Paul de Brem

► Pierre-Benoît Joly a remis en décembre 2011 un rapport au ministre de la Recherche sur les conditions du dialogue autour de la biologie de synthèse. Il est l'un des artisans de l'organisation des colloques « Sciences de la vie en société ». Il a supervisé l'étude dont il va être question, présentée par Benjamin Rimbault. ◀

Caractérisation du processus d'émergence de la biologie synthétique à partir d'une approche scientométrique

Benjamin Rimbault, Jean-Philippe Cointet, Pierre-Benoît Joly

La biologie de synthèse est un espace hétérogène dont la définition n'est pas stabilisée. Elle se situe à l'intersection de différents champs disciplinaires : chimie, biologie moléculaire, bio-informatique, et est classiquement abordée par la littérature selon trois angles [1-3] :

- une approche « *DNA based device construction* » (*biobrick*), ou approche *bottom-up*, focalisée sur l'ingénierie de circuits de gènes synthétiques ;
- une approche « *genome-driven cell engineering* », ou approche *top-down*, focalisée sur la synthèse d'un génome dans son intégralité ;
- une approche « *proto-cell creation* » qui vise à recréer des systèmes cellulaires minimaux, et en vertu de laquelle les recherches sont focalisées sur le principe de l'émergence, ouvertement présenté comme un obstacle par les tenants de la première approche.

Méthodologie de l'analyse

L'analyse scientométrique que nous avons menée avait pour objectif de caractériser le processus d'émergence *in vivo*, c'est-à-dire alors même qu'il était à l'œuvre. Pour cela, nous avons travaillé à partir des hypothèses suivantes :

- il est possible de suivre la dynamique d'une science à partir des publications scientifiques, qui permettent de retracer la morphogenèse et la dynamique du domaine considéré. Cette hypothèse communément admise est issue du concept d'inscription littéraire développé par Latour et Woolgar [4] ;

Inra/SenS - IFRIS, Institut francilien recherche, innovation, société (IFRIS), Inra, Unité de recherche 1326 Science en société (SenS), Champs-sur-Marne, 77454 Marne-la-Vallée, France.
rimbault.benjamin6@gmail.com

¹ B. Rimbault est étudiant en Master 2, J.-P. Cointet est chercheur, et P.B. Joly est directeur de l'IFRIS et de SITES. L'exposé a été présenté par B. Rimbault.

- la seconde hypothèse porte sur le fait de considérer la production scientifique comme une activité sociotechnique. Appliqué à notre étude, cela signifie que dans la constitution même de ce domaine émergent, on retrouve des interactions fortes avec l'environnement disciplinaire mais aussi avec l'environnement sociétal ;

- enfin, la dernière hypothèse considère l'émergence d'un nouveau domaine scientifique comme étant le théâtre d'une renégociation des frontières disciplinaires préexistantes.

Pour réaliser cette analyse, nous avons pris en considération les dynamiques propres à l'émergence :

- dans un premier temps, nous avons considéré que les idées, quoique nouvelles, doivent être suffisamment crédibles pour être acceptées par les autres champs scientifiques [5] ;
- dans un second temps, nous avons pris en compte le fait que tout champ émergent est soumis à une tension liée au fait de devoir accepter de très nombreux nouveaux arrivants, tout en construisant des frontières autour de normes et de pratiques pour forger une identité au domaine (*boundary work*) [6] ;
- enfin, dans un domaine en émergence comme l'est la biologie de synthèse, il faut être attentif aux « petits événements » dont les conséquences peuvent être très importantes par la suite (dépendance de sentier).

Figure 1. Croissance du réseau de collaboration et évolution du nombre de publications.

l'émergence (Figure 1). Le nombre de publications croît très rapidement. Les frontières de la biologie de synthèse sont extrêmement ouvertes puisque le pourcentage de nouveaux publiants dans le champ est de l'ordre de 75 à 80 %, ce qui est tout à fait remarquable même pour un champ en émergence. Parallèlement à cet afflux très important

de nouveaux arrivants, nous observons, à partir de 2008, la constitution d'une communauté centrale constituée d'auteurs possédant des liens privilégiés de collaboration (représentée par la courbe jaune). Malgré le taux important de nouveaux arrivants, cette communauté continue de croître, ce qui permet de conclure que ces derniers viennent renforcer la communauté centrale.

