

Identification des éléments clefs du métabolisme des lipides et de leurs régulateurs

Pierre BLAVY : ASC



12 mars 2010

Contexte multidisciplinaire



**Biologie,
métabolisme
génétique et
physiologie**

UMR SENAH : Systèmes d'Elevage
Nutrition Animale et Humaine
(Département PHASE, INRA)
Florence Gondret, Jaap Van Milgen

UMR GA : Génétique Animale
(Département Génétique Animale, INRA)
Sandrine Lagarrigue

UR66 : Pharmacologie-Toxicologie
(Département Santé Animale, INRA)
Hervé Guillou, Pascal Martin



**Mathématiques
et modélisation**

Equipe SYMBIOSE : Systèmes et
modèles biologiques, bioinformatique
et séquences (IRISA, INRIA)
Anne Siegel, Ovidiu Radulescu

Plan

- 1 Contexte
- 2 Stratégie
- 3 Exploitation de données cinétiques
- 4 Exploitation de données transcriptomiques haut débit
- 5 Conclusions et perspectives

Importance des lipides

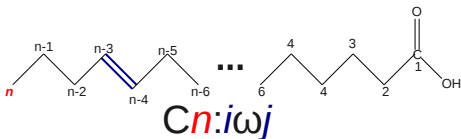
physiologie, pathologie et économie

	Pathologies	Qualité des produits	Physiologie
Quantité	<ul style="list-style-type: none">• Obésité• Diabète• maladies cardiovasculaires• dyslipidémie• stéatose hépatique ...	<ul style="list-style-type: none">• Goût• Palatabilité• Tendreté	<ul style="list-style-type: none">• Réserves• ...
Composition	<ul style="list-style-type: none">• Carence $\omega 3$, $\omega 6$<ul style="list-style-type: none">⇒ Problèmes cérébraux• Mauvais rapport $\omega 3/\omega 6$<ul style="list-style-type: none">⇒ Inflammation⇒ Maladies cardiovasculaires	<ul style="list-style-type: none">• Rancissement• Aptitude à la transformation	<ul style="list-style-type: none">• Propriétés des membranes cellulaires
Localisation		<ul style="list-style-type: none">• Choix du consommateur	

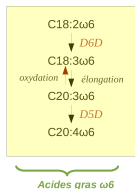
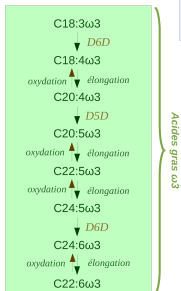
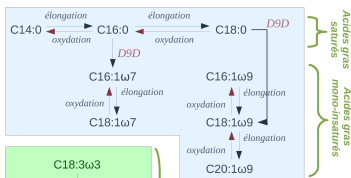
Conclusion

La **quantité totale**, la **composition** en acides gras et la **localisation** sont importants.

Composition en acides gras (Murray et al., 2006)



- n Nombre de carbones
- i Nombre d'insaturations
- j Position de la première insaturation à partir du CH₃



- Acides gras non essentiels
 - ▶ saturés, ω₉ et ω₇
 - ▶ synthétisés par l'animal
- Acides gras essentiels
 - ▶ Acides gras ω₃ et ω₆
 - ▶ Nécessitent les précurseurs C18 :3ω₃ et C18 :2ω₆
 - ▶ Ne changent pas de famille.

Quantité de lipides et localisation

- La quantité/composition en acides gras dépend des fonctions :

- ▶ Synthèse (*Murray et al., 2006*)
- ▶ Lieu de stockage (*Murray et al., 2006*)
- ▶ Dégradation
 - ★ Oxydation (*S. Eaton. 2002*)
 - ★ Cétogenèse (*Murray et al., 2006*)

- Les principaux organes de ces fonctions **dépendent de l'espèce**

Organes principaux Fonctions principales	Foie	Tissu Adipeux
Synthèse	<ul style="list-style-type: none">● Homme (<i>F.Diraison et al., 2003</i>)● Rat (<i>G.Gandemer et al., 1983</i>)● Poulet (<i>G.A. Leville et al., 1975</i>)	<ul style="list-style-type: none">● Porc (<i>E.K. O'Hea et al., 1969</i>)● Bovins
Stockage	<ul style="list-style-type: none">● Oiseaux migrateurs	<ul style="list-style-type: none">● Homme● Rat● Porc● Autres oiseaux (poulet)
Dégradation (la plus part des organes oxydent)	<ul style="list-style-type: none">● ∇ espèce : Cétogenèse	

Les connaissances relatives au métabolisme des lipides

- Données expérimentales
 - 😊 Nombreuses (ex *microarray*)
 - 😞 Difficiles à interpréter
 - 😞 Non disponibles sur les espèces d'intérêt
- Littérature : réactions biochimiques
 - 😊 Réactions bien connues (ex : *Murray et al., 2006*)
 - 😊 Connaissances accessibles (ex : KEGG, manuels)
 - 😞 Cinétique mal connue (ex : MetaCyc)
- Littérature : régulations
 - 😊 De nombreuses régulations
 - 😊 De nombreux éléments
 - ▶ Ex : Pubmed "*Regulation PPAR alpha*" ⇒ 1 655 publications
 - 😞 Information éparpillée (experts, littérature, bases de données)

Conclusion

😞 Le métabolisme des lipides est vaste et complexe.

Les connaissances relatives au métabolisme des lipides

- Données expérimentales
 - 😊 Nombreuses (ex *microarray*)
 - 😞 Difficiles à interpréter
 - 😞 Non disponibles sur les espèces d'intérêt
- Littérature : réactions biochimiques
 - 😊 Réactions bien connues (ex : *Murray et al., 2006*)
 - 😊 Connaissances accessibles (ex : KEGG, manuels)
 - 😞 Cinétique mal connue (ex : MetaCyc)
- Littérature : régulations
 - 😊 De nombreuses régulations
 - 😊 De nombreux éléments
 - ▶ Ex : Pubmed "*Regulation PPAR alpha*" ⇒ 1 655 publications
 - 😞 Information éparpillée (experts, littérature, bases de données)

Conclusion

😞 Le métabolisme des lipides est vaste et complexe.

Les connaissances relatives au métabolisme des lipides

- Données expérimentales
 - ☺ Nombreuses (ex *microarray*)
 - ☹ Difficiles à interpréter
 - ☹ Non disponibles sur les espèces d'intérêt
- Littérature : réactions biochimiques
 - ☺ Réactions bien connues (ex : *Murray et al., 2006*)
 - ☺ Connaissances accessibles (ex : KEGG, manuels)
 - ☹ Cinétique mal connue (ex : MetaCyc)
- Littérature : régulations
 - ☺ De nombreuses régulations
 - ☺ De nombreux éléments
 - ▶ Ex : Pubmed "*Regulation PPAR alpha*" ⇒ 1 655 publications
 - ☹ Information éparpillée (experts, littérature, bases de données)

Conclusion

☹ Le métabolisme des lipides est vaste et complexe.

Les connaissances relatives au métabolisme des lipides

- Données expérimentales
 - 😊 Nombreuses (ex *microarray*)
 - 😞 Difficiles à interpréter
 - 😞 Non disponibles sur les espèces d'intérêt
- Littérature : réactions biochimiques
 - 😊 Réactions bien connues (ex : *Murray et al., 2006*)
 - 😊 Connaissances accessibles (ex : KEGG, manuels)
 - 😞 Cinétique mal connue (ex : MetaCyc)
- Littérature : régulations
 - 😊 De nombreuses régulations
 - 😊 De nombreux éléments
 - ▶ Ex : Pubmed "*Regulation PPAR alpha*" ⇒ 1 655 publications
 - 😞 Information éparpillée (experts, littérature, bases de données)

Conclusion

😞 Le métabolisme des lipides est vaste et complexe.

Problématique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier les régulateurs de ces voies

Difficulté

Le métabolisme des lipides est **vaste** et **complexe**

Stratégie

Approche **systemique** réalisée par **modélisation**

Difficulté

Peu de données sur les espèces d'intérêt

Stratégie

Hypothèse de **généricité**
⇒ extrapoler les voies biochimiques connues

Difficulté

Le métabolisme des lipides est **multi-échelles**

Stratégie

Choix du niveau de détail des données expérimentales

Problématique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier les régulateurs de ces voies

Difficulté

Le métabolisme des lipides est **vaste** et **complexe**

Stratégie

Approche **systemique** réalisée par **modélisation**

Difficulté

Peu de données sur les espèces d'intérêt

Stratégie

Hypothèse de **généricité**
⇒ extrapoler les voies biochimiques connues

Difficulté

Le métabolisme des lipides est **multi-échelles**

Stratégie

Choix du niveau de détail des données expérimentales

Problématique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier les régulateurs de ces voies

Difficulté

Le métabolisme des lipides est **vaste** et **complexe**

Stratégie

Approche **systemique** réalisée par **modélisation**

Difficulté

Peu de données sur les espèces d'intérêt

Stratégie

Hypothèse de **généricité**
⇒ extrapoler les voies biochimiques connues

Difficulté

Le métabolisme des lipides est **multi-échelles**

Stratégie

Choix du niveau de détail des données expérimentales

Problématique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier les régulateurs de ces voies

