

# **Génomique évolutive**

**M2 Graduate School « Biologie évolutive »**

**M2 Méthodes informatiques et statistiques pour les omiques (MISO)**

**Resp : Eléonore Durand (eleonore.durand@univ-lille.fr)**

**Co-resp : Sylvain Legrand (sylvain.legrand@univ-lille.fr)**

**Évolution - Écologie – Paléontologie (UMR CNRS 8198),**

**Bureau 204, Bâtiment SN2**

**Université de Lille**

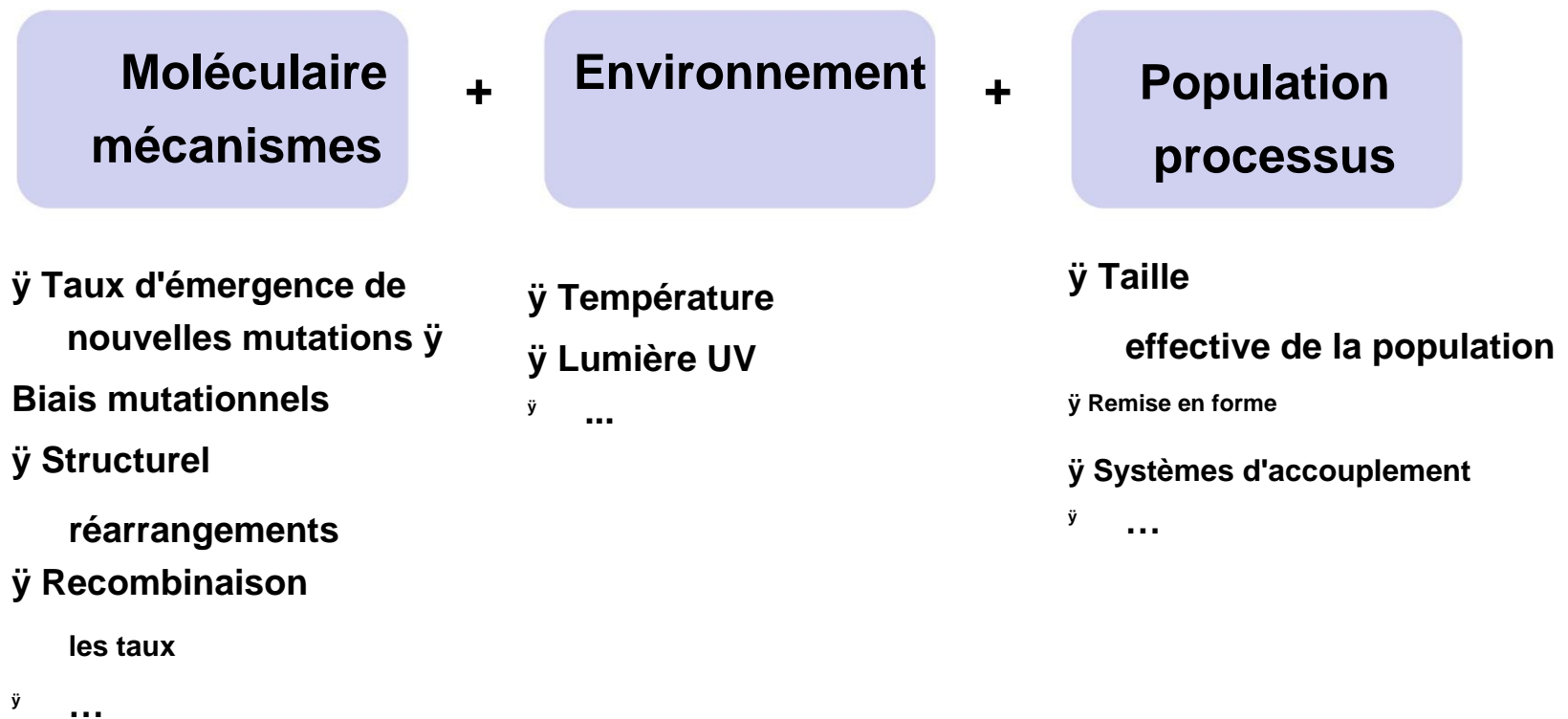
# Génomique évolutive

**Comment les génomes évoluent-ils dans le temps ?**

- Quelle est la taille ou la taille des génomes en termes de nombre de bases d'ADN qu'ils contiennent ? Les organismes plus « complexes » ont-ils plus de gènes ?**
- Comment l'information dans un génome est-elle organisée ?**
- D'où viennent les gènes et les séquences non géniques et comment évoluent-ils ?**
- Quels sont les processus moléculaires et évolutifs (à la fois adaptatifs et non adaptatifs) à l'origine de l'évolution des génomes ?**
- Et bien d'autres questions...**

# Génomique évolutive

Comment les génomes évoluent-ils dans le temps ?



# **Contenu et organisation du génome en tant que traits moléculaires**

• **Taille du génome**

• **Proportions codantes et non codantes**

• **Taille et densité des gènes**

• **Teneur en ARN non codants**

• **Contenu GC**

• **Contenu des éléments transposables**

• **Expression (niveaux transcriptionnel et traductionnel)**

• ...

**Comment ces traits varient-ils et  
évoluent-ils ?**

## **ÊTRE CONSCIENT!**

- ÿ Il existe de nombreux génomes disponibles de qualités différentes (assemblage, nb de lacunes, annotation) selon les méthodes**
- ÿ Les études génomiques comparatives ont besoin de génomes de haute qualité ! (surtout en se concentrant sur des régions intergéniques, plus difficile à faire l'assemblage)**
- ÿ Les annotations géniques et intergéniques dépendent fortement des données, des méthodes et des seuils choisis !**

# Évaluation

**ÿ Examen final 70 % : évaluation écrite sur les cours théoriques et appliqués et analyse de documents**

**ÿ Examen continu 30 % :**

**Travail en équipe de 2-3 étudiants**

**Analyse d'une publication scientifique**

**Mémoire de recherche ou de méthode**

**Présentation orale le 3 octobre (MISO) ou le 6 octobre (Evobio)**

# Planification

	Conférences			TD EvoBio	TD MISO
La semaine			Conférencier	Jeudi 13h-16h15	Lundi 13h-16h15
36	lundi 09/05 14h45	CM1. Introduction, évolution de la taille du génome	DE		
36	Mercredi 09/07 10h15	CM2. Origines de nouveaux gènes	DE		
37	lundi 09/12 14h45	CM3. Taux de mutation : spectre et évolution	DE		
37	Mercredi 14/09 10h15	CM4. Éléments transposables : processus et évolution	SL		
38	Mercredi 21/09 10h15	CM5. Origine et évolution des petits ARN	ED	Ouvrage bibliographique 22/09	Ouvrage bibliographique 09/19
39	Mercredi 28/09 10h15	CM6. Introns : origine et évolution	CV		
40		Pas de conférence EG		Présentation orale 10/06	Présentation orale 10/03
41	Mercredi 10/12 10h15	CM7. Détection des signatures de sélection	XV		
42	Mercredi 10/19 10h15	CM8. Détection des signatures de sélection	XV		
50		Examen final (évaluation écrite) : 14 décembre			

ED : Eléonore Durand, SL : Sylvain Legrand, VC : Vincent Castric, XV : Xavier Vekemans

## **Objectifs de l'UE Génomique Évolutive**

**ÿ Acquérir des connaissances sur l'architecture et la diversité du génome à travers divers organismes (espèces modèles et non modèles)**

**ÿ Être capable de décrire et de discuter des processus évolutifs impliqués dans la génération de l'architecture du génome**

**ÿ S'initier aux méthodes/concepts utilisés dans les études de génomique comparative**



# **Présenter**

## **Introduction à la génomique évolutive**

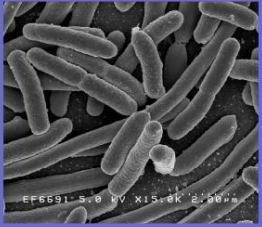
**I. Variation de la taille du génome, structure et paradoxe de la valeur C**

**II. Évolution de la taille du génome: moléculaire et évolutive  
processus**

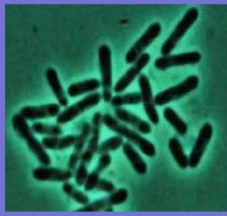
**III. Modèles d'évolution de la taille du génome**

