

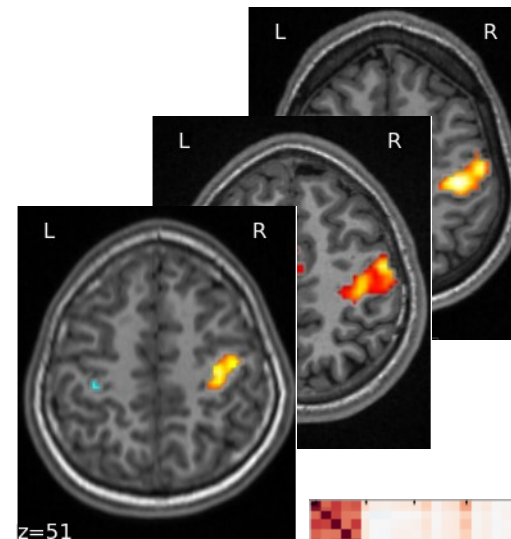
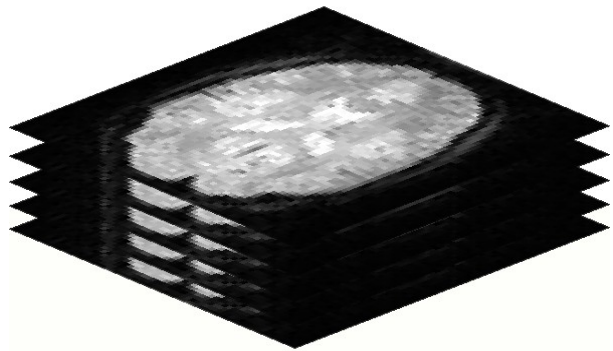
Phénotype, fonction et génotype

Bertrand Thirion, bertrand.thirion@inria.fr

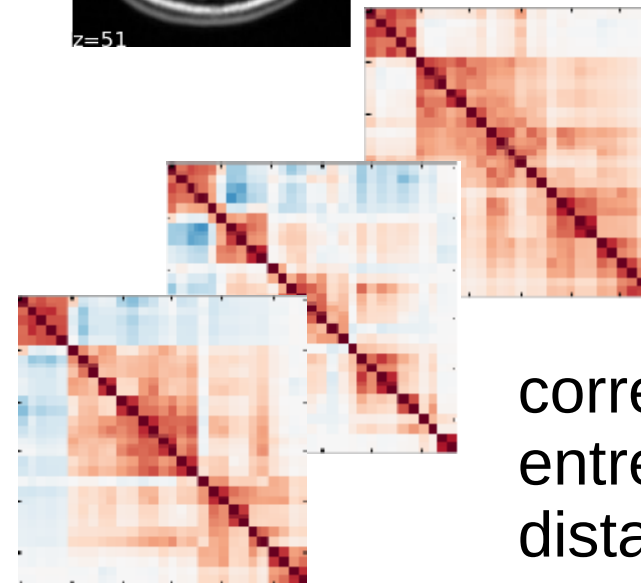
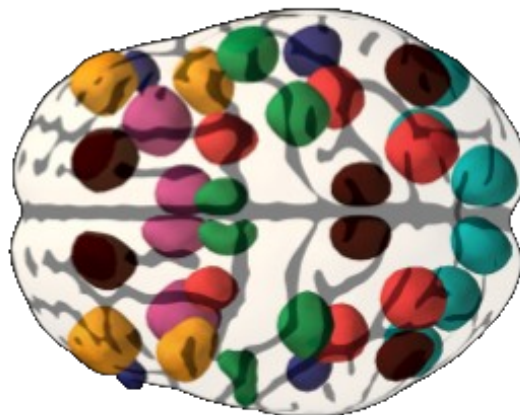


Une approche quantitative du fonctionnement cérébral

Images fonctionnelles

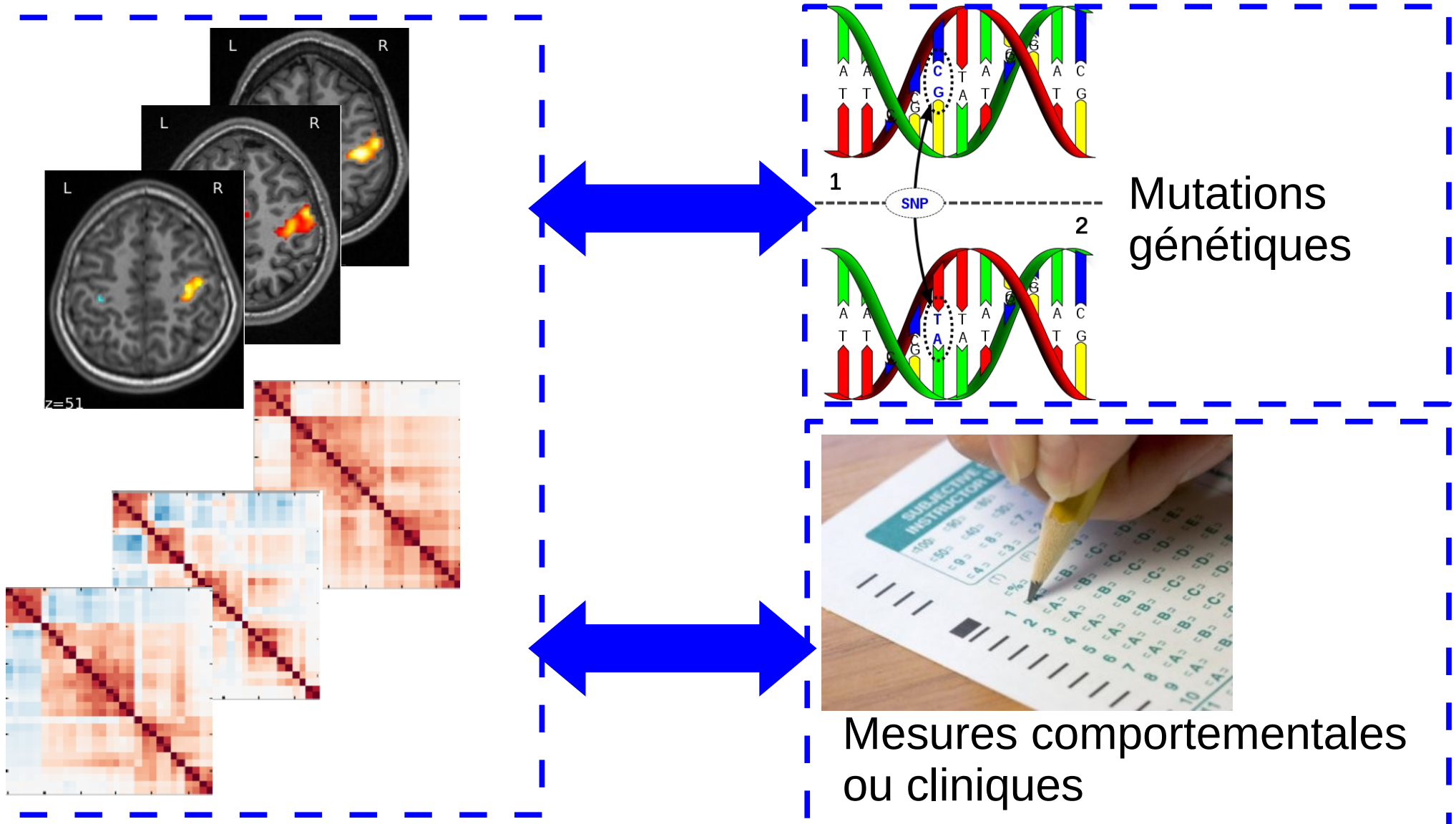


Amplitude
des réponses
évoquées



corrélations
entre régions
distantes

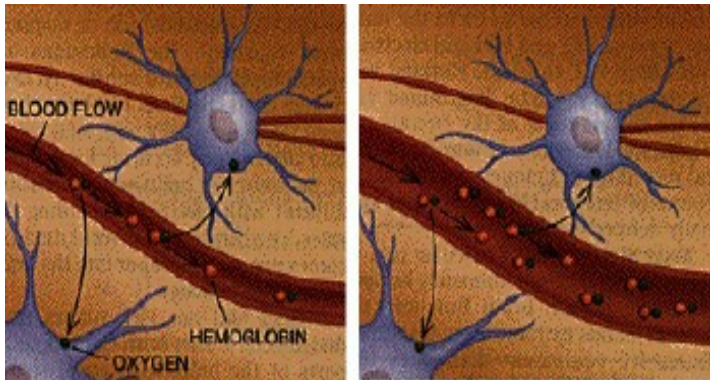
Une approche quantitative du fonctionnement cérébral



Plan

- **Biomarqueurs d'imagerie fonctionnelle**
- Le défi génétique - neuroimagerie fonctionnelle
- Améliorer la modélisation statistique pour l'analyse de groupe

Imagerie BOLD chez l'homme



Paradigme Experimental



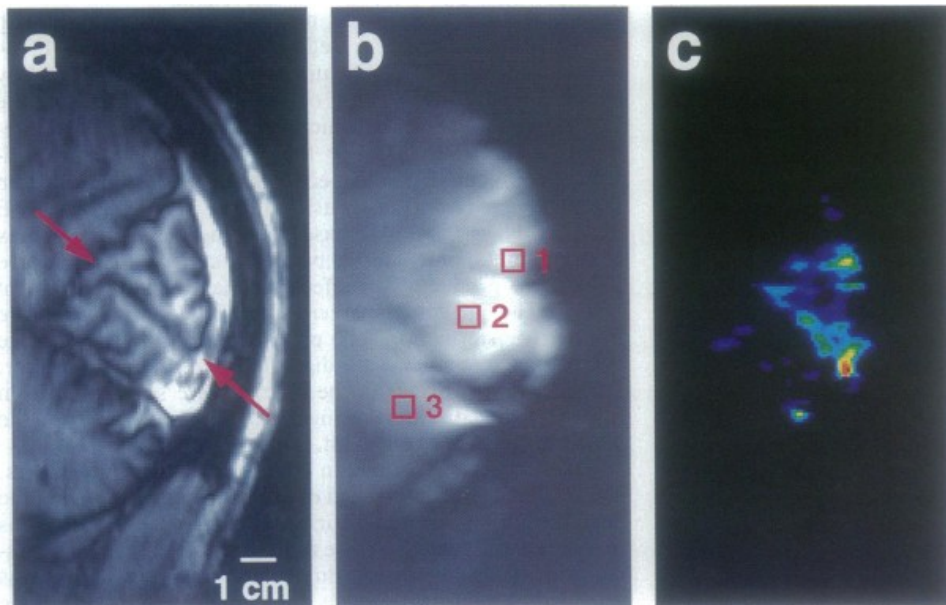
Activité neuronale



Massif Afflux local de sang oxygéné (5s)

