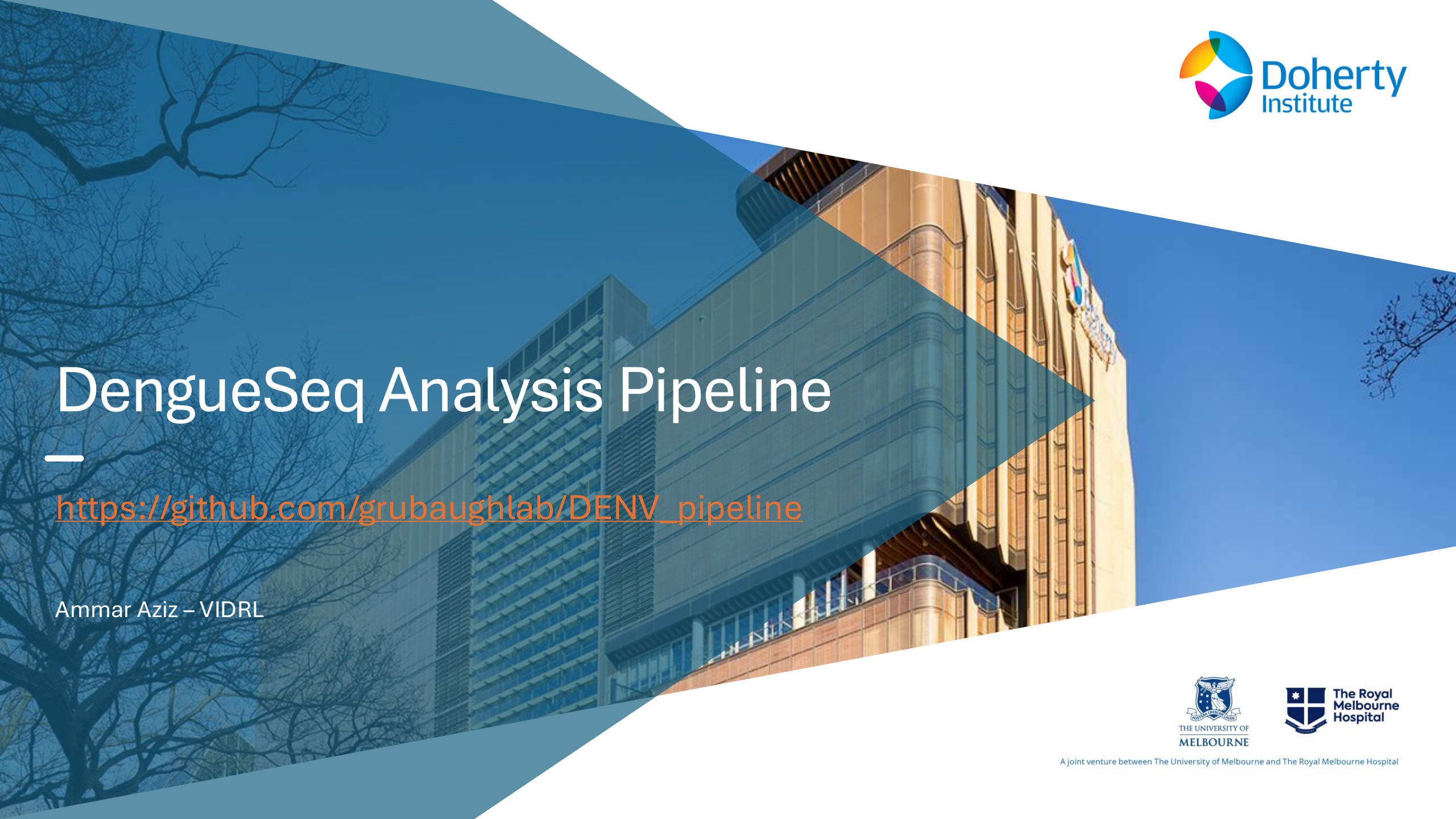


DengueSeq Analysis Pipeline

https://github.com/grubaughlab/DENV_pipeline

Ammar Aziz – VIDRL

Doherty Institute

THE UNIVERSITY OF MELBOURNE

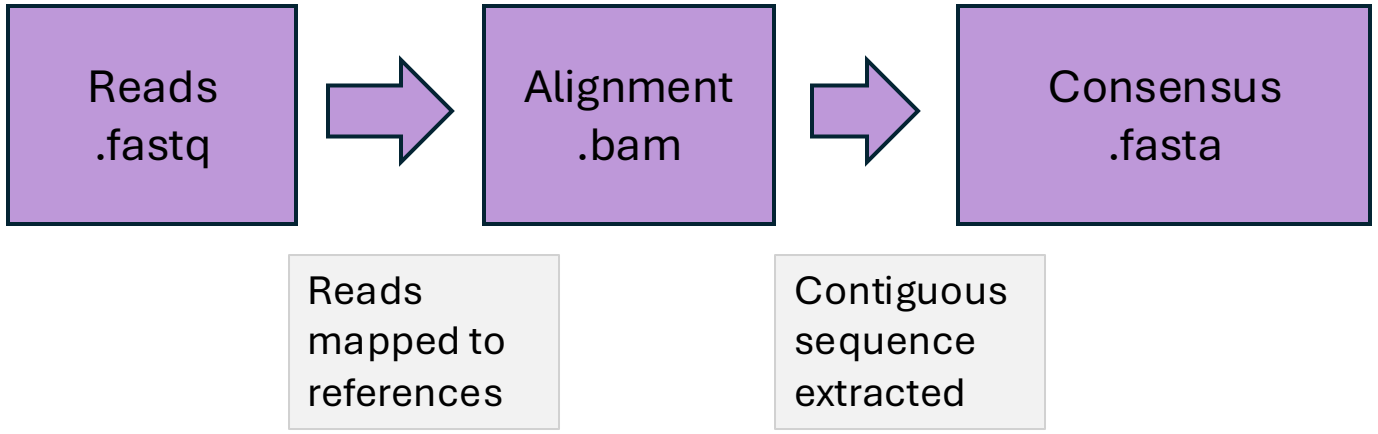
The Royal Melbourne Hospital

A joint venture between The University of Melbourne and The Royal Melbourne Hospital

Objectives

- Demonstrate how to download and install the DENV_Analysis pipeline
- Understand how the pipeline operates
- Understand how to run the pipeline
- Interpret pipeline outputs

How does DengueSeq analysis pipeline work?

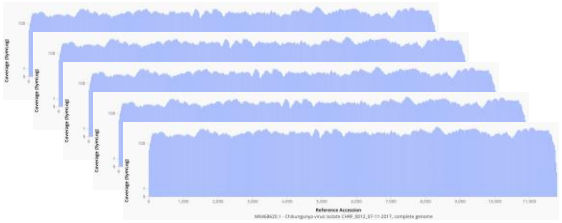


How does DengueSeq analysis pipeline work?

