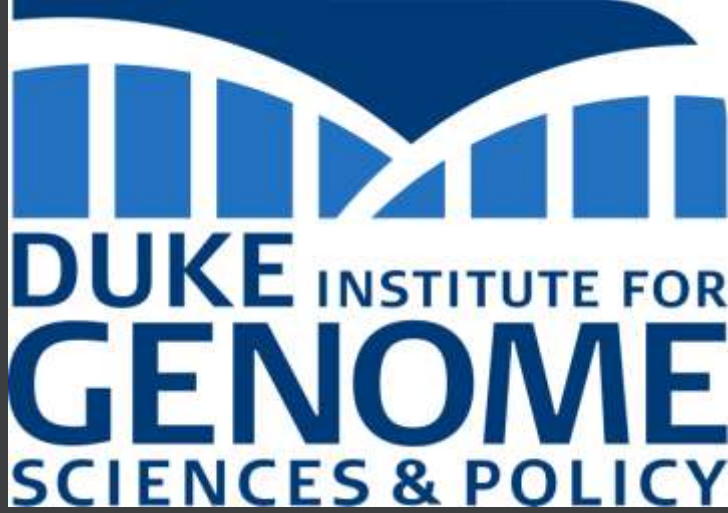


The Genome and the Internet: Growing Up Together

Robert Cook-Deegan, MD,

OECD/OCDE 12 November 2012

Who really did the work?

Michael McGeary, PhD
Subhashini Chandrasekharan, PhD
Ilse Wiechers, MD, MPP
Noah Perin, MPP + MBA
Sapna Kumar, JD
Jennifer Pohlhaus, PhD
Colin Crossman, JD

Alessandra Colaianni (U)
Joe Fore (U)
Whitney Laemmli (U)
Anupama Kotha (U)
Nancy Wang (U)
Suparna Salil (U)
Daidree Tofano (U)
Phebe Ko, BA
Molly Nicholson, BA
Cindy Wang, MPP
Matt DeCamp, MD/PhD (Philosophy)
Britt Rusert, PhD cand. (English)
Stacy Lavin, PhD cand. (English)
Marie Hicks, PhD cand. (History)
Marjorie Gurganus, JD

Common features

- Government-driven “first uses”
  - Census, military, space
  - Many inventions in Bell Labs or firms, but gestation in academe
  - Medical market for biotech
- Intricate government-industry-academic mutualism
- Big projects that push the envelope
- Selling a vision of how technology can rock our world

