

# Modèle d'interaction de Gibbs pour l'organisation spatiale d'un tissu biologique

## Application à la cancérogénèse

Mathieu Emily

Laboratoire TIMC - Institut de l'Ingénierie de l'Information de Santé - Grenoble

22 Mai 2006

# Présentation - Objectif

## ► Présentation

- Equipe TIMB (Techniques de l'Informatique et de l'Information en Biomédecine)
- 3<sup>ème</sup> Année de thèse sous la direction d'Olivier François
- Collaborations avec Jean-Michel Billiot et Rémy Drouilhet (Laboratoire LabSAD - Grenoble)

# Présentation - Objectif

- ▶ Présentation
  - ▶ Equipe TIMB (Techniques de l'Informatique et de l'Information en Biomédecine)
  - ▶ 3<sup>ème</sup> Année de thèse sous la direction d'Olivier François
  - ▶ Collaborations avec Jean-Michel Billiot et Rémy Drouilhet (Laboratoire LabSAD - Grenoble)
- ▶ Objectif de ce travail :
  - ▶ Quantifier les forces d'adhésion au sein d'un tissu biologique

# Plan

Problème Biologique

Modélisation

Simulation

Estimation

# Plan

Problème Biologique

Modélisation

Simulation

Estimation

# Organisation spatiale d'un tissu

- ▶ Motifs biologiques typiques :
  - ▶ Damier
  - ▶ Cluster
  - ▶ Engloutissement

# Organisation spatiale d'un tissu

- ▶ Motifs biologiques typiques :
  - ▶ Damier
  - ▶ Cluster
  - ▶ Engloutissement
- ▶ Lien avec de nombreux processus biologiques :
  - ▶ Embryogénèse
  - ▶ Morphogénèse
  - ▶ Régénérescence par agrégats (Hydra)

# Differential Adhesion Hypothesis [Steinberg, 1962]



# Differential Adhesion Hypothesis [Steinberg, 1962]

- ▶ Un tissu est configuration de cellules

# Differential Adhesion Hypothesis [Steinberg, 1962]

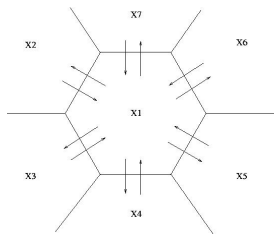
- ▶ Un tissu est configuration de cellules
- ▶ Une configuration résulte d'une minimisation d'Énergie

# Differential Adhesion Hypothesis [Steinberg, 1962]

- ▶ Un tissu est configuration de cellules
- ▶ Une configuration résulte d'une minimisation d'Énergie
- ▶ Énergie =  $f$ (Adhésion entre les cellules)

# Differential Adhesion Hypothesis [Steinberg, 1962]

- ▶ Un tissu est configuration de cellules
- ▶ Une configuration résulte d'une minimisation d'Energie
- ▶ Energie =  $f(\text{Adhésion entre les cellules})$



*Interactions entre une cellule et ses voisines*

# Niveau moléculaire

- ▶ Réseau de signalisation très complexe

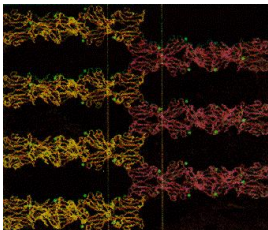
# Niveau moléculaire

- ▶ Réseau de signalisation très complexe
- ▶ Caractérisation de certaines molécules
  - ▶ Cadhérines E
  - ▶  $\alpha$ ,  $\beta$  et  $\gamma$  caténines

# Niveau moléculaire

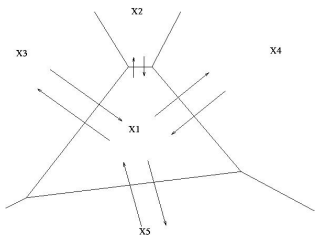
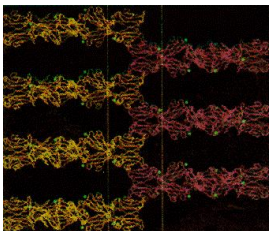
- ▶ Réseau de signalisation très complexe
- ▶ Caractérisation de certaines molécules
  - ▶ Cadhérines E
  - ▶  $\alpha$ ,  $\beta$  et  $\gamma$  caténines
- ▶ Rôle de suppresseur de tumeurs

