



[Course](#) > [Section 1: Biology b...](#) > [A Review of What ...](#) > Gene signature ass...

Gene signature assessment

The 70-gene diagnostic signature for breast cancer progression risk

1/1 point (graded)

The genefu package has been a component of Bioconductor since 2011. Its DESCRIPTION file says it is an "R package providing various functions relevant for gene expression analysis with emphasis on breast cancer". You can use this package to get acquainted with aspects of the mammaprint diagnostic test that Rafa mentioned in the lecture. Install the genefu package with biocLite.

```
library(BiocInstaller)
biocLite("genefu")
```

A data.frame with information on the 70 gene signature used in the mammaprint algorithm is in the `sig.gene70` data.frame. You can have a look at this:

```
library(genefu)
data(sig.gene70)
dim(sig.gene70)
head(sig.gene70)[,1:6]
```

You can see from this that there are 70 records in the data frame, and that there are diverse ways of describing the "genes" in the signature.

How many components of the signature have a missing value for the associated NCBI gene symbol? (Remember to use `is.na`, never `== NA`.)



You have used 2 of 5 attempts

Kinases in the 70 gene signature

1/1 point (graded)

Kinases are important for cell-cell communications; see the Wikipedia entry on Kinase for some background.

You can use `grep` on the Description field of the `sig.gene70` data.frame to search for substrings of long gene names. How many of the members of the 70-gene signature are genes coding for kinases?



You have used 2 of 5 attempts

