

Section 1: What we measure and why

Mammaprint Gene Signature

- Exploring genes used in the Mammaprint gene signature - assess risk of breast cancer
- Diagnostic signature using gene expression levels of 70 genes
- Information about the 70 gene signature used in the Mammaprint algorithm

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

Assessment: Phenotypes

- COPDSexualDimorphism.data package - phenotypes (cols) individuals (rows)
- Data to assess incidence of COPD and emphysema by gender and smoking status
- The pkys variable in the expr.meta data.frame represents pack years smoked. Other variables include gender and diagmaj (disease status). These variables correspond to phenotypes.

```
library(COPDSexualDimorphism.data)
data(lgrc.expr.meta)
```

Assessment: Chromosomes and SNPs

- GWAS (Genome-wide association studies)
- Comparing individuals with disease vs. controls using SNP chips or DNA sequencing.
- SNPs with association are investigated for disruption of gene regulation or function
- Bioconductor *gwascat* package

```
library(gwascat)
data(ebicat_2020_04_30)
ebicat_2020_04_30
```