



HAL
open science

Nous descendons d'un ancêtre polyploïde

Hervé Tostivint

► **To cite this version:**

Hervé Tostivint. Nous descendons d'un ancêtre polyploïde. Les amis du Muséum National d'Histoire Naturelle, 2017, 269, pp.5-7. hal-04012740

HAL Id: hal-04012740

<https://hal.science/hal-04012740>

Submitted on 3 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nous descendons d'un ancêtre polyploïde

(vraisemblablement octoploïde)

Hervé Tostivint

Professeur du Muséum National d'Histoire Naturelle

Evolution des Régulations Endocriniennes

CNRS UMR 7221, MNHN UMS 501

htostivi@mnhn.fr

La plupart des organismes vivants connus aujourd'hui sont diploïdes. Cela signifie qu'ils sont constitués de cellules dont les chromosomes sont en deux exemplaires (soit $2n$ chromosomes au total¹). Chez certaines espèces toutefois, surtout des plantes mais aussi quelques animaux, ce nombre d'exemplaires peut être beaucoup plus important: quatre (Fig. 1A), six, ou huit et même parfois davantage. Elles sont alors qualifiées de polyploïdes.

Au cours de ces dernières années, les chercheurs ont rassemblé toute une série d'indices suggérant que nous descendons d'un ancêtre octoploïde (c'est-à-dire dont les cellules possédaient $8n$ chromosomes). Comme nous le verrons, cet ancêtre n'est pas uniquement le nôtre, à nous humains, il est en fait celui de tous les vertébrés. Pour aller à sa rencontre, comprendre comment il a pu apparaître, mais aussi découvrir ce que nous en avons hérité, il nous faut nous plonger dans un passé vieux de plus de 500 millions d'années (Ma) (Fig. 1B).

Comment les espèces polyploïdes prennent-elles naissance?

Le cas le plus commun de polyploïdie est la tétraploïdie. On appelle tétraploïdisation le passage d'un état diploïde à un état tétraploïde, c'est-à-dire le doublement complet d'un génome. On pense que les espèces octoploïdes se forment lorsque se produisent successivement deux événements de tétraploïdisation.

La tétraploïdisation peut survenir par deux mécanismes principaux (Fig. 2A). Dans le cas de l'auto-tétraploïdisation, il y a doublement du stock chromosomique d'une même espèce, soit par méiose anormale produisant des gamètes diploïdes, soit par mitose anormale (appelée endomitose) affectant le zygote. Dans le cas de l'allo-tétraploïdisation, il y a association des stocks chromosomiques provenant de deux espèces différentes (mais en général très proches sur le plan évolutif) suivie d'un doublement par endomitose de chacun de ces stocks.

Comment détecte-t-on des événements anciens de polyploïdisation?

Prouver que le génome d'une espèce a été entièrement dupliqué par tétraploïdisation est relativement aisé si cette duplication s'est produite récemment car, dans ce cas, son résultat est directement visible au niveau du caryotype : tous les chromosomes des cellules sont en quatre exemplaires au lieu de deux (Fig. 1A). En revanche, cette mise en évidence est beaucoup plus difficile si l'évènement est plus ancien. En effet, sur le long terme, les chromosomes ne sont pas les objets stables que l'on croit, et inéluctablement, ils vont subir des remaniements tels que des cassures, des fusions, des duplications de certaines portions, des délétions d'autres (Fig. 2B). La conséquence est que, non seulement le nombre global des chromosomes est modifié, mais surtout la tétraploïdie est perdue. En effet, tous ces réarrangements survenant de manière aléatoire, il n'y a quasiment aucune chance qu'ils affectent de la même manière chacune des paires de chromosomes nouvellement dupliqués. Au final, généralement en seulement quelques Ma, ces chromosomes voient leur structure diverger et un retour à l'état diploïde initial s'opère

¹ Leurs gamètes sont les seules cellules à ne pas être diploïdes mais haploïdes, c'est à dire avec un seul exemplaire de chaque chromosome.

inévitablement (on parle alors de rediploïdisation secondaire). Dans ces conditions, les indices exploitables pour démontrer la survenue d'un évènement de tétraploïdisation ancien sont forcément beaucoup moins spectaculaires.

