GENETICS, GENOMICS AND EPILEPSY
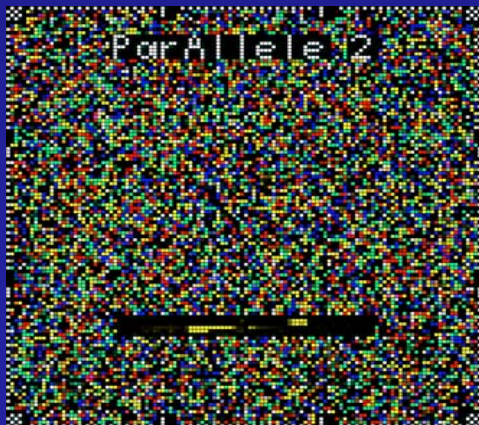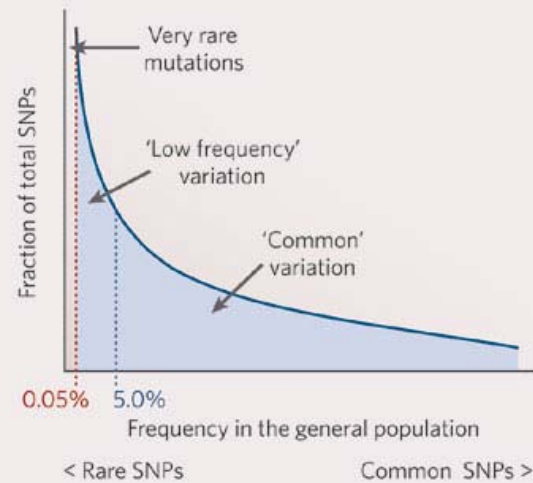
An alternative and accessible version of this presentation is available at 9:30 am in the

GENETIC VARIATION IN HUMANS
Variation is measured by single nucleotide polymorphisms (SNPs).

Very rare mutations

'Low frequency' variation

'Common' variation

Fraction of total SNPs

0.05%   5.0%

Frequency in the general population

< Rare SNPs   Common SNPs >

>75 Genes Linked to Monogenic Epilepsy

SCN1A	SLC9A1	GABBR1
SCN2A1	SLC1A2	KCNJ6
SCN1B	SLC2A1	NEUROD1
KCNA1	KCC2	MECP2
KCNC2	ATP1A2	EPM2A
KCNQ2	NPY	FLN1
KCNQ3	GAD2	PPT1
KCNMA	ITPR1	ALPL
KCNMB4	CAMK2A	TRK1
CACNA1A	PLCB1	LAMR1P11
CACNB4	SYN1+2	RORA
CACNG4	SV2A	PTEN
CACNA2D2	BSN	CBP-B
CICN2	AP3D1	AMT
HCN2	DCX	UBE3a
GABRA1	DCLX1	CIT
GABRB3	OTX	CYSTB
GABRG2	EMX2	MYO5A
CHRNA4	SOX1	TSC1, 2
CHRNB2	FCN2	NHLRC1
HTR2C	UPAR	LGi1
GRIA2	ARX	Caspr2

Human

Mouse

Both

Etiology of Epilepsy

Inherited

Mendelian

Non-Mendelian

+ multiple genes (complex inheritance)

de novo

single gene (sporadic) mutations
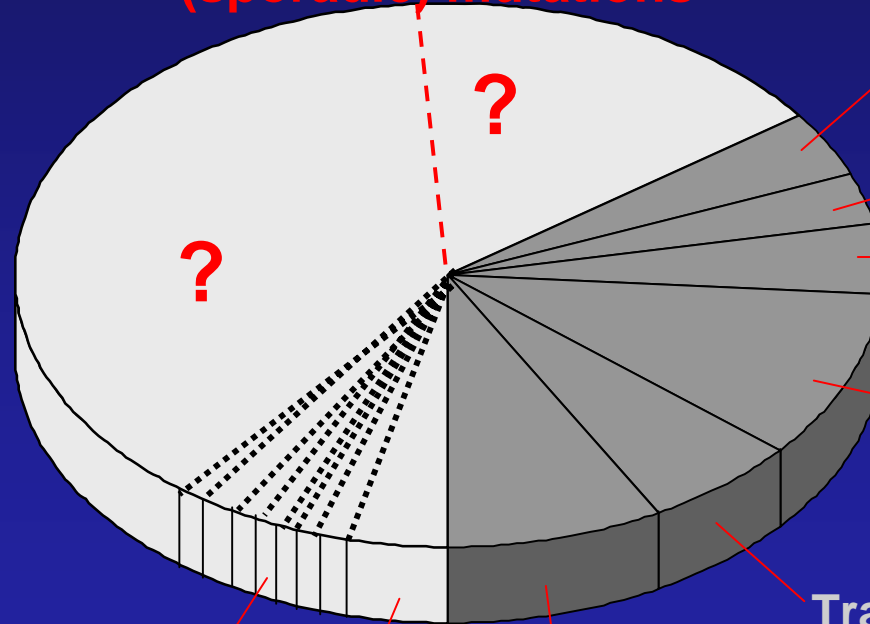
Acquired

Many Pathologies

Degenerative

Infection

Neoplastic

Vascular

Trauma

Congenital

Single genes Ion Channels

?

?

(HETEROZYGOUS) MUTATION DETECTION

FIG. 2. Multiplex amplification of the HPRT locus. All nine exons of the human HPRT gene were amplified on eight separate DNA fragments using 16 oligonucleotide primers in a single PCR
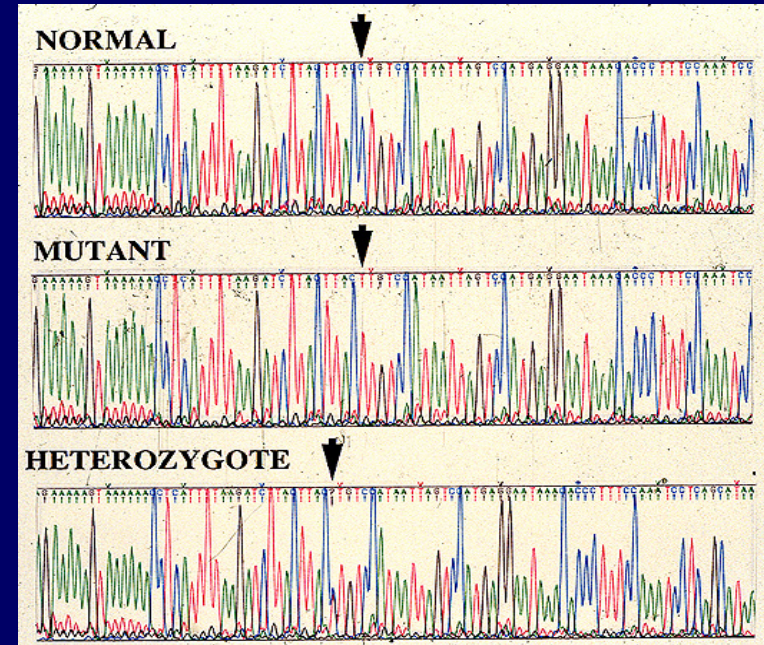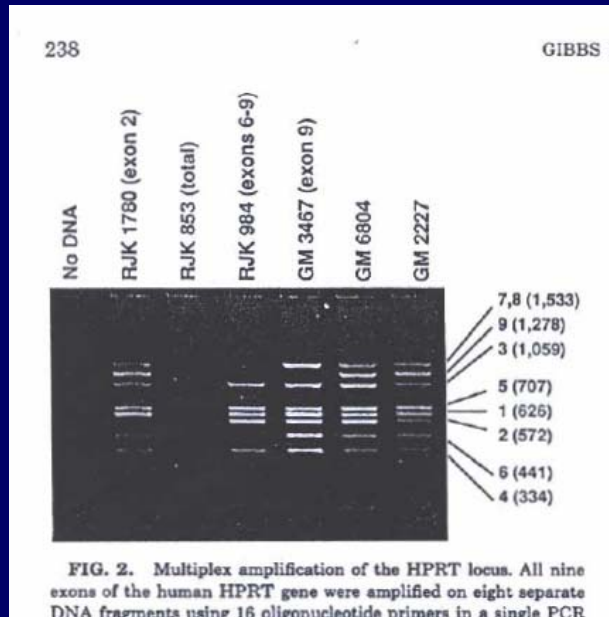
ABI 370: Universal Primer Tag PCR

Complete **Gene Ascertainment!**

THE FINISHED HGP 1991-2004: Lessons

Finishing the euchromatic sequence of the human genome

International Human Genome Sequencing Consortium*

* A list of authors and their affiliations appears in the Supplementary Information

The sequence of the human genome encodes the genetic instructions for human physiology, as well as rich information about

The DNA sequence of the human X chromosome

The DNA sequence, annotation and analysis of human chromosome 3

Donna M. Muzny[1], Steven E. Scherer[1], Rajinder Kaul[2], Jing Wang[3], Jun Yu[3], Ralf Sudbrak[4,5], Christian J. Buhay[1], Rui Chen[1], Andrew Cree[1], Yan Ding[1], Shannon Dugan-Rocha[1], Rachel Gill[1], Preethi Gunaratne[1], R. Alan Harris[1], Alicia C. Hawes[1], Judith Hernandez[1], Anne V. Hodgson[1], Jennifer Hume[1], Andrew Jackson[1], Ziad Mohid Khan[1], Christie Kovar-Smith[1], Lora R. Lewis[1], Ryan J. Lozado[1], Michael L. Metzker[1], Aleksandar Milosavljevic[1]

The finished DNA sequence of human chromosome 12

Steven E. Scherer[1], Donna M. Muzny[1], Christian J. Buhay[1], Rui Chen[1], Andrew Cree[1], Yan Ding[1], Shannon Dugan-Rocha[1], Rachel Gill[1], Preethi Gunaratne[1], R. Alan Harris[1], Alicia C. Hawes[1], Judith Hernandez[1], Anne V. Hodgson[1], Jennifer Hume[1], Andrew Jackson[1], Ziad Mohid Khan[1], Christie Kovar-Smith[1], Lora R. Lewis[1], Ryan J. Lozado[1], Michael L. Metzker[1], Aleksandar Milosavljevic[1], George R. Miner[1], Kate T. Montgomery[2], Margaret B. Morgan[1], Lynne V. Nazareth[1], Graham Scott[1], Erica Sodergren[1], Xing-Zhi Song[1], David Steffen[1], Ruth C. Lovering[3], David A. Wheeler[1], Kim C. Worley[1], Yi Yuan[1], Zhengdong Zhang[1], Charles Q. Adams[1], M. Ali Ansari-Lari[1], Mulu Ayele[1], Mary J. Brown[1], Guan Chen[1], Zhijian Chen[1], Kerstin P. Clerc-Blankenburg[1], Clay Davis[1], Oliver Delgado[1], Huyen H. Dinh[1], Heather Draper[1], Manuel L. Gonzalez-Garay[1], Paul Havlak[1], Laronda R. Jackson[1], Leni S. Jacob[1], Susan H. Kelly[1], Li Li[2], Zhangwan Li[1], Jing Liu[1], Wen Liu[1], Jing Lu[1], Manjula Maheshwari[1], Bao-Viet Nguyen[1], Geoffrey O. Okwuonu[1], Shiran Pasternak[1], Lesette M. Perez[1], Farah J. H. Plopper[1], Jireh Santibanez[1], Hua Shen[1], Paul E. Tabor[1], Daniel Verduzco[1], Lenee Waldron[1], Qiaoyan Wang[1], Gabrielle A. Williams[1], JingKun Zhang[1], Jianling Zhou[1], Baylor College of Medicine Human Genome Sequencing Center Sequence Production Team*, David Nelson[1], Raju Kucherlapati[2], George Weinstock[1] & Richard A. Gibbs[1]

Immediate Findings:

Even Less genes? (~22,000)

