sequence verification

Molecular cloning to Sanger sequencing

Gulzhanay Kamshybayeva,
PhD
Senior Lecturer

The question sequencing answers

DNA sequence = the order of A, C, G and T

A read covers a particular region

Alignment compares the read with an expected sequence

Verification requires reliable coverage of important features



Capillary electrophoresis instrument

Sanger chain termination

A primer starts DNA synthesis

dNTPs support extension

A ddNTP lacks the 3′-OH needed for further extension

Fragments terminate at different positions

PCR and Sanger reactions

After the sequencing reaction

A capillary array separates labelled extension fragments by length.

Fluorescence signals become a chromatogram.

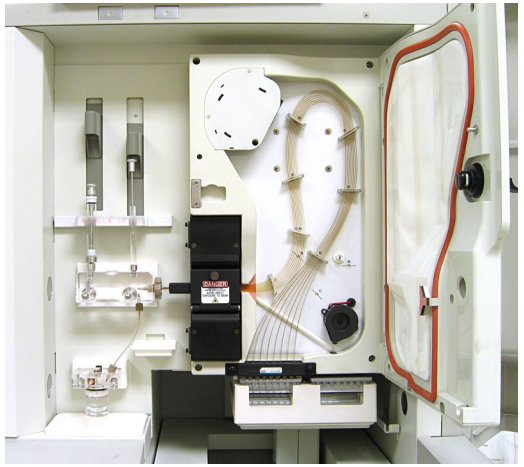
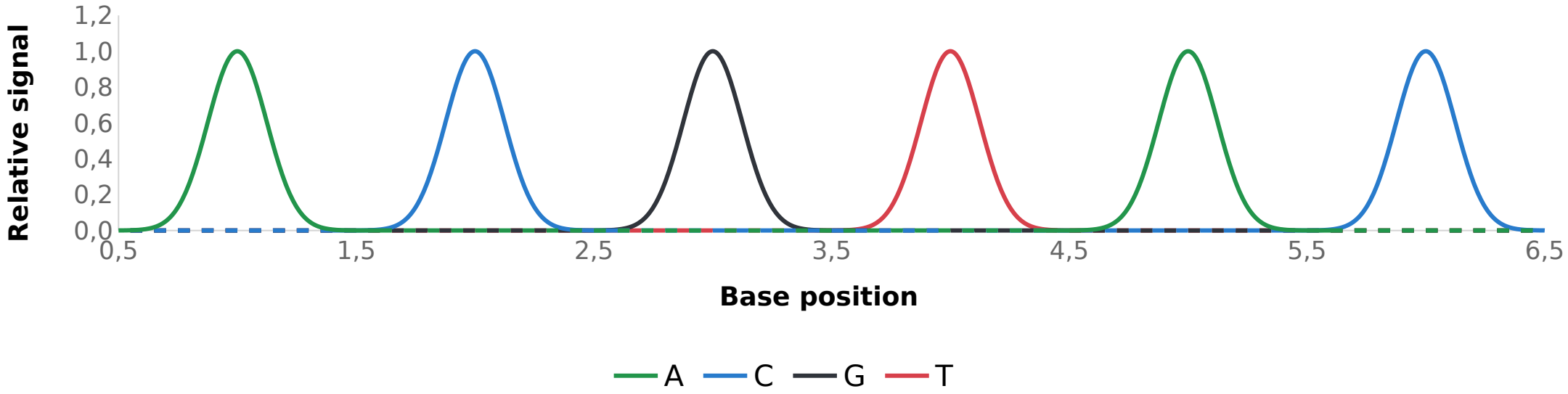


Photo: Michael Pereckas, CC BY-SA 2.0

Feature	Conventional PCR	Standard Sanger reaction
Typical primer setup	Two primers	One primer per reaction
Main output	Amplified DNA interval	Terminated fragments and a read
Main question	Can this target amplify?	What bases occur in this region?

