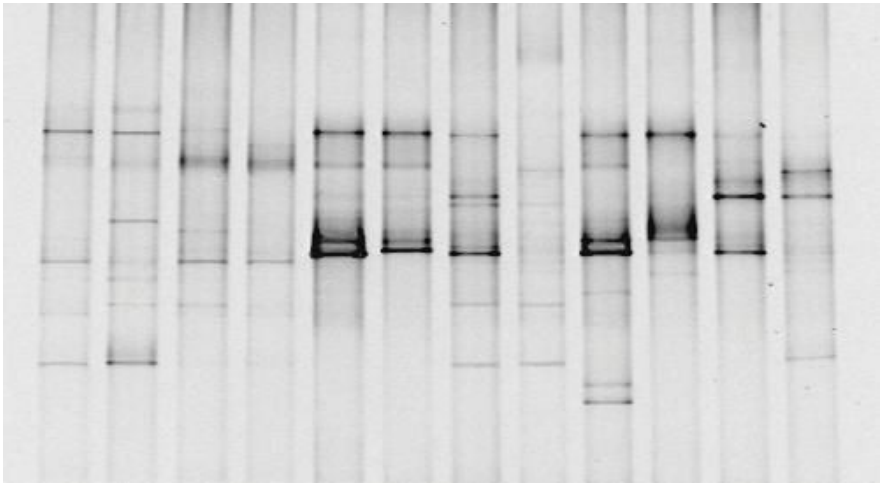
Bioinformatics Applications

Lucas Bishop, M.S.
Nevada Bioinformatics Center

- The need for bioinformatic (BioX) analysis is present in most life sciences lab
- Graduate degree = likely will need to learn some BioX
- Reveals a lot of information that would not otherwise be available from basic science.
    - ex) How does a low- or high- calorie diet affect the life cycle of a fly? We know it will shorten/lengthen lifespan but what genes are up or down regulated?
- Lots of labwork that used to be dangerous/ involve lots of time or effort can be done computationally.

    - ex) Denaturing Gradient Gel Electrophoresis (DGGE) to estimate diversity of a microbial population

