# DNA sequence questions

1. From which databases could you obtain a DNA sequence that was from a gene you are interested in?
   1. What might be be important information to record about your DNA sequence and why?
2. What is:
   1. A complete genome?
   2. cDNA?
   3. Expressed Sequence Tag (ESTs)?
3. What is the most recent version of the human reference genome? Which version is used in the human genomics laboratory?
4. What is meant by non-coding sequences and do they have any functions?
5. Where could we get the annotations (such as name of the gene, population frequency etc) that we might want to interpret variation found in our sample?
   1. Where does it come from in the bioinformatics pipeline?
   2. Where might a genomics clinical scientist get this information?
6. If the same genes are present in the DNA over the whole body, then why when someone has a pathogenic variant might some tissues be affected, and others less obviously affected or apparently unaffected?

## Resources

### NCBI

<https://www.ncbi.nlm.nih.gov/>

### Ensembl

<https://grch37.ensembl.org/Homo_sapiens/Info/Index>

### Variant Annotation

<https://www.ensembl.org/info/docs/tools/vep/index.html>

<https://www.interactive-biosoftware.com/alamut-visual/>