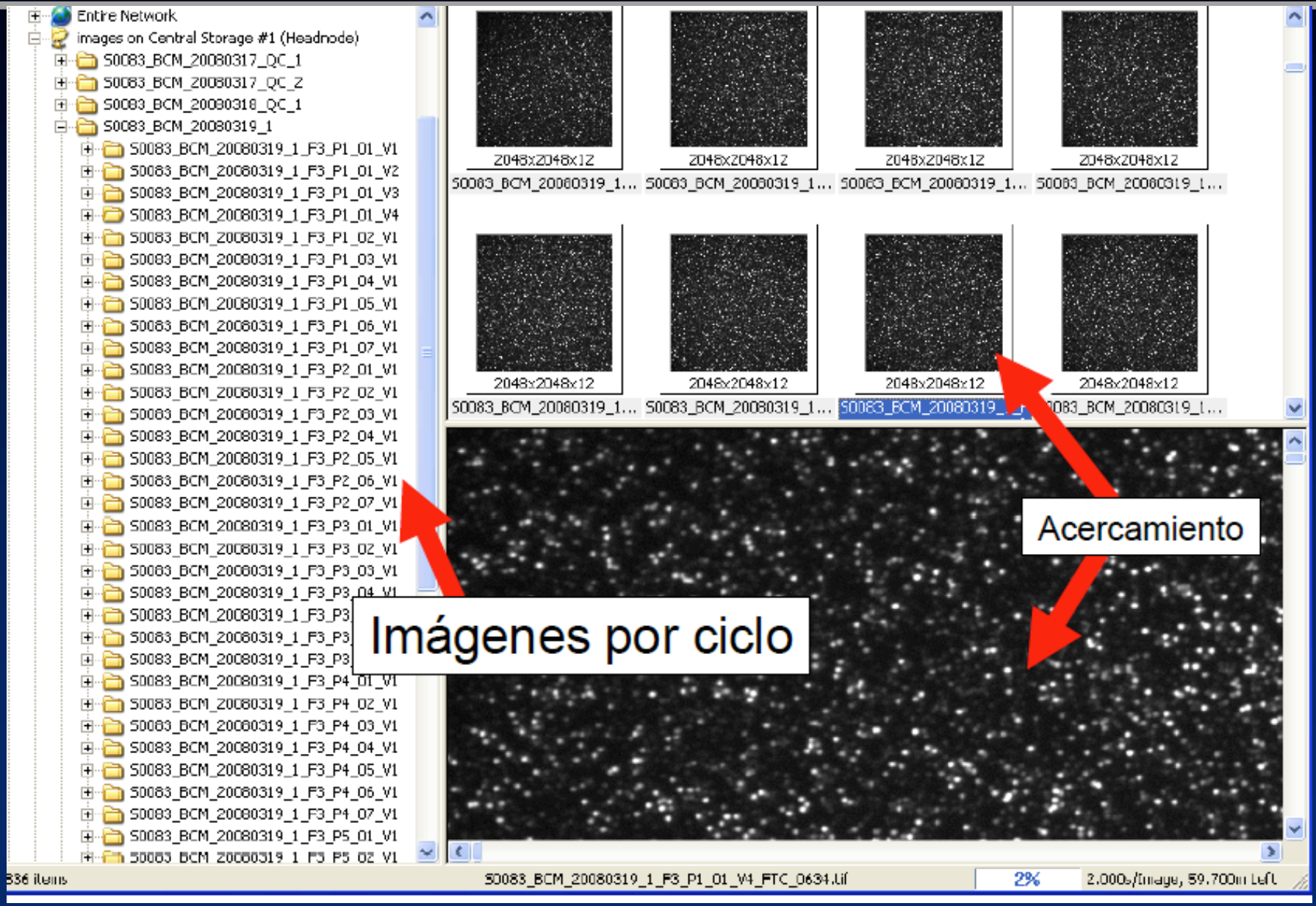

Cambio de color aislado
Basecall incorrecta
Referencia

Resultados

Ejemplo de SNPs identificados en SOLiD™ Alignment Brow



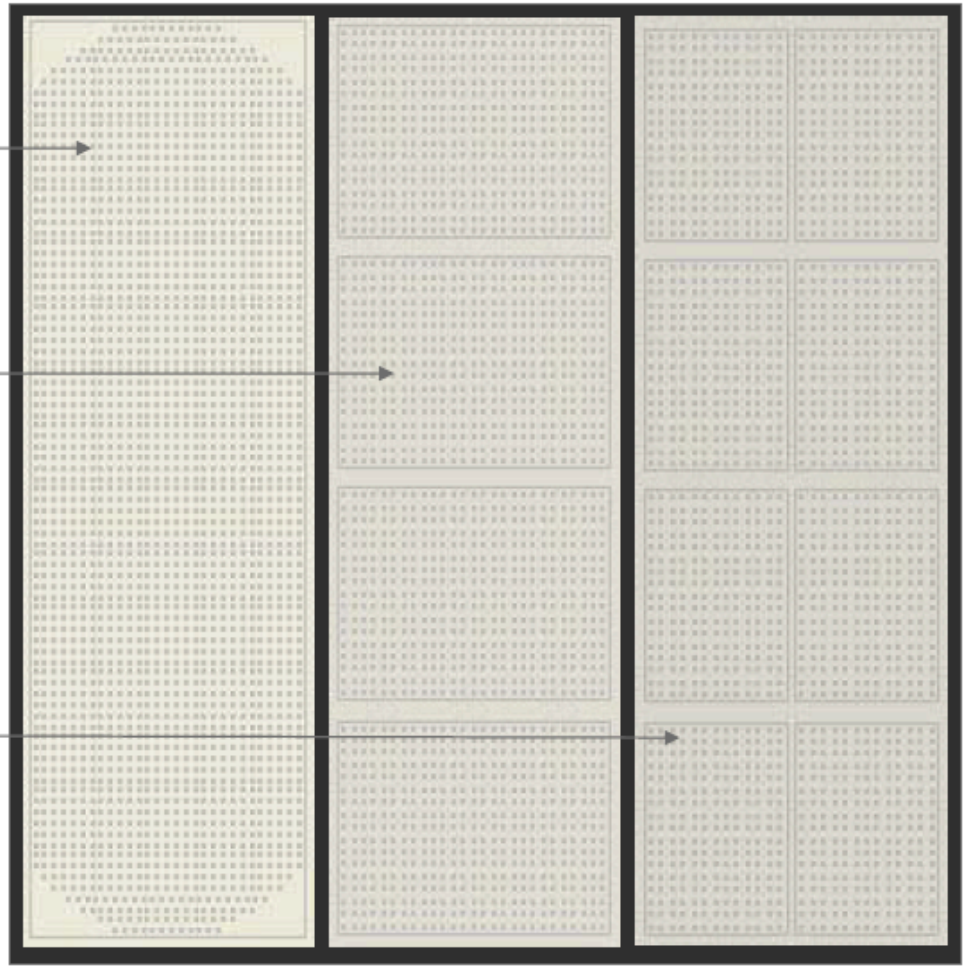

Flowcell 2
S0015_20080113_2_AccT_2
Run Log
Cycle Scans
Sample Slide
1 sample
- in -
1_spot_mask_sf
Show Samples >>
Protocol
SOLiD
F3
35 bases
Pre-Scan
F3
F3 - Primer 3 - Ligation 4
Ligate
Wash
SOLiD 25
Scan Slide
Executing Chase-Phos-Cleave
Fluidics
Imaging
In Progress: Primer 3 [of F3]
Pause Run...
Running
Cooling
Used Disk Space:
Manage...

