

L'émergence de nouvelles formes d'invention collective : réseaux et consortia de recherche dans le domaine des biotechnologies

La recherche en consortium permet de partager les coûts et les risques de la R&D, de réduire le volume des investissements, de créer de nouvelles techniques, de générer de nouvelles connaissances grâce à la réunion de collections d'échantillons privés et dispersés. La constitution et la gestion des consortia se heurtent toutefois à des problèmes d'organisation et de coordination : la difficulté réside, notamment, dans l'équilibre à trouver entre le partage des connaissances au sein du réseau, afin de favoriser l'apprentissage collectif, et un certain degré de protection individuelle, afin d'assurer des retours pour les participants qui ont investi des ressources privées dans le projet.

**par Maurice Cassier
Institut pour le management
de la recherche (IMRI),
Université Paris-Dauphine**

Si les acteurs économiques ont utilisé des formes d'invention collective, au moins dès le XIX^e siècle, par exemple via des réseaux d'échange de connaissances techniques dans des

districts industriels, dans l'industrie du fer et de l'acier (1), on observe actuellement une montée des consortia et des réseaux de recherche coordonnés, entre firmes ou entre firmes et institutions scientifiques (2). Ces alliances peuvent être spontanées, à l'initiative d'un club de firmes ou de laboratoires publics, ou encore aidées, sinon organisées, par les pouvoirs publics (consortium japonais sur le génome humain impulsé par le MITI ou réseaux de recherche européens soutenus par la CEE). Les bénéfices économiques de la recherche en consortium sont multiples : elle permet de partager les coûts et les risques de la R&D, de réduire le volume des investissements, de réunir des compétences complémentaires pour créer de nouvelles connaissances ou techniques, de générer de nouvelles connaissances grâce à la réunion de collections d'échantillons privés et dispersés. La constitution et la gestion des consortia de recherche posent toutefois deux types de problèmes que nous nous proposons d'examiner dans cet article : il s'agit, en premier lieu, d'organiser une certaine convergence technique entre les participants et de favoriser des dynamiques d'apprentissage collectif au sein du réseau ; il s'agit, en second lieu, d'inventer de subtils compromis entre le nécessaire partage des informations et des savoir-faire entre les participants et l'assurance d'un certain degré de protection individuelle. Des difficultés de coordination technique nuisent à l'efficacité du consortium. La crainte d'une faible appropriabilité des données de la recherche conduisent les firmes à minimiser leur participation au sein du consortium ou à refuser de s'y engager. Nous montrerons que les acteurs

locaux sont susceptibles d'inventer spontanément de nouvelles règles et de nouvelles pratiques pour résoudre les problèmes posés par l'invention collective et qu'ils peuvent échanger des « bonnes pratiques de coopération » entre les différents réseaux. Nous fondons notre argument sur l'étude de plusieurs consortia de recherche dans le domaine des biotechnologies : le réseau de laboratoires européens constitué pour décrypter le génome de la levure ; le projet Bridge sur les lipases qui associe vingt-deux participants publics et privés pour obtenir des connaissances de base sur une famille d'enzymes d'un grand intérêt industriel ; les consortia international et européen, mis en place pour accélérer l'identification des gènes de prédisposition au cancer du sein.

Exploiter plusieurs configurations d'objets techniques

La recherche collective peut s'organiser autour de trois configurations d'objets techniques et, sous certaines conditions et dans certaines limites, permet d'en exploiter les économies d'échelle ou les effets de collection.

Une recherche en consortium peut, tout d'abord, tirer avantage de la divisibilité et la modularité extrêmes de certains objets scientifiques comme les chromosomes ou les génomes. Les génomes

(1) Allen R, 1982, « Collective invention », *Journal of Economic Behavior*, 4.

