PCR Primers Design

Azita Zadeh-Vakili, PhD
Medical Biotechnologist

Modern molecular Biology

- Molecular biology methods has been revolutionized by these tools and thechniques:
  - Restriction endonucleases
  - Gene cloning
  - Oligonucleotide synthesis

Polymerase Chain Reaction

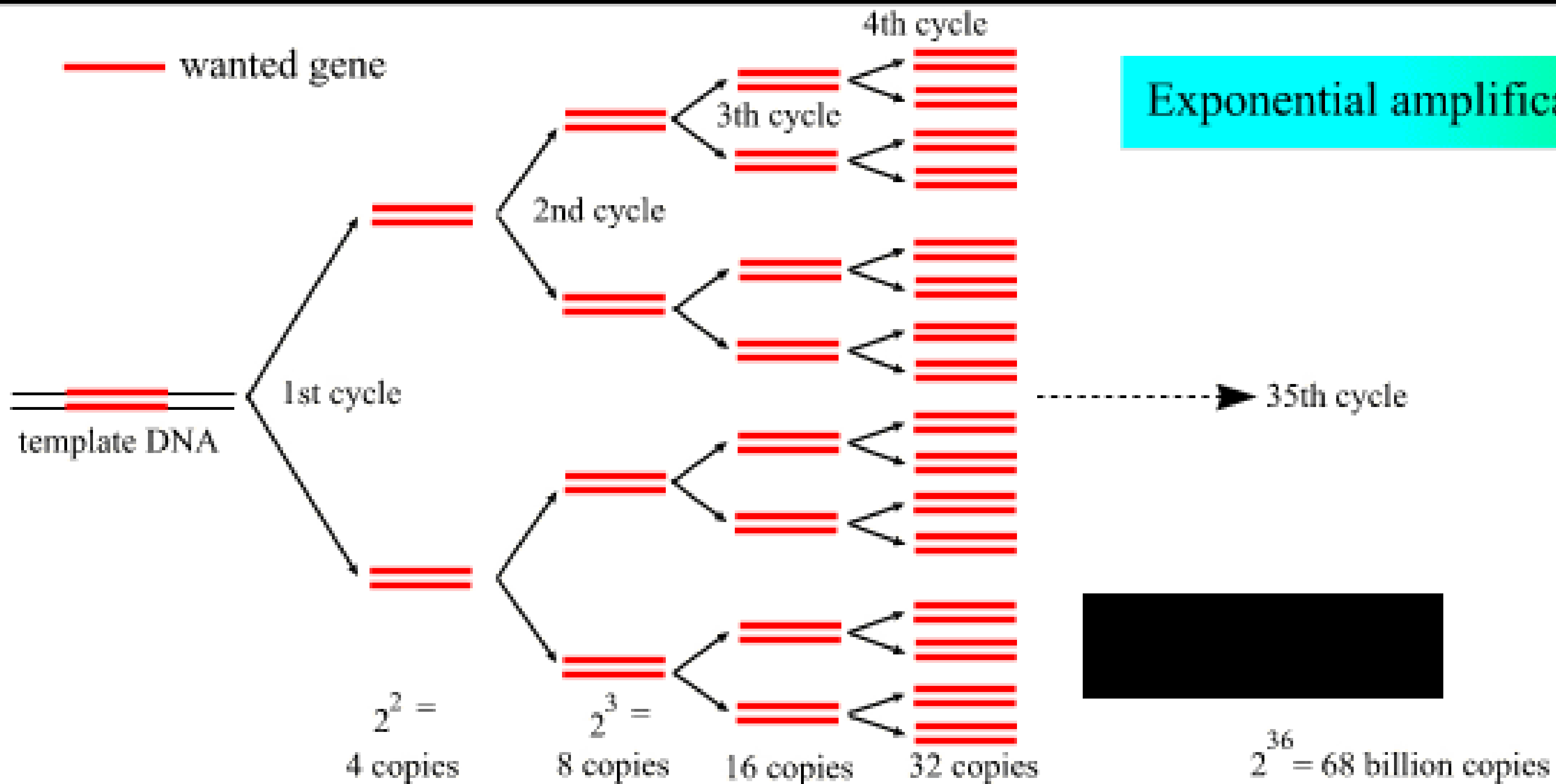
- Discovered by Kary *Mullis* in 1985.

