

Modélisation de réseaux biologiques et médecine de précision en oncologie

Biological network modelling and precision medicine in oncology

Laurence Calzone^{1,2,3*}, Inna Kuperstein^{1,2,3*}, David Cohen^{1,2,3}, Luca Grieco^{1,2,3,4,5,6,7}, Eric Bonnet^{1,2,3}, Nicolas Servant^{1,2,3}, Philippe Hupé^{1,2,3,8}, Andrei Zinovyev^{1,2,3*}, Emmanuel Barillot^{1,2,3*}

¹ Institut Curie, 26 rue d'Ulm, 75248 Paris, France

² INSERM U900, 75248 Paris, France

³ Mines ParisTech, 77300 Fontainebleau, France

⁴ Ecole Normale Supérieure, IBENS, 46 rue d'Ulm, Paris, France

⁵ CNRS UMR8197, 75005 Paris, France

⁶ INSERM U1024, 75005 Paris, France

⁷ Present address : Clinical Operational Research Unit, University College London, London, UK

⁸ CNRS UMR144, 75248 Paris, France

< Emmanuel.Barillot@curie.fr > < Andrei.Zinovyev@curie.fr >

*Contributions égales

Pour citer cet article : Calzone L, Kuperstein I, Cohen D, Grieco L, Bonnet E, Servant N, Hupé P, Zinovyev A, Barillot E. Modélisation de réseaux biologiques et médecine de précision en oncologie. *Bull Cancer* 2014 ; 101 (suppl. 1) : S18-S21.

doi : 10.1684/bdc.2014.1973

*Médecine
personnalisée,
médecine
de précision :
réalités
et perspectives*

Résumé. La médecine de précision en oncologie devient aujourd'hui réalité grâce au séquençage haut débit des tumeurs et au développement d'inhibiteurs ciblés, qui ouvrent la voie à des thérapies sur mesure. Beaucoup d'essais cliniques basent leur stratégie sur l'identification des mutations de la tumeur pour choisir l'inhibiteur ciblé qui contrecarrera vraisemblablement l'effet du gène muté. Des résultats récents ont montré que cette stratégie centrée sur les gènes peut être couronnée de succès, mais qu'elle peut aussi s'avérer infructueuse pour stopper la prolifération. La raison en est l'existence de mécanismes de compensation, de « cross-talks », et de boucles de rétroaction qui permettent à la cellule tumorale d'échapper au traitement. La prise en compte du réseau de régulation est nécessaire pour établir quel inhibiteur ou quelle combinaison d'inhibiteurs aboutiraient à la meilleure réponse thérapeutique. La modélisation mathématique des réseaux biologiques, ainsi que des bases de données de réseaux biologiques de haute qualité, regroupant notre connaissance des circuits moléculaires des cellules normale et tumorale, portent l'espoir d'un futur meilleur pour la médecine de précision en oncologie. ▲

Mots-clés : biologie des systèmes, bases de données de voies de signalisation, profils moléculaires de tumeur, analyse génomique basée sur les réseaux, inhibiteur ciblé.

Abstract. Precision medicine in oncology is becoming reality thanks to the next-generation sequencing of tumours and the development of targeted inhibitors enabling tailored therapies. Many clinical trials base their strategy on the identification of mutations to deliver the targeted inhibitor that counteract supposedly the effect of a mutated gene. Recent results have shown that this gene-centered strategy can be successful, but can also fall short in stopping progression. This is due to the many compensation mechanisms, cross-talks and feedback loops that enable the tumoral cell to escape treatment. Taking into account the regulatory network is necessary to establish which inhibitor or combination of inhibitors would achieve the best therapeutic results. Mathematical modelling of biological networks, together with high-quality pathway databases collecting our knowledge of the molecular circuitry of normal and tumoral cells, hold the hopes of an enhanced future for precision medicine in oncology. ▲

Key words: systems biology, pathway databases, molecular profiles of tumour, network-based genomic analysis, targeted inhibitor.