Difficulté

Le métabolisme des lipides est **vaste** et **complexe**

Stratégie

Approche **systemique** réalisée par **modélisation**

Difficulté

Peu de données sur les espèces d'intérêt

Stratégie

Hypothèse de **généricité**
⇒ extrapoler les voies biochimiques connues

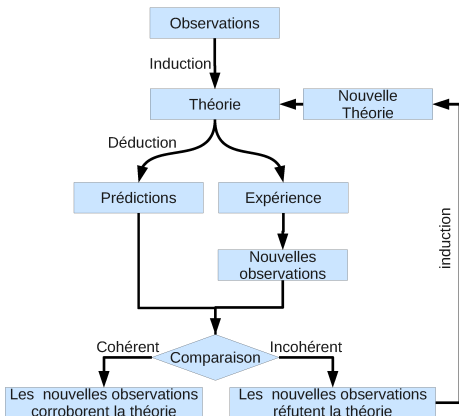
Difficulté

Le métabolisme des lipides est **multi-échelles**

Stratégie

Choix du niveau de détail des données expérimentales

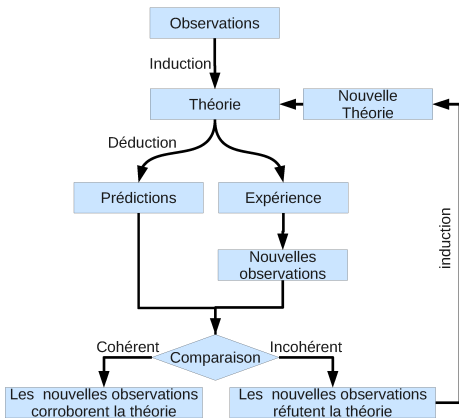
La complexité limite la démarche expérimentale



Difficulté

- Vaste et complexe
- ⇒ Impossible de tester la cohérence de la théorie
- ⇒ Impossible de réaliser des prédictions
- ⇒ Impossible de comparer la théorie aux observations

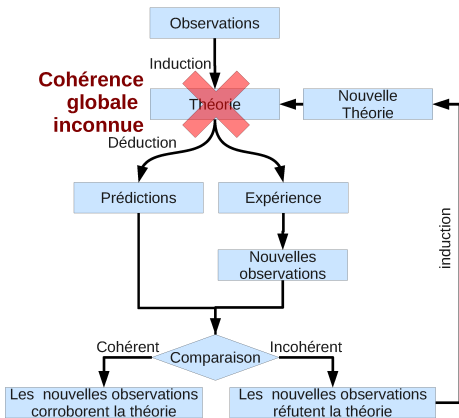
La complexité limite la démarche expérimentale



Difficulté

- Vaste et complexe
 - ⇒ Impossible de tester la cohérence de la théorie
 - ⇒ Impossible de réaliser des prédictions
 - ⇒ Impossible de comparer la théorie aux observations

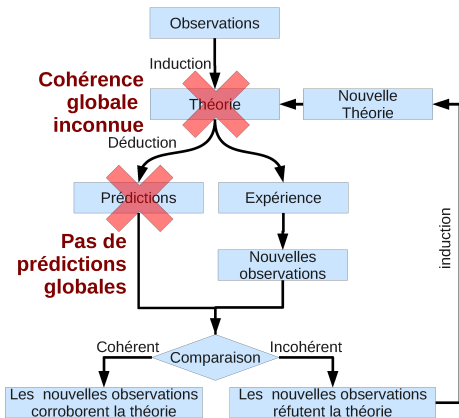
La complexité limite la démarche expérimentale



Difficulté

- Vaste et complexe
- ⇒ Impossible de tester la cohérence de la théorie
- ⇒ Impossible de réaliser des prédictions
- ⇒ Impossible de comparer la théorie aux observations

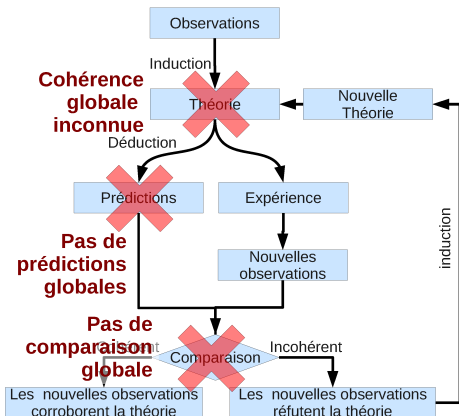
La complexité limite la démarche expérimentale



Difficulté

- Vaste et complexe
- ⇒ Impossible de tester la cohérence de la théorie
- ⇒ Impossible de réaliser des prédictions
- ⇒ Impossible de comparer la théorie aux observations

La complexité limite la démarche expérimentale



Difficulté

- Vaste et complexe
- ⇒ Impossible de tester la cohérence de la théorie
- ⇒ Impossible de réaliser des prédictions
- ⇒ Impossible de comparer la théorie aux observations

Appréhender la complexité par une approche systémique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier les régulateurs de ces voies
- ⇔ Rechercher un ensemble d'éléments parmi les éléments internes qui sont responsables d'une propriété globale.

Stratégie

- **Approche systémique** : lien entre la mécanique interne et les propriétés globales
- **Systeme**
 - ▶ Elements
 - ▶ Relations entre éléments
 - ▶ Frontière
- **Modèles** : moyen pour représenter les systèmes

Appréhender la complexité par une approche systémique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier **les régulateurs de ces voies**
- ⇔ Rechercher un **ensemble d'éléments** parmi les éléments internes qui sont responsables d'une propriété globale.

Stratégie

- **Approche systémique** : lien entre la mécanique interne et les propriétés globales
- **Systeme**
 - Elements
 - Relations entre éléments
 - Frontière
- **Modèles** : moyen pour représenter les systèmes

Appréhender la complexité par une approche systémique

Problématique

- Identifier la hiérarchie **d'importance des voies métaboliques**
- Identifier les régulateurs de ces voies
- ⇔ Rechercher un ensemble d'éléments parmi les éléments internes qui sont responsables d'une **propriété globale**.

Stratégie

- **Approche systémique** : lien entre la mécanique interne et les propriétés globales
- **Systeme**
 - Elements
 - Relations entre éléments
 - Frontière
- **Modèles** : moyen pour représenter les systèmes

Appréhender la complexité par une approche systémique

Problématique

- Identifier la hiérarchie d'importance des voies métaboliques
- Identifier les régulateurs de ces voies
- ⇔ Rechercher un ensemble d'éléments parmi les éléments internes qui sont responsables d'une propriété globale.

Stratégie

- **Approche systémique** : lien entre la mécanique interne et les propriétés globales
- **Systeme**
 - ▶ Elements
 - ▶ Relations entre éléments
 - ▶ Frontière
- **Modèles** : moyen pour représenter les systèmes

Exploiter la généricité

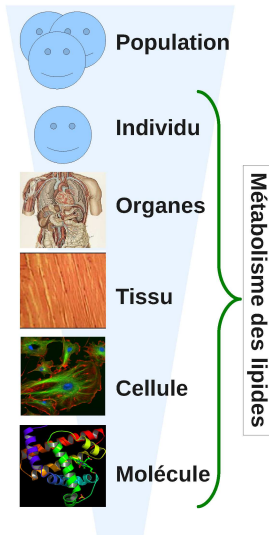
Problème

- ☹ Peu de données sur les espèces d'intérêt
- ☹ Aquisition de nouvelles données coûteuse
- 😊 De nombreuses données et connaissances sur d'autres espèces (souris)

Solution

- **Hypothèse** : les voies sont communes aux différents organes/espèces
- ⇒ Extrapolation du squelette des voies métaboliques
- ⇒ Transversalité : modèles indépendants du contexte (i.e. organe, espèce, conditions)
- ⇒ Comparaisons entre différents contextes

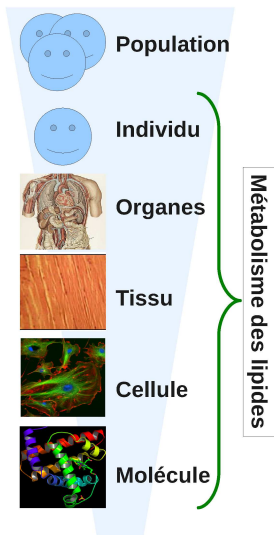
Le métabolisme des lipides est multi-échelles



Difficulté

- Le métabolisme des lipides est **multi échelle**
- Niveaux de détail **interdépendants**
 - ☹ Comportement **niveau actuel** = Propriétés émergentes + résultantes du **niveau inférieur**
 - ☹ Comportement **niveau actuel** ← influencé par le **niveau supérieur**

Choix du niveau de détail



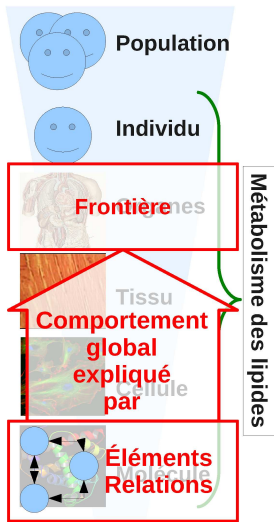
Stratégie

Niveau de détail du système = niveau de détail des principales données biologiques disponibles.

- Système
 - ▶ **Elements** : variations de quantité de molécules
 - ▶ **Relations** : activation, réaction ...
 - ▶ **Frontière** : niveau de mesure : organe
- Données principalement de deux types
 - ▶ **Cinétique** : mesure *quantitative au cours du temps de quelques métabolites*
 - ▶ **Microarray** : mesure qualitative entre *quelques conditions de nombreux ARNm*

⇒ Modélisation à ces deux niveaux de détail.

Choix du niveau de détail



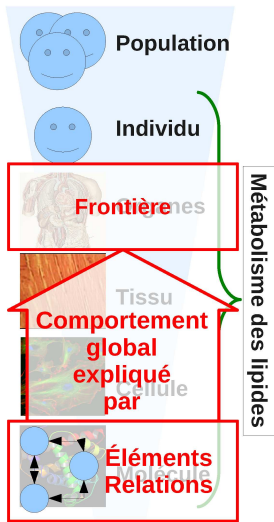
Stratégie

Niveau de détail du système = niveau de détail des principales données biologiques disponibles.