La composition des publications témoigne de la spécificité de l'émergence de la biologie de synthèse. Sur l'ensemble du corpus étudié, nous avons observé que 30 % des publications sont des revues (*reviews*) et des éditoriaux, ce qui signifie que les articles programmatiques occupent une place importante. Nous avons également établi la carte de la distribution des publications et collaborations existantes au niveau international (Figure 2). Le pôle fort de la biologie de synthèse se situe essentiellement aux États-Unis, mais l'Europe possède également un foyer assez actif. En Asie, le foyer est anecdotique mais affiche un taux de croissance très élevé.

de nouveaux arrivants, nous observons, à partir de 2008, la constitution d'une communauté centrale constituée d'auteurs possédant des liens privilégiés de collaboration (représentée par la courbe jaune). Malgré le taux important de nouveaux arrivants, cette communauté continue de croître, ce qui permet de conclure que ces derniers viennent renforcer la communauté centrale.

La composition des publications témoigne de la spécificité de l'émergence de la biologie de synthèse. Sur l'ensemble du corpus étudié, nous avons observé que 30 % des publications sont des revues (*reviews*) et des éditoriaux, ce qui signifie que les articles programmatiques occupent une place importante. Nous avons également établi la carte de la distribution des publications et collaborations existantes au niveau international (Figure 2). Le pôle fort de la biologie de synthèse se situe essentiellement aux États-Unis, mais l'Europe possède également un foyer assez actif. En Asie, le foyer est anecdotique mais affiche un taux de croissance très élevé.

Analyse des réseaux d'émergence de la biologie de synthèse

Représentation des réseaux

Le cœur de notre étude se trouve dans l'analyse des réseaux qui ont été cartographiés à l'aide de la plateforme CorText. La première carte (Figure 3) a été réalisée en prenant les 100 articles les plus cités du corpus qui ont été mis en réseau par la plate-forme CorText. Sur la carte ainsi obtenue, chaque article est symbolisé par un petit triangle dont l'aire est proportionnelle au nombre de citations reçues. Sur la carte, deux articles

Analyse scientométrique de l'émergence

Analyse statistique

L'analyse scientométrique commence par le traitement de statistiques descriptives afin de mettre en évidence des indicateurs de

Figure 2. Carte des publications par nationalité. La surface des cercles correspond au nombre de publications par pays. Plus la couleur des cercles est foncée et plus le taux de croissance sur les deux dernières années est élevé.

les trois derniers *clusters* de l'école *biobricks*, sur 47 publications, 31 sont signées par quatre chercheurs en position de dernier auteur (Keasling, Weiss, Voigt et Collins). Le profil de ces chercheurs – dont l'analyse scientométrique établit la centralité – est illustratif de la pluralité disciplinaire de la biologie

de synthèse : Ron Weiss (MIT, Boston, Ma) est un ingénieur informaticien, Chris Voigt (MIT, Boston, Ma) est un ingénieur chimiste, Jay D. Keasling (Berkeley, Ca) est chimiste et Jim Collins (Harvard University, Boston, Ma) est bio-ingénieur.

Représentation radar

Nous avons établi le même type de carte sous la forme d'une représentation radar (Figure 4), ce qui permet de suivre la dynamique de l'apparition des *clusters*, des écoles, et de replacer les articles selon leur année de parution. Nous pouvons ainsi revenir aux origines des trois écoles.

L'école *top-down*, focalisée sur les travaux de synthèse de génomes, semble avoir pour origine un article de Fraser en 1995 [10], cosigné par Craig Venter, présentant le séquençage complet du génome de *Mycoplasma genitalium*, qui sera ensuite synthétisé par le Craig Venter Institute à la fin des années 2000. Le développement de cette école se poursuit ensuite, de 2002 à 2010, et l'on retrouve les étapes clés du projet de Craig Venter, de la synthèse de la première bactérie au génome entièrement synthétique. On voit ainsi successivement apparaître les travaux de Glass en 2006 [11] qui visent à identifier un génome minimal, les travaux de Lartigue en 2007 sur la transplantation du génome [12], et, en 2008, les travaux de Gibson sur la synthèse du génome intégral [13], pour arriver à l'expérience finale en 2010 avec la synthèse du génome entièrement synthétique [14].