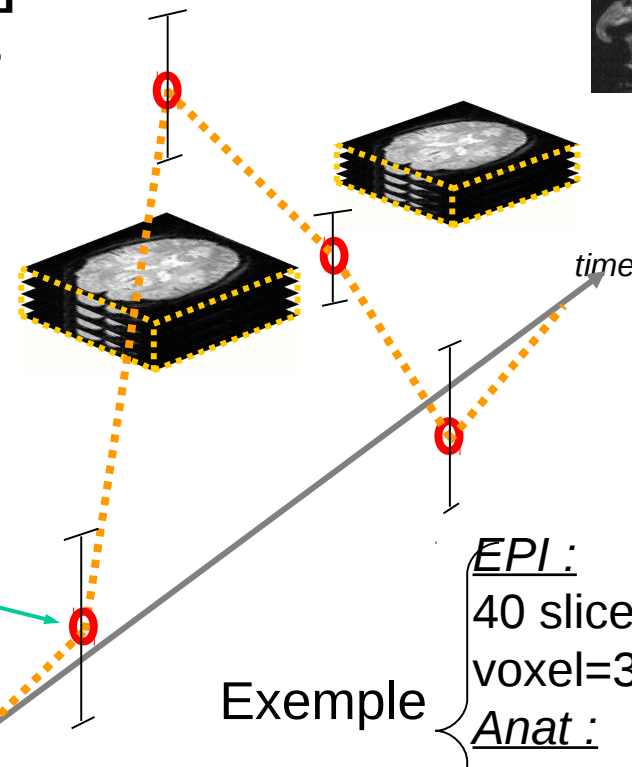
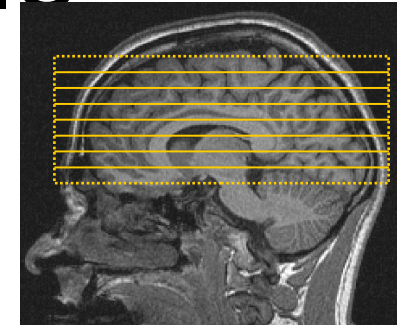
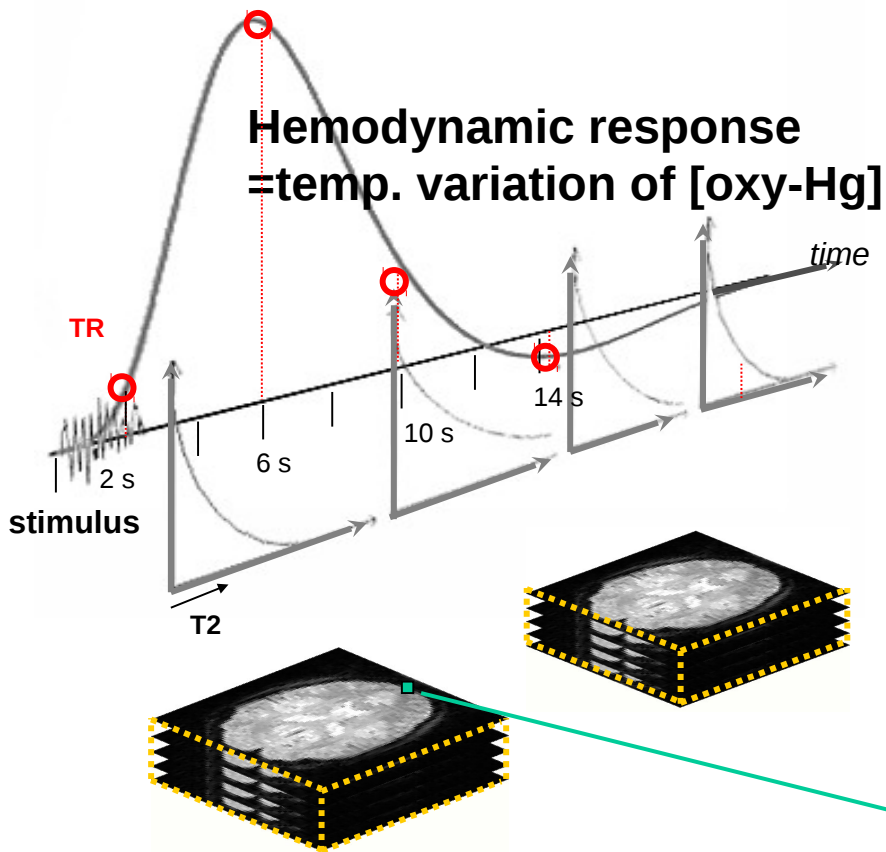


Augmentation du signal IRM



[Ogawa et al, PNAS, 1992]

Recueil de la réponse BOLD par l'IRM fonctionnelle

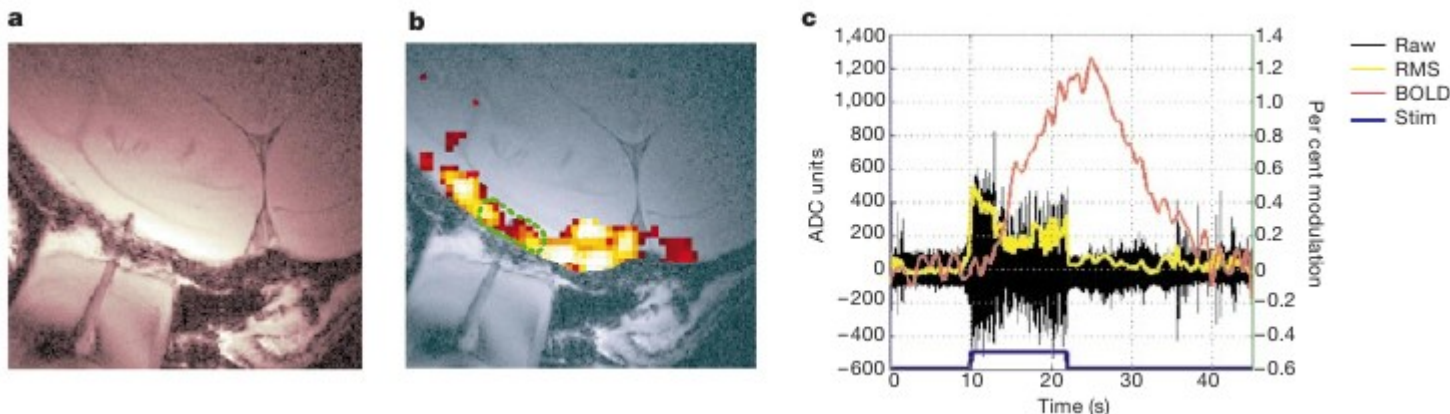


EPI :
 40 slices, TR=2000 ms, TE=40ms,
 voxel=3x3x3mm, 64x64 matrix

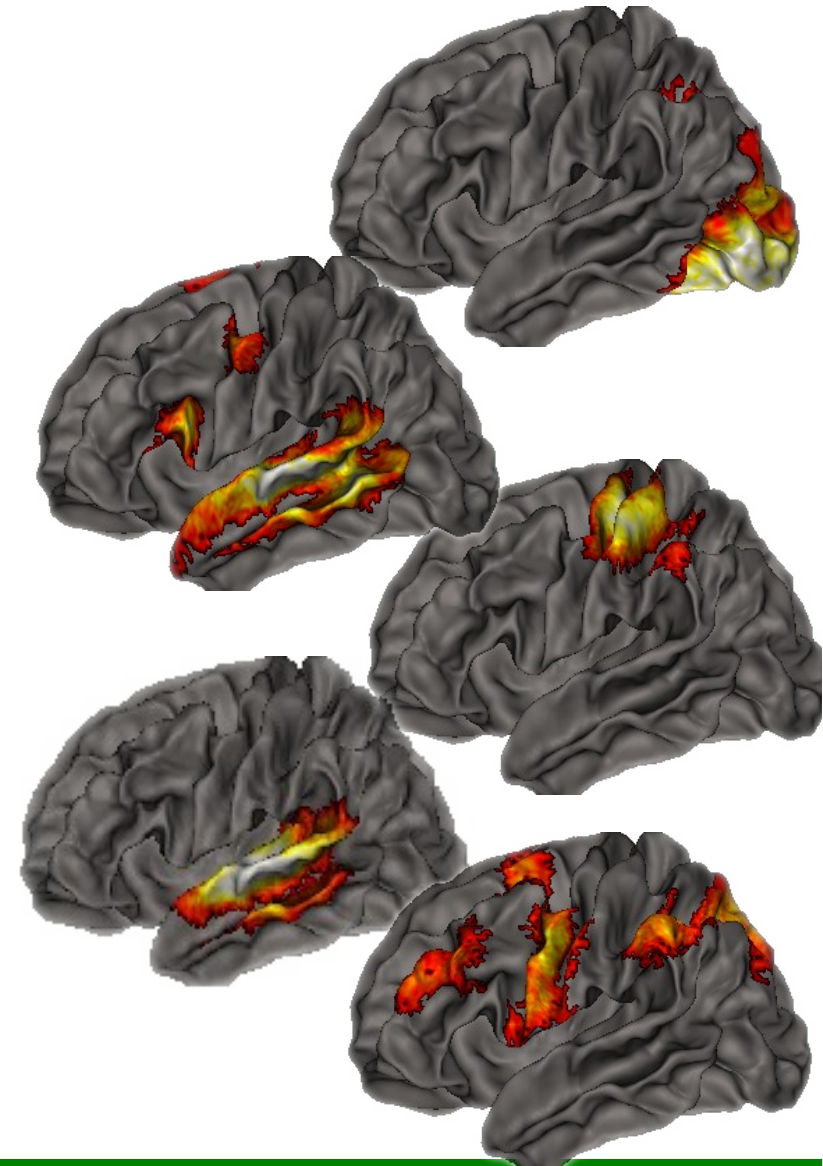
Anat :
 TR=1600ms, FOV=256x256mm²,
 256X256 matrix, 192 slices
 Slice thickness 1mm

20 ans d'imagerie BOLD humaine

- Réponse BOLD : approximativement linéaire / fonction de stimulation
 - Application de simples modèles linéaire pour l'analyse de données [Friston et al. 1995]
- Signal BOLD fortement corrélé avec les potentiels de champ locaux [Logothetis et al. Nature 2001]
- Bonne résolution spatiale (~2mm) [Ugurbil et al. NeuroImage 2007]
- Faible résolution temporelle, pas de modèle complet du signal

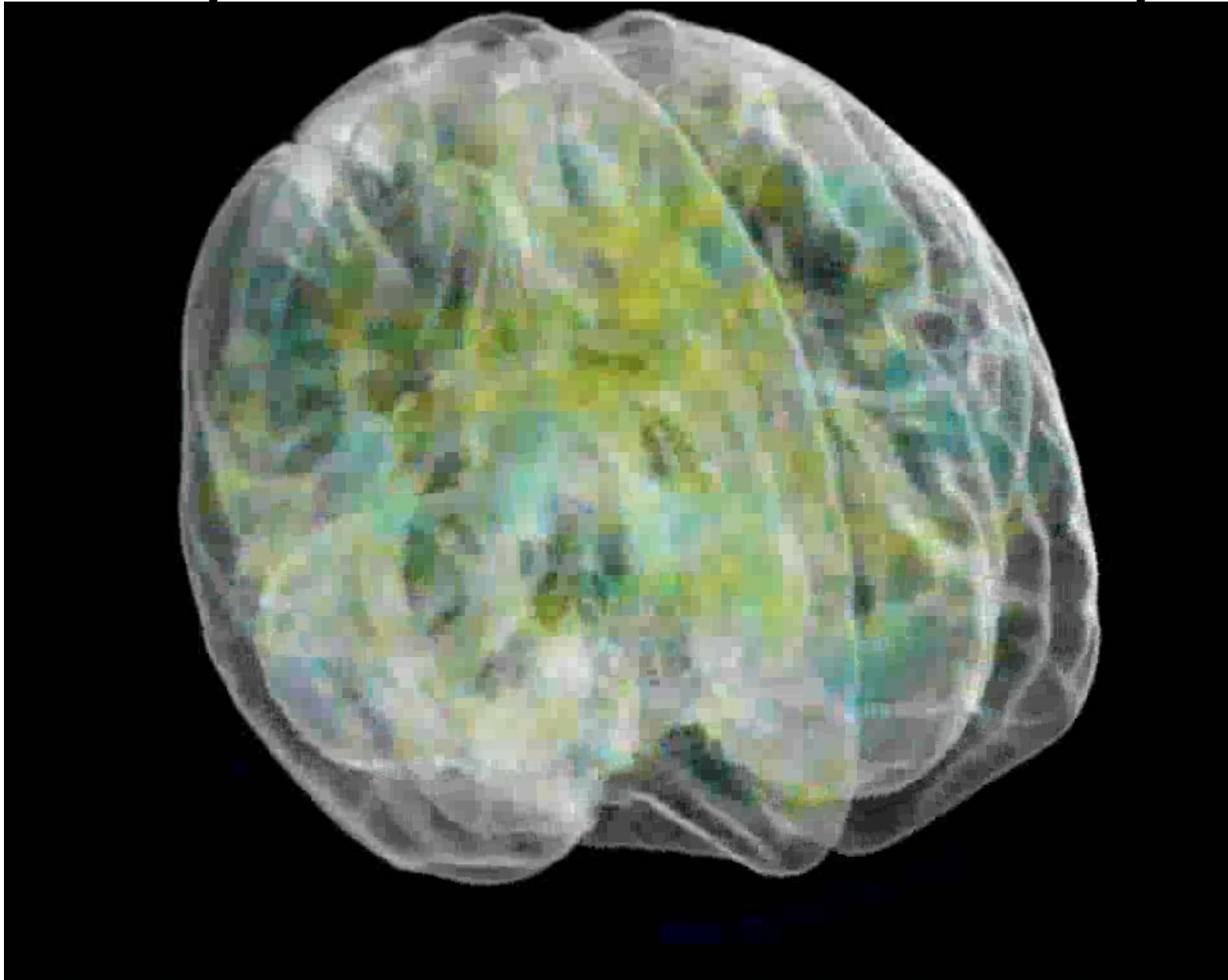


Imagerie BOLD chez l'homme



- Utilisation intensive pour cartographie le fonctionnement cérébral
 - Principe de ségrégation

