
DTU

DTU Health Technology
Bioinformatics

Association analysis

Gisle Vestergaard
Associate Professor
Section of Bioinformatics
Technical University of Denmark
gisves@dtu.dk

Cohort based studies

GWAS: Genome-wide association study

MWAS: Metagenomice-wide association study

Fig. by NHS:
https://www.genomicseducation.hee.nhs.uk

GWAS

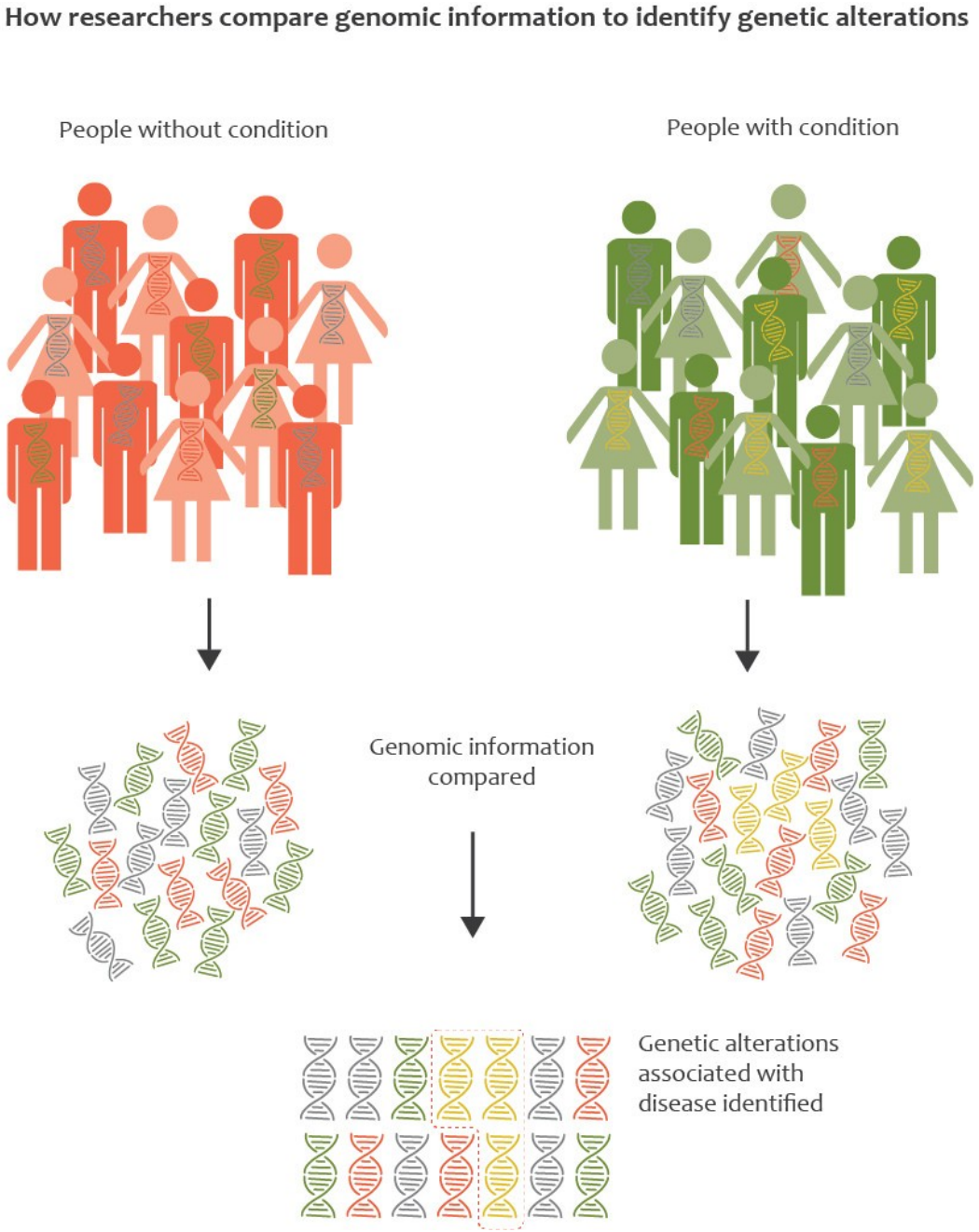
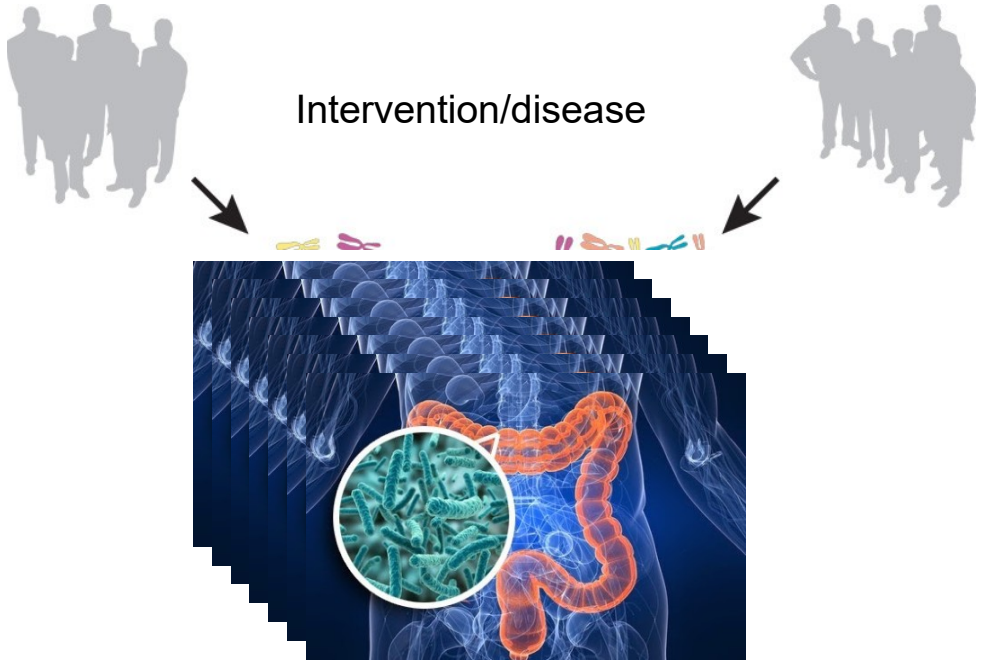