(2) Aldrich H.E., Sasaki T., 1995, « R&D consortia in the United States and Japan », *Research Policy* 24 ; Cassier M, Foray D, Steinmueller E, 1997, « Collective Invention and European Policies », Working Papers IMRI, Université Paris-Dauphine.

sont, en effet, des objets hautement divisibles en segments de taille variable, qu'il est possible de répartir entre un grand nombre de participants pour en accélérer le décryptage. Ainsi, dans le cadre du réseau européen sur le séquençage de la levure, les 300 kilobases du chromosome III

furent réparties entre trente-cinq laboratoires, chacun d'entre eux en prenant en charge une dizaine. La division du travail est organisée en parallèle : chaque laboratoire reçoit les morceaux de chromosome qu'il doit décrypter et envoie les données obtenues dans un laboratoire informatique central pour vérification et assemblage final. Cette organisation a été décrite comme une forme « d'industrie à domicile » (3) . Elle permet d'éviter la duplication des efforts de recherche et les phénomènes de course de vitesse entre laboratoires qui travaillent sur la même région du génome.

Des consortia sont également mis sur pied pour constituer et exploiter des collections d'objets et de données. La réunion d'une collection d'échantillons permet d'obtenir des connaissances d'une qualité supérieure ou impossible à obtenir autrement. La recherche collective permet ici d'exploiter les interdépendances qui existent entre les objets et les données rassemblées : une connaissance structurale obtenue sur une enzyme A permettra d'élucider la structure d'une enzyme B et des enzymes voisines. Dans le projet Bridge sur les lipases, la réunion d'une collection d'enzymes d'une taille suffisante (10 ou 15 enzymes de la famille des lipases) favorise l'obtention de nouvelles connaissances grâce aux comparaisons et aux synthèses qu'il sera possible d'entreprendre : « *New knowledge on a number of these enzymes will make it possible to understand why they are lipases and how they function as such* » (article 1 des *guidelines* du consortium) (*). La réunion de matériaux qui appartiennent à des propriétaires différents pose des problèmes de confidentialité (certaines firmes ne souhaitent pas communiquer des souches qu'el-

La réunion de matériaux qui appartiennent à des propriétaires différents pose des problèmes de confidentialité

les exploitent dans leurs procédés). Ces difficultés peuvent être résolues en organisant une circulation cloisonnée de ces matériaux entre les propriétaires qui les mettent à disposition et les laboratoires qui les étudient et les restituent au terme de leurs investigations. Les propriétaires peuvent

également communiquer des matériaux de référence moins confidentiels mais suffisamment proches des souches industrielles pour récupérer des connaissances transposables. Les recherches de gènes de prédisposition aux maladies supposent également la constitution de collections de données biologiques (prélèvements d'ADN) et de données médicales (pédigrés familiaux bien constitués). Ici aussi, c'est l'interdépendance entre les données qui joue : la confrontation d'un nombre suffisant de données familiales, au sein d'un petit nombre de familles bien constituées ou d'un grand nombre de familles éparses, permettra d'identifier un gène de prédisposition à une maladie. Les acteurs scientifiques s'associent en consortium pour mettre en commun leurs collections d'échantillons et faire des études statistiques sur les gènes : « *So it became a numbers game : the more families from diverse populations researchers can collect, the better their chance to find the culprit gene. Solid evidence of a gene location in the genome will emerge from looking a different data sets* » (4) (**). C'est la finalité du consortium européen sur les gènes de prédisposition au cancer du sein.

La recherche collective permet enfin d'exploiter les complémentarités qui lient les activités et les objets techniques. La combinaison des souches d'enzymes de plusieurs partenaires est susceptible de produire des mutants inédits. L'association de centres d'excellence publics en cinétique enzymatique et de centres d'excellence privés en biologie moléculaire au sein du projet lipases a permis de concevoir des hybrides, de les tester en milieu de laboratoire, puis en milieu industriel. La complémentarité est, à la fois, hori-

zontale entre des spécialités différentes et verticale entre différentes phases de l'activité de R&D. Le développement d'instruments de dosage biologique suppose souvent l'assemblage de ressources complémentaires : les uns apportent l'instrumentation de mesure, les autres les molécules spécifiques à doser. Quant à la mise au point de tests de diagnostic génétique, elle suppose des collaborations entre des laboratoires de recherche spécialisés en biologie moléculaire, des laboratoires cliniques et des sites hospitaliers - qui jouent un rôle en amont, pour réunir les collections de familles et d'échantillons d'ADN, et en aval, pour tester les prototypes - et, enfin, des firmes pharmaceutiques ou de biotechnologie, dès lors qu'il est question de produire des kits à l'échelle industrielle. Il en va ainsi des consortia régionaux mis sur pied en Grande-Bretagne pour coordonner le développement des tests génétiques.

