

Développement et Application de Méthodologies Statistiques pour Études Multi-Omiques dans le Diabète de Type 2

Au-delà de l'Ère des Études d'Association Pangénomiques

Soutenance de Thèse

MICKAËL CANOUIL

MICKAEL.CANOUIL@CNRS.FR

Direction de Thèse

Dr. Chislain Rocheleau & Pr. Philippe Froguel

29 Septembre 2017



Sommaire

1 Contexte

2 Glycémie et Diabète de Type 2 : un Effet
Génétique Conjoint?

3 Au delà des Études Pangénomiques

4 Apports & Perspectives

Contexte

Le Diabète dans le Monde

Quelques chiffres sur le diabète [Roglic and World Health Organization, 2016] :

En 1980

- **108 millions** de diabétiques
- prévalence mondiale de **4,7 %**

Le Diabète dans le Monde

Quelques chiffres sur le diabète [Roglic and World Health Organization, 2016] :

En 1980

- **108 millions** de diabétiques
- prévalence mondiale de **4,7 %**

En 2014

- **400 millions** de diabétiques
- prévalence mondiale de **8,5 %**

Les Formes de Diabète

Principalement deux formes de diabète :

Diabète de type 1 Maladie auto-immune (diabète “jeune”)

Diabète de type 2 Maladie métabolique complexe (diabète “adulte”)

Les Formes de Diabète

Principalement deux formes de diabète :

Diabète de type 1 Maladie auto-immune (diabète “jeune”)

Diabète de type 2 Maladie métabolique complexe (diabète “adulte”)

Le Diabète de Type 2

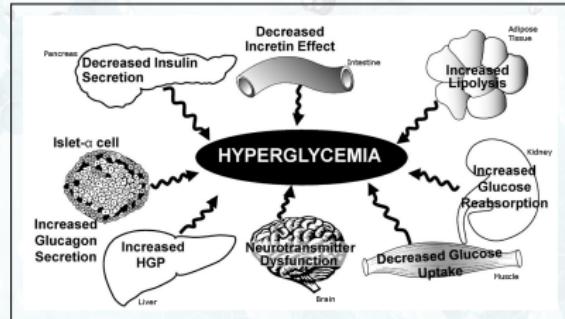
Le **diabète de type 2** (T2D) se caractérise par une **hyperglycémie** chronique, conséquence d'une production insuffisante d'**insuline**, et est défini selon deux critères de diagnostique :

OMS Organisation Mondiale de la Santé

- [Glycémie_{à jeun}] $\geq 7 \text{ mmol/L}$
- [Glycémie_{à 2 heures}] $\geq 11,1 \text{ mmol/L}$

ADA Association Américaine pour le Diabète

- [Glycémie_{à jeun}] $\geq 7 \text{ mmol/L}$
- [HbA1c] $> 6,5 \%$



Tissus et organes impliqués dans l'hyperglycémie et le diabète de type 2 (HGP : "hepatic glucose production", production hépatique de glucose).

La Génétique du Diabète de Type 2

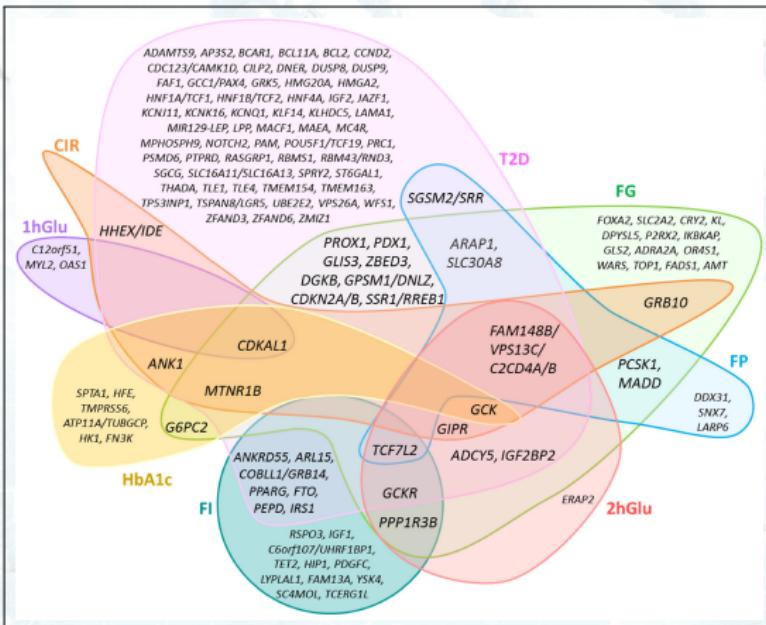


Diagramme de Venn des loci identifiés par études d'association pangénomiques pour leur effet sur différents traits glycémiques et le diabète de type 2 [Marullo et al., 2014].

T2D : diabète de type 2;

FG : glycémie à jeun;

FI : insulinémie à jeun;

FP : proinsulinémie à jeun;

2hGlu : glycémie à 2 heures;

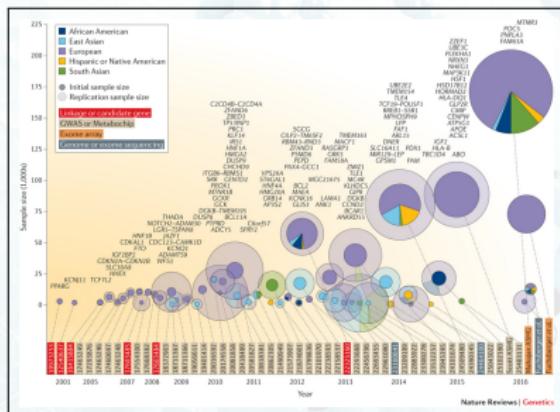
1hGlu : glycémie à 1 heure;

HbA1c : hémoglobine glyquée;

CIR : ratio carbohydrate-insuline.

Un Grand Nombre de Découverte, mais ...

Plus de **100 loci identifiés**, mais :

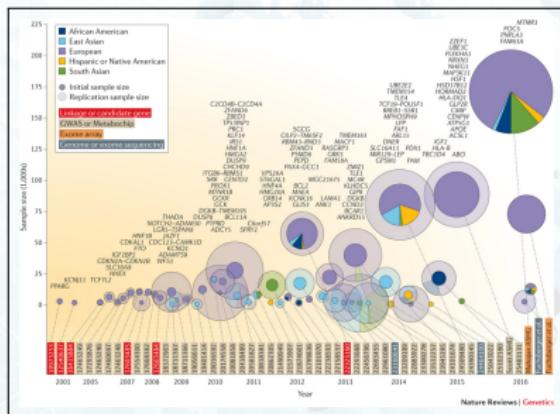


Historique des loci de susceptibilité au diabète de type 2 identifiés par études d'association pangénomiques [Flannick and Florez, 2016].

- des **effets** sur la glycémie à jeun (FG) et le risque de diabète de type 2 (T2D) **peu corrélés**
- une **fonction** biologique **méconnue** ou **inconnue** pour la plupart

Un Grand Nombre de Découverte, mais ...

Plus de **100 loci identifiés**, mais :



Historique des loci de susceptibilité au diabète de type 2 identifiés par études d'association pangénomiques [Flannick and Florez, 2016].

- des **effets** sur la glycémie à jeun (FG) et le risque de diabète de type 2 (T2D) **peu corrélés**
- ⇒ Effet génétique conjoint sur **FG** et le risque de **T2D**?
- une **fonction** biologique **méconnue** ou inconnue pour la plupart
- ⇒ Lien entre l'**expression** des gènes et la sécrétion d'**insuline**?

Glycémie et Diabète de Type 2 : un Effet Génétique Conjoint?

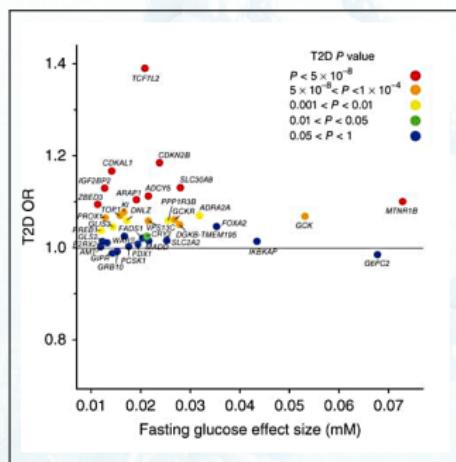
Glycémie et Diabète de Type 2 : un Effet Génétique Conjoint?

Variants Génétiques Associés à la Trajectoire de la Glycémie à Jeun et à l'Incidence du Diabète de Type 2 : Une Approche par Modèle Joint

Mickaël Canouï^{1,2,3}, Beverley Balkau^{4,5,6}, Ronan Rousset^{7,8,9}, Philippe Froguel^{1,2,3,10} & Ghislain Rocheleau^{1,2,3}

Contexte & Objectifs

"In all cases, the glucose-raising allele was associated with increased risk of T2D, yet fasting glucose effect sizes and T2D ORs were weakly correlated"



Taille d'effet β et Odd Ratio respectivement pour l'augmentation de la concentration de glucose et du risque de T2D (Scott et al. [2012]).

Objectif

Recherche de variants (SNPs) associés à la fois à la glycémie à jeun et au diabète de type 2

Données

Cohorte D.E.S.I.R.

Méthode

Modèle Joint

Modèle Joint : Présentation

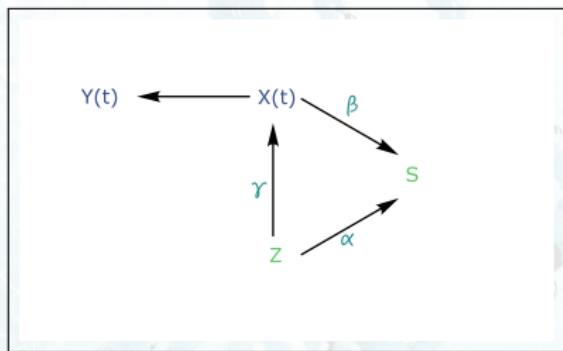
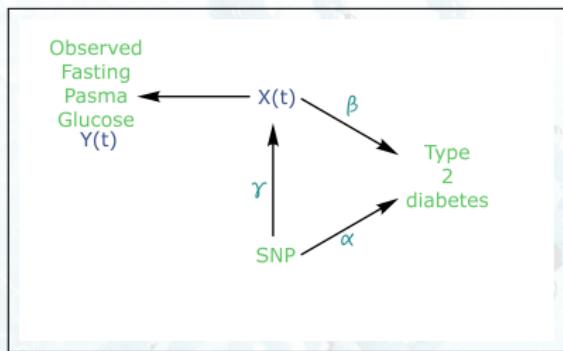


Diagramme causal du Modèle Joint
(adapté de Ibrahim et al. [2010]).

- $X(t)$ trajectoire du biomarqueur générant les observations longitudinales $Y(t)$
- $Y(t)$ données observées avec erreur de mesure
- T événement (survie) au temps T
- Z traitement (covariable)
- α effet du traitement sur l'événement
- β effet de la trajectoire sur l'événement
- γ effet du traitement sur la trajectoire

Modèle Joint : Présentation



- $X(t)$ trajectoire de la glycémie générant les observations longitudinales $Y(t)$
- $Y(t)$ glycémie observée avec erreur de mesure
- S survenue du diabète au temps T
- α effet du SNP sur le diabète
- β effet de la glycémie sur le diabète
- γ effet du SNP sur la glycémie

Modèle Joint : Motivation

Objectifs des études cliniques :

- identifier un biomarqueur prédictif d'un événement clinique;
- effet d'un traitement sur un événement clinique
(évenement indépendant du traitement conditionnellement au biomarqueur).

Exemple

Biomarqueur Glycémie à jeun

Événement Survenue d'un diabète de type 2

Modèle de Cox à risque proportionnel étendu

$$h(t) = h_0(t) \exp(\beta Y(t)) \quad (1)$$

Modèle Joint : Motivation

Objectifs des études cliniques :

- identifier un biomarqueur prédictif d'un événement clinique;
- effet d'un traitement sur un événement clinique
(évenement indépendant du traitement conditionnellement au biomarqueur).

Exemple

Biomarqueur Glycémie à jeun

Événement Survenue d'un diabète de type 2

Modèle de Cox à risque proportionnel étendu

$$h(t) = h_0(t) \exp\{\beta Y(t) + \alpha Z\} \quad (1)$$

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation ⇒ **Introduction de biais**

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation ⇒ **Introduction de biais**
- Variabilité des mesures du biomarqueur

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation ⇒ **Introduction de biais**
- Variabilité des mesures du biomarqueur
- ⇒ Mauvaise représentation de la “vraie” trajectoire ($Y_i(t_{ij}) \neq X_i(t_{ij})$)

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation ⇒ **Introduction de biais**
- Variabilité des mesures du biomarqueur
- ⇒ Mauvaise représentation de la “vraie” trajectoire ($Y_i(t_{ij}) \neq X_i(t_{ij})$)
- Nature “endogène” du biomarqueur

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation ⇒ **Introduction de biais**
- Variabilité des mesures du biomarqueur
- ⇒ Mauvaise représentation de la “vraie” trajectoire ($Y_i(t_{ij}) \neq X_i(t_{ij})$)
- Nature “endogène” du biomarqueur
- ⇒ La trajectoire du biomarqueur est affectée par l'événement

Modèle Joint : Limitations

- Biomarqueur non mesurable en tout temps ou au temps de survenue de l'événement
 - Mesure à des temps spécifiques (t_{ij})
 - Valeur manquante
- ⇒ Imputation ⇒ **Introduction de biais**
- Variabilité des mesures du biomarqueur
- ⇒ Mauvaise représentation de la “vraie” trajectoire ($Y_i(t_{ij}) \neq X_i(t_{ij})$)
- Nature “endogène” du biomarqueur
- ⇒ La trajectoire du biomarqueur est affectée par l'événement ⇒ **Introduction de biais**

Modèle Joint : Avantages

- Tester l'effet d'un SNP simultanément sur :
 - le biomarqueur (γ);
 - la survenue de l'événement (α);
 - le biomarqueur et la survenue de l'événement ($\beta\gamma + \alpha$).

Modèle Joint : Avantages

- Tester l'effet d'un SNP simultanément sur :
 - le biomarqueur (γ);
 - la survenue de l'événement (α);
 - le biomarqueur et la survenue de l'événement ($\beta\gamma + \alpha$).
- Gain de puissance statistique :
 - si $\beta \neq 0$ (pour détecter un effet conjoint du SNP $\beta\gamma + \alpha \neq 0$)
 - ◆ [Chen et al., 2011];
 - par rapport à un modèle de “Cox étendu” (modèle de Cox avec covariable dépendante du temps).

Formulation

Dans le Modèle Joint, deux composantes (modèles) sont utilisées :

- un modèle longitudinal est utilisé pour modéliser la “vraie” trajectoire du biomarqueur (non-observable),

Formulation

Dans le Modèle Joint, deux composantes (modèles) sont utilisées :

- un modèle longitudinal est utilisé pour modéliser la “vraie” trajectoire du biomarqueur (non-observable),
- et est incorporé en tant que covariable (latente) dépendante du temps dans un modèle de survie.

Formulation : Composante Longitudinale

Composante longitudinale : modèle (linéaire) mixte

Avec T_i , le temps d'événement et C_i , le temps de censure à droite :

$$\tilde{T}_i = \min(T_i, C_i), \quad (2)$$

avec $t_{ij} \leq \tilde{T}_i, j = 1, \dots, n_i$, où n_i est le nombre de mesures.

Pour un individu i :

$$Y_i(t_{ij}) = X_i(t_{ij}) + \epsilon_i(t_{ij}) \quad (3)$$

Où :

$$\epsilon_i(t_{ij}) \sim \mathcal{N}(0, \sigma^2) \quad (4)$$

Formulation : Composante Longitudinale

Composante longitudinale : modèle (linéaire) mixte

Pour un individu i :

$$Y_i(t_{ij}) = X_i(t_{ij}) + \epsilon_i(t_{ij}) \quad (5)$$

La trajectoire $X_i(t_{ij})$ peut être définie par un polynôme, fonction du temps t_{ij} :

$$X_i(t_{ij}) = \theta_{0i} + \theta_{1i}t_{ij} + \cdots + \theta_{pi}t_{ij}^p, \quad \boldsymbol{\theta}_p \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (6)$$

Formulation : Composante Longitudinale

Composante longitudinale : modèle (linéaire) mixte

Pour un individu i :

$$Y_i(t_{ij}) = X_i(t_{ij}) + \epsilon_i(t_{ij}) \quad (5)$$

La trajectoire $X_i(t_{ij})$ peut être définie par un polynôme, fonction du temps t_{ij} :

$$X_i(t_{ij}) = \theta_{0i} + \theta_{1i}t_{ij} + \cdots + \theta_{pi}t_{ij}^p, \quad \boldsymbol{\theta}_p \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (6)$$

Exemple : D.E.S.I.R.

