



[Course](#) > [Section 3: Manage...](#) > [Organizing microar...](#) > Assessment: Expre...

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: ExpressionSet

If you have not done so already, install the following library which is available from our GitHub repository:

```
BiocManager::install("genomicsclass/GSE5859Subset")
```

Load the library, load the data it contains, and note that three tables are loaded:

```
library(GSE5859Subset)
data(GSE5859Subset)
dim(geneExpression)
dim(sampleInfo)
dim(geneAnnotation)
```

This data represents a subset of a larger [microarray experiment](#) with GEO accession number GSE5859. We used it extensively in course 525.3.

Note that these three tables are related, although nothing in our R session tells us they are. We are going to create a single R object to keep them organized together.

Question 1

1/1 point (graded)

Which Bioconductor class should we use to do this?

☐ *GRanges*

☒ *ExpressionSet*

☐ *geneExpressionClass*

☐ *AnnotatedDataFrame*



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 2

1/1 point (graded)

The `geneExpression` object contains the microarray measurements, `sampleInfo` provides sample information and `geneAnnotation` provides feature information. Read the help file for the *ExpressionSet* class and determine which of the following best describes the best way to use the data in this context:

- ☐ `geneExpression` will be the `phenoData`, `sampleInfo` will be the `assayData`, and `expressionAnnotation` will be the `featureData`
- ☐ `geneExpression` will be the `assayData`, `sampleInfo` will be the `featureData`, and `expressionAnnotation` will be the `phenoData`
- ☐ `geneExpression` will be the `assayData`, `sampleInfo` will be the `phenoData`, and `expressionAnnotation` will be the `annotation`
- ☒ `geneExpression` will be the `assayData`, `sampleInfo` will be the `phenoData`, and `expressionAnnotation` will be the `featureData`



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Question 3

1/1 point (graded)

Before creating this object we must make sure that the rows of `sampleInfo` match the columns of `geneExpression` and that the rows of `geneAnnotation` match the rows of `geneExpression`. This is an important step in which the data creators and the analysts must work closely together to get right. In this particular case, the tables have identifiers that can be used:

```
identical(colnames(geneExpression), sampleInfo$filename)
identical(rownames(geneExpression), geneAnnotation$PROBEID)
```

However, we only know these are the correct columns/rows to check because the data creator (in this case the instructor) told us so. In the *ExpressionSet* object, we ensure this connection is established by forcing the `rownames()` of the `assayData` to match the `rownames()` of `featureData` and the `rownames()` of `phenoData` to match the `colnames()` of `assayData`.

Now let's create the objects we will use in the `phenoData` slot by applying the `AnnotatedDataFrame()` function to the appropriate table. Call this object `pd`, make sure to give it `rownames()` that match the `colnames()` of `geneExpression`.

Once you are done creating these objects, find the experiment date:

```
pData(pd)[ "GSM136530.CEL.gz", "date" ]
```

What is the date of this experiment?

Input the date in YYYY-MM-DD format, as shown in R.



Submit

You have used 1 of 5 attempts

✓ Correct (1/1 point)

Question 4

1/1 point (graded)

Running `varLabels()` on our *AnnotatedDataFrame* `pd` will list the variable names for the phenotype data:

```
varLabels(pd)
```

What is the first variable name?

ethnicity



Submit

You have used 1 of 5 attempts

✓ Correct (1/1 point)

Question 5

1/1 point (graded)

Note that the *ExpressionSet* class has a slot for storing a table describing the features (rows). In our example, annotations for the features are included in the `geneAnnotation` object. Here we will create an *AnnotatedDataFrame* version of `geneAnnotation` to add to the `featureData` slot of an *ExpressionSet* at a later step. Use the `AnnotatedDataFrame` function to create an object `fd`. Make sure the `rownames()` match the `rownames()` of `geneExpression`.

After creating this object, run the command:

```
pData(fd)["204810_s_at", "CHR"]
```

What chromosome is this gene from?

Give your answer in this form: chrZ

chr19



Submit

You have used 1 of 5 attempts

✓ Correct (1/1 point)

Question 6

1/1 point (graded)

Now we are ready to create our *ExpressionSet* object. To do this we simply call the `ExpressionSet()` function on the `geneExpression`, `pd` and `fd` objects. The latter two were created in question 3.3.2. Call the new object `eset`.

After creating `eset`, run the following lines to explore the expression difference on the Y chromosome in males versus females.

```
ind1 <- which(featureData(eset)$CHR == "chrY")
ind2 <- pData(eset)$group == 1
femaleY <- colMeans(exprs(eset)[ind1, ind2])
maleY <- colMeans(exprs(eset)[ind1, !ind2])
boxplot(maleY, femaleY)
median(maleY) - median(femaleY)
```

What is the difference in median expression of genes on the Y chromosome between `maleY` and `femaleY`?



You have used 1 of 5 attempts

✓ Correct (1/1 point)

Question 7

1/1 point (graded)

Suppose that you now want to subset on just the first 10 feature and first 5 samples of this dataset.

Which of the following options is the most efficient way to create such an object?



```
eset <- ExpressionSet(assayData = geneExpression[1:10, 1:5],  
                      phenoData = pd[1:10, ],  
                      featureData = fd[1:5, ],  
                      annotation = "hgfocus")
```



```
geneExpression <- geneExpression[1:10, 1:5]  
sampleInfo <- sampleInfo[1:10, ]  
geneAnnotation <- geneAnnotation[1:5, ]
```

`eset[1:10, 1:5]`

It is not possible

[Submit](#)

You have used 1 of 5 attempts

✓ Correct (1/1 point)

Discussion

[Hide Discussion](#)**Topic:** Section 3 / ExpressionSet[Add a Post](#)

Show all posts ▼

by recent activity ▼

There are no posts in this topic yet.

[Learn About Verified Certificates](#)

© All Rights Reserved