

Modélisation in silico de la régulation d'un métabolisme (le métabolisme des acides gras) :

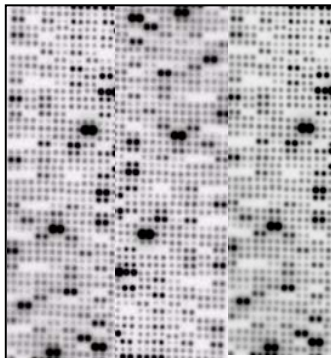
Intérêt, méthodes/outils, limites

Sandrine Lagarrigue - UMR GA - INRA Agrocampus, Rennes
Anne Siegel – IRISA - CNRS, Rennes

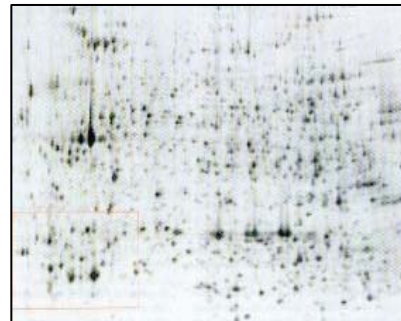
Objectifs

Contribuer à une « meilleure » interprétation
i.e. « plus globale », « plus intégrée »
des données expérimentales de type « omique »

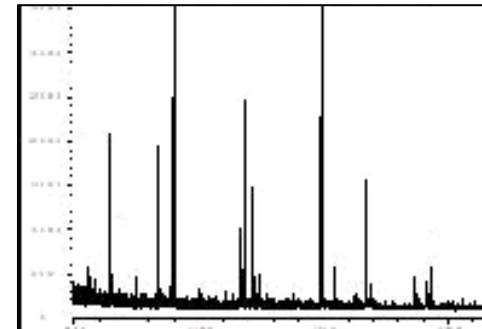
transcriptome



protéome



métabolome



Comment ?

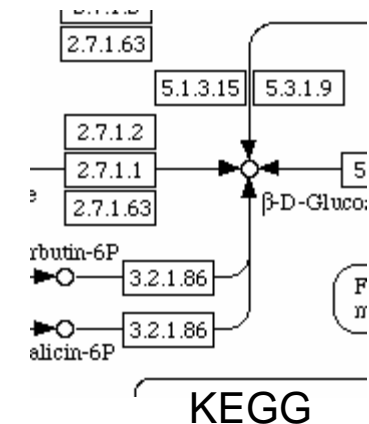
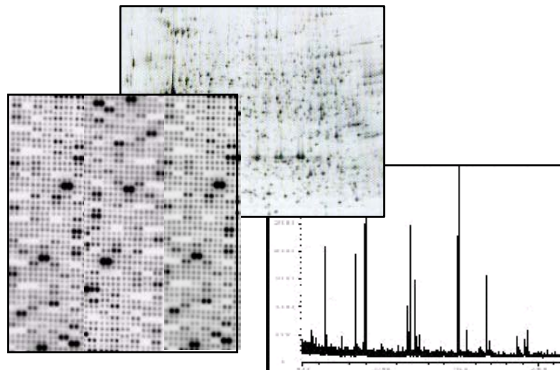
Développer des méthodes et outils
permettant de trouver les cohérences/incohérences

entre

données expérimentales
(en « omique »)

et

connaissances du moment
(bibliographie + autres sources)