Segmental duplication polymorphisms

Publicly available data is important.
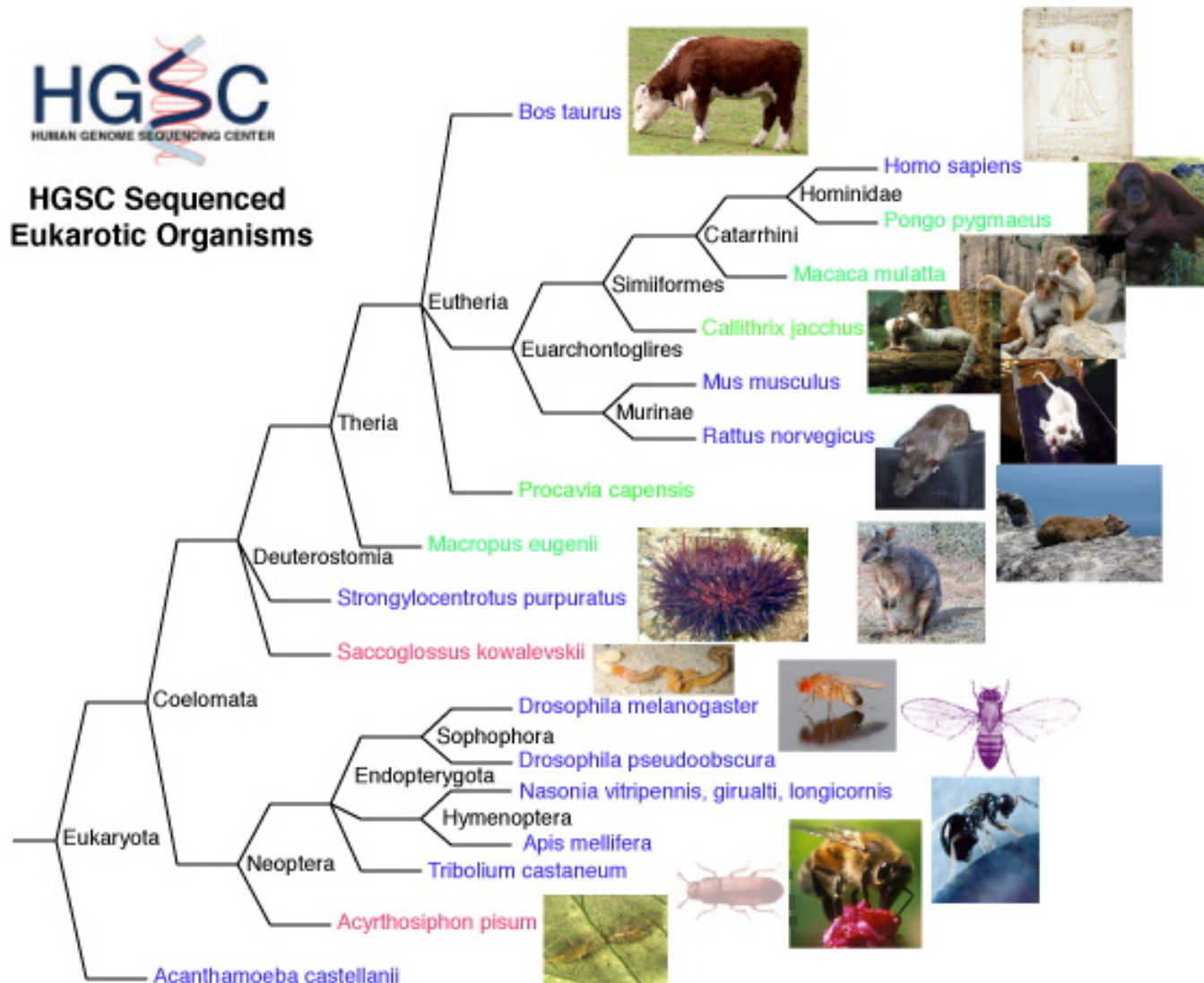
Feb '96, 97, 98
Bermuda Meetings

nature
Tetraodon to human

Bermuda

'Genomes -> Mutation Discovery

HGSC
HUMAN GENOME SEQUENCING CENTER

HGSC Sequenced Eukarotic Organisms
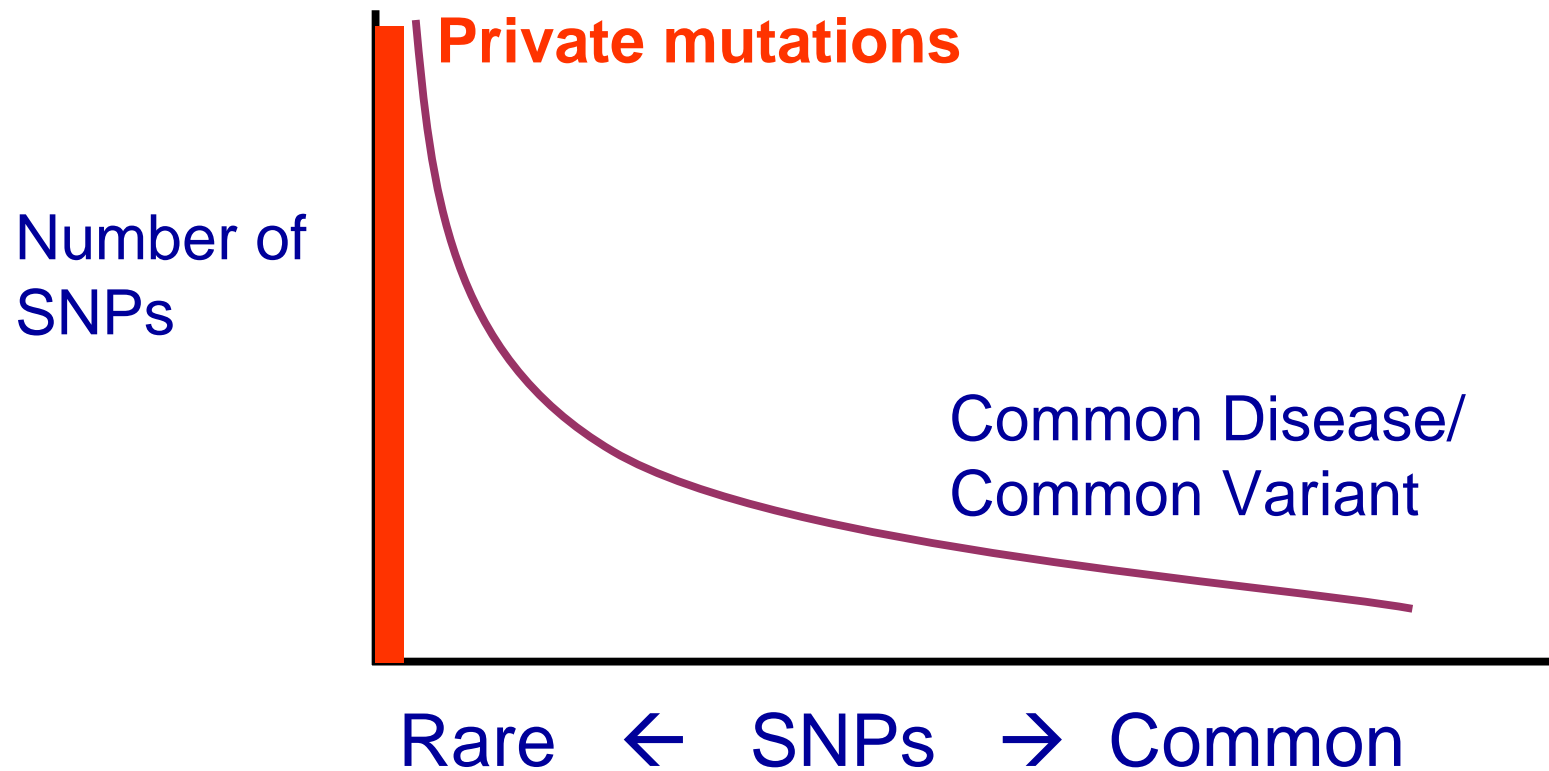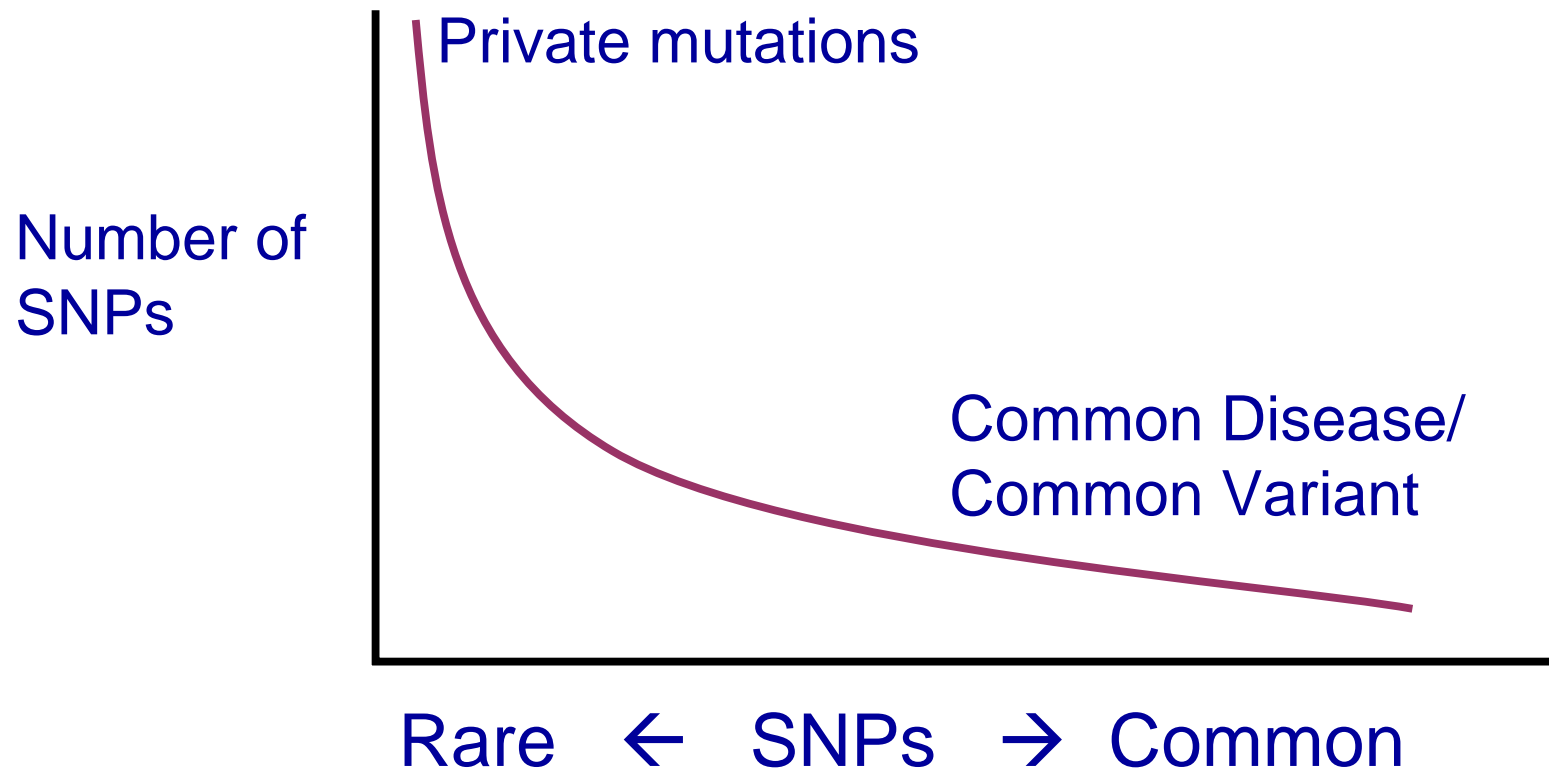
Dolphin

Armadillo

Kangaroo Rat

Gibbon

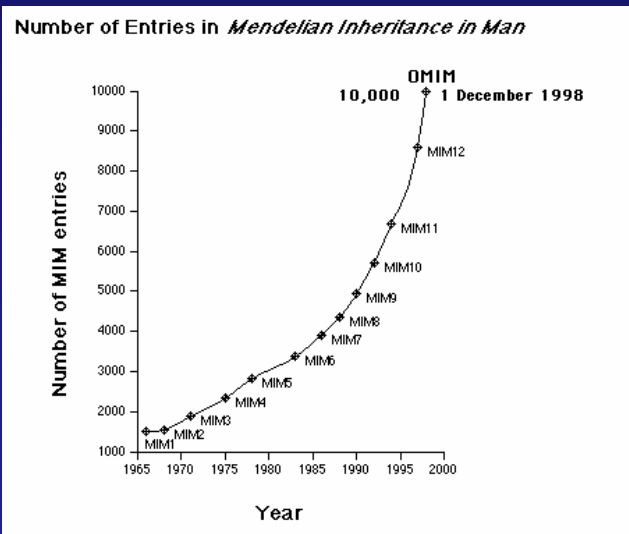**And soon ??**

Baboon

Y-Chromosomes

Overall Distribution of Variation

Number of
SNPs

Private mutations

Common Disease/
Common Variant

Rare ← SNPs → Common

OMIM - 13,000 Entries

Number of Entries in *Mendelian Inheritance in Man*

OMIM
10,000 1 December 1998

> 6,000 Mendelian Traits,
~ 1,800 solved
All possible by large scale
Re-sequencing