Une des conséquences attendues d'un évènement d'une tétraploïdisation est un doublement du nombre de tous les gènes. Parmi les nouveaux gènes ainsi apparus, nombreux sont appelés à disparaître, car ils accumulent rapidement des mutations délétères. Certains peuvent néanmoins être préservés en particulier lorsqu'ils acquièrent de nouvelles fonctions car si celles-ci sont avantageuses, elles peuvent être conservées par la sélection naturelle.

Sur quels arguments fait-on reposer l'existence de cet ancêtre polyploïde ?

Si nous dérivons d'un ancêtre qui est passé par un état octoploïde, il est clair, pour les raisons évoquées ci-dessus, que cet état n'a pas pu être conservé très longtemps. On peut d'ailleurs penser que cet ancêtre n'a été réellement octoploïde que si les deux évènements de tétraploïdisation qu'il a subis sont survenus très rapidement l'un à la suite de l'autre, car dans le cas contraire, il aurait forcément rebasculé dans un état diploïde avant la seconde duplication. En revanche, puisque cet ancêtre a possédé beaucoup plus de gènes que son ancêtre diploïde immédiat (quatre fois plus en théorie), on peut s'attendre à ce que certains de ces gènes surnuméraires ont été préservés jusqu'à aujourd'hui et qu'ils sont donc identifiables dans notre propre génome. De fait, à partir des années 1990, diverses études ont montré qu'un certain nombre de gènes ou groupes de gènes étaient présents en quatre exemplaires dans notre génome (et plus généralement celui des vertébrés) contre un seul chez la plupart des autres espèces animales (Fig. 3A). Pour certains chercheurs, ces gènes ont été rapidement considérés comme des témoins d'un état passé d'octoploïdie.

Parmi ces gènes figurent ceux d'une famille très étudiée, la famille des gènes homéotiques (également appelés gènes Hox) (Fig. 3B). Les gènes Hox sont des gènes dont l'expression est impliquée dans la polarisation du corps et des appendices des animaux bilatériens. Tous ces gènes présentent en commun un domaine très conservé de 180 nucléotides appelé homéoboîte qui signe leur origine commune. Les gènes Hox sont généralement organisés en complexes, chaque complexe regroupant un nombre variable de gènes disposés en tandem. La plupart des espèces animales examinées possèdent un complexe Hox unique alors que les vertébrés en ont généralement quatre. Les analyses phylogénétiques suggèrent que ces quatre complexes proviennent d'un complexe originel unique qui se serait dupliqué « en bloc », à deux reprises successivement.

De manière intéressante, beaucoup d'autres blocs de ce type, c'est-à-dire présents en quatre exemplaires à travers le génome et que l'on qualifie de tétra-paralogons, ont également été mis en évidence (Fig. 3A). En 2005, une recherche exhaustive portant sur l'ensemble du génome humain, alors tout juste séquencé, en a recensé plus de 450, dispersés sur l'ensemble des chromosomes. En vertu du principe de parcimonie et sur la base d'analyses réalisées sur le génome de plusieurs autres espèces, en particulier des espèces de cordés appartenant à des groupes très proches mais distincts de celui des vertébrés, comme la cione (Fig. 1B), il a été progressivement proposé que ces différents blocs de gènes ont été tous engendrés simultanément au tout début de l'histoire des vertébrés à la faveur de deux évènements de tétraploïdisation. Ce scénario, qui constitue l'hypothèse dite des « 2R » (pour « *two rounds of whole genome duplication* ») a suscité une assez forte opposition pendant plusieurs années mais de nouvelles analyses publiées en 2008 à partir de données comparant le génome humain à celui de l'amphioxus (Fig. 1B) ont finalement conduit à son adoption quasi unanime par la communauté scientifique.

Qu'avons-nous hérité de notre ancêtre octoploïde ?

On estime que 20% seulement des nouveaux gènes engendrés par les 2R ont été préservés et se retrouvent encore aujourd'hui dans notre génome. Bien connaître ces gènes est important car on peut légitimement penser qu'ils sont à l'origine d'innovations qui ont contribué de manière significative au succès évolutif des vertébrés.