- **Système**
 - ▶ **Elements** : variations de quantité de molécules
 - ▶ **Relations** : activation, réaction ...
 - ▶ **Frontière** : niveau de mesure : organe
- **Données principalement de deux types**
 - ▶ **Cinétique** : mesure *quantitative au cours du temps de quelques métabolites*
 - ▶ **Microarray** : mesure *qualitative entre quelques conditions de nombreux ARNm*

⇒ Modélisation à ces deux niveaux de détail.

Choix du niveau de détail



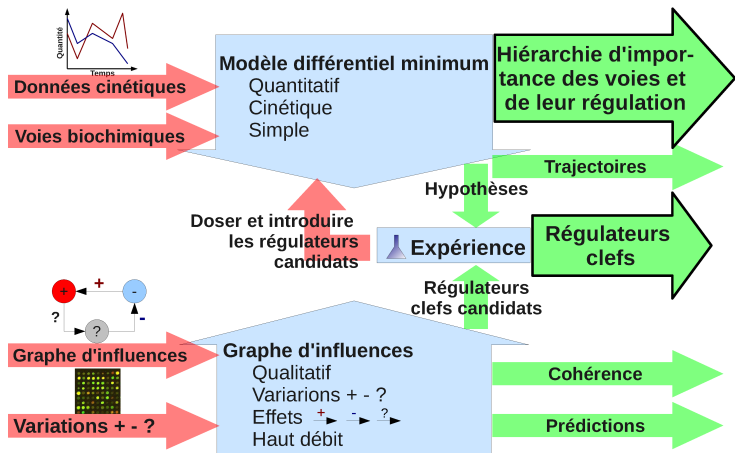
Stratégie

Niveau de détail du système = niveau de détail des principales données biologiques disponibles.

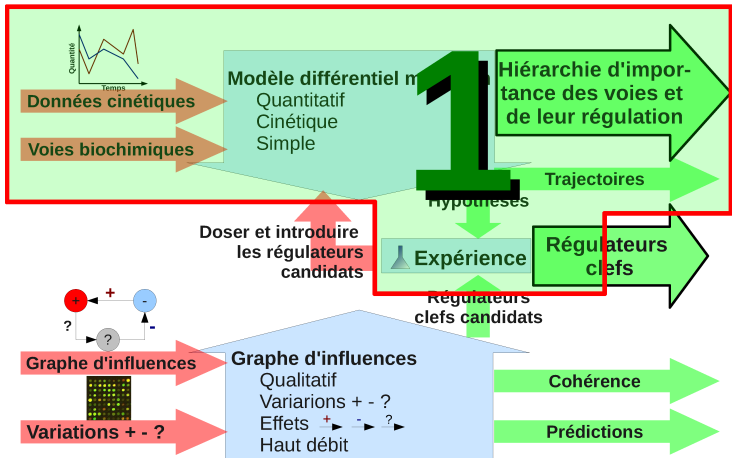
- **Système**
 - ▶ **Elements** : variations de quantité de molécules
 - ▶ **Relations** : activation, réaction ...
 - ▶ **Frontière** : niveau de mesure : organe
- **Données principalement de deux types**
 - ▶ **Cinétique** : mesure *quantitative au cours du temps de quelques métabolites*
 - ▶ **Microarray** : *mesure qualitative entre quelques conditions de nombreux ARNm*

⇒ Modélisation à ces deux niveaux de détail.

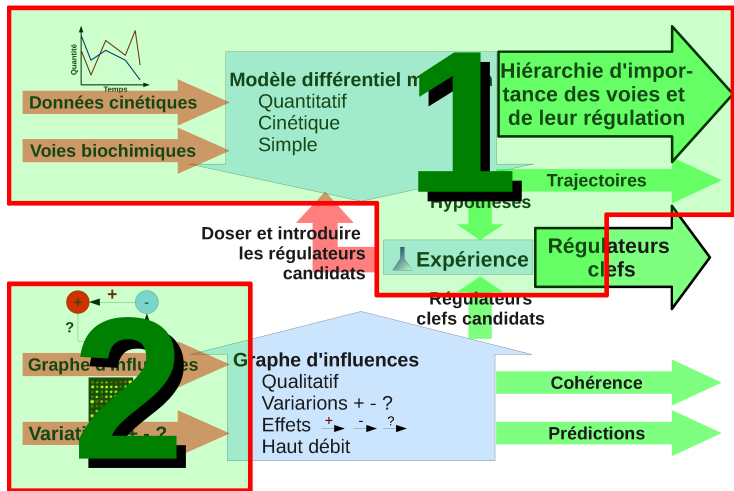
Deux échelles complémentaires



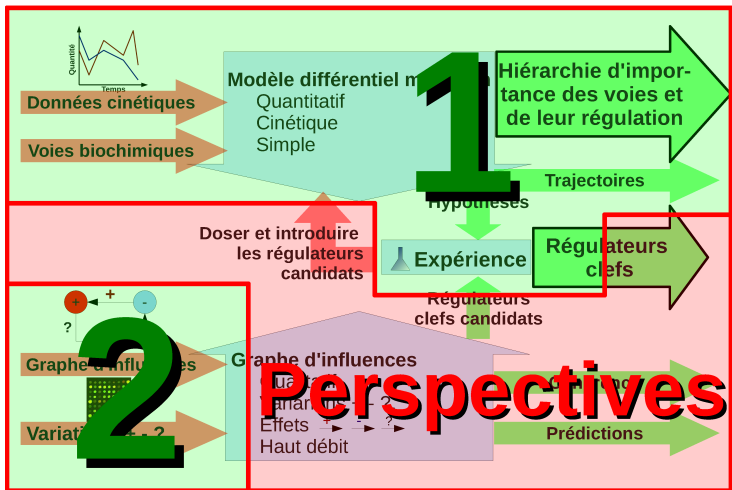
Deux échelles complémentaires



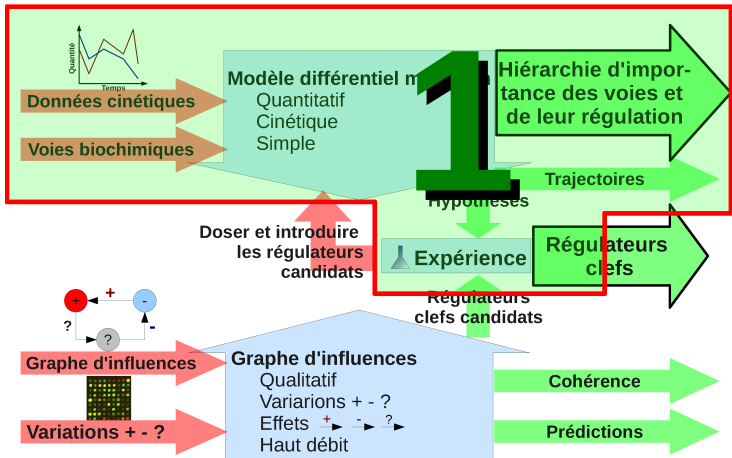
Deux échelles complémentaires



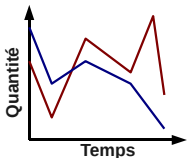
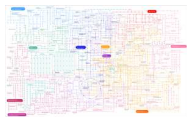
Deux échelles complémentaires



Plan : Exploitation des données cinétiques



Contexte : les informations disponibles pour expliquer les données cinétiques



- Informations bibliographiques

- ☺ Voies biochimiques
- ☺ Principaux régulateurs
- ☺ Différents enzymes

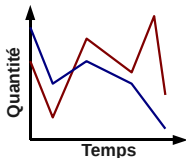
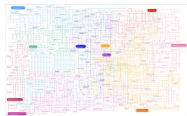
- Données cinétiques quantitatives

- ☹ Forte variabilité
- ☹ Peu de données sur les espèces d'intérêt
- ☺ Données dans d'autres espèces (ici souris, UR66)

- Informations inconnues

- ☹ Les fonctions importantes.
- ☹ L'effet des régulateurs connus.
- ☹ Les vitesses des réactions biochimiques.

Contexte : les informations disponibles pour expliquer les données cinétiques



- Informations bibliographiques

- ☺ Voies biochimiques
- ☺ Principaux régulateurs
- ☺ Différents enzymes

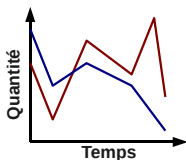
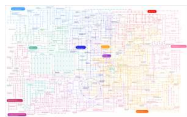
- Données cinétiques quantitatives

- ☹ Forte variabilité
- ☹ Peu de données sur les espèces d'intérêt
- ☺ Données dans d'autres espèces (ici souris, UR66)

- Informations inconnues

- ☹ Les fonctions importantes.
- ☹ L'effet des régulateurs connus.
- ☹ Les vitesses des réactions biochimiques.

Contexte : les informations disponibles pour expliquer les données cinétiques



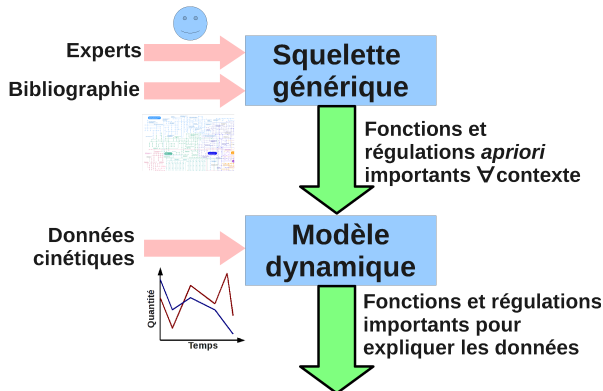
?

