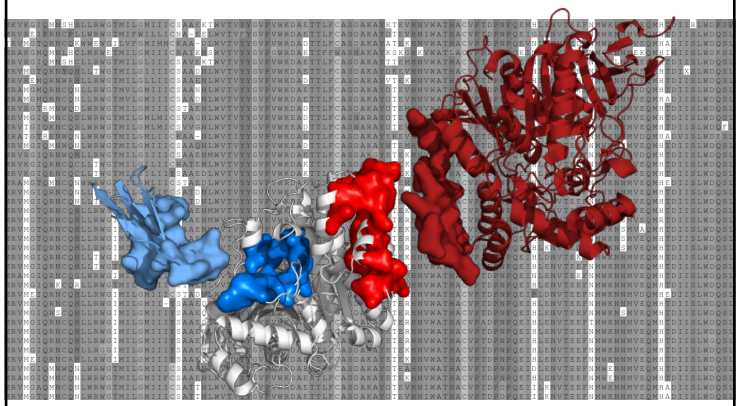
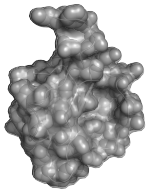
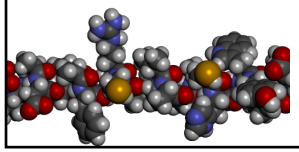
La co-évolution des protéines et le monde des virus



Alessandra Carbone, Laboratoire de Biologie Computationnelle et Quantitative,
Sorbonne Université - CNRS, France

Protéines

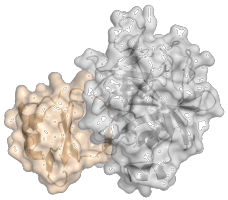




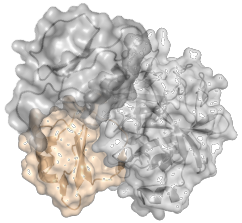
YTVPSFTFSRSKELAKLLRLHAK....

Dans un alphabet de 20 lettres

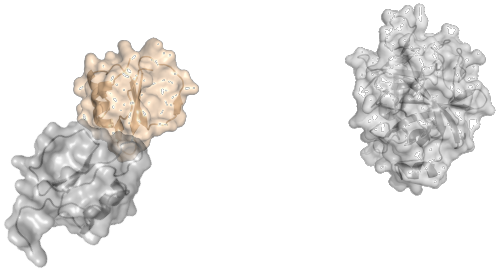
Les protéines interagissent par contacte physique



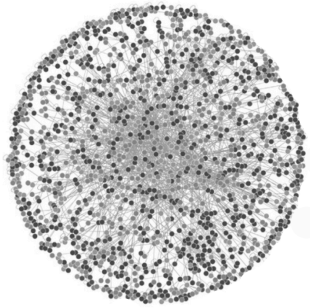
Les protéines interagissent par contacte physique



Les protéines interagissent par contacte physique



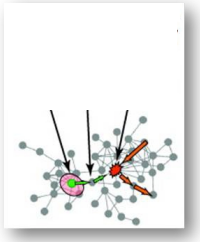
Reconstruction d'un réseaux



Génomes bactériens/eucaryotes

Qu'est-ce que les biologistes aimeraient connaître?

Réseaux

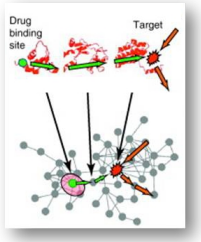


Pour cibler les interactions

Qu'est-ce que les biologistes aimeraient connaitre?

Structures

Réseaux

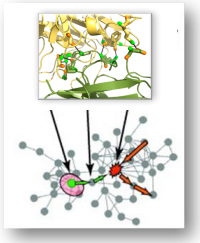


Pour cibler les interactions

Qu'est-ce que les biologistes aimeraient connaitre?