Old in theory and highly versatile in practice

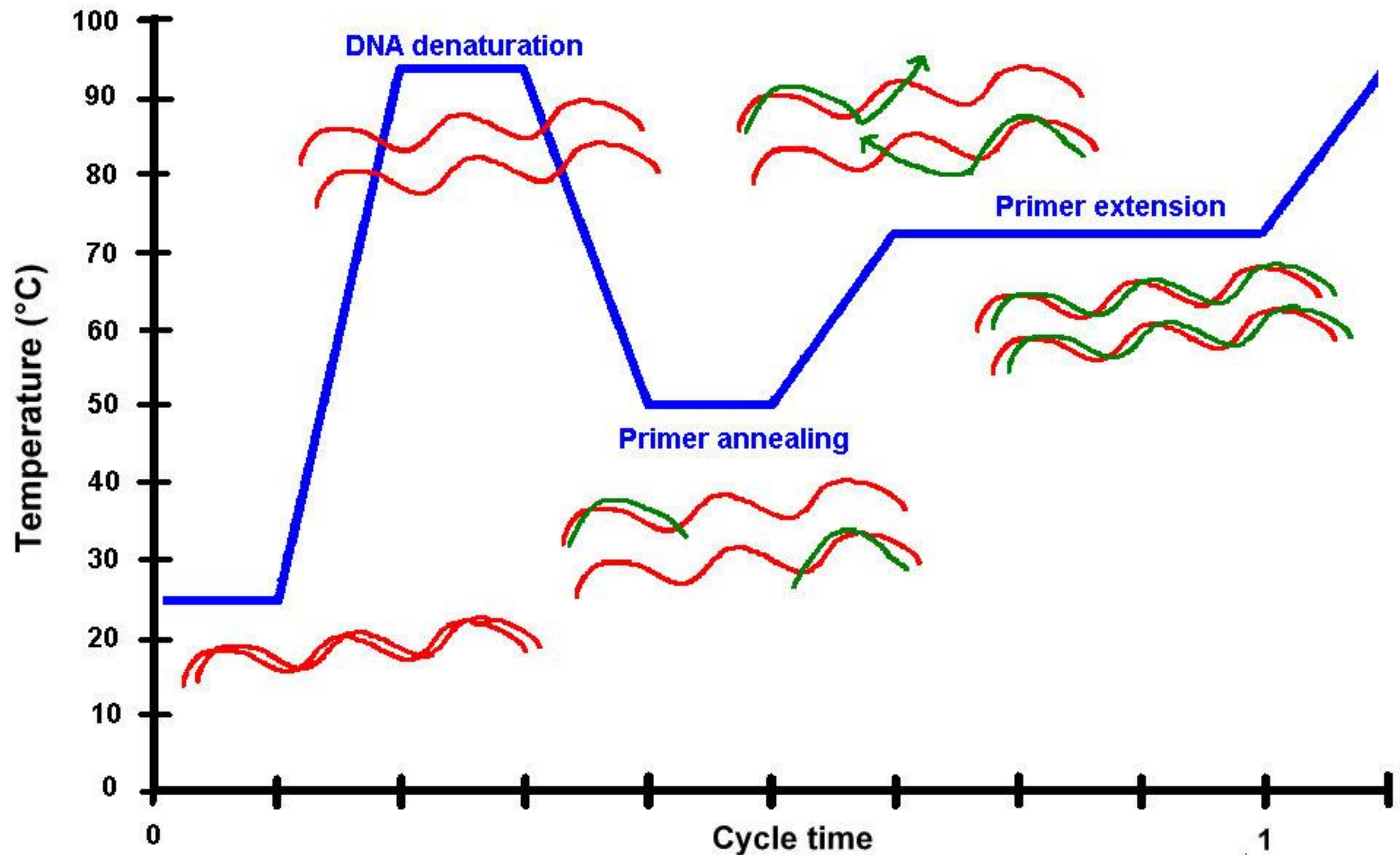
- "Amplify" large quantities of DNA (µg quantities) from small quantities (fg quantities) [billion fold amplification]