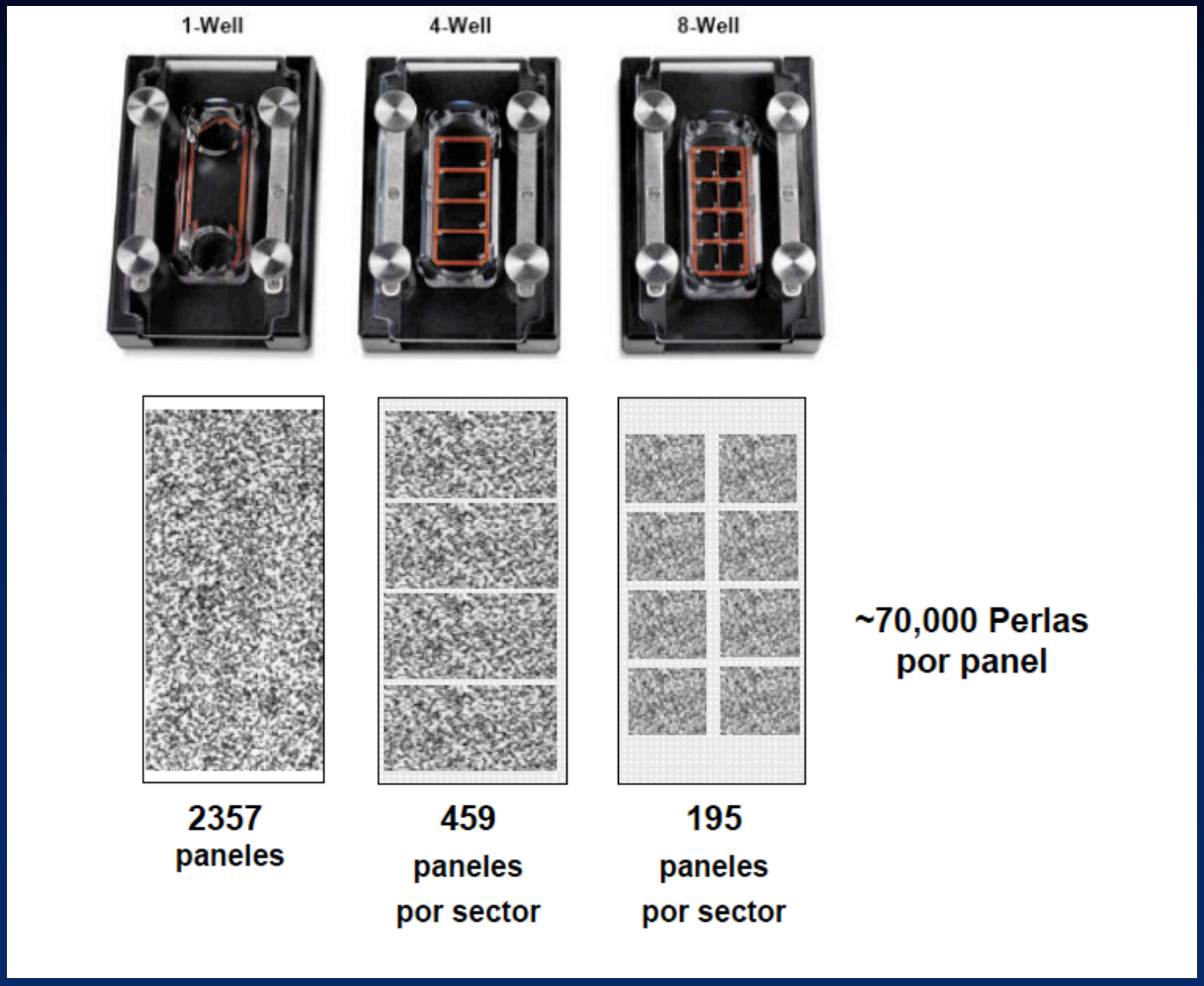

Entire Network
images on Central Storage #1 (Headnode)
S0083_BCM_20080317_QC_1
S0083_BCM_20080317_QC_2
S0083_BCM_20080318_QC_1
S0083_BCM_20080319_1
S0083_BCM_20080319_1_F3_P1_01_V1
S0083_BCM_20080319_1_F3_P1_01_V2
S0083_BCM_20080319_1_F3_P1_01_V3
S0083_BCM_20080319_1_F3_P1_01_V4
S0083_BCM_20080319_1_F3_P1_02_V1
S0083_BCM_20080319_1_F3_P1_03_V1
S0083_BCM_20080319_1_F3_P1_04_V1
S0083_BCM_20080319_1_F3_P1_05_V1
S0083_BCM_20080319_1_F3_P1_06_V1
S0083_BCM_20080319_1_F3_P1_07_V1
S0083_BCM_20080319_1_F3_P2_01_V1
S0083_BCM_20080319_1_F3_P2_02_V1
S0083_BCM_20080319_1_F3_P2_03_V1
S0083_BCM_20080319_1_F3_P2_04_V1
S0083_BCM_20080319_1_F3_P2_05_V1
S0083_BCM_20080319_1_F3_P2_06_V1
S0083_BCM_20080319_1_F3_P2_07_V1
