

Genome assembly

Shortest Common Superstring

Substrings may contain errors

Some models proven to be NP-complete

Double-stranded

Original genome may be longer than the shortest common superstring

Substrings are exact

NP-complete

Single-stranded

Superstring must be minimum length

Given a set of short strings, find a long string that (roughly speaking) contains them.