# I. Variation de la taille du génome

## Procaryotes



Escherichia coli 5Mb



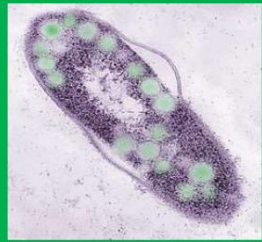
Sorange  
cellulose 13 Mo



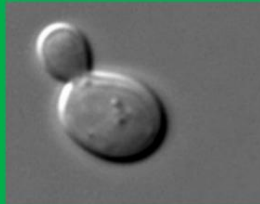
Genlisea tuberosa 61 Mo



Paris japonica  
~150 000 Mo



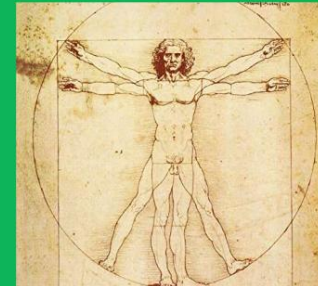
Encephalitozoon  
intestinalis 2.25 Mo



Saccharomyces  
cerevisiae 12 Mo



Drosophile  
melanogaster  
118 Mo



Homo sapiens ~3100 Mo



Mus musculus 2700 Mo

## Eucaryotes

# I. Variation de la taille du génome

**Grande variation de la taille du génome  
au sein et entre les principaux groupes  
de vie!**

**Valeur C<sub>y</sub>: quantité totale d'haploïdes  
d'ADN dans une cellule**

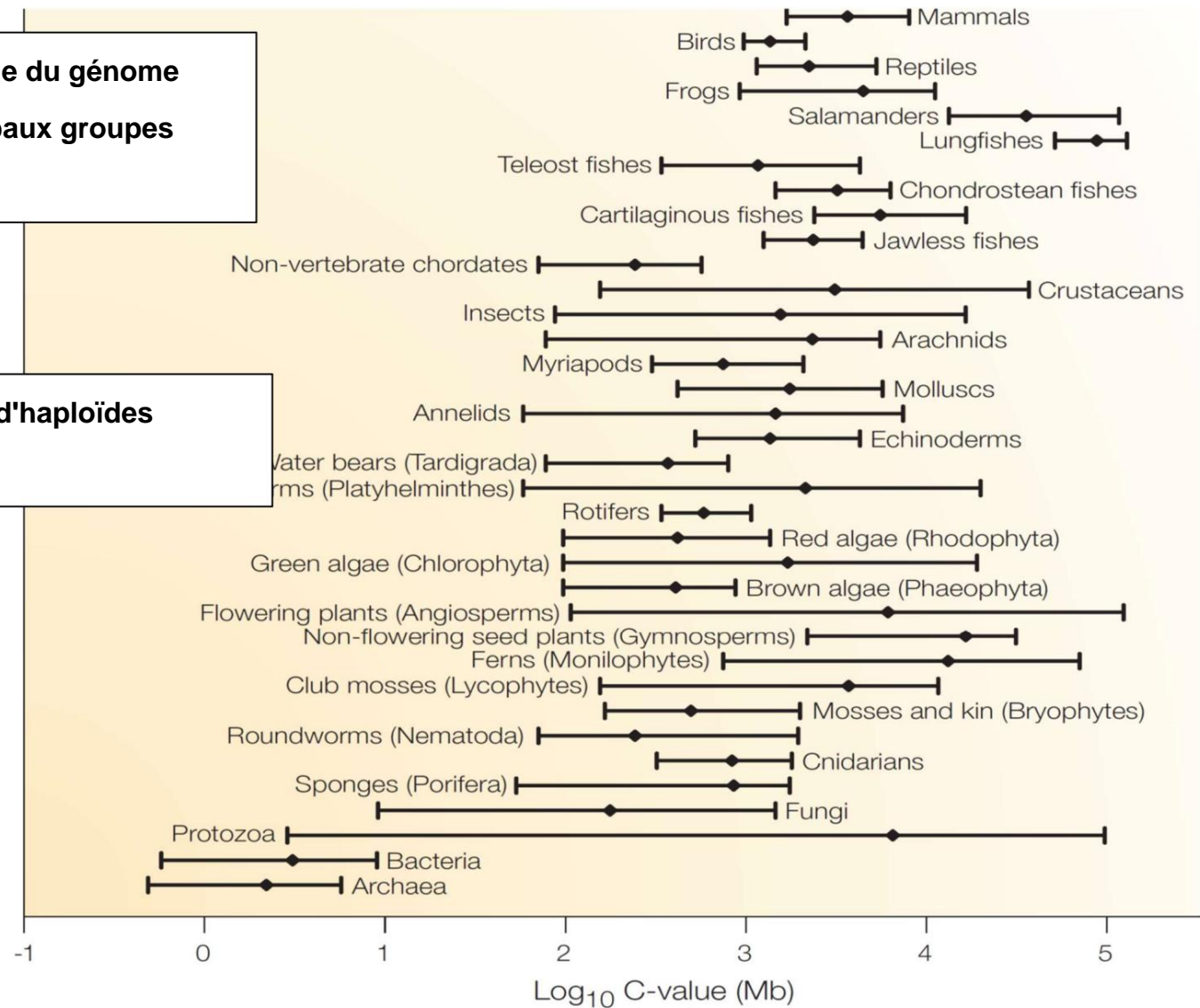
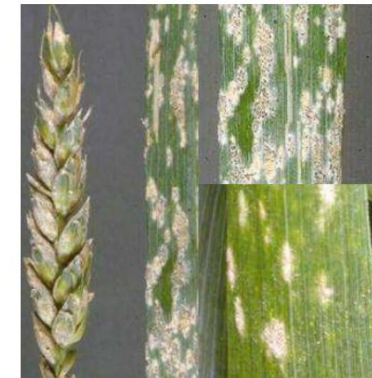
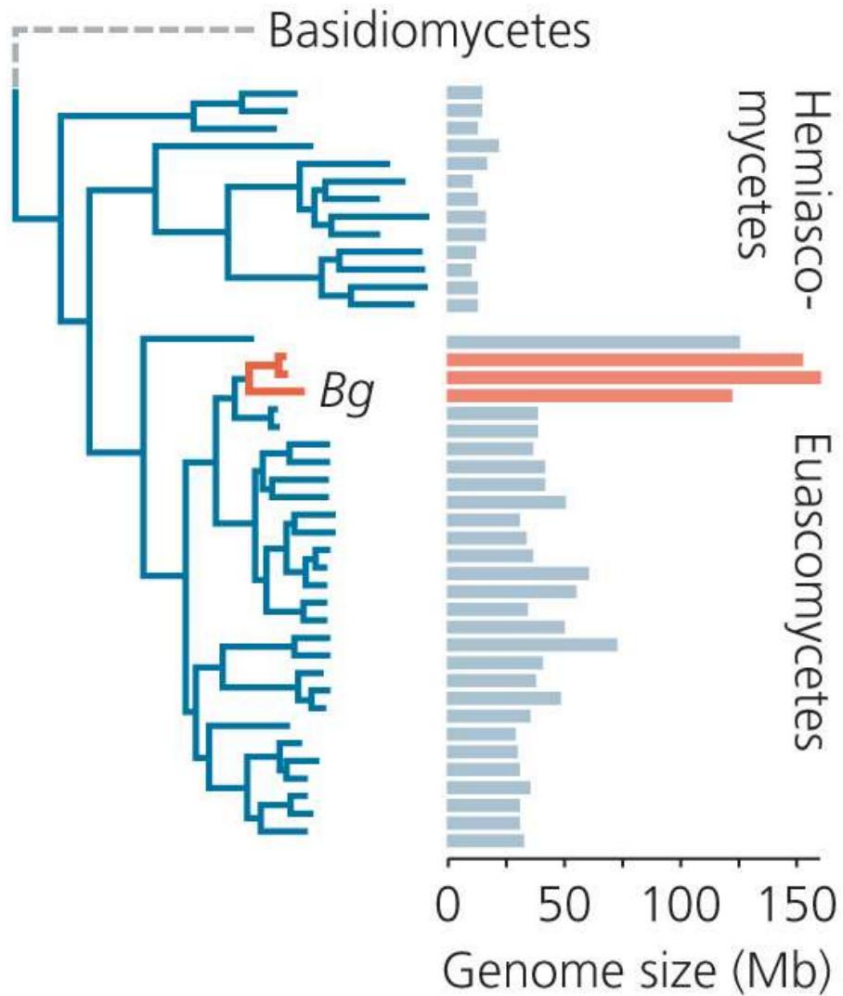


Figure de Grégory 2011

# I. Variation de la taille du génome

Variation même entre des espèces étroitement apparentées



**Blumeria graminis (Bg)**  
**Oïdium des céréales**

# I. Taille du génome et complexité de l'organisme

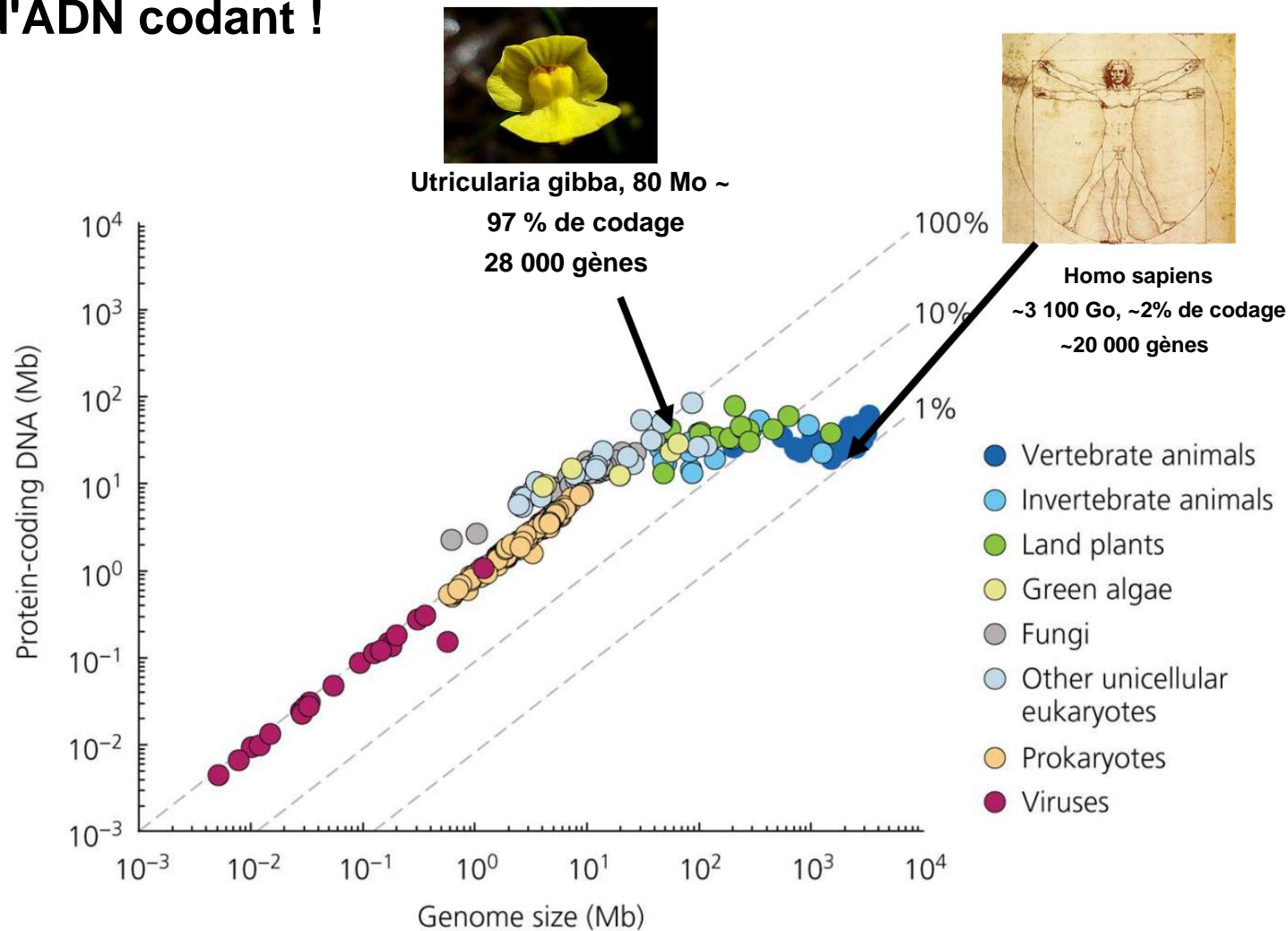
La taille du génome des eucaryotes multicellulaires n'est pas liée à :

