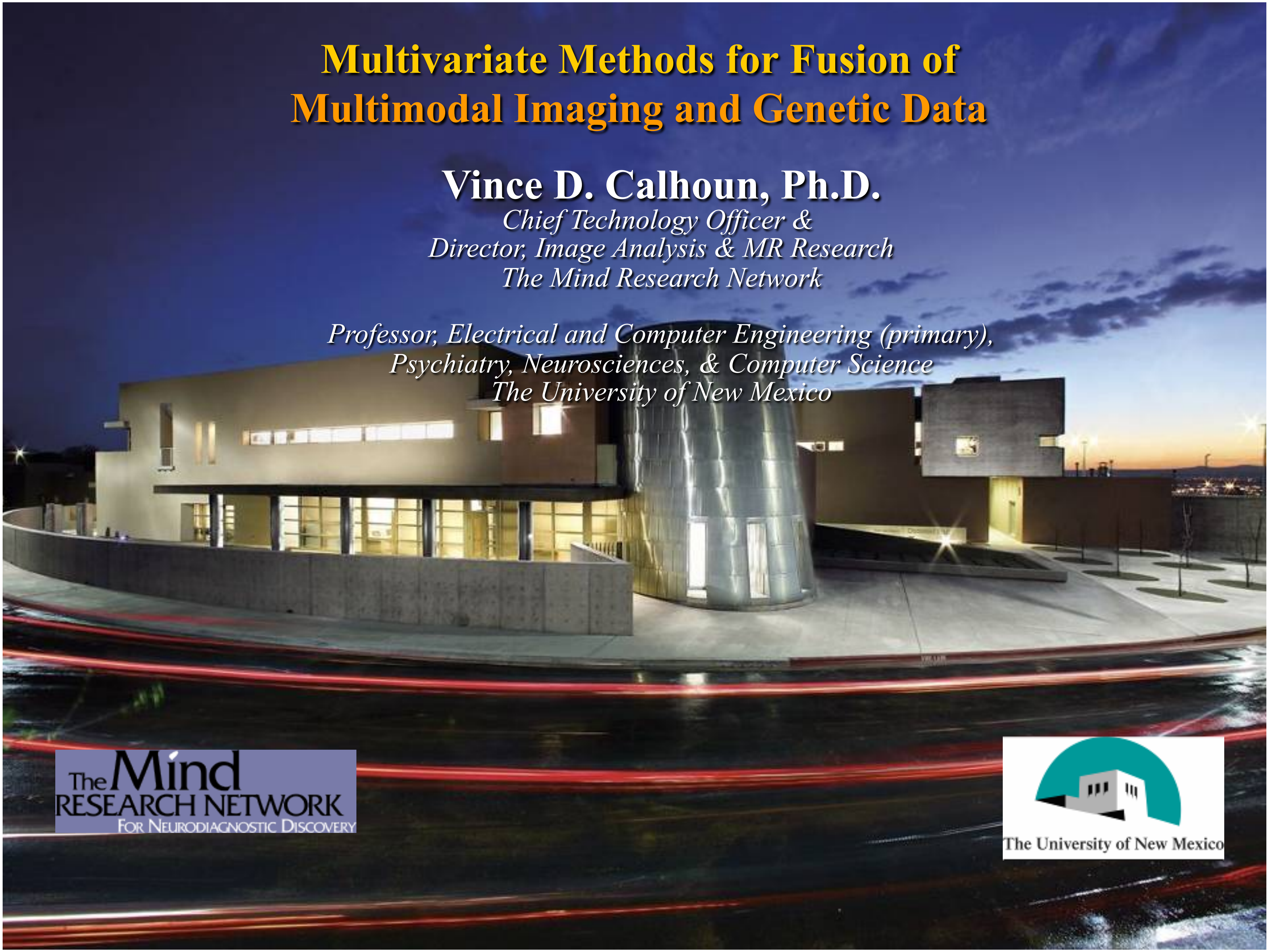


Multivariate Methods for Fusion of Multimodal Imaging and Genetic Data

Vince D. Calhoun, Ph.D.
Chief Technology Officer &
Director, Image Analysis & MR Research
The Mind Research Network

Professor, Electrical and Computer Engineering (primary),
Psychiatry, Neurosciences, & Computer Science
The University of New Mexico

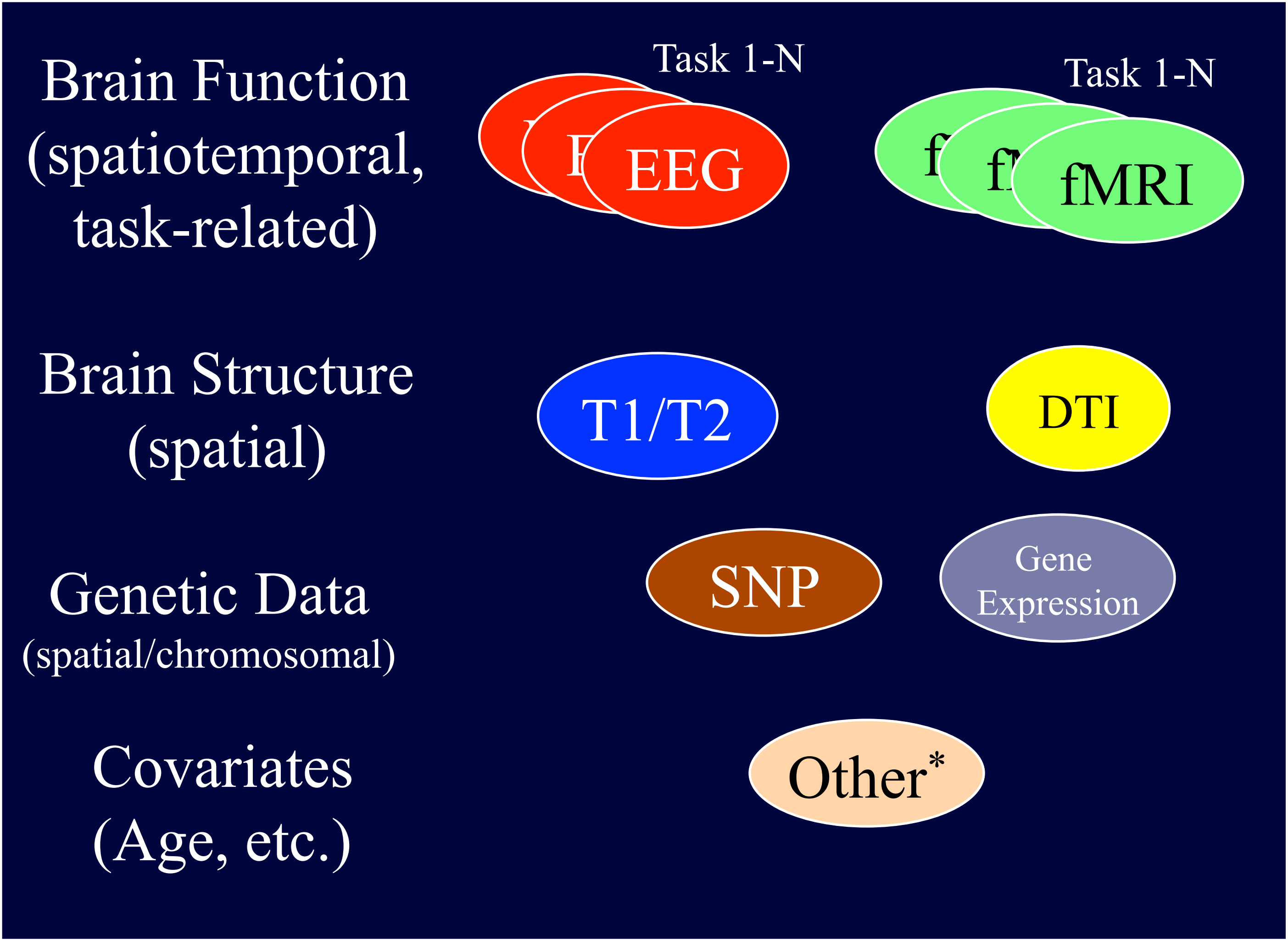
The Mind Research Network
For Neurodiagnostic Discovery

The University of New Mexico

Overview

- Motivation
- Joint ICA
    - ERP/FMRI
    - Multitask-fMRI
    - sMRI/fMRI
    - CC-ICA/Biomarker Identification
- mCCA
    - sMRI/fMRI/ERP
- mCCA+jICA
    - DTI+fMRI
- Parallel ICA
    - SNPs
    - Epigenetics
- Conclusions

Multimodal Data Collection



Brain Function (spatiotemporal, task-related)

Task 1-N

EEG

Task 1-N

fMRI

Brain Structure (spatial)

T1/T2

DTI

Genetic Data (spatial/chromosomal)

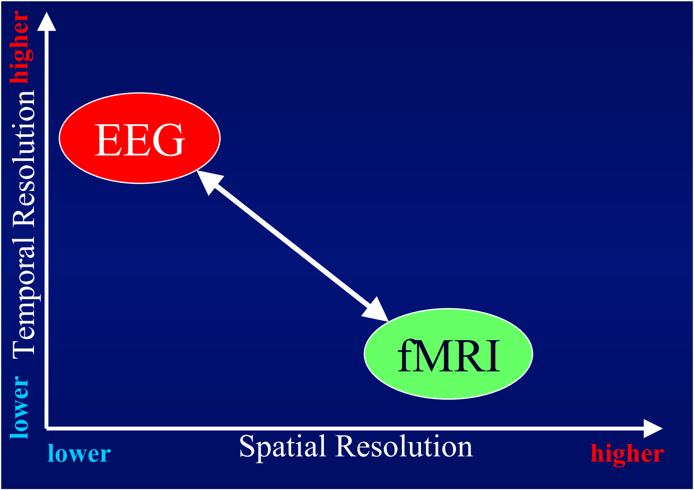
SNP

Gene Expression

Covariates (Age, etc.)

Other*

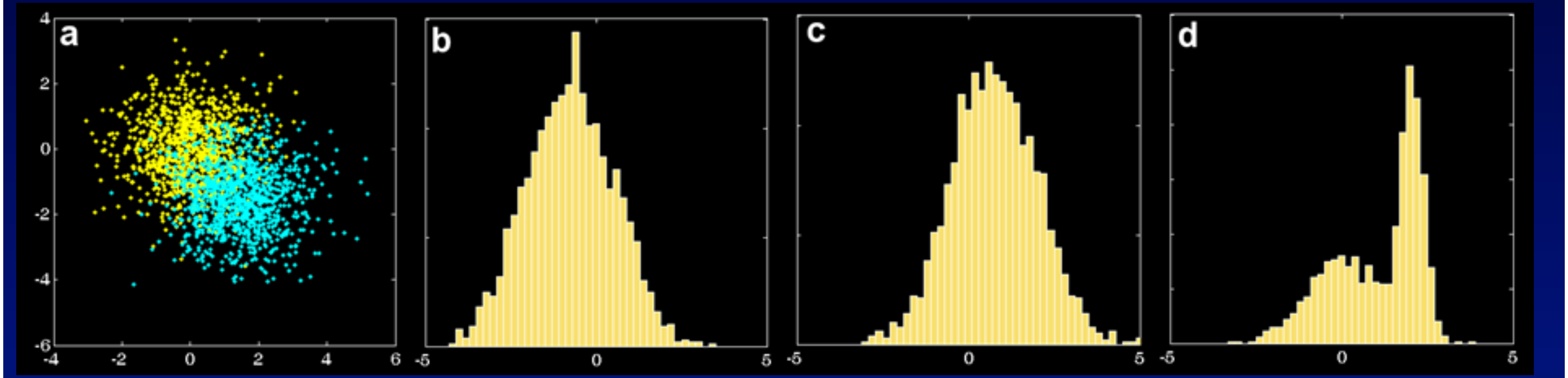
A Role For Data Fusion?:



Possible approaches for joint analyses

- Voxel-based
  - Correlation [Worsely 1998]
  - *Straightforward, but difficult to visualize*
- Region-based
  - Interregional correlation [Horwitz, et al, 1984]
  - Structural equation modeling [McIntosh and Gonzalez-Lima 1994; Friston et al., 2003, McIntosh & Gonzalez-Lima, 1991, Buchel & Friston, 1997]
  - Multiple regression and extensions [e.g., Kalman filters, Buchel & Friston, 1998]
  - Bayes networks, Dynamic Causal Modeling [Dynamic Causal Modeling, Friston, Penny, et al, 2003]
  - *Useful for model testing, does not take into account all brain regions*
- Transformation-based
  - *A natural set of tools for this problem include those that transform data matrices into a smaller set of modes or components*
  - Singular value decomposition [Friston et al., 1993; Friston et al., 1996]
  - Partial Least Squares [McIntosh, Bookstein, et al, 1996]
  - Canonical Variates Analysis [Strother et al, 1995]
  - Independent Component Analysis [McKeown et al, 1998, Calhoun et al, 2001]

Why Joint/Multimodal?



In a non-joint analysis, we maximize the likelihood functions for each modality separately…

$$w_1^* = \arg\max_{w_1} \log p\left(x^{(E)}; w_1\right)$$

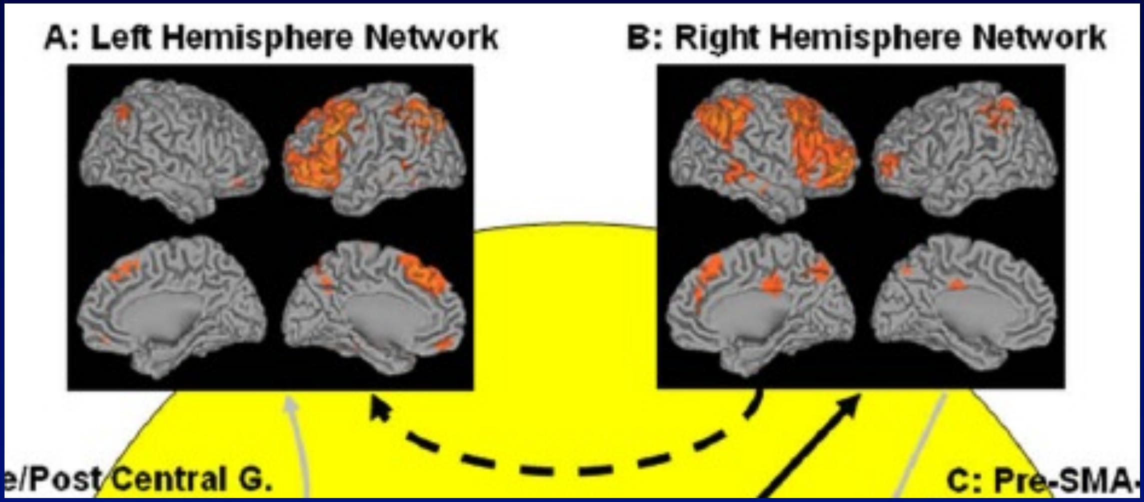
Resulting in two unmixing parameters, that then have to somehow be fused together

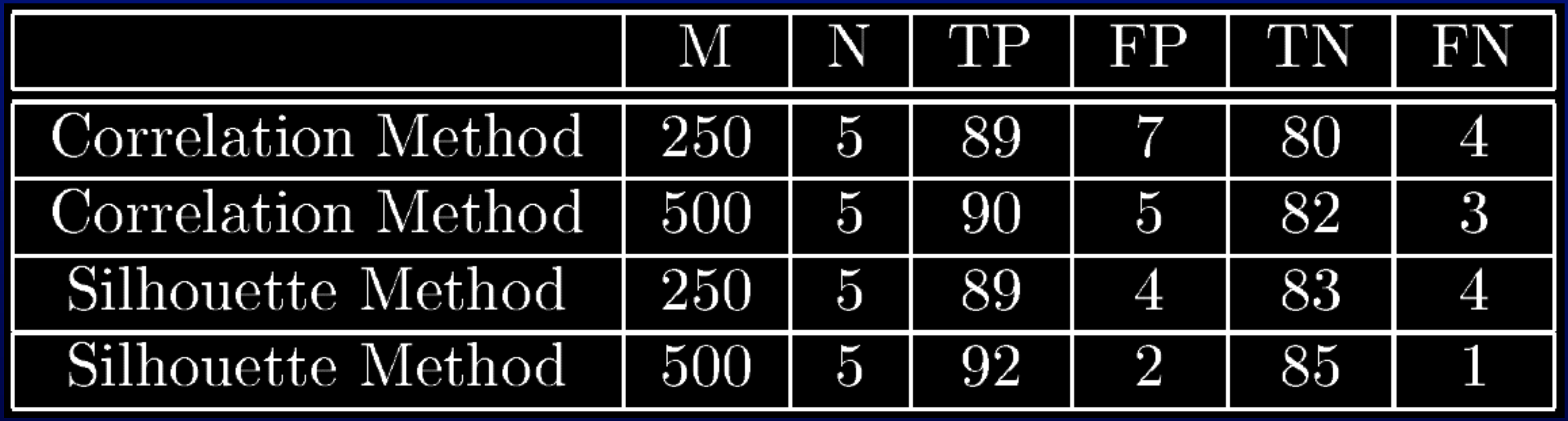
$$w_2^* = \arg\max_{w_2} \log p\left(x^{(F)}; w_2\right)$$

In contrast: for a joint analysis we maximize the joint likelihood function, resulting in a single fused unmixing parameter

$$w^* = \arg\max_{w} \log p\left(x^{(E)}, x^{(F)}; w\right)$$

Why Multivariate?



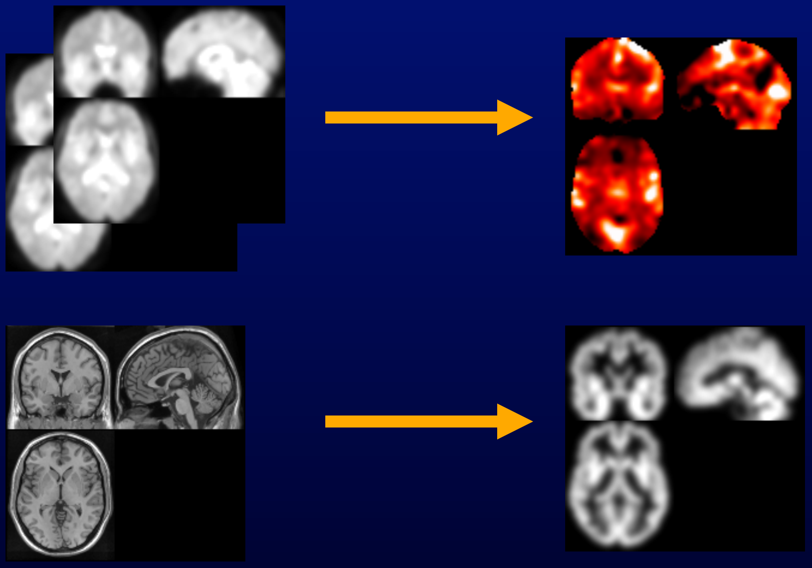


	M	N	TP	FP	TN	FN
Correlation Method	250	5	89	7	80	4
Correlation Method	500	5	90	5	82	3
Silhouette Method	250	5	89	4	83	4
Silhouette Method	500	5	92	2	85	1

Cross–validation performance using the top M SNP’s selected via two different methods as the basis for an N-factor principle component model.

Why Features?

- What is a feature?
  - Lower dimensional data containing information of interest
  - Examples: An image of activation amplitudes, A gray matter segmentation image, fractional anisotropy image
- Advantages
  - Less-computationally complex/easier to model
  - Takes advantages of existing analytic approaches
  - Can be used to examine inter-relationships between multiple data types at the subject level



fMRI

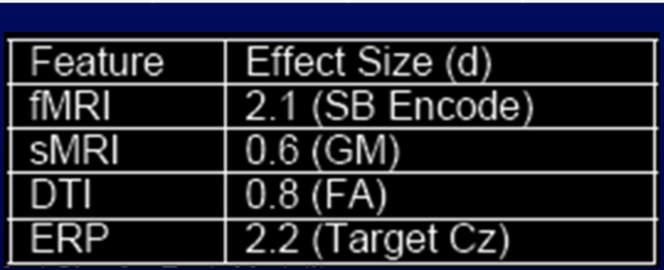
sMRI

Feature Data

- What are the features?
  - Examples: An image of activation amplitudes, A gray matter segmentation image, fractional anisotropy image
- Advantages
  - Less-computationally complex/easier to model
  - Takes advantages of existing analytic approaches
  - Examines inter-subject covariation of multiple data types

Modality	***Core-Feature***
fMRI SB task	Recognition related activity
	Encode-related activity
fMRI AOD task	Target-related activity
	Novel-related activity
sMRI	GM concentration
	WM concentration
	CSF concentration
EEG AOD task	Target-related ERP
	Novel-related ERP

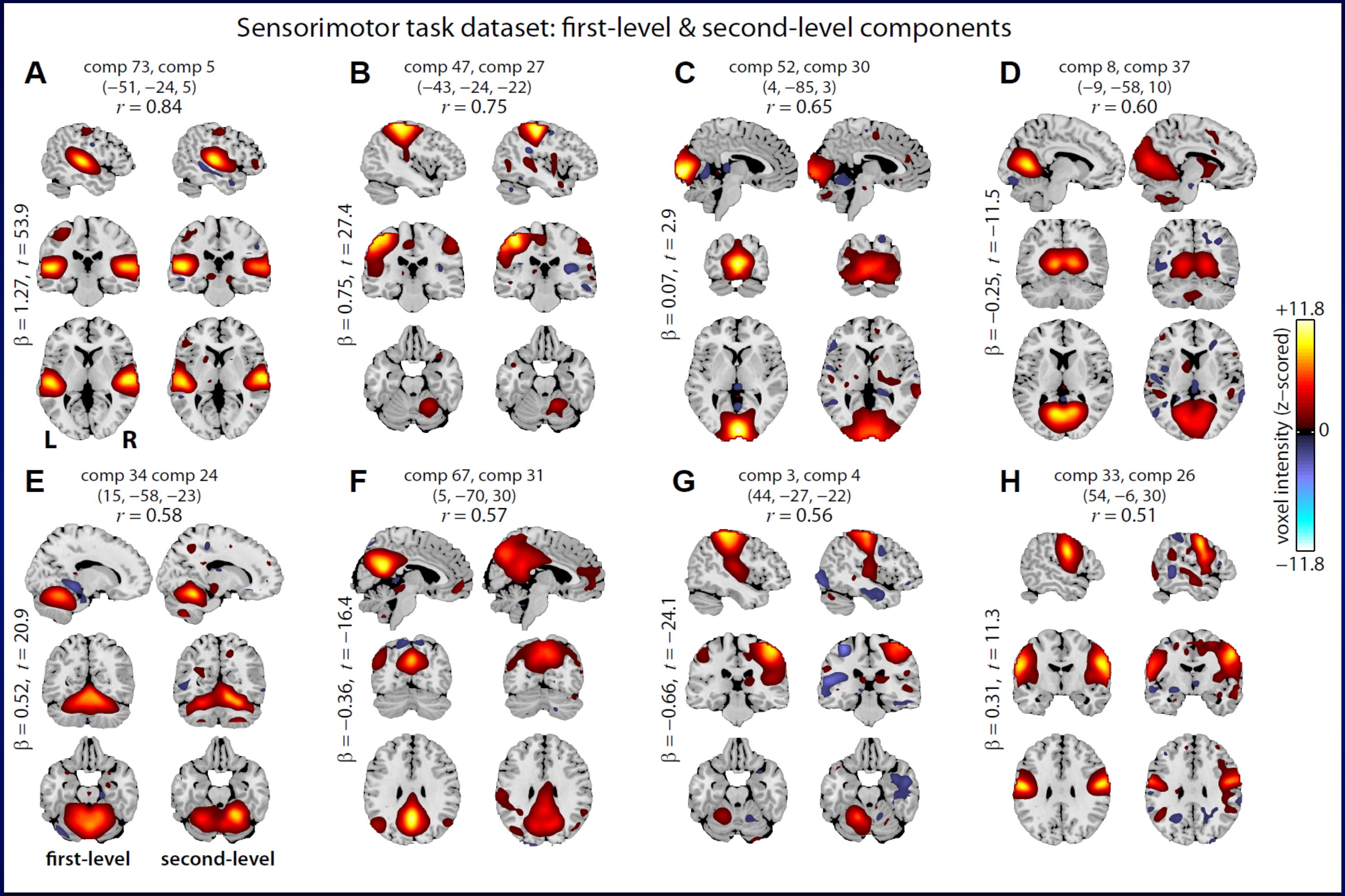
Modality	*Raw data size*	*Feature size*	*Compression ratio*
fMRI	260 MB	1.3 MB	1:200
DTI	200 MB	10 MB	1:20
sMRI	46 MB	28 MB	1:1.6
EEG	25MB	.45MB	1:56



Feature	Effect Size (d)
fMRI	2.1 (SB Encode)
sMRI	0.6 (GM)
DTI	0.8 (FA)
ERP	2.2 (Target Cz)

V. D. Calhoun and T. Adali, "Feature-based Fusion of Medical Imaging Data," *IEEE Trans. Inf. Tech. in Biomedicine,* vol. 13, pp. 1-10, 2009