Prenons un exemple

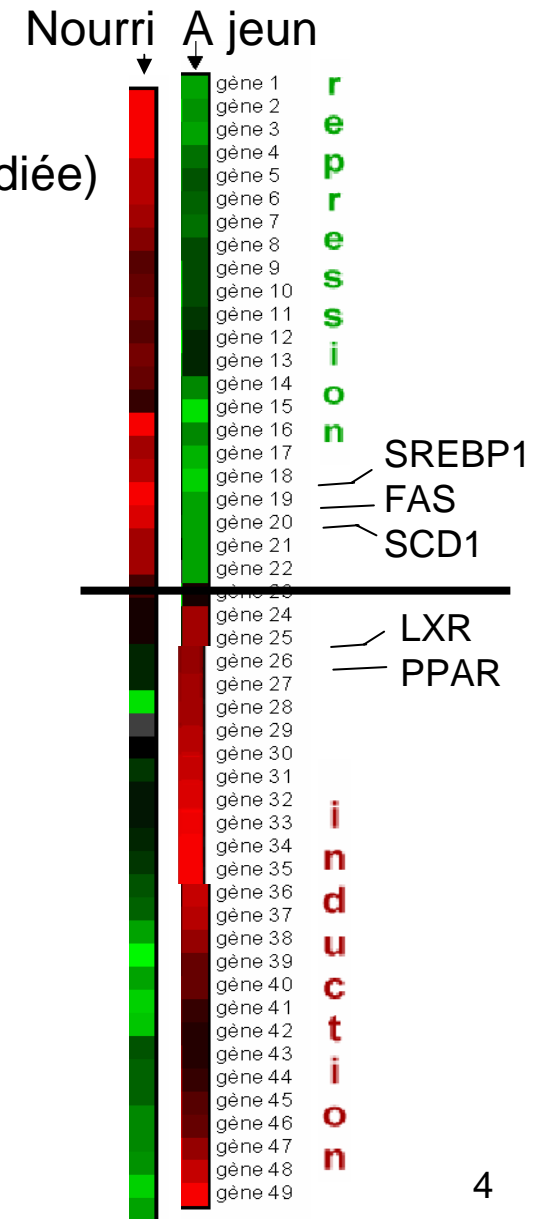
- **Données expérimentales utilisées :**
 - Profils transcriptomiques d'~300 gènes (puce à ADN dédiée)
 - Foies d'animaux Nourri *versus* A jeun

→ Une cinquantaine de gènes
sur/sous exprimés entre les 2 états

- **Comment interpréter l'ensemble de ces variations ?**

→ sont-elles cohérentes / incohérentes
par rapport aux connaissances du moment

• **Voyons la démarche**
au travers d'un petit exemple d'un réseau de 5 gènes

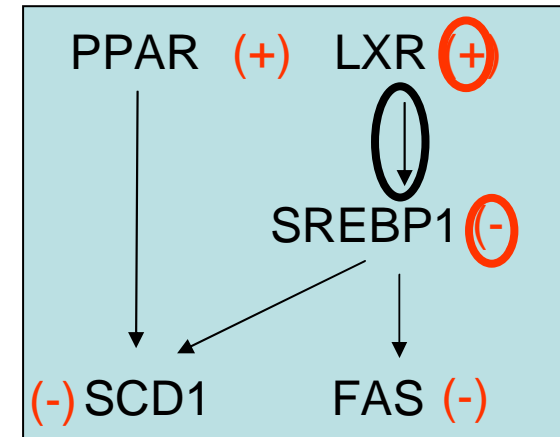


Petit exemple autour d'un réseau de 5 gènes

- **Rassembler les connaissances du moment**

- LXR induit la transcription de SREBP1
- SREBP1 induit la transcription de FAS et de SCD1 (Enzyme de la lipogénèse)
- PPAR induit la transcription de SCD1

→1) **Créer un modèle de connaissances**



→2) **Confronter les données expérimentales**

et le modèle de connaissances

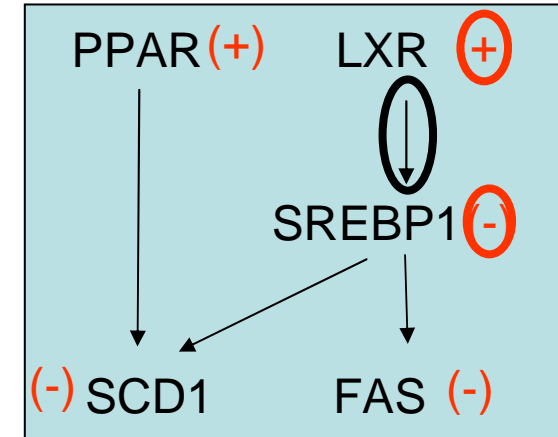
- Entre les états nourri & à jeun,
- LXR, PPAR augmentent (+)
 - SREBP, FAS, SCD1 diminuent (-)

2.1- Identifier les incohérences

entre **modèle** et **données expérimentales**

Identifier les incohérences et les corriger

2.1- Identifier les incohérences entre modèle et données expé.