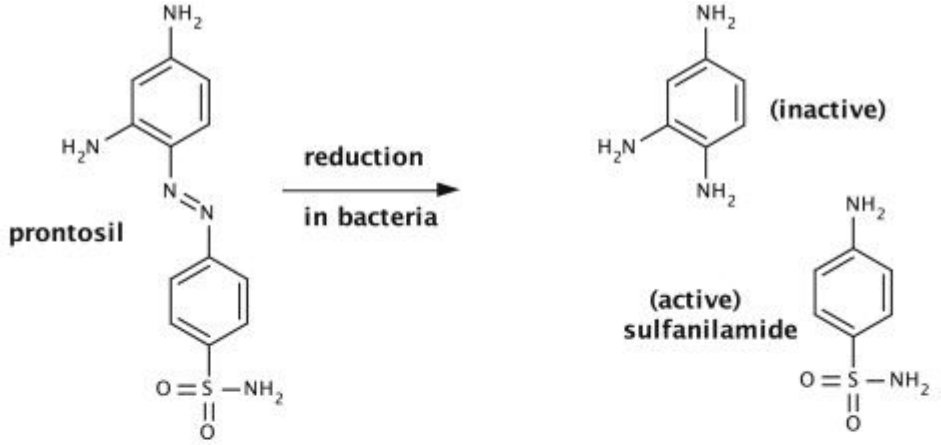


Diagnosis through bioinformatics

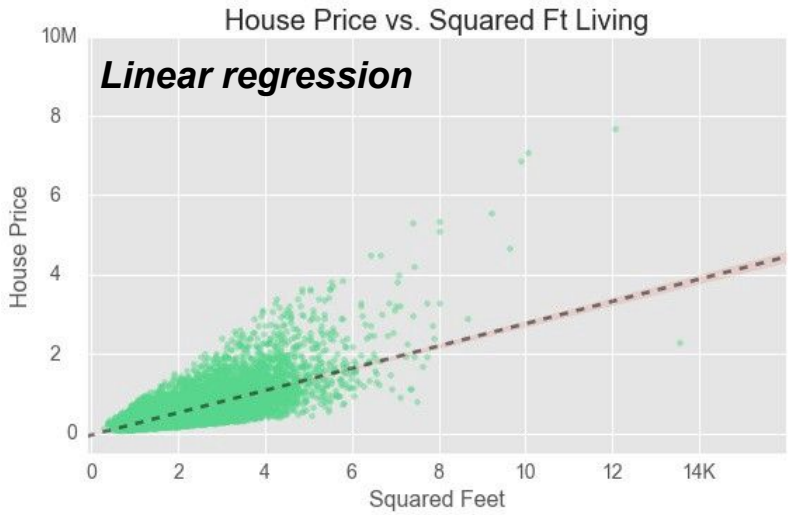
- BRCA1
  - Blood test to identify susceptibility due to mutation.
- Prenatal screens
  - Determine risk for genetic disorders and reduce risk of miscarriage.
  - NIPT: non-invasive prenatal screening; uses cell-free fetal DNA found in maternal plasma
- Fecal bioinformatic screens for colon cancer
  - Alternative to fecal immunochemical test (FIT)
  - Analyzes bacterial community and/or host cell expression
  - FAP: familial adenomal polyposis. Screens allow for non-invasive alternative to colonoscopies
- Identifying variants of infectious disease
  - Show the importance of genomic surveillance. Hospitals, wastewater, etc.

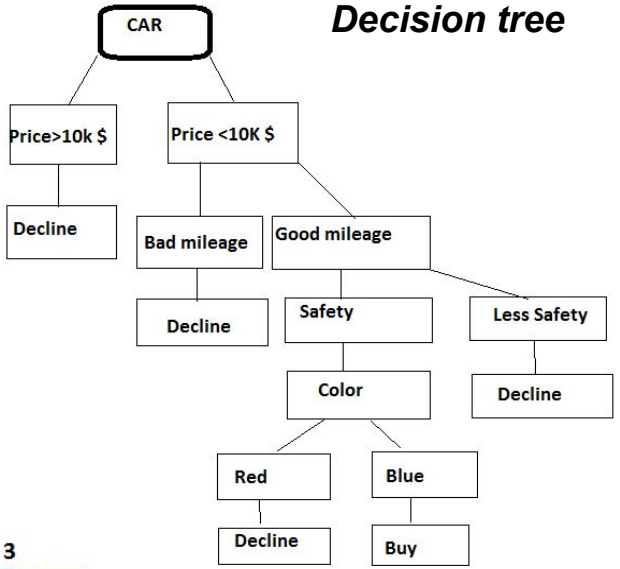
Bioinformatically-informed treatment

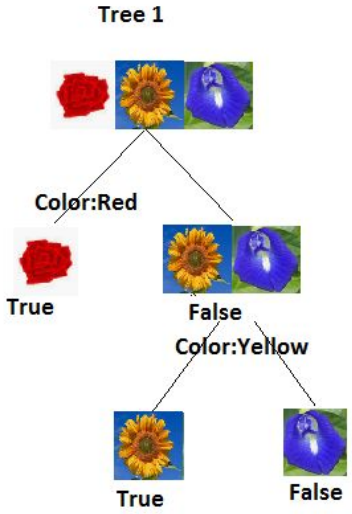
- By looking at an individual's genome, metabolome, microbiome, etc. can inform medications prescribed, treatments used.
- Xenobiotic metabolism: alternative metabolism of external compounds due to microbiota. (Abx, statins, probiotics, supplements, etc.)
- Example: prontosil inactivated by *E. lenta*
- Sequencing informs efficacy

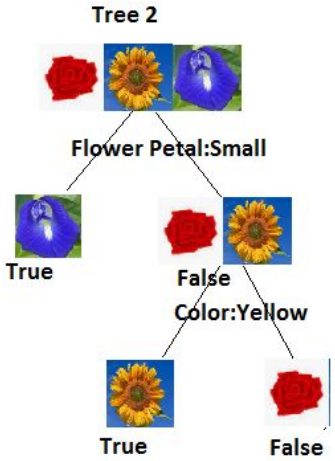


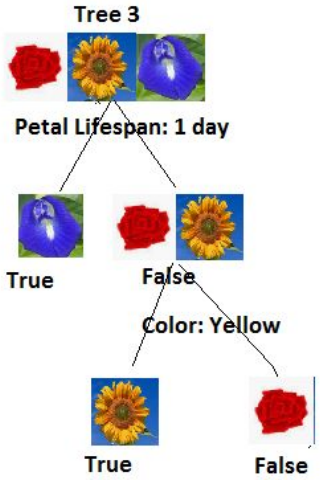
Modeling:









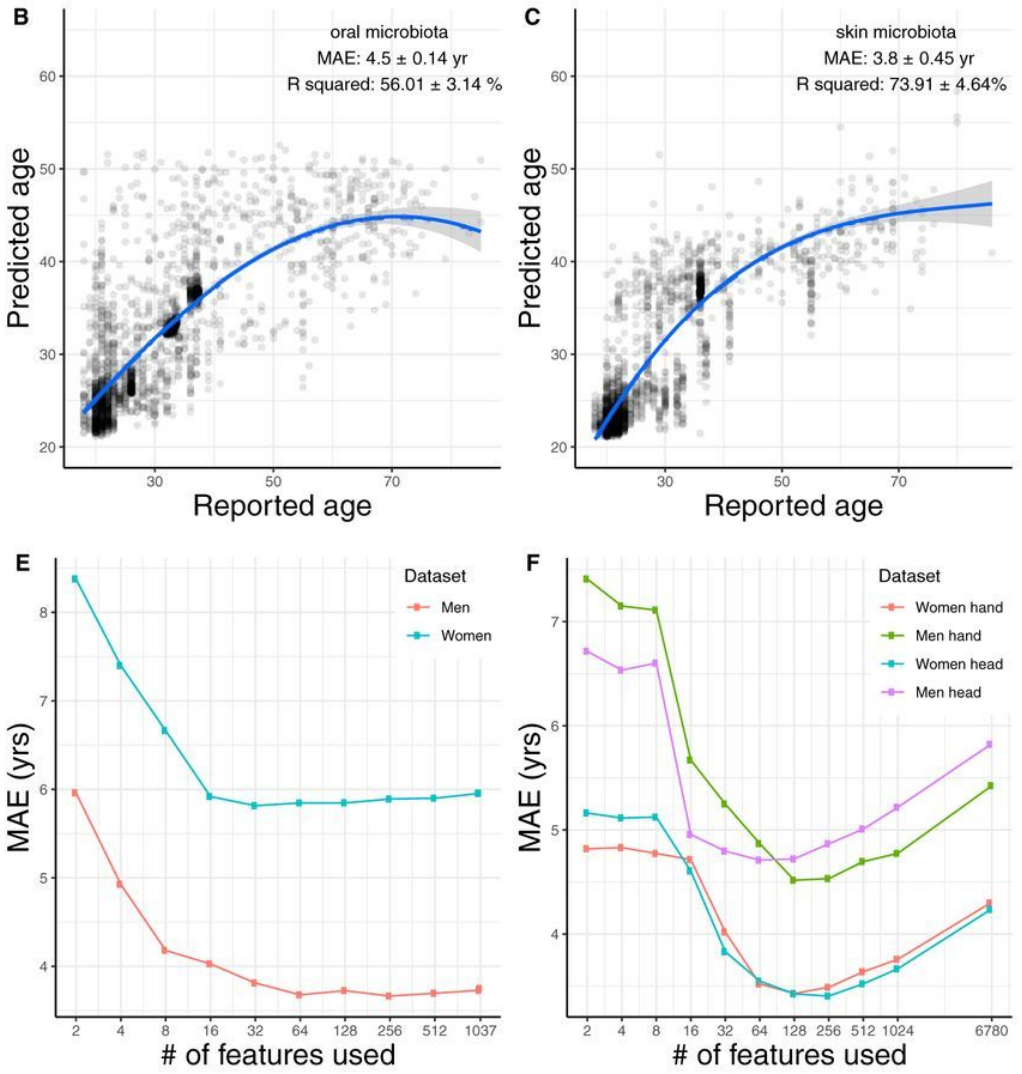


Random Forest

Images: courtesy of Aditya Kumar

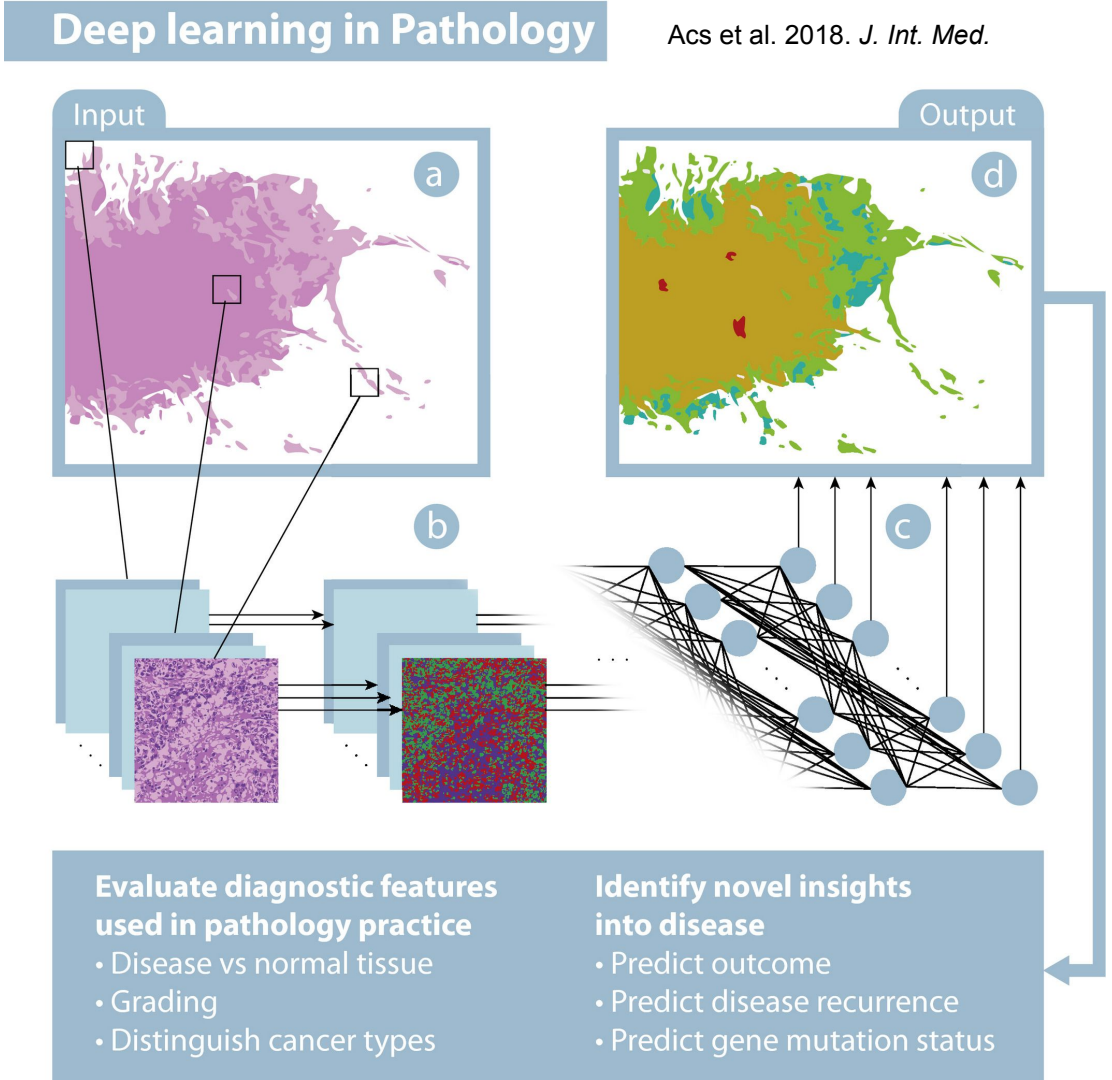
Modeling - Machine Learning

- By using a dataset with known outcomes (dx, age, etc.) can train a computer model to use features (genes, microbes, etc.) to predict future outcomes.
- Huang et al. 2018. *mSystems:*
  - Each feature is a species of bacteria
  - MAE = mean absolute error
  - A linear line, R2 = 1.0 would be perfect prediction
  - Skin is best predictor of age, within 4 years on avg.



