Genome wide association in AD: will they ever yield new therapy?

Cornelia M van Duijn

Genetic Epidemiology Unit
Erasmus University Medical Centre
Rotterdam, The Netherlands

Outline

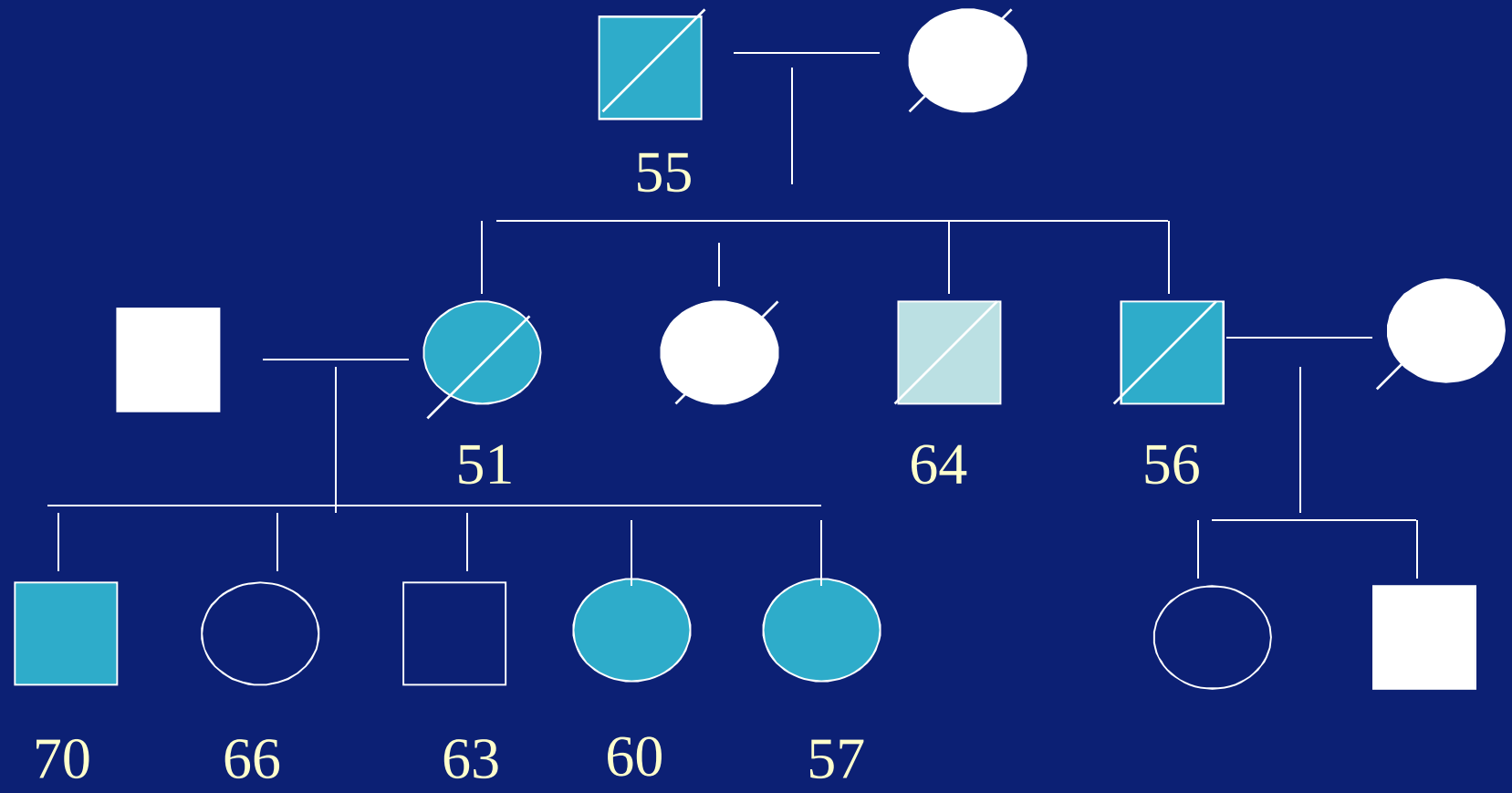
- Complexity of complex disease
- From explaining to predicting
- Challenges

Heritability Alzheimer's disease (AD)

§ < 70 years: 40%-80%

§ > 70: 20%-40%

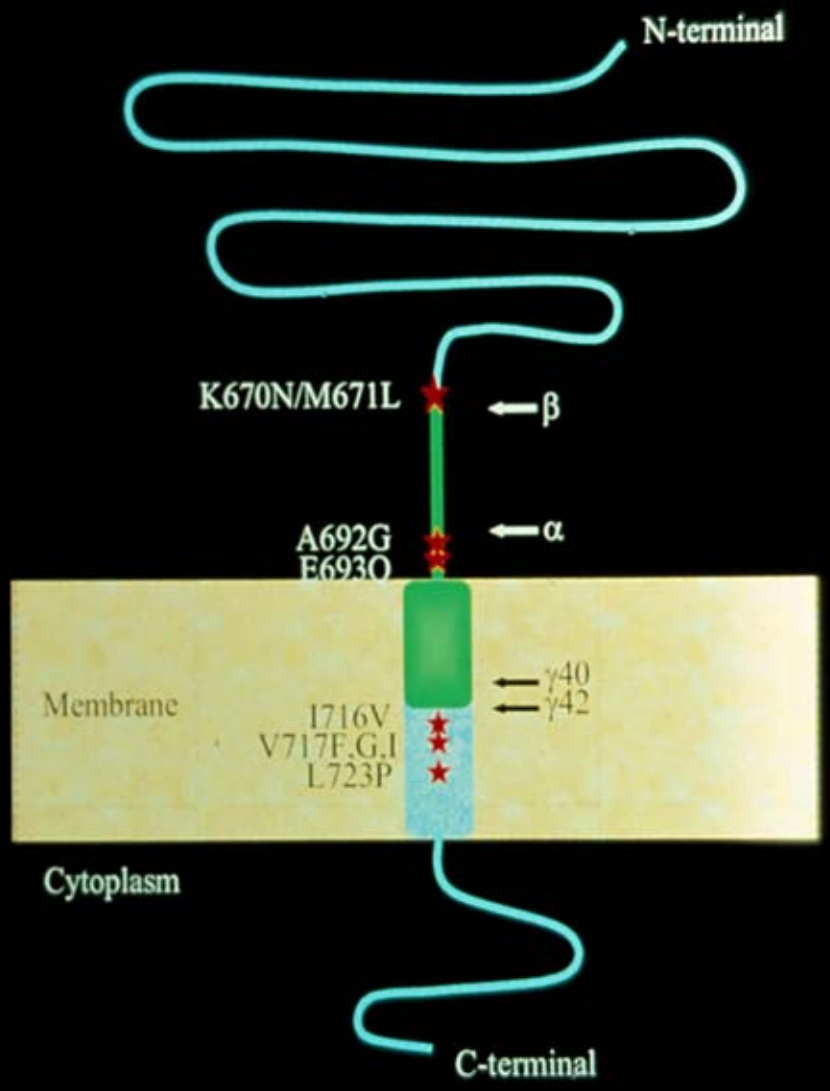
Dominant families



Alzheimer genes AD

Locus	Chromosome	Gene	References
AD1	21q21.2	APP	St George-Hyslop, 1987 Goate et al, 1991
AD2	19q13.2	APOE	Pericak-Vance et al, 1988 Corder et al, 1993
AD3	14q14.3	PSEN1	Schellenberg et al, 1993 Sherrington et al, 1995
AD4	1q31-42	PSEN2	Levy-Lahad et al, 1995 Levy-Lahad et al, 1995

Amyloid Precursor Protein

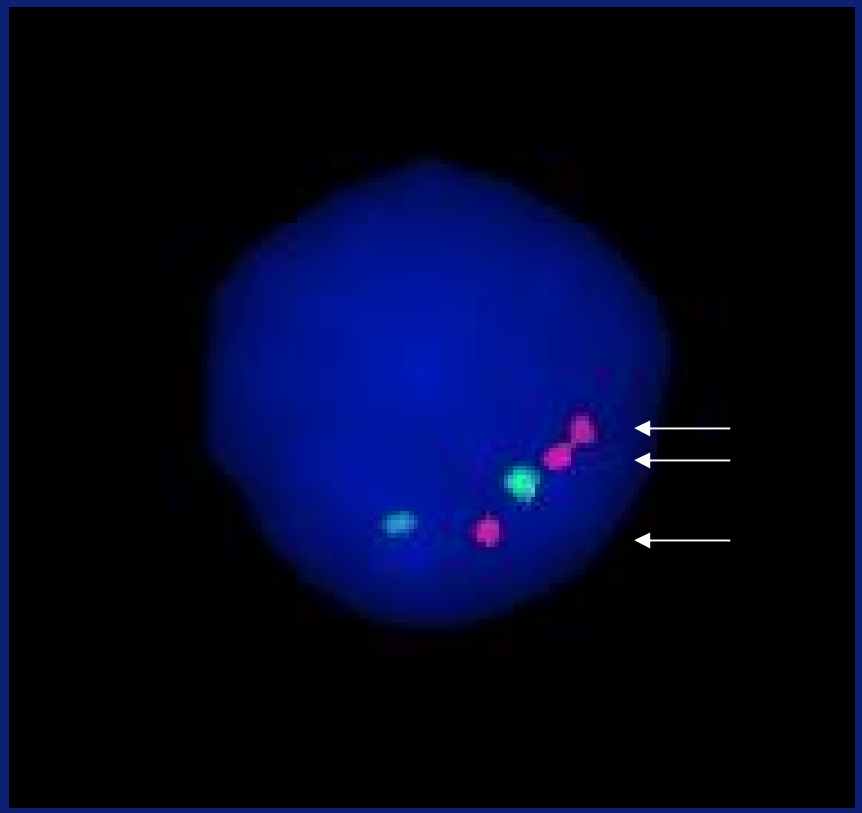


APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy

Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A, Dumanchin C, Feuillette S, Brice A, Vercelletto M, Dubas F, Frebourg T, Campion D

Nat Genet 2006;38:24-26

Fish of interphase nuclei of a patient with an APP duplication



Sleegers et al, Brain 2006

Dominant mutations in early-onset AD (<65 years)

- Amyloid Precursor Protein (APP) 2.5%
- Presenilin 1 (PSEN1) 6%
- Presenilin 2 (PSEN2) 1%

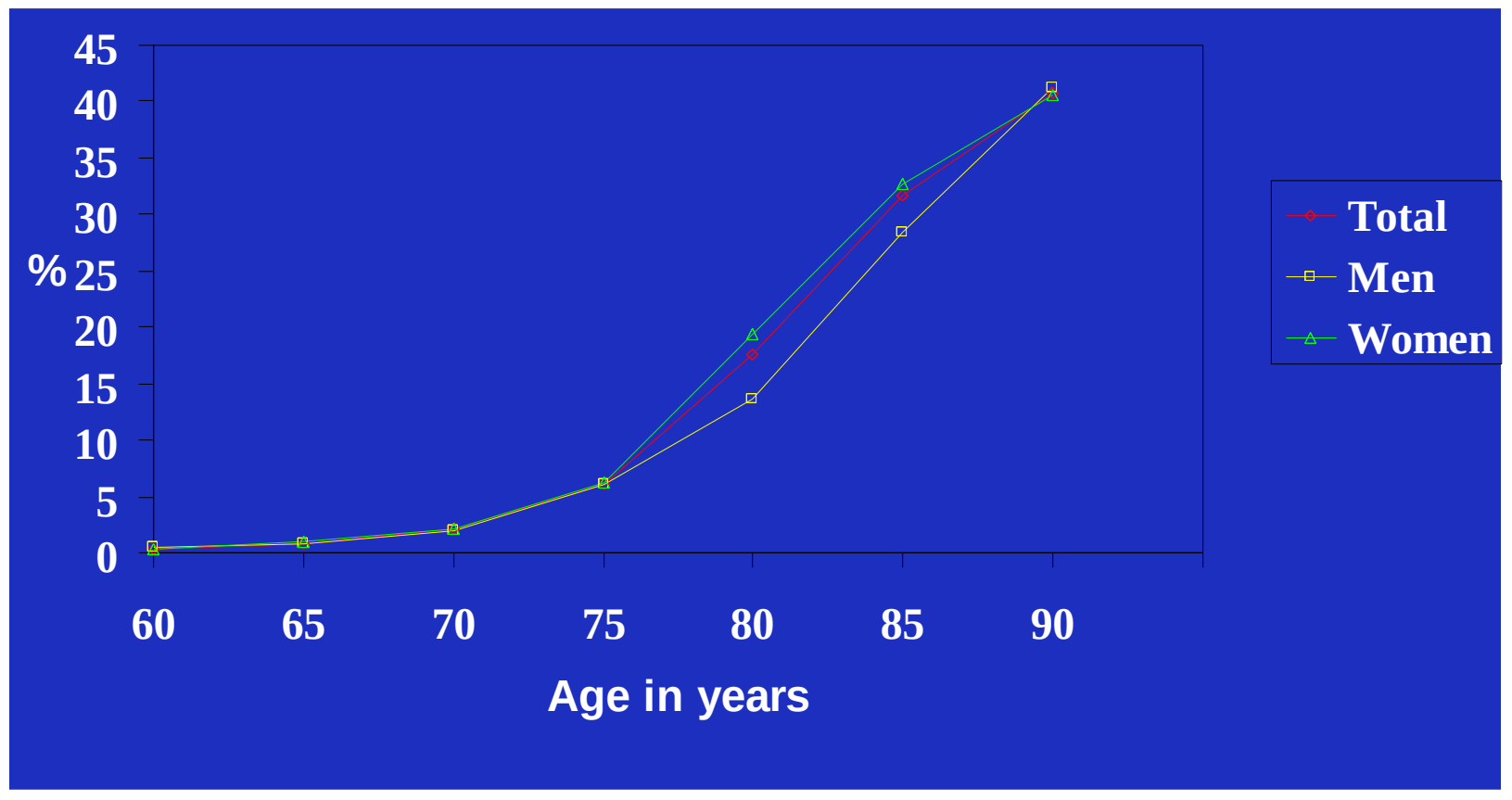
Sleegers et al, Brain 2006

Population attributable risk

§ Proportion of cases that can be prevented if the risk factor is neutralized (e.g., treatment in LDL-receptor mutation carriers with statins)

