Analytical and Clinical Validation Standards

- Selection of SNP genotyping platform in a CLIA regulated laboratory with high accuracy
  - The DNA chips used covering 500,000 to 1M SNP markers have 4 to 20 fold redundancy for each SNP measured resulting in high accuracy defined by proficiency testing
- SNPs chosen for annotation of risk must be replicated in multiple powered studies and their OR derived from large datasets (typically thousands of patients and controls)
- All three companies use methodologies the convert from the reported allelic OR (or genotype-specific OR) to risk compared to the general population
- All three companies assume a multiplicative model for both the allelic risk at each marker and when combining markers to define overall risk unless there are data supporting a better model

Example: all 3 companies include 9p21 variants for MI/CHD - the only region to show significant association in the 4 GWA studies published to date

-deCODE Study (*Science*, May 2007)

(rs10757278 marker)

-5 populations (4589 pts vs 12,768 controls)

-Ottawa Heart/ US Study (*Science*, May 2007)

(rs10757274 : correlation with rs10757278 r2=0.86)

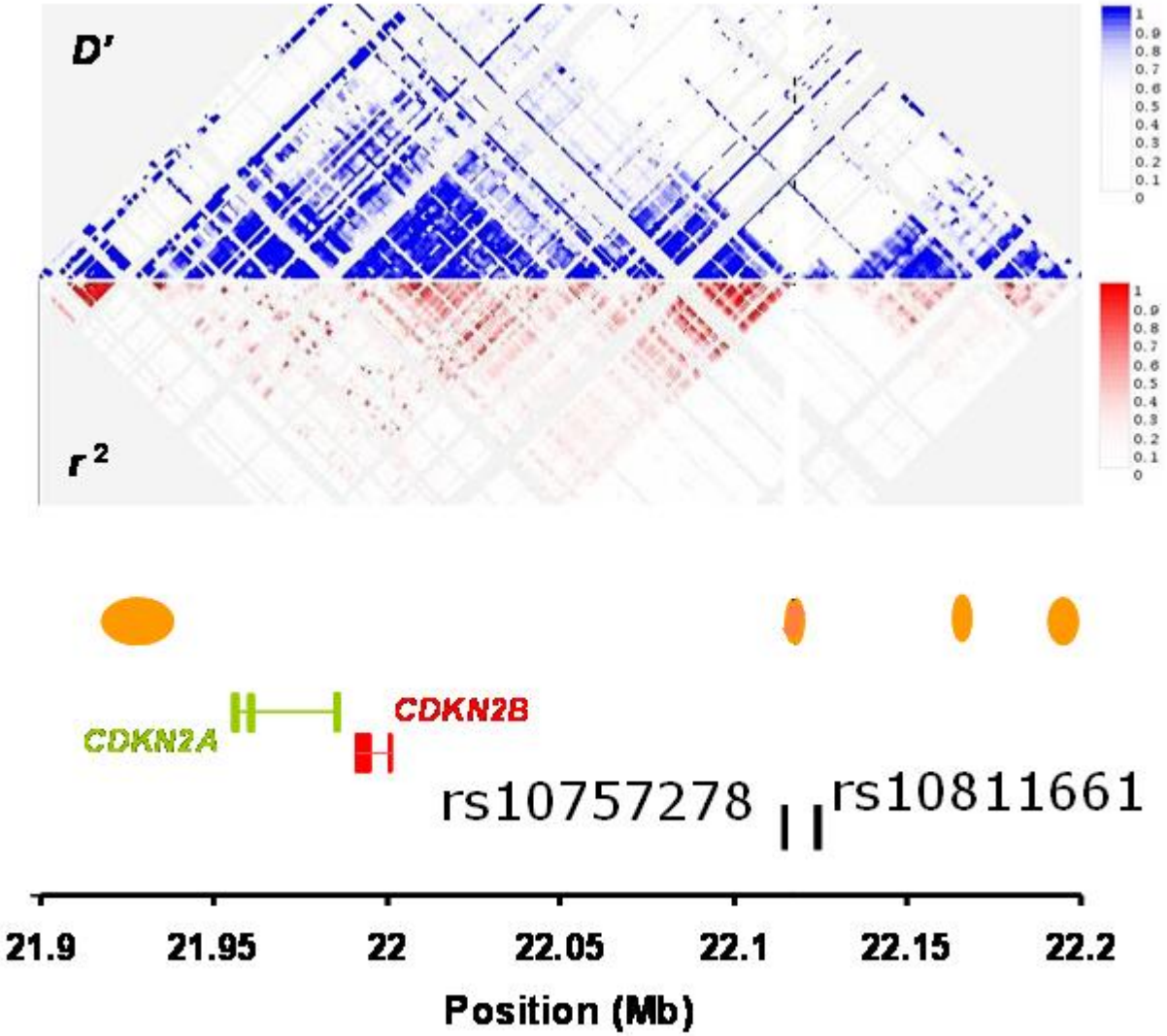
–6 populations including the prospective ARIC study (3500 pts vs 12,500 controls)

