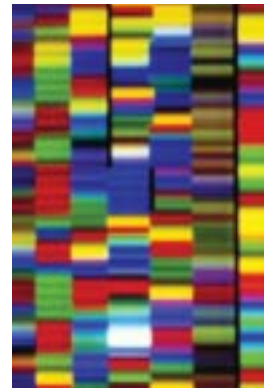


> La cartographie comparée est une analyse comparative de l'organisation chromosomique des génomes, qui repose sur l'exploitation de la biodiversité dans le double objectif de comprendre le fonctionnement des structures biologiques et d'en inférer l'évolution. La popularité actuelle de cette approche relève cependant d'une autre démarche, plus technique, qui résulte de la structure très particulière des connaissances actuelles sur les génomes. Le coût expérimental de l'analyse d'un génome restant très élevé, les informations disponibles sont le plus souvent concentrées sur quelques organismes « modèles » (principalement l'homme et la souris chez les mammifères). Déduire des informations sur des organismes d'intérêt, à partir de l'étude d'organismes modèles, est donc une démarche très précieuse, et la cartographie comparée entre dans ce cadre. Il faut également souligner que les comparaisons entre génomes peuvent être plus complexes et profiter des possibilités expérimentales offertes par certaines espèces. Ainsi, bien que l'homme soit l'un des organismes les mieux connus, la cartographie comparée est largement utilisée pour effectuer des allers-retours entre l'information génomique humaine et celle d'autres organismes dans lesquels l'observation, voire l'expérimentation génétique, est plus accessible. <

Bio-informatique (4)

Cartographie génomique comparée chez les mammifères

Gisèle Bronner, Bruno Spataro, Christian Gautier



Laboratoire de Biométrie et de Biologie Évolutive, Université Claude Bernard Lyon I, 43, boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex, France.

Les différentes cartes génomiques

1. Les cartes génétiques reposent sur l'existence, au cours de la méiose, de recombinaisons qui aboutissent à des échanges réciproques de matériel génétique (*crossing-over*) entre les chromosomes homologues. Lorsque deux marqueurs sont suffisamment proches, la probabilité d'une recombinaison entre eux se comporte comme une distance au sens mathématique et permet ainsi la construction de cartes des chromosomes. Ces cartes génétiques résultent donc de la prise en compte d'un mécanisme biologique, ce qui n'est pas réellement le cas pour les cartes des deux groupes suivants.
2. De nombreuses cartes sont construites à partir de l'appartenance de marqueurs à un même segment génomique. Différentes approches expérimentales permettent de découper les génomes en segments, soit en les colorant (bandes chromosomiques), soit en les fragmentant physiquement (digestion de séquences d'ADN par des enzymes de restriction, fragmentation par irradiation). Des techniques expérimentales reposant sur l'hybridation entre acides nucléiques permettent de repérer l'appartenance de marqueurs à un même segment. Dans le cas de la fragmentation du génome, les

La cartographie génomique peut être définie comme l'identification de la position de tout élément du génome pouvant être caractérisé de façon non ambiguë. Ces éléments, appelés marqueurs, sont très divers. Il peut s'agir de simples fragments d'ADN sans caractérisation fonctionnelle (segments anonymes), d'éléments répétés (microsatellites en particulier), ou de gènes. La nature du positionnement des marqueurs et la définition de la distance qui les sépare ont conduit à l'existence de plusieurs types de cartes que l'on peut schématiquement présenter en trois groupes.



segments obtenus sont le plus souvent chevauchants, ce qui permet d'affiner le positionnement des marqueurs. Enfin, les cartes qui utilisent la coloration en bandes des chromosomes occupent une place particulière, d'une part parce qu'il s'agit des plus anciennes et d'autre part, parce qu'elles mettent en jeu une propriété intrinsèque des chromosomes, même si la signification biologique de cette propriété est loin d'être claire.

- Enfin, la séquence complète du génome constitue en elle-même une carte particulière. C'est la plus fine de toutes les cartes.

L'ensemble des cartes qui utilisent, directement (séquence) ou indirectement (banques génomiques*), une distance correspondant au nombre de nucléotides séparant deux marqueurs constitue ce qu'on appelle les cartes physiques. Les cartes d'hybrides d'irradiation constituent un troisième groupe de cartes. Issues de la fragmentation par irradiation de séquences génomiques, ces cartes sont calculées sur la base de modèles de liaison entre marqueurs, semblables aux modèles des cartes génétiques. Contrairement à ces dernières, les cartes d'hybrides d'irradiation ne reflètent cependant aucune propriété biologique associée au génome. Les différentes cartes, génétiques, physiques et les cartes d'hybrides d'irradiation ne sont pas interchangeables car elles reposent sur des notions de distances différentes. En outre, les relations entre cartes physiques et génétiques sont difficiles à établir notamment en raison de la nature différente des marqueurs utilisés. Les cartes physiques contiennent, en effet, une majorité de marqueurs uniques et conservés, par exemple des gènes. En revanche, l'établissement des cartes génétiques nécessite l'utilisation de marqueurs hautement polymorphes comme les microsatellites. Les cartes de Marey (Figure 1) qui permettent une représentation simultanée des

marqueurs communs à deux cartes d'un même chromosome, comme par exemple des gènes à la fois polymorphes et conservés, montrent que distances physique et génétique ne sont pas proportionnelles, mettant ainsi en lumière l'existence de régions génomiques dont les taux de recombinaison sont plus ou moins élevés. Cette observation conduit à prendre en compte une variabilité du « paysage génomique » suivant des régions de tailles diverses. Les plus connues de ces régions sont les isochores (voir glossaire) caractérisés par la variabilité de critères divers : fréquences des bases, structure des gènes, densité en gènes, nature des éléments répétés [1, 2].