§ Is developed to evaluate the impact of the risk factor in the *population*

Prevalence of AD by age



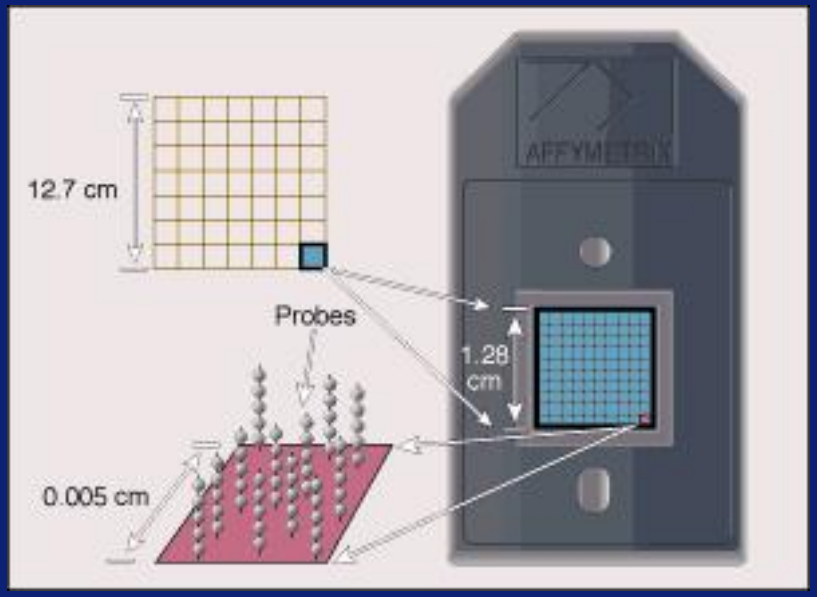
Impact dominant mutations in population

	Relative Risk Carrier	Population Attributable Risk
APP	100%	0.10 %
PSEN1	100%	0.24 %
PSEN2	100%	0.04 %

Sleegers et al, Brain 2006

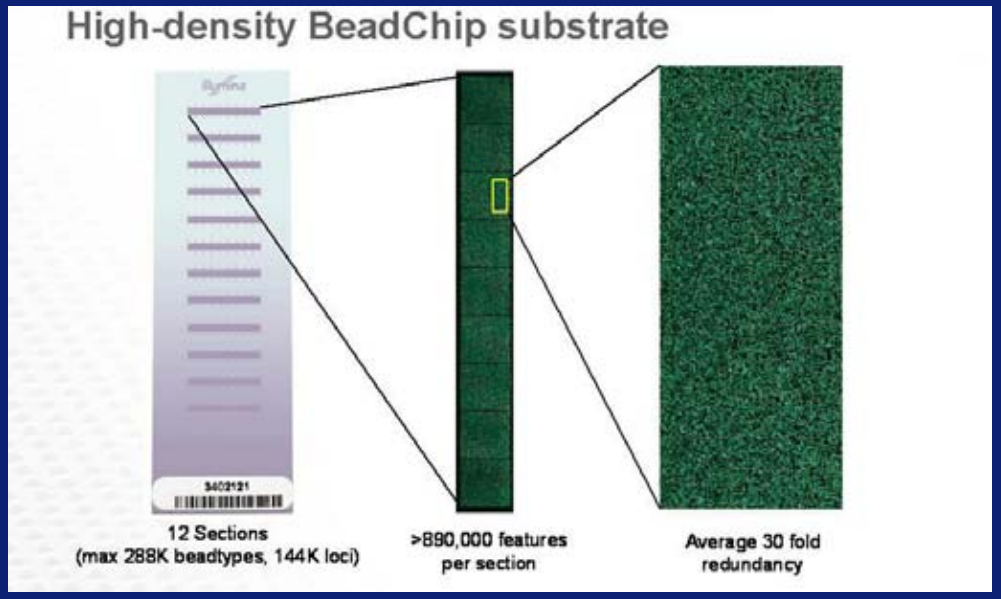
Genome wide association (GWAS)

Affymetrix arrays



250,000 to 1,000,000 SNPs

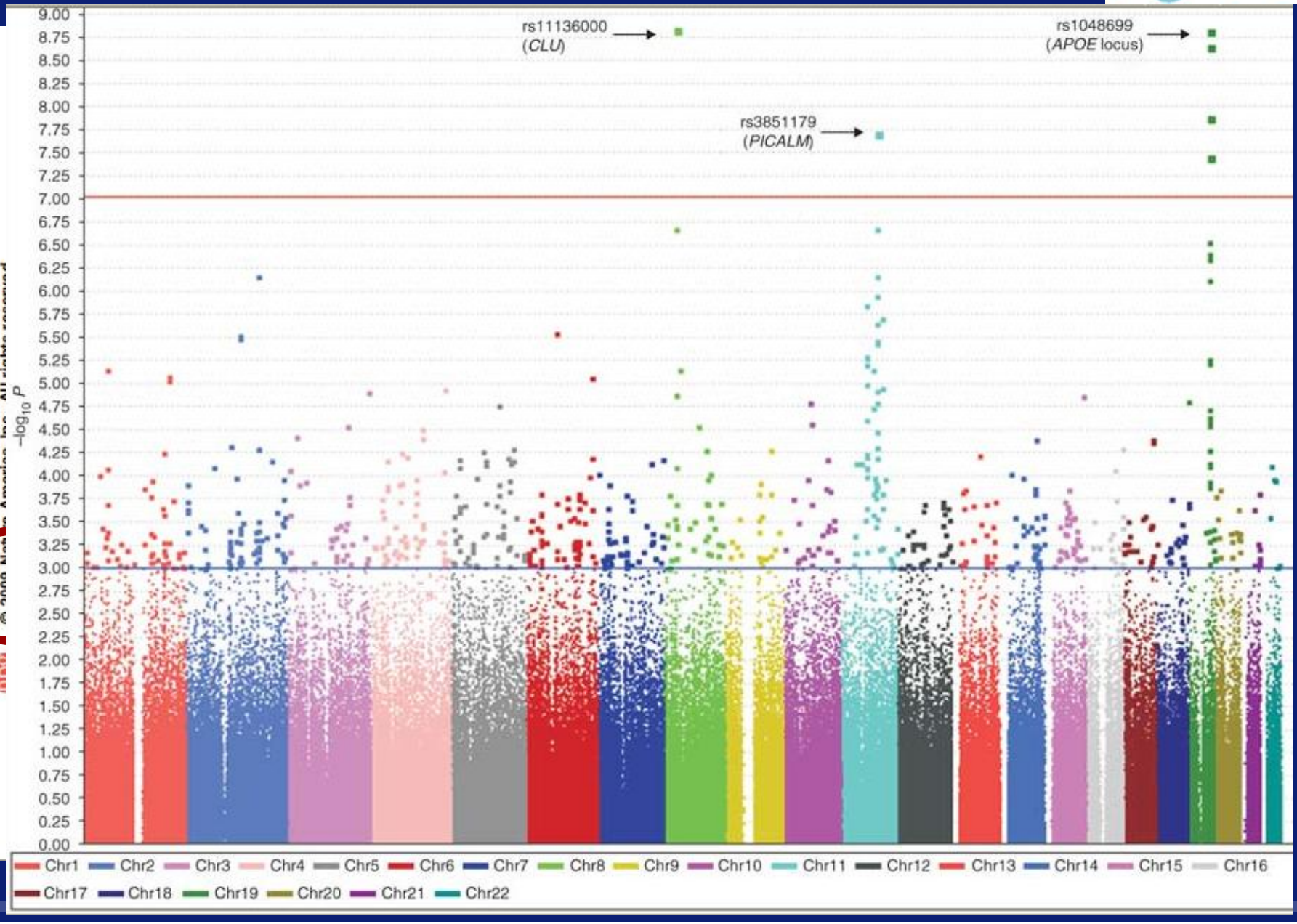
Illumina arrays



317,000; 550,000; 1,000,000 SNPs

Capture > 95% of common SNPs (frequency >0.05)

Finding the next APOE …



Genome-wide association study identifies variants at *CLU* and *CR1* asso

Jean-Charles Lambert[1–3], Si
Onofre Combarros[9], Diana
Claudine Berr[13], Florence P
Peter De Deyn[7,15], Ignacio I
the European Alzheimer's D
Carole Dufouil[22,23], Céline
Michelangelo Mancuso[27], F
Giorgio Annoni[32], Davide S
Hilkka Soininen[8], Karen Rit
Ivo Gut[4], Christine Van Bro

The gene encoding apolipoprot 19 is the only confirmed suscep Alzheimer's disease. To identify large genome-wide association France with Alzheimer's disease Markers outside *APOE* with sug ($P < 10^{-5}$) were examined in co Finland, Italy and Spain totaling cases and 3,297 controls. Two l of association: one within *CLU* clusterin or apolipoprotein J, o OR = 0.86, 95% CI 0.81–0.90, data) and the other within *CR1* component (3b/4b) receptor 1, OR = 1.21, 95% CI 1.14–1.29, data). Previous biological studi *CR1* in the clearance of β amyl constituent of amyloid plaques brain lesions of individuals witl

Genome-wide Analysis of Genetic Loci Associated With Alzheimer Disease

Sudha Seshadri, MD; Annette L. Fitzpatrick, PhD; M. Arfan Ikram, MD, PhD; Anita L. DeStefano, PhD; Vilmundur Gudnason, MD, PhD; Merce Boada, MD, PhD; Joshua C. Bis, PhD; Albert V. Smith, PhD; Minerva M. Carrasquillo, PhD; Jean Charles Lambert, PhD; Denise Harold, PhD; Elisabeth M. C. Schrijvers, MD; Reposo Ramirez-Lorca, PhD; Stephanie Debette, MD, PhD; W. T. Longstreth Jr, MD; A. Cecile J. W. Janssens, PhD; V. Shane Pankratz, PhD; Jean François Dartigues, PhD; Paul Hollingworth, PhD; Thor Aspelund, PhD; Isabel Hernandez, MD; Alexa Beiser, PhD; Lewis H. Kuller, MD; Peter J. Koudstaal, MD, PhD; Dennis W. Dickson, MD; Christophe Tzourio, MD; Richard Abraham, PhD; Carmen Antunez, MD; Yangchun Du, PhD; Jerome I. Rotter, MD; Yurii S. Aulchenko, PhD; Tamara B. Harris, MD; Ronald C. Petersen, MD; Claudine Berr, MD, PhD; Michael J. Owen, MB, ChB, PhD; Jesus Lopez-Arrieta, MD; Badri N. Vardarajan, MS; James T. Becker, PhD; Fernando Rivadeneira, MD, PhD; Michael A. Nalls, PhD; Neill R. Graff-Radford, MD; Dominique Campion, MD, PhD; Sanford Auerbach, MD; Kenneth Rice, PhD; Albert Hofman, MD, PhD; Palmi V. Jonsson, MD; Helena Schmidt, MD, PhD; Mark Lathrop, PhD; Thomas H. Mosley, PhD; Rhoda Au, PhD; Bruce M. Psaty, MD, PhD; Andre C. Uitterlinden, PhD; Lindsay A. Farrer, PhD; Thomas Lumley, PhD; Agustin Ruiz, MD, PhD; Julie Williams, PhD; Philippe Amouyel, MD, PhD; Steve C. Younkin, PhD; Philip A. Wolf, MD; Lenore J. Launer, PhD; Oscar L. Lopez, MD; Cornelia M. van Duijn, PhD; Monique M. B. Breteler, MD, PhD
for the CHARGE, GERAD1, and EADI1 Consortia

Context Genome-wide association studies (GWAS) have recently identified *CLU*, *PICALM*, and *CR1* as novel genes for late-onset Alzheimer disease (AD).