S0083_BCM_20080319_1_F3_P3_01_V1
S0083_BCM_20080319_1_F3_P3_02_V1
S0083_BCM_20080319_1_F3_P3_03_V1
S0083_BCM_20080319_1_F3_P3_04_V1
S0083_BCM_20080319_1_F3_P4_01_V1
S0083_BCM_20080319_1_F3_P4_02_V1
S0083_BCM_20080319_1_F3_P4_03_V1
S0083_BCM_20080319_1_F3_P4_04_V1
S0083_BCM_20080319_1_F3_P4_05_V1
S0083_BCM_20080319_1_F3_P4_06_V1
S0083_BCM_20080319_1_F3_P4_07_V1
S0083_BCM_20080319_1_F3_P5_01_V1
2048x2048x12
Imágenes por ciclo
Acercamiento
S0083_BCM_20080319_1_F3_P1_01_V4_FTC_0634.tif
2%

Configuración de Paneles

3 opciones

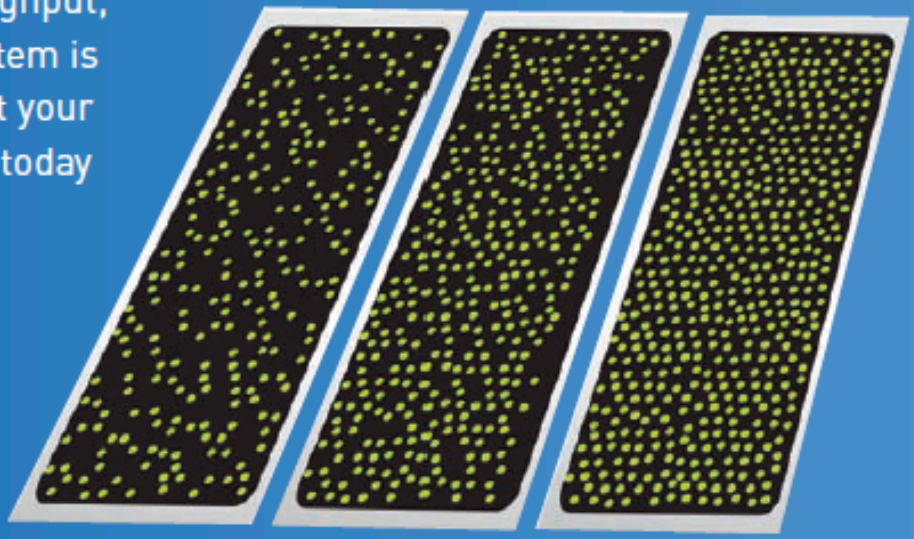
- Plantilla completa
  - 2357 paneles
- 4 Plantillas
  - 435 paneles/spot
  - Opción única para corrida de titulación/QC
- 8 Plantillas
  - 195 paneles/spot