Comparaison entre cartes génomiques de différentes espèces

Afin de réaliser une analyse comparative, il faut à la fois qu'une diversité existe dans les structures étudiées et que cette diversité n'excède pas un seuil au-delà duquel les structures biologiques n'ont plus de point commun. Le

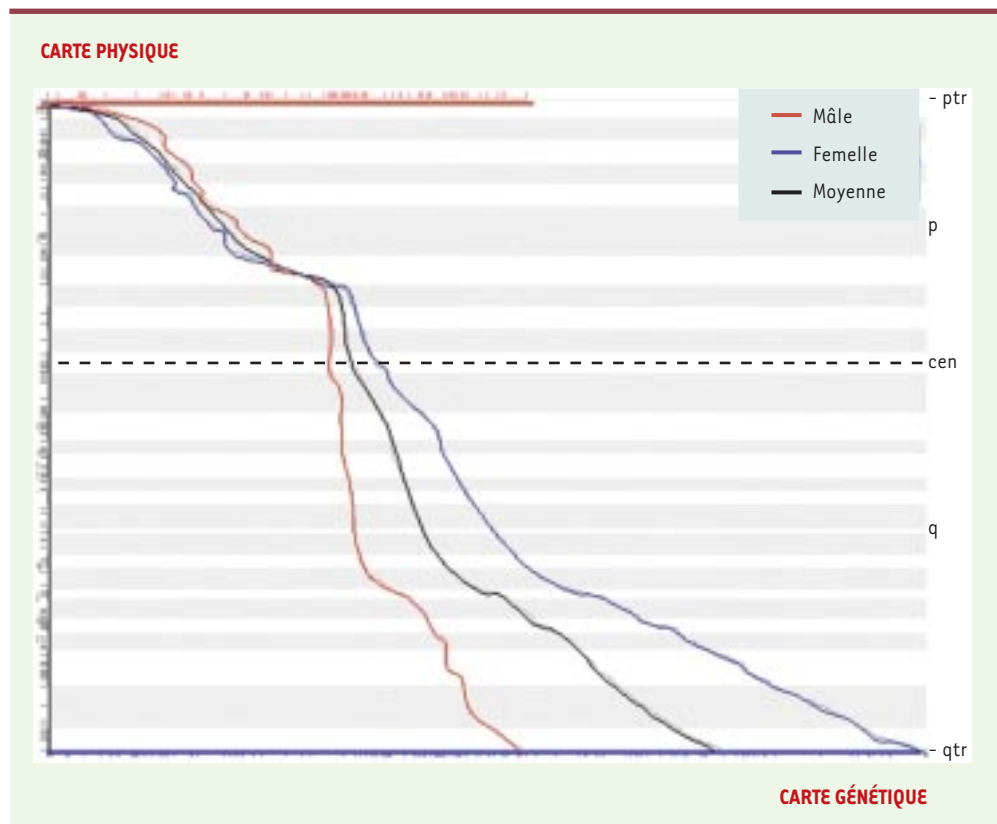


Figure 1. Carte de Marey. La représentation dans un espace à deux dimensions des cartes génétiques (abscisse) et physique (ordonnée), permet de comparer ces deux types de cartes. Les gènes cartographiés communs aux deux cartes sont reportés dans le plan, leurs coordonnées correspondant à leur position sur les cartes génétique et physique respectivement. Les courbes liant ces points représentent la relation entre la distance physique et génétique mâle (courbe rouge) et femelle (courbe bleue). Ces courbes, très différentes d'une droite, illustrent clairement que les distances physiques et génétiques ne sont pas proportionnelles.

* L'ADN génomique est fragmenté et inséré dans des vecteurs pour former une banque génomique, image fragmentée de la séquence du génome.



mécanisme évolutif clé de la cartographie comparée est le réarrangement chromosomique, qui consiste dans le déplacement de segments d'ADN, soit de manière locale (modification de l'ordre des gènes), soit à l'échelle des chromosomes eux-mêmes (translocations). La première étude évolutive systématique de ces modifications a fait l'objet d'un numéro spécial de *Mammalia* en 1986 [3]. Fondée sur la comparaison des chromosomes et l'estimation, par comparaison de leur coloration en bandes, des événements mettant en jeu un réarrangement, cette étude a clairement montré la pertinence phylogénétique de la cartographie comparée. Ce travail pionnier a été suivi de nombreuses études sur l'évolution de l'organisation des

génomés, études qui ont intégré d'autres types de données, en particulier l'ordre des gènes le long des chromosomes (voir en particulier [4, 5]).

