Genetics of Alcoholism

Howard J. Edenberg
Indiana University School of Medicine

Outline

- Alcoholism is a complex disease
- Genes with strong effects-the exception
- Strategies & results: COGA
- Functional studies
- Where are we?

Alcoholism (like diabetes) is a complex genetic disease

- Runs in families, but no simple pattern
  - Children of alcoholics are at 2- to 4-fold higher risk
  - But fewer than half become alcoholic
- Risk is affected by genes
- Risk is affected by choice

It is hard to find genes affecting risk for complex diseases

- Phenotypic complexity, heterogeneity
- Multiple genes, each with small effect
- Environmental variability
- Gene-gene interactions
- Gene-environment interactions

Outline

- Characteristics of alcoholism that shaped our strategy
- Genes with strong effects - the exception
- Strategies & results: COGA
- Functional studies
- Where are we?

Metabolism matters!

- Strong protective effects of
  - high-activity ADH enzymes
  - nearly inactive ALDH2 enzyme
    - "endogenous disulfiram"
  - 1/2 to 1/8 the risk.

ADH ALDH

Alcohol **Acetaldehyde** Acetate

ADH1B variations affect risk

- *ADH1B*2* (His48; rs1229984) encodes more active ADH
  - High frequency in East Asians (~70%)
    - Strongly protective against alcohol dependence (~10^{-41}); OR 2-4
  - Low prevalence in Europeans (<5%)
    - Strongly protective (7 x 10^{-10}); OR ~3
      - Not found in GWAS; coverage, frequency

Li et al, in press; Bierut et al, in press

Metabolism (pharmacogenetics) is not everything

- No other genes with as large an effect have been found
- There is a large fraction of the risk that *ADH* and *ALDH* don't explain, particularly in European populations
- So…

How can we identify other genes that contribute to the risk of alcoholism?

Outline

- Characteristics of alcoholism that shaped our strategy
- Genes with strong effects-the exception
- Strategies & results: COGA
- Functional studies
- Where are we?

COGA: Collaborative Study on the Genetics of Alcoholism

- Principal Investigators: B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut
  - Univ. of Connecticut — V. Hesselbrock
  - Indiana University — H. Edenberg, J. Nurnberger Jr., T. Foroud
  - University of Iowa — S. Kuperman, J. Kramer
  - SUNY Downstate — B. Porjesz
  - Washington University — L. Bierut, A. Goate, J. Rice, K. Bucholz
  - Univ. of Calif. (UCSD) — M. Schuckit
  - Rutgers University — J. Tischfield
  - Southwest Foundation — L. Almasy
  - Howard University — R. Taylor
  - VCU — D. Dick
  - NIAAA Staff Collaborators: A. Parsian, M. Reilly

This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA).

COGA

- Large, family-based, genetic study
  - 14,000 interviewed
- Detailed subject characterization
  - SSAGA, Electrophysiology
- Now following adolescents & young adults prospectively

Hypotheses that have shaped our strategies

- Most of the variation that underlies complex genetic disease leads to subtle regulatory differences, not major coding differences- so look across genes.
- Most variations will have a small effect.
- Broad linkage peaks probably harbor several genes that affect the phenotype

COGA strategies

- Family studies: linkage and candidate genes
- Case-control Genome-Wide Association Study (GWAS)
  - Family follow-up
- Family GWAS with follow-up
- Rare variants - next-gen sequencing
- Functional studies

Primary discovery sample

- Families densely affected by alcohol dependence
  - Probands recruited from treatment facilities
  - Families with at least 3 alcohol dependent first degree relatives