Médecine de précision en oncologie

Le terme médecine de précision a été introduit récemment pour préciser celui de médecine personnalisée ou de médecine guidée par la génomique. Dans cet article, nous en adoptons la définition proposée par Garraway [1] en oncologie, et qui retient trois critères, en plus de l'adaptation au patient individuel : une résolution moléculaire accrue, une clarté mécanistique et une puissance thérapeutique. En pratique, on commence par établir les profils moléculaires d'un échantillon tumoral du patient à l'aide des technologies de la génomique, de séquençage haut débit (SHD) et/ou de puces à ADN, le plus souvent afin d'en déterminer les mutations ou l'expression de certains gènes. L'approche s'appuie typiquement sur l'identification de mutations connues, dont on connaît éventuellement l'effet fonctionnel (clarté mécanistique). Par exemple, on peut citer le cas de la mutation activatrice d'une kinase contrôlant une voie de prolifération, comme le gène *BRAF* dans le mélanome [2]. Une fois les mutations identifiées, des stratégies thérapeutiques ciblées utilisant un inhibiteur spécifique du gène, voire de la mutation identifiée, comme le vémurafénib dans l'exemple ci-dessus, peuvent être mises en œuvre. De nombreux inhibiteurs ciblant des oncogènes connus sont actuellement en développement dans l'industrie pharmaceutique, conférant à la médecine de précision en oncologie un potentiel thérapeutique croissant.

Essais cliniques de médecine de précision

Le principe de la médecine de précision en oncologie est actuellement mis en pratique dans de nombreux centres de soins, en particulier dans le cadre d'essais cliniques [3-5] dans lesquels sont inclus des patients aux tumeurs réfractaires, toutes localisations confondues. Le plus souvent, l'investigation explore les mutations ponctuelles ou les amplifications de gènes connus pour être impliqués dans le cancer (par SHD et éventuellement puces, suivies d'une confirmation en immuno-histochimie) et pour lequel un inhibiteur est disponible afin de déterminer le traitement, et le succès est jugé à l'aune de la survie sans progression. C'est le cas par exemple de l'essai SHIVA [6] coordonné par l'Institut Curie et impliquant huit partenaires français, qui a démarré en octobre 2012 et a déjà inclus 700 patients, toutes localisations confondues.

On ne sait pas encore dans quel cas cette stratégie, qui ignore l'organe et privilégie le génome, s'avèrera payante, ni dans quel cas elle se montrera infructueuse. Sa logique repose sur de francs succès, comme l'utilisation de l'imatinib dans le traitement de la leucémie myéloïde chronique [7]. Néanmoins, certains résultats ont aussi montré que considérer les mutations individuelles et non leurs interactions, et en ignorer le contexte de localisation, pouvait aussi conduire à l'échec : par exemple, si le vémurafénib, inhibiteur spécifique de *BRAF*, est efficace dans le traitement des mélanomes montrant la mutation V600E, il ne l'est absolument pas dans celui des carcinomes du colon affectés par la même mutation [8]. La raison en est une boucle de rétroaction négative de *BRAF*

vers *EGFR*, que l'inhibition de *BRAF* invalide, qui est sans effet dans le cas du mélanome où l'expression de *EGFR* est très faible, contrairement au colon où la prolifération cellulaire s'en trouve accrue. En d'autres termes, un même gène peut avoir des activités oncogénique ou suppresseur de tumeur selon le contexte, et même concomitamment. Plus encore, on sait que, souvent, la cellule soumise à un inhibiteur spécifique sensé interrompre une voie de signalisation finit par trouver une voie d'échappement, grâce à un mécanisme de compensation ou de rétroaction.

Les exemples ci-dessus montrent que la médecine de précision en oncologie ne peut pas se limiter à considérer des mutations individuelles, mais doit les placer dans le contexte des réseaux de signalisation qui gouvernent la tumorigenèse et la progression tumorale. Une modélisation plus réaliste et plus complexe doit être utilisée pour orienter le choix thérapeutique, et cette modélisation doit intégrer la connaissance du réseau biologique. Par ailleurs, les études de profils moléculaires de séries de tumeurs, qu'elles soient à visée cognitive ou clinique, ont fleuri depuis une douzaine d'années, comme par exemple les projets internationaux TCGA (<http://cancer-genome.nih.gov/>) et ICGC (www.icgc.org). Les études d'interprétation de ces profils ont donné lieu au développement de méthodes d'analyse intégrant les réseaux biologiques. Ces méthodes recherchent typiquement les gènes et régions du réseau différemment exprimés, ou particulièrement mutés entre deux conditions, par exemple deux types tumoraux, ou bien un type tumoral et un tissu sain. Elles s'appuient sur des bases de données d'interactions entre protéines ou de voies moléculaires, comme par exemple STRING [9], HPRD [10], PathwayCommons [11], KEGG [12], SPIKE [13], Reactome [14] ou ACSN (<https://acsn.curie.fr>). Ces méthodes se divisent en trois différents types principaux, qui exploitent respectivement les notions de module, de voisinage dans le réseau, et de dynamique du réseau (figure 1).