Avec W_i , la matrice des covariables d'ajustement (Âge, Sexe et IMC) et Z_i , le génotype du SNP d'intérêt :

$$Y_i(t_{ij}) = \theta_{0i} + \theta_{1i}t_{ij} + \gamma Z_i + \delta W_i + \epsilon_i(t_{ij}) \quad (7)$$

Formulation : Composante de Survie

Composante de survie : modèle de Cox à risque proportionnel

Interrelation entre $X_i(t)$ et T_i :

$$h_i(t) = \lim_{dt \rightarrow 0} \frac{P\{t \leq T_i < t + dt | T_i \geq t, \bar{Y}_i(t)\}}{dt} \quad (8)$$

$$= h_0(t) \exp\{\beta X_i(t)\} \quad (9)$$

Où $\bar{Y}_i(t) = \{Y_i(u), 0 \leq u \leq t\}$ est l'historique de la trajectoire jusqu'au temps t .

Formulation : Composante de Survie

Composante de survie : modèle de Cox à risque proportionnel

Interrelation entre $X_i(t)$, T_i et W_i :

$$h_i(t) = \lim_{dt \rightarrow 0} \frac{P\{t \leq T_i < t + dt | T_i \geq t, \bar{Y}_i(t), W_i\}}{dt} \quad (8)$$

$$= h_0(t) \exp\{\beta X_i(t) + \eta W_i\} \quad (9)$$

Où $\bar{Y}_i(t) = \{Y_i(u), 0 \leq u \leq t\}$ est l'historique de la trajectoire jusqu'au temps t .

Formulation : Composante de Survie

Composante de survie : modèle de Cox à risque proportionnel

Interrelation entre $X_i(t)$, T_i et W_i :

$$h_i(t) = \lim_{dt \rightarrow 0} \frac{P\{t \leq T_i < t + dt | T_i \geq t, \bar{Y}_i(t), W_i\}}{dt} \quad (8)$$

$$= h_0(t) \exp\{\beta X_i(t) + \eta W_i\} \quad (9)$$

Où $\bar{Y}_i(t) = \{Y_i(u), 0 \leq u \leq t\}$ est l'historique de la trajectoire jusqu'au temps t .

Exemple : D.E.S.I.R.

Avec W_i , la matrice des covariables d'ajustement (Âge, Sexe et IMC) et Z_i , le génotype du SNP d'intérêt :

$$h_i(t_{ij}) = h_0(t_{ij}) \exp\{\beta X_i(t_{ij}) + \alpha Z_i + \eta W_i\} \quad (10)$$

Test d'Hypothèse

Pour tester l'hypothèse nulle :

$$H_0 : \theta = \theta_0 \quad (H_1 : \theta \neq \theta_0)$$

Test du Rapport de Vraisemblance $LRT = -2\{\ell(\hat{\theta}_0) - \ell(\hat{\theta})\}$

Test de Wald $W = (\hat{\theta} - \theta_0)^\top \mathcal{I}(\hat{\theta})(\hat{\theta} - \theta_0)$ (Univarié : $(\hat{\theta}_j - \theta_{0j})/\widehat{s.e.}(\hat{\theta}_j)$)

Test du Score $U = S^\top(\hat{\theta}_0)\{\mathcal{I}(\hat{\theta}_0)\}^{-1}S(\hat{\theta}_0)$

Étude de Simulation : Objectifs

- Précision des estimateurs : RMSE (Root-Mean Square Error)

Root-Mean Square Error

$$\text{MSE}(\hat{\theta}) = \text{Biais}(\hat{\theta})^2 + \text{Var}(\hat{\theta})$$

$$\begin{aligned}\text{RMSE}(\hat{\theta}) &= \sqrt{\text{MSE}(\hat{\theta})} \\ &= \sqrt{E\{(\hat{\theta} - \theta)^2\}}\end{aligned}$$

- Faisabilité à l'échelle du génome : temps de calcul

Étude de Simulation : Objectifs

⇒ Comparaison à une approche en “deux-étapes” (TS) [Tsiatis et al., 1995] :

“Two-Step”

- Première étape :

$$Y_i(t) = X_i(t) + \epsilon_i(t)$$

$$X_i^*(t) = E\{X_i(t)|\bar{Y}_i(t), T_i \geq t\}$$

- Deuxième étape :

$$h_i(t) = h_0(t) \exp\{\beta X_i^*(t)\}$$

Étude de Simulation : Génération des Données

Modèle longitudinal : $Y_i(t) = \theta_{0i} + \theta_{1i}t + \gamma Z_i + \epsilon_i(t)$

Modèle de survie : $h_i(t) = h_0(t) \exp\{\beta X_i(t) + \alpha Z_i\}$

Génération des temps d'événement : distribution exponentielle

$$\lambda > 0$$

$$h_0(t) = \lambda$$

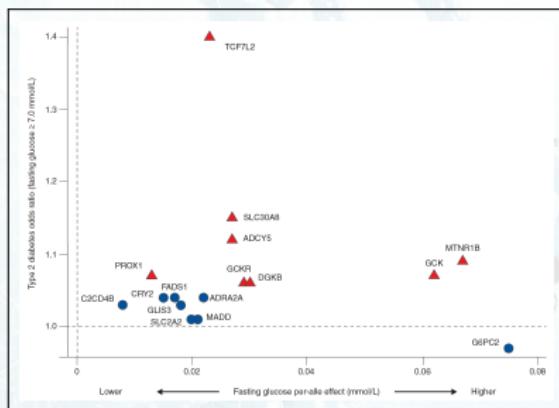
$$H_0(t) = \lambda t$$

$$H_i(T_i) = \int_0^{T_i} \lambda \exp(\beta X_i(t) + \alpha Z_i) dt$$

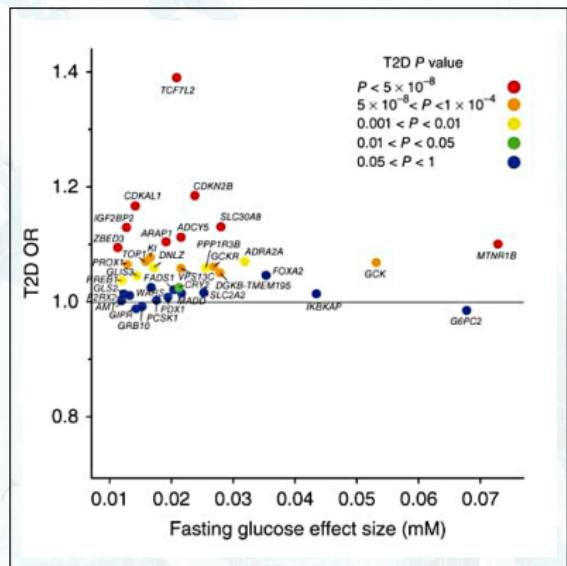
$$T_i = \frac{1}{\beta \theta_{1i}} \log \left(1 - \frac{\beta \theta_{1i} \times \log(1-u)}{\lambda \exp(\beta \theta_{0i} + (\beta \gamma + \alpha) Z_i)} \right), u \sim \mathcal{U}(0,1)$$

Étude de Simulation : Paramètres de Simulation

Yaghootkar and Frayling [2013]



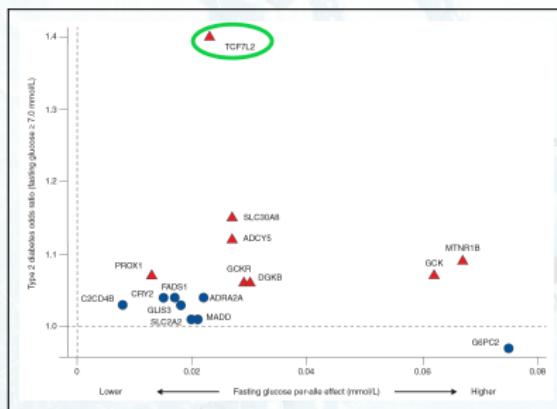
Scott et al. [2012]



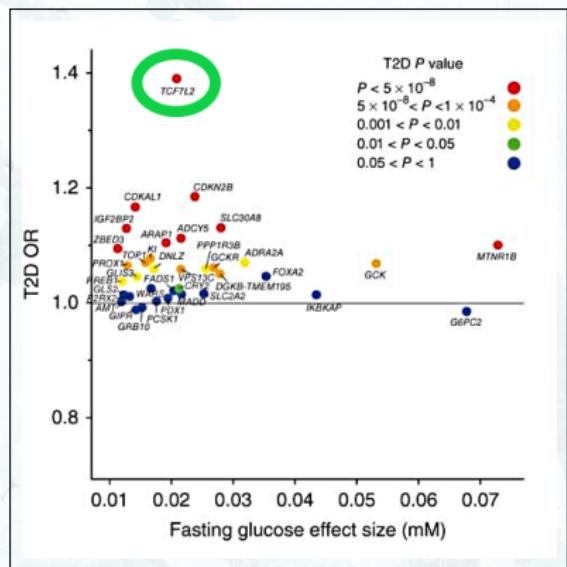
Taille d'effet β et Odd Ratio respectivement pour la concentration de glucose et le diabète de type 2.

Étude de Simulation : Paramètres de Simulation

Yaghootkar and Frayling [2013]



Scott et al. [2012]



Taille d'effet β et Odd Ratio respectivement pour la concentration de glucose et le diabète de type 2.

Étude de Simulation : Paramètres de Simulation

Parameters	Values
Number of participants (n)	4,352
Number of measures (m)	4
Diabetes incidence rate (d)	0.0384
Minor allele frequency (f)	0.244
Random effects (θ)	$\sim \mathcal{N}_2 \left(\begin{bmatrix} 4.55 \\ 0.0108 \end{bmatrix}, \begin{bmatrix} 0.143 & -0.00109 \\ -0.00109 & 6.8 \times 10^{-4} \end{bmatrix} \right)$
SNP effect on $Y_i(t_{ij})$ (γ)	0.0229
SNP effect on T_i (α)	0.265
Association between $Y_i(t_{ij})$ and T_i (β)	3.17
Error term (ϵ)	$\sim \mathcal{N}(0, 0.305^2)$

Paramètres de simulation des données, basés sur rs17747324 (TCF7L2).

Étude de Simulation : Paramètres de Simulation

Puissance statistique pour détecter un effet joint du SNP rs17747324 (*TCF7L2*), à partir de la formule développée par [Chen et al., 2011] :

$$H_0 : \beta\gamma + \alpha = 0$$

$$d = \frac{(z_{\tilde{\beta}} + z_{1-\tilde{\alpha}})^2}{f(1-f)(\beta\gamma + \alpha)^2}$$

$$z_{\tilde{\beta}} = \pm \sqrt{df(1-f)(\beta\gamma + \alpha)^2} + z_{1-\tilde{\alpha}}$$

avec :

d le nombre de T2D incidents,

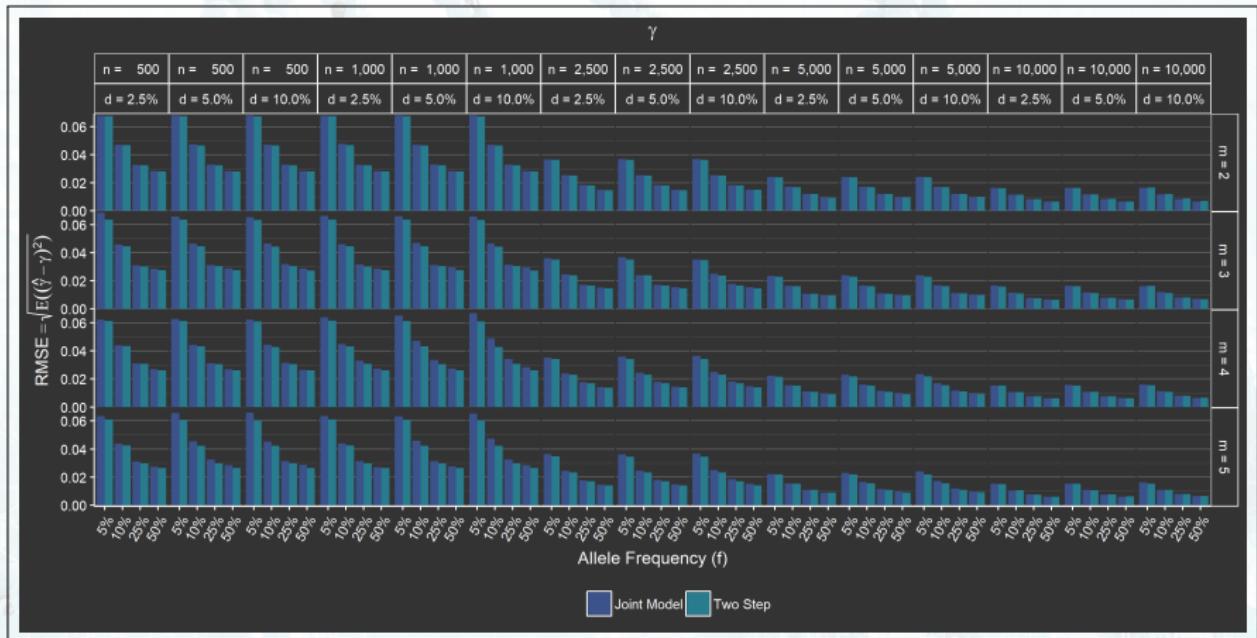
f la fréquence de l'allèle à risque,

α̃ le seuil de significativité,

$\tilde{\beta}_{\tilde{\alpha}}$ la puissance statistique au seuil $\tilde{\alpha}$.

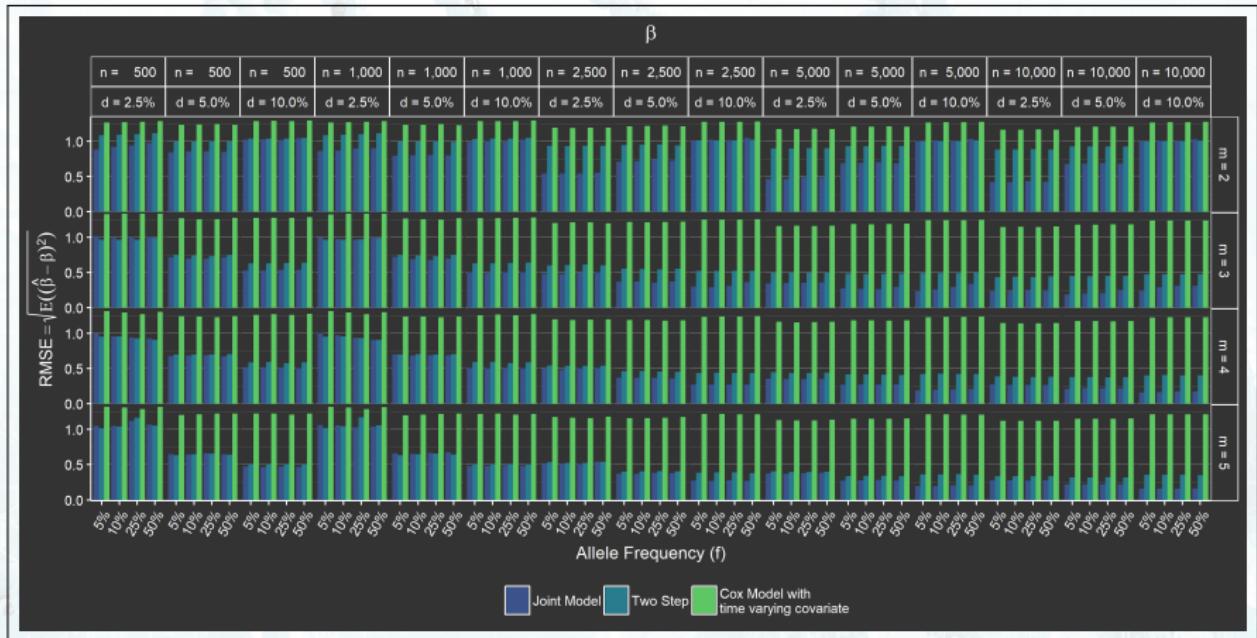
$$\Rightarrow \tilde{\beta}_{\tilde{\alpha}} = 46,56\% \text{ pour rs17747324 (TCF7L2)}$$

Étude de Simulation : Résultats



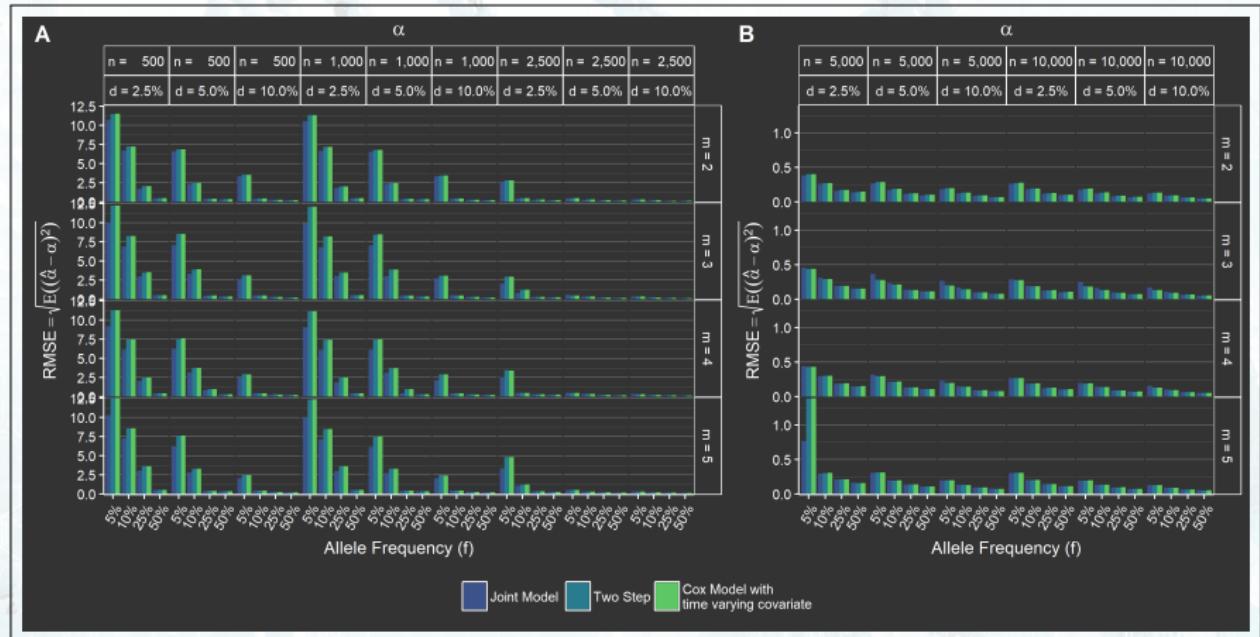
Étude de simulation de l'estimation $\hat{\gamma}$ du Modèle Joint (extension JM) et de l'approche en "deux étapes" (extension nlme).