Initial strategy

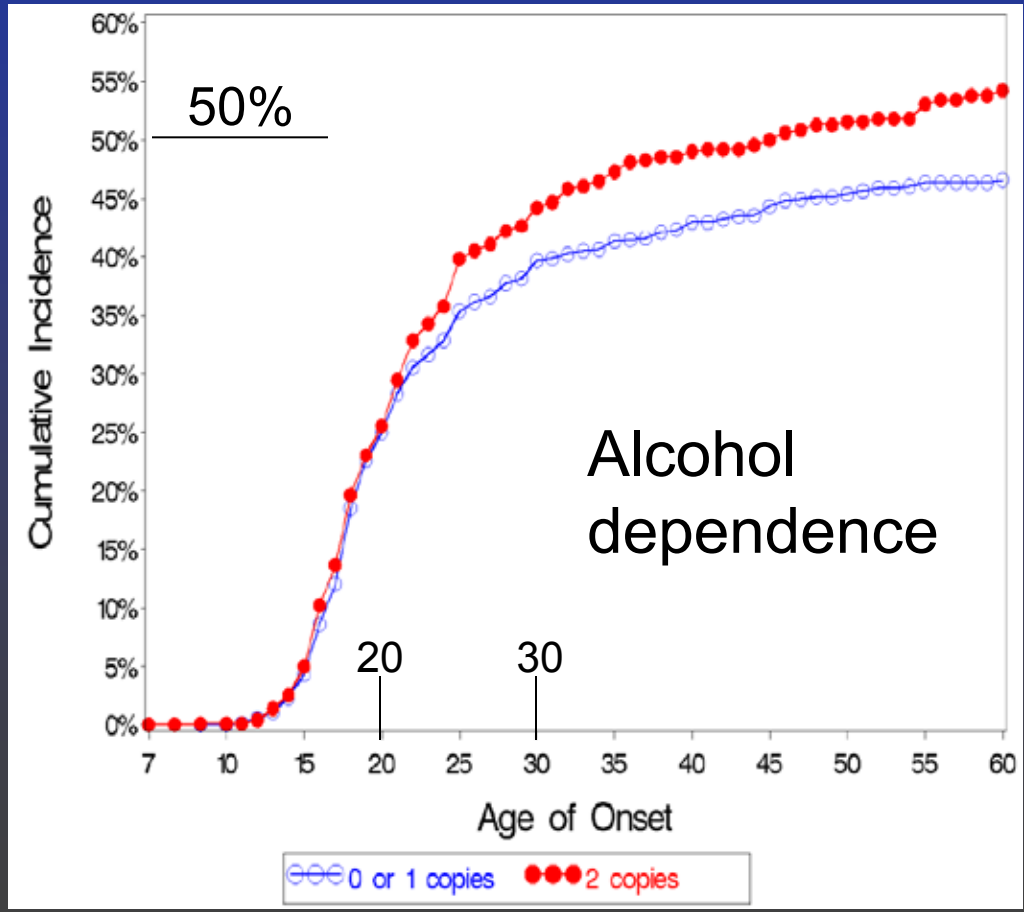
- Linkage studies of the densely-affected families
- Follow-up genotyping
  - Candidate genes in regions of linkage
  - SNPs across regions of linkage
  - Variations across genes- not just coding region
  - Additional candidate genes
- Endophenotypes also analyzed

Linkage and family follow-up: $GABA_A$ Cluster and Alcoholism