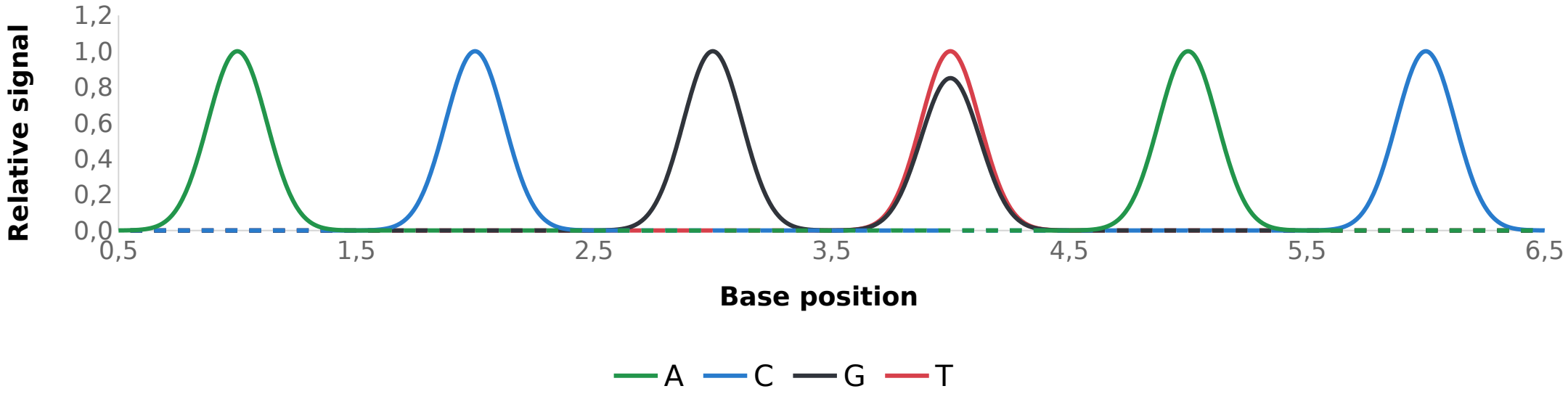
A clean chromatogram



Simplified simulated trace, not experimental data

Read the dominant peaks from left to right
Base order describes the newly synthesized strand, 5′ to 3′
Peak colors follow the displayed legend

An ambiguous chromatogram



Simulated mixed signal at position 4

Two prominent signals do not justify one confident base call
Inspect neighboring peaks, template purity and primer specificity
Repeat or obtain independent sequence evidence

Reverse complements and alignment

Forward-oriented reference: 5′-ACGTTA-3′

Reverse-direction read: 5′-TAACGT-3′

Reverse-complement the reverse read before comparison

Compare matching coordinates after quality review

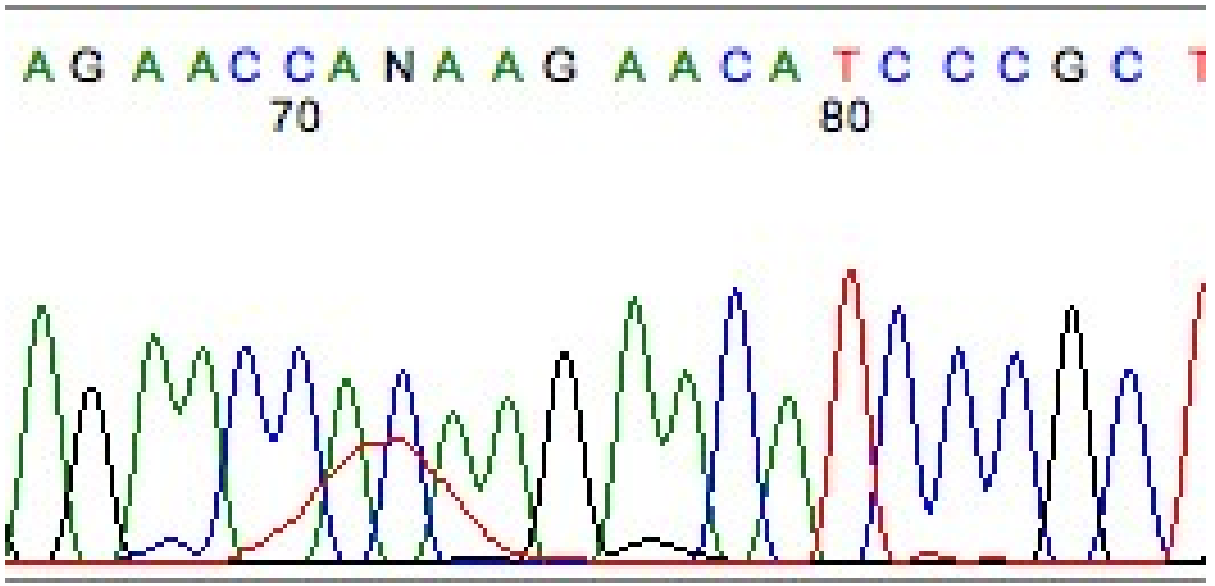
Sequence disagreement: what follows?

Expected: ACGTTA

Observed: ACGCTA

A mismatch is a finding to evaluate

A sequence change does not establish a functional effect



External trace: an uncertain base call

Separate example from the sequences on the left

Trace image: Addgene.