Étude de Simulation : Résultats



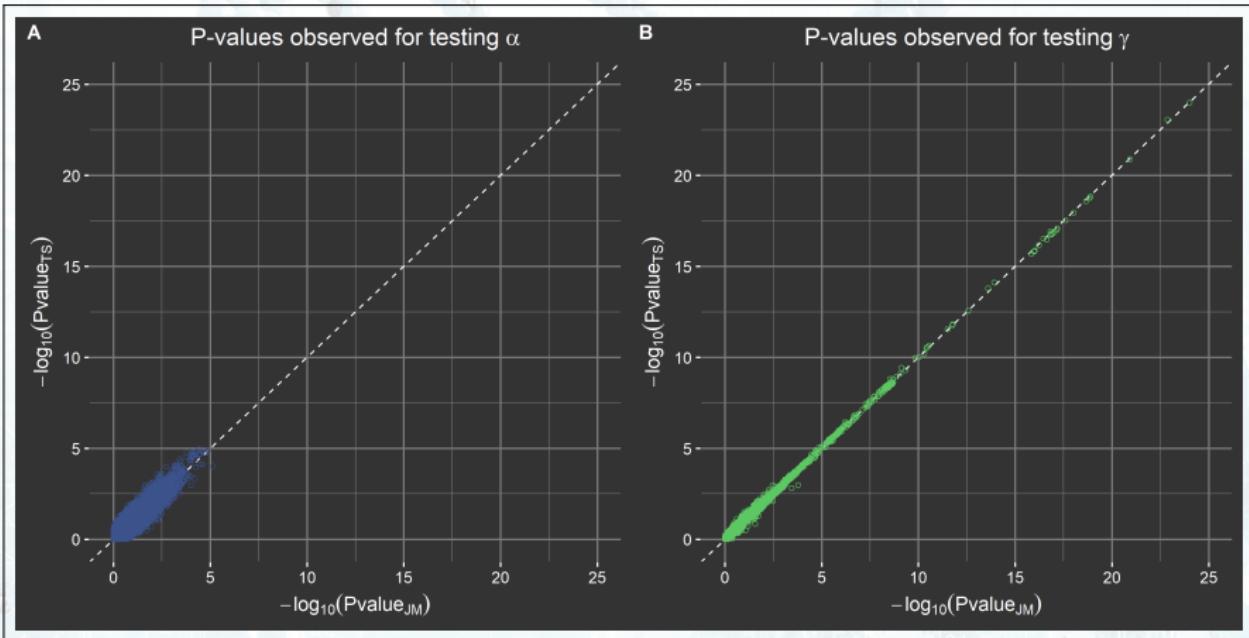
Étude de simulation de l'estimation $\hat{\beta}$ du Modèle Joint (extension JM) et de l'approche en "deux étapes" (extension nlme).

Étude de Simulation : Résultats



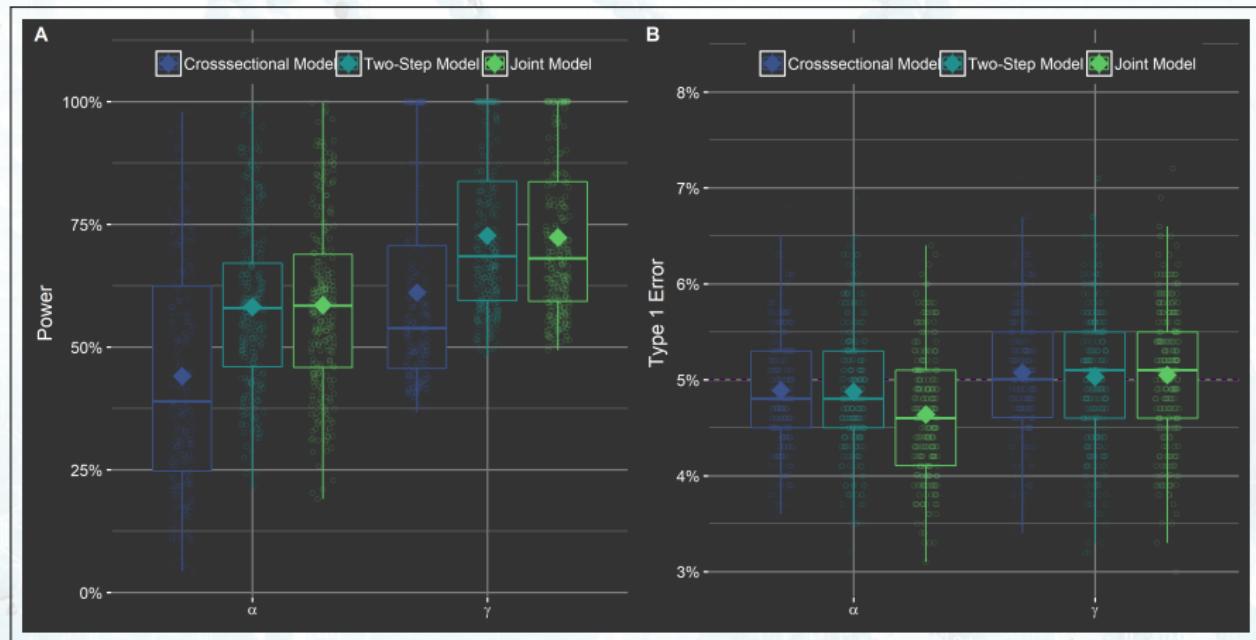
Étude de simulation de l'estimation $\hat{\alpha}$ du Modèle Joint (extension JM) et de l'approche en "deux étapes" (extension nlme).

Étude de Simulation : Résultats



Comparaison des résultats des approches par Modèle Joint et en “deux étapes”.

Étude de Simulation : Résultats



Comparaison des résultats des approches par Modèle Joint et en “deux étapes”.

Étude de Simulation : Résultats

Nombre d'individu	Modèle Joint		Modèle “deux-étapes”	
	Moyenne (Écart-Type)	100K SNPs	Moyenne (Écart-Type)	100K SNPs
500	51 s (3,4)	59 j	0,71 s (0,066)	0,82 j
2 500	100 s (11)	120 j	3,1 s (0,092)	3,6 j
5 000	180 s (25)	210 j	6,3 s (0,17)	7,3 j
10 000	340 s (34)	400 j	9 s (0,22)	10 j

Temps de calcul (fonction system.time du logiciel R).

Étude de Simulation : Résultats

Nombre d'individu	Modèle Joint		Modèle “deux-étapes”	
	Moyenne (Écart-Type)	100K SNPs	Moyenne (Écart-Type)	100K SNPs
500	51 s (3,4)	59 j	0,71 s (0,066)	0,82 j
2 500	100 s (11)	120 j	3,1 s (0,092)	3,6 j
5 000	180 s (25)	210 j	6,3 s (0,17)	7,3 j
10 000	340 s (34)	400 j	9 s (0,22)	10 j

Temps de calcul (fonction system.time du logiciel R).

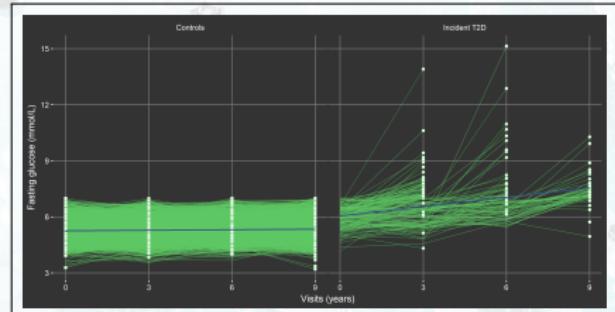
Application sur Données Réelles : D.E.S.I.R.

D.E.S.I.R.

- Données Épidémiologiques sur le Syndrome d'Insulino-Résistance
- 5 212 individus
- 1 visite tous les 3 ans, de 0 à 9 ans
- 4 500 individus génotypés

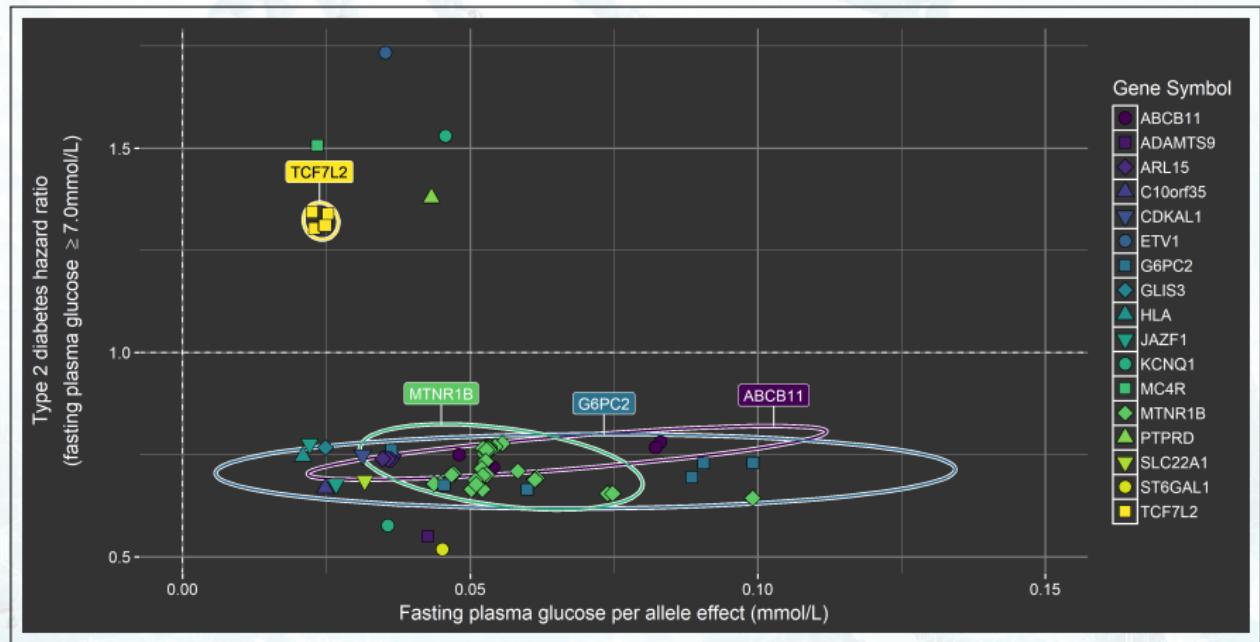
	Témoins	Cas Incidents
Sexe	♂2 293; ♀2 483	♂142; ♀64
Âge (années)	46 (10)	51 (9)
Glycémie (mmol/L)	5,2 (0,51)	6 (0,57)
BMI	24 (3,7)	28 (4,4)
HbA1c (%)	5,4 (0,39)	5,8 (0,48)

Caractéristiques des individus D.E.S.I.R. selon le statut diabète de type 2.



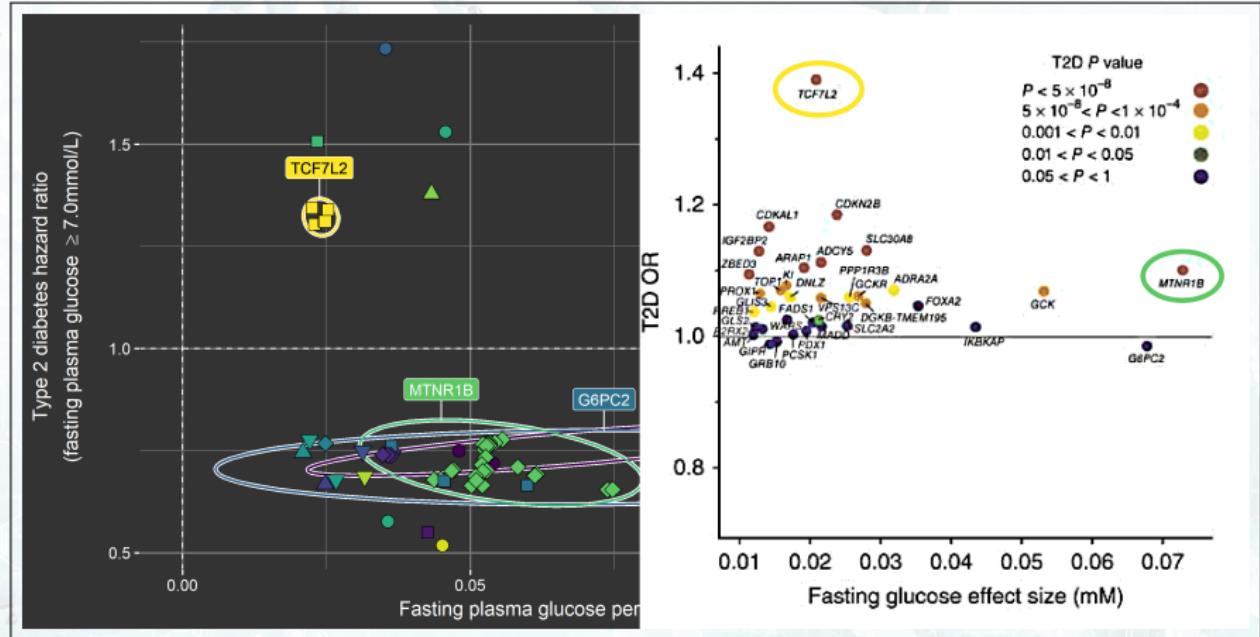
Trajectoires de la glycémie à jeun des individus D.E.S.I.R.

Application sur Données Réelles : D.E.S.I.R.



Taille d'effet β et Hazard Ratio respectivement pour la glycémie et le diabète de type 2. Gènes identifiés dans des études pangénomiques pour une association à la glycémie et/ou au diabète de type 2.

Application sur Données Réelles : D.E.S.I.R.



Taille d'effet β et Hazard Ratio respectivement pour la glycémie et le diabète de type 2. Gènes identifiés dans des études pangénomiques pour une association à la glycémie et/ou au diabète de type 2.

Application sur Données Réelles : D.E.S.I.R.

SNP (gène)	β (p-valeur)	
	JM (D.E.S.I.R.)	TS (D.E.S.I.R.)
rs10830963_G (MTNR1B)	3,25 ($3,6 \times 10^{-42}$)	3,52 ($2,7 \times 10^{-54}$)
rs17747324_C (TCF7L2)	3,17 ($8,9 \times 10^{-42}$)	3,39 ($2,2 \times 10^{-52}$)

Coefficients β estimés par JM et TS.

Application sur Données Réelles : D.E.S.I.R.

