

# Comment la structure globale d'un réseau d'interactions de protéines évolue?

Par Joel Sandé and Jacques Diby

Basé sur un article de  
Andreas Wagner

# Plan

## 1 Notion d'interaction PPI et de réseaux d'interaction

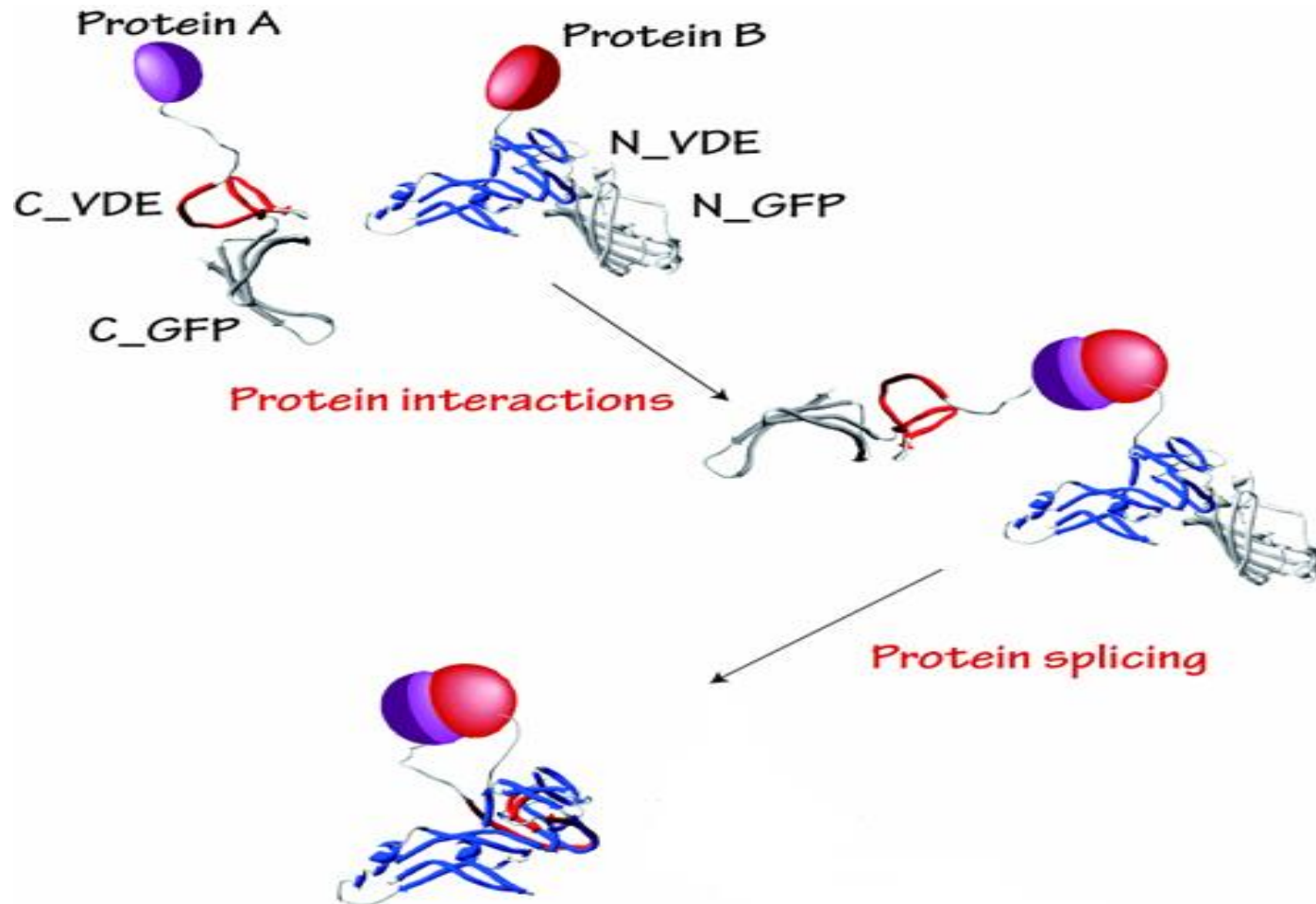
- Interaction protéine-protéine
- Modèles de réseaux

## 2 L'article

- L'hypothèse (Null)
- Evolution du réseau par duplication de gène
- Notion de divergence
- Evolution par addition ou délétion d'interactions
- Règles de la loi de distribution
- Limites
- Conclusion

# Notions de PPI et de réseaux

- Les interactions protéine-protéine:



Pourquoi étudier les interactions entre protéines

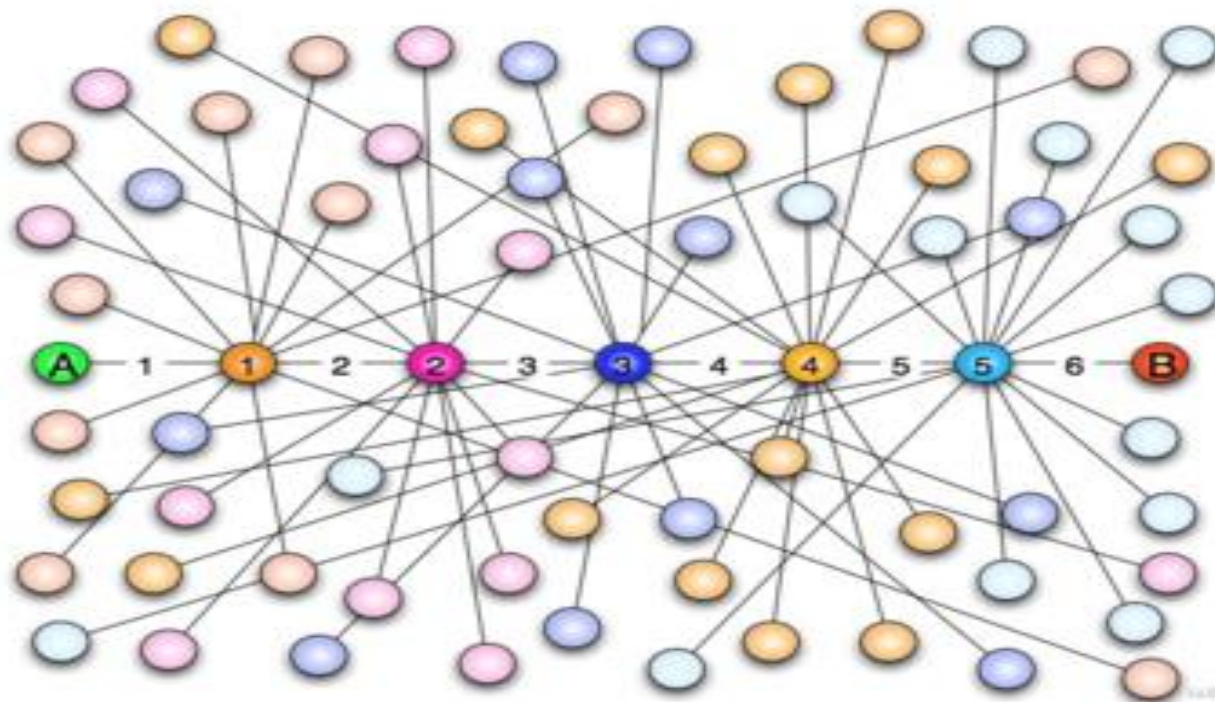


- ➔ Réalisent les réactions catalytiques, de transport etc.
- ➔ Transmettent l'information de l'ADN via l'ARN en traversant les membranes
- ➔ Forment des canaux de communications et règlent leurs comportements
- ➔ Rendent possible la synthèse de nouvelles protéines
- ➔ Sont responsables de la dégradation de protéines qui ne sont plus nécessaires
- ➔ Permettent les réponses immunitaires
- ➔ ...

# Notions de PPI et de réseaux

## Qu'est ce qu'un réseau ?

Un **réseau** est un ensemble de *nœuds* (ou *pôles*) reliés entre eux par des *liens* (ou *canaux*).



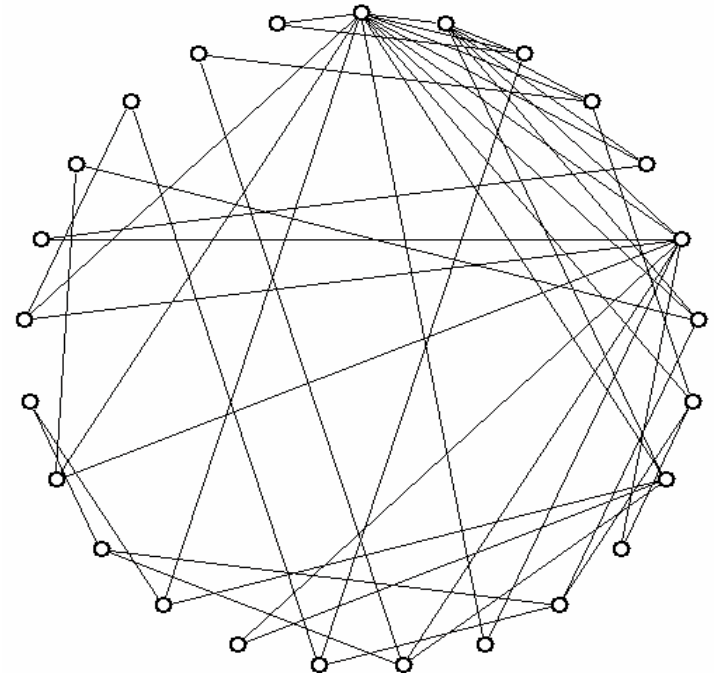
# Notions de PPI et de réseaux

- Modèles de réseaux existant
  - ➔ Réseaux aléatoires
  - ➔ Réseaux “sans échelle” (Scale free)
  - ➔ Réseaux hiérarchiques

# Notions de PPI et de réseaux

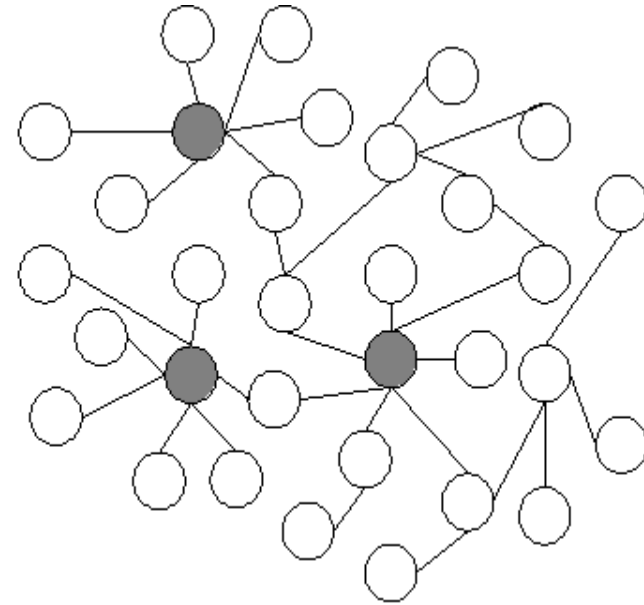
## ○ Réseaux aléatoires

- Le modèle de Erdős-Rényi (ER) d'un réseau aléatoire commence par  $N$  noeuds et relie chaque paire de noeuds avec une probabilité  $p$ , et cela génère un graphe d'approximativement,  $pN(N-1)/2$  liens placés aléatoirement.
- Les degrés des noeuds suivent une distribution de Poisson



# Notions de PPI et de réseaux

- Réseaux “sans échelle” (Scale free)
- Les propriétés du réseau sont déterminés par les **hubs**
- Le réseau est souvent généré par un processus de croissance appelé “modèle de Barabási-Albert”
- $P(d) \sim d^{-\gamma}$ , ou  $\gamma$  est le degré de l'exposant.
- Les réseaux sans échelle ont un exposants de degré  $2 < \gamma < 3$ , un intervalle qui est souvent observé dans les réseaux biologiques et non-biologiques comme l'Internet, le World Wide Web, le réseau des réactions métaboliques et le graphe des appels téléphoniques.