Les 2R sont reconnues pour être responsables de l'expansion d'un grand nombre de familles de gènes chez les vertébrés. Il n'est évidemment pas question ici d'en dresser un inventaire complet mais simplement d'en donner quelques exemples représentatifs.

Le premier exemple concerne la famille des globines (Fig. 3B). Il s'agit d'une très vaste et très ancienne famille, présente à la fois chez les bactéries, les plantes et les animaux. Chez les vertébrés, les globines les plus connues sont l'hémoglobine et la myoglobine, deux protéines spécialisées respectivement dans le transport de l'O₂ par les globules rouges et son stockage au niveau des muscles. Une troisième globine a été également identifiée. Appelée cytoglobine, en raison de son expression ubiquitaire, ses fonctions sont actuellement mal comprises. Des travaux récents ont établis que les trois régions chromosomiques portant les gènes de la myoglobine, de la cytoglobine et de l'hémoglobine appartiennent à l'un des 450 tétra-paralogons évoqués précédemment et sont par conséquent directement issues des 2R. Dans le même temps, ils suggèrent que le gène d'une quatrième globine a aussi été engendré à cette occasion mais a été rapidement perdu par la suite. L'ensemble de ces travaux montre donc que les fonctions de stockage et de transport de l'O₂ assurées chez les vertébrés par les hémoglobines et la myoglobine, respectivement, peuvent donc être considérées comme des innovations évolutives apparues grâce aux 2R. Notons toutefois qu'elles ne sont pas propres à ce groupe car elles ont été également développées par convergence dans d'autres lignées, comme celles des annélides ou de certains insectes.

Le deuxième exemple concerne les opsines, une famille également très ancienne représentée dans l'ensemble du monde vivant (Fig. 3B). Chez les vertébrés, on distingue deux grands types d'opsines, toutes exprimées dans les cellules visuelles de la rétine: la rhodopsine responsable de la vision nocturne et les opsines impliquées dans de la vision des couleurs. Ces dernières se distinguent essentiellement sur la base des longueurs d'onde auxquelles elles sont sensibles et leur nombre peut varier de façon assez importante d'une espèce à l'autre: deux chez la plupart des mammifères, mais trois chez l'homme et jusqu'à quatre chez les poissons téléostéens et les oiseaux. Une étude récente basée sur la localisation chromosomique de tous ces gènes a montré que l'ancêtre des vertébrés possédait probablement deux gènes d'opsines ancestrales (notés SW et LW) et que lorsque les 2R sont survenues, ce nombre s'est passé à huit. Par la suite, des pertes de gènes sont intervenues, en nombre variable selon les lignées, aboutissant à la situation actuelle. Dans cet exemple, on voit donc comment les 2R ont décuplé les capacités de vision des couleurs, même si, toutes les solutions n'ont finalement pas été toujours préservées.

Conclusion

Tous ces exemples démontrent que les 2R ont considérablement enrichi le répertoire de nos gènes. En réalité très peu de familles de gènes aujourd'hui présentes dans notre génome y ont échappé. Certes, beaucoup de travail reste encore à accomplir pour en saisir toutes les conséquences fonctionnelles mais les 2R nous fournissent d'ores et déjà une grille de lecture nouvelle pour comprendre la complexité de notre génome. A l'instar des roches qui ont enregistré dans leur composition et leur structure les événements subis par la Terre au cours de son histoire, notre génome apparaît donc aussi comme un miroir de notre passé

Bibliographie

- Lagman D et al. (2013) The vertebrate ancestral repertoire of visual opsins, transducin alpha subunits and oxytocin/vasopressin receptors was established by duplication of their shared genomic region in the two rounds of early vertebrate genome duplications. *BMC Evolutionary Biology* 13: 238.
- Tostivint H. (2010) Familles multigéniques et dynamique des génomes. *Biologie-Géologie* 3 : 145-166.
- Tostivint H. (2016) Avancées récentes sur l'évolution des gènes de la famille des globines chez les vertébrés. *Biologie-Géologie* 2 : 127-144
- Van de Peer Y et al. (2009) The evolutionary significance of ancient genome duplications. *Nature Reviews of Genetics* 10: 725-732.