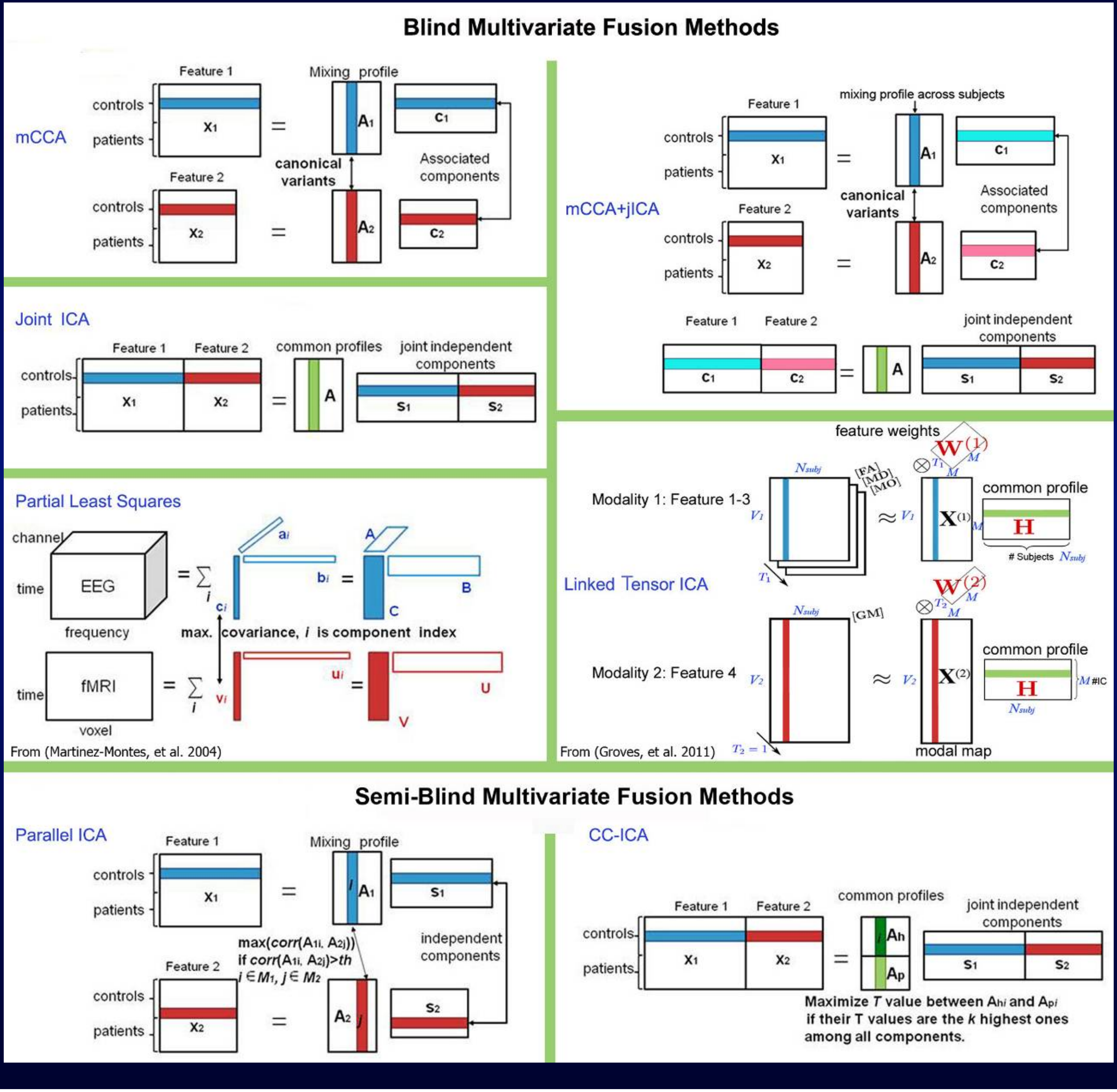
Feature-based ICA



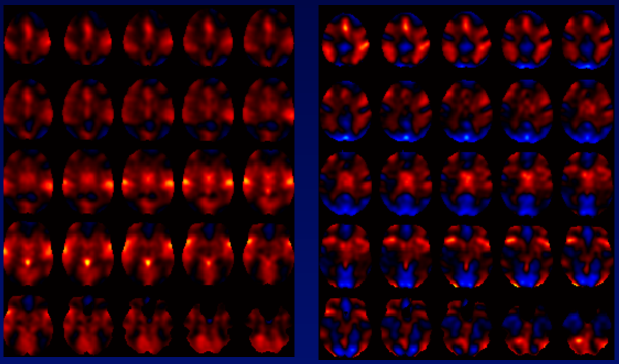
V. D. Calhoun, T. Adali, K. A. Kiehl, R. S. Astur, J. J. Pekar, and G. D. Pearlson, "A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia," *Hum.Brain Map.,* vol. 27, pp. 598-610, 2006

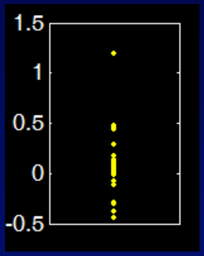
V. D. Calhoun and T. Adali, "Feature-based Fusion of Medical Imaging Data," *IEEE Trans. Inf. Tech. in Biomedicine*, vol. 13, pp. 1-10, 2009

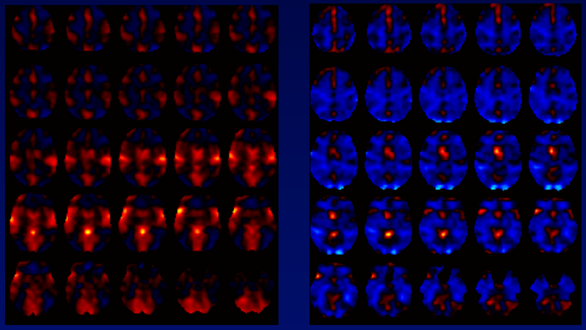
A Family of Multivariate Methods



Joint ICA







$$\left[\mathbf{x}_{Task1} \quad \mathbf{x}_{Task2}\right] = \mathbf{A} \times \left[\mathbf{s}_{Task1} \quad \mathbf{s}_{Task2}\right]$$

Generative Model:

$$\left[x_i^{F1} \quad x_i^{F2}\right] = \mathbf{A}\left[s_c^{F1} \quad s_c^{F2}\right]$$

Update Equation:

$$\Delta\mathbf{W} = \eta\left\{\mathbf{I} - 2\mathbf{y}^{F1}\left(\mathbf{u}^{F1}\right)^T - 2\mathbf{y}^{F2}\left(\mathbf{u}^{F2}\right)^T\right\}\mathbf{W}$$

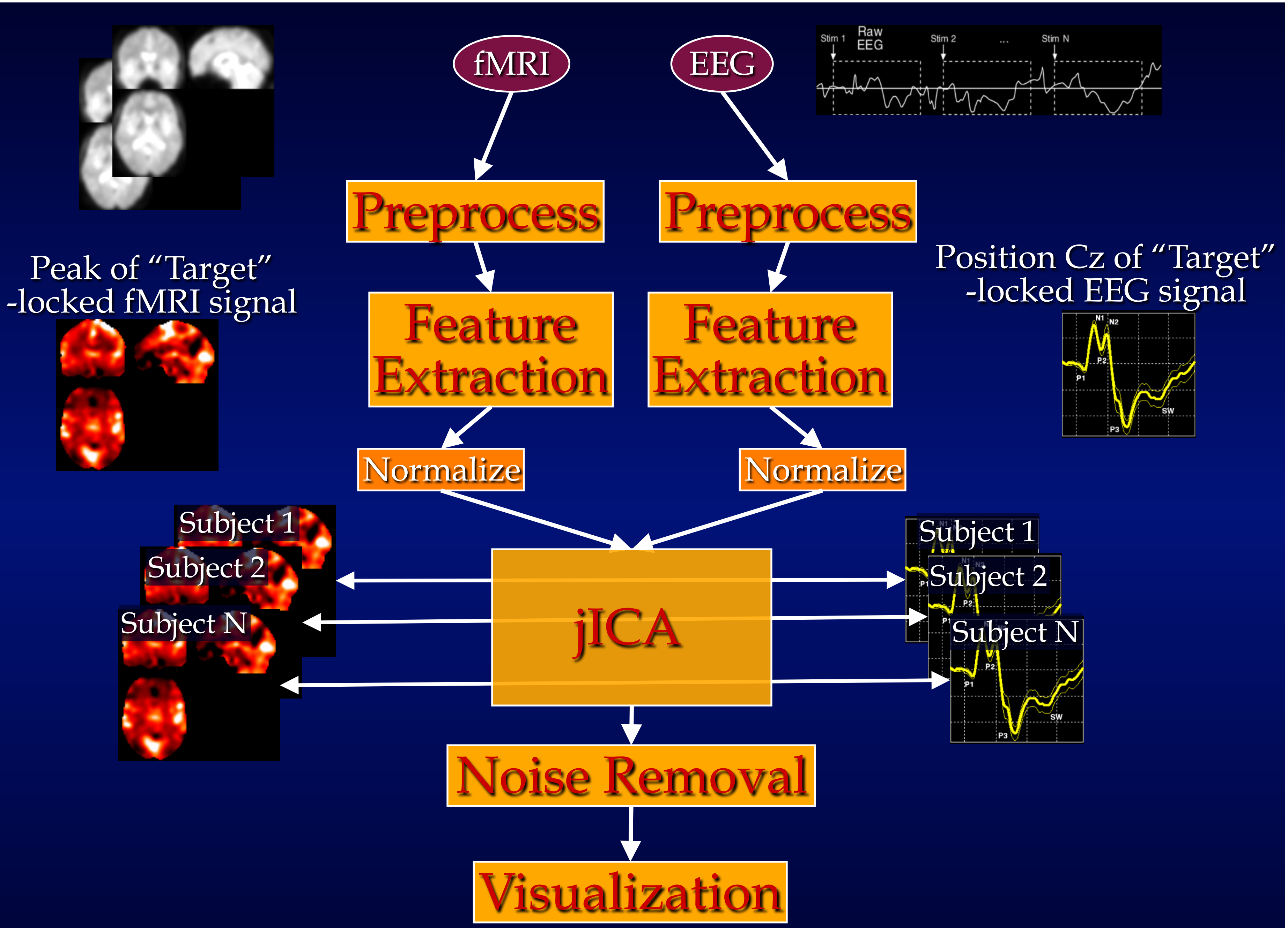
sMRI/fMRI: Calhoun VD, Adali T, Giuliani N, Pekar JJ, Pearlson GD, Kiehl KA. (2006): A Method for Multimodal Analysis of Independent Source Differences in Schizophrenia: Combining Gray Matter Structural and Auditory Oddball Functional Data. Hum.Brain Map. 27(1):47-62.

fMRI/fMRI: Calhoun VD, Adali T, Kiehl KA, Astur RS, Pekar JJ, Pearlson GD. (2006): A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia. Hum.Brain Map. 27(7):598-610.

EEG/fMRI: Calhoun VD, Pearlson GD, Kiehl KA. (2006): Neuronal Chronometry of Target Detection: Fusion of Hemodynamic and Event-related Potential Data. NeuroImage 30(2):544-553.

Overview

- Motivation
- Joint ICA
  - ERP/FMRI
  - Multitask-fMRI
  - sMRI/fMRI
  - CC-ICA/Biomarker Identification
- mCCA
  - sMRI/fMRI/ERP
- mCCA+jICA
  - DTI+fMRI
- Parallel ICA
  - SNPs
  - Epigenetics
- Conclusions



fMRI

EEG

Preprocess

Preprocess

Feature Extraction

Feature Extraction

Normalize

Normalize

jICA

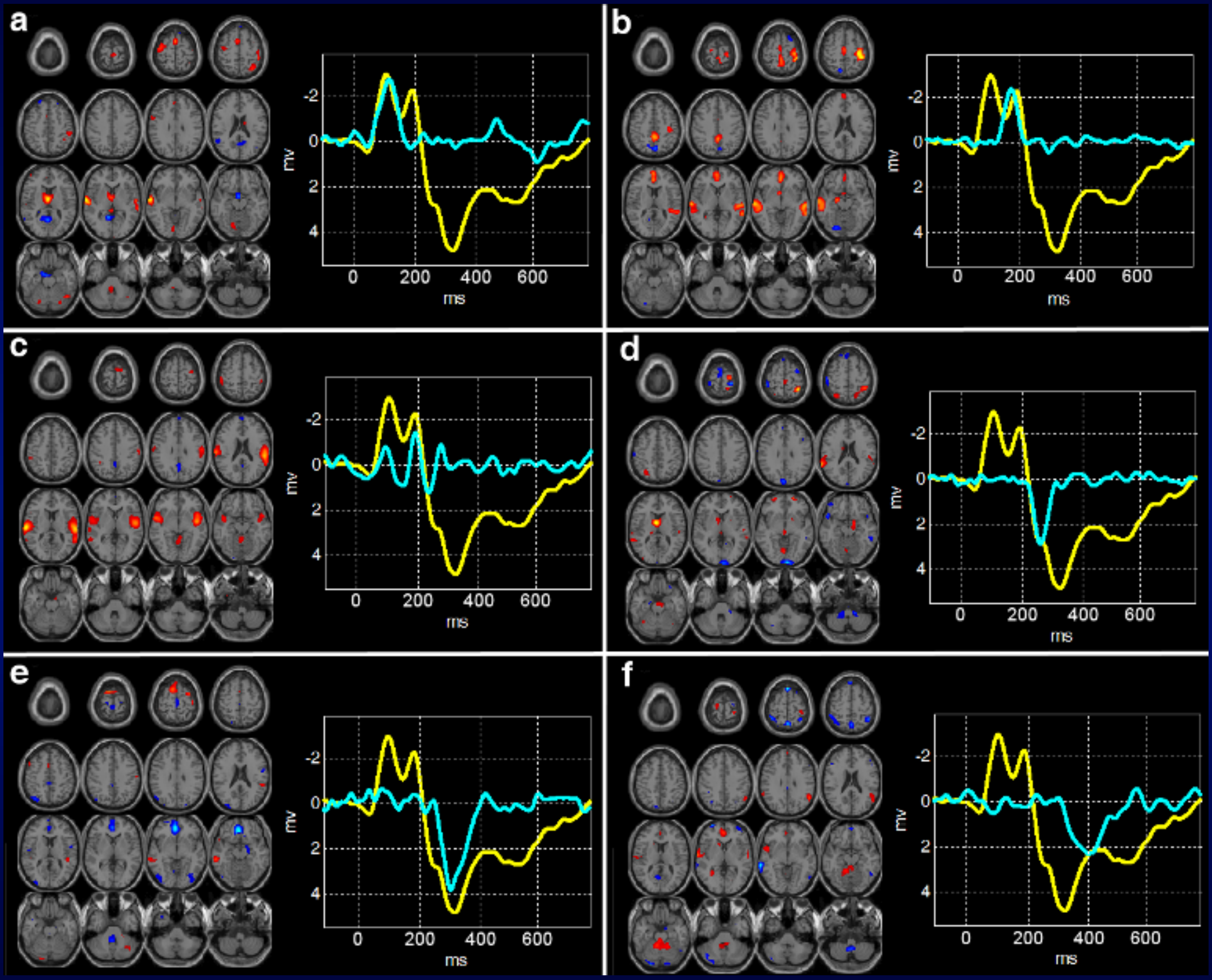
Noise Removal

Visualization

Peak of “Target” -locked fMRI signal

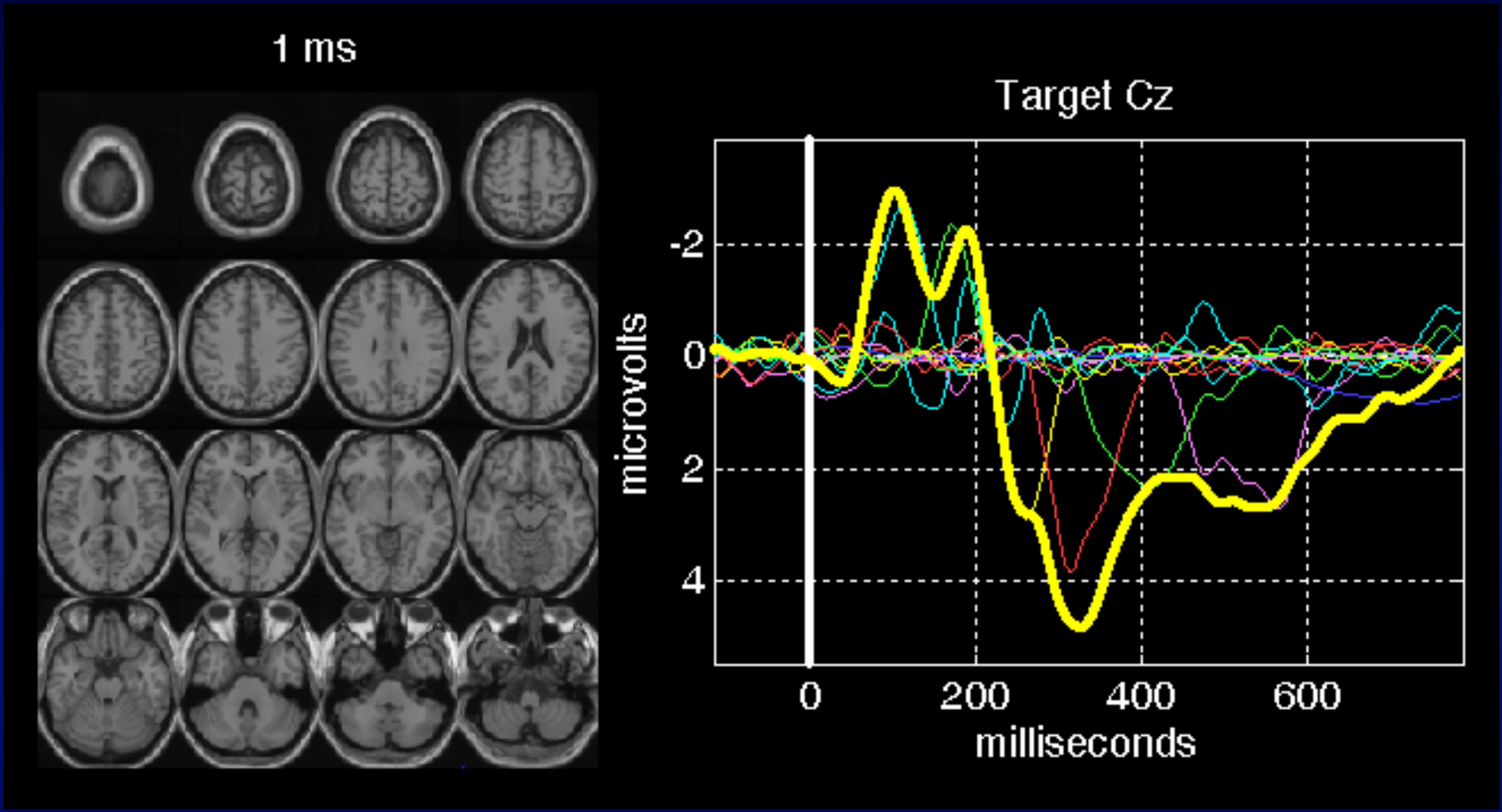
Position Cz of “Target” -locked EEG signal

Joint ERP/fMRI Components



Calhoun, V.D., Pearlson, G.D., and Kiehl, K.A. (2006). Neuronal Chronometry of Target Detection: Fusion of Hemodynamic and Event-related Potential Data. NeuroImage 30, 544-553.

FMRI Snapshots (movie)



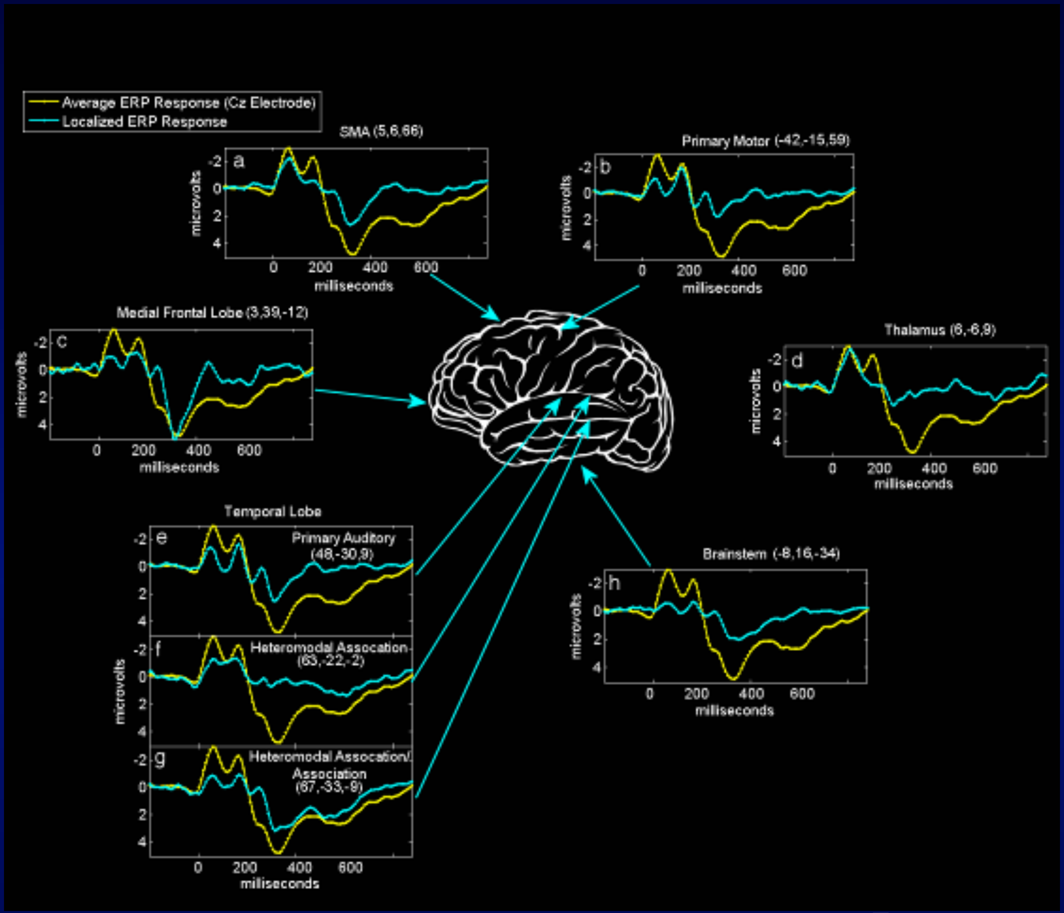
ERP (temporal) Components: $\mathbf{T}=\left[\mathbf{t}_1 \quad \ldots \quad \mathbf{t}_N\right]$

FMRI (spatial) Components: $\mathbf{S}=\left[\mathbf{s}_1 \quad \ldots \quad \mathbf{s}_N\right]$

ERP Timecourse Snapshot: $\mathbf{M}_E(v)=\mathbf{T}\times\left|\mathbf{S}\right|^T(v)$

Calhoun, V.D., Pearlson, G.D., and Kiehl, K.A. (2006). Neuronal Chronometry of Target Detection: Fusion of Hemodynamic and Event-related Potential Data. NeuroImage 30, 544-553.

ERP Snapshots



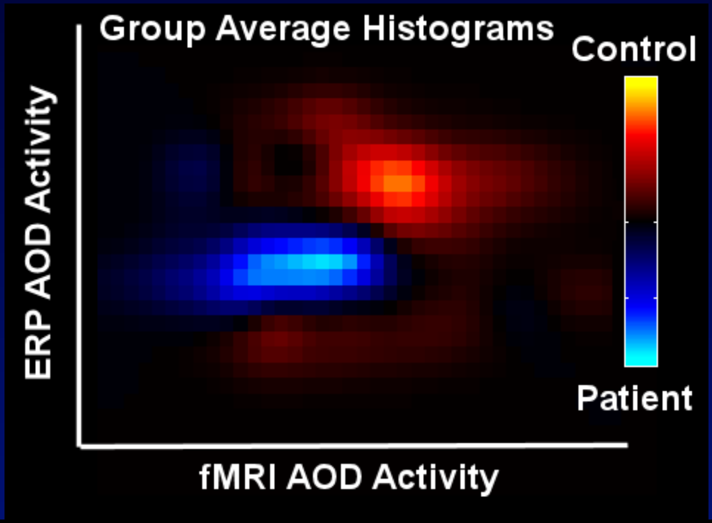
$$\text{ERP (temporal) Components: } \mathbf{T} = \left[\mathbf{t}_1 \quad \text{K} \quad \mathbf{t}_N\right]$$

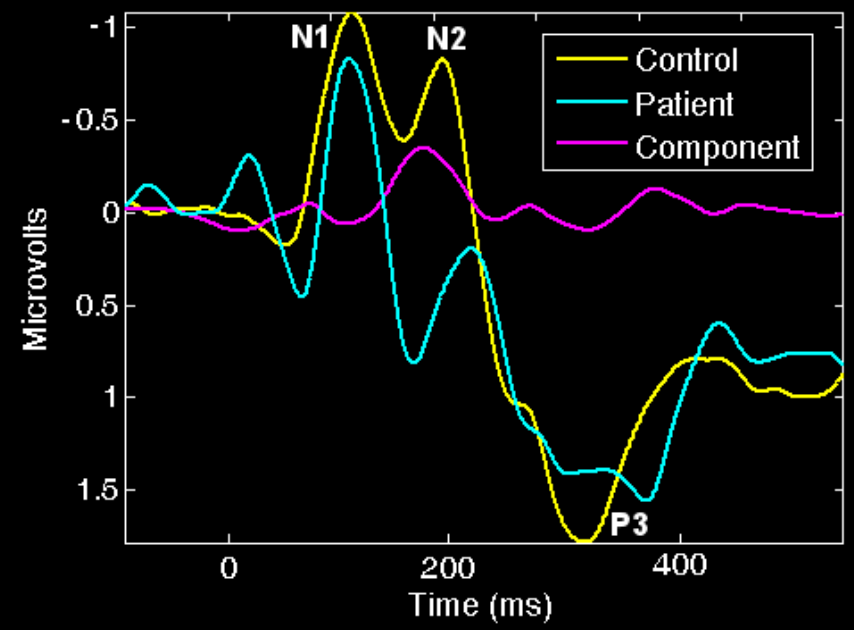
$$\text{FMRI (spatial) Components: } \mathbf{S} = \left[\mathbf{s}_1 \quad \text{K} \quad \mathbf{s}_N\right]$$

$$\text{ERP Timecourse Snapshot: } \mathbf{M}_E\left(v\right) = \mathbf{T} \times \left|\mathbf{S}\right|^T\left(v\right)$$

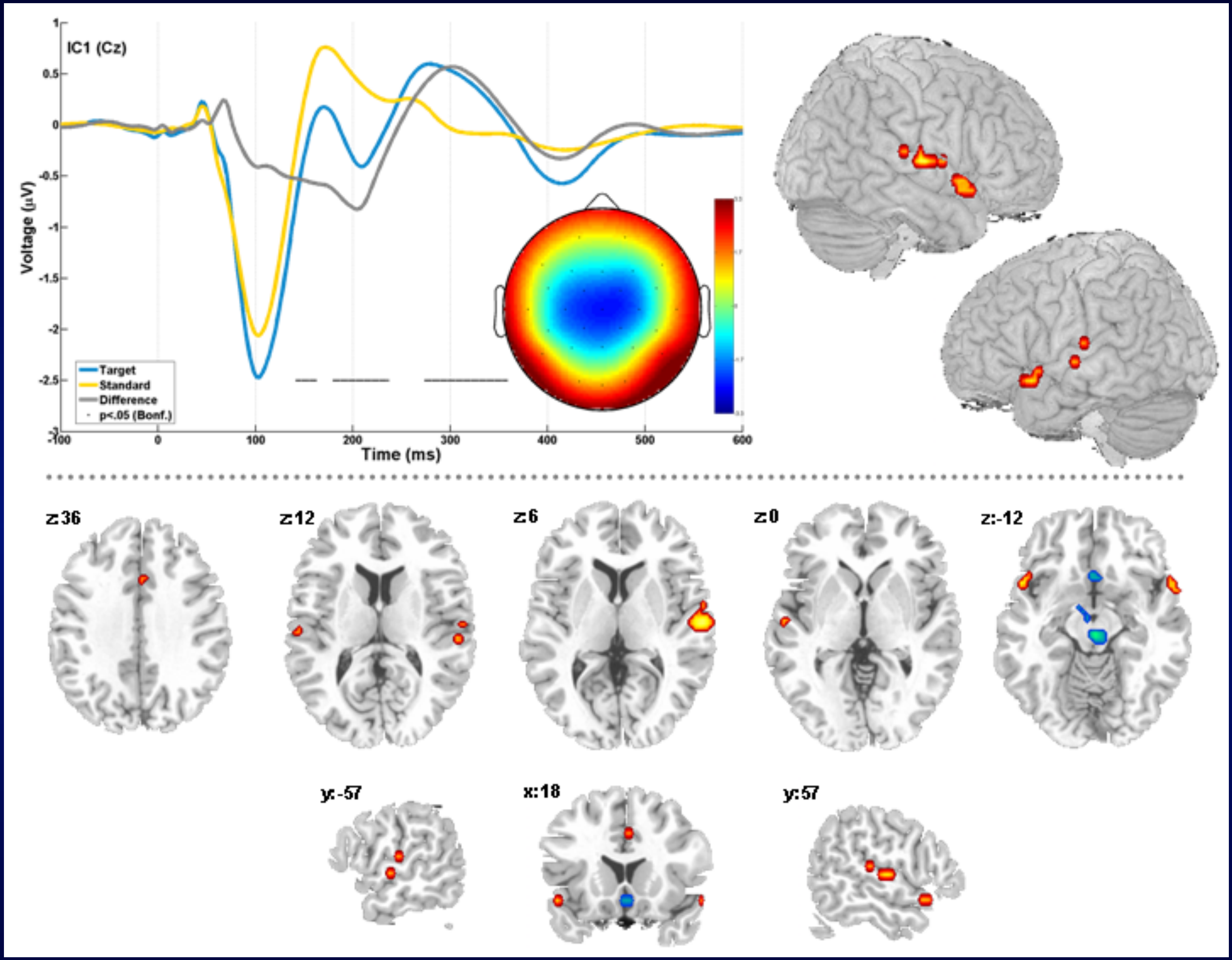
Calhoun, V.D., Pearlson, G.D., and Kiehl, K.A. (2006). Neuronal Chronometry of Target Detection: Fusion of Hemodynamic and Event-related Potential Data. NeuroImage 30, 544-553.