Objectives To identify and strengthen additional loci associated with AD and confirm these in an independent sample and to examine the contribution of recently identified genes to AD risk prediction in a 3-stage analysis of new and previously published GWAS on more than 35 000 persons (8371 AD cases).

Design, Setting, and Participants In stage 1, we identified strong genetic associations ($P<10^{-3}$) in a sample of 3006 AD cases and 14 642 controls by combining new data from the population-based Cohorts for Heart and Aging Research in Genomic Epidemiology consortium (1367 AD cases [973 incident]) with previously reported results from the Translational Genomics Research Institute and the Mayo AD GWAS. We identified 2708 single-nucleotide polymorphisms (SNPs) with $P<10^{-3}$. In stage 2, we pooled results for these SNPs with the European AD Initiative (2032 cases and 5328 controls) to identify 38 SNPs (10 loci) with $P<10^{-5}$. In stage 3, we combined data for these 10 loci with data from the Genetic and Environmental Risk in AD consortium (3333 cases and 6995 controls) to identify 4 SNPs with $P<1.7\times10^{-8}$. These 4 SNPs were replicated in an independent Spanish sample (1140 AD cases and 1209 controls). Genome-wide association analyses were completed in 2007-2008 and the meta-analyses and replication in 2009.

Main Outcome Measure Presence of Alzheimer disease.

Results Two loci were identified to have genome-wide significance for the first time: rs744373 near *BIN1* (odds ratio [OR], 1.13; 95% confidence interval [CI], 1.06-1.21 per copy of the minor allele; $P=1.59\times10^{-11}$) and rs597668 near *EXOC3L2/BLOC1S3/MARK4* (OR, 1.18; 95% CI, 1.07-1.29; $P=6.45\times10^{-9}$). Associations of these 2 loci plus the previously identified loci *CLU* and *PICALM* with AD were confirmed in the Spanish sample ($P<.05$). However, although *CLU* and *PICALM* were confirmed to be associated with AD in this independent sample, they did not improve the ability of a model that included age, sex, and *APOE* to predict incident AD (improvement in area under the receiver operating characteristic curve from 0.847 to 0.849 in the Rotterdam Study and 0.702 to 0.705 in the Cardiovascular Health Study).

Conclusions Two genetic loci for AD were found for the first time to reach genome-wide statistical significance. These findings were replicated in an independent population. Two recently reported associations were also confirmed. These loci did not improve AD risk prediction. While not clinically useful, they may implicate biological pathways useful for future research.

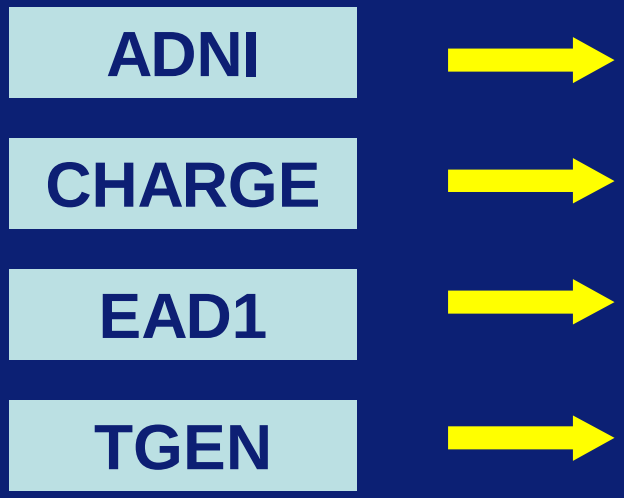
JAMA. 2010;303(18):1832-1840 www.jama.com

Effects of new genes are small!

GENE	*Odds ratio (95%CI)*	*p-value*
APOE	*2.53 [2.41-2.66]*	*$1.04*10^{-235}$*
CLU	*0.85 [0.81-0.90]*	*$1.62*10^{-16}$*
PICALM	*0.87 [0.84-0.91]*	*$3.16*10^{-12}$*
BIN1	*1.15 [1.11-1.20]*	*$1.59*10^{-11}$*
CRI	*1.20 [1.11-1.25]*	*$1.40*10^{-11}$*

Note: odds ratio of 0.85 implies a 1.18 reduction in risk!

Mega meta-analyses of the Internatonal Genetics of AD consortium!



MEGA Meta analyses 25,000 AD cases and 30,000 controls

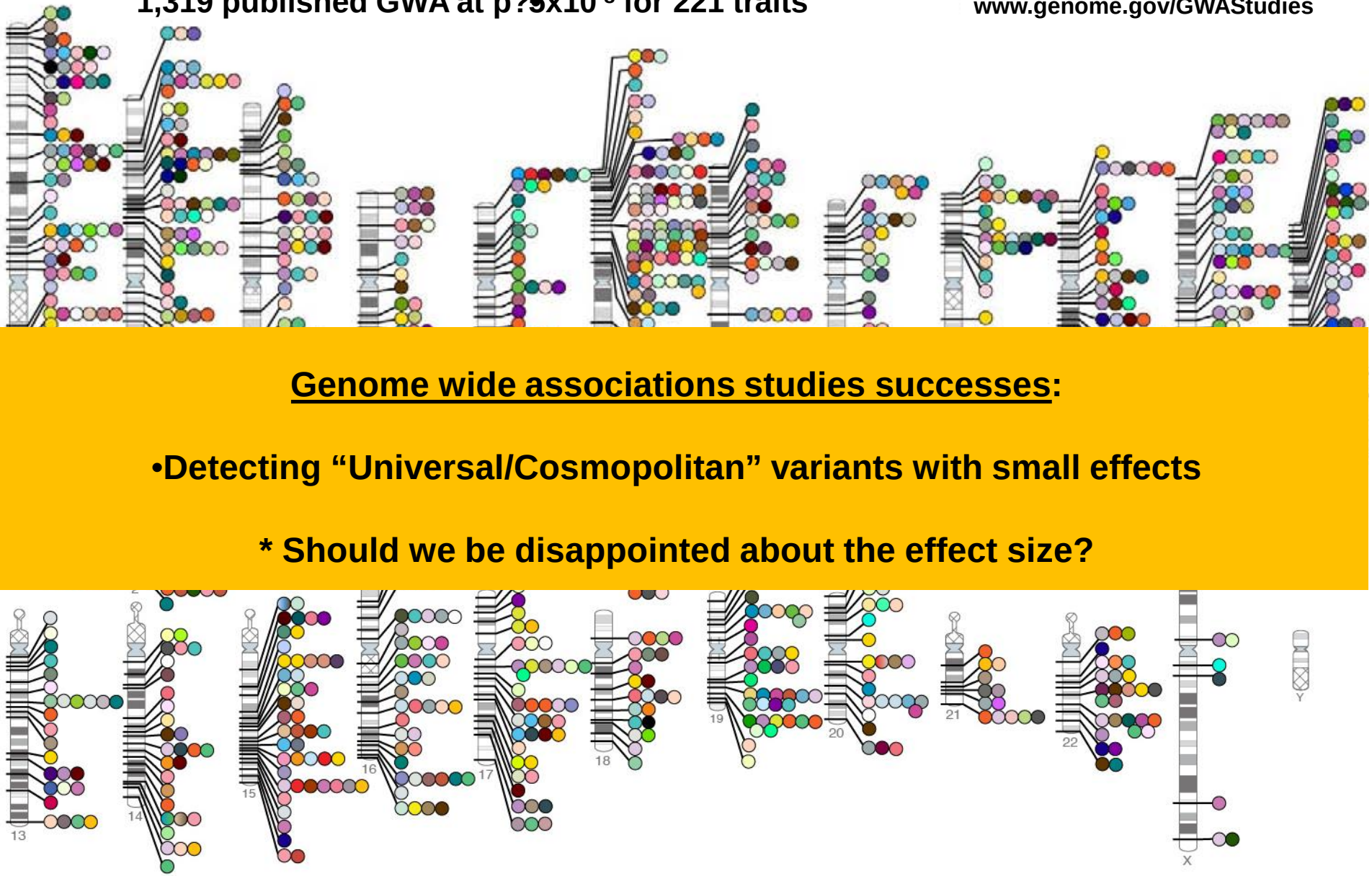
GWAS: 19 AD genes

No evidence we reached the limit of GWAS

The limit appears to be the number of patients with GWAS

Published Genome-Wide Associations through 03/2011,
1,319 published GWA at p?5x10^{-8} for 221 traits

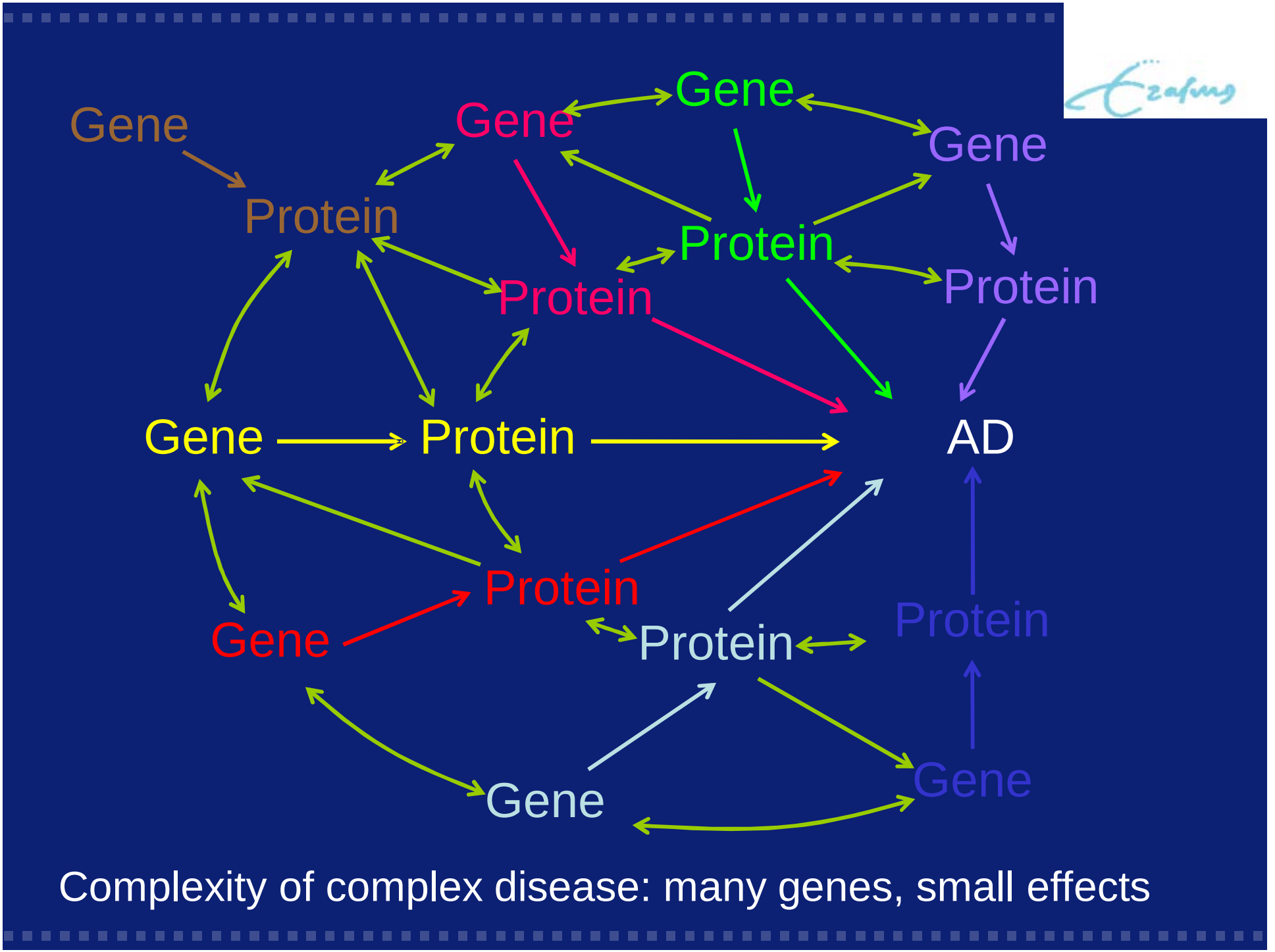
NHGRI GWA Catalog
www.genome.gov/GWAStudies



