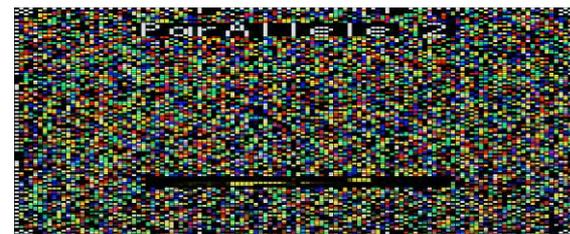




Presidential Commission for
the Study of Bioethical Issues

Genome Data: Research to Clinic to Consumer

**Richard Gibbs, Baylor College of Medicine
Human Genome Sequencing Center, Houston Tx.**



COI:

Co-investment with Life Technology

Co-Founder SeqWright

A lot has changed since 1986...

-Personal computers, internet,
cell phones etc...

Genetics:

- Human Genome Project
- Personal genomes
- Third –party DNA providers
- ‘Genomes’ in our lives

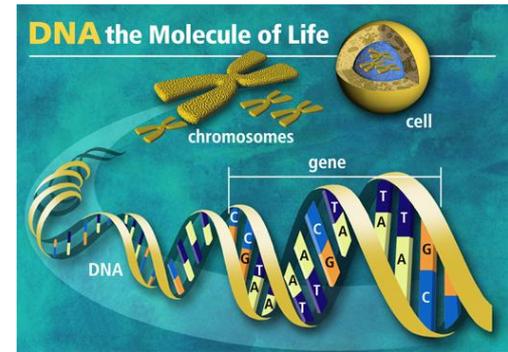
How will most people interact with their genome data?



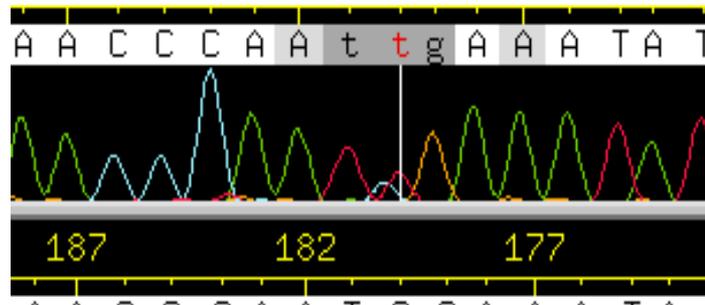
Armchair futurism:

History – The ROAD MAP – 50,000 ft view:

(i) Sequence Genome:

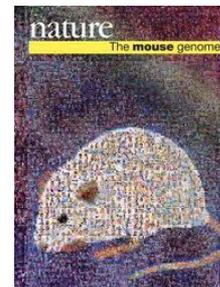


(ii) Large Scale Genetic Mapping:

