Quote

• “It may seem odd that scientists in the Internet age spend years on a line of research, even bet their careers on it, without having first determined that their mountain had not already been climbed.”


Overview

• Presentation
• Bioinformatics
  – NCBI
• Practice problems

Presentation

• Question: Which side of the class?
  – Left or Right

• Section 01 – Tue 11:00 – 306 DeBartolo
  – Graham Lappin
• Section 02 - Tue 12:30 – 331 DeBartolo
  – Graham Lappin
• Section 03 - Thu 11:00 – 306 DeBartolo
  – Graham Lappin
• Section 04 – Thu 12:30 – 228 DeBartolo
  – Graham Lappin
Presentation

• Presenters
  – Prior to presentation
    • Before 8am the day of your presentation
      – Email your bibliography
      – Email your powerpoint presentation
    – Use the clicker
    – ~20% faster when live than in practice
    – No need to say everything written on slides
    – Use a pointer
    – Talk to / Look at the audience
    – Why is the subject interesting
    – Be prepared to answer questions

• Audience
  – Become familiar with criteria on Peer Evaluation forms
    • Forms will be provided in class
  – Read abstracts prior to class
    • Identify three things you expect to hear
    • Current edition on the web
  – Come to class on time
  – Some sessions may go long
  – Ask questions
    • Your questions are easier
    • Polite & expected behavior

Informatics

• Informatics is a combination of
  – Subject expertise
  – Computing systems
  – Knowing how to find information

Types of Informatics

• Astroinformatics
• Bioinformatics
• Cheminformatics
• Ecoinformatics
• Geoinformatics
• Health Informatics

• Medical Informatics
• Neuroinformatics
• Nursing Informatics
• Pharmacy Informatics
• Quantum Informatics

Bioinformatics

• Bioinformatics includes
  – the development of methods to search databases quickly
  – analyzing DNA and protein sequence information and protein structures, and
  – predicting protein sequence and structure from DNA sequence data.

Bioinformatics Databases

- 1,436 databases

Major Databases

- NCBI - National Center for Biotechnology Information
- EMBL - European Molecular Biology Laboratory
- DDBJ - DNA Data Bank of Japan
- PDB - Protein Data Bank
- Swiss-Pro

NCBI

- National Center for Biotechnology Information
- Created in 1988 as a national resource for molecular biology information
- Part of the National Library of Medicine
- 4.4 million RefSeq DNA & RNA sequences from 10,700+ organisms

- NCBI has ~40 databases
- Each database has a search engine
- Unified search engine for all NCBI databases
“Peer Review”

- Journal articles
  - Most articles have been vetted by experts
  - Not magazine articles, encyclopedia articles, newspaper articles
- Patents
  - All issued patents vetted by a patent examiner
- What about data?

Data “peer review”

- Archival - non-peer reviewed
- Curated - peer reviewed
  - RefSeq data
  - Curated records begin with
    - NM - Nucleotide Database
    - NP - Protein Database

Conclusion

- Presentation details
- Bioinformatics
  - NCBI
    - Archival vs Curated data
- Practice
  - Purpose:
    - Intro to what can be done using NCBI databases