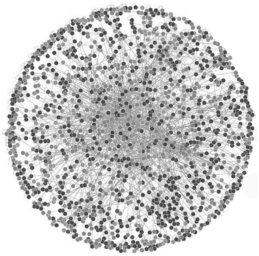
Structures

Réseaux

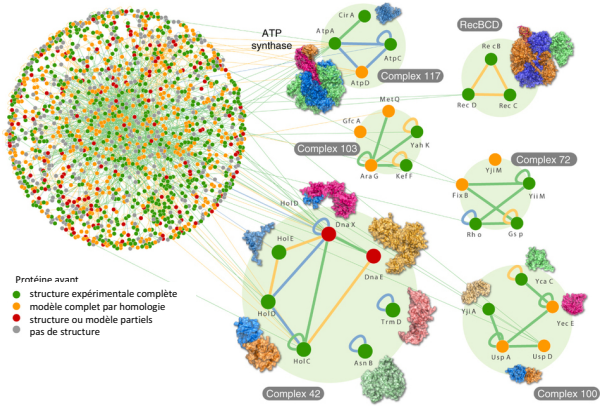


Pour cibler les interactions

Connaissance structurale partielle

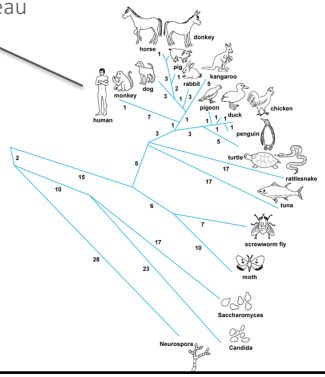


Connaissance structurale partielle

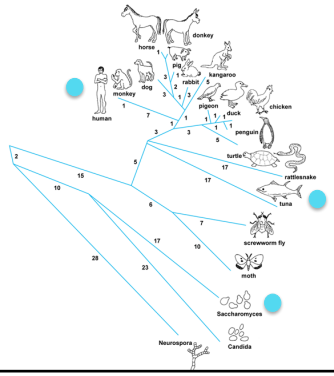


Elles ont une histoire

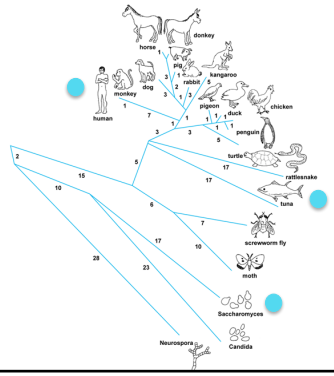
Nous sommes interessés à construir un réseau chez l'homme



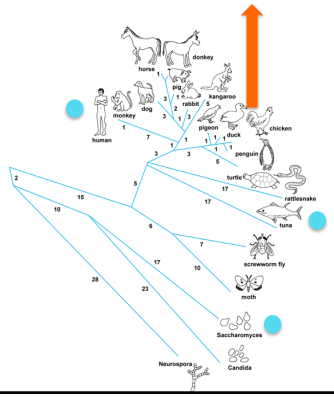
levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDAVHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AGGS--FVK--



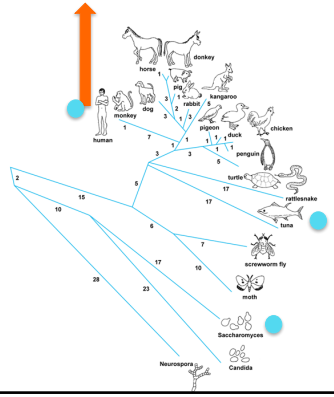
levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDAVHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AGGS--FVK--



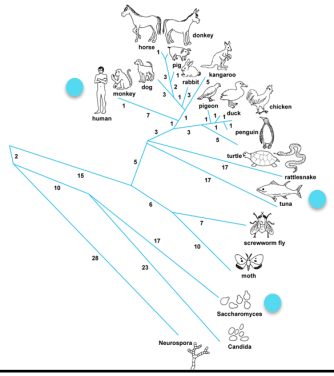
levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDAVHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AGGS--FVK--



levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDAVHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AGGS--FVK--



levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDAVHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AGGS--FVK--