- *GABRA2* is associated with alcoholism and with β-EEG (endophenotype)
  - Association concentrated among the more severely affected (e.g. early onset, dependent on other drugs)
  - Effects differ across life-cycle
- Recent evidence (NIAAA, Yale): *GABRG1* - *GABRA2* region also

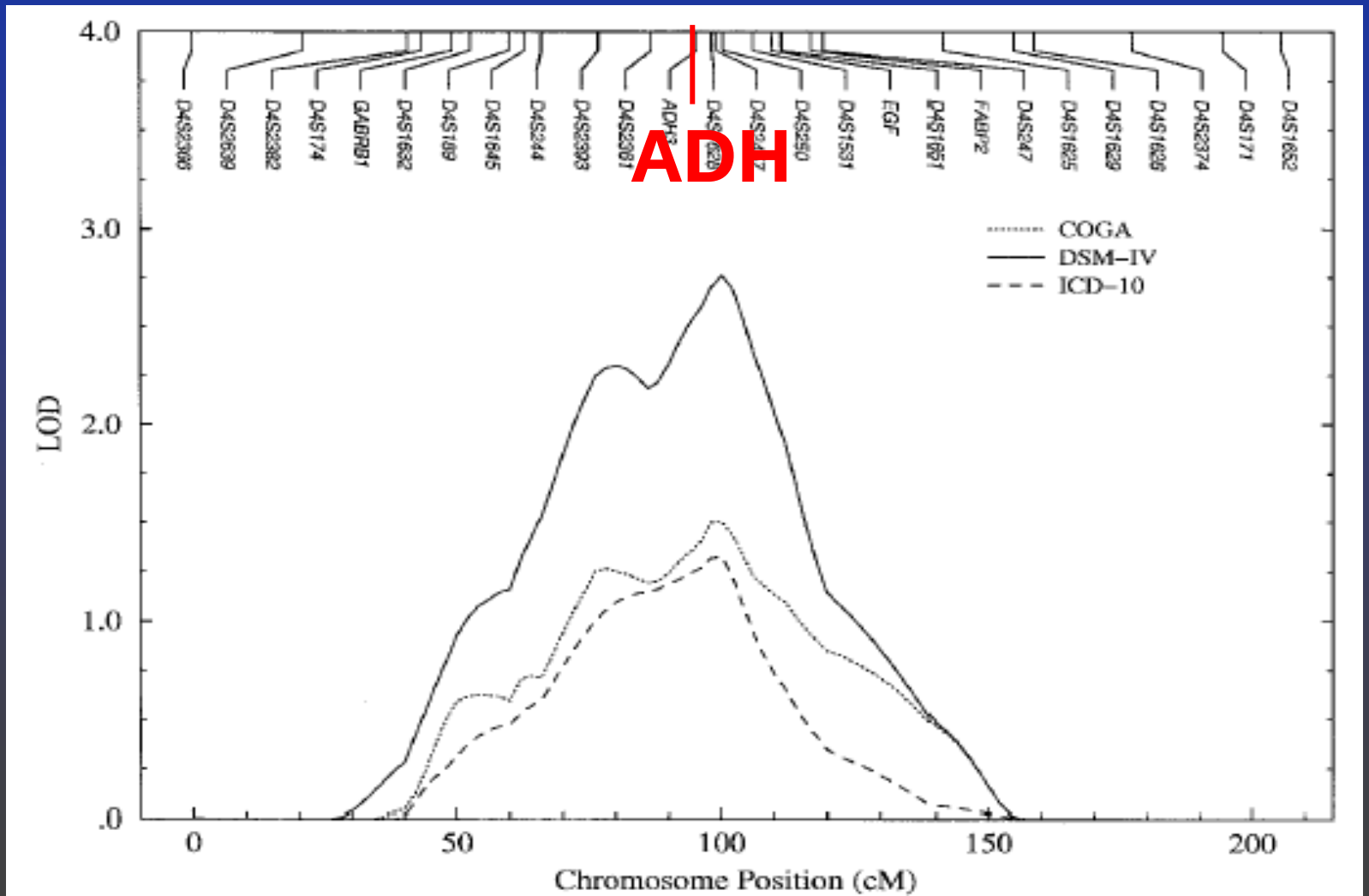
GABRA2: effects of high risk allele differ across the life cycle