Linked EEG/fMRI Results





V. D. Calhoun, L. Wu, K. A. Kiehl, T. Eichele, and G. D. Pearlson, "Aberrant Processing of Deviant Stimuli in Schizophrenia Revealed by Fusion of FMRI and EEG Data," *Acta Neuropsychiatria,* vol. 22, pp. 127-138, 2010

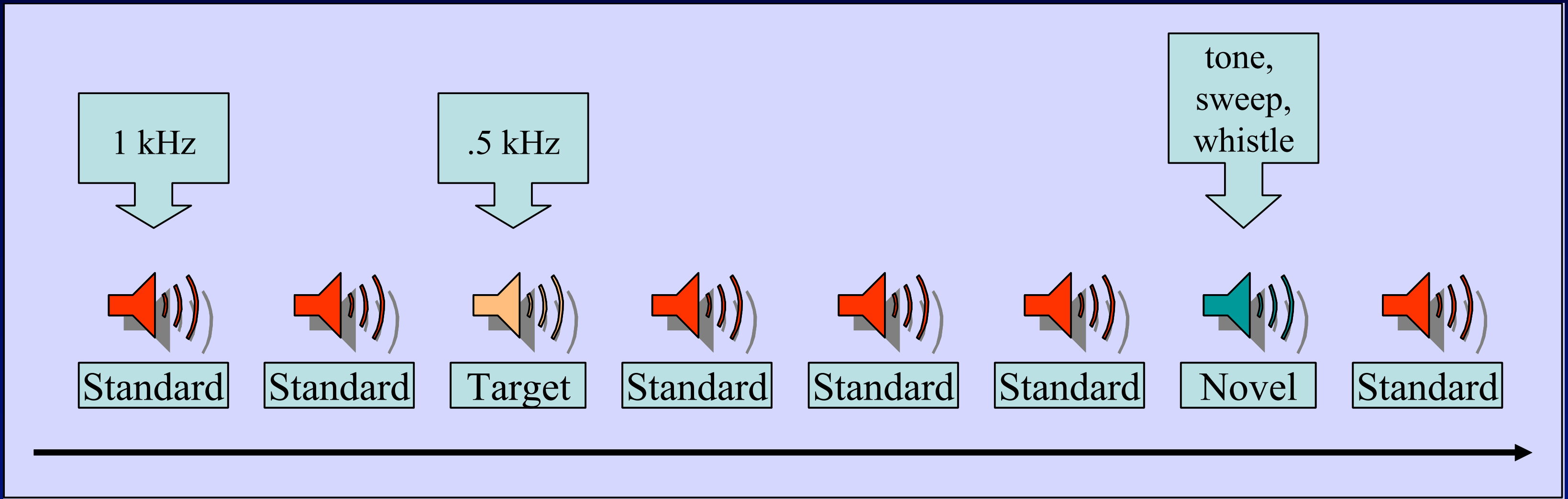


T. Eichele, V. D. Calhoun, M. Moosmann, K. Specht, M. Jongsma, R. Quiroga, H. Nordby, and K. Hugdahl, "Unmixing concurrent EEG-fMRI with parallel independent component analysis," *Int. J. Psych., vol. 67, pp. 222-234, 2008.*

Overview

- Motivation
- Joint ICA
  - ERP/FMRI
  - Multitask-fMRI
  - sMRI/fMRI
  - CC-ICA/Biomarker Identification
- mCCA
  - sMRI/fMRI/ERP
- mCCA+jICA
  - DTI+fMRI
- Parallel ICA
  - SNPs
  - Epigenetics
- Conclusions

Auditory Oddball & Sternberg Tasks



1 kHz

.5 kHz

tone, sweep, whistle

Standard

Standard

Target

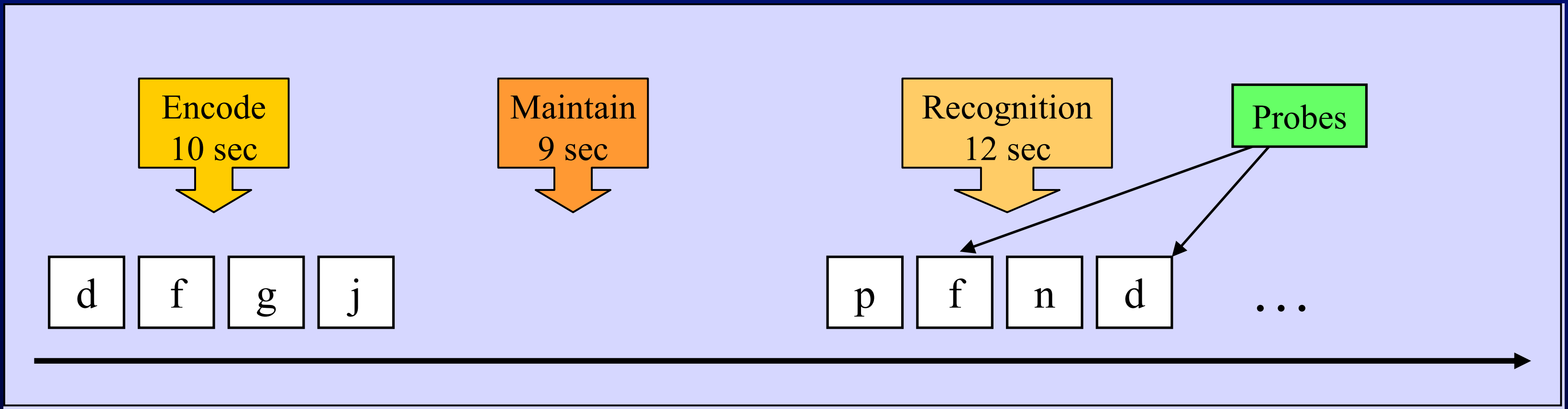
Standard

Standard

Standard

Novel

Standard



Encode 10 sec

Maintain 9 sec

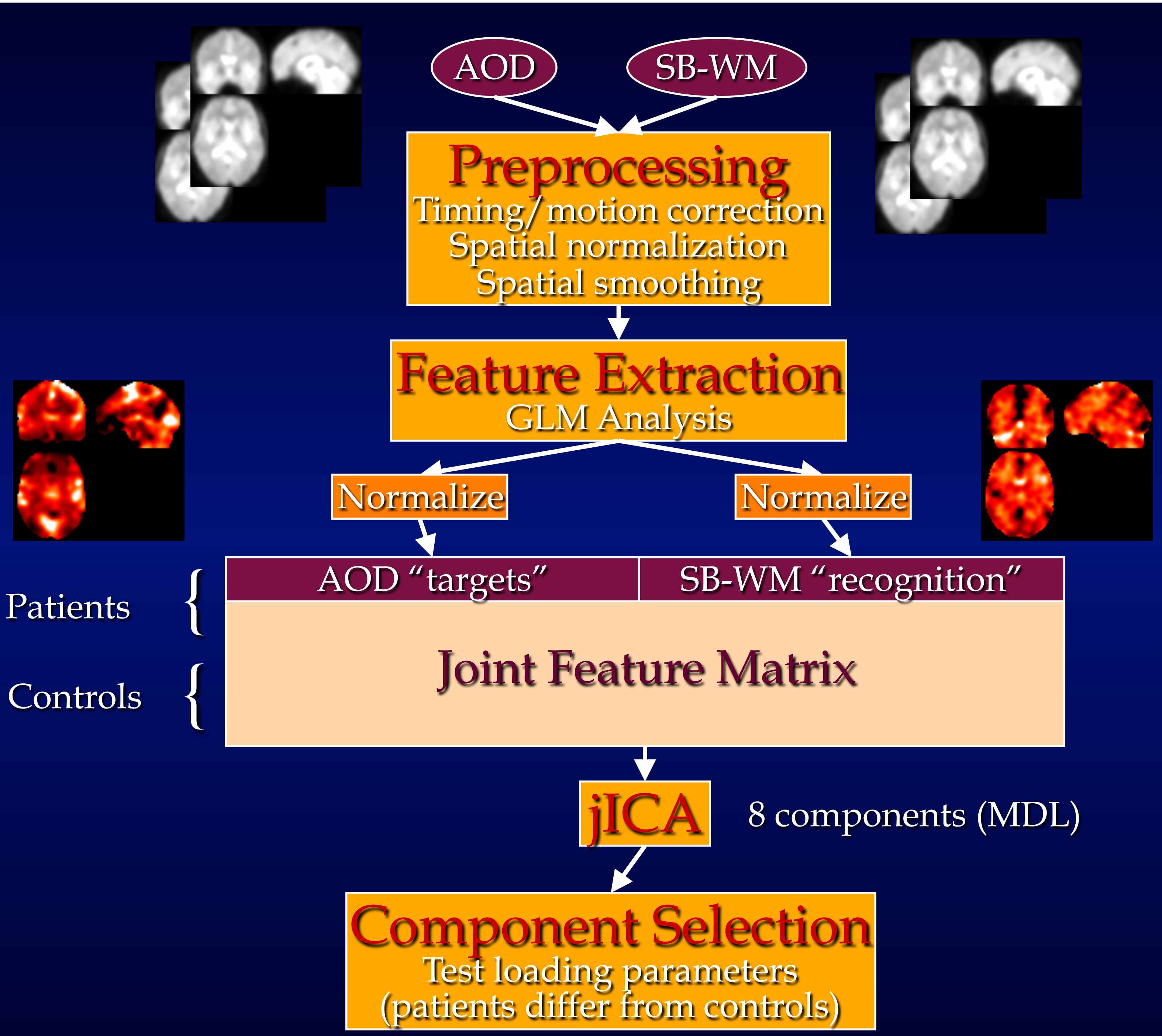
Recognition 12 sec

Probes

d f g j

p f n d …

V. D. Calhoun, T. Adali, K. A. Kiehl, R. S. Astur, J. J. Pekar, and G. D. Pearlson, "A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia," Hum.Brain Map., vol. 27, pp. 598-610, 2006.



AOD

SB-WM

Preprocessing

Timing/motion correction
Spatial normalization
Spatial smoothing

Feature Extraction

GLM Analysis

Normalize

Normalize

AOD “targets”	SB-WM “recognition”

Patients

Controls

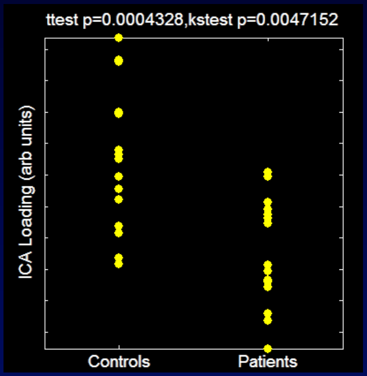
Joint Feature Matrix

jICA

8 components (MDL)

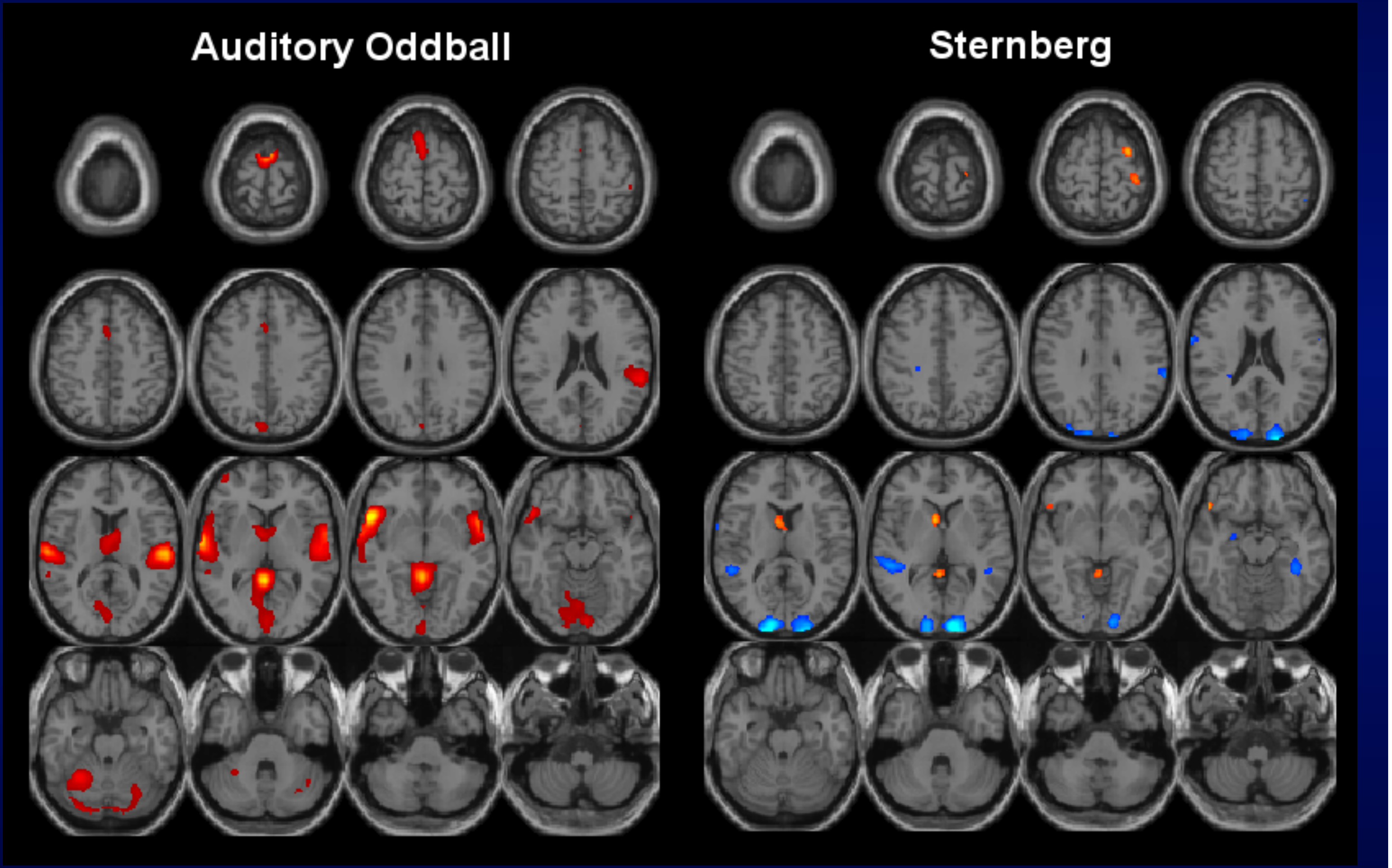
Component Selection

Test loading parameters
(patients differ from controls)



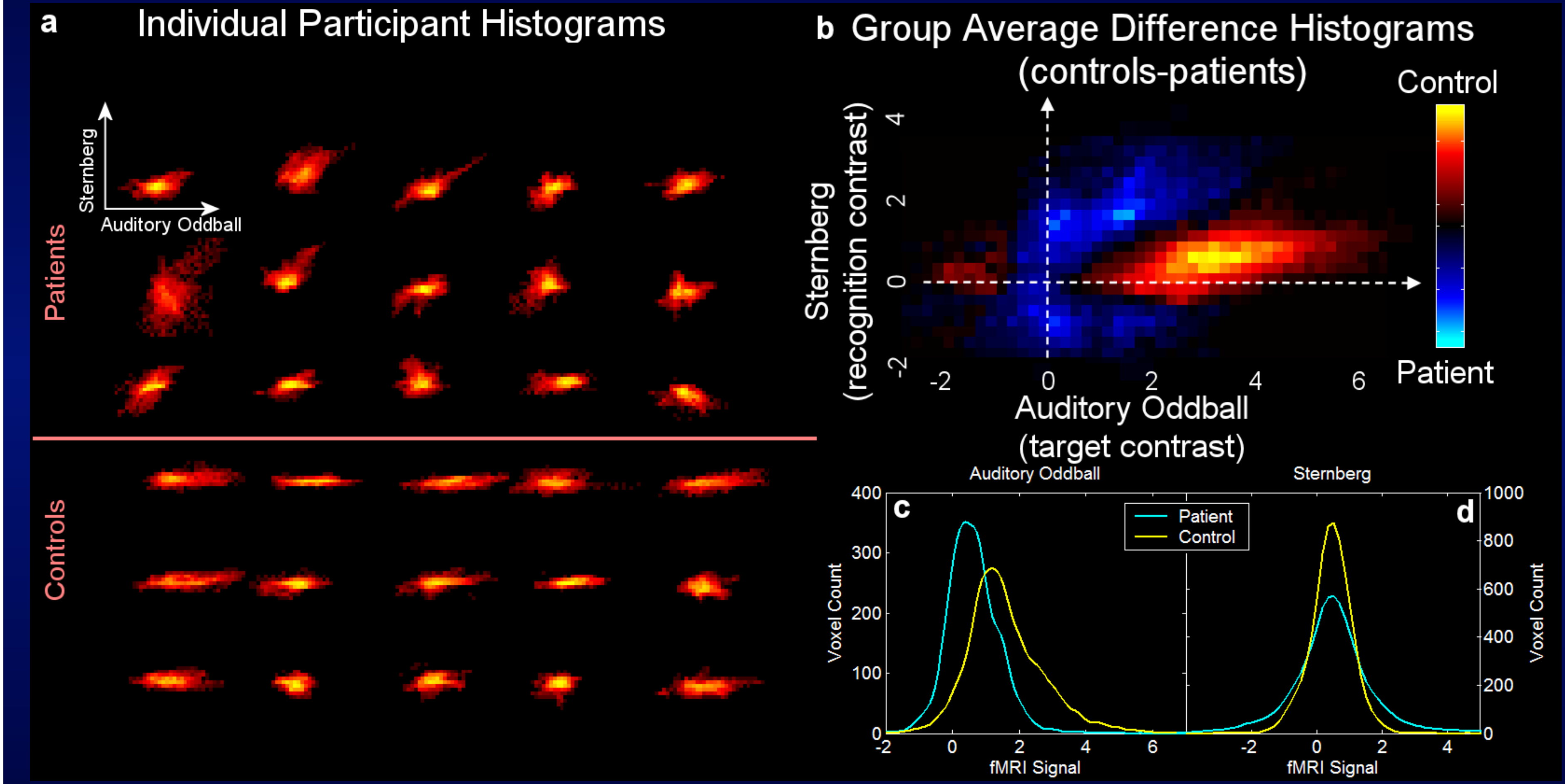
jICA Results

One component
Significant at $p<0.01$



V. D. Calhoun, T. Adali, K. A. Kiehl, R. S. Astur, J. J. Pekar, and G. D. Pearlson, "A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia," Hum.Brain Map., vol. 27, pp. 598-610, 2006.

Joint Histograms: fMRI

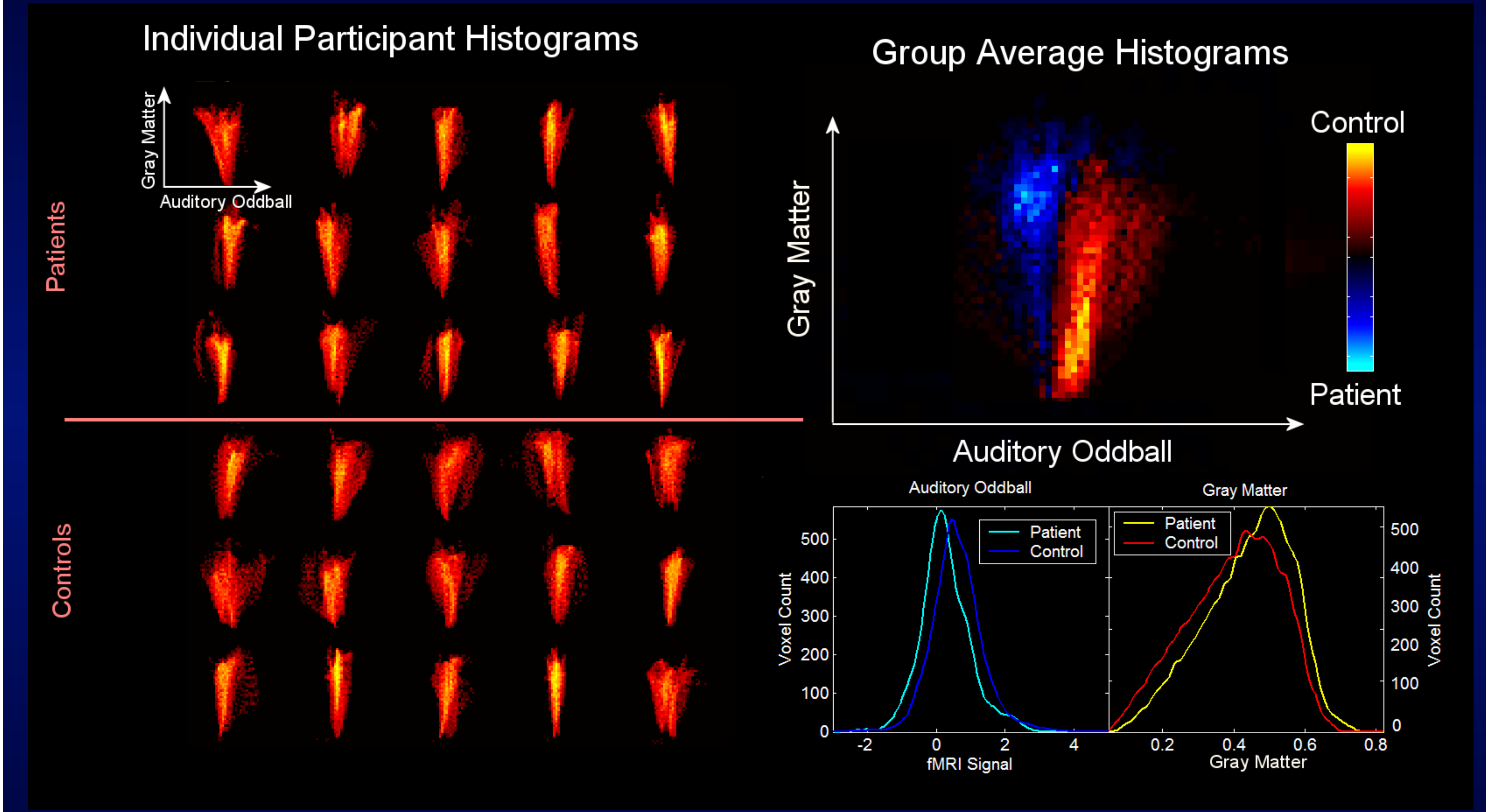


V. D. Calhoun, T. Adali, K. A. Kiehl, R. S. Astur, J. J. Pekar, and G. D. Pearlson, "A Method for Multi-task fMRI Data Fusion Applied to Schizophrenia," Hum.Brain Map., vol. 27, pp. 598-610, 2006.