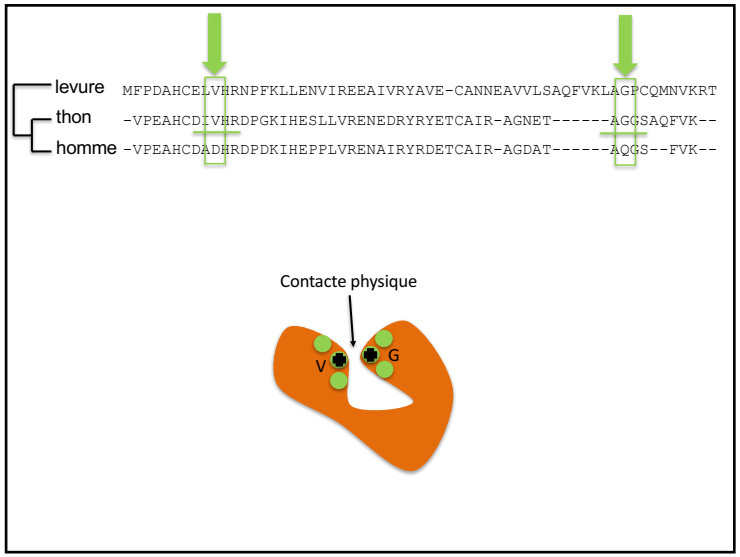


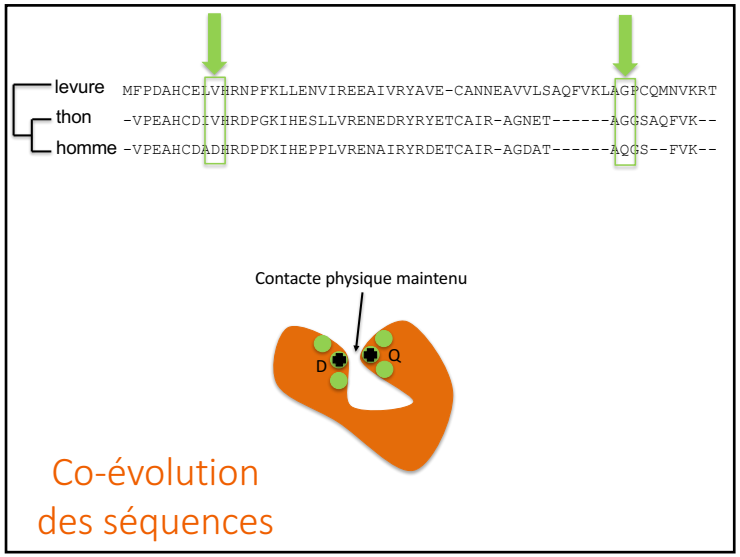
levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDAVHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AGGS--FVK--

Évolution des séquences

levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDADHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AQGS--FVK--

levure MFPDAHCELVHRNPFKLLENVIREEAIVRYAVE-CANNEAVVLSAQFVKLAGPCQMNVKRT
thon -VPEAHCDIVHRDPGKIHESLLVRENEDRYRYETCAIR-AGNET------AGGSAQFVK--
homme -VPEAHCDADHRDPDKIHEPPLVRENAIRYRDETCAIR-AGDAT------AQGS--FVK--





Co-évolution des séquences

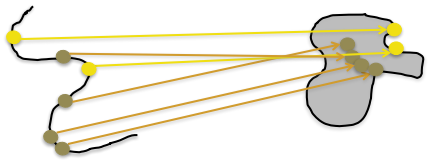
Elles ont aussi une histoire à deux

Co-évolution des séquences

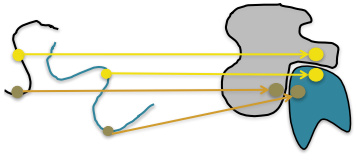
- Changements conformationnels
- contactes structuraux

- interaction physique
- Interaction fonctionnelle

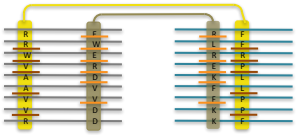
Identification les positions clé



Changements conformationnels et/ou des contactes structuraux

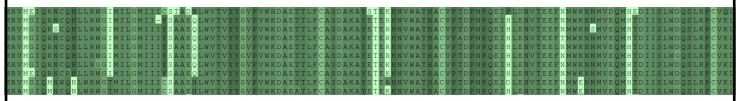


Marqueurs d'une interaction physique ou fonctionnelle

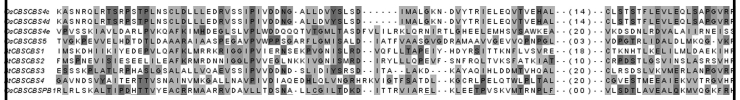


Inférence des positions clé par analyse de co-évolution

Séquences de « vertébrés » et de « virus »

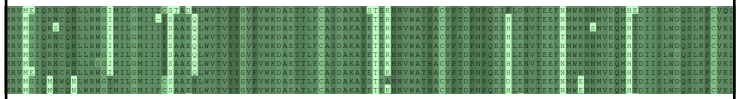


Alignements de séquences



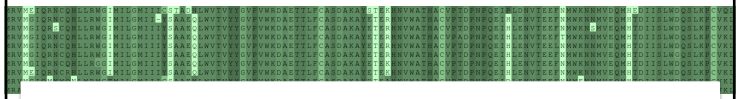
Séquences « bactériennes »

On mesure la "regularité" de la distribution et sa distance d'une "conservation complète" par rapport à très peu de mutations



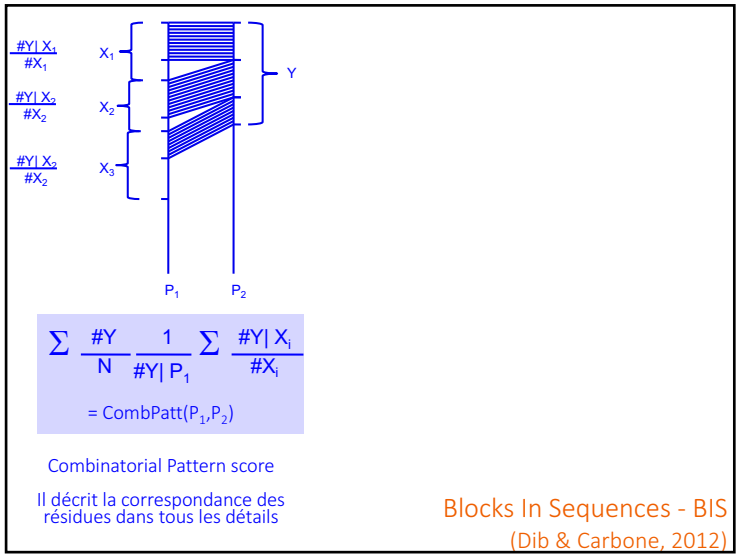
Les méthodes statistiques mesurent la distance de la distribution des acides-aminés du bruit de fond.