- Conduct disorder symptoms in young people
  - Odds Ratio for ≥3 symptoms = 2.0
- Alcohol dependence by mid 20s

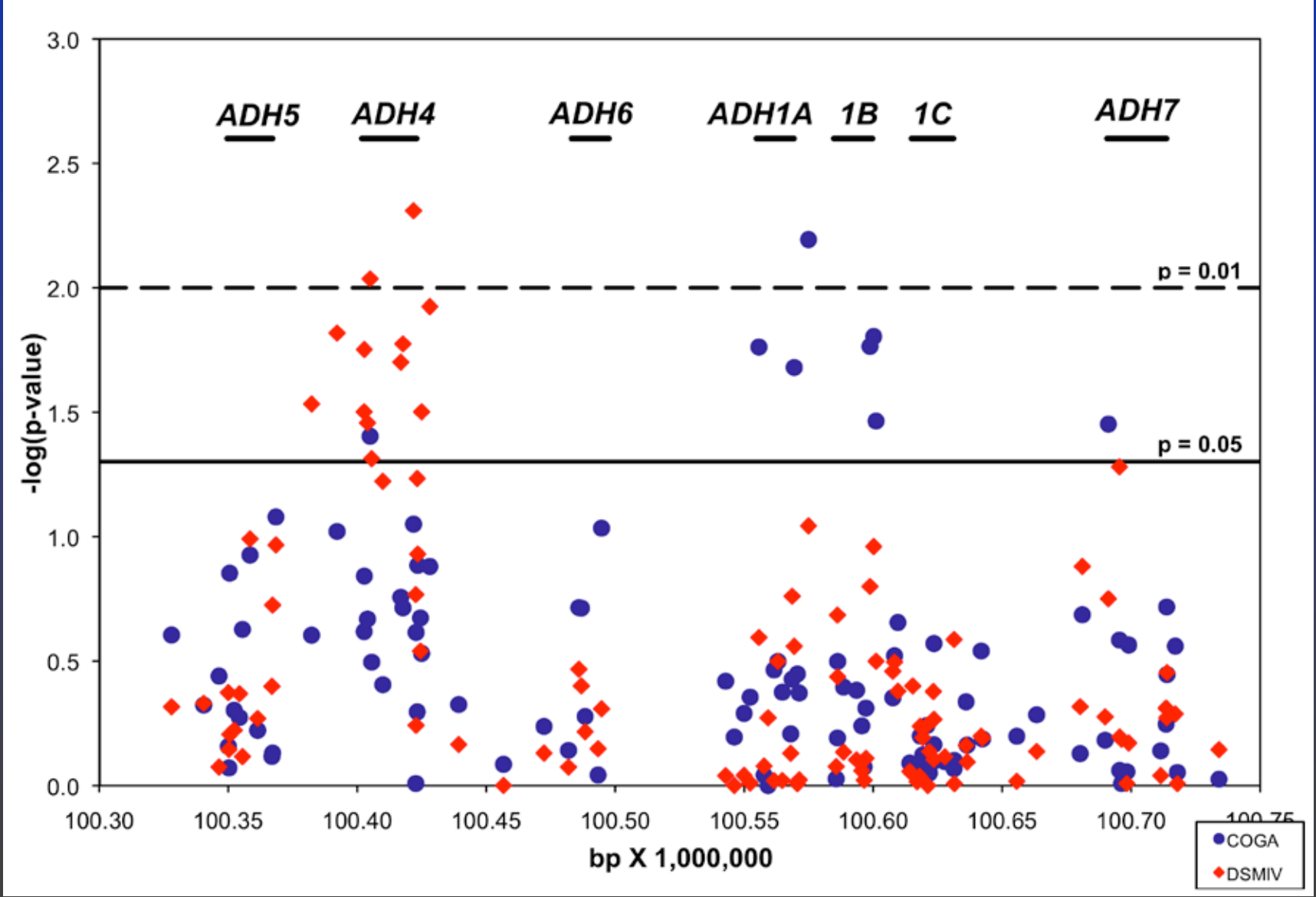


Dick et al. 2006, Behav Genet 36, 577

Broad linkage peak on chromosome 4: multiple genes associated



ADH Gene Cluster

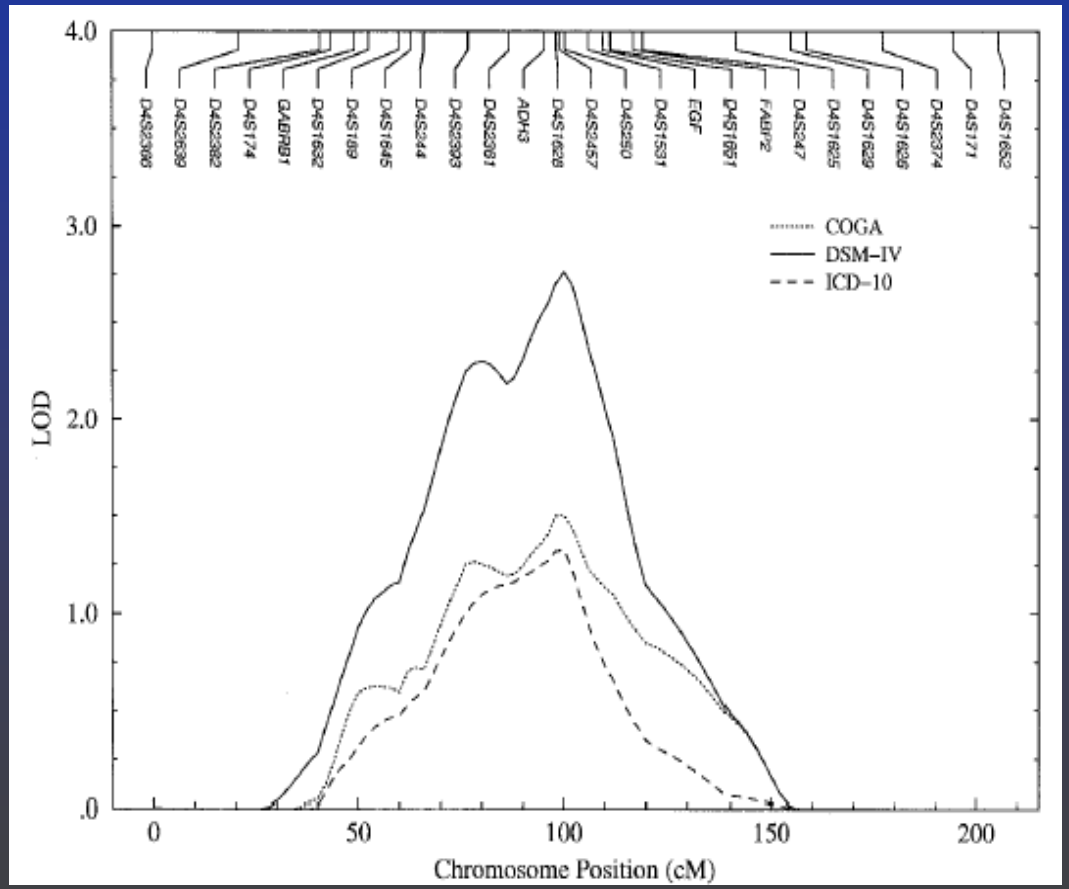


Two regions of association in *ADH* cluster

- *ADH4* driven by MORE severe
  - Early age of regular drinking
  - Early first drunkenness
  - Early onset of dependence
- *ADH1B-ADH1A* driven by LESS severe
- Pharmacogenetics makes sense:
  - ADH1A, ADH1B, ADH1C at low alcohol
  - ADH4 at intoxicating levels