Légendes des figures

Figure 1.

- A. Caryotype de deux espèces parentes, l'une diploïde et l'autre tétraploïde. Chez l'espèce diploïde, les chromosomes peuvent être regroupés par paires. Chez l'espèce tétraploïde, ils sont chacun en quatre exemplaires. Le nombre total de chromosomes est donc doublé. n désigne le nombre de chromosomes d'un jeu unique (dans cet exemple, $n=18$).
- B. Arbre phylogénétique simplifié des animaux à symétrie bilatérale (bilatériens). Les deux flèches symbolisent les deux événements de tétraploïdisation (1R et 2R) survenus chez le dernier ancêtre commun des vertébrés (représenté par un point rouge), entre -500 à -550 Ma

Figure 2.

- A. Détail des mécanismes pouvant conduire à un état tétraploïde. Dans chaque cas, la tétraploïdisation est la conséquence d'un accident touchant soit la méiose à l'origine des gamètes, soit de la première mitose de l'œuf fécondé.
- B. Mécanisme de rediploïdisation d'un génome tétraploïde. Par souci de simplification, le génome initial est représenté sous sa forme haploïde, c'est-à-dire avec un seul jeu de chromosomes (ici $n=6$). Noter qu'après rediploïdisation, en raison du caractère aléatoire des remaniements chromosomiques, le nombre de chromosomes présent dans le nouveau jeu peut être modifié par rapport à ce qu'il était initialement (ici $n'=8$).

Figure 3.

- A. Trois exemples de tétra-paralogons parmi les 450 environ que compte le génome humain, interprétés comme le résultat de deux événements successifs de tétraploïdisation subis par le dernier ancêtre des vertébrés (hypothèse des 2R; voir explications dans le texte).
- B. Détail de la structure des trois tétra-paralogons représentés en A (en utilisant le même code couleur). Les nombres correspondent aux numéros des chromosomes chez l'homme. Chaque tétra-paralogon est considéré comme provenant d'un segment ancestral unique qui s'est ensuite dupliqué en bloc deux fois successivement. L'histoire des trois familles de gènes portées par ces tétra-paralogons, familles des gènes Hox, des globines et des opsines, est détaillée dans le texte. Les gènes représentés en pointillés sont des gènes perdus. Les gènes représentés par un petit rectangle de couleur pâle sont des gènes voisins des gènes des familles de la globine et de l'opsine mais sans parenté directe avec eux. Cgb, cytoglobine; Mb, myoglobine; Hb, hémoglobine; Gb, globine ancestrale. *NB. Le locus de l'hémoglobine contient plusieurs gènes engendrés à la suite de duplications locales récentes. Chez les mammifères et les oiseaux, certains d'entre eux ont été secondairement transloqués sur un autre chromosome (11 chez l'homme). Rh, rhodospine; SW, opsine sensible aux ondes courtes; LW, opsine sensible aux ondes longues. *NB. Les gènes Rh2 et SW2 sont toujours présents chez les oiseaux et les pois-

sons télésootéens. Chez les primates du Nouveau-Monde, le gène LW s'est dupliqué en tandem pour former un gène non représenté ici, appelé MW car sensible aux longueurs d'ondes moyennes.