# Complexe Cadhérine-Caténine





# Complexe Cadhérine-Caténine



*Schéma d'une membrane intercellulaire et de son implication sur la modélisation*

# Modèles Mathématiques sous l'hypothèse DAH

- ▶ Modèles sur grille [Mochizuki *et al.*, 1996]

# Modèles Mathématiques sous l'hypothèse DAH

- ▶ Modèles sur grille [Mochizuki *et al.*, 1996]
- ▶ Modèles centrés [Honda *et al.*, 1996]

# Modèles Mathématiques sous l'hypothèse DAH

- ▶ Modèles sur grille [Mochizuki *et al.*, 1996]
- ▶ Modèles centrés [Honda *et al.*, 1996]
- ▶ Modèles sur sommets [Ngai *et al.*, 2001]

# Modèles Mathématiques sous l'hypothèse DAH

- ▶ Modèles sur grille [Mochizuki *et al.*, 1996]
- ▶ Modèles centrés [Honda *et al.*, 1996]
- ▶ Modèles sur sommets [Ngai *et al.*, 2001]
- ▶ Modèles hybrides :

Le modèle de Graner et Glazier [1992]

# Le modèle de Graner et Glazier

- ▶ Chaque cellule, notée  $\sigma$ , est un ensemble de plusieurs pixels

# Le modèle de Graner et Glazier

- ▶ Chaque cellule, notée  $\sigma$ , est un ensemble de plusieurs pixels
- ▶ Chaque cellule a un type  $\tau_\sigma$  (3 types possibles)

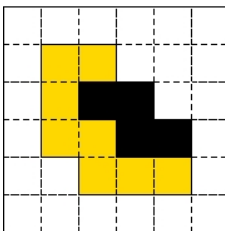
# Le modèle de Graner et Glazier

- ▶ Chaque cellule, notée  $\sigma$ , est un ensemble de plusieurs pixels
- ▶ Chaque cellule a un type  $\tau_\sigma$  (3 types possibles)
- ▶ L'énergie ou Hamiltonien d'une configuration de cellules est décrite par

$$H_{GG} = \sum_{(i,j) \sim (i',j')} J(\tau(\sigma_{ij}), \tau(\sigma_{i'j'})) (1 - \delta_{\sigma_{ij}, \sigma_{i'j'}}) + \lambda \sum_{\sigma} (a(\sigma) - A_{\tau(\sigma)})^2 \Gamma(A_{\tau(\sigma)}),$$



# Le modèle de Graner et Glazier



*Exemple de configuration à 2 cellules actives*

# Le modèle de Graner et Glazier

La discrétisation de l'espace limite l'étude :

# Le modèle de Graner et Glazier

La discrétisation de l'espace limite l'étude :

- ▶ Perte de connexité des cellules

# Le modèle de Graner et Glazier

La discrétisation de l'espace limite l'étude :

- ▶ Perte de connexité des cellules
- ▶ Algorithme sensible au pas de discrétisation

# Le modèle de Graner et Glazier

La discrétisation de l'espace limite l'étude :

- ▶ Perte de connexité des cellules
- ▶ Algorithme sensible au pas de discrétisation

Dans la suite de cette présentation, une version continue du modèle de Graner et Glazier est présentée.

# Plan

Problème Biologique

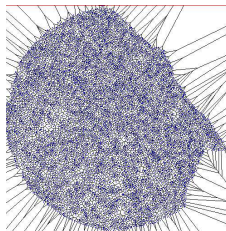
**Modélisation**

Simulation

Estimation

# Modélisation Géométrique

- ▶ D'après Honda [1978, 1983], les cellules peuvent se modéliser par un diagramme de Voronoï centré sur les noyaux des cellules biologiques.



*Exemple de tissu modélisé par un diagramme de Voronoï*

# Notations

- ▶ Une configuration marquée s'écrit :

$$\varphi = \{(x_1, \tau_1), \dots, (x_n, \tau_n)\},$$

où  $x_i$  est le centre de la cellule  $i$  et  $\tau_i$  le type de la cellule  $i$



# Notations

- ▶ Une configuration marquée s'écrit :

$$\varphi = \{(x_1, \tau_1), \dots, (x_n, \tau_n)\},$$

où  $x_i$  est le centre de la cellule  $i$  et  $\tau_i$  le type de la cellule  $i$

- ▶  $|\text{Dir}(x_i)|$  est l'aire de la cellule de Dirichlet modélisant la cellule  $i$