- Informations bibliographiques
 - ☺ Voies biochimiques
 - ☺ Principaux régulateurs
 - ☺ Différents enzymes
- Données cinétiques quantitatives
 - ☹ Forte variabilité
 - ☹ Peu de données sur les espèces d'intérêt
 - ☺ Données dans d'autres espèces (ici souris, UR66)
- Informations inconnues
 - ☹ Les fonctions importantes.
 - ☹ L'effet des régulateurs connus.
 - ☹ Les vitesses des réactions biochimiques.

Stratégie

Objectif

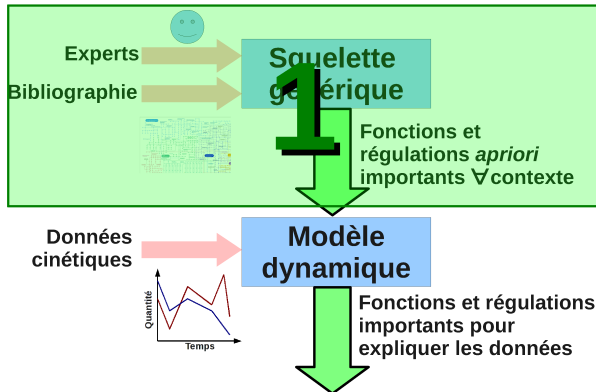
- Identifier la hiérarchie d'importance des voies biochimiques
 - 1 *A priori*, indépendamment du contexte
 - ⇒ Base de construction de modèles
 - 2 Dans un contexte particulier



Stratégie

Objectif

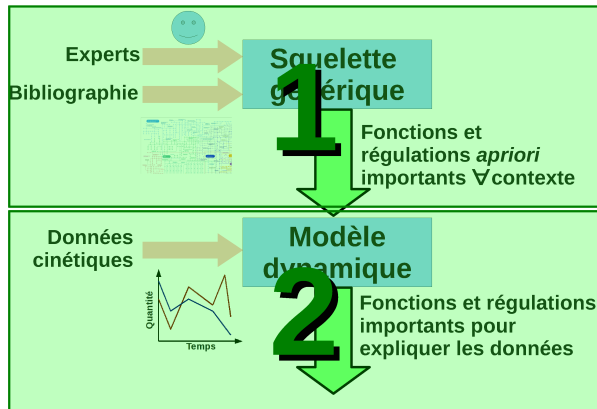
- Identifier la hiérarchie d'importance des voies biochimiques
 - 1 **A priori, indépendamment du contexte**
 - ⇒ Base de construction de modèles
 - 2 Dans un contexte particulier



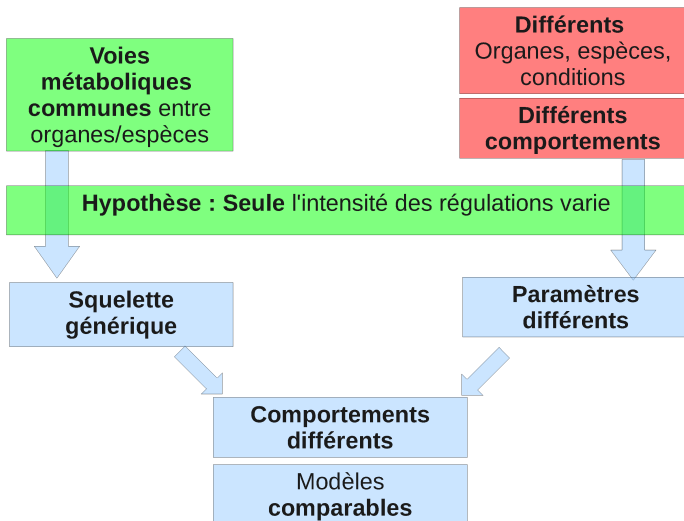
Stratégie

Objectif

- Identifier la hiérarchie d'importance des voies biochimiques
 - 1 *A priori*, indépendamment du contexte
 - ⇒ Base de construction de modèles
 - 2 **Dans un contexte particulier**



Stratégie : exploiter la généricité dans le modèle dynamique



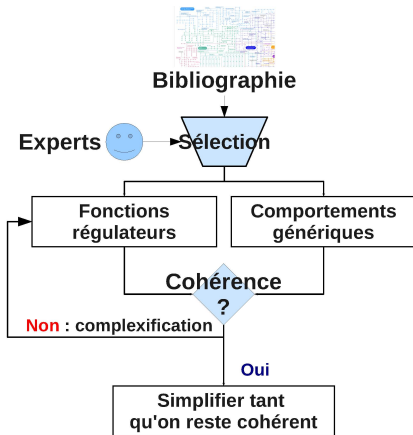
Méthode : construction d'un squelette générique

Difficulté

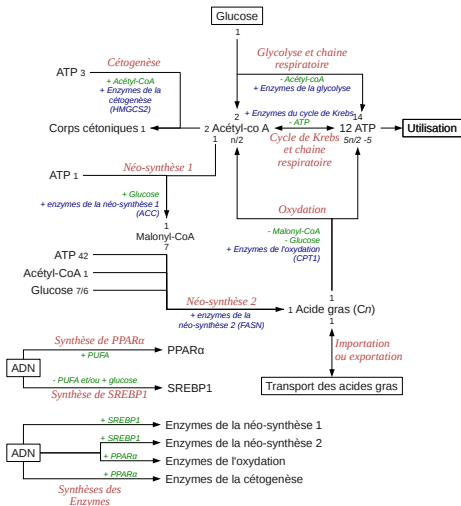
Les voies et régulateurs importants ne sont pas connus *a priori*

Solution

Définir comme importants les éléments nécessaires et suffisants pour expliquer des comportements génériques



Résultats : Le squelette générique construit (Blavy et al., Jobim 2008)

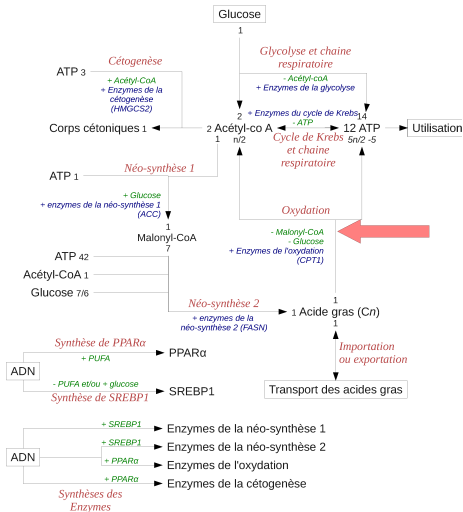


Résultats

- **Base de construction simple des futurs modèles**
- Une régulation dont l'importance n'était pas évidente *a priori* dans un modèle minimal
- Une **synergie** entre la régulation de la transcription et de l'activité enzymatique ⇒ observer à différentes échelles pour dissocier les deux

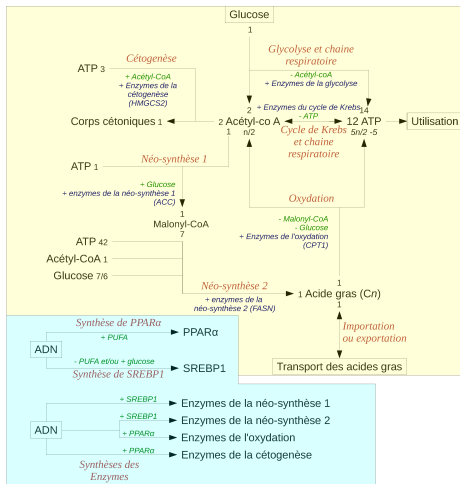
Résultats : Le squelette générique construit (Blavy et al., Jobim 2008)

Résultats



- Base de construction simple des futurs modèles
- **Une régulation dont l'importance n'était pas évidente a priori dans un modèle minimal**
- Une **synergie** entre la régulation de la transcription et de l'activité enzymatique ⇒ observer à différentes échelles pour dissocier les deux

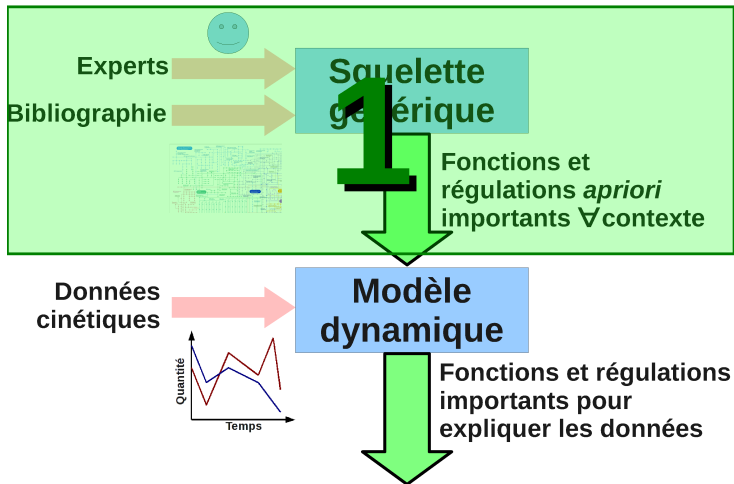
Résultats : Le squelette générique construit (Blavy et al., Jobim 2008)



