

Módulo: Expresión diferencial

Bioinformática y Estadística 2

Dra. Evelia Coss

Dra. Alejandra Medina

21 al 24 de Febrero, 2023

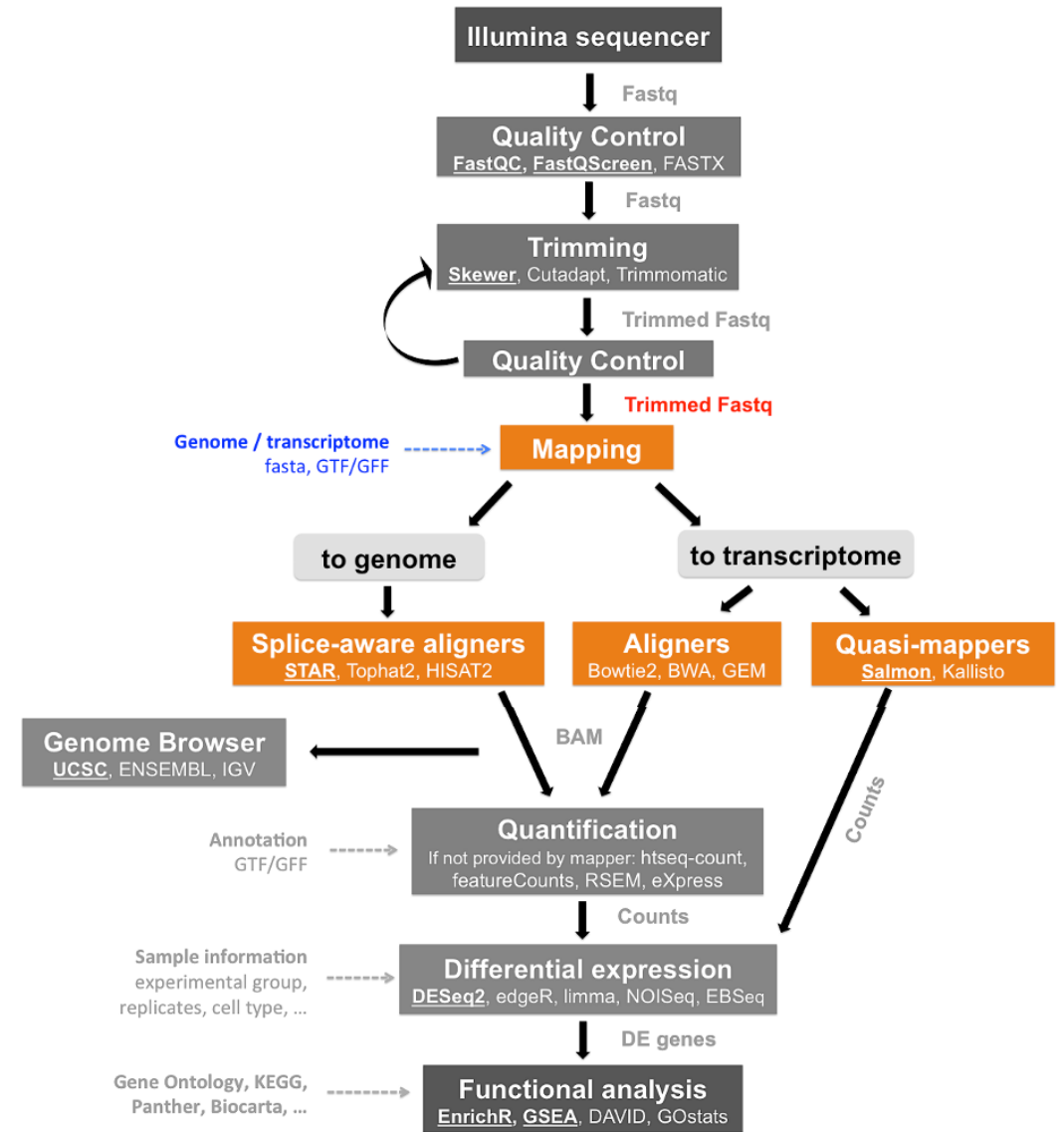
Día 2

- Diversos pipelines bioinformáticos:
 - Alineamiento al genoma de referencia
 - Ensamblaje con el transcriptoma de referencia
 - Ensamblaje *de novo*
- Ejercicio con Kallisto
 - Pseudoalineamiento de kallisto