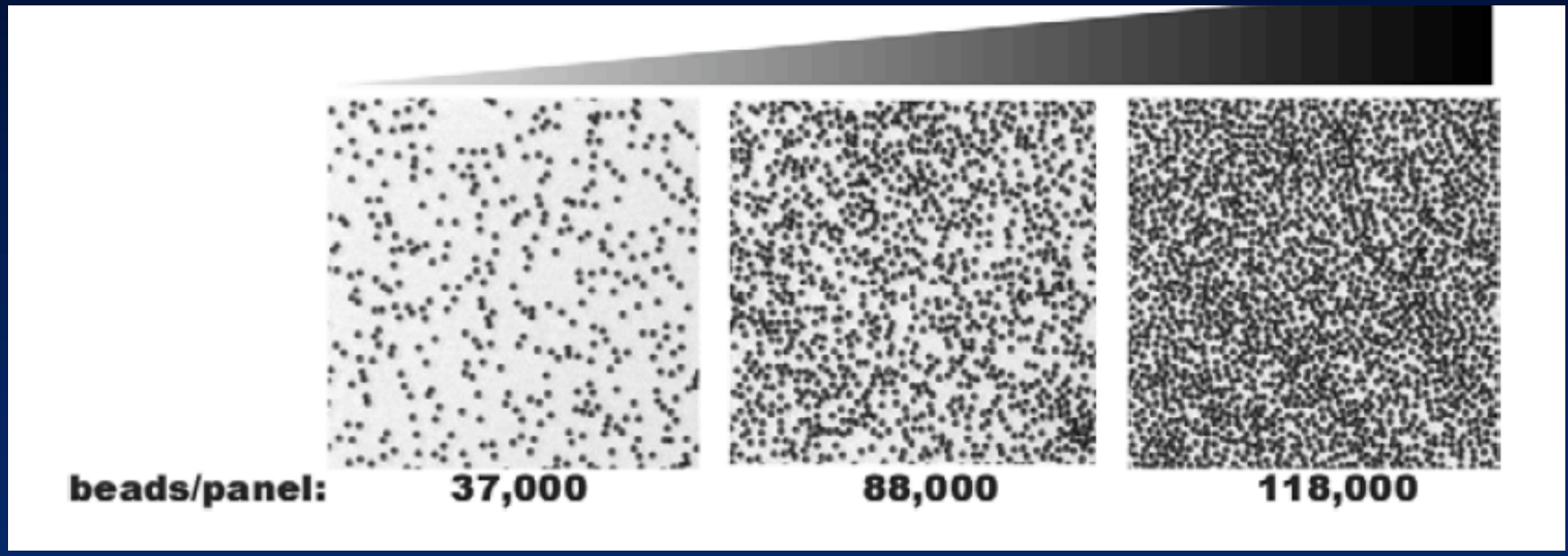



1-Well
4-Well
8-Well
2357
paneles
459
paneles
por sector
195
paneles
por sector
~70,000 Perlas
por panel

System Scalability

The SOLiD 3 System's open slide format and flexible bead densities enable increases in throughput with modest analysis and chemistry optimizations. While competitive technologies are already near maximum throughput, the SOLiD 3 System is scalable to meet your research needs today and tomorrow.




beads/panel:
37,000
88,000
118,000

Ventajas

- **Química de ligación**
  - No se observan problemas con homopolímeros
- **Etapa de Defosforilación elimina desfases**
  - Incrementa el Control de señal de ruido en ciclos posteriores
- **Primer reset**
  - Remoción de ligaciones acumuladas que estuvieran fuera de fas
  - Permite lecturas mas largas (Recuperación de señal)

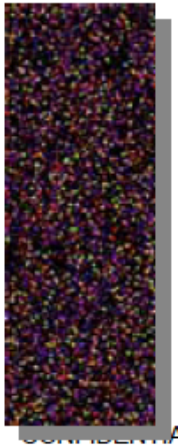
Ventajas

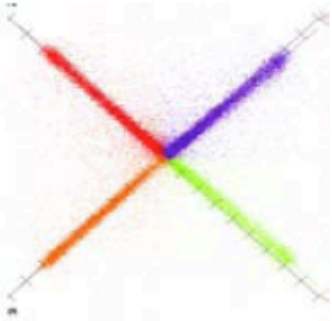
- **Ligación**
  - Alta fidelidad
  - Elevada calidad de lectura

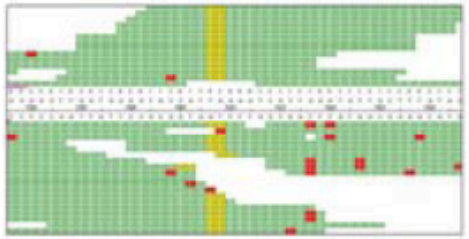
- **Doble interrogación de cada base**
  - Exclusivo del método de ligación
  - Incrementa la precisión en la asignación de bases
  - SNP resulta en el cambio de dos colores

