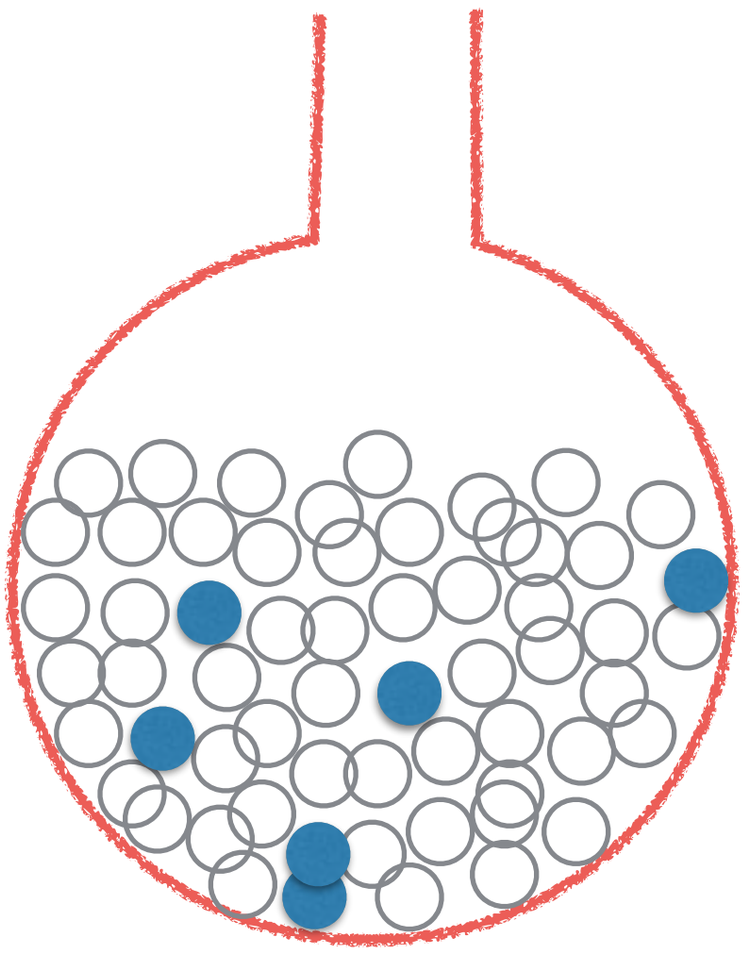
Over-representation analysis

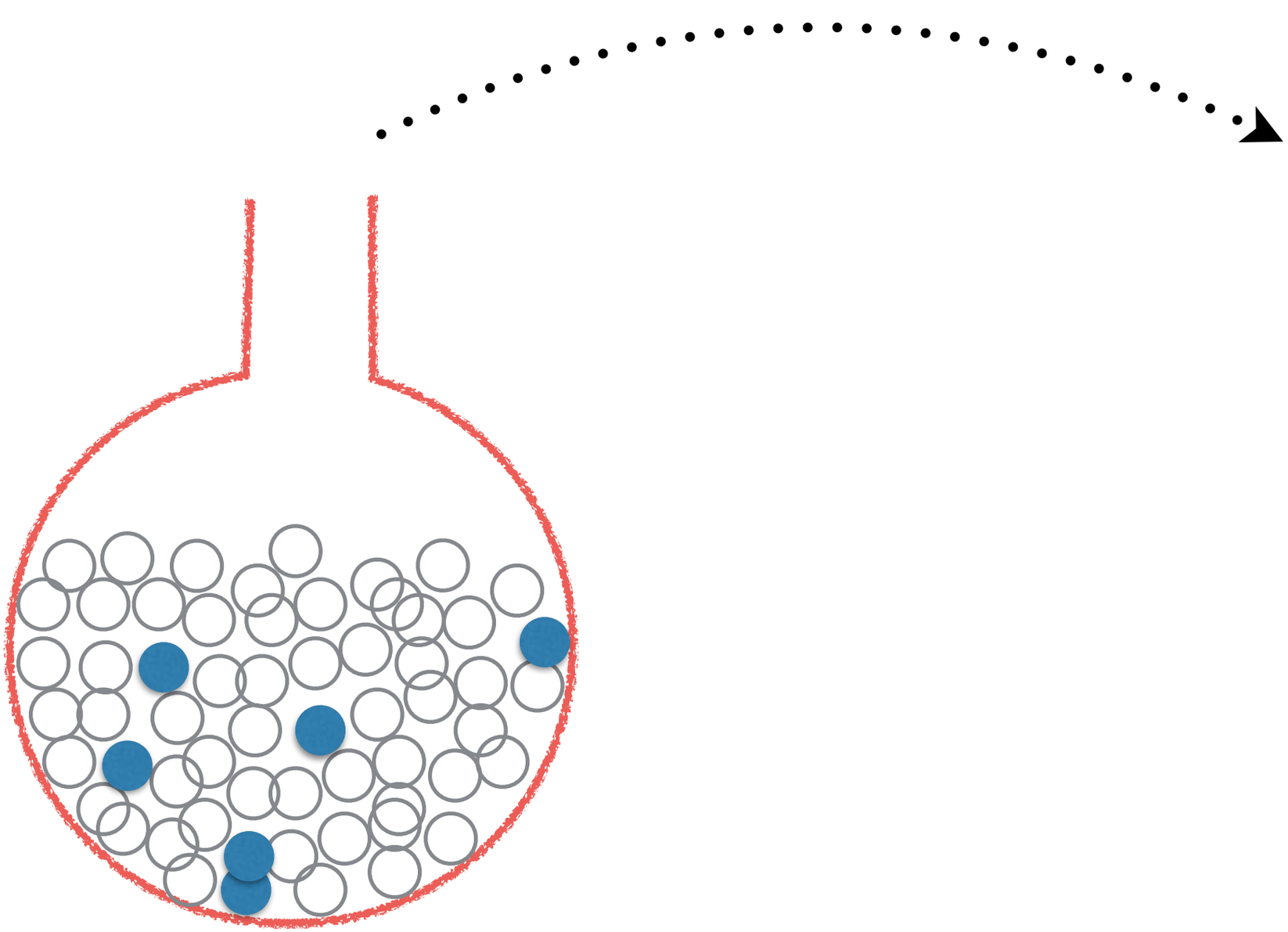


Find 1,000 genes differentially expressed in response to shade.

If instead of expectation, you find that 50 of these 1,000 have the GO term “Cell wall modification”, then what do you conclude?

Genome with 34,000 genes
340 have GO term:
“Cell wall modification”

Over-representation analysis: additional considerations



1) The Jar (aka Universe) should only contain **expressed genes.** (Why?)