Nous avons également identifié les références historiques de l'école *Protocell*. Elles commencent avec l'article de Szostak en 2001 [15], et se terminent avec un article de Luisi en 2006 sur la question de l'origine de la vie [16]. Cela montre une profonde asymétrie entre les différentes écoles. Depuis 2006, aucune publication de l'école *Protocell* n'est entrée dans les 100 publications les plus citées.

sont d'autant plus proches qu'ils ont tendance à être cités dans un même article. CorText permet de passer à une étape de groupement en détectant automatiquement certains groupes de référence, ou *clusters*, au sein desquels existe une structure cohésive.

L'interprétation de ces cartes, après lecture des articles les plus cités et de la totalité des *abstracts*, et après l'analyse des auteurs, montre que l'on retrouve les trois écoles mentionnées précédemment : l'approche des *proto-cells* (articles du cluster *minimal cell/sequence* avec des auteurs comme Luisi ou Szostak, voir ci-dessous), les travaux tournés vers une approche du génome intégral (articles du cluster *synthetic genome* avec un auteur comme Craig Venter), et la dernière école, qui est la plus importante, celle des *biobricks*. Cette dernière comporte cinq *clusters* avec des thématiques propres :

- Le *cluster* que nous avons intitulé *stochasticity and cellular noise*, dont les articles abordent les thématiques liées à la stochasticité, d'un point de vue très mathématique ou lié à des fonctionnements biologiques et au bruit cellulaire ;
- le *cluster design/programmatic discourse*, qui comporte les articles fondateurs de l'approche des circuits de gènes synthétiques ou de réseaux de gènes (le *toggle switch* de Gartner et Collins [7] et le *repressilator* de Elowitz et Leibler [8], mais également d'importants discours programmatiques, notamment l'article de Drew Endy [9] ;
- le *cluster* annoté *mammalian cell* qui regroupe des travaux portant sur les cellules de mammifères ;
- le *cluster control*, dont les publications traitent principalement des questions de contrôle de la régulation (notamment à partir de molécules d'ARN) ;
- enfin le *cluster applications*, où l'on trouve des travaux avec des réseaux de gènes de plus en plus complexes dans un but annoncé d'applications.

Cette carte permet en outre d'identifier les *principal investigators* (PI). Certains spécialistes nous ont fait remarquer que la carte ne faisait apparaître que les premiers auteurs. Or en biologie, il est important de souligner que c'est le dernier auteur qui fait souvent autorité. Par exemple, dans le cluster *synthetic genome*, Craig Venter est dernier auteur de plus de la moitié des articles. De même, dans

Figure 3. Carte de cocitation des références. Chaque école a été manuellement étiquetée. Les articles coécrits par les PI (*principal investigators*) et listés ci-dessus sont colorés en conséquence.

L'origine de l'approche des biobricks est buissonnante et peut être repérée à la fois dans des travaux issus de la modélisation (de nature profondément mathématique) et traitant de la stochasticité, et dans des travaux sur la synthèse des premiers circuits synthétiques (*toggle switch*, *repressilator*) dans les années 2000. Elle se diversifie par la suite avec les *clusters* abordant les thématiques de contrôle (*cluster control*), notamment pour diminuer l'influence de processus biologiques émergents, venant perturber le fonctionnement des circuits ingénierés, les travaux sur les cellules de mammifères (*cluster mammalian cell*) et le *cluster* le plus récent, dirigé par une volonté d'application (*cluster application*).

Cette carte montre nettement la visibilité bien moindre de l'approche des *Protocells*, qui a tendance à se stabiliser puis à n'être plus visible à partir de 2006, et la tendance du *cluster* des applications, né en 2004, à fortement croître depuis. Il est intéressant de rapprocher cette observation des propos de Luis Campos, qui fixe la date de l'apparition de la biologie de synthèse moderne en 2004, avec l'implication des ingénieurs dans ce domaine. Cela correspond à la naissance de ce *cluster* très porté sur l'application.

Figure 4. Carte de co-citation de références (représentation radar).

Représentation lexicale

Nous avons également travaillé à partir d'un troisième type de carte : la carte lexicale (Figure 5). Nous avons pour cela utilisé CorText pour identifier dans les titres, abstracts ou mots clés des publications du corpus des groupes de mots spécifiques à la biologie de synthèse.

La carte se lit de la même manière que la précédente. Elle permet de détecter les communautés et d'isoler les *clusters*. Nous avons ainsi identifié un *cluster* « *stochasticity* », avec des termes parlants comme *molecular noise*, ou *non linear stochastic*. Un *cluster* traite des problèmes de régulation, de contrôle et de modélisation, avec des termes tels que *positive feedback*. On trouve également des regroupements originaux, à l'image du *cluster* « *biological parts* », qui associe des termes comme *registry of standard biological parts* rappelant l'approche « *biobrick* ».

