
RAPPORT DE CONJONCTURE

DU COMITÉ NATIONAL DE LA RECHERCHE SCIENTIFIQUE

ÉDITION 2014

Extrait



CNRS ÉDITIONS

15, rue Malebranche – 75005 Paris

SECTION 21

ORGANISATION, EXPRESSION, ÉVOLUTION DES GÉNOMES. BIOINFORMATIQUE ET BIOLOGIE DES SYSTÈMES

Extrait de la déclaration adoptée par le Comité national de la recherche scientifique réuni en session plénière extraordinaire le 11 juin 2014

La recherche est indispensable au développement des connaissances, au dynamisme économique ainsi qu'à l'entretien de l'esprit critique et démocratique. La pérennité des emplois scientifiques est indispensable à la liberté et la fécondité de la recherche. Le Comité national de la recherche scientifique rassemble tous les personnels de la recherche publique (chercheurs, enseignants-chercheurs, ingénieurs et techniciens). Ses membres, réunis en session plénière extraordinaire, demandent de toute urgence un plan pluriannuel ambitieux pour l'emploi scientifique. Ils affirment que la réduction continue de l'emploi scientifique est le résultat de choix politiques et non une conséquence de la conjoncture économique.

L'emploi scientifique est l'investissement d'avenir par excellence

Conserver en l'état le budget de l'enseignement supérieur et de la recherche revient à prolonger son déclin. Stabiliser les effectifs ne suffirait pas non plus à redynamiser la recherche : il faut envoyer un signe fort aux jeunes qui intègrent aujourd'hui l'enseignement supérieur en leur donnant les moyens et l'envie de faire de la recherche. On ne peut pas sacrifier les milliers de jeunes sans statut qui font la recherche d'aujourd'hui. Il faut de toute urgence résorber la précarité. Cela suppose la création, sur plusieurs années, de plusieurs milliers de postes supplémentaires dans le service public ainsi qu'une vraie politique d'incitation à l'emploi des docteurs dans le secteur privé, notamment industriel.

Composition de la section

Frédéric BARRAS (président de section); Lucas WALTZER (secrétaire scientifique); Catherine ANDRÉ; Valérie BORDE; Élisabeth BUGNARD; Pierre Antoine DEFOSSEZ; Guiseppina GIGLIA; Clotilde GIMOND; Philippe GLASER; Aline HUBER; Olivier JEAN-JEAN; Daniel KAHN; Christophe MILLIEN; Cécile NEUVEGLISE; Béatrice PY; Hughues ROEST CROLLIUS; Pierre RUSTIN; Bernd SCHUTTENGRUBER; René TOCCI; Chantal VAURY ZWILLER.

Résumé

La section 21 couvre les domaines de la génétique et de la génomique. Les chercheurs et Unités associés ont pour la plupart élu l'ADN soit comme molécule-objet d'étude soit comme molécule-outil pour analyser les mécanismes moléculaires du vivant. Les thématiques couvertes par la section parcourent donc la plupart des questionnements liés aux organismes vivants, de leur évolution à leur métabolisme. L'identité de la section 21 réside d'une part, dans son objectif de déchiffrement des mécanismes au niveau moléculaire et d'autre part, dans son exigence de comprendre les processus moléculaires dans leur contexte cellulaire. Les niveaux atomiques ou populationnels peuvent être occasionnellement intégrés mais ne font explicitement pas partie du cœur de la section 21. Dans les dernières années, les recherches couvertes par la section 21 ont été particulièrement influencées par le développement des technologies omiques et les développements de nouvelles méthodes de séquençage (Next Generation Sequencing, NGS). Ces approches affichent comme ambition le passage d'une biologie réductrice à une biologie plus intégrée embrassant l'étude de voies et réseaux biologiques. Ces changements méthodologiques se voient aussi accompagnés du passage d'une approche causale des événements à une approche corrélative. La section 21 participe à cette transition notamment par son implication dans les recherches en bioinformatique et biologie systémique mais ce changement profond dans l'approche du vivant et son interfaçage avec l'informatique, les mathématiques et la physique, doivent donner lieu à une véritable réflexion dépassant très largement les contours de la seule section 21 sur ses enjeux épistémologiques, son insertion dans les sciences du vivant et son évaluation.

