



UNU
BIOLAC

Cursos Internacional .
Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica
de Enfermedades Transmitidas por Alimentos y Aguas



Ascanio Rojas A.

CeCalCULA – ULA

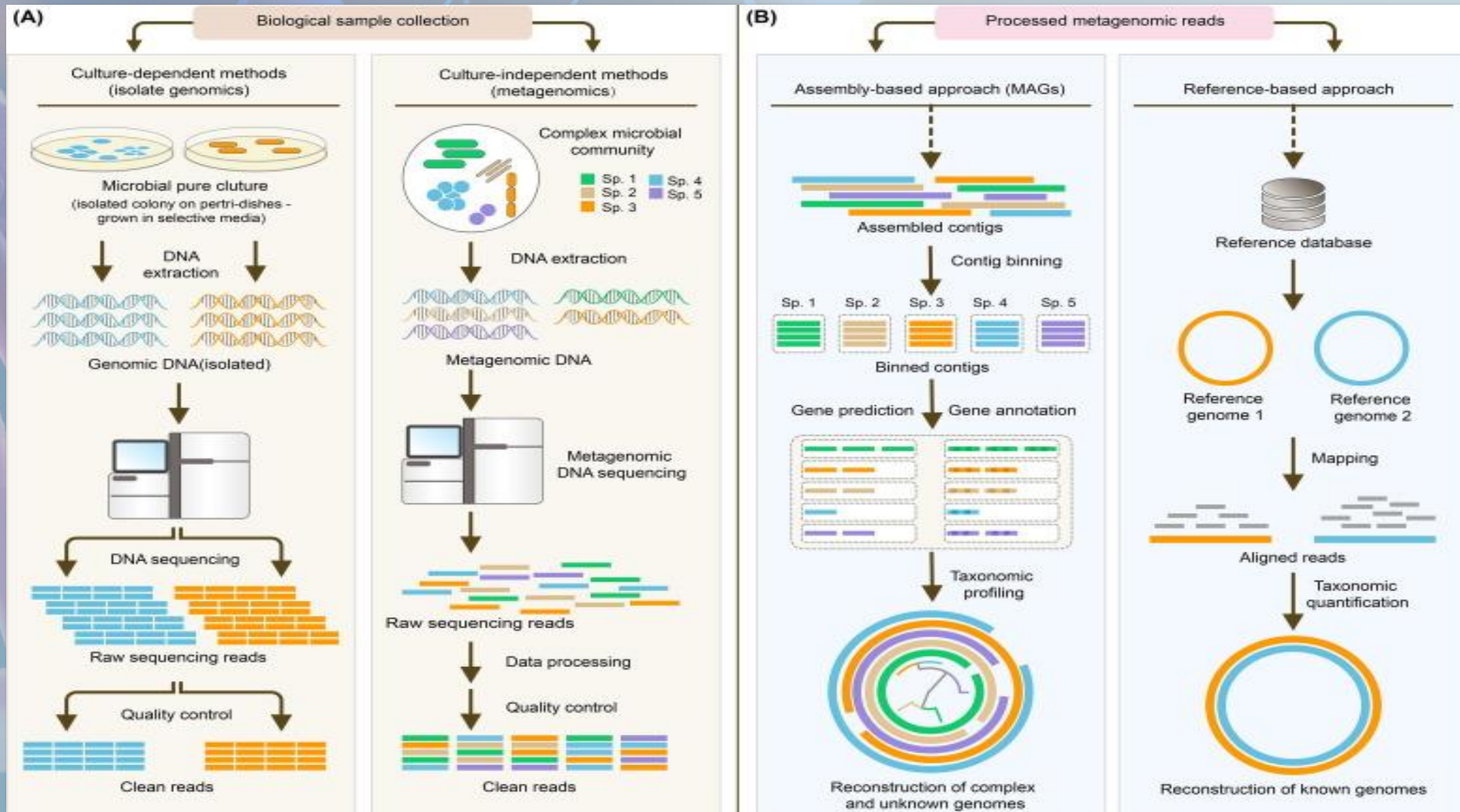
ascanio@ula.ve



