Survey Results

Clinical areas of emphasis (check all that apply):

Aging	9
Alzheimer's disease	15
Vascular dementias	4
Lewy body dementias	8
Fronto-temporal dementias	10

Focus areas (number all that apply in order of priority):

Causative (major) genes	1, 1.5, 2, 2, 2, 3
Modifying/susceptibility genes	1, 1, 1, 1, 1, 1, 1, 2
Genes affecting response to treatment	2
Other:	3

Survey Results

Predominant study designs (check all that apply):

Family-based study	11
Isolated or special population(s)	9
Population association study	6
Other	2

Comments:

Multiplex and sib-pair families
Case control series
Native American risk factors for AD
Discordant sib-pairs being collected as well as concordant multiple sibs

Major methodologies in use (check all that apply):

Traditional linkage study	8
SNP study	8
Gene expression/microarray study	10
Other:	3

Bioinformatics tool development (Y/N):

Yes	2
No	8

Respondents

Case Western
Columbia
Duke University
Harvard-MGH
Mayo
Northwestern
OHSU
Rush
University of Arizona
UCSD
UCLA
University of Kentucky
University of Michigan
University of Pennsylvania
University of Texas SW
Washington University

Early Lessons from Human AD Gene Expression Studies

Patricia Kramer, Ph.D.
Departments of Neurology and Molecular & Medical Genetics
Oregon Health & Science University

Hypotheses

- Gene Expression Differs Significantly with Aging and in AD:
  - by severity of disease/health status
  - by course of disease
  - by intrinsic genetic characteristics
  - by environmental characteristics

- Microarray analysis as a hypothesis-*generating* tool
- Pilot project:
  - Explore the potential of gene expression analysis to understand brain aging and neurodegeneration

OADC Microarray Working Group

Jeff Kaye
Patti Kramer
Barnaby Sheridan
Geoff Murdoch
Joe Quinn

Beth Wilmot
Sri Nagalla
Jeanne O'Malley
DJ Fife
Mark Turner

Extraction of sample

ADC and OBAS

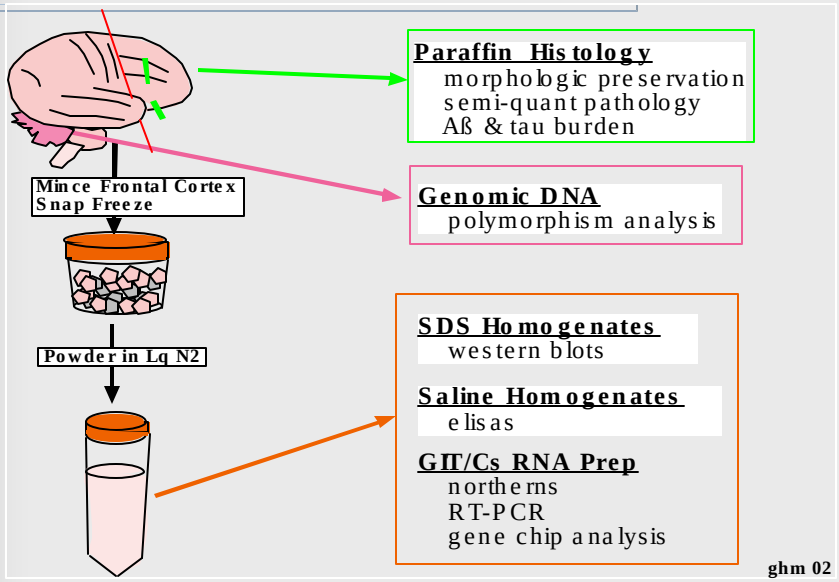
PMI < 24 hours
Neuropathology dx: AD, AP, or healthybrain
Frozen brain tissue prepared for genetic studies
Extensive clinical characterization

50 subjects

High quality RNA
Representation across range of age of onset
Cost

19 subjects

Brain Tissue Preparation



General Description of sample

- Sample composition:
  - 6 controls
    non-demented; average healthy
    age at death: 80-104 yrs
  - 13 AD cases
    onset 41-94 years
    41-64 yrs: 4
    65-84 yrs: 5
    > 85 yrs: 4

Sample characterization

- Cognitive measures
  - MMSE
  - CDR, CDR_sob
  - NCSE
  - SIB
- Functional measures
  - ADL
  - IADL
- Genetics
  - ApoE
  - HLA
  - Family hx
- Neuropathology
  - 2F plaques
  - 2F tangles NT
  - last Braak stage
- Other
  - age onset/disease duration
  - sequential MRI data
  - cell lines
  - Clinical Diagnosis
  - Pathology Diagnosis