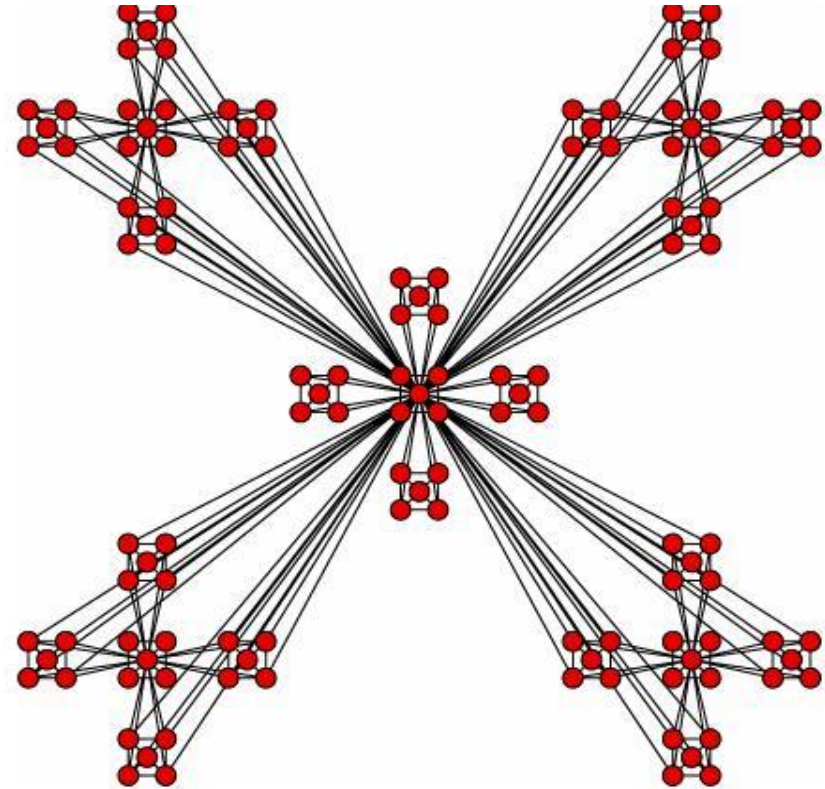




# Notions de PPI et de réseaux

## ○ Réseaux hiérarchiques

- Pleins de petits clusters très connectés se combinent dans des grands clusters, peu nombreux et moins connectés qui se combinent dans des clusters même plus grands, et même moins connectés
- Le coefficient de clusterisation d'un nœud avec  $k$  liens suit la loi d'échelle:



## Comment la structure globale d'un réseau d'interaction entre protéines évolue?

Deux processus peuvent influencer la structure de ce type de réseau:

1. La duplication de gène augmentant le nombre de protéines et leurs interactions
2. Addition et élimination des interactions entre les protéines

Les taux de ces processus seront estimés à partir de données du Saccharomyces cerevisiae (100 interactions → 1 millions d'années)

# L'article

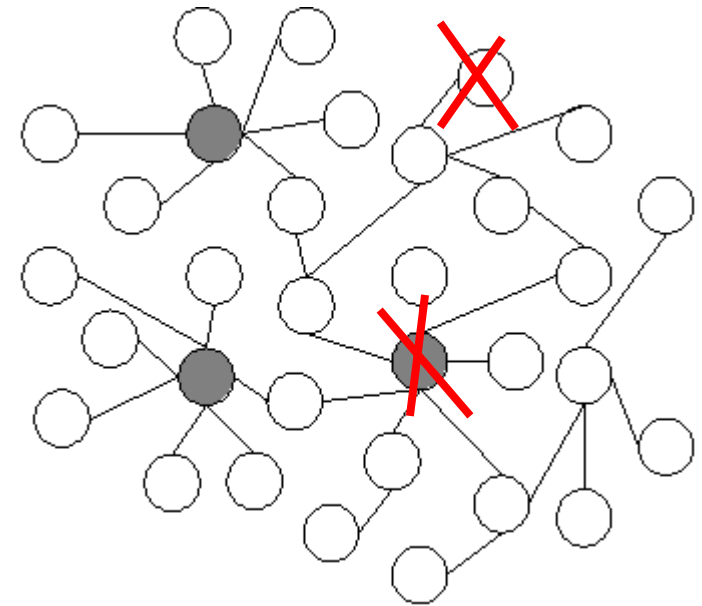
Une fois la structure d'un réseau connue, questions sur origine évolutive et connexion fonctionnelle:

1. Est-ce que la structure du réseau nous renseigne sur la fonction du réseau ?
2. Comment la sélection naturelle serait arrivée à façonner la structure globale ?
3. Une distribution de degrés en loi de puissance transmet-elle un quelconque avantage à un organisme?
4. La sélection naturelle aurait-elle favorisé uniquement la survie des organismes qui respectaient cette loi de puissance?

# L'article

## Propriétés des réseaux d'interaction entre protéines:

- Très peu de nœuds fortement connectés
- Lors du retrait d'un nœud choisi au hasard
- Petit réseau VS grand réseau
- $P(d) \sim d^{-\gamma}$



# L'article/ l'hypothèse

## Une hypothèse « nulle » sur l'origine de la structure des réseaux:

Le terme « nulle » ne veut pas dire que l'hypothèse n'est pas valable!