La comparaison des « paysages génomiques » est également apparue fructueuse. Malgré les remaniements importants subis par les génomes au cours de l'évolution, le contexte génomique des gènes reste largement conservé au sein des mammifères. La classe d'isochore d'un gène est ainsi, sauf rare exception, conservée au sein des mammifères et, au-delà, chez la plupart des vertébrés. De plus, l'établissement de liens entre les cartes de deux organismes permet l'étude de la régionalisation de paramètres évolutifs. Ainsi, on peut montrer que des gènes voisins dans les cartes de l'homme et de la souris présentent des vitesses de substitutions

silencieuses (glossaire) proches : schématiquement, la vitesse d'évolution d'un gène dépend de sa position dans le génome.

De manière encore plus globale, on peut comparer la longueur totale des génomes. Si la longueur physique des génomes de mammifères présente peu de différences, il n'en est pas de même de leur longueur génétique. Par exemple, la longueur physique des génomes de l'homme et de la souris est de l'ordre de 3×10^9 pb, alors que leurs longueurs génétiques sont d'environ 3 700 et 1 600 cM respectivement. L'analyse de ces différences de longueurs génétiques a conduit à deux hypothèses très différentes. La première repose sur l'observation que le facteur le mieux corrélé à la longueur génétique d'un génome est le nombre total de bras chromosomiques ($R^2 = 0,79$, voir [6] pour une revue). Cela plaide en faveur de la nécessité d'au moins un chiasma par bras chromosomique - et non par chromosome - comme il est classique-

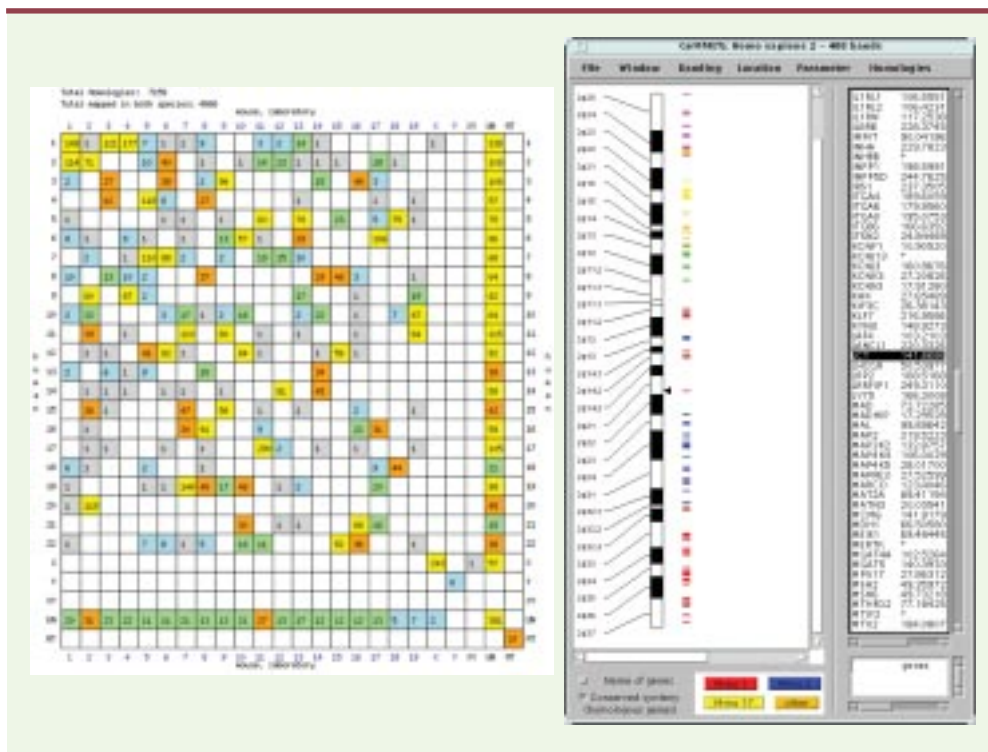


Figure 2. Comparaison homme/souris : grille d'Oxford et analyses des données. La partie gauche de la figure représente une « grille d'Oxford », montrant les similitudes globales entre les chromosomes de chacune des deux espèces. Les niveaux de couleur indiquent le nombre de marqueurs communs aux chromosomes humains et murins considérés : gris : 1 ; bleu : 2-10 ; vert : 11-25 ; orange : 26-50 ; jaune : > 50. Un signifie qu'il n'y a pas d'assignation chromosomique pour le marqueur dans l'espèce considérée. MT signifie que le marqueur est localisé dans le génome mitochondrial. La partie droite de la figure est obtenue par interrogation du système GEM (*genomic mapping*) dédié à la cartographie comparée des génomes de mammifères et développé au laboratoire. Elle représente la carte des gènes du chromosome 2 humain déduite de la séquence du génome. Les coordonnées des gènes (en Mégabase) apparaissent dans le cadre de droite. Le panneau de gauche représente la carte cytogénétique du chromosome 2 humain. Les gènes possédant un orthologue connu chez la souris sont indiqués au niveau de la carte cytogénétique par des tirets dont la couleur illustre la localisation chromosomique du gène chez la souris (rouge : chromosome 1 ; bleu : chromosome 2 ; jaune : chromosome 17, orange : autre). La visualisation simultanée de la position d'un gène sur la carte cytogénétique et celle issue du séquençage du chromosome 2 de l'homme est illustrée ici par la sélection dans le cadre de droite du gène LCT, et la représentation par un triangle noir de sa position par rapport à la carte cytogénétique dans le panneau de gauche.