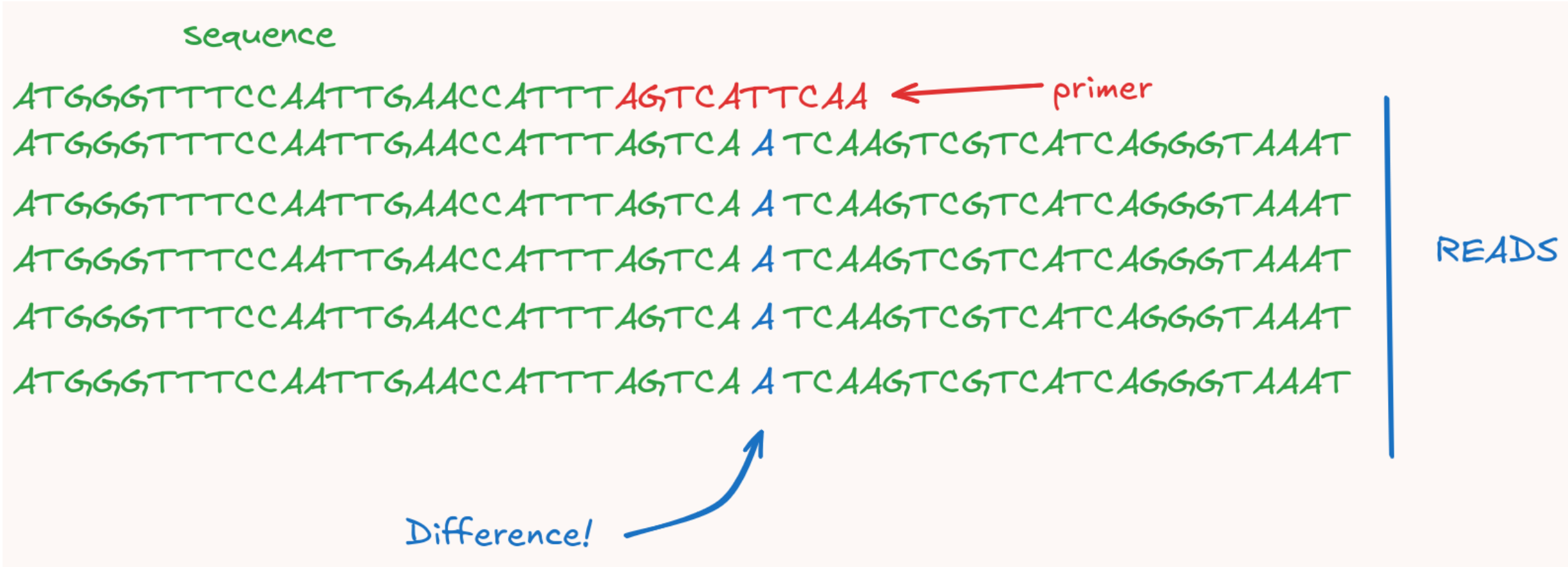
Align to reference using BWA-MEM → Convert SAM to BAM using samtools → Sort bam file → Trim primers using iVar → Create consensus using iVar by applying SNPs to reference → Calculate coverage using samtools

DENV1
DENV2
DENV2-Sylvatic
DENV3
DENV4
DENV4-Sylvatic

Finally – select consensus genome with highest coverage



Why trim primers?