Focusing on Mendelian Traits is Guaranteed High Yield!!
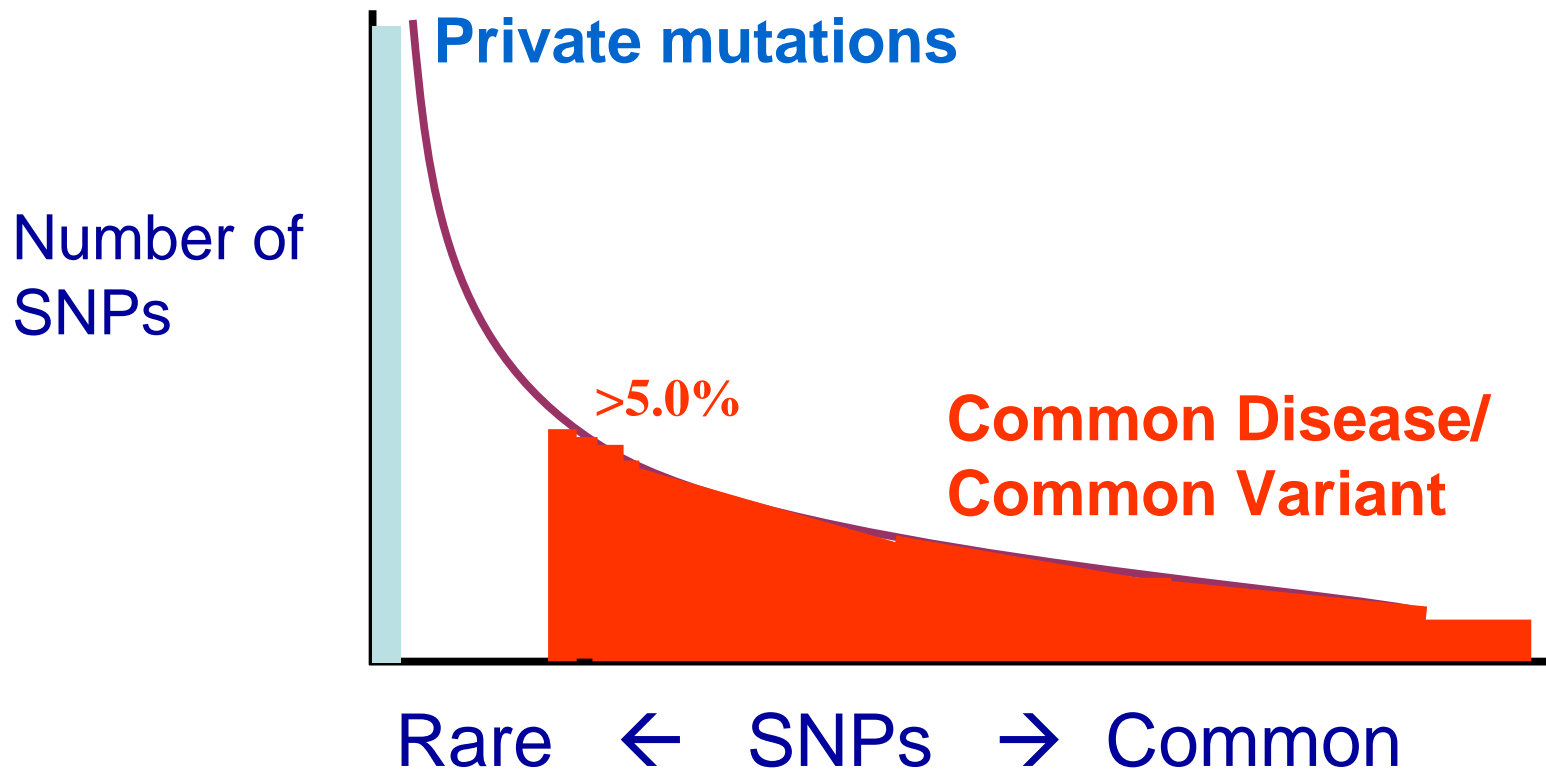
Common Disease/Common Variant
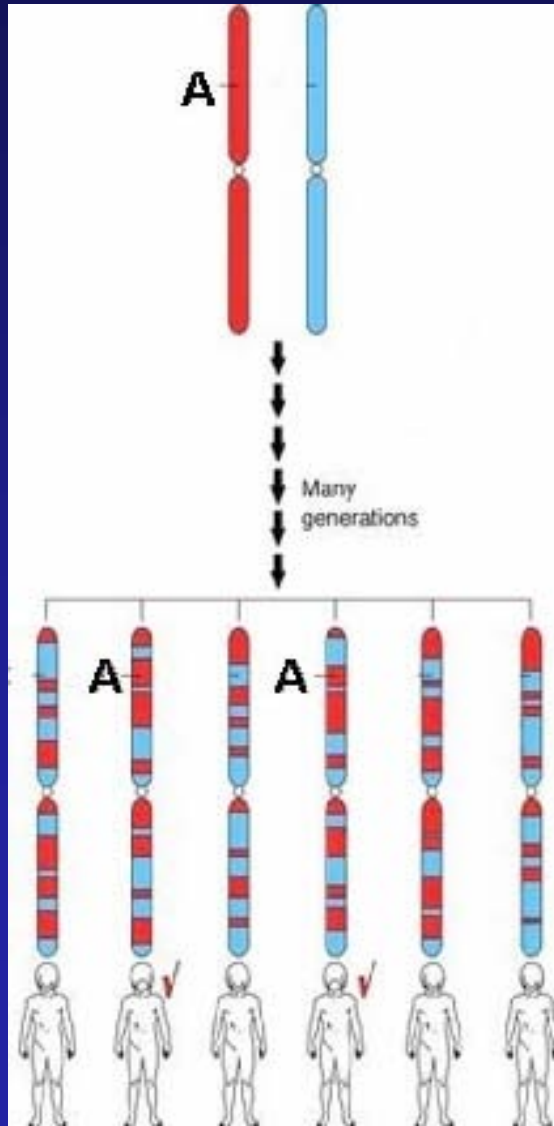
The human population is

young enough that

disease causing mutations

Will be linked to common

variants:

Also – can 'Tag'.

nature

THE HAPMAP PROJECT

**Chapter and verse on
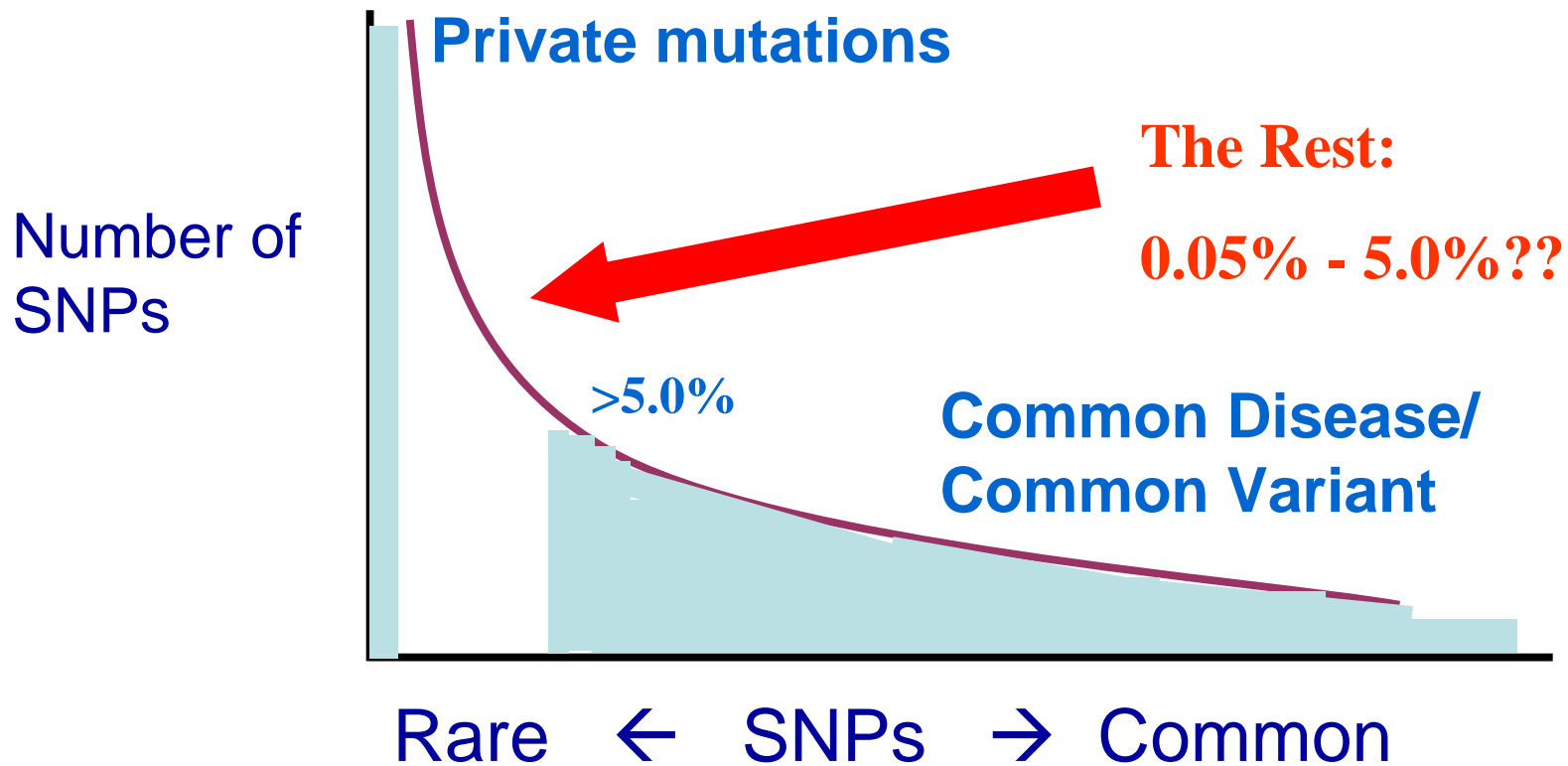human genetic variation**

HapMap Shows the Landscape:
e.g. Genomic Distribution of LD (CEU)
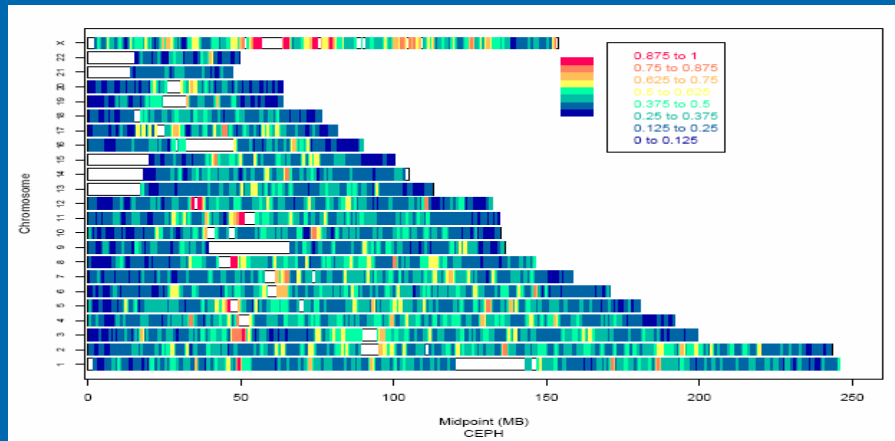
**Expected r² at 30kb
Bright Red > 0.88
Dark Blue < 0.12**

Some early disease successes and many

WGA maturing in the next few months

Overall Distribution of Variation

Contribution of multiple rare alleles

- "Nonsynonymous variants, although individually rare, are cumulatively frequent and influence

Multiple Rare Alleles Contribute to Low Plasma Levels of HDL Cholesterol

Jonathan C. Cohen,[1,2,3][†] Robert S. Kiss,[5] Alexander Pertsemlidis,[1] Yves L. Marcel,[5][†] Ruth McPherson,[5] Helen H. Hobbs[1,3,4]

Science 305:869-72 (2004)

