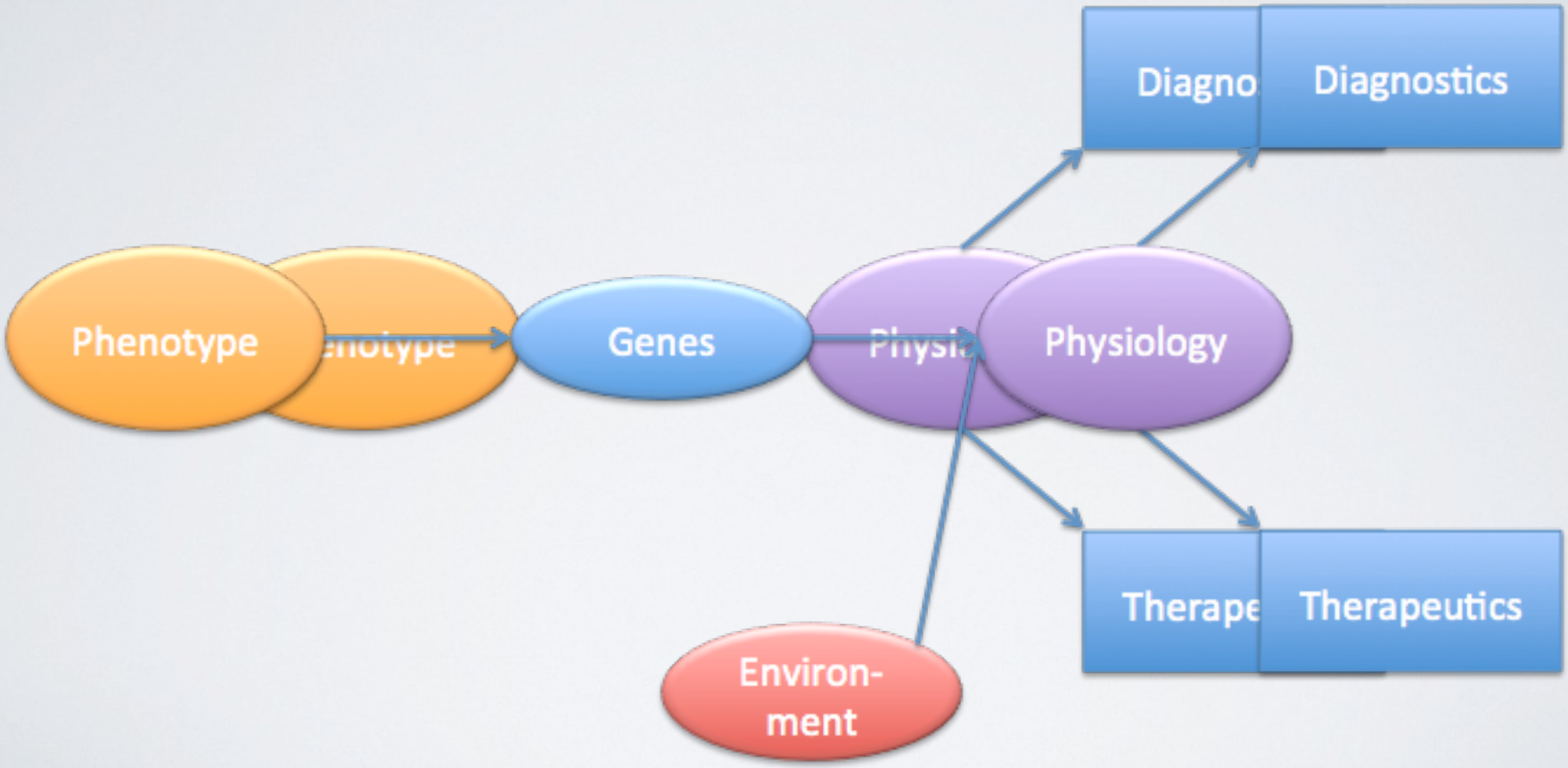
GENE IDENTIFICATION

Bruce R. Korf, M.D., Ph.D.
Department of Genetics

GENOMIC APPROACHES

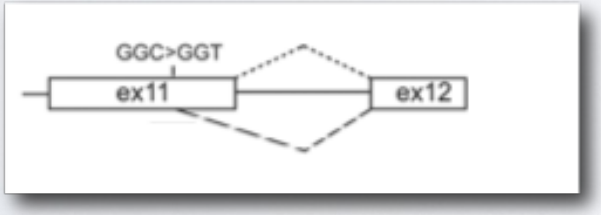
Level	Approaches
Population	Case-control association study
Family	Genetic linkage study
Individual	Genome sequencing
Tissue/Cell	Gene expression analysis, epigenetics, genome sequencing
Organism	Microbiome analysis

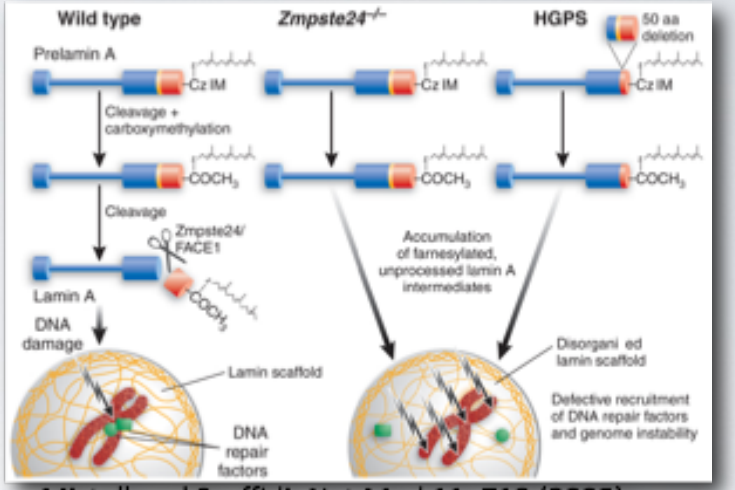
APPROACH TO GENETIC DISORDERS



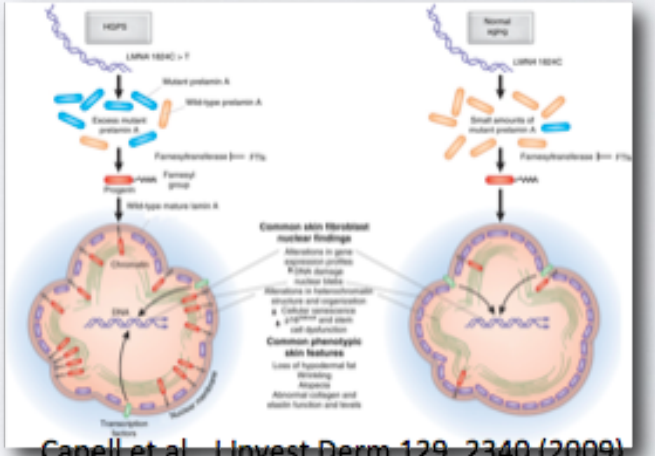
HUTCHINSON-GILFORD PROGERIA

Scaffidi and Mistell, Nat Med 11, 440 (2005)



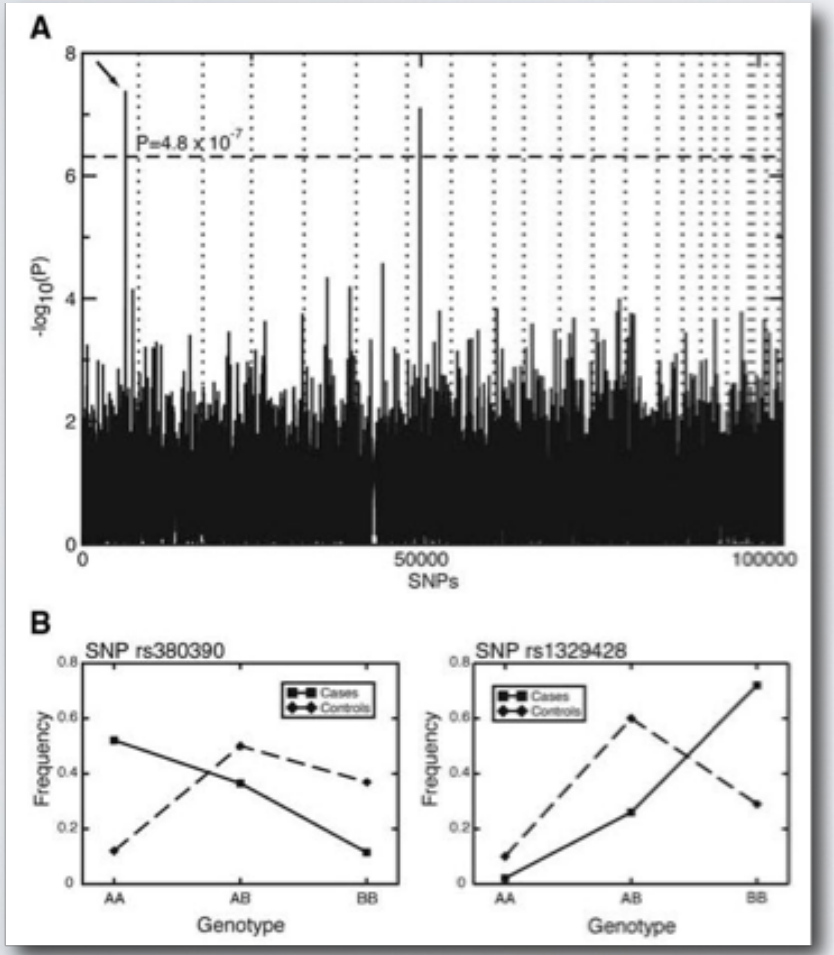


Mistell and Scaffidi, Nat Med 11, 718 (2005)

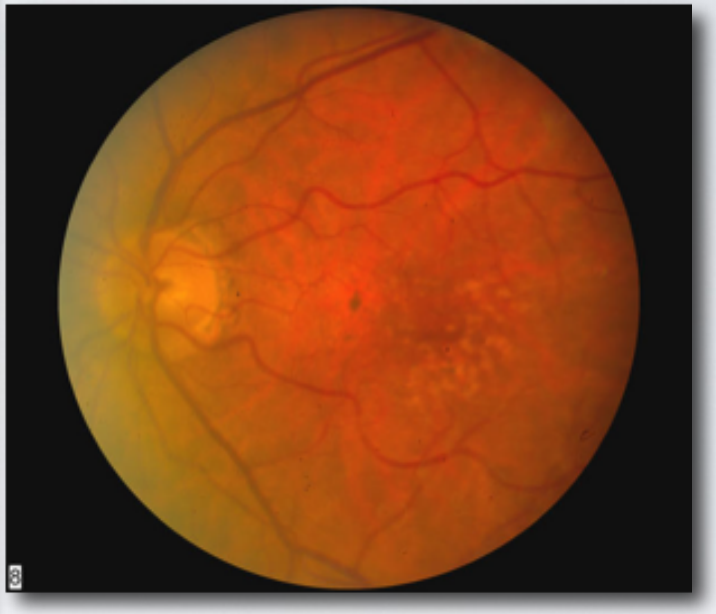


Capell et al., J Invest Derm 129, 2340 (2009)

AGE-RELATED MACULAR DEGENERATION

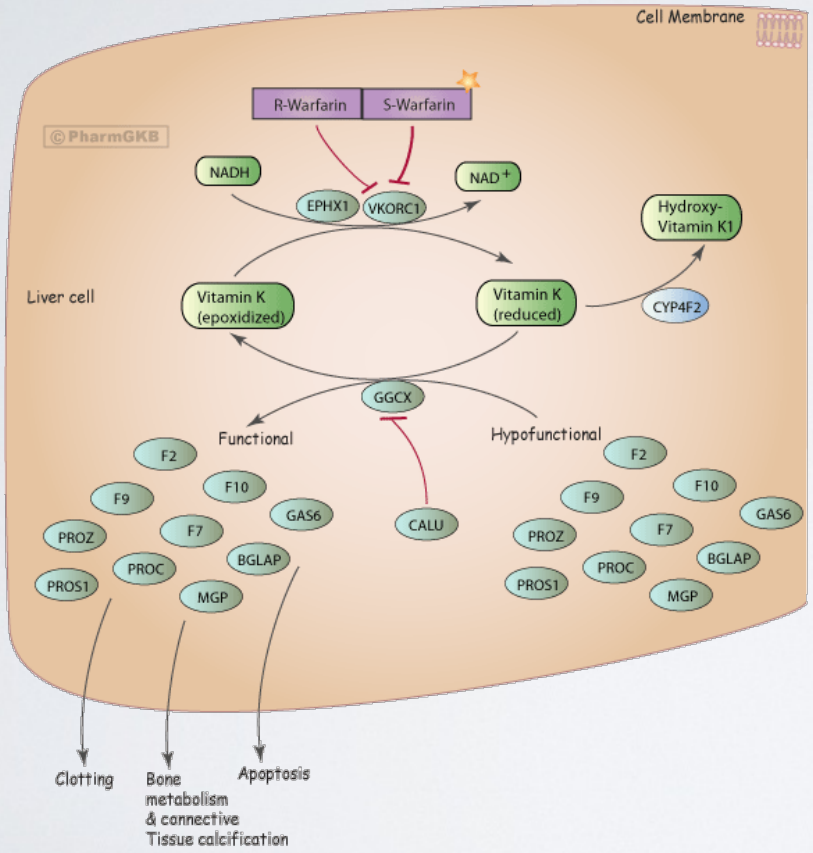


Klein et al., Science (80-). 2005 April 15; 308(5720): 385–389.



http://www.medrounds.org/amd/uploaded_images/fig2-757825.JPG

PHARMACOGENETICS



Show results for Bruce Korf

Print summary of elevated risks

« Return to Overview | Disease Risks | Carrier Status | Traits | Drug Response | Recently Updated

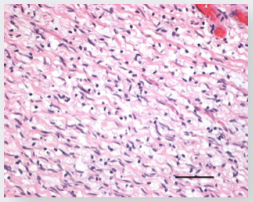
Name	Status	Last Updated
Warfarin (Coumadin®) Sensitivity	Increased	Mar 19, 2009
Abacavir Hypersensitivity	Typical	Oct 8, 2009
Clopidogrel (Plavix®) Efficacy	Typical	May 7, 2009
Drinking, Smoking, and Risk of Esophageal Cancer new	Typical	Jan 14, 2010
Fluorouracil Toxicity	Typical	Oct 1, 2009
Pseudocholinesterase Deficiency	Typical	Nov 19, 2009
Response to Hepatitis C Treatment new	Typical	Jan 14, 2010
Oral Contraceptives, Hormone Replacement Therapy and Risk of Venous Thromboembolism ♀ new	n/a	Feb 11, 2010

The genotyping services of 23andMe are performed in LabCorp's CLIA-certified laboratory. The tests have not been cleared or approved by the FDA but have been analytically validated according to CLIA standards.