- sa complexité morphologique perçue (c'est-à-dire le nombre de tissus) : **paradoxe de la valeur C**
- le nombre de gènes : **paradoxe de la valeur G**

Organisme		Taille du génome Mbp	Nombre approximatif de protéines codant gènes
Plante	Utricularia cornuta	82	28 000
Nématode	Caenorhabditis elegans	100	22 000
Insecte	Drosophila melanogaster	118	16 500
Plante	Arabidopsis thaliana	120	27 000
Poisson	Tetraodon nigroviridis	390	19 000
Mammifère	Homo sapiens	3 100	20 000
Plante	Zea mais	5 000	50 000
Poisson	Protopterus aethiopicus	130 000	
Plante	Paris japonica	150 000	

# I. Variation de la taille du génome et de la quantité d'ADN codant

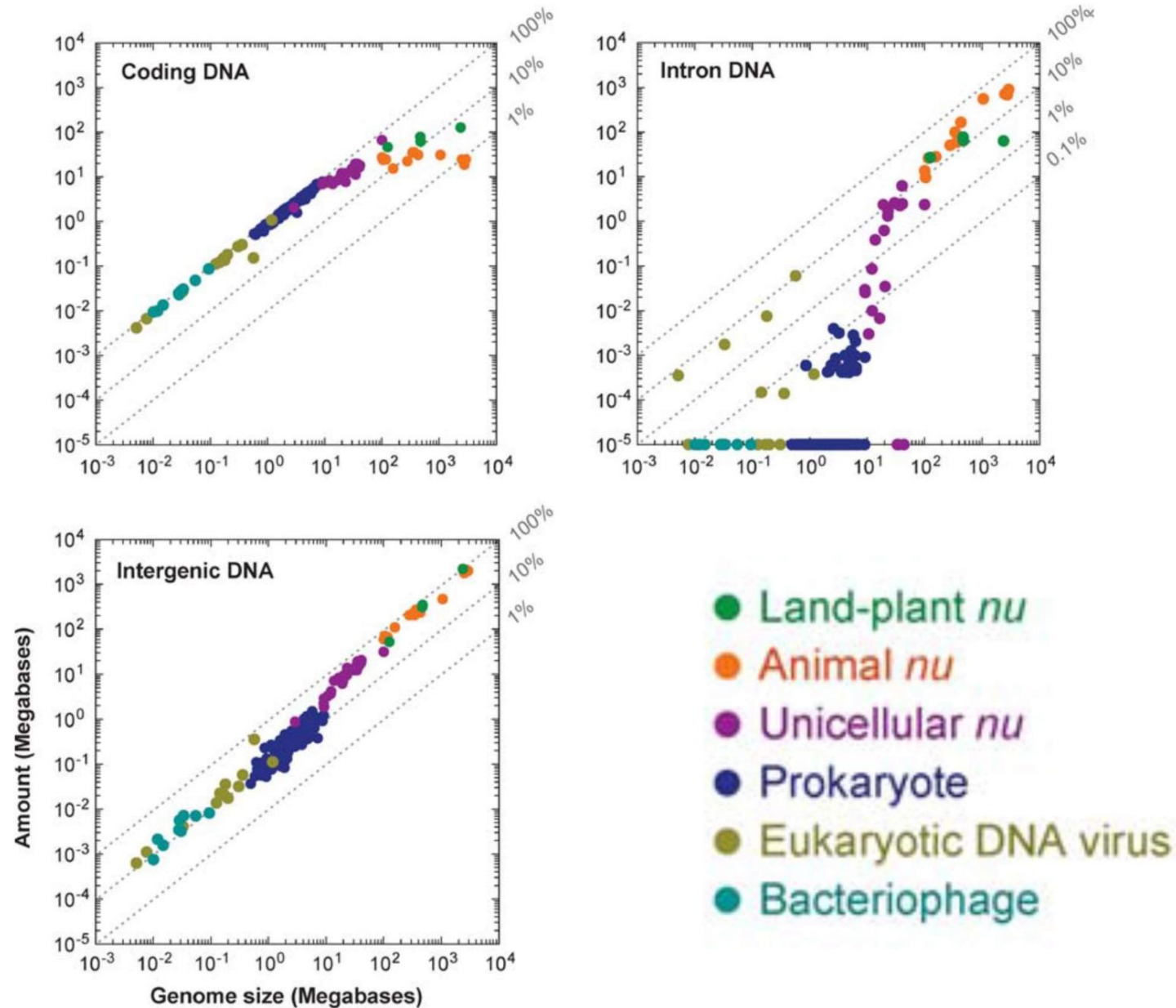
La taille du génome n'est pas forcément corrélée à la quantité d'ADN codant !