La sélection naturelle ne doit être invoquée que si toutes les hypothèses nulles sont rejetées!

Pour fournir une hypothèse nulle:

- Examiné les processus qui influencent la structure de l'interaction des protéines du réseau chez la levure
- Estimé leurs taux à partir de données empiriques et tenté d'apporter des explications sur la distributions des degrés sans invoquer ...

# L'article

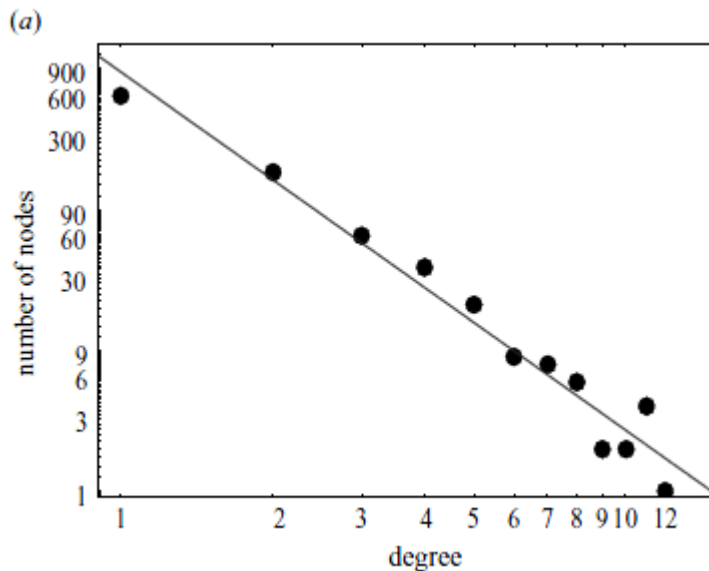
Caractéristique du réseau global sont indépendante de l'approche expérimentale:

L'interprétation biologique des réseaux d'interaction protéique est entravée par plusieurs facteurs :

- Elle réduit les informations spatiales et temporelles en un arrêt sur image statique du réseau
- Les cartes générées montrent un nombre d'interaction inférieur à ce qu'on peut observer expérimentalement

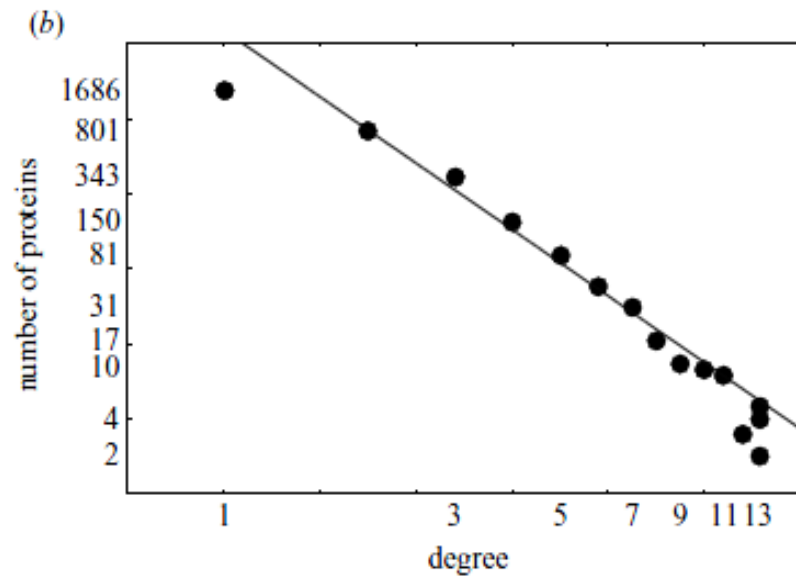
# L'article

La loi de puissance est un robuste principe dans la fonctionnalité d'un réseau d'interaction indépendant de l'approche expérimentale (Uetz *et al.* 2000)



(a):  $n=985$  protéines  
 $k=899$  interactions

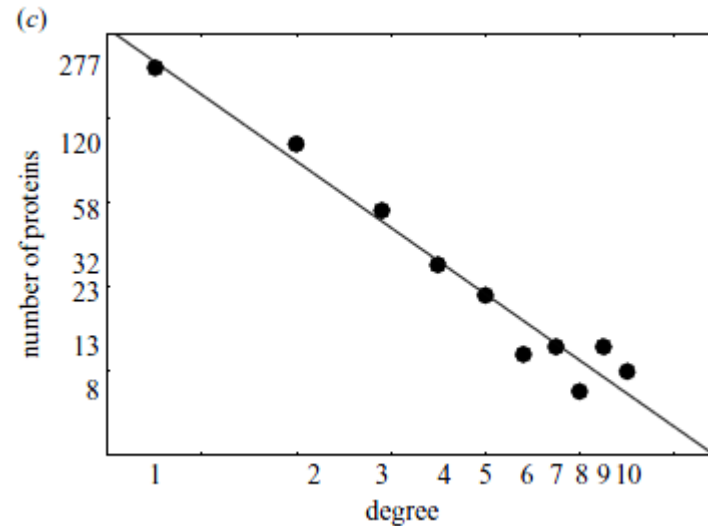
$(P(d) \propto d^{-2.55 \pm 0.35})$



(b):  $n=780$  protéines  
 $k=747$  interactions

$(P(d) \propto d^{-2.43 \pm 0.35})$

# L'article



Munich Information Center for Protein Sequences (MIPS)

