

# 22

---

## **ORGANISATION, EXPRESSION ET ÉVOLUTION DES GÉNOMES**

*Président de la section*  
Dominique GIORGI

*Membres de la section*  
Frédéric BARRAS  
Catherine BISBAL  
Agamemnon CARPOUSIS  
Christophe CAZAUX  
Jean-Michel CLAVERIE  
Laura CORBO  
Bertrand DAIGNAN-FORNIER  
Sam DUKAN  
Olivier GANDRILLON  
Thierry GRANGE  
Annie JACQ  
Domenico LIBRI  
Régine LOSSON  
Julie MAALOUN  
Éric MICHEL  
Bernard MIGNOTTE  
Henry NEEL  
Geneviève PAYEN-GROSCHENE  
Isabelle RICHARD  
Dominique WEIL

### **PRÉAMBULE**

De nombreux documents concernant les orientations stratégiques, la conjoncture, le bilan des mandats des sections ou le rapport «les génomes et la génétique» ont été écrits ces dernières années.

La section 22 (ex-23) intitulée «Organisation, expression et évolution des génomes», depuis ce changement de mandature, a noté une orientation de ses thèmes de prédilection vers des approches plus globales en fonction des questions posées et des innovations méthodologiques. Cette section qui historiquement était la section de «génétique» se tourne de plus en plus vers la compréhension de l'évolution et la plasticité des génomes et des fonctions qui en découlent ainsi que sur la régulation de ces fonctions. La section 22 est consciente des recouvrements très forts avec d'autres sections du fait de l'étude moléculaire des régulations présidant aux grandes fonctions du vivant, en particulier les sections 21 et 23 à 26. D'autres aspects touchant plus particulièrement à l'évolution des génomes sont communs avec la section 29.

## 1 – LES ENJEUX SCIENTIFIQUES

Dans les domaines d'intervention de la section 22, il est clair que l'analyse des données issues du séquençage des génomes constitue une base importante. Si la France n'a pas pu, faute de moyens, se positionner pour être compétitive dans les grands programmes de séquençage de beaucoup d'organismes (néanmoins, les efforts de séquençage obtenus par le Centre National de Séquençage et le réseau des Génopoles ont permis d'obtenir des résultats de grande qualité dans le séquençage d'organismes à génomes plus petits que ceux des mammifères), elle peut tirer partie de l'analyse des génomes et se positionner dans « l'après-gènes » et aux retombées que cela entraîne dans la connaissance des grandes fonctions et de la plasticité des systèmes. L'analyse *in vivo* ou *in vitro* (*in cellulo*) des données obtenues constitue l'activité de prédilection de nombreuses équipes rattachées à la section 22.

### 1.1 ORGANISATION ET EXPRESSION DES GÉNOMES

Ces dernières années, l'analyse globale des génomes (transcriptomes, protéomes, métabolomes etc.) a pu se développer grâce à des innovations technologiques (puces ADN par exemple, bioinformatique) mais aussi grâce à des découvertes majeures (l'inactivation fonctionnelle de gènes des cellules animales par les siRNA). Un bon nombre de laboratoires se concentre sur la régulation génétique et épigénétique de l'expression des génomes par des études classiques (métabolisme des ARN codants par exemple) mais aussi avec le développement croissant de recherches centrées sur le rôle de la chromatine, et sur les ARN non-codants dont le nombre et les rôles deviennent de plus en plus importants.

Il semble que les petits ARNs non-codants aient un rôle quasi universel et très conservé au

cours de l'évolution. Ils ont des fonctions classiques (biogénèse des ribosomes, spliceosome) et contrôlent également l'expression génique par le système d'ARN interférence qui cible la stabilité des ARN messagers. Ils sont également impliqués dans certains cas dans des modifications de la structure chromatinienne. Ils sont impliqués dans des processus aussi variés que le développement embryonnaire, le contrôle des réarrangements géniques, le contrôle de la transposition et autres. Chez les procaryotes ce domaine est également en plein essor et intervient dans l'expression génique même s'il n'utilise pas les mêmes voies que chez les eucaryotes. Il est important de soutenir ces approches pour maintenir le niveau des laboratoires français dans le domaine.