2.2- Rendre le modèle et les données cohérentes

Identifier les données expérimentales responsables des incohérences

→ **Les vérifier expérimentalement**

Ici, vérifier l'expression de LXR et SREBP par une autre technique (RT-PCR)

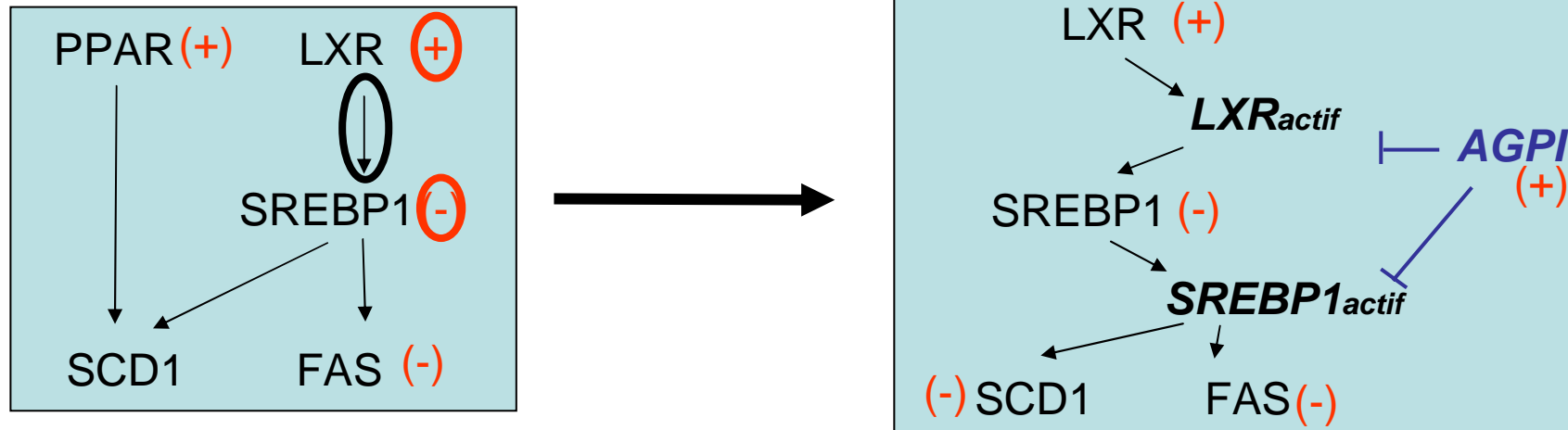
Identifier les interactions dans le modèle de connaissances posant problème

→ **Proposer de corrections du modèle**

Ici, problème au niveau de l'interaction LXR → SREBP

→ Permet d'orienter la recherche bibliographique en vue de corriger le modèle

Corriger le modèle (exemple)



→ **Si solution trouvée dans la biblio**
alors Proposition d'un nouveau modèle

Les AGPI (connus pour augmenter à l'état de jeûne) inhibent l'activité de LXR et de SREBP 1

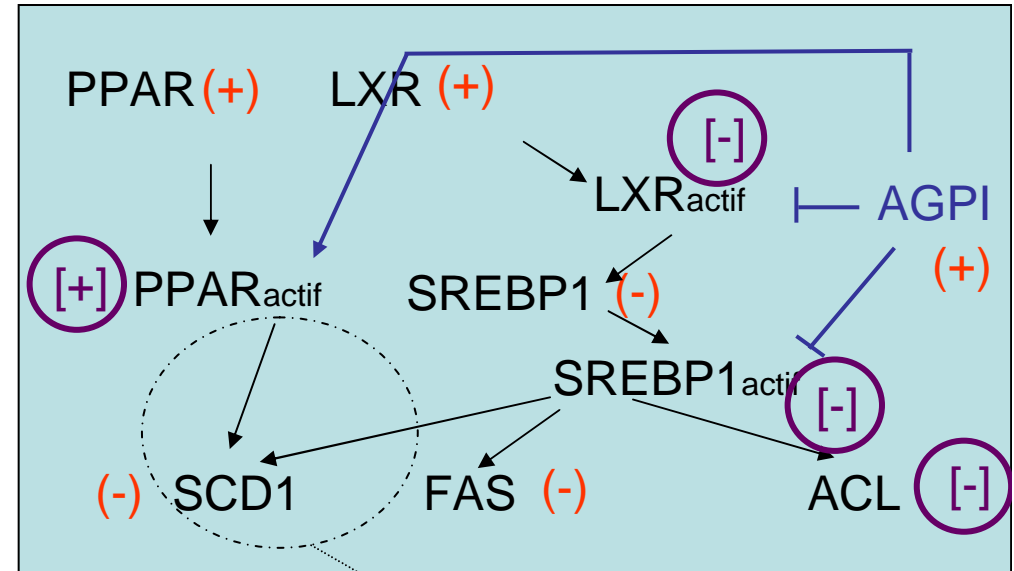
→ **Si pas de solution trouvée dans la biblio**
Alors Permet de proposer des hypothèses de travail