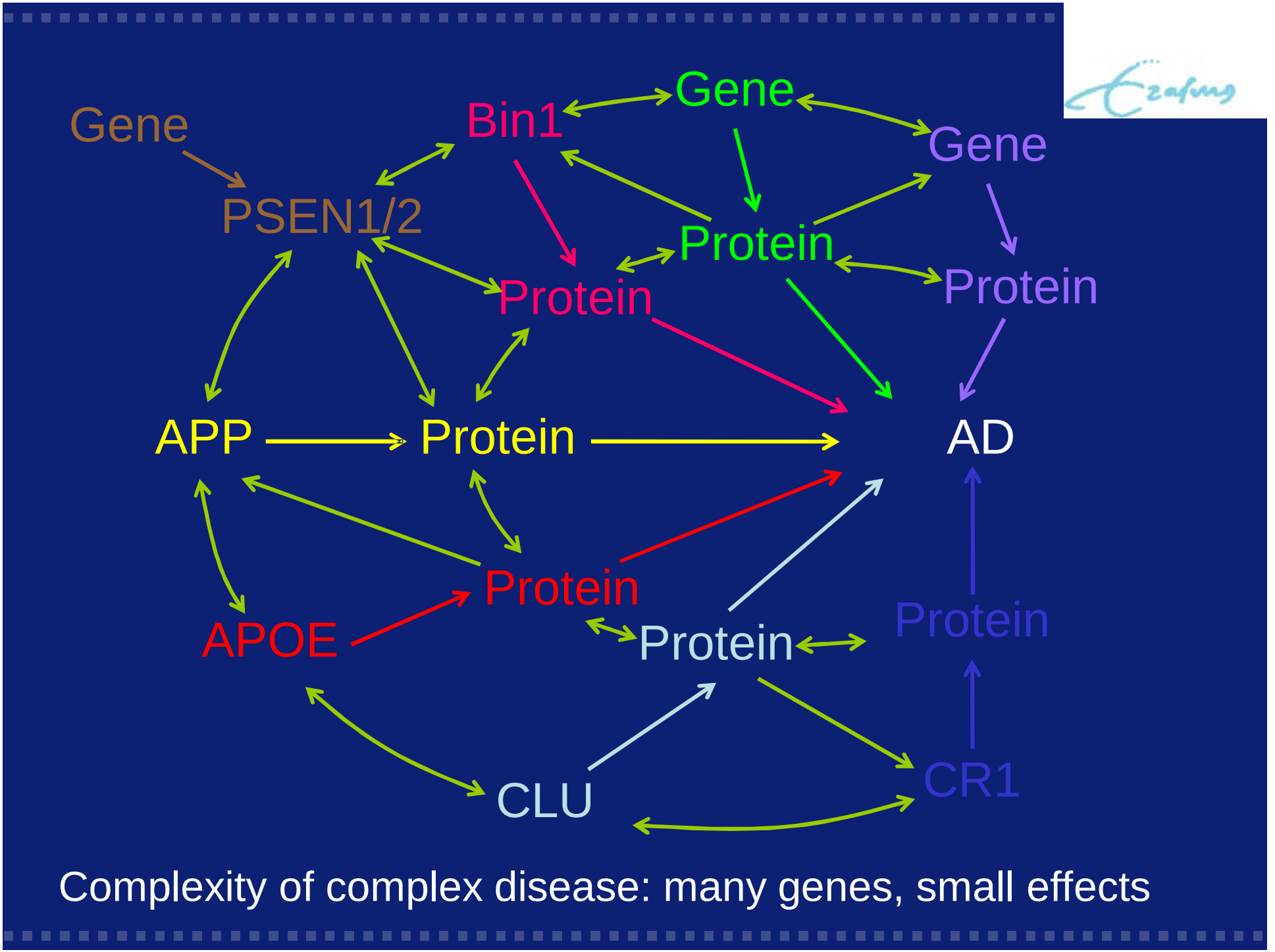
<u>Genome wide associations studies successes</u>:

•Detecting "Universal/Cosmopolitan" variants with small effects

*** Should we be disappointed about the effect size?**



Complexity of complex disease: many genes, small effects



Complexity of complex disease: many genes, small effects

Why did we catch only common with relatively small effects?

1) It is the entity of a complex multifactorial disease (Fisher)
2) Common variant (50% carrier) cannot have large relative risk (<2)
3) Hundreds of variant should have relative risks <<<<<<<2

What do the new AD genes tell?

- Clusterin (APOJ): underscores the role of apolipoproteins
- CR1(encodes encodes receptor of the complement C3b protein): complement activation
- PICALM (phosphatidylinositolbinding clathrin assembly protein):

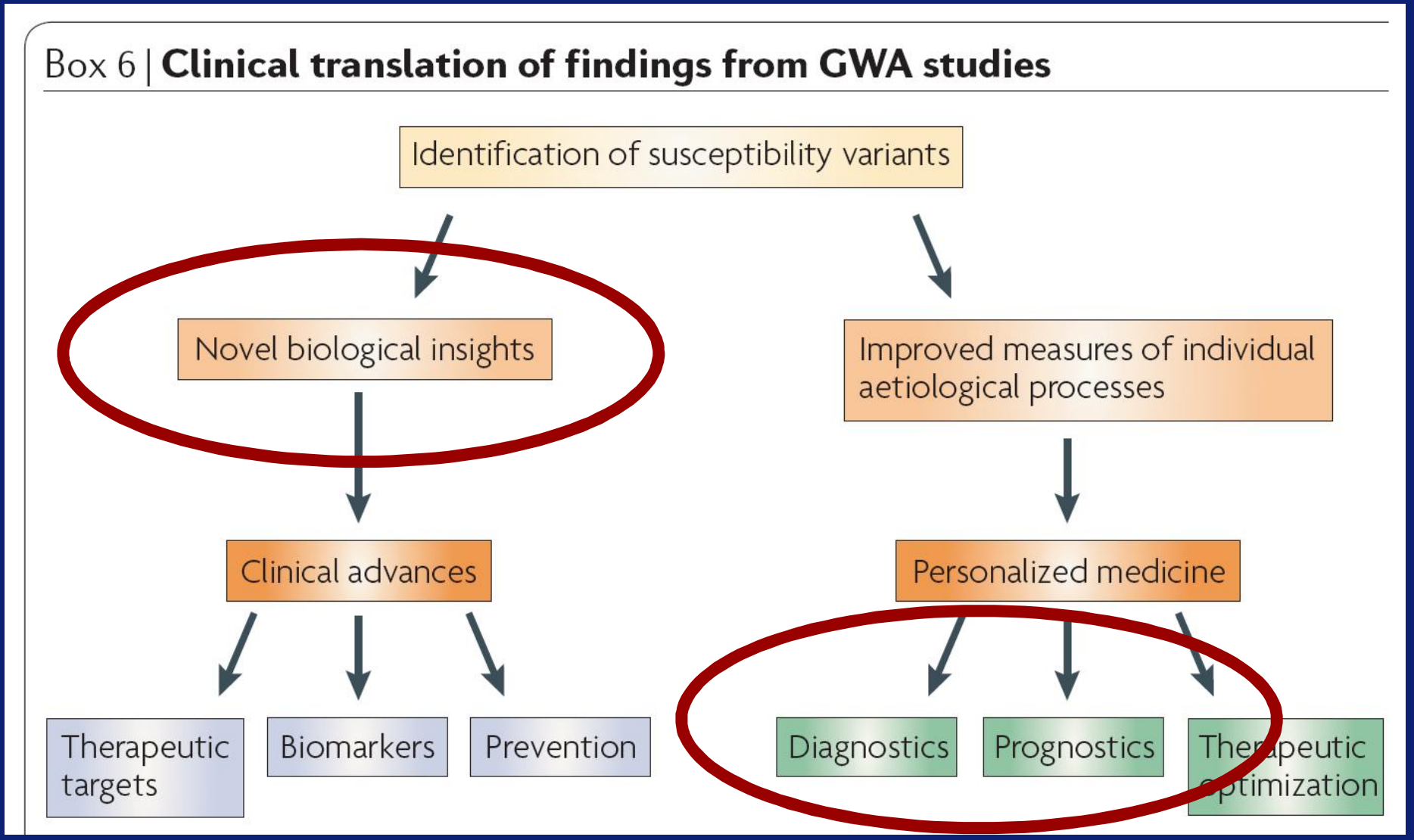
  (1) intracellular trafficking of proteins and lipids

  (2) clathrin-mediated endocytosis
- BIN1 (bridging integrator 10): clathrin-mediated endocytosis

Pathways identified

- *Ubiquitination*
- *Endocytosis*
- *Cholesterol transport*
- *Immunity*

Taking GWAS to new knowledge



McCarthy *Nat Rev Genet 2008*

Outline

- Complexity of complex disease
- From explaining to predicting
- Challenges

Explaining versus predicting: a different cup of tea!

- All popes have been men, predicting that the pope will be a man
- Which man will be the next pope?
- Many patients share a HLA haplotype, but 50% of the population carries that haplotype
- Which person will be the next patient?

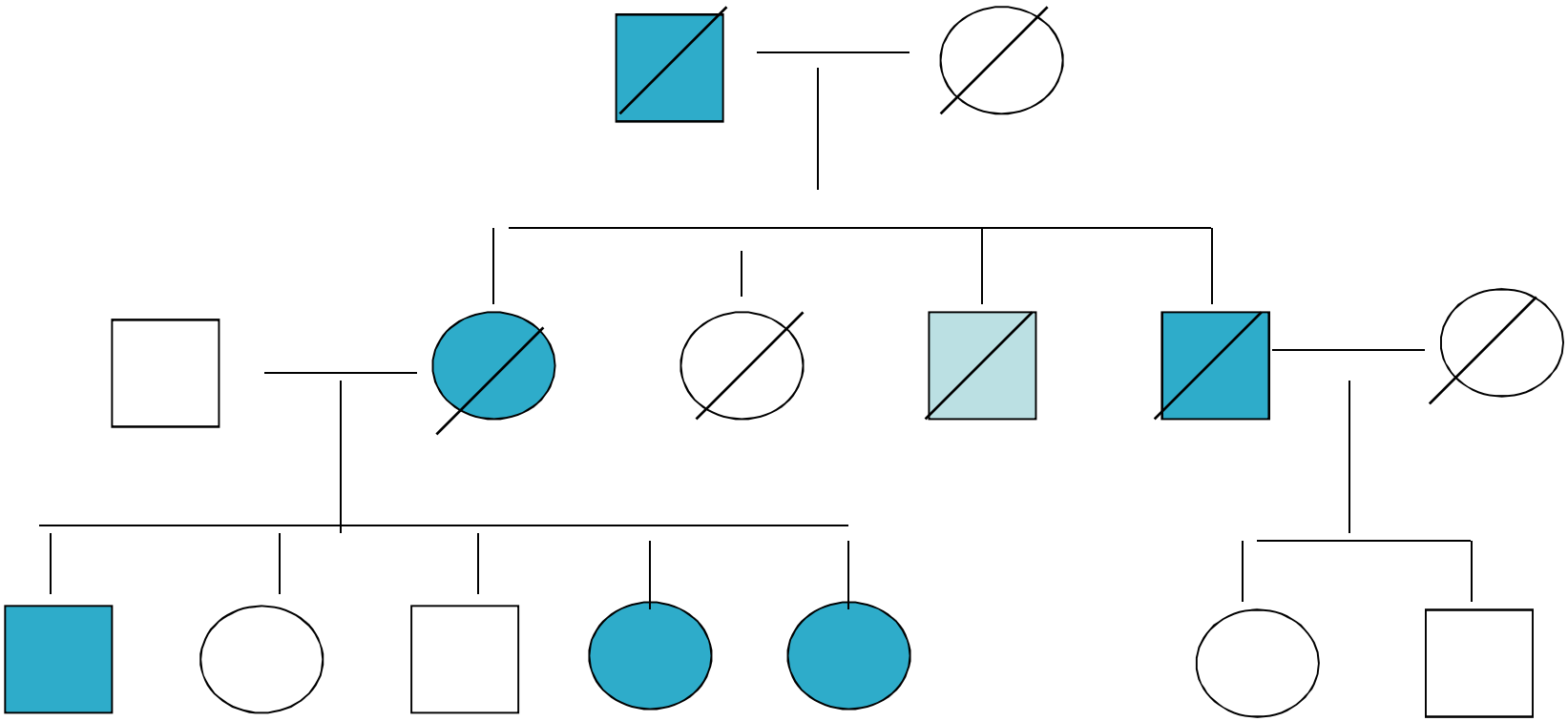
Population attributable risk

- Proportion of cases that can be prevented if the risk factor is neutralized (e.g., treatment in LDL-receptor mutation carriers with statins)

