

Détecter la coévolution au niveau des sites: le cas des protéines

Julien Dutheil

¹ CNRS UMR 5171 "Génome, Populations, Interactions, Adaptation",
Université Montpellier II, France

Julien.Dutheil@univ-montp2.fr

24 Janvier 2006

Coévolution moléculaire au niveau des sites

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

Mutations compensatoires :

- ARN : paires Watson-Crick au sein des tiges : une mutation $G \rightarrow A$ peut être compensée par une mutation $C \rightarrow U$ sur le brin opposé
- Protéines : Une mutation "petit \rightarrow gros" peut être compensée par une mutation "gros \rightarrow petit" dans son voisinage

Coévolution moléculaire au niveau des sites

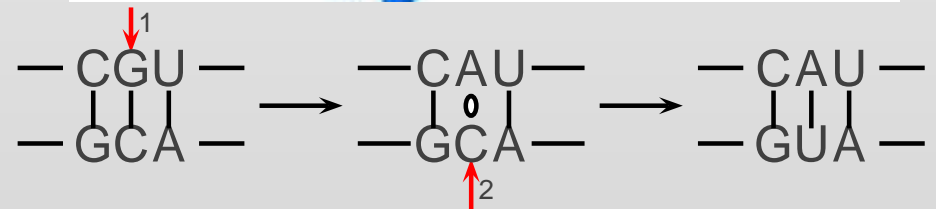
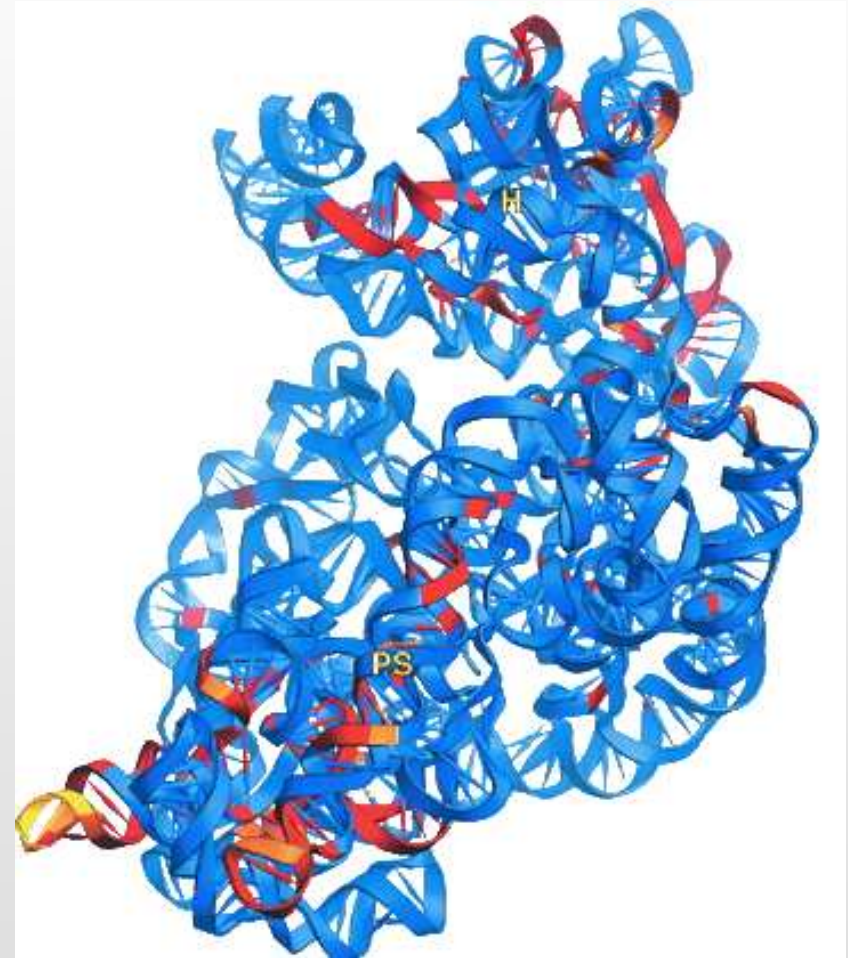
● Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

Mutations compensatoires :

- ARN : paires Watson-Crick au sein des tiges : une mutation G→A peut être compensée par une mutation C→U sur le brin opposé
- Protéines : Une mutation "petit→gros" peut être compensée par une mutation "gros→petit" dans son voisinage



Coévolution moléculaire au niveau des sites

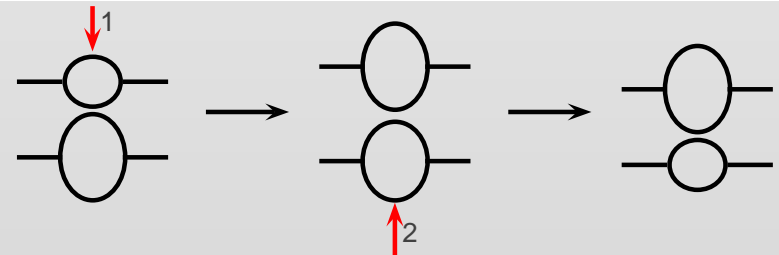
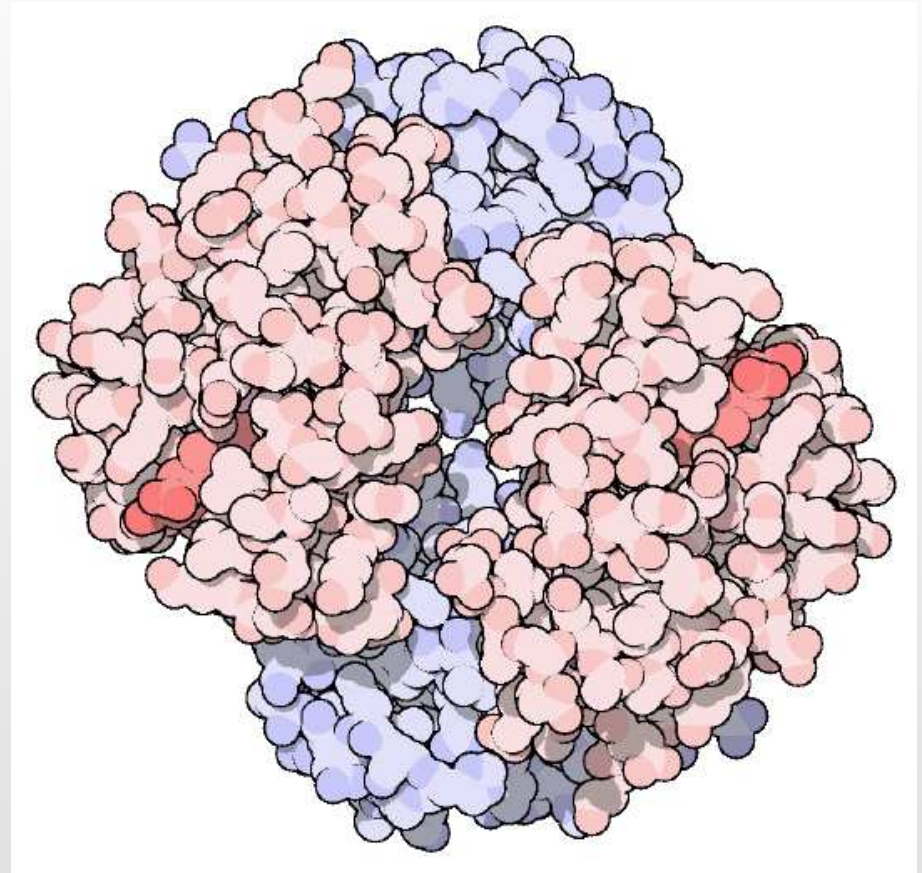
● Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

Mutations compensatoires :

- ARN : paires Watson-Crick au sein des tiges : une mutation $G \rightarrow A$ peut être compensée par une mutation $C \rightarrow U$ sur le brin opposé
- Protéines : Une mutation "petit \rightarrow gros" peut être compensée par une mutation "gros \rightarrow petit" dans son voisinage



- Coévolution moléculaire au niveau des sites

Cartographie de substitution

- Vecteurs de substitutions
- Mapping probabiliste

Application de la méthode aux protéines

Cartographie de substitution

Vecteurs de substitutions

- Coévolution moléculaire au niveau des sites

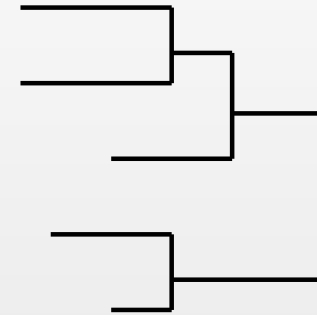
Cartographie de substitution

- Vecteurs de substitutions

- Mapping probabiliste

Application de la méthode aux protéines

| | |
|---------|---------------------------|
| Taxon 1 | A A D A D A I G I E D Q A |
| Taxon 2 | A V D A E - P G I A D Q A |
| Taxon 3 | A V E A A - I Q T V A - A |
| Taxon 4 | A L D A R - L R S D - - V |
| Taxon 5 | A V E S V A T G T G D - V |



Vecteurs de substitutions

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

- Vecteurs de substitutions

- Mapping probabiliste

Application de la méthode aux protéines

| | D_i | |
|---------|---------------------------|--|
| Taxon 1 | A A D A D A I G I E D Q A | |
| Taxon 2 | A V D A E - P G I A D Q A | |
| Taxon 3 | A V E A A - I Q T V A - A | |
| Taxon 4 | A L D A R - L R S D - - V | |
| Taxon 5 | A V E S V A T G T G D - V | |

$$V_i = \begin{pmatrix} v_{i,0} \\ \vdots \\ v_{i,b} \\ \vdots \\ v_{i,m} \end{pmatrix}, \hat{r}_i$$

Vecteurs de substitutions

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

- Vecteurs de substitutions

- Mapping probabiliste

Application de la méthode aux protéines



$$V_i = \begin{pmatrix} v_{i,0} \\ \vdots \\ v_{i,b} \\ \vdots \\ v_{i,m} \end{pmatrix}, \hat{r}_i \quad V_j = \begin{pmatrix} v_{j,0} \\ \vdots \\ v_{j,b} \\ \vdots \\ v_{j,m} \end{pmatrix}, \hat{r}_j$$

$$\rho_{i,j} = \frac{cov(V_i, V_j)}{sd(V_i) \times sd(V_j)}$$

Mapping probabiliste

- On veut estimer le nombre attendu $v_{i,b}$ de changements qui sont apparus au site i sur la branche b
- Ce nombre dépend des états aux noeuds supérieur et inférieur :

$$v_{i,b} = \sum_x \sum_y Pr(x, y | D, \Theta) \times n_{x,y}(t)$$

- $Pr(x, y | D, \Theta)$ est la probabilité jointe pour la paire
- Substitutions multiples : nombre de changements attendu le long d'une branche de longueur t sachant l'état initial x et l'état final $y \Rightarrow$ transformé de Laplace

$$n_{x,y}(t) = \frac{m_{x,y}(t)}{p_{x,y}(t)}, \quad M(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n-1} Q^p (Q + \Lambda) Q^{n-p-1}$$

avec $M(t) = \{m_{x,y}(t)\}$, $Q = \{q_{x,y}\}$, $P(t) = \exp(Qt)$, $\Lambda = \text{diag}(\{q_{x,x}\})$.