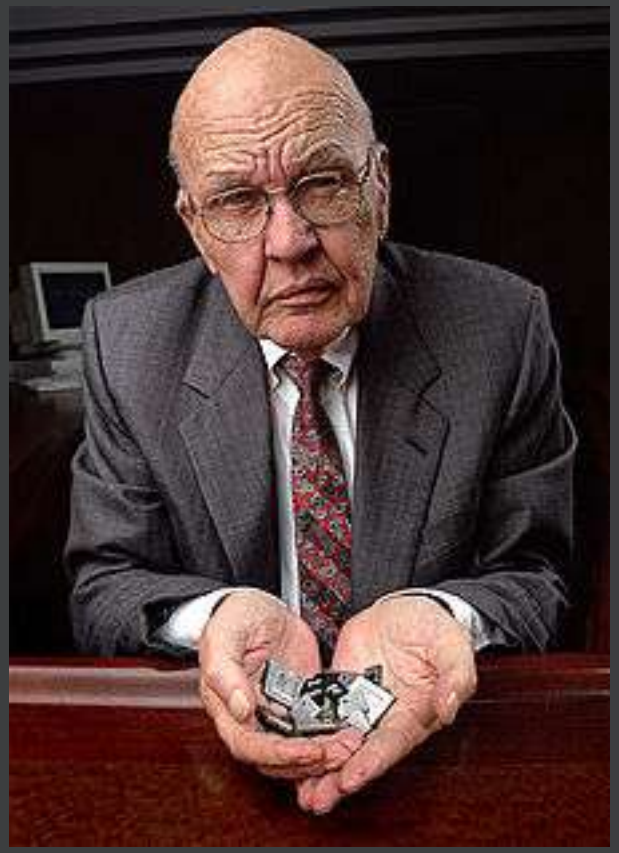

Microsoft

The transistor

John Bardeen, Walter Brattain, and William Shockley 1948

The Chip

Robert Noyce, Intel



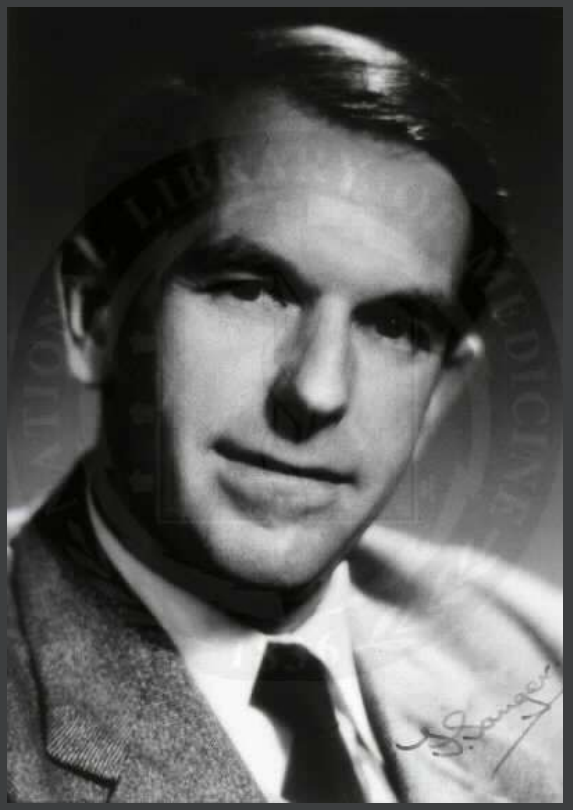
Jack Kilby, Texas Instruments

The World Wide Web

Two Unlikely Revolutionaries

- JCR Licklider

- Frederick Sanger



Twin Revolutions

Computing

- ENIAC
- Census
- Whirlwind/MAC
- "Human-computer Symbiosis"
- Personal computer
- Networking
- Mobile networking

Genetics

- Mendel's elements
- Population genetics
- Genes are DNA
- Recombinant DNA
- Sequencing DNA
- Automation
- Scale-up

Shared features: Rapidly accelerating power, drop in price.

Ubiquity. Pervasiveness. Disruption. Unanticipated uses. Network effects

Computing *versus* Biotech

- Price competition
- Month-years cycle
- Sales to consumers
- Prices low in USA
- Patents somewhat useful sometimes
- Keep costs down

- R&D competition
- Years-decades cycle
- Influence physicians
- Prices high in USA
- Patents essential
- FDA approval
- Keep prices up

Science-Dependent Innovation

- Academic research base
- Public funding
- Capital for technical opportunities

Human Genome Project

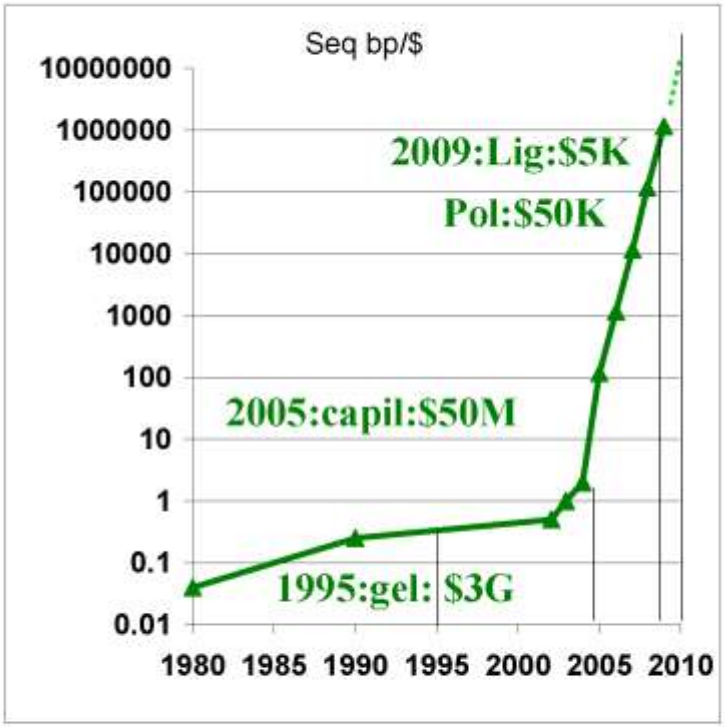
Key enabling technologies

- DNA detection and handling
- Recombinant DNA
- DNA sequencing
- Computing
- Networks and data-sharing

George Church's graph of DNA Sequencing Costs

4 logs in 4 years

(Moore's law) 1.5x/yr for electronics vs 10x/yr for DNA Sequencing



Costs of sequencing whole genome

1986	**Est. $3 billion**
2000-2004	$300M-$1 billion (Celera-HGP)
2007	$3-5 million (454 Watson's Jimome; Craig's Venterome)
2008	$350,000-$1 million (Knome, Yoruba and Chinese)
2009-10	$50,000 (Illumina); $68,000-$99,500 (Knome) $5,000-$20,000 (Complete Genomics)
?	$100-1000 (nanopore or Pacific Biosciences single-molecule)?