Overview

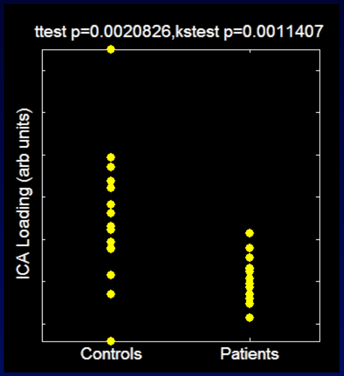
- Motivation
- Joint ICA
    - ERP/FMRI
    - Multitask-fMRI
    - sMRI/fMRI
    - CC-ICA/Biomarker Identification
- mCCA
    - sMRI/fMRI/ERP
- mCCA+jICA
    - DTI+fMRI
- Parallel ICA
    - SNPs
    - Epigenetics
- Conclusions

Joint sMRI/fMRI Histograms

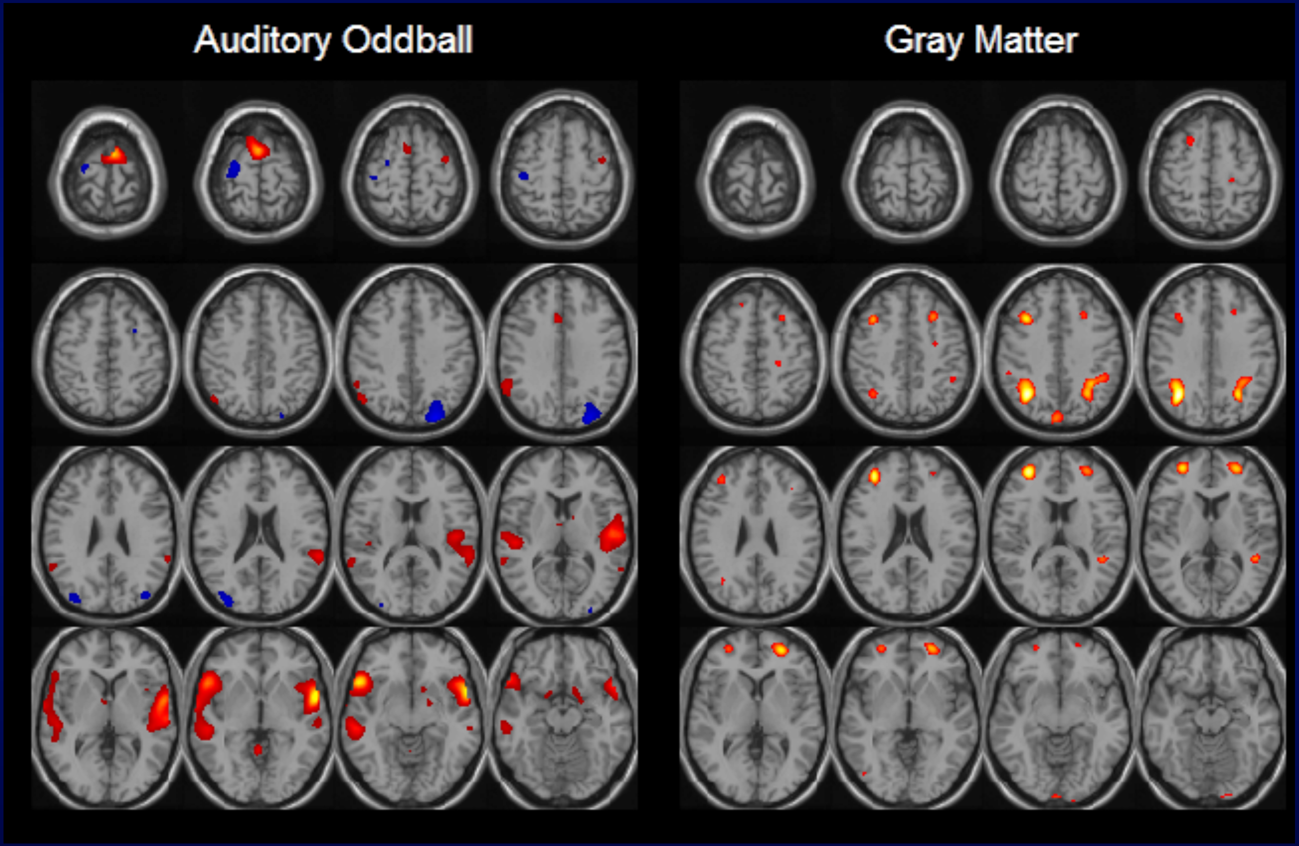


V. D. Calhoun, T. Adali, N. Giuliani, J. J. Pekar, G. D. Pearlson, and K. A. Kiehl, "A Method for Multimodal Analysis of Independent Source Differences in Schizophrenia: Combining Gray Matter Structural and Auditory Oddball Functional Data," *Hum. Brain Map.*, vol. 27, pp. 47-62, 2005.

jICA Results

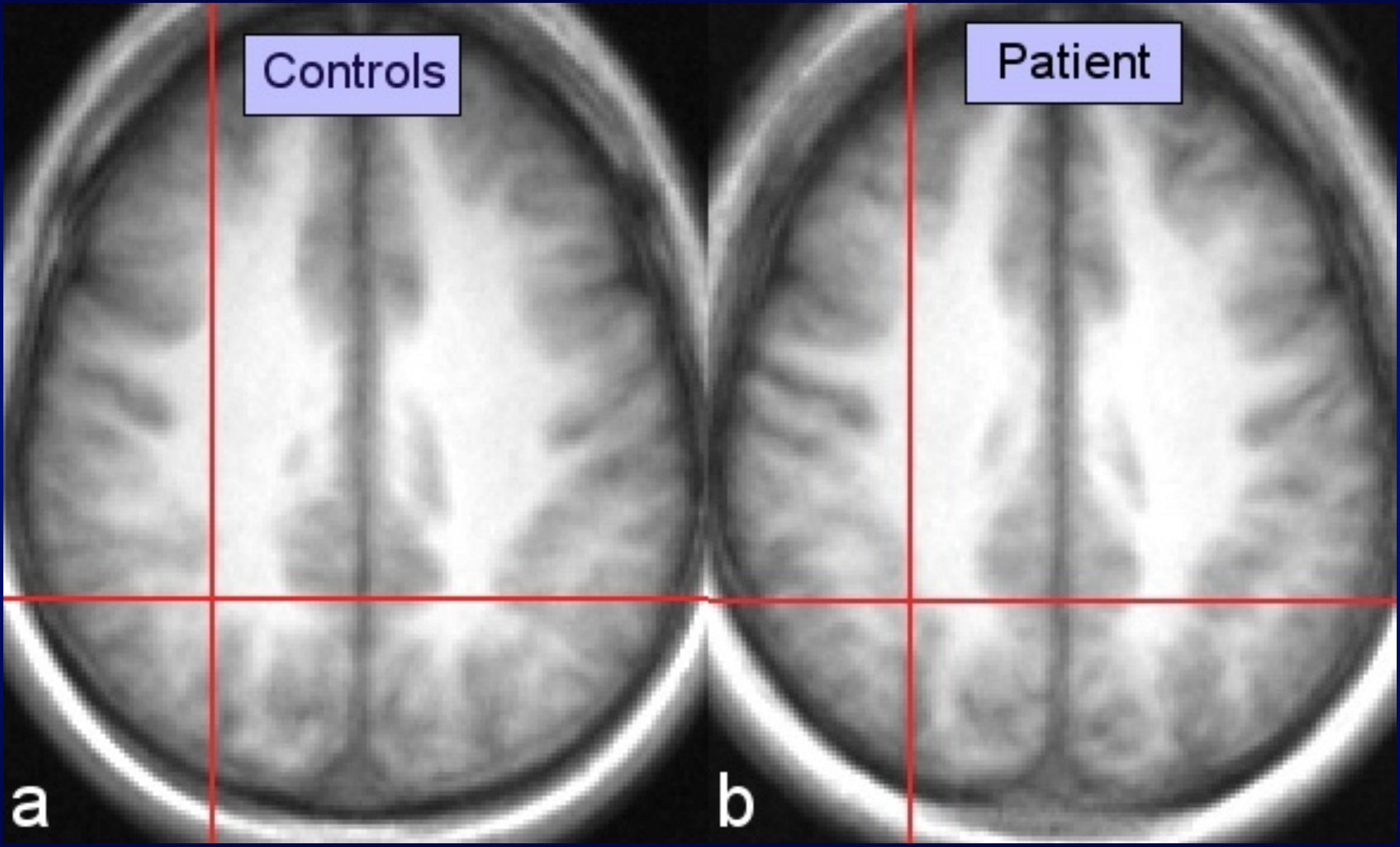


One component
Significant at p>0.01



V. D. Calhoun, T. Adali, N. Giuliani, J. J. Pekar, G. D. Pearlson, and K. A. Kiehl, "A Method for Multimodal Analysis of Independent Source Differences in Schizophrenia: Combining Gray Matter Structural and Auditory Oddball Functional Data," *Hum. Brain Map.*, vol. 27, pp. 47-62, 2005.

Group averaged Images



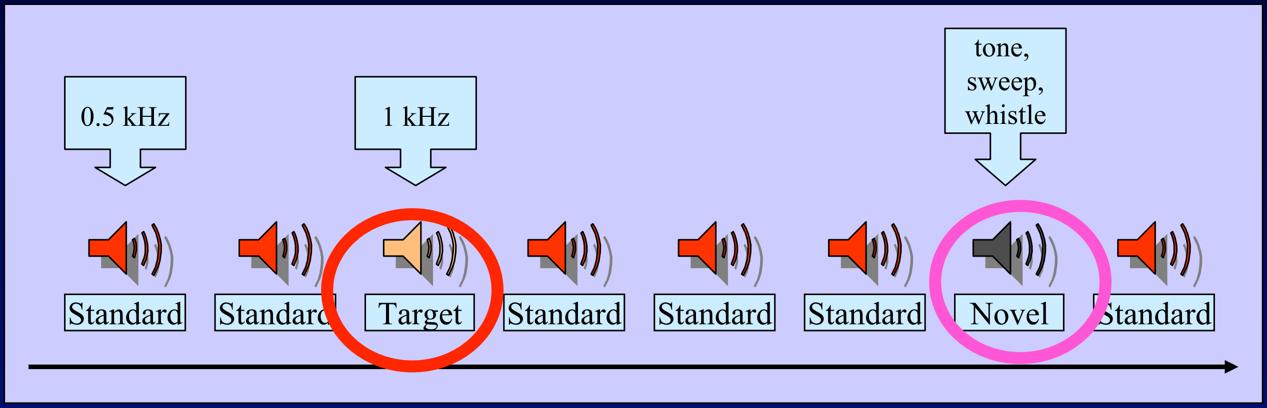
V. D. Calhoun, T. Adali, N. Giuliani, J. J. Pekar, G. D. Pearlson, and K. A. Kiehl, "A Method for Multimodal Analysis of Independent Source Differences in Schizophrenia: Combining Gray Matter Structural and Auditory Oddball Functional Data," *Hum. Brain Map.*, vol. 27, pp. 47-62, 2005.

Overview

- Motivation
- Joint ICA
  - ERP/FMRI
  - Multitask-fMRI
  - sMRI/fMRI
  - CC-ICA/Biomarker Identification
- mCCA
  - sMRI/fMRI/ERP
- mCCA+jICA
  - DTI+fMRI
- Parallel ICA
  - SNPs
  - Epigenetics
- Conclusions

Evaluating Multiple Features

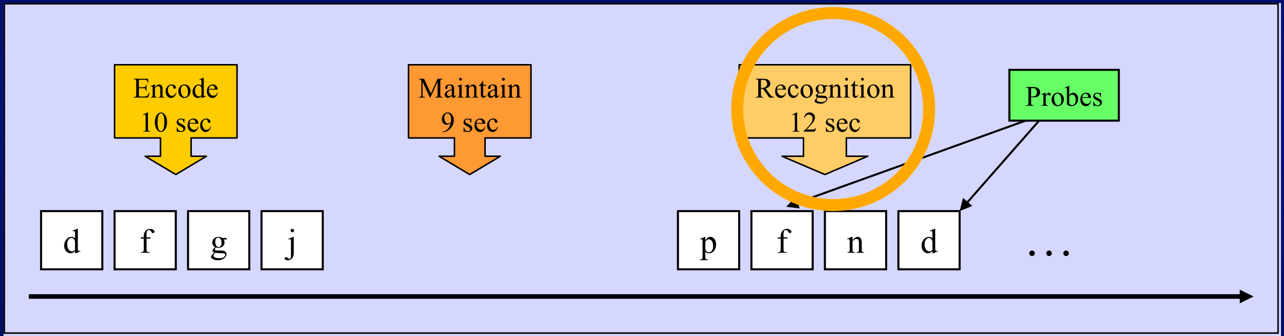
- Auditory Oddball fMRI Task
  - Target-related activity
  - Novel-related activity
- Sternberg fMRI Task
  - Recognition-related activity
- MPRAGE
  - Gray matter segmentation


0.5 kHz

1 kHz

tone, sweep, whistle

Standard Standard Target Standard Standard Standard Novel Standard


Encode 10 sec

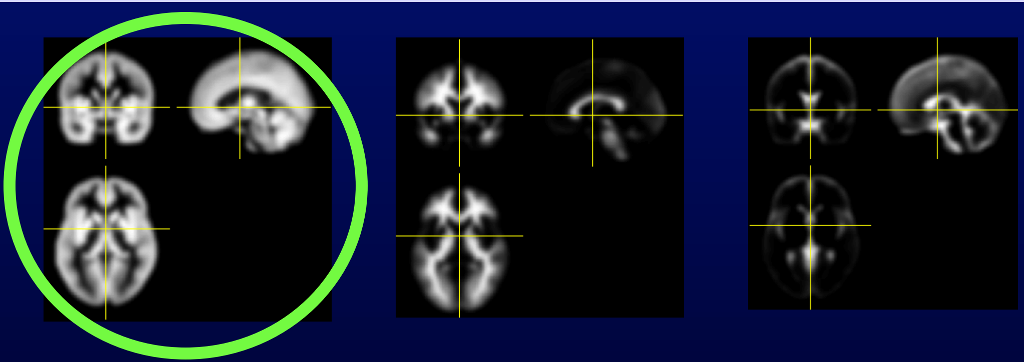
Maintain 9 sec

Recognition 12 sec

Probes

d f g j

p f n d …



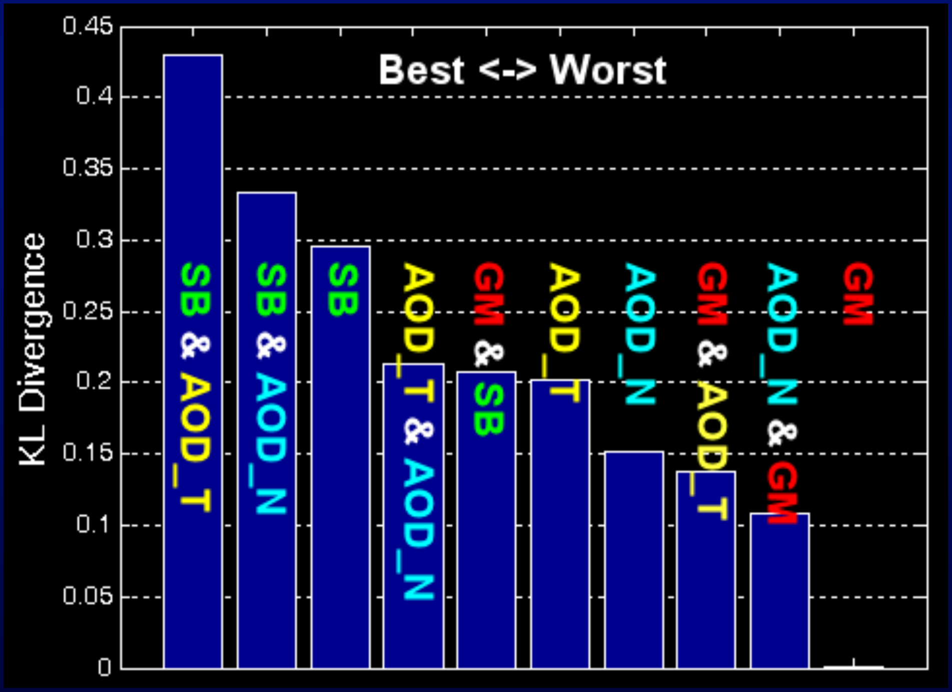
V. D. Calhoun and T. Adali, "Feature-based Fusion of Medical Imaging Data," *IEEE Trans. Inf. Tech. in Biomedicine*, vol. 13, pp. 1-10, 2009

Identifying the “best” combination

- Kullback-Leibler divergence….

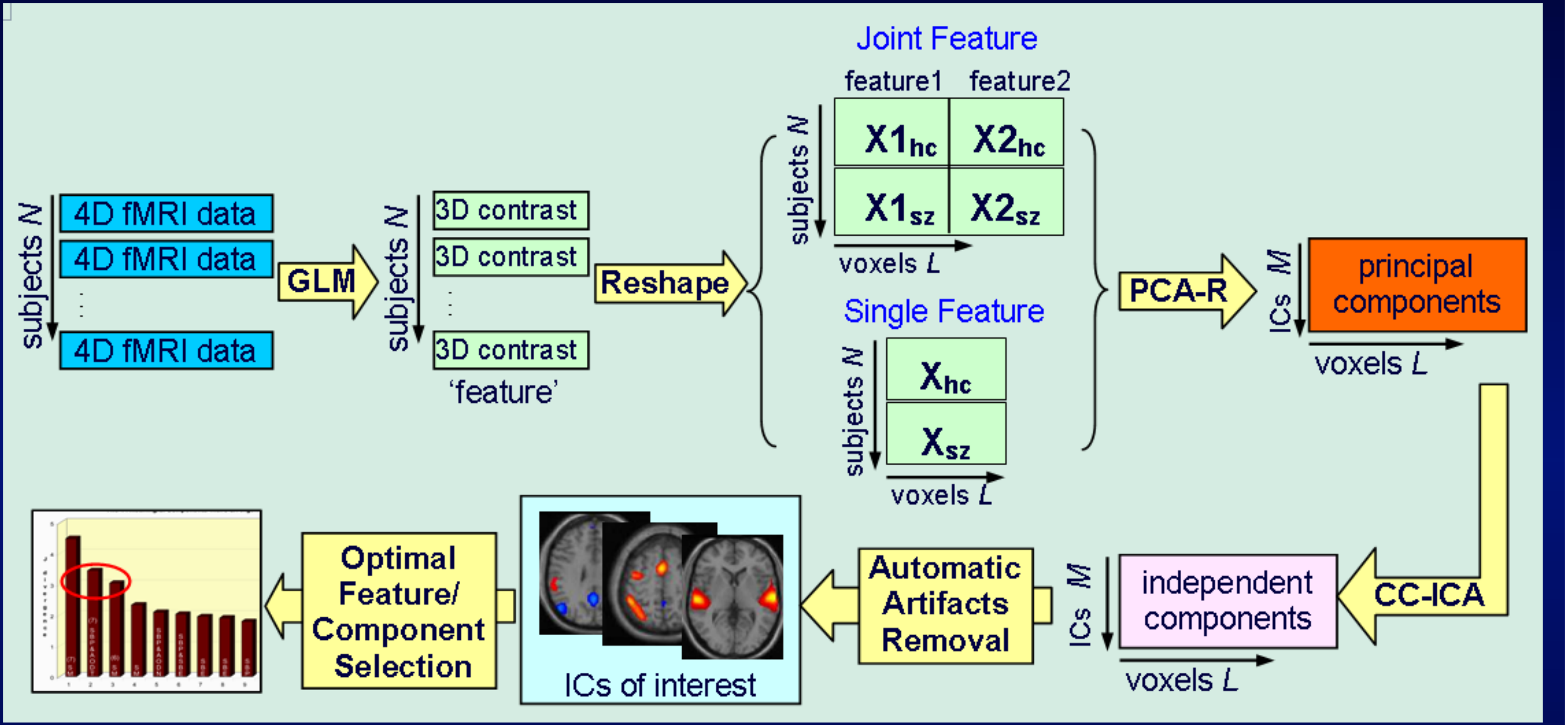
$$D(\mathbf{s}\|\mathbf{u}) = \int p_{\mathbf{s}}(\xi)\ln\left(\frac{p_{\mathbf{s}}(\xi)}{p_{\mathbf{u}}(\xi)}\right)d\xi$$

$$\mathbf{s},\mathbf{u} \in \mathbf{R}^2 \in \left\{\mathrm{F}_{\mathrm{SB}}, \mathrm{F}_{\mathrm{AOD\_N}}, \mathrm{F}_{\mathrm{AOD\_T}}, \mathrm{S}_{\mathrm{GM}}\right\}$$



V. D. Calhoun and T. Adali, "Feature-based Fusion of Medical Imaging Data," *IEEE Trans. Inf. Tech. in Biomedicine,* vol. 13, pp. 1-10, 2009

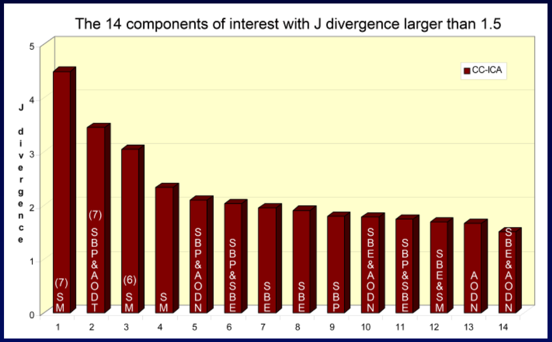
Optimal Selection of Discriminative Features

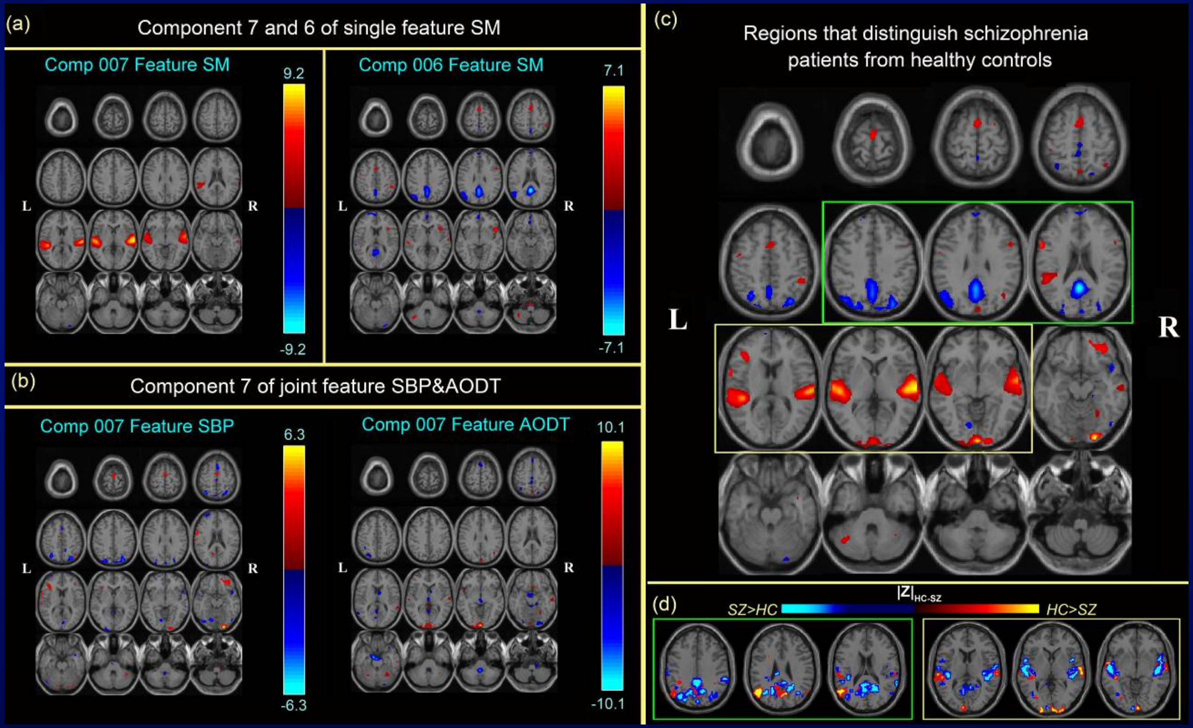


$$C = -E[\ln f(\mathbf{y})] + \lambda \cdot \sum T_i^2$$

J. Sui, T. Adali, G. Pearlson, and V. D. Calhoun, "An ICA-based Method for the Identification of Optimal FMRI Features and Components Using Combined Group-Discriminative Techniques," NeuroImage, vol. 46, pp. 73-86, 2009.

Optimal Selection of Discriminative Features

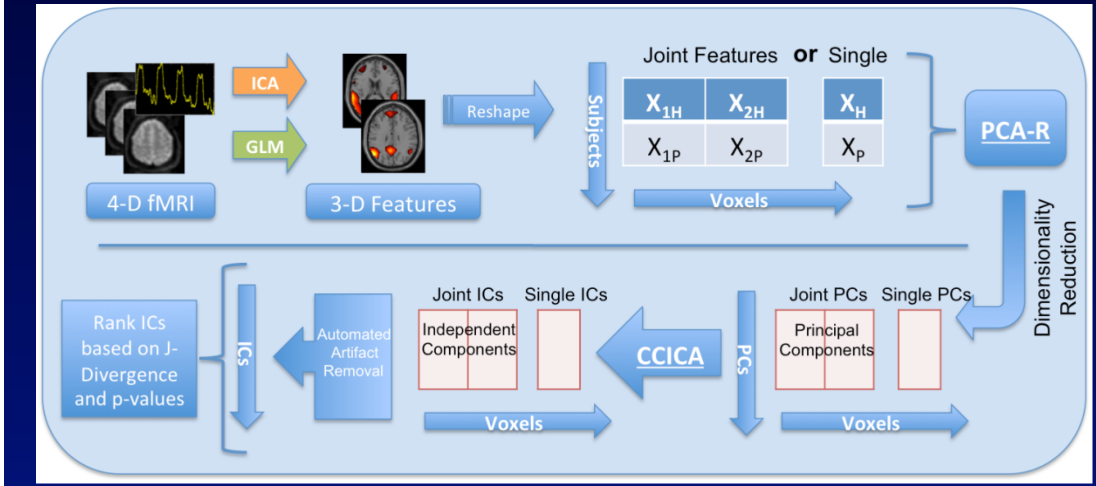




J. Sui, T. Adali, V. P. Clark, G. Pearlson, and V. D. Calhoun, "A Method for Accurate Group Difference Detection by Constraining the Mixing Coefficients in an ICA Framework," Human Brain Mapping, vol. 30, pp. 2953-2970, 2009.

J. Sui, T. Adali, G. Pearlson, and V. D. Calhoun, "An ICA-based Method for the Identification of Optimal FMRI Features and Components Using Combined Group-Discriminative Techniques," NeuroImage, vol. 46, pp. 73-86, 2009.