Analyse par modules du réseau biologique

L'approche la plus simple pour dépasser les études gène par gène consiste à agréger les gènes par modules. Ceux-ci peuvent être définis comme des sous-parties du réseau global correspondant à des voies biologiques canoniques décrites dans la littérature, ou bien ils peuvent être construits par voisinage dans le réseau. Pour chaque module, on détermine s'il contient plus de gènes mutés (ou si le niveau d'expression de ses gènes est plus altéré) que le hasard ne le permettrait. La méthode s'appuie donc sur une statistique prenant en compte le comportement de chaque gène et sur la définition d'ensembles de gènes fonctionnellement liés, pour en inférer les modules statistiquement impliqués dans le phénomène étudié. Cette méthode a le mérite d'être simple, mais ne tire pas parti de la structure du réseau car elle ne considère que des ensembles de gènes et ignore la topologie du réseau d'interactions. Le principe a été introduit dans l'algorithme GSEA [15], mais aussi par des tests hypergéométriques classiques.

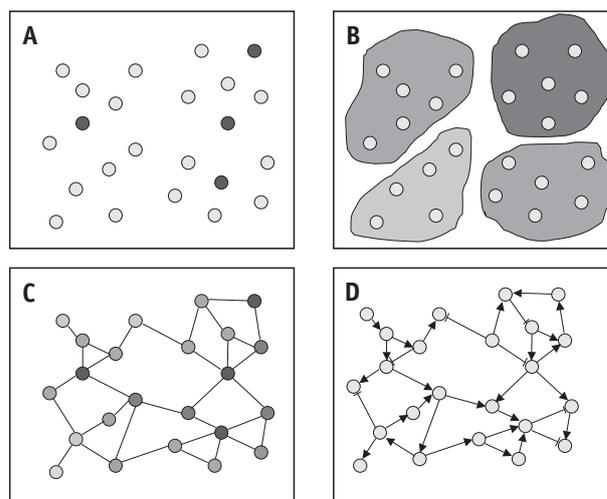
Figure 1. Différentes méthodes d'analyses des profils moléculaires de tumeurs, à complexité et présupposés biologiques croissants (voir texte).

A. Les gènes mutés ou différemment exprimés (en foncé) sont considérés, sans prendre en compte le contexte d'interaction.

B. L'analyse est conduite au niveau des modules (ensembles de gènes fonctionnellement reliés), les plus foncés étant statistiquement impliqués dans la pathologie.

C. La topologie du réseau est utilisée pour déterminer les régions impliquées dans la pathologie en supposant que les voisins de gènes mutés sont aussi impliqués ; la méthode est utilisée pour proposer des cibles thérapeutiques ou établir des signatures moléculaires de pathologie.

D. La modélisation dynamique du réseau s'appuie sur un modèle de régulation de chaque gène (les flèches symbolisent l'activation, les lignes barrées l'inhibition) et permet de prédire l'effet de perturbations comme le traitement par un inhibiteur ciblé.



Analyse du voisinage dans le réseau biologique

D'autres méthodes utilisent explicitement la topologie du réseau, et supposent donc une clarté mécanistique supérieure, c'est-à-dire des bases de données plus élaborées que les seuls ensembles de gènes : par exemple KEGG, PathwayCommons, Reactome ou ACSN. La première méthode adopte le principe de la culpabilité par association : sont considérés comme importants les gènes qui, sans être forcément eux-mêmes différemment exprimés ou mutés fréquemment, sont liés à de tels gènes dans le réseau, *a fortiori* s'ils partagent avec eux beaucoup de connexions. En effet, ces gènes coupables par association sont des cibles thérapeutiques potentielles particulièrement intéressantes en ce sens qu'ils touchent beaucoup de points de dérégulations. À ce jour, cette approche est surtout utilisée pour identifier de nouvelles cibles thérapeutiques ou des signatures pronostiques [16-19]. Pour autant, à notre connaissance, aucun agent ciblé n'a encore été développé à partir de cette approche, et la médecine de précision ne l'exploite donc pas encore.