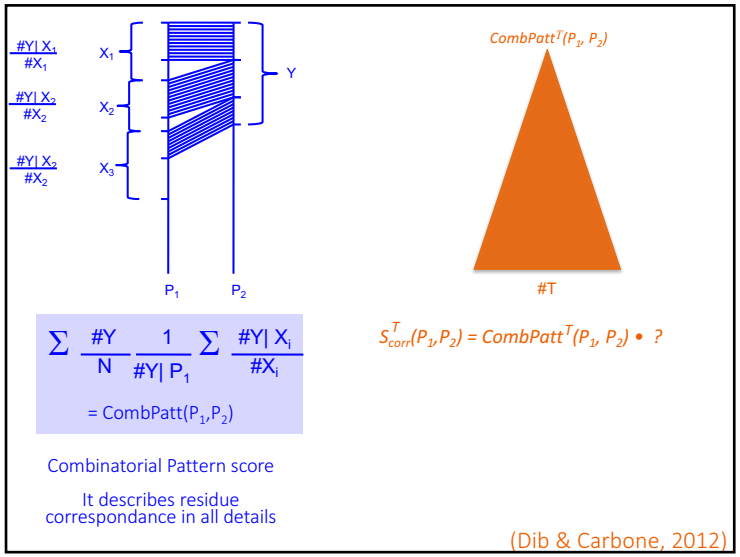
On mesure la "regularité" de la distribution et sa distance d'une "conservation complète" par rapport à très peu de mutations

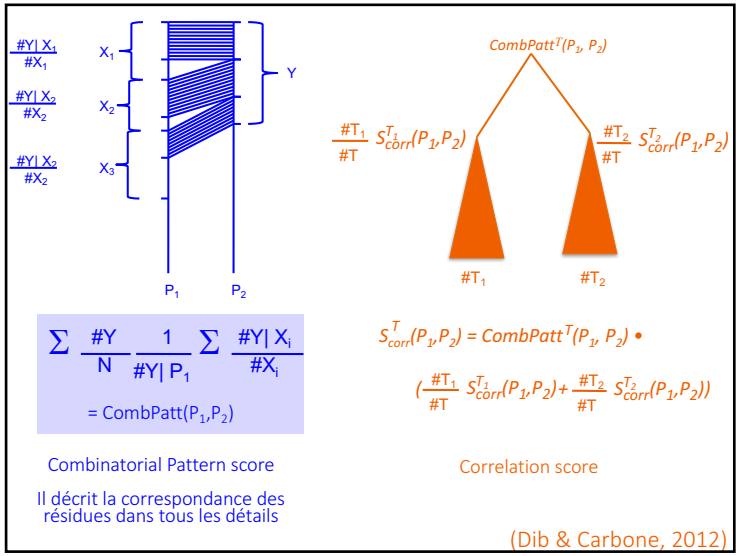


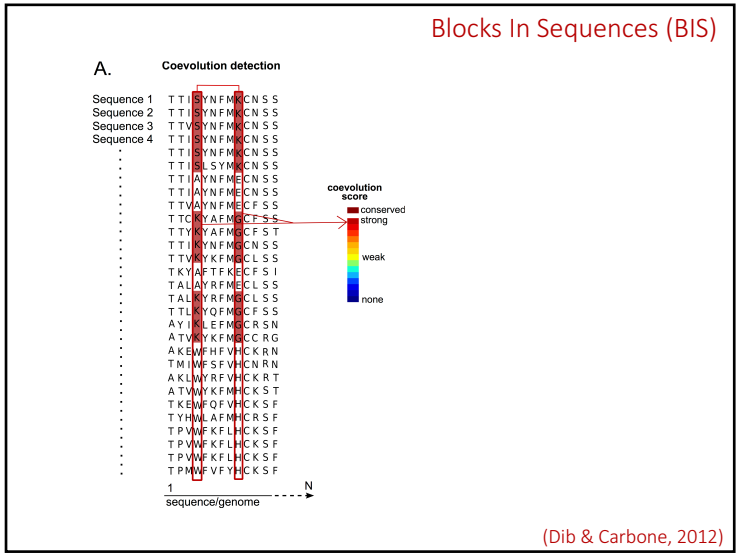
Un mot (rapide !) sur la méthode qui s'applique à des séquences conservées

Les méthodes statistiques mesurent la distance de la distribution des acides-aminés du bruit de fond.

