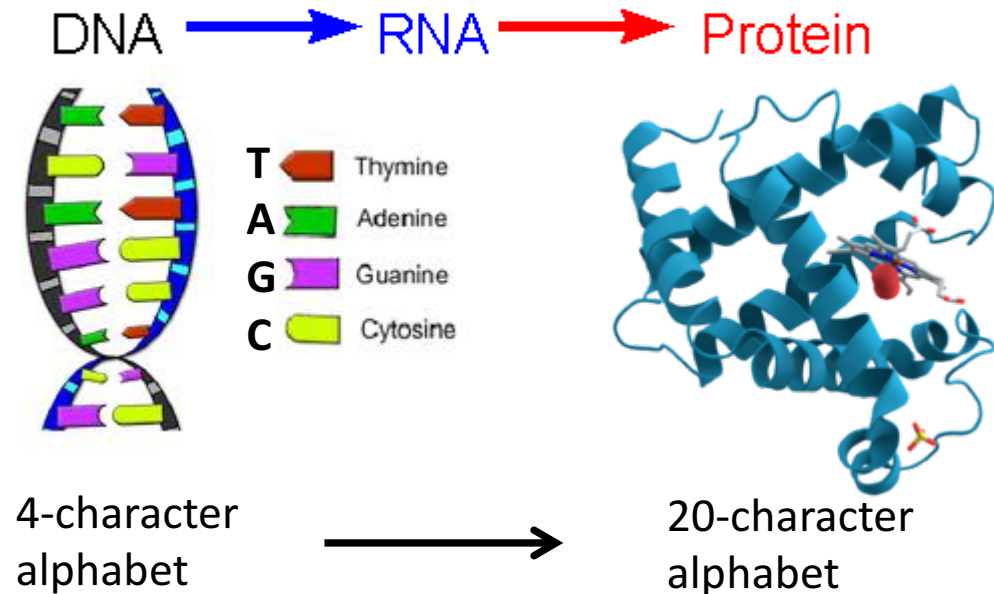


Bioinformatics Wrap-Up

Garrett Dancik, Ph.D.