Le *cluster* le plus intéressant reste celui que nous avons dénommé *cluster* « *Promises* ». Il rassemble des termes tels que *drug discovery*,

biotechnology industry ou *climate change*. Ce dernier terme, surprenant dans ce contexte, montre que même dans un espace comme celui de la publication scientifique, de grandes préoccupations et des enjeux globaux sont mobilisés. Il est également intéressant de voir que les termes de *synthetic biology application* se distinguent de ceux de *potential application in synthetic biology*. Notre méthode permet de distinguer les applications potentielles (*cluster promises*) des applications réelles (*molecular biology applications*). Ces deux *clusters* sont reliés par le terme *metabolic pathway*, ce qui illustre l'importance de l'ingénierie de voies métaboliques pour les applications potentielles ou réelles.

Le rôle de l'entrepreneur institutionnel

Revenons maintenant à la Figure 3 et à ce qui est l'un des points fondamentaux de notre étude, c'est-à-dire

le rôle de l'entrepreneur institutionnel. Il est assez étonnant de ne pas avoir entendu davantage parler depuis le début de la journée de Craig Venter et de Drew Endy. Cette carte montre en effet à quel point ces deux personnages, extrêmement médiatiques et controversés, ont joué un rôle structurant dans l'émergence de la biologie synthétique.

Nous avons déjà présenté le *cluster* contenant les publications de travaux majoritairement menés par Craig Venter et ses collaborateurs au sein du *Craig Venter Institute*¹. Le profil même de ce *cluster* correspond aux types de recherches et de collaborations menées au sein du *Craig Venter Institute*. Les recherches encadrées par Craig Venter sont en effet très orientées vers l'obtention de brevets et très liées à l'industrie, comme l'illustre le partenariat de l'ordre de \$600 millions passé avec Exxon Mobil en 2009, ce qui correspond à un modèle de développement assez classique dans le monde des biotechnologies.

Parallèlement, nous voulions montrer l'importance de Drew Endy dans la constitution de l'école dite *Biobricks* et donc du domaine en général. Nous sommes revenus sur le manque de diversité des investigateurs principaux dans les articles les plus récents de cette école. Nous nous sommes tout particulièrement intéressés à l'activité institutionnelle de Drew Endy dont le fameux credo est « *To make biology easier to engineer* ». En effet, entre 2004 et 2009, Drew Endy a mis en place ou dirigé un grand nombre d'institutions, notamment l'iGEM, qu'il a fondée, la *Biobrick Foundation*, qu'il a fondée et dont il est le président, le *BioFab*, qu'il a fondé, et Synberc, programme américain associant quatre universités, dont il est directeur stratégique.

Afin de montrer le rôle structurant de ces institutions, nous avons encadré en noir dans la *Figure 3* les triangles correspondant aux articles dont le dernier auteur est affilié à l'une de ces quatre institutions. On constate alors que 60 % de l'ensemble des 100 articles les plus cités sont concernés, et surtout qu'au sein des *clusters* les plus récents, quasiment l'ensemble des articles répondent à ce critère. Dans la totalité des cas, le directeur de recherche fait partie des institutions fondées ou dirigées par Drew Endy.

Conclusion

Notre étude permet de montrer que la biologie de synthèse est un domaine en émergence, soumis à la tension liée à l'arrivée de nouveaux entrants et à l'apparition de frontières. Nous avons mis en évidence le fait que ces frontières pouvaient être épistémiques et qu'il existait une forte interaction avec la société. Nous avons également montré que les promesses jouaient un rôle très structurant dans l'émergence de la biologie de synthèse, au même titre que les personnages porteurs d'une vision, à l'image de Drew Endy.

¹ Le premier fut établi à Rockville, MD, États-Unis, et un second existe à La Jolla, San Diego, California (États-Unis).