Les instruments de coordination technique et l'apprentissage collectif

La production de science et de technologie en collaboration suppose un certain degré de convergence technique entre les participants. La question est particulièrement importante dans les consortia comportant un grand nombre de membres qui travaillent chacun dans leurs murs (à la différence des consortia qui créent une installation de recherche commune). Les acteurs doivent s'entendre pour mettre sur pied une véritable infrastructure de recherche qui comprend la définition d'interfaces, afin de permettre l'assemblage

(*) « Les nouvelles connaissances acquises sur un certain nombre de ces enzymes permettront de comprendre en quoi il s'agit bien de lipases et comment elles fonctionnent, en tant que telles ».

(**) « Nous sommes amenés à jouer sur les grands nombres : plus les chercheurs pourront repérer de familles dans des populations les plus diverses, plus ils auront de chances de découvrir le gène coupable. Une certitude étayée quant à la localisation d'un gène sur son génome ne peut découler que de l'examen de corpus de données différents ».

(3) OCDE, 1995, « Le grand programme sur le génome humain », Paris.

(4) Science, 1996 « Gene Hunters close in on elusive Prey », vol 271, mars.

de données individuelles, la standardisation des méthodes ou des données de recherche, la création de supports matériels d'échange et d'accumulation des données, ce qui suppose l'adoption d'une nomenclature commune pour coder les informations, des procédures de contrôle de la qualité des données obtenues, la formation de ressources humaines spécifiques qui disposent d'une autorité particulière au sein du réseau (par exemple, les coordinateurs informatiques dans les réseaux européens de séquençage du génome).

Les problèmes de coordination furent particulièrement ardues au sein du consortium européen sur la levure, qui associait au départ trente-cinq laboratoires dispersés géographiquement en Europe et de niveaux de compétences très hétérogènes. On peut cependant remarquer que la réalisation collective d'une carte, qu'elle soit génétique ou géographique, exerce une forte contrainte de coordination sur les participants : « *Il faut que tous les bouts collent* » disait un coordinateur européen du projet Bridge. Le code de bonne conduite adopté par les membres du réseau, *The Perfect Gentleman Sequencer*, prévoit plusieurs instruments de coordination technique. En premier lieu, la préparation du matériel génétique est confiée à un laboratoire central, *The DNA coordinator*, chargé de construire une carte sur la totalité du chromosome avec une résolution donnée et de distribuer les segments d'ADN aux participants. Cette procédure garantit l'homogénéité des

objets étudiés au sein du réseau. En second lieu, l'assemblage des données individuelles sera facilité par la définition de chevauchements, les *overlaps*. En troisième lieu, le code de bonne conduite définit de manière très circonstanciée des standards de qualité des données. Ces standards de qualité jouent un rôle crucial pour faire converger les travaux des acteurs qui ont des méthodes d'obtention des données qui ne sont pas homogènes. Ils représentent également une sorte de certification

attribuée par le consortium. Les *guidelines* prévoient une progression du taux de validité des données entre le présent projet et le suivant : « *In reality, we propose to aim a confidence value of 99,8 %, which seems more realistic for the present Bridge program and reach 99,9 %*

for the next Biotech Programme » (***) . En

quatrième lieu, les *guidelines* s'attachent à fixer des

procédures de contrôle de la qualité des données produites, par le coordinateur informatique et par réexamen d'environ 15 % des séquences par des laboratoires anonymes. En cinquième lieu, la coordination informatique du réseau de laboratoires est essentielle pour la vérification des données produites, leur accessibilité, mais aussi pour l'assemblage final des données individuelles dans la carte du génome de la levure. A ce titre, le coordinateur informatique du projet a une responsabilité particulière dans le contrôle de la qualité des données. Il faut noter que le schéma de coordination informatique du consortium de la levure a inspiré d'autres consortia de biotechnologie. La mise en œuvre de ces différents instruments de coordination a favorisé un processus d'apprentissage collectif au

sein du réseau : la grande majorité des équipes remplirent leur contrat et le rendement du séquençage a notablement augmenté. « *The efficiency of the yeast genome sequencing network have increased appreciably, in term of flux of Mb/year final sequence* » (Bridge, Final report) (****).