Fig. by NHS:
https://www.genomicseducation.hee.nhs.uk

MWAS

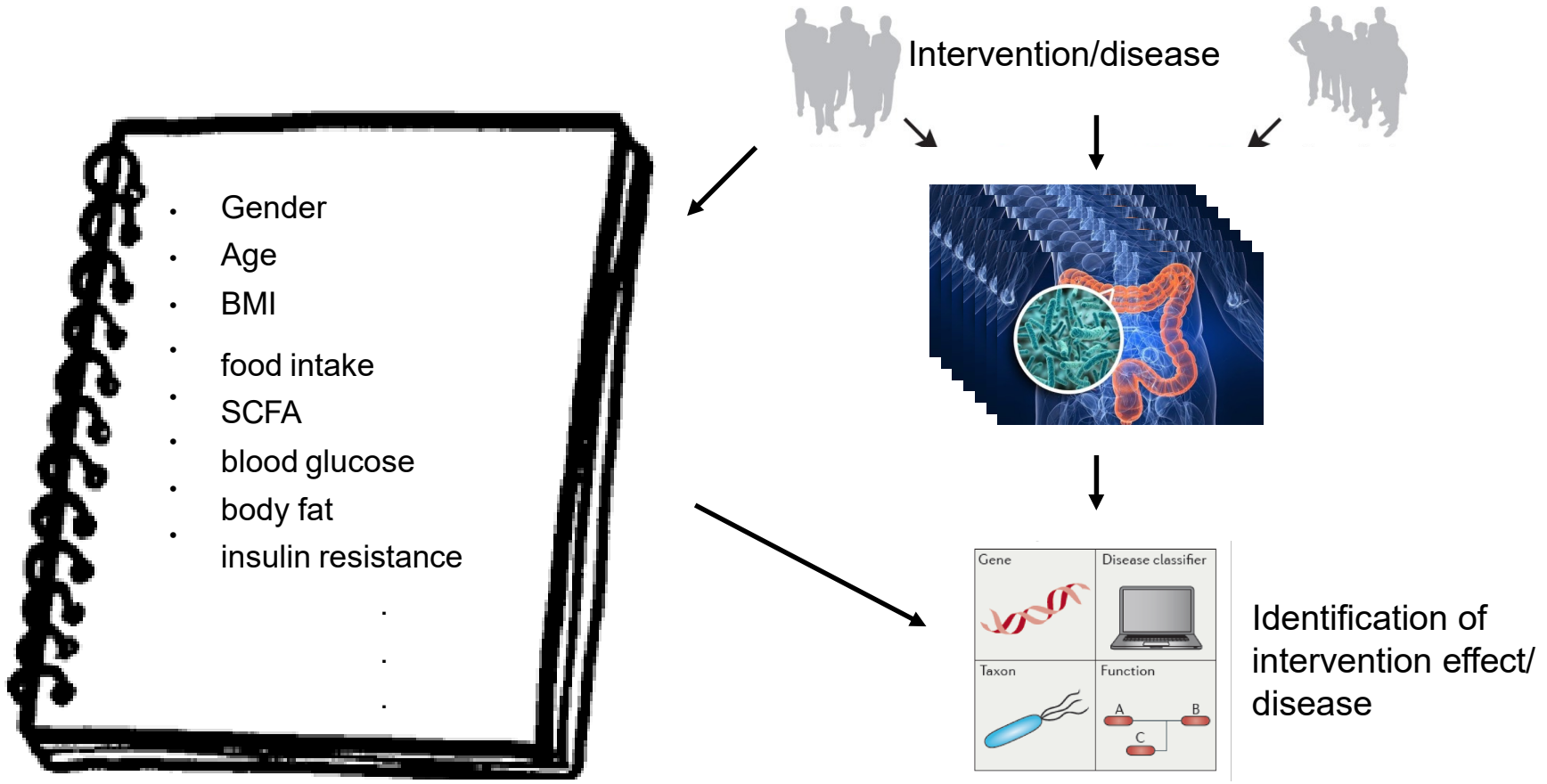


Identification of intervention/disease

MWAS

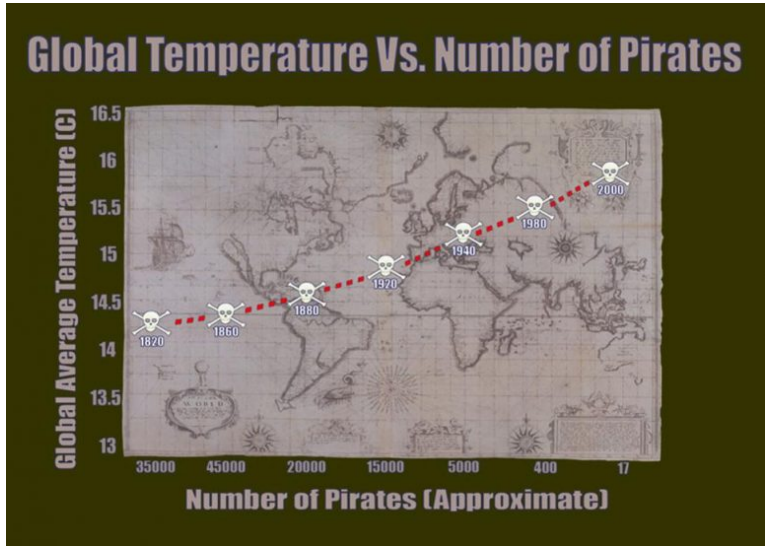
“The associations that can be identified by MWAS are not limited to the identification of taxa that are more or less abundant, as is the case with taxonomic approaches, but additionally include the **identification of microbial functions** that are enriched or depleted.” *Wang and Jia (2016)*