-UK and German MI study (*NEJM*, July 2007)

(rs1333049: correlation with rs10757278 r2=1)

(2801 pts vs 4582 controls)

All markers cluster within a single LD block



LD structure

Recombination hotspots

Variants associated with CAD and T2D

9p21 has been widely replicated in Caucasian and East Asian populations

- 21% of population are homozygous for variant and have 1.6 fold risk compared to non-carriers; 2.0 fold for early MI
  - Similar in magnitude and frequency to LDL cholesterol risk

- The 9p21 association has now been replicated in 25 Caucasian and 5 Asian populations (no effect in African populations)

- Replicated in over 30,000 patients and 60,000 controls, including several prospective studies

- Independent of known risk factors including family history, LDL, TG, hypertension, diabetes, obesity, smoking, and CRP

Conversion to risk relative to the general population

All 3 companies convert the allelic OR to risk relative to the general population

All 3 companies normalize OR by dividing by the total risk in the population (2 convert to relative risk before combining markers and one converts after combining marker ORs – generally with a multiplicative model)

9p21 variant with allelic OR of 1.28 (assuming multiplicative model and population controls):

Risk of heterozygote to non-carrier is 1.28

Risk of homozygote risk compared to non-carrier is 1.64

Total risk in population is:

0.21 X 1.64 + 0.53 X 1.28 + 0.26 X 1.0 = 1.28

Risk of double carrier GG is 1.64/1.28 or 1.3 relative to general population (1.6 for early MI)

Risk of GA is 1.0

Risk of noncarriers (AA) is 0.8

Addition of 9p21 variant to ARIC and NPHS prospective cohorts led to significant increase in accuracy of MI prediction

18% of patients in intermediate and intermediate-high categories are reclassified – change in LDL-C target

			Classification using ACRS + 9p21 allele				
			0–5%(%*)	5–10%(%*)	10–20%(%*)	>20%(%*)	
Category							
Total number reclassified for category (%*)							
10-year risk 0–5%	**Low**	3,428	3,237	191 (5.6)	0	0	191 (5.6)
Observed event rate†			*2.3*	*3.9*	*0*	*0*	*2.4*
10-year risk 5–10%	**Intermediate**	2,328	165 (7.1)	1,878	285 (12.2)	0	450 (19.3)
Observed event rate			*4.98*	*6.1*	*10.6*	*0*	*6.7*
10-year risk 10–20%	**Intermediate-high**	2,641	0	184 (7)	2,194	263 (10)	447 (17)
Observed event rate			*0*	*9.3*	*12.6*	*16.2*	*12.76*
10-year risk >20%	**High**	1,607	0	0	135 (8.4)	1,472	135 (8.4)
Observed event rate					*13.7*	*22.61*	*21.86*
	TOTAL	10,004	3,402	2,253	2,614	1,735	
Observed event rate		*1349*	*2.5*	*6.2*	*12.5*	*22*	*9.2*