Le consortium européen a fortement contribué au séquençage de la levure, en avance sur les plans américains. D'une modeste « *European cottage industry* » (5), qui rassemblait trente-cinq laboratoires en 1989, la formule s'est étendue à plus de 100 laboratoires européens ainsi qu'à des centres de recherche américains, canadiens et japonais. On voit donc ici que la formule du réseau de

Le consortium européen a fortement contribué au séquençage de la levure, en avance sur les plans américains

laboratoires, sous réserve d'y consentir des coûts de coordination suffisants pour faire travailler ensemble des participants hétérogènes, représente une solution organisationnelle tout à fait intéressante. Une autre alternative réside dans la création d'une installation centrale tra-

vaillant à l'échelle industrielle (c'est la voie suivie par le consortium japonais sur le génome humain qui regroupe vingt

firmes et le MITI ; ce consortium a créé une société de biotechnologie chargée de développer les travaux et les instruments de séquençage).

Si des consortia sont susceptibles de nourrir une dynamique de convergence technique, via des échanges de savoir-faire, de techniques et la définition de standards de qualité, à l'instar du « groupe français des onze » qui se propose de standardiser les techniques d'étude des mutations des gènes de prédisposition au cancer du sein, d'autres reposent sur une stricte spécialisation des participants. Il en est ainsi du consortium européen sur le cancer du sein, qui associe un laboratoire central fortement équipé qui se charge du séquençage et de la réalisation des études statistiques et des laboratoires fournisseurs d'échantillons et de données familiales. Ce dispositif a parfois généré des tensions entre fournisseurs et laboratoire central (6). Il s'avère également que c'est ce dernier qui a déposé la demande de brevet sur le gène identifié.

Les formules de compromis sur la propriété et l'accès aux données des recherches

(***) « *En fait, nous ambitionnons de nous rapprocher d'un taux de fiabilité de 99,8%, ce qui semble le plus réaliste pour le programme actuel, "Bridge", et d'atteindre les 99,9% pour le programme à venir, Biotech* ».

(****) « *L'efficacité du réseau pour le séquençage du génome de la levure a progressé de manière notable, en termes de flux de séquences définitives par année (en millions)* ».

(5) Science, 1996 « *Yeast Genome Sequence Ferments New Research* », vol 272, avril.

(6) Cassier M., Gaudillière JP. 1997 « *Les relations entre science, médecine et marché dans le domaine du génome* », Programme Génome du CNRS.

réalisées en collaboration

La difficulté réside dans l'équilibre à trouver entre le partage des connaissances au sein du réseau, afin de favoriser l'apprentissage collectif, et un certain degré de protection individuelle, afin d'assurer des retours pour les participants qui ont investi des ressources privées dans le projet. Les acteurs locaux sont susceptibles d'inventer des arrangements originaux pour réguler la propriété et l'accès aux données.

Le consortium européen sur les lipases offre un terrain de choix pour étudier ces formules de compromis (7). Dans ce projet multilatéral qui associe vingt-deux participants, dont six firmes industrielles largement concurrentes et seize laboratoires publics de recherche, la difficulté est double : premièrement, il faut inciter les industriels concurrents à participer au projet et à jouer le jeu coopératif, au moins un minimum : « *On va quand même partager de l'information un peu privée, c'est ça l'idée* » (coordinateur administratif) ; deuxièmement, il faut concilier règles académiques de publication et l'appropriation privative des connaissances produites.

La première réponse consiste à remonter en amont pour éviter les conflits entre industriels sur les applications : « *Si le choix des enzymes a été fait sur des critères d'applications potentielles ou même existantes, le financement de l'étude doit porter sur des données fondamentales qui devront être partagées (...). On a éliminé tout ce qui était trop près du développement, trop près des applications, pour éviter toute interférence, toute possibilité de conflit d'intérêts entre industriels (...). A mon avis, le seul moyen de faire cohabiter pendant un temps des industries concurrentes, c'est de remonter vers l'amont* » (coordinateur scientifique). Les entreprises peuvent ainsi s'engager sans livrer d'informations relatives à leurs procédés confidentiels. Elles préfèrent garder dans leurs murs les aspects biotechnologiques liés à la mise en œuvre des enzymes étudiées par le consortium (8). Remarquons toutefois que la séparation entre amont et aval est parfois ténue : les travaux du projet lipases ont produit des enzymes

mutantes qui ont été immédiatement brevetées.