- The objective is to amplify a specific DNA segment without any nonspecific byproducts

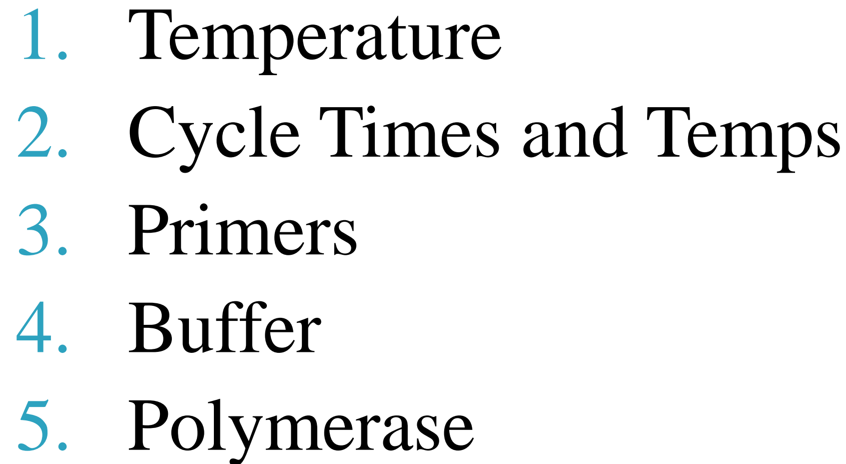
Exponential Amplification



Polymerase Chain Reaction



PCR Variables



Temperature

- Denaturation
  - Trade off between denaturing DNA and DNA Polymerase half-life
    - Taq half-life 40min at 95 °, 10min at 97.5°
- Annealing
  - Trade off between efficient annealling and specificity

(2-5 ° below Tm)
- Extension
  - Optimum Temperature for polymerase

Cycle Times and Temps

Typical PCR Run

Step Time/Temp

3-5 min at 94°

x20-30

- 15-120 Sec at 94°
- 15-60 Sec at 50-70°
- 1-2 min at 72°

5 min 72°

5 min 4°

End

The Melting Temperature

- **T_m (°C)= 4(G+C) + 2(A+T)**

This formula (the Wallace rule), is valid for oligos <15 bases and assumes that the reaction is carried out in the presence of 50mM monovalent cations

The temperature at which one-half of a particular DNA duplex will dissociate and become single strand DNA

- Annealing Temp: Tm -5

- T_A = 0.3(Tm primer)+ 0.7(Tm Product) -25

Basic Tm Calculations

example primer	GC+AT=length	Wallace rule	Santa Lucia
50/50 mixed: AGAGAGAGAGAGAGAGAGAG	10+10=20	60	47.7
50/50 separated: AAAAAAAAAAGGGGGGGGGG	10+10=20	60	52.7

Nearest-Neighbor method (SantaLucia equation)

- The most sophisticated T_m calculations take into account the exact sequence and **base stacking** parameters, not just the base composition

$$T_m(^oC) = (\Delta H / (\Delta S + R \ln[C])) - 273.15$$

Where:

- ΔH (kcal/mole) is the enthalpy of base stacking
- ΔS (kcal/mole) is the entropy of base stacking
- R is the universal gas constant (1.987Cal/°C*Mol)
- [C] is oligonucleotide strand concentration

Primer design

Critical to both the efficiency and the accuracy of the PCR

Consideration	Comment
Primer Melting Temperature (T_m)	55°-72°C
Primer Annealing Temperature (T_a)	~5°C < the lowest T_m of the of primers
Tm difference between forward and reverse primers	≤ 5°C
Primer Annealing Temperature (T_a)	~5°C < the lowest T_m of the of primers
Max 3′ Stability	ΔG value for five bases from 3′ end
Percentage GC content	40-60%
No Secondary Structures	Identify primer pairs which do not assume secondary structure
No self-complementarity	< 4 contiguous bases
No complementarity to other primer(s)	< 4 contiguous bases
No long runs with the same base	< 4 contiguous bases
Distance between two primers on target sequence	< 2000 bases apart
Plateau Effect	accumulation of product ≤0.3 to 1 pmol

Primers

- Paired flanking primers
- Length (17-28bp)
- GC content 50-60%
- GC Clamp

(The presence of 1-3 G or C bases within the last five bases from the 3' end of primers)

- Tm’s between 55-80
- Avoid simple sequences – e.g. strings of G’s
- Avoid primer self complementary
  - e.g. hairpins, homodimers, heterodimers

PCR Specificity

- The specificity of PCR is determined by the specificity of the PCR primers

Avoid the similarity

- Prevention of primers nonspecific binding (binding to more than one locus, e.g. paralog or common domain),effectively increases PCR specificity.