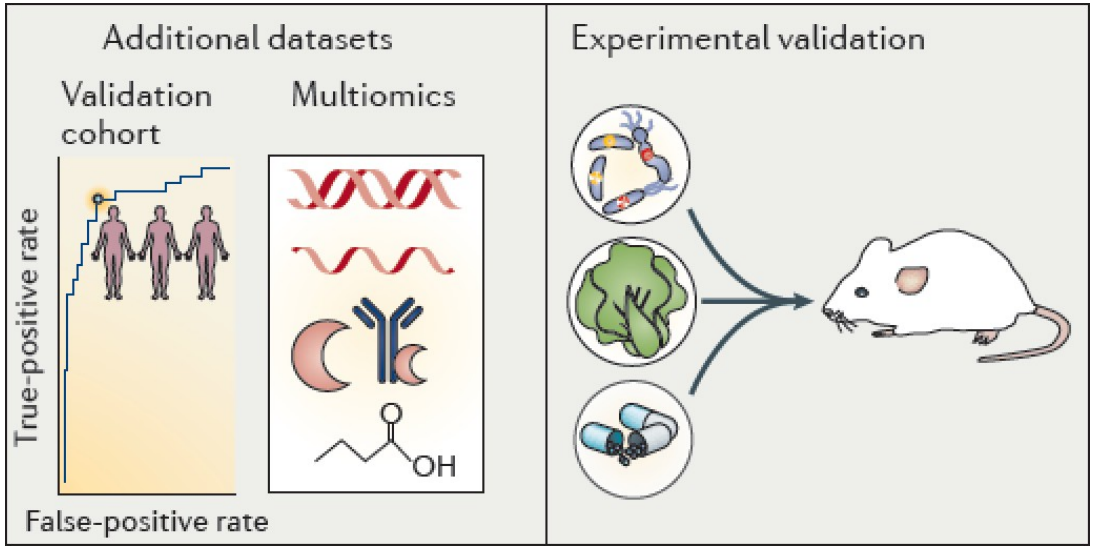
Metadata association



Remember

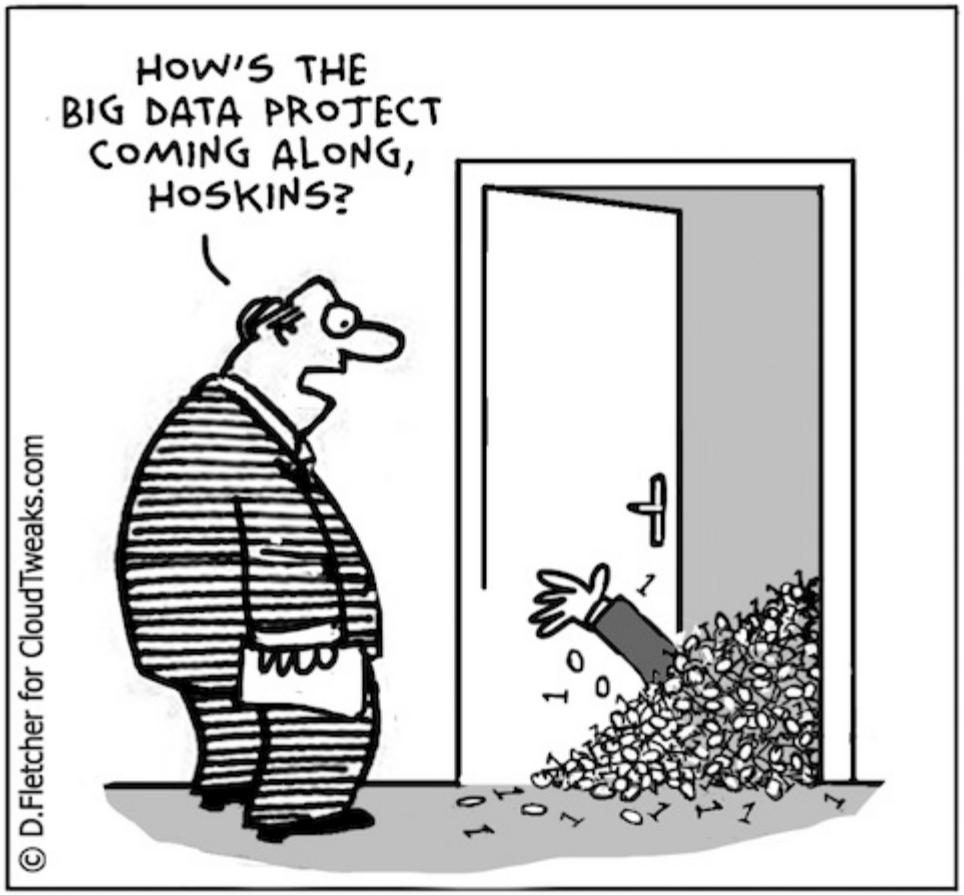
- **Correlation** does **not imply causation**
- Validation
  - Literature
  - Follow-up studies
  - Otheromics: metatranscriptomics, metaproteomics, metabolomics
  - Longitudinal studies
  - Experimental evidence such as animal models