- On applique la méthode successivement pour chaque branche \Rightarrow mapping probabiliste

Mapping probabiliste

- On veut estimer le nombre attendu $v_{i,b}$ de changements qui sont apparus au site i sur la branche b
- Ce nombre dépend des états aux noeuds supérieur et inférieur :

$$v_{i,b} = \sum_x \sum_y Pr(x, y | D, \Theta) \times n_{x,y}(t)$$

- $Pr(x, y | D, \Theta)$ est la probabilité jointe pour la paire
- Substitutions multiples : nombre de changements attendu le long d'une branche de longueur t sachant l'état initial x et l'état final $y \Rightarrow$ transformé de Laplace

$$n_{x,y}(t) = \frac{m_{x,y}(t)}{p_{x,y}(t)}, \quad M(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n-1} Q^p (Q + \Lambda) Q^{n-p-1}$$

avec $M(t) = \{m_{x,y}(t)\}$, $Q = \{q_{x,y}\}$, $P(t) = \exp(Qt)$, $\Lambda = \text{diag}(\{q_{x,x}\})$.

- On applique la méthode successivement pour chaque branche \Rightarrow mapping probabiliste

Mapping probabiliste

- On veut estimer le nombre attendu $v_{i,b}$ de changements qui sont apparus au site i sur la branche b
- Ce nombre dépend des états aux noeuds supérieur et inférieur :

$$v_{i,b} = \sum_x \sum_y Pr(x, y | D, \Theta) \times n_{x,y}(t)$$

- $Pr(x, y | D, \Theta)$ est la probabilité jointe pour la paire
- Substitutions multiples : nombre de changements attendu le long d'une branche de longueur t sachant l'état initial x et l'état final $y \Rightarrow$ transformé de Laplace

$$n_{x,y}(t) = \frac{m_{x,y}(t)}{p_{x,y}(t)}, \quad M(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n-1} Q^p (Q + \Lambda) Q^{n-p-1}$$

avec $M(t) = \{m_{x,y}(t)\}$, $Q = \{q_{x,y}\}$, $P(t) = \exp(Qt)$, $\Lambda = \text{diag}(\{q_{x,x}\})$.

- On applique la méthode successivement pour chaque branche \Rightarrow mapping probabiliste

Mapping probabiliste

- On veut estimer le nombre attendu $v_{i,b}$ de changements qui sont apparus au site i sur la branche b
- Ce nombre dépend des états aux noeuds supérieur et inférieur :

$$v_{i,b} = \sum_x \sum_y Pr(x, y | D, \Theta) \times n_{x,y}(t)$$

- $Pr(x, y | D, \Theta)$ est la probabilité jointe pour la paire
- Substitutions multiples : nombre de changements attendu le long d'une branche de longueur t sachant l'état initial x et l'état final $y \Rightarrow$ transformé de Laplace

$$n_{x,y}(t) = \frac{m_{x,y}(t)}{p_{x,y}(t)}, \quad M(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n-1} Q^p (Q + \Lambda) Q^{n-p-1}$$

avec $M(t) = \{m_{x,y}(t)\}$, $Q = \{q_{x,y}\}$, $P(t) = \exp(Qt)$, $\Lambda = \text{diag}(\{q_{x,x}\})$.

- On applique la méthode successivement pour chaque branche \Rightarrow mapping probabiliste

Mapping probabiliste

- On veut estimer le nombre attendu $v_{i,b}$ de changements qui sont apparus au site i sur la branche b
- Ce nombre dépend des états aux noeuds supérieur et inférieur :

$$v_{i,b} = \sum_x \sum_y Pr(x, y | D, \Theta) \times n_{x,y}(t)$$

- $Pr(x, y | D, \Theta)$ est la probabilité jointe pour la paire
- Substitutions multiples : nombre de changements attendu le long d'une branche de longueur t sachant l'état initial x et l'état final $y \Rightarrow$ transformé de Laplace

$$n_{x,y}(t) = \frac{m_{x,y}(t)}{p_{x,y}(t)}, \quad M(t) = \sum_{n=1}^{\infty} \frac{t^n}{n!} \sum_{p=0}^{n-1} Q^p (Q + \Lambda) Q^{n-p-1}$$

avec $M(t) = \{m_{x,y}(t)\}$, $Q = \{q_{x,y}\}$, $P(t) = \exp(Qt)$, $\Lambda = \text{diag}(\{q_{x,x}\})$.

- On applique la méthode successivement pour chaque branche \Rightarrow mapping probabiliste

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqués dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

Application de la méthode aux protéines

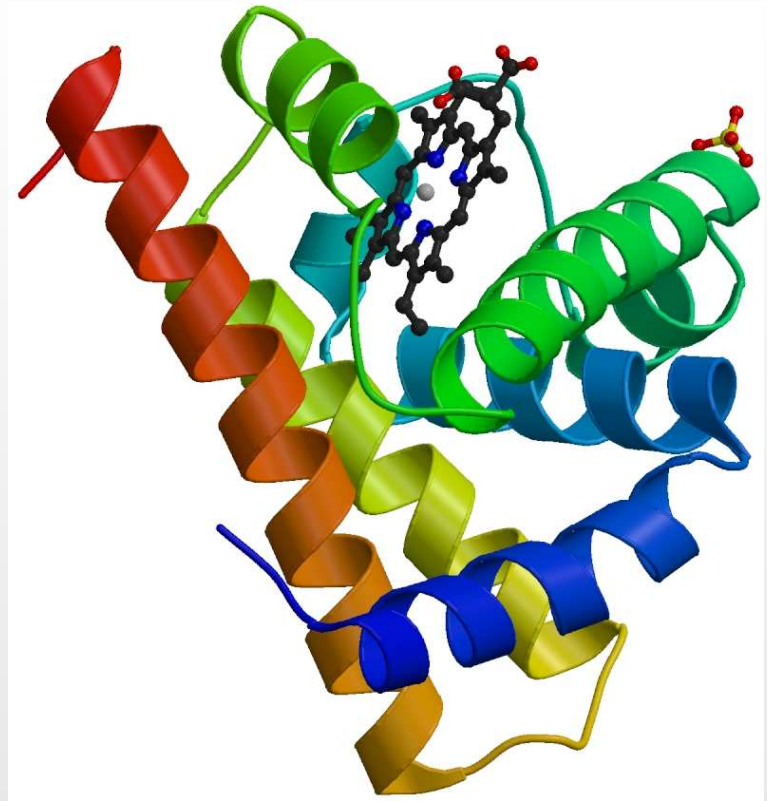
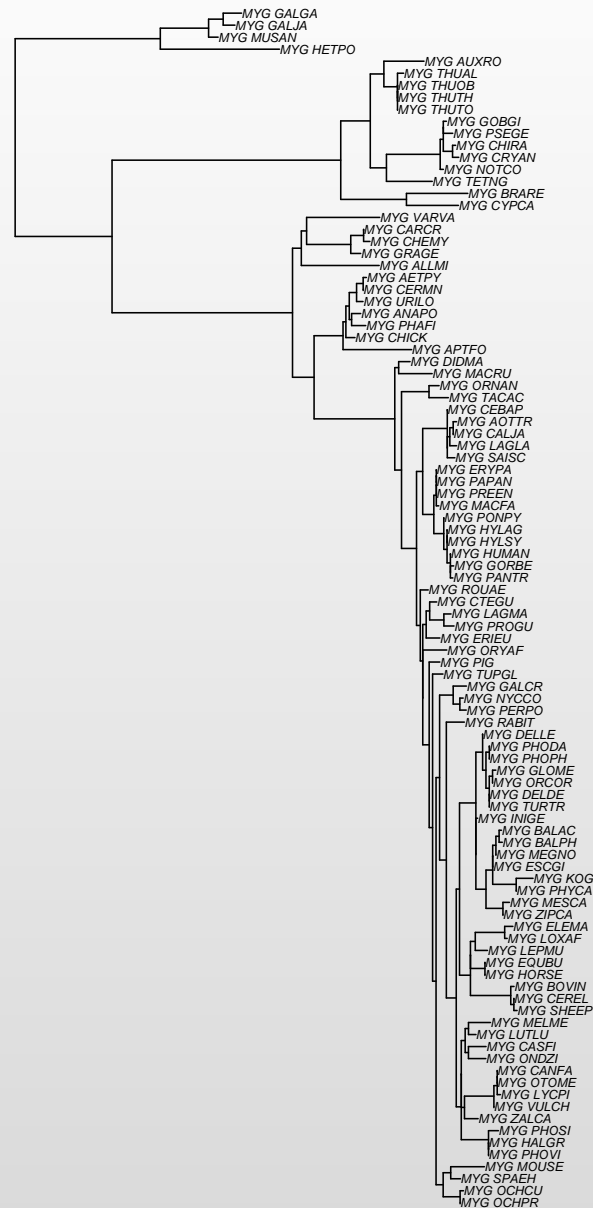
Myoglobine

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- **Myoglobine**
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqués dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives



- 100 espèces (vertébrés), 128 sites complets
- Nature physico-chimique des substitutions ?
- Coévolution impliquant plus de 2 sites ?

Prise en compte de la nature biochimique des changements

- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,

Prise en compte de la nature biochimique des changements

- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,
- On peut calculer des vecteurs de substitutions pondérés avec la formule :

$$w_{i,b} = \sum_x \sum_y Pr(x, y | D_i, \Theta) \times d_{x,y}$$

[en ignorant la correction pour substitutions multiples]

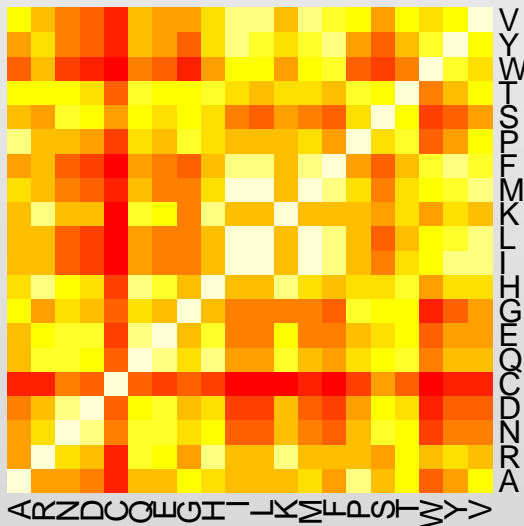
Prise en compte de la nature biochimique des changements

- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,
- On peut calculer des vecteurs de substitutions pondérés avec la formule :

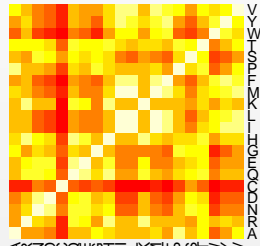
$$w_{i,b} = \sum_x \sum_y Pr(x, y | D_i, \Theta) \times d_{x,y}$$

[en ignorant la correction pour substitutions multiples]

- Exemples de matrices possibles :
 - Distance synthétique (Grantham, Miyata, ...)