Primer Specificity

- **Similarity Searching using**

Basic Local Alignment Search Tools

The family of **BLAST** programs provides a powerful way to compare a query sequence against a sequence database

BLAST

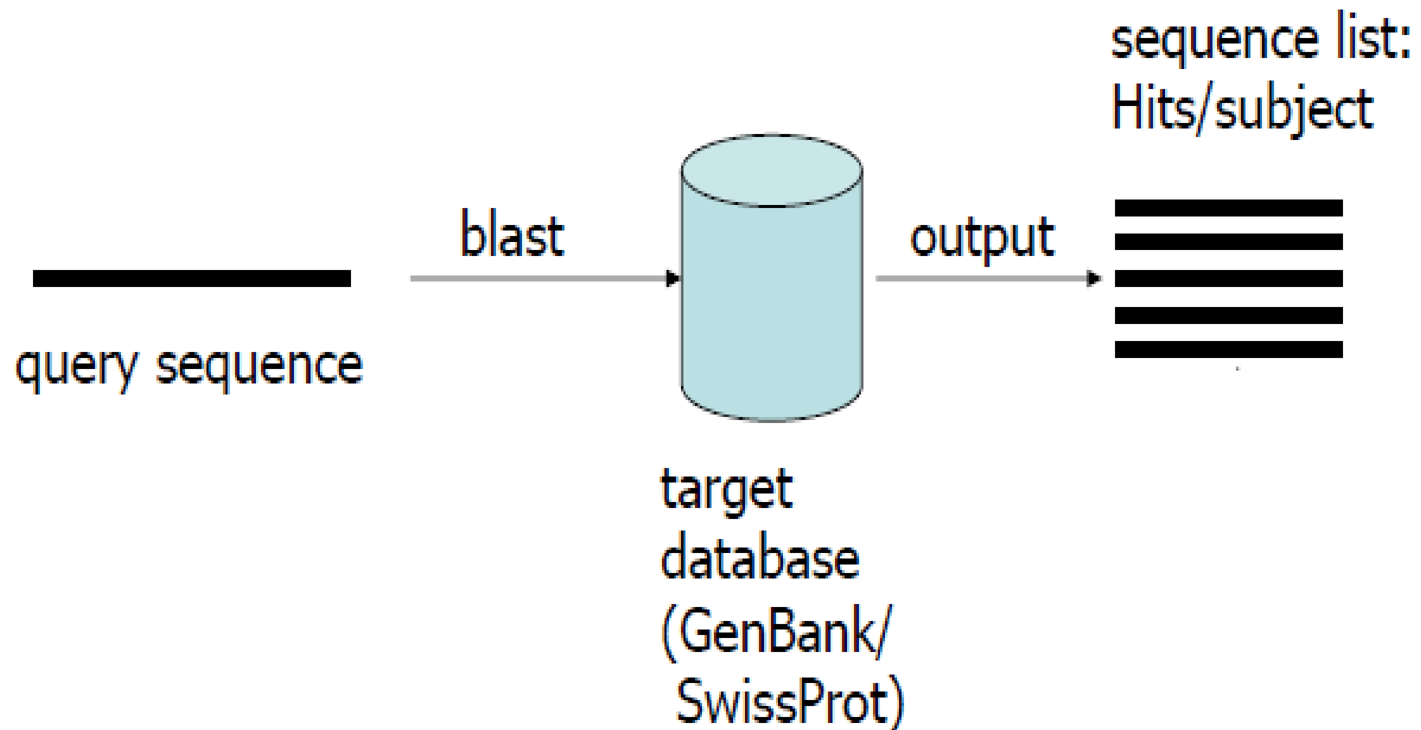
- BLAST improved overall speed of searches