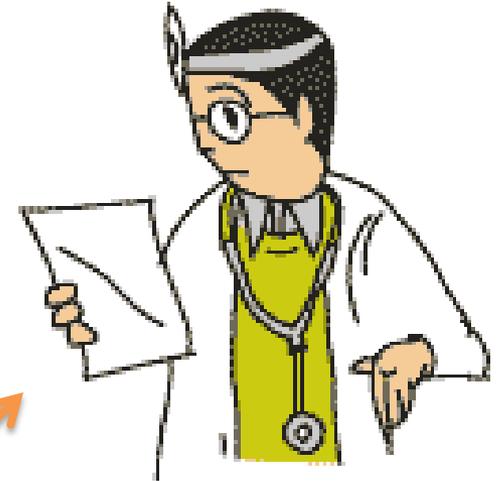


(iii) Sequence to discover alleles

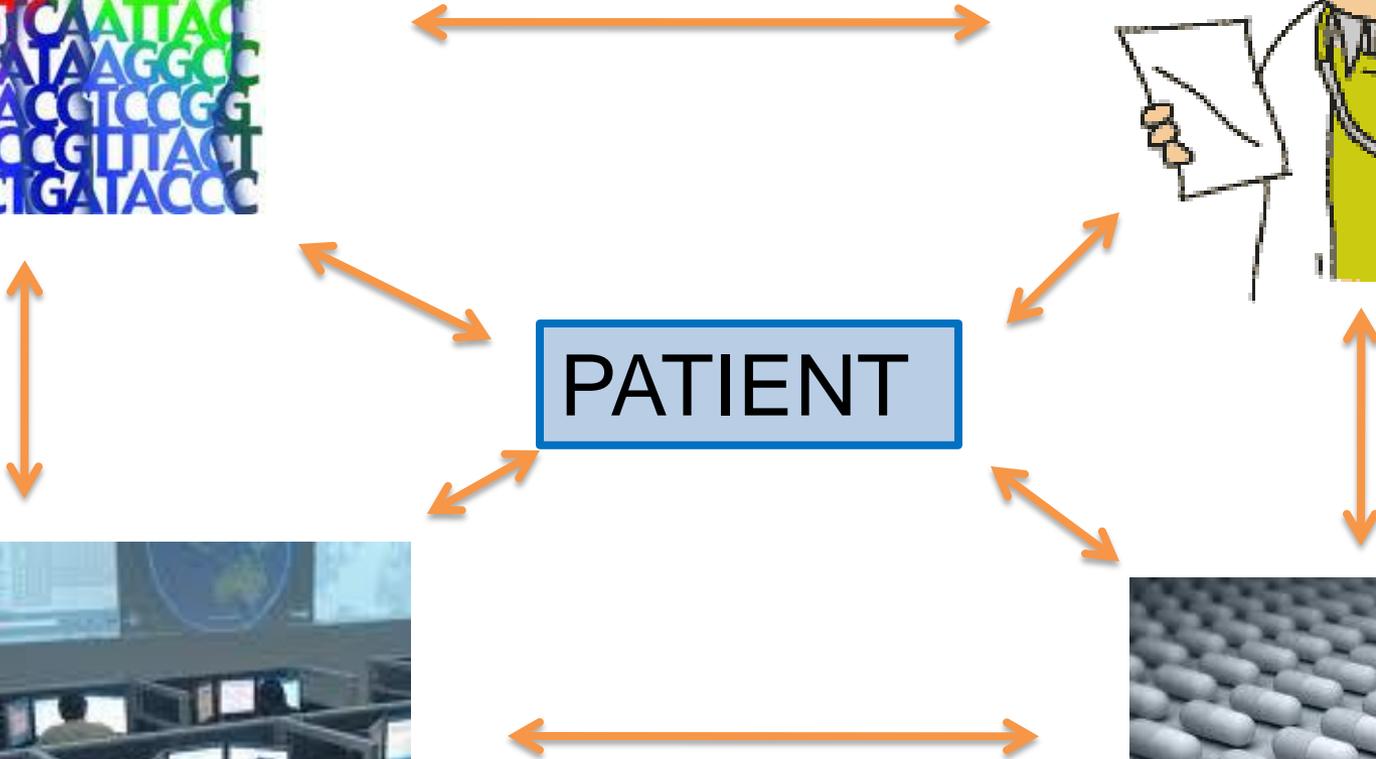
(iv) Function/Comparative studies



Genomic Medicine, 20??:



PATIENT



History:

The Bermuda Accord:

~ 1996

(then Fort Lauderdale;
Toronto)



- Free and unrestricted release of human genome data to all.

The basis of:

- Ubiquitous genome research
- Rapid post-genome development
- Healthier patenting
- Follow on projects:

The HapMap project; Mammalian gene collection; Most GWAS;
The Cancer Genome ATLAS

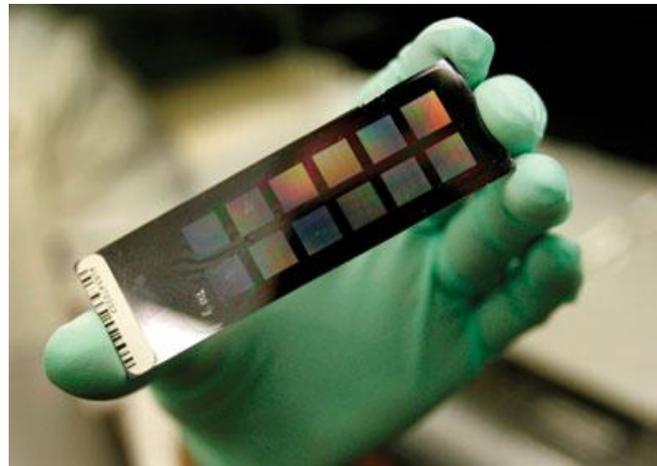
TENSION:

The 'benefits and risk' of access to genetic data:

- *Disease risk*
- *Diagnosis*
- *Treatments*
- *Ancestors*
- *Prognostics*

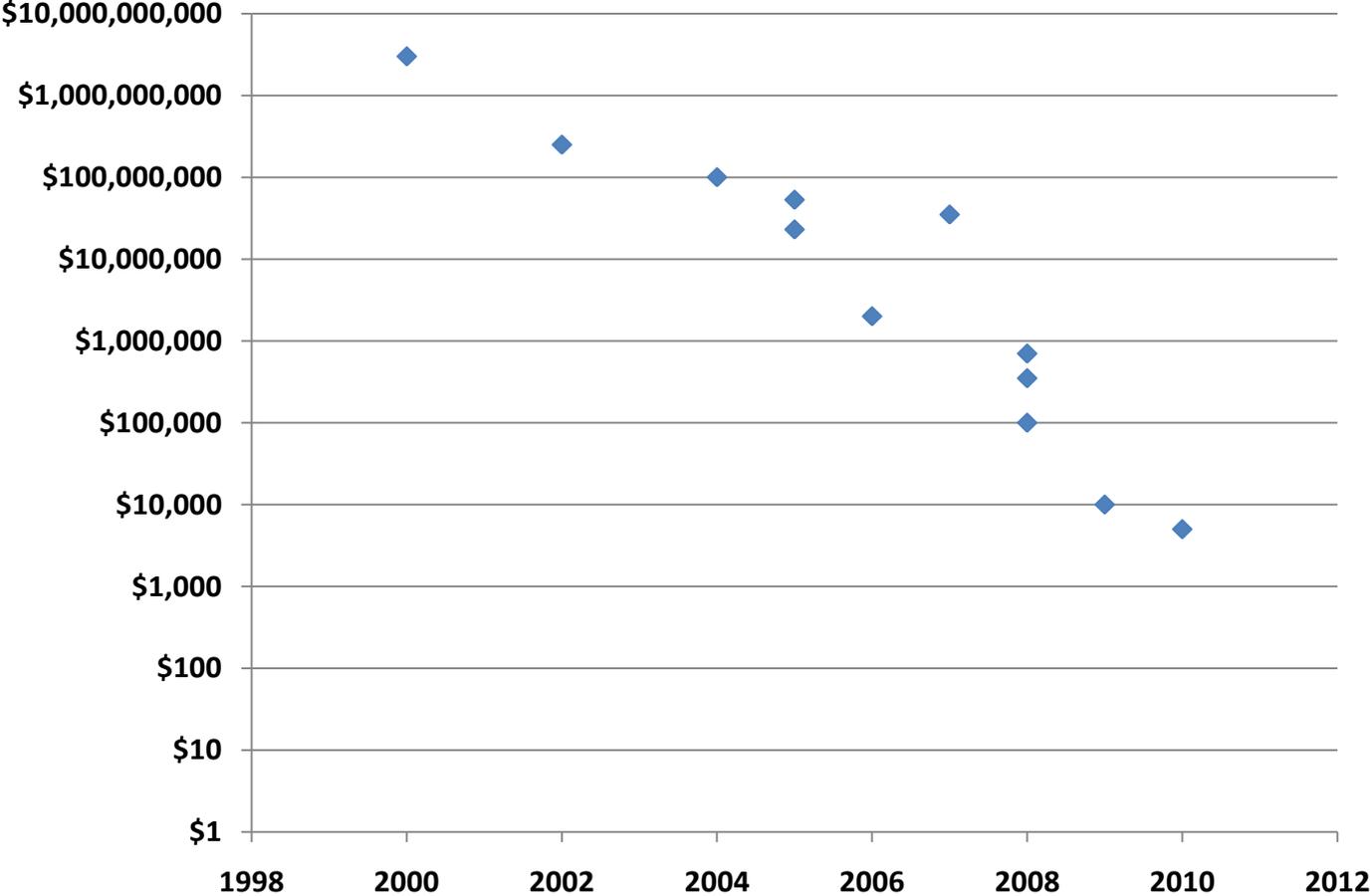


- *Insurance*