Etude plus fine de LXR et SREBP1
(étude au niv. protéique ou des activités)

Exploiter les données (exemple)

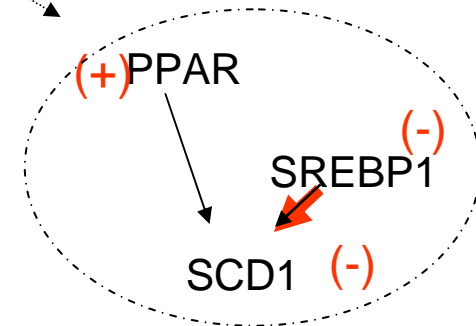
3- Aide à l'interprétation des données expérimentales

3.1- Proposer un modèle expliquant les données expé.
→ établir un réseau de régulation



3.2 Affiner l'interprétation

→ Identifier les balances/ compétitions et proposer les voies « dominantes »



3.3- Prédire les variations de produits non observés

Se rapprocher d'une vue intégrée du fonctionnement d'un système
En capitalisant les connaissances et les expérimentations successives⁸
en relation avec une thématique

Démarche générale

1. Construction d'un modèle de connaissances

Originalité du modèle : coupler connaissances métaboliques et autres (génétiques)

2. Validation du modèle en fonction des données expérimentales

2.1) Identifier les incohérences

2.2) Proposer des corrections au niveau des données expérimentales et/ou du modèle

→ 2.3) contribuer à une interprétation intégrée des données expér. (réseau de régulation = modèle)

Prédire les variations de variables non observées

3. Etudes des propriétés du modèle de connaissances

→ Identification les variables « importantes » du modèle dont la valeur influe sur l'ensemble du modèle

→ Recherche les variables sujet à des balances/compétitions

→ Identifier les variables à observer dans des futures expérimentations

→ Recherche de boucles positives et négatives de rétrocontrôle

Prédiction du comportement du système: états d'équilibre

Radulescu et al, Topology and linear response of interaction networks in molecular biology, Royal Society Interface 9(6), 2006)

Démarche générale

1. Construction d'un modèle à partir des connaissances

1- Caractéristiques du modèle

Quelques caractéristiques du modèle de connaissances

- Recenser des liens entre molécules
→ **Modèle d'interaction**

- Les molécules peuvent être des gènes/protéines ou des métabolites

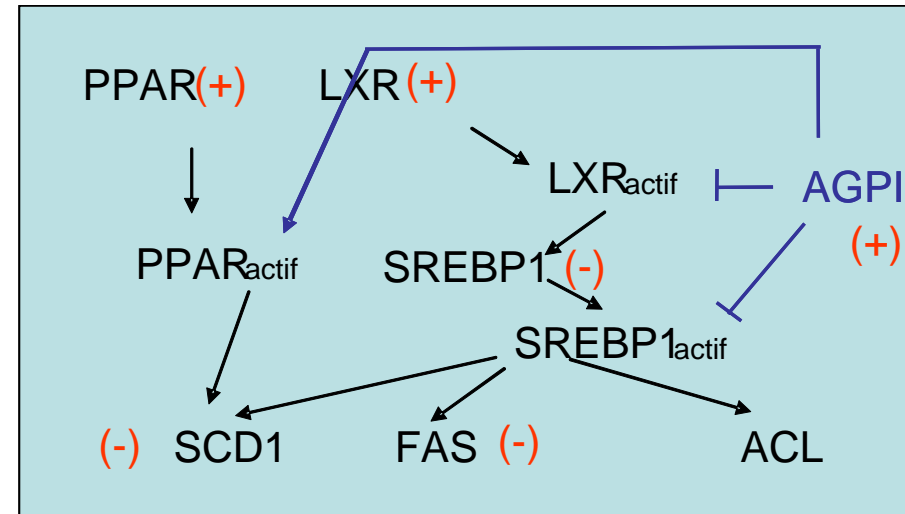
→ **Connaissances métaboliques ET connaissances de régulations**
(régulations « génétiques » et autres)