(→) m/s
2002, n°2,
p. 237

ment attendu - pour permettre une disjonction correcte des chromosomes lors de la méiose [6, 10]. Cependant, d'autres auteurs ont remarqué une forte corrélation de la longueur génétique du génome avec l'âge de première reproduction au sein d'une espèce. Ceci a conduit à émettre une seconde hypothèse sur l'existence d'un facteur en relation avec le brassage de l'information génétique qui serait sélectionné, permettant ainsi d'assurer une variabilité génétique suffisante des espèces lorsque celles-ci ont un temps de génération élevé. On retrouve ici un débat classique concernant l'évolution des structures génomiques qui peut résulter soit d'une sélection sur la valeur sélective des organismes, soit du fonctionnement même de la machinerie cellulaire.

Ces quelques résultats, qui concernent le fonctionnement global du génome, ne doivent pas occulter le fait que les efforts actuels portent essentiellement sur l'utilisation de la cartographie pour la localisation et l'interprétation fonctionnelle de l'information génétique.

L'approche expérimentale permettant de comparer les cartes génomiques d'espèces différentes repose largement sur la similarité entre des séquences homologues et donc sur les possibilités d'hybridations entre les séquences des deux génomes. L'approche la plus globale consiste à hybrider, après marquage fluorescent, des fragments d'un chromosome d'une espèce avec le génome complet de la seconde espèce. Ce dernier sera donc « peint » avec les zones similaires, et donc supposés homologues, du chromosome sonde. Cette technique, appelée « zoo-FISH », permet d'identifier rapidement des synténies entre espèces (→). Une des représentations classiques est

constituée par les grilles d'Oxford (Figure 2). Ce type d'approche présente cependant deux limitations importantes : la première est sa relative imprécision, due en particulier à la mauvaise définition des extrémités des segments génomiques « colorés » et la seconde, l'absence de détection des modifications locales des cartes, comme les inversions courtes.

L'identification de marqueurs orthologues (qui résultent d'un événement de spéciation, et non de duplication) entre cartes permet un ancrage plus précis. Mais une des difficultés est due à la nature même des processus évolutifs qui combinent, d'une part, l'accumulation de mutations le long des branches de l'arbre phylogénétique des espèces et, d'autre part, des événements de duplications des gènes. Une similarité de séquences entre marqueurs, en particulier des gènes, peut donc aussi être la conséquence de régions dupliquées (paralogie). La discrimination entre les orthologies (pas de duplication dans l'histoire évolutive reliant les deux gènes) et les paralogies est difficile (→). L'approche la plus efficace à ce jour consiste à reconstruire la phylogénie de la famille des gènes impliqués (par exemple en utilisant la base Hovergen <http://pbil.univ-lyon1.fr>). Lorsque ce problème est résolu, on peut alors envisager de construire des correspondances entre les cartes - éventuellement de types différents - d'espèces différentes. La Figure 2 illustre cette possibilité en présentant une correspondance entre la carte génomique humaine, déduite de la séquence du génome, et les synténies chez la souris.

Les données actuellement disponibles concernent essentiellement, d'une part l'homme et/ou la souris et d'autre

Base	Espèce	Adresse WWW	Nombre de locus cartographiés	Nombre de gènes cartographiés
GDB	<i>Homo sapiens</i>	http://www.gdb.org	415 295	10 418
MGD	<i>Mus musculus</i>	http://www.informatics.jax.org http://www.informatics.jax.org/menus/homology_menu.shtml	25 533	8 282
Ratmap	<i>Rattus sp.</i>	http://ratmap.ge.gu.se	12 605	1 872
Bovmap	<i>Bos taurus</i>	http://locus.jouy.inra.fr/cgi-bin/lgbc/mapping/common/intro2.pl?BASE=cattle	3 592	1 157
Pigmap	<i>Sus scrofa</i>	http://locus.jouy.inra.fr/cgi-bin/lgbc/mapping/common/intro2.pl?BASE=pig	2 148	668
Goatmap	<i>Capra hircus</i>	http://locus.jouy.inra.fr/cgi-bin/lgbc/mapping/common/intro2.pl?BASE=goat	621	263
ArkDB	Organismes de rente	http://www.thearkdb.org		
AGDJ	Généraliste	http://ws4.niai.affrc.go.jp/		

Tableau 1. Quelques sites proposant des données de cartographies chez les mammifères (partie supérieure) et sites « généralistes » (partie inférieure). GDB : genome data base ; MGD : mouse genome database ; AGDJ : animal genome database in Japan.