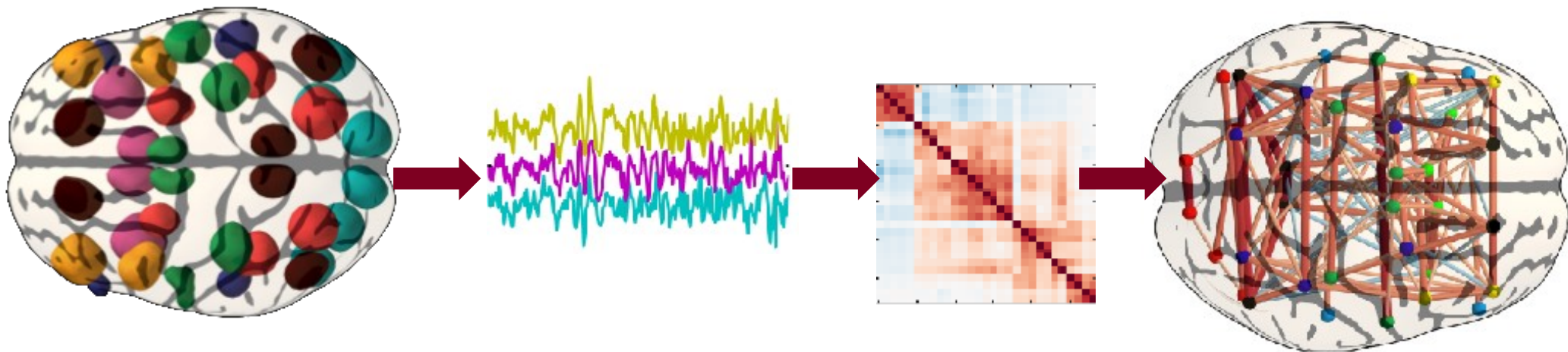
Même quand le cerveau 'se repose'



La covariance des signaux IRMf: interactions régionales

Problème:

- Signaux de p régions
- Connectivité **fonctionnelle** = dépendance statistique des signaux entre régions

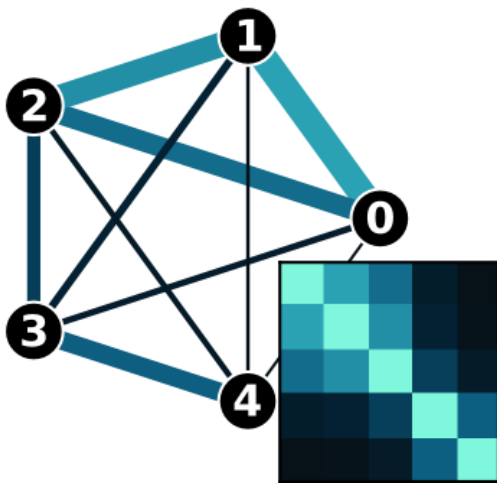


La covariance des signaux IRMf: modèle

$$\hat{\mathbf{K}} = \operatorname{argmin}_{\mathbf{K} \succ 0} \operatorname{tr}(\hat{\Sigma} \mathbf{K}) - \log \det(\mathbf{K})$$

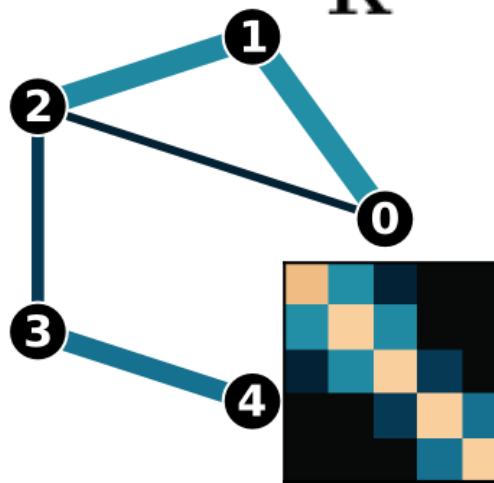
Corrélations observées

$\hat{\Sigma}$



Interactions sous-jacentes
(=corrélations partielles)

\mathbf{K}



[Fransson et al. 2008]
Structure des réseaux cérébraux

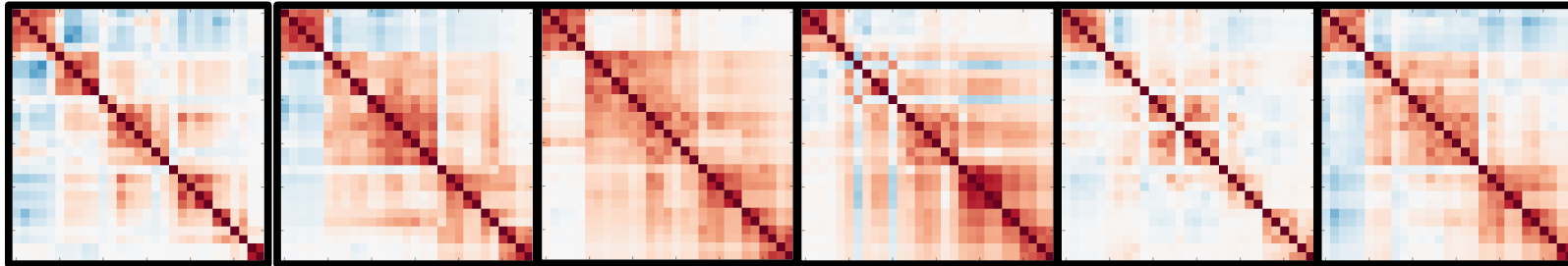
La covariance des signaux IRMf: parcimonie

- La **parcimonie** est une hypothèse naturelle
 - Peu de paires de régions sont effectivement connectées
 - Améliorer l'estimation de covariance: **sélection de covariance** [Dempster et al. 1972]
 - Approche moderne: régularisation L1 des coefficients non-diagonaux → procédure d'estimation convexe [Banejee et al. ICML 2006]

$$\hat{\mathbf{K}}_{\ell_1} = \operatorname{argmin}_{\mathbf{K} \succ 0} \operatorname{tr}(\mathbf{K} \hat{\Sigma}_{\text{sample}}) - \log \det \mathbf{K} + \lambda \|\mathbf{K}\|_1$$

La covariance des signaux IRMf: parcimonie de groupe

- On s'intéresse à des données **multi-sujet**



Exemple: corrélations chez 6 sujets, 30 régions

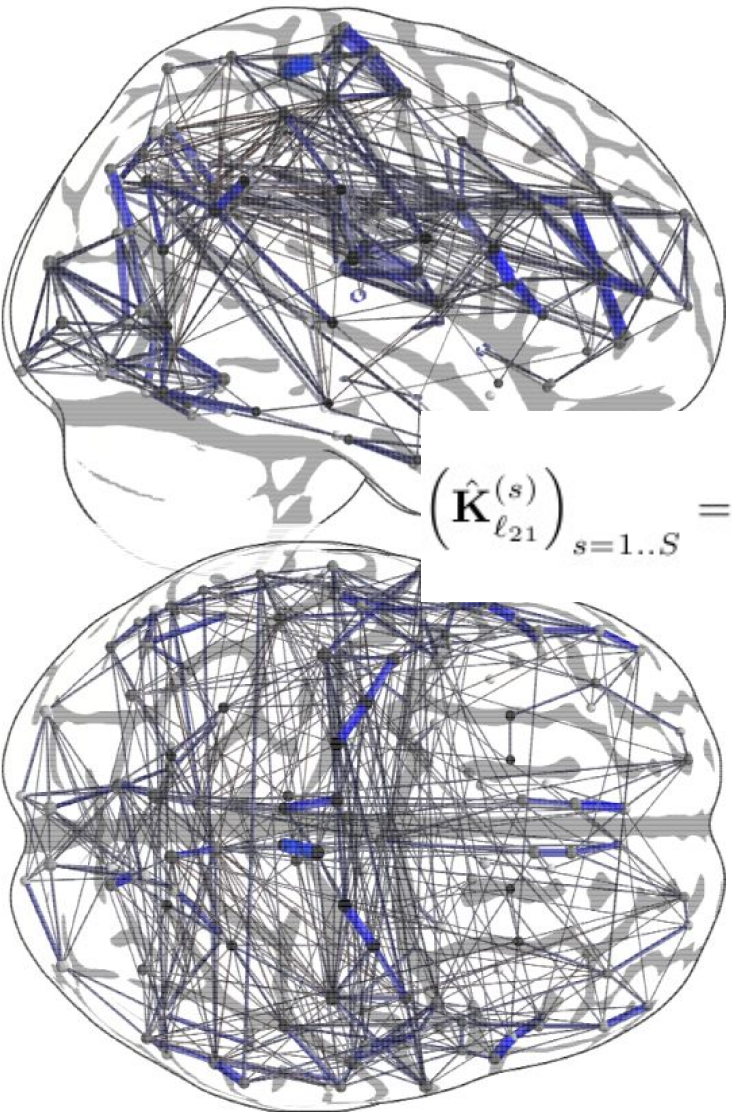
Motifs parcimonieux cohérents au sein d'un groupe imposés par des pénalités structurées, e.g. group graph lasso [Honorio et al. ICML 2010] [Varoquaux et al. NIPS 2010]

$$\left(\hat{\mathbf{K}}_{\ell_{21}}^{(s)} \right)_{s=1..S} = \operatorname{argmin}_{\mathbf{K}^{(s)} \succ 0} \left(\sum_{s=1}^S \left(\operatorname{tr}(\mathbf{K}^{(s)} \hat{\Sigma}_{\text{sample}}^{(s)}) - \log \det \mathbf{K}^{(s)} \right) + \lambda \sum_{i \neq j} \|\mathbf{K}_{ij}^{(\cdot)}\|_2 \right)$$

λ choisi par validation croisée

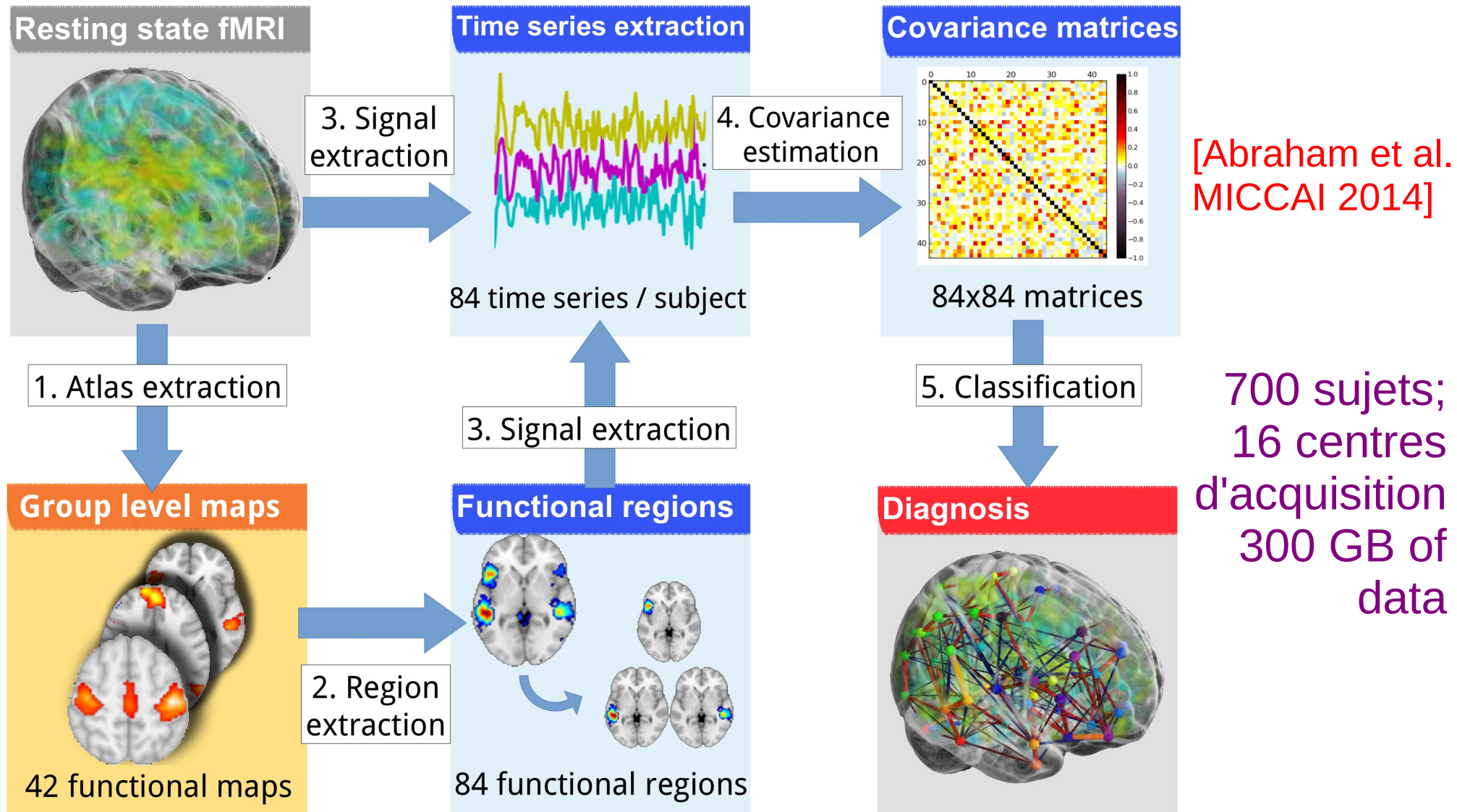
La covariance des signaux IRMf: parcimonie de groupe