- Informations en majorité qualitatives

→ **Modèle qualitatif (pas de variables quantitatives)**

- Formalisation d'une interaction (but: exploiter des données d'expression différentielle (up/down))

→ **Interaction signée : Agent → Cible (+/-)**



Une variation de la quantité de l'agent agit sur la production de la cible

1- Alimentation d'un modèle d'interactions

- **Sources de données**

- Banques de données (Bind, Kegg) :
 - Peu de connaissances sur les régulations génétiques
- Données bibliographiques (PubMed) :
 - Données massives
 - Diffuses, non standardisées (Langage naturel)

- **1^{ère} source utilisée (en attendant le remplissage des banques):**

Données bibliographiques

Comment Extraire des informations de la biblio ?

1. **Formaliser** le type d'information à retenir
2. Créer une **base de données**
3. Extraction manuelle des publications selon une **grille de lecture**

Base de données

Description d'une interaction

- Type d'interaction (Activation/inhibition)
- Agent (Factor)
- Cible(Target)
- niv d'observ:ARN-protéine-activité
- Différents champs relatifs aux conditions expérimentales
(espèce, tissus, type cellulaire, in vivo-exvixo-invitro, localisation)

grille de lecture

Interaction - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils Aide

http://bionix.insa.fr/litterature/genetan/litterature/new/interaction_details?pmid%3Aint=0&main_product_id=0&linked_to%3Aint=0&

Agenda of Anne Siegel... Google Infos Perso Pages Webs Symbiose Cours Insa Biblio Dicos BDD Interactions CNI

User: Test user
[Main menu](#) > [Interaction](#) > [Details](#)

Interaction: Activation

Details

Enter the details of this interaction.

Article ID	<input type="text"/>
Species	<input type="text"/> Search: <input type="text"/>
Input	Target <input type="text"/> products list
	Species <input type="text"/> Search: <input type="text"/>
	Factor <input type="text"/> products list
	+ Species <input type="text"/> Search: <input type="text"/>
Output	
Regulation mode	Indirect <input type="text"/>
Observation Level	DNA <input type="text"/>
Confidence	High <input type="text"/>
Comment	<input type="text"/>
Tissue	<input type="text"/> New: <input type="text"/>
Cell type	<input type="text"/> New: <input type="text"/>
Context	<input type="text"/> New: <input type="text"/>
Location	unspecified <input type="text"/> New: <input type="text"/>
Link	<input type="checkbox"/> Check this box to enter next a related biochemical interaction

1- Un exemple d'un modèle d'interaction : régulation du métabolisme des lipides

- 4 mois de vacations
- 300 publications sélectionnées par l'expert
- 600 produits
- 1800 interactions (génétiques, métabolique...)
- Défaut : modèle très partiel
 - sélection des articles biaisée ?
 - validation de la base ?

→ Ref: Y. Bastide and al., *Une méthodologie pour l'analyse qualitative des réseaux biologiques: de la base de données à la vérification formelle*, Poster session, Jobim 2005



View Interaction - Mozilla Firefox

Fichier Edition Affichage Aller à Marque-pages Outils Aide

http://bionix.irisa.fr/litterature/genetar/

Agenda of Anne Siegel... Google Infos Perso Pages W

User: Test user
[Main Menu](#) > [Interaction](#) > [View](#)

View Interaction

PMID: [1905534](#) Volume 177 (3) 1991 -Jun -28 1056-61
Arachidonic acid suppression of fatty acid synthase gene
M K Armstrong, W L Blake, S D Clarke

Activation (behavioral)
Created on 2004-08-30 20:21:29 by Utilisateur 2 GA

Target	fatty acid synthase (FASN)
Factor	(C20:1) n-9
Regulation mode	indirect
Observation level	RNA
Confidence	High
Comment	
Species	Norway rat (<i>Rattus norvegicus</i>)
Tissue	liver
Cell type	cultured cells
Context	ex vivo