- Is developed to evaluate the impact of the risk factor in the *population*

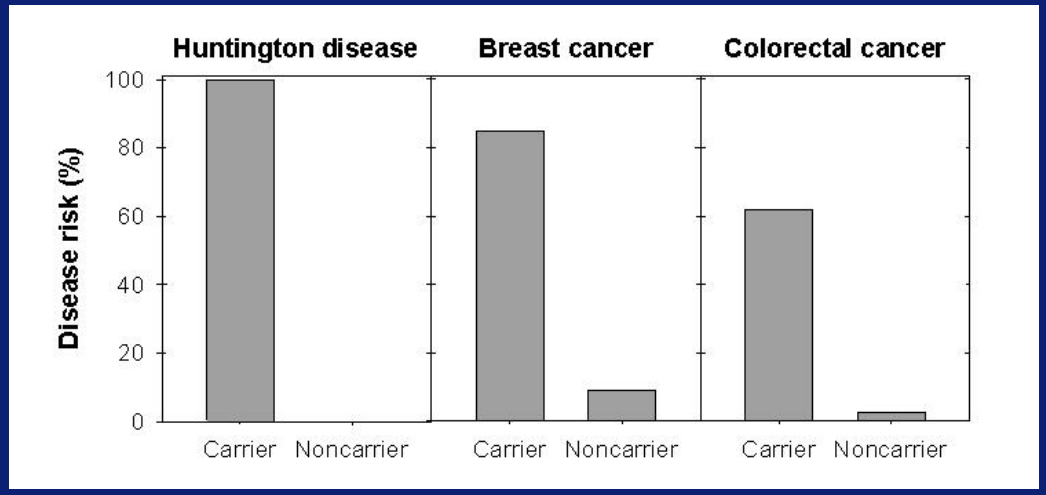
- *A low population attributable risk does NOT imply the gene is NOT valuable for INDIVIDUAL risk prediction*

Clinical relevance of APP and PSEN1/2



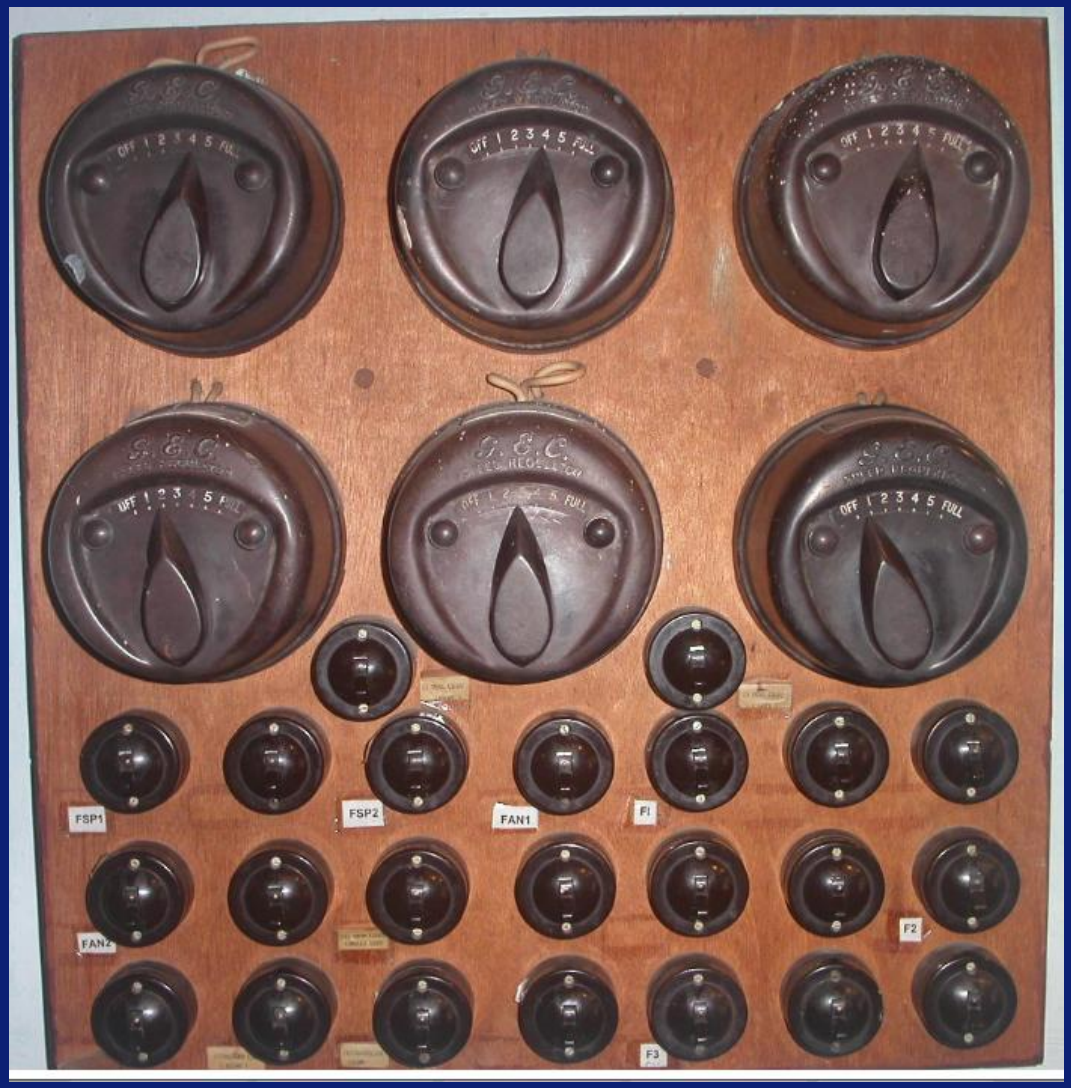
Personalized medicine: clinical counseling

Translation monogenic disorders starts with modeling the origin



Mutation = sufficient cause. We may not understand the mechanisms determining severity or age at onset but we have the crucial ‘switch’ setting off the pathogenesis jusitifying personalize medicine e.g. screening for cancer at early age, preventive amputations and statin treatment (childhood) to prevent irreverisable pathology of dyslipidemia

Translation complex disorders starts with defining the cause(s)

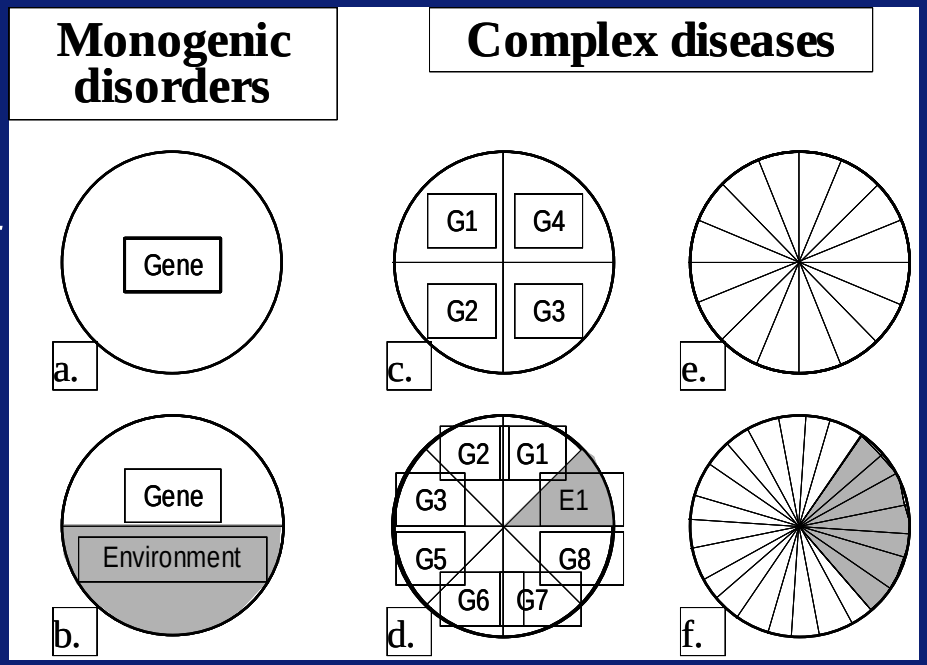


- There may be mutations with large effects leading to monogenic forms of disease (e.g. Alzheimer, breast or colon cancer)
- For most/many patients here is not a single master switch as is the case with monogentic (forms of a) disease
- Different combinations of mutations make up the sufficient cause to sett off the pathogenesis

Sufficient causes in genetics

APP, PSEN1 PSEN2

PRNP/BSE



Alzheimer's disease (AD)

Rothman & Greenland, 2005

Complex diseases may be caused by different causal mechanisms (different combinations of component causes)

Janssens & van Duijn *Hum Mol Genet* 2008, *BMC Genome Med* 2009

The risk factors for AD

	Description of study	Main outcomes
Lifestyle		
Obesity[109]	Meta-analysis of ten studies. All prospective studies with at least 2 years follow-up and participants over 40 years old	Dementia RR 1·42 (95% CI 0·93–2·16); Alzheimer's disease 1·80 (1·00–3·29)
Smoking[110]	Meta-analysis of four prospective studies with 2–25 years follow-up in over 17 000 people. In the four studies the dementia ORs were 3·17 (95% CI 1·37–7·35), 1·42 (1·07–1·89), 1·60 (1·00–2·57), and 1·63 (1·00–2·67)	Dementia RR 2·2 (1·3–3·6)
Physical activity[111]	13 prospective studies focusing on Alzheimer's disease, dementia, or both, with at least 150 000 participants	Dementia RR 0·72 (95% CI 0·60–0·86); Alzheimer's disease 0·55 (0·36–0·84)
Cognitive reserve (intelligence, occupation, and education)[112]	22 prospective studies with at least 29 000 participants followed up for a median of 7·1 years	Dementia OR 0·54 (95% CI 0·49–0·59)
Alcohol[113]	15 longitudinal studies with 2–8 years follow-up and at least 14 000 participants	Dementia RR 0·74 (95% CI 0·61–0·91); Alzheimer's disease 0·72 (0·61–0·86)

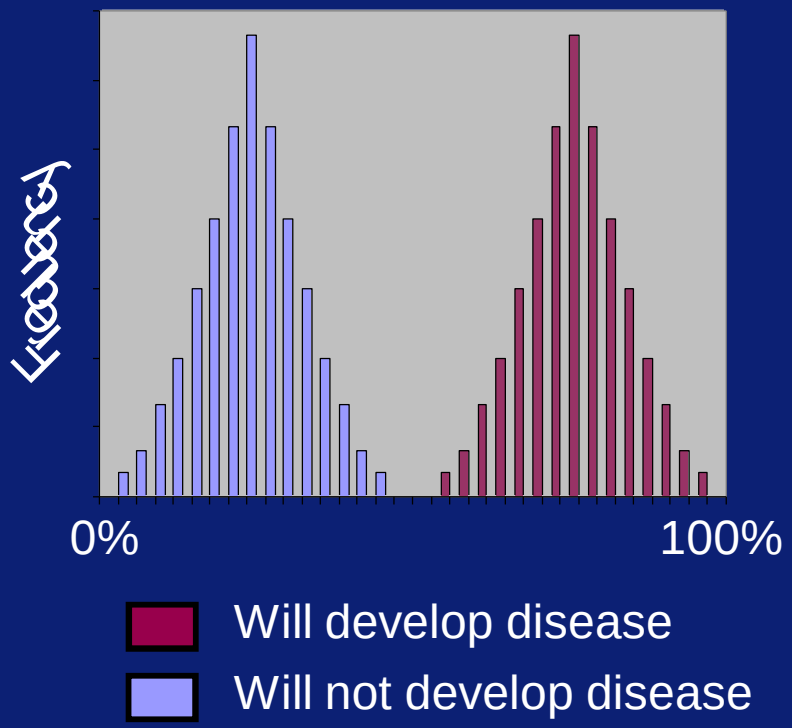
Beyond genetic risk scores in the JPND program PERADES

- *Ubiquitination risk scores * environment*
- *Endocytosis * environment*
- *Cholesterol transport * environment*
- *Immunity * environment*
- *All genes * environment*

What can we achieve?

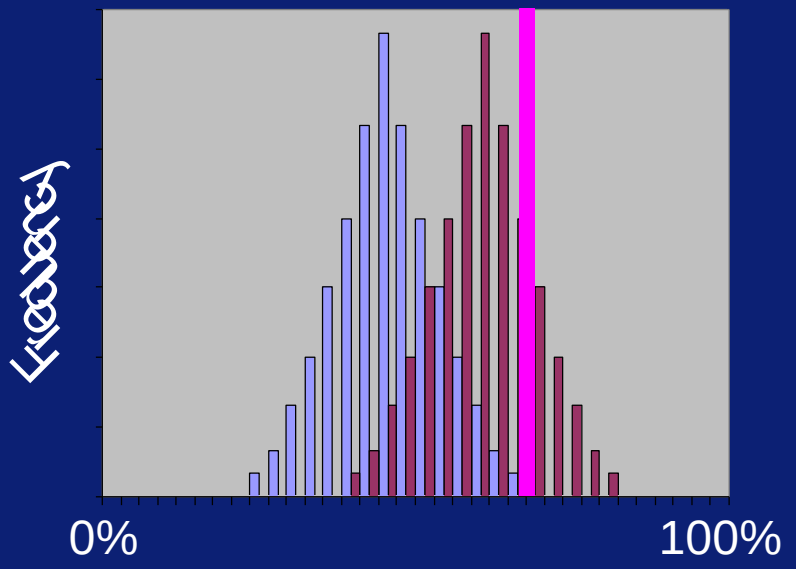
Risks of complex diseases are never 0.0!

Perfect discrimination



Unhappy few: Aggressive intervention!