Pertinent price points

- Cost of full-sequence single-condition testing ($3,000-7,000)
- Cost of mutation testing ($500 – 1,000)
- Cost of newborn screening ($100)

As price drops, full-genome sequencing will move closer to universal screen, and other uses will proliferate

The uses we don’t think of

- Ancestry and genealogy
- Paternity
- Forensics
- Who you’re related to
- Your pet
- Your microbiome
- “It’s just cool to know”

Network Effects

- “To many eyes, every bug is shallow”
- I learn from your data
- Just as important, you learn from my data
- Value of information grows with network size (squared)
- Higher speed of aggregating information
- Low coordination cost

~Think wikipedia, word processing files

Data Costs:
Genome sequence data that cost $1M to produce

1990

- $10 worth of disk space to store
- 1 second to transmit over a network

2010

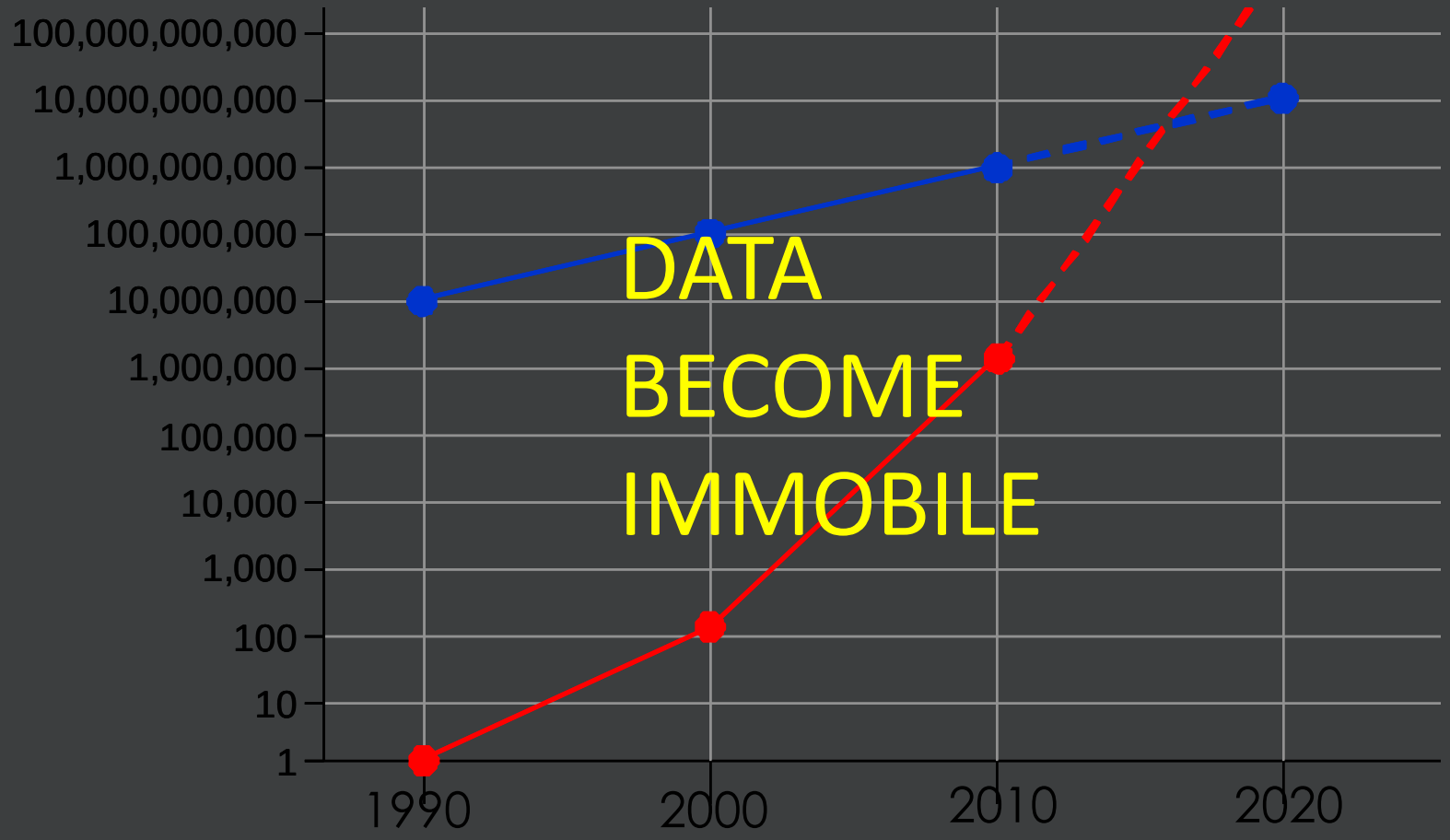
- cost $200,000 - $400,000 to store
- several hours to move over a network

2020

- Several million dollars to store
- Years to transmit over a network

Estimates from Leonard “Kris” Kristalka, University of Kansas

Data costs: Genomics



Blue = 10 X/decade improvemens in network bandwidth, bits/second

Red = 100 X/decade improvements in sequencing efficiency, bases/$

Graph and estimates courtesy of Leonard “Kris” Kristalka, University of Kansas

Policy dilemmas

- Genomic data are not just about you, but also your family
- They feel like they should be “private”
- We don’t know if sharing can be harmful
- We don’t have the infrastructure or procedures for you to use your data even if we could get them to you
- We don’t have interpretive services