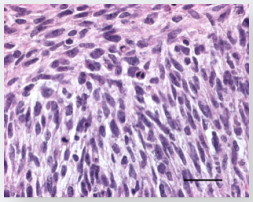
23andMe Name	Genotype	Combination
rs1799853	CC	
rs1057910	AA	CYP2C9 *1/*1, VKORC1 -1639/3673 AG
rs9923231	CT	

CANCER GENOMES

Normal



Tumor

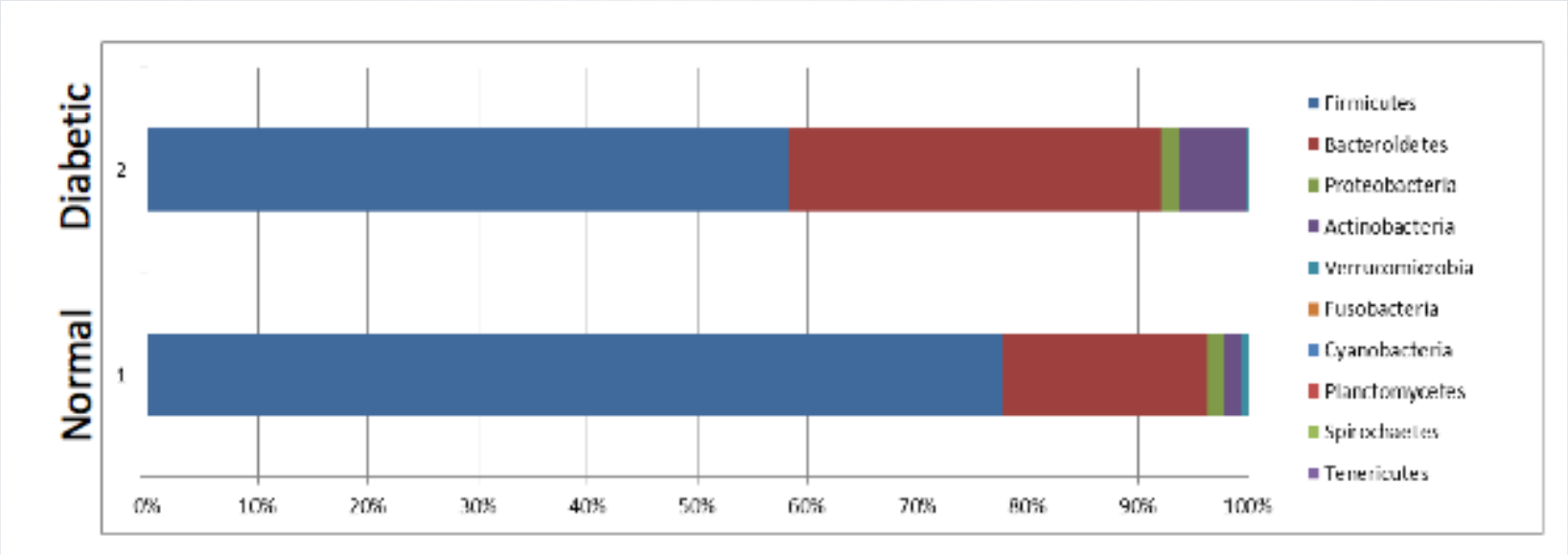


Sequence

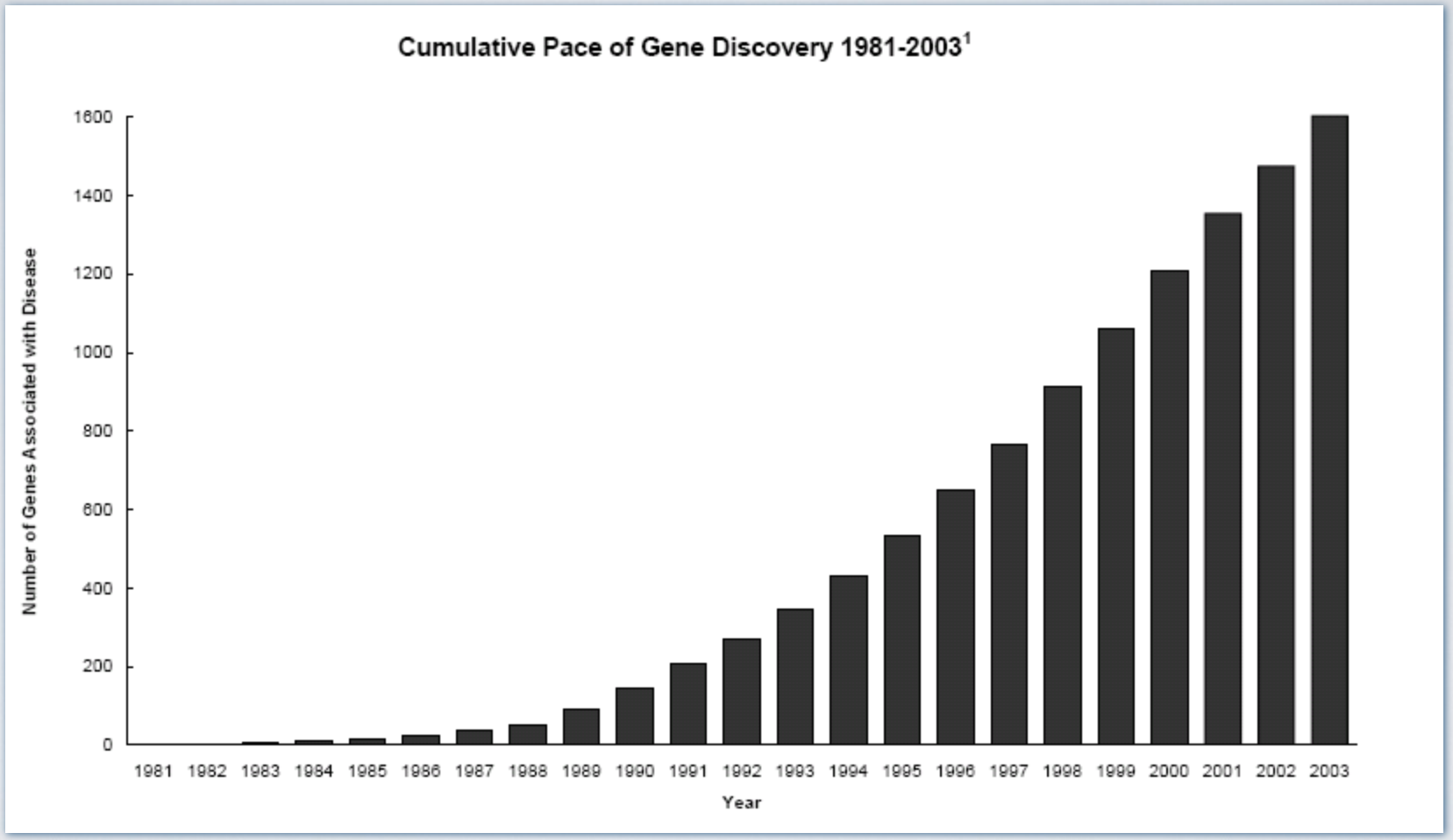
Difference =
cancer-specific genetic
changes

FUNCTIONAL GENOMICS

MICROBIOME

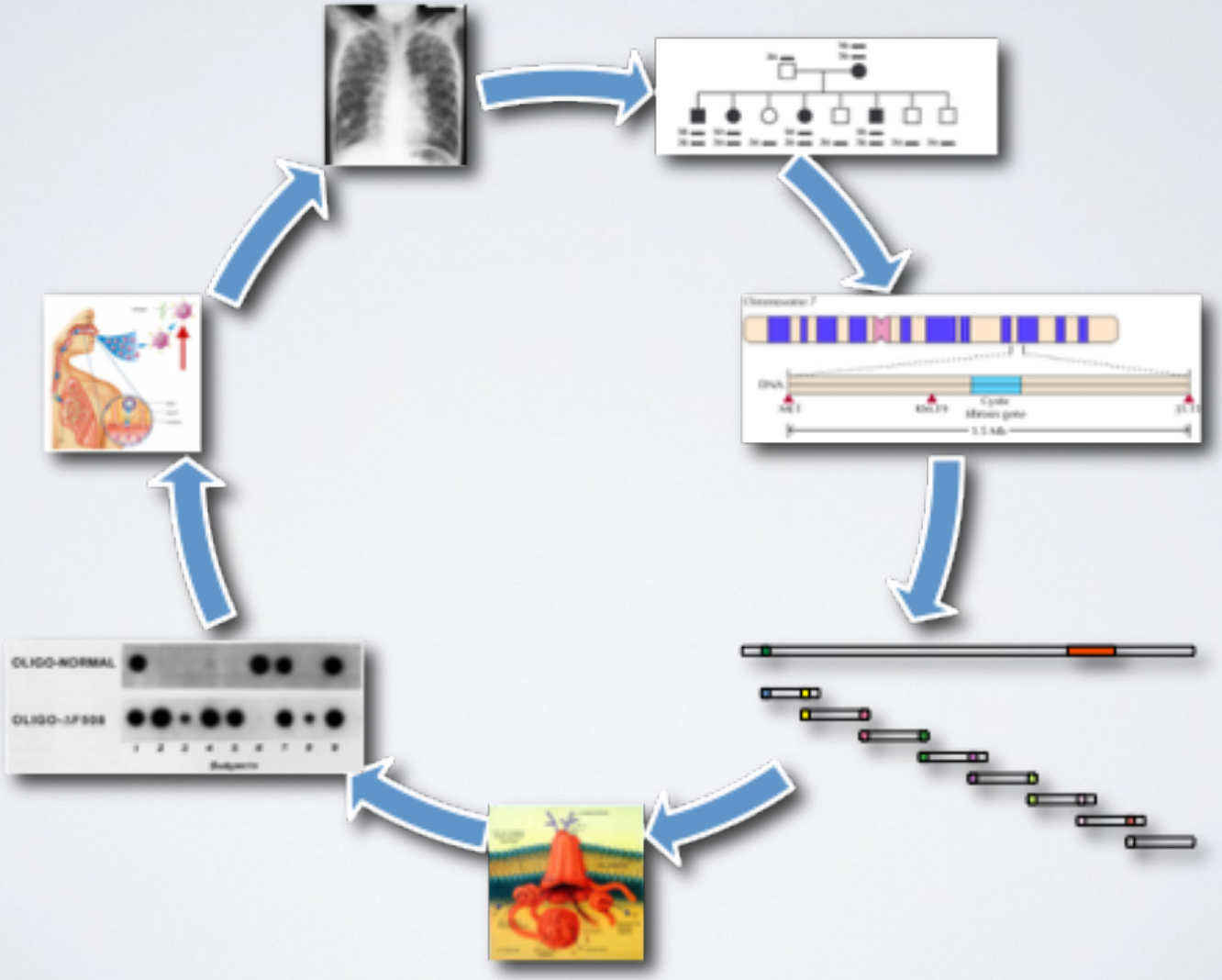

Diabetic
Normal
2
1
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Firmicules
Bacteroidetes
Proteobacteria
Actinobacteria
Verrucomicrobia
Fusobacteria
Cyanobacteria
Planctomycetes
Spirochaetes
Tenericules