Gente, Ciencia y Tecnología al Servicio de la Salud



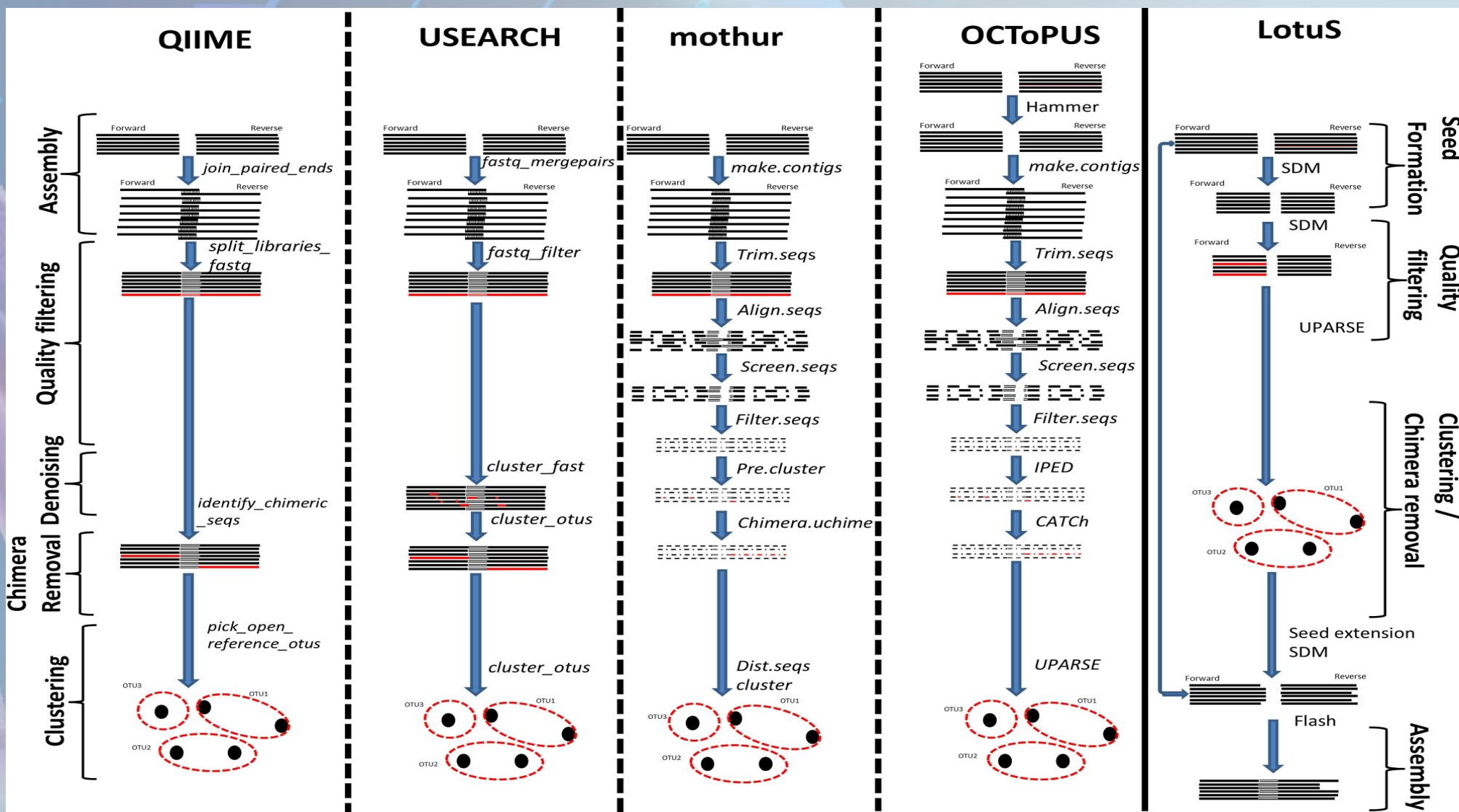
Cursos Internacional . Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica de Enfermedades Transmitidas por Alimentos y Aguas





Cursos Internacional .