Imperfect but useful

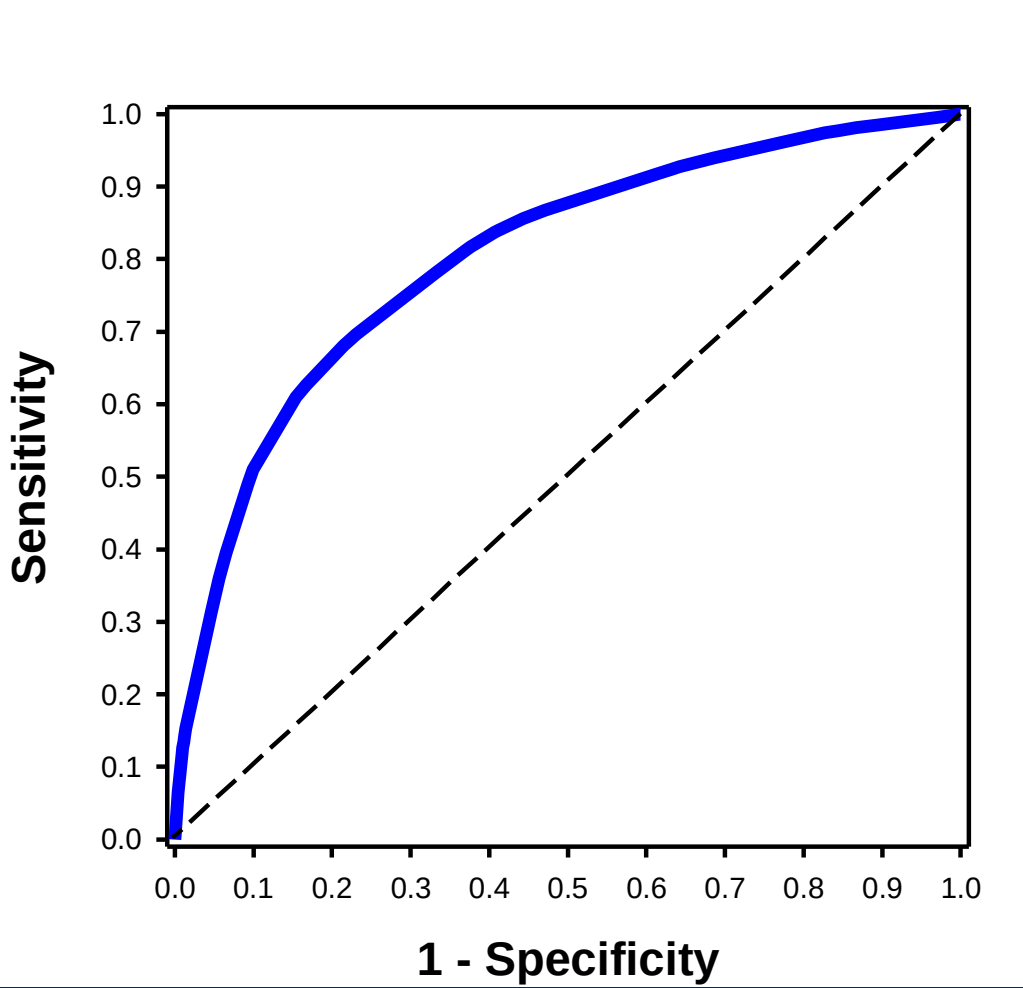


Find the cut-off point!

Ideally the cut-off is such that

- All people who will develop disease receive the intervention (sensitivity = 100%)
- Those who will not develop disease do not get intervention (specificity = 100%)
- Where to place the cut-off if curves are overlapping depends on the ‘aggressiveness’ of the intervention
- A first step to answer: is there risk differentiation?

Area under Curve (AUC) of Receiver Operator Curve (ROC)



AUC = Plot of all sensitivity-specificity combinations for ALL possible cut-off values of the predicted risks

Reference values AUC:

§ Tossing coin: AUC = 0.50

§ Perfect prediction: AUC = 1

Classical factors in coronary heart disease prediction: AUC ~ 0.75

For which diseases can we discriminate risks?

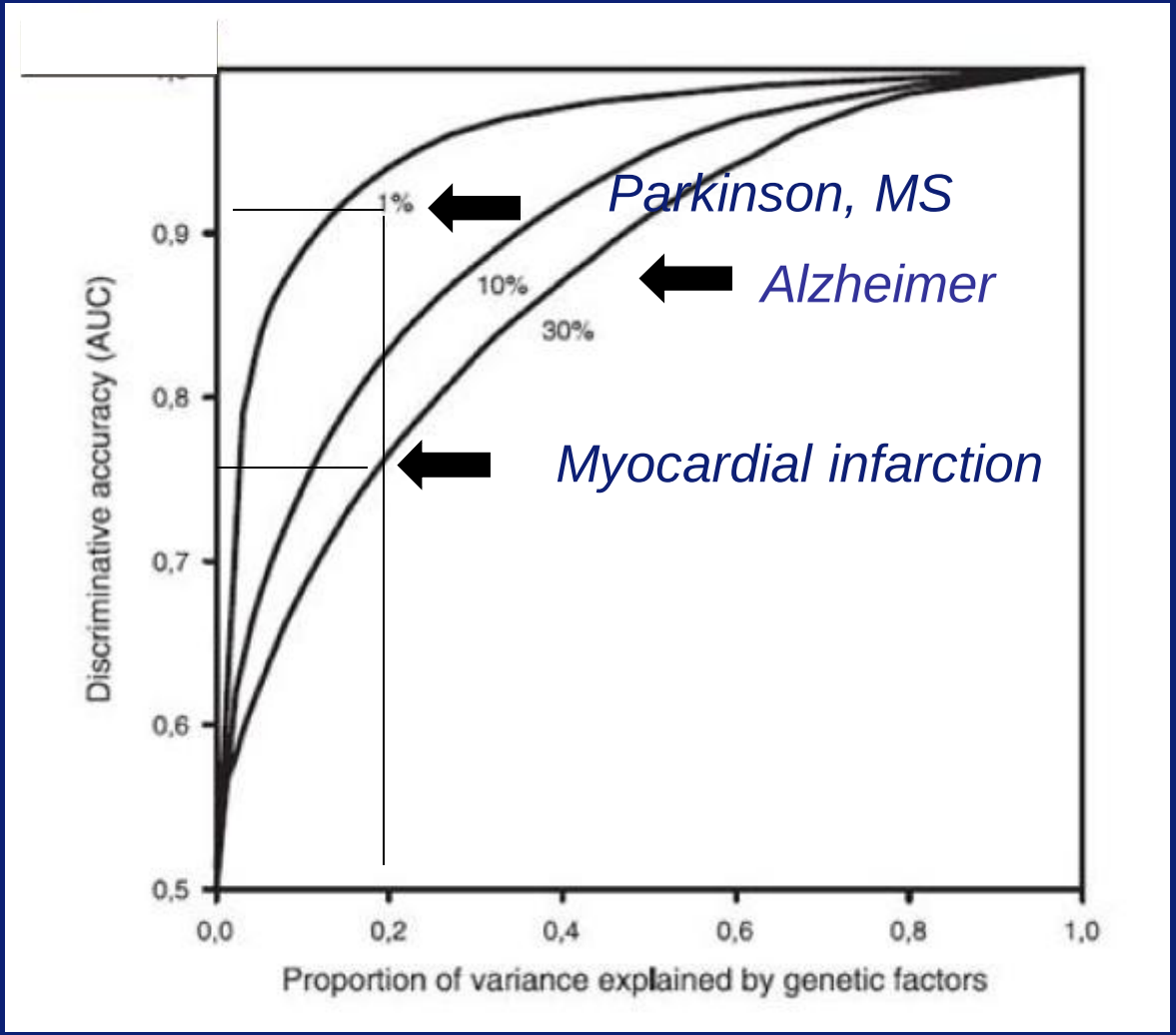
Scenario studies using simulated data

For example, we create population of 1 million individuals, of whom 10% has disease x. Disease results from effect of 400 genes (and many other unknown factors). Frequency of each risk allele is 10% and strength of association (odds ratio) varies from 2.0 to 1.05.

What is expected discriminative accuracy (AUC) of a prediction model that is based on these 400 genes?

Janssens et al. *Genet Med* 2006

Predictability of a complex disease by genes by disease prevalence



Common outcomes are more difficult to predict than rare ones

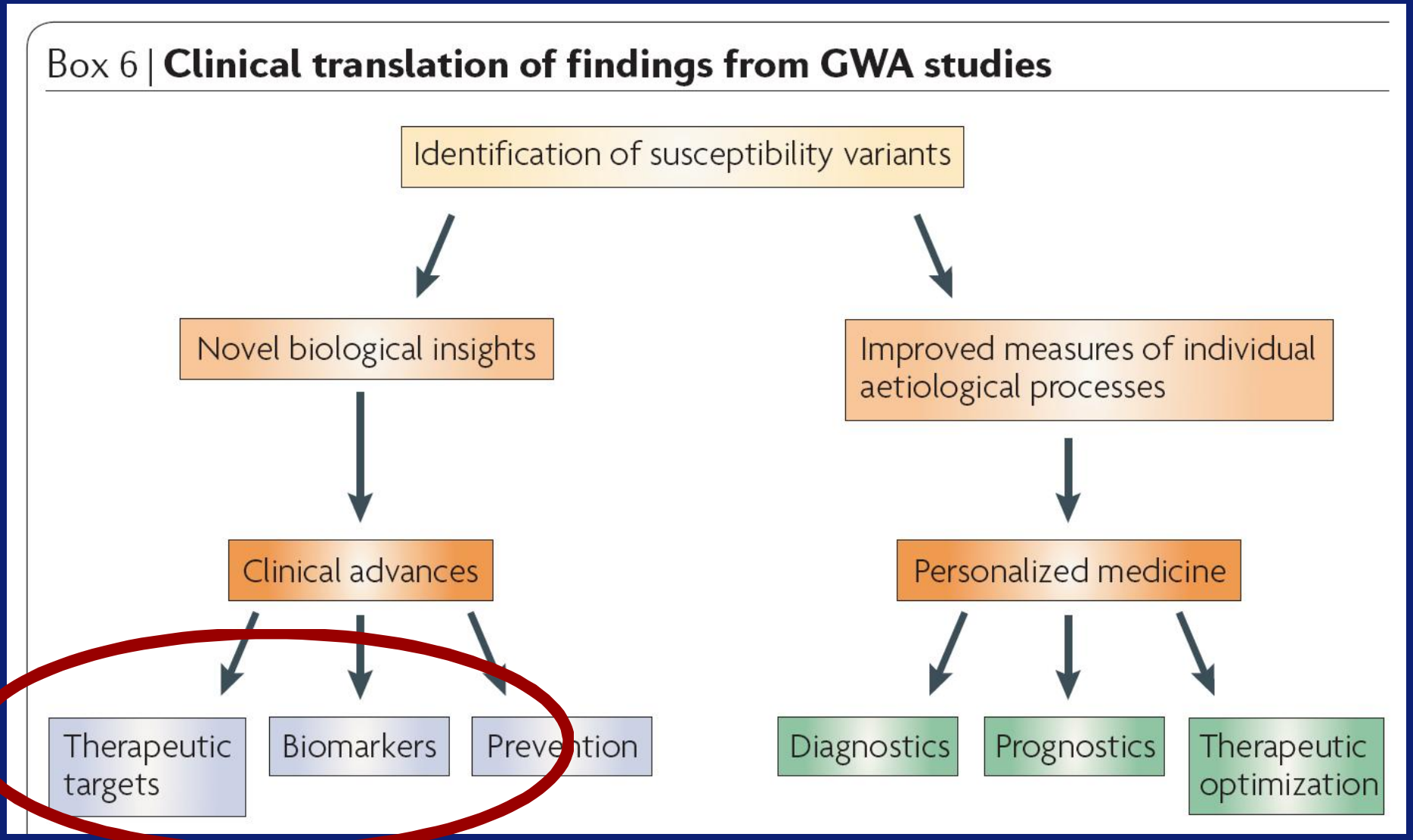
← *'Heritability"*

Janssens & van Duijn, Investig Genet. 2010; 1: 10.