(c):  $n=680$  protéines  
 $k=899$  interactions

$(P(d) \propto d^{-2.67 \pm 0.20})$

Les trois figures montrent que la loi de distribution est suivie



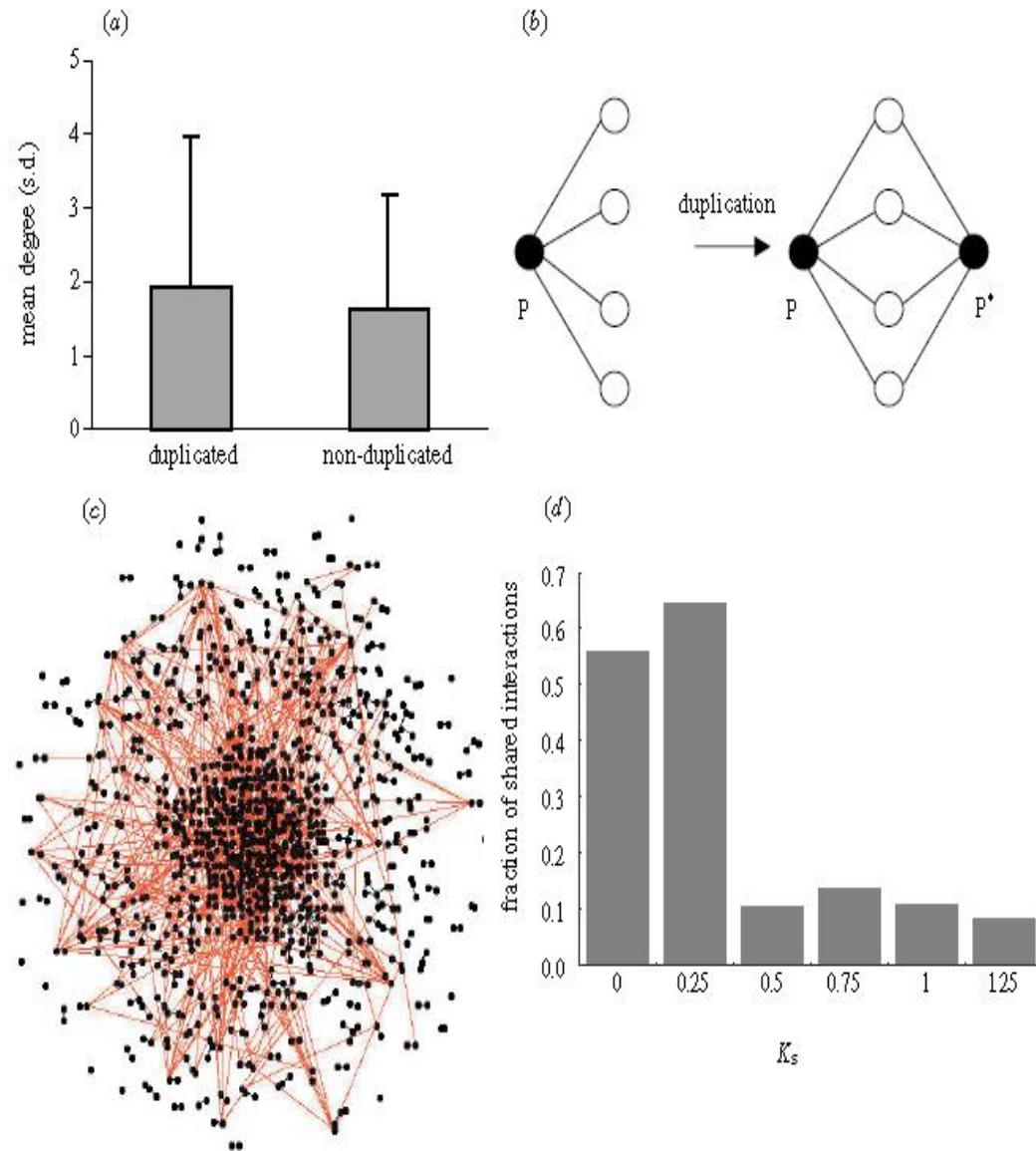
# L'article/ duplication de gènes

Est-ce que les gènes fortement liés ont peu de duplications dans le génome?

→ Duplication du gène serait lié à la structure du réseau

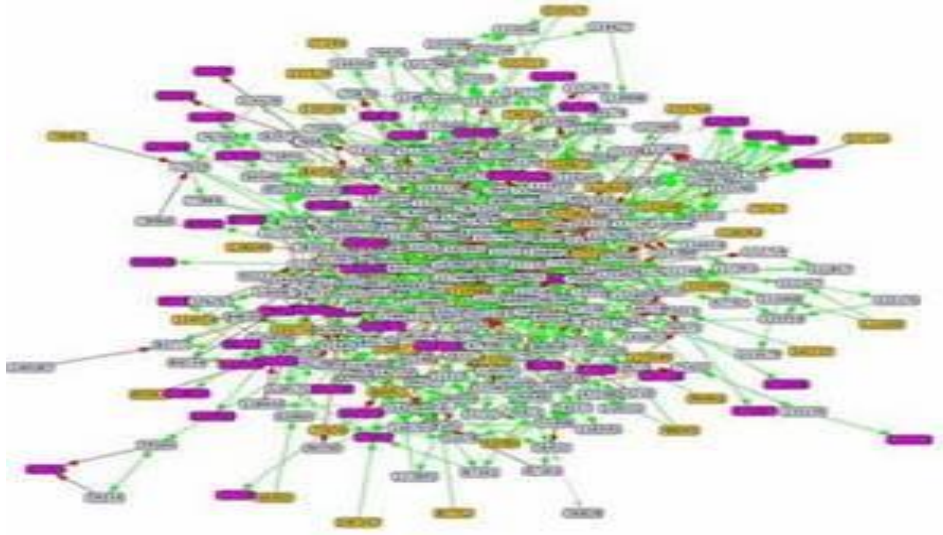
→ Les protéines dont le degré d'interaction est augmenté sont les partenaires d'interaction de la protéine double. Or seuls les hub ont cette tendance