I. Organisation, expression et stabilité des génomes

Les mécanismes fondamentaux contrôlant l'intégrité, l'organisation, l'expression des génomes et la création de la diversité génétique constituent autant de champs disciplinaires extrêmement dynamiques, compétitifs et essentiels dans la compréhension des propriétés du vivant. Ces dernières années ont vu des avancées spectaculaires dans la caractérisation fine de l'organisation chromosomique et chromatinienne à l'échelle du génome, et les rôles régulateurs variés que joue cette organisation pour contrôler de façon parfois dynamique, parfois métastable, l'activité du génome et la manière dont il détermine les propriétés du vivant. De façon générale, des ponts sont établis entre biologie de la chromatine et fonctions du génome étudiées traditionnellement : transcription, épissage, réplication, réparation.

Des avancées majeures ont été réalisées grâce à l'investissement de consortiums importants (ENCODE, modENCODE), focalisés sur l'étude du génome humain ou sur celui d'organismes modèles, dont la souris et la drosophile. Mais notre compréhension s'enrichit également des découvertes réalisées chez des organismes eucaryotes moins étudiés, ainsi que chez les procaryotes et les archées.

On assiste à une dissection de plus en plus précise et riche des éléments constituant le génome, qu'ils soient éléments régulateurs (enhancers, promoteurs...), transcrits (ARNs codants ou non codants...), ou structuraux (télomères, centromères...). Une emphase importante a été mise sur l'étude des phénomènes épigénétiques, terme dont l'acception fait débat, mais par lequel nous entendons ici l'ensemble des événements touchant la chromatine : nature et position des nucléosomes, modifications des histones, modifications chimiques de l'ADN (méthylation, hydroxyméthylation, et leurs dérivés). L'existence d'un « code des histones » est désormais

bien acceptée. Ce code s'enrichit perpétuellement de modifications chimiques nouvellement découvertes et par conséquent se complexifie. L'épigénétique est un champ d'investigation d'importance stratégique dans le domaine de la génétique, et la France figure parmi les leaders dans ce domaine.

Aujourd'hui, l'état des connaissances épigénétiques et épigénomiques n'en est qu'à ses prémices. L'analyse mécanistique et fonctionnelle de l'organisation chromatinienne, chromosomique et nucléaire des génomes a progressé très significativement grâce à l'accès facilité aux techniques de séquençage à très haut débit (RNA-seq, ChIP-seq, MethylC-Seq, HiC...) qui permettent d'augmenter considérablement la précision des résultats et leur contenu informatif. En particulier, des travaux fondateurs d'équipes françaises ont récemment déterminé les principes d'organisation tridimensionnelle du génome. Le séquençage de nouvelle génération couplé à la technique de Capture de Conformation Chromosomique (des approches de « 5C » ou « HiC ») permet l'analyse à l'échelle du génome de la conformation et la topologie des chromosomes et des domaines chromatiniens. La compréhension des mécanismes fondamentaux impliqués nécessite de plus d'intégrer ces données à l'échelle cellulaire en y adjoignant l'imagerie cellulaire à haute résolution qui permet d'analyser la diversité des comportements à l'échelle de la cellule individuelle évitant ainsi une vision moyennée d'une population de cellules (voir aussi section 3). De la même manière, un défi majeur dans le futur sera d'adapter les approches à l'échelle génomique sur une seule cellule (single cell HiC). L'intégration de l'ensemble de ces données représente une nouvelle frontière de ce domaine scientifique et ouvre des perspectives passionnantes de modélisation des processus biologiques et moléculaires.

L'organisation chromatinienne joue un rôle clé pour l'ensemble des « 3R » (réparation, recombinaison, réplication). De nombreuses études menées par des équipes françaises montrent, en effet, comment la réparation des lésions de l'ADN est influencée par la structure chromatinienne et l'architecture nucléaire.