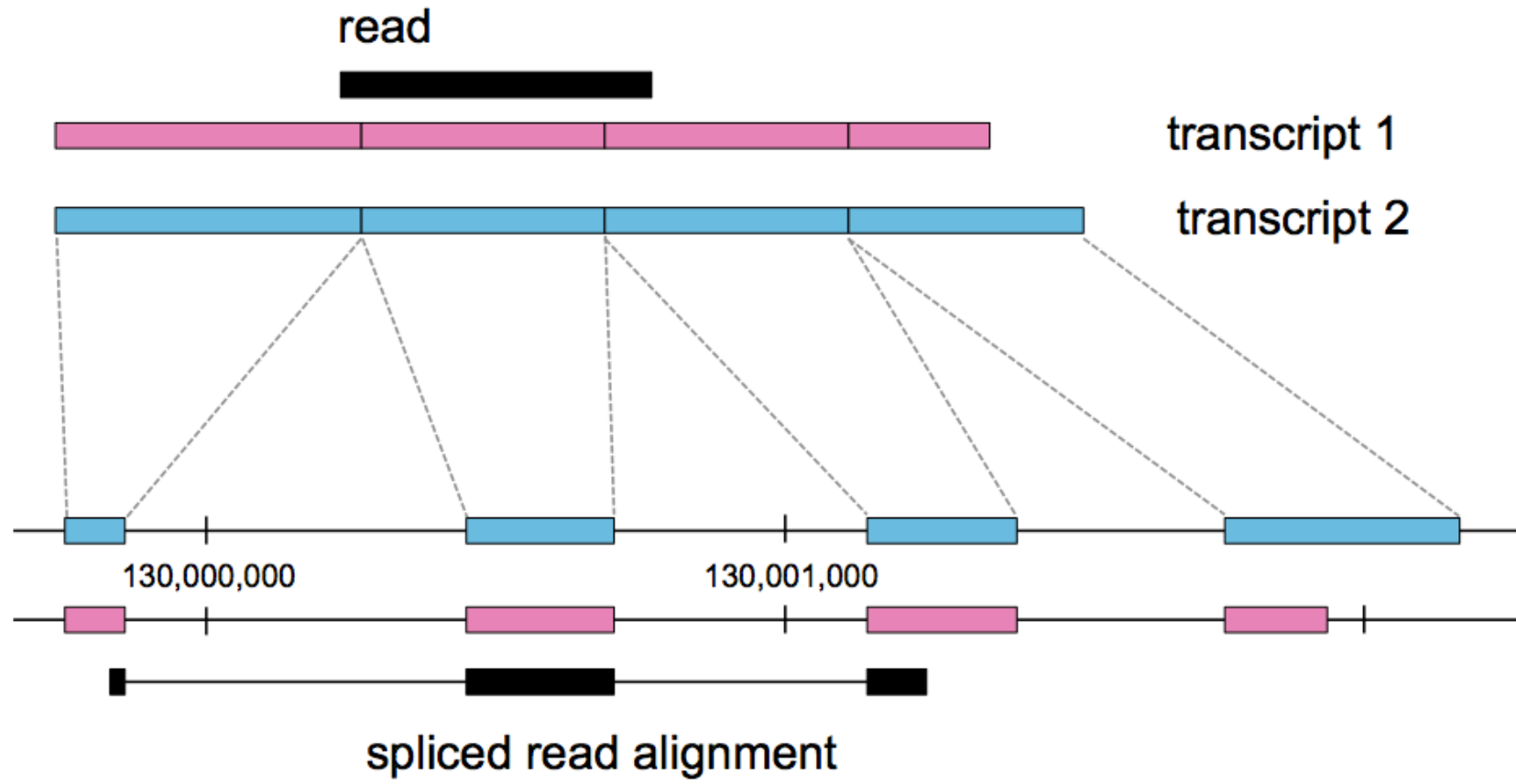
Pipeline bioinformática

Dónde estamos...