# Notations

- ▶ Une configuration marquée s'écrit :

$$\varphi = \{(x_1, \tau_1), \dots, (x_n, \tau_n)\},$$

où  $x_i$  est le centre de la cellule  $i$  et  $\tau_i$  le type de la cellule  $i$

- ▶  $|\text{Dir}(x_i)|$  est l'aire de la cellule de Dirichlet modélisant la cellule  $i$
- ▶  $|\text{Dir}(x_i \cap x_j)|$  est la longueur de l'arête de Dirichlet commune aux cellules  $i$  et  $j$

# Hamiltonien

Nous posons :

$$H(\varphi) = \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i})$$

# Hamiltonien

Nous posons :

$$H(\varphi) = \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i})$$

La densité de la configuration s'écrit :

$$f(\varphi) = \frac{\exp(-\theta H(\varphi))}{Z}$$

où  $\theta$  quantifie la force d'adhésion du tissu

# Etude du modèle

Notre modèle s'inscrit dans le contexte des Processus Ponctuels Marqués Markoviens [Van Lieshout, 2000]

**Théorème [Stabilité locale]**

Soit  $\varphi = \{(x_1, \tau_1), \dots, (x_n, \tau_n)\}$ , une configuration marquée. Supposons  $J$  bornée ( $|J| \leq K_J$ ) et qu'il existe  $d$  et  $D$  ( $0 < d < D < \infty$ ) pour lesquelles :

$$x_i \sim x_j \implies d \leq |x_i - x_j| \leq D$$

Sous ces hypothèses, il existe, pour tout  $x_{n+1} \in \mathbb{R}^2$  et tout  $\tau_{n+1} \in M$ , une constante  $C(K_J, \lambda, d, D)$  telle que :

$$|H(\varphi \cup (x_{n+1}, \tau_{n+1})) - H(\varphi)| \leq C(K_J, \lambda, d, D)$$

# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$H(\varphi') - H(\varphi)$$

# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$\begin{aligned} H(\varphi') - H(\varphi) \\ = \sum_{i \sim j} |\text{Dir}'(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \end{aligned}$$

# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$\begin{aligned} H(\varphi') - H(\varphi) &= \sum_{i \sim j} |\text{Dir}'(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\ &\quad - \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) - \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \end{aligned}$$



# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$\begin{aligned}
 & H(\varphi') - H(\varphi) \\
 &= \sum_{i \sim j} |\text{Dir}'(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &\quad - \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) - \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &= \sum_{i \sim n+1} |\text{Dir}'(x_i \cap x_{n+1})| J(\tau_i, \tau_{n+1})
 \end{aligned}$$

# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$\begin{aligned}
 & H(\varphi') - H(\varphi) \\
 &= \sum_{i \sim j} |\text{Dir}'(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &\quad - \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) - \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &= \sum_{i \sim n+1} |\text{Dir}'(x_i \cap x_{n+1})| J(\tau_i, \tau_{n+1}) \\
 &\quad + \lambda (|\text{Dir}'(x_{n+1})| - A_{\tau_{n+1}})^2 \Gamma(A_{\tau_{n+1}})
 \end{aligned}$$

# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$\begin{aligned}
 & H(\varphi') - H(\varphi) \\
 &= \sum_{i \sim j} |\text{Dir}'(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &\quad - \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) - \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &= \sum_{i \sim n+1} |\text{Dir}'(x_i \cap x_{n+1})| J(\tau_i, \tau_{n+1}) \\
 &\quad + \lambda (|\text{Dir}'(x_{n+1})| - A_{\tau_{n+1}})^2 \Gamma(A_{\tau_{n+1}}) \\
 &\quad + \sum_{\text{Triangles}(i,j,n+1)} (|\text{Dir}'(x_i \cap x_j)| - |\text{Dir}(x_i \cap x_j)|) J(\tau_i, \tau_j)
 \end{aligned}$$