part, des mammifères soit d'importance économique (porc, bœuf, mouton), soit permettant la recherche de gènes responsables de maladies génétiques humaines. Dans ce dernier cadre se trouvent notamment les animaux dits compagnons (chien, chat, cheval) qui ont été soumis à une forte sélection conduisant à l'existence de races multiples fortement homozygotes où peuvent s'exprimer des gènes récessifs. La démarche classique, utilisée par exemple en agronomie, consiste à localiser génétiquement des gènes intervenant pour des traits quantitatifs écono-

miquement intéressants, à associer à la région génomique obtenue la région homologue chez l'homme et à utiliser les connaissances (séquences, EST) pour déterminer des gènes candidats. Un exemple typique de cette démarche est fourni par la caractérisation d'un gène (RN) impliqué dans la qualité de la viande de porc [7]. Dans le cas des animaux compagnons, la cartographie comparée peut également aider à associer aux gènes impliqués dans des maladies génétiques du chien leurs homologues humains (voir la très complète revue sur le chien dans [8] et l'article de Lyons [9]).

Bio-informatique et cartographie comparée

La bio-informatique joue un rôle central dans le domaine de la cartographie comparée. Elle y intervient essentiellement à deux niveaux : le premier concerne la représentation des connaissances et le second l'analyse des données proprement dites.

Représentation des connaissances

L'accès aux données est désormais largement facilité par l'utilisation d'Internet. Le *Tableau 1* donne quelques points d'accès ainsi que des indications sur l'état des connaissances pour chacun des génomes concernés. Ces bases peuvent être interrogées sur le réseau au travers de formulaires de requêtes par le biais d'un navigateur. Dans la plupart des cas, ces formulaires permettent de réaliser des requêtes sur des critères génomiques ou cartographiques. De telles requêtes permettent d'obtenir des listes d'éléments biologiques répondant aux critères précisés ainsi qu'un lien hypertexte donnant accès à la des-

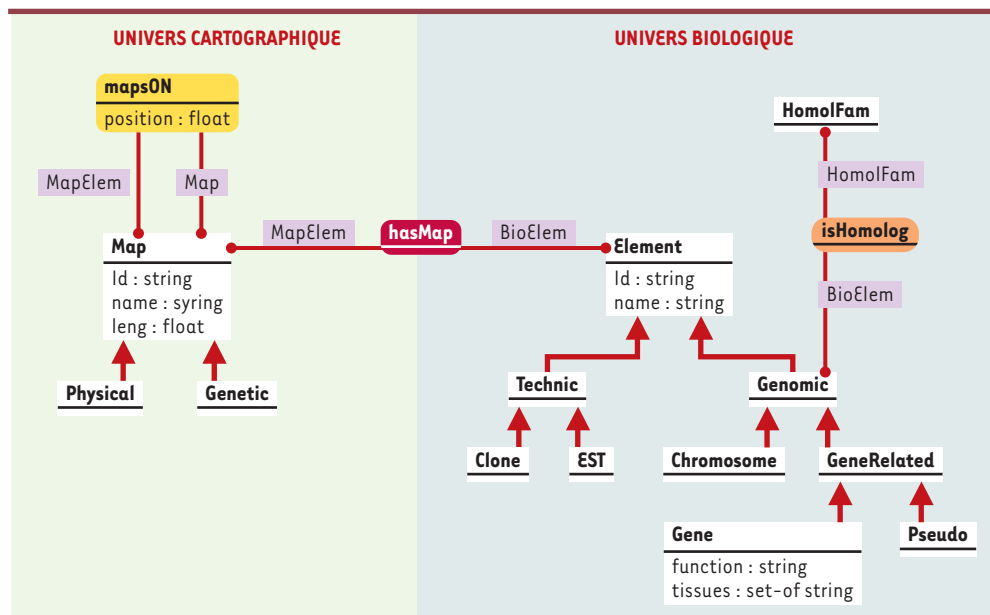


Figure 3. Modélisation des connaissances sur la cartographie comparée : modèle conceptuel. Les entités manipulées en cartographie comparée sont de deux ordres, d'une part les cartes et marqueurs cartographiques et d'autre part les entités biologiques. Ces dernières peuvent être considérées selon leurs propriétés biologiques ou cartographiques lorsqu'elles sont manipulées en tant que marqueurs cartographiques. Dans ce modèle, les classes et associations s'organisent autour des « univers » cartographique et biologique. Les entités manipulées sont décrites selon leurs propriétés biologiques à travers la classe *Element*, selon leurs aspects cartographiques à travers la classe *Map*. Pour retrouver l'intégrité de l'entité décrite (qui est éclatée dans deux classes), une association, *hasMap*, a été définie. Dans l'« univers biologique », la classe *Element* constitue la racine d'une hiérarchie de classes qui permet de définir des sous-classes de la classe *Element* dont elles héritent les propriétés, mais qu'elles spécifient également en fonction des propriétés particulières des éléments biologiques représentés. On peut ainsi distinguer les entités biologiques du génome (*Genomic*) comme les chromosomes (*Chromosome*), les gènes (*Gene*) ou les pseudogènes (*Pseudo*) et les structures issues des techniques moléculaires (*Technic*) comme les clones (*Clone*) ou les EST (*EST*). Les relations d'homologies sont décrites à travers la définition d'une classe de famille d'homologie (*HomolFam*) qui est liée aux éléments génomiques (classe *Genomic*) par l'association *isHomolog*. La relation d'homologie existant entre différents éléments génomiques est donc vue ici comme l'appartenance de ces différents éléments à une même famille de relation d'homologie. L'univers de la cartographie est décrit à travers la classe *Map*, racine d'une hiérarchie de classes dont les sous-classes (*Physical*, *Genetic*) représentent les différents types de cartes manipulées. Les entités localisées sur une carte peuvent être des marqueurs ou des cartes. Elles sont décrites à travers l'attribut *position* de l'association *mapsOn* qui lie l'élément cartographié (rôle *MapElem* de l'association *mapsOn*) à la carte qui le porte (rôle *Map* de la même association). Il est possible à travers une telle représentation de décrire la position d'un élément sur plusieurs cartes, ainsi que l'inclusion récursive de cartes localisées dans d'autres cartes.