SOLiD™ Software Suite Data Analysis and Management

	Primary Analysis	Secondary Analysis	Tertiary Analysis
Analysis	• Image acquisition and bead processing • Quality metrics • Color calls	• Filtering • Alignment to reference genome • Consensus calling • Difference report	• Visualization • Disease/ Gene annotations • Application specific • Gene Expression • Methylation • ChIP
Tools	SOLiD™ Analysis Tools SOLiD™ Experimental Tracking System	SOLiD™ Analysis Tools SOLiD™ Experimental Tracking System	SOLiD™ Alignment Browser 3rd Party Software Providers







GenomeQuest
One Source for Sequence Search, Content, and Analysis

Geospiza®

SOLID Analysis

Análisis Primario

- SOLID Experimental Tracking System (SETS)
  - Adquisición de imagen y procesamiento de perlas
  - Color calling
  - Medida de Calidad
  - Análisis en tiempo real con monitoreo remoto

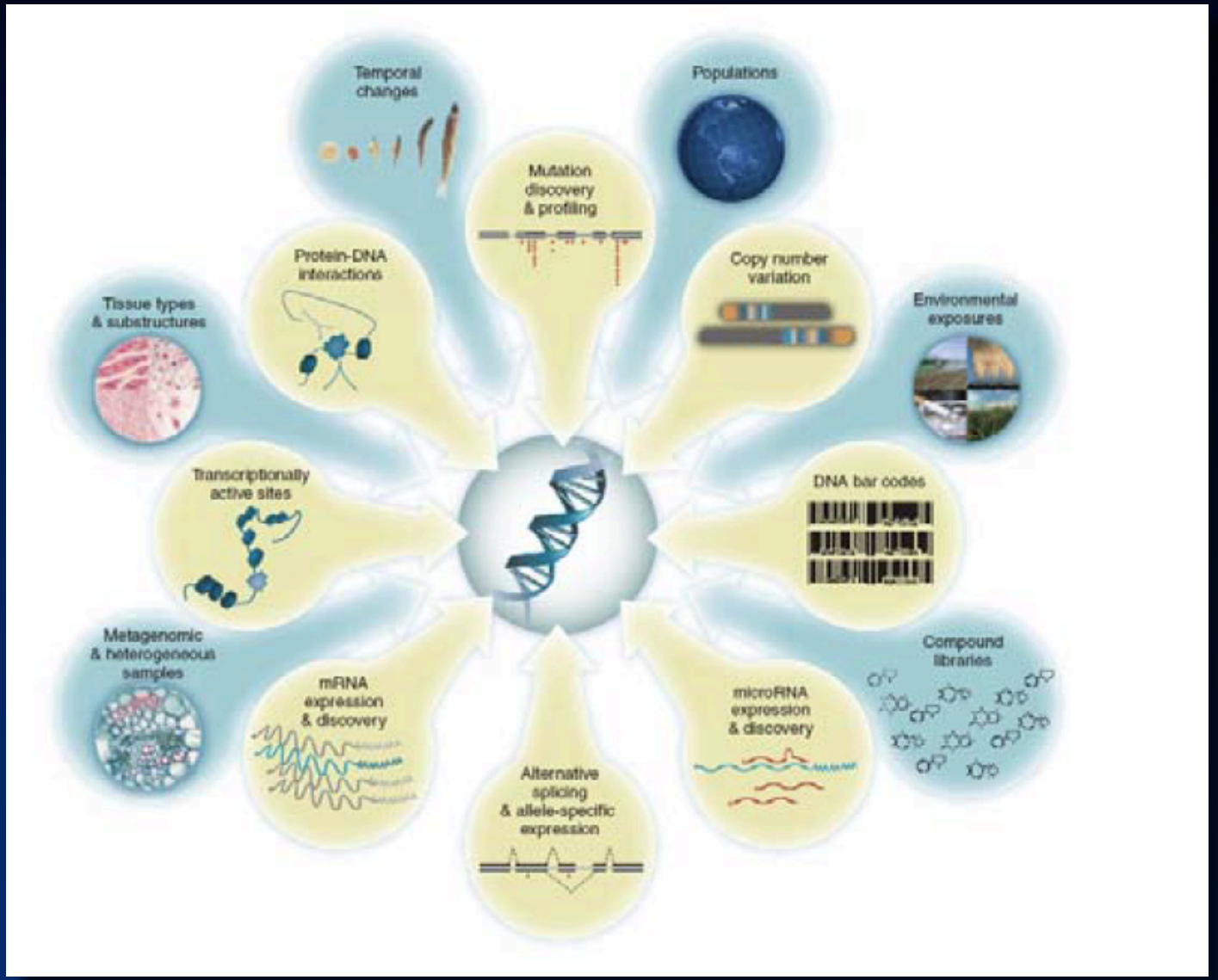
Análisis Secundario

- SOLID Experimental Tracking System (SETS)
  - Filtrado
  - Análisis de Mate pair / fragmentos
  - Alineación y mapeo al genoma de referencia
  - Consensus calling
  - Conversión de color a bases
  - SOLID Alignment Browser
  - Visualización