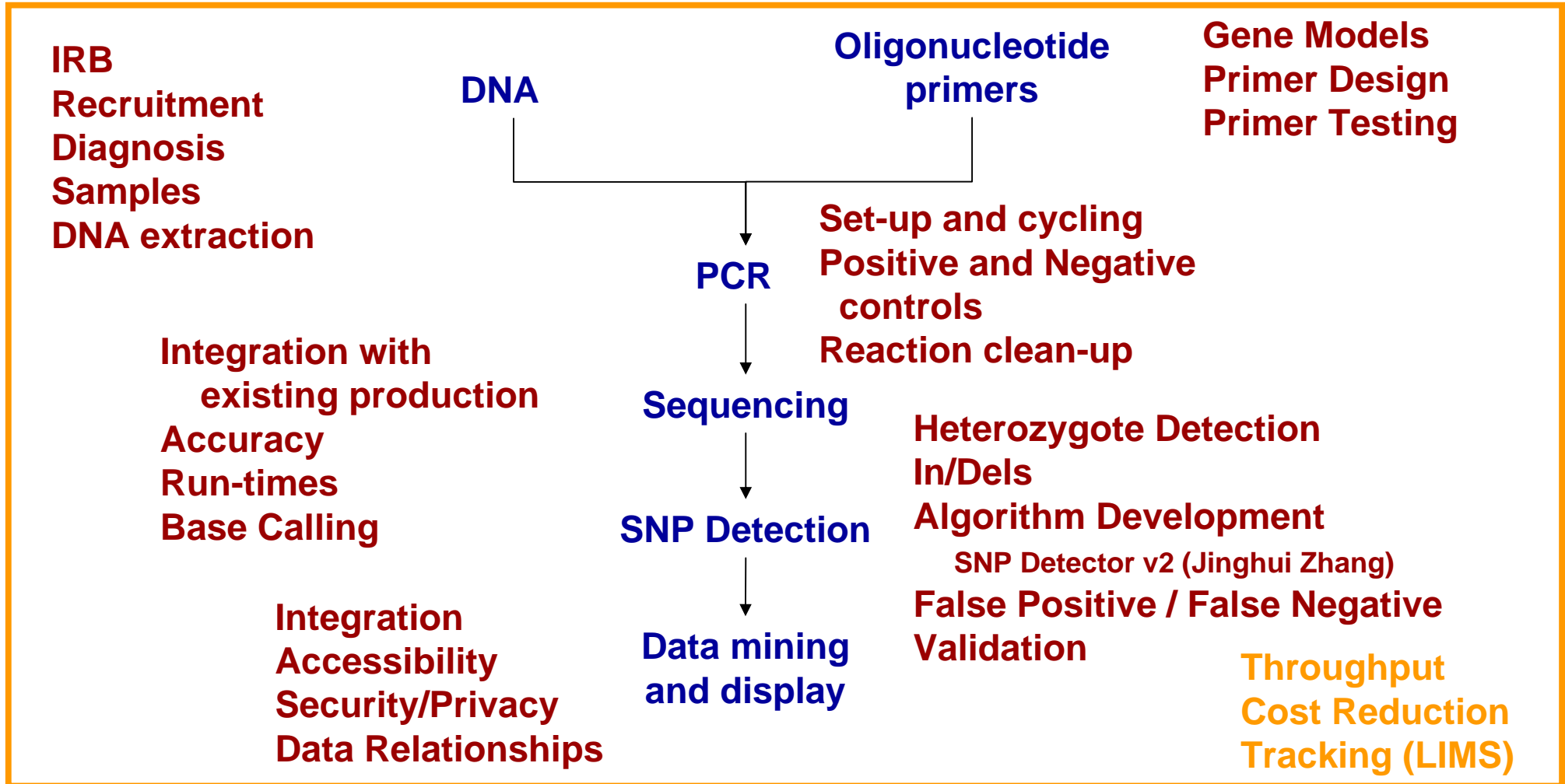**Multiple rare variants in *NPC1L1* associated with reduced sterol absorption and plasma low-density lipoprotein levels**

(cholesterol absorption | complex trait | genetic architecture | mutation | plant sterol)

Jonathan C. Cohen *[†‡§], Alexander Pertsemlidis [‡], Saleemah Fahmi *, Sophie Esmail [¶], Gloria L. Vega *[†], Scott M. Grundy *[†], and Helen H. Hobbs *[‡¶||]

PNAS 103:1810-1815 (2006)
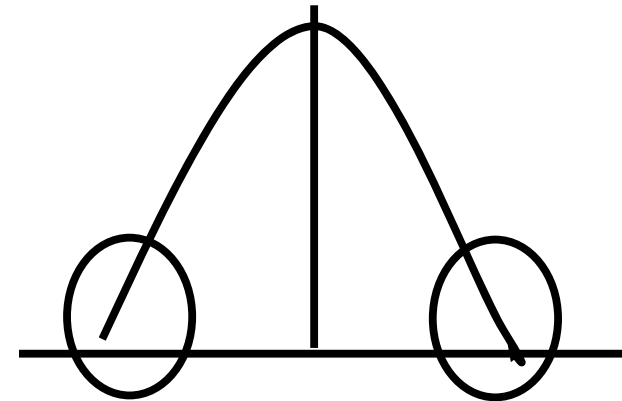
Medical Resequencing at BCM-HGSC

- HPRT (1988)
- CCR2, Noonans Syndrome,
- MGC clone validation
- West Nile Virus susceptibility
- Schizophrenia (2 studies)
- HapMap ENCODE regions
- Bipolar (50 genes)
- Parkinson's Disease
- Premature birth
- Idiopathic Generalized Epilepsy (IGE)

- Juvenile rheumatoid arthritis
- Prostate cancer
- NSC lung cancer (TSP)
- Bladder cancer
- Pediatric cancer
- Brain cancer
- Alpha1 anti-trypsin deficiency
- Juvenile obesity
- Leber's congenital amaurosis
- Mouse (ENU)

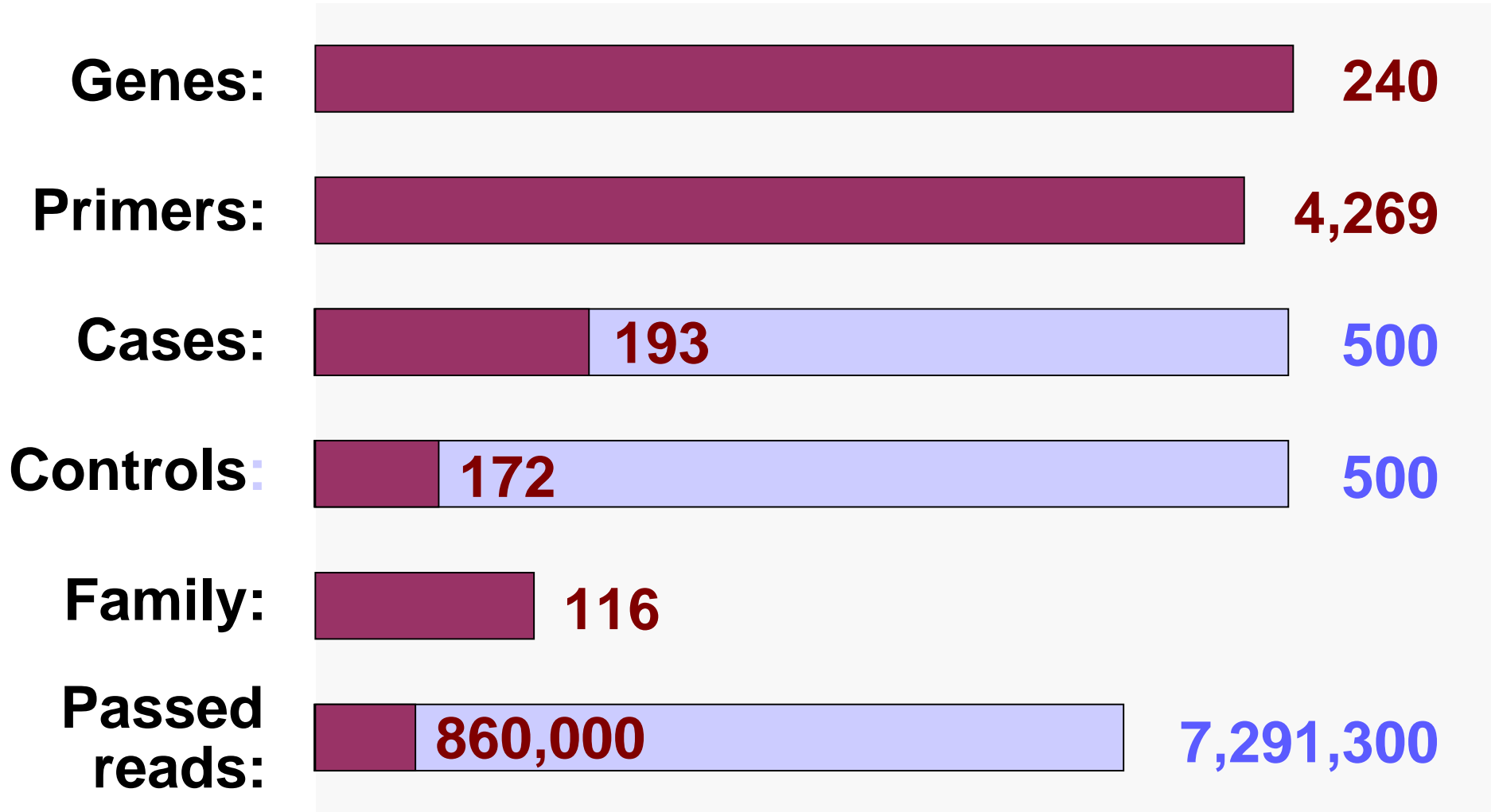~1,000 genes in play in all current projects

Baylor Human Channelopathy Project

NINDS (Jeff Noebels) and NHGRI (HGSC)

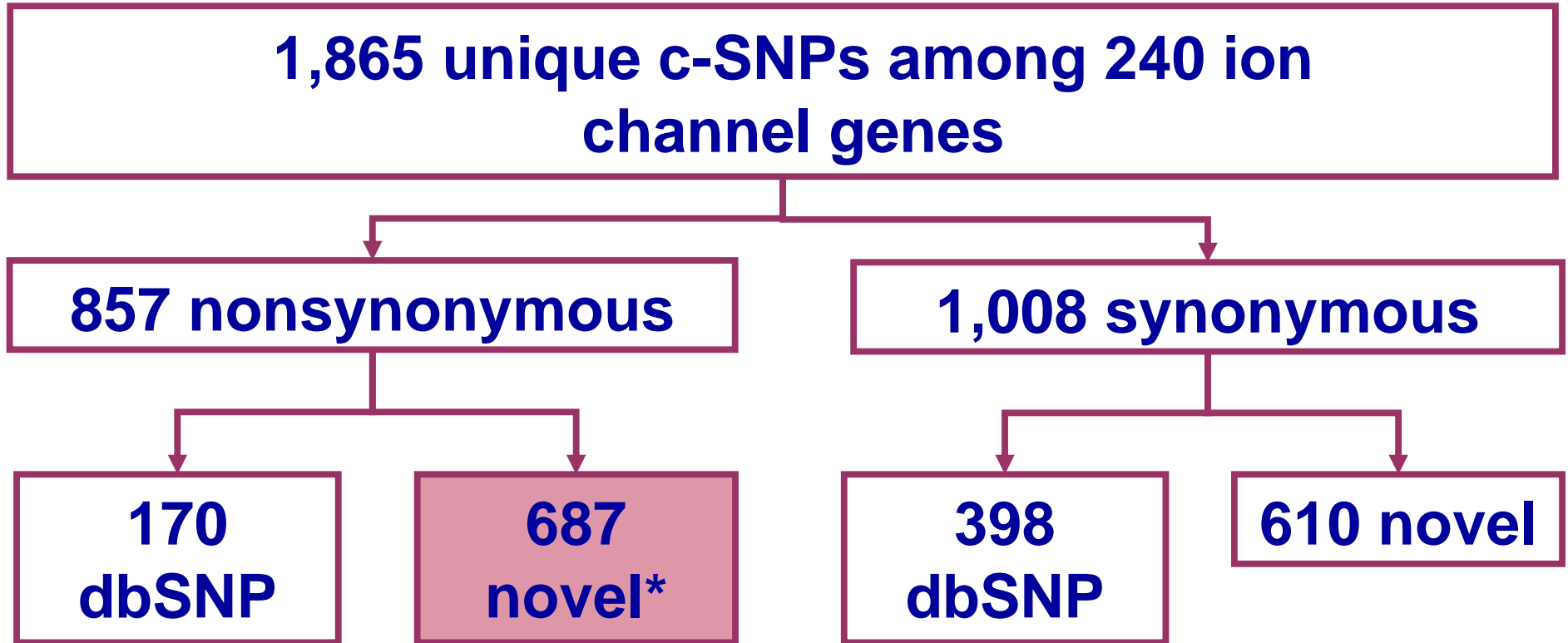# Baylor Human Channelopathy Project

Goals:

- Sequence ~250 ion channel genes in 1000 individuals
    (500 IGE patients and 500 matched controls)

- Develop tools for large scale medical sequencing and build a public database of human ion channel SNPs

- Test hypotheses on role of ion channel gene variation in sporadic epilepsy: common allele -vs- rare variant models, oligogenic and *de novo* mutations

- Use clinical genomics to develop individualized information on risk, clinical outcome, and response to therapy in common single index cases.

Progress (January 2007)

Genes:	240	
Primers:	4,269	
Cases:	193	500
Controls:	172	500
Family:	116	
Passed reads:	860,000	7,291,300

Interim Results

A lot of Variation!!