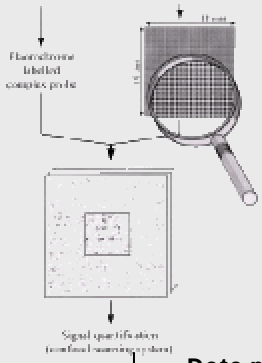
- *Clinical* diagnosis was consistent with *Neuropathological* diagnosis in 18 of 19 subjects

  94-yr-old “control” (high cognitive and functional scores; many plaques and tangles)

Methods

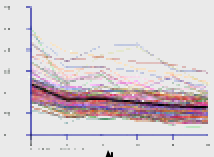
19 RNA samples: "probe"
3 replicates, ea at 2 laser intensities

Research Genetics clone set (13,000 genes/ESTs): "target"



Cluster analysis, PCA; confirmation studies

Visualization



Data processing:
Image analysis, normalization, **significance filtering**

Data Analysis, Stage 1

1. Image analysis on array data
2. Noise assessment
3. Average technological replicates
   a. $\log_2$ transformation applied
   b. global centering within triplicate sets
   c. statistical outliers excluded within triplicates
   d. mean expression level calculated for each spot within triplicate sets
4. Normalize across arrays
   a. median-centered

Data Analysis, Stage 2

1. **Define groups (phenotypes) for comparison**

1. **Identify set of genes that show a "significant difference" in expression**

SAM: Significance Analysis of Microarrays*

- Gene expression differences can be gene-specific
- d-statistic: ratio of change in gene expression to the s.d. of expression values for that gene
- Permutation method used to determine genes identified by chance
    - False Discovery Rate (FDR)
    - user-adjustable

*Tusher VG, Tibshirani R, Chu G. (2001) PNAS 98 (9), 5116-5121.

SAM

- Allows correlation of gene expression levels to multiple types of clinical variables
  - 2-class (paired, unpaired)
  - multiclass
  - quantitative
  - survival time

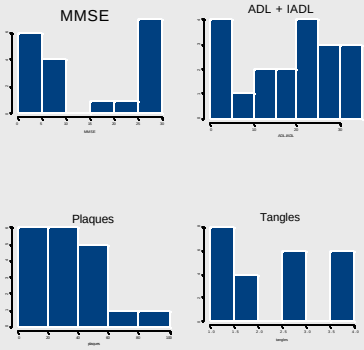
WE HAVE LOTS OF GENES

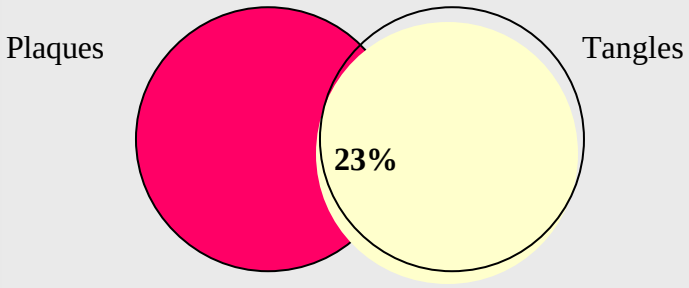
Finding genes is not the problem

➔ *The **challenge** is to relate these genes to AD and aging*

SAM: results

			# Significant Genes		
		Type of variable	15% FDR	5% FDR	lowest %FDR
Disease State					
AD		2 class	1139	557	0.8
Cognitive Tests					
MMSE<24 & CDR>0.5		2 class	2849	1291	0.2
Functional Tests					
ADL + IADL		quantitative	1247	16	4.1
Neuropathological Data					
Plaques		quantitative	240	0	9.0
Tangles		multiclass	2177	354	2.1





Evaluation of Significant Genes

Consider

- "intersection" of genes from related analyses
- linked regions
- PATHWAY relationships

Prioritize

- compelling face validity
- some validity
- no obvious connection
- ESTs

Evaluation of significant genes

- Visualization techniques
- Cluster analysis, Principal Components Analysis, etc
- Confirm results by other means
  - RT-PCR, Northerns
  - proteomic studies
  - functional studies (transgene or knockout animal studies)

Next Steps

Plan new analyses

- Construct compound variables
  - Discriminant analysis
- Refine hypotheses with more specifically defined phenotypes

Lessons

- Phenotype matters
- How one chooses a set(s) of “significant genes” for subsequent analysis and confirmation studies is critical