On observe aussi une progression spectaculaire des approches globales d'analyse des chromosomes et plus seulement des génomes (approches dites de ChIP-chip). Une cartographie systématique des variations le long du génome des composants protéiques du chromosome est en cours (par exemple, programme ENCODE aux USA). Il se réalise actuellement sur la localisation des modifications des histones, des variantes d'histones, des autres protéines chromatiniques, ainsi que des facteurs régulateurs, de la machinerie de transcription et d'autres événements affectant l'expression du génome tel que la réplication : détermination des origines et du moment de la réplication. Cette connaissance va modifier considérablement notre compréhension de l'organisation chromosomique ainsi que notre image des processus épigénétiques affectant le contrôle de l'expression des gènes.

L'analyse globale des génomes requiert le développement d'outils informatiques importants. A ce titre, la biologie s'est rapprochée récemment des sciences formelles (mathématiques, physique, informatique) pour traiter les énormes masses de données produites, mais des efforts sont à poursuivre pour que la concertation entre ces disciplines conduise à un meilleur rendement de données exploitable en analyse et/ou en modélisation, notamment dans la reconstruction dynamique des réseaux de régulation.

## 1.2 BACTÉRIOLOGIE, MICROORGANISMES EUCARYOTES ET GÉNOMES

La collection de génomes bactériens séquencés (on peut s'attendre à un multiple de 1 000 dans les années à venir) apporte des informations de premier plan sur l'évolution et la structuration des génomes. Outre les aspects phylogénétiques visant notamment à déterminer les caractéristiques de l'ancêtre commun procaryote-eucaryote, ces analyses permettent d'évaluer l'apport des transferts horizontaux et celui des éléments transposables.

De plus, ces dernières années la bactériologie a suscité un regain d'intérêt de la part de la société, non seulement par son impact dans le domaine de la santé mais également dans celui de l'environnement. Les bactéries sont un élément essentiel dans la constitution et la régulation de la biodiversité. L'hétérogénéité spatiale, la prédation, la richesse nutritive ou les perturbations extrinsèques de l'environnement peuvent affecter la biodiversité. Une des réponses bactériennes à ces divers processus est la formation de biofilms. En contraste avec la majorité des études microbiologiques *in vitro* où les bactéries sont considérées en tant qu'individus cellulaires isolés, en conditions naturelles les bactéries semblent majoritairement présentes sous forme de biofilms. Ceux-ci constituent un modèle des plus simples d'assemblages biologiques multicellulaires et des premiers stades d'évolution de la socialité et permettent ainsi d'aborder expérimentalement des questions majeures de biologie évolutive. Il est donc important de soutenir les projets porteurs développés sur cette thématique.

De plus, les microorganismes eucaryotes, levures notamment, présentent des intérêts techniques et biologiques importants pour l'étude de l'évolution. Leurs génomes, relativement petits et simples, sont disponibles. La génétique est abordable, permettant d'accéder aux mécanismes moléculaires de l'évolution. L'ensemble doit ouvrir la possibilité d'intégrer

les différents niveaux depuis les phénomènes moléculaires primaires jusqu'aux effets à long terme sur les populations, les espèces, voire d'autres taxons plus élevés.

## 1.3 BIOLOGIE DES SYSTÈMES

Dans les domaines scientifiques couverts par la section 22 un changement de fond est en train de s'opérer. En effet, la vision de plus en plus claire des limites du réductionnisme pour résoudre les problèmes complexes (relation génotype-phénotype, métabolisme et physiologie, etc.) implique la nécessité d'aborder ces problématiques avec des outils nouveaux.

La « system biology » semble être l'alternative émergeante mais cette discipline n'a pas encore pleinement fait ses preuves ni atteint sa maturité. Les années à venir seront déterminantes pour l'essor de cette discipline et pour le positionnement des groupes français qui participeront à ce glissement méthodologique. Cette discipline est en cours d'installation et il est important de soutenir les programmes de qualité l'utilisant.

## 1.4 LES EUCARYOTES SUPÉRIEURS

Ces dix dernières années ont vu l'amélioration des techniques de cytogénétique par utilisation de la FISH (hybridation fluorescente *in situ*) et des techniques de coloration chromosomique. Ces progrès techniques ont permis de repérer des remaniements beaucoup plus fins entre espèces (insertions, délétions, inversions). En parallèle, la progression des connaissances au niveau moléculaire a connu un essor fulgurant grâce au séquençage automatique, la PCR et la bioinformatique. Le séquençage récent de génomes complets (homme, souris, etc.), associé à la puissance d'outils bioinformatiques de plus en plus performants et à la constitution de bases de don-

nées intégrées, a permis de mettre en évidence une quantité insoupçonnée de remaniements conduisant à l'évolution et à la constitution des génomes. Le séquençage complet du génome de primates associé à la bioinformatique et à la génomique (puces à ADN) va conduire dans un futur proche à la caractérisation quasicomplète des différences génétiques inter-espèces. Ce courant de recherche est une porte ouverte sur les études évo-dévo et la génomique moderne et permettra de relier des traits phénotypiques (morphologiques, développementaux) à la fonction des gènes ou réseaux de gènes et d'expliquer les différences observées entre deux espèces. Comme évoqué plus haut, l'utilisation d'organismes modèles sera certainement nécessaire pour compléter ces approches.