1,865 unique c-SNPs among 240 ion channel genes

857 nonsynonymous

1,008 synonymous

170 dbSNP

687 novel*

398 dbSNP

610 novel

Top 10 nsSNP

Gene	nsSNP
RYR3	35
CACNA1H	33
RYR1	32
KCNJ12	28
ANK2	27
SCN10A	20
RYR2	19
CACNA1G	17
CLCNKB	16
GRIN3A	16
CACNA1A	14
CLCNKA	14
CACNA1E	13
SCN5A	13
CACNA1I	12
CACNA1S	12
GRM5	12
KCNH6	12
SCN2A2	12
SCN4A	12
SCN9A	12
CACNA1D	11
KCNA10	11
CACNA1F	10

Gene	nsSNP/kb
KCNJ12	21.51
KCNE1	10.03
HTR3D	9.52
CLCNKB	7.75
KCNA10	7.16
KCNMB1	6.94
CLCNKA	6.78
KCNJ1	6.63
KCNG4	6.35
HTR3B	6.03
KCNK16	5.38
KCNE2	5.38
KCNK6	5.31
HTR1B	5.12
GRIN3A	4.78
KCNA1	4.70
CACNA1H	4.67
HTR5A	4.66
GRINL1A	4.52
HTR3A	4.51
KCNJ15	4.43
KCNA7	4.38
KCNV2	4.27
KCNJ11	4.26

Gene	nsSNP/kb/person
CLCNKB	0.19
KCNJ12	0.19
CLCNKA	0.15
CHRNA3	0.13
KCNE1L	0.11
KCNJ4	0.10
KCNC2	0.09
KCNG1	0.08
DRD4	0.08
SCN1B	0.08
KCNE1	0.08
KCNG4	0.07
CACNG2	0.07
SCN4B	0.07
KCNE2	0.07
CHRNA2	0.06
GABRA6	0.06
DRD3	0.06
CHRNA5	0.06
CACNB2	0.06
CACNG1	0.06
DRD1	0.06
CACNG4	0.05
GLRB	0.05

raw count   normalized gene len   normalized gene len & people

Top 10 genes

Gene	All	Gene	cSNP	Gene	nsSNP
RYR1	199	RYR1	87	RYR1	32
RYR3	151	KCNJ12	57	KCNJ12	28
RYR2	125	CACNA1H	53	CACNA1H	26
CACNA1H	102	RYR3	41	ANK2	19
GABRA4	79	RYR2	37	RYR2	19
KCNT1	76	ANK2	30	CLCNKB	16
KCNQ2	70	CACNA1G	28	CACNA1G	15
CACNA2D3	66	KCNH6	27	CLCNKA	14
KCNJ12	66	SCN10A	26	RYR3	14
KCNJ5	63	CACNA1A	25	SCN10A	14

Novel nonsense mutations

Gene	Chr	Exon	AA	Codon	Allele Frequency			AA	Codon	Phenotype
KCNH2	chr7	13	W	T**G**G	GG.25	AG.1	AA.0	Term	T**A**G	Control
CLCN1	chr7	24	R	**C**GA	CC.58	CT.1	TT.0	Term	**T**GA	Proband
CLCN2	chr3	15	W	TG**G**	GG.60	AG.1	AA.0	Term	TG**A**	Proband
GABRR1	chr6	10	E	**G**AG	GG.49	GT.1	TT.0	Term	**T**AG	Proband
GRIN2A	chr16	14	C	TG**C**	CC.50	AC.1	AA.0	Term	TG**A**	Proband
HTR3A	chr11	2	L	T**T**G	TT.128	AT.1	AA.0	Term	T**A**G	Control
KCNK16	chr6	6	Q	**C**AG	CC.122	CT.1	TT.0	Term	**T**AG	Proband
KCNK7	chr11	3	Q	**C**AG	CC.128	CT.1	TT.0	Term	**T**AG	Control
KCNQ3	chr8	9	E	**G**AA	GG.116	GT.3	TT.0	Term	**T**AA	N/A
KCNT1	chr9	13	R	**C**GA	CC.133	CT.1	TT.0	Term	**T**GA	Proband
KCNV2	chr9	2	S	TC**G**	CC.18	AC.1	AA.0	Term	T**A**G	Proband
RYR1	chr19	35	Q	**C**AG	CC.29	CT.0	<u>TT.1</u>	Term	**T**AG	Control
SCN10A	chr3	15	R	**C**GA	CC.44	CT.2	TT.0	Term	**T**GA	Proband
SCN10A	chr3	27	R	**C**GA	CC.45	CT.1	TT.0	Term	**T**GA	Proband
SCN1A	chr2	11	E	**G**AG	GG.17	GT.1	TT.0	Term	**T**AG	Proband
SCN2A2	chr2	14	R	**A**GA	AA.40	AT.1	TT.0	Term	**T**GA	Proband

Novel, NS variants were found in ion channel genes associated with several disorders

Brain
Epilepsy
Cerebellar ataxia
Familial Migraine

Heart
Long QT syndrome,
    types 2, 3, 4
Ventricular tachycardia

Eye, Ear
Congenital stationary
    night blindness
Nonsyndromic sensorineural
    hearing loss

Skeletal Muscle
Myotonia
Malignant hyperthermia
Hypokalemic periodic paralysis
Slow-channel myasthenic syndrome

Kidney
Bartter syndrome, types 3 & 4
Bartter syndrome with deafness

Multisysytem
Infantile malignant osteopetrosis

Functional Variants?

Chr. 20

KCNQ2
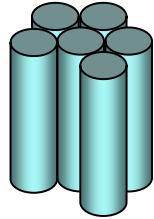
KCNQ2 + KCNQ3
function as a
heterotetramer
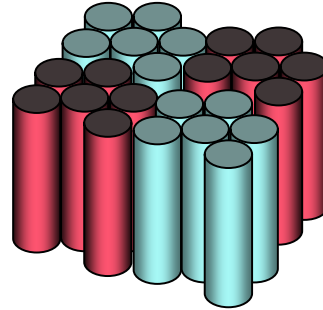
KCNQ3
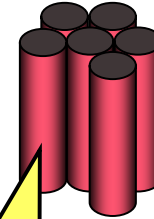
Chr. 8

KCNQ2, *KCNQ3* are two
related potassium channels
known to be mutated in
monogenic epilepsy:

P ►T

The functional effect of
this novel variant
observed in our study
is unknown

The "M" current
hyperpolarization *reduces*
neuronal excitability

BCM
Baylor College of Medicine
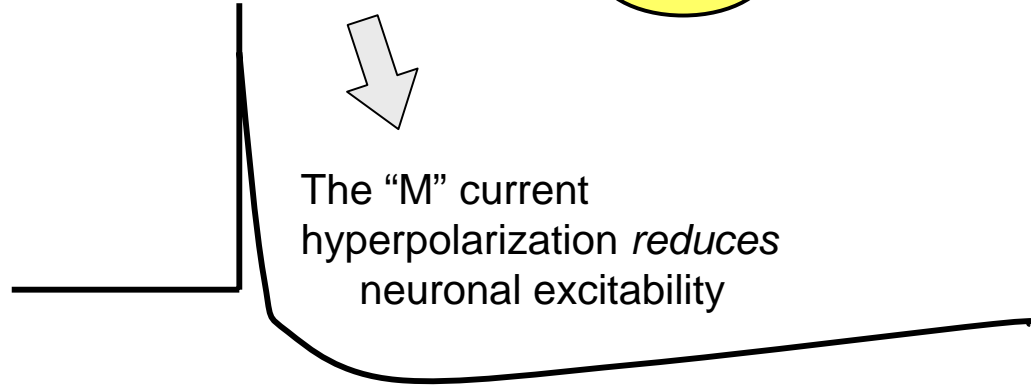
Statistical Power Not Yet Adequate:
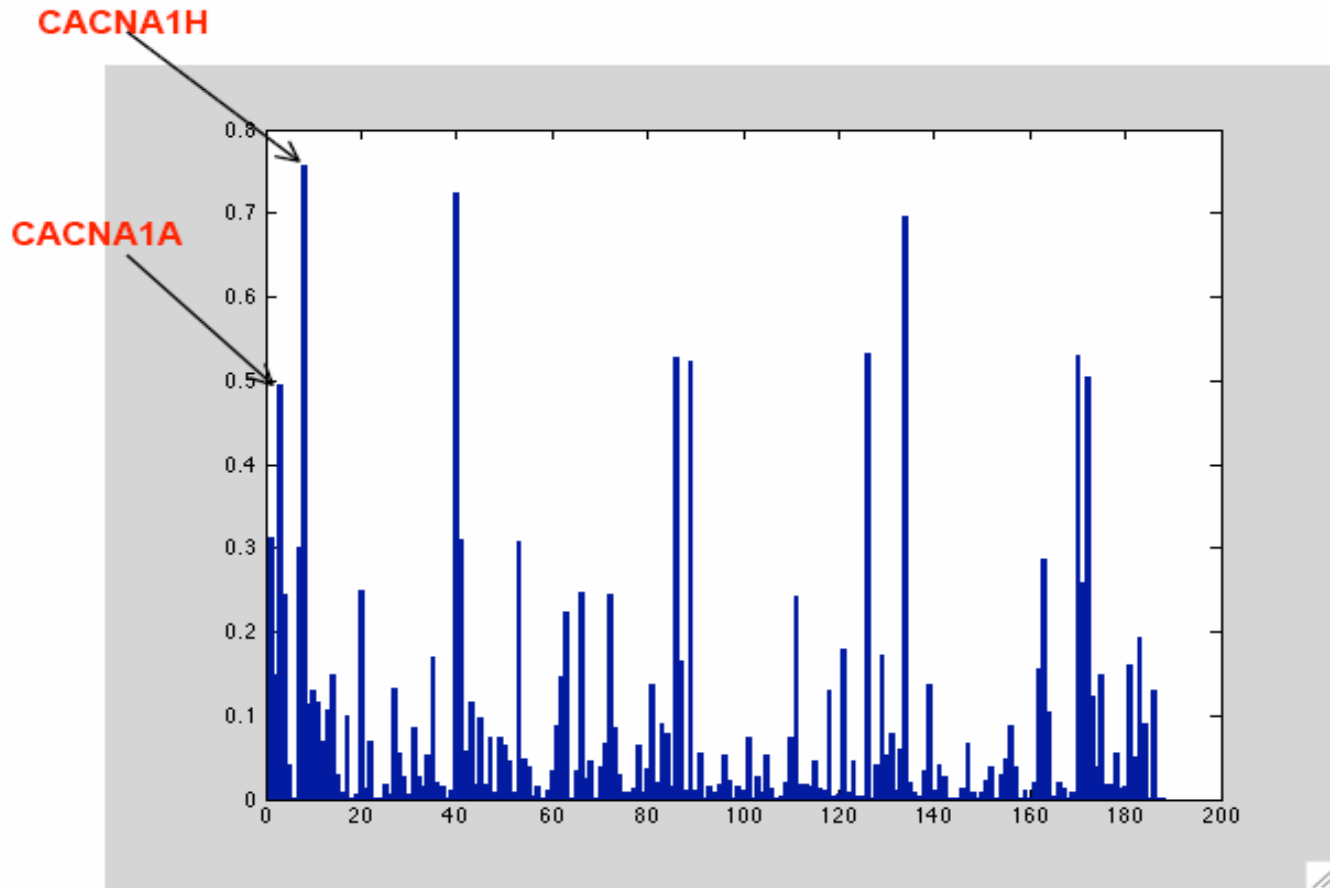
The discrimination power of CACNA1A and CACNA1H

The Validation Cycle:

Pyro- Sequencing OR

Molecular Inversion Probes (1.5 – 10K-plex)
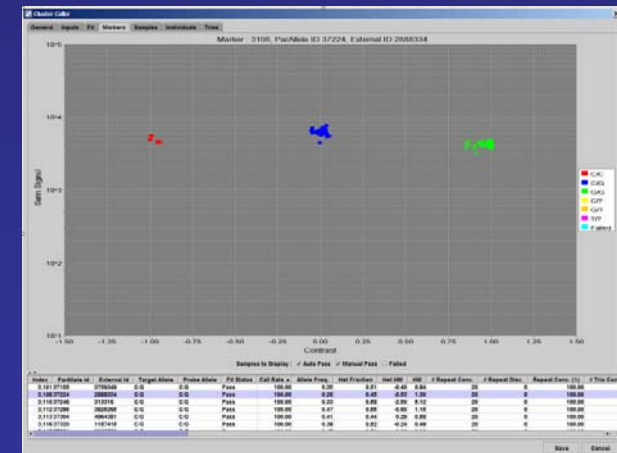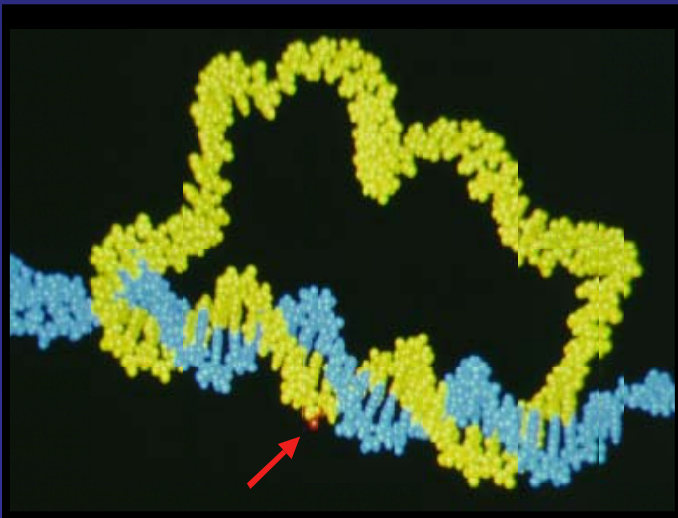
Minor allele frequency distribution
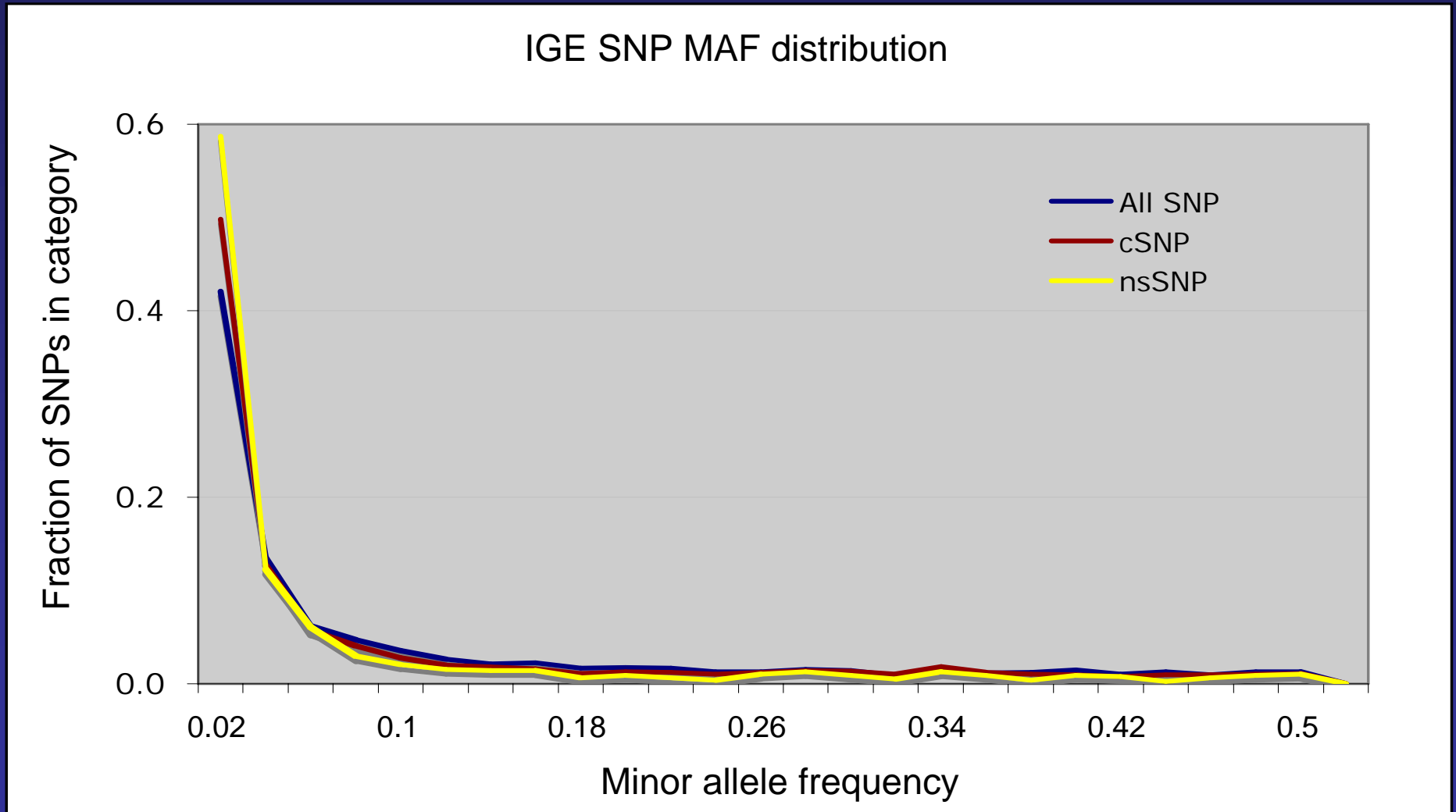
IGE SNP MAF distribution
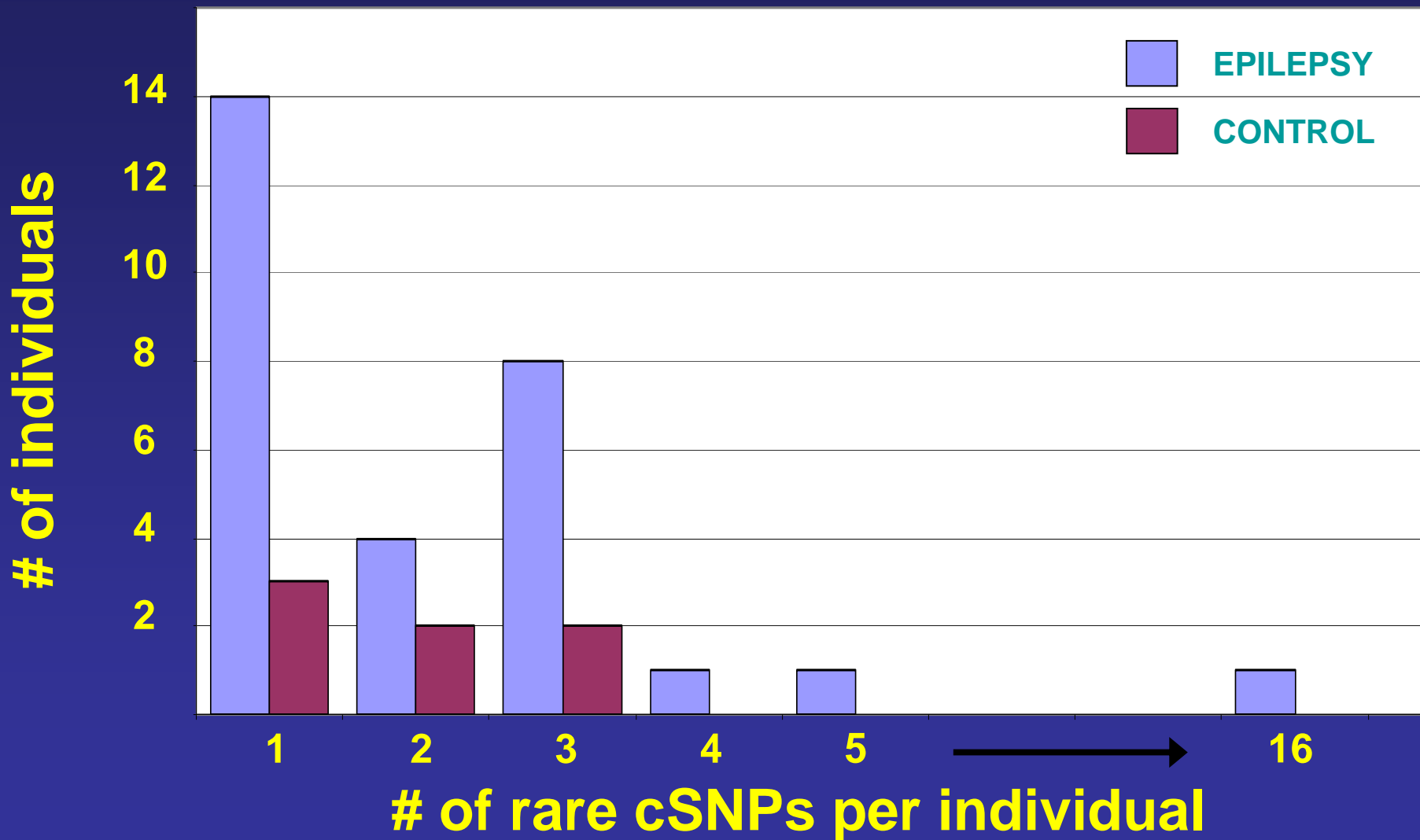
Functional Analysis – High Throughput?
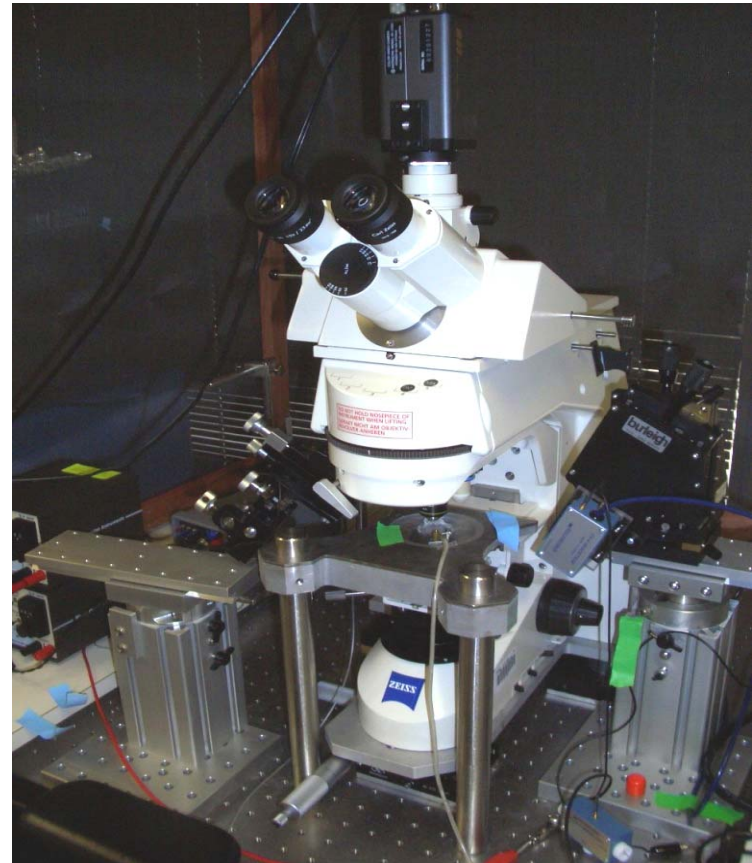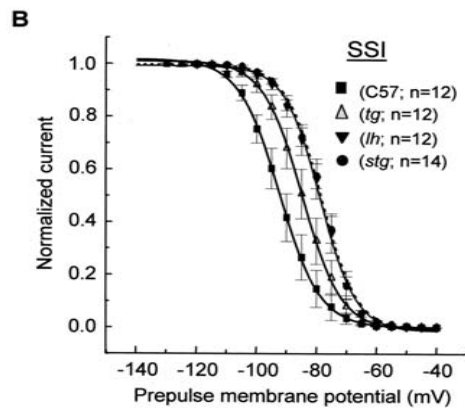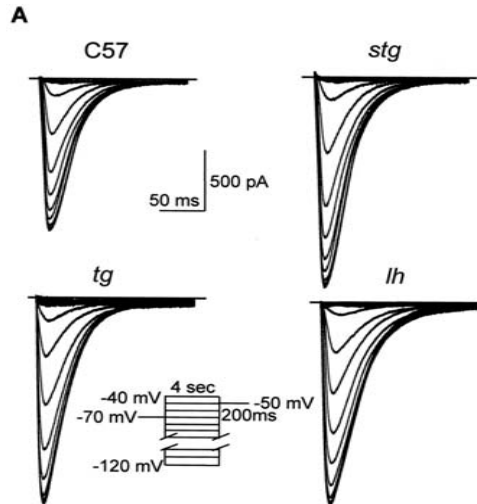
Detecting all rare variants

Private mutations

Find all between 0.05% and 5.0%???

Number of
SNPs

Common Disease/
Common Variant
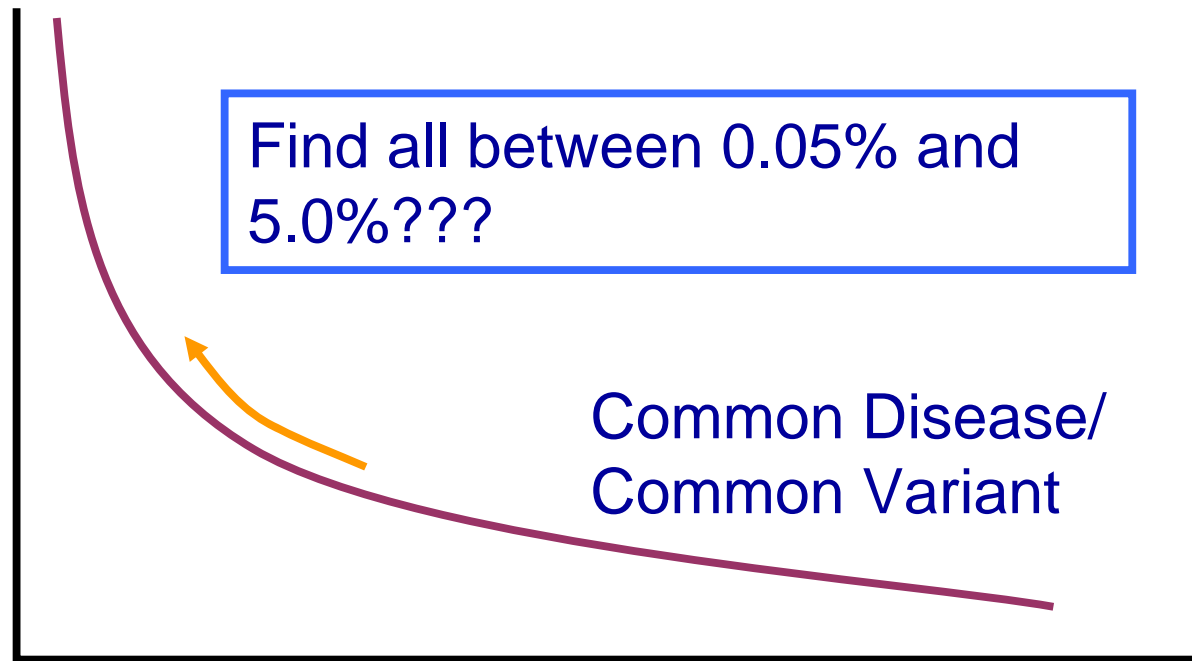
Rare ← SNPs → Common

Technology drives the realities …
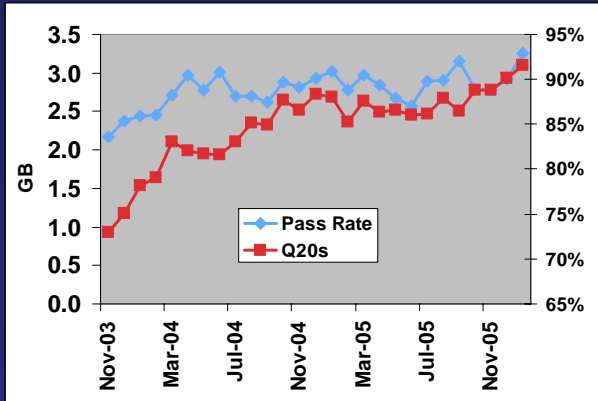
AB Cost ~0.55/kb

One 10x genome ~ $18M

OR

All exons by PCR ~ $400,000

Personalized genomes – expensive!!