→ les duplications pourraient très bien matcher avec les protéines périphériques



# L'article/ duplication de gènes

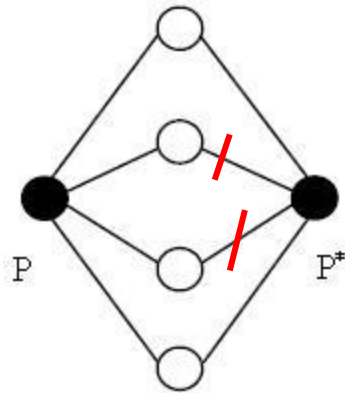
Dans un réseau, la fréquence des protéines dotées d'un seul partenaire d'interaction va lentement se rapprocher de zéro mais les aspects de la fonction d'un gène se perdent rapidement après la duplication; les interactions entre protéines dérivées divergent très rapidement



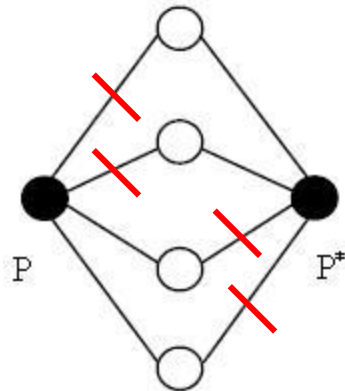
Quel est l'effet de cette divergence sur l'évolution des degrés de distribution ?

# L'article/Divergences

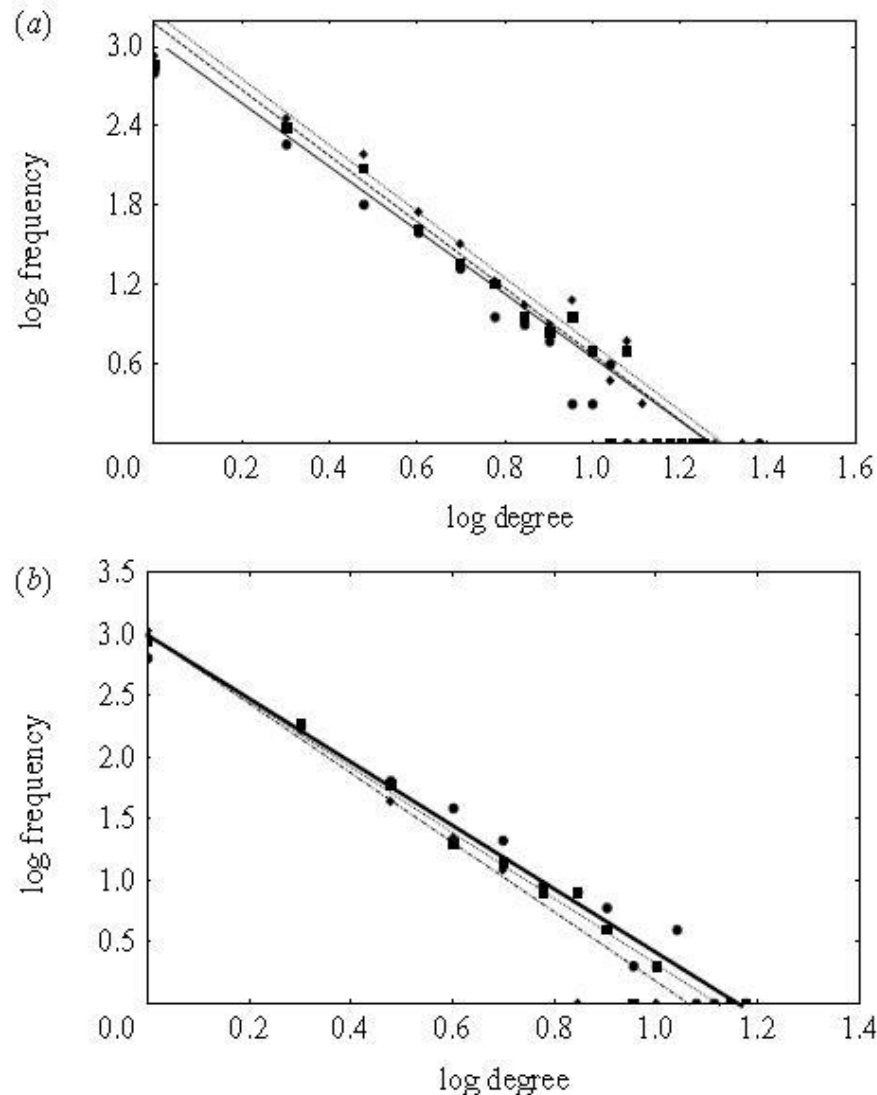
- Divergence asymétrique



- Divergence symétrique



# L'article/Divergences



La duplication et la divergence, pris ensemble, n'affectent pas le degré de distribution

- **divergence asymétrique**

cercles ( $\gamma = 2.55$ );

1000 duplications (1200 Myr ; 2.25);

2000 duplications (2400 Myr ; 2.29).

- **divergence symétrique**

cercles ( $\gamma = 2.55$ );

1000 duplications (1200 Myr ; 2.58);