A



B

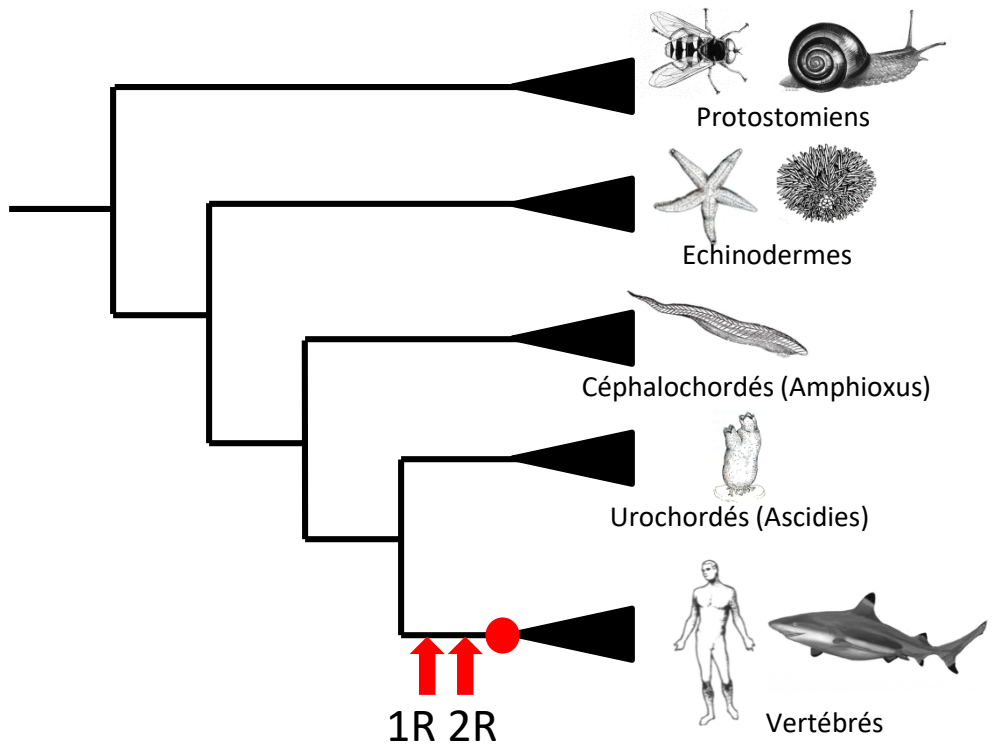
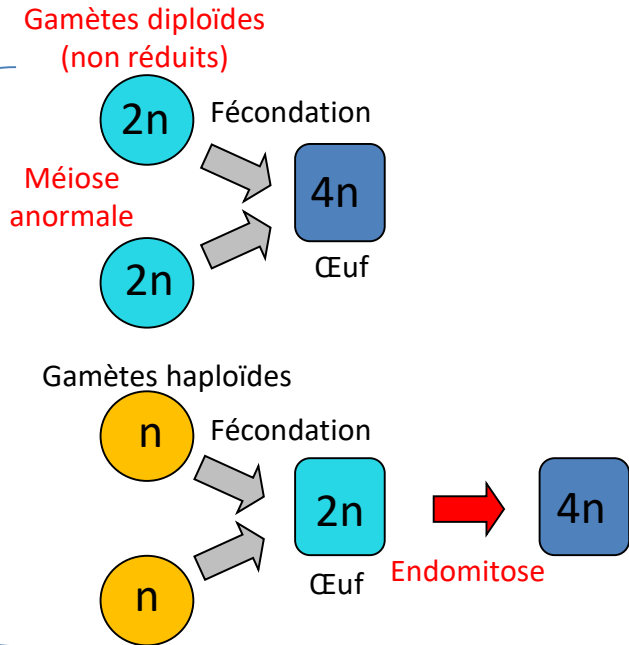
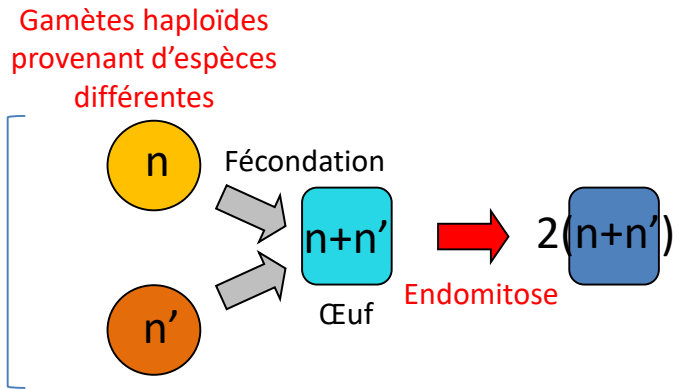