Prise en compte de la nature biochimique des changements



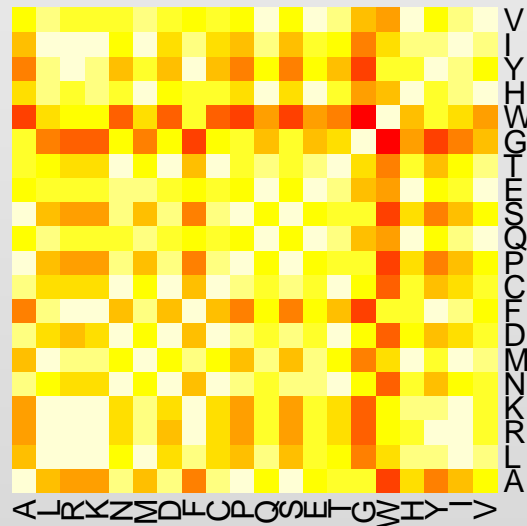
Grantham

- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,
- On peut calculer des vecteurs de substitutions pondérés avec la formule :

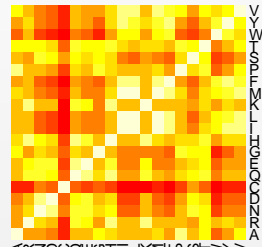
$$w_{i,b} = \sum_x \sum_y Pr(x, y | D_i, \Theta) \times d_{x,y}$$

[en ignorant la correction pour substitutions multiples]

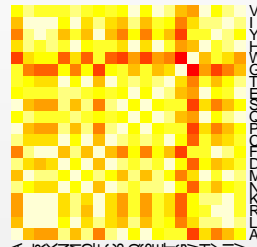
- Exemples de matrices possibles :
 - Distance synthétique (Grantham, Miyata, ...)
 - Différence de **volume**



Prise en compte de la nature biochimique des changements



Grantham



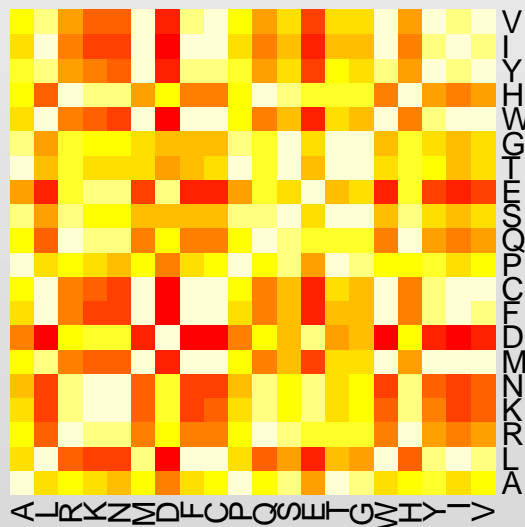
Volume

- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,
- On peut calculer des vecteurs de substitutions pondérés avec la formule :

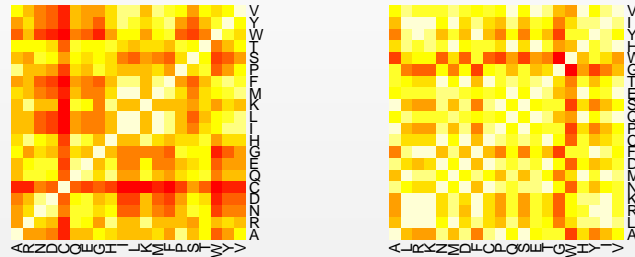
$$w_{i,b} = \sum_x \sum_y Pr(x, y | D_i, \Theta) \times d_{x,y}$$

[en ignorant la correction pour substitutions multiples]

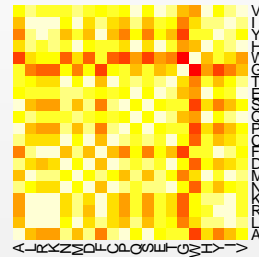
- Exemples de matrices possibles :
 - Distance synthétique (Grantham, Miyata, ...)
 - Différence de volume, de **polarité**



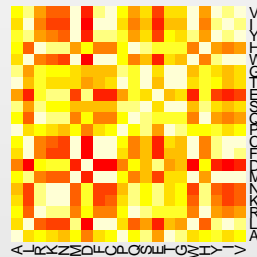
Prise en compte de la nature biochimique des changements



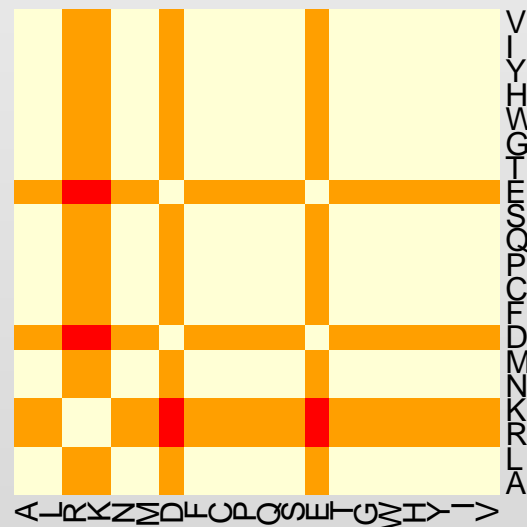
Grantham



Volume



Polarité



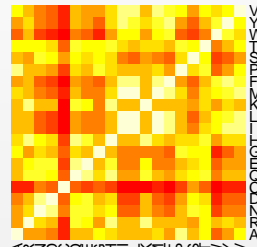
- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,
- On peut calculer des vecteurs de substitutions pondérés avec la formule :

$$w_{i,b} = \sum_x \sum_y Pr(x, y | D_i, \Theta) \times d_{x,y}$$

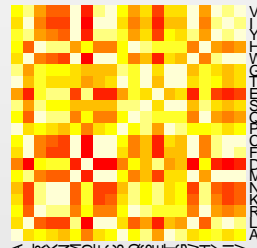
[en ignorant la correction pour substitutions multiples]

- Exemples de matrices possibles :
 - Distance synthétique (Grantham, Miyata, ...)
 - Différence de volume, de polarité, de **charge**, ...

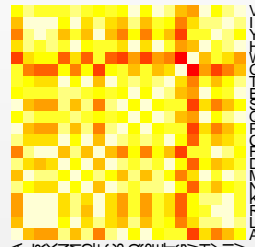
Prise en compte de la nature biochimique des changements



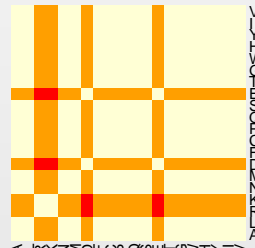
Grantham



Polarité



Volume



Charge

- Soit D une matrice de poids, ou $d_{x,y}$ est la pondération pour le changement $x \rightarrow y$,
- On peut calculer des vecteurs de substitutions pondérés avec la formule :

$$w_{i,b} = \sum_x \sum_y Pr(x, y | D_i, \Theta) \times d_{x,y}$$

[en ignorant la correction pour substitutions multiples]

- Exemples de matrices possibles :
 - Distance synthétique (Grantham, Miyata, ...)
 - Différence de volume, de polarité, de charge, ...
 - Descriptive : 0 si les 2 AA ont la même taille, 1 sinon, ... (équivalent à l'utilisation de 'sous-alphabets')

Site-clustering

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements

- **Site-clustering**

- Mesure de soutien

- Résultat

- Simulations
- Résultats sur la myoglobine

- Quelques prédictions sur la myoglobine: Un quadruplet

- Quelques prédictions sur la myoglobine: Un pentuplet

- Les récepteurs impliqués dans l'autoincompatibilité

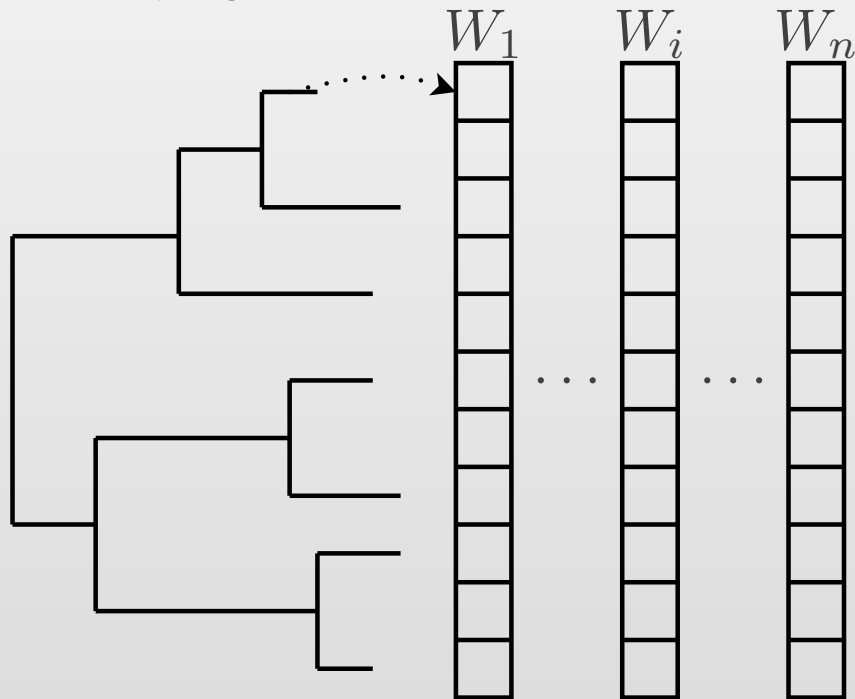
- Sites coévoluant chez SRK

- Conclusions et Perspectives

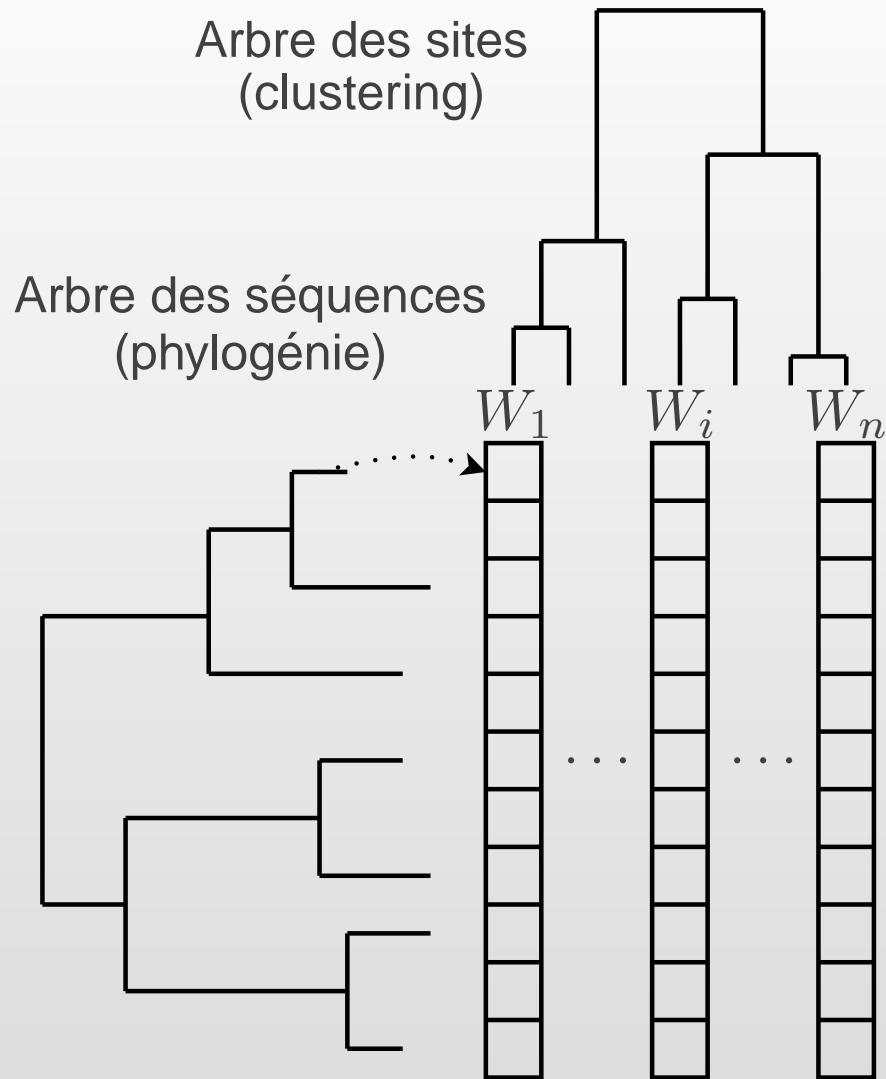
- Clustering hiérarchique agglomératif,
- Distance $1 - cor(y_i, y_j)$,
- Chaque nombre de substitution est divisé par le nombre total de substitution sur la branche,
- Plusieurs méthodes ont été comparées (distances, clustering),
- ... Pour chaque type de vecteur :
 - Substitutions multiples,
 - Pondération par distance de Grantham,
 - Pondération par différence de volume,
 - Pondération par différence de polarité,
 - Pondération par différence de charge.