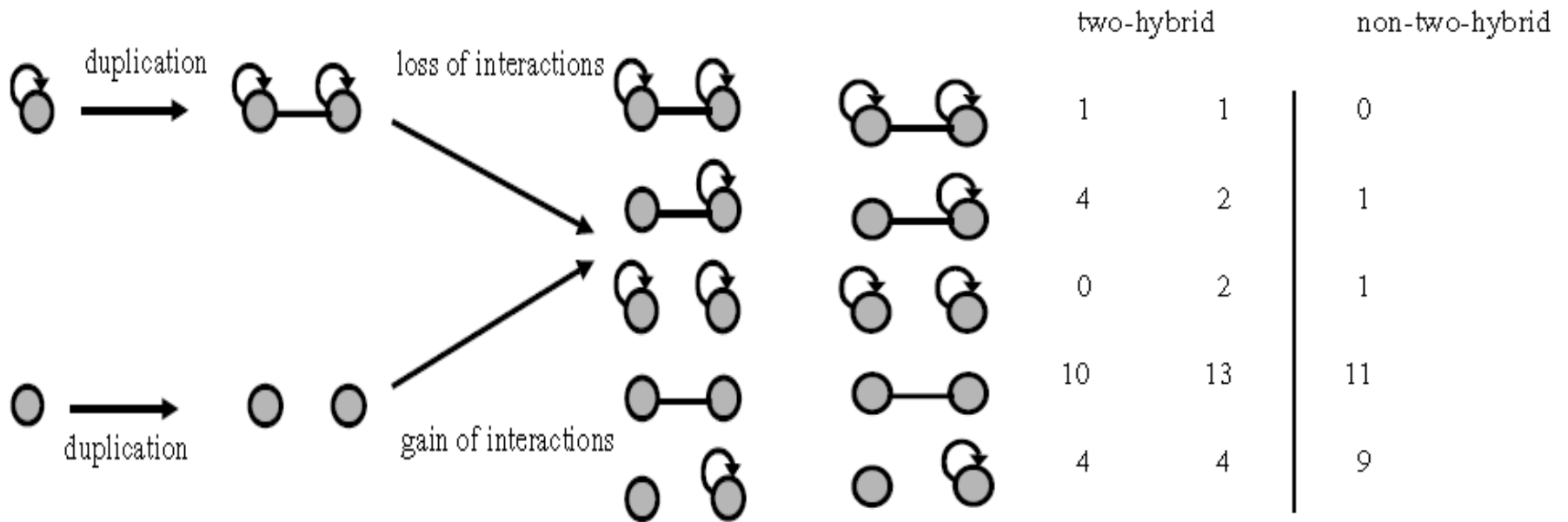
2000 duplications (2400 Myr ; 2.82).

# L'article/ Addition et suppression

- L'ajout et la suppression des interactions physiques entre protéines est causée par des mutations des protéines engendrant des changements de surfaces
- La vitesse d'interaction est-elle suffisamment grande pour influencer le réseau ?
- Ce taux peut-il être estimé ? Ce qui suit est une tentative d'estimation, basée sur la vitesse à laquelle les interactions sont ajoutées au réseau

# L'article/ Addition et suppression

comment l'auto-et la cross-interactions peuvent évoluer après duplications de gènes?



# L'article/ Addition et suppression

Supposons qu'il y avait un déséquilibre considérable entre les taux:

$c+$ : les interactions sont ajoutées

$c-$ : les interactions sont éliminées

Si par exemple  $c+ / c- = 2/1$ , ou Inversement, si  $c+ / c- = 1: 2$ ,

le nombre d'interactions par nœud pourrait soit disparaître ou exploser pendant un court laps de temps.

Il n'existe aucune preuve d'un changement radical de la sorte. Par exemple, l'interaction des protéines chez le procaryote *Helicobacter pylori*, indique que le taux de gain et la perte de l'interaction est à peu près égales

$$(c+ \approx c-).$$

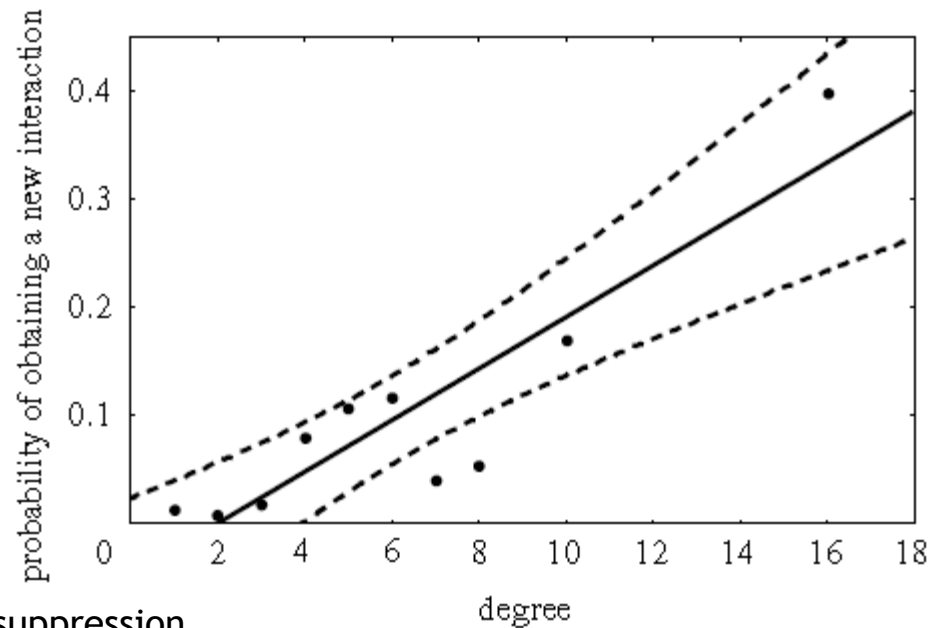
→ Les additions sont compensées par les suppressions.

# L'article/ Addition et suppression

- Basé sur les travaux de (Dorogovtsev & Mendes 2000; Albert & Barabasi 2002)

$c+e$  = probabilité qu'un nouveau nœud s'attache préférentiellement à une protéine périphérique ( $c+e = 0,3$ )

$c+n$  = probabilité qu'un nouveau nœuds s'attache à un hub ( $c+n = 0,7$ )



Il en est de même pour la suppression



# L'article/ Règles de la loi

$$C+ = C- \quad ?$$

Deux principes généraux sont suffisants pour obtenir des réseaux respectant la loi de puissance de degré distributions (Albert & Barabasi 2002).

