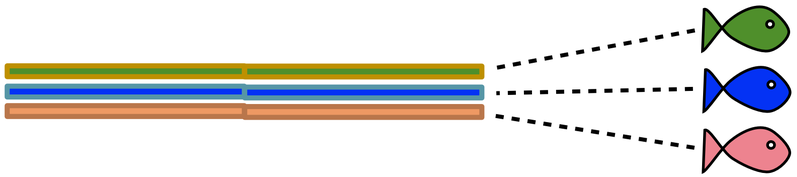
Next-generation DNA sequencing



Calling genotypes for diploid organisms

Regardless of sequencing platform, going from data to genotypes is similar:

1) Determine reference genome to map to, or create a denovo assembly for your data

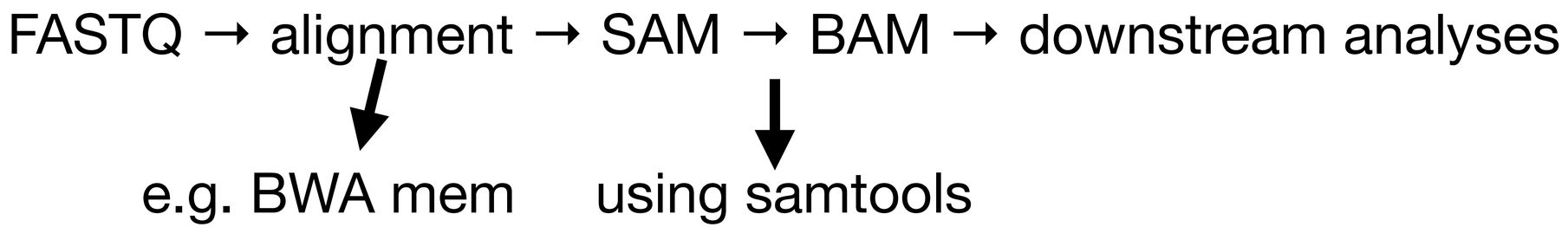
2) Map reads

3) Call genotypes

4) *Filter, analyze and check, filter again, check again, analyze, filter some more, check some more…*

mapping

Common algorithms used for mapping include BWA and Bowtie2



SAM = Sequence Alignment/Map (human readable, text)

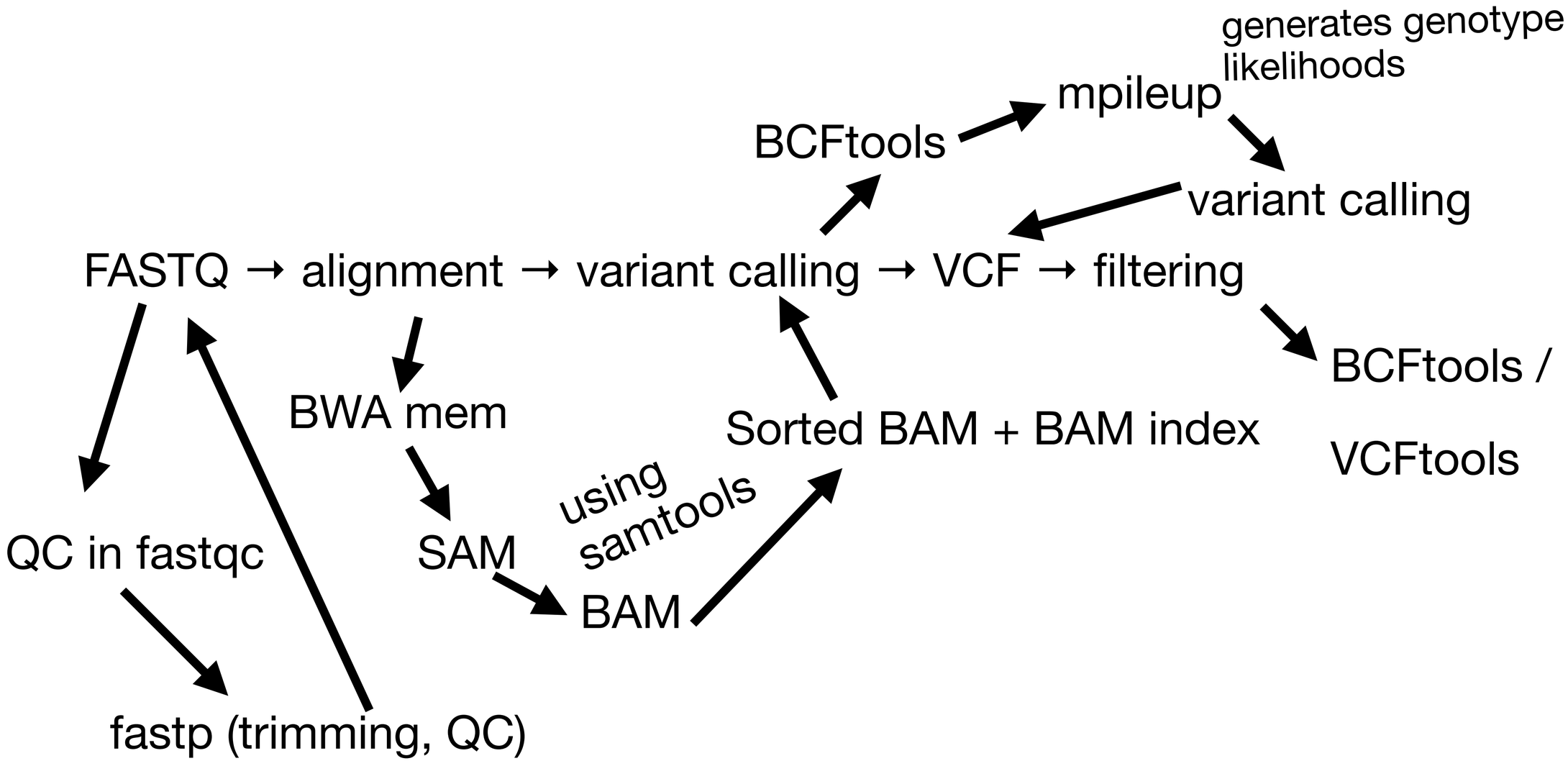
BAM = Binary Alignment/Map (binary, compressed version of SAM)