Estimation
parcimonieuse
niveau groupe

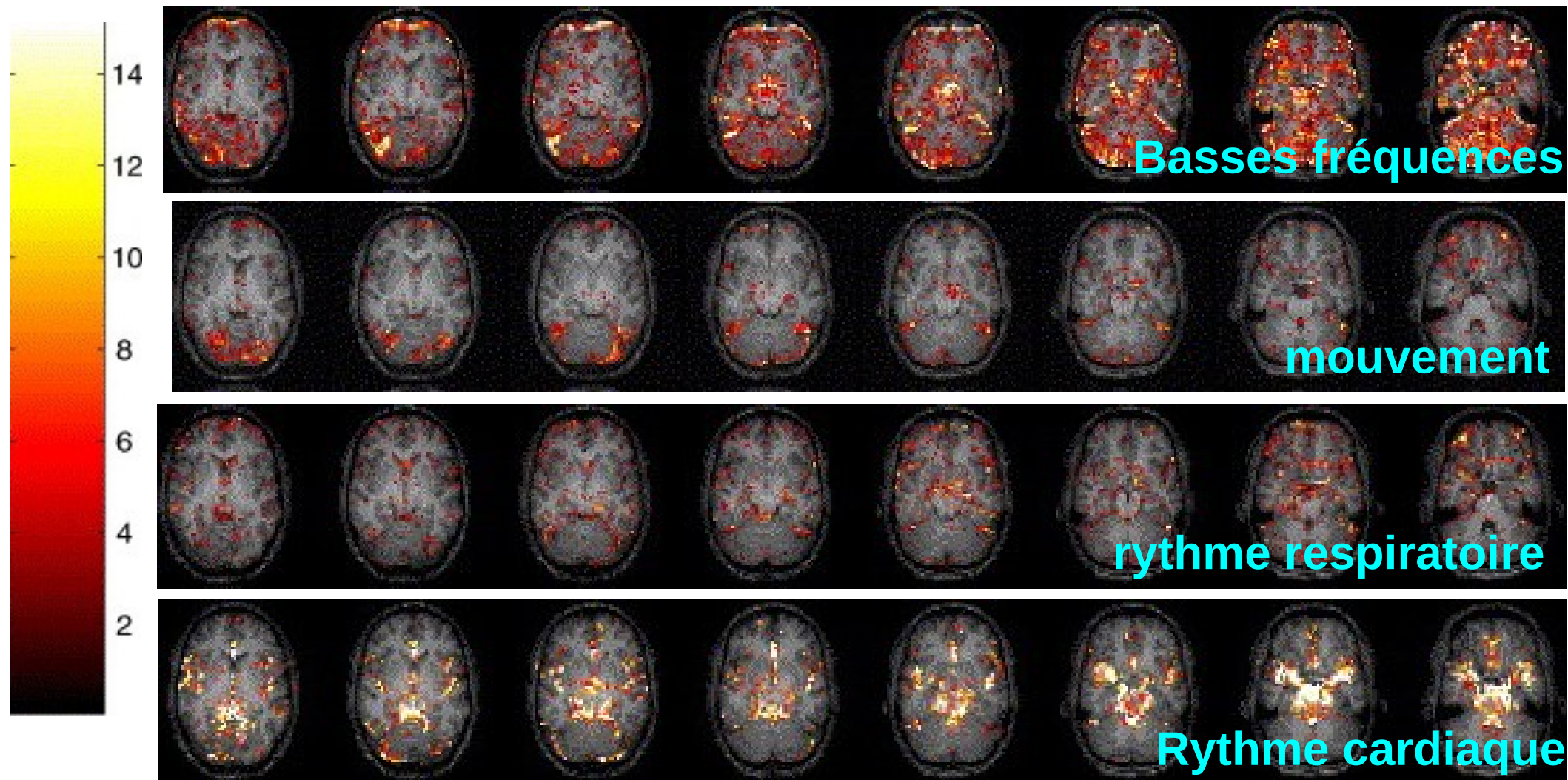

$$\left(\hat{\mathbf{K}}_{l_{21}}^{(s)} \right)_{s=1..S} = \operatorname{argmin}_{\mathbf{K}^{(s)} \succ 0} \left(\sum_{s=1}^S \left(\operatorname{tr}(\mathbf{K}^{(s)} \hat{\Sigma}_{\text{sample}}^{(s)}) - \log \det \mathbf{K}^{(s)} \right) + \lambda \sum_{i \neq j} \|\mathbf{K}_{ij}^{(\cdot)}\|_2 \right)$$

Prend en compte
l'information de tous les
sujets

Traitement des images fonctionnelles

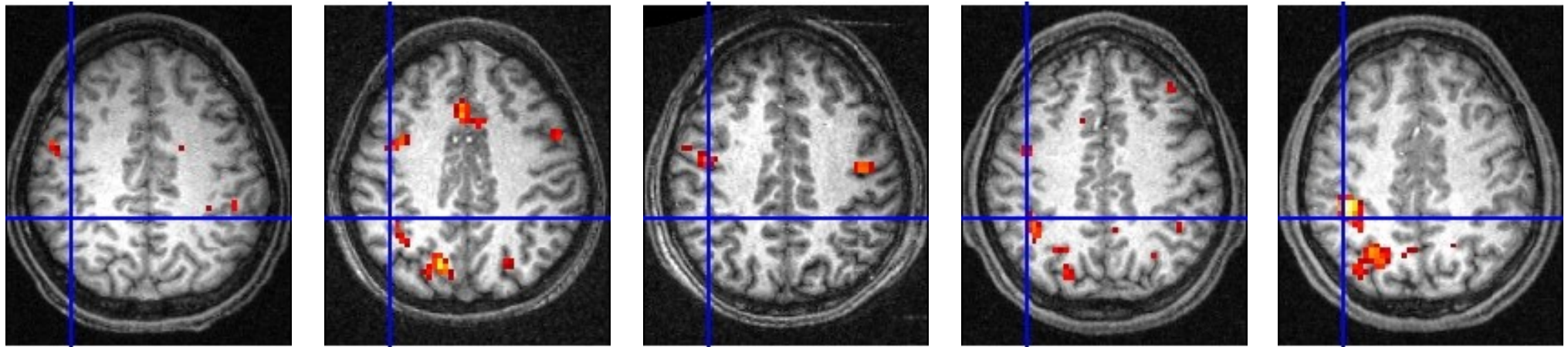


Limitations de l'IRM fonctionnelle



Limitations de l'IRM fonctionnelle

- Variabilité inter-individuelle lors d'une tâche de calcul



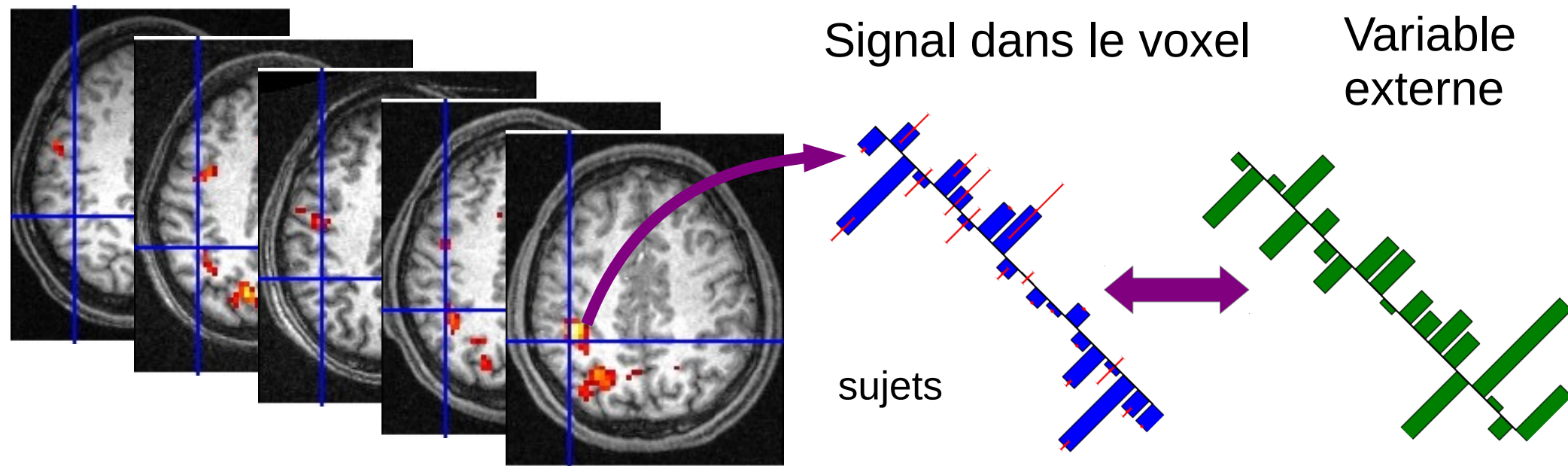
- Localisation imprécise (~10mm de variabilité)
- Quantité d'activation variable
- Organisation des réseaux variable

Études de cohorte en Neuroimagerie

- Passage à plusieurs centaines/milliers d'individus
 - Études multi-centriques
 - Données hétérogènes



[Thyreau et al. 2012]

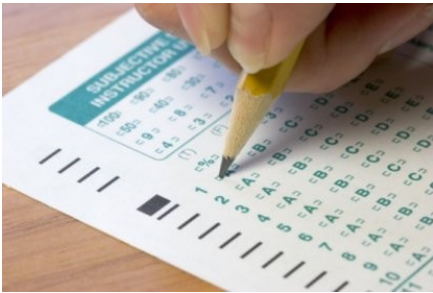


Plan

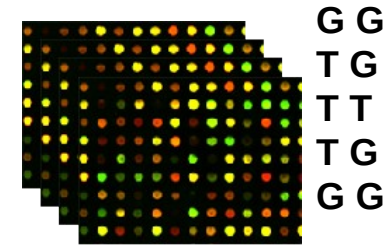
- Biomarqueurs d'imagerie fonctionnelle
- **Le défi génétique - neuroimagerie fonctionnelle**
- Améliorer la modélisation statistique pour l'analyse de groupe