- Tout d'abord, des nœuds doivent être ajoutés au réseau, même si ce n'est que de temps en temps.
- Deuxièmement, les nouvelles interactions doivent être susceptibles d'impliquer d'avantage les nœuds fortement connectés que des nœuds avec peu de connexions → "les riches deviennent plus riches "

$$C+ > C-$$

# L'article/ Règles de la loi

- Pour montrer la dépendance entre  $\gamma$  et les probabilités  $c_{+n}$ ,  $c_{+e}$ , et  $c_{-}$

$$\frac{\partial d(s,t)}{\partial t} = (c_{+n} + 2c_{+e} + 2c_{-}) \frac{d(s,t)}{\int_0^t d(s,t) du}.$$

a combien peut-on estimer  $\gamma$  ?

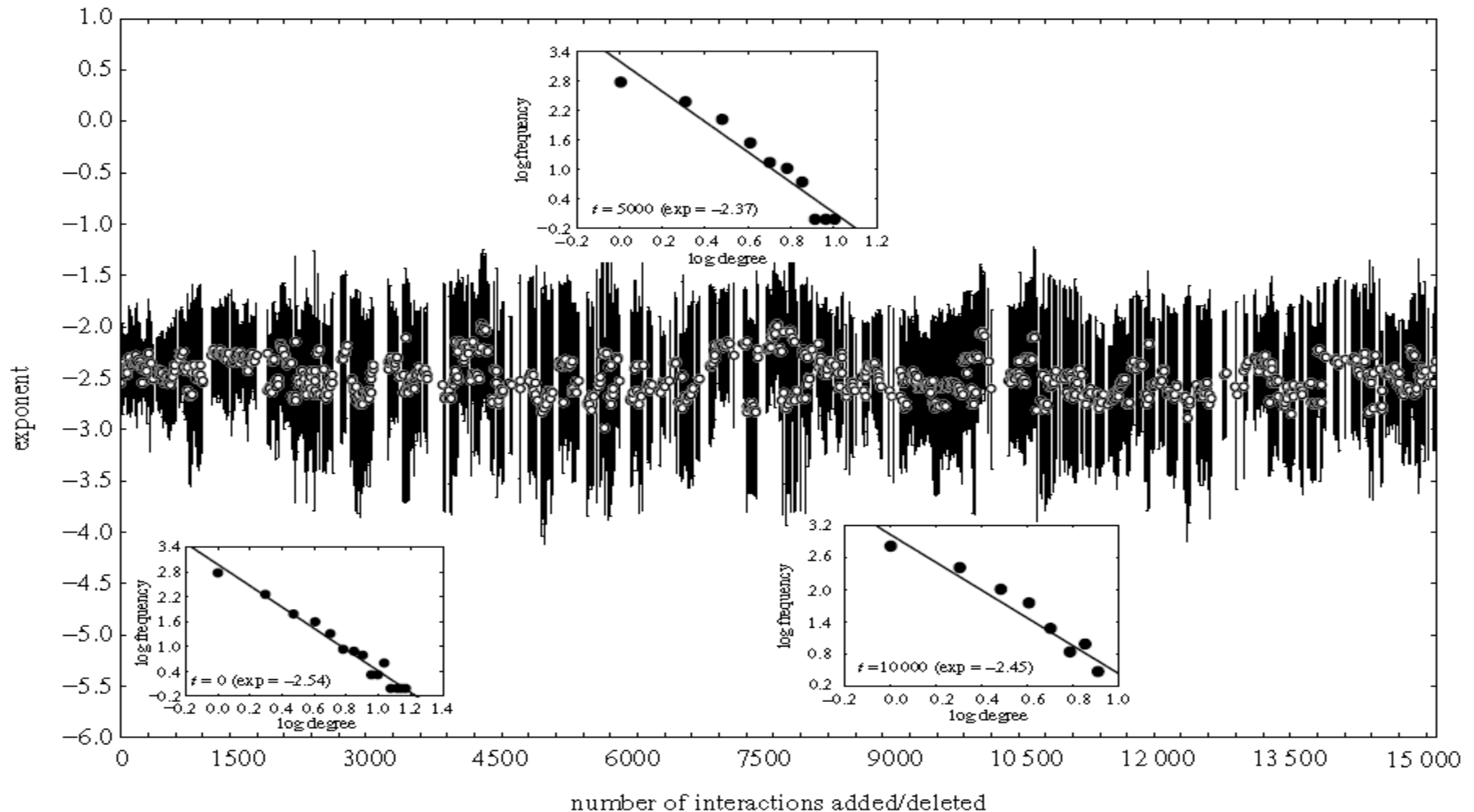
$$\gamma = 2 + c_{+n} / [c_{+n} + 2c_{+e} - 2c_{-}]$$

$$P(d) = d^{-\gamma}$$

$$\gamma > 2$$

# L'article/ Règles de la loi

## Constance de gamma



# L'article/ Limites

- Premièrement, les données disponibles sont des interactions protéiques de qualité limitée. Cependant, la structure du réseau est robuste face aux variations expérimentales .
- Deuxièmement, une incertitude est impliquée dans l'estimation de la divergence des gènes dupliqués (pour  $K_s > 1$ ).
- Troisièmement, Bien que les duplications des gènes puissent avoir dominés l'évolution des génomes à court et moyen des échelles de temps, le réarrangement des exons peut avoir dominé l'évolution du génome au début de l'évolution de la vie.
- Quatrièmement, on ne peut pas dire avec certitude que les divergences d'interactions après duplications de gènes évoluent seulement en perte de leurs interactions communes.

# Conclusion

- Le lien entre la structure et la fonction des réseaux PPI est au cœur des débats philosophiques entre les biologistes (squelette, conformation des protéines)
- Un test expérimental pour juger de la distribution semble presque impossible
- Il faut être prudent lorsqu'on parle de sélection naturelle pour caractériser un réseau
- A la question de savoir quel rôle joue la sélection naturelle dans l'évolution de la structure d'un réseau il est probable que toutes les hypothèses émises soient même l'objet de cette sélection naturelle