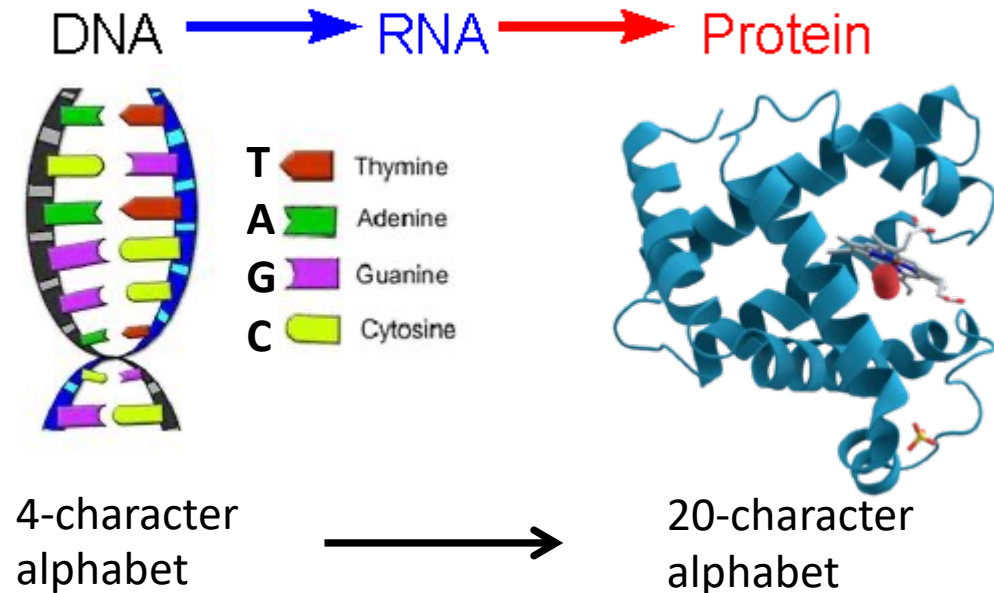


Bioinformatics Wrap-Up

Garrett Dancik, Ph.D.

What is bioinformatics

- Bioinformatics:
 - Biology + information
 - the study and utilization of methods for storing, retrieving and analyzing biological data