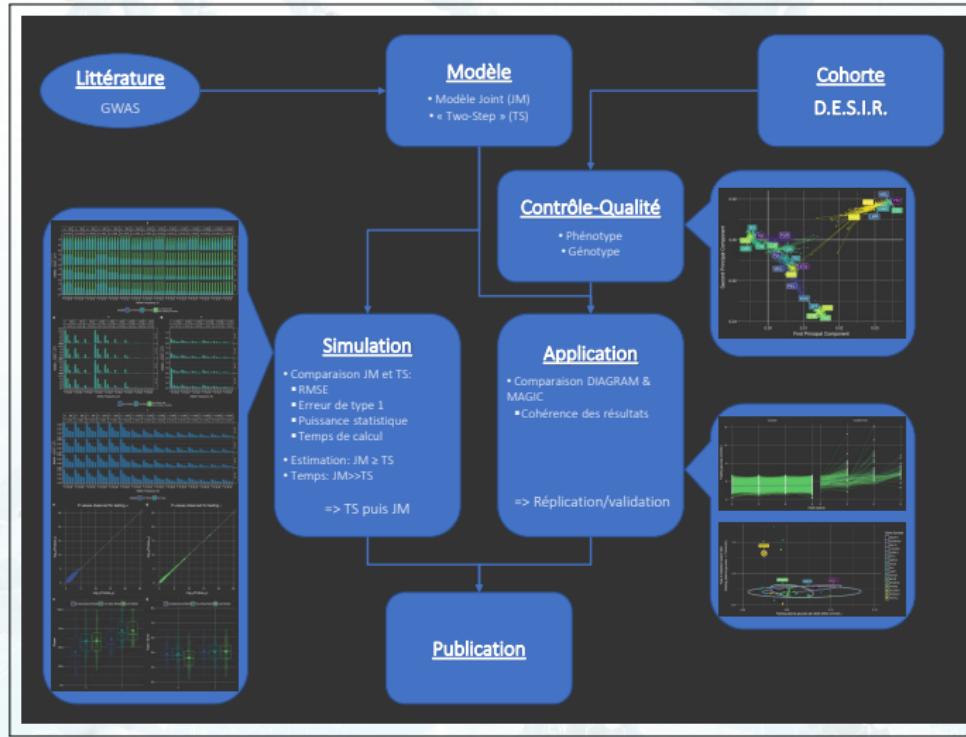
SNP (gène)	γ (p-valeur)		
	JM (D.E.S.I.R.)	TS (D.E.S.I.R.)	MAGIC
rs10830963_G (MTNR1B)	0,0991 ($1,3 \times 10^{-23}$)	0,0992 ($8,9 \times 10^{-24}$)	0,079 ($1,3 \times 10^{-68}$)
rs17747324_C (TCF7L2)	0,0229 ($3,0 \times 10^{-02}$)	0,0218 ($3,8 \times 10^{-02}$)	0,025 ($6,5 \times 10^{-08}$)

Coefficients γ estimés par JM et TS, et reportés dans le consortium MAGIC.

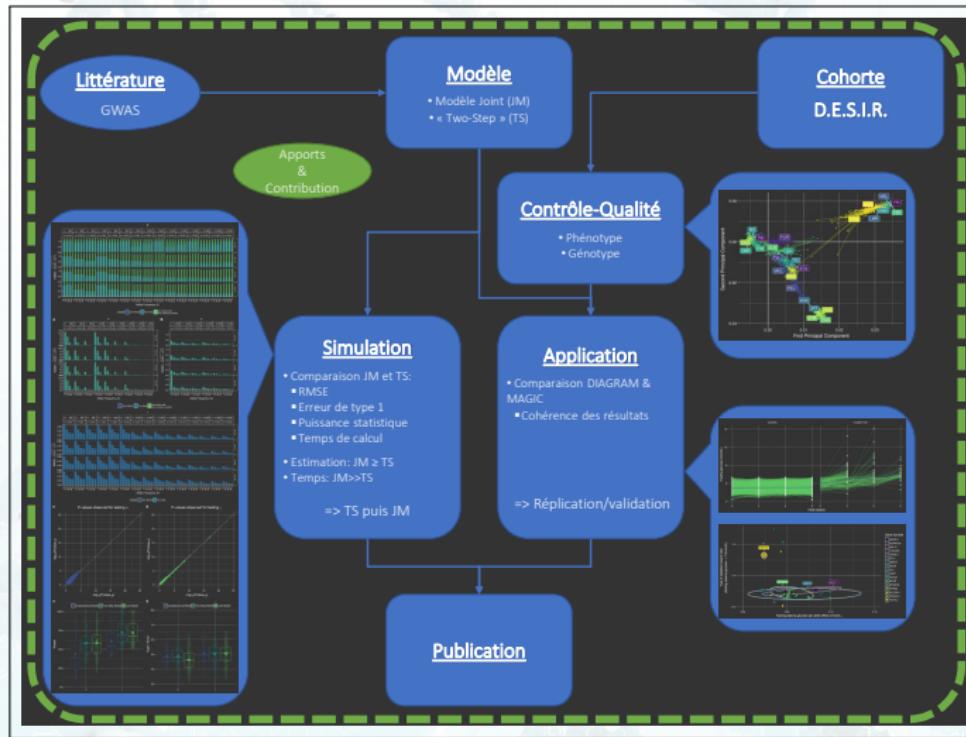
SNP (gène)	α (p-valeur)		
	JM (D.E.S.I.R.)	TS (D.E.S.I.R.)	DIAGRAM
rs10830963_G (MTNR1B)	-0,440 ($9,4 \times 10^{-04}$)	-0,443 ($5,0 \times 10^{-04}$)	0,104 ($7,3 \times 10^{-07}$)
rs17747324_C (TCF7L2)	0,265 ($4,1 \times 10^{-02}$)	0,284 ($2,2 \times 10^{-02}$)	0,358 ($8,5 \times 10^{-55}$)

Coefficients α estimés par JM et TS, et reportés dans le consortium DIAGRAM.

Conclusion : Modèle Joint



Conclusion : Modèle Joint



Au delà des Études Pangénomiques

Loci de Diabète et Sécrétion d'Insuline

L'Expression et l'Évaluation Fonctionnelle des Gènes de Susceptibilité au Diabète de Type 2 Identifient Quatre Nouveaux Gènes Contribuant à la Sécrétion d'Insuline Humaine

Fatou K Ndiaye^{1,*}, Ana Ortalli^{1,*}, **Mickaël Canouil^{1,†}**, Marlène Huyvaert¹, Clara Salazar-Cardozo¹, Cécile Lecoer¹, Marie Verbanck¹, Valérie Pawlowski¹, Raphaël Boutry¹, Emmanuelle Durand¹, Landry Rabearivelo¹, Olivier Sand¹, Lorella Marselli², Julie Kerr-Conte³, Vikash Chandra⁴, Raphaël Scharfmann⁴, Odile Poulain-Godefroy¹, Piero Marchetti², François Pattou³, Amar Abderrahmani^{1,5}, Philippe Froguel^{1,5,†} & Amélie Bonnefond^{1,5,†}

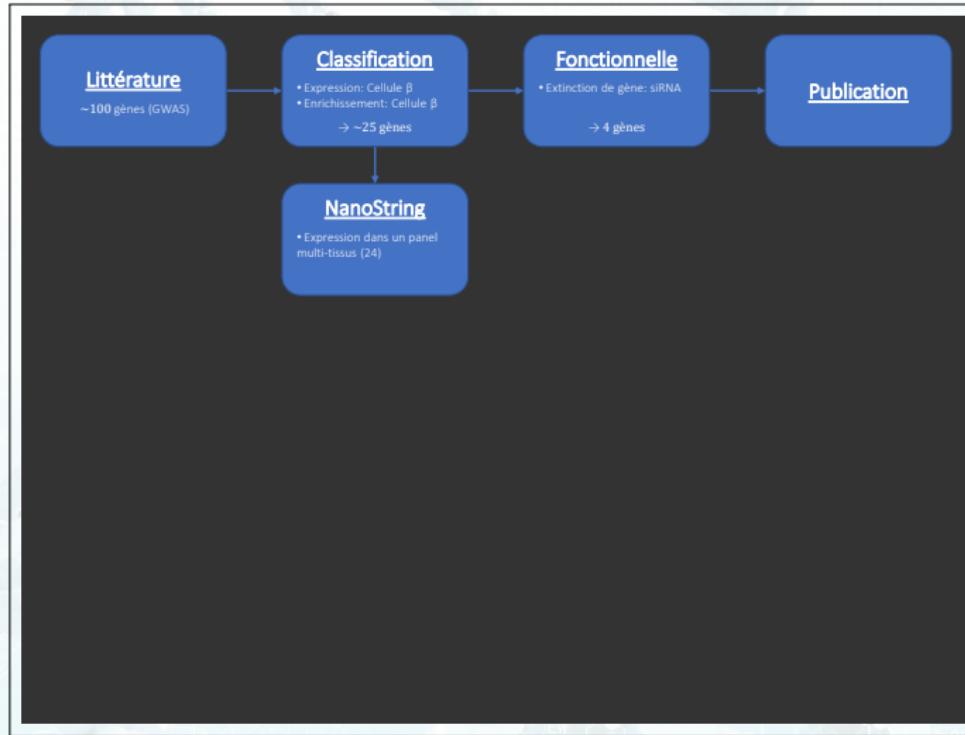
* Co-premier auteurs.

† Co-dernier auteurs.