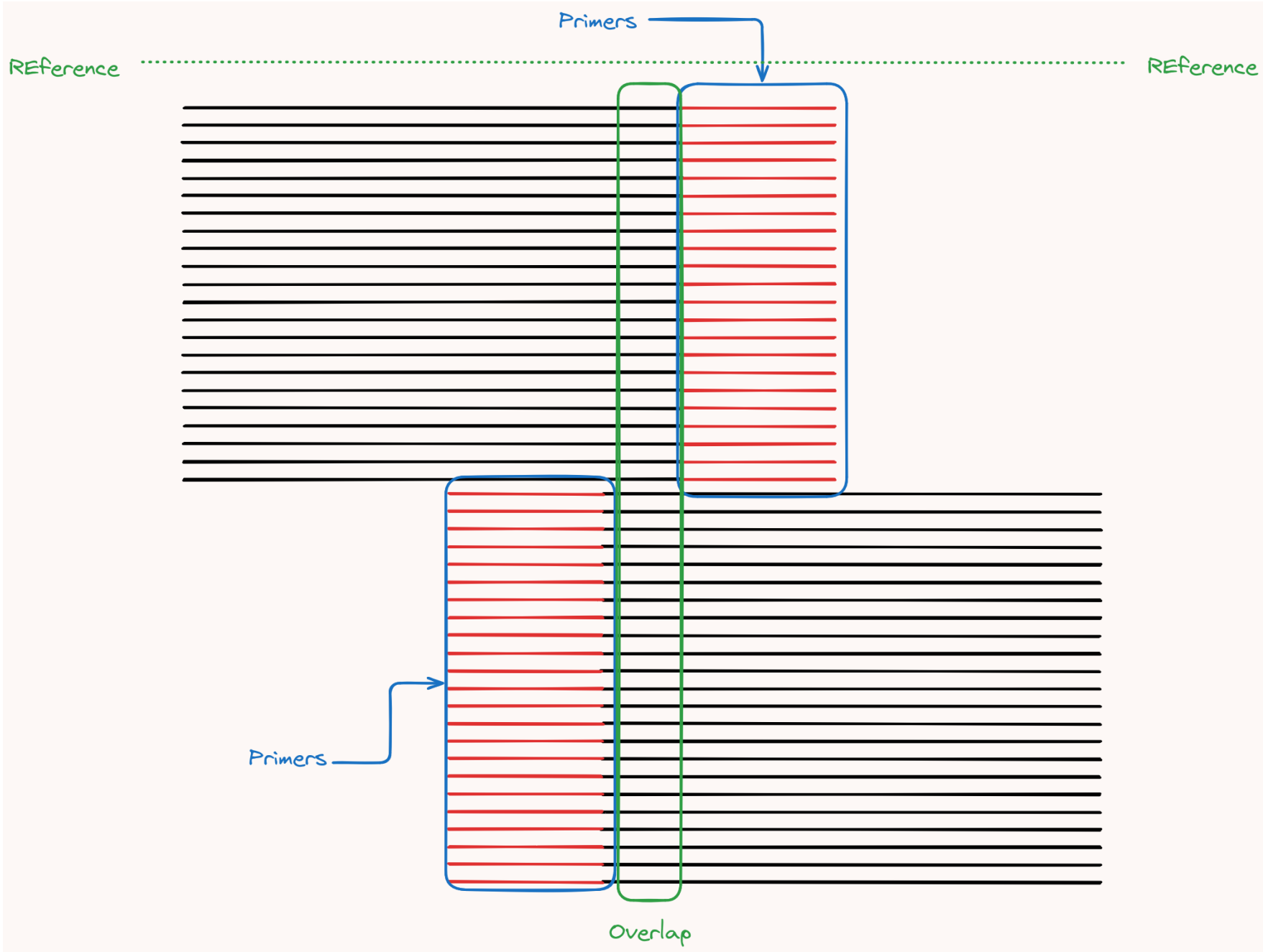


Primer sequence can **not** be incorporated into consensus sequence

Why trim primers?

No information loss because of amplicon design.

Primers overlap



How is a consensus created?

Create consensus using iVar by applying SNPs to reference

- Two methods – essentially the same:
  1. Use a dedicated variant caller to identify mutations (SNPs, INDELs), then apply mutations to reference
  2. Use the samtools mpileup format – take the majority call
     - iVar uses the mpileup format

Sequence	Position	Reference Base	Read Count	Read Results	Quality
seq1	272	T	24	,. $.....,,.,.,.,.,.,.,.^+.	<<<+;<<<<<<<<<<=<;<;7<&
seq1	273	T	23	,.....,.,.,.,.,.,.,..A	<<<;<<<<<<<<<3<=<<<;<<+
seq1	274	T	23	,.$.....,.,.,.,.,.,.,...	7<7;<;<<<<<<<<<=<;<;<<6
seq1	275	A	23	,$.....,.,.,.,.,.,.,.,...^l.	<+;9*<<<<<<<<<=<<:;<<<<
seq1	276	G	22	...T,,.,.,.,.,.,.,.,.,...	33;+<<7=7<<7<&<<1;<<6<
seq1	277	T	22	....,,.,.,.C.,,,,,,,.G.	+7<;<<<<<<<&<=<<:;<<&<
seq1	278	G	23	....,,.,.,.,.,.,.,.,....^k.	%38*<<;<7<<7<=<<<;<<<<<
seq1	279	C	23	A..T,,.,.,.,.,.,.,.,.,...	75&<<<<<<<<<=<<<9<<:<<<

https://en.wikipedia.org/wiki/Pileup_format

Next steps

- Analysis pipeline is only available through command line interface.
- We must learn CLI + Conda to install and run the pipeline

Questions?