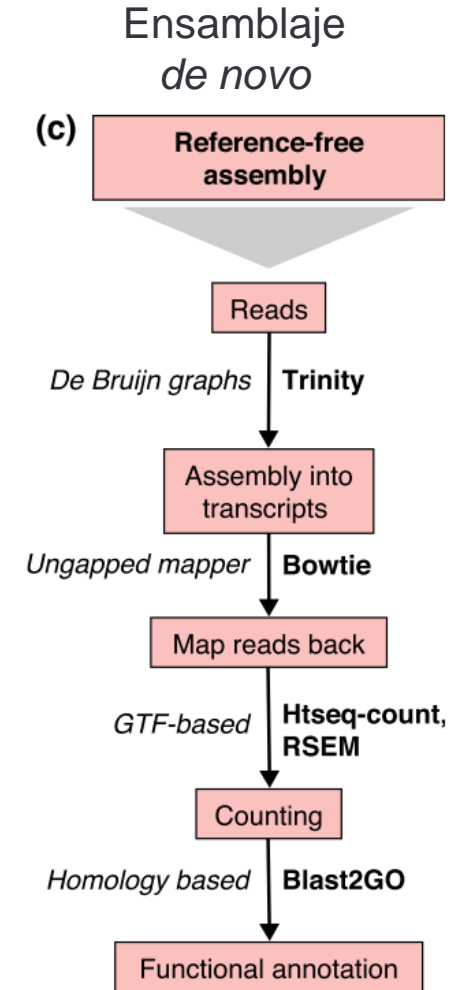
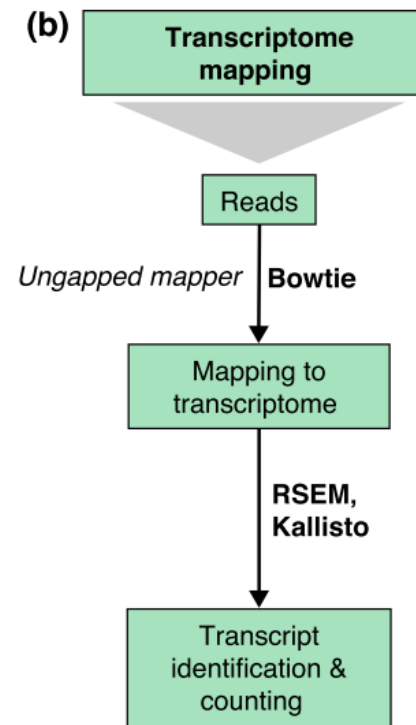
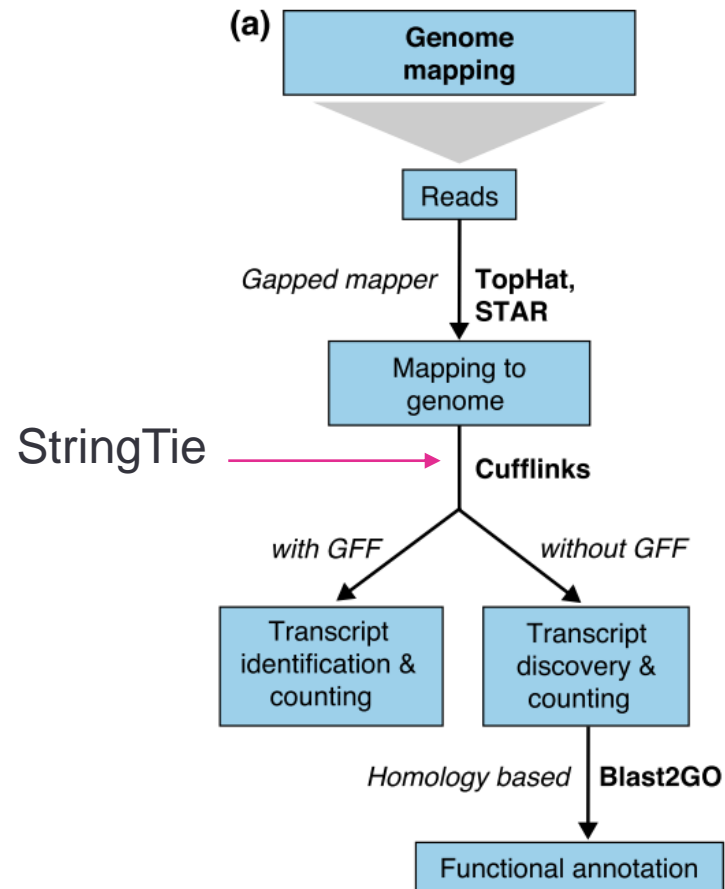
mRNA-Seq data analysis workflow
“https://biocorecrg.github.io/RNAseq_course_2019/workflow.html”