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. O'Malley A. Towards evidence based emergency medicine: best BETs from the Manchester Royal Infirmary. BET 4. Is a normal C-reactive protein sensitive enough to rule out septic arthritis. *Emerg Med J* 2008 ; 25 : 767-8.
2. Malaterre V, Ogorka J, Loggia N, Gurny R. Oral osmotically driven systems: 30 years of development and clinical use. *Eur J Pharm Biopharm* 2009 ; 73 : 311-23.
3. Heams T, Bed'hom B, Rebours E, et al. Insights into gene expression profiling of natural resistance to coccidiosis in contrasting chicken lines. *BMC Proc* 2011 ; 5 (suppl 4) : S26.
4. Latour B, Woolgar S. *Laboratory life: the construction of scientific facts*. Princeton, NJ : Princeton University Press, 1986.
5. Frickel S, Gross N. A general theory of scientific/intellectual movements. *Am Soc Rev* 2005 ; 70 : 204-32.
6. Gieryn TF. Boundary-work and the demarcation of science from non-science: strains and interests in professional ideologies of scientists. *Am Soc Rev* 1983 ; 48 : 781-95.
7. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 2000 ; 403 : 339-42.
8. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature* 2000 ; 403 : 335-8.
9. Endy D. Foundations for engineering biology. *Nature* 2005 ; 438 : 449-53.
10. Fraser CM, Gocayne JD, White O, et al. The minimal gene complement of *Mycoplasma genitalium*. *Science* 1995 ; 270 : 397-403.
11. Glass JI, Assad-Garcia N, Alperovich N, et al. Essential genes of a minimal bacterium. *Proc Natl Acad Sci USA* 2006 ; 103 : 425-30.
12. Lartigue C, Glass JI, Alperovich N, et al. Genome transplantation in bacteria: changing one species to another. *Science* 2007 ; 317 : 632-8.
13. Gibson DG, Benders GA, Axelrod KC, et al. One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc Natl Acad Sci USA* 2008 ; 105 : 20404-9.
14. Gibson DG, Glass JI, Lartigue C, et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 2010 ; 329 : 52-6.
15. Szostak JW, Bartel DP, Luisi PL. Synthesizing life. *Nature* 2001 ; 409 : 387-90.
16. Luisi PL. *The emergence of life: from chemical origins to synthetic biology*. Cambridge : Cambridge University Press, 2006.

TIRÉS À PART

B. Raimbault

DISCUSSION

Paul de Brem

Je propose à Pierre Monsan et Jacques Haiech de nous rejoindre pour cette discussion. Ces travaux de scientométrie sont fascinants. Ils montrent que sur les 100 articles les plus cités, 60 sont liés directement ou indirectement à Drew Endy, et qu'il existe également un véritable pôle autour de Craig Venter.

Pierre Monsan

Je suis fasciné par la puissance de cette analyse. Je serai tout à fait preneur de cet article lorsqu'il sera publié.

Morgan Meyer

Vous avez évoqué l'apparition de frontières. Pour Thomas Gieryn², la frontière renvoie à l'expulsion, à l'expansion et à la protection de l'autonomie. Vous montrez plutôt l'apparition de nouvelles frontières. Mais quelle est la matérialité de ces frontières ?

Ma deuxième question renvoie aux propos de Thomas Heams ce matin sur la notion d'émergence. Vos cartes sont explicites, mais je m'interroge sur le contexte politique, économique, financier et culturel qui fait émerger une science. Cette notion d'émergence peut également être critiquée car il n'est pas toujours facile de la caractériser.

Pierre-Benoît Joly

Notre référence aux travaux de Thomas Gieryn est assez indirecte et nous nous inspirons principalement de son concept de *boundary work*. L'idée est que les frontières entre les disciplines ne sont pas des données, mais qu'elles doivent être expliquées, particulièrement dans le cadre d'un processus d'émergence. La théorie que nous développons dans l'article concerne la qualité d'une frontière dans le contexte de l'émergence, car même en fin de période, il apparaît 75 % de nouveaux auteurs chaque année. La constitution d'un nouveau champ se fait par attraction d'auteurs qui s'intéressent à un domaine. Mais le champ doit également posséder une consistance, laquelle passe par l'élaboration de normes permettant de distinguer ce qui relève de la biologie de synthèse et ce qui lui est étranger. Il nous a semblé intéressant de pointer ce caractère contradictoire des frontières qui doivent être à la fois perméables et bien marquées. Cela passe par des définitions scientifiques, mais aussi par des normes de production dans lesquelles les membres de la communauté se retrouvent. Les entrepreneurs institutionnels jouent à ce titre un rôle très important.