## 1.5 ÉVOLUTION DES GÉNOMES ET PALÉOGÉNÉTIQUE

On a observé ces dernières années une évolution spectaculaire des technologies de séquençage qui permettent maintenant de déterminer, en une journée et avec une machine, la séquence de plusieurs millions de bases à partir d'un échantillon tout en s'affranchissant des étapes de construction et de caractérisation d'une banque génomique (technologie 454). Cette approche va permettre d'analyser un encore plus grand nombre de génomes différents et d'étudier leur évolution. Pour les génomes bactériens, on peut espérer reconstituer des génomes entiers à partir des fragments courts de séquences déterminés par cette approche (actuellement, environ 100 bases). Pour les génomes eucaryotes supérieurs, une telle hypothèse paraît encore hors d'atteinte. Par contre, on pourra utiliser cette approche pour étudier les évolutions de portions de génomes sur des distances évolutives variables (aussi bien évolution au sein des cellules d'un même individu, des individus au sein des populations, qu'entre des espèces proches ou éloignées). Ces études d'évolution des génomes sont parfois un peu limitées car on recons-

titue des événements passés par interpolation à partir des génomes actuels. La possibilité d'apporter des données complémentaires permettant de tester la validité des modèles d'interpolation utilisés en déterminant des portions de séquence des génomes anciens sort maintenant du cadre de la science-fiction. Il a ainsi été possible d'obtenir de nombreux fragments du génome du mammouth et un projet d'analyse du génome de l'homme de Néanderthal est en cours. Les approches paléogénétiques émergentes s'intéressent à l'évolution des génomes méritent d'être soutenues et pourraient rentrer dans le cadre de la section 22. Ces approches paléogénétiques se heurtent encore souvent aux section 22 Organisation, expression et évolution des génomes problèmes posés par la transformation de l'ADN dans les échantillons anciens due aux processus de dégradation qui fausse les séquences obtenues. Ces problèmes pourraient être traités par des approches exploitant la réparation *in vitro* de l'ADN qui rentreraient elles aussi dans les champs couverts par la section 22.

## 1.6 GÉNÉTIQUE HUMAINE, BIOLOGIE DE LA CELLULE

Génétique humaine et maladies génétiques ont toujours été parmi les secteurs d'intervention de la section 22-ex23. Ces thématiques touchent à beaucoup de disciplines telles que clinique, épidémiologie, biologie cellulaire, physiologie, etc et font donc appel aux expertises d'autres sections des sciences de la vie et de l'INSERM notamment.

Les efforts dans ces domaines sont très largement soutenus par des fondations (LNCC, ARC, AFM, FRM, etc.), et l'industrie pharmaceutique. Actuellement ces recherches vont de la recherche des gènes morbides et de leurs mutations à la caractérisation moléculaire et cellulaire des mécanismes impliqués jusqu'aux applications thérapeutiques (par transfert de gènes ou mise en évidence de cibles médicamenteuses par exemple). Ces efforts

impliquent dans la plupart des cas des réseaux collaboratifs importants. À titre d'exemple, des progrès notables ont été obtenus dans de grands groupes de maladies tels que maladies neuromusculaires (myopathies), neurosensoryielles (surdités), diabète et obésité.

L'étude des mécanismes moléculaires associés nécessite l'utilisation de modèles animaux (bactéries, levures, drosophile, souris, etc.) et de lignées cellulaires animales. Les techniques d'invalidation géniques (knock-out, knock-in, siRNA) sont des outils puissants couramment utilisés. La biologie cellulaire ainsi que les nombreux développements en techniques d'imagerie connaissent actuellement un essor important (peignage moléculaire, FRET, IRM etc.). Il faut donc soutenir les axes utilisant des organismes modèles et assurer un soutien financier fort dans la mise en place des plateaux techniques et d'animaleries (souris notamment) de grande capacité, pour que la France puisse rester compétitive.