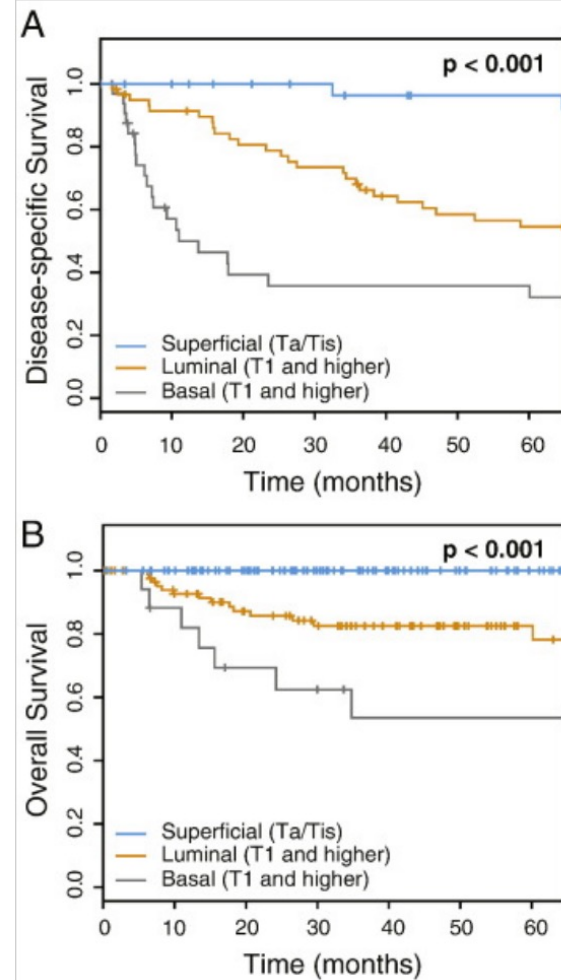
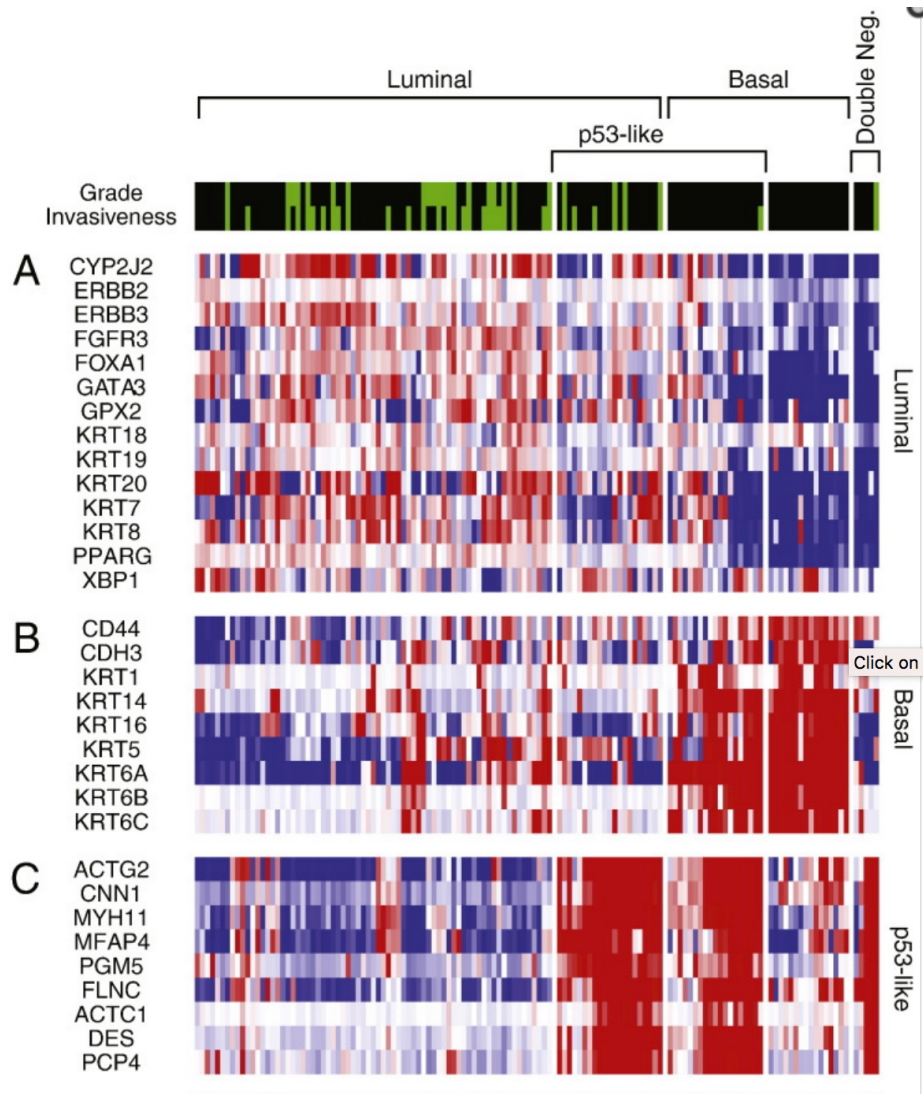


- How much information:
 - Human genome: 3 billion nucleotides
 - ~20,000 genes
 - many more when considering “junk DNA” and alternative splicing
 - >10 million sites of DNA variation
 - Countless possible interactions between DNA, RNA, and proteins

Why do we need bioinformatics?

- To identify genetic mechanisms of diseases and other inherited (or acquired) conditions
 - Nature via nurture
- For personalized treatment of disease

Bladder cancer subtypes



Additional Examples

- Gene expression analysis:
 - Development of a RNA-Seq Based Prognostic Signature in Lung Adenocarcinoma (<https://pubmed.ncbi.nlm.nih.gov/27707839/>)
 - RNA-seq transcriptome analysis of breast cancer cell lines under shikonin treatment (<https://pubmed.ncbi.nlm.nih.gov/29422643/>)
 - Expression of the SARS-CoV-2 cell receptor gene *ACE2* in a wide variety of human tissues (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7186534/>)
 - RNA-Seq and molecular docking reveal multi-level pesticide resistance in the bed bug (<https://pubmed.ncbi.nlm.nih.gov/22226239/>)
- Additional data analysis examples:
 - Facebook Data: <http://www.cnn.com/2013/03/11/tech/social-media/facebook-likes-study/>
 - Netflix Prize: <http://www.wired.com/2009/09/how-the-netflix-prize-was-won/>
 - And many more...

Amazon Movie Review Dataset

- Let's analyze data from about 8 million Amazon movie reviews
 - Why not?
 - <http://snap.stanford.edu/data/web-Movies.html>
 - Processed version is posted on Blackboard