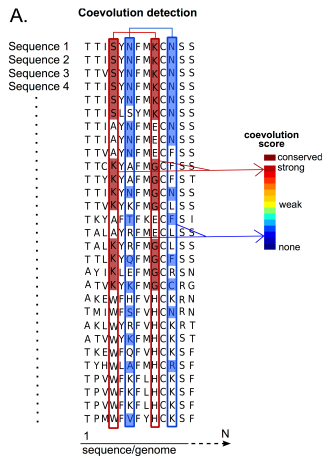

#Y| X1 / #X1
#Y| X2 / #X2
#Y| X2 / #X2
X1
X2
X3
Y
P1
P2
Σ #Y/N 1/#Y| P1 Σ #Y| Xi/#Xi
= CombPatt(P1,P2)
Combinatorial Pattern score
Il décrit la correspondance des résidues dans tous les détails
Blocks In Sequences - BIS
(Dib & Carbone, 2012)


#Y| X1 / #X1
#Y| X2 / #X2
#Y| X2 / #X2
X1
X2
X3
Y
P1
P2
CombPatt^T(P1, P2)
#T
Σ #Y/N 1/#Y| P1 Σ #Y| Xi/#Xi
= CombPatt(P1,P2)
S^T_corr(P1,P2) = CombPatt^T(P1, P2) • ?
Combinatorial Pattern score
It describes residue correspondance in all details
(Dib & Carbone, 2012)


#Y| X1 / #X1
#Y| X2 / #X2
#Y| X2 / #X2
X1
X2
X3
Y
P1
P2
CombPatt^T(P1, P2)
#T1/#T S^T1_corr(P1,P2)
#T2/#T S^T2_corr(P1,P2)
#T1
#T2
Σ #Y/N 1/#Y| P1 Σ #Y| Xi/#Xi
= CombPatt(P1,P2)
S^T_corr(P1,P2) = CombPatt^T(P1, P2) •
(#T1/#T S^T1_corr(P1,P2) + #T2/#T S^T2_corr(P1,P2))
Combinatorial Pattern score
Il décrit la correspondance des résidues dans tous les détails
Correlation score
(Dib & Carbone, 2012)

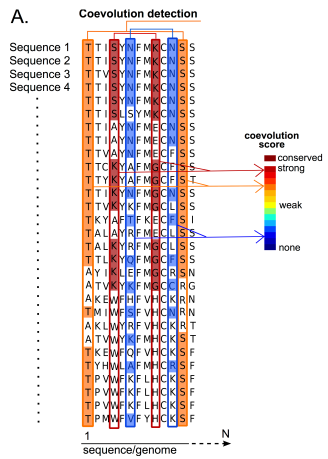

Blocks In Sequences (BIS)
A.
Coevolution detection
Sequence 1
Sequence 2
Sequence 3
Sequence 4
coevolution score
conserved
strong
weak
none
1
N
sequence/genome
(Dib & Carbone, 2012)

Blocks In Sequences (BIS)



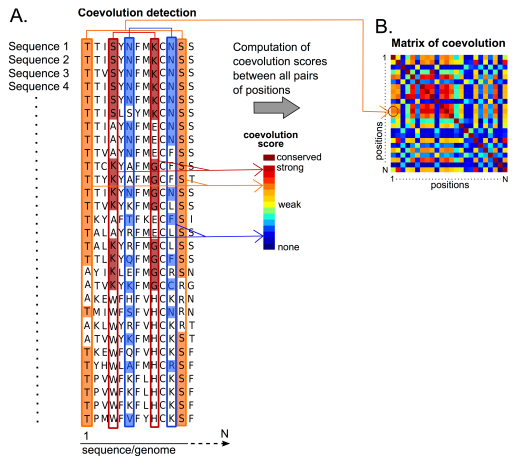
(Dib & Carbone, 2012)

Blocks In Sequences (BIS)



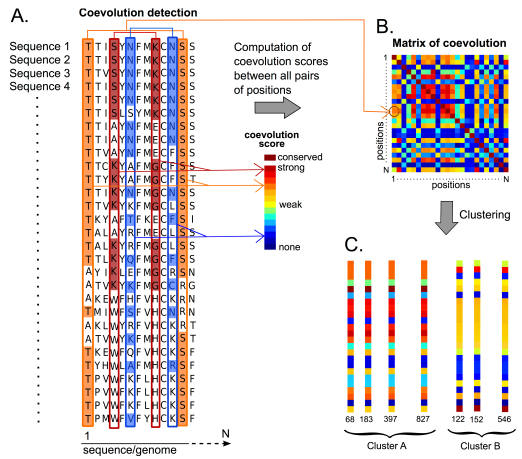
(Dib & Carbone, 2012)

Blocks In Sequences (BIS)