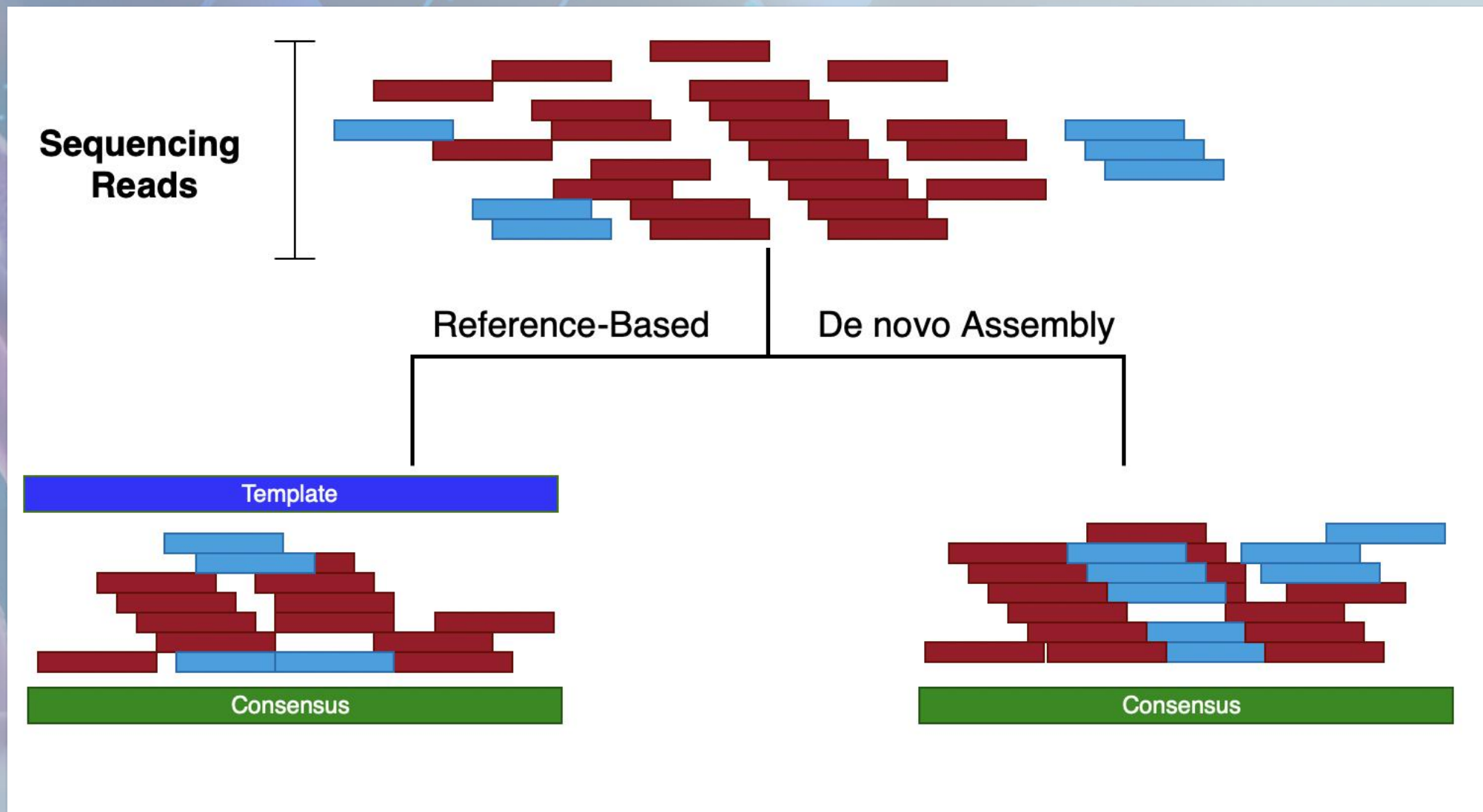
Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica de Enfermedades Transmitidas por Alimentos y Aguas





UNU
BIOLAC

Cursos Internacional .
Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica
de Enfermedades Transmitidas por Alimentos y Aguas



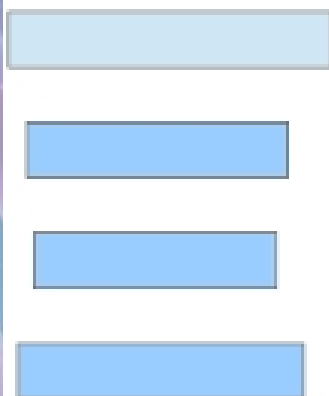


UNU
BIOLAC

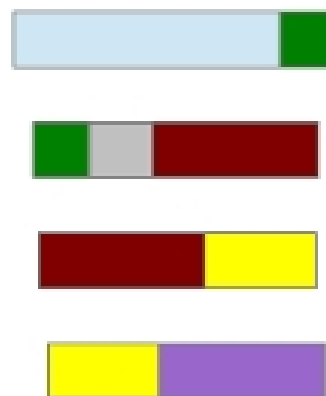
Cursos Internacional .
Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica
de Enfermedades Transmitidas por Alimentos y Aguas