- *Employment*
- *Social stigma*
- *Personal angst*

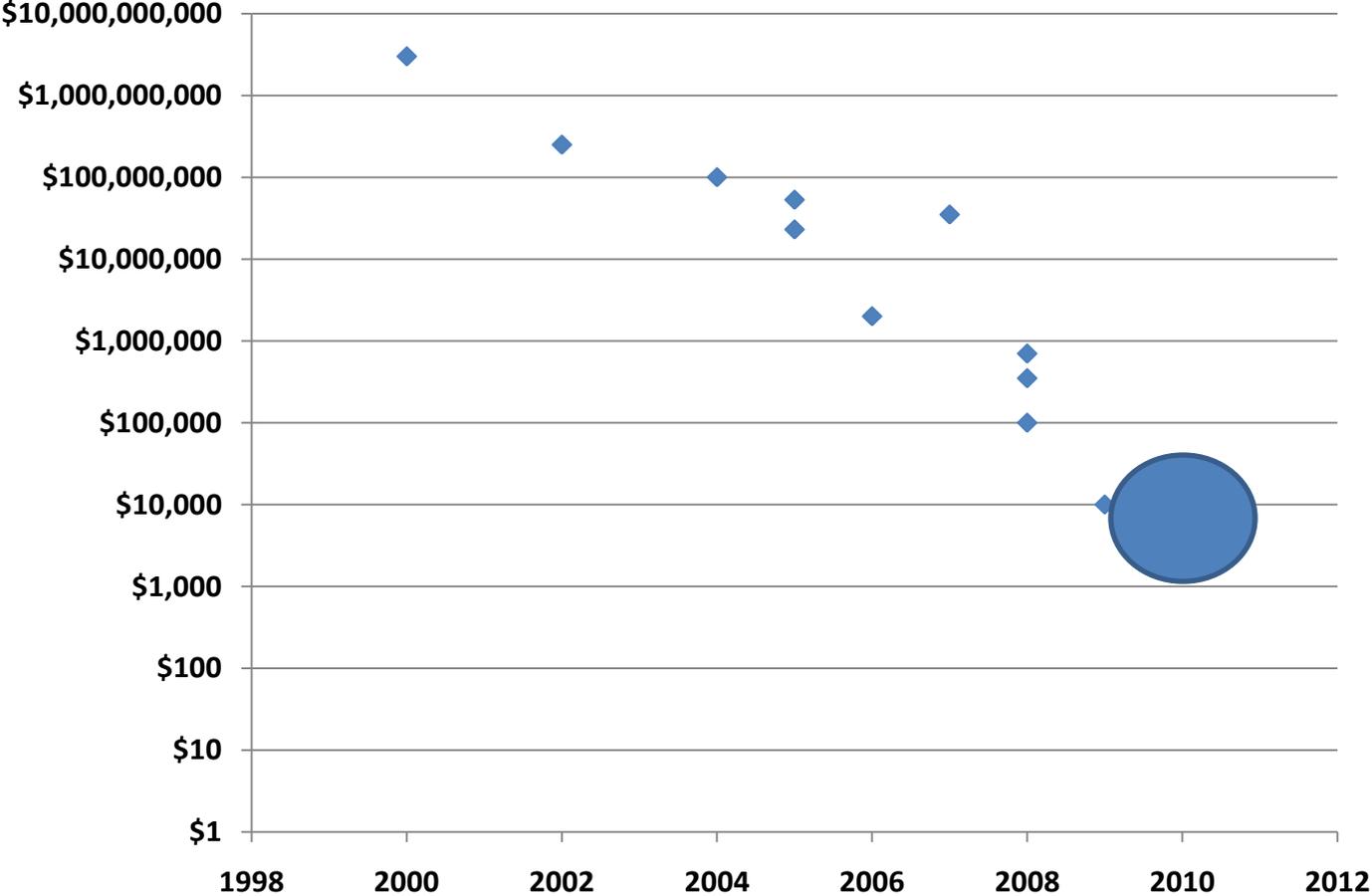


Meanwhile
Technology
Exploded.....