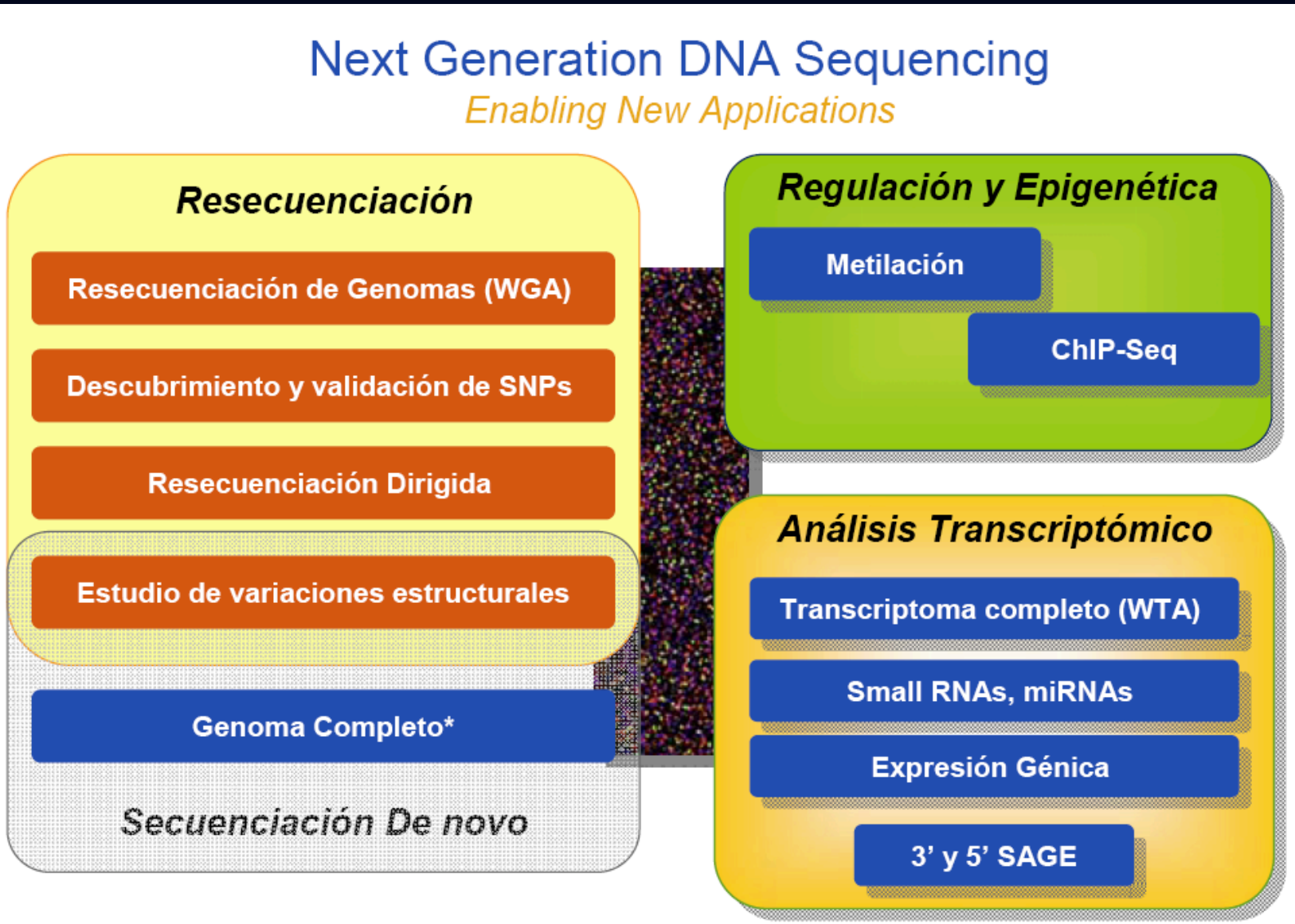
Análisis Terciario

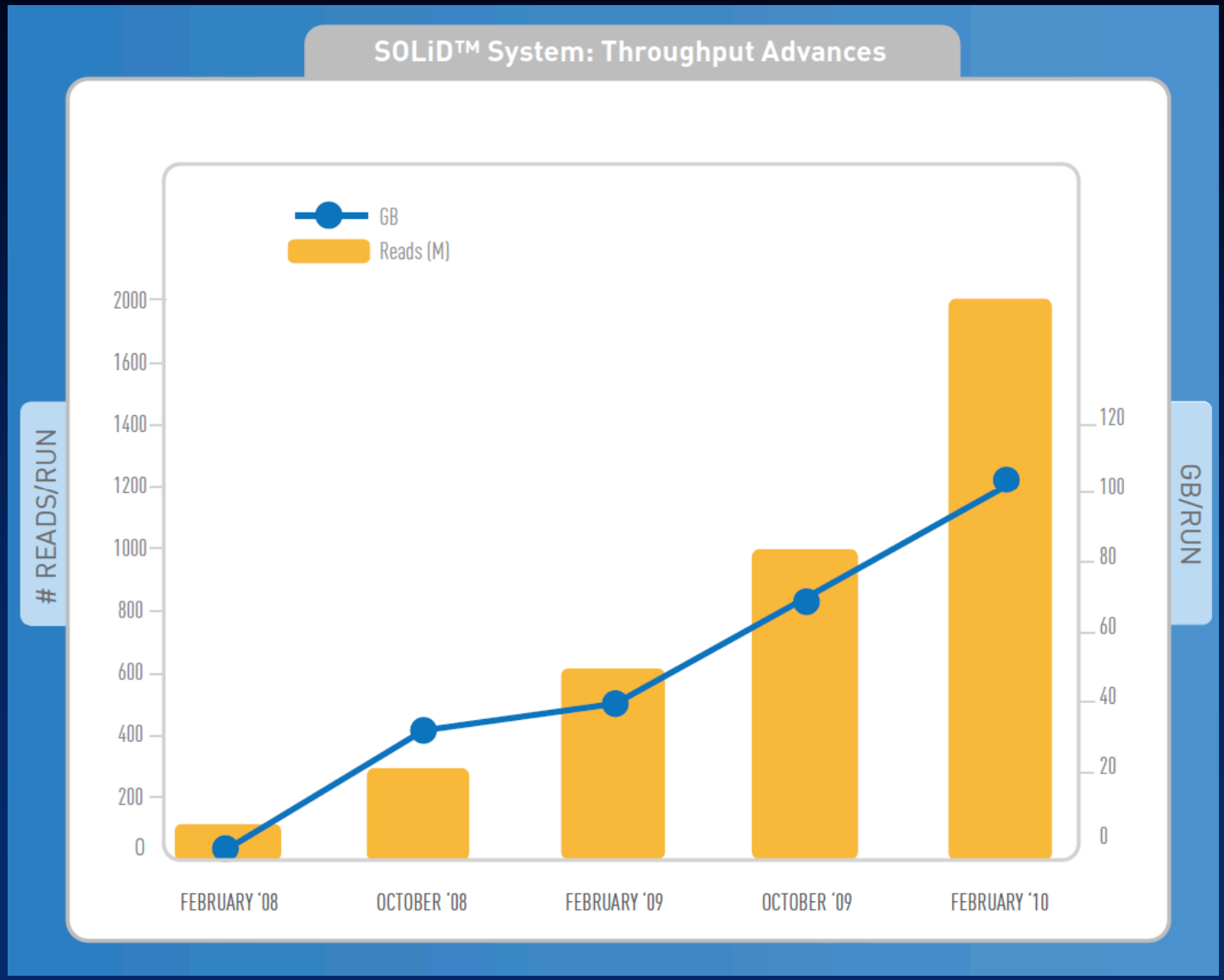
- 3rd party tools – application specific
  - Disease/ Gene Annotations
  - Gene Expression
  - Methylation
  - ChIP

Aplicaciones de secuenciación masiva



Aplicaciones SOLiD




SOLiD™ System: Throughput Advances
GB
Reads (M)
READS/RUN
GB/RUN
2000
1600
1400
1200
1000
800
600
400