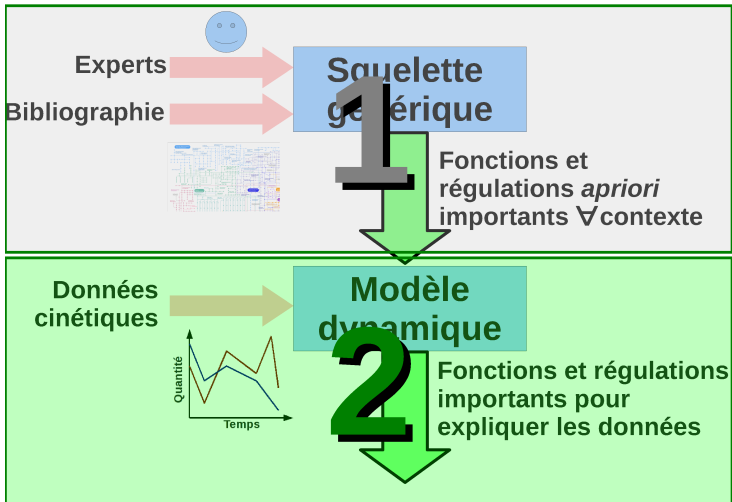
Résultats

- Base de construction simple des futurs modèles
- Une régulation dont l'importance n'était pas évidente *a priori* dans un modèle minimal
- Une synergie entre la régulation de la transcription et de l'activité enzymatique ⇒ observer à différentes échelles pour dissocier les deux

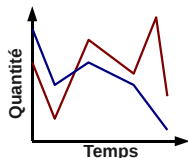
Plan : Exploitation des données cinétiques



Plan : Exploitation des données cinétiques



Choix : modèle d'équations différentielles ordinaires

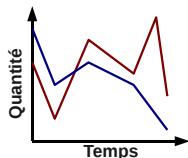


- Observations
 - ▶ **Données cinétiques**
 - ▶ Données quantitatives
 - ▶ Modèle déterministe
 - ▶ ⇒ Equations différentielles ordinaires (EDO)

$$\frac{dx}{dt}$$

- Les modèles d'équations différentielles ordinaires fournissent
 - ▶ 😊 **Trajectoires**
 - ▶ 😊 Trajectoires quantitatives
 - ▶ 😊 Simulation déterministe

Choix : modèle d'équations différentielles ordinaires

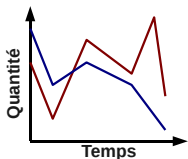


- Observations
 - ▶ Données cinétiques
 - ▶ **Données quantitatives**
 - ▶ Modèle déterministe
 - ▶ ⇒ Equations différentielles ordinaires (EDO)

$$\frac{dx}{dt}$$

- Les modèles d'équations différentielles ordinaires fournissent
 - ▶ 😊 Trajectoires
 - ▶ 😊 Trajectoires **quantitatives**
 - ▶ 😊 Simulation déterministe

Choix : modèle d'équations différentielles ordinaires

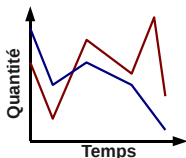


- Observations
 - ▶ Données cinétiques
 - ▶ Données quantitatives
 - ▶ **Modèle déterministe**
 - ▶ ⇒ Equations différentielles ordinaires (EDO)

$$\frac{dx}{dt}$$

- Les modèles d'équations différentielles ordinaires fournissent
 - ▶ 😊 Trajectoires
 - ▶ 😊 Trajectoires quantitatives
 - ▶ 😊 **Simulation déterministe**

Choix : modèle d'équations différentielles ordinaires



- Observations

- ▶ Données cinétiques
- ▶ Données quantitatives
- ▶ Modèle déterministe
- ▶ ⇒ **Equations différentielles ordinaires (EDO)**

$$\frac{dx}{dt}$$

- Les modèles d'équations différentielles ordinaires fournissent

- ▶ ☺ Trajectoires
- ▶ ☺ Trajectoires quantitatives
- ▶ ☺ Simulation déterministe

Méthode : Construction pas à pas

Difficulté



- ☹ Paramètres et équations inconnus
- ⇒ Surparamétrisation

Solution

- Simplifier le squelette en fonction du contexte expérimental
- Introduire pas à pas les éléments nécessaires pour expliquer les données
- ☺ Limite la surparamétrisation
- ☺ Identification **empirique** des éléments importants

Méthode : Compatibilité ascendante des modèles

Problème

- De nombreux modèles sont construits
 - ▶ Exploiter les anciens modèles.
 - ▶ Réaliser des comparaisons.

Solution : construction ascendante

- Lorsque cela est possible, **tout ancien modèle est considéré comme un cas particulier du nouveau**
 - ☺ Changer la valeur d'un paramètre, rajouter un flux, utiliser un cas limite.
 - ☹ Changer la forme des équations

Conclusion

- La construction ascendante permet
 - ▶ De réaliser des comparaisons
 - ▶ De modéliser un système pas à pas

Méthode : Compatibilité ascendante des modèles

Problème

- De nombreux modèles sont construits
 - ▶ Exploiter les anciens modèles.
 - ▶ Réaliser des comparaisons.

Solution : construction ascendante

- Lorsque cela est possible, **tout ancien modèle est considéré comme un cas particulier du nouveau**
 - ☺ Changer la valeur d'un paramètre, rajouter un flux, utiliser un cas limite.
 - ☹ Changer la forme des équations

Conclusion

- La construction ascendante permet
 - ▶ De réaliser des comparaisons
 - ▶ De modéliser un système pas à pas

Méthode : Compatibilité ascendante des modèles

Problème

- De nombreux modèles sont construits
 - ▶ Exploiter les anciens modèles.
 - ▶ Réaliser des comparaisons.

Solution : construction ascendante

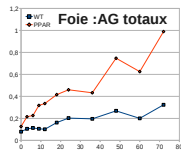
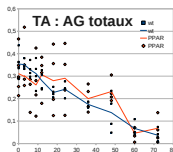
- Lorsque cela est possible, **tout ancien modèle est considéré comme un cas particulier du nouveau**
 - ☺ Changer la valeur d'un paramètre, rajouter un flux, utiliser un cas limite.
 - ☹ Changer la forme des équations

Conclusion

- La construction ascendante permet
 - ▶ De réaliser des comparaisons
 - ▶ De modéliser un système pas à pas

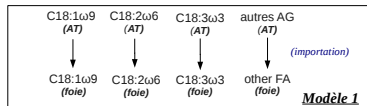
Modele 1 : construction (Blavy et al., 2009)

- Les données disponibles (UR66)
 - ▶ Foie et tissus adipeux
 - ▶ Mesure quantitative de 14 acides gras
 - ▶ 11 points, 0 à 72 heures de jeûne
 - ▶ Souris sauvages et *knockout* PPAR α
 - ▶ Deux échelles de temps



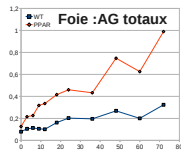
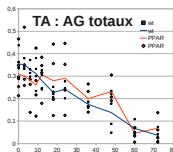
• Modèle1 : construction

- ▶ Importation_{foie} = $K_{in} \cdot \text{Exportation}_{\text{tissu adipeux}}$
- ▶ K_{in} est identique \forall acide gras
- ▶ K_{in} évalué chez *knockout* PPAR α puis extrapolée aux sauvages



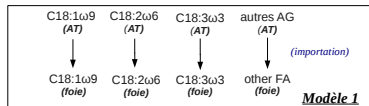
Modele 1 : construction (Blavy et al., 2009)

- Les données disponibles (UR66)
 - ▶ Foie et tissus adipeux
 - ▶ Mesure quantitative de 14 acides gras
 - ▶ 11 points, 0 à 72 heures de jeûne
 - ▶ Souris sauvages et *knockout* PPAR α
 - ▶ Deux échelles de temps

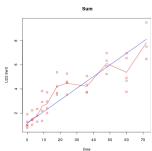


- Modèle1 : construction

- ▶ Importation_{foie} = $K_{in} \cdot \text{Exportation}_{\text{tissu adipeux}}$
- ▶ K_{in} est identique \forall acide gras
- ▶ K_{in} évalué chez *knockout* PPAR α puis extrapolée aux sauvages

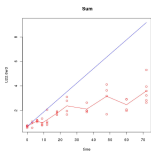


Modèle 1 : résultats (jeûne, foie, souris) (Blavy et al., 2009)



AG totaux *knockout*
PPAR α

- ☺ 10/14 acides gras **bien ajustés** chez *knockout* PPAR α
 - ⇒ Importation phénomène majeur
 - ⇒ Importation_{foie} proportionnelle exportation_{Tissu adipeux}



AG totaux sauvage

- ☹ PPAR α : 4 acides gras dont C16 :1 ω 9 non ajustés
- ☹ Sauvage : 13/14 acides gras surestimés chez
 - ⇒ L'importation n'explique pas tout
 - ⇒ Introduire l'oxydation (Murray et al., 2006), même chez *knockout* PPAR α

Conclusion

Complexifier le modèle en rajoutant l'oxydation des acides gras

Modèle 2 : Modéliser l'oxydation (Blavy et al., 2009)

Difficulté

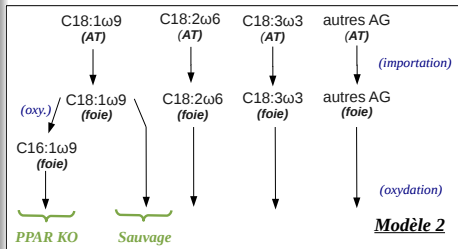


- Rajouter l'oxydation

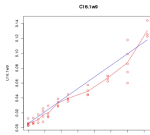
- ▶ ∃ Oxydation totale ou partielle
- ▶ La vitesse d'oxydation dépend de l'acide gras oxydé
- ⇒ Beaucoup de réactions, beaucoup de paramètres