Enfin, les séquences répétées et éléments transposables se sont révélés être des composants majeurs de tous les génomes et jouer un rôle aussi bien structural que fonctionnel. Mis sous silence par des voies de régulation impliquant de petits ARN, notamment dans la lignée germinale, les éléments transposables peuvent occasionnellement échapper à cette répression et amorcer des cycles de mobilisation qui vont conduire à de fortes instabilités génétiques. Si ces transpositions peuvent être délétères, les données de séquençage obtenues ces dernières années ont également mis en lumière leur rôle positif comme source de plasticité génétique grâce aux motifs fonctionnels qu'ils apportent au site de leur insertion. Ils sont également source de variations épigénétiques au sein des génomes, leur insertion pouvant conduire à la mise en place de structures chromatiniennes capables de modifier l'expression de gènes voisins. L'impact sur l'expression d'un génome où comme chez l'Homme ils représentent près de 50% des séquences, est certainement considérable et reste largement incomprise.

Les dysfonctionnements de la stabilité et la plasticité des génomes jouent un rôle important dans un grand nombre de maladies dont les cancers. En effet, de nombreuses données montrent que la transformation tumorale est associée à l'instabilité génomique. Stress réplicatif, déficit de réparation de l'ADN, recombinaison désordonnée, mobilisation de séquences transposables jouent tous des rôles importants dans l'étiologie des anomalies génétiques associées au cancer et sont observables dans les tissus pré-néoplasiques. Pour étudier en détail les mécanismes de la stabilité, de la plasticité génomique et de la surveillance du génome, les liens avec la clinique et les patients, mais aussi l'utilisation des systèmes modèles, sont indispensables.

Le développement et l'amélioration du NGS a modifié nos approches au plan quantitatif (celui de la vitesse de génération des données), mais également au plan qualitatif, en permettant la mise en œuvre d'approches expérimentales radicalement différentes de celles existantes jusque-là. L'application de ces appro-

ches a été fructueuse dans différents domaines, parmi lesquels celui concernant l'étude de l'expression des génomes. L'hypothèse selon laquelle l'environnement modifie la fonction du génome par le biais de l'épigénome est particulièrement suivie. Un autre axe retenant l'intérêt de la communauté est la question de l'héritabilité transgénérationnelle portée non pas par le génome mais par l'épigénome. Cette héritabilité ne fait aucun doute chez *C. elegans*, *D. melanogaster* et *A. thaliana*, et les arguments s'accroissent pour soutenir son existence chez la souris, suggérant qu'elle pourrait également exister chez l'humain.

La facilité de séquencer des génomes, dont des génomes humains, permet de mieux discerner les différences génétiques inter-individus et d'arriver à terme à établir un outil prédictif personnalisé. Le catalogage des transcrits a mis en évidence la diversité fonctionnelle des ARNs produits par le génome, et l'éventail de leurs rôles s'élargit constamment tant chez les eucaryotes que chez les procaryotes.

Il existe en France un nombre important d'équipes dynamiques et performantes étudiant ces différents aspects fondamentaux, et il est essentiel de soutenir les efforts dans ce domaine pour pouvoir maintenir cet avantage compétitif, assurer la vitalité du domaine et disposer d'un tissu de recherche qui permettra les retombées en terme de santé et d'applications qu'apportera une meilleure compréhension des liens entre génotype et phénotype.