Joint/Single GLM/ICA Imaging Biomarkers



4-D fMRI

ICA

GLM

3-D Features

Reshape

Subjects

Joint Features **or** Single

X_{1H} X_{2H} X_H

X_{1P} X_{2P} X_P

Voxels

PCA-R

Dimensionality Reduction

Joint PCs Single PCs

Principal Components

PCs

Voxels

CCICA

Joint ICs Single ICs

Independent Components

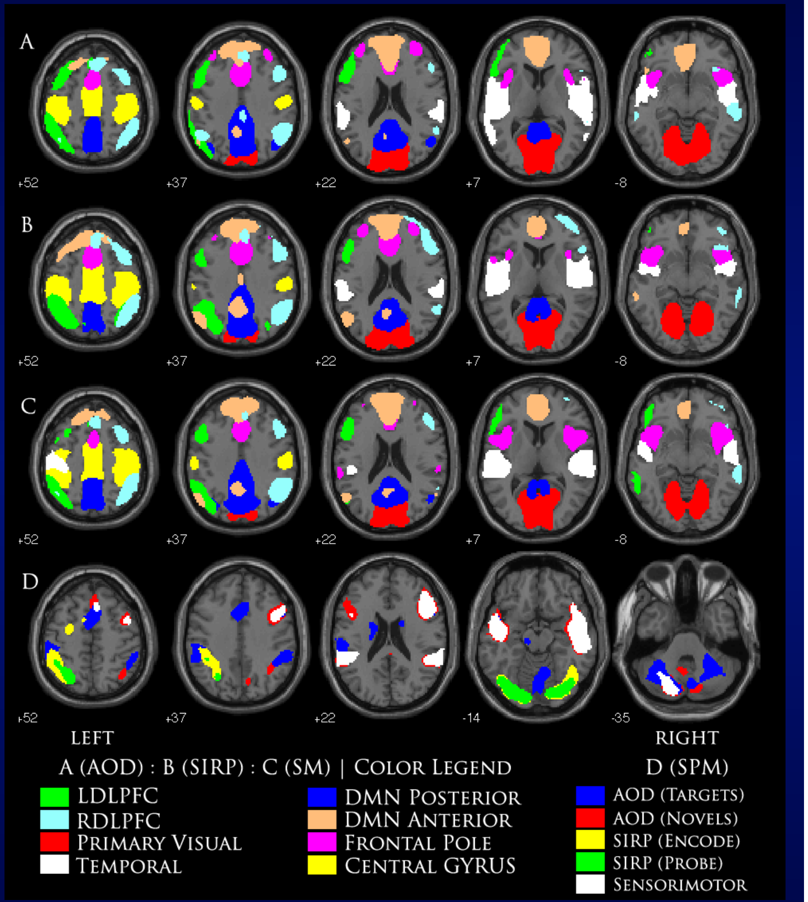
Voxels

Automated Artifact Removal

ICs

Rank ICs based on J-Divergence and p-values

Feature Description	Labels	Feature Type	Tasks
Left dorsal lateral prefrontal cortex	LDLPFC	ICA Component Map	AOD,SIRP,SM
Right dorsal lateral prefrontal cortex	RDLPFC	ICA Component Map	AOD,SIRP,SM
Primary Visual	V1	ICA Component Map	AOD,SIRP,SM
Bilateral Temporal	Temporal	ICA Component Map	AOD,SIRP,SM
Default Mode Network Posterior	DMN1	ICA Component Map	AOD,SIRP,SM
Default Mode Network Anterior	DMN2	ICA Component Map	AOD,SIRP,SM
Bilateral Frontal Pole	FPOLE	ICA Component Map	AOD,SIRP,SM
Pre/Post Central Gyrus	Central	ICA Component Map	AOD,SIRP,SM
Targets vs. Standards	Targets	SPM Contrast Map	AOD
Novels vs. Standards	Novels	SPM Contrast Map	AOD
Encode Block Average	Encode	SPM Contrast Map	SIRP
Probe Block Average	Probe	SPM Contrast Map	SIRP
Motor Tapping Block Average	Motor	SPM Contrast Map	SM



A (AOD) : B (SIRP) : C (SM) | COLOR LEGEND

D (SPM)

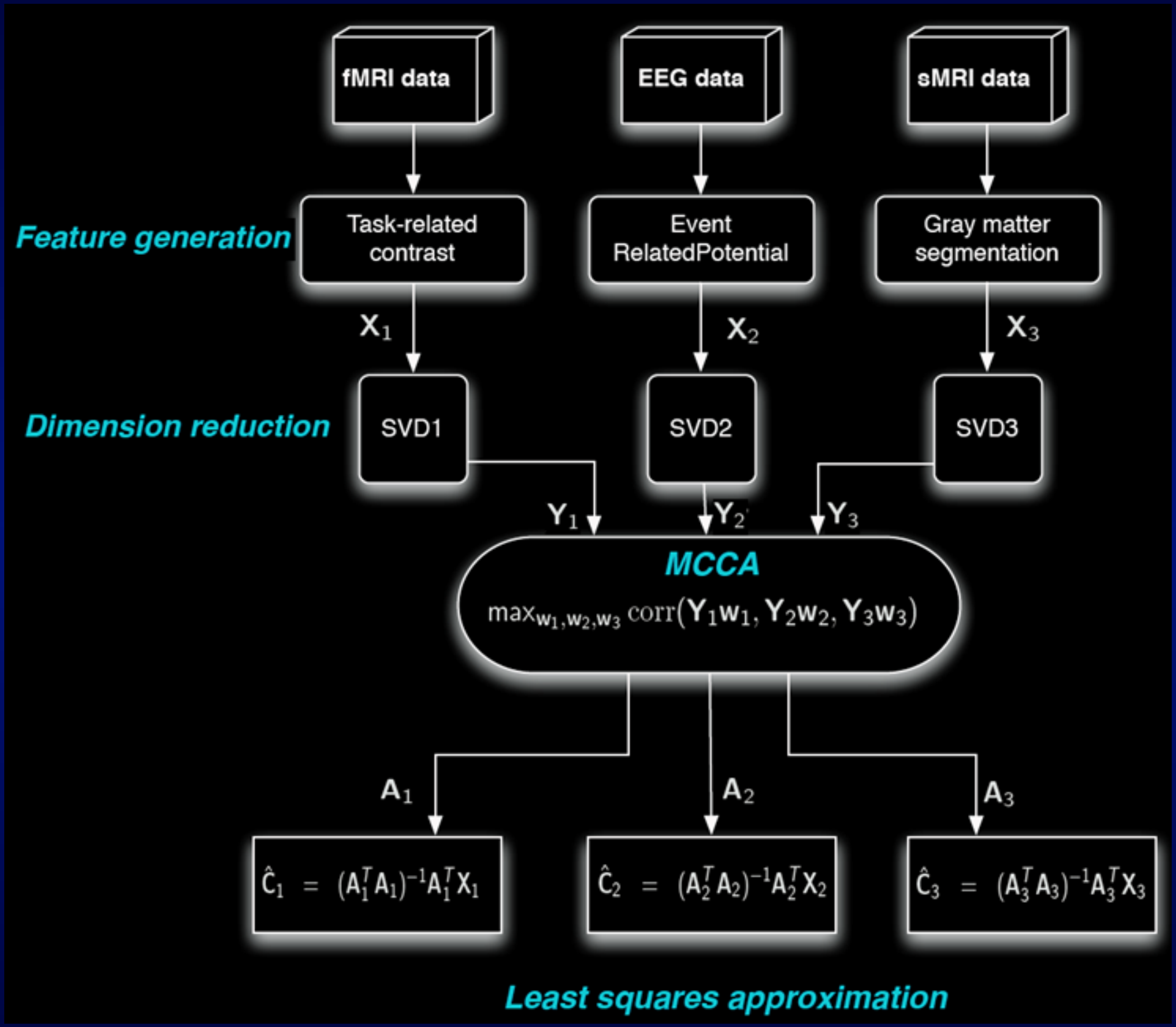
🟩 LDLPFC	🟦 DMN POSTERIOR	🟦 AOD (TARGETS)
⬜ RDLPFC	🟧 DMN ANTERIOR	🟥 AOD (NOVELS)
🟥 PRIMARY VISUAL	🟪 FRONTAL POLE	🟨 SIRP (ENCODE)
⬜ TEMPORAL	🟨 CENTRAL GYRUS	🟩 SIRP (PROBE)
		⬜ SENSORIMOTOR

D. Kim, J. Sui, S. Rachakonda, T. White, D. S. Manoach, V. P. Clark, B. C. Ho, S. C. Schulz, and V. D. Calhoun, "Identification of imaging biomarkers in schizophrenia: A coefficient-constrained independent component analysis of the Mind multi-site schizophrenia study," Journal of NeuroInformatics, vol. 8, pp. 213-229, 2010.

Overview

- Motivation
- Joint ICA
    - ERP/FMRI
    - Multitask-fMRI
    - sMRI/fMRI
    - CC-ICA/Biomarker Identification
- mCCA
    - sMRI/fMRI/ERP
- mCCA+jICA
    - DTI+fMRI
- Parallel ICA
    - SNPs
    - Epigenetics
- Conclusions

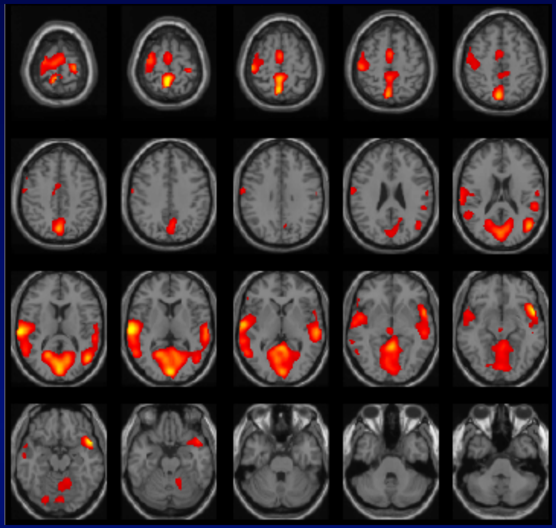
Multi-set Canonical Correlation Analysis



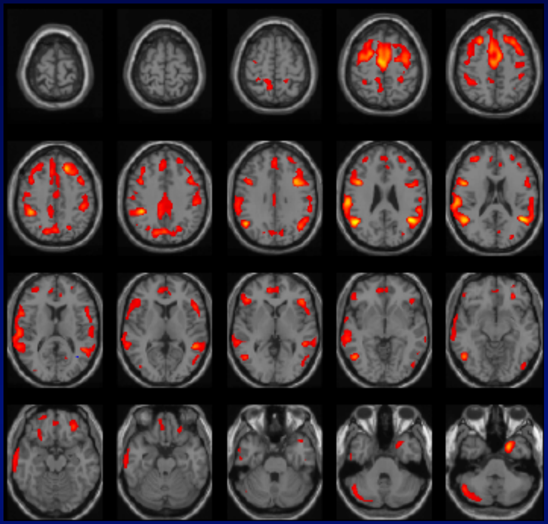
N. Correa, T. Adali, Y. Li, and V. D. Calhoun, "Canonical Correlation Analysis for Data Fusion and Group Inferences: Examining applications of medical imaging data," IEEE Signal Proc. Magazine, vol. 27, pp. 39-50, 2010.

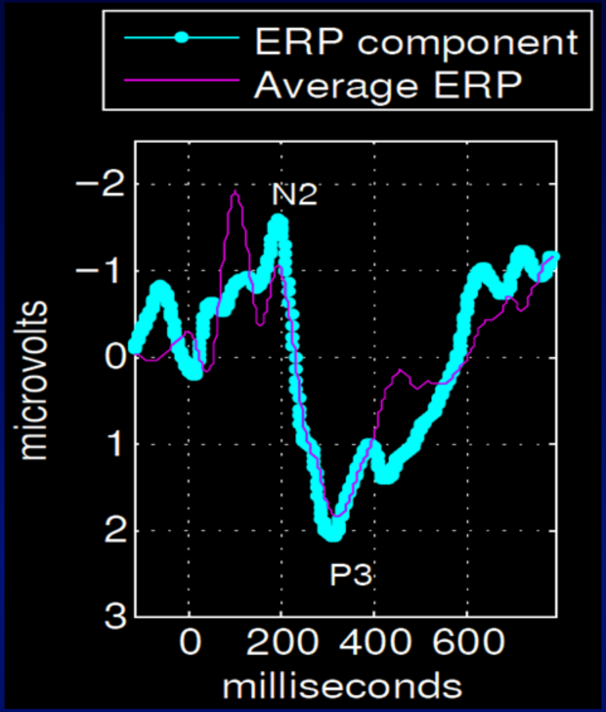
Three-way MCCA Results….

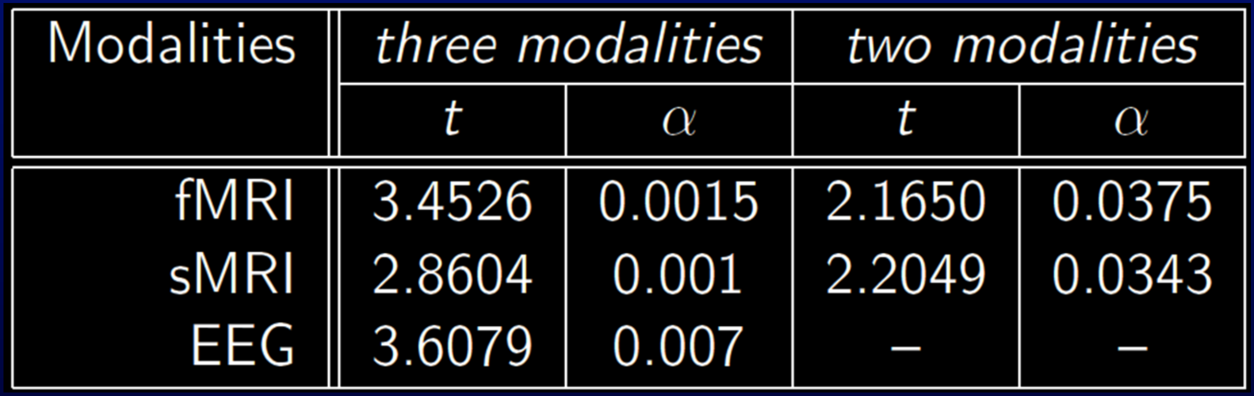
fMRI (Target)



sMRI (Gray Matter)







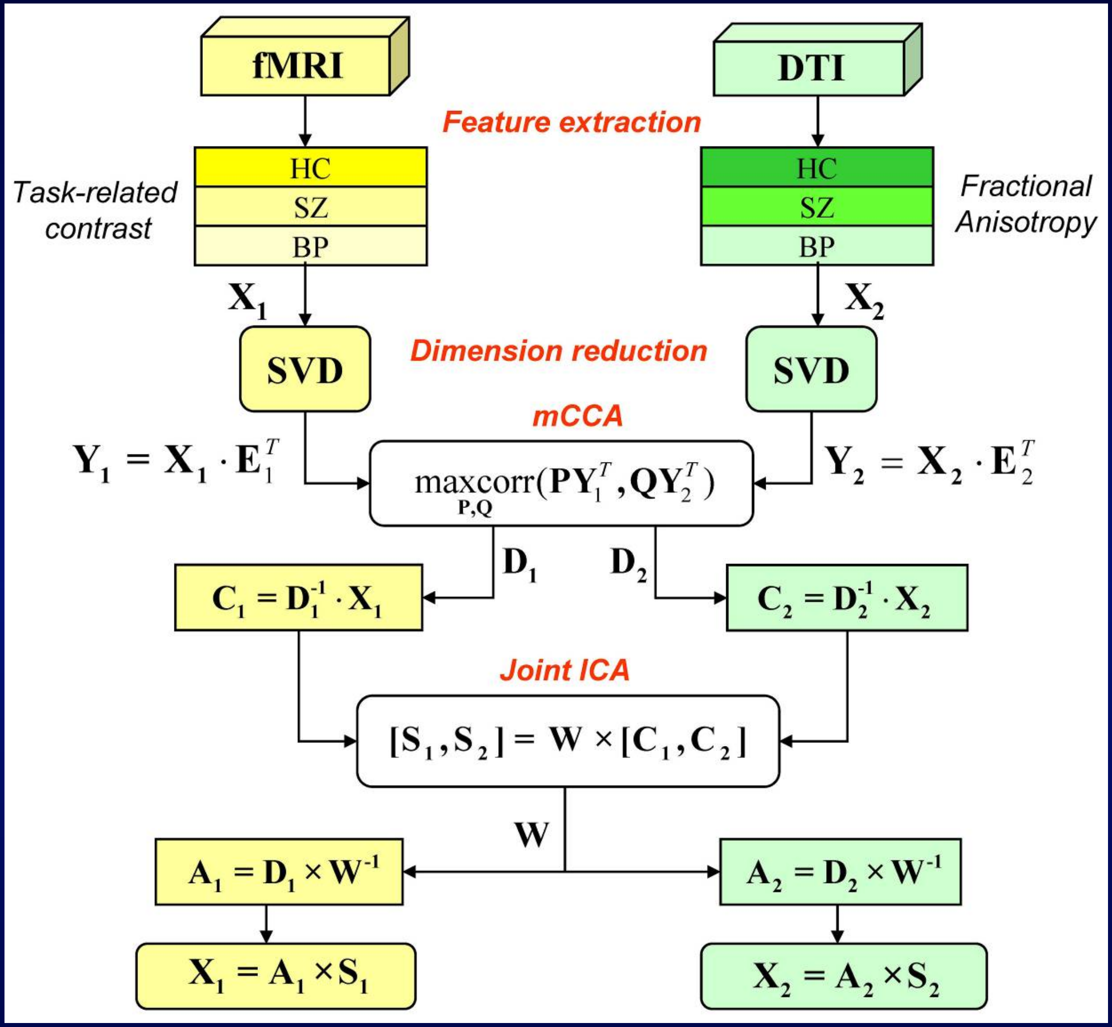
Modalities	*three modalities*		*two modalities*	
	t	α	t	α
fMRI	3.4526	0.0015	2.1650	0.0375
sMRI	2.8604	0.001	2.2049	0.0343
EEG	3.6079	0.007	–	–

N. Correa, T. Adali, Y. Li, and V. D. Calhoun, "Canonical Correlation Analysis for Data Fusion and Group Inferences: Examining applications of medical imaging data," IEEE Signal Proc. Magazine, vol. 27, pp. 39-50, 2010.

Overview

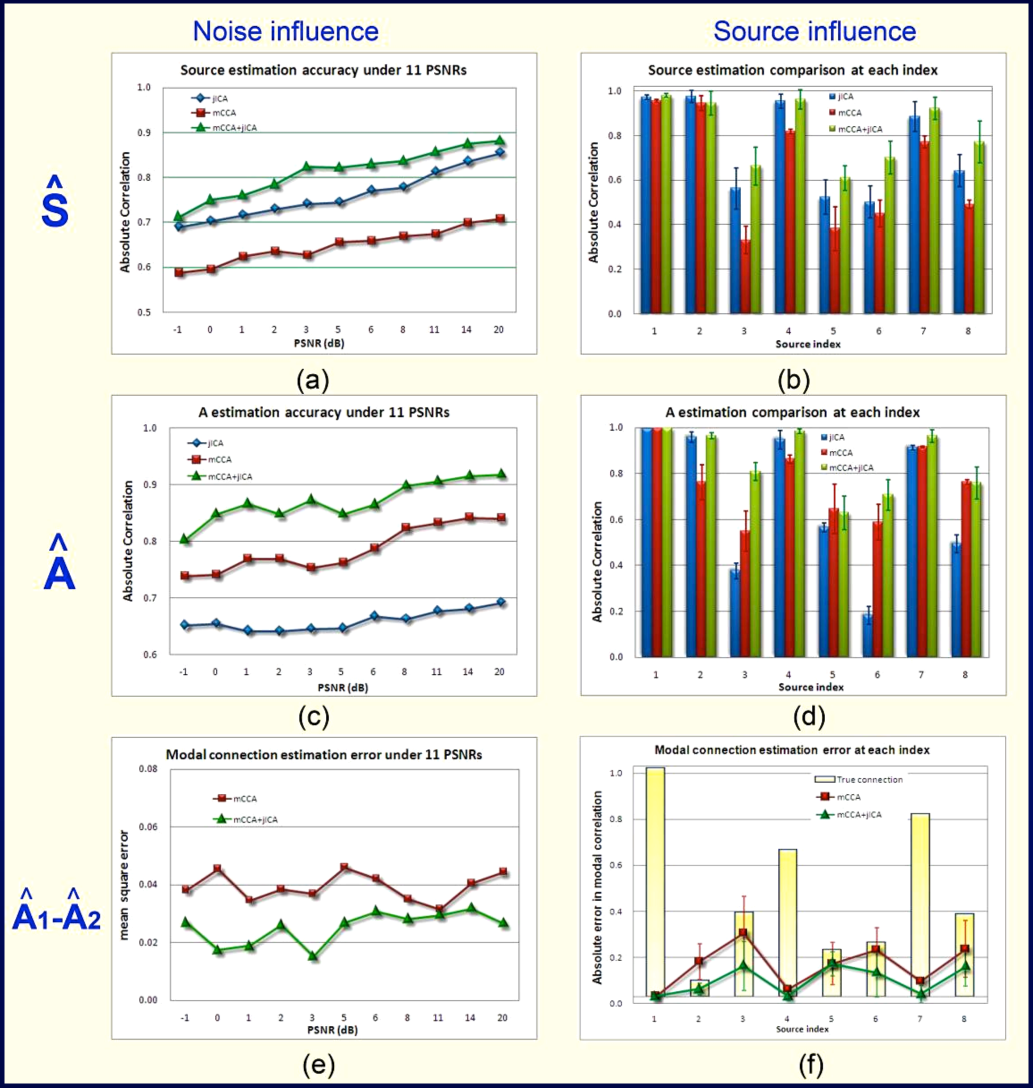
- Motivation
- Joint ICA
  - ERP/FMRI
  - Multitask-fMRI
  - sMRI/fMRI
  - CC-ICA/Biomarker Identification
- mCCA
  - sMRI/fMRI/ERP
- mCCA+jICA
  - DTI+fMRI
- Parallel ICA
  - SNPs
  - Epigenetics
- Conclusions

Fusion of DTI & fMRI via mCCA + Joint ICA



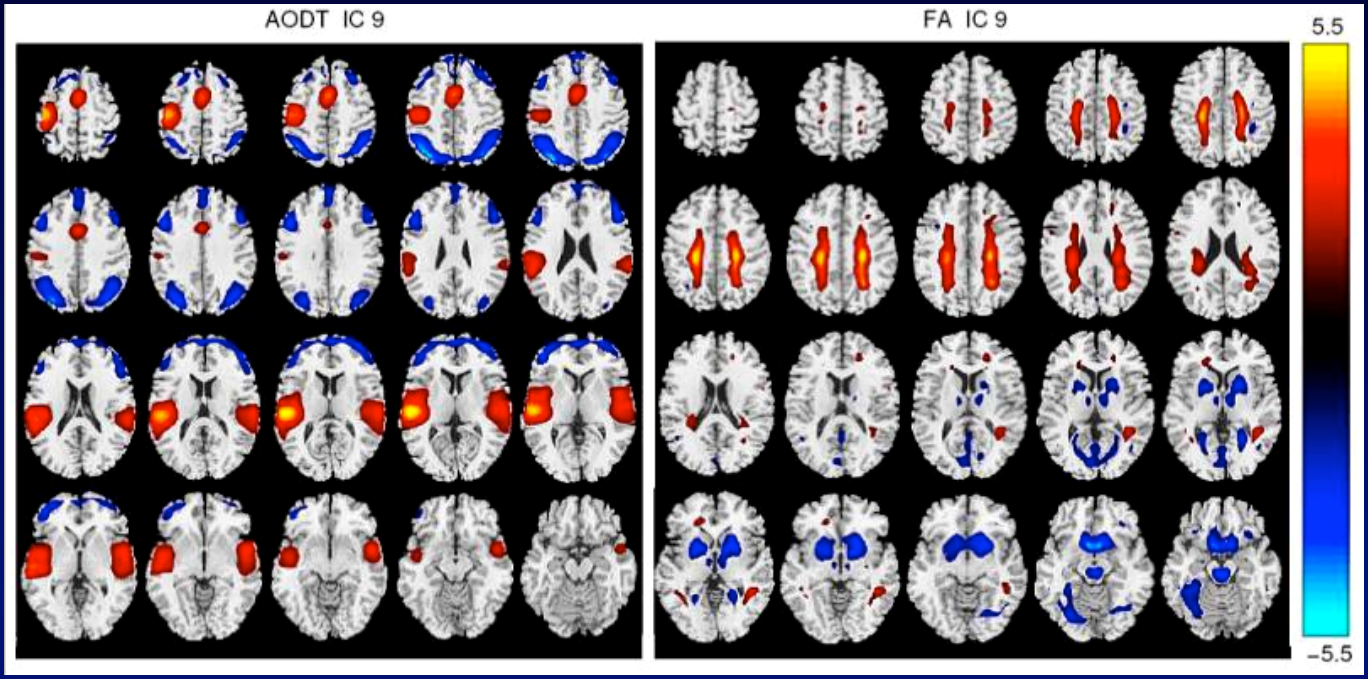
J. Sui, G. D. Pearlson, T. Adali, K. A. Kiehl, A. Caprihan, J. Liu, J. Yamamoto, and V. D. Calhoun, "Discriminating Schizophrenia and Bipolar Disorder by Fusing FMRI and DTI in A Multimodal CCA+ Joint ICA Based Model," NeuroImage, In Press

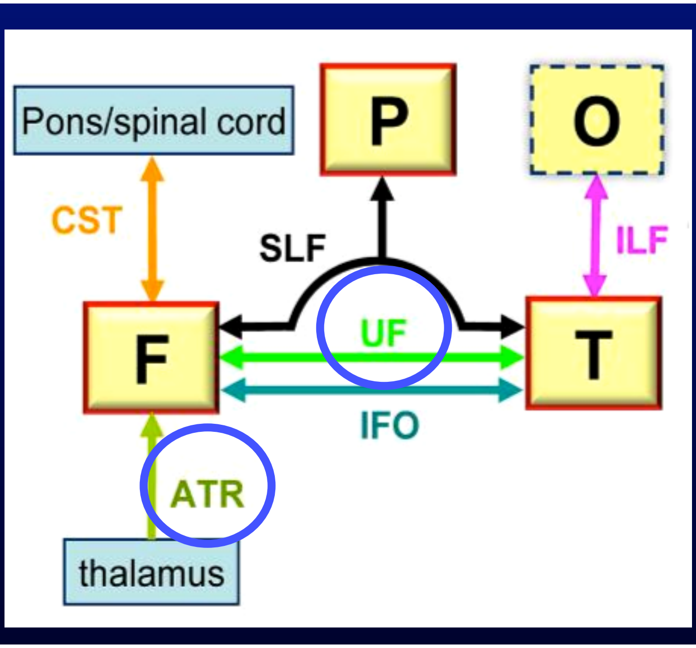
Comparison of jICA, mCCA, & mCCA+jICA



J. Sui, G. D. Pearlson, T. Adali, K. A. Kiehl, A. Caprihan, J. Liu, J. Yamamoto, and V. D. Calhoun, "Discriminating Schizophrenia and Bipolar Disorder by Fusing FMRI and DTI in A Multimodal CCA+ Joint ICA Based Model," NeuroImage, In Press