La méthode de propagation de réseau généralise le principe précédent : l'idée directrice est que la mutation d'un gène (ou sa sur-expression) impacte ses voisins immédiats mais aussi plus lointains de façon décroissante avec l'éloignement dans le réseau. Récemment, cette approche a permis de proposer une stratification des tumeurs, en utilisant les bases de données STRING, PathwayCommons et HumanNet pathway [20].

Analyse par modélisation dynamique du réseau biologique

Les méthodes ci-dessus, bien qu'elles considèrent la topologie du réseau, n'intègrent pas les mécanismes de régulation proprement dits. Une approche plus ambitieuse s'appuie sur la modélisation mathématique des relations de régulation entre protéines, sur la base d'un système d'équations, par exemple différentielles, ou d'un formalisme logique, par

exemple Booléen, explicitant comment se combinent les influences des protéines régulatrices contrôlant l'expression de chaque protéine. Ces modèles permettent de prédire l'effet d'une perturbation, qu'elle soit intrinsèque à la cellule cancéreuse (par mutation), ou extrinsèque (par introduction d'une drogue à effet thérapeutique). Ils requièrent une description précise du réseau de régulation, et donc des bases de données de haute qualité comme Reactome ou ACSN. Celui-ci est généralement construit pour reproduire d'abord le comportement d'une cellule générique et est ensuite adapté à la tumeur considérée, en fonction des mutations ou altérations génomiques qui les caractérisent. Le modèle mathématique est alors modifié, chaque variable ou paramètre du modèle étant fixé à une valeur qui correspond à l'observation : typiquement, une mutation activatrice, une surexpression ou amplification correspondront à une activité forcée de la variable et une mutation délétère ou une délétion hétéro- ou homozygote correspondra à une perte d'activité. Selon le formalisme du modèle, les méthodes utilisées pour arriver à ce résultat varieront. Dans le cadre Booléen, par exemple, la variable correspondant au gène affecté est fixée à 1 et à 0 respectivement.

Ainsi, la liste des mutations de chaque patient viendra modifier le modèle qui sera ensuite simulé. Selon la structure du réseau modifié et les résultats de la simulation (sous la forme de probabilité de phénotypes atteints par rapport au modèle de cellule saine par exemple), on pourrait prédire le pronostic et l'effet d'une drogue sur un patient en particulier [21].

Divers efforts ont déjà été faits dans ce sens, mais l'approche spécifique à chaque patient reste encore à parfaire. Ainsi, Sahin *et al.* ont étudié, grâce à un modèle Booléen, les combinaisons possibles de drogues, ou cibles thérapeutiques, qui pourraient être utilisées pour traiter les tumeurs du sein surexprimant le récepteur *ERBB2* dans les cas de résistance *de novo* au traitement par trastuzumab [22]. La modélisation de leur réseau, ainsi que le test sur des lignées cellulaires, explique l'inefficacité d'un traitement combiné avec un inhibiteur de *EGFR*, mais suggère l'efficacité d'inhibiteur de *c-MYC*.

Fumiã et Martins ont, de leur côté, identifié des gènes conducteurs, par opposition aux gènes passagers, qui mènent à une tumeur maligne, et ce à l'aide d'un modèle Booléen. L'identification de ces gènes et de la séquence des mutations au cours de la tumorigenèse permet de proposer des associations de thérapies ciblées plutôt que des monothérapies mises en place dans les traitements conventionnels [23]. D'autres travaux, qui consistent à utiliser des données de phosphoprotéomiques [24] pour adapter un modèle Booléen à un cancer en particulier, permettent d'anticiper les progrès à prévoir dans les prochaines années en matière de médecine de précision à l'aide de modèles mathématiques.