Concernant la seconde question, nous n'avons pas mené d'analyse systématique du lien entre la publication scientifique et les instruments institutionnels, les programmes qui soutiennent les recherches, ou bien du lien avec les utilisateurs potentiels, notamment avec les industriels. Il est possible d'approfondir le sujet avec les outils scientométriques, mais il s'agit de techniques particulières que nous n'avons pas encore utilisées et qui s'inscrivent dans le prolongement de nos travaux.

Pierre Monsan

Assiste-t-on réellement à l'émergence d'une nouvelle science ou bien constate-t-on l'arrivée à maturité d'une science qui débouche sur une technologie ?

Jacques Haiech

Je sais ce qu'est une discipline mais j'ignore ce qu'est une science. La science se définit selon moi par une question générale qui est « comment ça marche ? ». La discipline se définit par la création d'un club, l'apparition d'un réseau. Il s'agit d'un réseau hiérarchique, sans échelle, semblable aux réseaux sociaux, avec des personnalités plus connectées que d'autres.

De nouvelles façons d'analyser le vivant apparaissent. Je ne suis pas certain qu'elles soient issues simplement d'entrepreneurs institutionnels. Si l'on

avait posé la même question à l'époque d'Apple 2 contre IBM, et si l'on s'était fondé sur les seules publications, Steve Jobs aurait été totalement invisible. Dans le domaine de la biologie de synthèse, le DIY bio, iGEM, et un certain nombre de philosophies *open source* sont invisibles au niveau des publications. Il serait intéressant d'utiliser l'outil scientométrique à partir des rapports iGEM pour déterminer si ce n'est pas à ce niveau qu'apparaît l'émergence ou s'il existe des idées différentes de celles des entrepreneurs institutionnels.

Benjamin Rimbault

Le concours iGEM donne lieu à de la biologie de synthèse, et je rappelle que son fondateur est Drew Endy. En outre, DIY bio se revendique de l'*open source* tout comme Drew Endy le revendique également. Dans les publications scientifiques, on retrouve cette littérature grise qui naît en marge des institutions. Il existe certainement des biais dans la constitution du corpus, mais l'influence de l'iGEM est bel est bien présente.

Pierre-Benoît Joly

Nous nous basons sur la littérature concernant l'émergence des champs scientifiques. La question est moins de savoir s'il s'agit d'une nouvelle science que d'un nouveau champ. La différence porte sur des éléments qui permettent de distinguer un nouveau champ. Ces éléments relèvent effectivement de la densité des relations entre chercheurs, mais également des structurations de la communauté du champ (écoles, revues, associations spécialisées, etc.). Un article de Mullins sur la constitution de la biologie moléculaire a montré en 1972 [1] comment, à partir d'un certain stade, on pouvait parler d'un champ possédant une relative autonomie.

De la salle

Ma question est méthodologique et concerne l'apparition de nouveaux auteurs, qui contraste avec le nombre relativement faible de PI (*principal investigator*) restant dans le champ. S'agit-il d'un artéfact du fonctionnement des laboratoires ? Ne peut-on pas le contrôler par comparaison avec un champ qui ne serait pas caractérisé comme en émergence ?

Jean-Philippe Cointet

Nous nous sommes posé la même question. Nous avons effectivement fait la comparaison. Nous avons calculé exactement les mêmes statistiques - taux de nouveaux entrants, taille du cœur structurel dans lequel l'ensemble des auteurs sont connectés les uns aux autres - sur un *corpus* définissant la communauté des biologistes du développement spécialisés dans le poisson zébré, un animal modèle, et nous avons constaté qu'à taux de maturité égal,

² Rudy professor of sociology and adjunct professor in the department of history and philosophy of science, Indiana University.

c'est-à-dire à nombre de publications égal, le taux de nouveaux entrants était de 60 % chez les biologistes du développement, contre 75 % pour la biologie de synthèse. Actuellement, la part de chercheurs appartenant à la plus grande composante connexe dans la communauté du poisson zébré dépasse 80 %, nombre à mettre en perspective avec la communauté de la biologie de synthèse encore en cours de maturation.

Jean Weissenbach

Le génie génétique est apparu il y a 30 ou 40 ans. Le phénomène était à peu près similaire : il tournait autour de l'ingénierie. Il a donné lieu aux mêmes discussions sur l'apparition d'un nouveau domaine. Or finalement, le génie génétique n'est jamais devenu un champ. Il s'agit d'une technologie appliquée à la fois à des projets d'intérêt fondamental en tant qu'outil d'exploration, et à un certain nombre d'applications. Il serait peut-être intéressant d'établir une comparaison avec les premières expériences de recombinaison *in vitro* dans les années 1970.