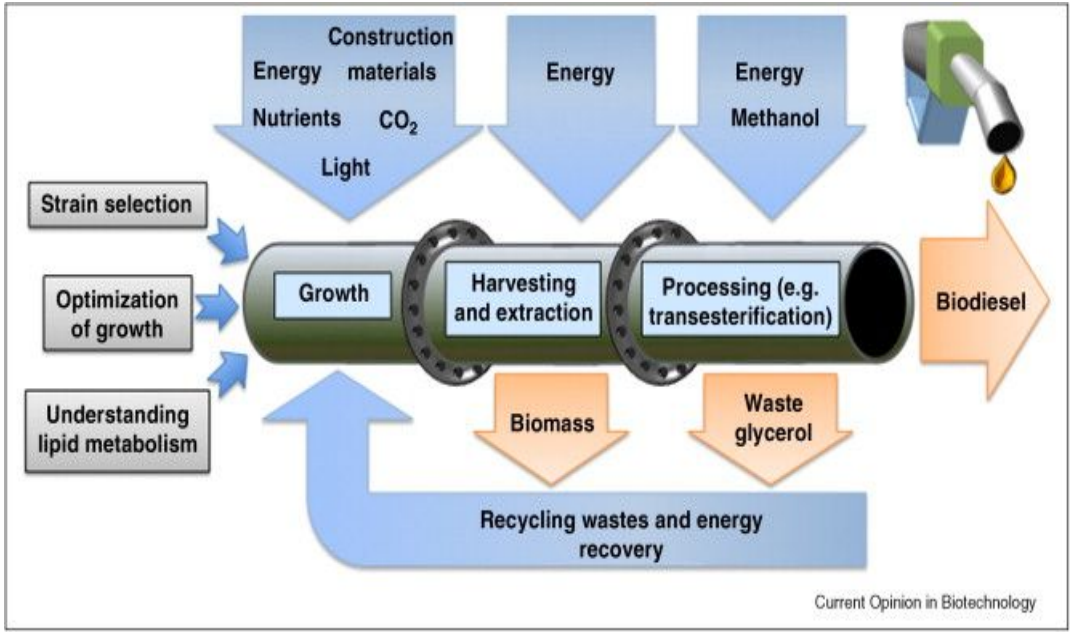
Modeling - visual ML

(not always sequencing data)