- BLAST maintains good sensitivity

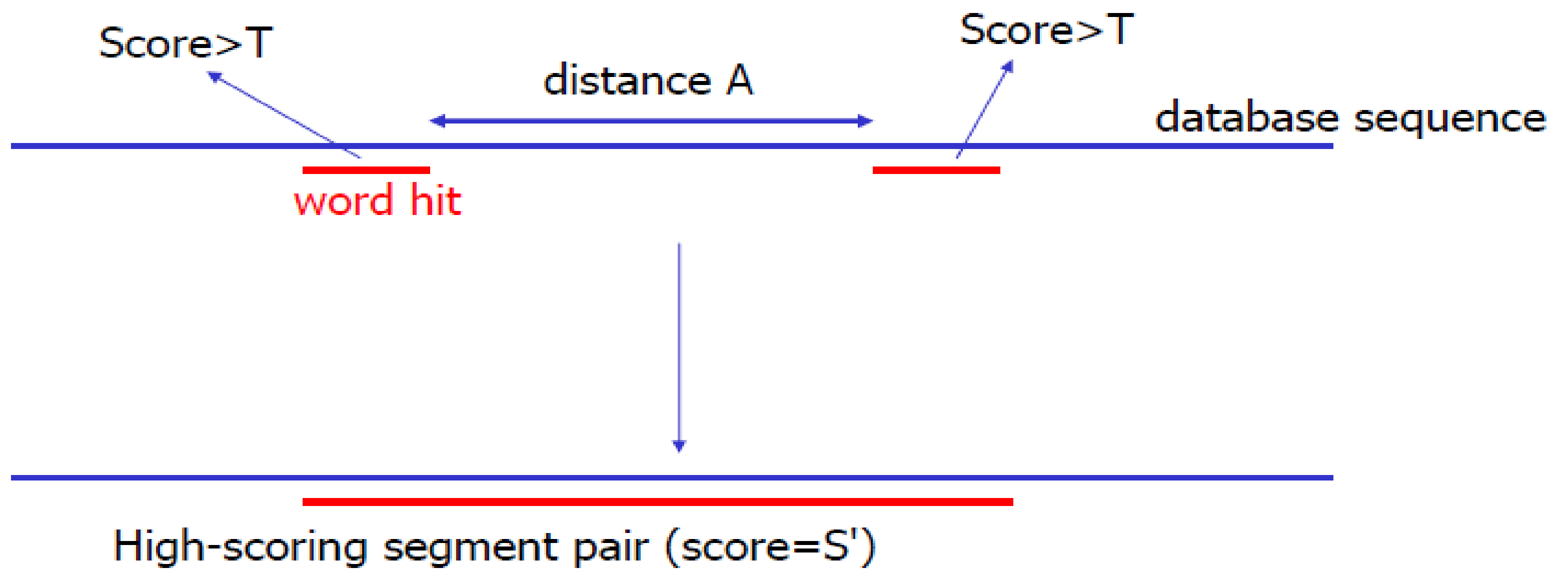
BLAST Algorithm

- Scoring of matches done using scoring matrices
- Sequences are split into words (default n=3)
  - Speed, computational efficiency
- BLAST algorithm extends the initial "seed" hit into an HSP
  - HSP = high scoring segment pair = Local optimal alignment