Adapté de Lynch et al. 2007, Herron & Freeman 5e édition

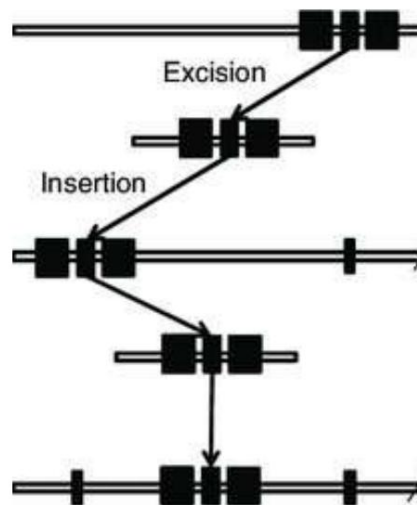


# I. Expansion de l'ADN non codant avec la taille du génome

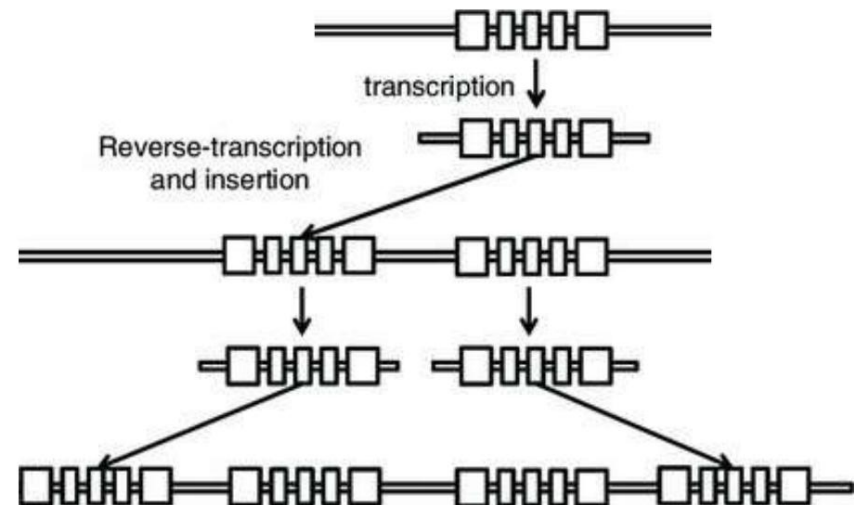


# I. Génomes étendus et éléments transposables (TE)

## Transpositions d'ADN coupées et collées

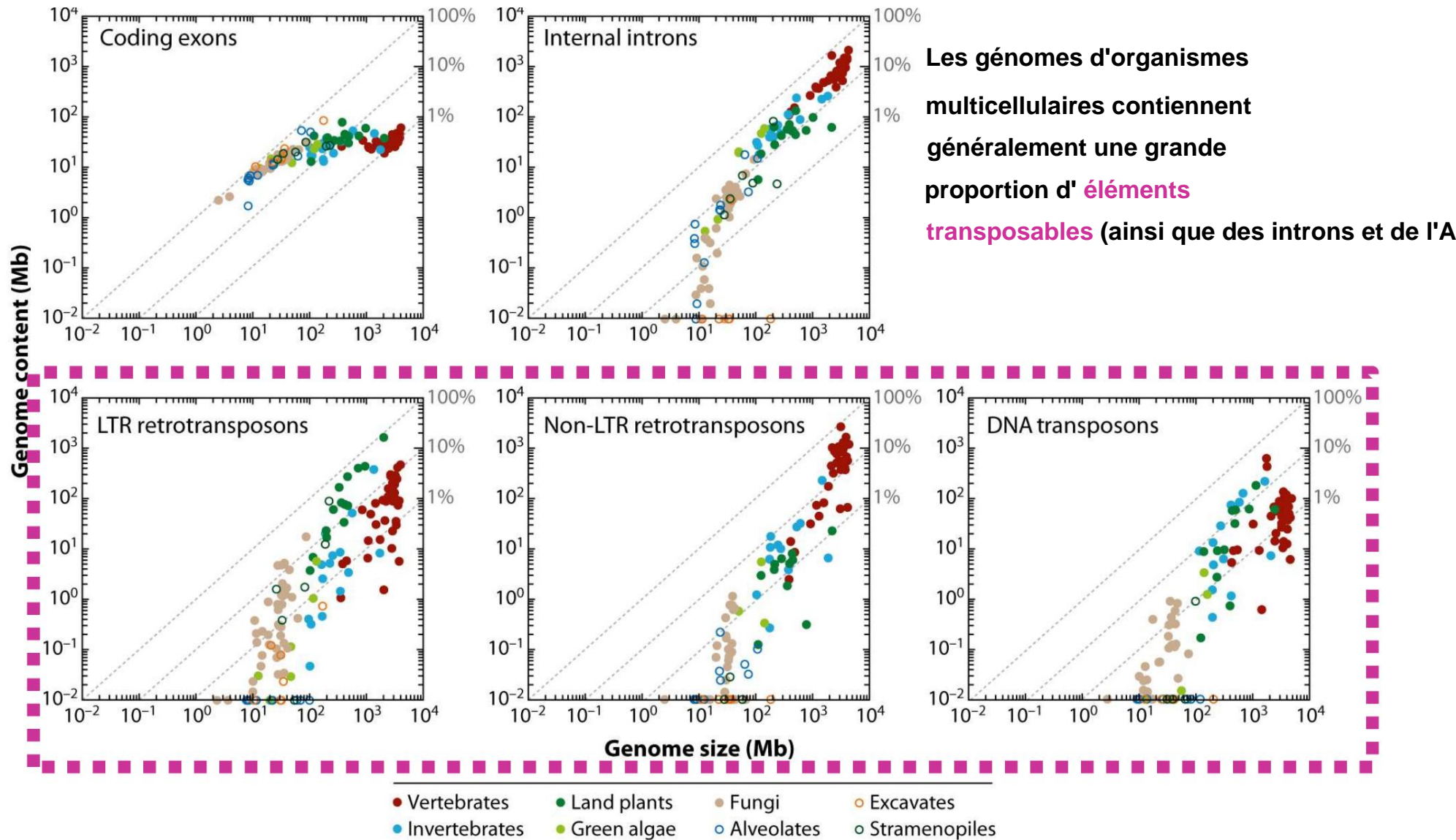


## Rétrotransposons Copier et coller



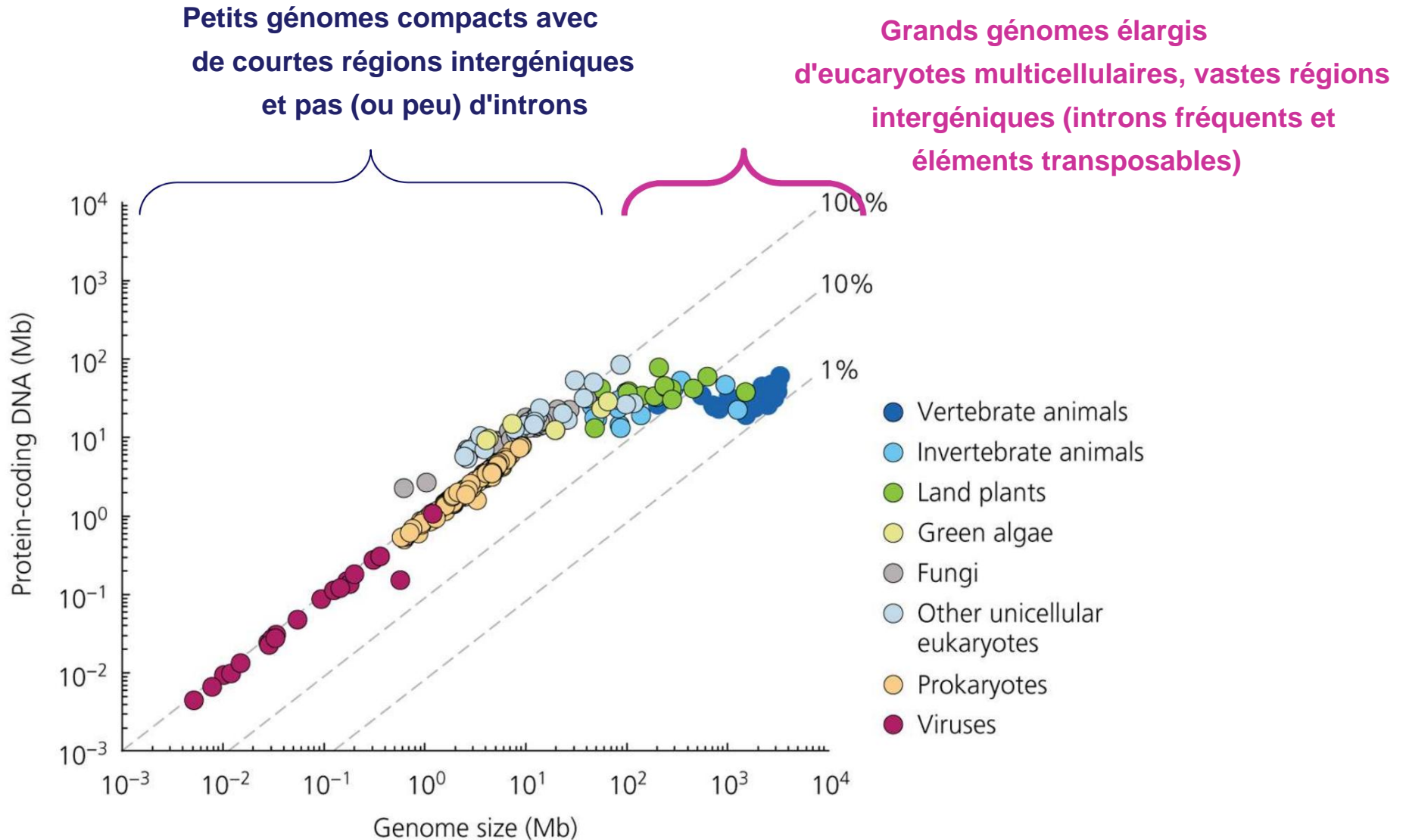


# I. Expansion de l'ADN non codant avec la taille du génome



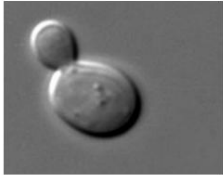
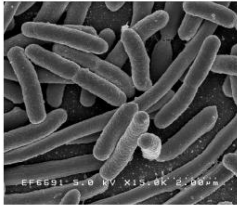
AR Lynch M, et al. 2011.  
Annu. Rev. Genomics Hum. Genet. 12:347–66

# I. Taille du génome et architecture générale



Adapté de Lynch et al. 2007, Herron & Freeman 5e édition

# I. Organisation typique des génomes



(a) Prokaryote



(b) Unicellular eukaryote, compact genome



(c) Multicellular eukaryote, expanded genome



Complete genes



Exons



Intergenic regions



Introns

© 2014 Pearson Education, Inc.

**Génomes compacts**

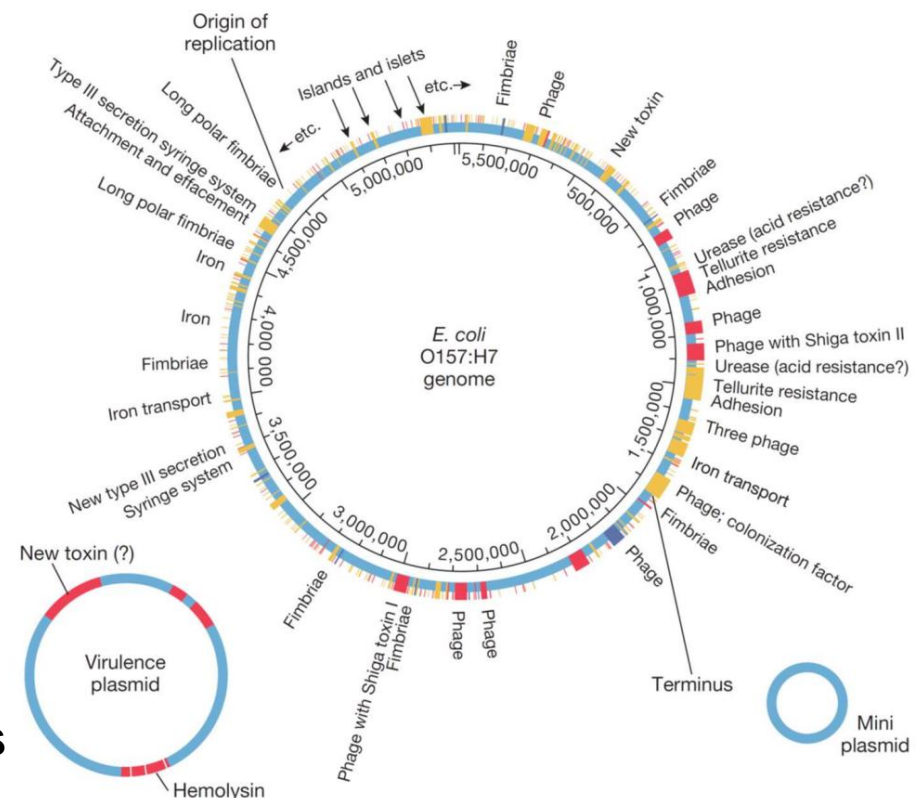
**Génomes étendus**

# I. Contenu et structure des génomes bactériens

- Chromosome circulaire unique  
+ molécules d'ADN circulaires  
plus petites (plasmides)
- 85 à 95% de codage des protéines  
séquence
- Structure génétique simple  
(introns très rares)
- Généralement une origine de  
réplication (bidirectionnelle)