II. Évolution des génomes et métagénomique

L'accélération du séquençage de génomes, bénéficiant de la chute des coûts de productions de séquence par des technologies en plein essor, a conduit à un changement d'échelle sans précédent. Il est désormais possible, par exemple, d'envisager le séquençage de toutes les espèces d'un genre pour de nombreux orga-

nismes procaryotes comme eucaryotes. On observe ainsi une forte densification des génomes séquencés par phylum, ainsi que de nombreux efforts pour obtenir des génomes de phylum non explorés qui ne présentaient jusqu'alors aucun intérêt autre que taxonomique. De gros projets internationaux de séquençages exploratoires ont ainsi vu le jour : 1KFG, 1 000 génomes fongiques ; i5K, 5 000 génomes d'insectes... Passée la première phase de séquençage massif de génomes qui a parfois généré des données génomiques de qualité très médiocre aussi bien pour leur séquence que pour l'annotation, une prise de conscience s'est faite sur la nécessité de produire des données de meilleure qualité. La possibilité de séquençage en lectures « pairées » de plus en plus distantes ainsi que le développement de nouveaux algorithmes d'assemblage et de prédiction de gènes prenant en compte des données de transcriptomique issues de RNA-seq contribuent désormais à un décryptage plus précis des génomes.

L'obtention d'autant de données génomiques engendre un renouveau des recherches en évolution moléculaire, basées non plus sur une poignée de marqueurs, mais sur l'intégralité des composants d'un génome. De nouveaux concepts pour reconstruire l'arbre de la vie, basés sur des événements évolutifs tels que présence/absence de gènes, transferts horizontaux ou ruptures de synténie sont autant de pistes prometteuses qui devraient enrichir les approches actuelles de phylogénomique.

L'accès à la génomique comparative à des échelles taxonomiques variées a eu de nombreuses conséquences sur la connaissance des mécanismes d'évolution et ce, grâce à de nouvelles méthodes telles que la reconstruction possible *in silico* du génome des ancêtres communs. Ainsi, les duplications complètes, duplications segmentales, et transferts horizontaux sont mieux caractérisés. L'accumulation de mutations sur des séquences dupliquées dégagées de pression de sélection, peut conduire à la production de protéines différentes de celles codées par les gènes originaux. Elles jouent également un rôle essentiel dans la naissance de nouvelles espèces. La duplication de

génomomes entiers est un événement rare mais qui a souvent contribué à la diversification des lignages. Les transferts horizontaux notamment des éléments transposables participent largement à la plasticité des génomes. Ces mécanismes et leurs conséquences restent cependant encore mal connus. Ainsi la duplication du génome de la truite arc-en-ciel, il y a environ 100 millions d'années, n'a pas été accompagnée de la réorganisation rapide du génome ou de la délétion massive de gènes, comme attendu. D'autres études, utilisant la levure comme modèle, suggèrent que la duplication du génome entier conduit généralement à la subfonctionnalisation des fonctions protéiques, tandis que les duplications segmentales permettraient l'apparition de nouvelles fonctions des produits du gène. La comparaison des génomes de levure a également révélé une fréquence insoupçonnée de génomes hybrides issus d'espèces proches ainsi que d'introgessions de matériel génétique dans des espèces distantes. Ces événements ont été particulièrement mis en évidence chez des espèces d'intérêt biotechnologique soumises à de fortes pressions de sélection. La contribution des équipes françaises dans ce domaine a été majeure.

Le séquençage à grande échelle des génomes conduit également à la résolution de grandes énigmes de l'évolution des espèces, telles le « big bang pré-cambrien », qui a vu l'apparition soudaine des grands groupes d'animaux modernes en seulement 20 millions d'années. C'est l'analyse conjointe des génomes d'arthropodes actuels et de leurs caractéristiques morphologiques qui a ainsi permis la construction de modèles mathématiques concluant à une vitesse d'évolution cinq fois plus rapide lors de l'explosion cambrienne que lors des époques ultérieures.