BLAST: discover sequence omology

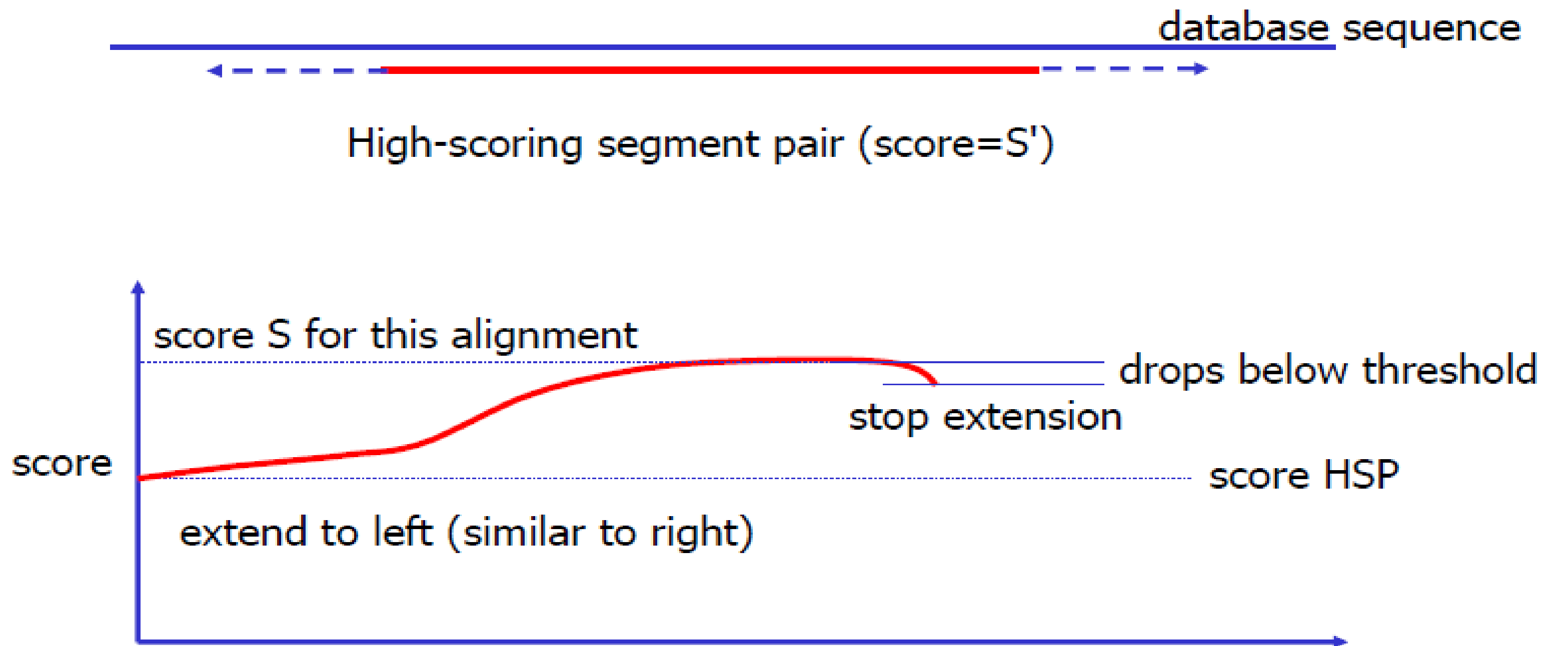


Join words on same diagonal (ungapped)



Join words on same diagonal

Extend HSP until score drops small amount below highest score of shorter alignment



If S>threshold (based on random sequences) then keep HSP

Where does the score (S) come from?

- The quality of each pair-wise alignment is represented as a score and the scores are ranked.
- Scoring matrices are used to calculate the score of the alignment base by base (DNA) or amino acid by amino acid (protein).
- The alignment score will be the sum of the scores for each position.

What's a scoring matrix?

- Substitution matrices are used for amino acid alignments.
  - each possible residue substitution is given a score
- A simpler unitary matrix is used for DNA pairs (+1 for match, -2 mismatch)

	A	C	D	E	F	G	H
A	4	0	-2	-1	-2	0	-2
C	0	9	-3	-4	-2	-3	-3
D	-2	-3	6	2	-3	-1	-1
E	-1	-4	2	5	-3	-2	0
F	-2	-2	-3	-3	6	-3	-
G	0	-3	-1	-2	-3	6	
H	-2	-3	-1	0			

BLOSUM 62

BLAST programs

Program	Description
blastp	Compares an amino acid query sequence against a protein sequence database.
blastn	Compares a nucleotide query sequence against a nucleotide sequence database.
blastx	Compares a nucleotide query sequence translated in all reading frames against a protein sequence database. You could use this option to find potential translation products of an unknown nucleotide sequence.
tblastn	Compares a protein query sequence against a nucleotide sequence database dynamically translated in all reading frames.
tblastx	Compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.