

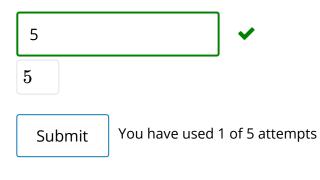
Course > Section 4: Genomic ... > Annotation of gene... > Assessment: Refere...

Assessment: Reference genomes

Q1: Reference genome discovery

1/1 point (graded)

How many Bioconductor packages provide reference genomic sequence for zebrafish (*Danio rerio*)? Exclude the packages with suffix ".masked", which we will discuss later.



✓ Correct (1/1 point)

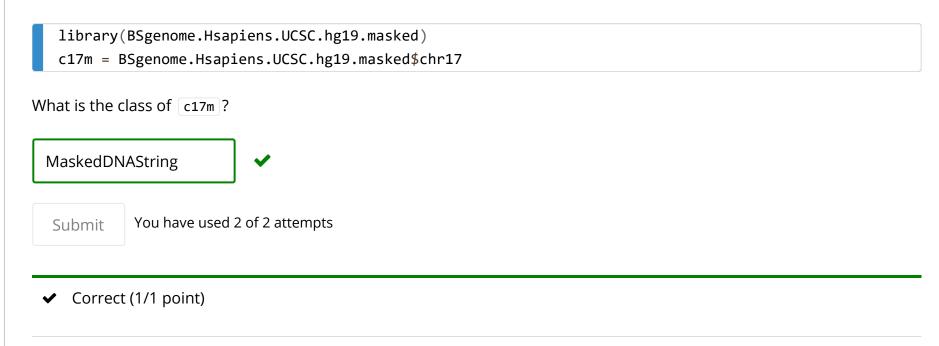
Q2: Masking structures for genome gaps and repetitions

1/1 point (graded)

We have noted that the reference genome builds for complex organisms are works in progress. Genomic sequence "mask" structures have been defined to isolate ambiguous, unmappable, and low-complexity segments of genomes so that sequence analysis research can be targeted to reflect current knowledge of sequence regions that are more likely to be functionally informative.

Obtain BSgenome.Hsapiens.UCSC.hg19.masked (it is only a 20MB transfer.)

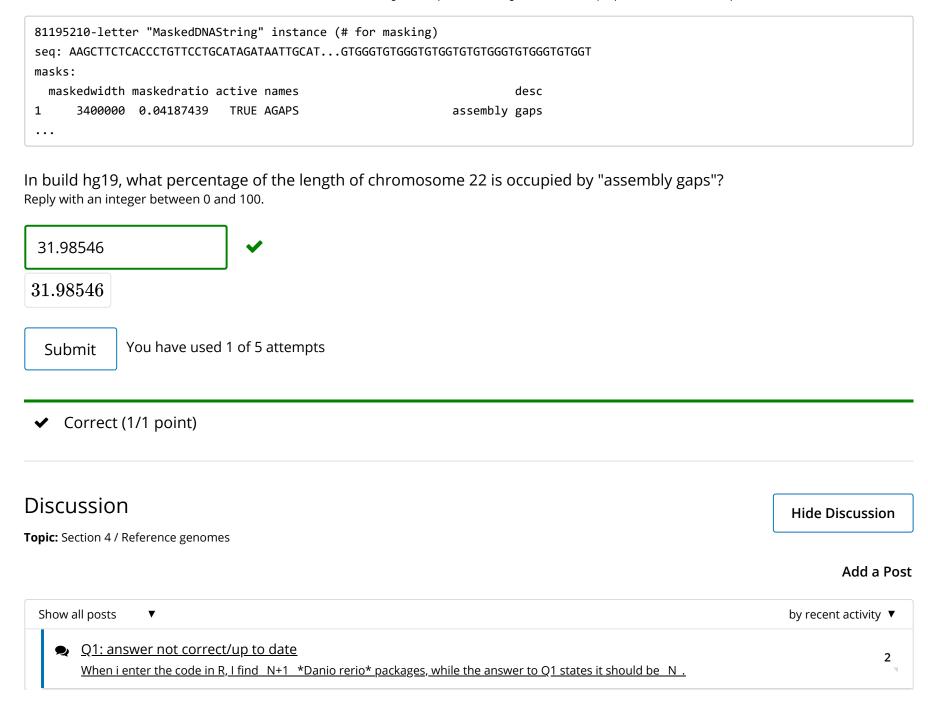
Run the commands:



Q3: Quantifying assembly gaps

1/1 point (graded)

When we print out the value of a *MaskedDNAString*, we get a report on types of mask present. Part of the report for chromosome 17 in hg19 is:



© All Rights Reserved