200
0
120
100
80
60
40
20
0
FEBRUARY '08
OCTOBER '08
FEBRUARY '09
OCTOBER '09
FEBRUARY '10

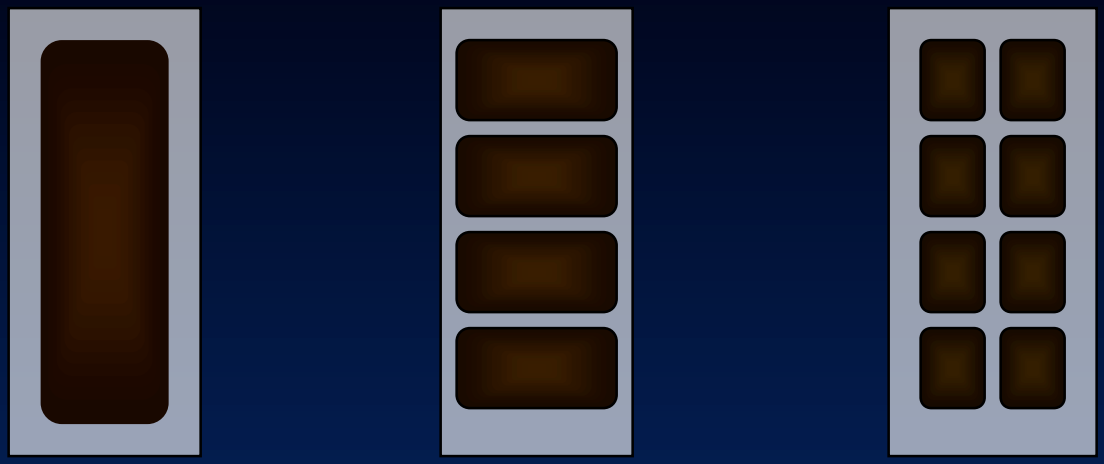
	FRAGMENT LIBRARY Using two full slides/run				MATE-PAIRED LIBRARY Using two full slides/run			
	2007	2008	2009	2010	2007	2008	2009	2010
Read Length	35	35	50	100	25	35	50	100
Total Output*	~2.5 GB	~4 GB	~30 GB	>50 GB	~3.5 GB	~12 GB	~60 GB	>100 GB