cription détaillée des éléments individuels. L'exploration d'informations complémentaires se fait généralement par l'intermédiaire de liens hypertextes qui permettent d'atteindre les bases de données appropriées (séquences, homologies, etc.). Les données cartographiques sont disponibles sous forme textuelle, graphique ou au travers d'outils de visualisation spécifiques.

La plupart des bases de données publiques distribuent leurs données sous la forme de fichiers dits « plats », c'est-à-dire sans structure de données explicite. Ceci permet aux bio-informaticiens de constituer aisément des jeux de données nécessaires à leurs analyses. En revanche, les données sont faiblement structurées et généralement dispersées dans de nombreux fichiers.

Force est de constater qu'en pratique, les bases de données dédiées à la cartographie comparée des génomes sont rares. On peut citer la *Animal Genome Database of Japan* (Tableau 1) qui permet de naviguer au sein des cartes génomiques de différentes espèces de mammifères. De même, la base de données ArkDB pour la génomique comparative propose une intégration, au sein d'une structure commune, de données génomiques pour différentes espèces de vertébrés. La MGD consacrée à la souris permet en outre de visualiser, sur des cartes et sous forme de grilles d'Oxford, des homologies avec d'autres mammifères. Enfin, l'INRA propose également des outils de comparaison impliquant les génomes d'organismes de rente (<http://www.toulouse.inra.fr/lgc/pig/compare/compare.htm>).

Le problème particulier de la modélisation des données biologiques a fait l'objet récemment d'un article dans *médecine/sciences* (→). On peut souligner que dans le cas des génomes d'invertébrés (*D. melanogaster*, *C. elegans*) ou des plantes (*A. thaliana*), le système de ges-

tion de bases de données génomiques AceDB (<http://www.acedb.org/>), basé sur un modèle à objets, est le plus largement utilisé. En revanche, la majorité des bases de données dédiées à la cartographie génomique des vertébrés repose sur le modèle relationnel, à l'exception de la base IGD consacrée au génome humain qui utilise également AceDB (<ftp://genome.dkfzheidelberg.de/igd>). Les systèmes de gestion de bases de données qui reposent sur le modèle relationnel offrent à la fois un langage d'interrogation puissant sur des systèmes robustes et des démarches de construction des bases largement validées. Le revers de la médaille est que le modèle mathématique sous-jacent reste relativement pauvre et ne permet pas la modélisation des structures complexes. Or, les protéines, séquences, cartes et citations de la littérature forment un corps de données cohérent dont la complexité doit être explicitée au même titre que les entités du domaine