# Preuve

Nous posons  $\varphi' = \varphi \cup (x_{n+1}, \tau_{n+1})$

$$\begin{aligned}
 & H(\varphi') - H(\varphi) \\
 &= \sum_{i \sim j} |\text{Dir}'(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda \sum_i (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &\quad - \sum_{i \sim j} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) - \lambda \sum_i (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) \\
 &= \sum_{i \sim n+1} |\text{Dir}'(x_i \cap x_{n+1})| J(\tau_i, \tau_{n+1}) \\
 &\quad + \lambda (|\text{Dir}'(x_{n+1})| - A_{\tau_{n+1}})^2 \Gamma(A_{\tau_{n+1}}) \\
 &\quad + \sum_{\text{Triangles}(i,j,n+1)} (|\text{Dir}'(x_i \cap x_j)| - |\text{Dir}(x_i \cap x_j)|) J(\tau_i, \tau_j) \\
 &\quad + \lambda \sum_{i; i \sim n+1} (|\text{Dir}'(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}) - (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i})
 \end{aligned}$$

# Plan

Problème Biologique

Modélisation

**Simulation**

Estimation

# Algorithme de Métropolis-Hastings

## Algorithme

- ▶ Si  $U < 1/2$  : Insertion

# Algorithme de Métropolis-Hastings

## Algorithme

- ▶ Si  $U < 1/2$  : Insertion
  - ▶ Choix uniforme de  $x_{n+1}$  et  $\tau_{n+1}$

# Algorithme de Métropolis-Hastings

## Algorithme

- ▶ Si  $U < 1/2$  : Insertion
  - ▶ Choix uniforme de  $x_{n+1}$  et  $\tau_{n+1}$
- ▶ Sinon : Suppression



# Algorithme de Métropolis-Hastings

## Algorithme

- ▶ Si  $U < 1/2$  : Insertion
  - ▶ Choix uniforme de  $x_{n+1}$  et  $\tau_{n+1}$
- ▶ Sinon : Suppression
  - ▶ Choix uniforme d'un point dans la configuration

# Algorithme de Métropolis-Hastings

## Algorithme

- ▶ Si  $U < 1/2$  : Insertion
  - ▶ Choix uniforme de  $x_{n+1}$  et  $\tau_{n+1}$
- ▶ Sinon : Suppression
  - ▶ Choix uniforme d'un point dans la configuration
- ▶ Proba. d'acceptation :  $p = \min[1, \exp(-\theta(\Delta H))]$

# Algorithme de Métropolis-Hastings

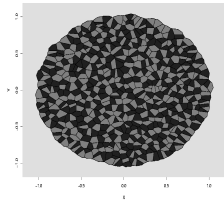
## Algorithme

- ▶ Si  $U < 1/2$  : Insertion
  - ▶ Choix uniforme de  $x_{n+1}$  et  $\tau_{n+1}$
- ▶ Sinon : Suppression
  - ▶ Choix uniforme d'un point dans la configuration
- ▶ Proba. d'acceptation :  $p = \min[1, \exp(-\theta(\Delta H))]$

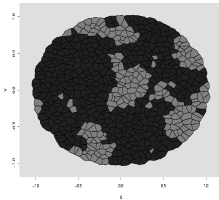
## Théorème

La chaîne de Markov générée par l'algorithme de Métropolis-Hastings associé à  $H$  est Harris-Récurrente

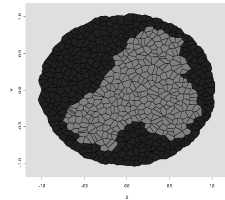
# Exemple de Simulations



(a)



(b)



(c)

*Exemples de configurations biologiques classiques : (a) Damier, (b) Aggrégation, (c) Engloutissement*

# Plan

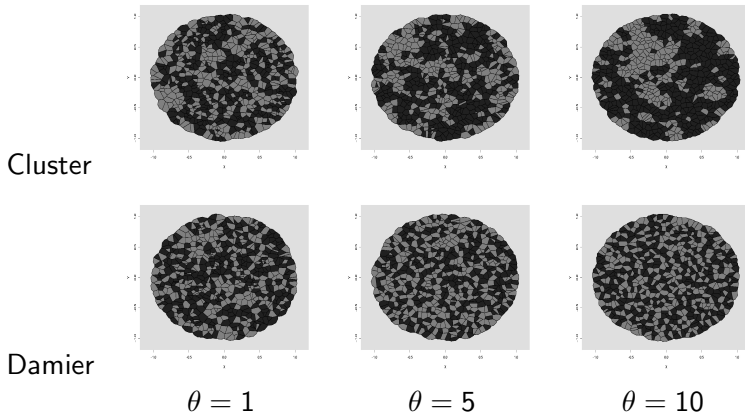
Problème Biologique

Modélisation

Simulation

Estimation

# Influence du paramètre de force d'adhésion $\theta$



# Pseudo-Vraisemblance

Conditionnellement à la réalisation, nous avons :

$$PL(\theta) = \prod_{(x_i, \tau_i) \in \varphi} \text{Prob}((x_i, \tau_i) | \varphi, \theta)$$

où

$$\text{Prob}((x_i, \tau_i) | \varphi, \theta) = \frac{\exp(-\theta H_\varphi(x_i, \tau_i))}{\sum_{m \in M} \exp(-\theta H_\varphi(x_i, m))}.$$