La deuxième réponse réside dans l'organisation du consortium. Le réseau est segmenté en cinq sous-projets au sein desquels les entreprises sont réparties. Cinq contrats séparés ont été signés avec cinq groupes distincts (trois sont pilotés par une grande société associée à deux ou trois laboratoires publics, un autre regroupe cinq laboratoires industriels et six laboratoires publics, le dernier, à dominante publique, rassemble huit laboratoires académiques et une PME de biotechnologie). Cette organisation délimite des espaces de collaboration séparés et protégés de la concurrence inter-firme. Chaque industriel peut travailler sur ses propres matériels en collaboration avec les laboratoires publics sans risque de divulgation (9). Il reste toutefois à organiser les échanges entre les différents sous-projets, sinon les bénéfices de la coopération risquent d'être très réduits.

Les *guidelines* du projet lipases prévoient une diffusion des résultats par cercles concentriques qui définit plusieurs degrés d'accessibilité en fonction du temps : « *To guarantee maximum transfer of information within the entire project, a series of times periods are defined for distribution of structural and functional data obtained with the lipases studied in the project* » (*Rules and Guidelines, article 7*). Ce système différencie ainsi quatre degrés de confidentialité ou de publicité : premier cercle, les « *team data* », qui sont uniquement accessibles aux membres du sous-projet (un industriel et quelques laboratoires publics). Elles doivent être transférées dès que possible dans l'espace collectif des vingt-deux participants, au plus tard six mois après leur entrée dans le sous-projet ; deuxième cercle, les « *pooled data* » qui sont accessibles à tous les membres du consortium. Les participants sont tenus de ne pas les communiquer à des tierces parties sans l'autorisation de l'auteur ou du propriétaire des données. Les *guidelines* fixent un nouveau délai de six mois pour la

publication des *pooled data* ; troisième cercle, celui de la plateforme industrielle, qui regroupe des industriels qui ne font pas partie du consortium mais qui acquittent un droit d'accès pour bénéficier d'une information avant publication ; quatrième cercle, les « *public data* » qui sont accessibles à tous les utilisateurs potentiels, via une publication ou le dépôt des résultats dans une base de données publique.

Ce système de diffusion contrôlé introduit plusieurs nouveautés en matière d'appropriation et de circulation des résultats. Il crée en premier lieu un droit de réservation temporaire sur des connaissances qui ne sont généralement pas brevetables en l'état, ce qui donne le temps au détenteur des données d'évaluer leur valeur d'usage industrielle et de décider ou non de faire des développements pour les protéger. En tout état de cause, il bénéficie d'une avance sur les autres utilisateurs. Il organise en second lieu un espace de recherche collectif entre tous les membres du consortium. Le partage des données entre les équipes permet d'accélérer les recherches par le fait qu'une connaissance relative à une protéine A permet d'avancer dans la caractérisation d'une enzyme B. Cet accès privilégié aux données pendant une période de six mois renforce l'inté-

Le projet multilatéral du consortium européen sur les lipases associe vingt-deux participants, dont six firmes industrielles largement concurrentes et seize laboratoires publics de recherche

rets des industriels à participer au consortium plutôt que de rester en dehors. Il facilite en troisième lieu la diffusion des externalités de la recherche, grâce à

la création de la plate-forme industrielle. On observe toutefois que dans le cas du projet lipases, les industriels participants ont refusé la création de ce club d'utilisateurs qui aurait donné une information

(7) Cassier M., 1996 « Les règles de bonne conduite du projet Bridge sur les lipases : appropriation et partage des connaissances dans un réseau coopératif », colloque du CSI de l'Ecole des mines, « Coordonner, Représenter, Attribuer », Paris.

(8) C'est également la solution choisie par le consortium japonais sur le génome : « The consortium will focus on very basic work, leaving drug development in the hands of individual companies » (Science, 1995, vol 270).

(9) Cette organisation décentralisée est adoptée par de nombreux consortia, notamment au Japon (Aldrich et Sasaki, 1995).

privé aux industriels non membres, ce qui aurait réduit l'avantage compétitif qu'ils retirent de leur participation au projet. Enfin, ce système incite à une publication rapide des résultats des recherches, ce qui satisfait les universitaires participant au projet (ceux-ci sont même parvenus à négocier un délai de publication de six mois au lieu d'une année).