Mesure de soutien

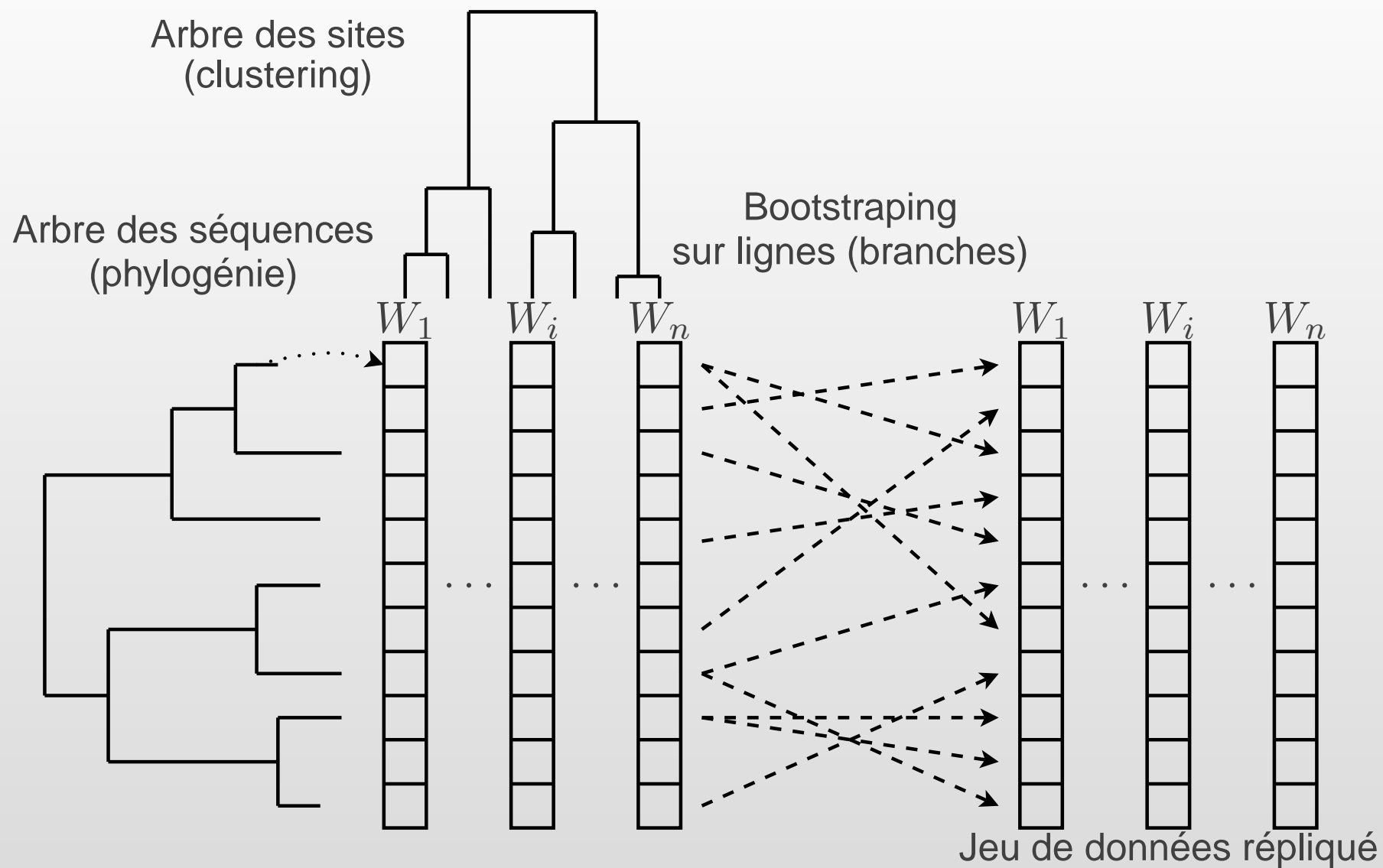
Arbre des séquences
(phylogénie)