Neuroimagerie-génétique

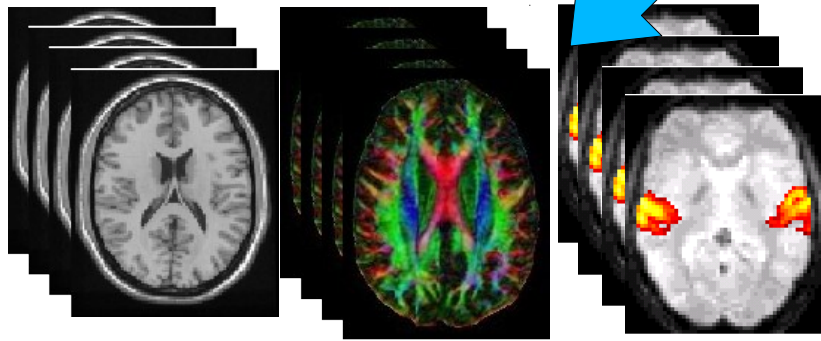
Clinique / Comportement



Information génétique: SNPs



Variabilité et
plasticité cérébrale



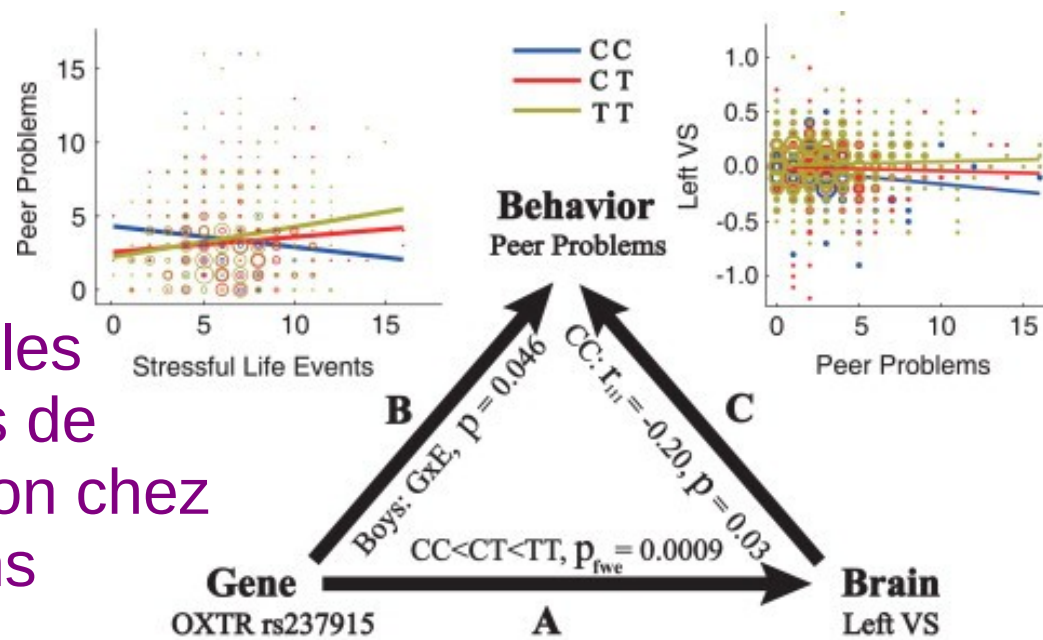
Images cérébrales



Facteurs environnementaux

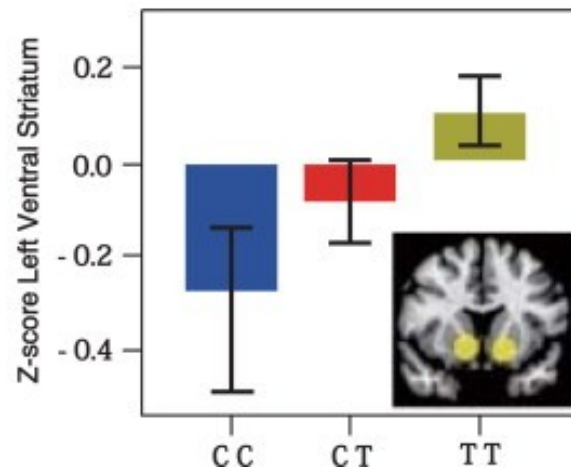
Fonction, génétique et comportement

Effet sur les problèmes de socialisation chez les garçons



Les porteurs de CC → corrélation entre l'activité du SV et les problèmes de socialisation

Effet du SNP rs237915 sur l'activité du striatum ventral

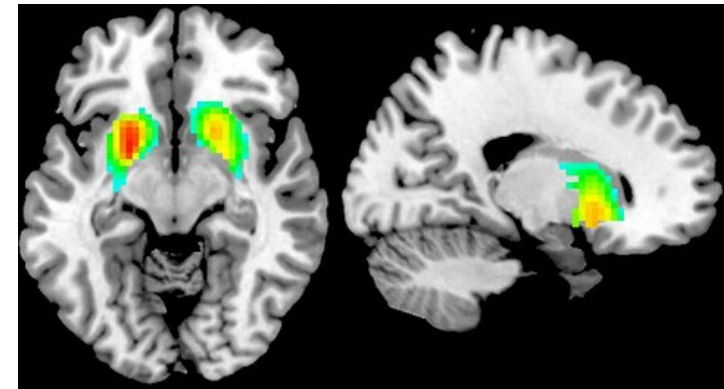


Gène (OXTR) : récepteur de ocytocine
CC = allèle mineur

[Loth et al. Mol. Psy 2013]

Fonction, génétique et comportement

Striatum ventral (SV): motivation



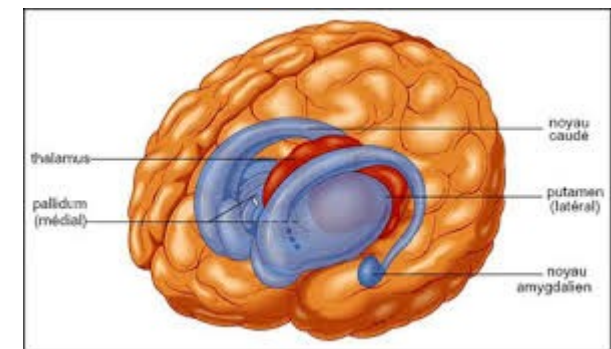
mémoire de travail

Activation **SV** et **caudé** pendant une récompense

Activation **SV** pendant anticipation récompense

rs1800497

Noyau caudé : apprentissage



[Nymberg et al. NN 2014]

Gène DRD2/ANKK1 : récepteur dopamine

Neuroimagerie-génétique: fragilités

- Sélection arbitraire de variables génétique ou neuroimagerie
- Règles strictes pour le contrôle statistique : $p < 10^{-8}$ en génétique, $p < 10^{-12}$ en neuroimagerie génétique !
- Reproduction sur une deuxième cohorte
- **Manque de puissance statistique**: cohortes plus vastes
 - UK Biobank : 500 000 individus !

Nature Reviews Neuroscience | AOP, published online 10 April 2013; doi:10.1038/nrn3475



Power failure: why small sample size undermines the reliability of neuroscience

Katherine S. Button^{1,2}, John P. A. Ioannidis³, Claire Mokrysz¹, Brian A. Nosek⁴, Jonathan Flint⁵, Emma S. J. Robinson⁶ and Marcus R. Munafò¹

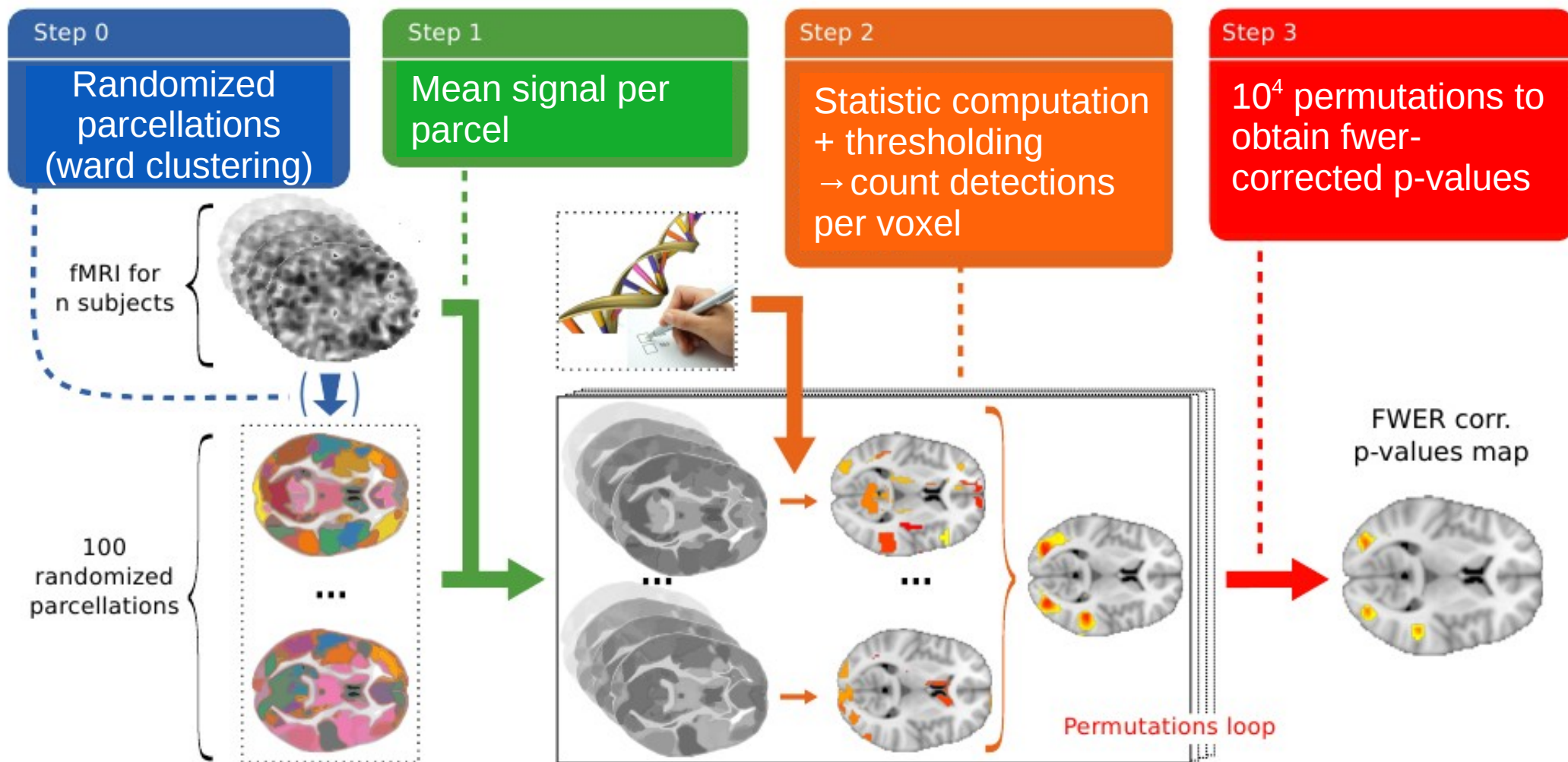
Plan

- Biomarqueurs d'imagerie fonctionnelle
- Le défi génétique - neuroimagerie fonctionnelle
- **Améliorer la modélisation statistique pour l'analyse de groupe**

Quelques approches pour résoudre le problème

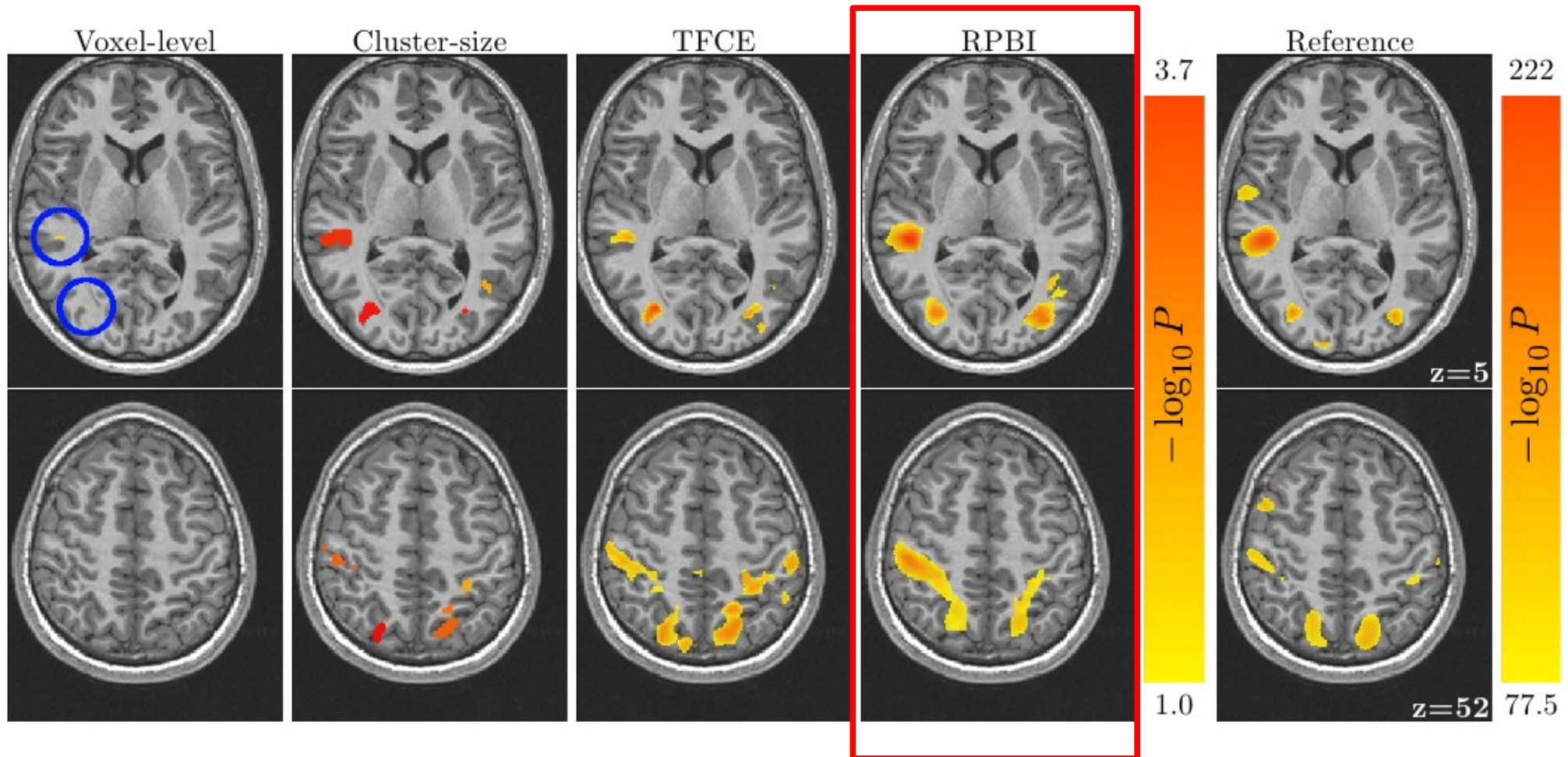
- Introduire des a-priori pertinents : continuité spatiale
- Modélisation statistique (robustesse)
- Regrouper les variables génétiques
- Approche prédictive : effet global de l'ensemble des variables d'imagerie sur le comportement