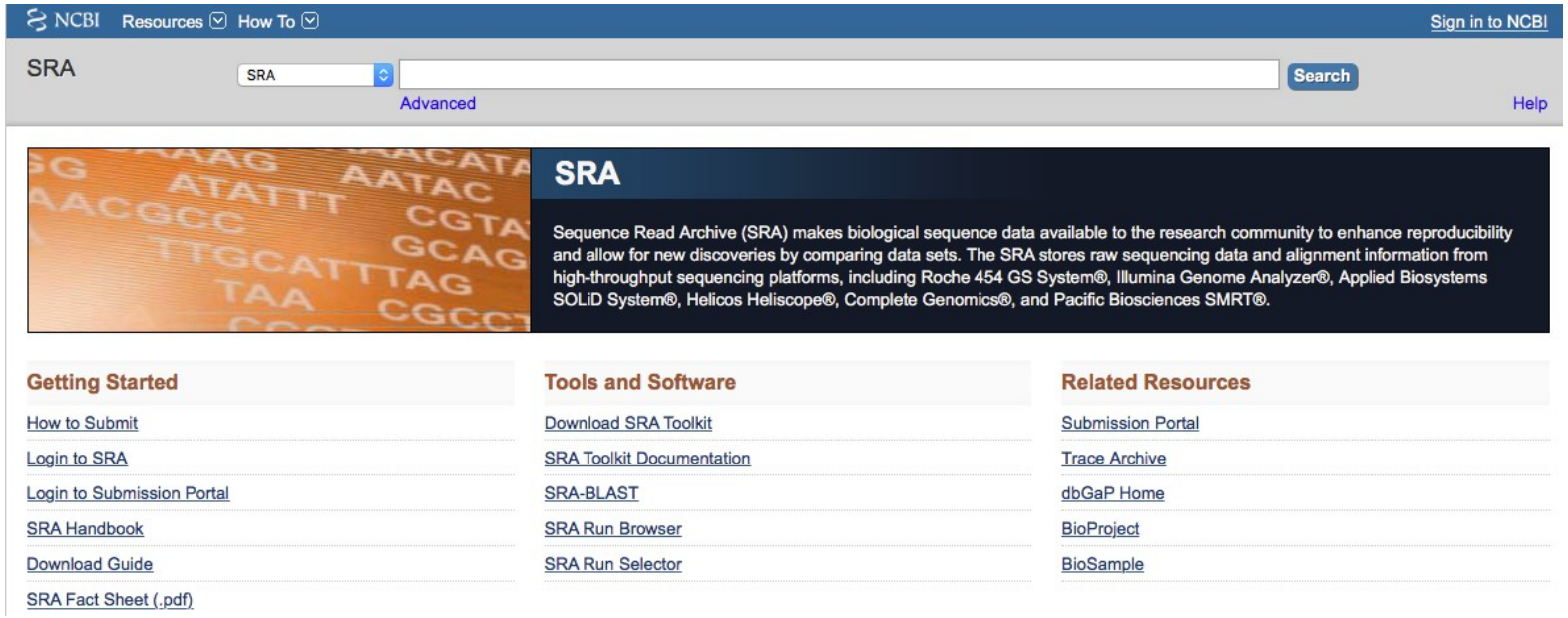
Good to know: Functions and Public Databases