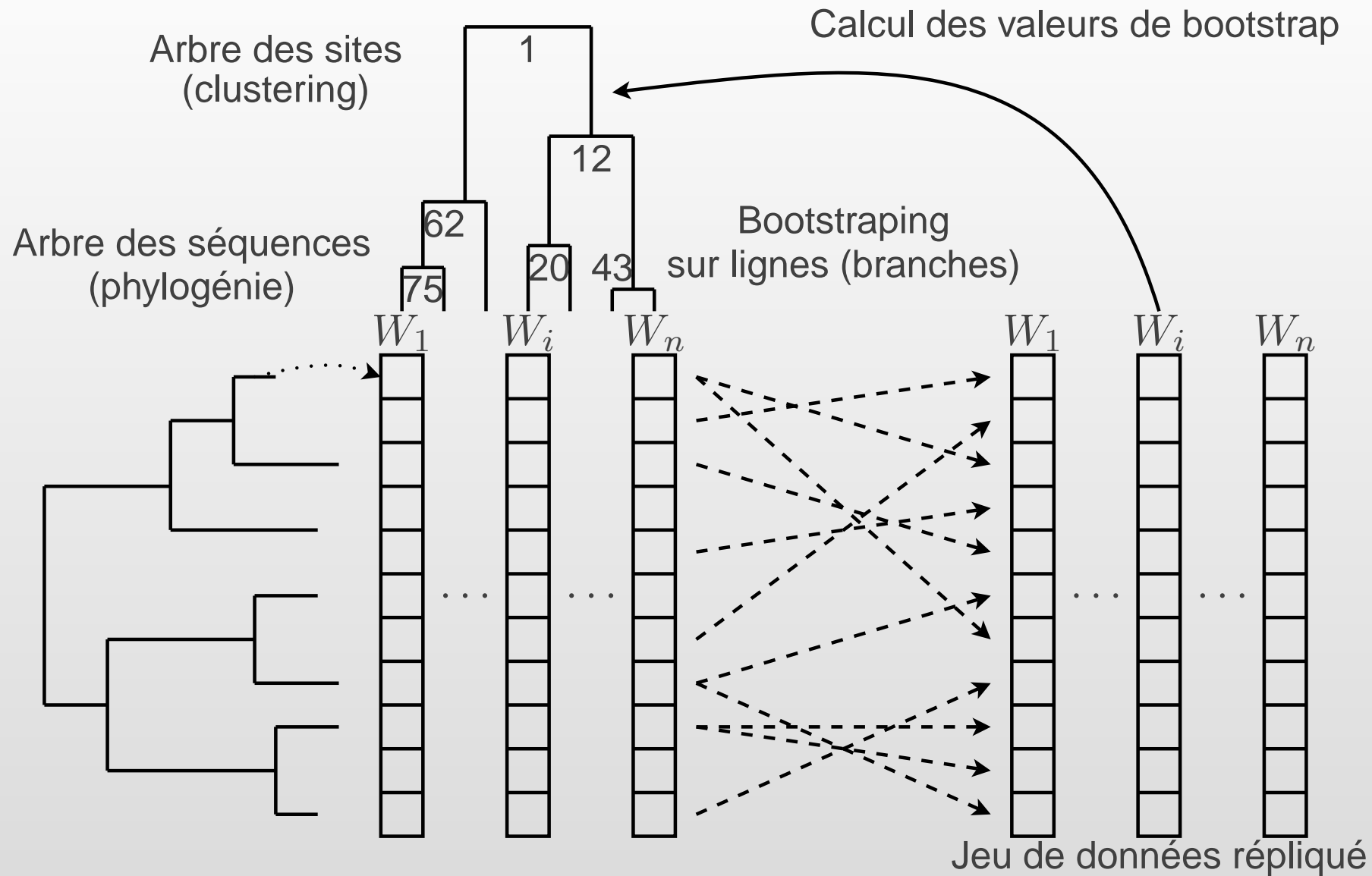
Mesure de soutien



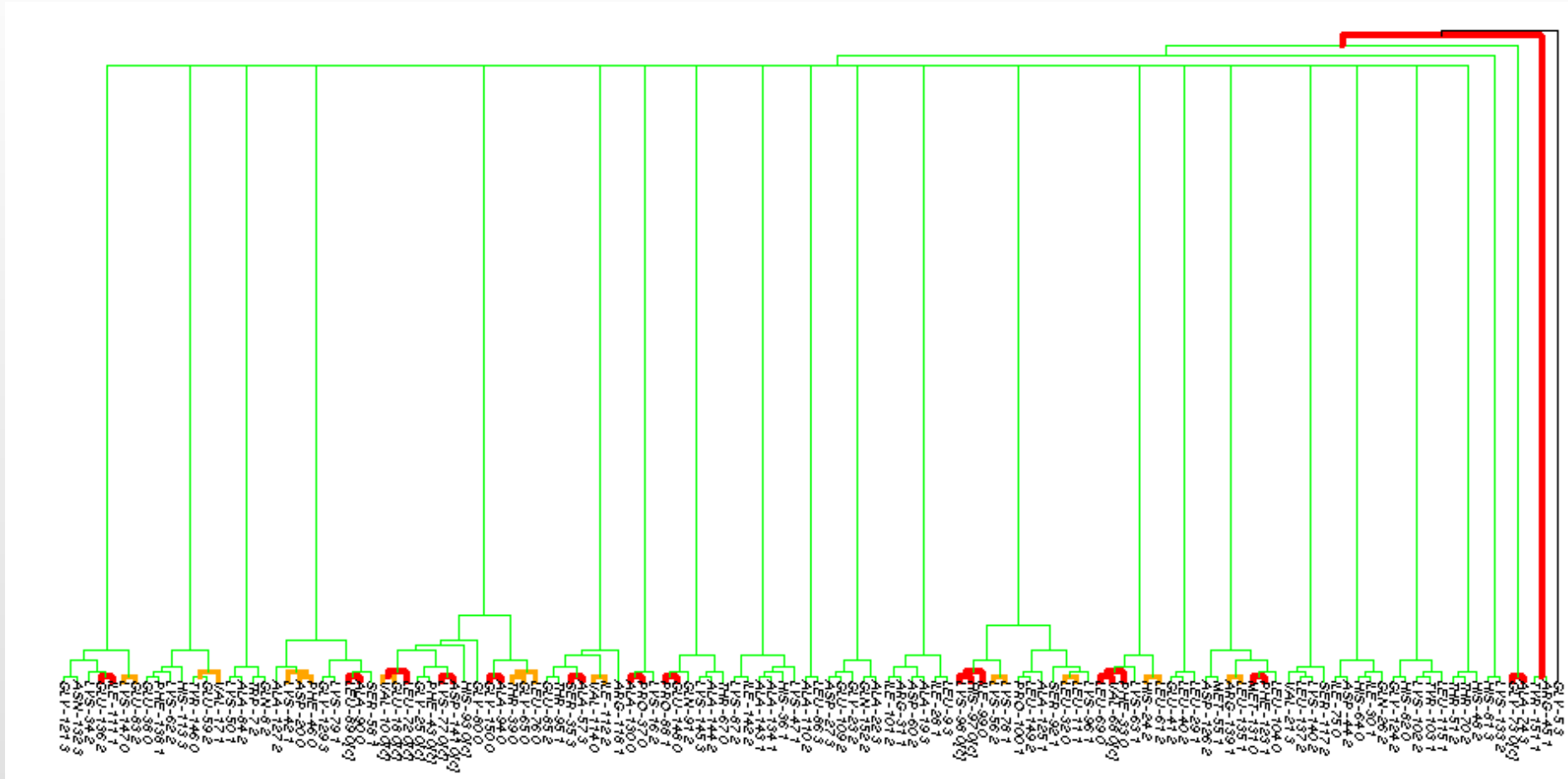
Mesure de soutien



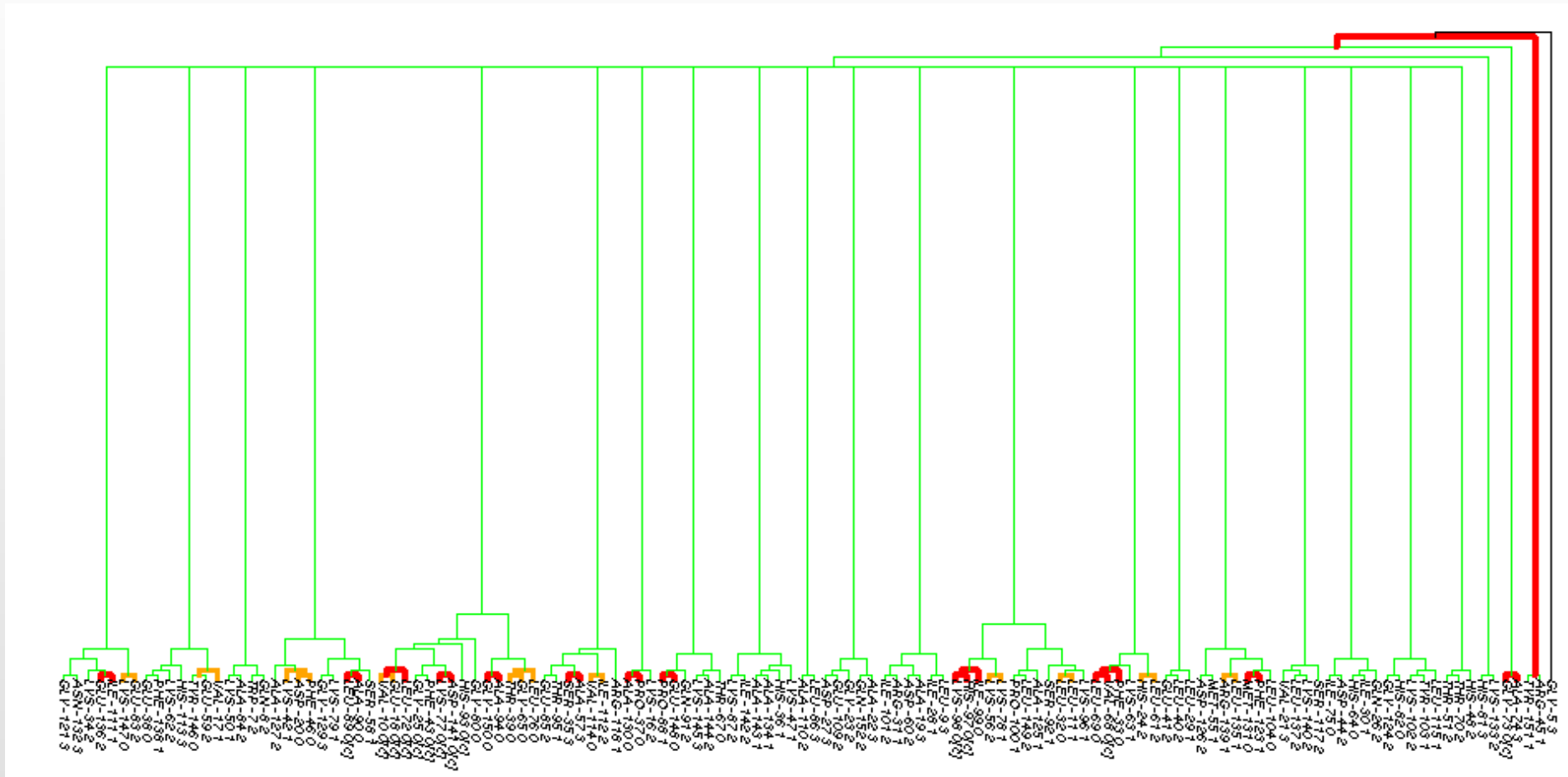
Mesure de soutien



Résultat

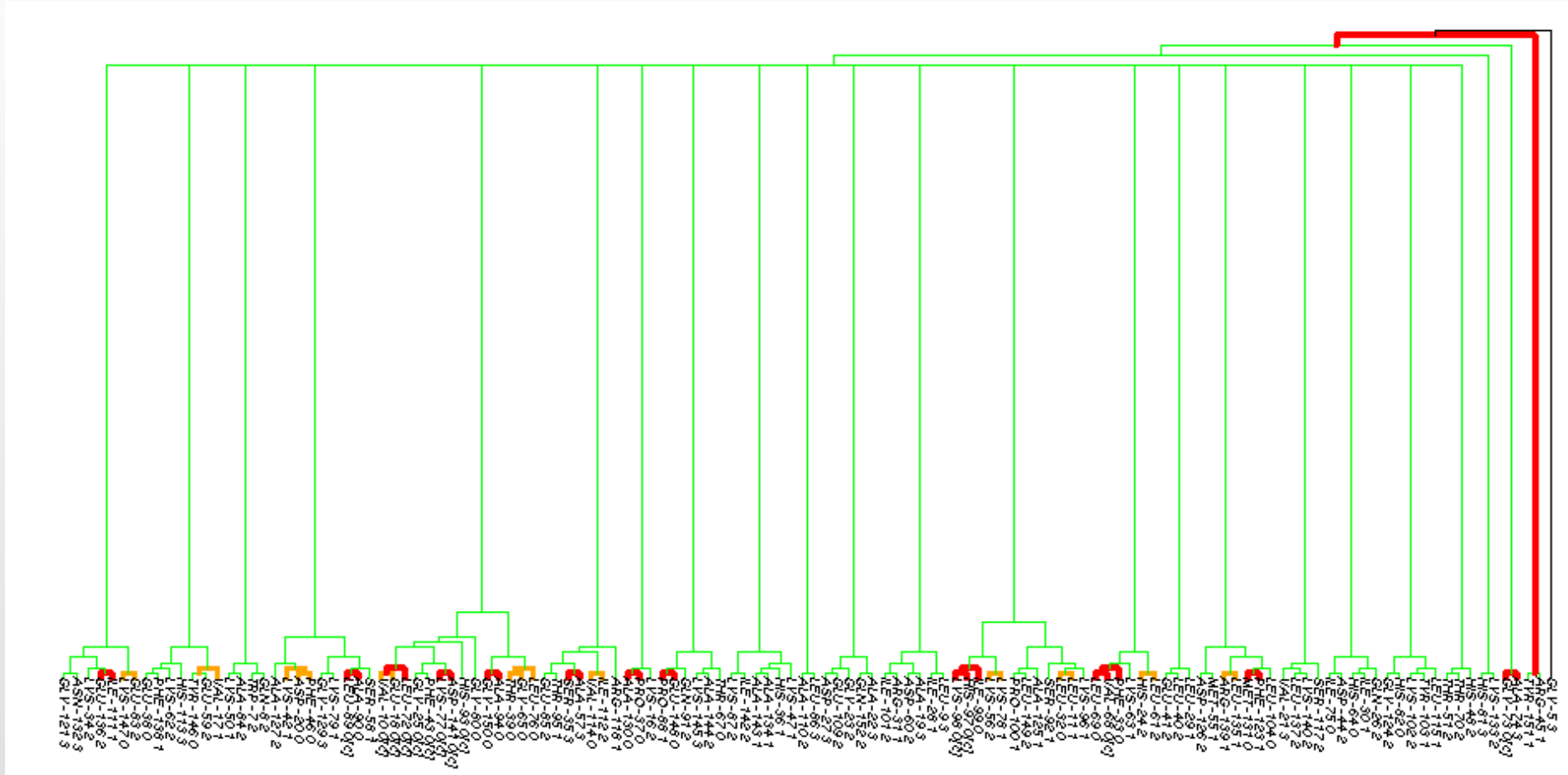


Résultat



- Les sites peu informatifs sont regroupés artificiellement avec un fort support,

Résultat



- Les sites peu informatifs sont regroupés artificiellement avec un fort support,
- Probabilités d'obtenir de telles valeurs sous l'hypothèse nulle ?

Simulations

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- **Simulations**
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqué dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

1. On simule 50 jeux de données selon les paramètres estimés sur le jeu de données, avec **indépendance entre les sites**,

Simulations

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- **Simulations**
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqué dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

1. On simule 50 jeux de données selon les paramètres estimés sur le jeu de données, avec **indépendance entre les sites**,
2. On calcule les vecteurs sur chaque jeu de données, sans réestimer les paramètres,

Simulations

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- **Simulations**
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqué dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

1. On simule 50 jeux de données selon les paramètres estimés sur le jeu de données, avec **indépendance entre les sites**,
2. On calcule les vecteurs sur chaque jeu de données, sans réestimer les paramètres,
3. On effectue la procédure de bootstrap (1000 répétitions)

Simulations

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements

- Site-clustering

- Mesure de soutien

- Résultat

- **Simulations**

- Résultats sur la myoglobine

- Quelques prédictions sur la myoglobine: Un quadruplet

- Quelques prédictions sur la myoglobine: Un pentuplet

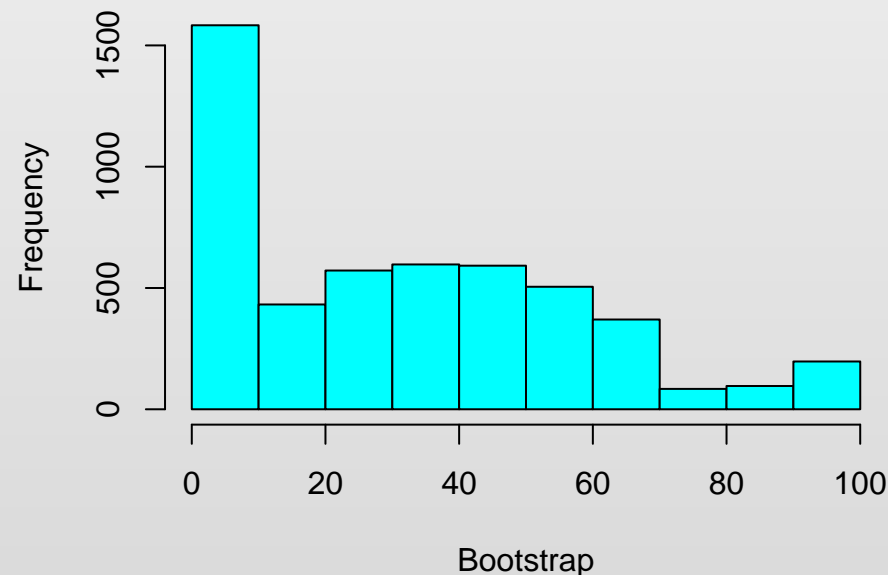
- Les récepteurs impliqués dans l'autoincompatibilité

- Sites coévoluant chez SRK

- Conclusions et Perspectives

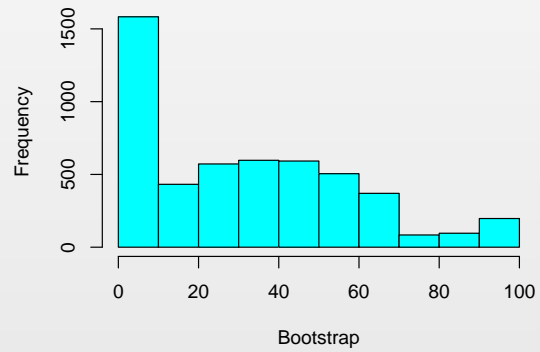
1. On simule 50 jeux de données selon les paramètres estimés sur le jeu de données, avec **indépendance entre les sites**,
2. On calcule les vecteurs sur chaque jeu de données, sans réestimer les paramètres,
3. On effectue la procédure de bootstrap (1000 répétitions)
4. On regarde les valeurs de bootstrap obtenues :