La stabilité et la plasticité des génomes interviennent dans un grand nombre de maladies dont les cancers, les résultats montrant que la transformation tumorale est associée à l'instabilité génomique. Pour étudier en détail les mécanismes de la stabilité et de la plasticité génomique l'utilisation des systèmes modèles est là aussi indispensable. Depuis quelques années, les «3R» (réplication, réparation, recombinaison) sont officiellement une prérogative de la section 22. Avec la démonstration récente que les tissus pré-néoplasiques sont le siège de stress réplicatif dû à un déficit de réplication et de réparation de l'ADN, l'étude mécanistique de ces processus de surveillance du génome, relativement peu explorés historiquement en comparaison de ceux gouvernant l'expression génique, doit être encouragée et soutenue.

La machinerie d'épissage des ARN pré-messagers nucléaires joue un rôle fondamental dans l'expression des gènes eucaryotes et tout particulièrement chez les vertébrés, où le nombre d'introns contenus dans la plupart des gènes de protéines est considérable. Des progrès importants ont permis l'identification et la caractérisation de nombreux facteurs de cette

machinerie. Néanmoins, beaucoup reste encore à faire pour comprendre comment cette machinerie extrêmement complexe s'assemble de manière spécifique au niveau des sites d'épissage et catalyse l'élimination des introns. De nouvelles pistes thérapeutiques sont actuellement développées (ciblage des facteurs d'épissage). Cet axe doit être aussi soutenu.

## 2 – RECOMMANDATIONS

Il est important de dynamiser la recherche en France, par des soutiens financiers importants non seulement pour la constitution de plate-formes technologiques (imagerie, animaleries, etc.), mais aussi pour des programmes spécifiques (génomique, bactériologie, par exemple). Les programmes blancs (ATIPE, ANR) répondent en partie à ces objectifs. Il est cependant à noter que la dotation de l'ANR blanche est insuffisante par rapport aux autres programmes thématiques, et que cette carence s'ajoute à l'absence de possibilité de politique propre au CNRS par manque de moyens. Le soutien en personnel (ITA, chercheurs) des équipes doit être absolument une priorité, et être pensé en même temps que l'attractivité des métiers de la recherche visant à attirer des acteurs de qualité dans ces filières (voir plus loin).

La recherche fondamentale est garante d'une recherche appliquée fructueuse. À ce titre, choisir les thèmes de recherche à partir de critères «finalisés» serait une grossière erreur d'appréciation. Il est important que la France, en particulier les politiques détenant les moyens de décision, soient comme c'est le cas dans d'autres pays industrialisés, convaincus par la démarche.

Durant ces dernières années on a pu constater un soutien accru aux jeunes équipes (ATIPE, Avenir, contrats jeunes chercheurs, etc.). Il est important que cette aide soit poursuivie après la période de financement initial, souvent assez courte, pour assurer la compétitivité

tivité de l'équipe. Néanmoins, il faut continuer à soutenir les équipes constituées ne rentrant pas dans le cadre de «jeunes équipes», celles-ci pouvant ne pas être forcément dans des projets «à la mode» ou pouvant amorcer des changements thématiques. On observe ainsi une dérive évidente et une «starisation» du système qui s'oriente vers un financement très important d'un tout petit nombres d'individus au détriment du financement de la majeure partie des équipes performantes. C'est particulièrement flagrant en ce qui concerne les jeunes chercheurs avec le programme EURYI qui offre 1 250 000 euros à un tout petit nombre de jeunes chercheurs sélectionnés dont le profil et les accomplissements diffèrent à peine du profil moyen de ceux recrutés actuellement en CR2 au CNRS, compte-tenu de l'intense compétition à l'entrée de notre organisme. Cette différence de niveaux con-

traste particulièrement avec la différence de financements mis à disposition (1 soutien EURYI correspond à 8 soutiens ATIPE). Il paraît douteux qu'une telle politique d'extrêmes inégalités de financement soit plus efficace qu'une politique visant à un meilleur équilibre entre ressources accordées et compétences démontrées qui permettrait le financement d'un plus grand nombre de jeunes chercheurs prometteurs et la formation d'une véritable pépinières de chercheurs. Une évaluation qualitative rapide de la productivité de tels déséquilibres financiers paraît essentielle.

Enfin, il est important, au-delà des sources de financement telles que l'ANR, que le CNRS puisse conserver des moyens d'action pour soutenir ses priorités, la direction du CNRS s'appuyant sur les travaux des différentes sections du Comité National pour les définir.