- Finding data that can help answer your question



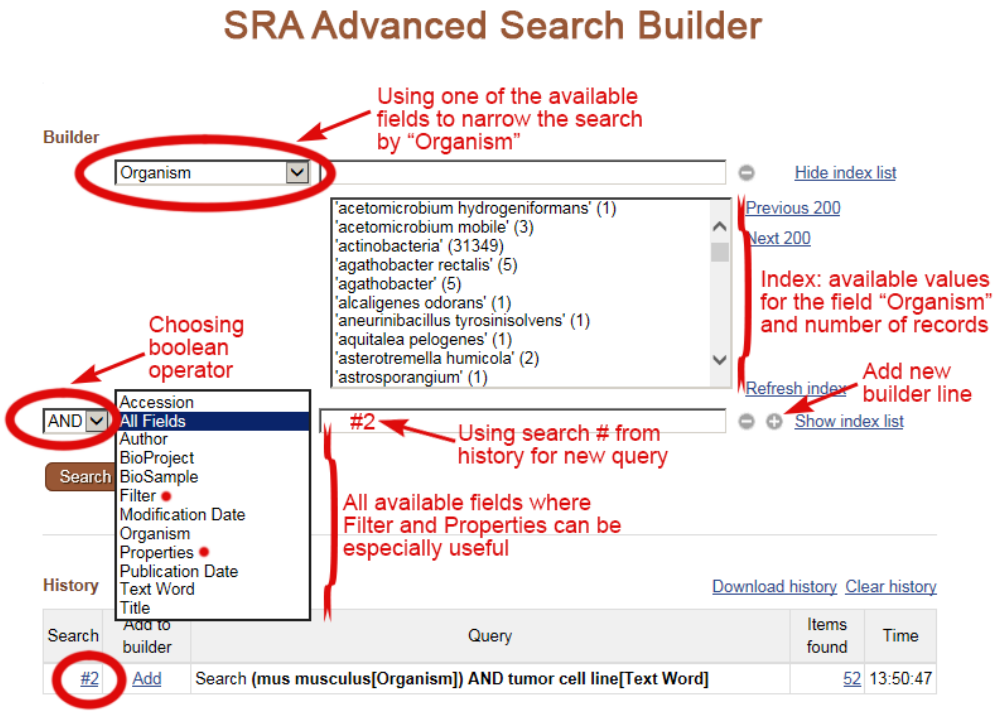
SRA

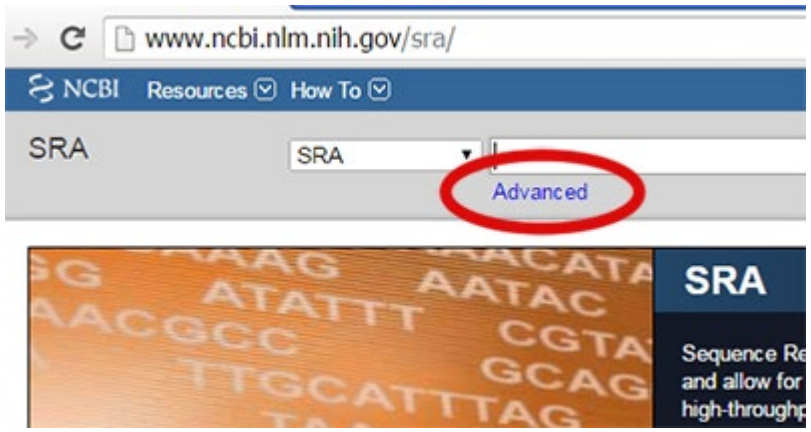
- SRA:
  - short read archive (part of NCBI)
  - https://www.ncbi.nlm.nih.gov/sra/



Searching SRA

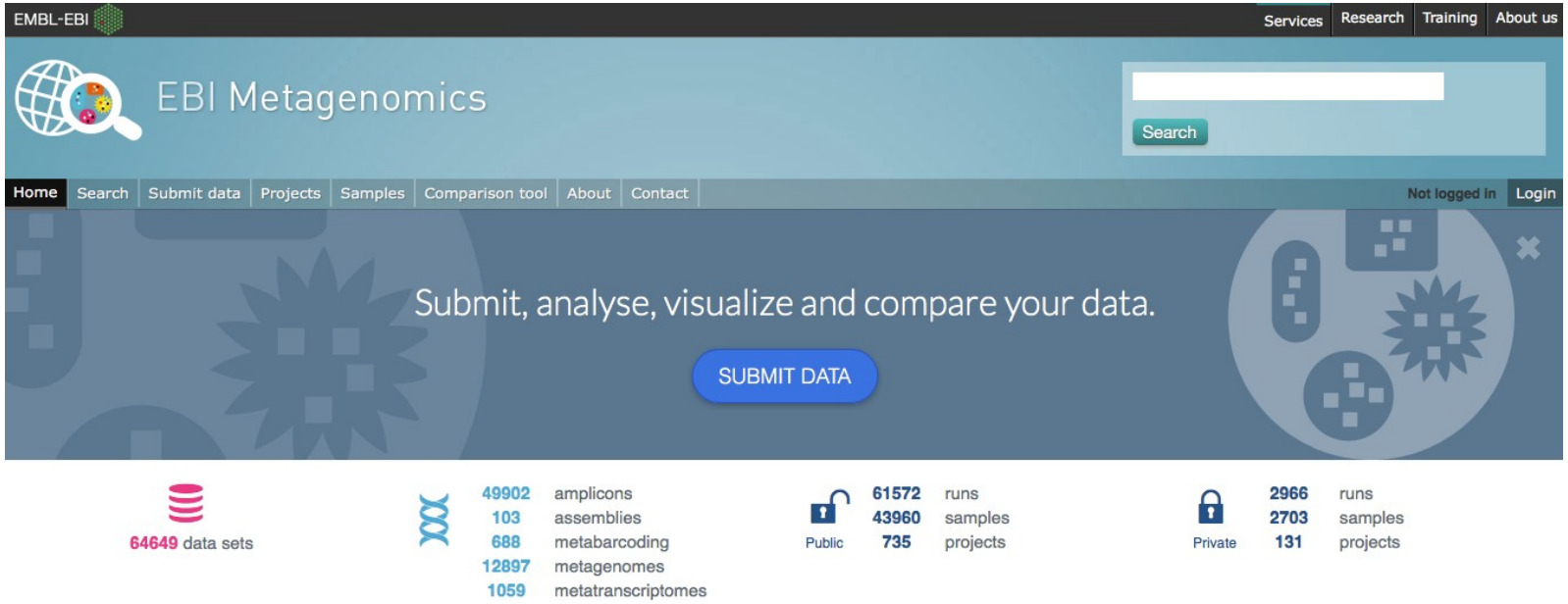
https://www.ncbi.nlm.nih.gov/sra/advanced