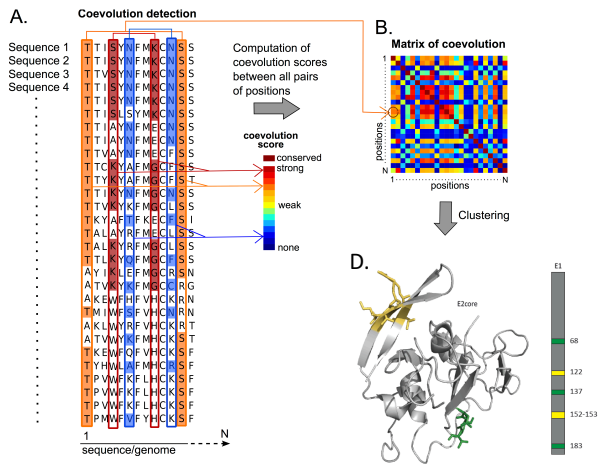
(Dib & Carbone, 2012)

Blocks In Sequences (BIS)



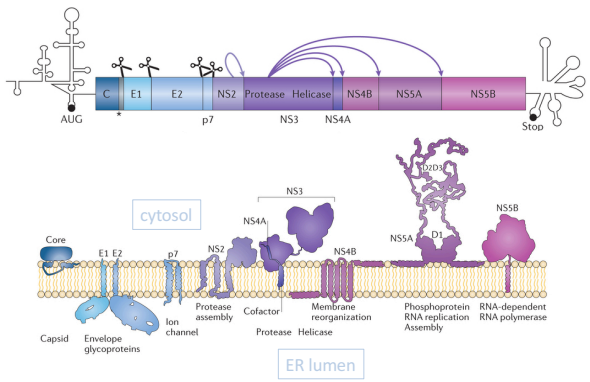
(Dib & Carbone, 2012)

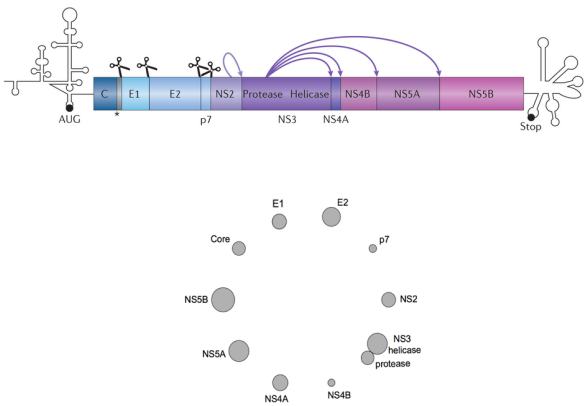
Blocks In Sequences (BIS)

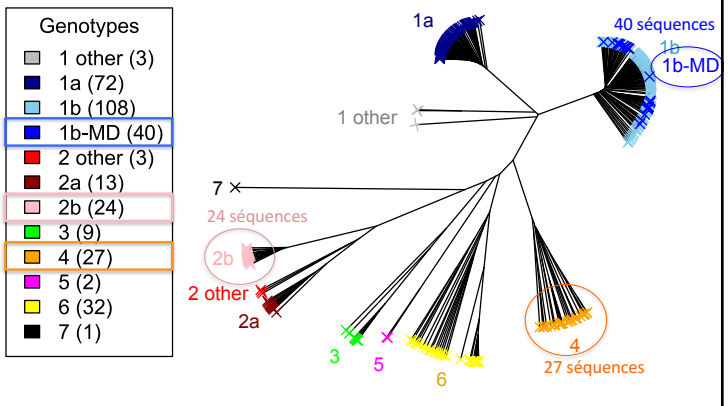


(Dib & Carbone, 2012)

Inférence du réseau d'interaction protéine-protéine du Virus de l' Hépatite C (VHC)

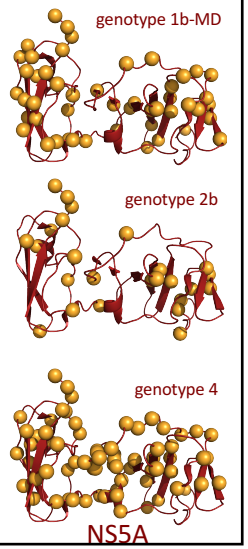




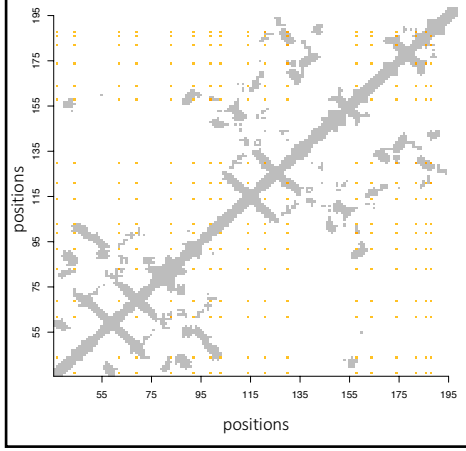


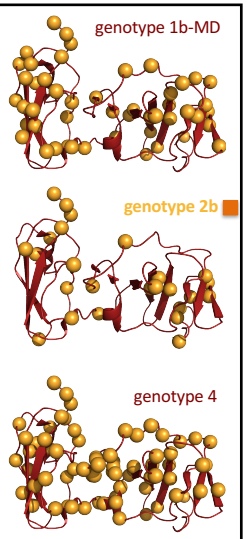
Les génotypes ont plus que 85% d'identité de séquence