Outline

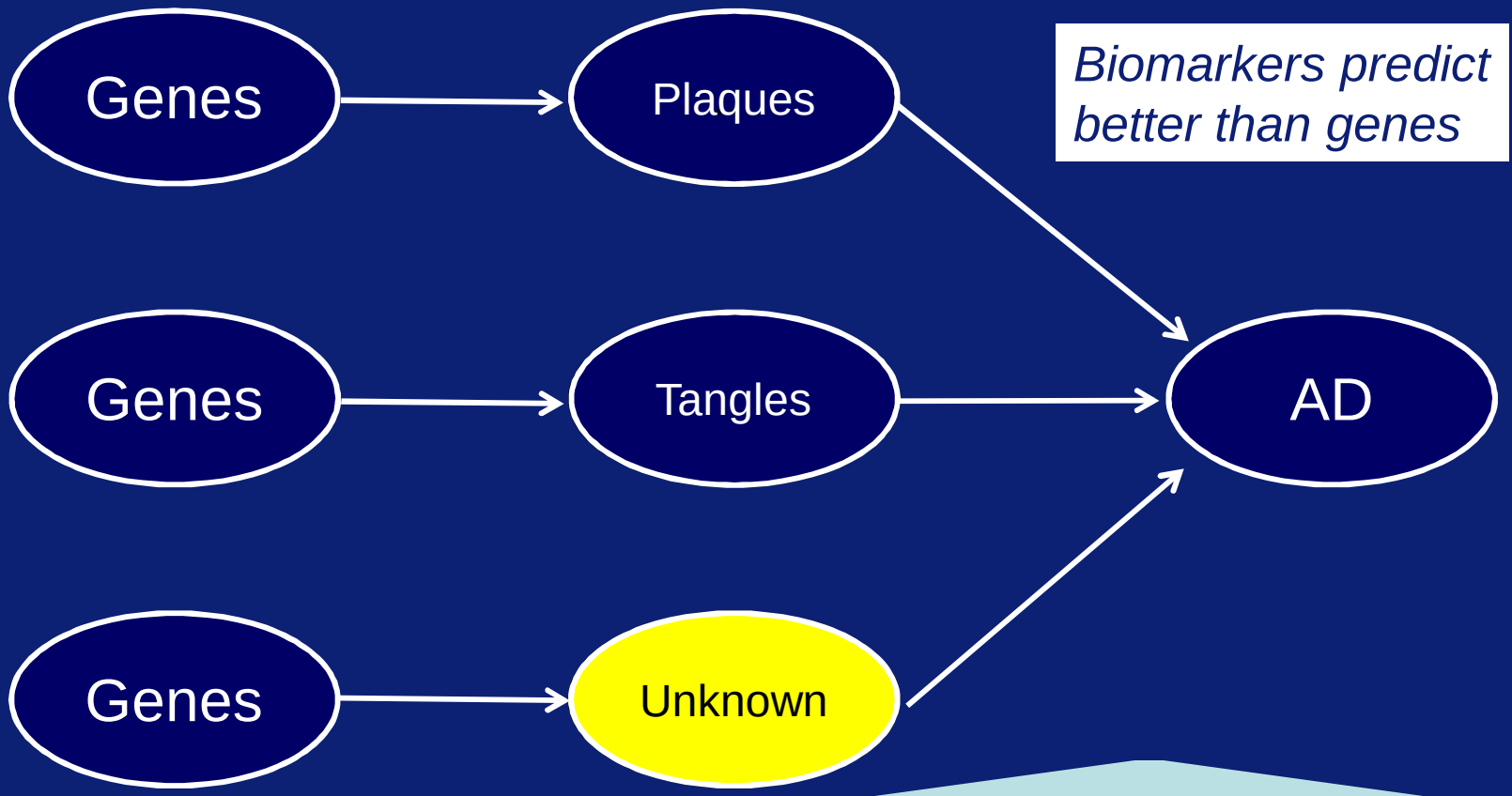
- Complexity of complex disease
- From explaining to predicting
- Challenges

Taking GWAS to new knowledge



McCarthy *Nat Rev Genet 2008*

Next step – beyond genetics!



Explaining versus predicting: the example of the next pope!

The outfit predicts non causally

High-throughput approaches metabolomics

Metabolomics

1

2

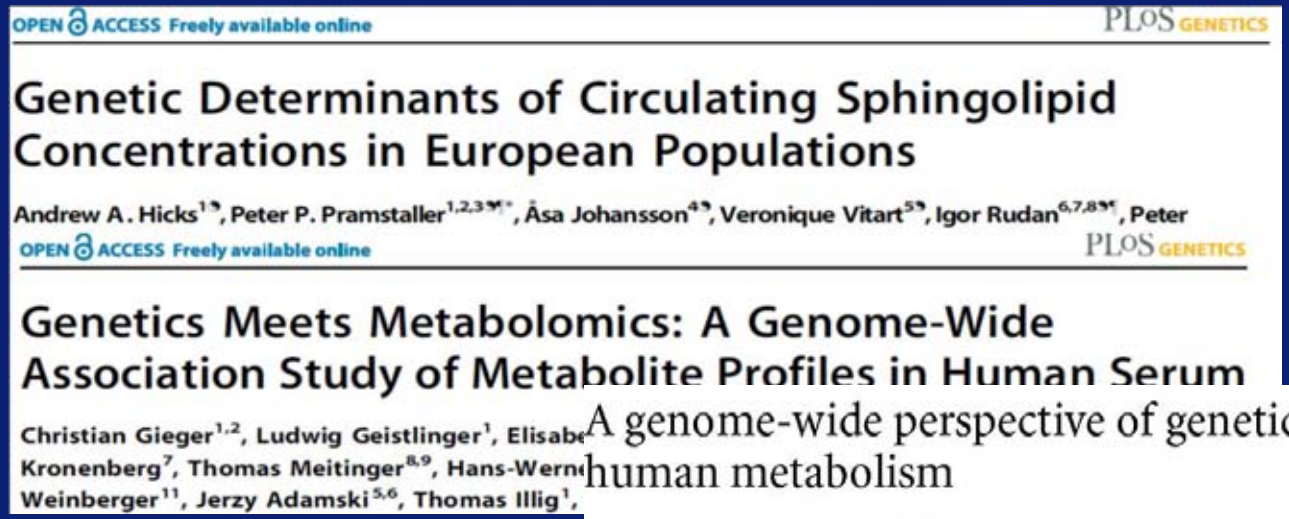
3

High throughput/ high resolution NMR: 190 known and many unknown metabolites

Mass Spectronomy: Biocrates array platform (181 metabolites)

Mass Spectronomy: Targeted Metabolomics Centre, (oxo) lipid platform

Metabolomics is ready for large scale cohorts



OPEN ACCESS Freely available online PLoS GENETICS

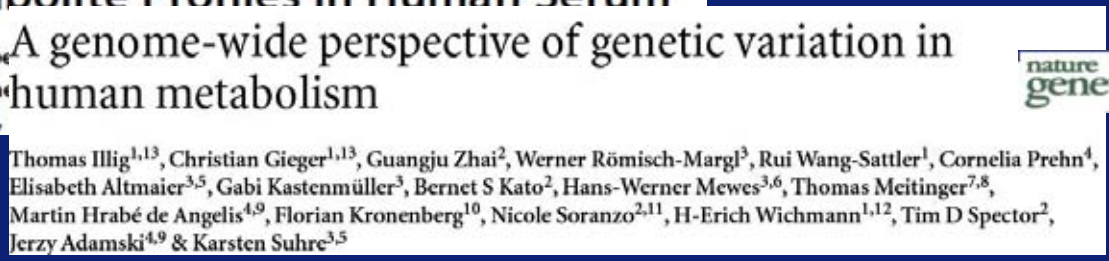
Genetic Determinants of Circulating Sphingolipid Concentrations in European Populations

Andrew A. Hicks[1☯], Peter P. Pramstaller[1,2,3☯*], Åsa Johansson[4☯], Veronique Vitart[5☯], Igor Rudan[6,7,8☯], Peter

OPEN ACCESS Freely available online PLoS GENETICS

Genetics Meets Metabolomics: A Genome-Wide Association Study of Metabolite Profiles in Human Serum

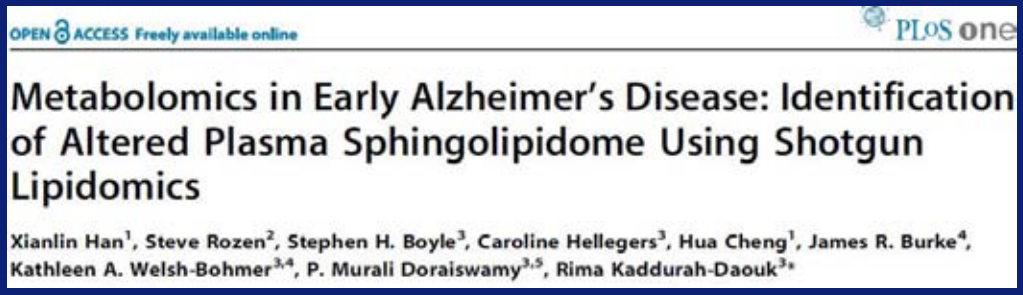
Christian Gieger[1,2], Ludwig Geistlinger[1], Elisabe
Kronenberg[7], Thomas Meitinger[8,9], Hans-Werne
Weinberger[11], Jerzy Adamski[5,6], Thomas Illig[1],



A genome-wide perspective of genetic variation in human metabolism

nature genetics

Thomas Illig[1,13], Christian Gieger[1,13], Guangju Zhai[2], Werner Römisch-Margl[3], Rui Wang-Sattler[1], Cornelia Prehn[4], Elisabeth Altmaier[3,5], Gabi Kastenmüller[3], Bernet S Kato[2], Hans-Werner Mewes[3,6], Thomas Meitinger[7,8], Martin Hrabé de Angelis[4,9], Florian Kronenberg[10], Nicole Soranzo[2,11], H-Erich Wichmann[1,12], Tim D Spector[2], Jerzy Adamski[4,9] & Karsten Suhre[3,5]

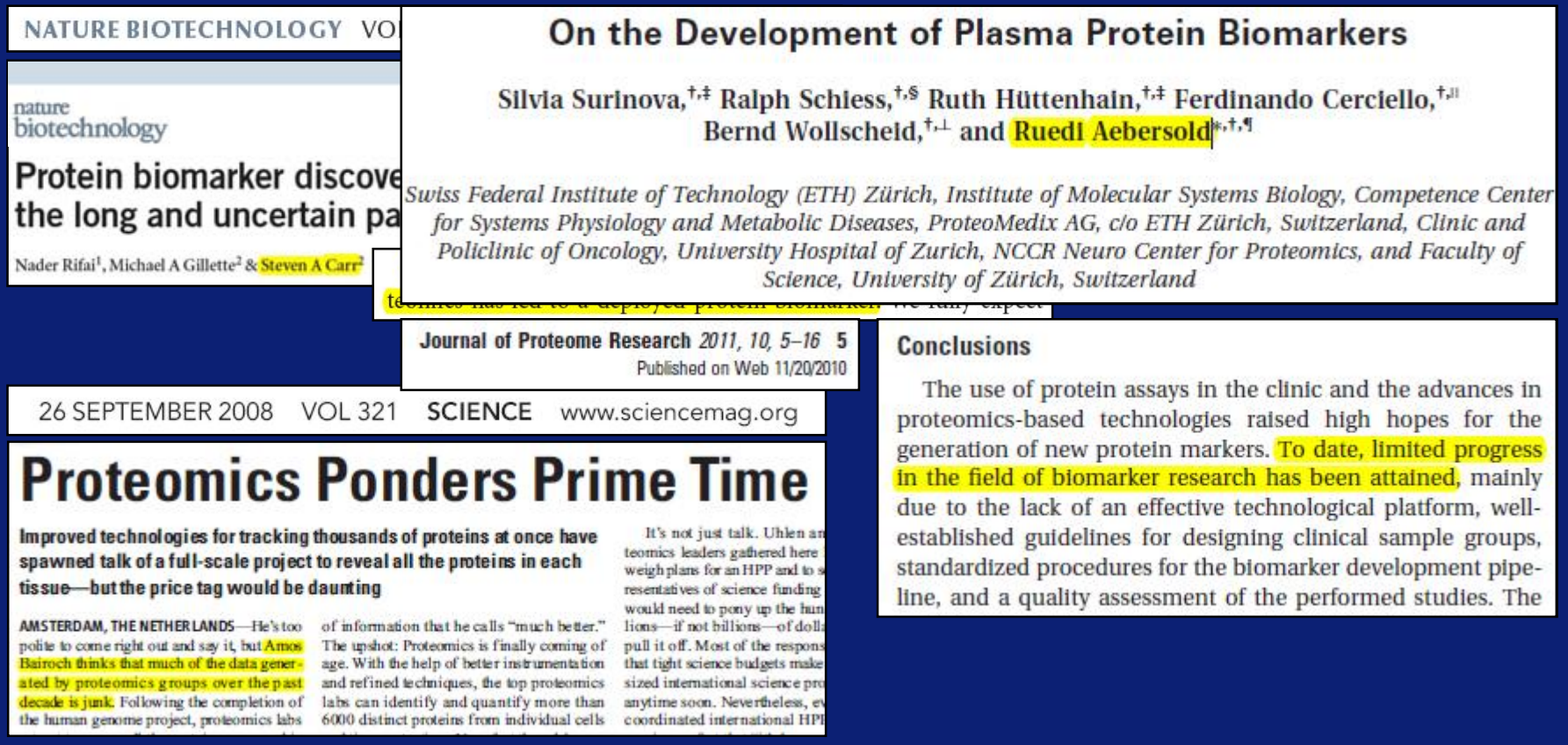


OPEN ACCESS Freely available online PLoS one

Metabolomics in Early Alzheimer's Disease: Identification of Altered Plasma Sphingolipidome Using Shotgun Lipidomics