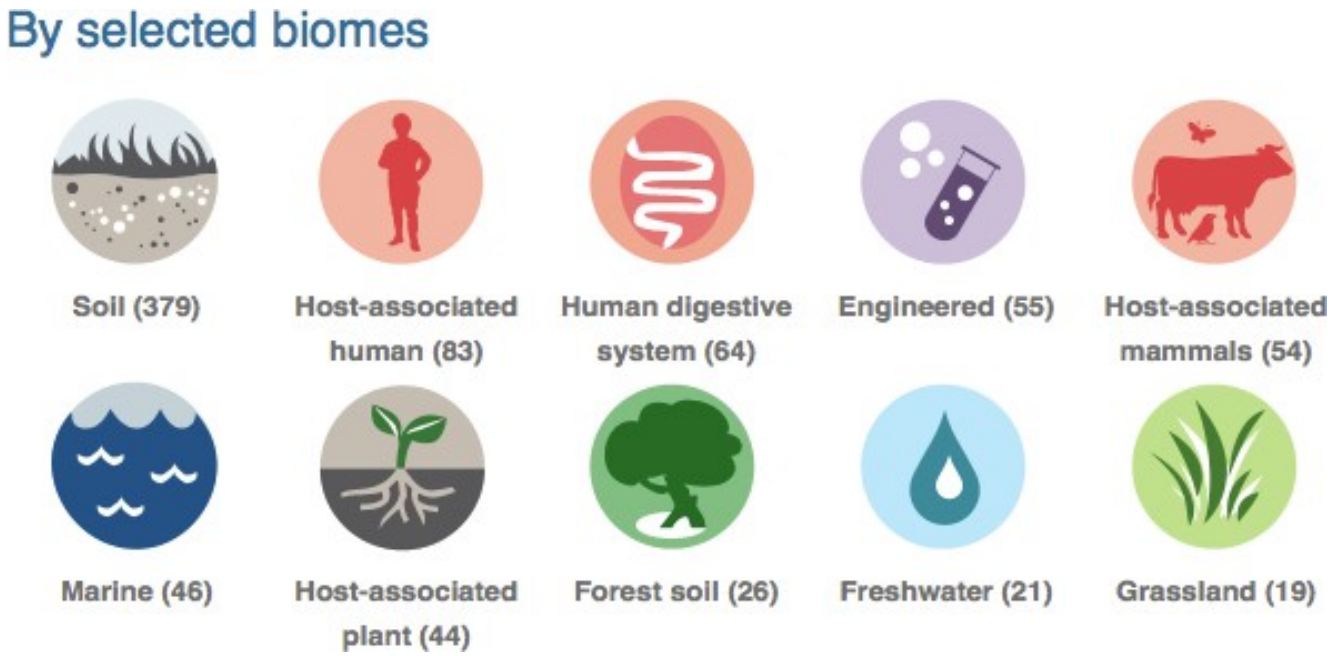
ENA

- ENA:
  - European Nucleotide Archive
  - https://www.ebi.ac.uk/metagenomics/ (metagenomic datasets)



Searching ENA

- ENA:
  - European Nucleotide Archive
  - https://www.ebi.ac.uk/metagenomics/ (metagenomic datasets)



Functional annotations

GO: Gene onology

- defines concepts/classes used to describe gene function, and relationships between these concepts

KEGG: Kyoto Encyclopedia of Genes and Genomes

- database resource for understanding high-level functions and utilities of the biological system