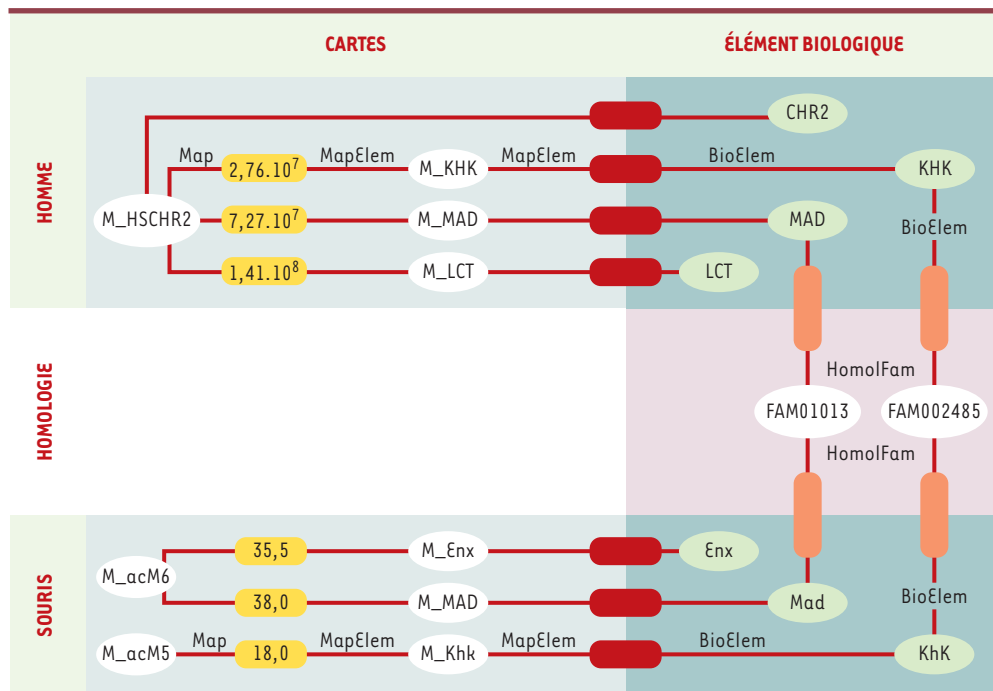


Figure 4. Modélisation des connaissances sur la cartographie comparée : diagramme d'instances. Ce diagramme d'instance illustre l'organisation au sein du système GeM (*genomic mapping*) des informations relatives à la description de la carte comparée du chromosome 2 de l'homme (Figure 2, partie de droite).

Les objets (instances de classes) sont représentés par leur nom et sous forme d'ovales. Les tuples (instances d'associations) sont représentés sous forme de rectangles à bords arrondis. Le même code couleur que celui de la Figure 3 permet d'identifier la nature des associations. M_HSCHR2 est la carte du chromosome 2 de l'homme, chromosome dont les aspects biologiques sont décrits dans l'objet CHR2. Sur cette carte, sont localisés les gènes KHK, MAD et LCT (M_KHK, M_MAD, M_LCT) qui sont décrits biologiquement dans KHK, MAD, LCT. Les gènes MAD et KHK de l'homme appartiennent respectivement aux familles d'homologie FAM001013 et FAM002485. Mad est l'homologue murin du gène MAD, il est localisé sur le chromosome 6 de la souris (M_acM6). Khk est homologue à KHK et est localisé sur le chromosome 5 de la souris (M_acM5).



elles-mêmes. Cette complexité peut en revanche être explicitée par les modèles à objet, comme le montre la *Figure 3* qui présente un fragment de la modélisation réalisée pour le système GeM (*Genomic Mapping*) développé dans notre laboratoire. Cette modélisation-objet utilise un formalisme dérivé d'UML (*Unified Modeling Language*), qui fait jouer un rôle central aux relations existant entre les entités biologiques et est implémenté en utilisant le système de représentation des connaissances AROM (Allier Relations et Objets pour Modéliser) (<http://www.inrialpes.fr/romans/pub/arom/>). La *Figure 4* illustre un exemple de l'organisation des données du système GeM suivant ce modèle.

Analyse des données

L'analyse des données de cartographies implique à la fois la mise en œuvre de méthodes statistiques classiques et des développements informatiques originaux, en particulier pour la comparaison de permutations. La mise en évidence de l'aspect régional de caractéristiques génomiques fait appel à des approches statistiques qui dépendent de la nature même des données. Si des méthodes de corrélation entre séquences voisines ont été longtemps l'approche statistique privilégiée (par exemple dans l'analyse des isochores), la disponibilité de longues séquences, voire de chromosomes entiers, ouvre la possibilité d'utiliser des méthodes qui parcourent ces longues séquences afin de rechercher de changements de structures statistiques. Actuellement les processus de Markov cachés constituent les méthodes les plus populaires du domaine. Enfin, lorsque les analyses portent sur des gènes, des méthodes d'auto-corrélation spatiale ont également été mises en œuvre [11].

Comparer la disposition de marqueurs homologues sur deux cartes revient à comparer des permutations d'objets. L'objectif est de construire, entre ces permutations, une distance qui ait un sens en tant que distance évolutive, par exemple qui puisse s'exprimer en nombre d'événements permettant d'atteindre ces deux permutations à partir d'une configuration ancestrale commune. De très nombreuses études ont été réalisées dans le cadre de petits génomes (bactéries, organelles). Une très bonne revue pourra être trouvée dans [12]. L'utilisation de ces méthodes dans le cadre de la cartographie comparée de mammifères reste néanmoins à développer, mais il convient de remarquer que les données sont encore peu nombreuses. Il est raisonnable, cependant, de penser que ces approches pourront rapidement se développer et permettront de donner un nouvel éclairage sur la variabilité, le long du génome, des différents mécanismes de réarrangement et de leur relation avec les caractéristiques fonctionnelles du génome.

Conclusions

Les cartes génomiques sont des objets biologiques complexes qui positionnent de plusieurs façon, sur une même structure spatiale, des entités biologiques très différentes (gènes, marqueurs anonymes mais aussi propriétés statistiques, évolutives ou structurales). La comparaison de tels objets entre espèces implique la manipulation de concepts très divers, et il ne paraît plus possible désormais d'échapper à une modélisation sophistiquée permettant une manipulation « intelligente » de ces données. De nombreux projets sont actuellement en cours, mais aucun n'a encore le degré d'achèvement permettant de tirer au mieux parti de la richesse et l'abondance des données qui s'accumulent très rapidement (séquences génomiques, cartes physiques, génétiques, EST, etc.). D'une manière plus générale, dans l'optique d'une modélisation intégrée de l'ensemble des connaissances liées aux génomes, la cartographie comparée jouera certainement un rôle majeur par les possibilités qu'elle offre de relier information génétique, fonctionnement cellulaire et évolution. ♦