PCR-AB
Reads
/month

All Exons,
1 person
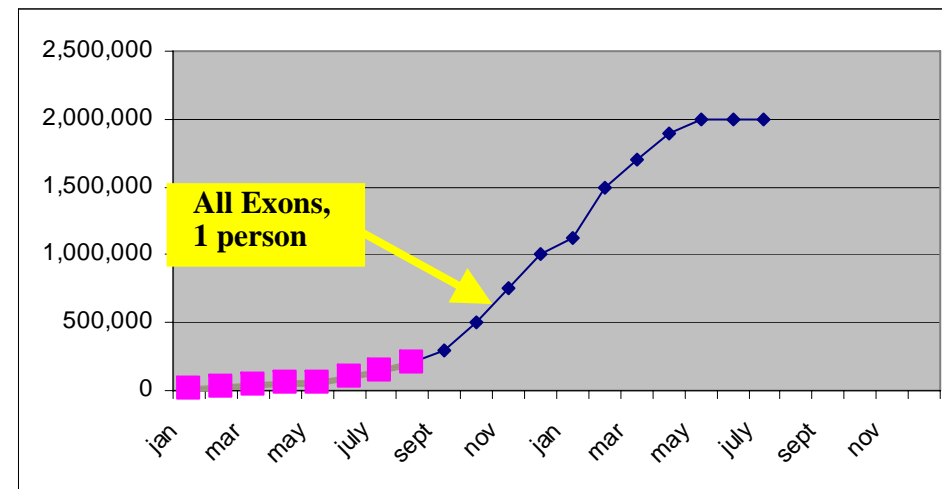
454 DNA Sequencing:
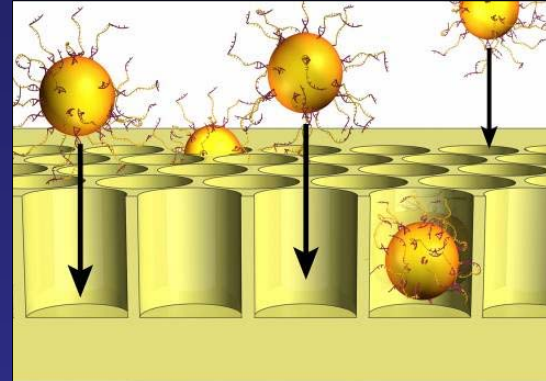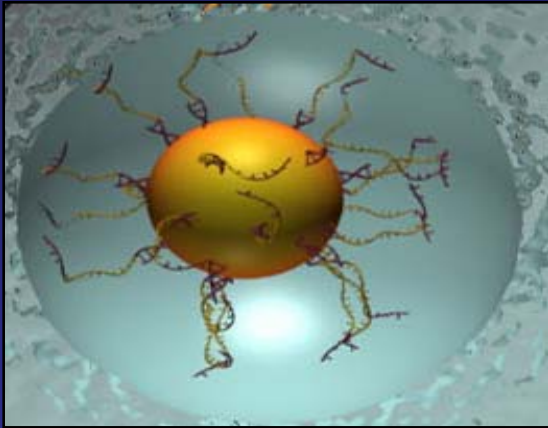
[RG is a member of the
454 Life Sciences SAB.
This conflict is managed
by the BCM COI policies.]
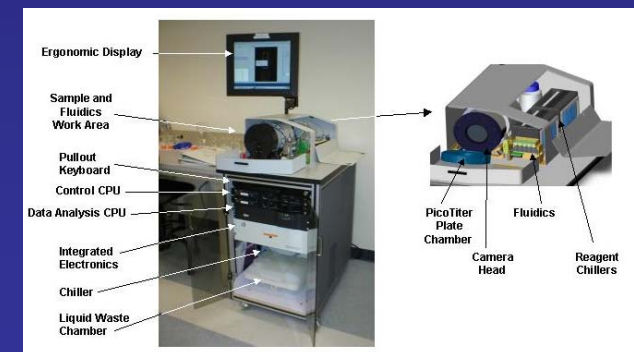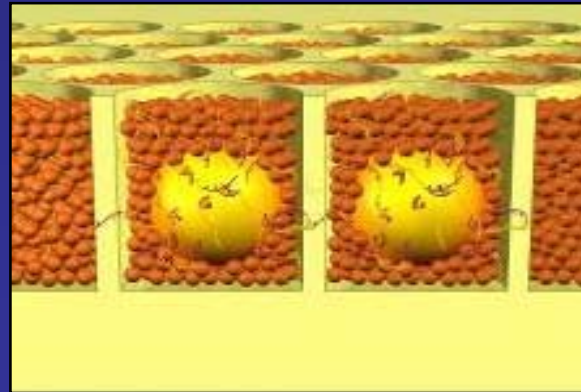
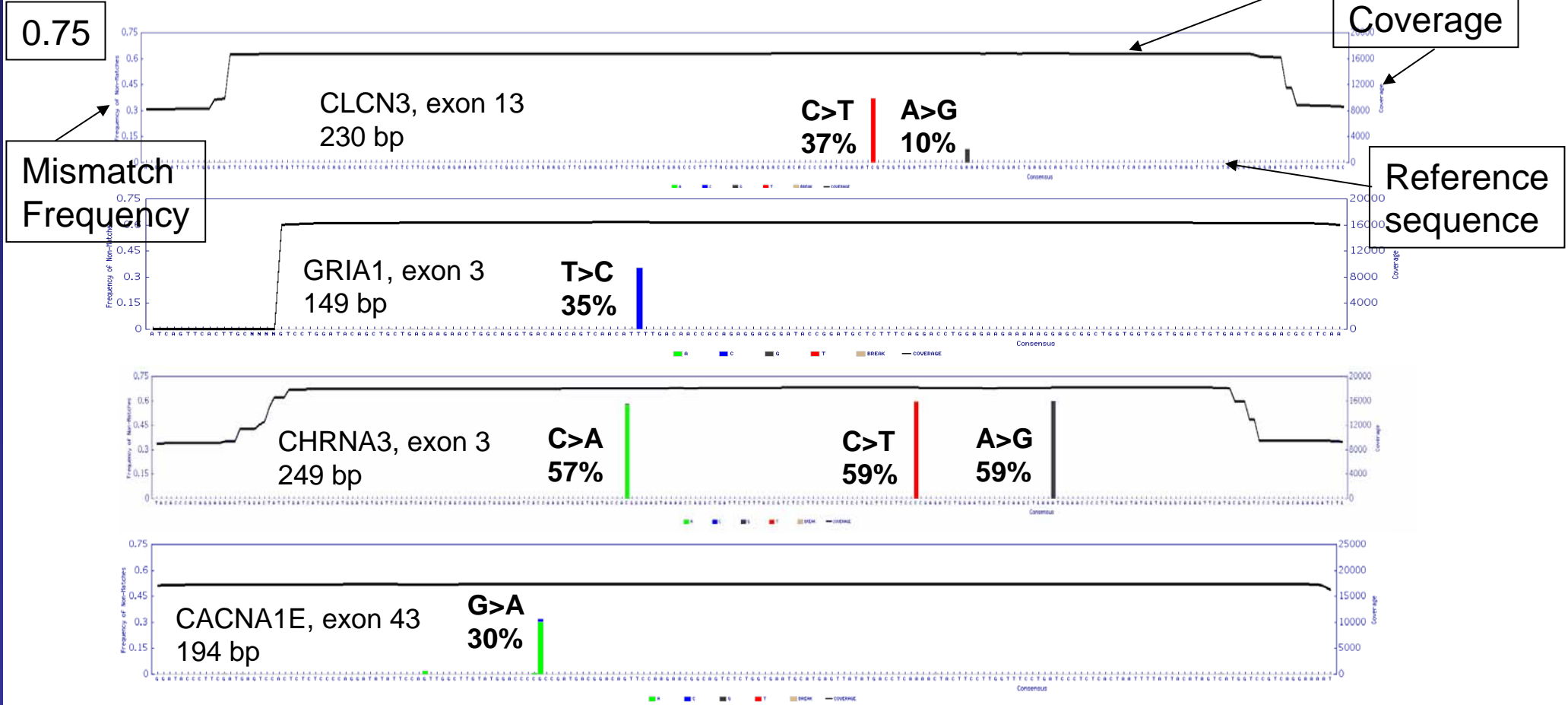# 454 Mutation Discovery Pipeline

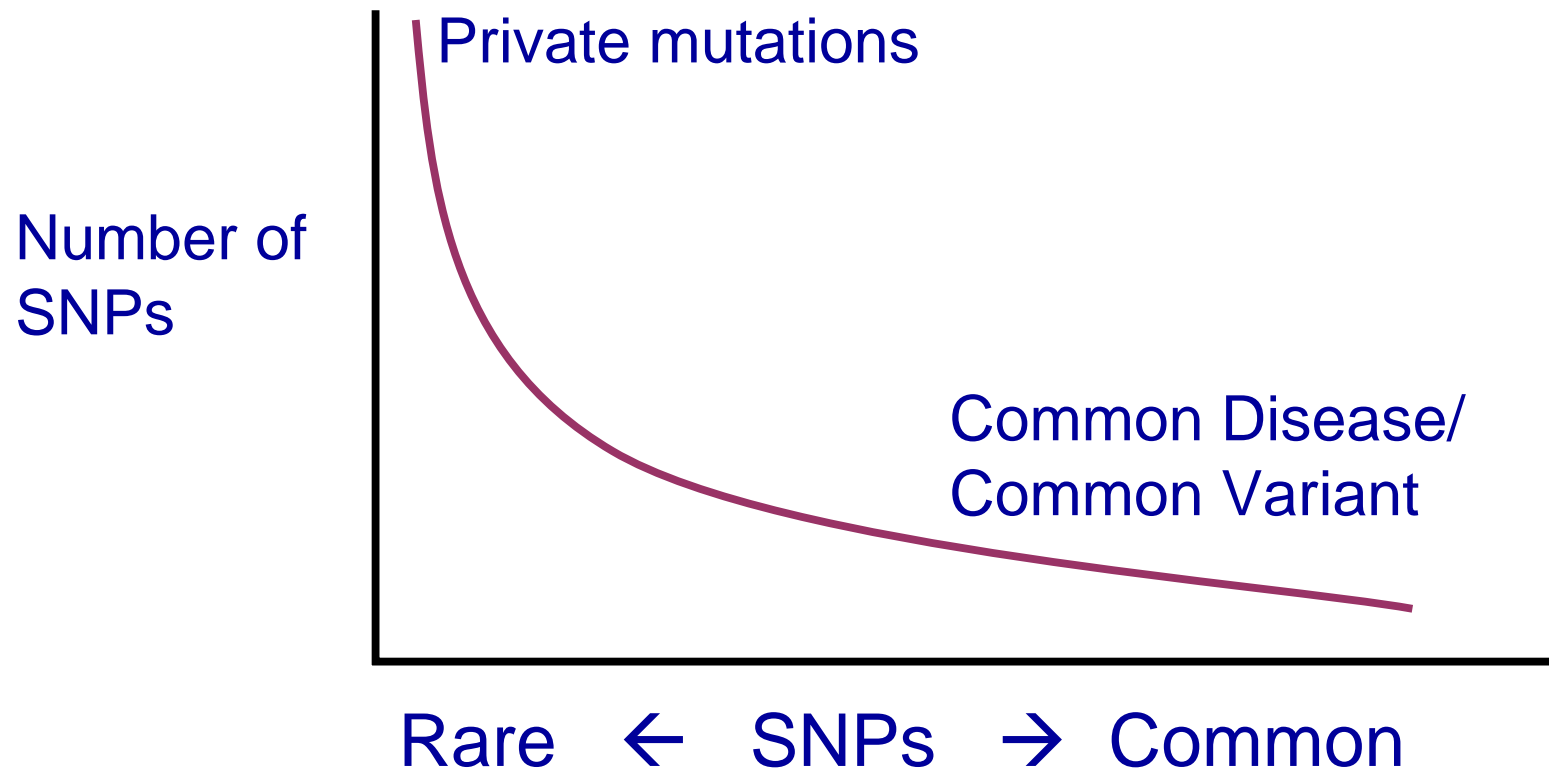Testing 50,000 haplicons at a time??
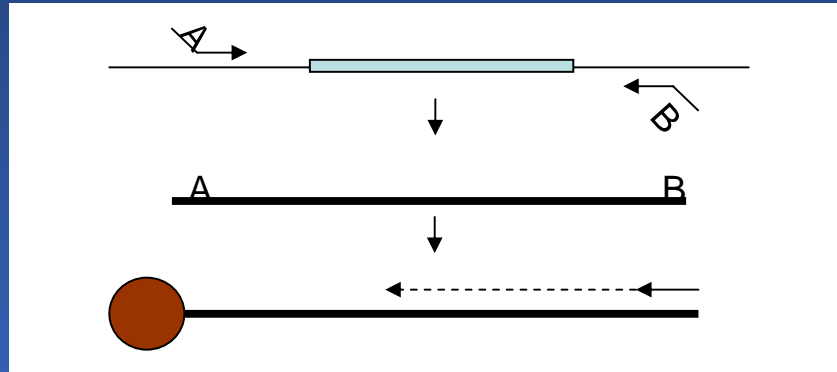
Best for Pooled PCR Products….