-> densité de gènes élevée  
et régions intergéniques minimales

Schéma du génome d'une souche  
pathogène d'*E. coli*



# I. Contenu et structure des **génomés nucléaires\*** eucaryotes

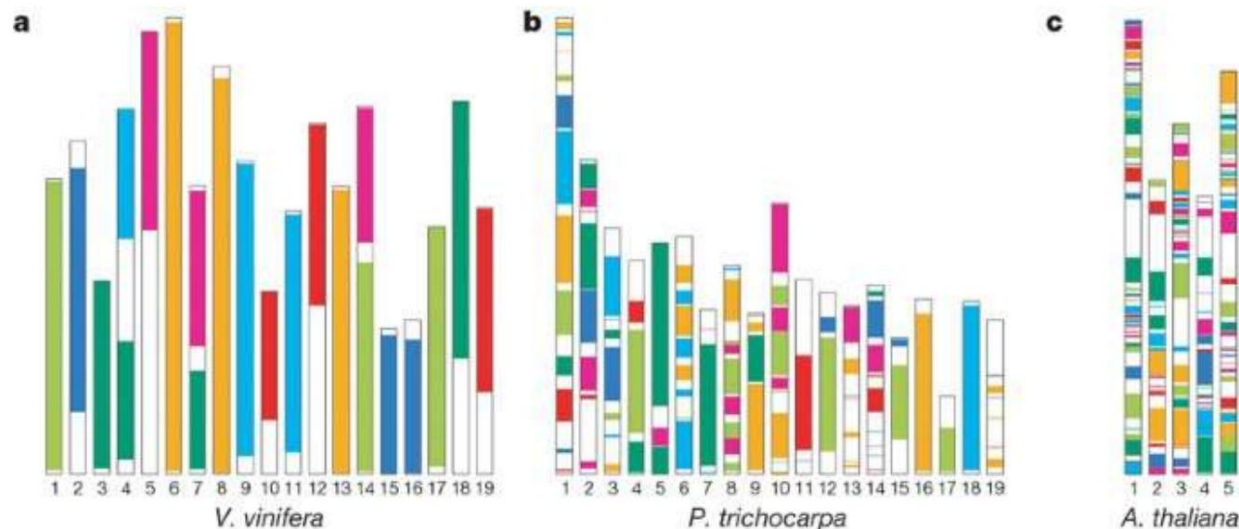
• Multiples chromosomes linéaires (avec centromères et télomères) • Multiples origines de réplication

• Points chauds de recombinaison

• Structure génique complexe avec introns communs •

Densité génique variable et proportion codant/non codant (de 3 à 98% de non ADN codant)

\*Les cellules eucaryotes peuvent également avoir des mitochondries et des chloroplastes avec leurs propres génomes! Non traité ici.



## **II. Évolution de la taille du génome**

**ÿ Quels processus pourraient produire une grande quantité d'ADN non codant?**

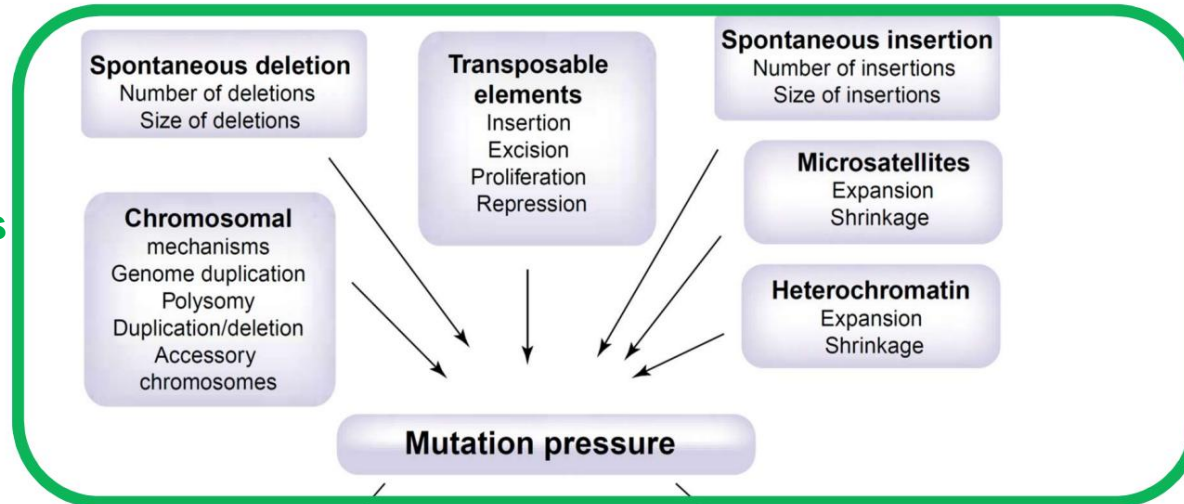
**ÿ Quelle est la fonction adaptative, le cas échéant, de l'ADN non génique ?**

**ÿ Si l'ADN non génique n'a pas de rôle adaptatif, pourquoi la sélection naturelle tolérerait-elle le fardeau de l'ADN supplémentaire ?**