GENE DISCOVERY

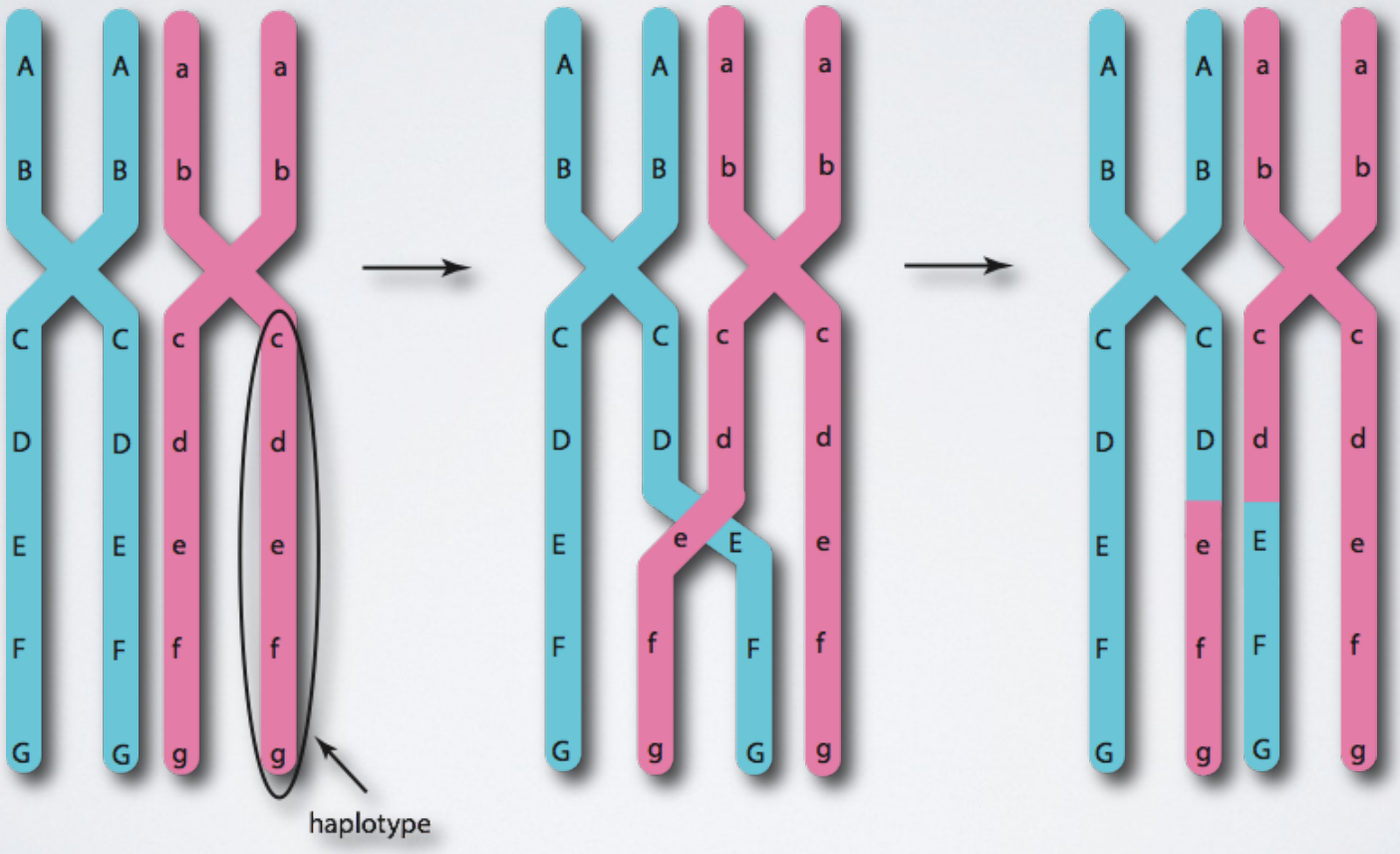


http://www.genome.gov/Pages/News/PaceofDiseaseGeneDiscovery.pdf

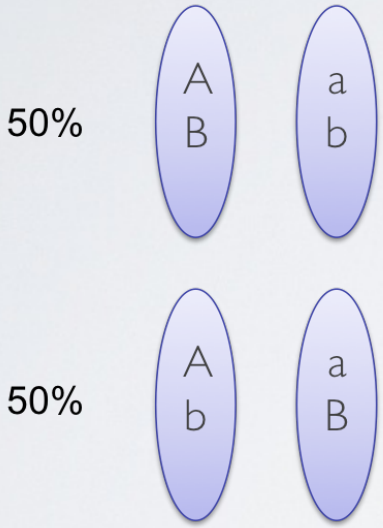
POSITIONAL CLONING

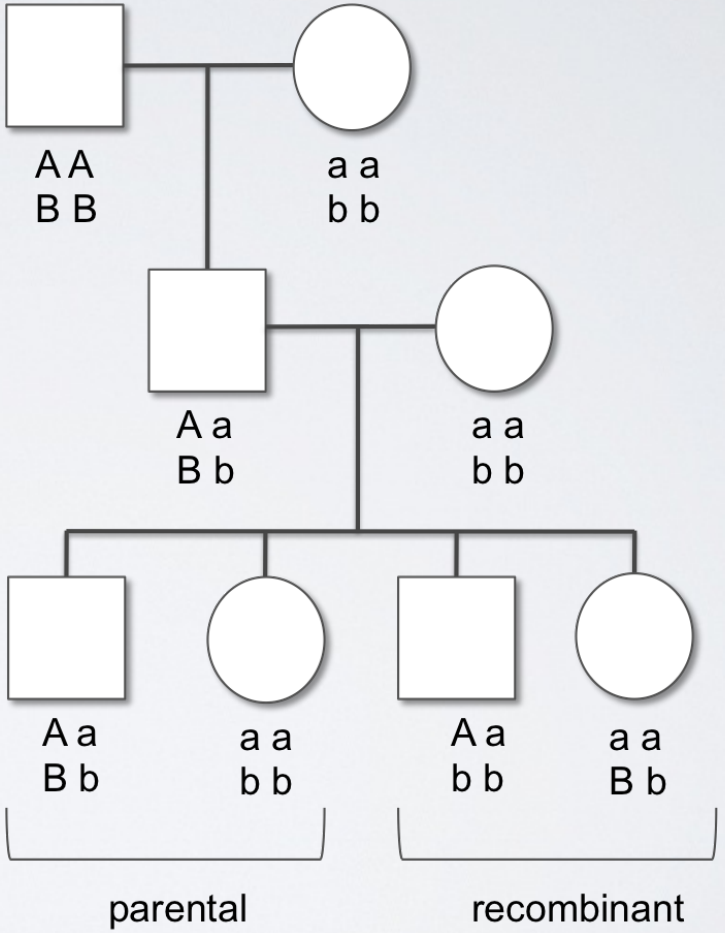


LINKAGE

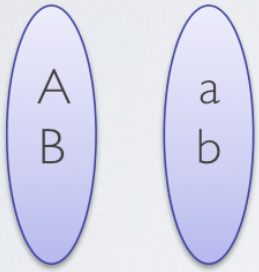


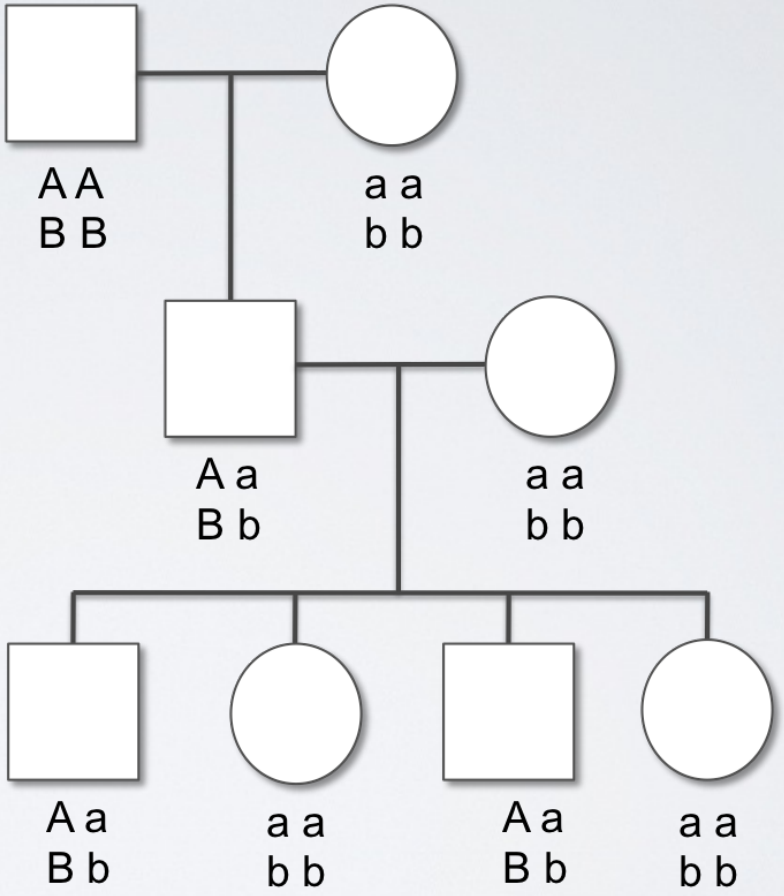
INDEPENDENT ASSORTMENT



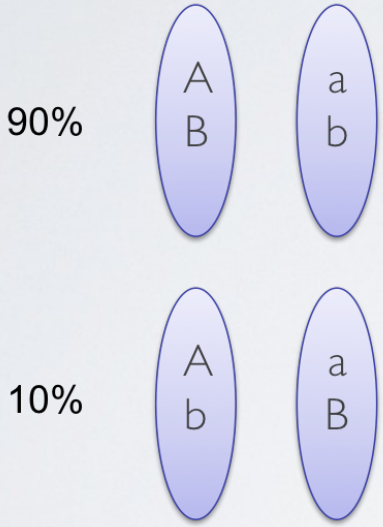


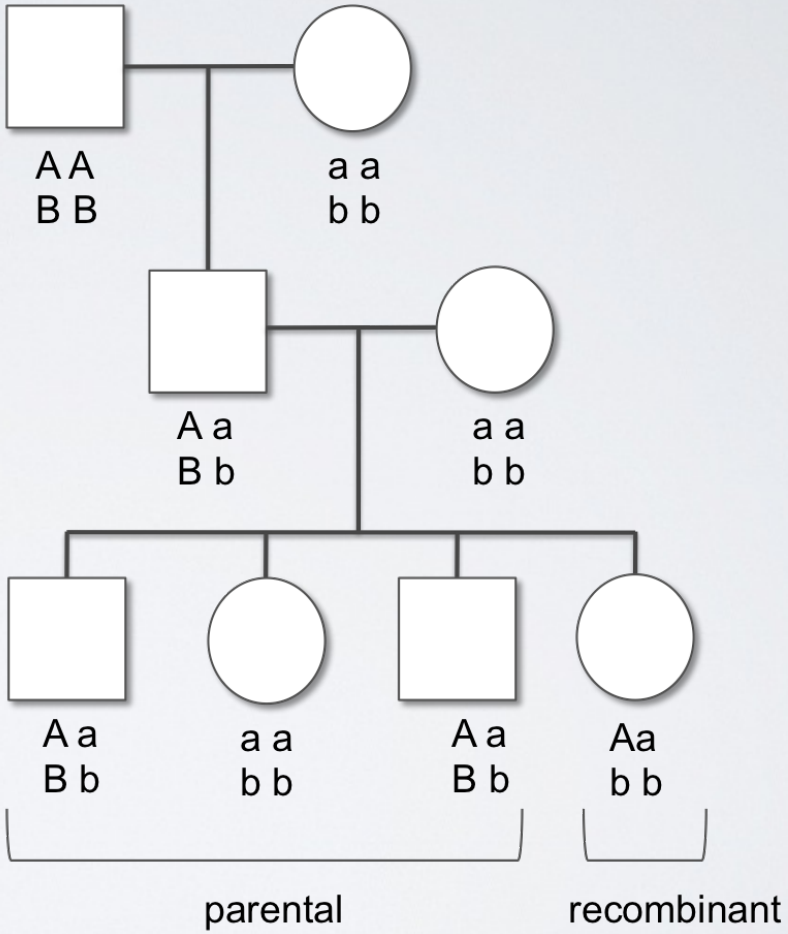
"COMPLETE" LINKAGE



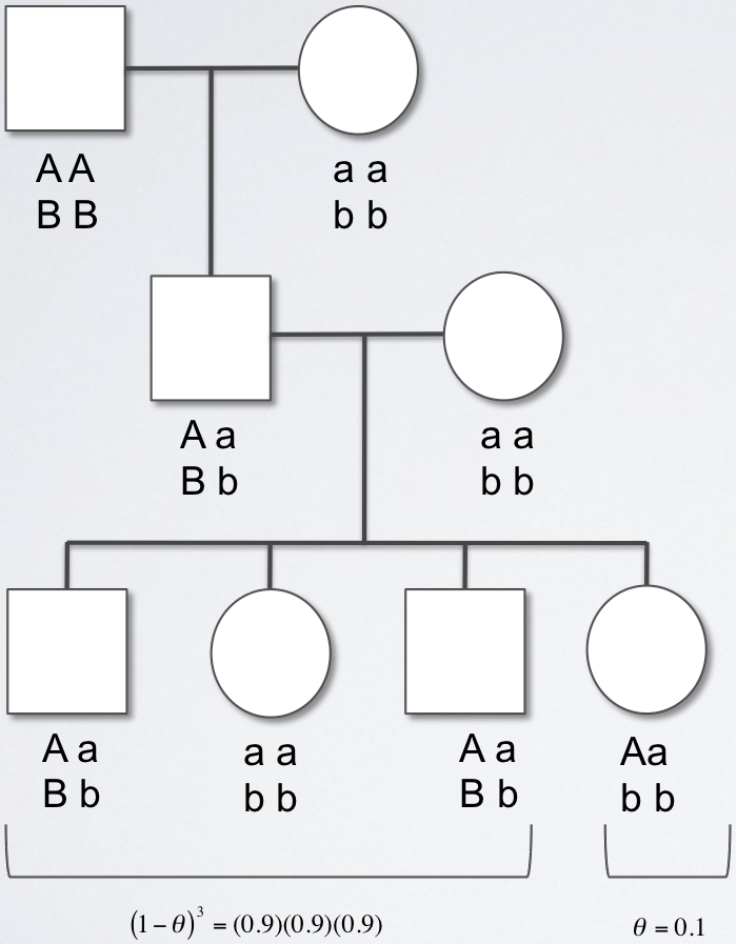


10% RECOMBINATION





LIKELIHOOD RATIO



odds ratio = $\dfrac{(1-\theta)^n(\theta)^r}{(1/2)^{n+r}}$

n = number non-recombinants
r = number recombinants

LINKAGE ANALYSIS

Family 1

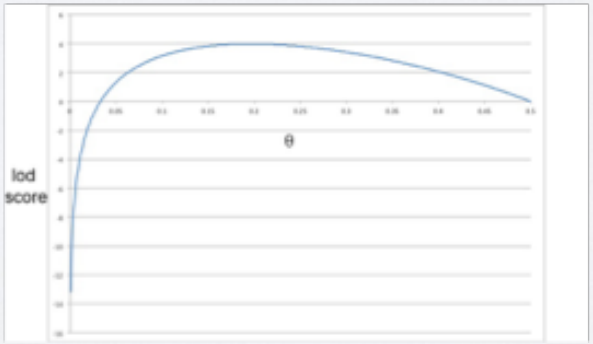
Family 2

Family 3

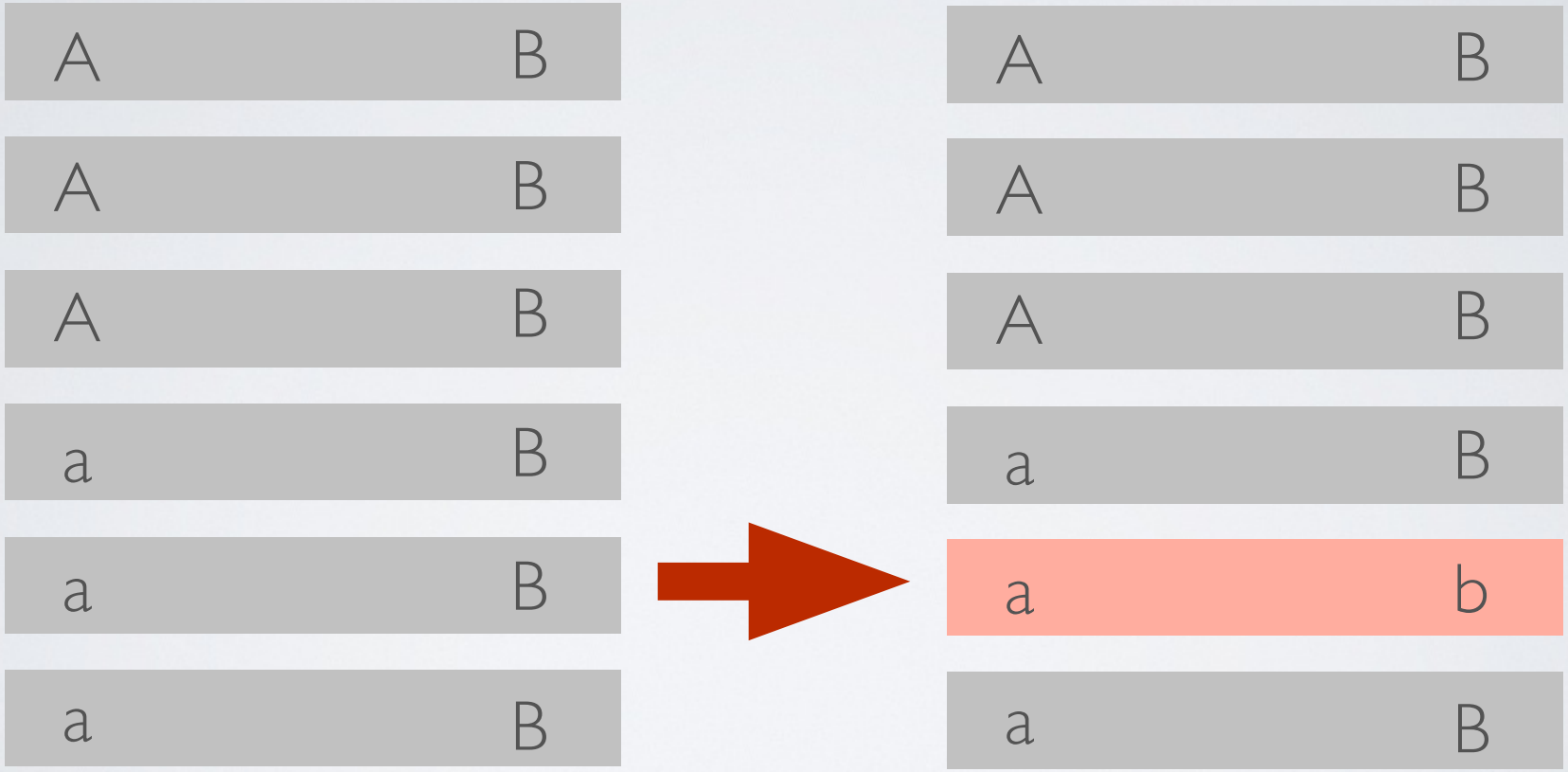
Family 4

Family 5

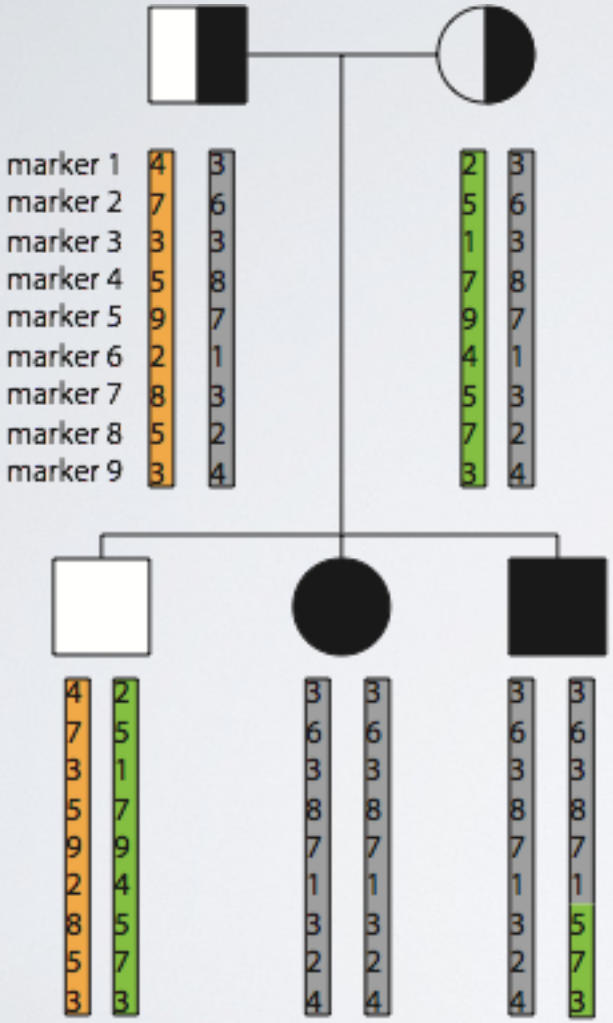
				θ				
Family	Sibs	Recombinants	Nonrecombinants	0	0.1	0.2	0.3	0.4
1	12	2	10	$-\infty$	1.15	1.25	1.02	0.60
2	9	2	7	$-\infty$	0.39	0.96	0.58	0.35
3	8	2	6	$-\infty$	0.13	0.43	0.43	0.28
4	10	2	8	$-\infty$	0.64	0.84	0.73	0.44
5	7	1	6	$-\infty$	0.83	0.83	0.65	0.38
Total	46	9	37	$-\infty$	3.14	4.31	3.41	2.05