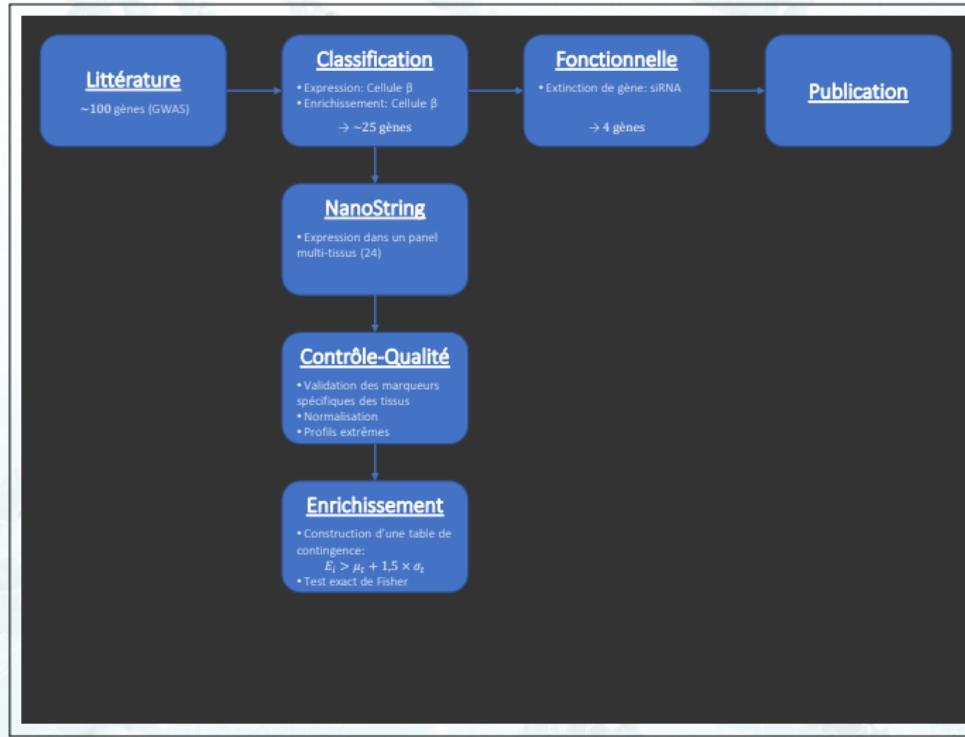
Loci de Diabète et Sécrétion d'Insuline



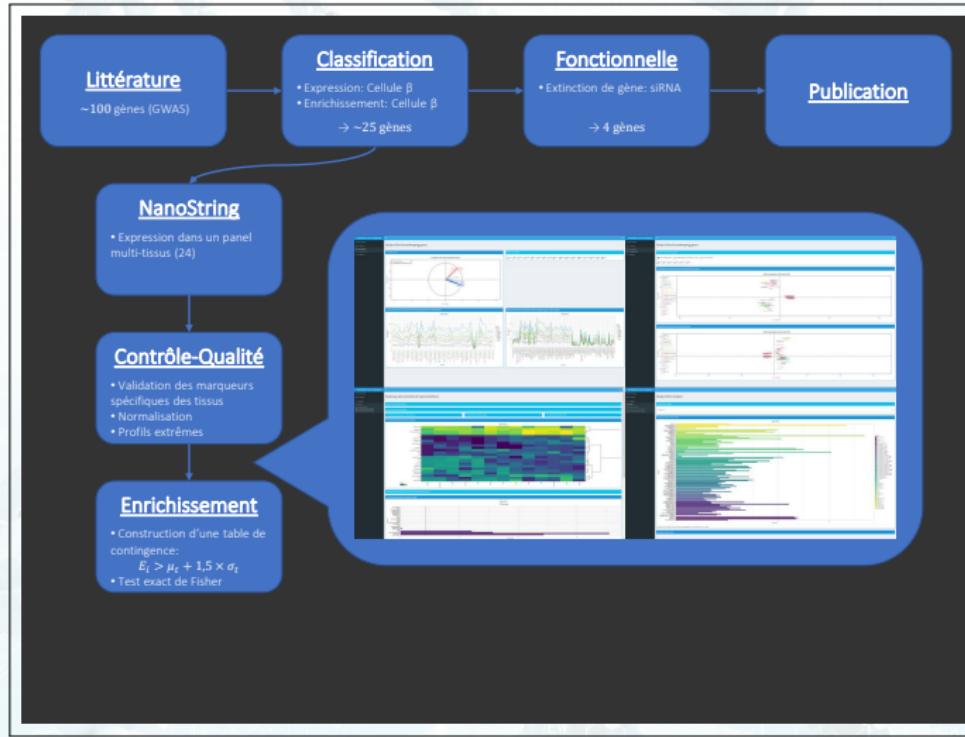
Loci de Diabète et Sécrétion d'Insuline



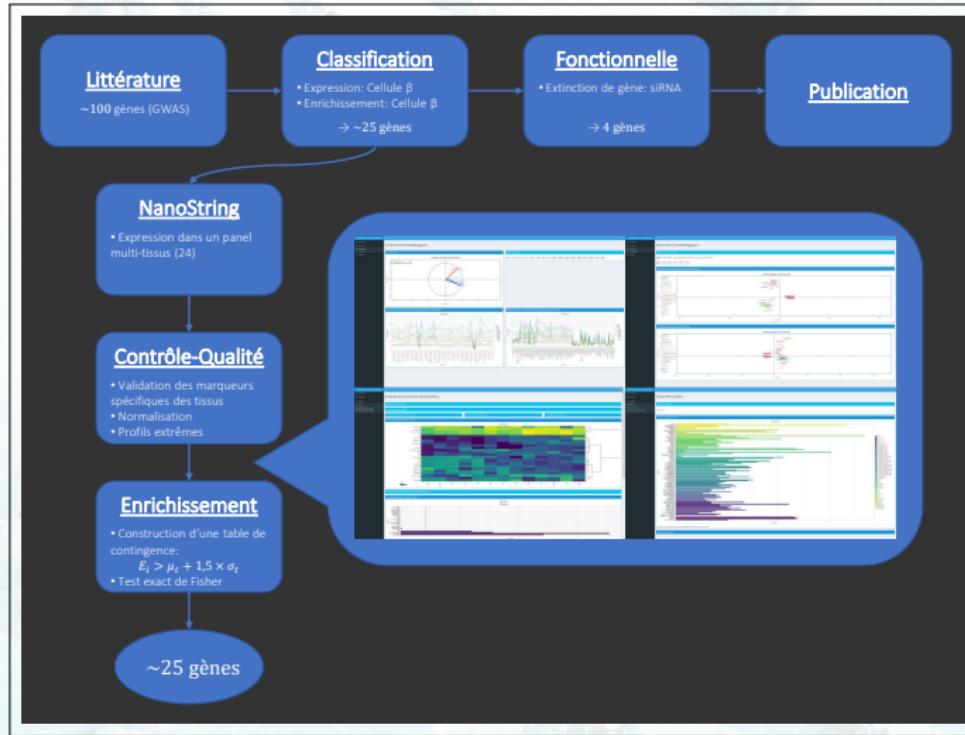
Loci de Diabète et Sécrétion d'Insuline



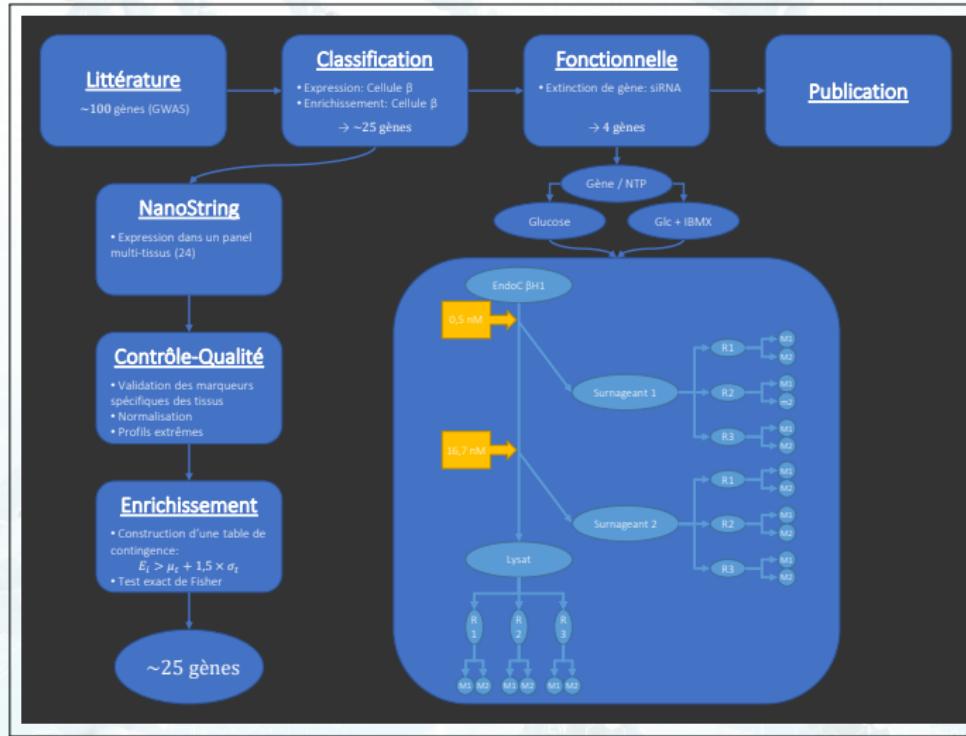
Loci de Diabète et Sécrétion d'Insuline



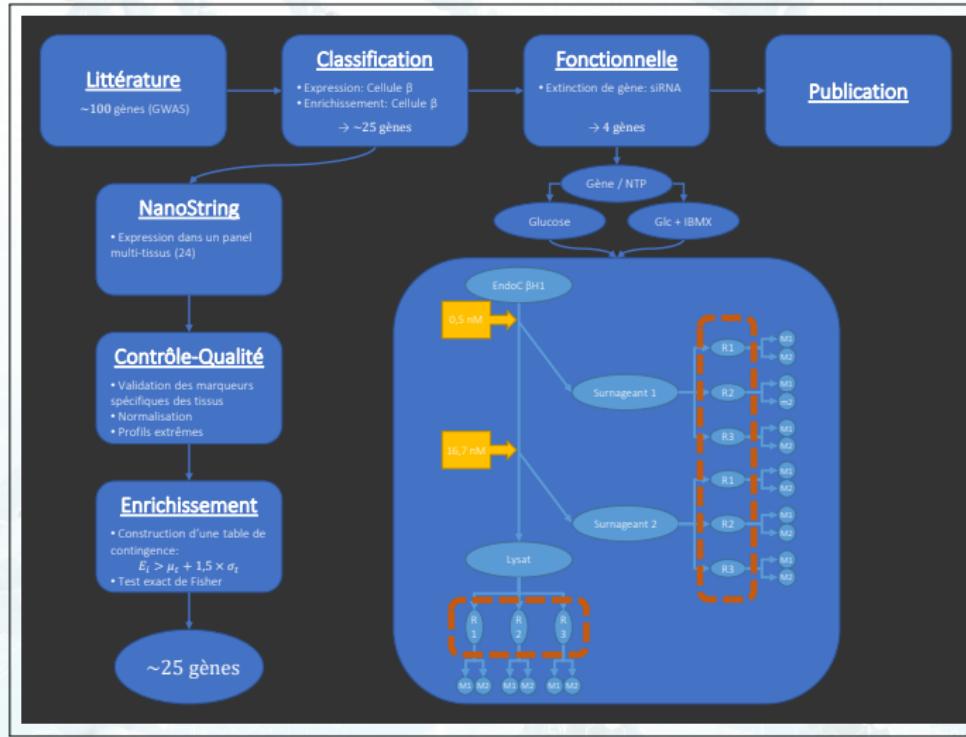
Loci de Diabète et Sécrétion d'Insuline



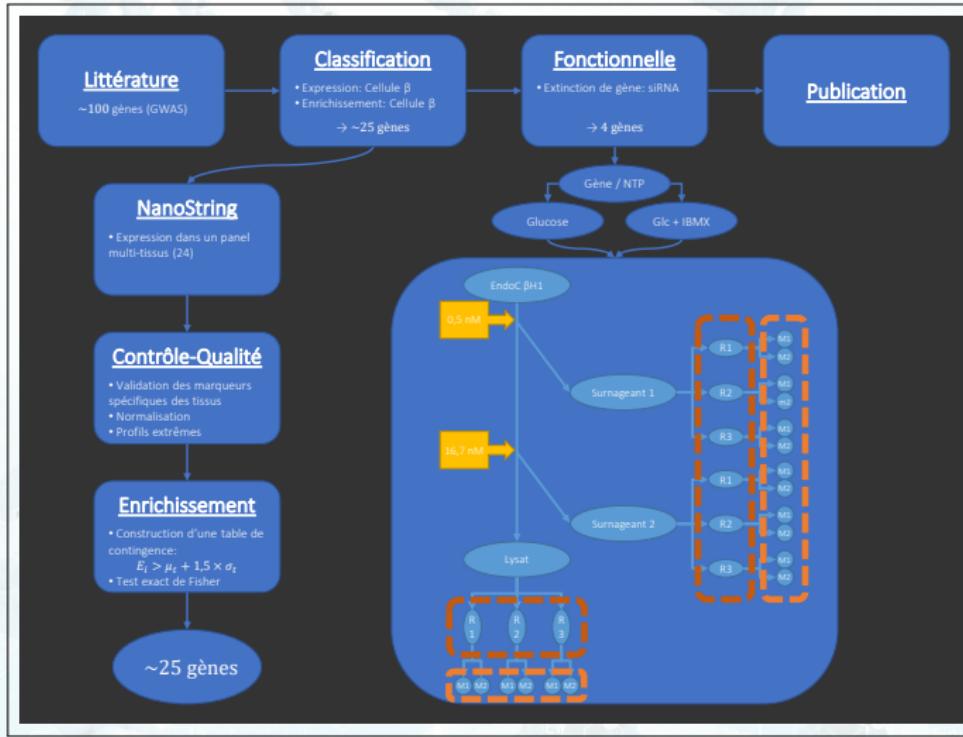
Loci de Diabète et Sécrétion d'Insuline



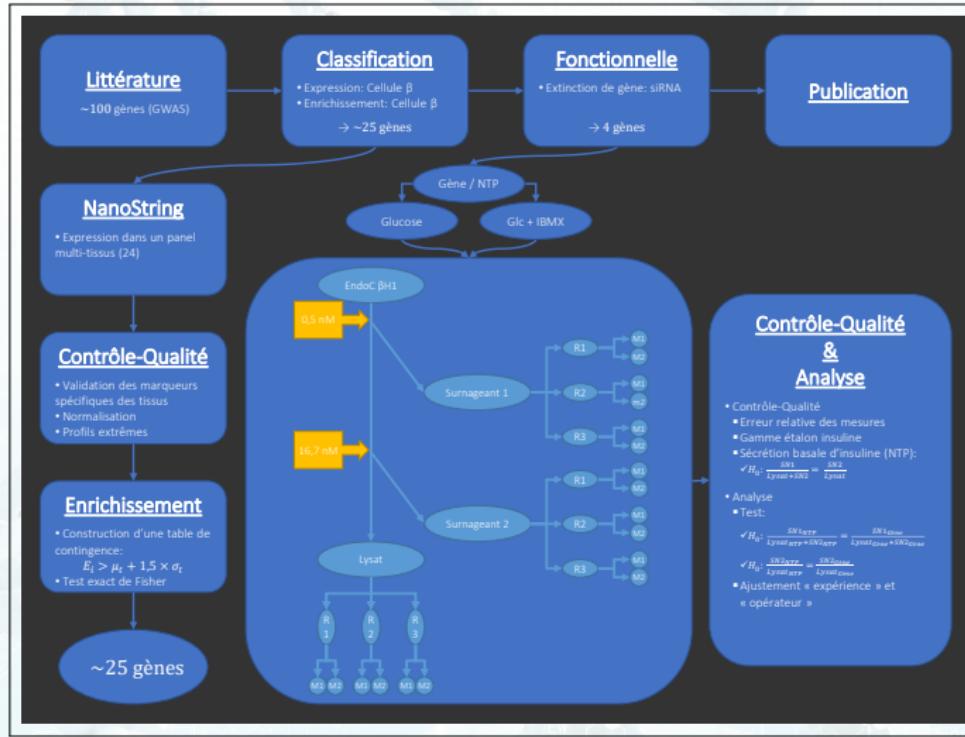
Loci de Diabète et Sécrétion d'Insuline



Loci de Diabète et Sécrétion d'Insuline



Loci de Diabète et Sécrétion d'Insuline



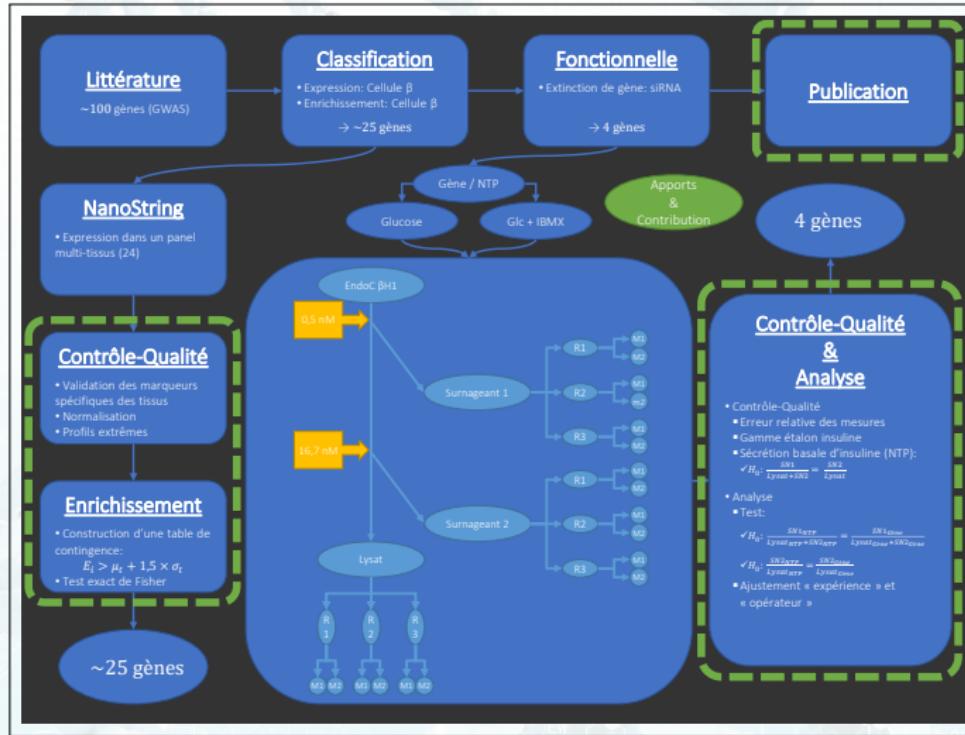
Loci de Diabète et Sécrétion d'Insuline



Loci de Diabète et Sécrétion d'Insuline



Loci de Diabète et Sécrétion d'Insuline



Méthylation et Expression : le Cas de PDGFA

La Surexpression Hépatique de PDGF-AA Affaiblit la Signalisation de l'Insuline dans le Diabète

Amar Abderrahmani^{1,2*}, Loïc Yengo^{1*}, Robert Caiazzo^{3*}, Mickaël Canouil^{1*}, Stéphane Cauchi¹, Violeta Raverdy², Valérie Plaisance¹, Stéphane Lobbens¹, Julie Maillet¹, Laure Rolland¹, Raphael Boutry¹, Maxime Kwapich¹, Mathie Tenenbaum¹, Julien Bricambert¹, Sophie Saussenthaler⁴, Elodie Anthony⁵, Pooja Jha⁶, Julien Derop¹, Olivier Sand¹, Landry Rabearivelo¹, Audrey Leloir¹, Marie Pigeyre², Martine Daujat-Chavanieu⁷, Sabine Gerbal-Chaloin⁷, Tasnim Dayeh⁸, Guillaume Lassailly², Philippe Mathurin⁹, Bart Staels¹⁰, Johan Auwerx⁵, Annette Schürmann⁴, Catherine Postic⁵, Clemens Schafmayer¹¹, Jochen Hampe¹², Amélie Bonnefond^{1,2}, François Pattou^{3†} & Philippe Froguel^{1,2†}

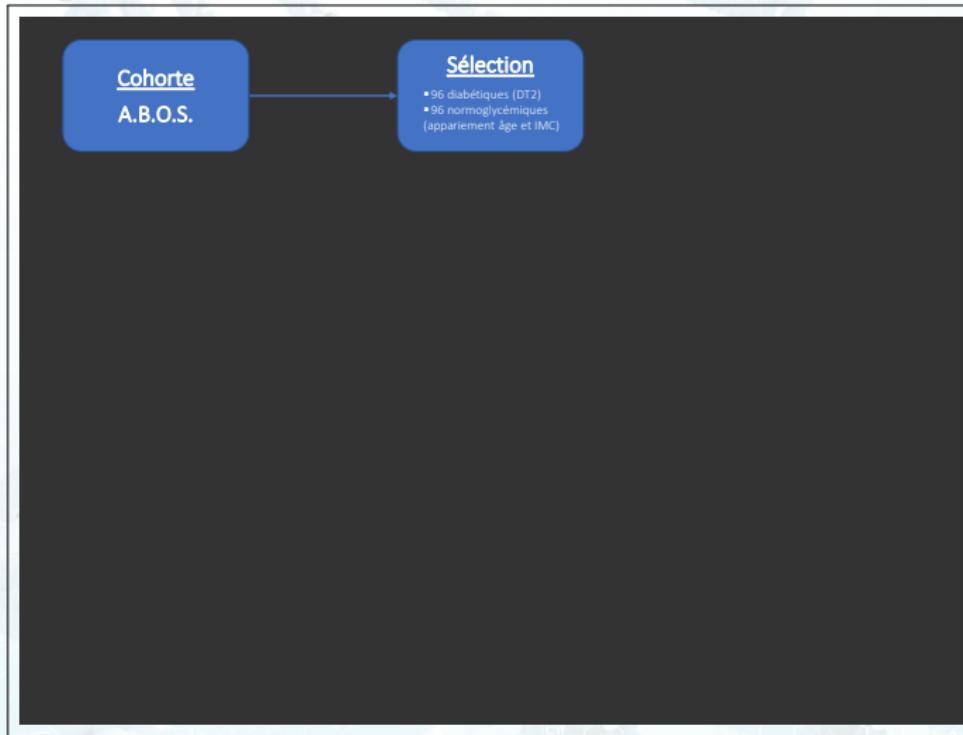
* Co-premier auteurs.

† Co-dernier auteurs.

Méthylation et Expression : le Cas de PDGFA

Cohorte
A.B.O.S.

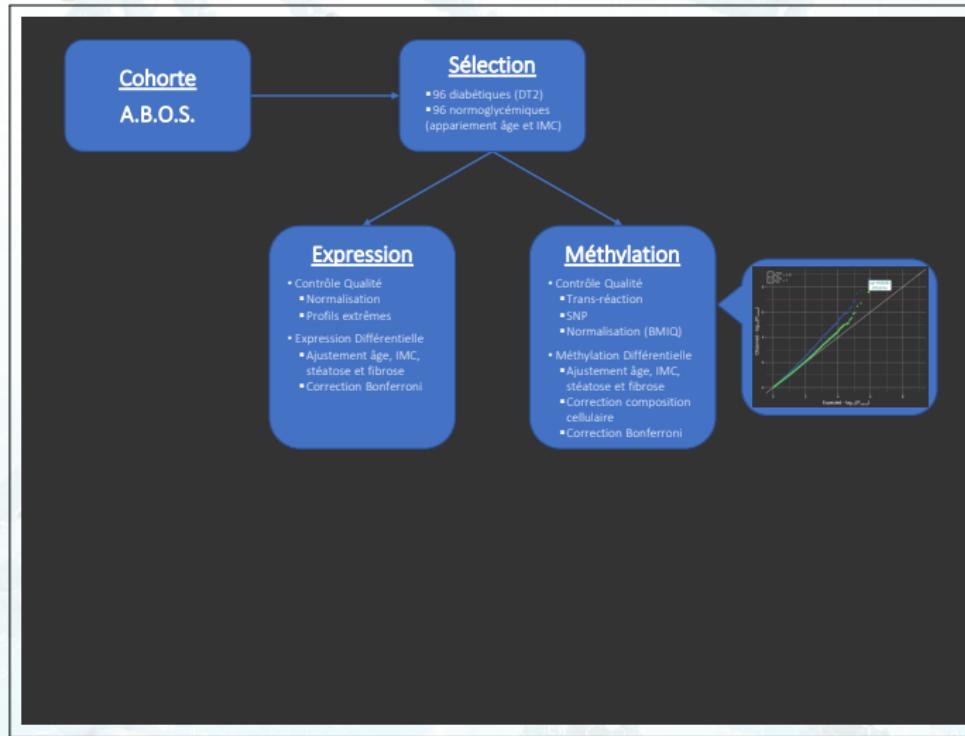
Méthylation et Expression : le Cas de PDGFA



