



[Course](#) > [Section 2: Biocond...](#) > [Biostrings](#) > Assessment: Gettin...

### Audit Access Expires Jun 2, 2020

You lose all access to this course, including your progress, on Jun 2, 2020.

Upgrade by May 19, 2020 to get unlimited access to the course as long as it exists on the site. [Upgrade now](#)

## Assessment: Getting sequences

In this assessment we will get the sequence for the regions created like this:

```
library(ERBS)
library(GenomicRanges)
data(HepG2)
data(GM12878)
res = findOverlaps(HepG2,GM12878)
erbs = HepG2[queryHits(res)]
erbs = granges(erbs)
```

We will use these sequences to compare the GC content of ESRRA binding sites to control sites.

## Q1: Genome build annotation

1/1 point (graded)

What genome build was used to create these regions?

☐ hg18

☒ hg19

☐ GRCh38

☐ hg17



Submit

You have used 1 of 5 attempts

✓ Correct (1/1 point)

## Q2: GC content of binding regions

1/1 point (graded)

Now load the human genome data:

```
library(BSgenome.Hsapiens.UCSC.hg19)
```

Use the `getSeq()` function to extract the sequence of each region in `erbs`. Then compute the GC-content (the number of C's + the number of G's divided by the length of sequence) of each.

What is the median GC-content?



You have used 1 of 5 attempts

---

✓ Correct (1/1 point)

---

### Q3: GC content of a shifted set of regions

1/1 point (graded)

Now create a control set of regions by shifting `erbs` by 10000.

What is the median GC-content of these control regions?



You have used 1 of 5 attempts

✓ Correct (1/1 point)

## Discussion

[Hide Discussion](#)

**Topic:** Section 2 / Getting sequences

[Add a Post](#)

Show all posts ▼

by recent activity ▼

There are no posts in this topic yet.

✕

[Learn About Verified Certificates](#)

© All Rights Reserved