Other genes in the broad linkage peak on chromosome 4

- *ADH genes*
- *NFKB1*
- *SNCA*
- *TACR3*
- *NPY2R*



Systems approach: Links within and between systems

- Given *GABRA2*, examine other $GABA_A$ receptor genes
  - *GABRG3*, *GABRA1*
- Given literature, examine opioid system
  - Kappa system: both *OPRK1*, *PDYN*
- Linkage between systems: *PDYN* is regulated by NF•B (Bakalkin)
  - site near significant SNPs

Our initial strategy (linkage/candidate genes) has been successful

- Genes that influence risk for alcoholism
  - *GABRA2, CHRM2, ADH4, ADH1A, ADH1B, CHRNA5, GABRG3, OPRK1, PDYN, NFKB1, ANKK1, ACN9, NPY2R, CRHR1 …*
- Genes that influence related traits
  - *SNCA, CHRM2, CHRNA5, CHRNA3, CNR1*
- Genes that influence neurophysiology:
  GABRA2, CHRM2, GRM8
- Replications. Continuing work

But…

- Many more genes to find

COGA GWAS: case-control design

- Cases: DSM-IV alcohol dependence
- Controls: not alcohol dependent, not dependent on illicit drugs
- Multiple ethnicities

High density Families 1º discovery sample	Lower density families	Comparison families

GWAS Genotyping

- Funding: NIAAA, NIH GEI (U01HG004438), NIH contract (HHSN268200782096C)

Results

- No SNP was genome-wide significant (similar to most GWAS: underpowered)
  - $<10^{-5}$: 11
  - $<10^{-4}$: 97 (27 also with early onset)
- Regions with multiple SNPs $\leq 10^{-4}$

GWAS: interesting genes

- *BBX* - bobby sox homolog
- *CARS* - cysteinyl-tRNA synthetase
- *NAP1L4* - nucleosome assembly
- *SLC2A14* - glucose transporter
- *SLC37A3* - glycerol-3-P transporter
- *OSPBL5* - oxysterol-binding

How can we prioritize genes and regions?

- Replication (difficult: 'winner's curse')
- Support from
  - clustering of SNPs
  - related phenotypes (early onset)
  - follow-up in families (PDT)
  - gene expression studies

Support for SNPs from prior GWAS (Treutlein et al., 2009)

- Replicated with same risk allele:
  - *GATA4** (transcription)
  - *ID4* (transcription)
  - *ADCY3** (second messenger)
  - *PRKCA* (second messenger)
  - *SYNE1** (neurological disease)
  - *ARL6IP5* (inhibits Glu transporter)

*among top in our early onset analysis

Pathway analysis

Ingenuity Canonical Pathways	pvalue	Molecules
Amyotrophic Lateral Sclerosis Signaling	0.0020	GRIN3B,HECW1, GRIN2C, CAT,GRIA4,CASP7
GABA Receptor Signaling	0.0041	GABBR2,GABRR2,GPHN, GABRP
Glutamate Receptor Signaling	0.0091	GRIN3B,GRIN2C,SLC1A1, GRIA4
Calcium Signaling	0.0363	GRIN3B,CAMK2D,ITPR2, GRIN2C,CHRNA7,CHRNB3
Neuropathic Pain Signaling In Dorsal Horn Neurons	0.0468	GRIN3B,CAMK2D,ITPR2, GRIN2C