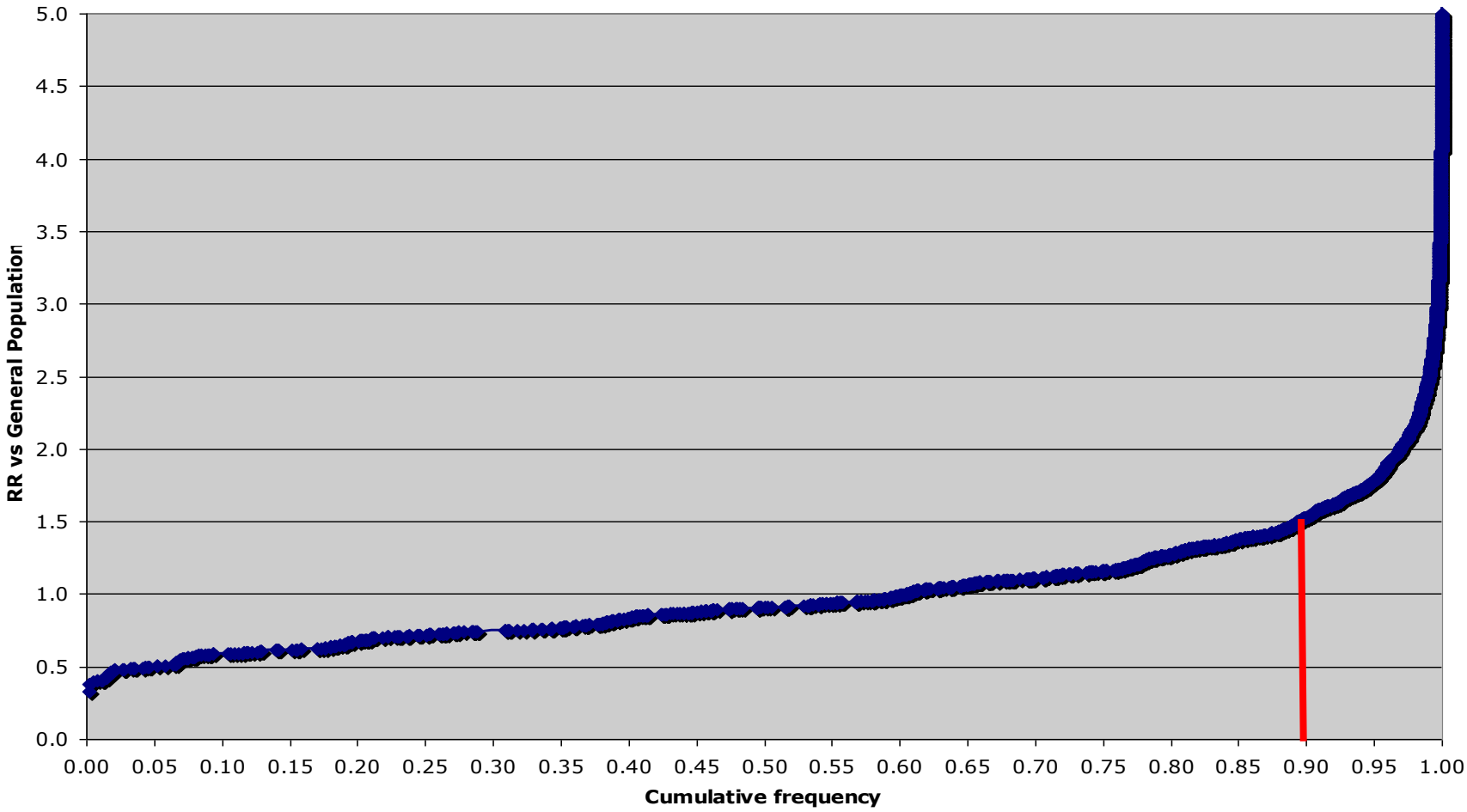
*** Percentage of individuals reclassified from ACRS based risk model after adding 9p21 allele to risk calculation. † Observed event rate have been extrapolated to 10-year rate (number of events per 100 people per 10 years of observation) from a follow up time of 14.6 years. Conclusion:** The addition of the 9p21 allele to traditional risk factors, in the white population of the ARIC study, improved CHD risk prediction and reclassified a number of subjects, especially in the intermediate and intermediate-high risk categories. For the majority of the reclassified individuals, target LDL-C levels would be changed, thus altering therapy