Figure 1

Auto-tétraploïdisation



Allo-tétraploïdisation



B

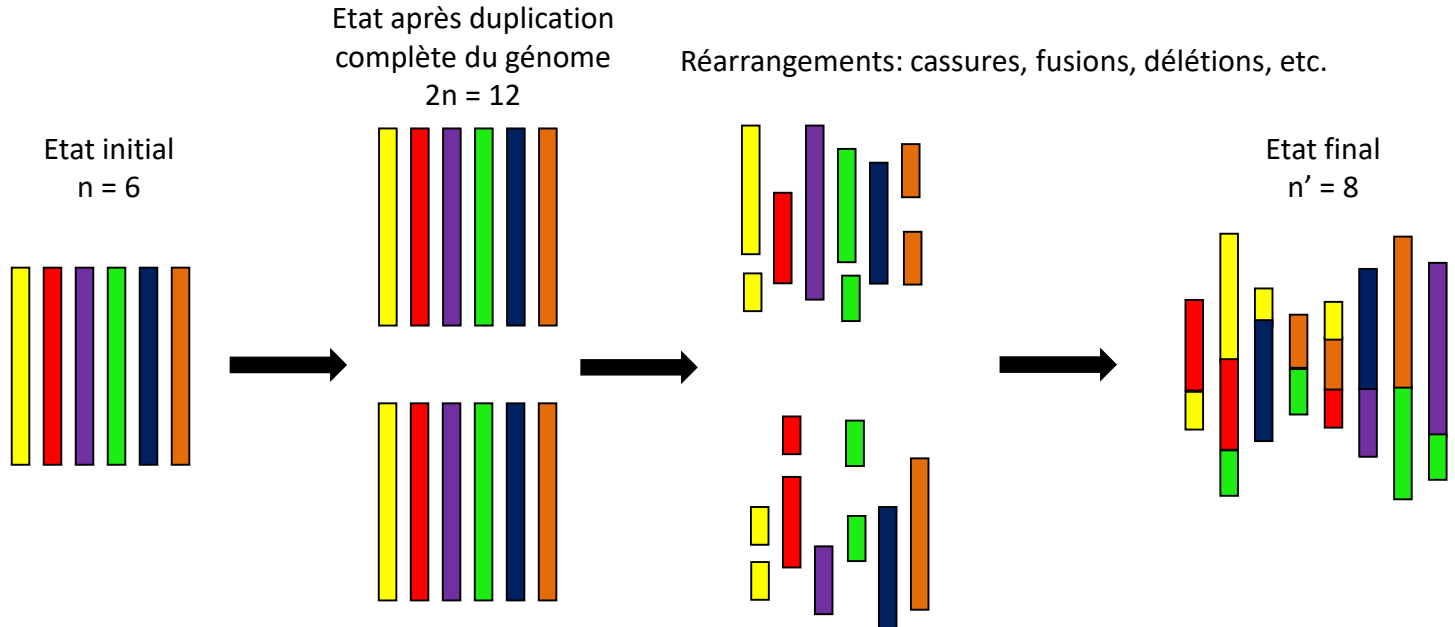
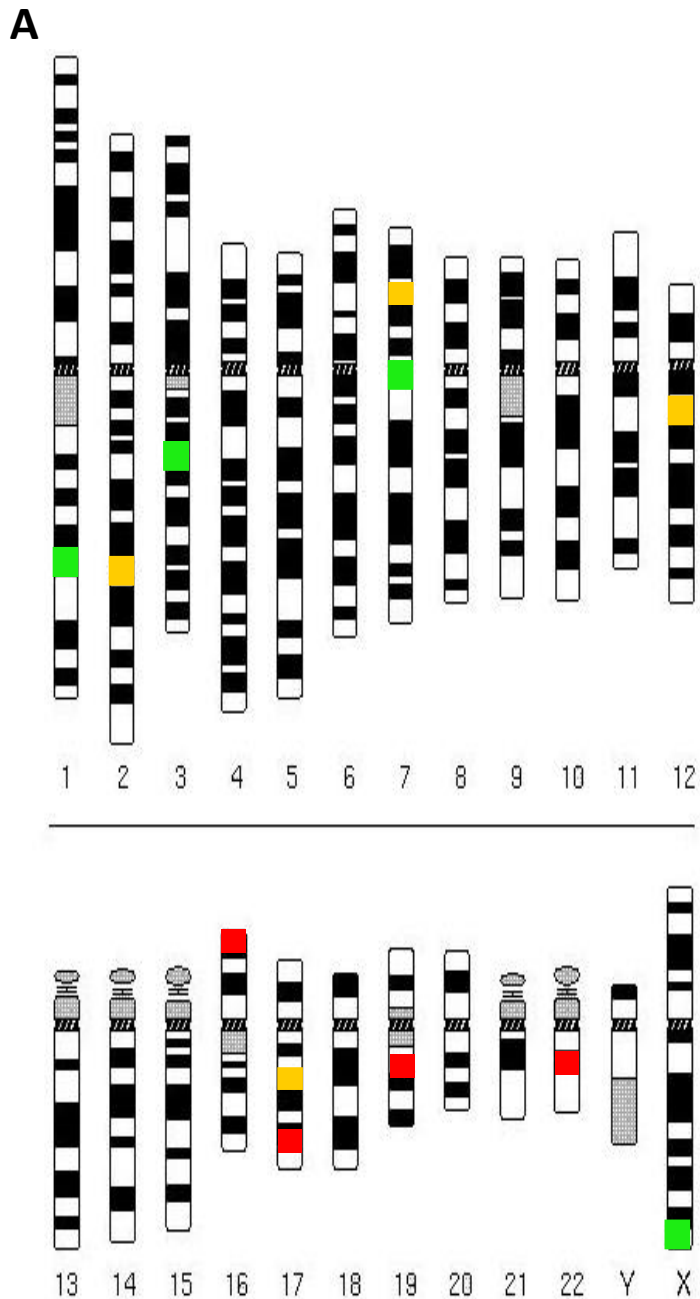
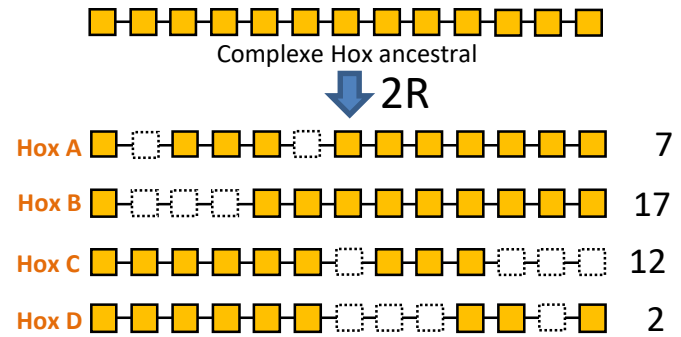