L'éspace des positions non-conservées

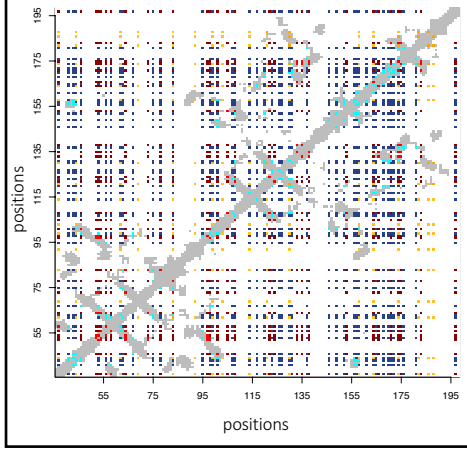


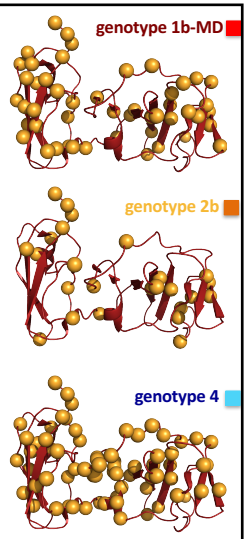
L'éspace des positions non-conservées

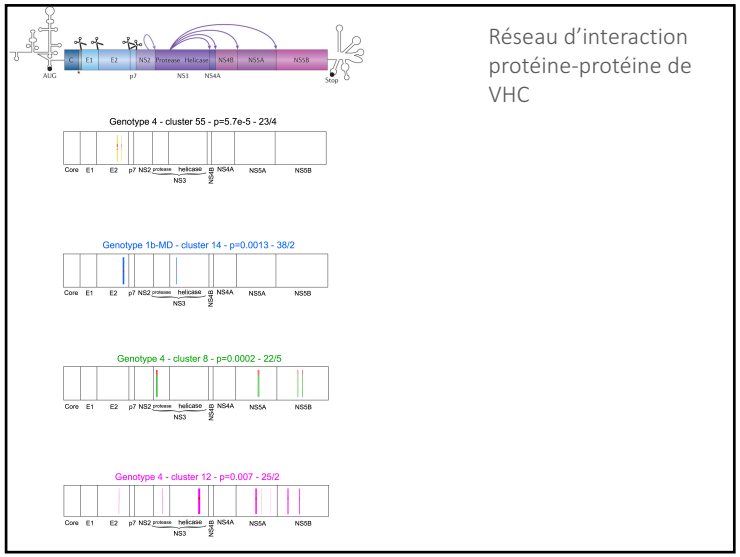




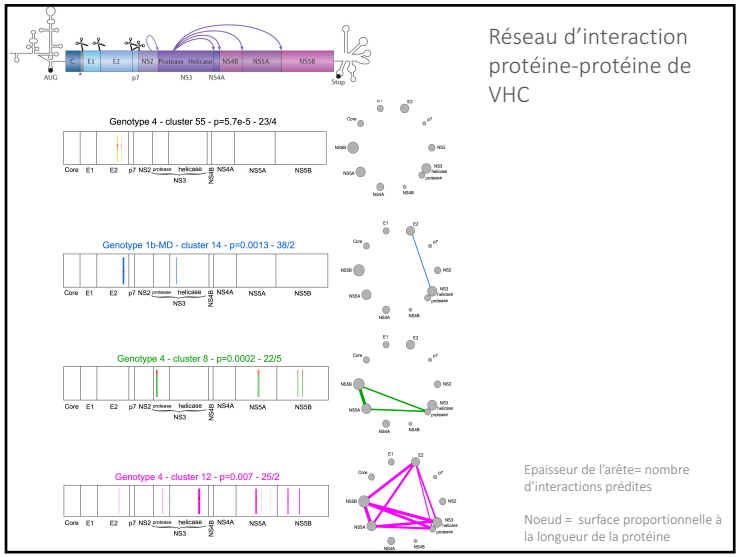
L'éspace des positions non-conservées







Réseau d'interaction protéine-protéine de VHC

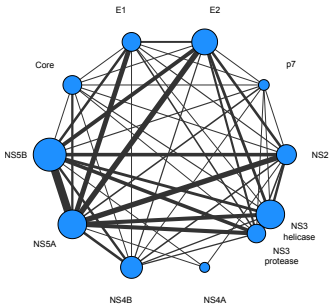


Réseau d'interaction protéine-protéine de VHC

Epaisseur de l'arête= nombre d'interactions prédites

Noeud = surface proportionnelle à la longueur de la protéine

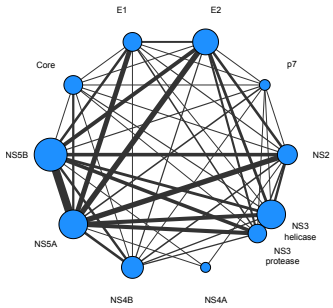
Réseau d'interaction protéine-protéine du VHC