Xianlin Han[1], Steve Rozen[2], Stephen H. Boyle[3], Caroline Hellegers[3], Hua Cheng[1], James R. Burke[4], Kathleen A. Welsh-Bohmer[3,4], P. Murali Doraiswamy[3,5], Rima Kaddurah-Daouk[3*]

Trials and tribulations in metabolomics



NATURE BIOTECHNOLOGY VO

nature biotechnology

Protein biomarker discove
the long and uncertain pa

Nader Rifai[1], Michael A Gillette[2] & Steven A Carr[2]

On the Development of Plasma Protein Biomarkers

Silvia Surinova,[†,‡] Ralph Schiess,[†,§] Ruth Hüttenhain,[†,‡] Ferdinando Cerciello,[†,∥] Bernd Wollscheid,[†,⊥] and Ruedi Aebersold[*,†,¶]

Swiss Federal Institute of Technology (ETH) Zürich, Institute of Molecular Systems Biology, Competence Center for Systems Physiology and Metabolic Diseases, ProteoMedix AG, c/o ETH Zürich, Switzerland, Clinic and Policlinic of Oncology, University Hospital of Zurich, NCCR Neuro Center for Proteomics, and Faculty of Science, University of Zürich, Switzerland

Journal of Proteome Research *2011, 10, 5–16* 5
Published on Web 11/20/2010

Conclusions

The use of protein assays in the clinic and the advances in proteomics-based technologies raised high hopes for the generation of new protein markers. To date, limited progress in the field of biomarker research has been attained, mainly due to the lack of an effective technological platform, well-established guidelines for designing clinical sample groups, standardized procedures for the biomarker development pipeline, and a quality assessment of the performed studies. The

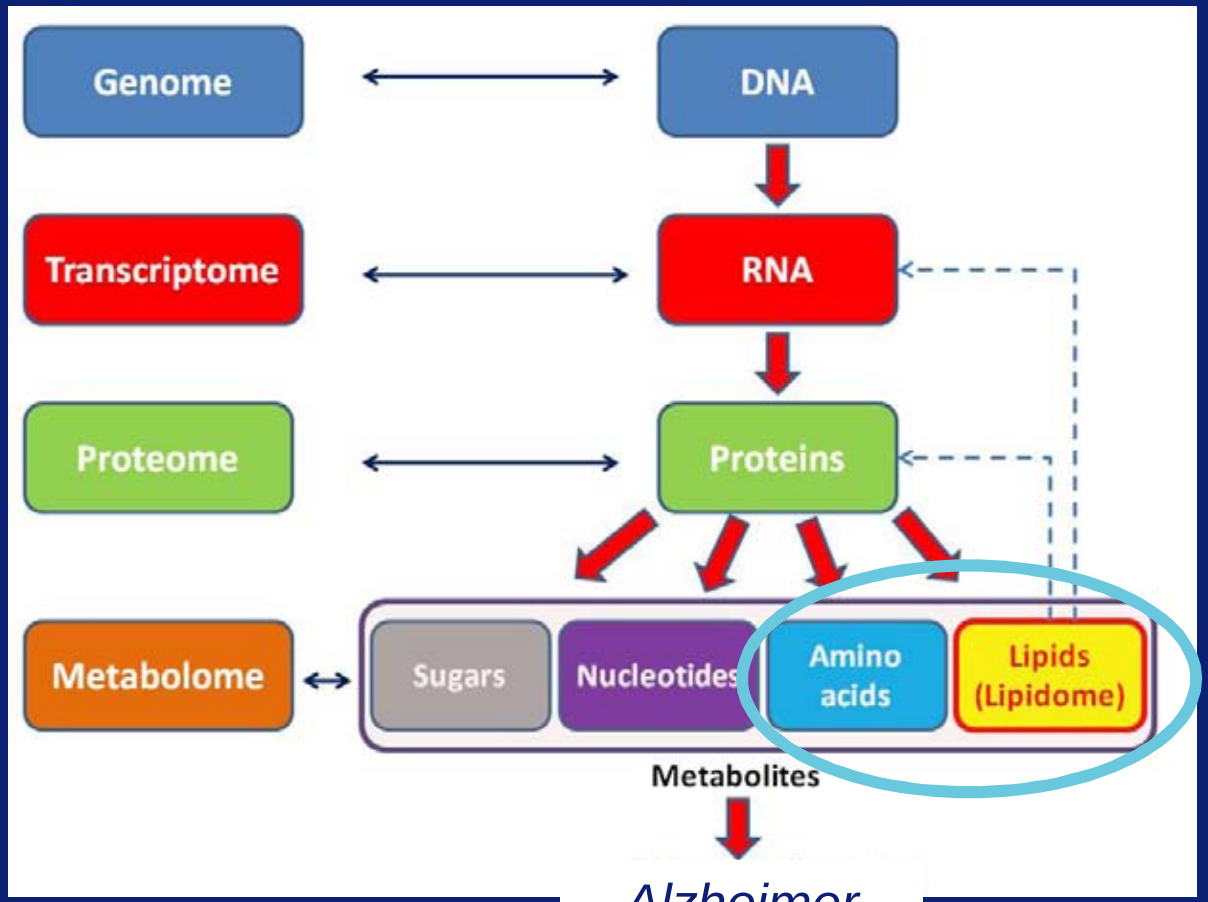
26 SEPTEMBER 2008 VOL 321 SCIENCE www.sciencemag.org

Proteomics Ponders Prime Time

Improved technologies for tracking thousands of proteins at once have spawned talk of a full-scale project to reveal all the proteins in each tissue—but the price tag would be daunting

AMSTERDAM, THE NETHERLANDS—He's too polite to come right out and say it, but Amos Bairoch thinks that much of the data generated by proteomics groups over the past decade is junk. Following the completion of the human genome project, proteomics labs of information that he calls "much better." The upshot: Proteomics is finally coming of age. With the help of better instrumentation and refined techniques, the top proteomics labs can identify and quantify more than 6000 distinct proteins from individual cells

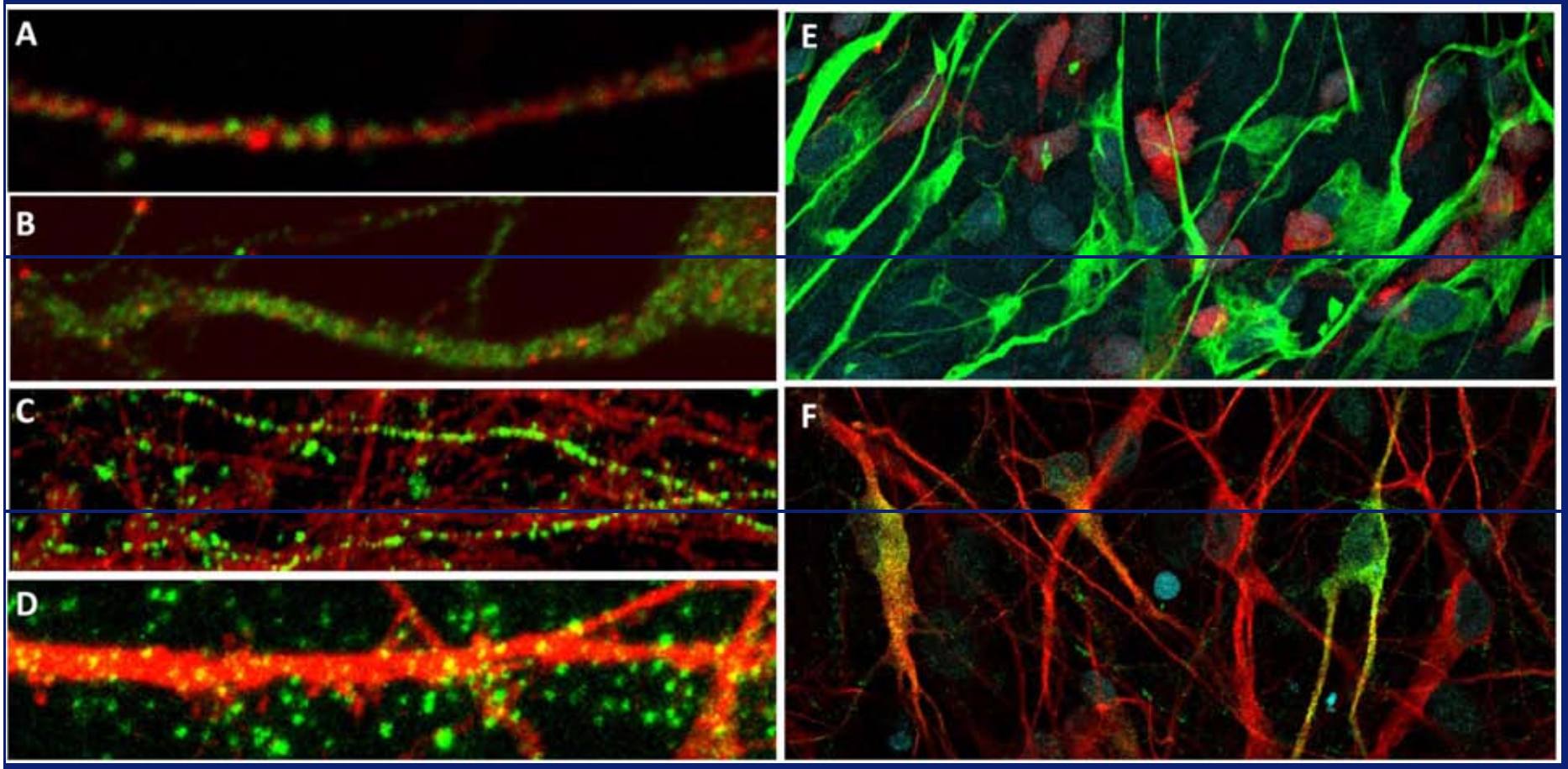
Discovering causal predictors



Genetic contribution to biomarker discovery using metabolomics?

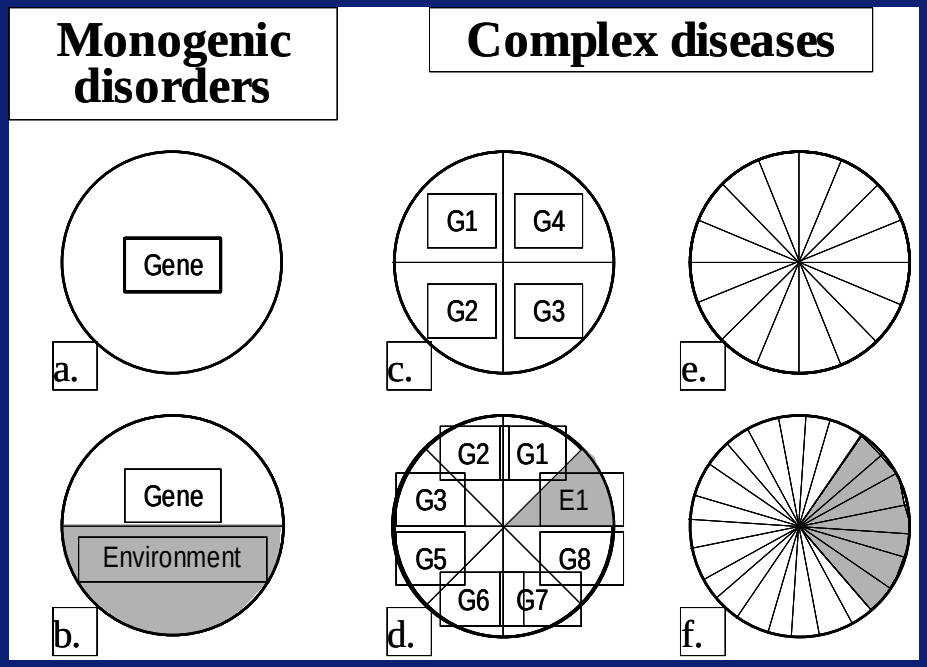
- Find the genes involved in the metabolites
- Relate the metabolites to the disease of interest
- Use Mendelian Randomisation to evaluate causal and non-causal relationships
- But gene expression and metabolites may be tissue specific : here induced pluriformic stem cells (iPSC) may help

What do the new AD genes tell about the metabolome iPSC (neurons, astrocytes, microglia)?



JPND – Metabolomic studies

APP,
PSEN1
PSEN2,
APOE44,
TREM2



*Select patients who carry APOE*4, CLU, PICALM and unknown genes who developed Alzheimer's disease (early)*

Compare the metabolic profile to that of controls who at old age are normal cognitive and at MRI at old age

Take home messages

- The effect of single common variants are irrelevant in complex diseases: you have to consider them in concert
- For a subset of patients and complex diseases, risk prediction based on genes is possible
- The next step in –omics is to translate genomic findings to metabolomic biomarkers and drug targets