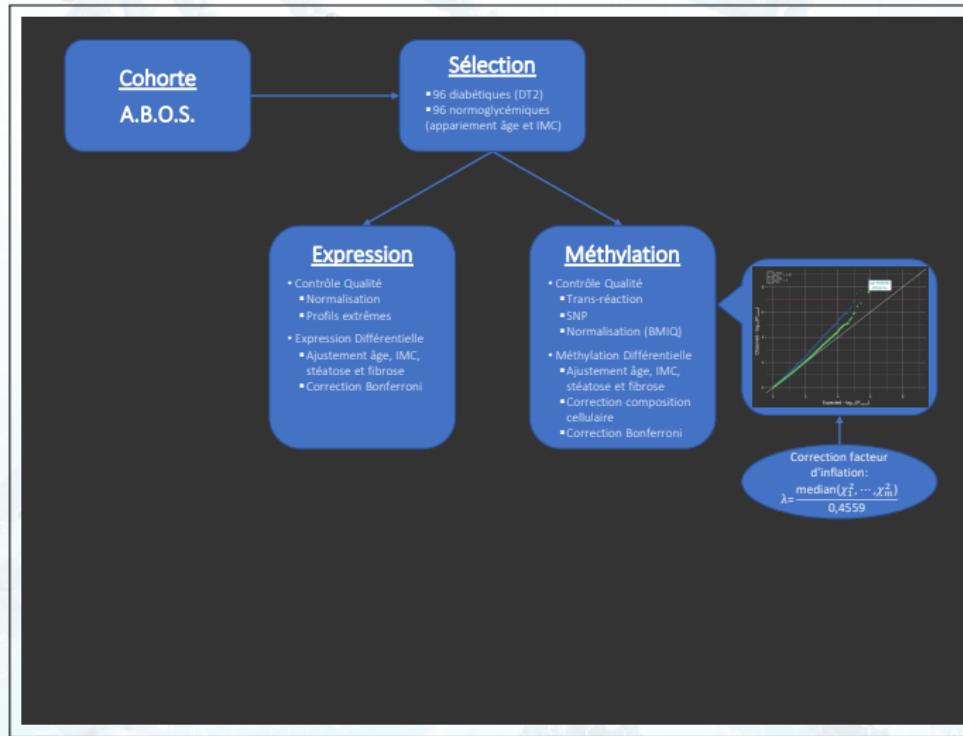
Méthylation et Expression : le Cas de PDGFA



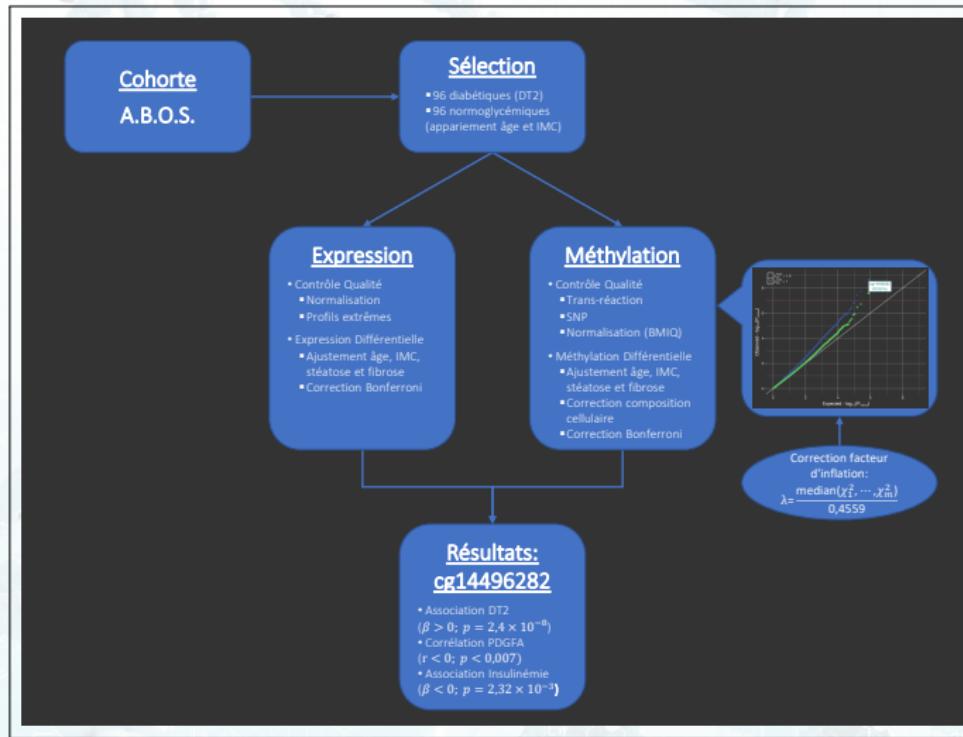
Méthylation et Expression : le Cas de PDGFA



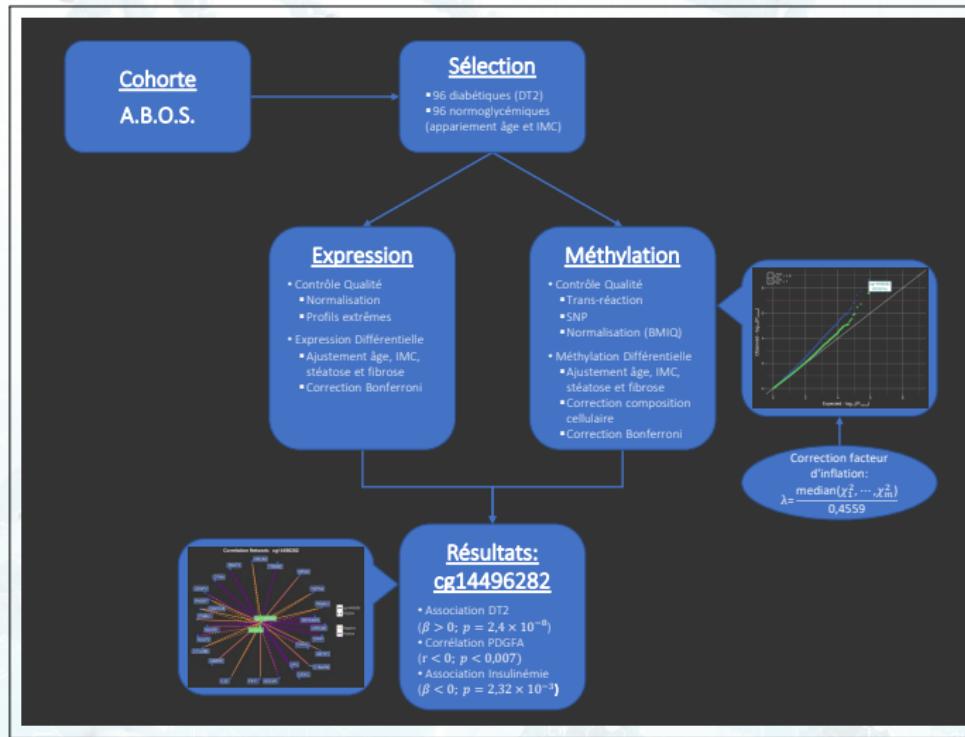
Méthylation et Expression : le Cas de PDGFA



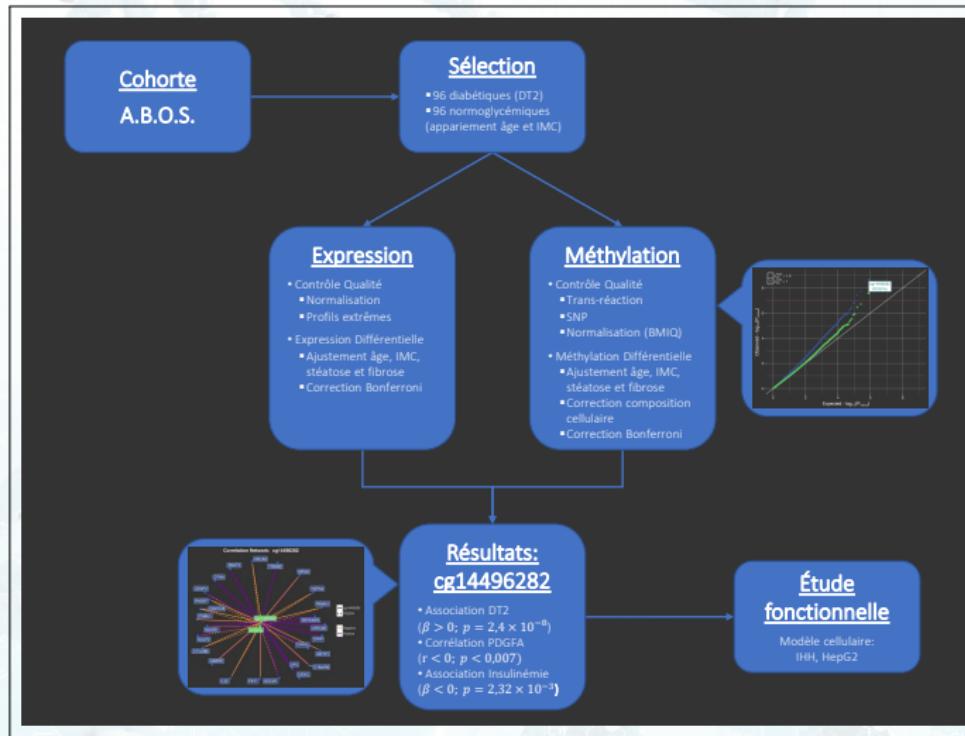
Méthylation et Expression : le Cas de PDGFA



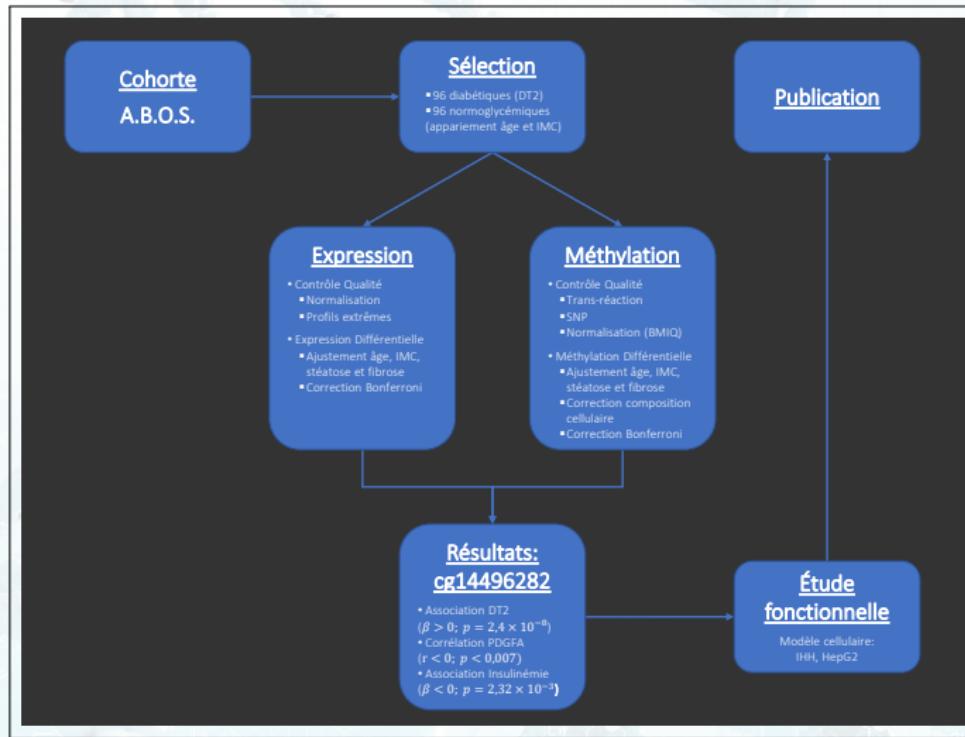
Méthylation et Expression : le Cas de PDGFA



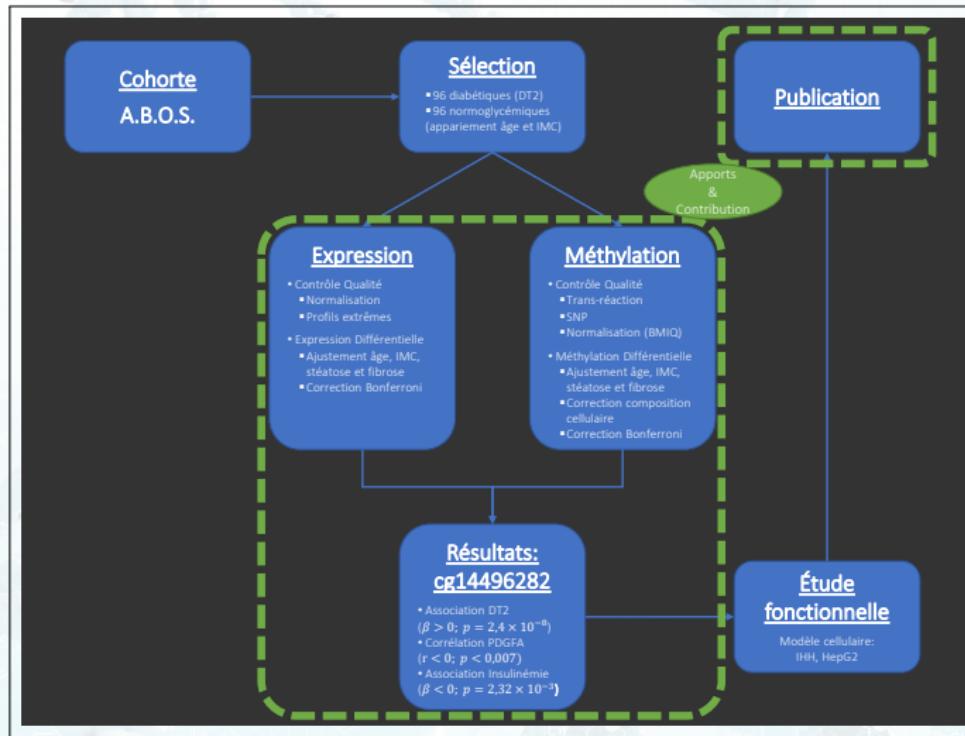
Méthylation et Expression : le Cas de PDGFA



Méthylation et Expression : le Cas de PDGFA



Méthylation et Expression : le Cas de PDGFA



Méthylation et Expression : le Bisphénol et ses Substituants

L'Exposition à Faible Dose aux Bisphénols A, F et S des Adipocytes Primaires Humains Modifie les Profils d'ARN Codant et Non-Codant

Marie Verbanck^{1,*}, Mickaël Canouil^{1,*}, Audrey Leloir¹, Véronique Dhennin¹, Xavier Coumoul², Loïc Yengo¹, Philippe Froguel^{1,3,†} & Odile Poulain-

Codefroy^{1,‡}

* Co-premier auteurs.

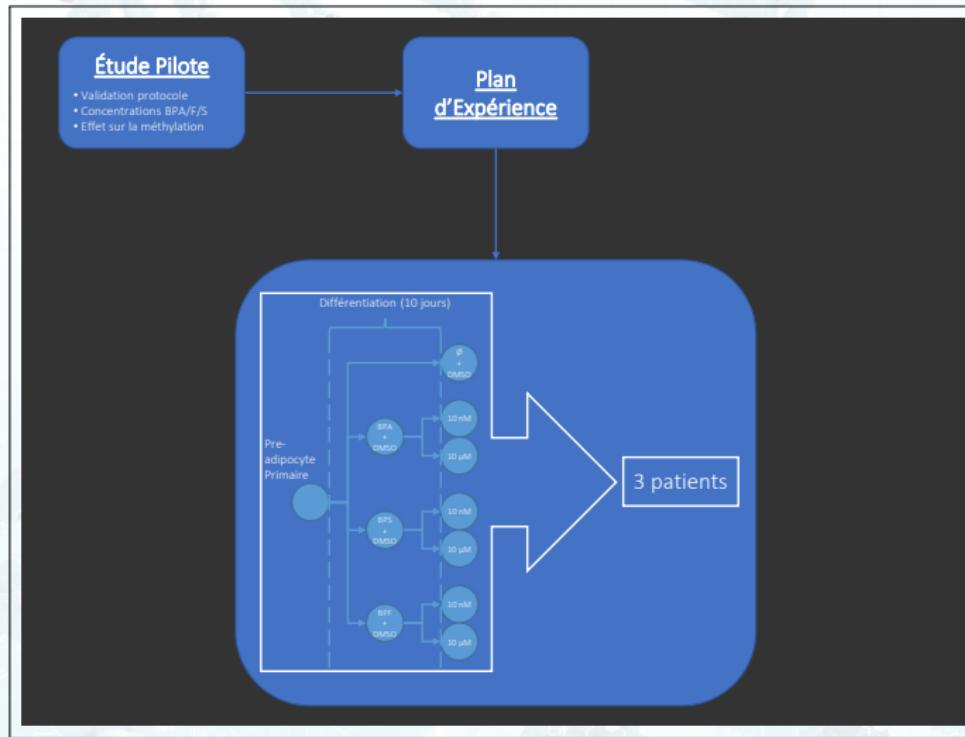
† Co-dernier auteurs.