Alineamiento genómico

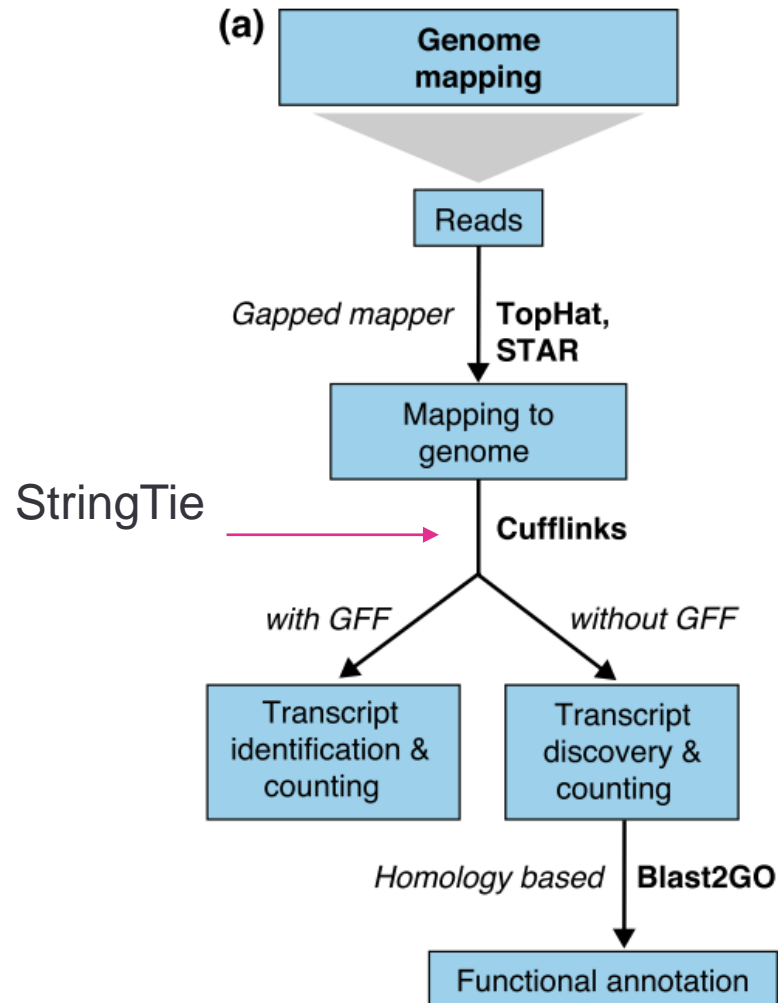


Diversos pipelines

¿Cómo saber qué tipo de algoritmo usar?