Distribution sous l'hypothèse d'indépendance

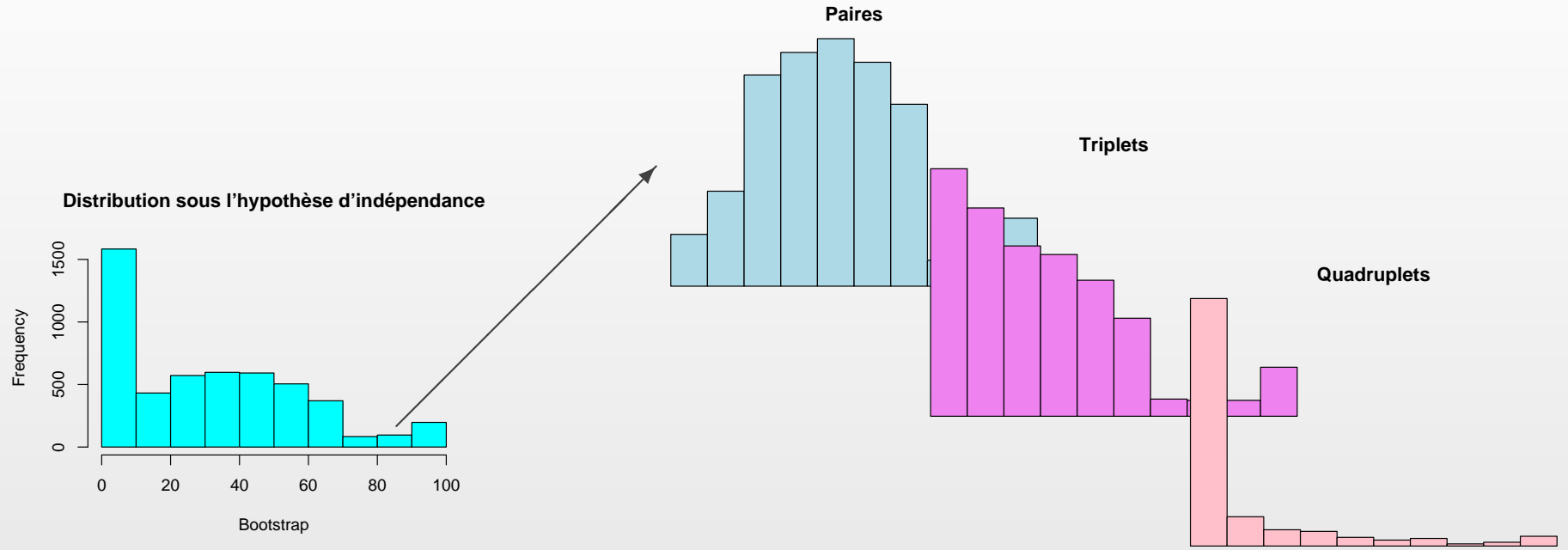


Simulations (2)

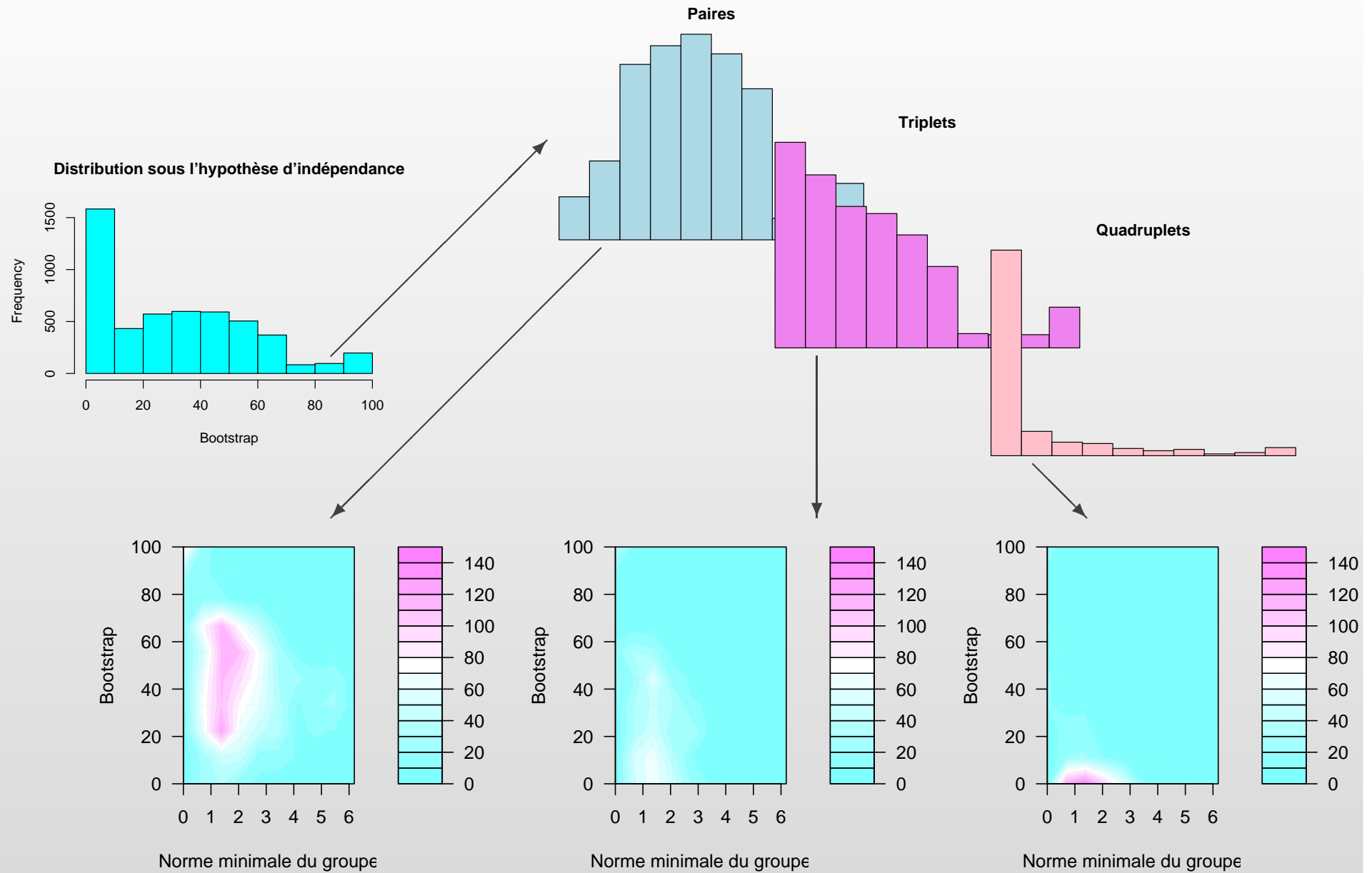
Distribution sous l'hypothèse d'indépendance



Simulations (2)



Simulations (2)



Résultats sur la myoglobine

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqué dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

- p-value dans 3 dimensions : taille du groupe, norme minimale du groupe, bootstrap,
- On ajuste une enveloppe de telle sorte qu'elle contienne 95% des valeurs obtenues par simulation,
- Les valeurs observées hors de cette enveloppe sont considérées comme significativement co-évoluant, on peut les lister :

Résultats sur la myoglobine

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements

- Site-clustering
- Mesure de soutien
- Résultat

- Simulations
- Résultats sur la myoglobine

- Quelques prédictions sur la myoglobine: Un quadruplet

- Quelques prédictions sur la myoglobine: Un pentuplet

- Les récepteurs impliqué dans l'autoincompatibilité

- Sites coévoluant chez SRK

- Conclusions et Perspectives

- p-value dans 3 dimensions : taille du groupe, norme minimale du groupe, bootstrap,
- On ajuste une enveloppe de telle sorte qu'elle contienne 95% des valeurs obtenues par simulation,
- Les valeurs observées hors de cette enveloppe sont considérées comme significativement co-évoluant, on peut les lister :
 - 25 groupes significatifs à 5% (attendu : 6.25, soit 4× plus),
 - 8 groupes significatifs à 1% (attendu : 1.25, soit 6.4× plus),
 - 4 groupes significatifs à 0.1% (attendu : 0.125, soit 32× plus),
 - Par méthode :

| | Multiple | Grantham | Volume | Polarité | Charge |
|------|----------|----------|--------|----------|--------|
| 5% | 9 | 3 | 8 | 0 | 5 |
| 1% | 2 | 0 | 3 | 0 | 3 |
| 0.1% | 2 | 0 | 0 | 0 | 2 |

Quelques prédictions sur la myoglobine : Un quadruplet

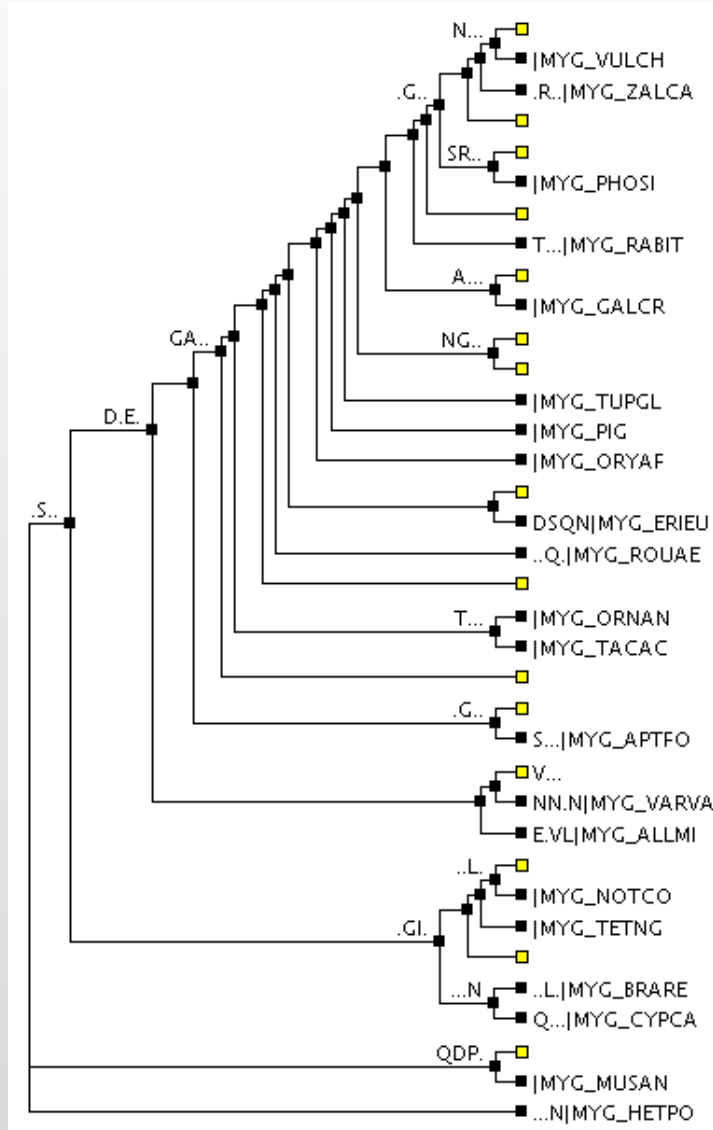
- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqués dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