Temporal Lobe

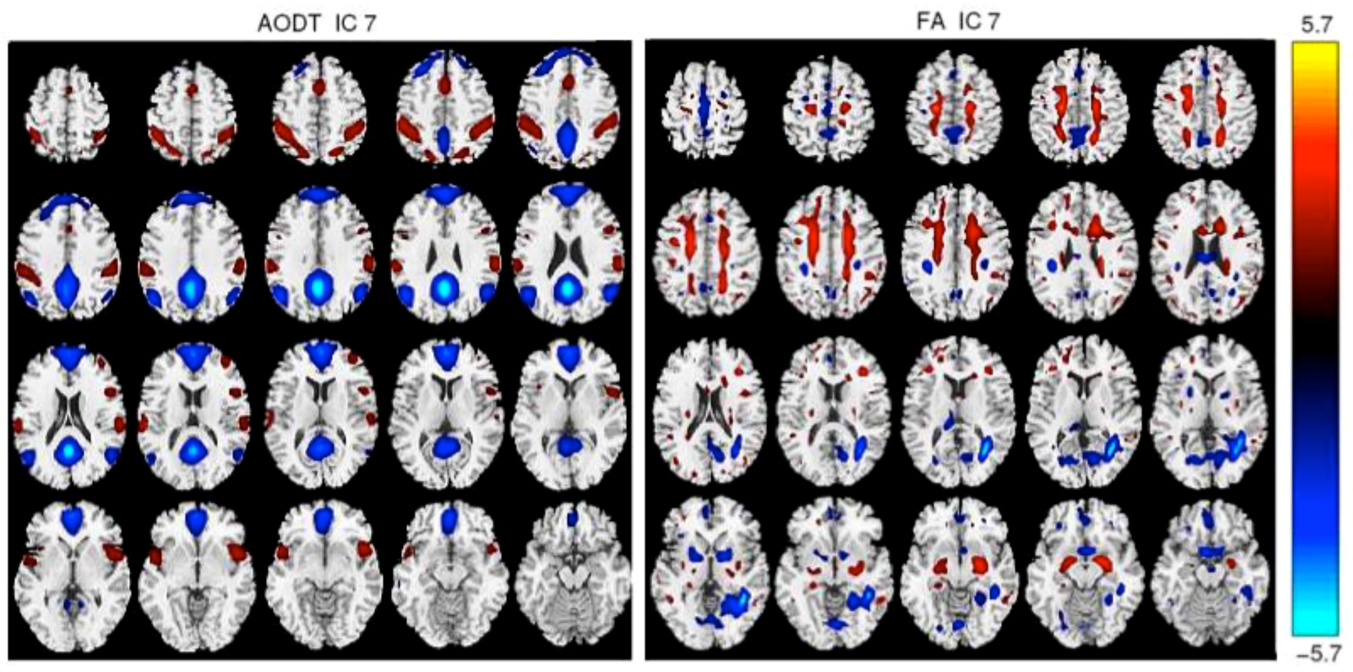


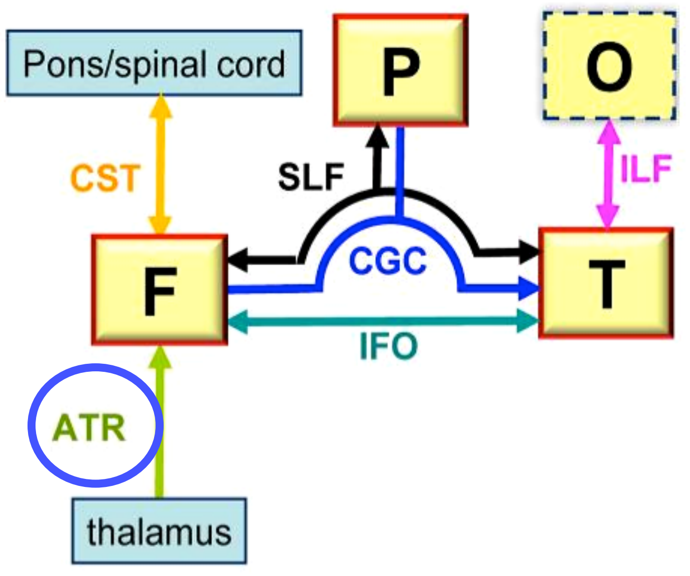


F: frontal lobe
P: parietal lobe
T: temporal lobe
O: occipital lobe

SLF: Superior longitudinal fasciculus
CGC: Cingulum
CST: Corticospinal tract
UF: Uncinate fasciculus
IFO: Inferior fronto-occipital fasciculus
ILF: Inferior longitudinal fasciculus
ATR: Anterior thalamic radiation
PH: Parahippocampal Gyrus
FMIN/MAJ: Forceps Major/Minor

Default Mode



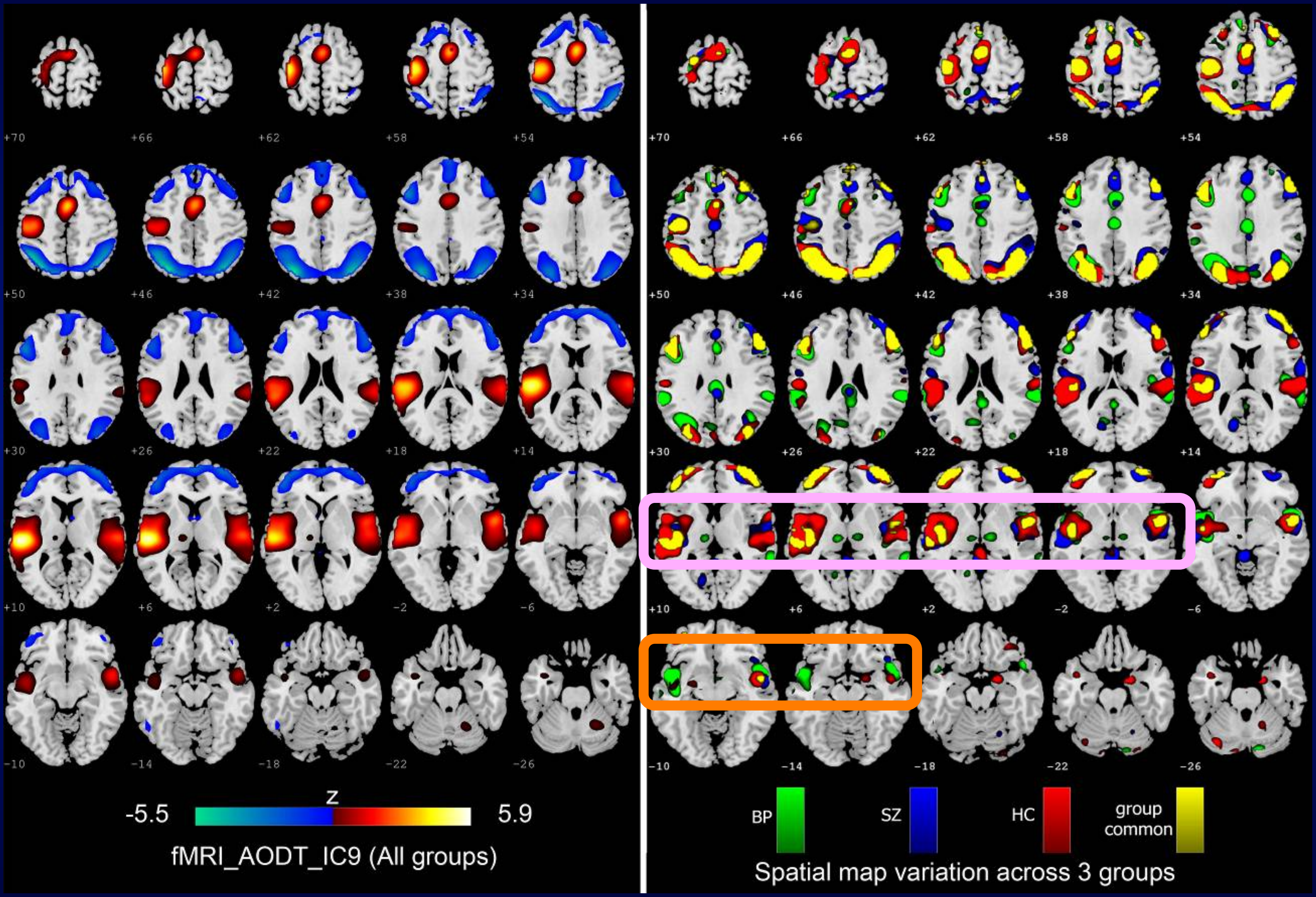


Abbreviations :

F: frontal lobe
P: parietal lobe
T: temporal lobe
O: occipital lobe

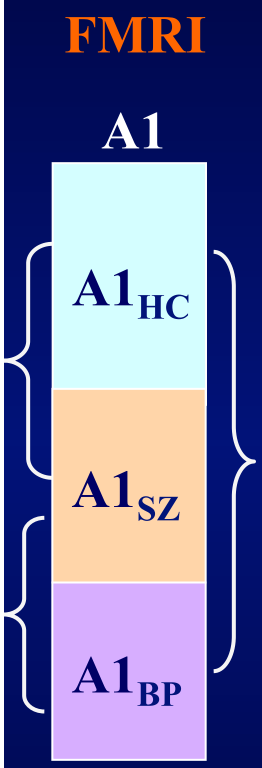
SLF: Superior longitudinal fasciculus
CGC: Cingulum
CST: Corticospinal tract
UF: Uncinate fasciculus
IFO: Inferior fronto-occipital fasciculus
ILF: Inferior longitudinal fasciculus
ATR: Anterior thalamic radiation.

Variation of functional spatial maps



J. Sui, G. D. Pearlson, T. Adali, K. A. Kiehl, A. Caprihan, J. Liu, J. Yamamoto, and V. D. Calhoun, "Discriminating Schizophrenia and Bipolar Disorder by Fusing FMRI and DTI in A Multimodal CCA+ Joint ICA Based Model," NeuroImage, In Press

Results from Schizophrenia, Bipolar, and Healthy Individuals



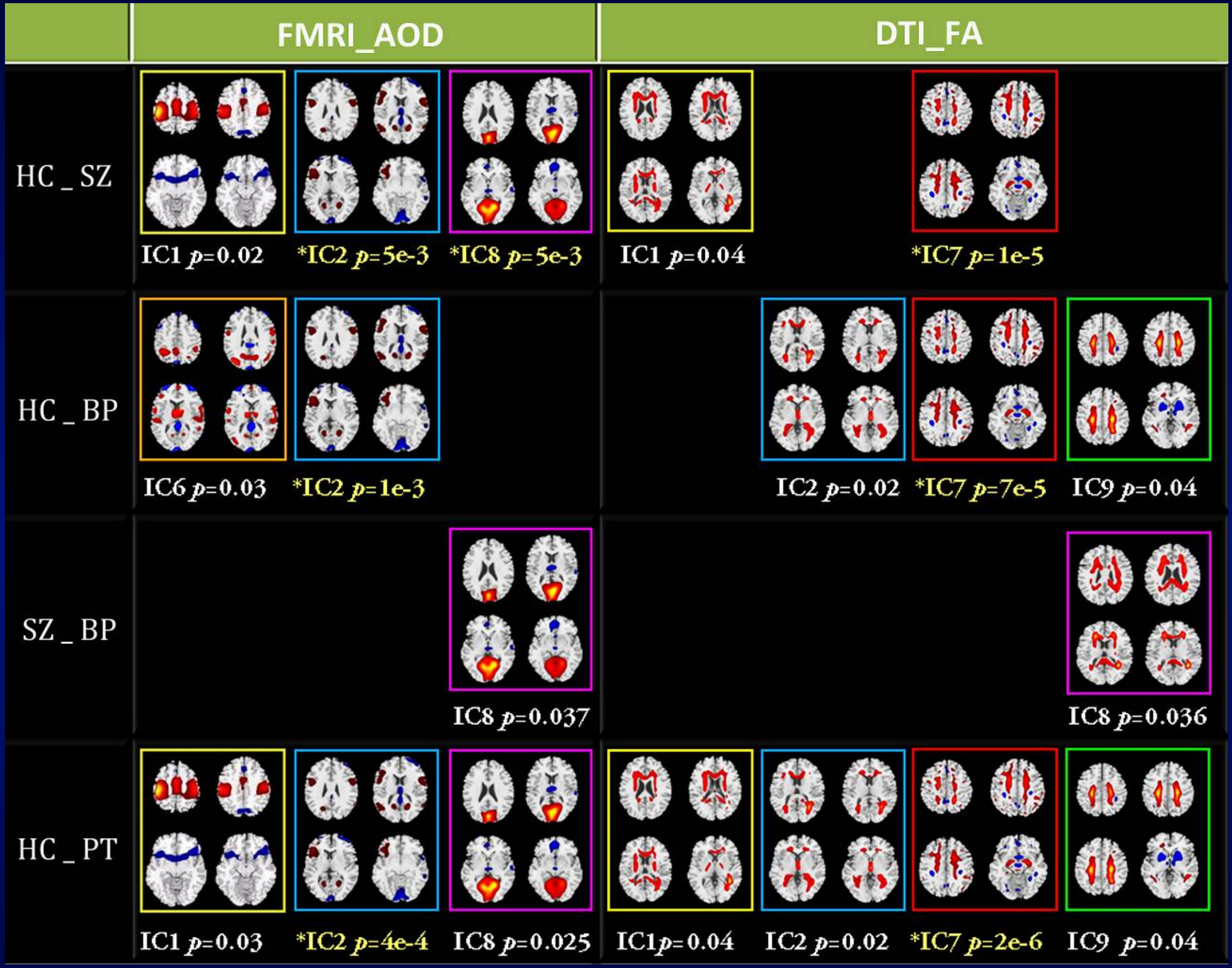
FMRI

A1

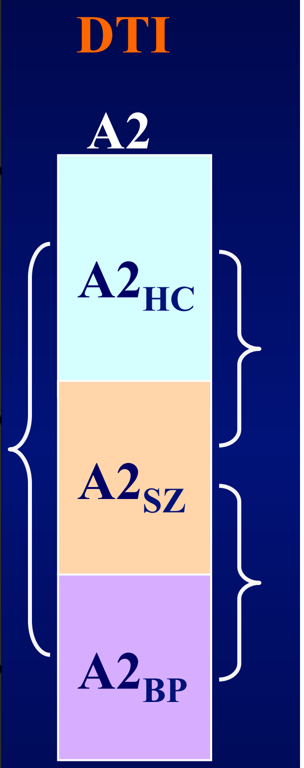
$A1_{HC}$

$A1_{SZ}$

$A1_{BP}$



	FMRI_AOD	DTI_FA
HC _ SZ	IC1 p=0.02 *IC2 p=5e-3 *IC8 p=5e-3	IC1 p=0.04 *IC7 p=1e-5
HC _ BP	IC6 p=0.03 *IC2 p=1e-3	IC2 p=0.02 *IC7 p=7e-5 IC9 p=0.04
SZ _ BP	IC8 p=0.037	IC8 p=0.036
HC _ PT	IC1 p=0.03 *IC2 p=4e-4 IC8 p=0.025	IC1p=0.04 IC2 p=0.02 *IC7 p=2e-6 IC9 p=0.04



DTI

A2

$A2_{HC}$

$A2_{SZ}$

$A2_{BP}$

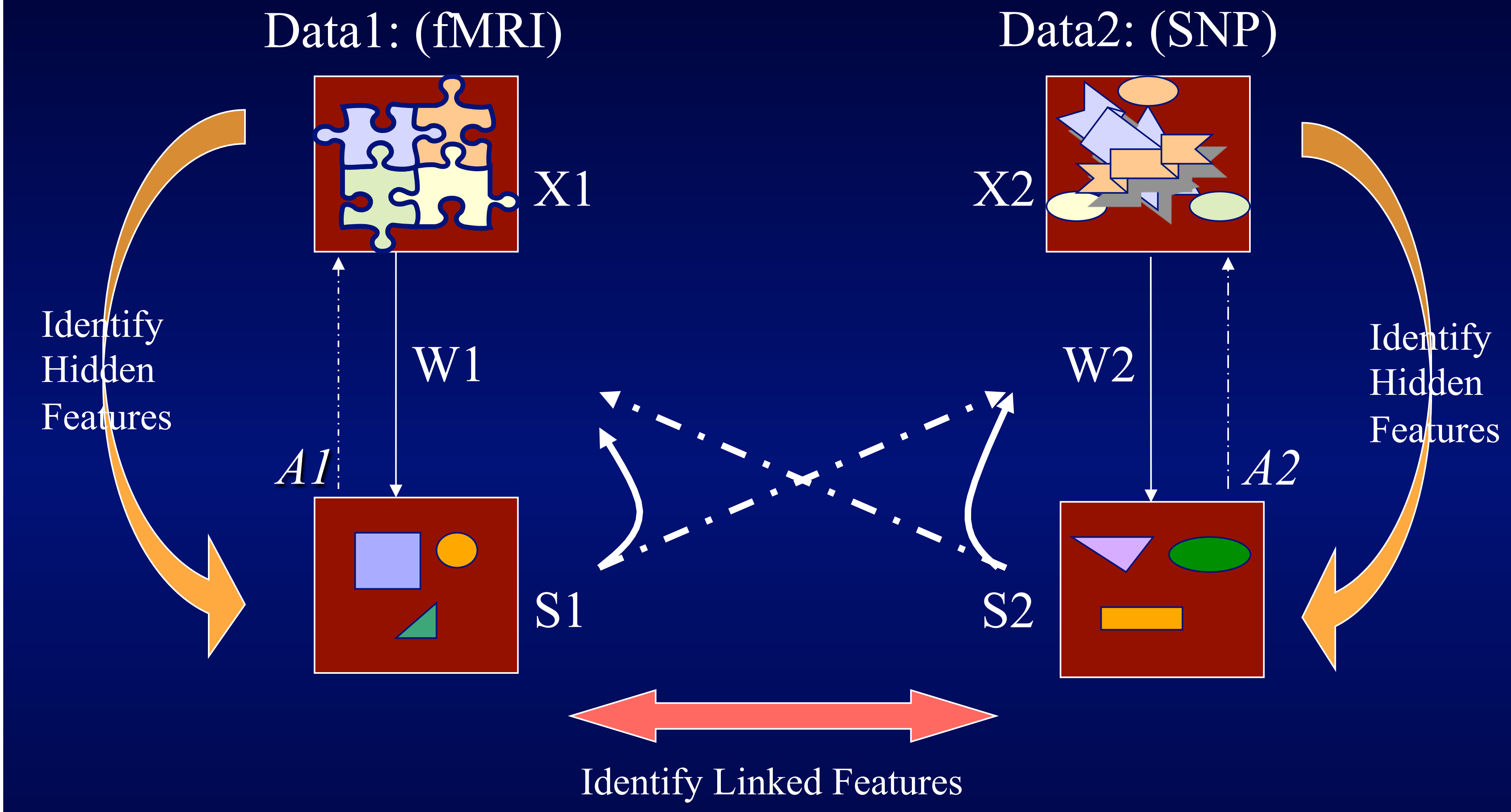
If the two ICs have the same frame color in two modalities, they are joint ICs

J. Sui, G. D. Pearlson, T. Adali, K. A. Kiehl, A. Caprihan, J. Liu, J. Yamamoto, and V. D. Calhoun, "Discriminating Schizophrenia and Bipolar Disorder by Fusing FMRI and DTI in A Multimodal CCA+ Joint ICA Based Model," NeuroImage, In Press

Overview

- Motivation
- Joint ICA
  - ERP/FMRI
  - Multitask-fMRI
  - sMRI/fMRI
  - CC-ICA/Biomarker Identification
- mCCA
  - sMRI/fMRI/ERP
- mCCA+jICA
  - DTI+fMRI
- Parallel ICA
  - SNPs
  - Epigenetics
- Conclusions

Parallel ICA: Two Goals



Identify Hidden Features

Identify Hidden Features

Identify Linked Features

MAX : $\{H(\mathbf{Y1}) + H(\mathbf{Y2})\}$, <Infomax>

Subject to: arg max $g\{W1, W2\ \hat{s}1, \hat{s}2\}$, $g(\cdot) = \text{Correlation}(A_1, A_2)^2 = \dfrac{\text{Cov}(a_{1i}, a_{2j})^2}{\text{Var}(a_{1i}) \times \text{Var}(a_{2j})}$

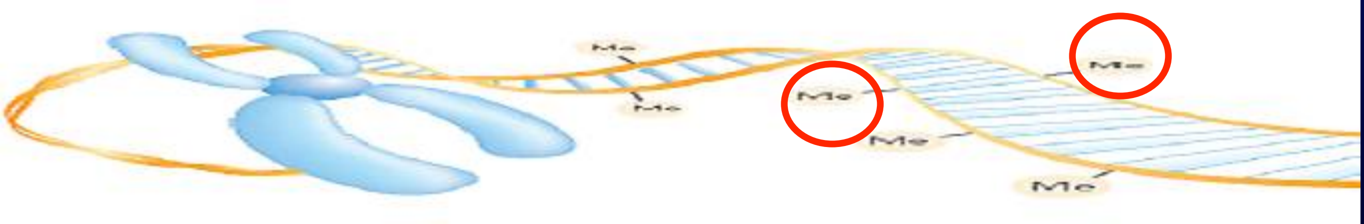
J. Liu, O. Demirci, and V. D. Calhoun, "A Parallel Independent Component Analysis Approach to Investigate Genomic Influence on Brain Function," *IEEE Signal Proc. Letters,* vol. 15, pp. 413-416, 2008.

Genetic Information

- **Genetic: single nucleotide polymorphism (SNP)**

- **Genetic: copy number variation (CNV)**

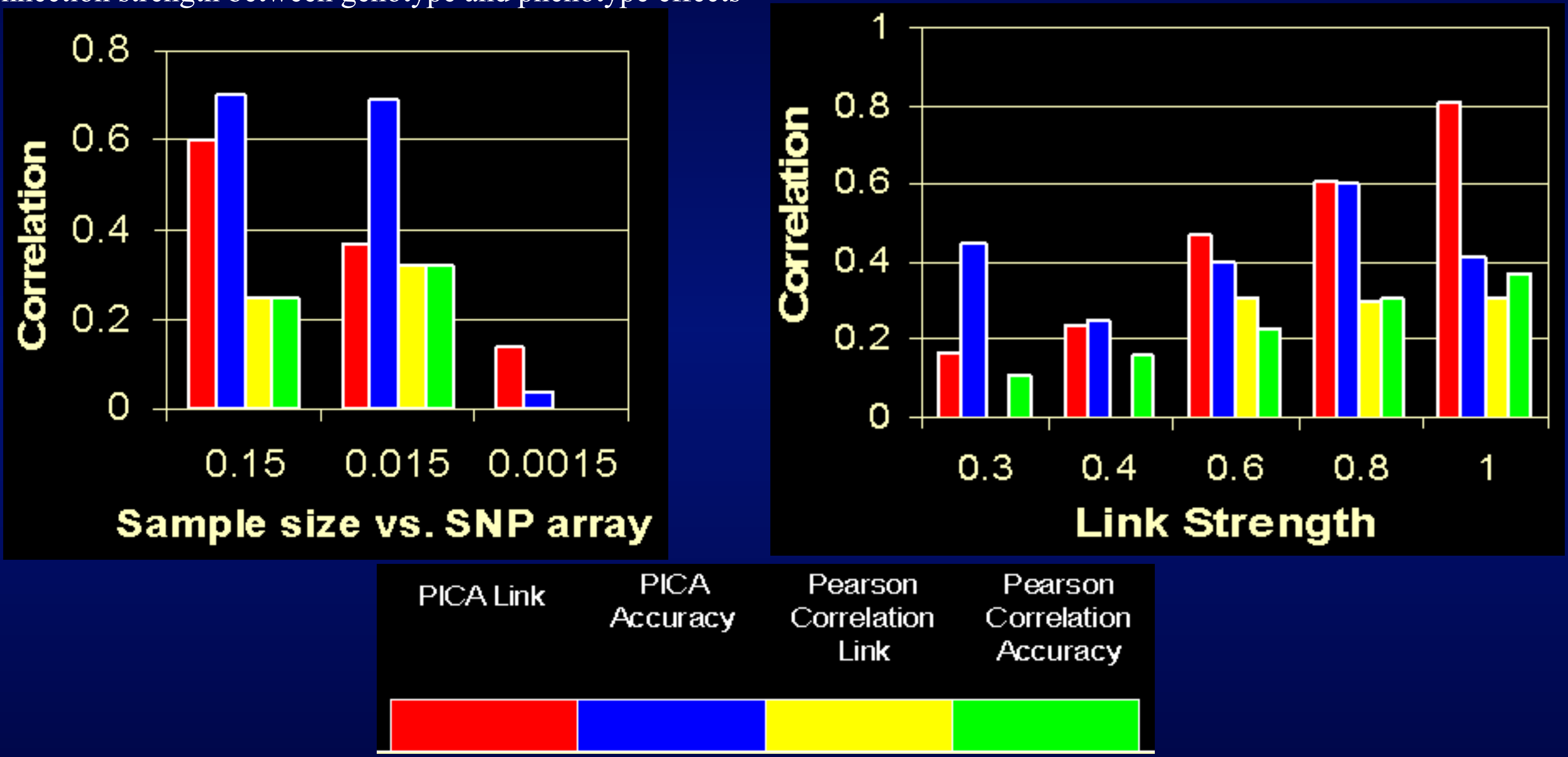
- **Epigenetic: methylation**



Simulation

Simulation: Designed to provide a more complete understanding of Parallel ICA while applied to genomic SNP array studies. We specified the parameters for each component and input them into PLINK, an open-source whole genome association analysis toolset [http://pngu.mgh.harvard.edu/purcell/plink/].