Méthylation et Expression : le Bisphénol et ses Substituants

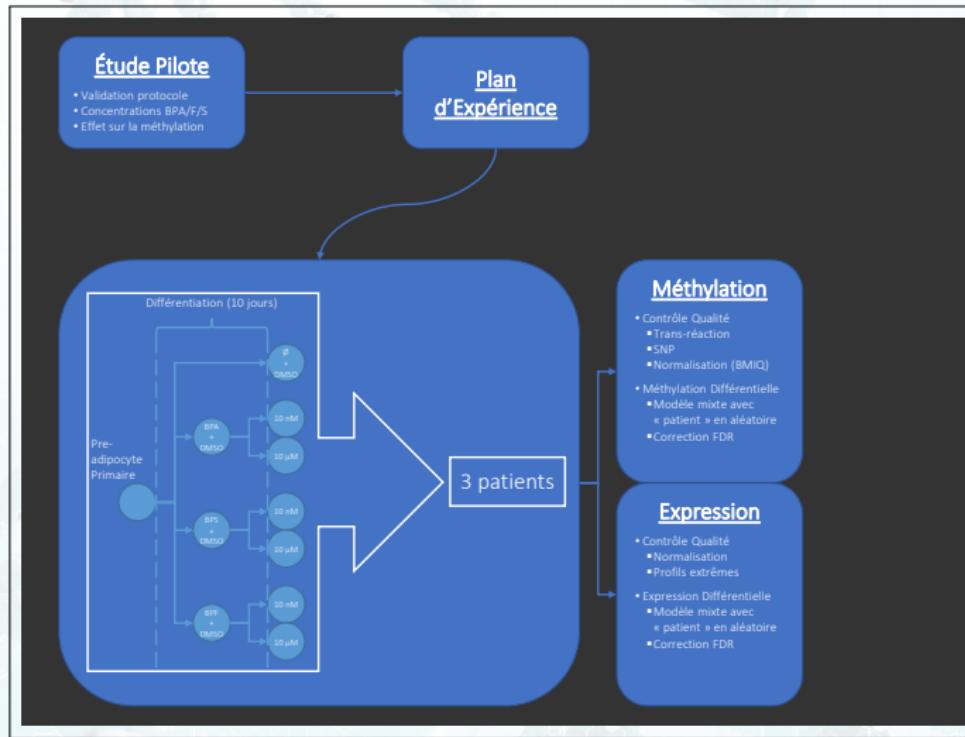
Étude Pilote

- Validation protocole
- Concentrations BPA/F/S
- Effet sur la méthylation

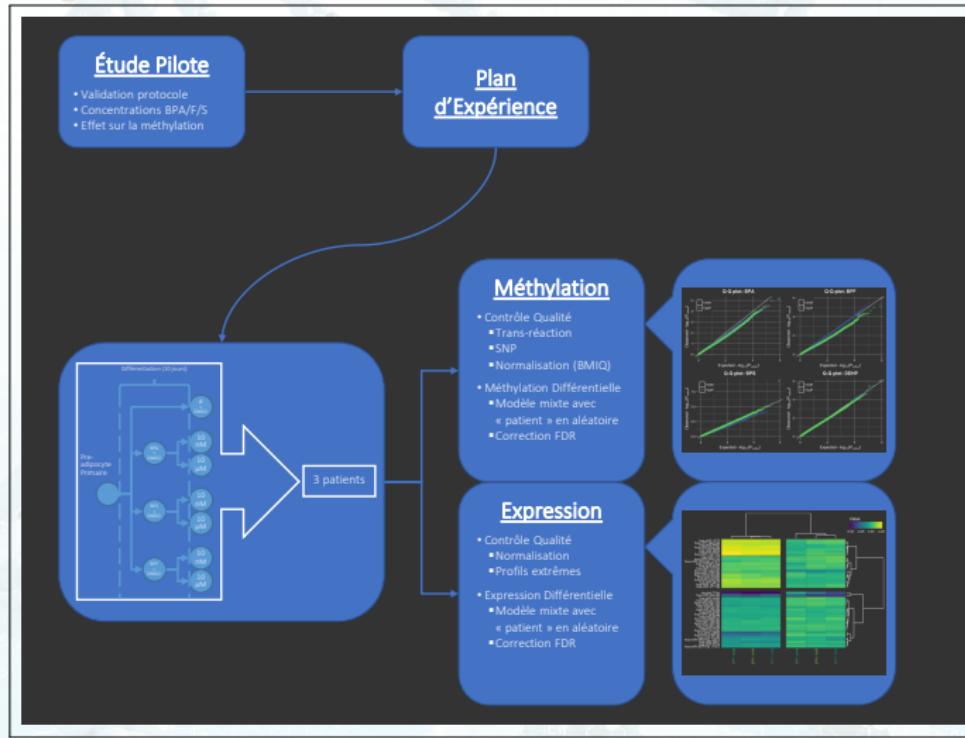
Méthylation et Expression : le Bisphénol et ses Substituants



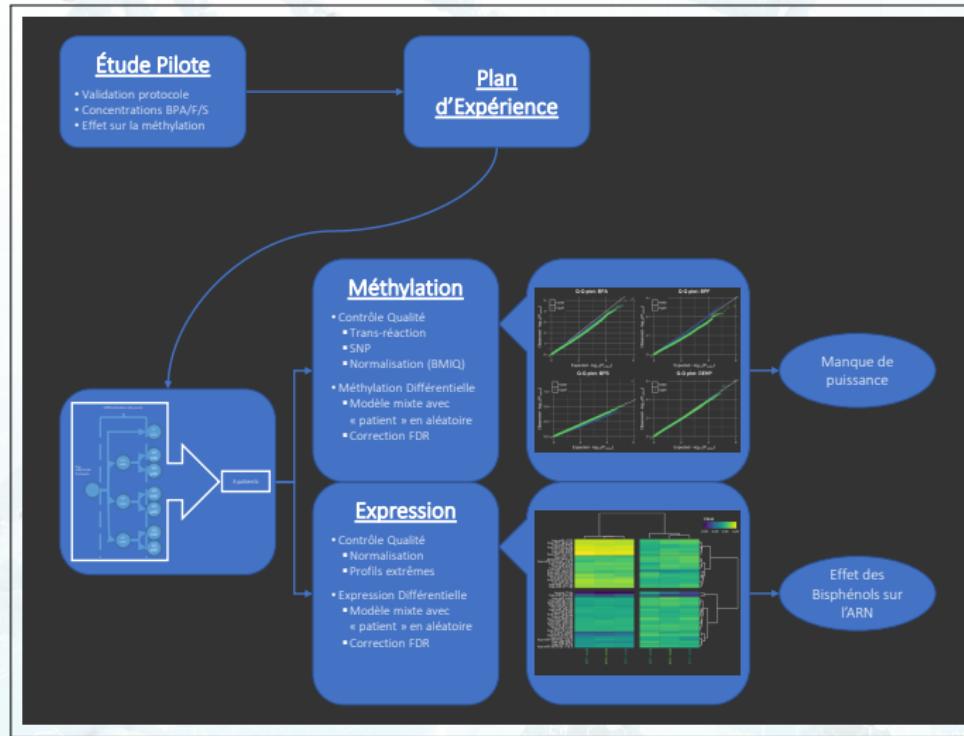
Méthylation et Expression : le Bisphénol et ses Substituants



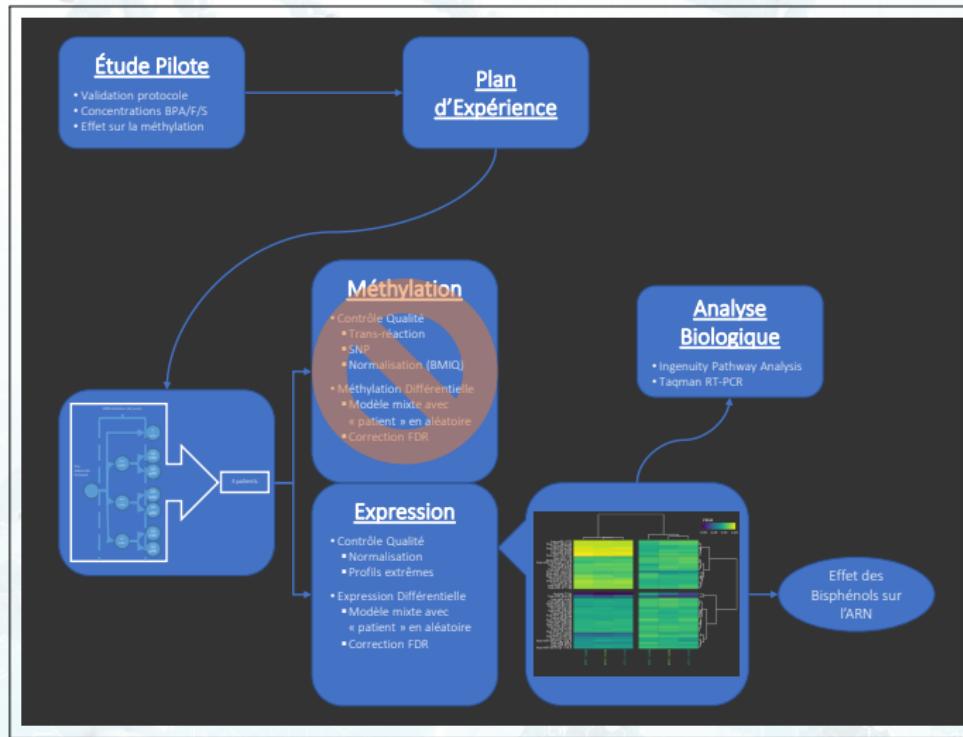
Méthylation et Expression : le Bisphénol et ses Substituants



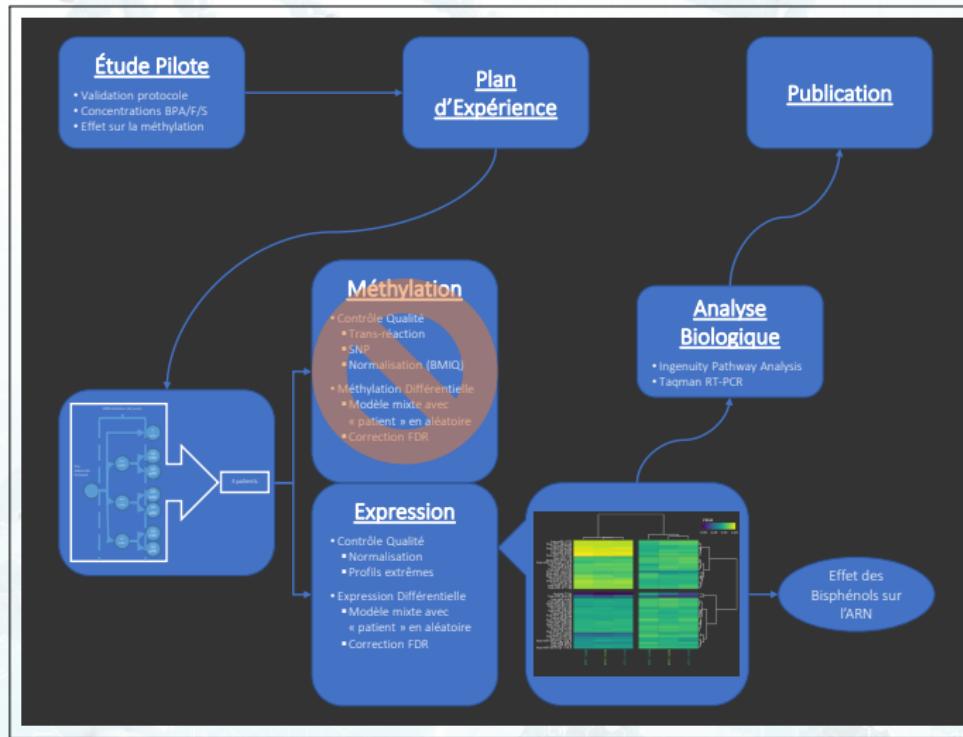
Méthylation et Expression : le Bisphénol et ses Substituants



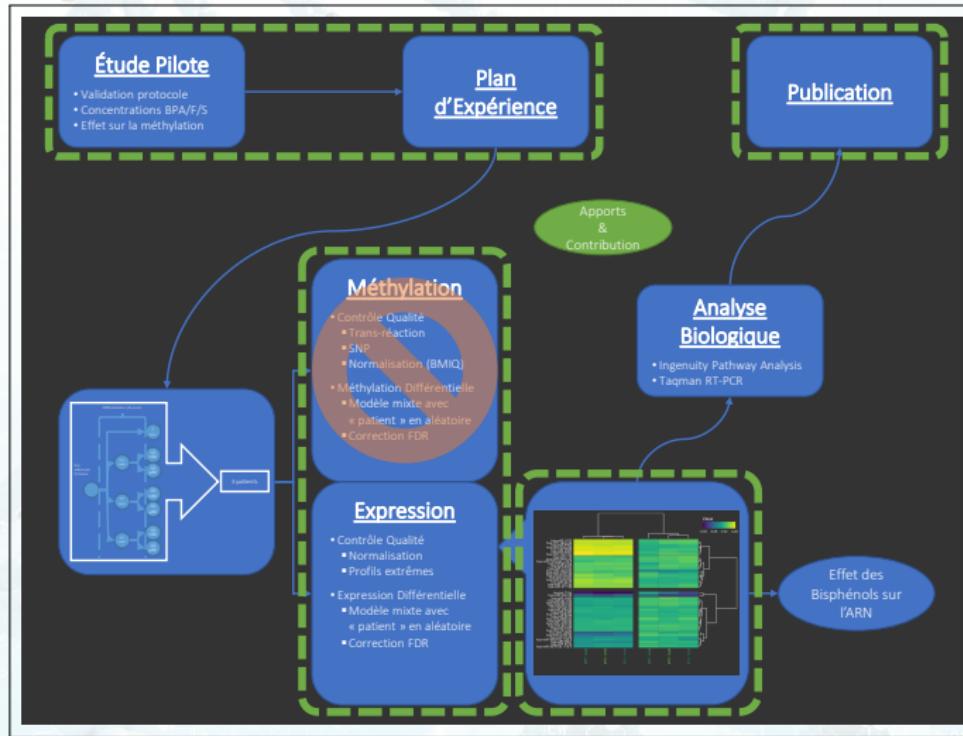
Méthylation et Expression : le Bisphénol et ses Substituants



Méthylation et Expression : le Bisphénol et ses Substituants



Méthylation et Expression : le Bisphénol et ses Substituants



Apports & Perspectives

Apports & Perspectives

- Nouvelle approche de modélisation dans la génétique et les études fonctionnelles.

Apports & Perspectives

- Nouvelle approche de modélisation dans la génétique et les études fonctionnelles.
 - ⇒ RéPLICATION des résultats et application à d'autres cohortes (Consortium RHAPSODY)

Apports & Perspectives

- Nouvelle approche de modélisation dans la génétique et les études fonctionnelles.
 - ⇒ RéPLICATION des résultats et application à d'autres cohortes (Consortium RHAPSODY)
- CrÉATION d'une application Shiny pour l'étude et l'analyse d'expÉrience cellulaire.

Apports & Perspectives

- Nouvelle approche de modélisation dans la génétique et les études fonctionnelles.
 - ⇒ RéPLICATION des résultats et application à d'autres cohortes (Consortium RHAPSODY)
- CrÉATION d'une application Shiny pour l'étude et l'analyse d'expÉrience cellulaire.
 - ⇒ GÉNÉRALISATION/diffusion sous forme d'une extension R et d'une publication

Apports & Perspectives

- Nouvelle approche de modélisation dans la génétique et les études fonctionnelles.
 - ⇒ RéPLICATION des résultats et application à d'autres cohortes (Consortium RHAPSODY)
- CrÉATION d'une application Shiny pour l'étude et l'analyse d'expÉrience cellulaire.
 - ⇒ GÉNÉRALISATION/diffusion sous forme d'une extension R et d'une publication
- Participation à l'élaboration, l'analyse et la diffusion de travaux de recherches.

Apports & Perspectives

- Nouvelle approche de modélisation dans la génétique et les études fonctionnelles.
 - ⇒ RéPLICATION des résultats et application à d'autres cohortes (Consortium RHAPSODY)
- CrÉATION d'une application Shiny pour l'étude et l'analyse d'expÉrience cellulaire.
 - ⇒ GÉNÉRALISATION/diffusion sous forme d'une extension R et d'une publication
- Participation à l'élaboration, l'analyse et la diffusion de travaux de recherches.
 - ⇒ Élaboration de processus de contrÔLE-qualité et gÉNÉRATION de rapport "automatisé" pour les donnÉES mÉthylomiques et gÉnomiques

Merci pour votre attention !



Remerciements

Chislain Rocheleau

Loïc Yengo

Philippe Froguel

CNRS UMR 8199

Financement:

Ce travail a été soutenu par des subventions de
"European Genomic Institute for Diabetes" (E.G.I.D., ANR-10-LABX-46),
"European Commission",
"Société Francophone du Diabète" and "Lilly".