SUMMARY

Comparative genomic mapping in mammals

Comparative genomic mapping is a comparative analysis of the chromosomal organization of genetic information. Making use of the biodiversity, it aims at understanding the biological structures, their functioning but also at inferring their evolution. The current popularity of this approach raised from another processes resulting from the peculiar structure of our current knowledge about genomes. Current cost of experimental analysis of genomes are still high, thus the information available is mostly concentrated on "model" organisms (as human or mouse for mammals). Inferring information relative to species of medical or economical interest from the analysis of model organisms is thus a promising processes. So is comparative genomic mapping. It should be also noticed that genome comparisons may be more complex and may benefit the experimental potential of some species. Although human is one of the most studied organism, comparative mapping is widely used to bridge the genomic information from human and the genetic knowledge from other species, in peculiar the ones for which genetic observations and experimentations are possible. ♦

RÉFÉRENCES

- Bernardi G. Isochores and the evolutionary genomics of vertebrates. *Gene* 2000 ; 241 : 3-17.
- D'Onofrio G, Mouchiroud D, Aïssani B, Gautier C, Bernardi G. Correlations between the compositional properties of human genes, codon usage and amino acid composition of proteins. *J Mol Evol* 1991 ; 32 : 504-10.
- Numéro spécial sous la direction de B. Dutrillaux. *Mammalia* 1986 ; 50 : 1-203. *Ce numéro spécial regroupe 9 articles consacrés à l'analyse chromosomique appliquée à la phylogénie des mammifères.*
- Anderson L, Archibald A, Ashburner M, et al. Comparative genome organization of vertebrates. *Mamm Genome* 1996 ; 7 : 717-34.
- Murphy WJ, Stanyon R, O'Brien SJ. Evolution of mammalian genome organization inferred from comparative gene mapping. *Genome Biol* 2001 ; 2 : 1-8.
- De Villena FPM, Sapienza C. Recombination is proportional to the number of chromosome arms in mammals. *Mamm Genome* 2001 ; 12 : 318-22.
- Robic A, Jeon TT, Rey V, et al. Construction of a high-resolution RH map of the human 2q35 region on TNG panel and comparison with a physical map of the porcine homologous region 15q25. *Mamm Genome* 2001 ; 12 : 380-6.
- Ostrander EA, Galibert F, Patterson DF. Canine genetics comes of age. *Trends Genet* 2000 ; 16 : 117-24.
- Lyons LA. Companion animals genetics. In : Sankoff D, Nadeau JH, eds. *Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families*. Computational Biology Series*. Dordrecht: Kluwer Academic Publishers, 2000 : 367-400.
- Faraut T, Demongeot J. Benefits of a model of segregation for the understanding of chromosomal evolution. In : Sankoff D, Nadeau JH, eds. *Comparative genomics: Empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families*. Computational Biology Series*. Dordrecht: Kluwer Academic Publishers, 2000 : 13-8.
- Matassi G, Sharp O, Gautier C. Chromosomal location effects on gene sequences evolution in mammals. *Curr Biol* 1999 ; 9 : 786-91.
- Sankoff D, Nadeau JH. *Comparative genomics: empirical and analytical approaches to gene order dynamics, map alignment and the evolution of gene families*. Computational Biology Series*. Dordrecht: Kluwer Academic Publishers, 2000 : 558 p.

* Ce livre est intégralement consacré à la cartographie comparée, avec une forte composante informatique. Outre les deux articles cités liés à la présente revue, de très nombreux articles sont consacrés à l'algorithmie et à la statistique de la comparaison de l'ordre des gènes entre espèces.

GLOSSAIRE

Isochore

Région (de taille variable) d'un chromosome caractérisée par une homogénéité du taux de G + C. On distingue ainsi traditionnellement les isochores « lourds » (fort taux de G + C), « moyens » et « légers » (faible taux G + C).

Substitution silencieuse

On considère traditionnellement deux types de mutations sur l'ADN (et affectant le produit des gènes) : les insertions/délétions (un ou plusieurs nucléotides ont été ajoutés/perdus) et les substitutions (un ou plusieurs nucléotides ont été remplacés). Dans le cas où la substitution ne modifie pas la séquence protéique codée par le gène (mutation vers un codon synonyme), elle est qualifiée de silencieuse. Très souvent, une substitution sur la troisième position d'un codon est silencieuse. Vitesse de substitution : vitesse à laquelle les substitutions s'accumulent.

Bon de commande

À retourner à EDK, 10 Villa d'Orléans - 75014 PARIS
Tél. : 01 40 64 27 49 - Fax : 01 40 64 27 47 - E-mail : editorial@edk.fr

NOM : Prénom :

Adresse :

Code postal : Ville :

Pays :

Fonction :

Je souhaite recevoir l'ouvrage : **Bien naître** : 16,03 € + 3 € de port = **19,03 € TTC**

en exemplaire, soit un total de €

- Par chèque, à l'ordre de EDK
 Par carte bancaire : Visa Eurocard/Mastercard American Express

Carte n° | | | | | | | | | | | | | | | | | | | | | |

Signature :

Date d'expiration : | | | | | | | |



ISBN : 2-84254-059-X 214 pages