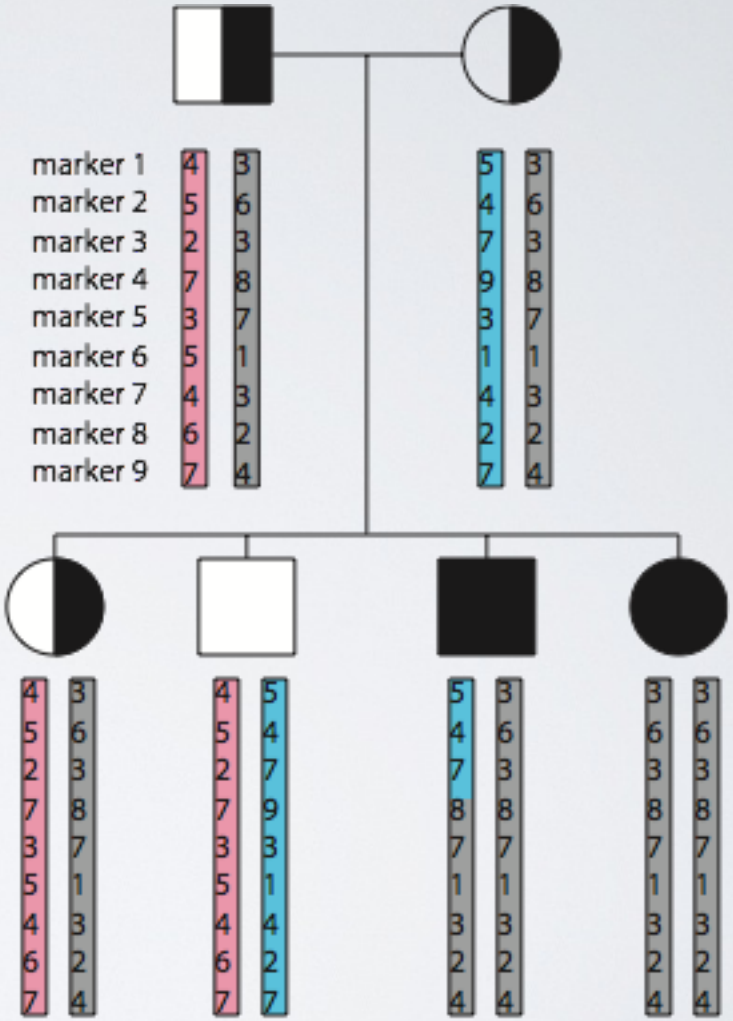


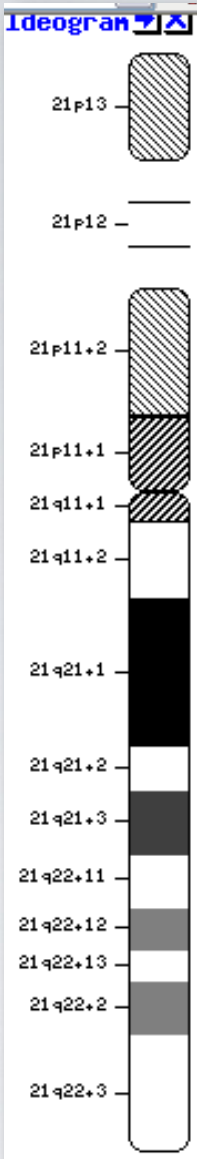
LINKAGE DISEQUILIBRIUM



HAPLOTYPE ANALYSIS

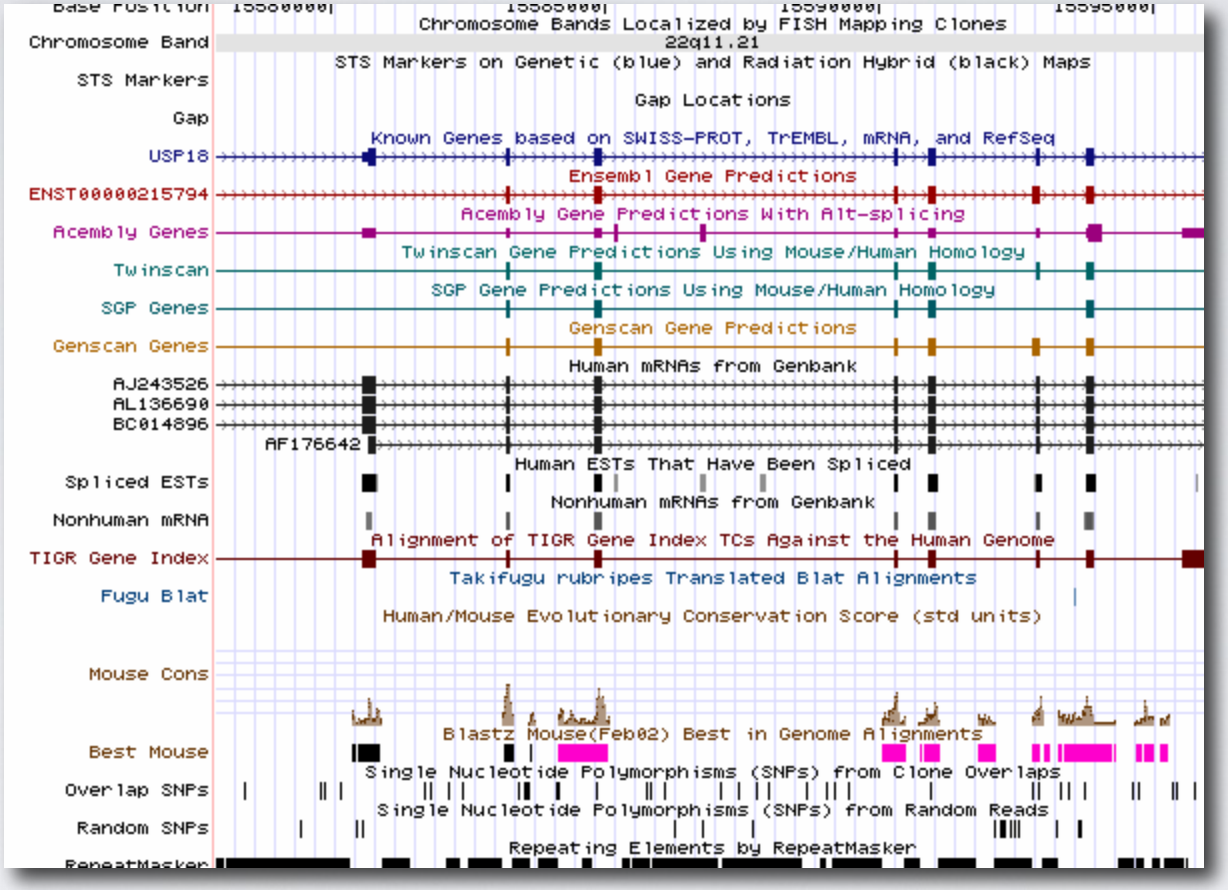






marker	cM(sex-average)	cM(female)	cM(male)	Heterozygo
GATA11C12	2.99	5.26	2.21	0.65
GATA31C01	9.72	17.19	2.21	0.83
AFM268xg1	13.05	24.06	2.21	0.49
AFMa116yg1	14.20	26.44	2.21	0.59
AFM294we1	17.67	30.05	5.42	0.81
AFM205wg9	20.45	33.37	7.56	0.65
D21S210	20.45	33.37	7.56	
GATA116E08	22.59	36.57	8.63	0.81
AFMa043ze5	25.80	42.98	8.72	0.37
AFM211zg9	27.40	44.05	11.03	0.74
UT706	28.48	44.35	11.84	
AFMa218wb1	31.26	47.29	15.09	0.86
AFMb280xd9	33.84	48.66	18.72	0.82
ATA27F01	36.77	54.84	18.72	0.74
UT409	38.65	54.84	22.36	
AFM238wc3	38.65	54.84	22.36	0.87
UT6930	40.47	55.81	24.84	
D21S231	45.55	60.79	30.46	
UT1355	51.49	64.59	38.61	
AFM071xa1	53.05	64.59	41.87	0.80

GENOME BROWSER

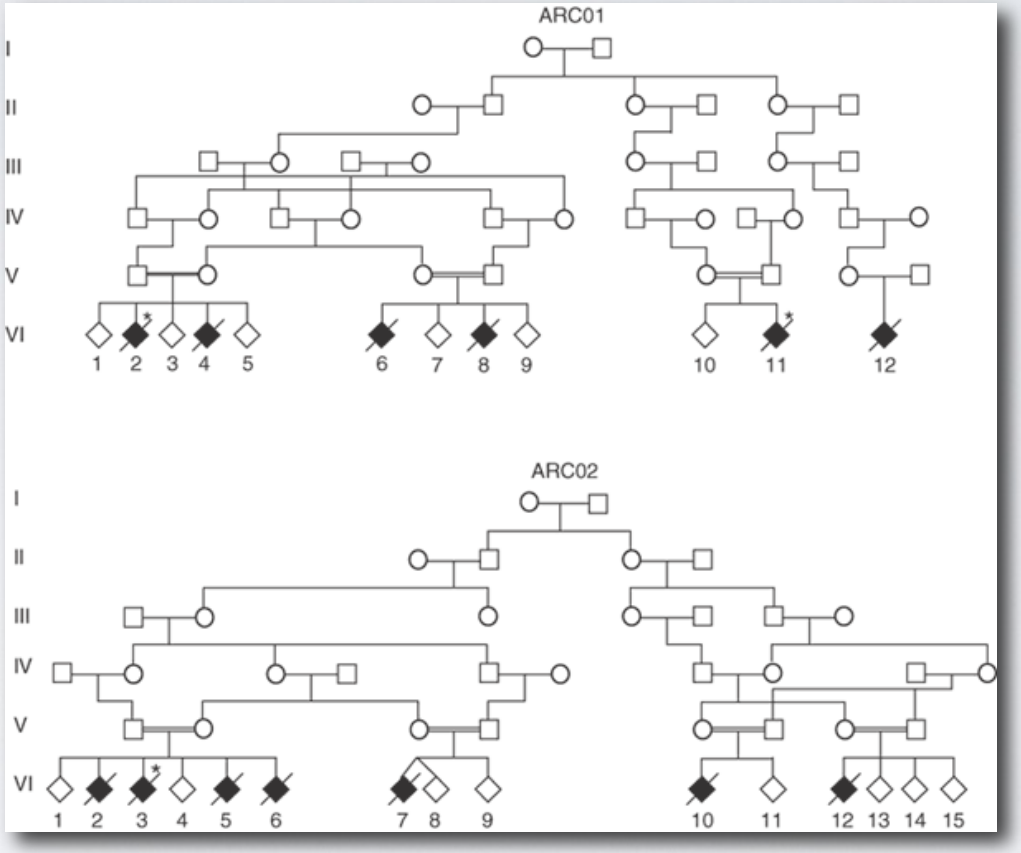


http://genome.ucsc.edu

Nature Genetics **36**, 400 - 404 (2004)

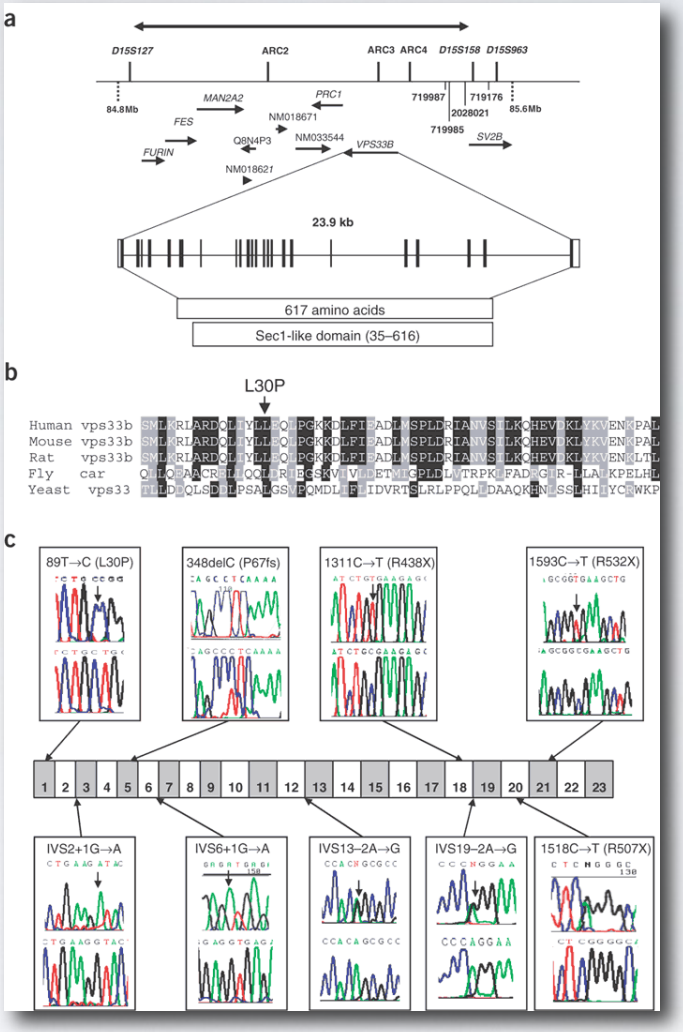
Mutations in *VPS33B*, encoding a regulator of SNARE-dependent membrane fusion, cause arthrogryposis–renal dysfunction–cholestasis (ARC) syndrome

Paul Gissen, et al.