Le projet de séquençage de 1 000 Génomes humains initié en 2008 a ouvert le champ de la génomique des populations. Le séquençage massif des génomes d'individus au sein d'une population ou issus de populations multiples d'une même espèce, permet, par diverses techniques comme la cartographie QTL (Quantitative Trait Locus) ou les analyses d'association

pangénomiques (Genome-wide association studies, GWAS), d'identifier des régions génomiques et des polymorphismes contrôlant l'expression de traits phénotypiques spécifiques de certaines populations ou même d'individus particuliers, que l'expression de ces traits soit constitutive ou dépendante de certaines conditions environnementales. Ces approches ont notamment été utilisées pour appréhender des réponses à traits phénotypiques multiples, définir leur degré de corrélation génétique et mieux connaître les architectures génétiques complexes, comprendre le fonctionnement des réseaux de gènes mis en jeu dans l'élaboration de réponses phénotypiques (en particulier, de réponses adaptatives très diverses), et enfin étudier des mécanismes de plasticité phénotypique, de leur variabilité au sein d'une même espèce ou entre espèces voisines (notion de génomique des populations comparative). En outre, la génomique des populations a fourni un jeu de données sans précédent pour appréhender les mécanismes d'évolution du contenu en nucléotides, pour évaluer les biais de mutations/substitutions en relation avec des processus biologiques tels que réplication ou recombinaison méiotique. La reconstruction récente du génome ancestral du chimpanzé et de l'homme couplée à l'analyse fine des zones de recombinaison a révélé une évolution plus rapide du génome humain par l'intermédiaire du « GC-biased gene conversion » (GBGC). *A contrario*, l'analyse des populations de *D. melanogaster* ne révèle aucune trace de GBGC. Le couplage de telles approches dans des phylums différents devrait permettre de comprendre l'évolution et l'impact des processus biologiques sur le modelage des génomes.

Les progrès des techniques de séquençage et de bioinformatique (algorithmes, analyses statistiques, gestion des « big data ») ouvrent de nouvelles voies pour étudier la biodiversité des populations dans un écosystème donné. La métagénomique, discipline en pleine expansion, touche des microbiomes extrêmement variés : intestinal, buccal, du sol, des eaux, des aliments, *etc.* Donnant accès à une biodiversité insoupçonnée (recensement exhaustif et quantitatif), constituée pour partie d'organis-

mes non cultivables, de virus, et de phages, la métagénomique a engendré un renouveau de la taxonomie et un intérêt pour l'étude de la structure et de l'évolution des populations microbiennes. Depuis le métagénome de la mer des Sargasses en 2004, les projets de métagénomique se sont multipliés de façon explosive ces dernières années, permettant également une meilleure compréhension du fonctionnement (et parfois des dérèglements...) des écosystèmes naturels ou des environnements colonisés par l'homme. La diversité des microorganismes et leurs interactions avec l'environnement ou leur hôte font partie des maillons primordiaux de la diversité biologique, de la qualité de l'eau, des sols ou encore de l'atmosphère, toutes choses qui influencent la santé des êtres vivants et plus particulièrement de l'homme. Ces dernières années ont vu une prise en compte de plus en plus importante de ces microbiomes, de leur diversité, de leur dynamique mais aussi de leur déséquilibre dans de nombreuses pathologies, souvent chroniques, affectant en premier lieu l'appareil digestif mais également la peau, les muqueuses, des voies respiratoires, et même certaines fonctions du système nerveux central. Le fonctionnement du corps humain est donc vu à présent comme le résultat des interactions de l'expression du génome humain, du fonctionnement du microbiome et de l'environnement (notion de métabiome). Mieux comprendre comment des modifications de ces populations microbiennes permettent l'émergence de ces maladies et réussir à modéliser leur dynamique permettra des diagnostics et des traitements mieux adaptés à chacun. Le développement de la métagénomique a déjà permis le séquençage de larges cohortes de génomes humains et de leurs microbiomes associés, au sein de populations en bonne santé ou atteintes de maladies chroniques, et a montré la formidable diversité de ces microbiomes. À l'avenir, ces études permettront de mieux comprendre comment les génomes humains et leurs microbiomes ont co-évolué à travers le temps et dans différentes régions du globe. Pour cela, il est important de coupler métagénomique, méta-transcriptomique, et métaprotéomique, afin

de collecter des données sur l'activité biologique du métagénome. Ce concept se développe également dans d'autres domaines que la génétique humaine. Permettant de cribler des activités enzymatiques d'intérêt, il est en plein essor pour étudier les mécanismes de résistance aux métaux dans le sol, les réponses des sols aux changements climatiques pour ce qui est de la dégradation de la matière organique, ou encore la diversité fonctionnelle des microorganismes des océans (projet Tara Oceans).