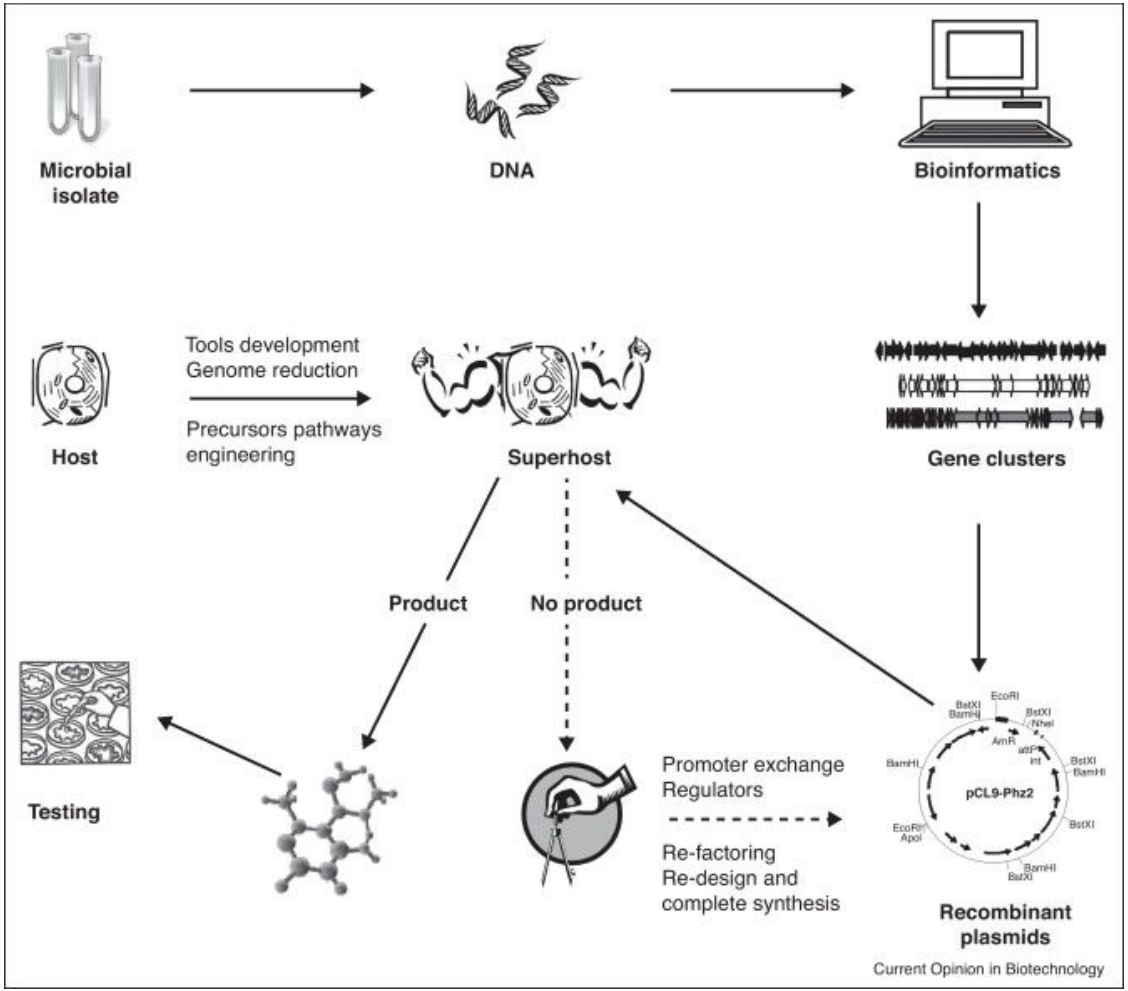
Application - biofuels

- Algae make triacylglycerides (TAGs) that can be turned into biodiesel.
- In future, we may be able to metabolically engineer algae to produce a wide variety of useful industrial polymers/chemicals such as other fuels, polymers, antibiotics, or health supplements.
- BioX is used to select the type of algae, optimize metabolic pathways



Application - bioprospecting

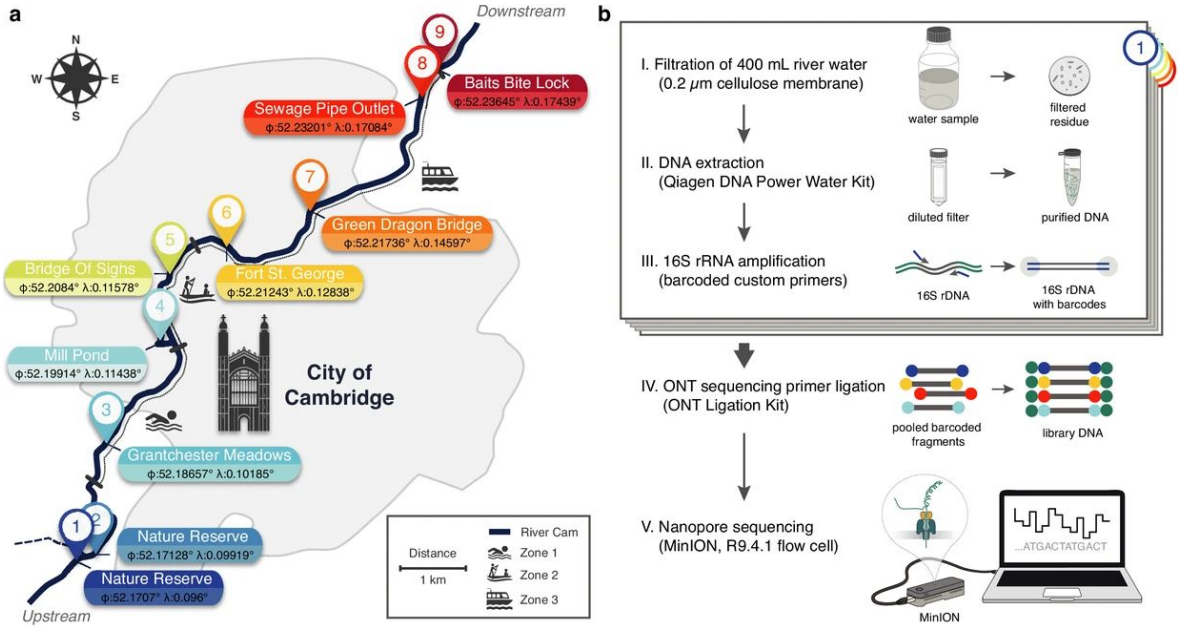
- Searching for plant-, animal-, or bacteria- derived medicinal drugs, biochemicals, and other commercially valuable material.
  - Alexander Fleming’s discovery of penicillin from *Penicillium* molds
  - Squalene, the adjuvant, is harvested from shark livers
  - Plants - caffeine, morphine, cacao, etc.
- Finding the organism of choice is informed by its ecological role, but actually identifying the gene clusters that are responsible for producing valuable biologics requires heavy computation.



Zotchev et al. 2012. *Curr. Opinion Biotech.*

Application - aquatic health

- By shotgun/targeted sequencing genetic material gathered from water sources - reclaim, runoff, lakes, oceans.
- Scientists can assess biodiversity of the sample and general aquatic health or pathogenic potential (Washoe county).



Urban et al. 2021. *eLife*