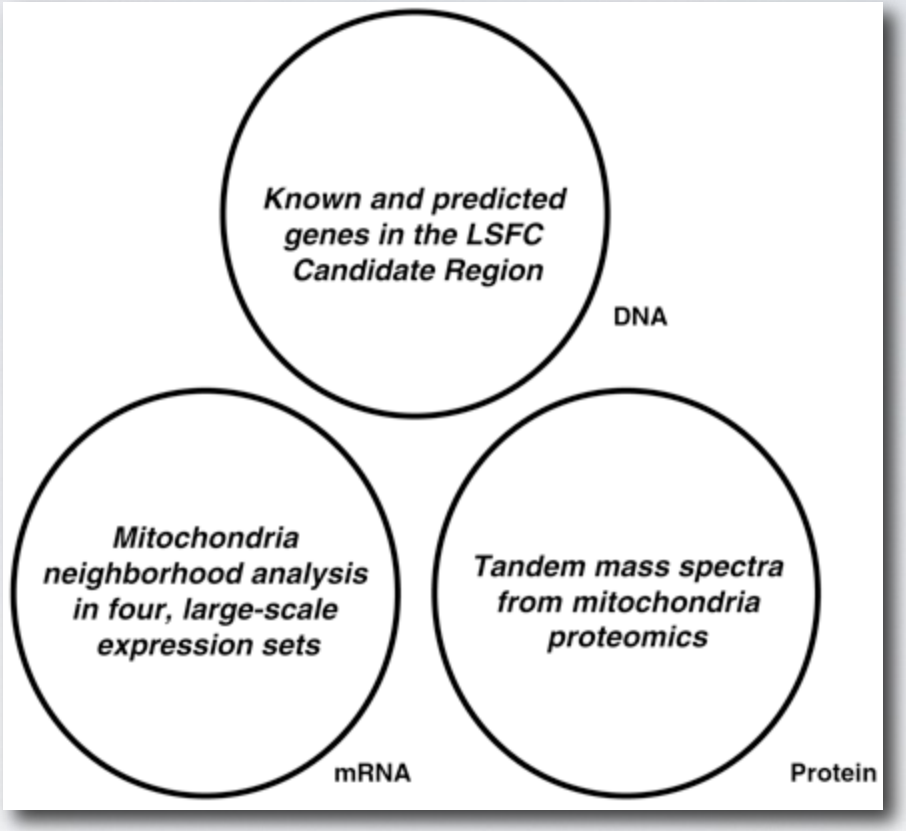
Nature Genetics **36**, 400 - 404 (2004)

CANDIDATE GENES



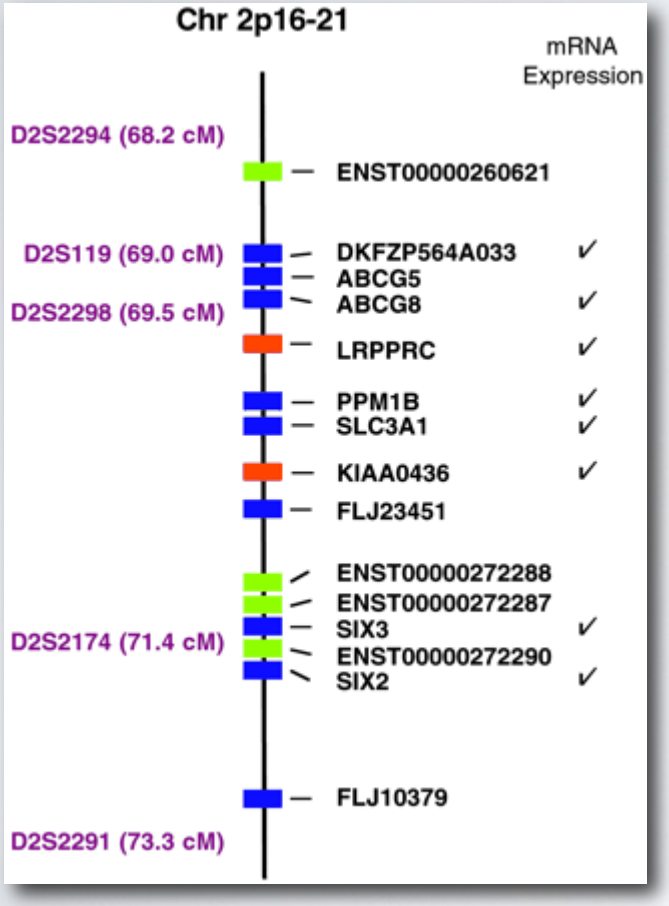
Nature Genetics **36**, 400 - 404 (2004)

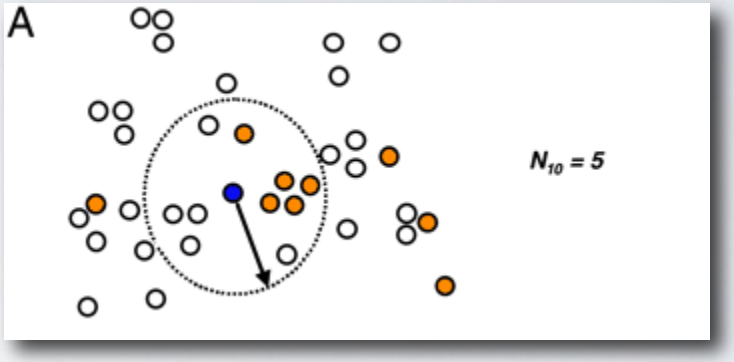
INTEGRATIVE GENOMICS

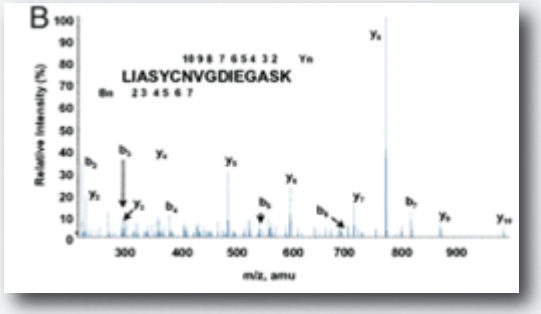


Proc. Natl. Acad. Sci. USA, Vol. 100, Issue 2, 605-610, January 21, 2003

MAPPING, GENOMICS, PROTEOMICS

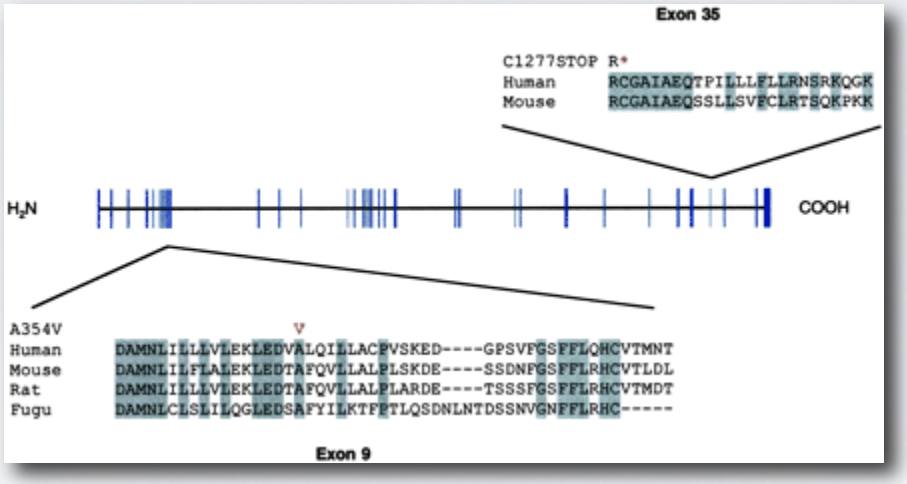






Proc. Natl. Acad. Sci. USA, Vol. 100, Issue 2, 605-610, January 21, 2003

LRPPRC MUTATIONS IN LEIGH DISEASE



Proc. Natl. Acad. Sci. USA, Vol. 100, Issue 2, 605-610, January 21, 2003

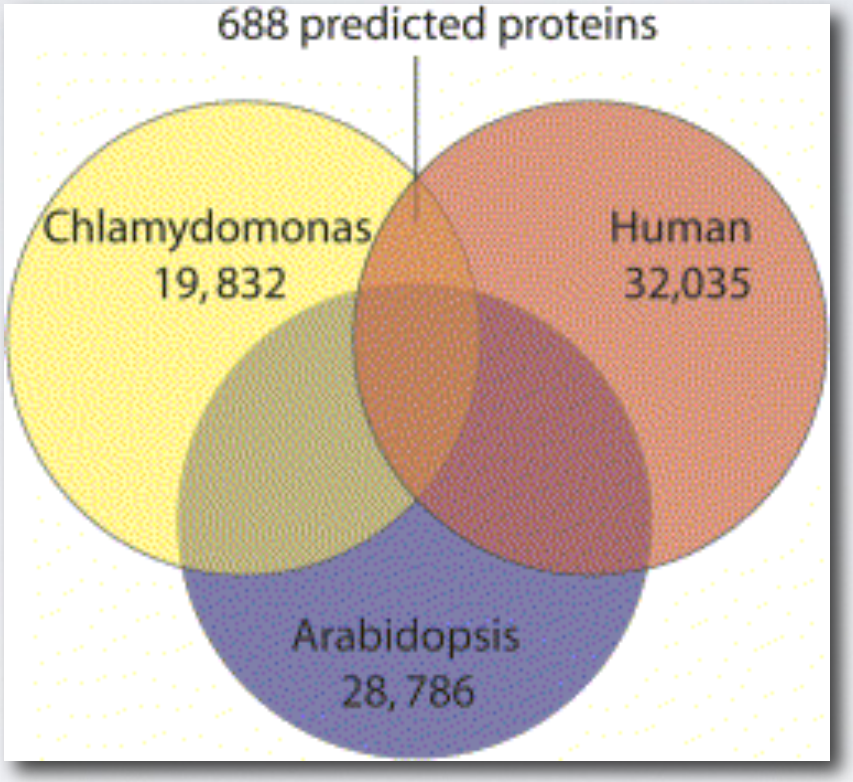
COMPARATIVE GENOMICS APPROACH

Comparative Genomics Identifies a Flagellar and Basal Body Proteome that Includes the *BBS5* Human Disease Gene

Jin Billy Li[1], Jantje M. Gerdes[2], Courtney J. Haycraft[4], Yanli Fan[5], Tanya M. Teslovich[2], Helen May-Simera[6], Haitao Li[7], Oliver E. Blacque[5], Linya Li[1], Carmen C. Leitch[2], Richard Allan Lewis[8], Jane S. Green[9], Patrick S. Parfrey[9], Michel R. Leroux[5], William S. Davidson[5], Philip L. Beales[6], Lisa M. Guay-Woodford[7], Bradley K. Yoder[4], Gary D. Stormo[1], Nicholas Katsanis[2, 3] and Susan K. Dutcher

Cell 117, Issue 4 , 14 May 2004, Pages 541-552

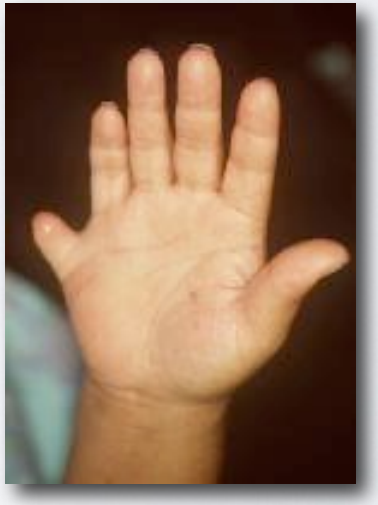
PROTEINS FOUND IN HUMAN AND CHLAMYDOMOMAS BUT NOT ARABIDOPSIS (FLAGELLAR PROTEINS)



BARDET-BIEDL SYNDROME

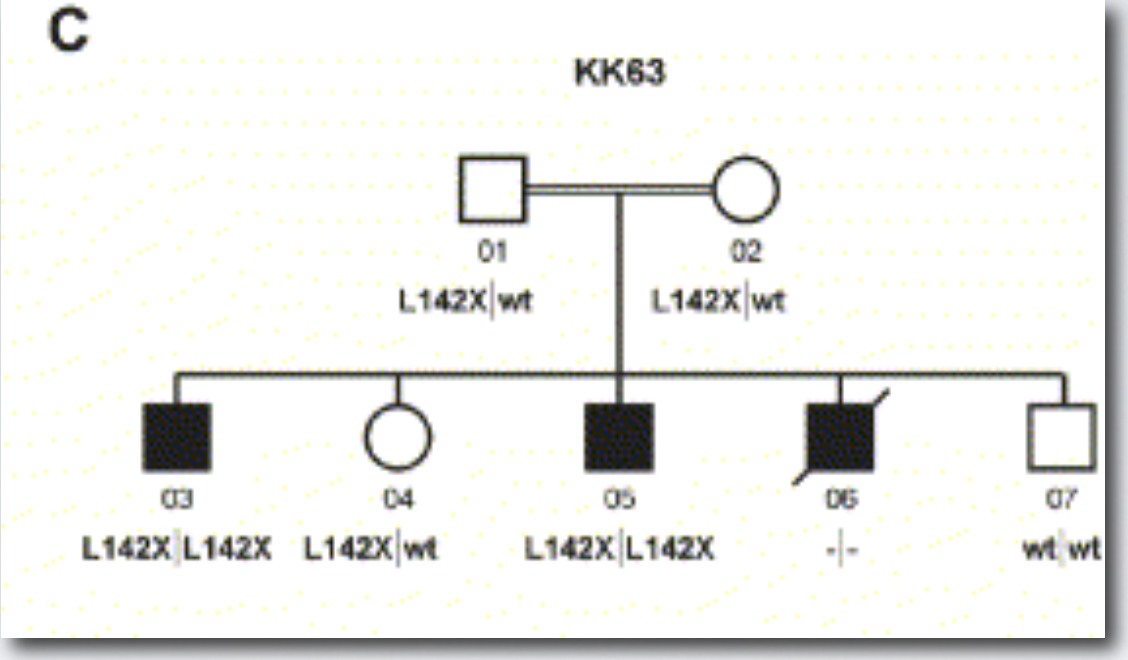
- Obesity
- Polydactyly
- Retinitis pigmentosum
- Renal failure
- Genetic heterogeneity
- Some forms associated with genes involved in cilia

www.makeachildsmile.org/ 2003/ prev_2003_dec.shtml



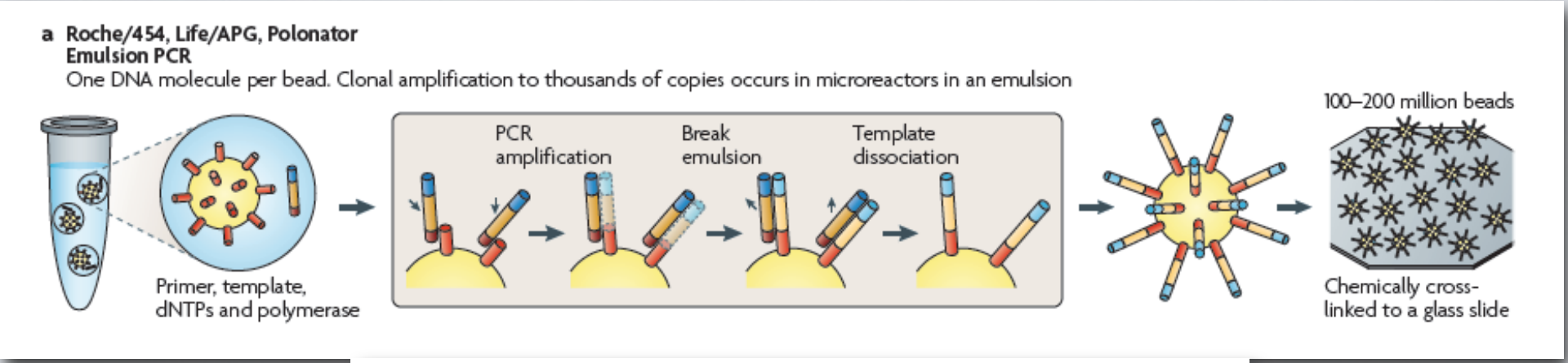
http://www.emedicine.com/oph/images/199OPH0704-04.jpg