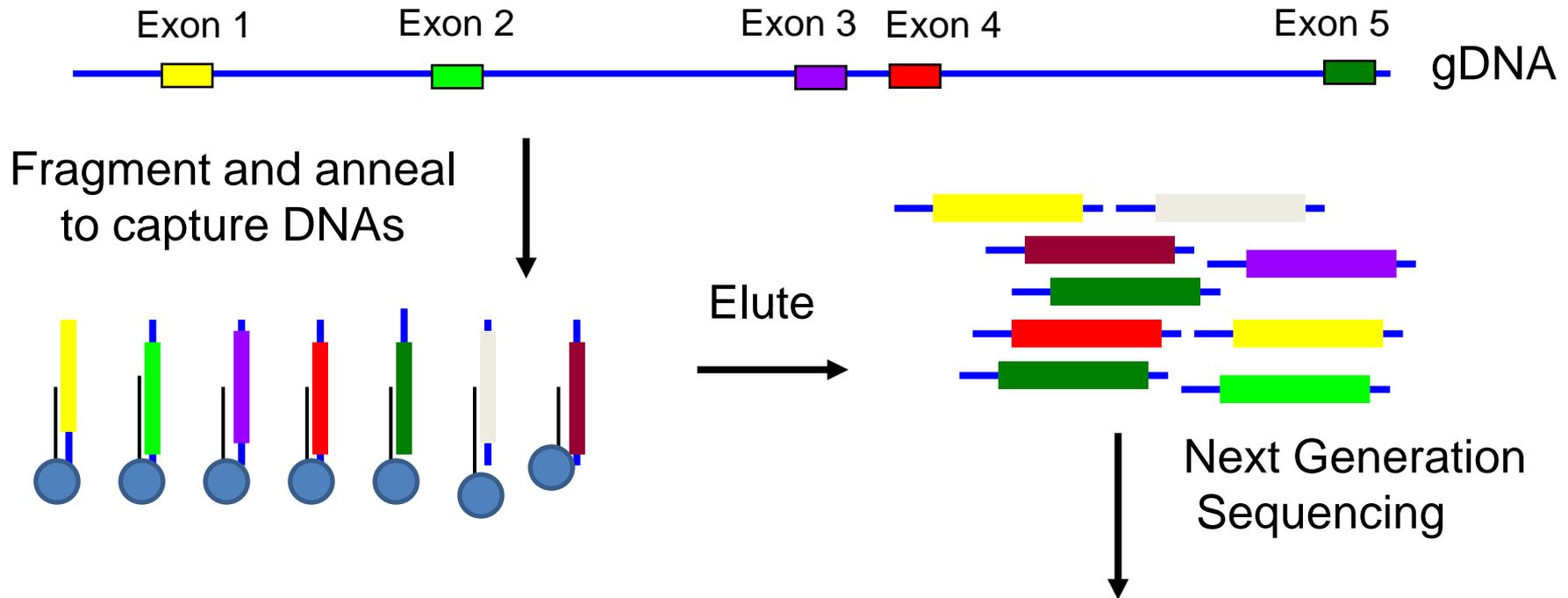
The Cost of Sequencing a Human:



The Cost of Sequencing a Human:



Shortcut Genomes: Baylor-Nimblegen Gene Capture Sequencing



Direct selection of human genomic loci by microarray hybridization

Thomas J Albert¹, Michael N Molla¹,
Donna M Muzny², Lynne Nazareth², David Wheeler²,
Xingzhi Song², Todd A Richmond¹, Chris M Middle¹,
Matthew J Rodesch¹, Charles J Packard¹,
George M Weinstock² & Richard A Gibbs²

We applied high-density microarrays to the enrichment of specific sequences from the human genome for high-throughput sequencing. After capture of 6.726 approximately 500-base

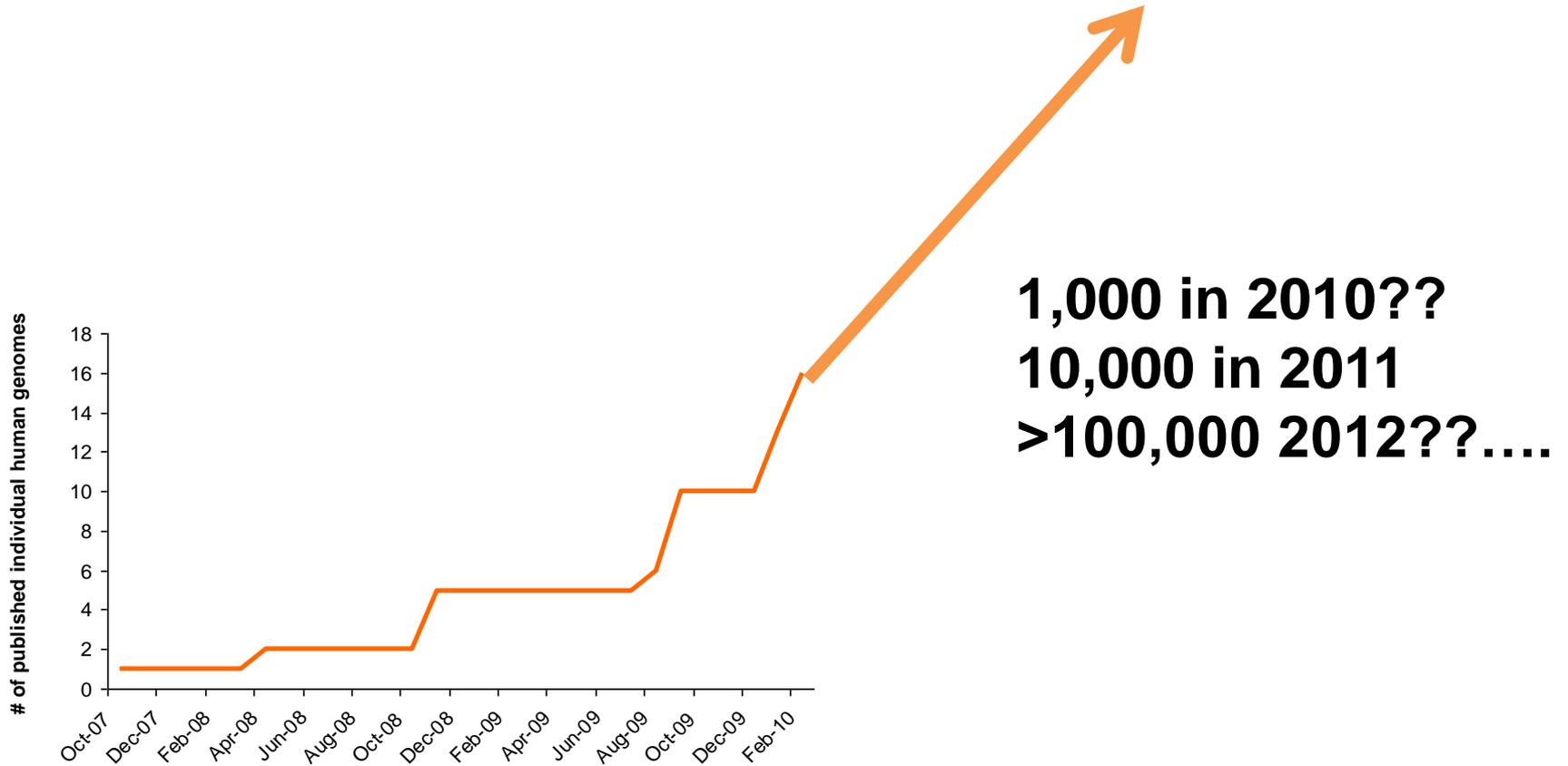
corresponding to represent genomic in addition to the presence of repeat the final enrichment