Solution

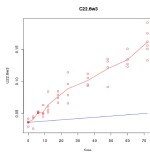
- Essayer les hypothèses par ordre de complexité
 - ▶ Oxydation totale de tous les acides gras avec des vitesses spécifiques chez le sauvage
 - ▶ Oxydation du C18 :1 ω 9 en C16 :1 ω 9 chez le *knockout* PPAR α



Modèle2 : Résultats (jeûne, foie, souris) (Blavy et al., 2009)



C16 :1 ω 9 *knockout*



C22 :6 ω 3 *sauvage*

- 😊 Sauvage : 13/14 acides gras ajustés
- 😊 Oxydation sauvage \gg *knockout* PPAR α
 - \Rightarrow Sauvage : β -oxydation phénomène majeur
 - 😊 Conforme aux *a priori* (Desvergne, B. et al., 1999)
 - 😞 Sauvage : Surparamétrisation? (Murray et al., 2006)
- 😊 *Knockout* : C16 :1 ω 9 ajusté
 - \Rightarrow **knockout** : \exists β -oxydation
- 😞 *Knockout* : C22 :6 ω 3 sous-estimé, C18 :3 ω 3 sur-estimé
- 😞 Sauvage : C22 :6 ω 3 sous-estimé
 - \Rightarrow **Rajouter élongation désaturation dans les deux génotypes**

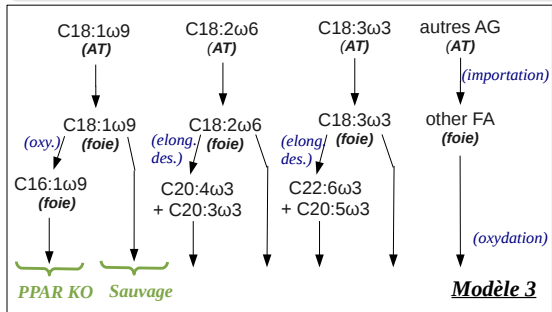
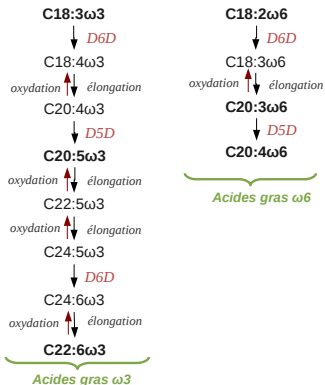
Modèle 3 : Modéliser l'élongation désaturation (Blavy et al., 2009)

Difficulté

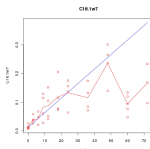
- Tous les intermédiaires ne sont pas mesurés

Solution

- D6D Etape limitante (Cho et al., 1999)
- ⇒ Ne pas modéliser les intermédiaires mais la somme intermédiaires + produit final.

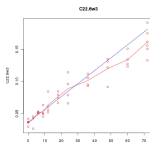


Modèle3 : Résultats (jeûne, foie, souris) (Blavy et al., 2009)



C16 :1 ω 7 knockout

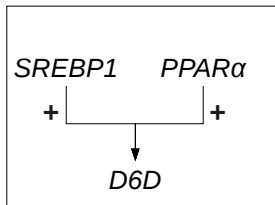
- ☹ *Knockout* : C16 :1 ω 7 non ajusté
 - ▶ Dynamique en deux temps spécifique
 - ? Capté par d'autres organes (Cao et al., 2008) ?
 - ⇒ D'autres expériences nécessaires



C22 :6 ω 3 sauvage

- Elongation désaturation sauvage = 1.8 x *knockout*
 - ☺ Conforme aux *a priori* (Matsuzaka et al., 2002)
- ☺ C22 :6 ω 3 et C18 :3 ω 3 sont correctement ajustés
 - ⇒ **Elongation désaturation active dans les deux génotypes**

L'activité de D6D à jeun chez le *knockout* PPAR α est surprenante (Blavy et al., 2009)

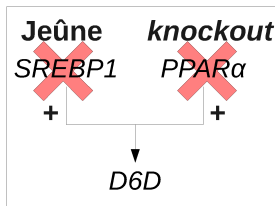


- **D6D est activé par SREBP1** (Matsuzaka et al., 2002)
 - ▶ SREBP1 inactif au bout de 3h de jeûne (Horton et al., 1998)
- **D6D est activé par PPAR α** (Matsuzaka et al., 2002)
 - ▶ PPAR α non présent chez le *knockout*
- Or Modèle : D6D active
 - ⇒ Surprenant

Conclusion

- ⇒ Dosage de D6D (UR66)

L'activité de D6D à jeun chez le *knockout* PPAR α est surprenante (Blavy et al., 2009)

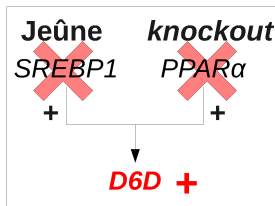


- D6D est activé par SREBP1 (Matsuzaka et al., 2002)
 - ▶ **SREBP1 inactif au bout de 3h de jeûne** (Horton et al., 1998)
- D6D est activé par PPAR α (Matsuzaka et al., 2002)
 - ▶ **PPAR α non présent chez le knockout**
- Or Modèle : D6D active
⇒ Surprenant

Conclusion

- ⇒ Dosage de D6D (UR66)

L'activité de D6D à jeun chez le *knockout* PPAR α est surprenante (Blavy et al., 2009)

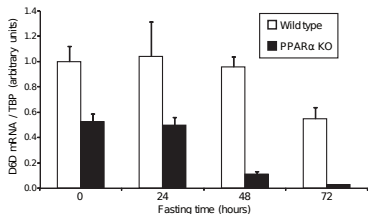


- D6D est activé par SREBP1 (Matsuzaka et al., 2002)
 - ▶ SREBP1 inactif au bout de 3h de jeûne (Horton et al., 1998)
- D6D est activé par PPAR α (Matsuzaka et al., 2002)
 - ▶ PPAR α non présent chez le *knockout*
- Or Modèle : **D6D active**
 - ⇒ Surprenant

Conclusion

- ⇒ Dosage de D6D (UR66)

Retour à l'expérience (UR66) (Blavy et al., 2009)



- Dosage des ARNm de D6D (UR66)
 - ▶ ARNm D6D présent y compris chez le *knockout* PPAR α
 - ▶ ARNm D6D stable les 24 premières heures
 - ▶ ARNm D6D décroît ensuite

Conclusion

- \exists Régulation de D6D, y compris en l'absence de PPAR α et de SREBP1.
- ☹ Régulateur inconnu

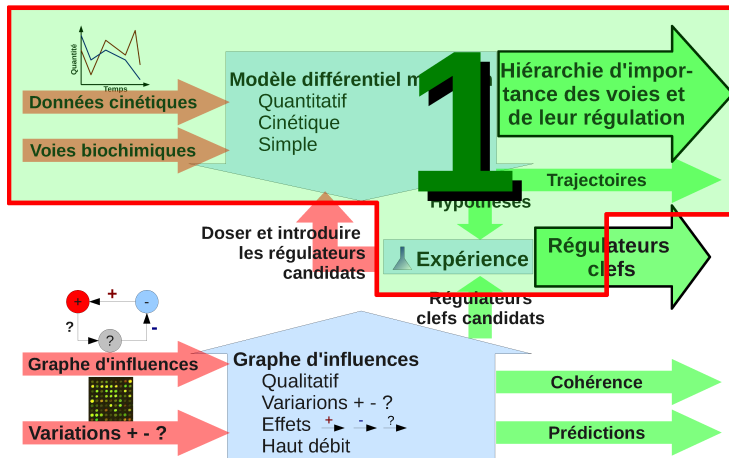
Conclusion

- Données multiorganes
 - ▶ Utiles pour construire et ajuster les modèles \Rightarrow Expérimentation futures multi-organes.
 - ▶ Mais : pas d'information sur l'importation et l'oxydation, ce qui était **nécessaire a priori** d'après le squelette \Rightarrow Réaliser un modèle en amont des premières expériences.
- Apports de la modélisation
 - ▶ Identification des voies importantes pour expliquer un jeu de données.
 - ▶ Identification d'une nouvelle régulation potentielle, puis validation expérimentale \Rightarrow Aller-retour modèle expérience.

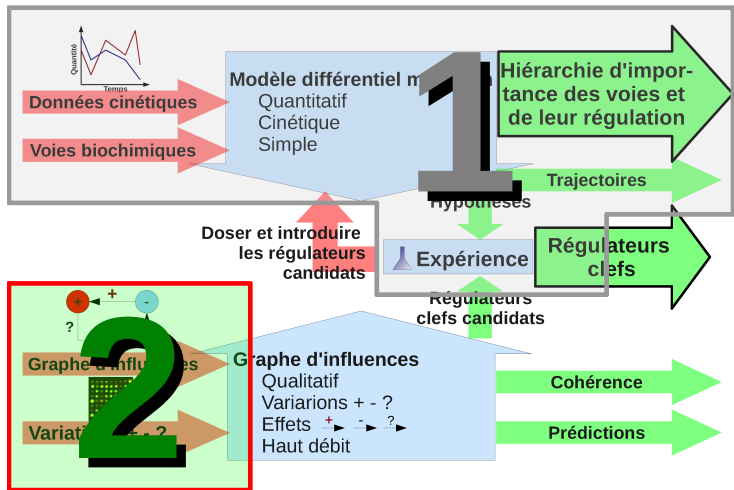
Conclusion

Modèles simples en amont et aval des expériences avec le biologiste

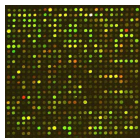
Plan : Exploitation des données transcriptomiques haut débit