COGA Family GWAS

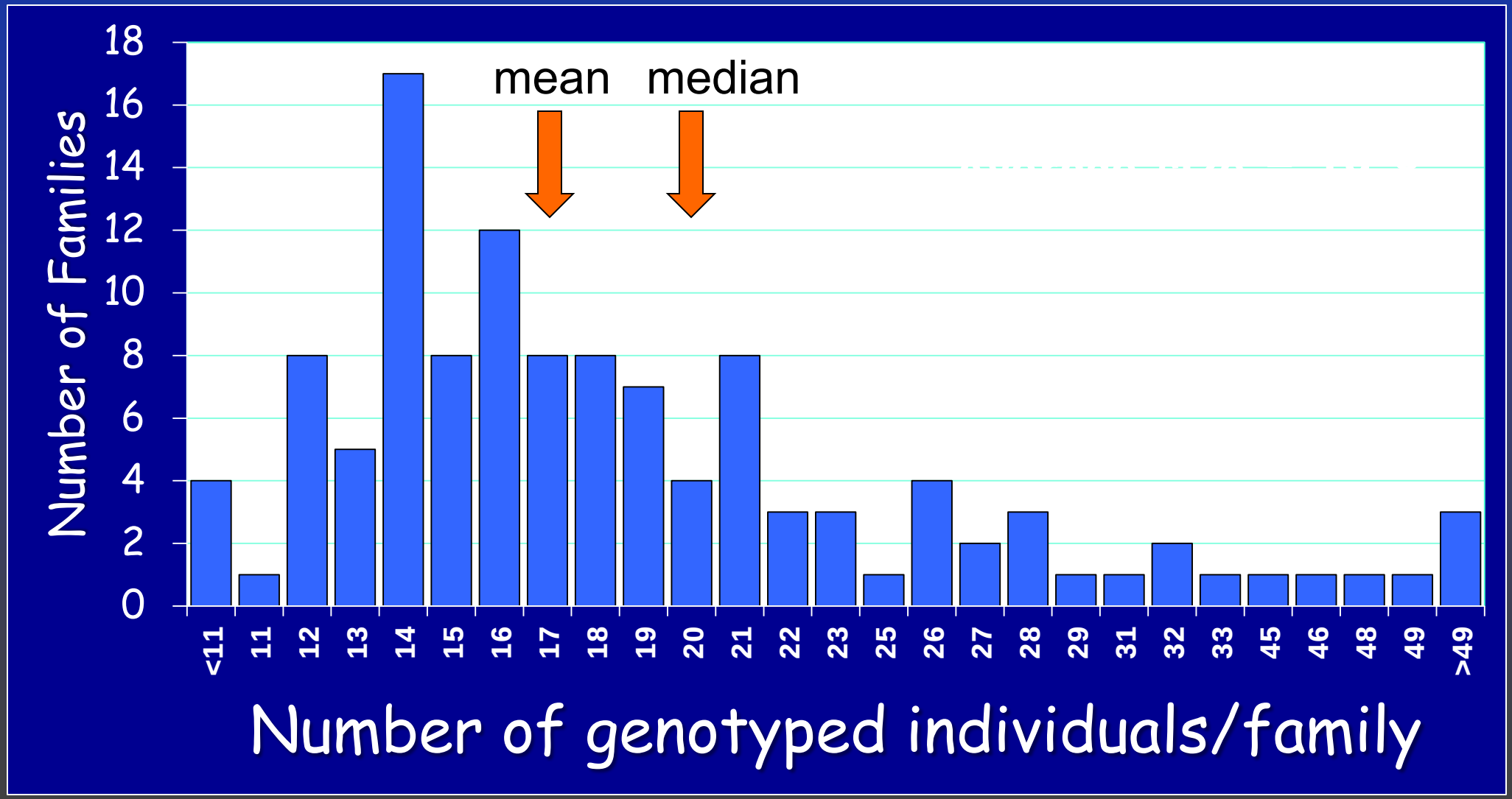
- Large families
- Many alcohol dependent individuals
- Electrophysiological measurements
- European-Americans (to reduce heterogeneity)

Large, high density
Families

COGA Family GWAS
118 large families

	Mean	Median	Range
Genotyped individuals	19.8	17	6 - 70
Alcohol dependent	6	5	1 - 31
EEG data	14.2	13	6 - 70

Distribution of Family Size



Association Analysis: work in progress

- Imputation still in progress
- Quantitative trait: symptom count (0-7)
  - Dependence diagnosis as 2º phenotype
- Applying several analytical methods
  - Covariates: Age at evaluation, Sex, Cohort effects
- Looks promising

Next steps

- Seek replication in other datasets
- Prioritize findings for followup
  - multiple SNPs, methods
  - interesting genes, variants
- Test in full COGA sample
  - many more families and individuals

GWAS are a powerful approach, but with limitations

- GWAS targets common variants
  - Expect most common variants to have small effects (natural selection)

Rare variants may also contribute

- Hypotheses:
  - Genes whose products are involved in pathways that affect risk for alcoholism are likely to have both common variants with small effect and rare variants with larger effect
  - Rare(r) variants that increase risk for disease more likely to be found in affected subjects (may not be in 1000 genomes)

Common and rare variants in same gene

- Cystic Fibrosis: a classic "simple Mendelian disorder" (autosomal recessive)
- CFTR gene identified 1989
  - Cases shared a relatively common polymorphism: F508del, ~66% of cases
- BUT: 1721 other mutations are known, and mutations not yet found for many cases
  - Allelic heterogeneity, rare mutations

Two strategies

- Targeted resequencing in regions with evidence for association
- Exome sequencing of extreme families

