

Mailman et al (2007).The NCBI dbGaP database of genotypes and phenotypes. Nature Reviews Genetics 39(10): 1181-.

1. What are the general features of the database dbGaP?
2. dbGaP accommodates studies of varying design. What are the four basic types of data it contains?
3. How does dbGaP ensure confidentiality of study subjects? What are the responsibilities of the PI (primary investigator) to ensure secure use of the data?
4. What does the term “authorized data” mean in the context of dbGaP?
5. What are some of the data cleansing and quality control measure taken by the maintainers of dbGaP?
6. Are all data provided by dbGap non-public? What are non-public data sources? What do they include? Do you think summary data are secure enough?

Flintoft (2005). From genotype to phenotype: a shortcut through the library. *Nature Reviews Genetics* 6: 1.

7. What is an alternative method to predict the functions of large sets of sequenced genes?
8. What do researchers generally conclude about shared orthologous genes, when a group of species also share a phenotypic trait?
9. Which search tools does the new method involve?
10. Can this method also be employed to link genes to disease-related phenotypes? If so, what would be the benefit of this method above the classical method to “model” the relation between disease-phenotype and genes?