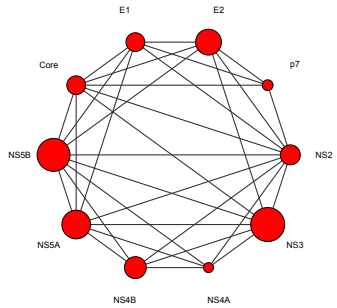
Analyse de co-évolution

(Champeimont et al 2015)

Réseau d'intéraction protéine-protéine du VHC



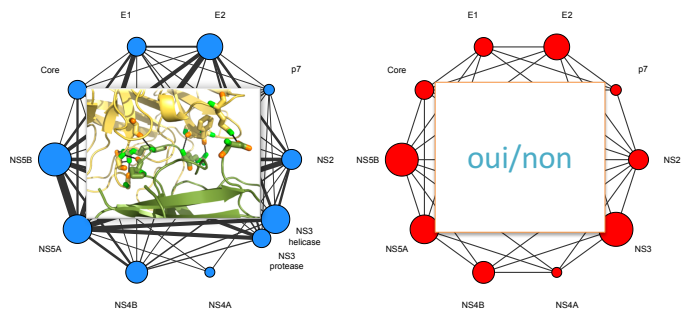
Analyse de co-évolution



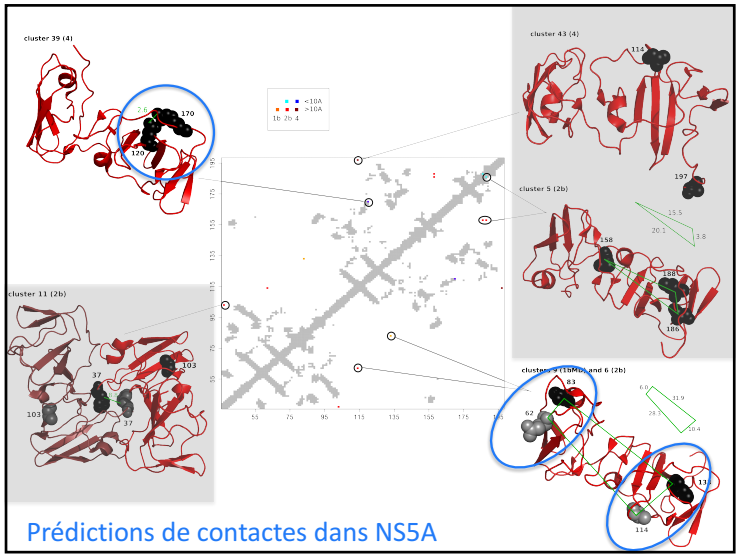
Experimentalement connu (FACS and FRET).

[N.Hagen, K.Bayer, K.Roesch, M.Schindler, The intra viral protein interaction network of hepatitis C virus, MCP, 2014]

(Champeimont et al 2015)

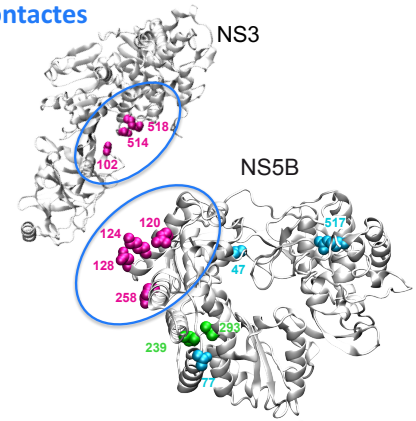


(Champeimont et al 2015)



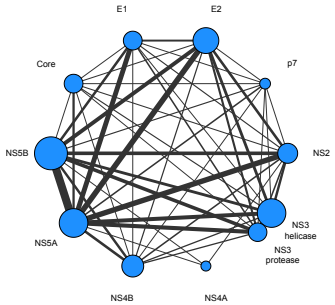
Prédictions de contactes dans NS5A

Prédictions de contactes entre protéines



Prédictions qui ont besoin de validations expérimentales

Le futur



Application systématique en biologie moléculaire des virus:

Hepatitis B Virus
Measles virus
Nipah virus
Crimean-Congo Hemorragic Fever Virus

Hepatitis C Virus (Champeimont et al 2016, Douam et al 2018)
Dengue Virus (Douam et al 2018)

Réseau d'interaction VHC

Francois Penin, IBCP, Université de Lyon

Interactions E1-E2

F.Douam CIRI, ENS de Lyon
D.Lavillette CIRI, ENS de Lyon
F-L. Cosset CIRI, ENS de Lyon

Co-évolution

F.Oteri F.Nadalin L.Dib E.Laine R.Champeimont S-w. Hu