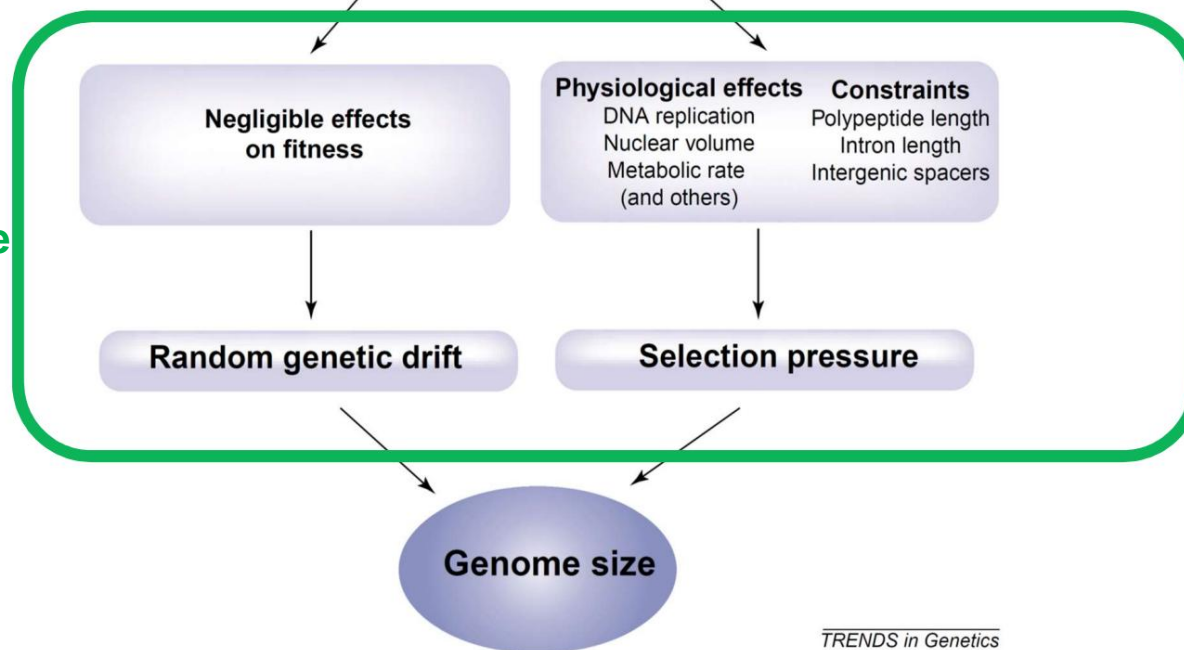


## II. Processus moléculaires et évolutifs affectant le génome évolution de la taille

### Mécanismes moléculaires



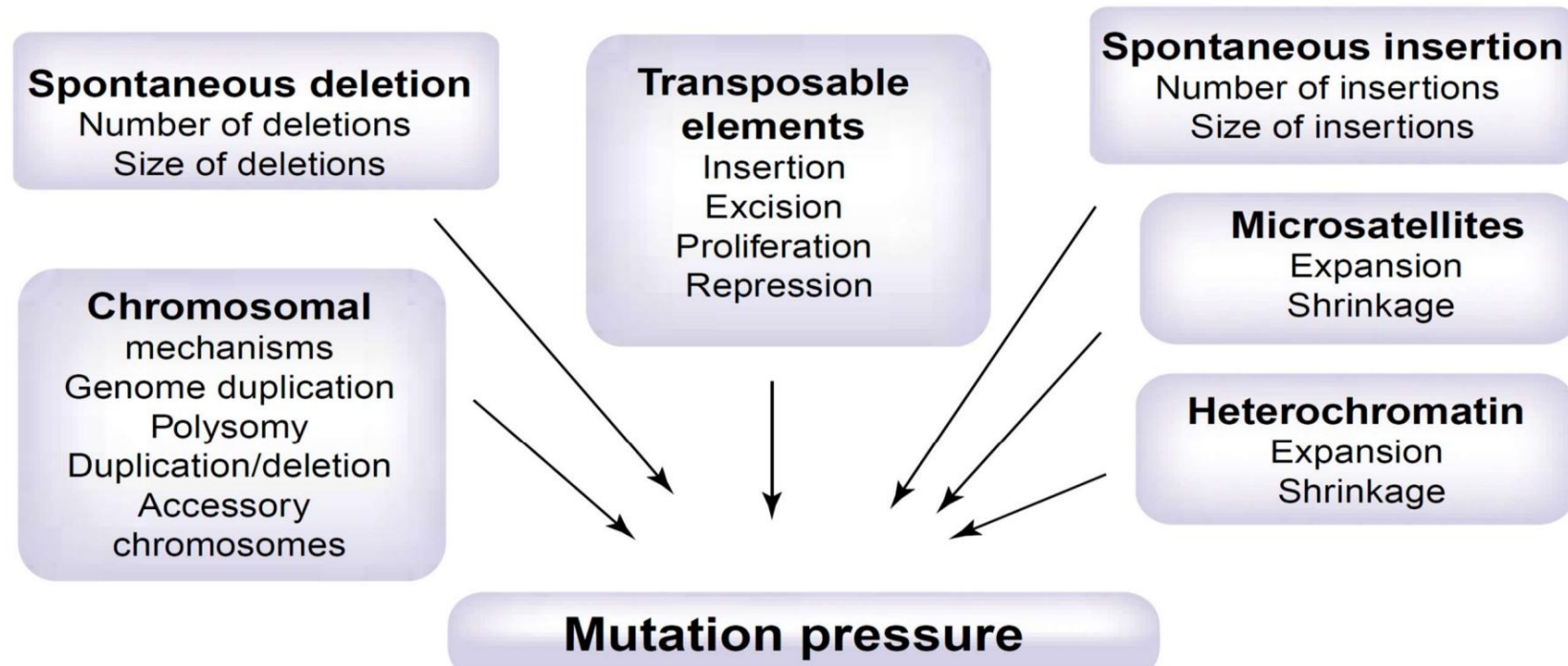
### Évolutionniste processus



*TRENDS in Genetics*

DA Petrov 2001, adapté par D. Mouchiroud

## II. Chiffre d'affaires génomique



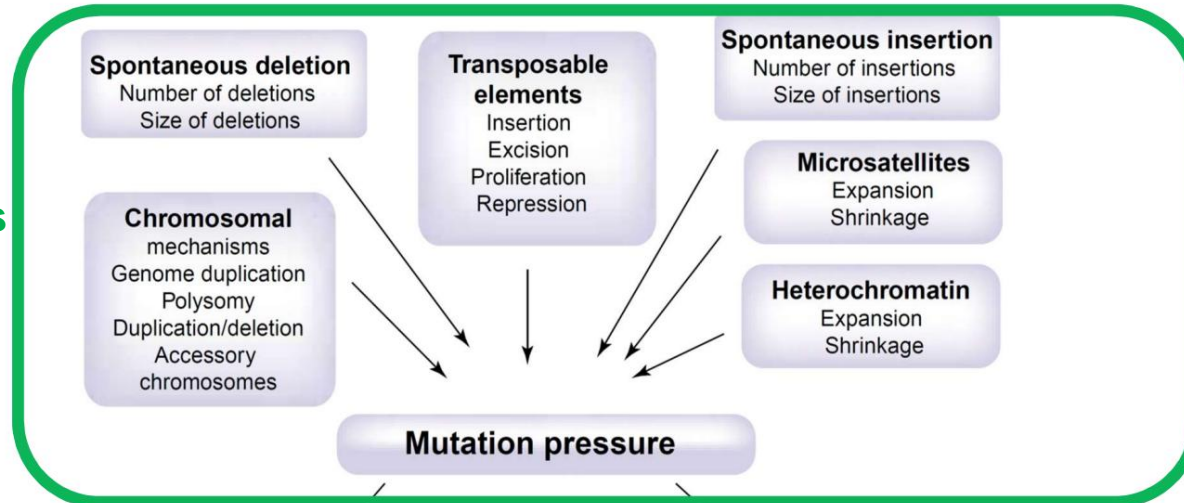
**Pressions mutationnelles majeures qui étendent ou contractent la taille du génome :**

- Éléments transposables : taux élevé d'auto-réplication et d'insertion
- Duplications du génome entier
- Taux et taille des insertions ou des délétions

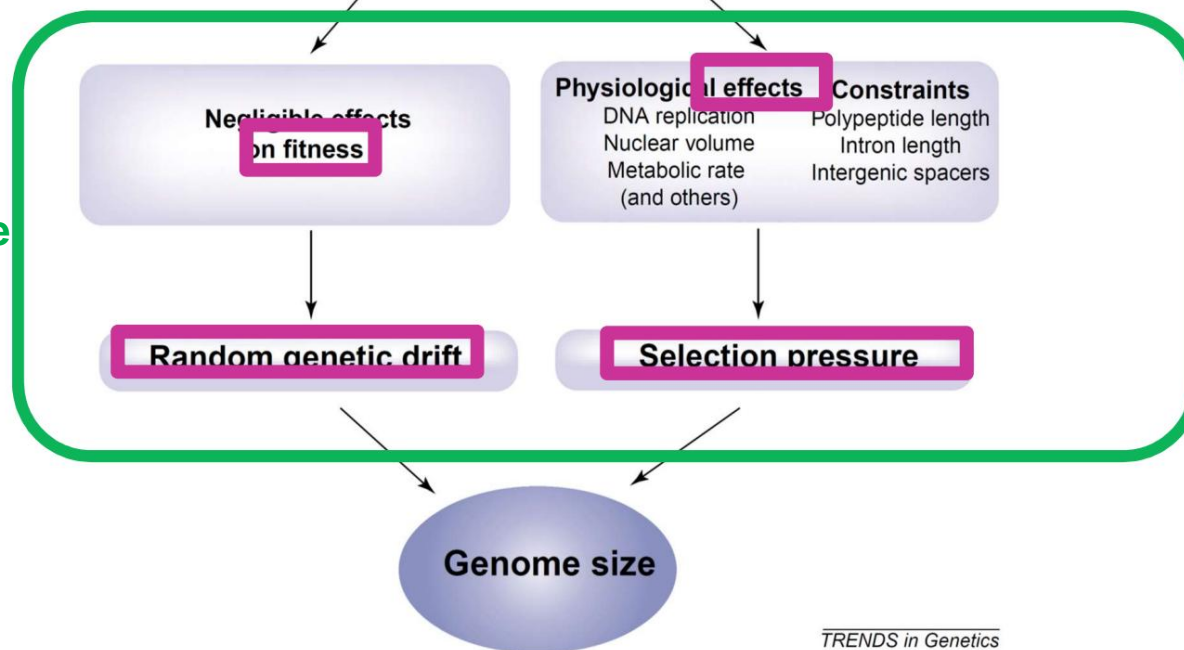


## II. Processus moléculaires et évolutifs affectant le génome évolution de la taille

Mécanismes  
moléculaires



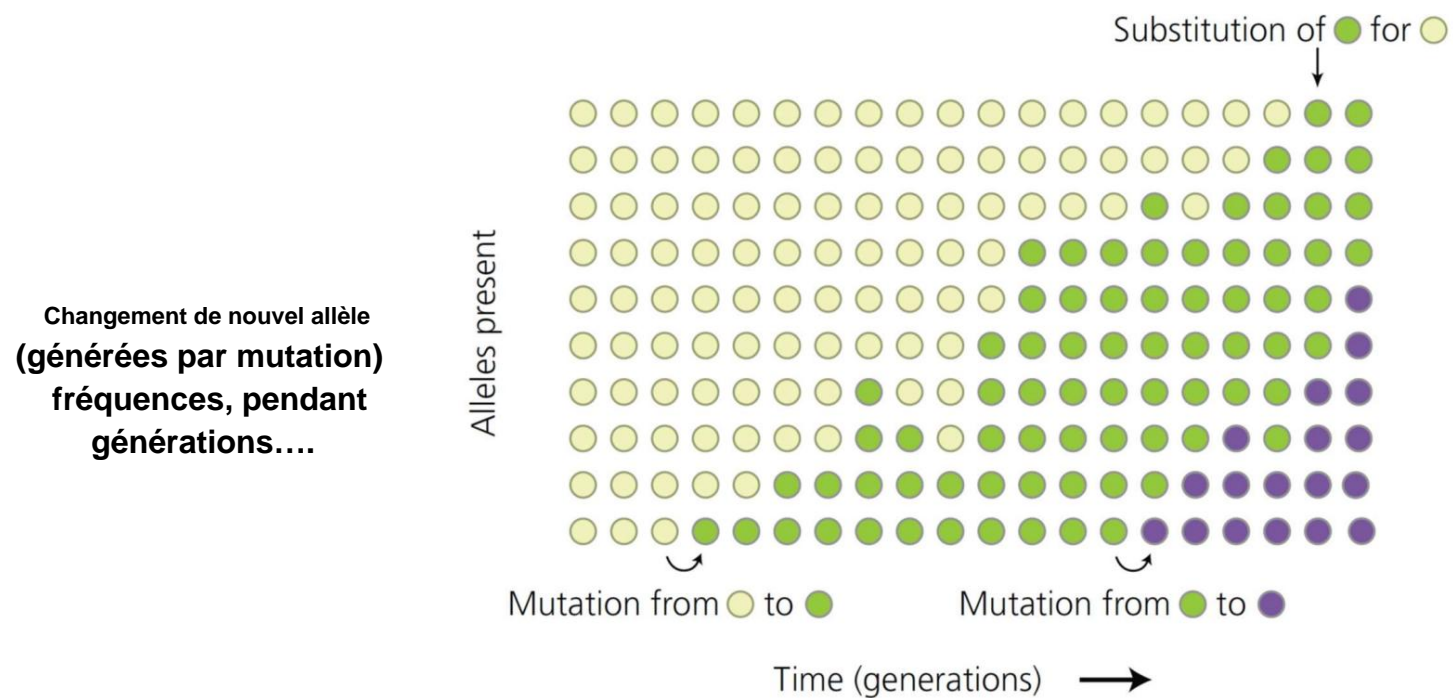
Évolutionniste  
processus



*TRENDS in Genetics*

DA Petrov 2001, adapté par D. Mouchiroud

## II. Processus évolutifs



## II. Processus évolutifs

**Les changements évolutifs impliquent une interaction complexe de mutation, de recombinaison, de sélection naturelle et d'échantillonnage aléatoire de gamètes de génération en génération**

**ÿ La capacité d'un organisme à survivre et à se reproduire, par rapport aux autres organismes de la même population, constitue son aptitude**

**ÿ La dérive génétique est le changement des fréquences alléliques dans une population d'une génération à l'autre qui se produit en raison d'un échantillonnage aléatoire de gamètes**

**ÿ Taille effective de la population ( $N_e$ ) : un concept visant à mesurer la quantité de dérive génétique se produisant dans une population à un moment donné**

## II. Processus évolutifs

I

**Mutation pressure**

Mutations germinales

**L'ADN supplémentaire profite-t-il à l'organisme?**

**Negligible effects  
on fitness**

Certains pourraient avoir  
négligeable  
effets de remise en forme

**Physiological effects**

DNA replication  
Nuclear volume  
Metabolic rate  
(and others)