SER35, ALA57, GLU85 et THR95



Quelques prédictions sur la myoglobine : Un quadruplet

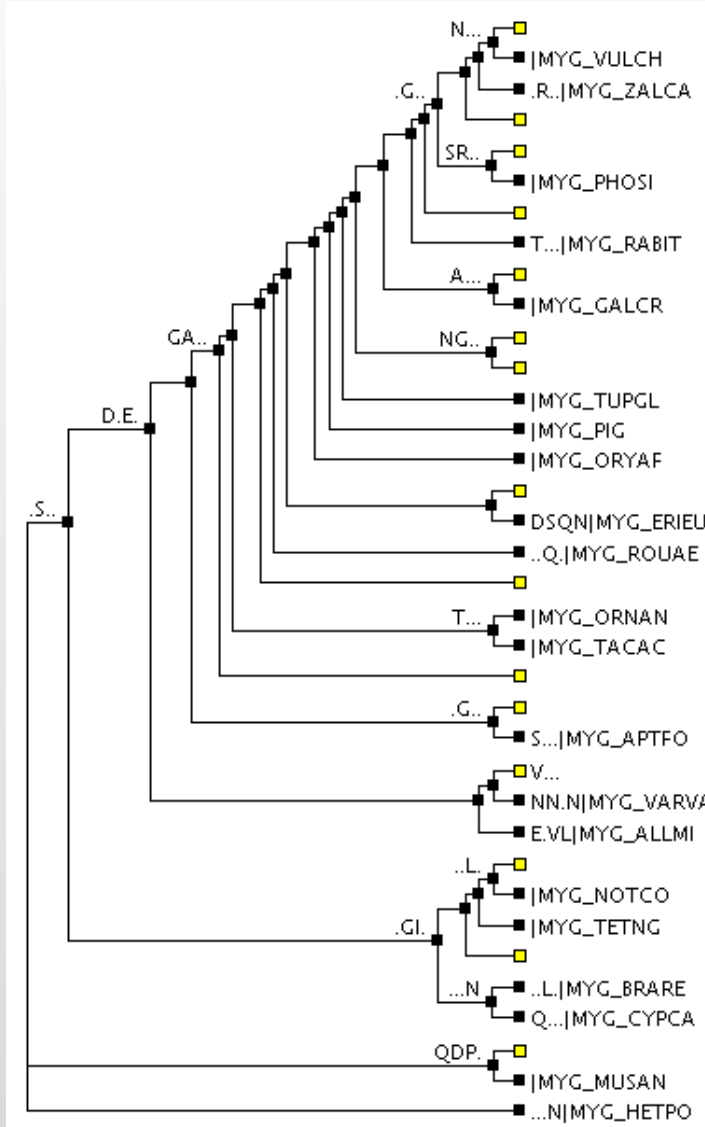
- Coévolution moléculaire au niveau des sites

Cartographie de substitution

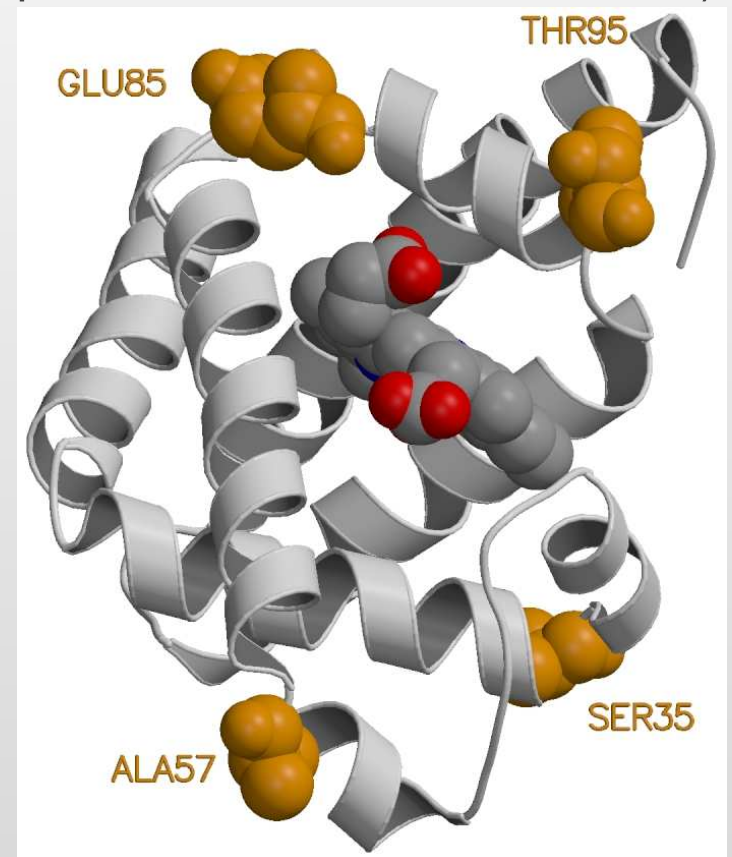
Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- **Quelques prédictions sur la myoglobine: Un quadruplet**
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqués dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

SER35, ALA57, GLU85 et THR95



DéTECTÉS avec la méthode "multiple", bootstrap de 31%*** (pour le quadruplet, 93%*** pour SER35 et ALA57).



Quelques prédictions sur la myoglobine : Un pentuplet

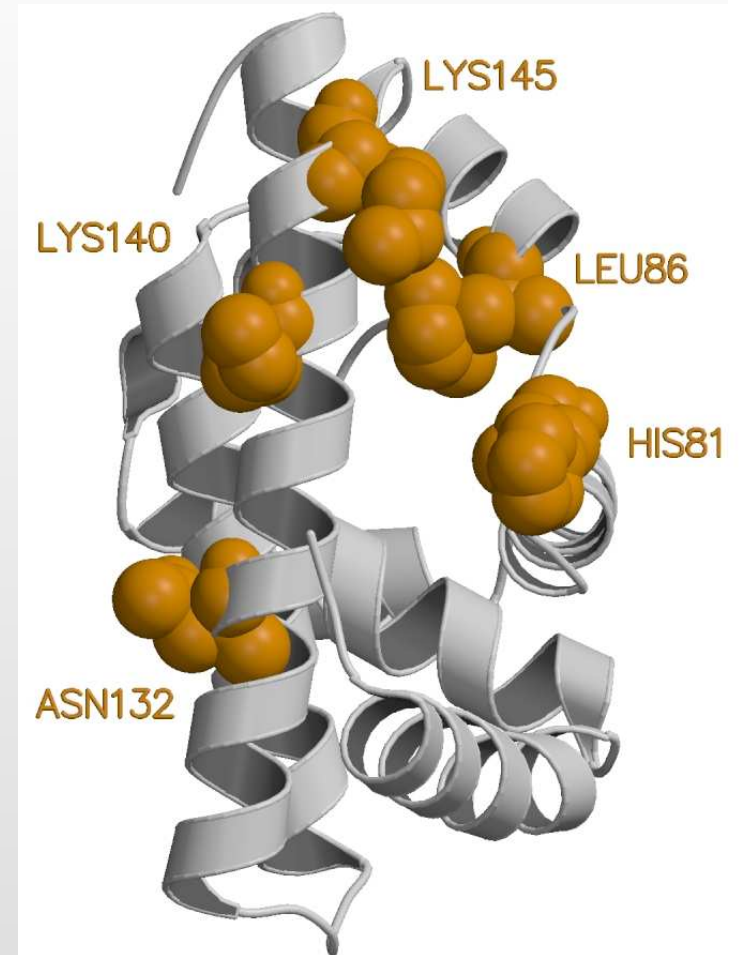
- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- **Quelques prédictions sur la myoglobine: Un pentuplet**
- Les récepteurs impliqué dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

- Les 5 : volume (73%**)
- Les lysines : volume (81%*) et charge (90%*)
- Les lysines + arginine : charge (86%***)



Test 3D significatif
(p.value=0.014).

Quelques prédictions sur la myoglobine : Un pentuplet

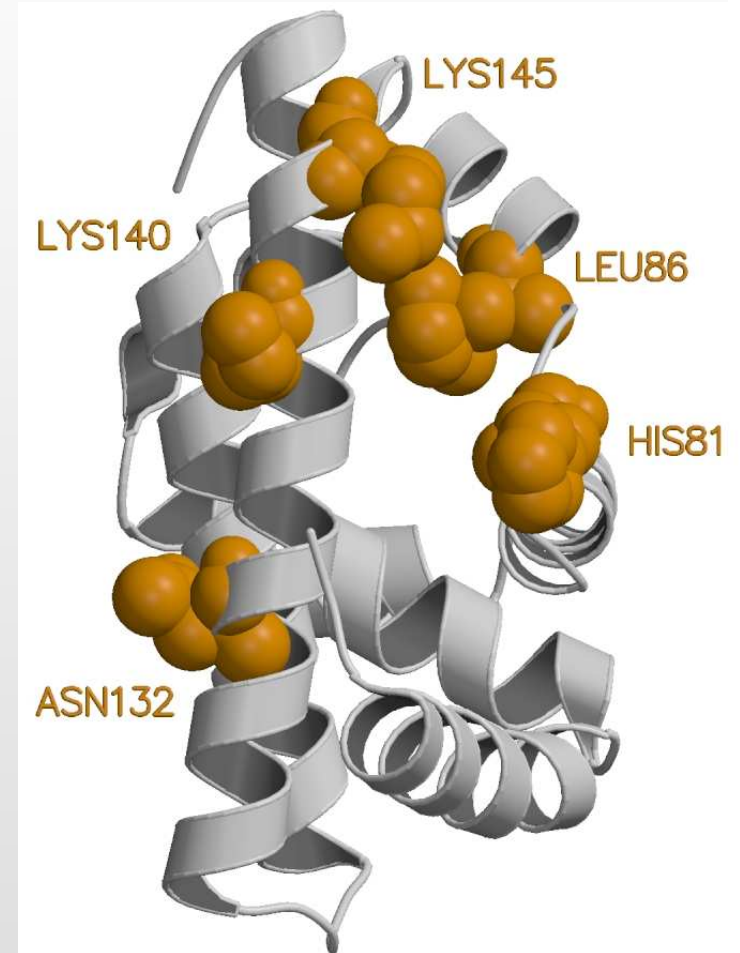
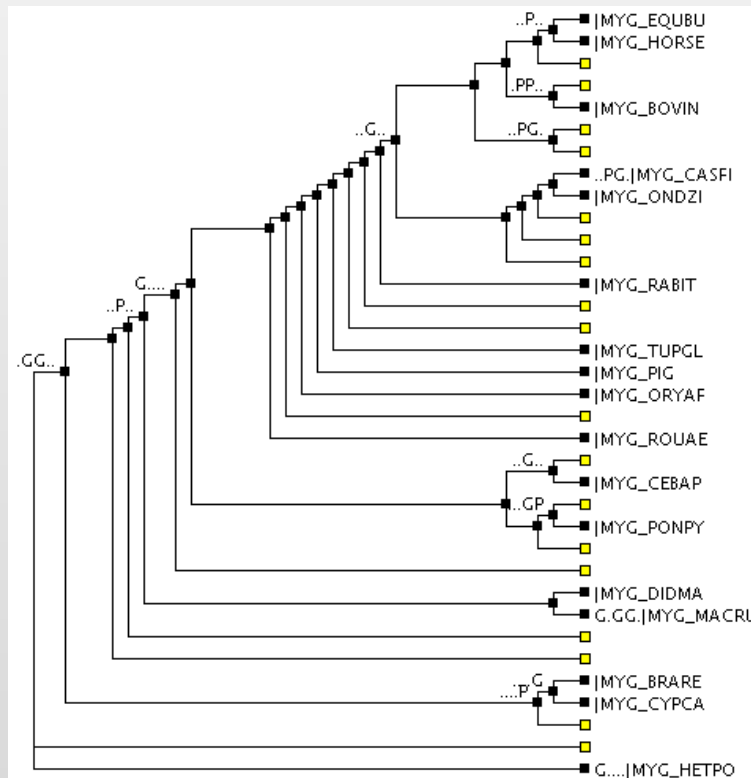
- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- **Quelques prédictions sur la myoglobine: Un pentuplet**
- Les récepteurs impliqués dans l'autoincompatibilité
- Sites coévoluant chez SRK
- Conclusions et Perspectives

- Les 5 : volume (73%**)
- Les lysines : volume (81%*) et charge (90%*)
- Les lysines + arginine : charge (86%***)



Test 3D significatif
(p.value=0.014).