$H_{\varphi(x_i, \tau_i)}$  est la contribution de  $i$  dans  $H$  :

$$H_\varphi(x_i, \tau_i) = \sum_{j \sim i} |\text{Dir}(x_i \cap x_j)| J(\tau_i, \tau_j) + \lambda (|\text{Dir}(x_i)| - A_{\tau_i})^2 \Gamma(A_{\tau_i}).$$

# Maximum de Pseudo-Vraisemblance

## Définition

Le logarithme de la pseudo-vraisemblance s'écrit :

$$L(\theta) = - \sum_{(x_i, \tau_i) \in \varphi} \left( \theta H_\varphi(x_i, \tau_i) + \log \sum_{m \in M} \exp(-\theta H_\varphi(x_i, m)) \right).$$

Un estimateur de la force d'adhésion (paramètre  $\theta$ ) est donné par :

$$\hat{\theta}(\varphi) = \operatorname{argmax}_\theta L(\varphi, \theta)$$



Performances de  $\hat{\theta}$ 

	Damier		Agrégation	
	Moyenne	Variance	Moyenne	Variance
$\theta = 1$	0.98	0.70	1.03	0.4
$\theta = 3$	3.14	0.66	3.01	0.51
$\theta = 5$	5.01	0.57	4.94	0.94
$\theta = 8$	8.20	1.07	8.01	0.81
$\theta = 10$	10.47	1.20	9.80	1.00
$\theta = 12$	12.28	1.81	12.05	1.09
$\theta = 15$	14.58	2.22	15.03	1.20
$\theta = 20$	20.44	3.55	20.08	2.98

*Moyenne et Variance obtenues pour 100 réplifications*

# Discussion - Conclusion

- ▶ Modèle spatial continu d'un tissu multi-cellulaire

# Discussion - Conclusion

- ▶ Modèle spatial continu d'un tissu multi-cellulaire
- ▶ Etude mathématique d'un algorithme de simulation

# Discussion - Conclusion

- ▶ Modèle spatial continu d'un tissu multi-cellulaire
- ▶ Etude mathématique d'un algorithme de simulation
- ▶ Estimateur du paramètre de la force d'adhésion  $\hat{\theta}$

# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles

# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles
  - ▶ Améliorer la fonction  $J$  grâce à la biologie

# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles
  - ▶ Améliorer la fonction  $J$  grâce à la biologie
  - ▶ Prospecter d'autres domaines d'applications

# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles
  - ▶ Améliorer la fonction  $J$  grâce à la biologie
  - ▶ Prospector d'autres domaines d'applications
- ▶ Amélioration du modèle
  - ▶ Estimer les coefficients  $J$



# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles
  - ▶ Améliorer la fonction  $J$  grâce à la biologie
  - ▶ Prospector d'autres domaines d'applications
- ▶ Amélioration du modèle
  - ▶ Estimer les coefficients  $J$
  - ▶ Proposer de nouveaux estimateurs de graphes aléatoires

# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles
  - ▶ Améliorer la fonction  $J$  grâce à la biologie
  - ▶ Prospecter d'autres domaines d'applications
- ▶ Amélioration du modèle
  - ▶ Estimer les coefficients  $J$
  - ▶ Proposer de nouveaux estimateurs de graphes aléatoires
  - ▶ Proposer un modèle continu sur les marques

# Perspectives

- ▶ Utilisation du modèle
  - ▶ Application à un jeu de données réelles
  - ▶ Améliorer la fonction  $J$  grâce à la biologie
  - ▶ Prospecter d'autres domaines d'applications
- ▶ Amélioration du modèle
  - ▶ Estimer les coefficients  $J$
  - ▶ Proposer de nouveaux estimateurs de graphes aléatoires
  - ▶ Proposer un modèle continu sur les marques
  - ▶ Etude de la transition de phase

Merci pour votre attention