* Mappable data from 2 slide run

APLICACIÓN	COVERTURA MAX RECOMENDADA POR ABI
Descubrimiento SNP, WGS	20X
Descubrimiento Small RNA	10 millones de Tags
Transcriptoma	50 millones de Tags
SAGE	50 millones de Tags
ChIP-Seq	5 millones de Tags
Metilacion	10 millones Tags

Cobertura Promedio V3 Plus 2 Laminillas

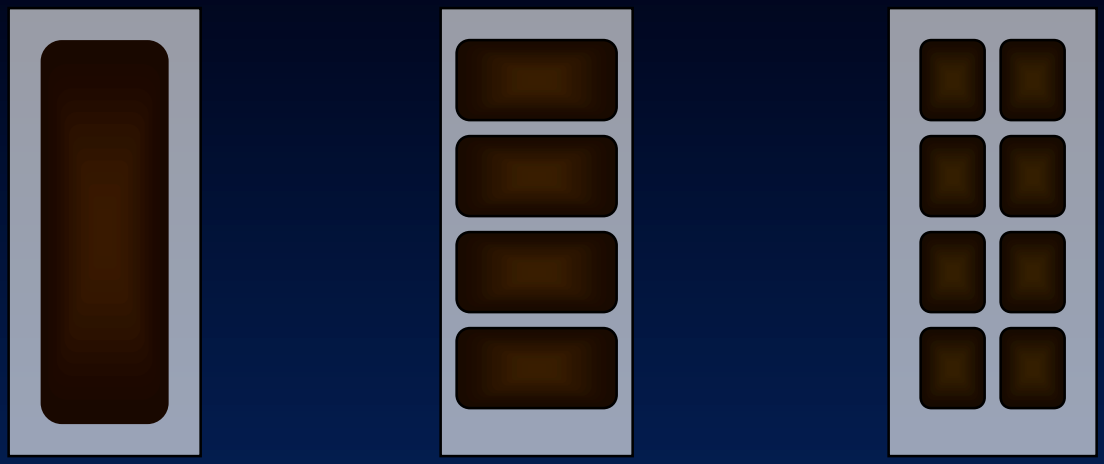
Por corrida de
2 laminillas



Fragmentos 50pb	27,500 Mb 550 Millones	5,000 Mb 100 Millones	2,200 Mb 43 Millones	**Bas** **Tag**
Mate-Pair 2 x 50pb	55,000 Mb 1,100 Millones	10,000 Mb 200 Millones	4,400 Mb 86 Millones	**Bas** **Ta**

Cobertura Promedio V3 Plus 1 Laminilla

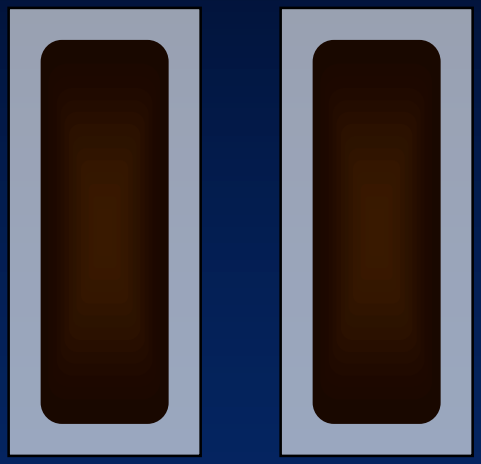
Por corrida de
1 laminilla



Fragmentos 50pb	13,750 Mb 275 Millones	2,500 Mb 50 Millones	1,100 Mb 21.5 Millones	**Bas** **Tag**
Mate-Pair 2 x 50pb	27,500 Mb 550 Millones	5,000 Mb 100 Millones	2,200 Mb 43 Millones	**Bas** **Ta**

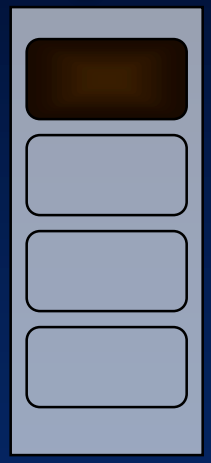
Genoma Humano
1 individuo

Mate-Pair



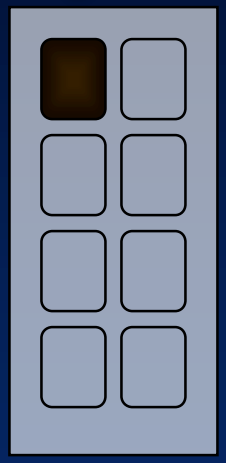
Trancriptoma
1 muestra

Fragmentos



Small RNA
2 muestras

Fragmentos



Genoma Bacter
20 muestras

Fragmentos