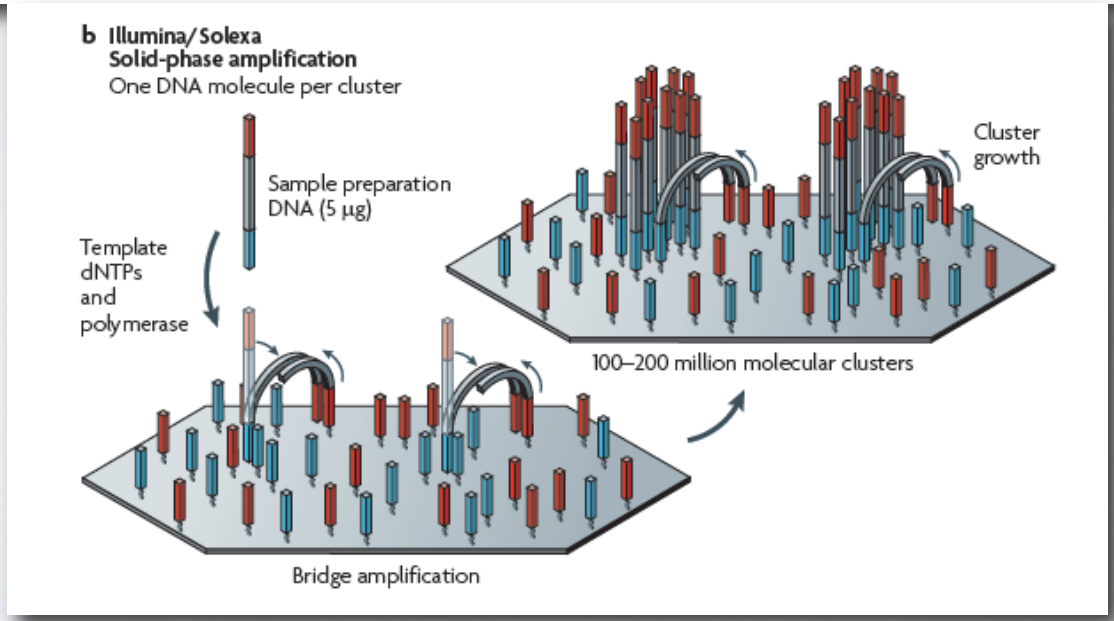
BBS5



2 proteins in database map to 2q31 where BBS5 known to map

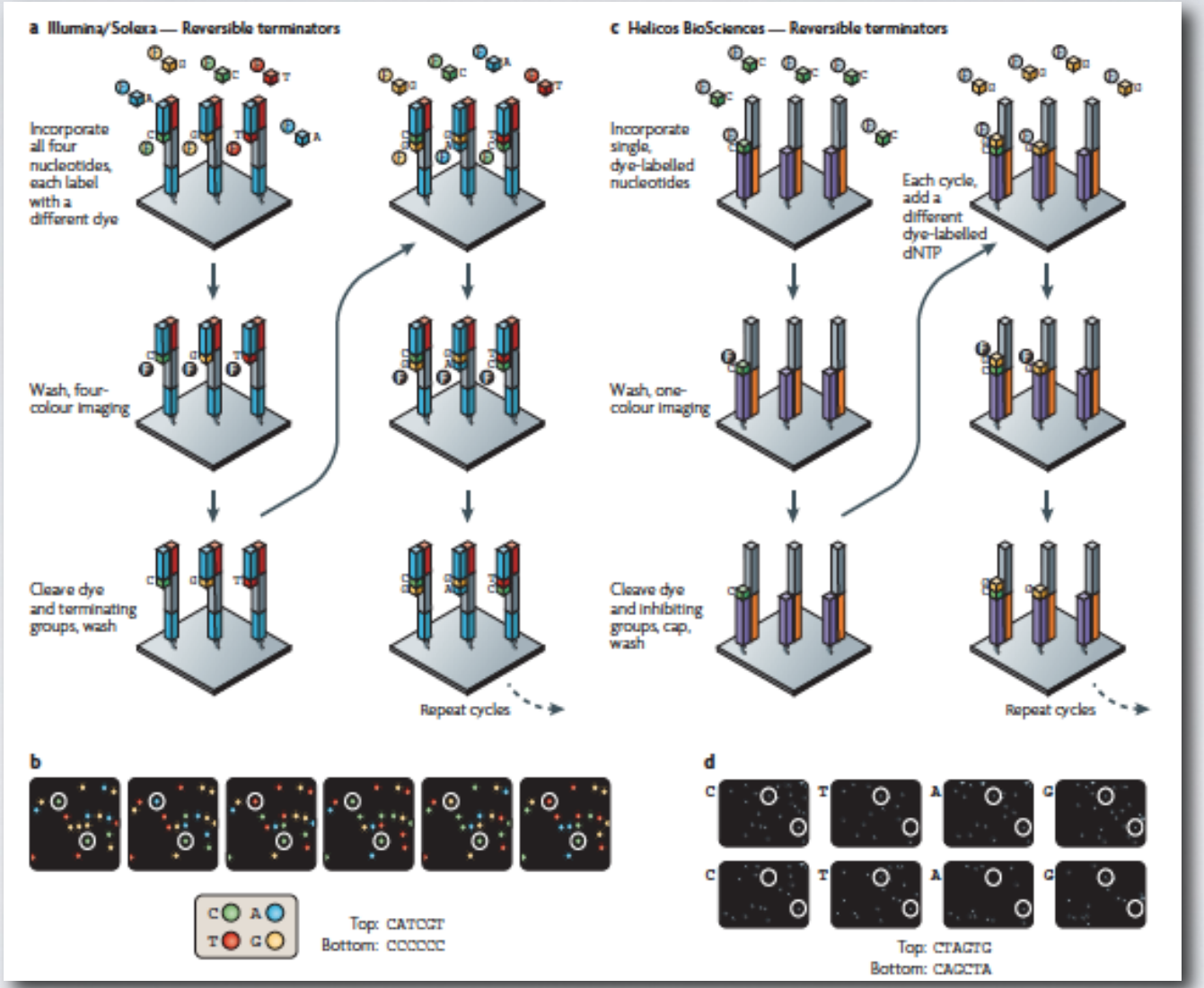
IMMOBILIZATION/ AMPLIFICATION





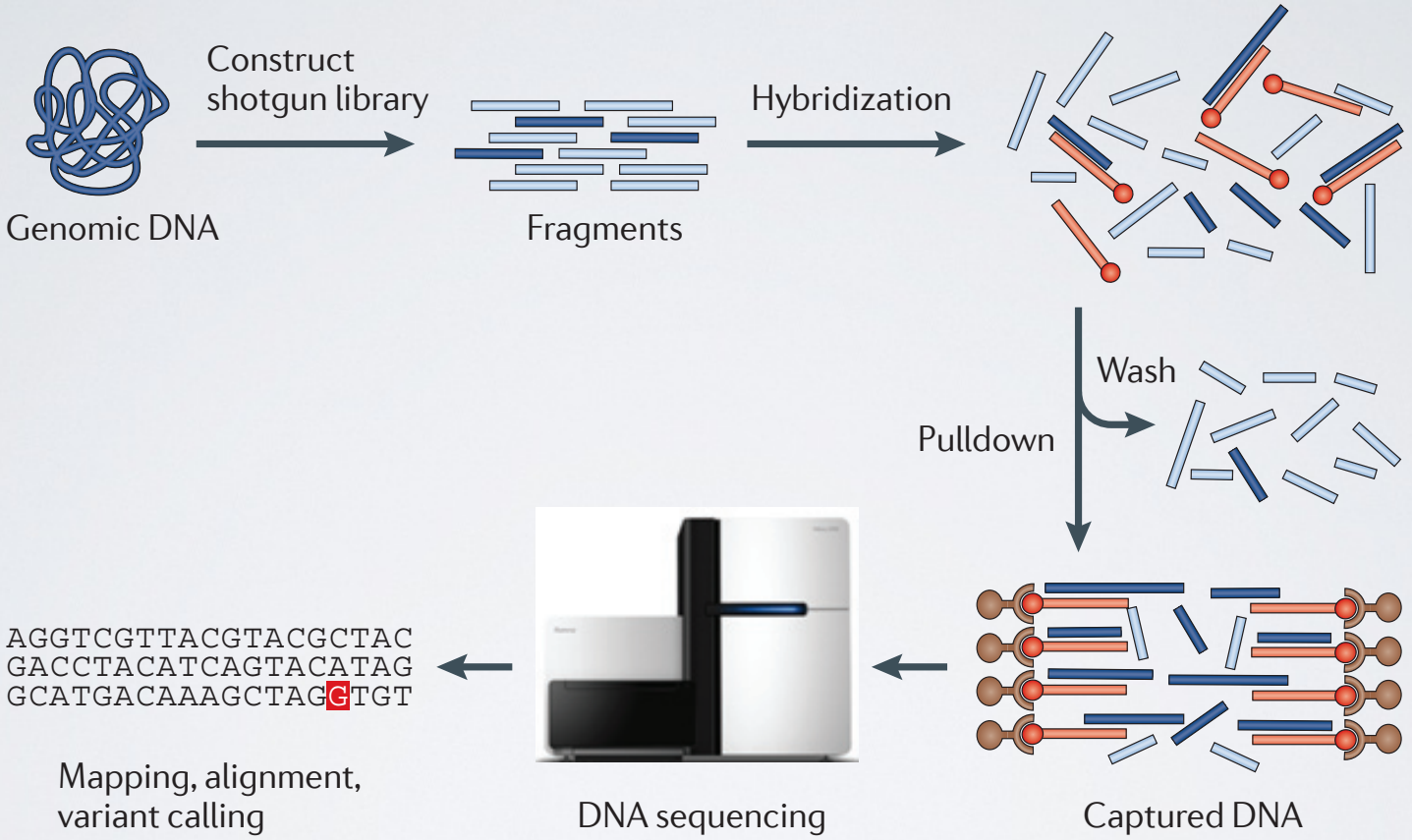
Metzker, M. Nature Reviews Genetics 2010;11:3146.

DNA SYNTHESIS/DETECTION



Metzker, M. Nature Reviews Genetics 2010;11:3146.

EXOME SEQUENCING



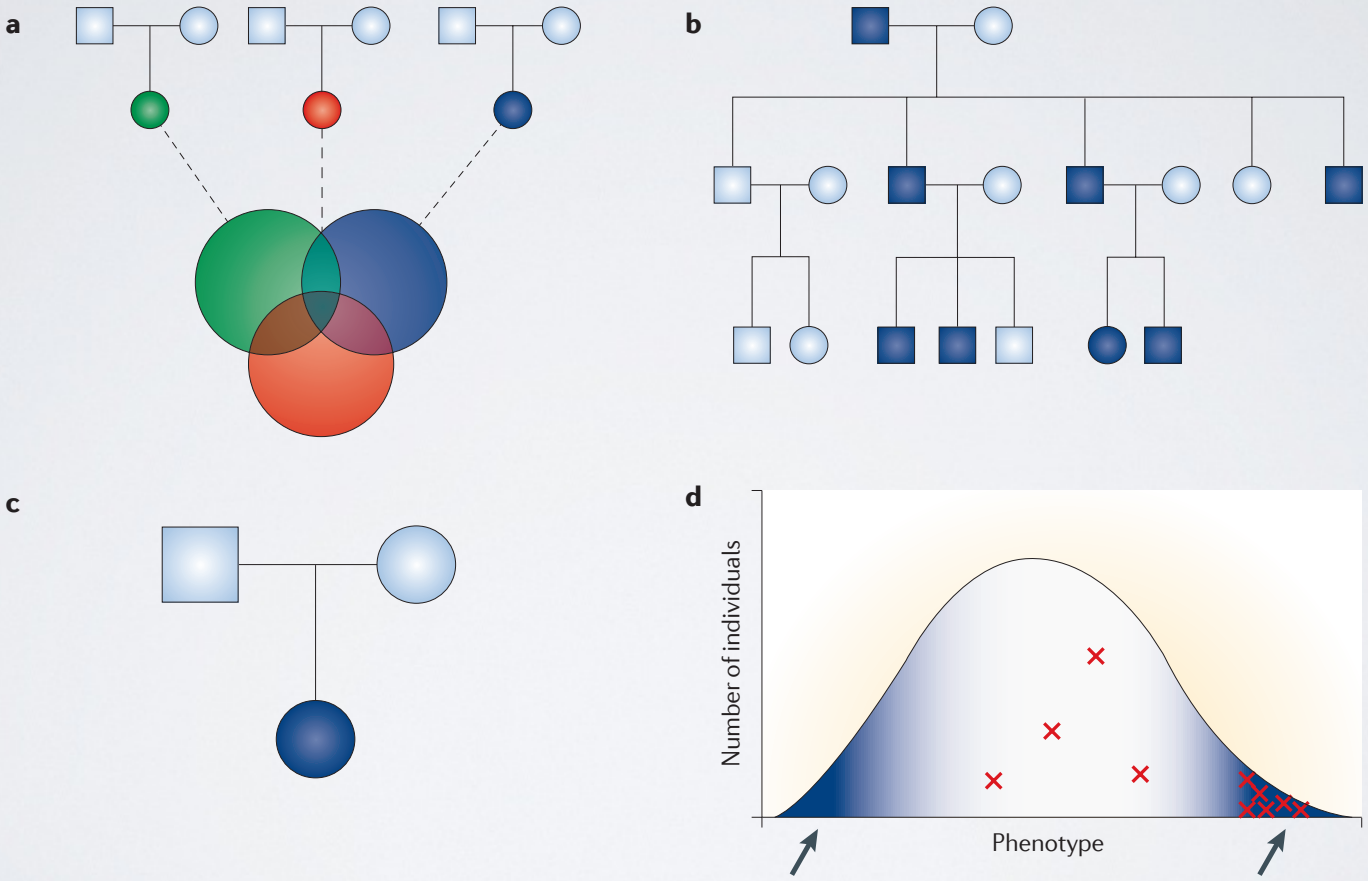
Bamshad et al. Nat. Rev. Genet. 2011; 12:745.