[Edit](#) [Use as template](#) [Delete](#)

[Back](#)

Démarche générale

2. Validation du modèle en fonction des données expérimentales

Recherche des incohérences

2- Recherche des incohérences

- Formalisation de la notion de **cohérence** entre modèle et données expérimentales

Solutions d'un système d'équations qualitatives

Modèle

SREBP1 = LXR

SCD1 = PPAR + SREBP1

FAS = SREBP1

Données

LXR = [+]

SREBP1 = [-]

PPAR = [+]

SCD1 = [-]

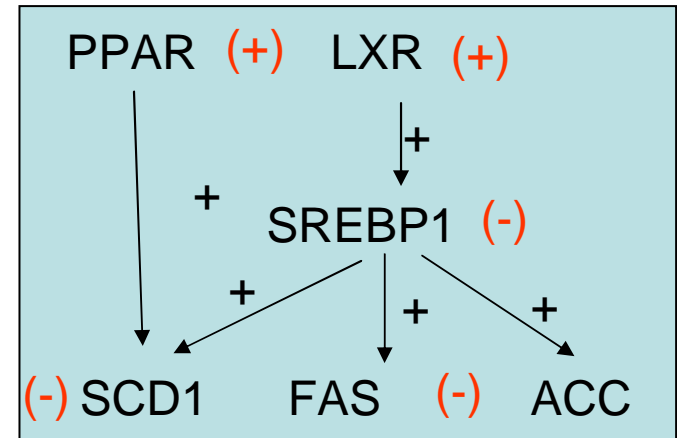
FAS = [-]

Règles de calcul

[+][+] = [+]

[-][-] = [-]

[+][-] = [?]



[+] = LXR = SREBP = [-]

Pas de solutions

- Formalisation de la notion de **compétition**
- Application à un modèle extrait de la régulation de la lipogénèse.

2- Exemple d'application

- régulation de la lipogénèse.
- 59 produits (génétiques et métaboliques)
- 110 interactions
- 16 observations

[+]	PUFA, CPT1A, LXR, RAET1C, PPAR, ACL, FACE, FDFT1
[-]	D6D, SREBF1, ACC, FAS, HMGCR, LDLR

Etape 1: incohérences modèle/données

Etape 2: utilisation des notions de compétitions pour proposer 3 corrections des données compatibles

[+]	PUFA, CPT1A, LXR, RAET1C, PPAR
[-]	D6D, SREBF1, ACC, FAS, HMGCR, LDLR, ACL, FACE, FDFT1

Etape 3: variations de 16 molécules non observées

[+]	PUFA, CPT1A, LXR, RAET1C, PPAR, ACS, ABCA1, D5D, SCD1, ABCD3, APOE, CD5L, ADAM11, CYP7A1, FBXO3, FABP1, HMGCS2, LPL, MTP, UCP2, VLDL.TG
[-]	D6D, SREBF1, ACC, FAS, HMGCR, LDLR, ACL, FACE, FDFT1

2- Implémentation

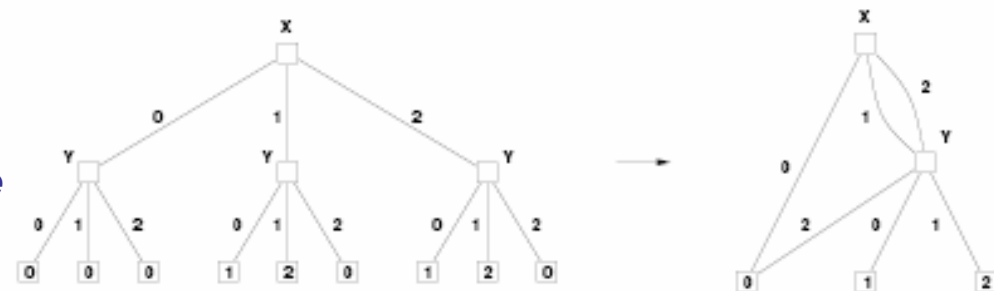
- Résolutions d'équations qualitatives: **NP-complet**
 - Trop long pour de gros modèles
- Méthode utilisée: codage des équations par des polynômes sur des corps finis

sign algebra		$\mathbb{Z}/3\mathbb{Z}$	sign algebra		$\mathbb{Z}/3\mathbb{Z}$
+	→	1	$e_1 + e_2$	→	$-\overline{e_1} \cdot \overline{e_2} \cdot (\overline{e_1} + \overline{e_2})$
-	→	-1	$e_1 \times e_2$	→	$\overline{e_1} \cdot \overline{e_2}$
?	→	0	$e_1 \approx e_2$	→	$\overline{e_1} \cdot \overline{e_2} \cdot (\overline{e_1} - \overline{e_2})$