Les récepteurs impliqués dans l'autoincompatibilité

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements

- Site-clustering
- Mesure de soutien
- Résultat

- Simulations
- Résultats sur la myoglobine

- Quelques prédictions sur la myoglobine: Un quadruplet

- Quelques prédictions sur la myoglobine: Un pentuplet

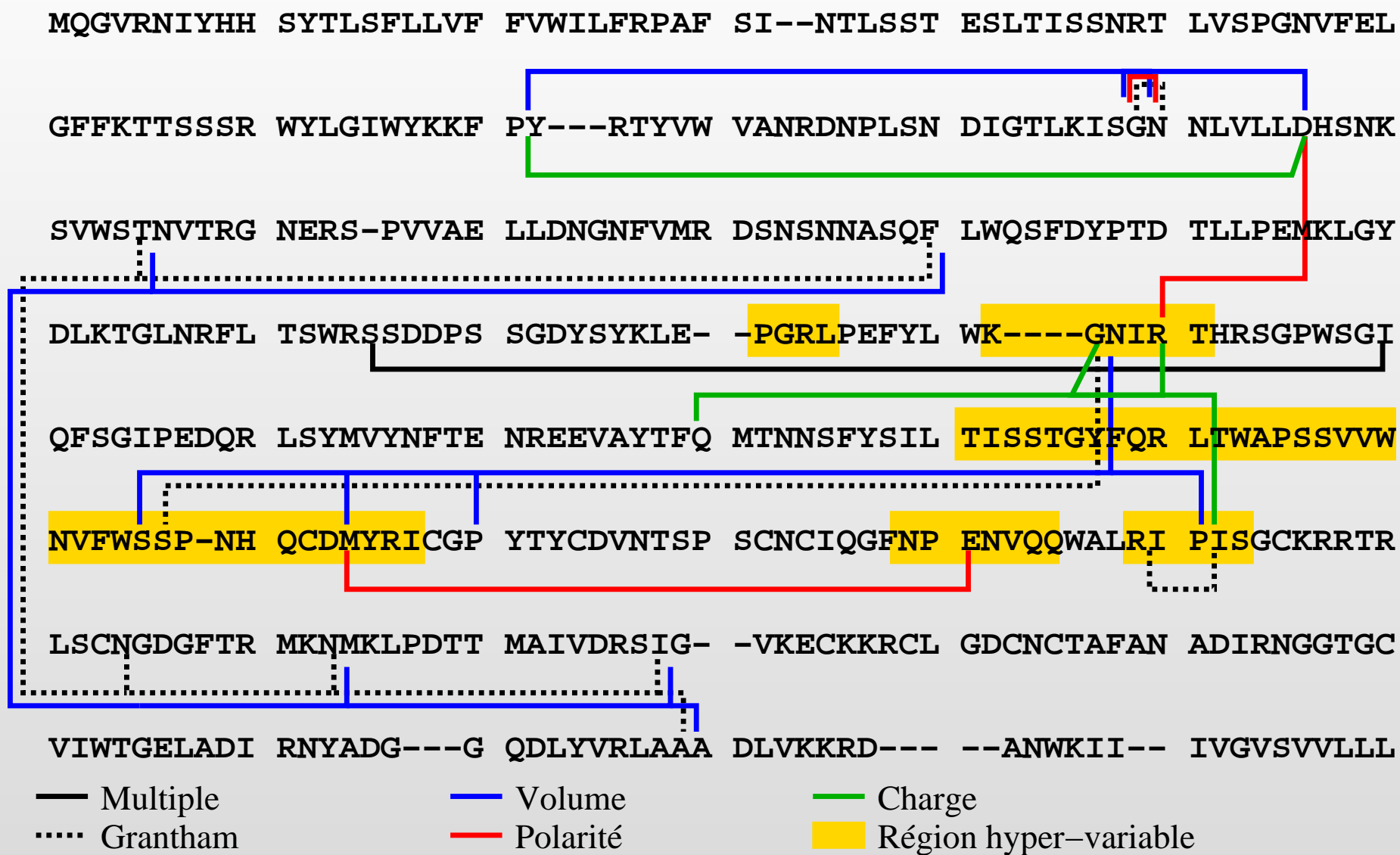
- **Les récepteurs impliqués dans l'autoincompatibilité**

- Sites coévoluant chez SRK

- Conclusions et Perspectives

- Coévolution entre le récepteur et son ligand,
- Chaque nouvel allèle est favorisé car il permet à son possesseur de se reproduire avec le plus grand nombre d'individus,
- Les différences entre allèles sont concentrées au sein de région hyper-variables,
- 2 grands types :
 - Incompatibilité sporophytique, SRK (*Brassicaceae*) : 53 séquences pour 386 sites complets (19 groupes à 1%).
 - Incompatibilité gamétophytique, RNase
 - (i) *Solanaceae* : 146 séquences pour 76 sites (5 groupes à 1%)
 - (ii) *Rosaceae* : 61 séquences pour 107 sites (13 groupes à 1%)
- Des sites sous sélection positive ont été mis en évidence : Sainudiin *et al.* 2005.

Sites coévoluant chez SRK



Conclusions et Perspectives

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements

- Site-clustering

- Mesure de soutien

- Résultat

- Simulations

- Résultats sur la myoglobine

- Quelques prédictions sur la myoglobine: Un quadruplet

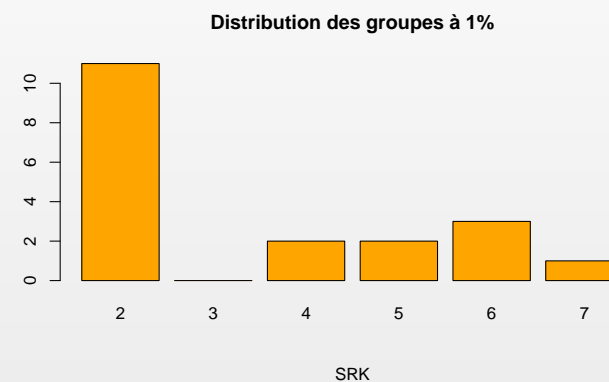
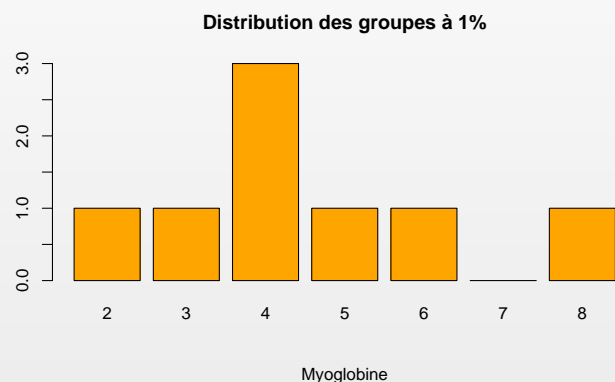
- Quelques prédictions sur la myoglobine: Un pentuplet

- Les récepteurs impliqué dans l'autoincompatibilité

- Sites coévoluant chez SRK

- Conclusions et Perspectives

- On détecte des groupes avec plus de 2 sites :



Conclusions et Perspectives

- Coévolution moléculaire au niveau des sites

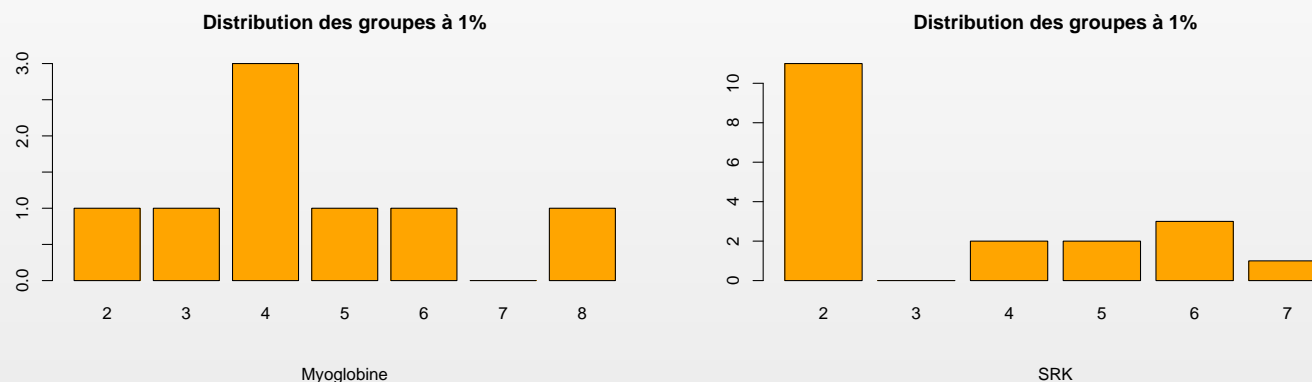
Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements
- Site-clustering
- Mesure de soutien
- Résultat
- Simulations
- Résultats sur la myoglobine
- Quelques prédictions sur la myoglobine: Un quadruplet
- Quelques prédictions sur la myoglobine: Un pentuplet
- Les récepteurs impliqué dans l'autoincompatibilité
- Sites coévoluant chez SRK

● Conclusions et Perspectives

- On détecte des groupes avec plus de 2 sites :



- Projets en cours :

- Utilisation de la méthode pour détecter des interactions entre gènes (application au complexe COX),
- Test à grande échelle des relations structure/coévolution

Conclusions et Perspectives

- Coévolution moléculaire au niveau des sites

Cartographie de substitution

Application de la méthode aux protéines

- Myoglobine
- Prise en compte de la nature biochimique des changements

- Site-clustering

- Mesure de soutien

- Résultat

- Simulations

- Résultats sur la myoglobine

- Quelques prédictions sur la myoglobine: Un quadruplet

- Quelques prédictions sur la myoglobine: Un pentuplet

- Les récepteurs impliqué dans l'autoincompatibilité

- Sites coévoluant chez SRK

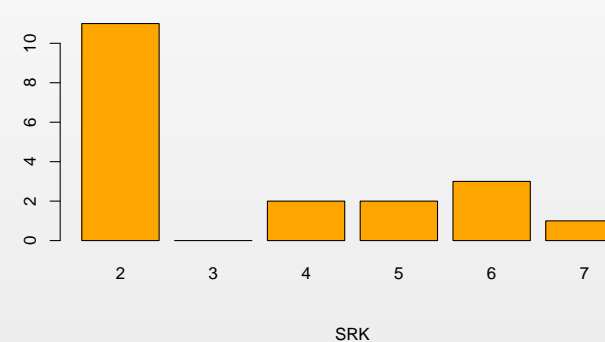
- Conclusions et Perspectives

- On détecte des groupes avec plus de 2 sites :

Distribution des groupes à 1%



Distribution des groupes à 1%



- Projets en cours :

- Utilisation de la méthode pour détecter des interactions entre gènes (application au complexe COX),
- Test à grande échelle des relations structure/coévolution

- Améliorations du programme de détection :

- Les distributions de bootstrap sous l'hypothèse nulle semblent assez similaires \Rightarrow détection rapide sans simulation,
- Accélérer le clustering en limitant la taille des groupes.