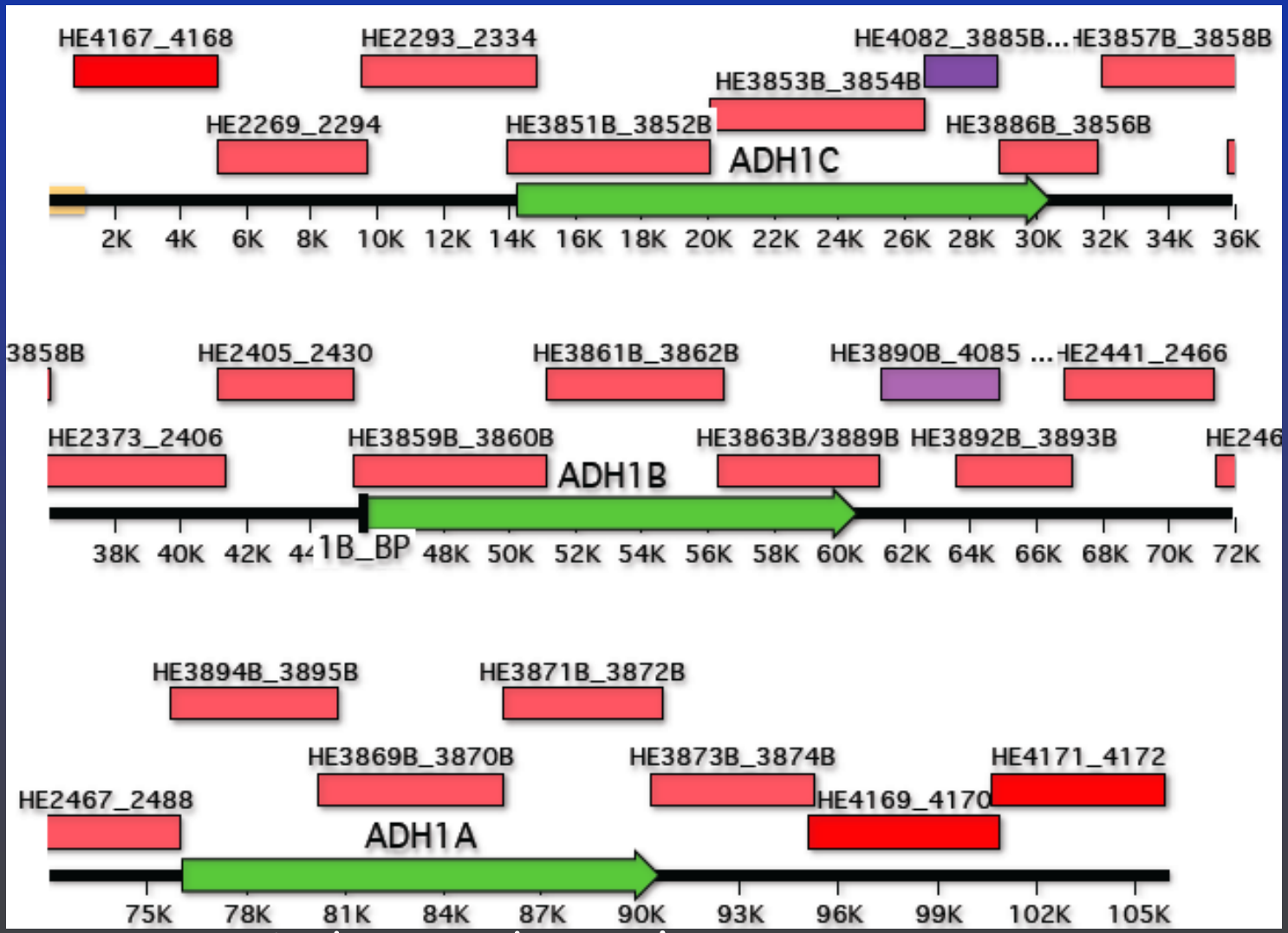
Targeted resequencing in pools

- Pool DNA from subjects (96/pool)
  - Organize pools by phenotype
- Amplify individual fragments (PCR)
- Combine equimolar amounts of each fragment
- Multiplex sequencing (barcoded)
- Statistical analysis to detect variants

Re-sequencing targets

- GABA-receptors
- *ADHs*
- Opioid system [already found functional variant with standard sequencing: *OPRK1*]
- Chromosome 11 region from GWAS
- Nicotinic receptors
- Muscarinic receptors

ADH1C-ADH1B-ADH1A



Coding and regulatory variants

Variant discovery: distinguishing rare alleles from noise (sequencing error)

- Idea: noise will show up as 3rd and 4th alleles (assuming 2 allele SNP)
  - Conservative modeling: is the number of hits for the 2nd allele greater than for the 3rd?
    - Set α (false positive rate)
    - Set $\beta = 0.1$ (power = 90%)
- Test whether SNP is replicated in second experiment

Error model works well as judged by technical replicates

# chromosomes detected	Average % confirmed	Min % confirmed	Max % confirmed
1	83.7%	80.5%	91.3%
2	99.9%	99.5%	100.0%
3	100.0%	99.6%	100.0%
4	100.0%	99.9%	100.0%

Confirmation by independent technique in progress

Challenge: prioritization

- Current estimates: 6-30 x 10^6 bp differences between individuals
- Which are related to the phenotype?
- Need bioinformatics to prioritize SNPs for follow-up
  - Function, position
- Test inheritance and association with alcoholism in families

Rare variants of larger effect: exome sequencing in families

- Coding changes are likely to have larger effects, easier to interpret
  - Exome sequencing
- Extreme families:
  - densely affected
  - early onset
  - extreme electrophysiology
  - Linkage: high lod score

Outline

- Characteristics of alcoholism that shaped out strategy
- Genes with strong effects-the exception
- Strategies & results: COGA
- Functional studies
- Where are we?