Proceso de ensamblaje

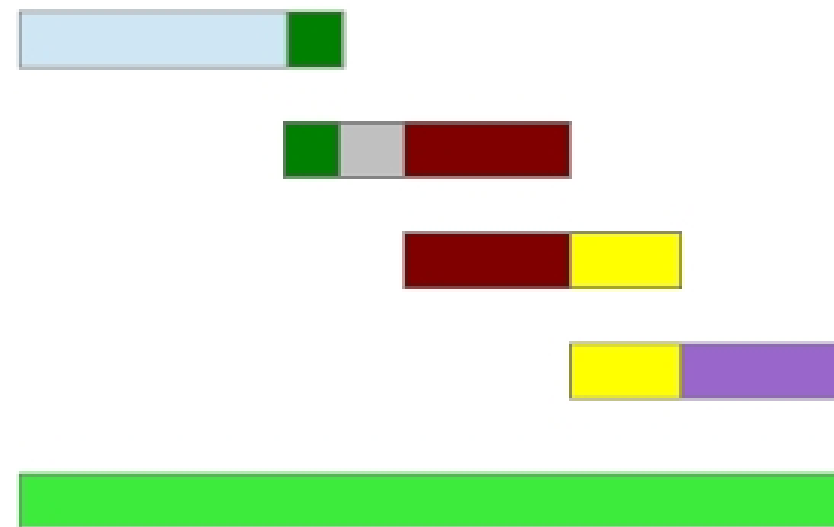
Limpieza



Comparación 2 a 2



Ensamblaje



Consenso



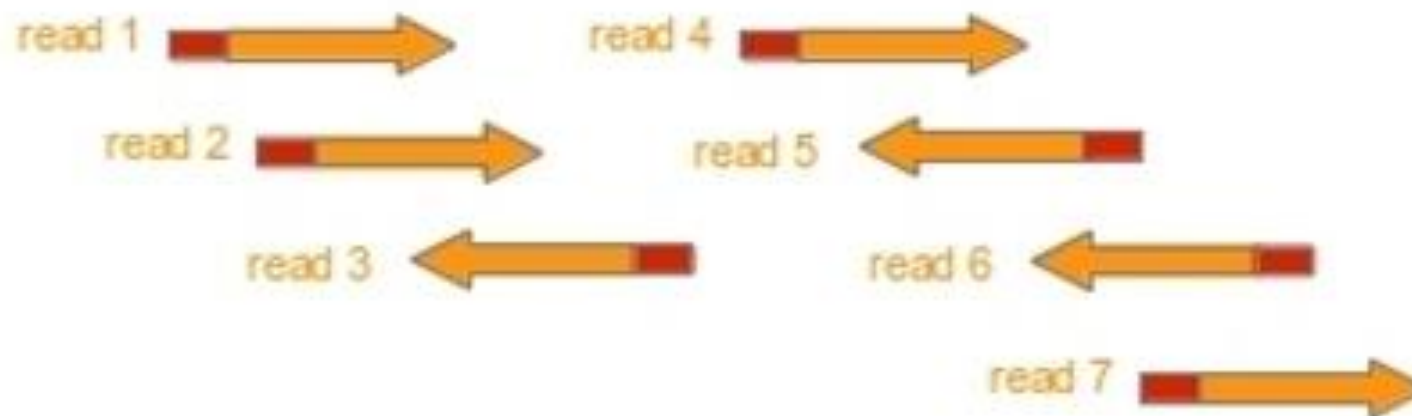
UNU
BIOLAC

Cursos Internacional .
Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica
de Enfermedades Transmitidas por Alimentos y Aguas