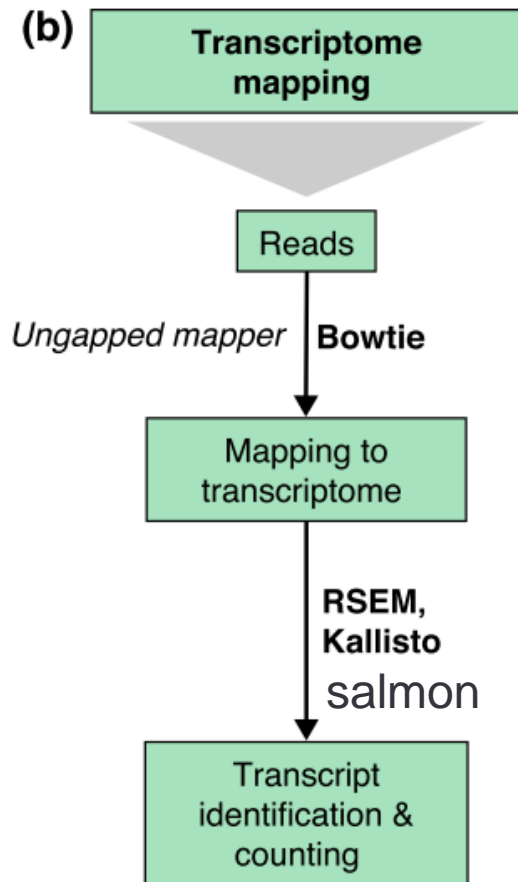


Alineamiento y ensamblaje de lecturas guiado por el genoma de referencia



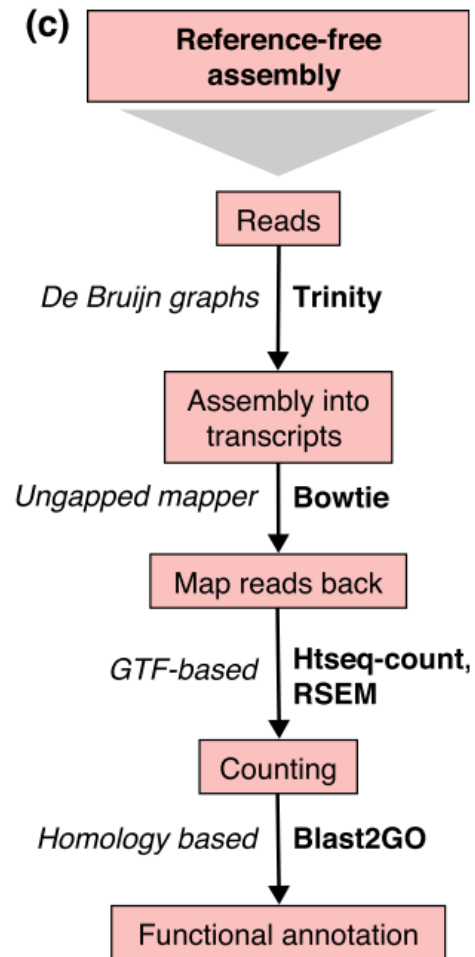
- Podemos anotar nuevos transcritos, así como cuantificarlos.
- Identificación de isoformas.
- Especie con genoma de buena calidad.
- De preferencia contar con un archivo de anotación.
- Empleado normalmente en organismo modelo.

Ensamblaje de transcriptoma guiado



- Expresión por genes, por lo que no vemos isoformas.
- NO hay anotación de nuevos transcritos.
- Si no esta en el archivo de anotación (tx2gene / kallisto) no lo veremos.
- Es necesario un archivo de anotación con buena calidad

Ensamblaje *de novo*



- Especie con genoma de mala calidad.
- Organismo no modelo.
- No contamos con un archivo de anotación de buena calidad.
- Emplear reads tipo PE.

