

At what offset does TAC occur in GATTACA?

3

The outer loop of the naive exact matching algorithm iterates over

- ☒ Alignments
- ☐ Characters
- ☐ Patterns
- ☐ Reads

The character I (capital i) has ASCII value 73. Assuming Phred+33 encoding, what is the corresponding base quality?

40

✓ **Correct**

Say the photograph of the cluster shows that all the light coming from the templates in the cluster is the same color. Then the base quality should be:

☒ High

☐ Low

Another purpose of the terminator is to

- ☒ glow a particular color corresponding to the base
- ☐ prevent non-complimentary bases from being integrated
- ☐ keep the template strand from breaking into pieces

 **Correct**

One purpose of the terminator is to

- ☐ remove bases from the template
- ☒ keep the templates in sync
- ☐ allow multiple bases to be incorporate per cycle



**Correct**

A typical second-generation sequencing read is

- ☐ 10s of bases long
- ☒ 100s of bases long
- ☐ 1,000s of bases long
- ☐ millions of bases long

In the puzzle analogy, using the picture of the completed puzzle as guide is analogous to using \_\_\_\_\_ for finding where reads originate:

- ☒ a reference genome
- ☐ a DNA sequencer
- ☐ a strand of DNA
- ☐ a gene

✓ **Correct**



How many different suffixes does a string of length  $n$  have? Assume empty string doesn't count as a suffix.

- ☐ 1
- ☒  $n$
- ☐  $n + 1$
- ☐  $2n$

Which of the following is not a suffix of GATTACA?

☐ CA

☐ TTACA

☒ GATT

☐ A

Second-generation sequencing started to be used

- ☐ before the human genome project
- ☒ after the human genome project

Second-generation sequencing started to be used

- ☐ before the human genome project
- ☒ after the human genome project

✓ **Correct**

The Human Genome Project ended around 2003, and second-generation sequencing arrived around 2005-2007 or so

How similar are the sequences of two genomes from unrelated humans?

- ☒ 99.8-99.9%
- ☐ 100%
- ☐ 95-99%
- ☐ 75-90%



**Correct**

The enzyme that incorporates complementary bases, making a single-stranded template into a double-stranded molecule, is called:

- ☐ kinase
- ☐ ligase
- ☐ helicase
- ☒ polymerase

If we're studying sequencing data from a genome that has never been sequenced, what kind of computational method is most appropriate?

- ☒ de novo assembly
- ☐ compression
- ☐ polymerization
- ☐ read alignment

✓ **Correct**

How many character comparisons occur when matching  $P = \text{'AAA'}$  to  $T = \text{'AAATAA'}$ ?

- ☒ 9
- ☐ 6
- ☐ 3
- ☐ 12



What is the base at offset 5 of GATTACA? Remember: the leftmost base has offset 0.

☐ A

☒ C

☐ G

☐ T