5 – 10 times less cost than current AB methods

# Medical Resequencing using the GS20

Pool of 92 individuals

Read Coverage

0.75

CLCN3, exon 13
230 bp

C>T 37% A>G 10%

Mismatch Frequency

Reference sequence

GRIA1, exon 3
149 bp

T>C 35%

CHRNA3, exon 3
249 bp

C>A 57% C>T 59% A>G 59%

CACNA1E, exon 43
194 bp

G>A 30%

Overall Distribution of Variation

Why Not 'Discover' all Rare Variants???

Private mutations

Find all between 0.05% and 5.0%???

Number of SNPs

Common Disease/ Common Variant
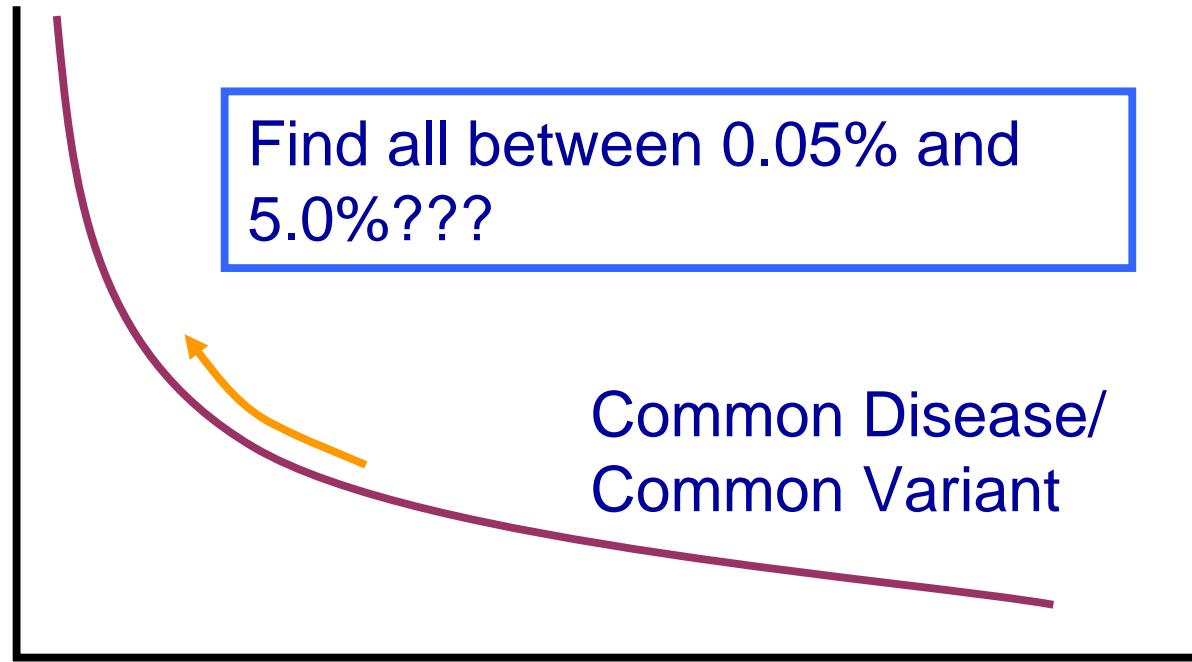
Rare ← SNPs → Common

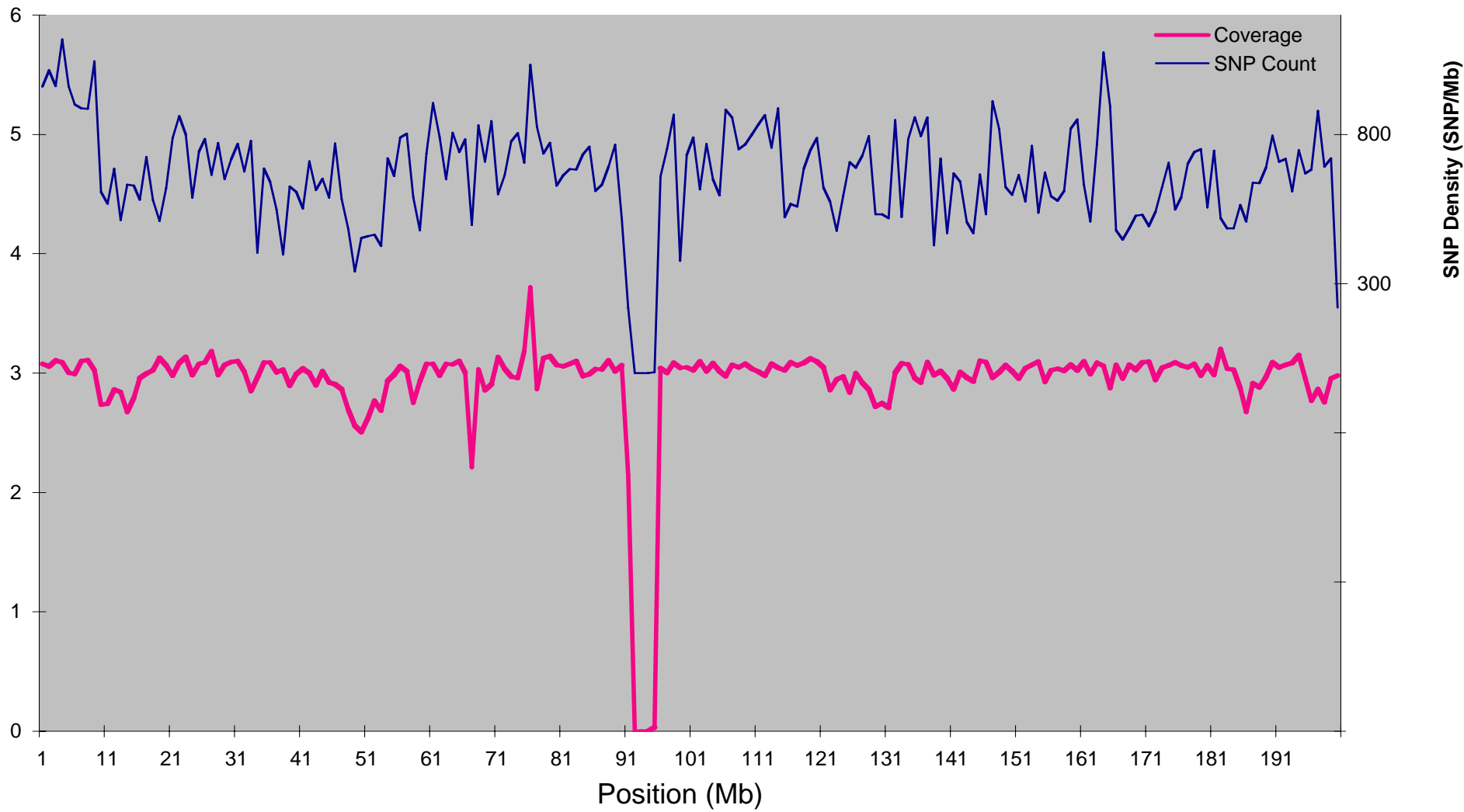# Goal: Sequence a Single Human's Genome

- Proof of concept
  - What can we learn from a single genome?
  - Can we get unbiased sequence?
  - What analytical issues arise?
- 454 Timeline
  - Dec 15 '06 10 million reads, ~2.5 Gbases
  - Jan 15 '07 20 million reads, ~5.0 Gbases
  - Jan 30 '07 40 million reads, ~10 Gbases

Distribution of Non-Repetitive Matches
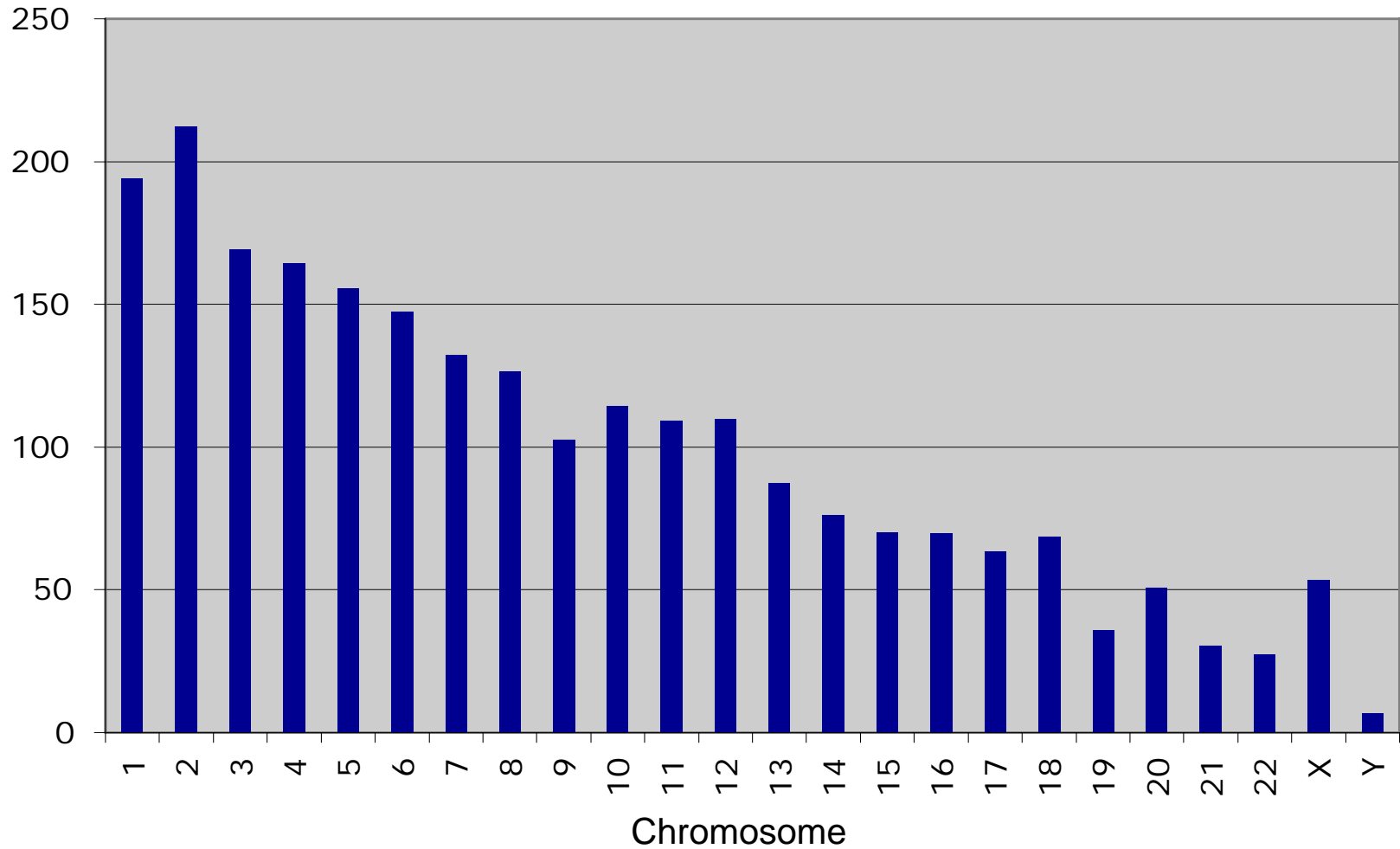
Watson: Chromosome 3 Base Coverage

Functional Alleles

- 50 SNPs match to database of human (disease) phenotypes
- 32 are recognized polymorphisms
- 18 of the 50 are "associated" with the phenotype base on population studies

> Get 6x coverage and deliver data to JW for releas

Conclusions:

- Determined many 'interesting' ion channel mutations,

- Ready for large scale functional analysis

- Poised for 'personal genomes'

HGSC Acknowledgements:

Acknowledgements