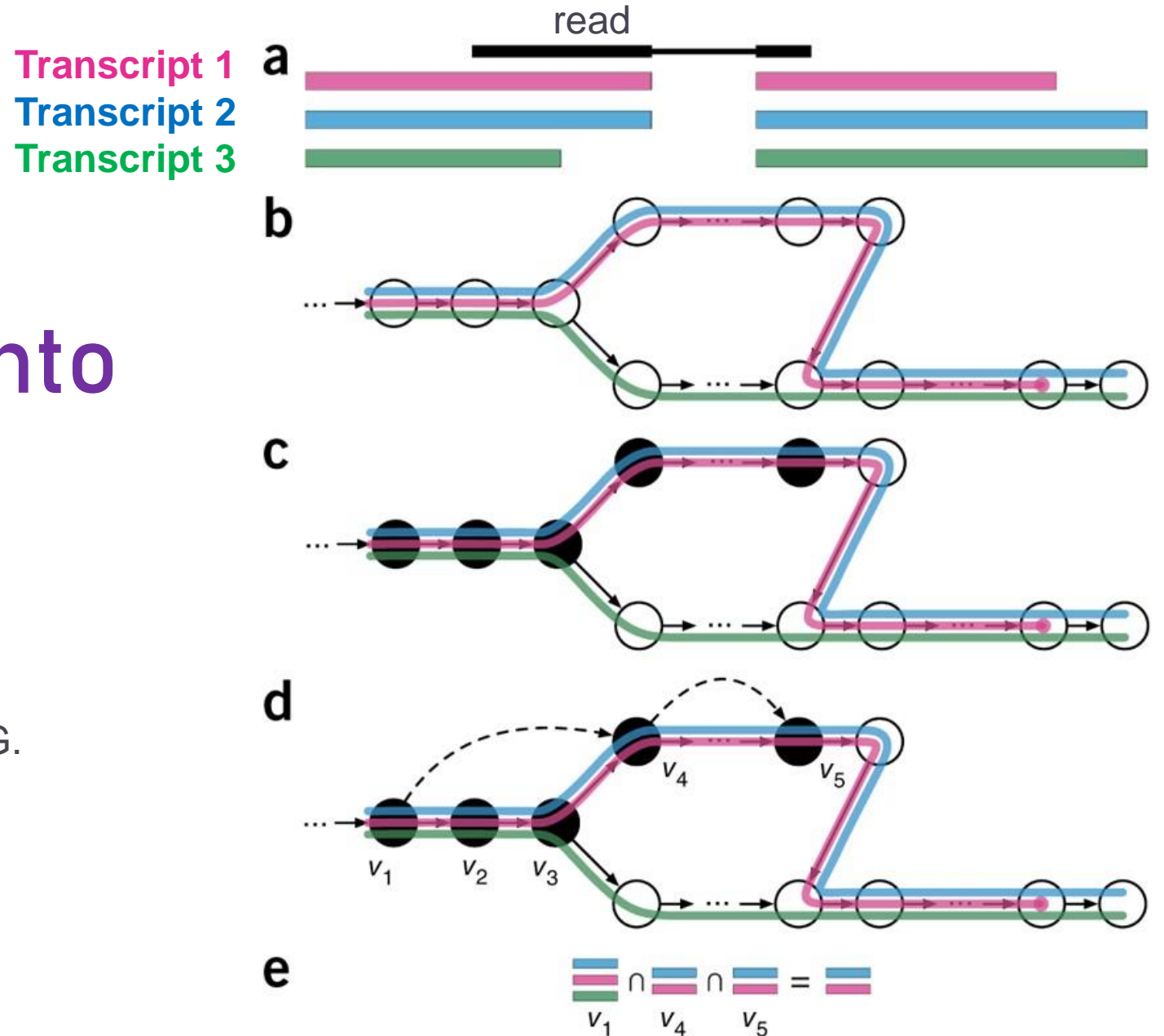
Práctica: Ensamblaje de transcriptoma guiado

Kallisto

Pseudoalineamiento de kallisto

- Brujin Graph (T-DBG)
- Los Nodos (v_1, v_2, v_3) son k -mers
- Omite pasos redundantes en el T-DBG.

Bray, *et al.* 2016. Nature



Práctica 2

- Input
 - fastq.gz

- Github

https://github.com/EveliaCoss/RNAseq_classFEB2023/tree/main/RNA_seq#practica2