Pierre-Benoît Joly

La biologie synthétique est-elle un champ scientifique ou bien relève-t-elle plutôt de la technologie et des outils ? Nous n'avons pas les moyens de répondre à cette question, mais il est clair qu'il s'agit d'une hypothèse à prendre en compte. Un champ se définit autour d'un certain nombre de questions fondamentales en biologie. Dans le même temps, l'applicatif joue un rôle très important dans la structuration du champ. Nous avons tendance à penser que ce champ se caractérise par son caractère hybride, portant à la fois sur des questions fondamentales, y compris l'origine de la vie, le génome minimal, etc., et sur des aspects applicatifs. J'ignore si nous pouvons trancher cette question à notre niveau. Certains chercheurs identifiés par notre analyse relèvent sans aucun doute de la recherche fondamentale.

Jean Weissenbach

Je ne veux pas mettre en avant l'opposition entre fondamental et appliqué mais l'analogie avec l'émergence du génie génétique, qui n'est pas devenu une discipline.

Jacques Haiech

Il se peut que le génie génétique ne soit pas devenu une discipline parce qu'il ne s'est pas figé. Si l'on définit un champ comme un domaine prenant une place dans l'histoire de la science et continuant à se reproduire, le génie génétique ne correspond pas à cette définition. Mais sans le génie génétique, il n'y aurait pas eu les « -omiques » voire la biologie de synthèse. Il faudrait définir quelque chose relevant davantage d'un flux que de la création d'un « lac ».

Pierre Monsan

Ne s'agit-il pas justement de la convergence de nombreux ruisseaux qui se rejoignent pour former un fleuve ?

Morgan Meyer

On peut tout de même parler de discipline dans la mesure où elle donne lieu à de nombreux *Masters*, à des revues contenant « *synthetic biology* » dans leur titre, et il existe une conférence sur la biologie de synthèse depuis 2004.

La biologie de synthèse est en cours d'institutionnalisation et on peut parler d'un champ en cours de création.

Paul de Brem

Vous m'avez indiqué que certains résultats de vos travaux étaient en contradiction avec des résultats préalablement publiés.

Pierre-Benoît Joly

J'ai effectivement indiqué qu'en partant à peu près des mêmes données, nous observons des phénomènes que les auteurs qui avaient travaillé sur ces cartes n'avaient pas du tout perçus, parce que nous disposons d'algorithmes qu'ils n'ont pas. En dépit d'un champ très fluide, avec un grand nombre de nouveaux entrants et une très forte hétérogénéité, les analyses de dépendance montrent qu'il existe en réalité deux pôles extrêmement importants dans la structuration du champ : le pôle Craig Venter et le pôle constitué autour de Drew Endy, fonctionnant sur des bases assez différentes. Ce résultat est relativement robuste. Un article récent montre très clairement qu'il existe deux façons de penser la propriété intellectuelle en matière de biologie de synthèse, l'une assez traditionnelle qui est celle de Craig Venter, et l'autre plutôt basée sur la métaphore *open source*, qui est celle de Drew Endy, avec un fonctionnement communautaire et une structuration très différente. Ce caractère quelque peu bipolaire est original.

Paul de Brem

Pouvez-vous rappeler en quoi consiste la plate-forme CorText ?

Jean-Philippe Cointet

CorText est une plate-forme de l'IFRIS, à la fois infrastructure technique, collectif de chercheurs en linguistique, sciences des systèmes complexes et modélisation et d'ingénieurs web. Son objectif consiste à développer des outils et des méthodologies d'analyse de grands *corpus* textuels, qui ne se limitent pas nécessairement à des *corpus* scientométriques. L'enjeu est d'accompagner des projets en sciences sociales au sein de l'IFRIS pour produire des connaissances sur les dynamiques sociosémantiques à partir de grands *corpus* de données définis par des chercheurs de l'IFRIS. L'ambition consiste à mêler cette approche méthodologique provenant des sciences dures à des questionnements venus, eux, des sciences sociales. C'est ce qui constitue l'originalité de la démarche. ♦

RÉFÉRENCE

1. Mullins NS. The development of a scientific specialty. The phage group and the origin of molecular biology. *Minerva* 1972 ; 19 : 52-82.