

Computing and Genomics

- ▶ Computing Power and Genomic Sequencing
- ▶ Sequencing Human Genome
 - Around 750 Megabytes
 - Human Genome Project took 8 years (1995–2003)
 - Most commercial companies take 2 days now
- ▶ Commercialization
 - Personal sequencing services
 - Personal sequencers
- ▶ But, analyzing takes more computing power!
 - Supercomputers in bioinformatics
 - High cost of infrastructure
 - Traditional supercomputer vs. distributed systems (farms)

Growing Role of Software

- ▶ Software issues in Bioinformatics Infrastructure
 - “Cloud” and distributed computing systems
 - Queueing
 - Data distribution
 - Networking
 - Logging and hardware/software failure
- ▶ Software in the Future Bioinformatics
 - In-house software vs. third-party software
 - Licensing the in-house software
 - Security
 - Genomics data are beginning to be stored in web servers

Reprogenetics

- ▶ Preimplantation Genetic Diagnosis (PGD)
 - high rate of success in analyzing and determining the genetic content of gametes
- ▶ "Designer Babies"
 - can currently identify gender chromosomes and a dozen or so X-linked genetic disorders (Fragile X Syndrome, Duchenne muscular dystrophy)
- ▶ Ethical Issues

Privacy: A traditional take

- ▶ Medical records pass through many hands within an organization
- ▶ WHO recommends:
 - Show-and-tell information to be shared with patients
 - No details over shared phone, fax, email, or by postcard
 - Not to store data on shared computers
 - Restricting info to patient and those granted access
 - Data handling by select individuals only



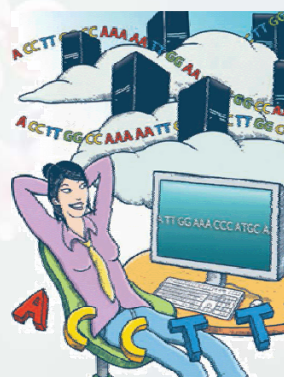
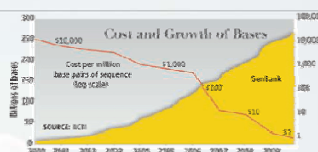
Privacy: A more modern view

- ▶ 23andMe removes middle-man i.e. the clinic
- ▶ Data is available via the web
- ▶ Data is protected by a simple email and password
- ▶ Susceptibility to social engineering attacks and hacking
- ▶ 23andMe maintains that protecting personal information is a shared responsibility
 - users must safeguard passwords, secret questions



Privacy: Behind the scenes

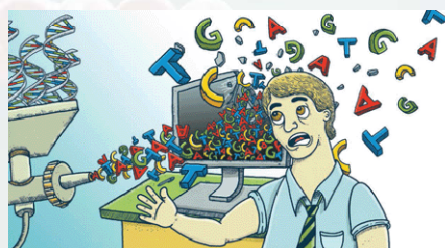
- ▶ Costs of sequencing are decreasing
- ▶ Cloud providers increasingly used for computation and data storage
- ▶ Cloud providers reduce cost and offer centralized repository of public genetic archives
- ▶ Shared data storage and virtual systems are less secure
- ▶ Security not a top priority for cloud providers



GINA

- ▶ Genetic Information Nondiscrimination Act
- ▶ Prohibits:
 - Discrimination in health insurance and employment decisions
 - Looking at genetic information before enrollment
 - Requiring genetic testing
 - Changing cost or coverage after testing
- ▶ Excludes:
 - Life and disability insurance

Conclusion



For further information, visit:
<http://stanford.edu/~bocon/cs181/>