Here we report tide microarrays (genome segments, long segments, on design aimed to ca (Supplementary 1 of total sequence genome, and a set kb, 500 kb, 1 Mb, gene locus. Each oligonucleotide pr 1 and 10 base an

RAPID SEQUENCING OF GENE PORTIONS ONLY

1/5 cost of whole genome sequencing

Human Personal Genomes



PERSONAL GENOMES INFORM THE SFS:



James D. Watson
Nature (2008) 452: 872-876.



Individual variation



Desmond Tutu
(Khoisan & Bantu genomes)
Nature (2010) 463:943-947.



Population variation



Jim Lupski
New Engl J Med (2010) 362:1181-1191.



Identification of medically actionable, disease-causing variants



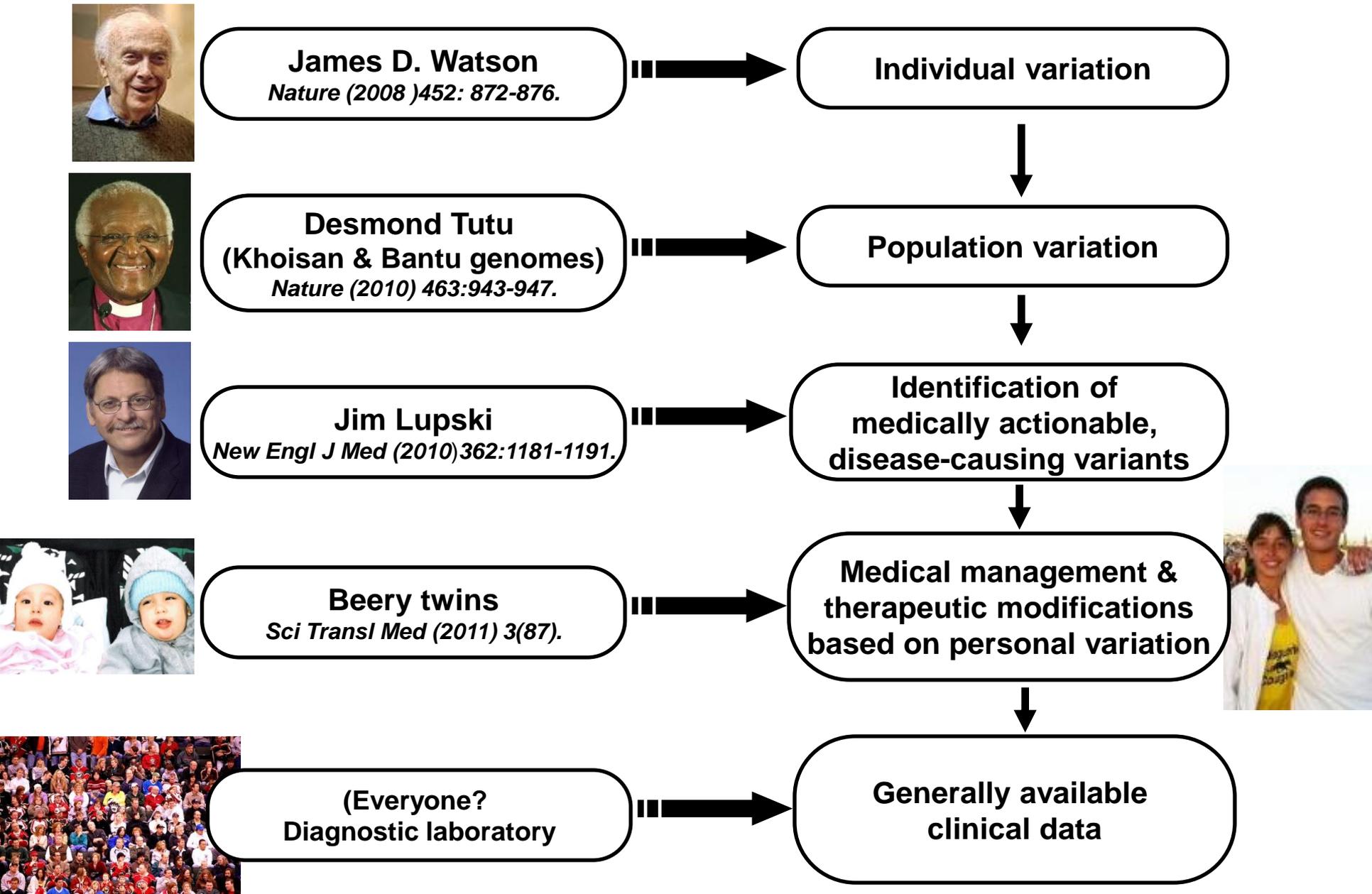
Beery twins
Sci Transl Med (2011) 3(87).