Figure 2

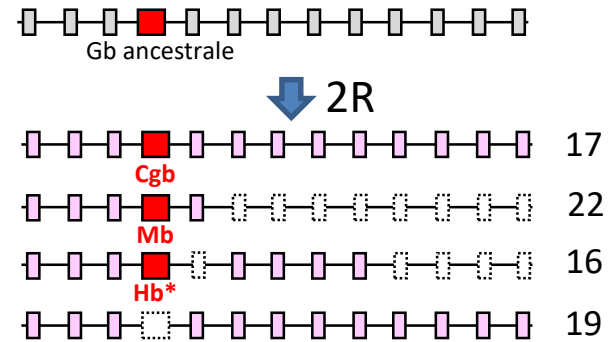


B

Famille des gènes Hox



Famille des globines



Famille des opsines

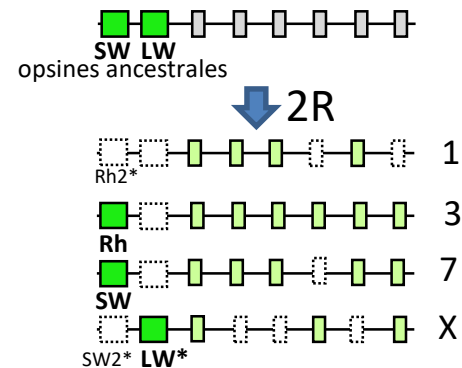


Figure 3