Example: Large datasets support use of multiplicative model for these independent risk factors for prostate cancer

Locus	Chromo-some	Variant / SNP	My Codes	Relative Risk	Genotype frequency	#Cases / #Controls
	2	rs2710646	AA	1.25	3.60%	10000 / 29000
	8	rs10505483	GG	0.96	93.90%	2600 / 5500
	8	rs1447295	CC	0.91	82.40%	2000 / 5000
	8	rs6983267	GG	1.25	25.00%	4300 / 4300
	11	rs10896449	GG	1.19	27.00%	5000 / 5000
TCF2	17	rs4430796	AA	1.21	23.80%	3500 / 14000
	17	rs1859962	GT	1.01	49.70%	3500 / 14000
	X	rs5945572	A	1.14	35.00%	10000 / 29000

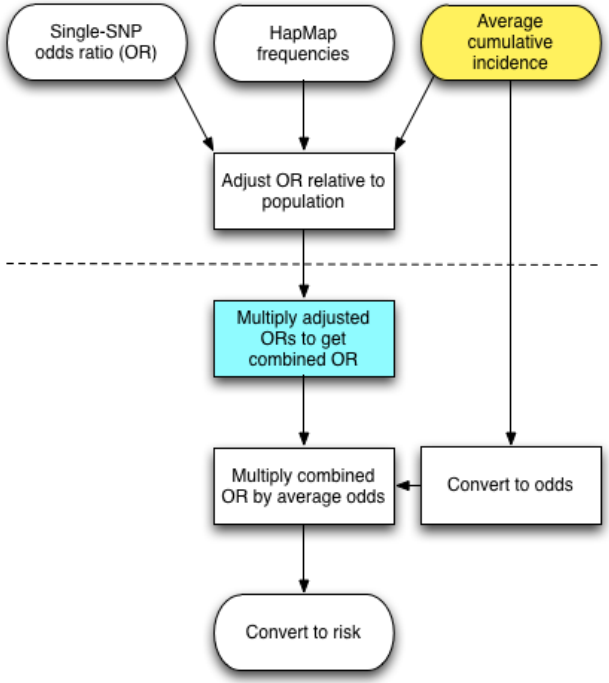
Total relative risk for this patient =
1.25 X 0.96 X 0.91 X 1.25 X 1.19 X 1.21 X 1.01 X 1.14 = 2.01

8 validated genetic markers define prostate cancer risk ranging from 0.4 to 5 fold