Inférence sur des découpages aléatoires



[Da Mota et al. NeuroImage 2013]

Inférence sur des découpages aléatoires

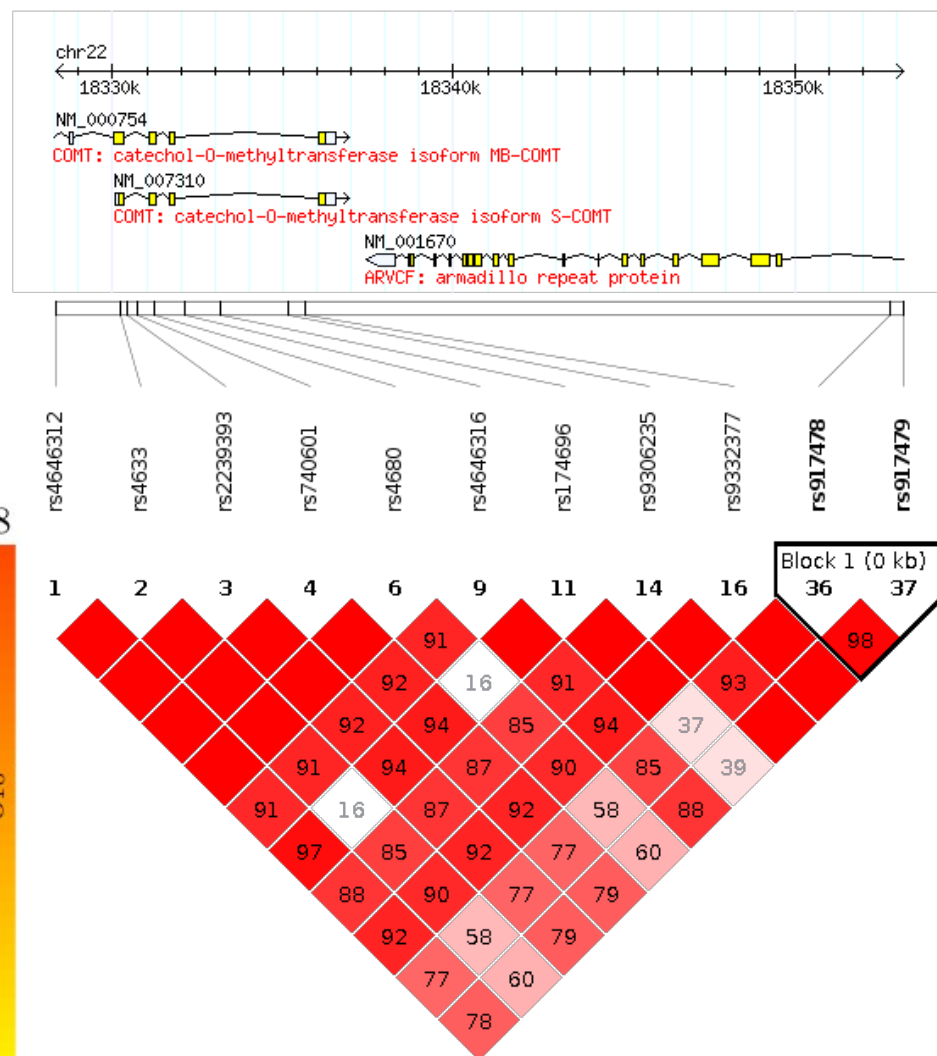
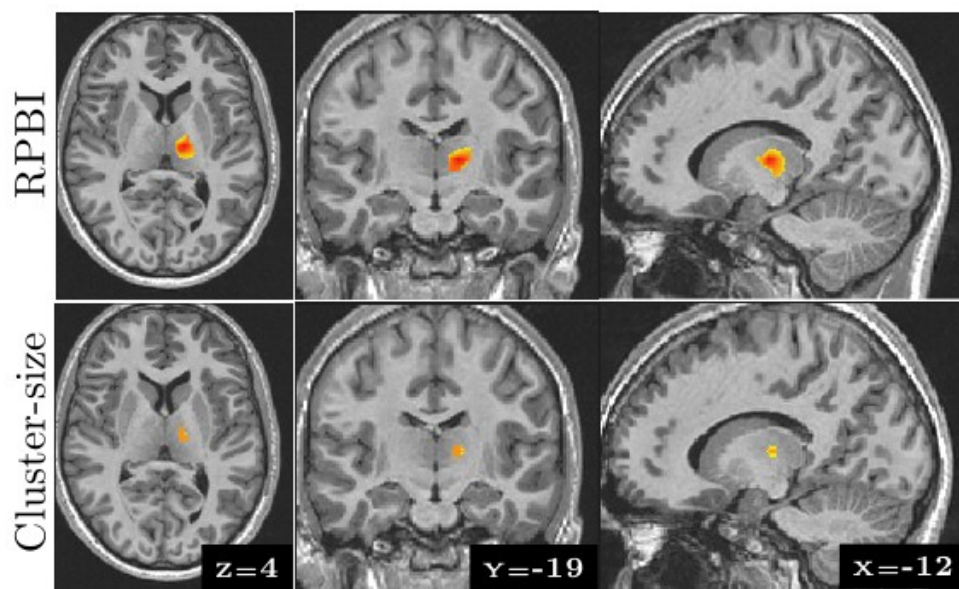


Test de l'effet principal de la tâche sans les effets (latéralité manuelle, site, sexe), sur contraste fonctionnel [angry faces - control]

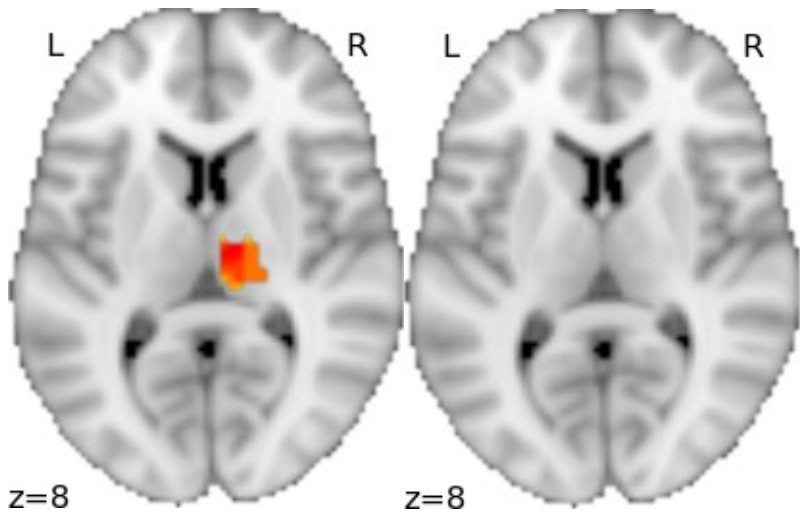
Inférence sur des découpages aléatoires

Expérience avec quelques SNPs sur le gène ARVCF (proche de COMT): signaux IRMf après des réponses erronées

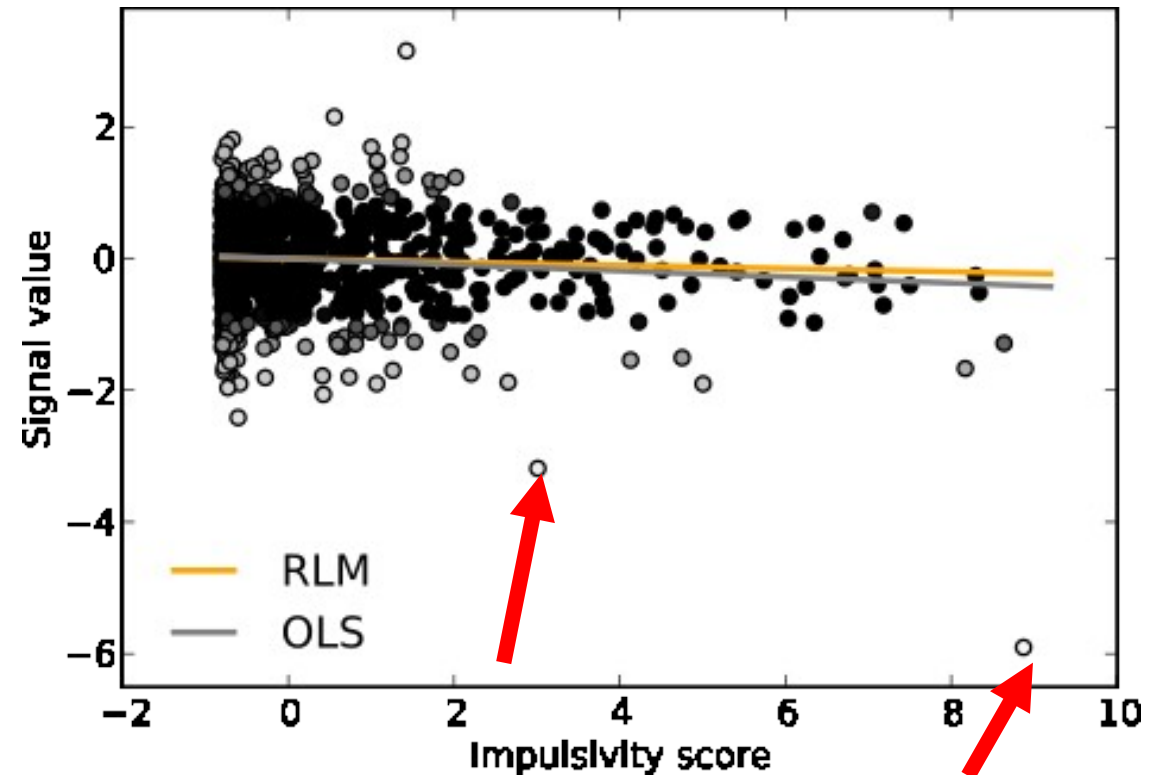
RPBI trouve une association statistique



Utilisation des statistiques robustes



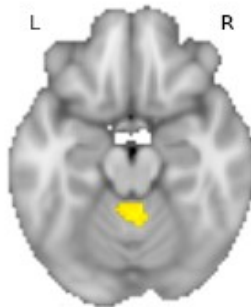
Un effet faux généré par des données aberrantes:
Un score d'impulsivité corrélé négativement avec une activation thalamique



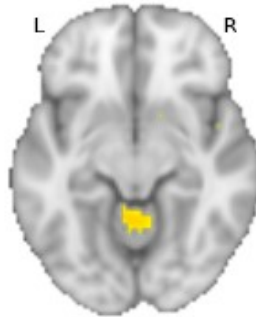
Statistiques robustes



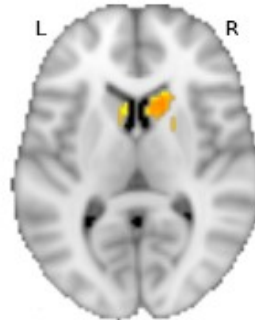
RPBI_{OLS}



$z = -18$



$z = -10$

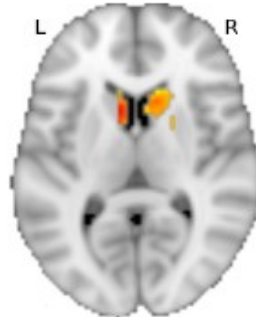
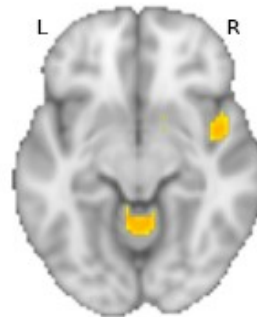
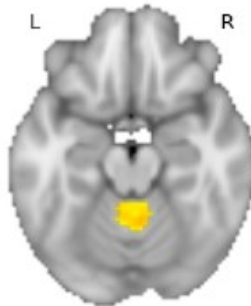