Conditions: sample size effect. case to control ratio, SNP array size effect, case-related SNP's vs. total SNP's, odds ratio, connection strength between genotype and phenotype effects



Simulation results suggest that parallel ICA, in general, is able to extract more accurately the components and connections than a correlation test, in particular for weak linkages. Results also indicate that the ratio of sample size to SNP size should be at least 0.02. However, when the data have a low odds ratio or cases vs. controls ratio, the correlation test provides results reliably, though with lower accuracy.

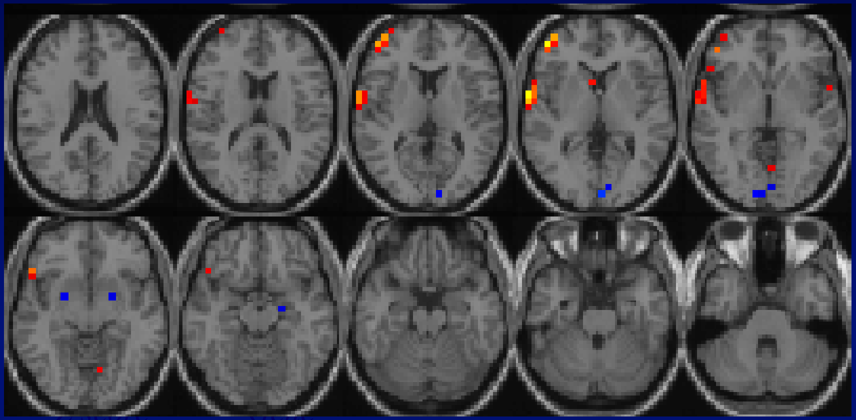
Liu, J, et al, *IEEE Bioinformatics and Biomedicine*. 2008: Philadelphia, PA.

Initial Proof of Concept: SNP/fMRI Fusion

Data Description: 20 Sz & 43 Healthy controls

fMRI: one image per subject (Target activation in AOD task)

SNP: one array per subject (384 SNP genotypes - -> 367 SNPs)



Control vs Patient
$p<0.001$

SNP	*Z score*	*Gene*
Rs1466163	-4.08	AADC: aromatic L-amino acid decarboxylase
Rs2429511	3.97	ADRA2A: alpha-2A adrenergic receptor gene
Rs3087454	**-3.09**	**CHRNA7: alpha 7 nicotinic cholinergic receptor**
Rs821616	**2.96**	**DISC1: disrupted in schizophrenia 1**
Rs885834	**-2.78**	**CHAT: choline acetyltransferase**
Rs1355920	**-2.77**	**CHRNA7: cholinergic receptor, nicotinic, alpha 7**
R4765623	2.73	SCARB1: scavenger receptor class B, member 1
Rs4784642	-2.71	GNAO1: guanine nucleotide binding protein (G protein), alpha activating activity polypeptide O
Rs2071521	2.58	APOC3: apolipoprotein C-III
Rs7520974	2.55	CHRM3: muscarinic-3 cholinergic receptor

J. Liu, G. D. Pearlson, A. Windemuth, G. Ruano, N. I. Perrone-Bizzozero, and V. D. Calhoun, "Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA," *Hum.Brain Map.*, vol. 30, pp. 241-255, 2009

Larger scale study of schizophrenia

- Schizophrenia patients and healthy controls
  - MCIC data: Boston, Iowa, Minnesota and New Mexico
  - Genome-wide 1M SNP data - [biallelic coding (AA, AB, or BB]
  - fMRI sensorimotor task- Block design motor response to auditory stimulation
- SNP data
  - Subject control: heterozygosity, nearest neighbor, 2nd degree or closer relatives, duplication
  - SNP control: missing genotyping ratio, minor allele frequency, Hardy-Weinberg equilibrium, linkage disequilibrium, etc.
  - Population stratification correction: using PCA
  - Coding (0 for ‘AA’, 1 for ‘AB, and 2 for ‘BB’)
- fMRI data
  - SPM preprocessing (alignment, normalization, filter, GLM) and contrast image
  - Outlier subject excluded
  - Datasets: 208 subjects have SNP (777365 SNPs) and fMRI data (52322 voxels)

SNP Selection and Approach

- SNP selection
  - SNPs differentiating schizophrenia patients and healthy controls are included (p-value < 0.005, 3318 SNPs)
  - SNPs related to schizophrenia risk genes are included (1843 SNPs selected, related to DISC1, COMT, etc.)
  - Combine selected SNPs: 5157 SNPs as final input
- Parallel ICA
  - Use subject type (SZ patients. vs. healthy control) as reference
  - Apply reference PCA to SNP and fMRI data to reduce dimension and select components of interest
  - Apply parallel ICA to identify linked components optimized to the correlation between the two modalities

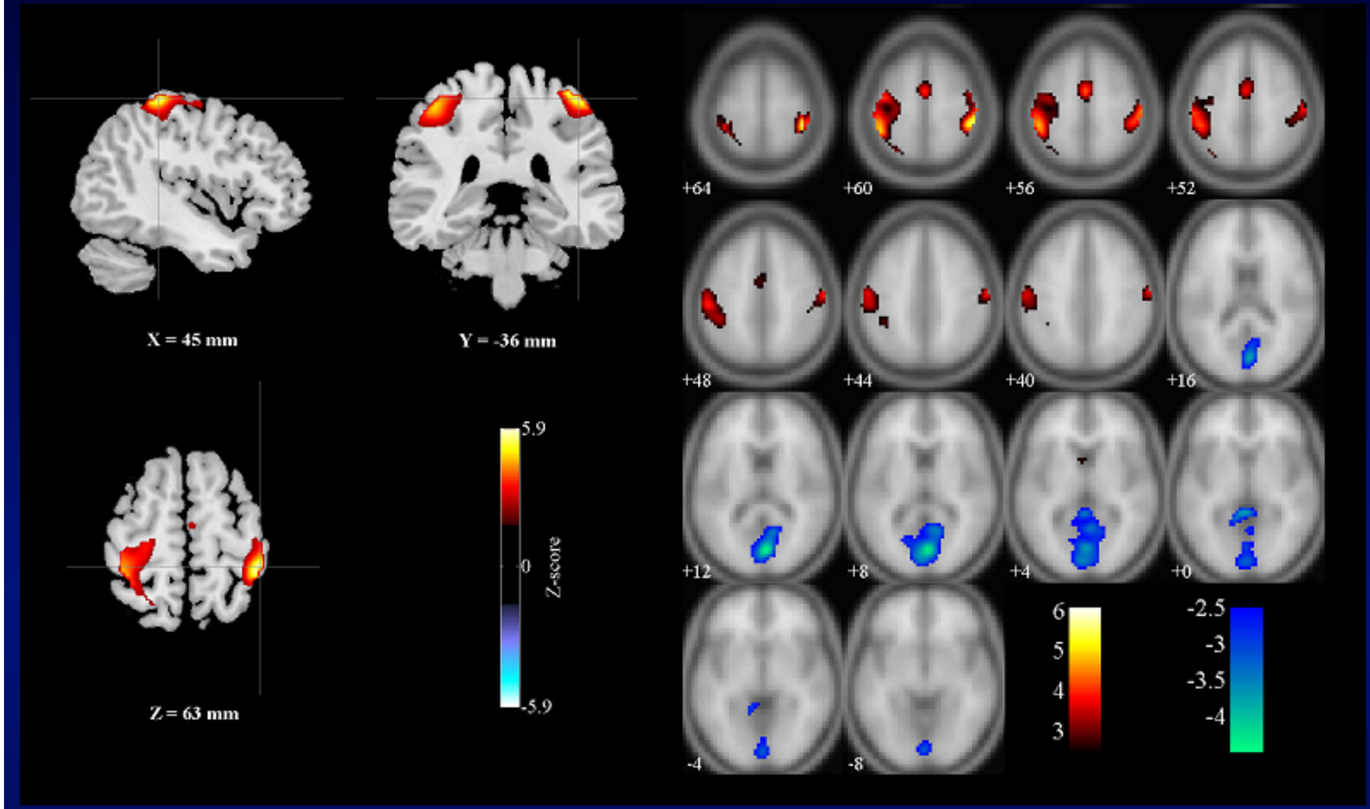
Resulting Linked Component

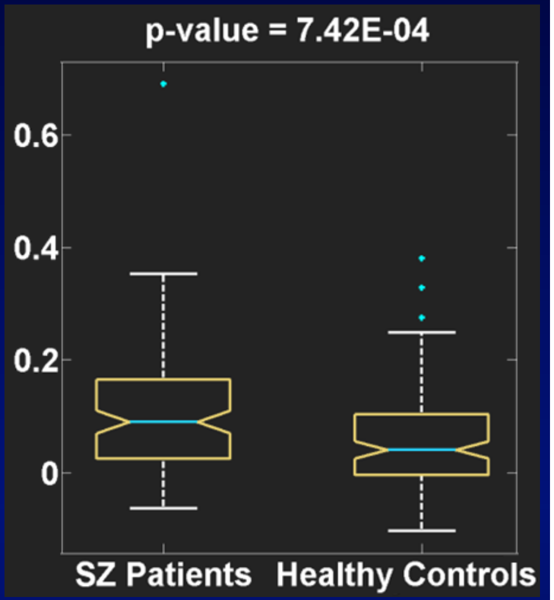
- fMRI component number = 8, SNP component number = 5
- One pair of linked components is identified, with p-value passing Bonferroni correction

fMRI component index	SNP component index	$r_{fMRI\text{-}SNP}$	P-value
1	3	-0.065	3.49E-01
2	2	0.042	5.48E-01
3	1	0.099	1.51E-01
4	2	-0.138	4.54E-02
5	1	0.141	4.16E-02
6	5	-0.128	6.44E-02
7	1	0.178	9.95E-03
8	**4**	**0.282**	**3.39E-05**

- Bootstrap: Multiple runs of parallel-ICA with 5157 randomly selected SNPs. The median correlation was 0.16

Identified fMRI component





Brain region	Brodmann area	Volume	Max z-score
Postcentral Gyrus	1, 2, 3, 5, 7, 40, 43	9.7/4.4	6.23(42,-32,60)/5.70(-39,-32,62)
Precentral Gyrus	4, 6	8.2/1.7	6.03(42,-12,56)/4.40(-59,-12,42)
Inferior Parietal Lobule	40	3.0/0.5	6.03(45,-35,57)/4.88(-45,-32,57)
Medial Frontal Gyrus	:6, 32	1.0/1.6	4.82(3,-3,53)/5.27(0,-3,53)
Superior Temporal Gyrus	21, 22, 38	1.4/0.5	3.24(62,-12,1)/2.93(-50,-3,-5)

Summary of SNP Results

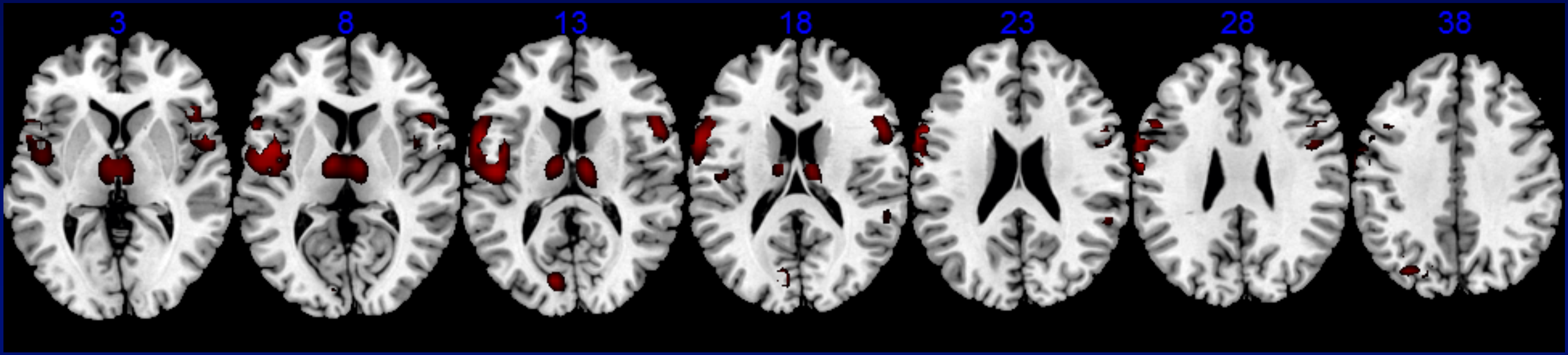
- Conduct pathway analysis and functional annotation clustering based on identified 94 genes
  - IPA (Ingenuity Pathway Analysis) identifies "Schizophrenia of humans" as one of the top biofunctions, involving 11 genes
  - IPA also identifies a number of significant canonical pathways, four of which are related to neurotransmitter signaling
  - David's Bioinformatics Resource reports the most significant cluster to be functionally related to synapse. A cluster annotated as "cell projection" is also identified

Disease and disorder	Gene	p-value
Schizophrenia of humans	BDNF, COMT, DISC1, DRD3, ERBB4, GAD1, GRIN2B, HTR7, NOTCH4, NRG1, NRG3	6.49E-09
Neurotransmitter signaling pathway	**Gene**	**p-value**
GABA receptor signaling	GABRA4, GABRG3, GAD1	2.13E-03
Dopamine receptor signaling	COMT, DRD3, PPP2R2C	7.66E-03
Neuregulin signaling	NRG1, NRG3, ERBB4	1.15E-02
Glutamate receptor signaling	GRIN2B, GRID2	3.74E-02
Functional annotation cluster	**Gene**	**p-value**
Synapse	GABRA4, GABRG3, GAD1, GRIN2B, GRID2, ERBB4, SHC4, OTOF, PSD3, CTBP2	5.20E-03
Cell projection	DRD3, GAD1, GRIN2B, MYCBP2, DNAH11, WNT2, ESR1, CDH13, ALCAM, MYO5A	9.70E-02

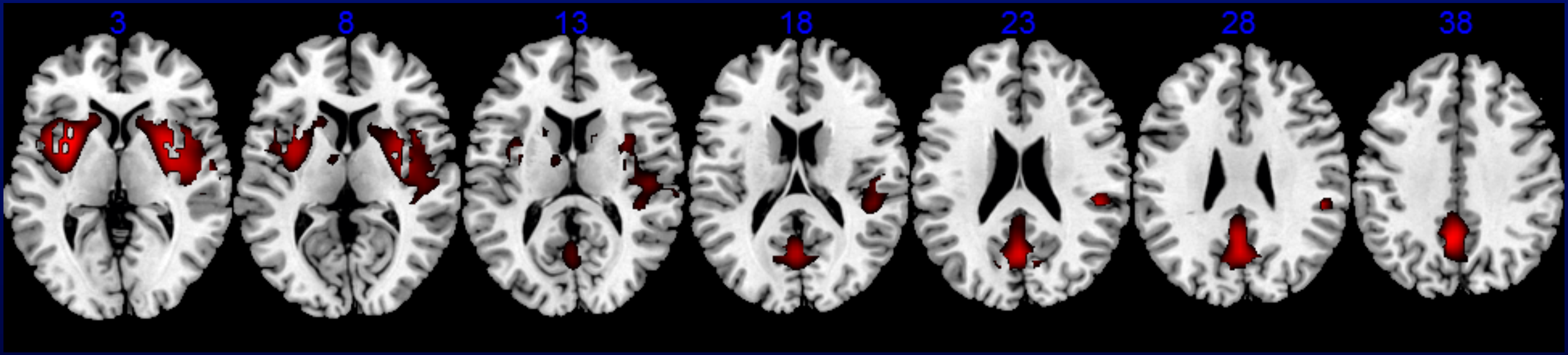
sMRI/SNP

Structural deficits in brain regions consistently implicated in previous schizophrenia reports, including frontal and temporal lobes and thalamus were related to SNPs from 16 genes, several previously associated with schizophrenia risk and/or involved in normal CNS development, including AKT, PI3k, SLC6A4, DRD2, CHRM2 and ADORA2A.

A. sMRI component –A (group difference)



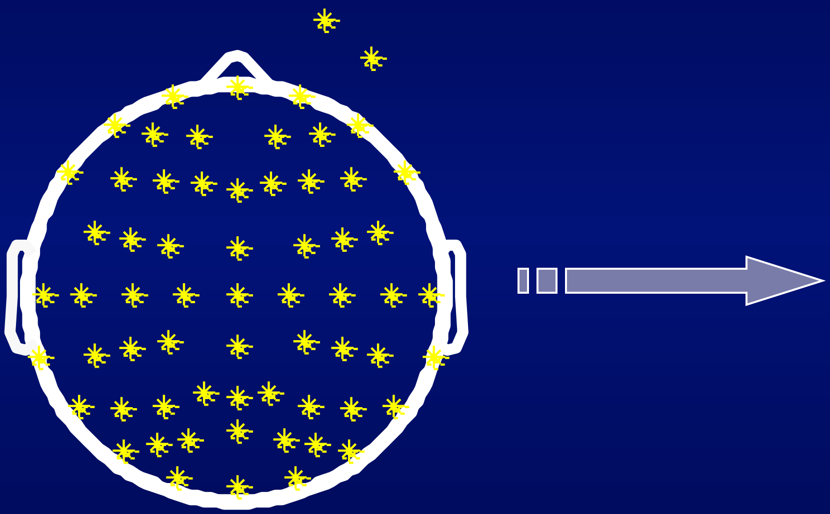
B. sMRI component –B (linked, but no group difference)

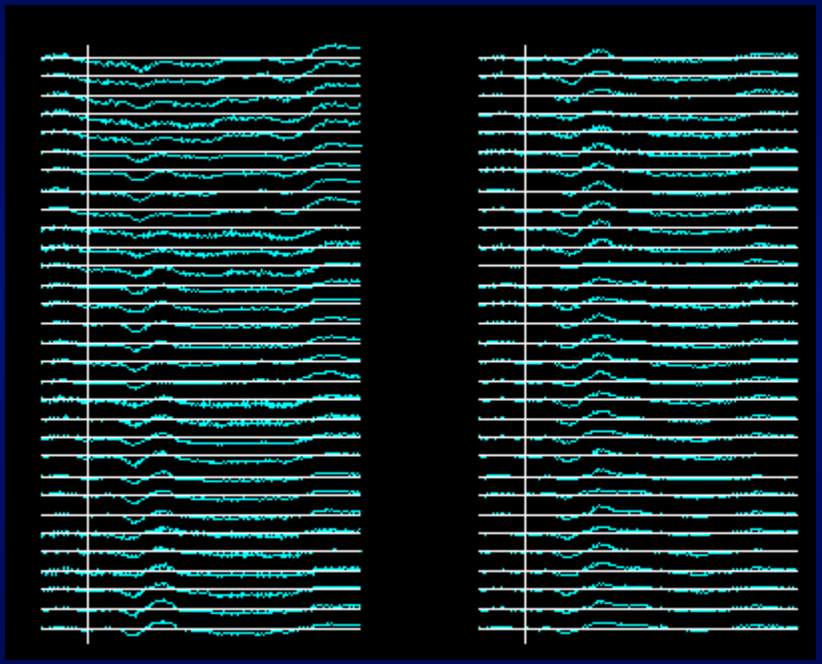


K. Jagannathan, V. D. Calhoun, J. Gelernter, M. Stevens, J. Liu, F. Bolognani, A. Windemuth, G. Ruano, and G. D. Pearlson, "Genetic associations of brain structural networks in schizophrenia: a preliminary study using parallel ICA," *Biological Psychiatry,* vol. 68, pp. 657-666, 2010

Genetics and P3 ERP generation

- Subjects: 41 healthy subjects(24 female, 17 male)
- EEG collected during AOD task,
  target/novel ERPs extracted

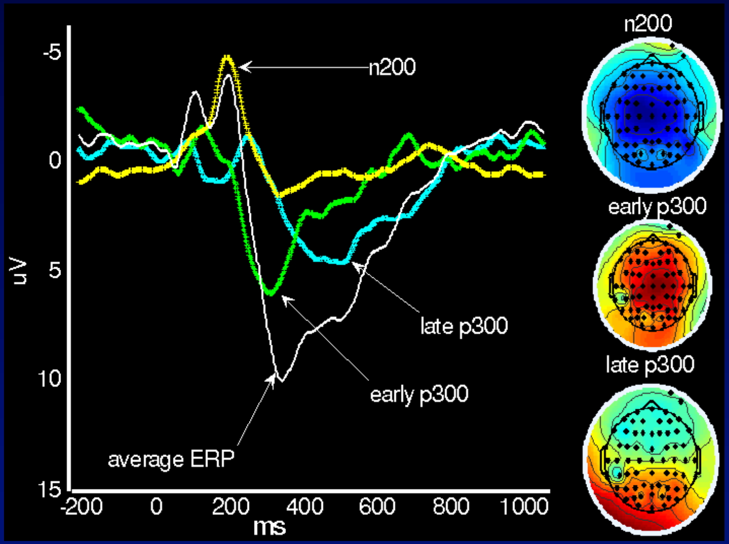




- Blood sample collected, genotyped 384 SNPs from 222 genes 6 physiological systems.

ERP Topography & SNP Associations

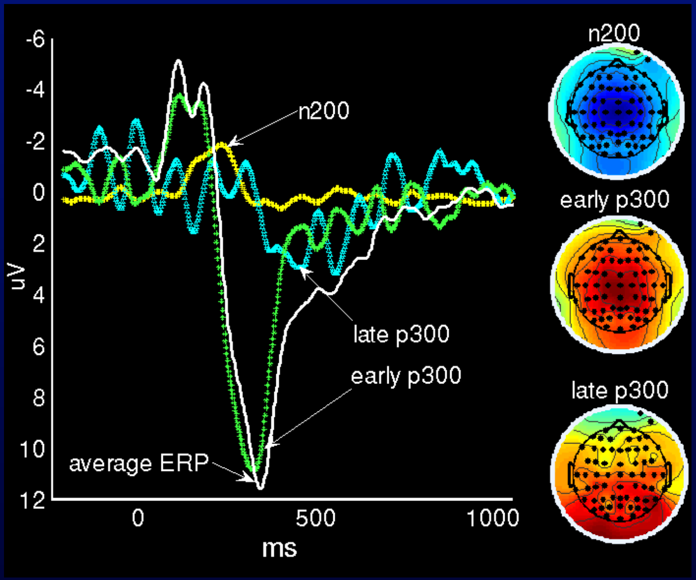
Target Stimuli



0.55

SNPs	Genes
rs1800545	ADRA2A
rs7412	APOE
rs1128503	ABCB1
rs6578993	TH
rs1045642	ABCB1
rs2278718	MDH1
rs4784642	GNAO1
rs521674	ADRA2A

Novel Stimuli

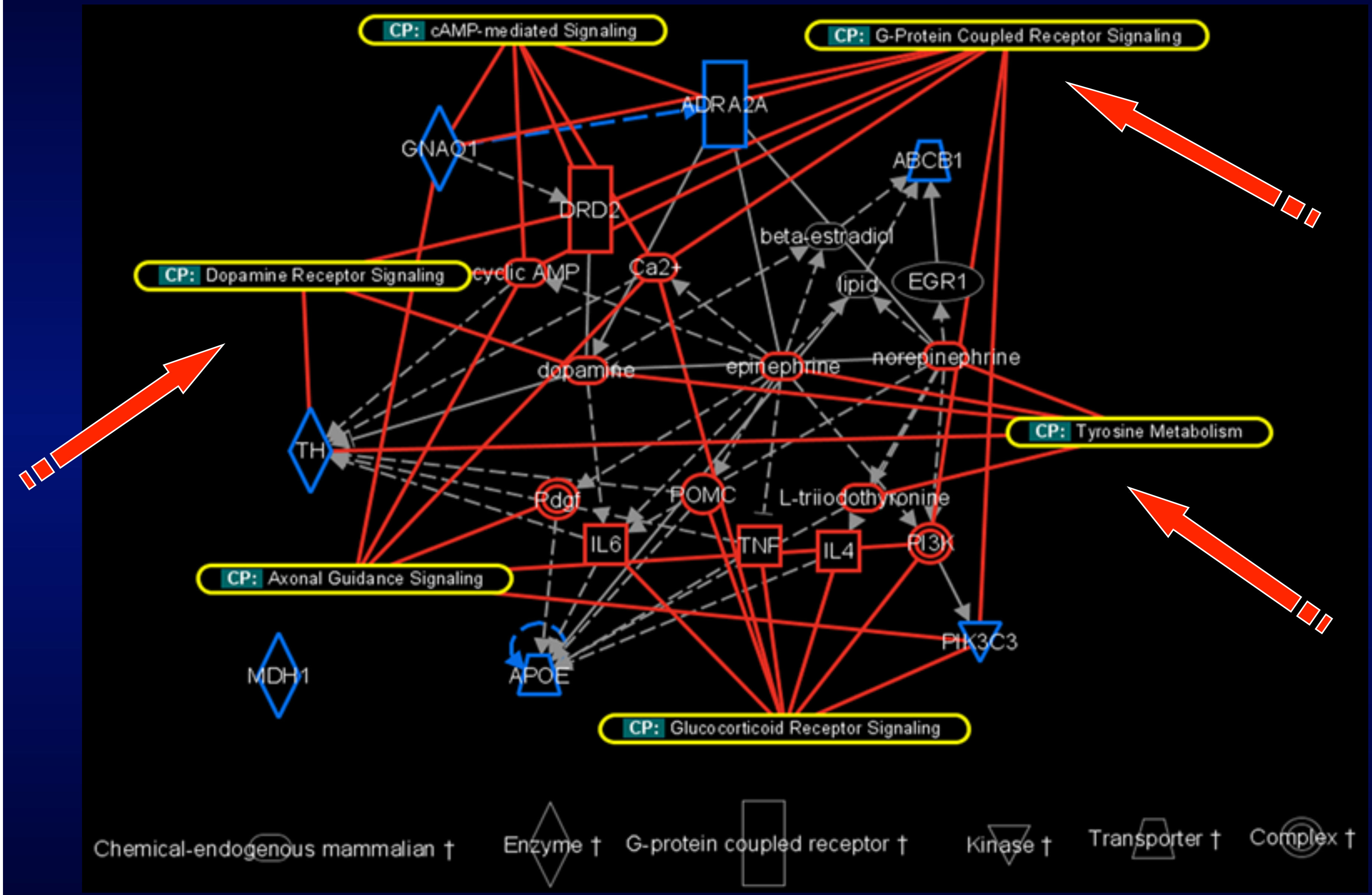


0.47

SNPs	Genes
rs1800545	ADRA2A
rs7412	APOE
rs6578993	TH
rs2278718	MDH1
rs1128503	ABCB1
rs429358	APOE
rs3813065	PIK3C3
rs4121817	PIK3C3
rs521674	ADRA2A

J. Liu, K. A. Kiehl, G. D. Pearlson, N. I. Perrone-Bizzozero, and V. D. Calhoun, "Genetic Determinants of Target and Novelty Processing," *NeuroImage,* vol. 46, pp. 809-816, 2009.