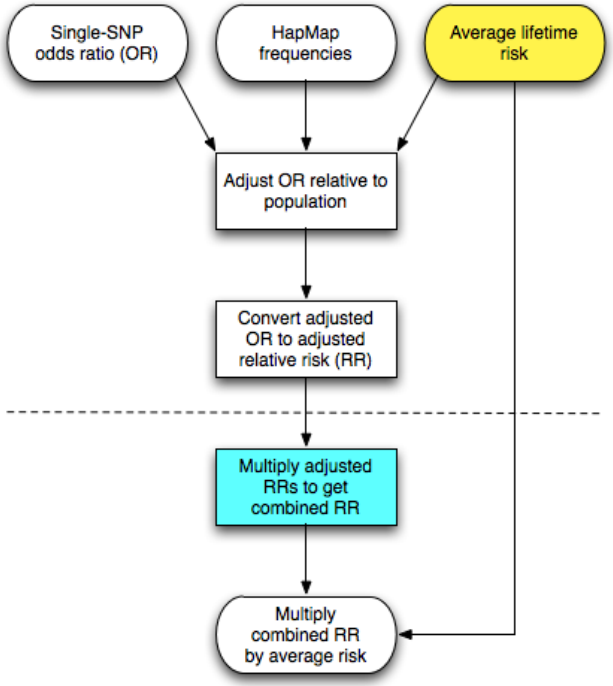


Calculating genotype-specific risk

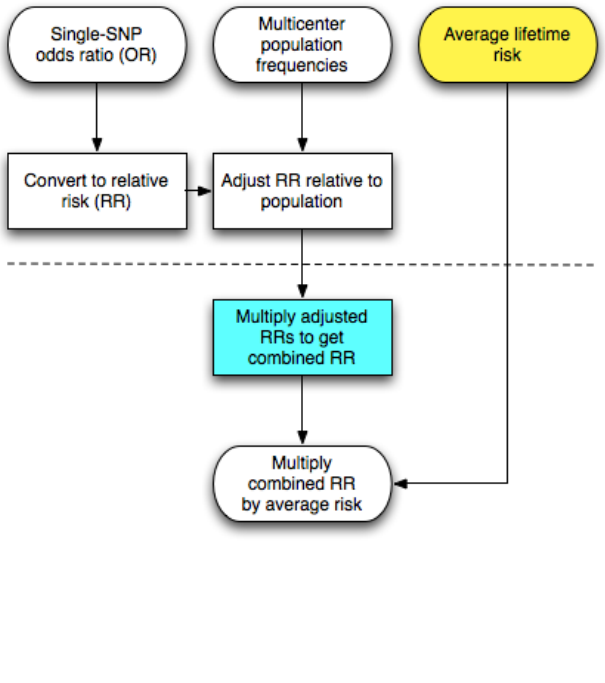
23andMe



Navigenics



deCODE



Baseline Epidemiological Data Comparison -conditions with similar numbers

	NAVIGENICS		23andMe	deCODE	
Condition	**male LTR**	**female LTR**	**cum. incidence (Euro)**	**male LTR (Euro)**	**female LTR (Euro)**
Alzheimer's Disease	0.0910	0.1720	NA	0.06	0.12
atrial fibrillation	0.2600	0.2300	NA	0.25	0.25
Breast Cancer	NA	0.1325	0.162	NA	0.12
Colorectal cancer	0.0579	0.0534	0.087	0.06	0.06
Crohn's Disease	0.0058	0.0054	0.0043	0.005	0.005
Diabetes, Type 2	0.2537	0.2964	0.219	0.25	0.28
Lung Cancer	0.0809	0.0647	0.073	smokers = 0.172, non-smokers 0.013	smokers = 0.116, non-smokers = 0.014
Multiple Sclerosis	0.0030	0.0077	0.0052	0.0023	0.0053
Prostate Cancer	0.1658	NA	0.178	0.16	NA
rheumatoid arthritis	0.0156	0.0334	0.042	0.01	0.01
systemic lupus erythamatosus	0.0003	0.0026	0.0025	NA	NA

Baseline Epidemiological Data Comparison -conditions with dissimilar numbers

	NAVIGENICS		23andMe		deCODE		
Condition	**male LTR**	**female LTR**	**cum. incidence (Euro)**	**age range**	**male LTR (Euro)**	**female LTR (Euro)**	**reason**
abdominal aneurysm	0.0305	0.0146	NA	NA	0.17	0.05	Ruptured only vs ruptured plus unruptured
Age Related Macular Degeneration	0.0310	0.0310	0.07	40-79	0.08	0.08	case definition, methodology
Body Mass Index, obesity endpoint (BMI≥30kg/m²)	0.3380	0.3240	0.575	17-59	0.395	0.395	cohort, case definition
Celiac Disease	0.0006	0.0011	0.0017	25-84	0.01	0.01	Underdiagnosed- diagnosed cases vs screened cases
exfoliation glaucoma	0.0110	0.0240	NA	NA	0.15	0.15	not well studied in US – European numbers
intracranial aneurysm	0.0064	0.0090	NA	NA	0.05	0.05	Ruptured vs total cases
Myocardial infarction	0.4240	0.2490	0.177	45-84	0.49	0.3	Stable angina added as CHD
Psoriasis	0.0400	0.0400	0.107	0-79	0.02	0.02	Different refs
Restless Leg Syndrome	0.0400	0.0400	0.04	30-89		0.13	Greek study(navi) vs meta-analysis (decode)

Baseline Epidemiological Standards Next Steps

Companies will investigate dissimilar numbers more thoroughly

Need for scientific community to establish standardized baseline numbers

Continue to have transparency on website with regards to references used, backend calculations if number is not reported in the text, explanatory text