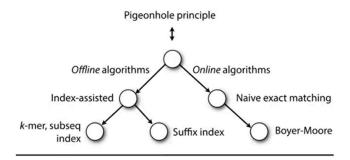
# **Read Alignment Methods**



Dynamic programming and edit distance

# **Number of Character Comparisons Example**

How many character comparisons in this example?

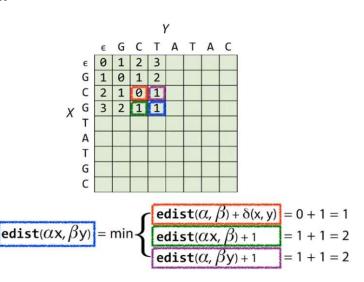
P: word

T: There would have been a time for such a word
word word word word word word word
word word word word word word
word word word word word word
word word word word word word
word word word word word word
word word word word word word
word word word word word word

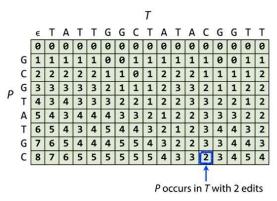
40 mismatches + 6 matches = 46 character comparisons

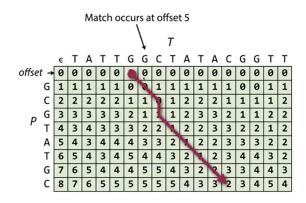
Much closer to minimum (41) than maximum (164)

#### **Edit Distance Matrix**



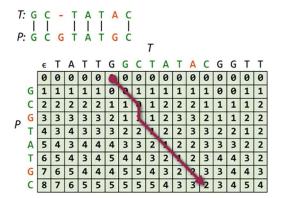
# **Alignment Matrix**



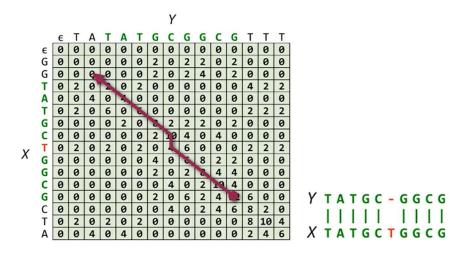


Shortest edit distance btwn the full strings

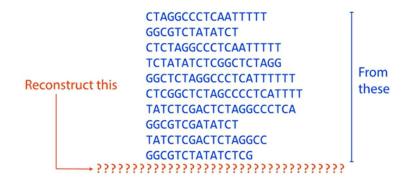
Path to the shortest distance



# **Local Alignment Matrix**



# de novo Shotgun Assembly



# Coverage

CTAGGCCCTCAATTTTT CTAGGCCCTCAATTTTT CTCTAGGCCCTCAATTTTT **CTCTAGGCCCTCAATTTTT GGCTCTAGGCCCTCATTTTTT GGCTCTAGGCCCTCATTTTTT** CTCGGCTCTAGCCCCTCATTTT CTCGGCTCTAGCCCCTCATTTT TATCTCGACTCTAGGCCCTCA **TATCTCGACTCTAGGCCCTCA** TATCTCGACTCTAGGCC TATCTCGACTCTAGGCC 177 bases **TCTATATCTCGGCTCTAGG TCTATATCTCGGCTCTAGG GGCGTCTATATCTCG GGCGTCTATATCTCG GGCGTCGATATCT GGCGTCGATATCT** 35 bases **GGCGTCTATATCT GGCGTCTATATCT** GGCGTCTATATCTCGGCTCTAGGCCCTCATTTTTT GGCGTCTATATCTCGGCTCTAGGCCCTCATTTTTT Average coverage = 177 / 35 ≈ 7-fold Coverage = 5

# **Overlap Graphs**

Nodes: all 6-mers from GTACGTACGAT

Edges: overlaps of length ≥4