III. Microbiologie

Avec une estimation de 5.10^{30} cellules, les procaryotes (bactéries et archaea) représentent le groupe le plus important sur la Terre. Ils ont la capacité de coloniser tous les environnements, même ceux qui peuvent nous sembler les plus hostiles à la vie comme les sources hydrothermales des fonds océaniques ou les déserts salés. Ces organismes présentent des propriétés métaboliques et phénotypiques d'une extrême diversité. L'analyse des communautés microbiennes par des techniques sans culture montre que les bactéries que nous savons cultiver ne représentent qu'une infime minorité des espèces bactériennes. Elles sont utilisées depuis le début des civilisations humaines pour la production des aliments ou le traitement des déchets. Les études récentes sur le microbiome humain ont révélé son rôle essentiel pour la santé qui l'apparente à un véritable organe (voir section 2). Les bactéries ont été étudiées depuis Pasteur et Koch pour leur capacité à provoquer des maladies chez l'homme, les animaux et les plantes. L'étude du monde bactérien présente donc des facettes multiples, cognitives et appliquées, et les équipes françaises ont su depuis de nombreuses années maintenir un niveau d'excellence en microbiologie moléculaire, qu'elle soit fondamentale, environnementale, biotechnologique ou infectieuse.

Les bactéries sont des modèles uniques d'étude moléculaire, de véritables micromachines sources d'innovations multiples en biologie moléculaire. Les enzymes de restriction en furent un des fleurons, la PCR un autre. Récemment, c'est le système CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) qui a ouvert de nouvelles méthodes dans l'ingénierie génomique. Ce mécanisme Lamarckien d'immunité bactérienne contre les phages et les plasmides conjuguatifs est un sujet d'intenses recherches pour caractériser les mécanismes moléculaires de la capture de fragments d'ADN des cibles (immunisation) et de dégradation de ces éléments génétiques lors d'attaques ultérieures (immunité). Ce système est maintenant utilisé comme un outil de manipulation génétique puissant nécessitant uniquement l'expression d'une protéine (Cas9) et d'un ARN guide déterminant le site de coupe. CRISPR-Cas9 est utilisé pour répondre à des questions de recherche fondamentale pour de très nombreux organismes vivants et en vue d'application comme pour la thérapie génique.

La microbiologie est une discipline globale allant de la molécule à l'écosystème. Elle est l'objet d'études tant réductionnistes que systémiques. Par leur apparente simplicité, les microorganismes peuvent désormais permettre une recherche à plusieurs niveaux d'intégration du vivant. Les enjeux y sont à la fois cognitifs, pour le développement et la validation de méthodes qui sont ensuite appliquées à des systèmes plus complexes et finalement applicatifs afin d'améliorer des processus biotechnologiques. Ainsi, les mécanismes moléculaires sont-ils analysés au sein de la cellule par des méthodes d'imagerie à super résolution et par reconstitution des architectures moléculaires dans des systèmes *in vitro* reproduisant des paramètres cellulaires. De même, l'organisation et la dynamique du génome sont maintenant étudiées en combinant des approches génétiques, d'imageries et d'étude de la conformation *in vivo* du chromosome (Hi-C) basée sur le séquençage haut débit. Enfin, la diversité métabolique et leur capacité d'adaptation font des bactéries les systèmes les plus performants pour le développement d'une approche systémique et la modélisation des processus bio-

logiques. La connaissance du métabolisme permet de redessiner les processus biologiques et d'en créer de nouveaux en combinant les analyses intégratives au génie génétique et à des expériences d'évolution contrôlée pour la sélection de nouveaux processus biologiques. Cette biologie synthétique permet de comprendre le fonctionnement d'une bactérie et est porteuse d'applications considérables dans la chimie de synthèse.