Pathway Analysis



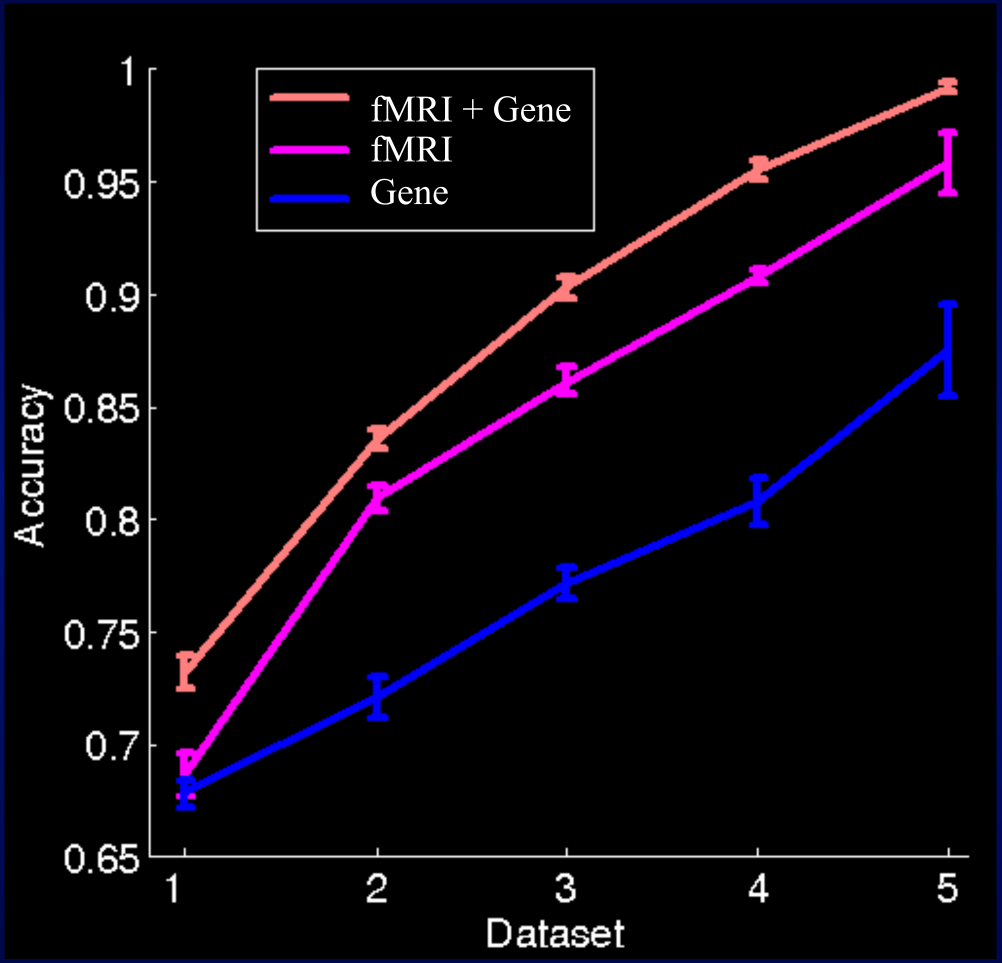
J. Liu, K. A. Kiehl, G. D. Pearlson, N. I. Perrone-Bizzozero, and V. D. Calhoun, "Genetic Determinants of Target and Novelty Processing," *NeuroImage,* vol. 46, pp. 809-816, 2009.

Classification with SNP & fMRI

Subjects	20 Patients 20 Healthy Controls
fMRI data	Auditory Oddball Task
Gene data	367 SNPs

Dataset	Number of Training Subjects	Number of Testing Subjects
1	4	36
2	12	28
3	20	20
4	28	12
5	36	4

SVM Classification



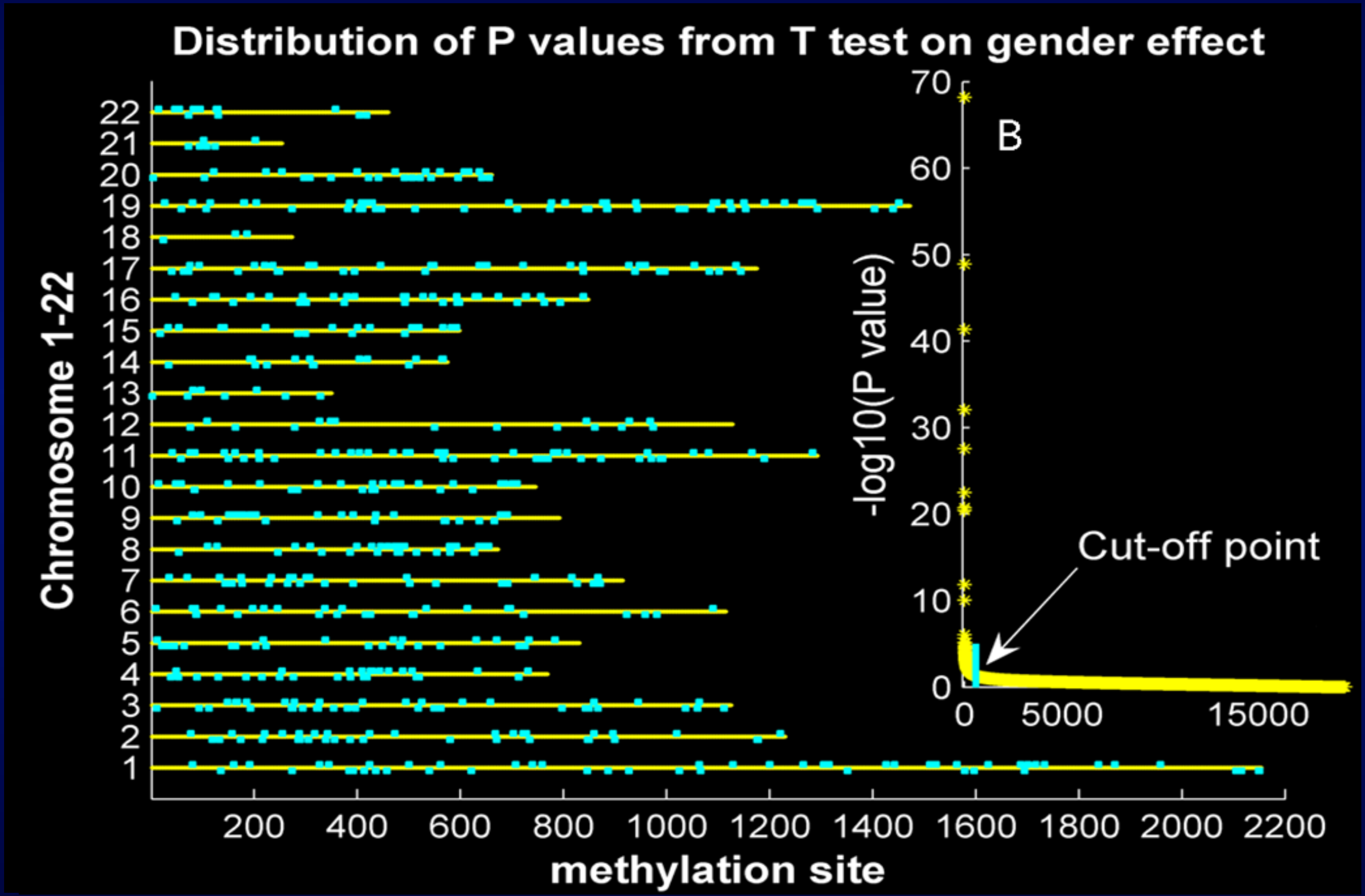
H. Yang, J. Liu, J. Sui, G. Pearlson, and V. D. Calhoun, "A Hybrid Machine Learning Method for Fusing fMRI and Genetic Data to Classify Schizophrenia," *Frontiers in Human Neuroscience,* vol. 4, pp. 1-9, 2010

Overview

- Motivation
- Joint ICA
  - ERP/FMRI
  - Multitask-fMRI
  - sMRI/fMRI
  - CC-ICA/Biomarker Identification
- mCCA
  - sMRI/fMRI/ERP
- mCCA+jICA
  - DTI+fMRI
- Parallel ICA
  - SNPs
  - Epigenetics
- Conclusions

Methylation Sex Correction

Sex difference:



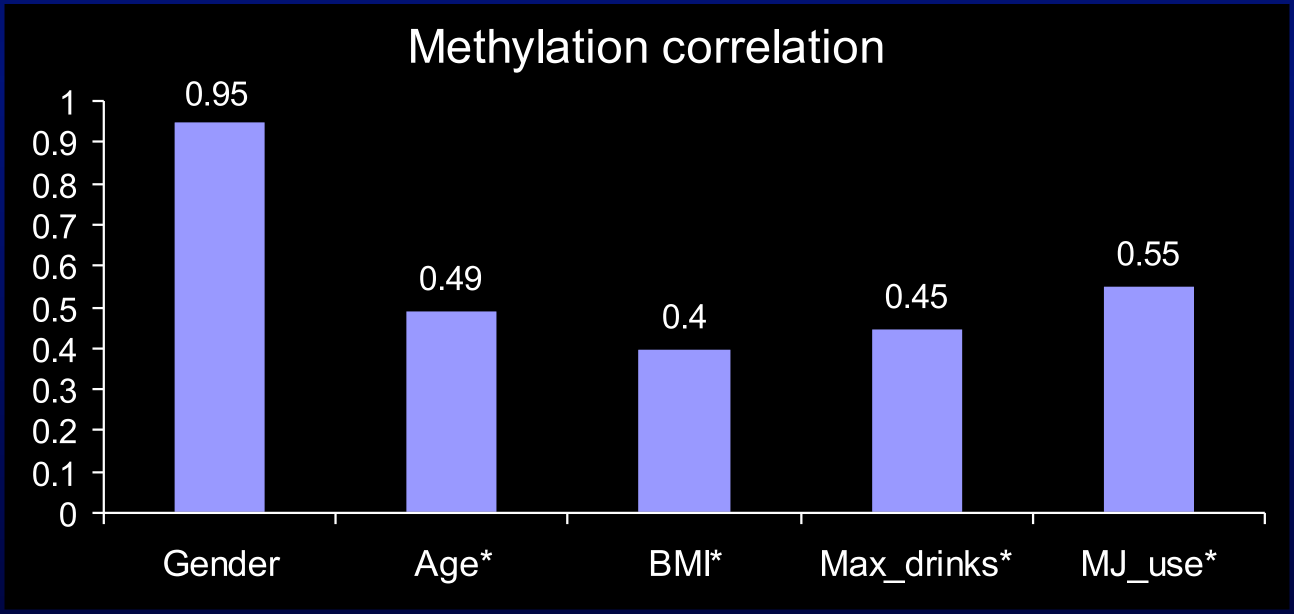
J. Liu, M. Morgon, K. Hutchison, E. Claus, and V. D. Calhoun, "A Study of the Influence of Sex on Genome Wide Methylation," PLoS ONE, vol. 5, pp. 1-8, 2010.

Methylation Sex Correction

- Genomic ~27,000 sites from 23 pairs of chromosomes
- 130 subjects (heavy drinker 33 females, 97 males, age 31.3±9.7)

Goal: Association with gender, age, BMI, alcohol use, cigarette use, marijuana use, depression, stress, etc.

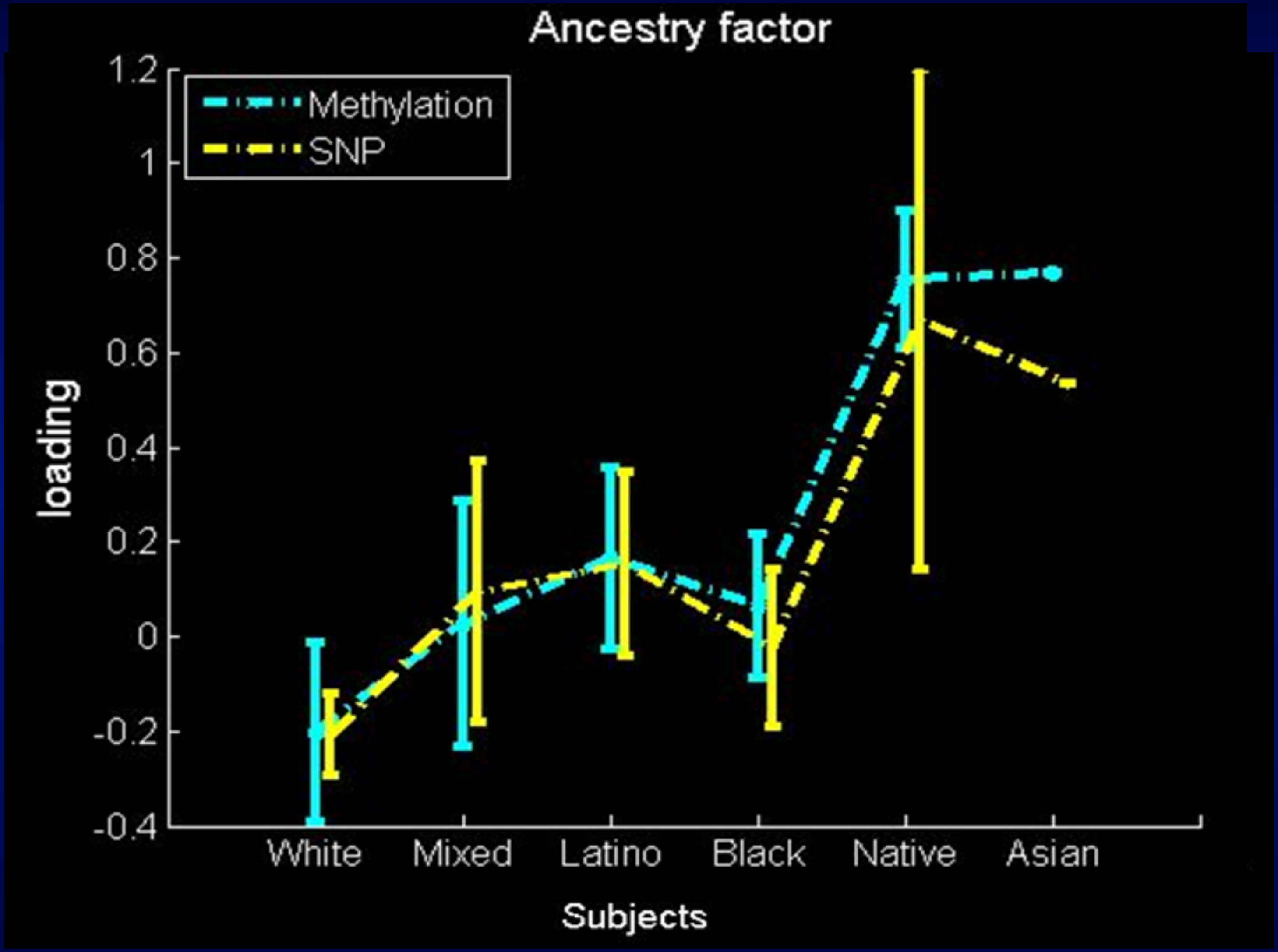
Results:



*: after sex effect correction

J. Liu, M. Morgon, K. Hutchison, E. Claus, and V. D. Calhoun, "A Study of the Influence of Sex on Genome Wide Methylation," PLoS ONE, vol. 5, pp. 1-8, 2010.

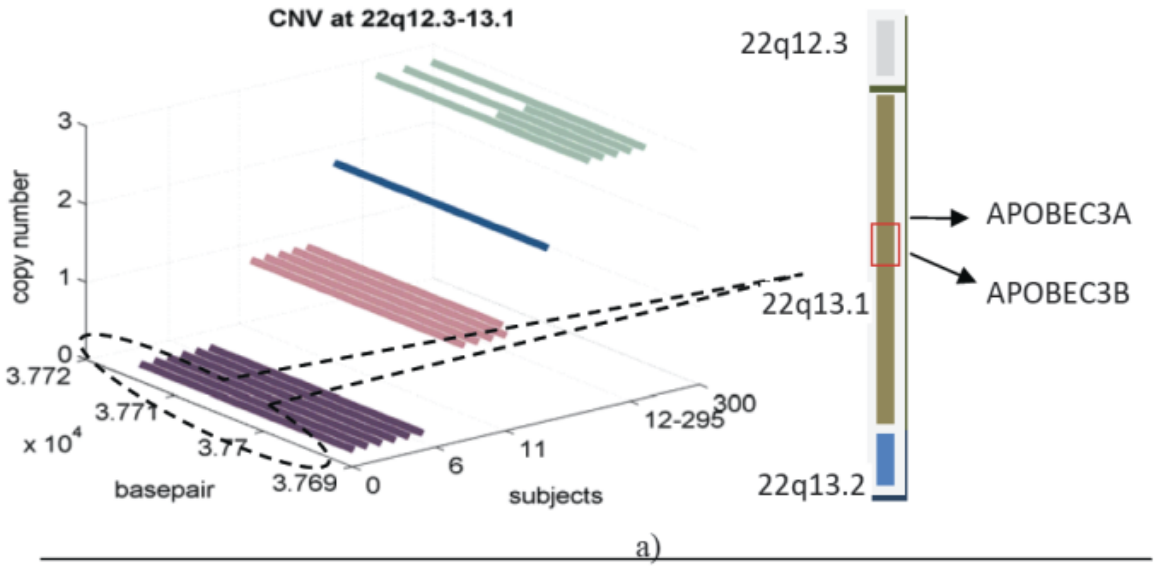
Ancestry effect on methylation

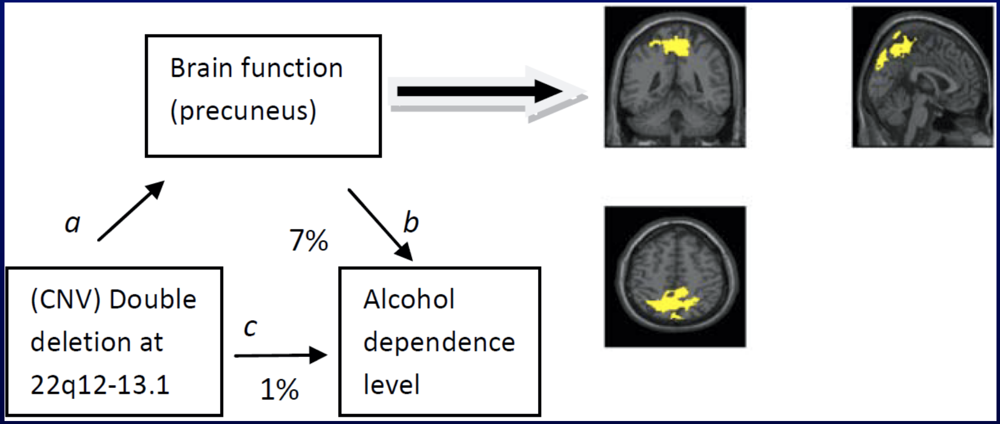


J. Liu, K. Hutchison, M. Morgan, N. I. Perrone-Bizzozero, J. Sui, and V. D. Calhoun, "Identification of Genetic and Epigenetic Factors Contributing to Population Structure," *PLoS ONE,* vol. 5, pp. 1-8, 2010.

CNV and substance abuse

- Effect of double deletion at 22q12.3-13.1 on alcohol use disorder severity and cue-elicited BOLD response in the precuneus





Brain function (precuneus)

a

b

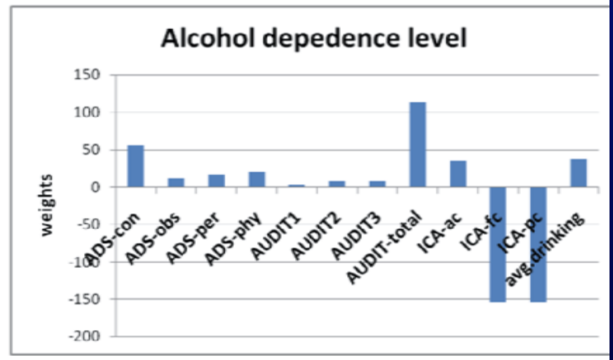
7%

(CNV) Double deletion at 22q12-13.1

c

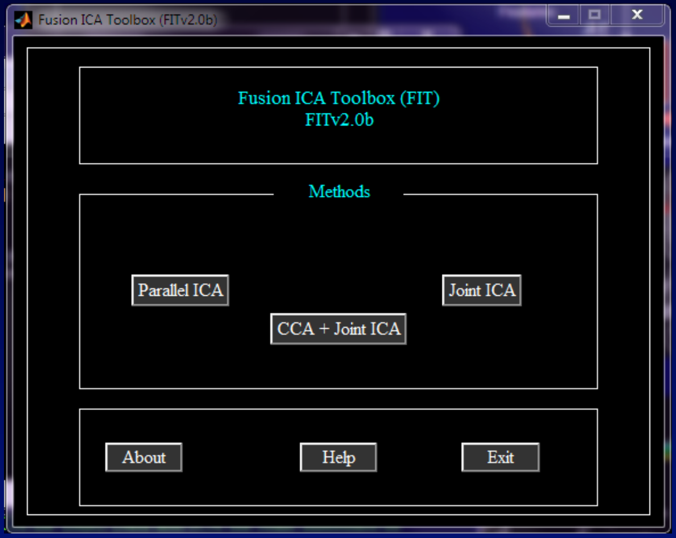
1%

Alcohol dependence level



R. A. Yeo et al PLoS ONE, vol. 6, 2011.
R. A. Yeo et al, Biological Psychiatry, In Press.
J. Liu et al. Plos One, Under Review.

Fusion ICA Toolbox (FIT)



Implements Joint ICA, CC-ICA, Parallel ICA

-fusion of EEG, fMRI, sMRI, DTI, Genetics

-over 900 unique downloads

http://icatb.sourceforge.net

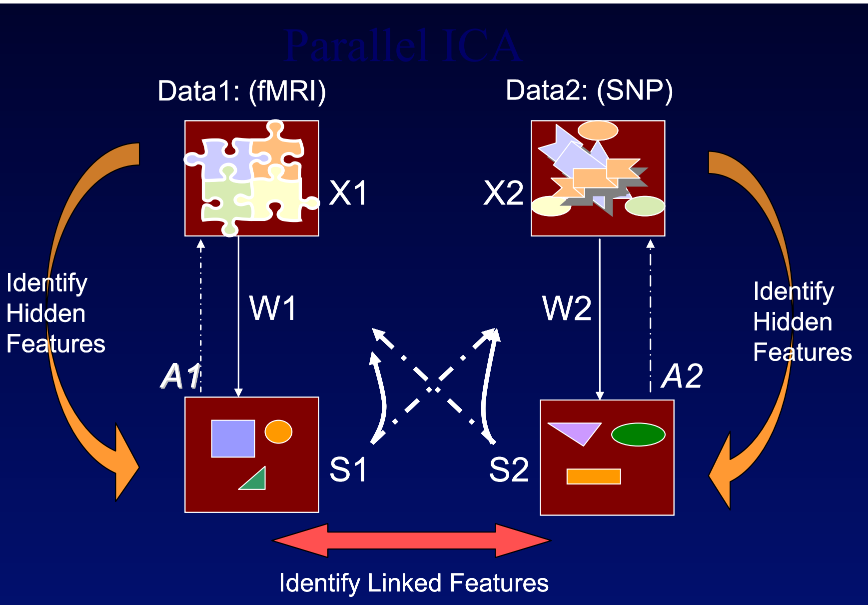
-funded by NIBIB R01EB005846

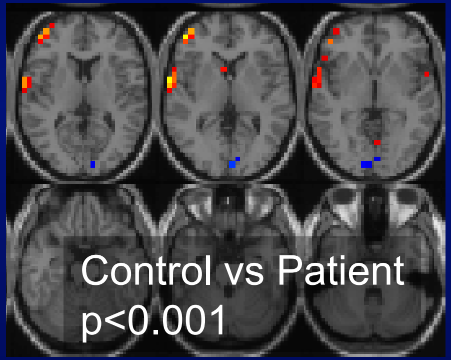
-biomarker identification through optimized ranking of imaging features

-multiple algorithms

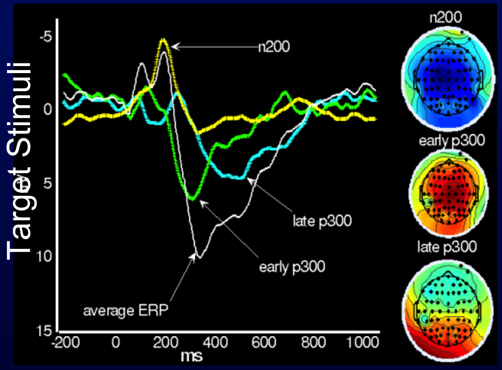
-leave-one-out stability evaluation

V. D. Calhoun, J. Liu, and T. Adali, "A Review of Group ICA for fMRI Data and ICA for Joint Inference of Imaging, Genetic, and ERP data," *NeuroImage,* vol. 45, pp. 163-172, 2009.





SNP	*Z score*	*Gene*
Rs1466163	-4.08	AADC
Rs2429511	3.97	ADRA2A
Rs3087454	**-3.09**	**CHRNA7**
Rs821616	**2.96**	**DISC1**
Rs885834	**-2.78**	**CHAT**
Rs1355920	**-2.77**	**CHRNA7**
R4765623	2.73	SCARB1
Rs4784642	-2.71	GNAO1
Rs2071521	2.58	APOC3
Rs7520974	2.55	CHRM3



0.55

SNPs	Genes
rs1800545	ADRA2A
rs7412	APOE
rs1128503	ABCB1
rs6578993	TH
rs1045642	ABCB1
rs2278718	MDH1
rs4784642	GNAO1
rs521674	ADRA2A

MIALAB

MEDICAL IMAGE ANALYSIS LAB

http://mialab.mrn.org