mapping to variant calling

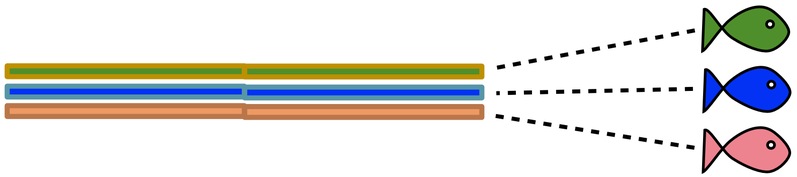
1. Align reads (e.g., bwa mem)
2. Convert SAM → BAM
3. Sort BAM by coordinate
4. Index BAM (.bai)

map between genomic coordinates and position in the BAM file needed by variant callers

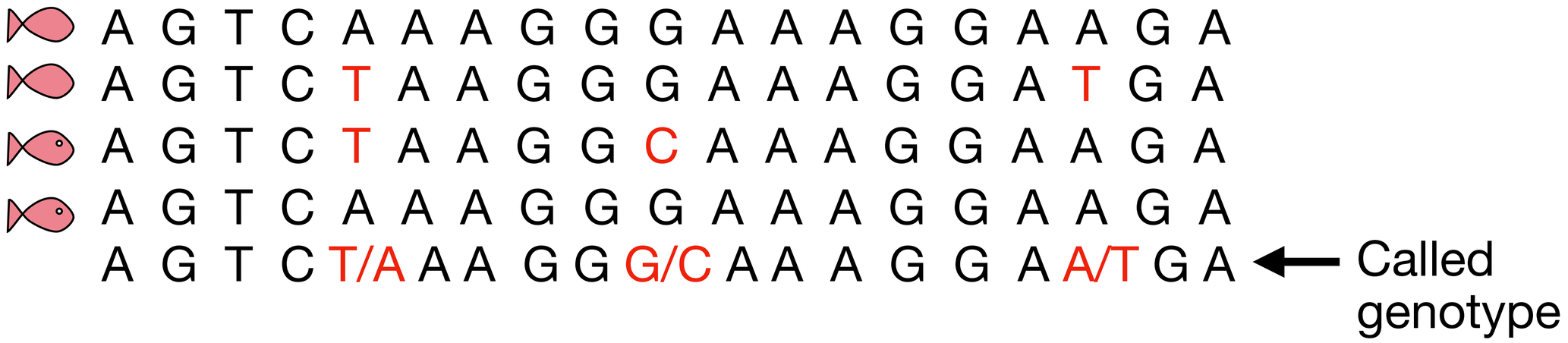
reads to variants…one approach



Next-generation DNA sequencing



Calling genotypes for diploid organisms



Next-generation DNA sequencing

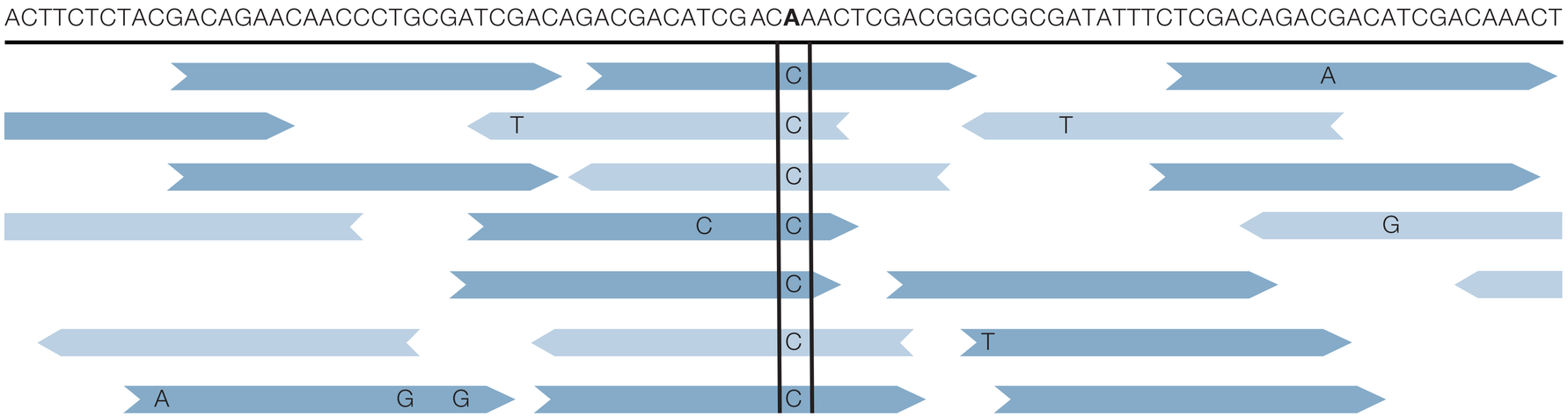


Figure 2.4