Inférence Statistique : EM

Espérance-Maximisation

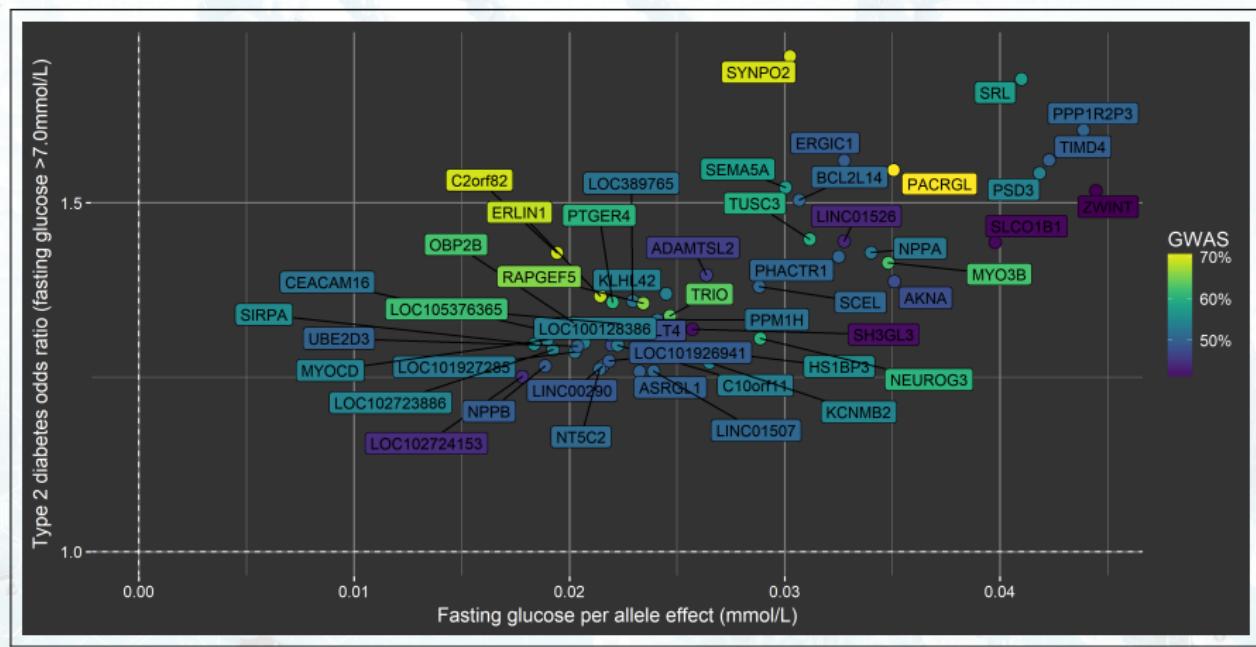
Espérance :

$$\begin{aligned} E\{g(b_i)|Y_i, \tilde{T}_i, \Delta_i, \Psi^{(m)}\} &= \int g(b_i) f(b_i|Y_i, \tilde{T}_i, \Delta_i, \Psi^{(m)}) db_i \\ &= \frac{\int (g(b_i) f(b_i, Y_i, \tilde{T}_i, \Delta_i | \Psi^{(m)}) db_i}{\int f(Y_i, \tilde{T}_i, \Delta_i | b_i, \Psi^{(m)}) f(b_i | \Psi^{(m)}) db_i} \end{aligned}$$

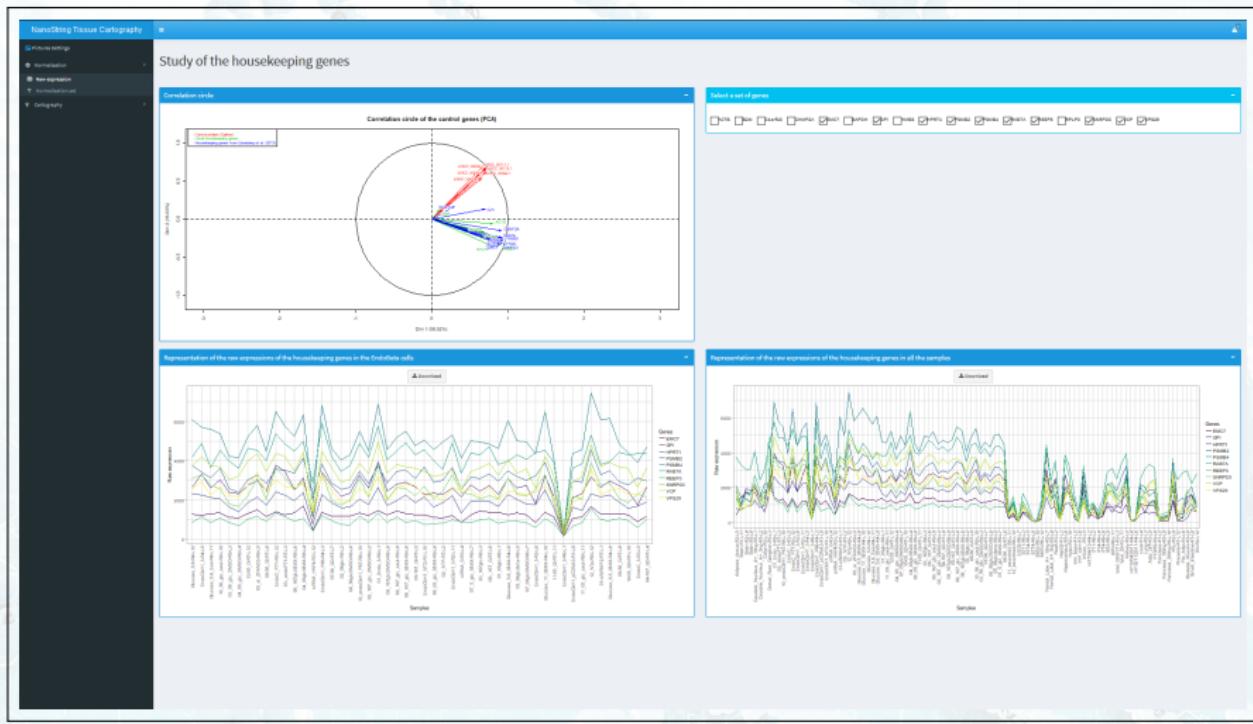
Maximisation :

$$\begin{aligned} \Psi^{(m+1)} &= \operatorname{argmax}_{\Psi} \mathcal{Q}(\Psi; \Psi^{(m)}) \\ &= \operatorname{argmax}_{\Psi} E_{b|Y, \tilde{T}, \Delta, \Psi^{(m)}} (\log L(\Psi; Y, \tilde{T}, \Delta, b)) \end{aligned}$$

Modèle Joint : Application



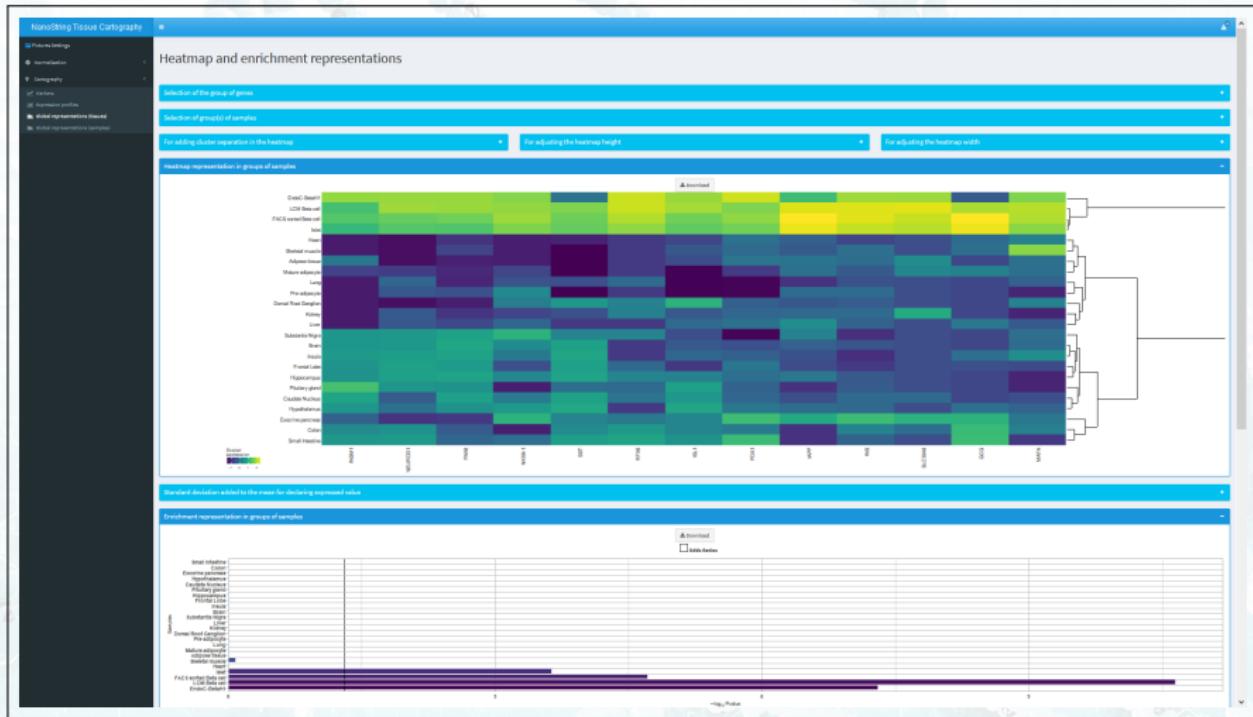
Application Shiny Enrichissement



Application Shiny Enrichissement



Application Shiny Enrichissement



Application Shiny EndoC

EndoC β (v1.1.3)

Amélie Bonnefond (Principal investigator)
Mickaël Canouil (Biostatistician)

Download Data
 Pictures Settings
 Import Excel Files

Technical Error Threshold
0 - 01 02 03 04 05 06 07 08 09 10

Quality Control
» Summary
» Technical Error (OO)
» Technical Error Correction

Analysis

File Input

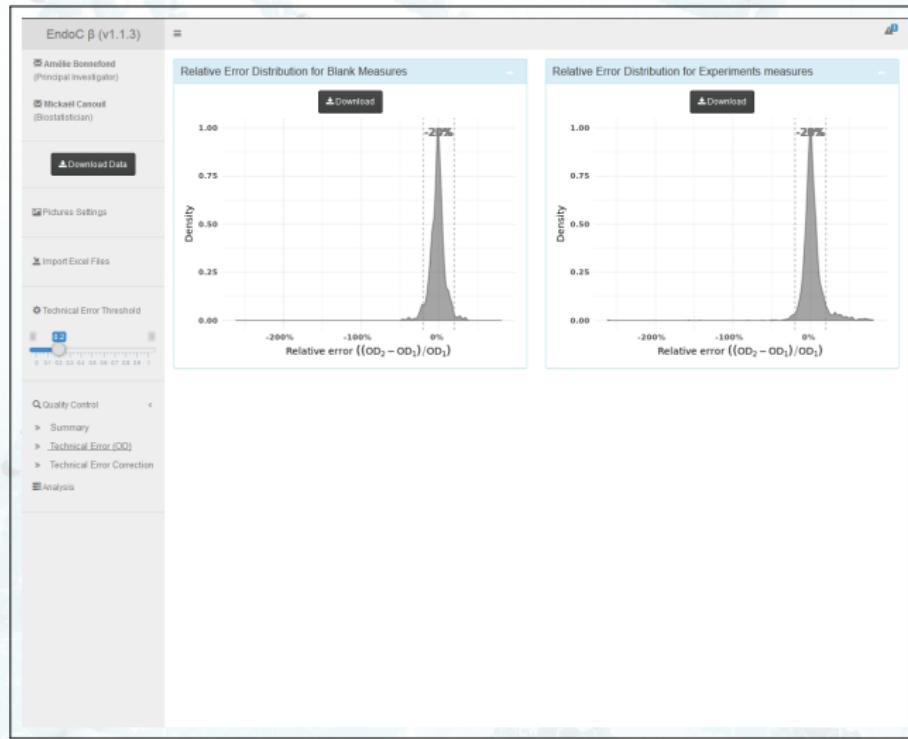
Browse... File Management
CT_stat_AnA_03022016_st24.xlsx Delete File

61 File(s) 187 Experiment(s) 146 / 41 Clear / Warning(s)

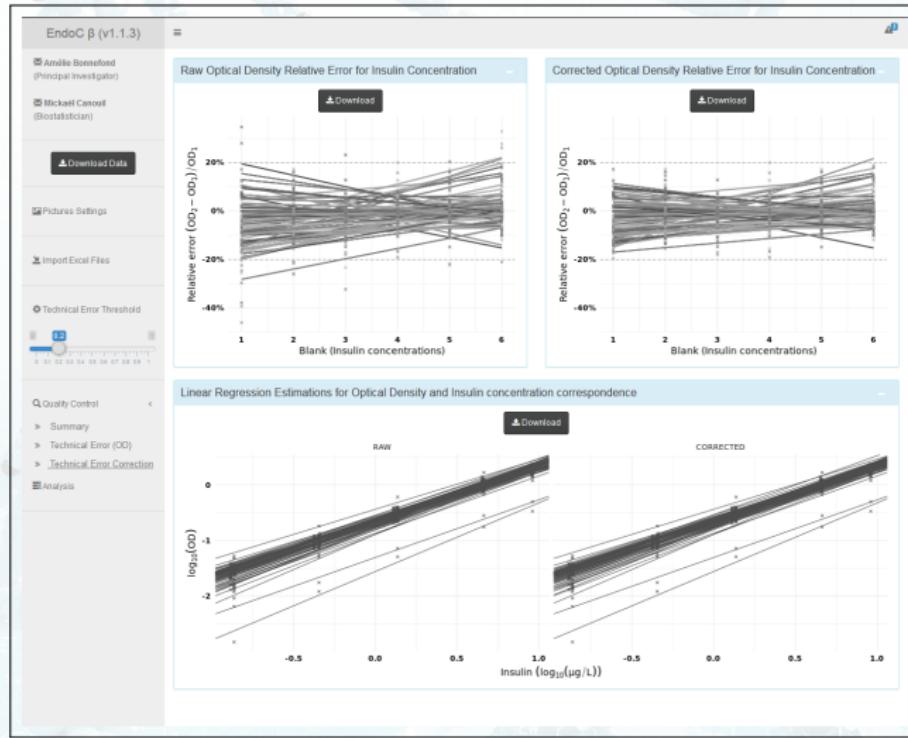
Files Already Imported

Name	Sheets	Operator	Check
CT_stat_AnA_03022016_st24.xlsx	3	Ana	2
CT_stat_AnA_04022016_st25.xlsx	3	Ana	3
CT_stat_AnA_06102016_st21.xlsx	3	Ana	3
CT_stat_AnA_08032016_st28.xlsx	3	Ana	2
CT_stat_AnA_09022016_st25.xlsx	3	Ana	0
CT_stat_AnA_16022016_st26.xlsx	3	Ana	2
CT_stat_AnA_20102015_st22.xlsx	3	Ana	3
CT_stat_AnA_21072015_st15.xlsx	3	Ana	2
CT_stat_AnA_22092015_st20.xlsx	3	Ana	2
CT_stat_AnA_23022016_st27.xlsx	3	Ana	3
CT_stat_AnA_24112015_st23.xlsx	2	Ana	2
CT_stat_Clara_02032016.xlsx	3	Clara	3
CT_stat_Clara_06102015.xlsx	3	Clara	2
CT_stat_Clara_08022016.xlsx	2	Clara	0
CT_stat_Clara_09052016.xlsx	2	Clara	1
CT_stat_Clara_10052016.xlsx	4	Clara	4
CT_stat_Clara_10402016.xlsx	2	Clara	2
CT_stat_Clara_14122015.xlsx	3	Clara	3
CT_stat_Clara_21072015.xlsx	2	Clara	2

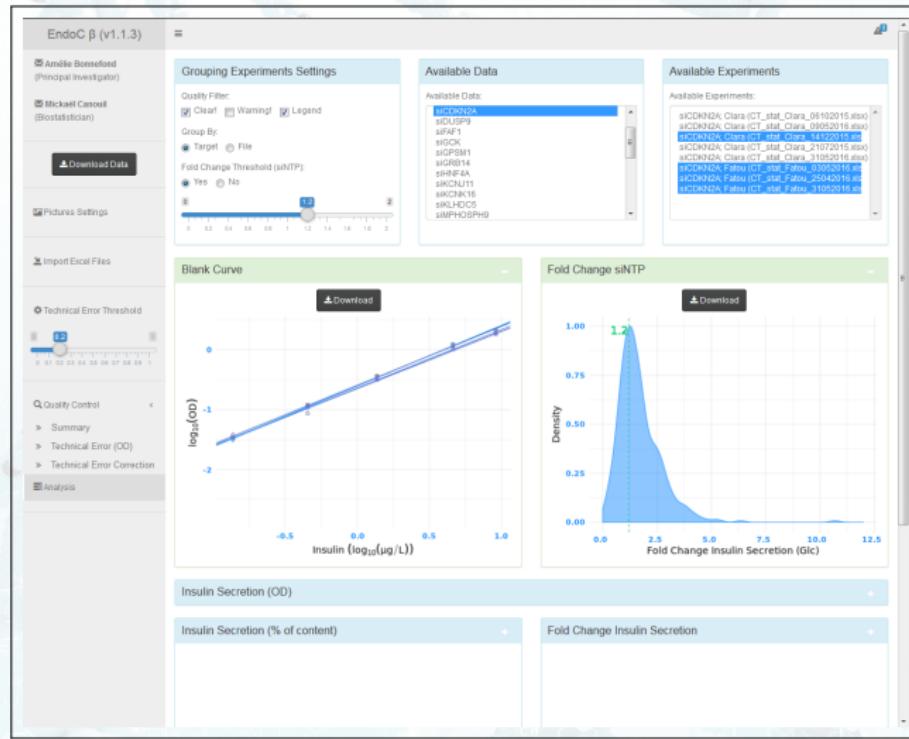
Application Shiny EndoC



Application Shiny EndoC



Application Shiny EndoC



Application Shiny EndoC