Unknown sequence



experimental
evidence

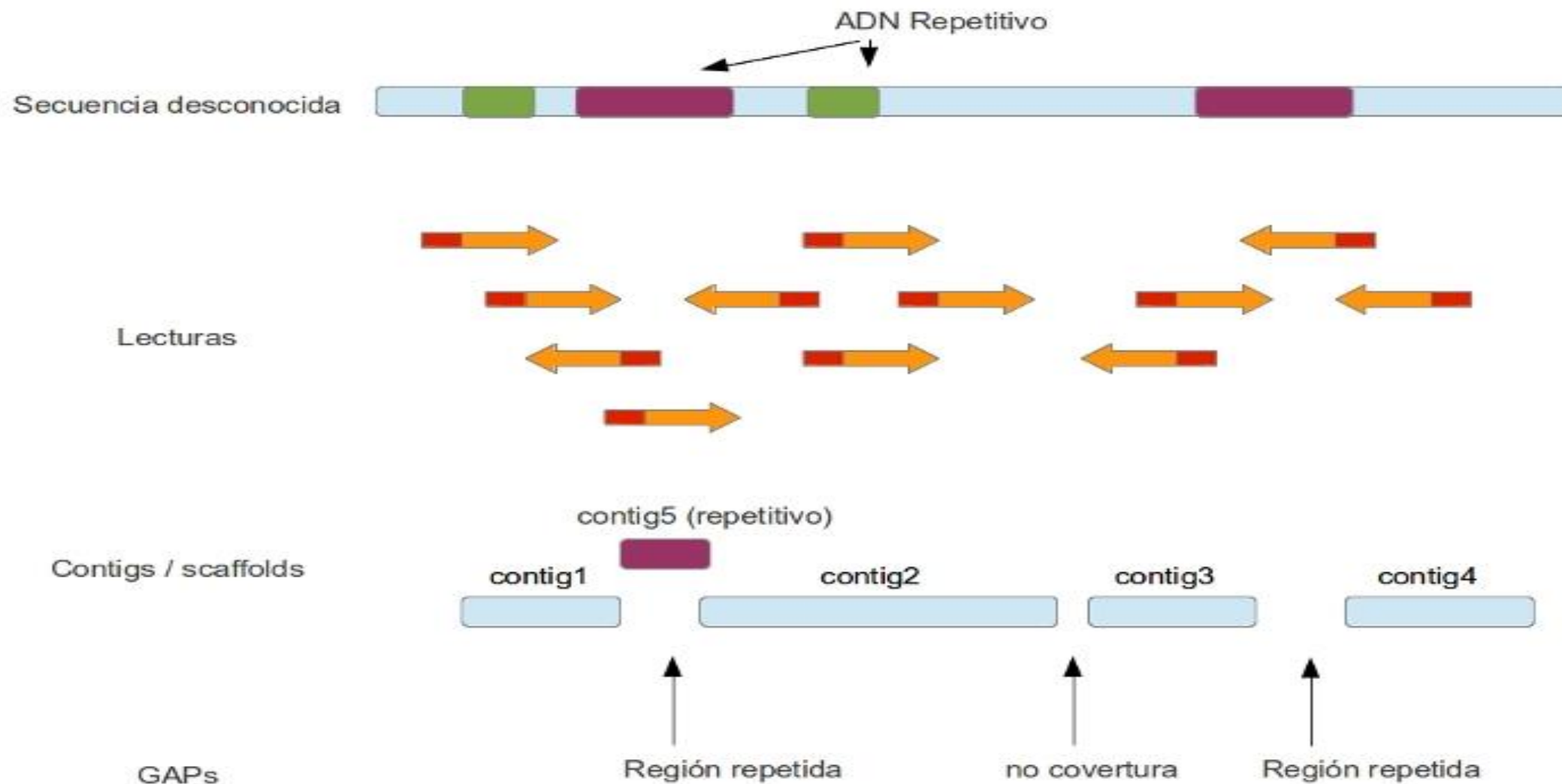


result





El problema del ensamblaje





UNU
BIOLAC

Cursos Internacional .
Secuenciación y Análisis de Datos Genómicos para la Detección Microbiológica
de Enfermedades Transmitidas por Alimentos y Aguas



FASTAQ

@HWI-ST330:304:H045HADXX:1:1101:1111:61397

CACTTGTAAGGGCAGGCCCCCTTCACCCTCCCGCTCCTGGGGGANNNNNNNNNNANNNCGAGGCCCTGGGGTAGAGGGNNNNNNNNNNNGATCTTGG

+

@?@DDDDDDHHH?GH:?FCBGGB@C?DBEGIIIIAEF;FCGGI#####

Línea Descripción

1 Comienza siempre con "@" y a continuación información sobre la lectura

2 La secuencia de ADN real

3 Siempre empieza por "+" y a veces la misma información de la línea 1

4 Tiene una cadena de caracteres que representan las puntuaciones de calidad; debe tener el mismo número de caracteres que la línea 2

Para cada posición se dibuja un gráfico de tipo BoxWhisker. Los elementos del gráfico son los siguientes:

- La línea roja central es la mediana
- La caja amarilla representa el rango intercuartílico (25-75%).
- Los bigotes superior e inferior representan los puntos del 10% y el 90%.
- El eje y del gráfico muestra las puntuaciones de calidad. Cuanto mayor es la puntuación, mejor es la llamada base.
- El fondo del gráfico divide el eje y en lecturas de muy buena calidad (verde), calidad razonable (naranja) y de mala calidad (rojo).
- La calidad de las lecturas en la mayoría de las plataformas se degrada a medida que avanza la posición