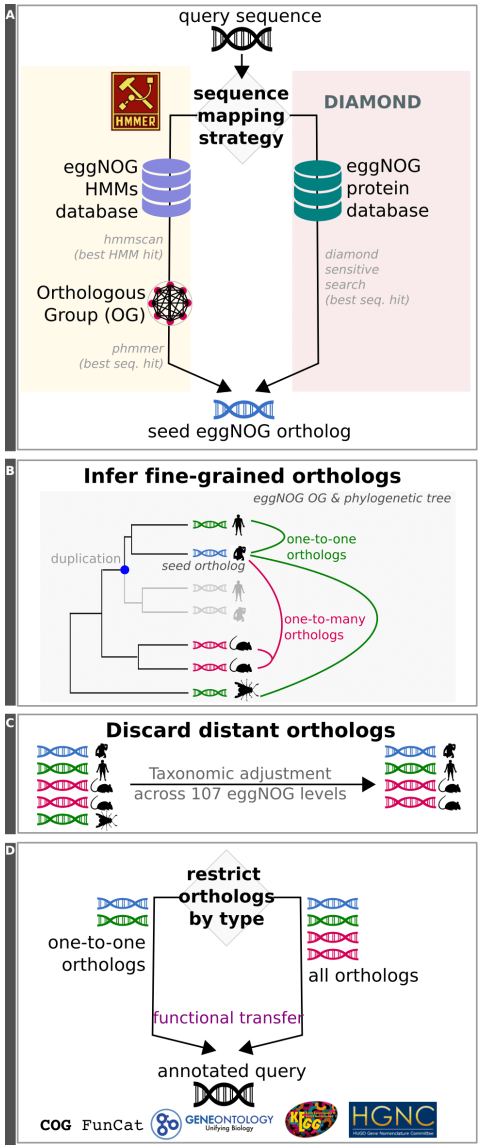
eggNOG

- A database of orthologous groups and functional annotation
- Cross-references to other databases

EggNOG-mapper

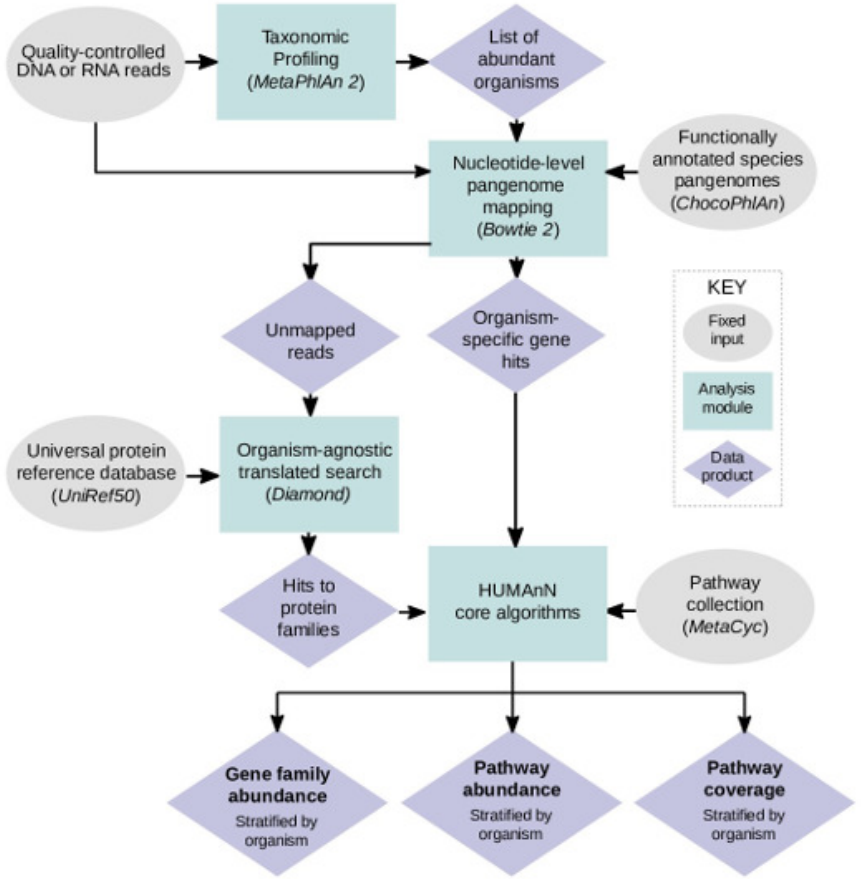
1. query_name
2. seed eggNOG ortholog
3. seed ortholog evalue
4. seed ortholog score
5. Predicted taxonomic group
6. Predicted protein name
7. Gene Ontology terms
8. EC number
9. KEGG_ko
10. KEGG_Pathway
11. KEGG_Module
12. KEGG_Reaction
13. KEGG_rclass
14. BRITE
15. KEGG_TC
16. CAZy
17. BiGG Reaction
18. tax_scope: eggNOG taxonomic level used for annotation
19. eggNOG OGs
20. bestOG (deprecated, use smallest from eggnog OGs)
21. COG Functional Category
22. eggNOG free text description

- Functional annotation of reads or ORFs
- Cross-references to other databases
- Complex tabulated output



HUMAnN2

- Very easy to install and use



How to hands on:
Association analysis exercise