**Constraints**

Polypeptide length  
Intron length  
Intergenic spacers

Certains d'entre eux  
les mutations peuvent  
affecter la condition physique

**Population  
processus**

**Random genetic drift**

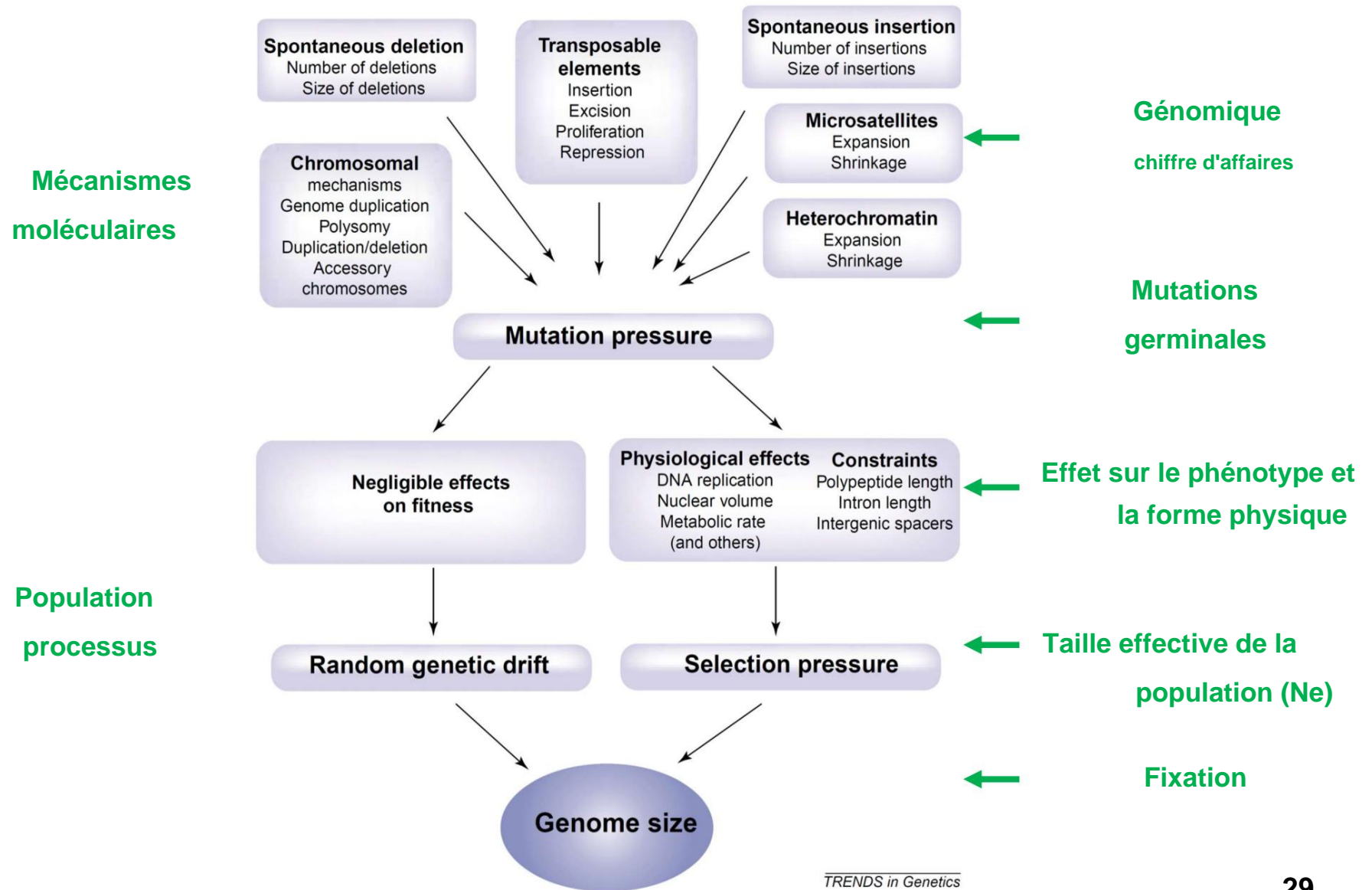
**Selection pressure**

Importance de la taille  
effective de la population ( $N_e$ )

**Genome size**

*TRENDS in Genetics*

## II. Sommaire



### **III. Modèles d'évolution de la taille du génome**

• Hypothèse d'ADN en vrac

• Hypothèse nucléotypique

• Hypothèse ADN égoïste

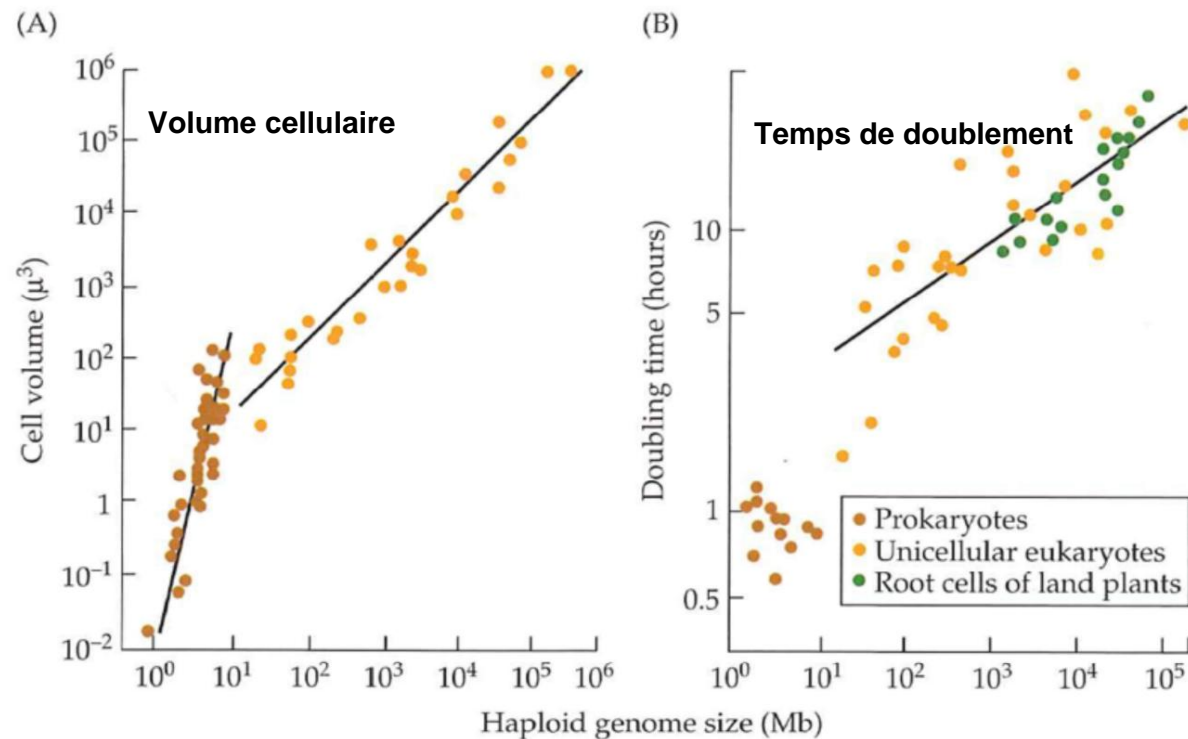
• Hypothèse d'équilibre mutationnel

• Hypothèse du hasard mutationnel

### III. ADN en vrac et hypothèses nucléotypiques

• Hypothèse de l'ADN en vrac: sélection pour la vitesse de réplication, la petite taille des cellules et l'efficacité énergétique -> peut favoriser **les réductions de la taille du génome**

• Hypothèse nucléotypique: une plus grande taille de génome pourrait être bénéfique pour un organisme (des volumes nucléaires et cellulaires plus importants, pourraient aider à amortir les fluctuations de la concentration de protéines régulatrices) -> peut favoriser **l'augmentation de la taille du génome**

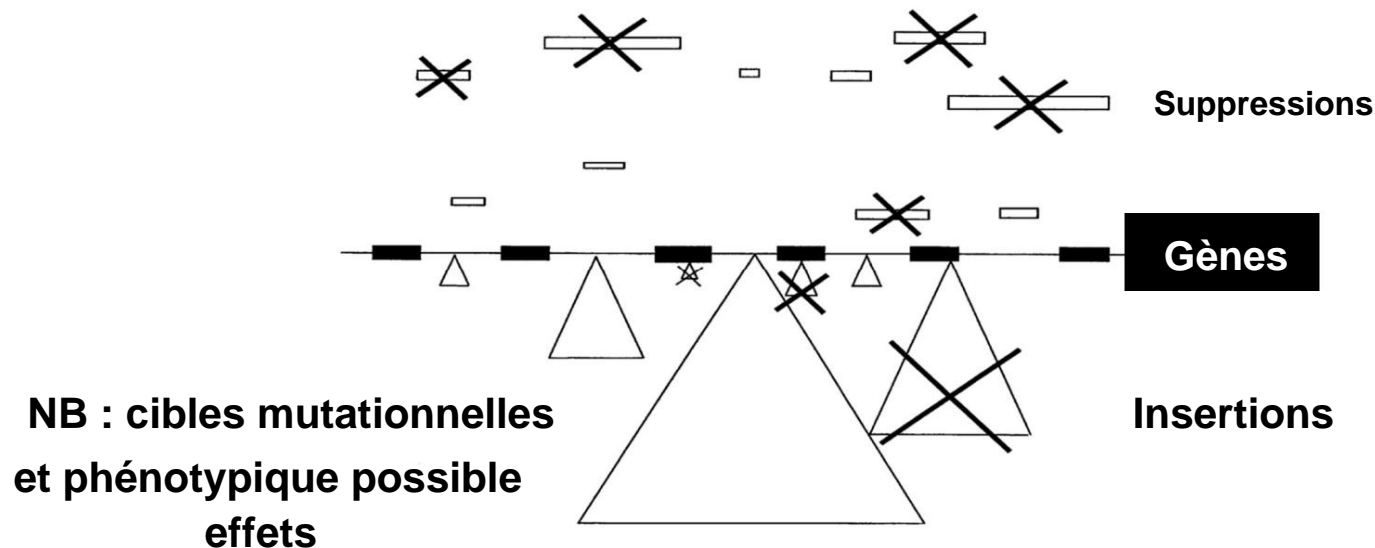


Vinogradov, Gregory, personnage de Lynch 2007

### III. ADN égoïste et hypothèses d'équilibre mutationnel

• ADN égoïste: prolifération d'éléments autoréplicatifs tels que les éléments transposables "égoïstes" -> **augmentation de la taille du génome**