What is bioinformatics

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

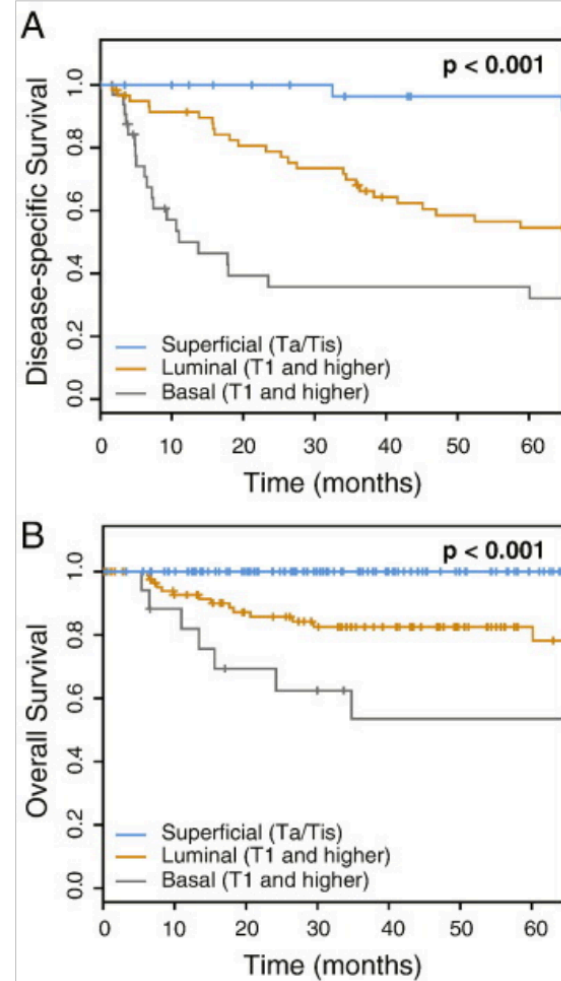
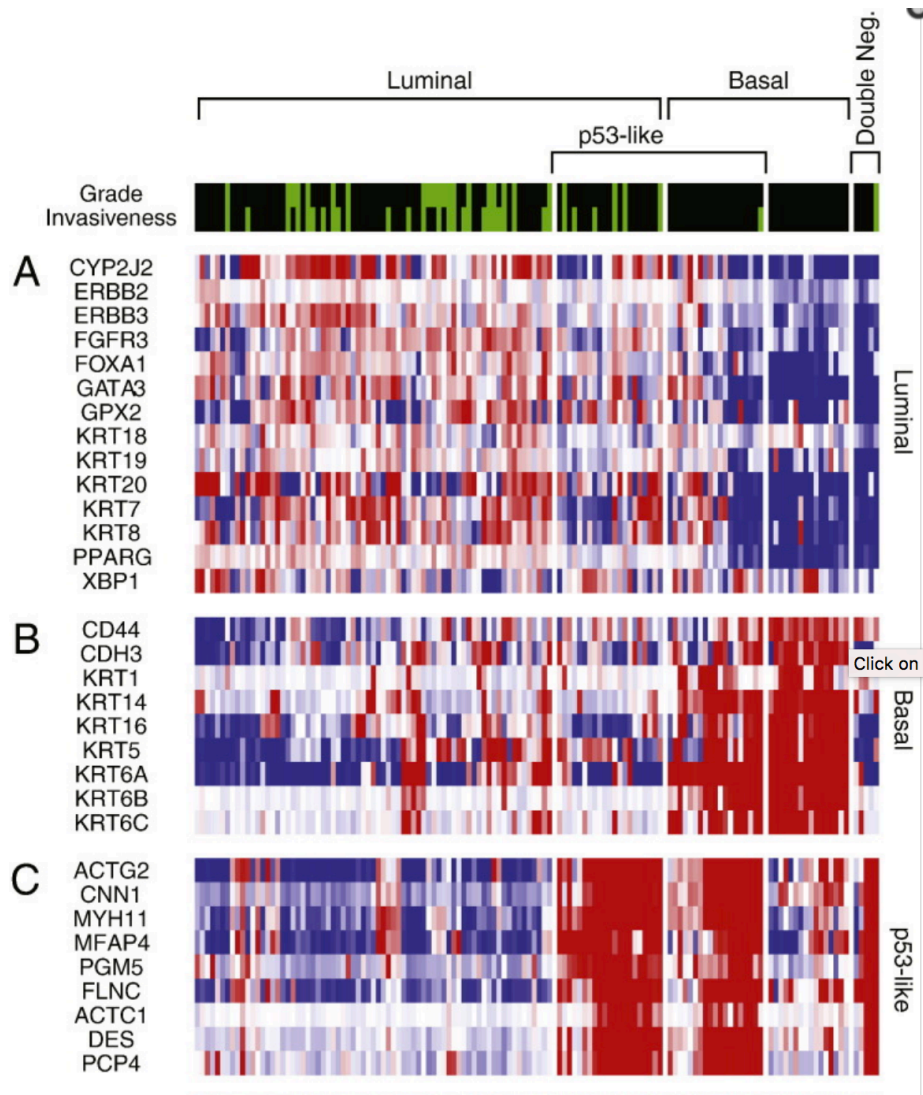


- How much information:
 - Human genome: 3 billion nucleotides
 - ~20,000 genes
 - many more when considering “junk DNA” and alternative splicing
 - >10 million sites of DNA variation
 - Countless possible interactions between DNA, RNA, and proteins

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:
 - Comprehensive Gene Expression Analysis of Prostate Cancer Reveals Distinct Transcriptional Programs Associated with Metastatic Disease (<http://cancerres.aacrjournals.org/content/62/15/4499.short>)
 - Gene Expression Signatures Diagnose Influenza and Other Symptomatic Respiratory Viral Infections in Humans (<http://www.sciencedirect.com/science/article/pii/S1931312809002510>)
 - Gene Expression Phenotypes of Arabidopsis Associated with Sensitivity to Low Temperatures (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC167029/>)
- 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>