L'étude de la diversité phénotypique et génétique au sein d'une population bactérienne constitue un champ de recherche aujourd'hui accessible et en pleine expansion. L'hétérogénéité populationnelle appréhendée par les approches de cellule unique et d'imagerie offre un regard nouveau sur l'ensemble des processus classiquement étudiés, de la régulation de l'expression à la réplication via l'adaptation métabolique et bien évidemment dans les études des communautés bactériennes et dans les biofilms. Les bactéries ont développé des moyens de communication intercellulaires des plus simples, qui leur permettent de se compter (le quorum sensing), aux plus compliqués, faisant intervenir de véritables processus développementaux. Les biofilms sont des structures multicellulaires qui procurent des propriétés nouvelles de colonisation et de résistance. L'existence de cellules persistantes, récemment décrites, au sein de ces structures a de très fortes implications pour la compréhension de l'adaptation des bactéries aux conditions défavorables et la résistance aux antibiotiques.

Les développements méthodologiques en génomique, en imagerie, en métabolomique et en modélisation mathématique permettent aujourd'hui une approche intégrée des écosystèmes microbiens afin de caractériser les relations entre les espèces et les propriétés de ces communautés impliquant des réseaux de mutualisme par exemple trophique mais aussi d'antagonisme.

L'émergence de bactéries multi-résistantes combinée à la raréfaction de nouveaux agents anti-bactériens sur le marché font craindre le retour à une ère pré-antibiotique. Les protocoles de criblage à haut débit robotisé per-

mettent le développement d'approches de « chemical genetics » visant à identifier des interactions entre drogues de manière génotype-dépendante et/ou espèce-dépendante. Véritable inventaire de combinaisons à grande échelle, ces nouvelles approches sont une source d'hypothèses et de questionnements nouveaux qui alimentent les approches mécanistiques de génétique moléculaire classique, créant un continuum salubre entre approches globales et réductionnistes.

IV. Génétiques des modèles et aspects translationnels

Avec l'émergence ces dernières années de techniques de plus en plus efficaces d'extinction/d'expression de gènes, un grand nombre d'organismes génétiquement modifiés ont été créés. Une première finalité réside dans une meilleure compréhension de la fonction des gènes et de l'organisation du vivant, du normal au pathologique. Cette approche a produit et continuera à produire des outils irremplaçables pour analyser les mécanismes. Une seconde finalité est la modélisation de situations pathologiques connues chez l'homme, avec bien souvent la thérapie en ligne de mire. Malheureusement la prise en compte tardive ou inexistante des spécificités relatives à chaque espèce, des sensibilités propres à chaque individu (les « modèles », en particulier murins, sont souvent des clones), associée à la complexité inattendue des réseaux de régulations mettant en jeu chromatine, gènes, ARN, protéines, métabolites (bien loin du schéma initial linéaire liant de façon simple l'ADN et la fonction) a conduit à l'obtention de modèles fragmentaires, bien souvent décevants, ne modélisant pas parfaitement la complexité des maladies telle qu'exprimée dans la population humaine. Néanmoins, il existe de notables exceptions de modèles murins ayant bien reproduit la pathologie humaine et plus récemment, des modèles spontanés de maladies

homologues entre l'Homme et des animaux domestiques, le chien par exemple, ont permis, non seulement la découverte de nouveaux gènes et de nouvelles fonctions dans les maladies homologues humaines, mais aussi la mise à disposition de modèles naturels de maladies monogéniques et de maladies complexes permettant de développer des essais thérapeutiques avec un double bénéfice pour la médecine humaine et vétérinaire.

L'absence d'expérimentation *sensu stricto* chez l'homme laisse intacte la nécessité de générer et étudier des organismes génétiquement modifiés ou des populations d'organismes modèles. Dans les cas où une modélisation transposable à l'homme est recherchée, le recours à des modèles naturels de maladies homologues est envisageable mais il convient d'effectuer un choix toujours plus rigoureux des organismes utilisés et cela en fonction des questions ciblées et des maladies : le caractère relativement aisé de l'étude des muridés ne saurait en faire disparaître la trop grande spécificité. Il apparaît dès lors justifié et recommandé d'utiliser la plus grande variété de modèles pour bénéficier des avantages de chaque organisme. Du point de vue méthodologique, lorsque la transposition à l'homme est l'objectif poursuivi, il est impératif de respecter la complexité du vivant. Il faut donc favoriser des projets qui prennent véritablement en compte la richesse génétique des populations.

