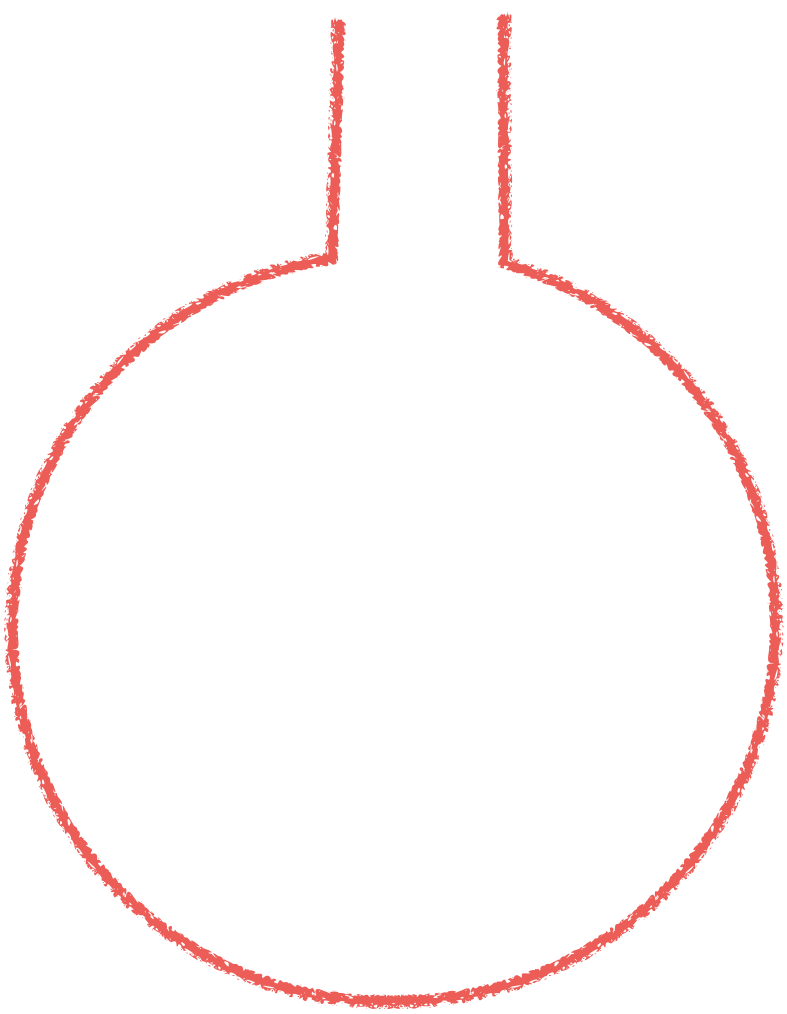


Over-representation analysis: additional considerations













































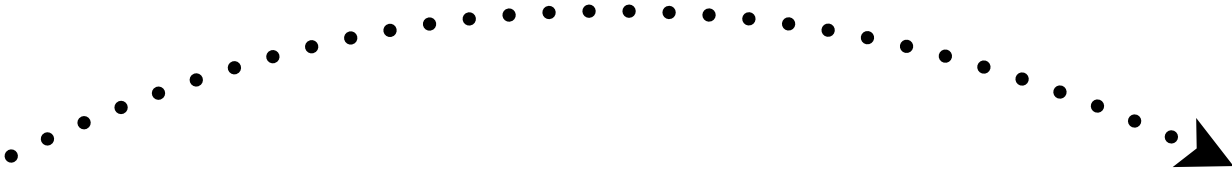








The Jar (aka Universe) should only
contain **expressed genes**. (Why?)



Biases in detecting differentially
expressed genes should
be considered

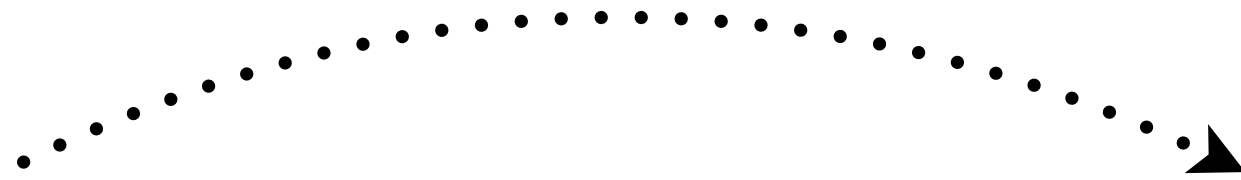
Statistical tests for over (or under) representation:

- Build a contingency table categorizing genes by D.E. and GO

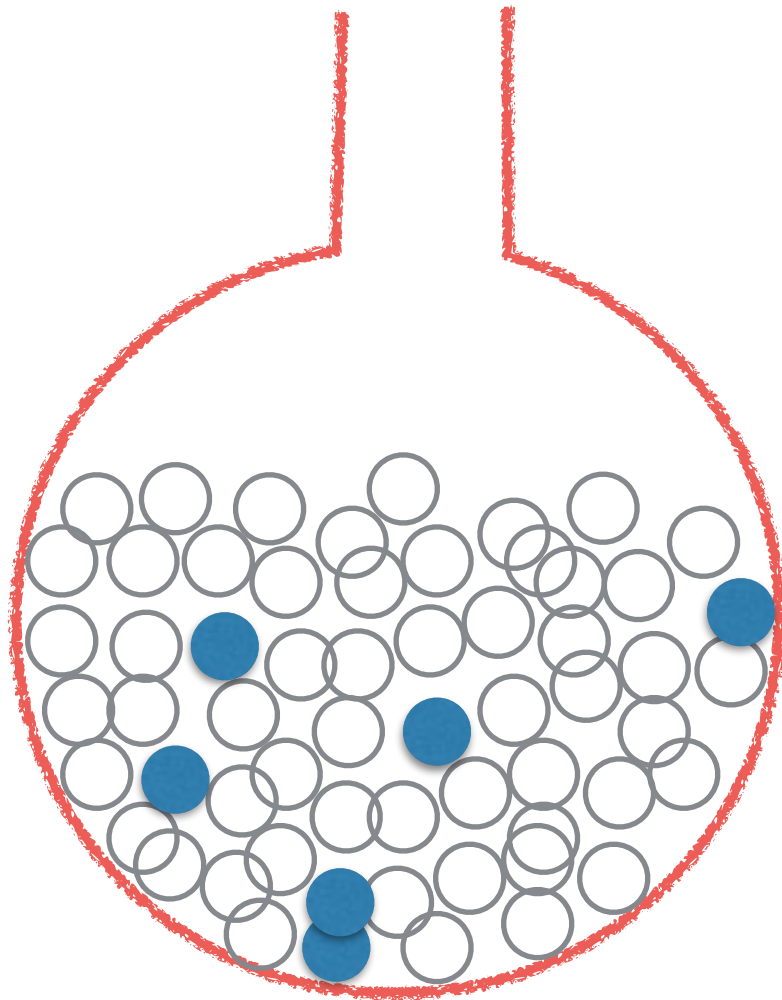
		Differentially Expressed	
		NO	YES
GO term “cell wall modification” present?	NO	32,710	950
	YES	290	50

- Use a (modified) Fisher’s Exact Test to look for unequal ratios (is 32710:950 different from 290:50)

Over-representation analysis: additional considerations



Biases in detecting differentially expressed genes should be considered



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