On observe également dans le consortium sur les lipases des pratiques d'échange des matériels biologiques, qui, au-delà des règles formulées par les participants, organisent de subtils compromis entre bien privé et bien public. Un laboratoire du CNRS et une firme de biotechnologie danoise ont produit, de manière fortuite puis concertée, des protéines mutantes. Les deux parties se sont entendues pour fabriquer deux types de mutants : des mutants publics à partir des collections du laboratoire, et des mutants privés à partir des souches industrielles. L'entreprise fabrique les deux types de mutants dans son laboratoire de biologie moléculaire. Elle communique les mutants publics au laboratoire du CNRS et elle conserve les siens dans ses murs. Elle bénéficie des observations que le laboratoire effectue sur ses mutants publics et les transpose sur ses propres protéines qui restent ainsi protégées : « *On teste l'activité des mutants de lipase pancréatique qu'ils nous envoient. Bien entendu, ils transmettent tout de suite l'information sur leur mutant de lipase*

de mucor, en parallèle, et il les testent dans leurs lessives. On a l'information en même temps et chacun l'utilise pour son usage » (responsable du laboratoire CNRS).

Le consortium européen sur la levure a quant à lui proposé un arrangement original pour conjuguer incitation individuelle et recherche collective :

il a attribué à chaque participant un droit de propriété temporaire sur la portion de génome qu'il a reçu pour étude. Ce droit de propriété est maintenu pendant la durée du programme. Pendant ce temps, le laboratoire a le droit de publier, de breveter ou de distribuer les données

qu'il a obtenues sur sa « région ». Ce droit de propriété s'éteint dès que le séquençage complet du chromosome est achevé et publié (à moins qu'il n'ait été breveté au préalable) : « *After publication of the entire sequence, each contractor automatically loses ownership of the sequence (unless previously patented)* ». Simultanément, les règles de bonne conduite prévoient pour les participants au consortium un droit d'accès limité et contrôlé aux données non publiées. Une liste de résultats résumés est mise à la disposition de chaque participant tous les trois mois par le coordinateur informatique. Soit une forme de partage des résultats au sein du collectif, pour favoriser l'avancement des travaux de chacun.

Le travail de régulation accompli par

On observe, dans le consortium sur les lipases, des pratiques d'échange des matériels, qui, au-delà des règles formulées par les participants, organisent de subtils compromis entre bien privé et bien public

les chercheurs en matière de propriété intellectuelle

Les expériences de recherche collective sont un lieu d'intense innovation institutionnelle. Si les contrats-types qui lient les participants des programmes européens à la CEE contiennent bien des clauses pour gouverner la

propriété et la confidentialité des données et des résultats des recherches, les chercheurs ont écrit des jeux de règles

pour trouver des compromis entre droit de réservation individuelle, partage des données au sein d'un collectif et publication. Dans le cas des consortiums européens sur la

levure et sur les lipases, ce sont les chercheurs qui ont négocié et rédigé les *rules et guidelines* adoptés par les participants, à l'écart des services juridiques des firmes et des organismes publics de recherche. Ce faisant, ils s'affirment comme des acteurs à part entière dans une nouvelle économie de la science. Ces accords privés entre professionnels expriment une tendance à la décentralisation de la régulation économique, particulièrement efficace pour traiter de nouveaux objets et de configurations inédites. Ces codes de bonne conduite peuvent également être interprétés comme un laboratoire du droit, comme du droit en train de se faire. Ainsi les *guidelines* du projet levure ont inspiré le code de bonne conduite du projet lipases. Certaines dispositions du *Perfect Gentleman Sequencer* pourraient enrichir les contrats-types de la CEE. La diffusion de certains modèles de règles dans les réseaux scientifiques permet d'en tester la robustesse. Dans les cas étudiés, les *guidelines* constituent une forme originale du droit d'accès aux connaissances dans un réseau, qui complète le droit de la propriété industrielle existant. Le droit de premier usage des connaissances, qui attribue un droit de réservation sur des données qui sont encore loin d'une invention brevetable, ou la catégorie de données collectives, partagées au sein d'un réseau le temps de la recherche, sont des nouveautés en matière de droit de la propriété intellectuelle. ●