Medical management & therapeutic modifications based on personal variation



PERSONAL GENOMES INFORM THE SFS:



The first 'whole genome' consultation: The J. D. Watson personal genome



- ~ 25 Mb of DNA missing from reference, in JDW
- Sequence reads reveal CNVs
- 16% of Watson SNPs are novel
- 15% of Venter SNPs are novel
- ~10,500 ns variants
- ~1,500 novel ns variants !!
- Overall...more *previously novel* functional variants than expected

		Total variation	Known	Novel
Watson	Raw	14,829,087	3,283,273	11,545,814
	1	4,427,488	2,815,322	1,612,166
	2	3,971,513	2,752,991	1,218,522
	3	3,325,725	2,704,029	621,696
Venter	4	3,470,669	2,726,935	743,734

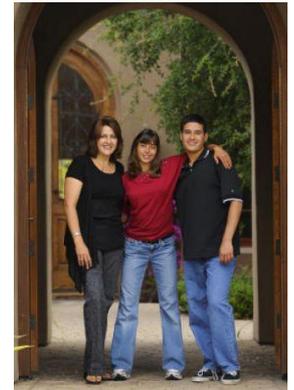
WHO WILL INTERPRET AND DELIVER THESE DATA?

Key societies leading the way, but task is greater.



Mrs. Beery is an expert!