Plan : Exploitation des données transcriptomiques haut débit



Les informations disponibles

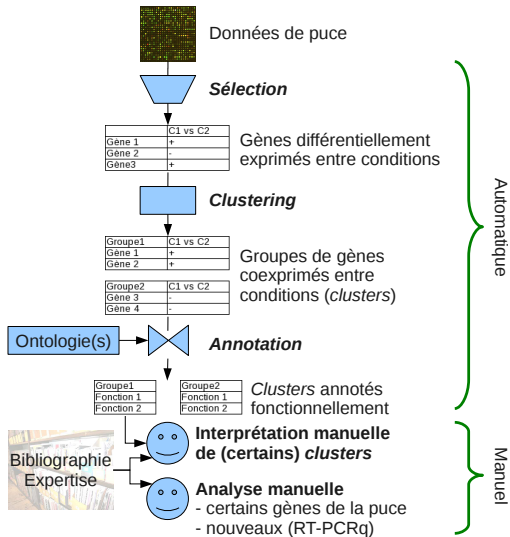


- Données transcriptomiques haut débit
 - ▶ Données nombreuses
 - ▶ Mesure simultanée des quantités d'ARNm
 - ▶ Données qualitatives
 - ▶ Comparaison entre deux (ou quelques) conditions



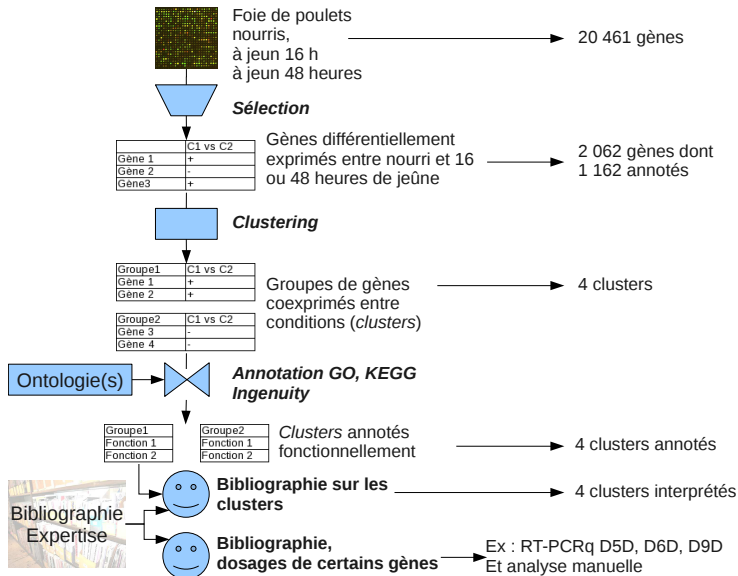
- Connaissances de type
 - ▶ Mécanismes de régulations
 - ▶ Effets d'une molécule
 - ▶ Annotation fonctionnelle
 - ▶ ...

Méthode d'analyse classique



Analyse transcriptomique de foie de poulets mis à jeûn

Desert et al., 2008



Analyse transcriptomique de foie de poulets mis à jeûn

Desert *et al.*, 2008

- Clustering

- ▶ **Cluster 1 : -0** : Synthèse des AG, D9D
- ▶ **Cluster 2 : +-** : Oxydation des AG
- ▶ **Cluster 3 : 0-** : Glycolyse, signalisation cellulaire, insuline D5D, D6D
- ▶ **Cluster 4 : 0+** : AMPc + autres

⇒ Identification des principales voies régulées

- Modèle dynamique ⇒ Analyse manuelle des désaturase

- ▶ Comportement différents
- ▶ D5D, D6D, D9D ont les mêmes régulateurs (SREBP1, PPAR α ,
(Matsuzaka *et al.*, 2002))

? Autres régulateurs? Lesquels?

⇒ Rechercher les régulateurs de D5D D6D dans les clusters

Conclusion

⇒ A long terme, peut on obtenir automatiquement des régulateurs?

Analyse transcriptomique de foie de poulets mis à jeûn

Desert *et al.*, 2008

- Clustering

- ▶ **Cluster 1 : -0** : Synthèse des AG, **D9D**
- ▶ **Cluster 2 : +-** : Oxydation des AG
- ▶ **Cluster 3 : 0-** : Glycolyse, signalisation cellulaire, insuline **D5D, D6D**
- ▶ **Cluster 4 : 0+** : AMPc + autres

⇒ Identification des principales voies régulées

- Modèle dynamique ⇒ Analyse manuelle des désaturase

- ▶ **Comportement différents**

- ▶ D5D, D6D, D9D ont les mêmes régulateurs (SREBP1, PPAR α ,
(*Matsuzaka et al.*, 2002))

? Autres régulateurs? Lesquels?

⇒ Rechercher les régulateurs de D5D D6D dans les clusters

Conclusion

⇒ A long terme, peut on obtenir automatiquement des régulateurs ?

Analyse transcriptomique de foie de poulets mis à jeûn

Desert *et al.*, 2008

- Clustering

- ▶ **Cluster 1** : -0 : Synthèse des AG, D9D
- ▶ **Cluster 2** : +- : Oxydation des AG
- ▶ **Cluster 3** : 0- : Glycolyse, signalisation cellulaire, insuline D5D, D6D
- ▶ **Cluster 4** : 0+ : AMPc + autres

⇒ Identification des principales voies régulées

- Modèle dynamique ⇒ Analyse manuelle des désaturase

- ▶ Comportement différents
- ▶ D5D, D6D, D9D ont les mêmes régulateurs (SREBP1, PPAR α ,
(*Matsuzaka et al.*, 2002))

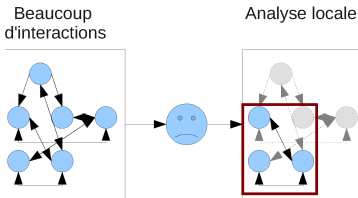
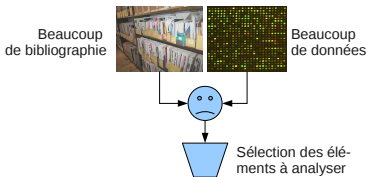
? Autres régulateurs? Lesquels?

⇒ Rechercher les régulateurs de D5D D6D dans les clusters

Conclusion

⇒ A long terme, peut on obtenir automatiquement des régulateurs ?

Les limites de l'approche classique



● Le système est vaste

- ▶ ☹️ L'expert ne peut pas lire toute la bibliographie relative à chaque molécule
- ▶ ⇒ L'expert sélectionne des éléments
- ▶ 😊 Analyse précise

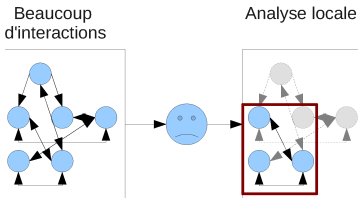
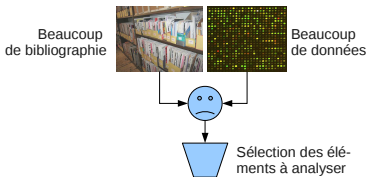
● Le système est complexe

- ▶ De nombreuses relations
- ▶ ⇒ Analyse globale complexe
- ▶ ⇒ ☹️ Analyse locale
- ▶ ⇒ ☹️ Perte de comportements

Conclusion

Le passage à large échelle nécessite d'automatiser l'exploitation de la bibliographie

Les limites de l'approche classique



● Le système est vaste

- ▶ ☹️ L'expert ne peut pas lire toute la bibliographie relative à chaque molécule
- ▶ ⇒ L'expert sélectionne des éléments
- ▶ 😊 Analyse précise

● Le système est complexe

- ▶ De nombreuses relations
- ▶ ⇒ Analyse globale complexe
- ▶ ⇒ ☹️ Analyse locale
- ▶ ⇒ ☹️ Perte de comportements

Conclusion

Le passage à large échelle nécessite d'automatiser l'exploitation de la bibliographie

Choix des graphes d'influences



	C1 vs C2
Gène 1	+
Gène 2	-
Gène 3	+

- Puce
- = Comparaison entre conditions
 - ▶ Augmente
 - ▶ Diminue
 - ▶ Variation non significative

- Graphe d'influence
- = Information qualitative
 - ▶ (+)
 - ▶ (-)
 - ▶ (?) = { (+) ou (-) }



- Biblio et raisonnement
- = Effets
 - ▶ Des effets
 - ▶ Des mécanismes souvent interprétables en terme d'effets

- Graphe d'influence
- = Raisonnement sur les effets
 - ▶ $A \xrightarrow{+} B$
 - ▶ $A \xrightarrow{-} B$
 - ▶ $A \xrightarrow{?} B$

Conclusion

Les graphes d'influences sont adaptés pour analyser des puces d'après la bibliographie

Choix des graphes d'influences



	C1 vs C2
Gène 1	+
Gène 2	-
Gène 3	+

- Puce
- = Comparaison entre conditions
 - ▶ Augmente
 - ▶ Diminue
 - ▶ Variation non significative

- Graphe d'influence
- = Information qualitative
 - ▶ (+)
 - ▶ (-)
 - ▶ (?) = { (+) ou (-) }



- Biblio et raisonnement
- = Effets
 - ▶ Des effets
 - ▶ Des mécanismes souvent interprétables en terme d'effets

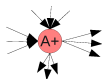
- Graphe d'influence
- = Raisonnement sur les effets
 - ▶ $A \xrightarrow{+} B$
 - ▶ $A \xrightarrow{-} B$
 - ▶ $A \xrightarrow{?} B$



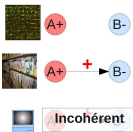
Conclusion

Les graphes d'influences sont adaptés pour analyser des puces d'après la bibliographie

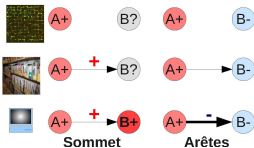
Méthodologie existante d'analyse des graphes d'influences



- Analyse de la topologie (exemples) (A.L. Barabasi et al., 1999)
 - ▶ Elements très connectés \Rightarrow régulateurs clefs potentiels
 - ▶ Composantes fortement connexes \Rightarrow Modules fonctionnels potentiels



- Test de cohérence (M.J. Herrgård et al., 2006, IRISA A. Siegel et al., 2006)
 - ▶ *Escherichia coli* (JP Shaffer et al., 2002)
 - ▶ *Saccharomyces cerevisiae* (N. Guelzim et al., 2002)



- Prédications qualitatives (IRISA A. Siegel et al., 2006)
 - ▶ Signe des sommets
 - ▶ Signe des interactions

Conclusion

- Raisonnement automatique chez la levure et E.Coli
- ? Application aux animaux?