D'un point de vue éthique, chaque analyse génétique nécessitant des modèles animaux doit débiter par une large réflexion sur la pertinence des modèles utilisés, respecter la règle des 3R (Réduire, Raffiner, Remplacer), réfléchir à leur complémentarité et élargir le champ des possibles vers des modèles spontanés de pathologies. Ces « modèles spontanés » sont des « patients » à part entière avec des analyses génétiques et des essais thérapeutiques contrôlés en milieu vétérinaire, avec suivi et autorisation du propriétaire. Ainsi ces « modèles » sont plus proches de l'Homme que de la souris pour laquelle les échecs de développement de thérapies sont malheureusement flagrants. Autant ces modèles spontanés sont pertinents et très informatifs en médecine humaine pour des

aspects génétiques et thérapeutiques, autant des modèles classiques génétiquement modifiables (souris, poisson zèbre, drosophile, *C. elegans*, *S. cerevisiae*...) demeurent nécessaires pour des études fonctionnelles et mécanistiques des processus biologiques sous-jacents aux maladies étudiées. La réflexion sur la nécessité, le choix et la complémentarité des modèles est donc vraiment à l'ordre du jour.

Pour les aspects translationnels, les termes « recherche translationnelle » (signifiant de la « paillasse au lit du malade »), et « médecine personnalisée » sont les fers de lance de la génétique médicale actuelle et à venir. La connaissance des génomes, de plus en plus fine dans leur organisation, leur expression, leur régulation/dérégulation, leurs modifications post-transcriptionnelle et post-traductionnelle, de leur régulation spatio-temporelle... constitue un immense bouleversement. L'étude des pathologies ne peut plus se limiter aujourd'hui à la recherche de mutations dans les régions codantes du génome mais doit prendre en compte l'impact d'altérations génétiques (et épigénétiques) dans des régions régulatrices et des régions anciennement considérées comme non codantes. Ces aspects sont maintenant analysables grâce aux progrès des méthodes NGS (cf. ci-dessus) pour lesquelles il est crucial de rappeler le besoin de nouvelles ressources et de développements en (bio)informatique et biostatistique. Ceci reste un enjeu stratégique majeur pour les années à venir et le caractère

multidisciplinaire du CNRS est un atout exceptionnel dans ces domaines.

Il apparaît donc qu'une « translation » depuis la recherche fondamentale dont elle se nourrit n'est évidemment possible qu'à condition de préserver l'intégrité du secteur de recherche fondamentale bien souvent mise à mal ces derniers temps par la finalisation imposée à court, voire très court, terme pour le financement des travaux.

Conclusion

En conclusion, les champs de recherche couverts par la section 21 ont été extrêmement dynamiques et ont connu des évolutions majeures au cours de ces dernières années, notamment grâce au dynamisme des équipes de recherche françaises. À un moment où les ressources financières et humaines nécessaires au maintien du niveau de compétitivité de ces équipes sont de plus en plus incertaines, il est essentiel que le CNRS en particulier apporte tout son soutien à ce domaine de la biologie par essence multidisciplinaire. Être à l'affût d'éventuels développements est nécessaire et légitime pour autant que soient respectés le caractère non prédictible du travail de recherche et l'indispensable sécurité financière des laboratoires.

Comité national de la recherche scientifique. « Section 21- Organisation, expression, évolution des génomes. Bioinformatique et biologie des systèmes ». *Rapport de conjoncture 2014*, [édition PDF en ligne]. ISBN : 978-2-271-08746-1. Disponible sur : <http://rapports-du-comite-national.cnrs.fr/>