NUMBER OF VARIANTS

Variant type	Mean number of variants (± sd) in African Americans	Mean number of variants (± sd) in European Americans
Novel variants		
Missense	303 (± 32)	192 (± 21)
Nonsense	5 (± 2)	5 (± 2)
Synonymous	209 (± 26)	109 (± 16)
Splice	2 (± 1)	2 (± 1)
Total	520 (± 53)	307 (± 33)
Non-novel variants		
Missense	10,828 (± 342)	9,319 (± 233)
Nonsense	98 (± 8)	89 (± 6)
Synonymous	12,567 (± 416)	10,536 (± 280)
Splice	36 (± 4)	32 (± 3)
Total	23,529 (± 751)	19,976 (± 505)
Total variants		
Missense	11,131 (± 364)	9,511 (± 244)
Nonsense	103(± 8)	93 (± 6)
Synonymous	12,776 (± 434)	10,645 (± 286)
Splice	38 (± 5)	34 (± 4)
Total	24,049 (± 791)	20,283 (± 523)

Bamshad et al. Nat. Rev. Genet. 2011; 12:745.

STRATEGIES

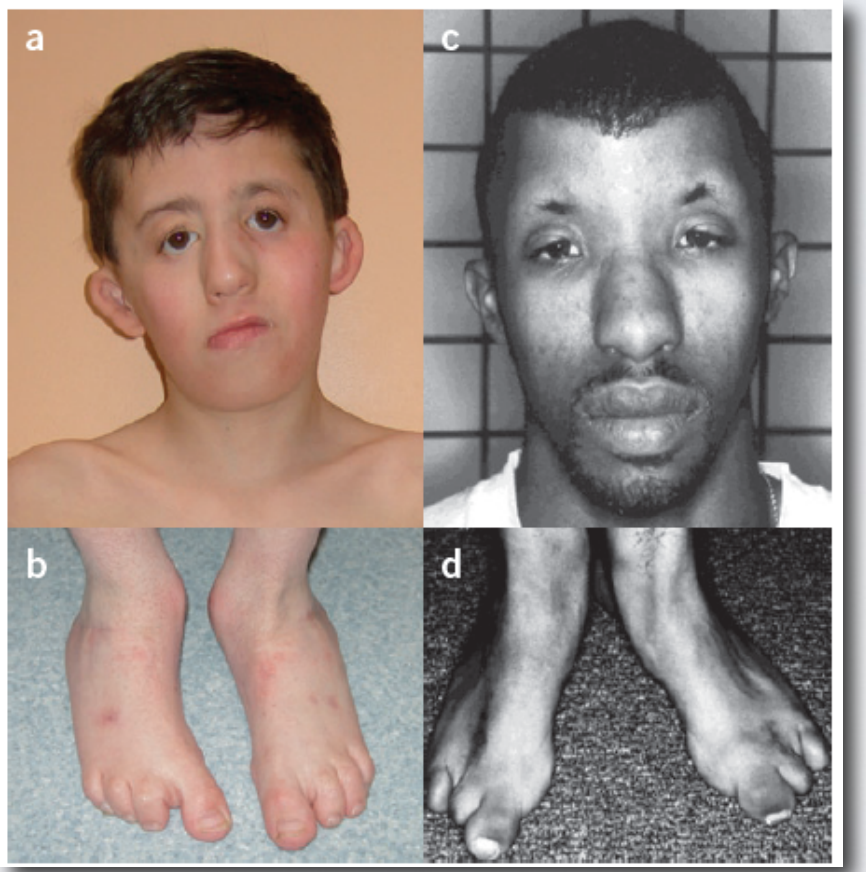


Bamshad et al. Nat. Rev. Genet. 2011; 12:745.

Exome sequencing identifies the cause of a mendelian disorder

Sarah B Ng[1,10], Kati J Buckingham[2,10], Choli Lee[1], Abigail W Bigham[2], Holly K Tabor[2,3], Karin M Dent[4], Chad D Huff[5], Paul T Shannon[6], Ethylin Wang Jabs[7,8], Deborah A Nickerson[1], Jay Shendure[1] & Michael J Bamshad[1,2,9]

Nature Genetics 2010;42:30-35.



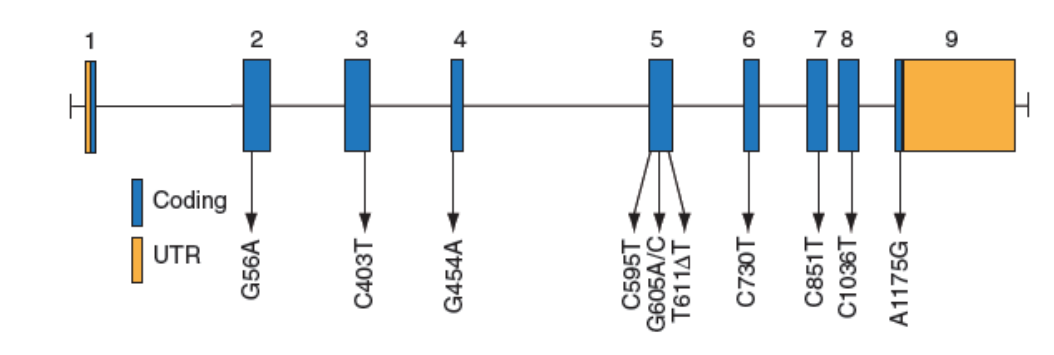
Miller syndrome

EXOME SEQUENCING

Table 1 Direct identification of the gene for a mendelian disorder by exome resequencing

	Kindred 1-A		Kindred 1-B		Kindred 1 (A+B)		Kindreds 1+2		Kindreds 1+2+3	
Filter	Dominant	Recessive	Dominant	Recessive	Dominant	Recessive	Dominant	Recessive	Dominant	Recessive
NS/SS/I	4,670	2,863	4,687	2,859	3,940	2,362	3,099	1,810	2,654	1,525
Not in dbSNP129	641	102	647	114	369	53	105	25	63	21
Not in HapMap 8	898	123	923	128	506	46	117	7	38	4
Not in either	456	31	464	33	228	9	26	1*	8	1*
Predicted damaging	204	6	204	12	83	1	5	0	2	0

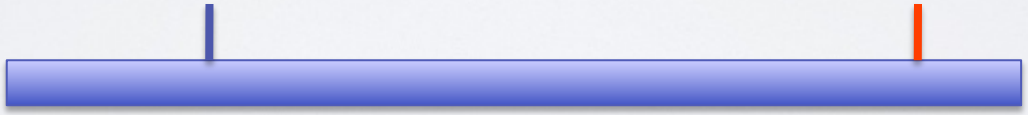
Each cell indicates the number of genes with nonsynonymous (NS) variants, splice acceptor and donor site mutations (SS) and coding indels (I). Filtering either by requiring the presence of NS/SS/I variants in siblings (kindred 1 (A+B)) or of multiple unrelated individuals (columns) or by excluding annotated variants (rows) identifies 26 and 8 candidate genes under a dominant model and only a single candidate gene, *DHODH*, under a recessive model (light gray cells). Exclusion of mutations predicted to be benign using PolyPhen (row 5) increases sensitivity under a dominant model but excludes *DHODH* under a recessive model because a variant in kindred 1 is predicted to be benign. A single candidate gene is identified in kindred 1 under a recessive model and excluding benign mutations (dark gray cell), but this candidate is excluded in comparisons with unrelated cases of Miller syndrome. Mutations in this candidate, *DNAH5*, were found to cause a primary ciliary dyskinesia in kindred 1. The asterisk indicates that a second gene, *CDC27*, was also identified as a candidate gene, but this is due to the presence of multiple copies of a processed pseudogene that recurrently gave rise to a false positive signal in exome analyses.



Nature Genetics 2010;42:30-35.

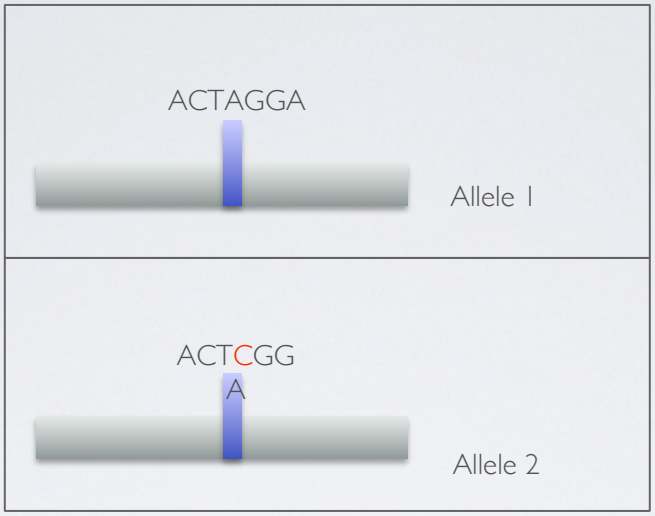
COMMON DISEASE – COMMON VARIANT HYPOTHESIS

Common diseases accounted for by genetic variants found in 1-5% of population



Linkage disequilibrium

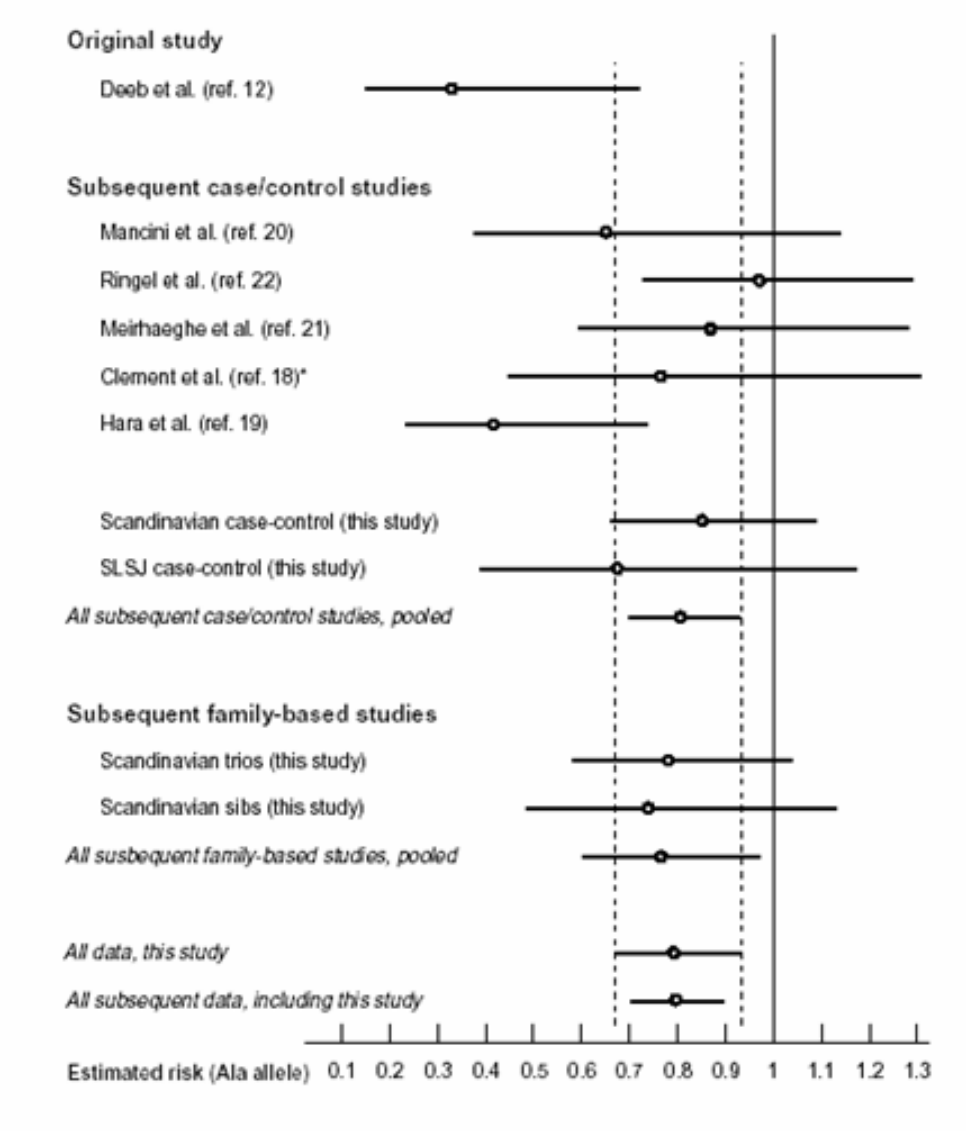
SNP ASSOCIATION



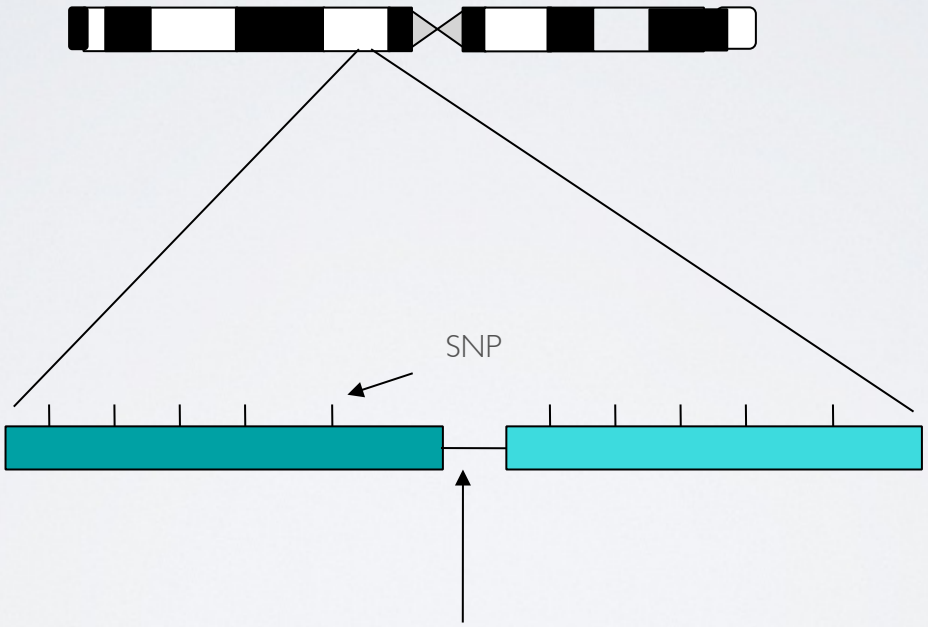
	Asthma	No Asthma
Allele 2 Present	30	10
Allele 2 Not Present	70	90