$z = 7$



$z = 56$

RPBI_{RLM}



2.15

$-\log_{10} P$

1.0

Corrélation entre

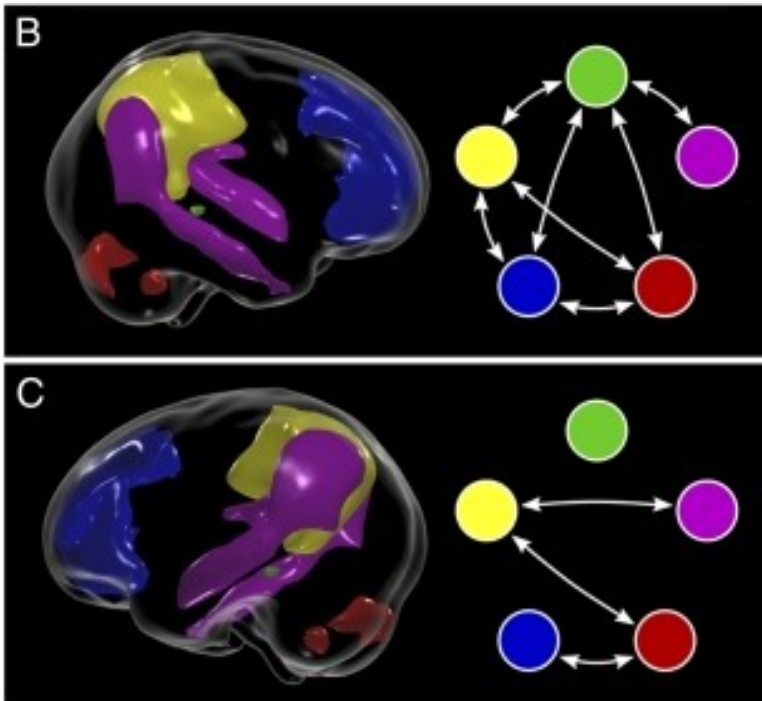
- Interaction (SNP du gène OXTR) * (# événements négatifs)
- activation à des visages en colère [Loth et al. 2013]

Régression robuste au lieu des moindres carrés → détections plus fiables et parfois plus sensibles [Fritsch et al PRNI 2013]

Neuroimagerie-génétique : hérabilité

Hérabilité: part de la variance phénotypique expliquée par la structure de parenté ($h^2 = \sigma_g^2 / \sigma_p^2$).

Hérabilité de la corrélation au sein du réseau du mode par défaut
 0.424 ± 0.17 (n=333)



Corrélations génétiques dans le réseau du mode par défaut.

Corrélations environnementales dans ce réseau.

[Glahn et al. PNAS 2010]

Héritabilité des activations fonctionnelles

$$\mathbf{Y} = \mathbf{X}\beta_1 + \mathbf{Z}\beta_2 + \epsilon$$

Phénotype d'imagerie \rightarrow \mathbf{Y} \leftarrow (full) Génotype \leftarrow covariables \leftarrow \mathbf{Z}

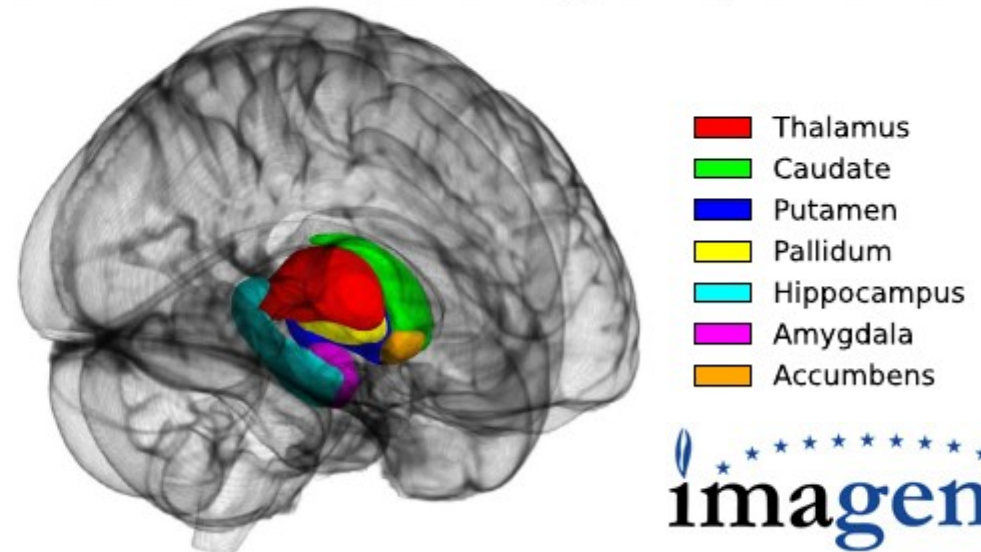
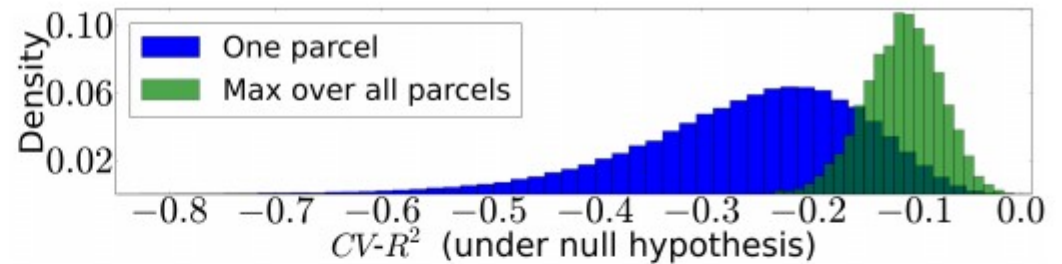
Estimation par régression ridge (voir [Lee et al., 2011; Lippert et al., 2011])

Évaluation:

$$CV-R^2 = 1 - \text{mean}_{(train, test) \in \text{split}(n)} \frac{\|\mathbf{Y}^{test} - \mathbf{X}^{test}\beta_1^{train} - \mathbf{Z}^{test}\beta_2^{train}\|^2}{\|\mathbf{Y}^{test} - \mathbf{Z}^{test}\beta_2^{train}\|^2}$$

Héritabilité des activation fonctionnelles

ROI name		$CV-R^2$	fwe corr. p-value
Thalamus	left	0.026	1.10^{-4}
	right	0.038	1.10^{-4}
Caudate	left	0.003	2.10^{-4}
	right	-0.012	3.10^{-4}
Putamen	left	0.019	1.10^{-4}
	right	0.006	2.10^{-4}
Pallidum	left	0.018	1.10^{-4}
	right	-0.010	3.10^{-4}
Hippocampus	left	0.010	2.10^{-4}
	right	0.020	1.10^{-4}
Amygdala	left	0.016	1.10^{-4}
	right	0.015	1.10^{-4}
Accumbens	left	0.022	1.10^{-4}
	right	-0.002	2.10^{-4}

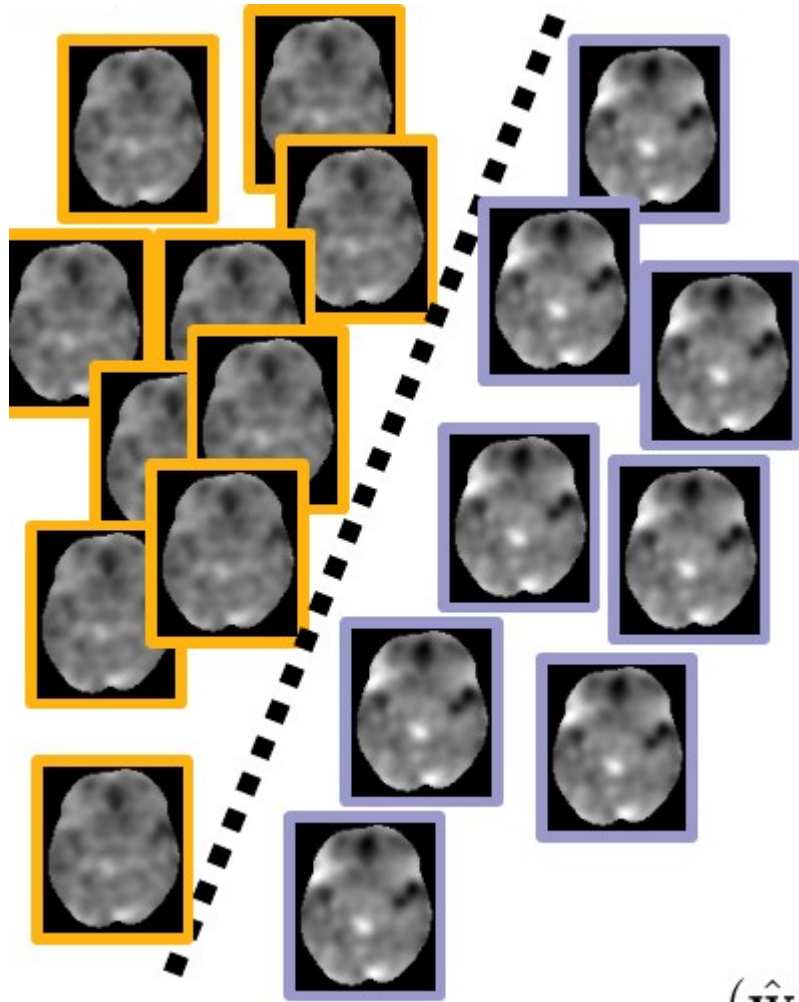


imagen

Héritabilité de l'activation pour une tâche d'échec d'arrêt dans les noyaux sous-corticaux: Les activations sont **plus héritables que la chance dans toutes les régions considérées**

[da Mota et al. 2014]

Classification de données IRMf



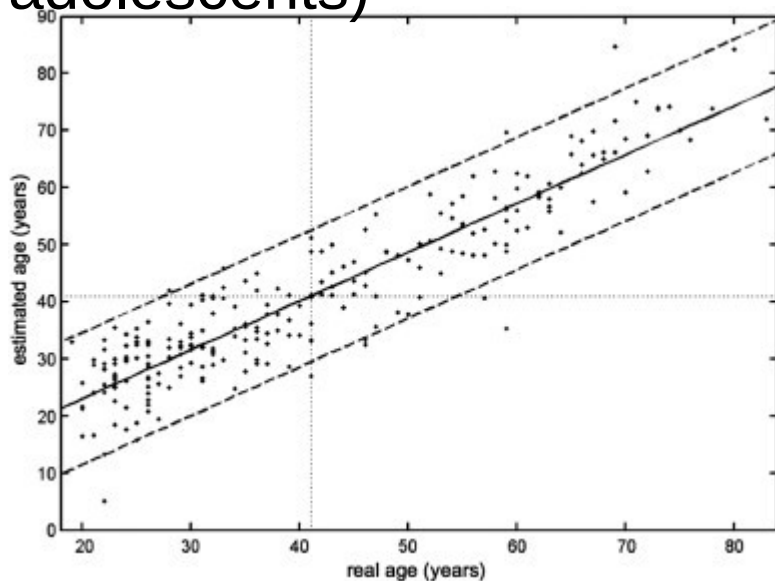
- Pour $x \in \mathbb{R}^p$, (volume IRMf de p voxels), prédire un label $y \in \{-1, 1\}$ i.e. ■ or ■
- or mieux la probabilité de classification $\text{Proba}(y = 1|x)$
- régression logistique régression: apprendre les poids w et biais b tq

$$(\hat{w}, \hat{b}) = \operatorname{argmin}_{w, b} \sum_{i=1}^n \log (1 + \exp (-y_i (\mathbf{x}_i^T \mathbf{w} + b)))$$