- Implémentations: diagrammes de décisions (utilisés pour la vérification de circuits et programmes). Logiciel SIGALI

*Diminue la complexité en espace
et en temps*

(testé sur réseau de 600 sommets)



Ref: P. Veber et al., *Complex Qualitative Models in Biology: a new approach*, Proceedings of ECCS, Paris, November 2005. To appear in *Complex Us*

Perspectives

- **Construction de modèles**

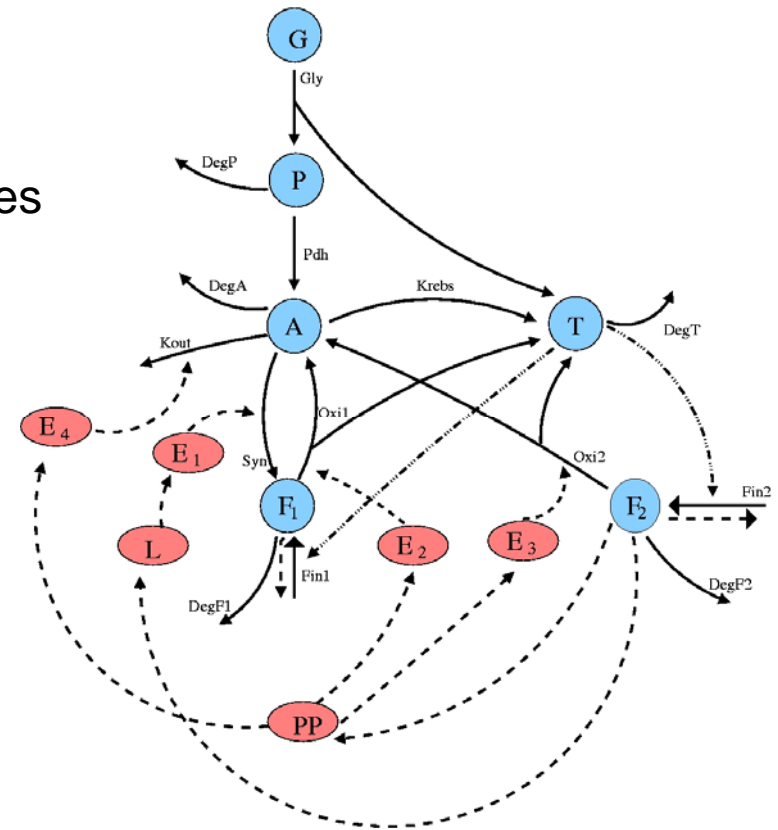
- Sources bibliographiques : Valider la méthode « manuelle »
- Inclure les banques de données en particulier KEGG, ...
- Exploiter Transfac couplé à Genomatix pour les interactions de type régulation transcriptionnelle (éléments cis prédits/validés)

- **Exploitation des modèles**

- Améliorer les heuristiques de corrections
- Augmenter la taille des réseaux interprétables

- **Enrichissement du modèle** (aller retour entre expérimentations et modélisation)

- **Abstraction du modèle.** Et méthodes d'analyse dédiées. **Inter-espèce ?**



ACI MathResoGen 2002 (IMPBio).
Collaboration IRISA. INRA/Agrocampus.
IMB Dijon. Inserm U456. Pitié-Salpêtrière.

3^{ème} Carrefour OUEST-genopole®



"OUEST-genopole® :
l'outil technologique de votre recherche"

Jeudi 26 janvier 2006
BREST | Faculté de Médecine

Programme et inscriptions : www.ouest-genopole.org



Avec le soutien de



Merci de votre attention

*Du gène à l'écosystème
Du laboratoire vers la société
De l'ouest à l'Europe*