Construction des graphes d'influences chez les animaux

Difficulté

Comment construire les graphes d'influences chez les animaux ?

Solution

- Exploiter les bases de données bibliographiques de relations entre molécules
 - ☹ Travail difficile
 - ☹ Qualité des bases inconnue
 - ⇒ On n'est pas allé plus loin

Construction à partir de bases de connaissances

- Gardon : base interne **experte dédiée**
 - ☹ Peu d'éléments
 - 😊 Fiable
 - 😊 Contient certains les éléments importants
 - ⇒ Base de référence

- TRANSPATH et Ingenuity : bases généralistes commerciales
 - 😊 Beaucoup d'éléments
 - ☹ Contenu inconnue
 - ⇒ Evaluation de leur contenu

Evaluer la qualité des bases

- Peut on extraire un graphe d'influences ?
- Quelle base exploiter en priorité pour exploiter un jeu de données relatif au métabolisme des lipides ?
 - ▶ Quantité de bibliographie relative à ce jeu
 - ▶ Quantité de sommets relative à ce jeu
 - ▶ Quantité d'influences relative à ce jeu
- Faut il exploiter plusieurs bases ?
 - ▶ Publications communes ?
 - ▶ Sommets communs ?
 - ▶ Influences communes ?

Extraction d'un graphe d'influence

- Extraction automatique à partir des trois bases de :
 - ▶ 9 667 influences et 5 903 sommets.
- Extraction laborieuse
 - ☹ Beaucoup de développement spécifique d'une base
 - ☹ Problème de synonymes
 - ☹ Interface d'Ingenuity inadaptée au haut débit
 - ☹ Ingenuity ne définit pas ses concepts

Conclusion

- L'extraction automatique de graphes d'influences est possible mais laborieuse
- L'extraction automatique d'Ingenuity est très difficile

Quelles bases exploiter en priorité ?

- Sélection sur un jeu de données relatif au métabolisme énergétique
 - ▶ Gènes différenciellement exprimés lors de la mise à jeûn chez le poulet
 - ▶ Gènes issus d'expertise
- ⇒ Bases relatives au métabolisme énergétique (notées [R])

Base	ID pubmed	Sommets	Influences
Gardon [R]	151	247	411
TRANSPATH [R]	1 979	4 101	6 426
Ingenuity [R]	3 156	2 427	6 033

Conclusion

- Ingenuity contient beaucoup de bibliographie
- Ingenuity et TRANSPATH \gg Gardon

Faut il exploiter plusieurs bases ?

- La base experte est indispensable pour la bibliographie et les influences
 - ▶ Gardon \cap TRANSPATH + Ingenuity
 - ▶ **ID pubmeds** 35%
 - ▶ **Sommets** 71%
 - ▶ **Influences** 11%
- Les sommets sont communs mais pas la bibliographie et les influences
 - ▶ TRANSPATH \cap Ingenuity
 - ▶ **ID pubmeds** 3.4%
 - ▶ **Sommets** 35%
 - ▶ **Influences** 1.5%

Conclusion

- Il faut construire une base experte
- Il faut exploiter plusieurs bases pour obtenir des influences.
- Si ce n'est pas possible il faut exploiter TRANSPATH en priorité

Conclusion

- L'extraction de graphes d'influences est possible mais difficile
 - L'extraction de graphes d'influences est dépendante des sources
- ⇒ Exploiter un maximum de sources
- ⇒ A long terme : construire des sources exploitables

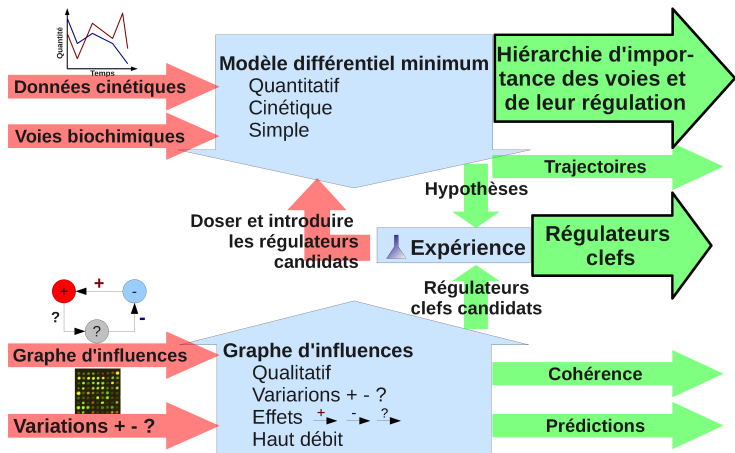
Conclusions

- Données cinétiques \Rightarrow Modèle simple dynamique
 - ▶ Squelette générique \Rightarrow Voies importantes génériques
 - ▶ Modèle dynamique
 - ★ Minimum \Rightarrow Voies importantes pour expliquer un jeu de données
 - ★ Ascendants \Rightarrow Modèles comparables (sauvage vs *knockout* PPAR α)
 - ★ Identification d'une nouvelle régulation \Rightarrow expérience
- Données transcriptomique \Rightarrow Analyse classique
 - ▶ Clustering \Rightarrow Principales fonctions annotés régulées au cours des conditions expérimentales
 - ▶ Analyse manuelle de quelques gènes \Rightarrow meilleure interprétation des clusters
 - ▶ L'exploitation manuelle de la bibliographie est l'étape limitante
- Exploitation automatique de la bibliographie \Rightarrow Graphes d'influences
 - ▶ Construction d'un graphe d'influence
 - ▶ Forte complémentarité entre les bases de données
 - ▶ \Rightarrow Exploiter plusieurs bases

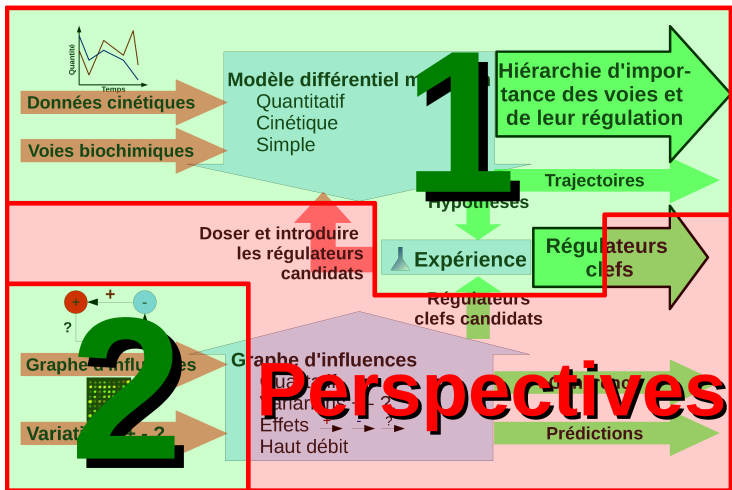
Conclusions : Construction de modèles

- Le niveau de détail du modèle correspond à celui des données
 - ☹ Adapter les modèles aux données
 - 😊 Obtenir des données adaptées aux modèles
 - ⇒ Les expérimentation doivent êtres pensées en parallèle de leur modélisation
- La formalisation des connaissances est un problème complexe
 - ☹ La construction de modèle haut débit dépend des données formalisées
 - ☹ Problème d'unicité des éléments
 - ☹ Problème de définition des concepts (Ingenuity)
 - ⇒ A long terme : investir dans la formalisation de connaissances

Plan : perspectives



Plan : perspectives



- Améliorer le graphe d'influences
 - ▶ Traiter les synonymes entre bases (CAS)
 - ▶ Exploiter Pathway Commons
- Exploiter le graphe d'influences
 - ▶ Appliquer les méthodes existantes chez la levure et *Escherichia coli*
 - ▶ Analyse croisée : topologie, gènes coexprimés, annotation fonctionnelle
- Intégration verticale
 - ▶ Identifier des éléments clef dans le graphe d'influence
 - ▶ Mesurer leur variation au cours du temps
 - ▶ Introduire ces éléments dans le modèle dynamique
 - ▶ Rechercher des régulateurs de D6D dans le graphe d'influences

Remerciements

- Jury
 - ▶ **Rapporteurs** Jean Pierre Mazat, David James Sherman
 - ▶ **Examineurs** David Causeur, Maela Kloareg, Sandrine Lagarrigue, Anne Siegel
- Collaborateurs
 - ▶ **Encadrement** Anne Siegel, Jaap Van Milgen, Florence Gondret, Sandrine Lagarrigue, Ovidiu Radulescu, Pascal Martin, Hervé Guillou
 - ▶ **Autres** Daniel Sauvart, Jean Sébastien Pierre, Sophie Vincent, François Moreews, Claire Nedellec, Pauline Gloaguen
- **Accueil** INRA : UR66, SENAH, PHASE, INRA-Agrocampus GARen, IRISA Symbiose