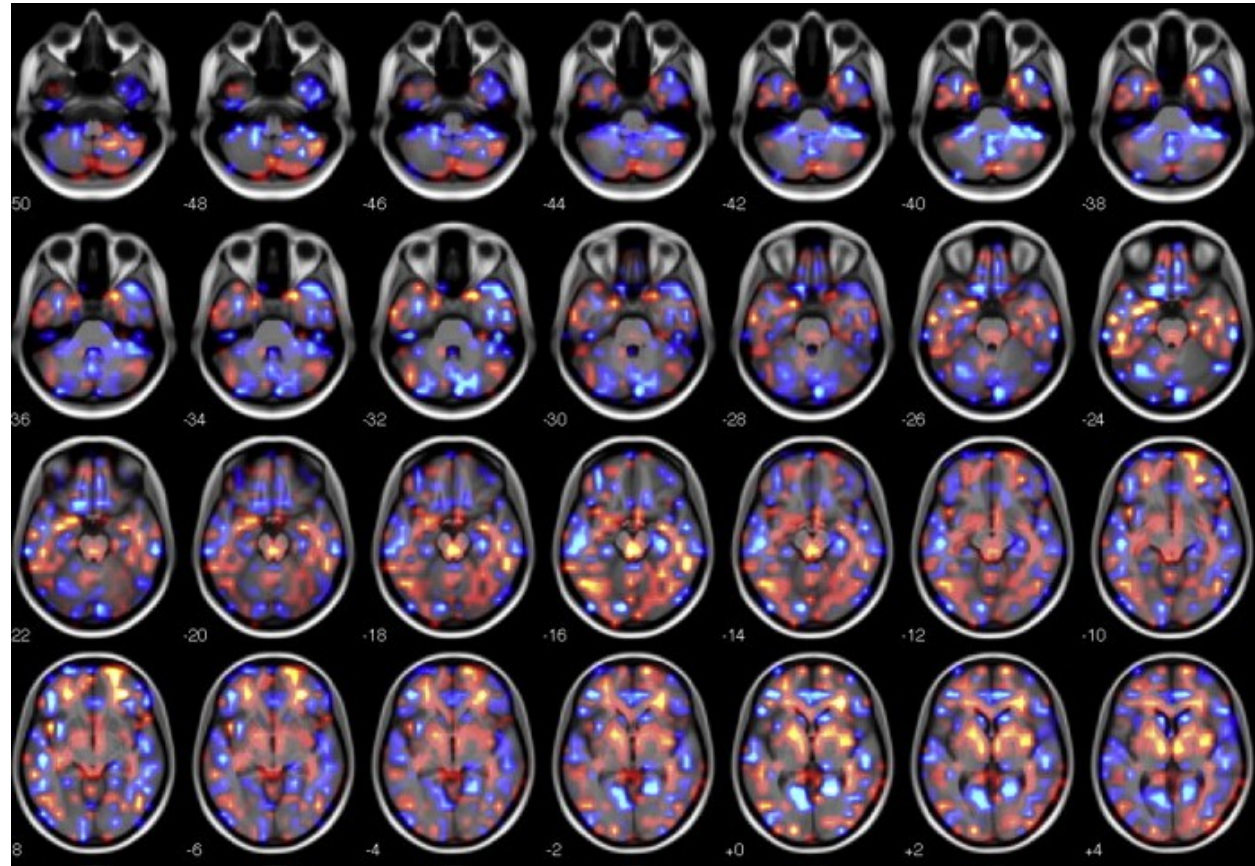
Prédire l'âge par la neuroimagerie...

... anatomique

L'épaisseur/la forme du cortex prédisent l'âge ± 5 ans (1 an pour des adolescents)



[Gaser et al. NIMG 2010]



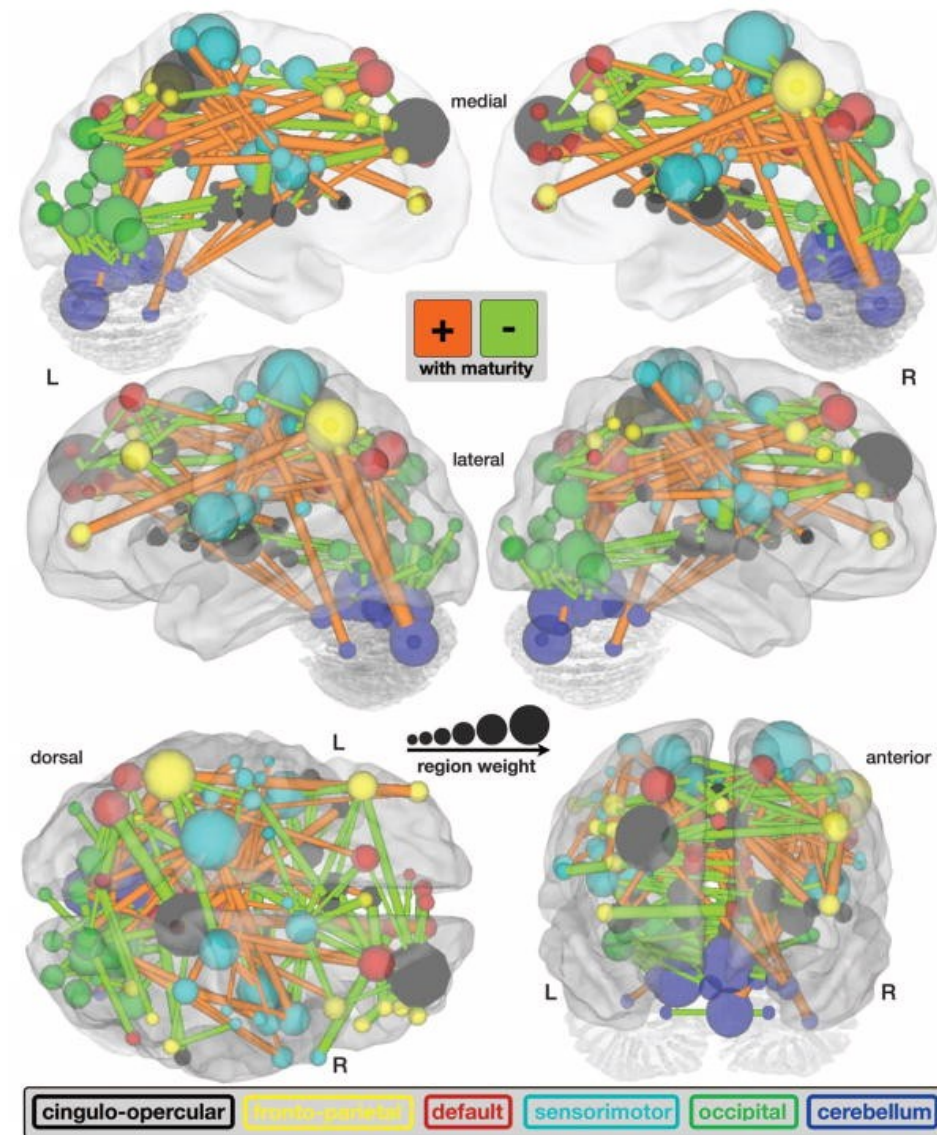
Intérêt : scénarios de vieillissement accéléré (AD), démences séniles, diabète

Prédire l'âge par la neuroimagerie...

... fonctionnelle

- Données: corrélation au repos entre 160 régions, n=120
- Les corrélations entre régions distantes reflètent l'âge de la personne [7-30 ans]
- La part des connexions «longues» augmente avec l'âge

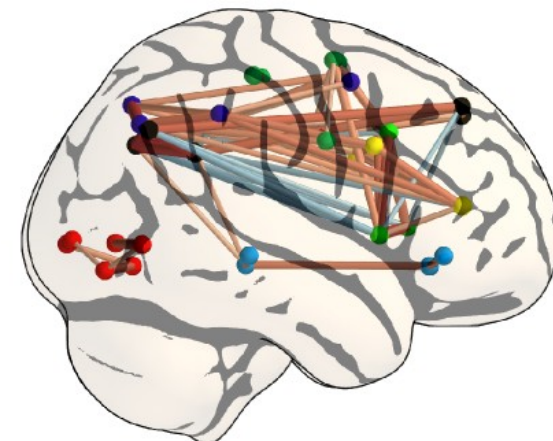
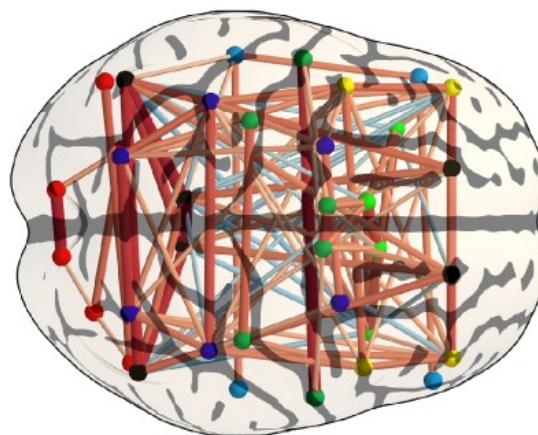
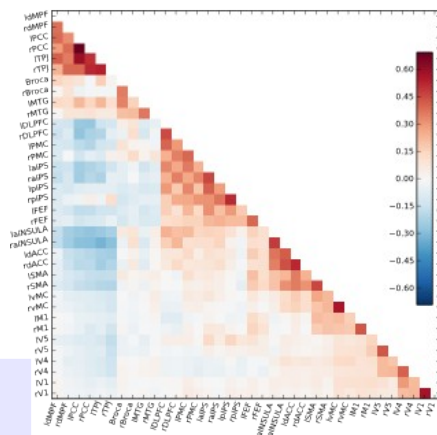
[Dosenbach et al science 2010]



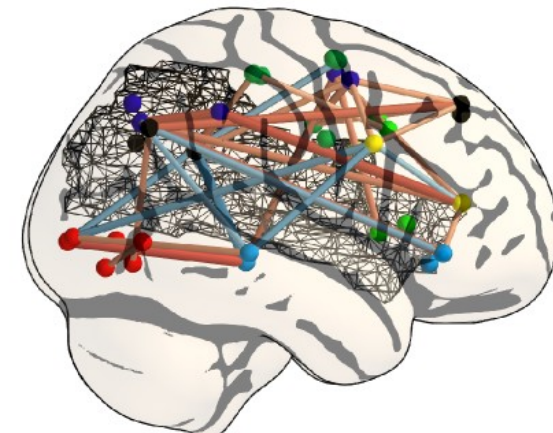
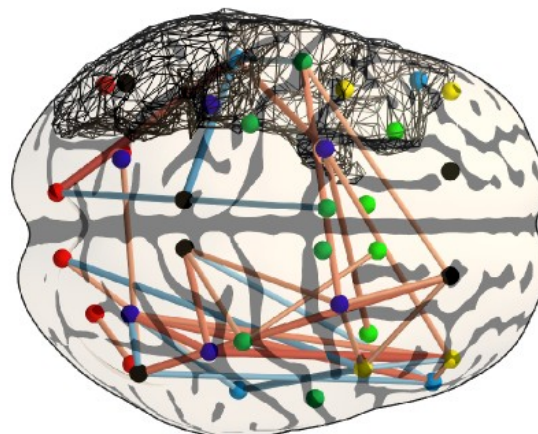
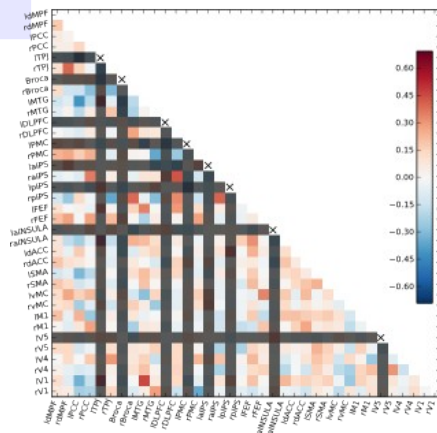
Biomarqueurs fonctionnels des déficits

Contrôles

[Varoquaux et al.
MICCAI 2010]



Perturbations distantes



Collaboration
with INSERM U992
A. Kleinschmidt,
F. Baronnet

Indication pour
la réhabilitation ?

Perspectives

- Développements méthodologiques
 - Beaucoup de **problèmes d'estimation** (régression multiple, covariance)
 - **Données manquantes**: stratégies d'imputation appropriées
 - Calcul massif : **méthodes efficaces** (e.g. stochastiques)
- Manque de puissance statistique: favoriser le **partage des données**

Remerciements

- Benoit da Mota
- Virgile Firtsch
- Vincent Frouin
- Jean-Baptiste Poline
- Gaël Varoquaux
- Kerdata (Azure experiment)
- Imagen consortium

- A-brain project



- Icogen project



- HiDiNim project



- Nicconnect project

