#### Sequence comparison

- Comparing two sequences
  Global comparison
  - Local comparison
  - Semi-local alignment
    - Prefix-suffix or containment
    - Application: fragment assembly

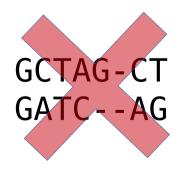
#### Global comparison

- <sup>1</sup> Sequences of enarly equal length
- Goal: to learnWhere they are alike
  - <sup>1</sup> Where they differ

### Alignments

- Add spaces to sequences
- Until sequences have the same length
- <sup>1</sup> Place one over the other
- Make sure no two spaces align
- Compute alignment score





### Scoring system

- <sup>a</sup> Match: usually a positive value
- Mismatch: usually a negative value
- Space: usually a negative value, worse than mismatch
- <sup>D</sup> Example:
  - I Match = +1
  - Mismatch = -1
  - <sup>1</sup> Space = -2

# Similarity

- Highest score among all alignments
  Notation:
- sim(s,t)

# Algorithm

- Construct a matrix
- Vertical: s
- Horizontal: t
- <sup>II</sup> Leave room for empty prefixes
- Meaning of a[i,j]: similarity of s[1,i], t[1,j]
- Filling matrix: spreadsheet
- Retrieving optimal alignments
- <sup>o</sup> O(mn) where m = |s| and n = |t|