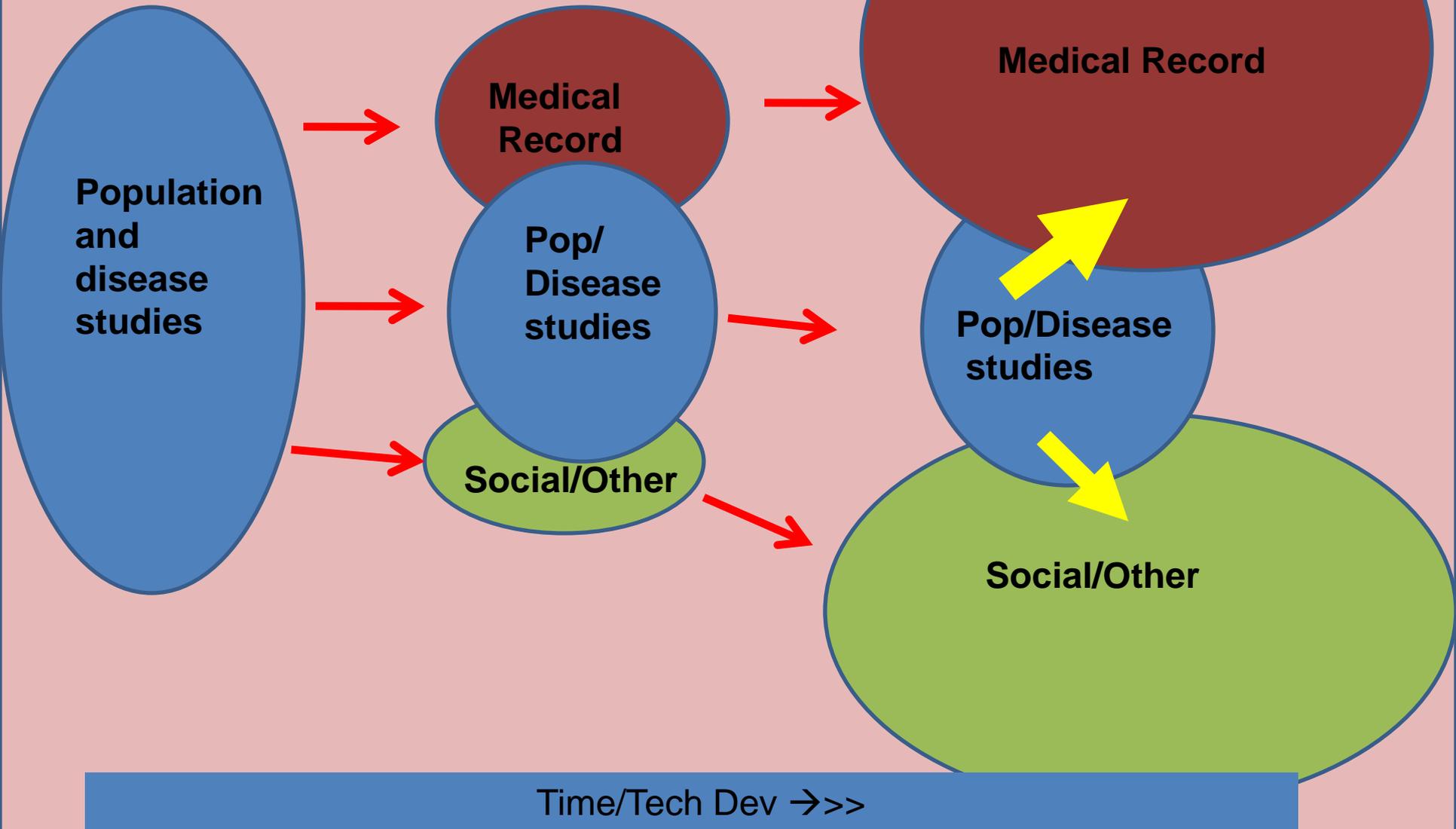
- **Informed**
- **Educated**
- **Connected**
- **Unusual!**



What about the 99% - who don't have genetics education...

- **How will they ...**
- **Have first contact with genetic data?**
- **Obtain genetic data?**
- **Interpret genetic data?**

Armchair Futurist's view of Genomic and Genetic Data



Complicated clinical reports -> simplified

REPORT THREE TIERS: (includes 'incidental findings')

1. Known, actionable
e.g. known alleles
2. Uncertain, not clear action
e.g. known locus, new allele
3. Unknown
e.g. *de novo*, new gene

1.1	1.1 DM_Focused	0		
2.1	2.1 DM_Focused	1		
3.11	3.11 UV_Focused	0		
7.11	7.11 UV_Focused	0		
1.3	1.3 Actionable Mutation_Focused	0		
2.3	2.3 Actionable Mutation_Focused	0		
1.4	1.4 Carrier Mutation_Focused	1		
2.4	2.4 Carrier Mutation_Focused	2		
8	Pharmacogenetic Alleles	0	4	
1.2	1.2 DM_Expanded	0		
2.2	2.2 DM_Expanded	0		
3.21	3.21 UV_Expanded	9		
7.21	7.21 UV_Expanded	5		
1.5	1.5 Mutation No Dis Asso._Expanded	6		
2.5	2.5 Mutation No Dis Asso._Expanded	4	24	
3.22	3.22 UV Unreported	14		
7.22	7.22 UV Reported	5		
5	5 Benign Reported	6		
0	FP	24		
-1	-1 FP Sanger	0		
U	No treatment	0		
P	Pending Sanger	13		
3.3	3.3 UV of Unknown Genes	0	62	
			90	
	variants in genes with risk alleles		808	
	Total		898	

SUMMARY:

- **Data acquisition relatively inexpensive**
- **Free access to genetic data has many positive benefits**
- **Tension with personal privacy**
- **Tide is rising – already seeing medical application**
- **Predict non-traditional paths will be utilized more often in the future**