• Modèle d'équilibre mutationnel : **équilibre** principalement basé sur biais mutationnels avec un équilibre entre la perte d'ADN (avec de fréquentes petites délétions) et le gain d'ADN obtenu par de grandes insertions (changements moins fréquents mais plus importants)





### III. Hypothèse du hasard mutationnel

Proposé par Lynch & Conery en 2003

**"Une grande partie de la complexité génomique eucaryote a initialement évolué comme une réponse indirecte passive à la taille réduite de la population"**



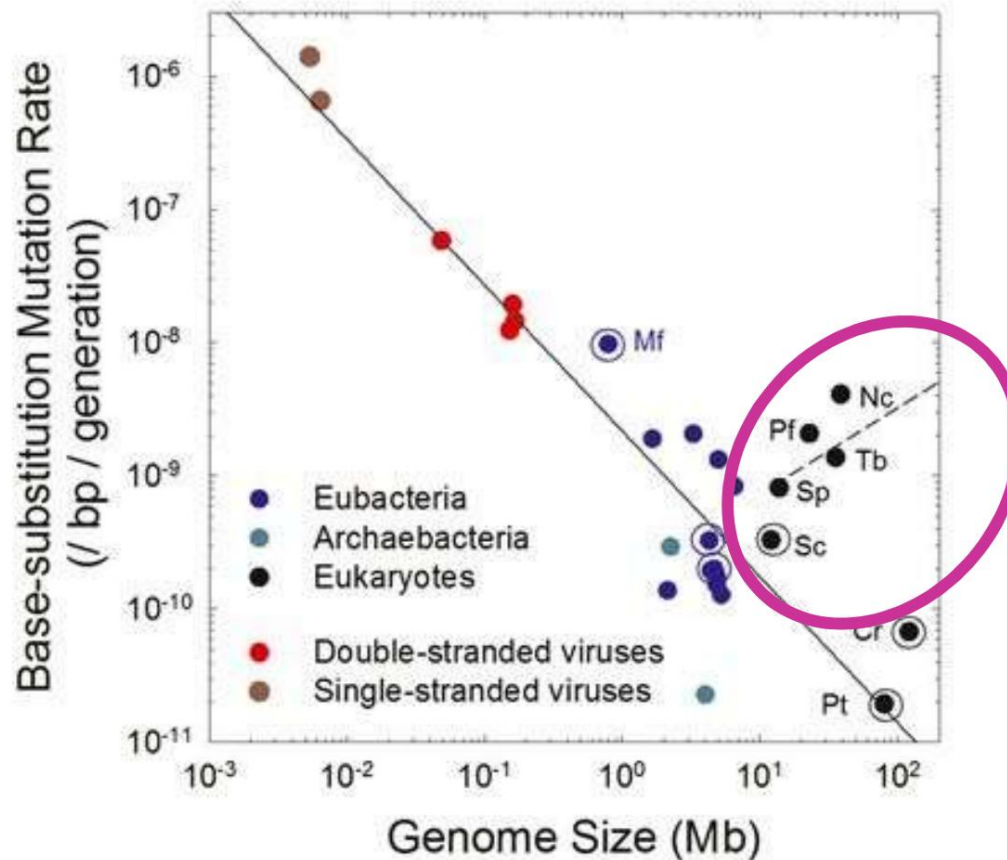
• Importance de la dérive génétique et de la taille effective de la population  $N_e$  dans la taille du génome évolution

• Rôle du taux de mutation : un trait moléculaire qui peut varier et évoluer

• Supposons que la sélection favorise de faibles taux de mutation pour limiter l'accumulation de mutations délétères

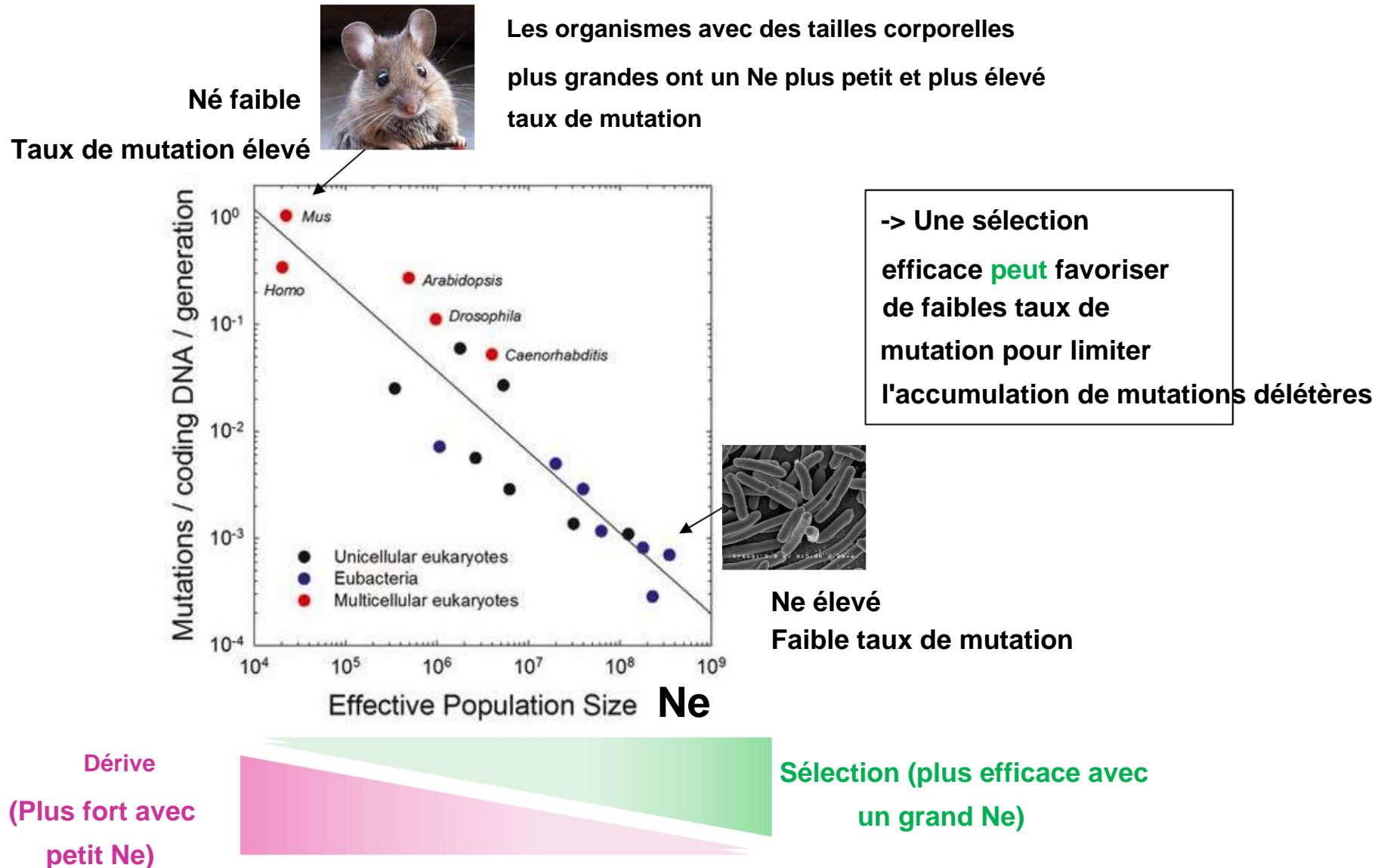
### III. Hypothèse du hasard mutationnel

• Chez les micro-organismes: les faibles taux de mutation ont tendance à être associés à des taille du génome • Tendance différente chez les eucaryotes multicellulaires



Lynch et Connery 2003, Sung et al 2012

### III. Hypothèse du hasard mutationnel



### III. Hypothèse du hasard mutationnel: résumé

Hypothèse : l'expansion globale de **la taille du génome**,  
des procaryotes aux eucaryotes multicellulaires, est liée à  
  
**une réduction importante de la taille effective de la population**,  
  
qui ont conduit à une **dérive génétique plus forte** (et donc à  
une **sélection moins efficace** contre les mutations délétères)  
  
a entraîné la **prolifération d'ADN supplémentaire**

# Conclusion

• Quels processus pourraient produire une grande quantité d'ADN non codant?

Turnover génomique, éléments transposables, ....

• Quelle est la fonction adaptative, le cas échéant, de l'ADN supplémentaire non génique?

Effet sur la forme physique de l'ADN supplémentaire (pas clair)? Hypothèse nucléotypique

• Si l'ADN non génique n'a pas de rôle adaptatif, pourquoi l'ADN naturel  
la sélection tolère-t-elle le fardeau de l'ADN supplémentaire?

Effet fitness d'une nouvelle mutation

Biais mutationnels

Interaction entre dérive et sélection (efficacité de la sélection naturelle et  $N_e$ )

-> Toutes les caractéristiques observées dans un génome et leur évolution ne sont pas dues à une  
évolution adaptative !!!

# Conclusion

ÿ Les forces évolutives qui influencent la taille des génomes dépendent de la biologie, de la physiologie et de l'écologie des organismes.

ÿ Poussé par l'interaction entre les forces moléculaires et évolutives

ÿ Dynamique du génome et échelles de temps de l'évolution adaptative (taux d'insertions/délétions, rafale rapide de transposition)

ÿ Importance des processus de population et  $N_e$

ÿ Taille et complexité du génome, nombre de gènes redondants et topologie des réseaux de régulation...

ÿ Coût de l'ADN supplémentaire ?

ÿ Effets environnementaux sur la dynamique du génome (ex. stress et transposition)