Functional analyses

- Promoter variations affecting gene expression [*ADH4, ADH1B, OPRK1*]
- Global gene expression studies
- Allele-specific gene expression
- Alternative transcripts [*GABRA2*]
- Epigenetic studies
- Model organisms: rats, mice, flies, worms

Conclusions

- Variations upstream of *ADH1B* affect gene expression
  - Associated with risk for alcoholism
- Variations proximal to *ADH4* affect expression
  - Promoter, 3' region
- Variations distal to *ADH4* (upstream enhancer ADH-4E) affect expression

Other functional studies

- Global gene expression differences related to alcoholism, alcohol exposure or preference for alcohol (arrays, seq)
  - Human autopsy brains
  - Human lymphoblastoid cells
  - Rat brains
- Allele-specific gene expression
- DNA methylation/epigenetics

Genes that differ in expression after ethanol exposure that were implicated in GWAS

- *BBX*
- *EPHB1*
- *AGPAT5*
- *CAMK2D*
- *PHLDA2*
- *PRKD2*
- *GPHN*
- *SOX6*
- *OXTR*

Pathways affected by alcohol

- Pro-inflammatory (especially NF-κB)
- IL-6 signaling
- Hepatic fibrosis/stellate cell activation
- PPAR signaling

Genes with expression differences in alcoholics

- 567 probe sets
  - 43% also affected by ethanol
- Some interesting ones:
  - *KCNA3*
  - *PRKCE*
  - *HDAC7*
  - *PDE4A*
  - *VDR*
  - *PNOC*

Outline

- Characteristics of alcoholism that shaped our strategy
- Genes with strong effects-the exception
- Strategies & results: COGA
- Functional studies
- Where are we?

Where are we?

- Tremendous progress
  - Multiple strategies- families, case-control studies, candidate genes, GWAS
  - Specific genes identified
  - Candidates awaiting confirmation
  - Exploring associations on many levels
    - Molecular, phenotypic

We're not looking for "the gene for alcoholism"

- There is no such gene!
- There are variations in many genes that lead to variations in physiology that affect the risk that we will become alcoholic (or depressed, diabetic…) and affect the course of disease

Where are we going?

- Finding more genes
  - Expanded family sample (more than doubled)
  - Meta-analyses
  - Confirmations in other datasets
  - Systems analyses

Where are we going?

- Exploring function on many levels
  - Molecular and cellular studies
  - Epigenetics
  - Endophenotypes and other disorders
  - Effects across the lifespan
    - Large adolescent sample
    - Prospective study (12-25)

Pharmacogenomics: can we predict which medications help specific individuals?

- Some initial studies (1 or a few variants)
  - Naltrexone: *OPRM1; OPRK1?*
  - Bromocriptine, olanzapine: dopamine
  - Acamprosate, topiramate: glu receptors
  - Ondansetron: *HTT* (serotonin)
  - Psychotherapy: *GABRA2*
- BUT: need more comprehensive approach.
- Need to bank samples in clinical trials!

Caution: complexity of mapping genotype to phenotype

(Dowell et al., 2010, Science 328:469)

- Yeast knockouts
  - *Saccharomyces cerevisiae* S288c (reference) *vs.* Σ1278b (close relative)
    - As similar as 2 humans
  - Test for conditional lethal genes
    - 5000 genes tested

COGA: Collaborative Study on the Genetics of Alcoholism

- Principal Investigators: B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut
  - Univ. of Connecticut V. Hesselbrock
  - Indiana University H. Edenberg, J. Nurnberger Jr., T. Foroud
  - University of Iowa S. Kuperman, J. Kramer
  - SUNY Downstate B. Porjesz
  - Washington University L. Bierut, A. Goate, J. Rice, K. Bucholz
  - Univ. of Calif. (UCSD) M. Schuckit
  - Rutgers University J. Tischfield
  - Southwest Foundation L. Almasy
  - Howard University R. Taylor
  - VCU D. Dick
  - NIAAA Staff Collaborators: A. Parsian, M. Reilly

This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA).

Acknowledgments

Jeanette McClintick
Xiaoling Xuei
Yunlong Liu
Sirisha Pocharredy
Jun Wang
Ronald Jerome
Sowmya Jairam
Hui-Ju Chen
Huijin Tian
Tatiana Foroud
Leah Wetherill
Dan Koller
Rachel Thowe
Sheena Becker
Jun Li
Praveena Yarru
Bill McBride
Rick Bell

National Institute of Alcohol Abuse and Alcoholism

ADH: R37AA006460

Gene expression in rat brain: U01AA016660