A verification plan for the 900 bp insert

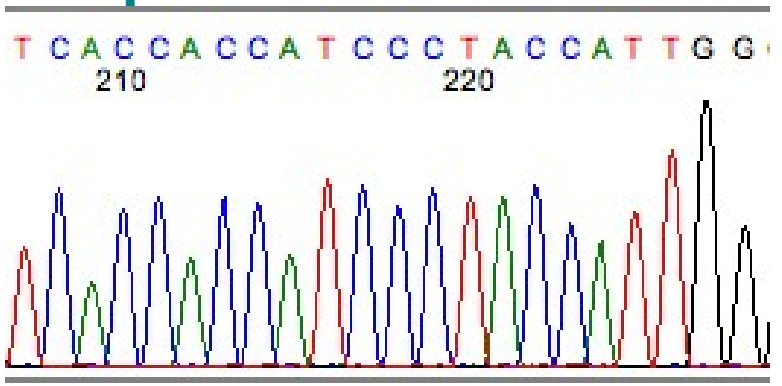
Screen candidate colonies

Check an informative restriction pattern if needed

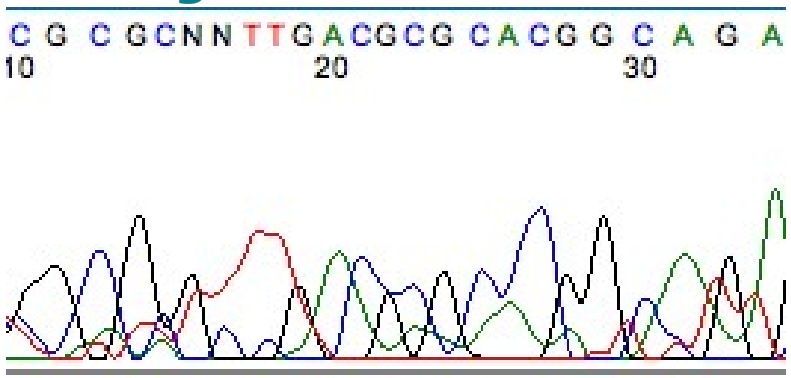
Sequence both junctions and the complete critical insert region

Add internal primers if usable reads leave a gap

Distinct peaks



High background



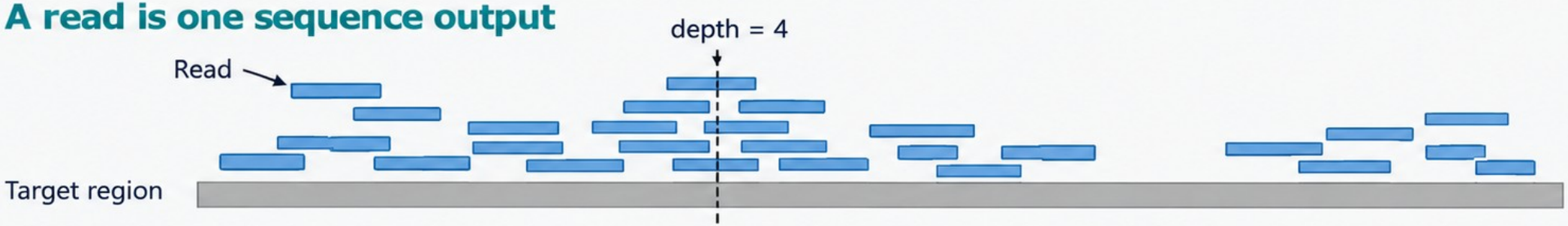
Trace images: Addgene. External examples.

Sanger and NGS: choosing the scale

Question	Reasonable starting choice
Verify a few defined insert regions	Sanger sequencing
Examine many loci or whole genomes	An appropriate NGS workflow
Resolve a specific technical limitation	Choose a method matched to that limitation

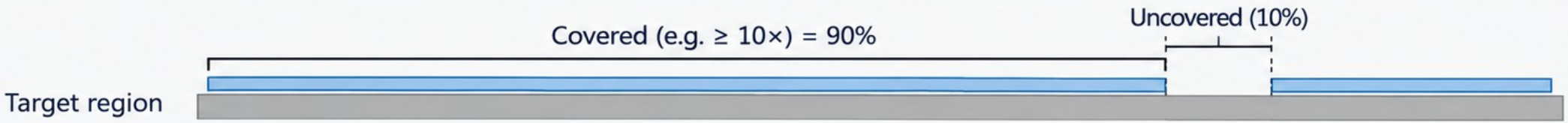
Reads and coverage: first interpretation

A read is one sequence output

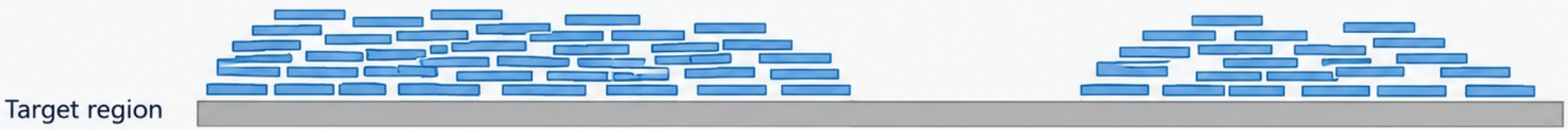


Depth counts reads covering a position

Breadth describes how much of the target meets a coverage criterion



High average depth can coexist with uncovered regions



Integrated case: the defensible conclusion

A reporter clone has the expected PCR and digest sizes

Both junctions match the design

Only 650 of 900 insert bases have reliable sequence

An expression assay has not been performed