Hypothesis: Allele 2 is associated with an increased risk of asthma

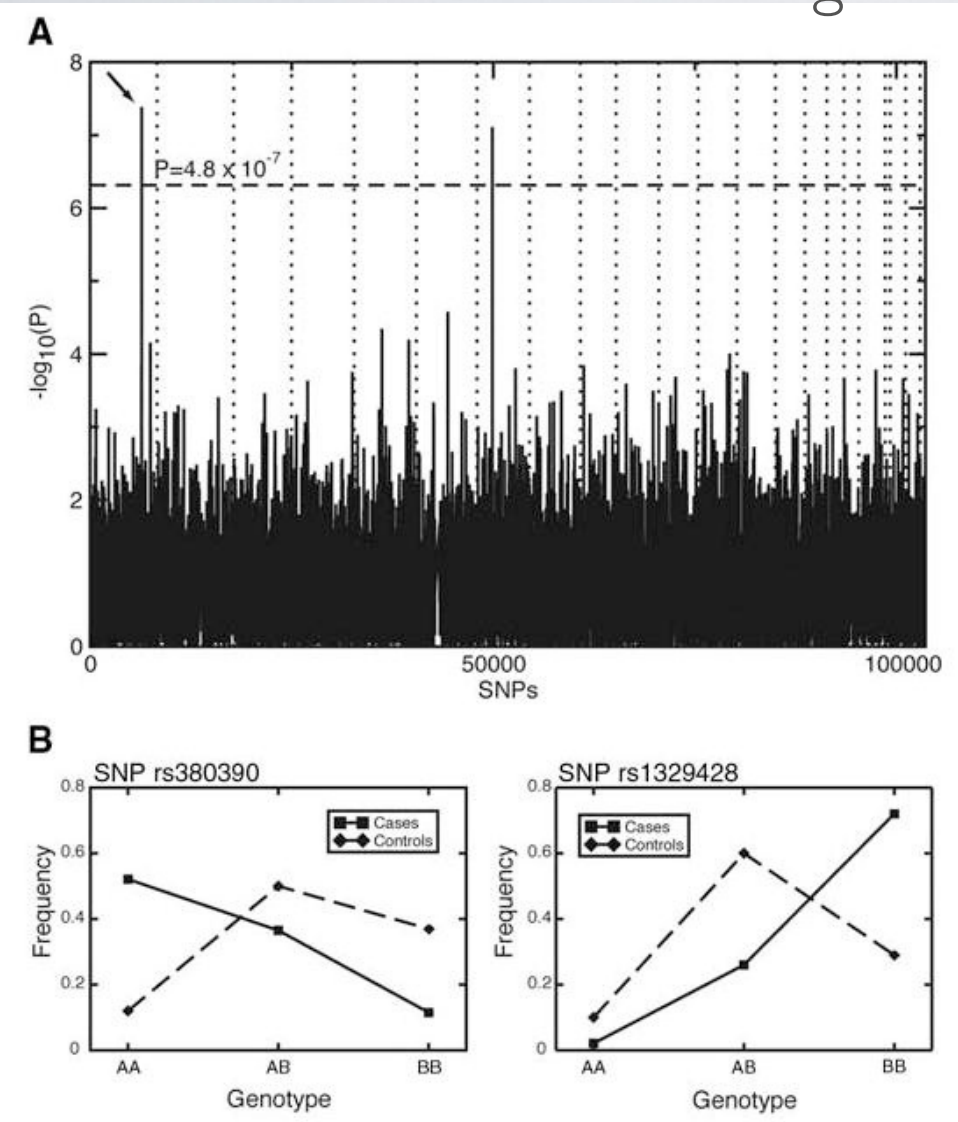
ASSOCIATION ANALYSIS

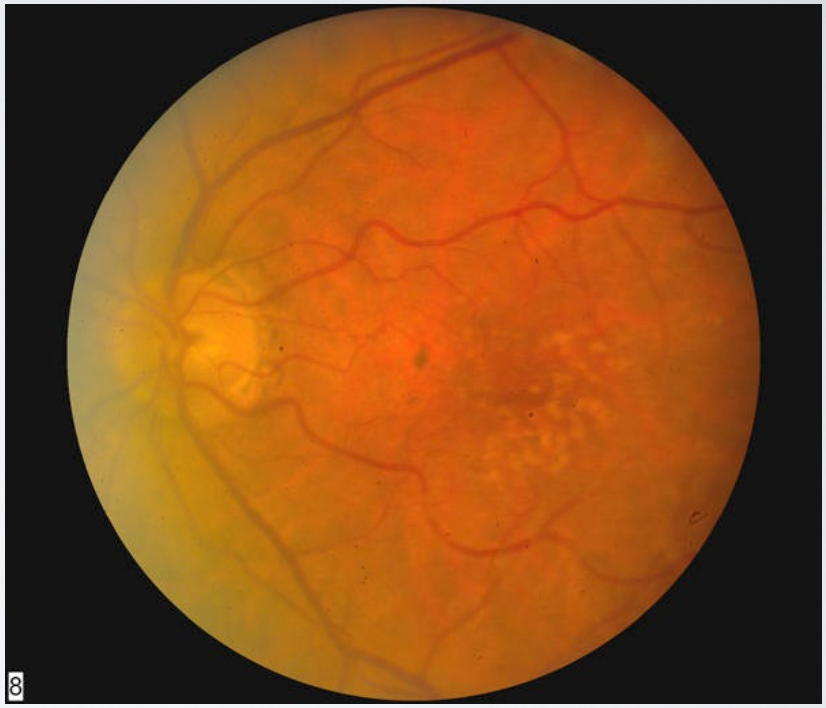



International
HapMap
Project


SNP
Recombination hot spot

Complement Factor H Polymorphism in Age-Related Macular Degeneration





http://www.medrounds.org/amd/uploaded_images/fig2-757825.JPG

Klein et al., Science (80-). 2005 April 15; 308(5720): 385–389.

ASSOCIATIONS FOR T2D

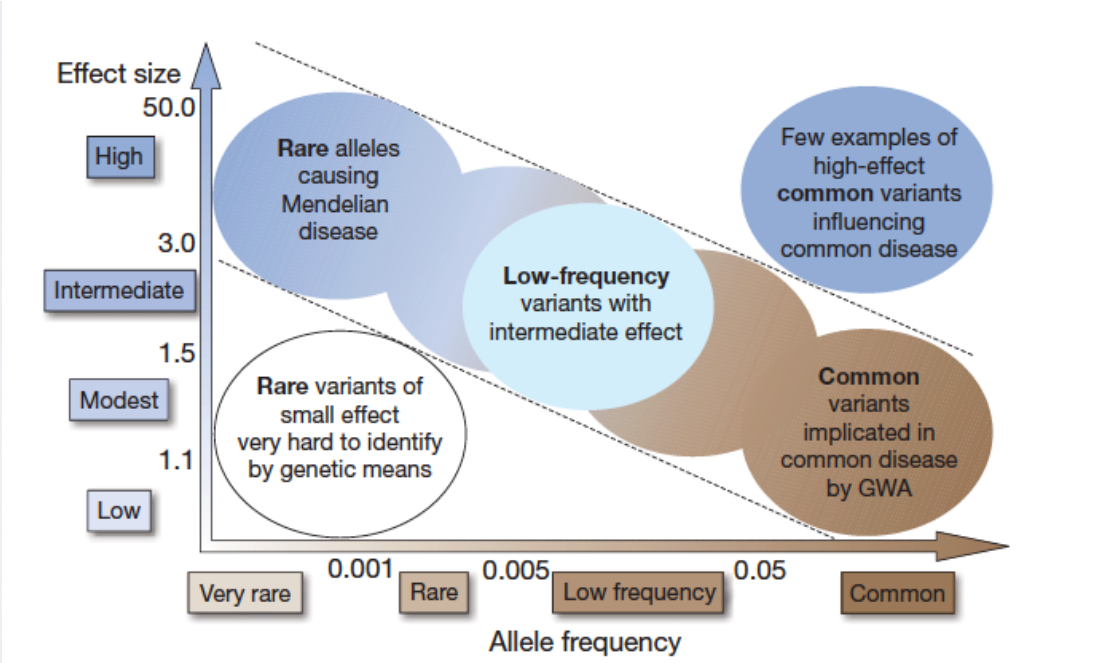
Gene	Symbol
Cyclin-dependent kinase inhibitor 2	*CDKN2A/B*
Insulin-like growth factor 2 mRNA binding protein	*IGF2BP2*
CDK5 regulatory subunit associated protein 1-like 1	*CDKAL1*
Hematopoietically-expressed homeobox	*HHEX*
S35-like splicing factor 30	*SCL30A8*
Transcription factor 7-like 2	*TCF7L2*
FATSO	*FTO*
Peroxisome proliferator-activated receptor gamma	*PPARG*
Potassium channel inwardly rectifying, subfamily J, member 11	*KCNJ11*

MISSING HERITABILITY

Table 1 | Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained	Heritability measure
Age-related macular degeneration[72]	5	50%	Sibling recurrence risk
Crohn's disease[21]	32	20%	Genetic risk (liability)
Systemic lupus erythematosus[73]	6	15%	Sibling recurrence risk
Type 2 diabetes[74]	18	6%	Sibling recurrence risk
HDL cholesterol[75]	7	5.2%	Residual* phenotypic variance
Height[15]	40	5%	Phenotypic variance
Early onset myocardial infarction[76]	9	2.8%	Phenotypic variance
Fasting glucose[77]	4	1.5%	Phenotypic variance

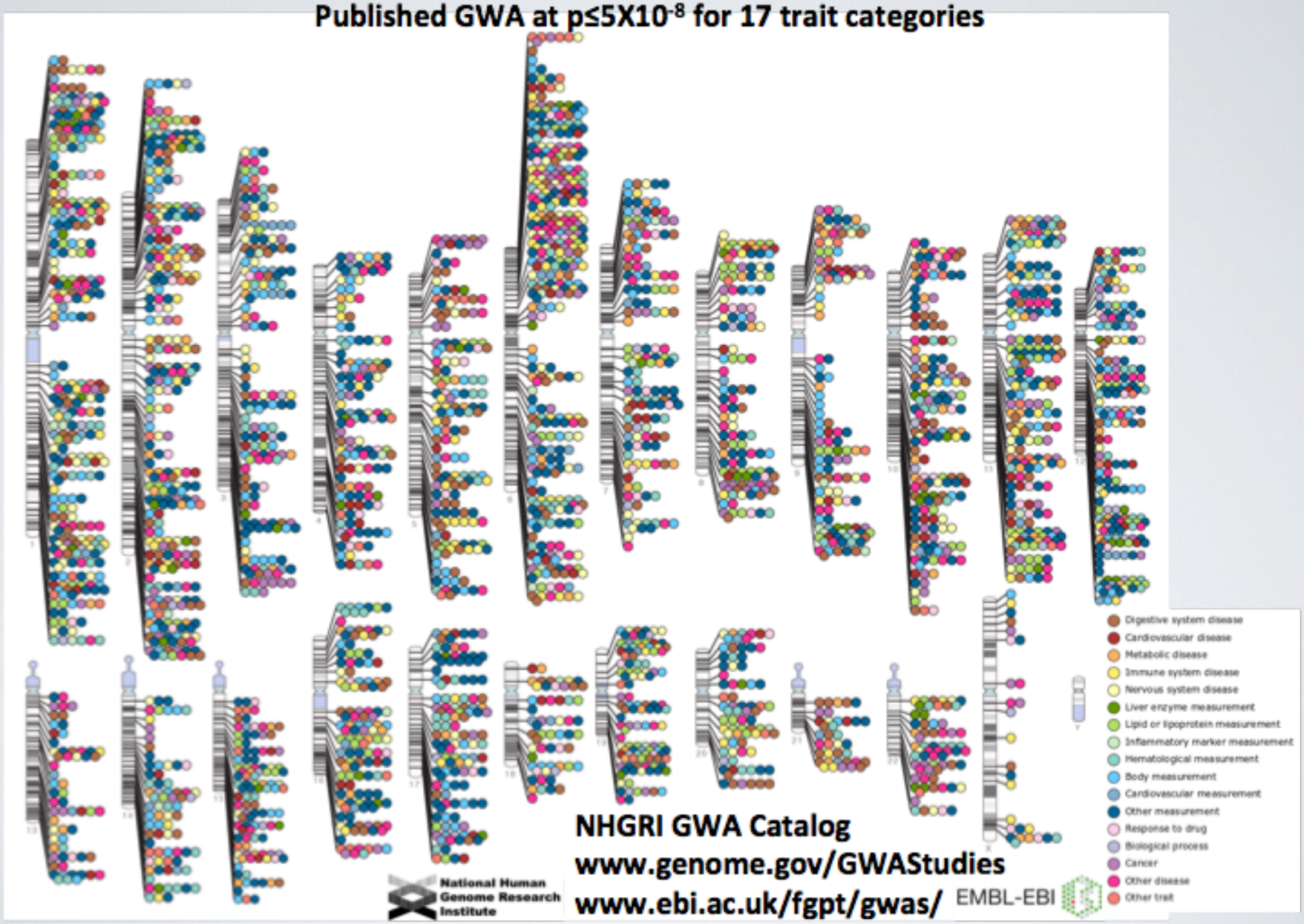
* Residual is after adjustment for age, gender, diabetes.



Manolio et al. Nature 461:747, 2009

Published Genome-Wide Associations through 12/2012

Published GWA at p≤5X10^{-8} for 17 trait categories



ODDS RATIO

	Asthma	No Asthma
Allele 2 Present	30	10
Allele 2 Not Present	70	90

Odds: ratio of probability event will happen over probability event will not happen – e.g., allele carrier getting asthma or not getting asthma

$$\text{odds of allele carrier having asthma } \frac{30/40}{10/40} = 3$$

$$\text{odds of non carrier having asthma } \frac{70/160}{90/160} = 0.77$$

$$\text{Odds ratio } \frac{3}{0.77} = 3.86$$

ODDS VS. ABSOLUTE RISK OF DISEASE

$$\text{52\% increase in odds} = 1.52 = \frac{3.5/1000}{2.3/1000}$$

CALCULATION OF RISK

Risk = (Population risk)(odds ratio)

Example:
Population risk is 3.5%
Odds ratio is 1.2
Risk = (3.5)(1.2) = 4.2%

UAB AND HUDSONALPHA

- comprehensive academic medical center
- 1300 faculty
- genetic testing and diagnosis
- expertise in health care and disease biology

- biotech research and development
- <20 faculty
- high-throughput next generation sequencing
- expertise in genomics and bioinformatics