Conclusion

Si la médecine de précision en oncologie est déjà possible grâce aux technologies moléculaires à haut débit comme le SHD, nos connaissances des mécanismes moléculaires de la cellule normale et tumorale ouvrent la voie à des approches analytiques intégrant la modélisation explicite du fonctionnement du réseau de régulation spécifique de la tumeur traitée. Dans leur version simple basée sur l'analyse de voisinage, ces approches permettent déjà une stratification des tumeurs. Plus exigeante, la modélisation dynamique du réseau de régulation constitue une perspective majeure de progrès thérapeutique, et doit permettre une utilisation rationnelle de la pharmacopée anticancéreuse, notamment de la puissance thérapeutique attendue des inhibiteurs ciblés. Les exemples ci-dessus montrent que la modélisation de réseau permet d'ores et déjà de guider le développement de molécules thérapeutiques dans des situations cliniques de résistance. En documentant les mécanismes contrôlant les processus de prolifération, de différenciation, de mort cellulaire et de survie, les bases de données de réseaux biologiques de haute qualité apporteront la résolution moléculaire et la clarté mécanistique indispensable à la réalisation des promesses de la médecine de précision en oncologie.

Références

1. Garraway LA, Verweij J, Ballman KV. Precision oncology: an overview. *J Clin Oncol Off J Am Soc Clin Oncol* 2013 ; 31 : 1803-5.
2. Prahallad A, Sun C, Huang S, et al. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* 2012 ; 483 : 100-3.
3. Von Hoff DD, Stephenson JJ Jr, Rosen P, et al. (2010). Pilot study using molecular profiling of patients' tumors to find potential targets and select treatments for their refractory cancers. *J Clin Oncol Off J Am Soc Clin Oncol* 2010 ; 28 : 4877-83.

4. Lacroix L, Boichard A, André F, Soria JC. Genomes in the clinic: the Gustave Roussy Cancer Center experience. *Curr Opin Genet Dev* 2014 ; 24C : 99-106.
5. Le Tourneau C, Kamal M, Trédan O, et al. Designs and challenges for personalized medicine studies in oncology: focus on the SHIVA trial. *Target Oncol* 2012 ; 7 : 253-65.
6. Le Tourneau C, Paoletti X, Servant N, et al. Randomised proof-of-concept phase II trial comparing targeted therapy based on tumour molecular profiling vs conventional therapy in patients with refractory cancer: results of the feasibility part of the SHIVA trial. *Br J Cancer* 2014 ; in press.
7. Stegmeier F, Warmuth M, Sellers WR, Dorsch M. Targeted cancer therapies in the twenty-first century: lessons from imatinib. *Clin Pharmacol Ther* 2010 ; 87 : 543-52.
8. Sun C, Wang L, Huang S, et al. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. *Nature* 2014 ; 508 : 118-22.
9. Franceschini A, Szklarczyk D, Frankild S, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 2013 ; 41 : D808-15.
10. Keshava Prasad TS, Goel R, Kandasamy K, et al. (2009). Human Protein Reference Database—2009 update. *Nucleic Acids Res* 2009 ; 37 : D767-72.
11. Cerami EG, Gross BE, Demir E, et al. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res* 2011 ; 39 : D685-90.
12. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2012 ; 40 : D109-14.
13. Paz A, Brownstein Z, Ber Y, et al. SPIKE: a database of highly curated human signaling pathways. *Nucleic Acids Res* 2011 ; 39 : D793-9.
14. Croft D, Mundo AF, Haw R, et al. The Reactome pathway knowledgebase. *Nucleic Acids Res* 2014 ; 42 : D472-7.
15. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005 ; 102 : 15545-50.
16. Bonnet E, Calzone L, Rovera D, Stoll G, Barillot E, Zinovyev A. BiNoM 2.0, a Cytoscape plugin for accessing and analyzing pathways using standard systems biology formats. *BMC Syst Biol* 2013 ; 7 : 18.
17. Zinovyev A, Viana E, Calzone L, Barillot E. BiNoM: a Cytoscape plugin for manipulating and analyzing biological networks. *Bioinformatics* 2008 ; 24 : 876-7.
18. Kairov U, Karpenyuk T, Ramanculov E, Zinovyev A. Network analysis of gene lists for finding reproducible prognostic breast cancer gene signatures. *Bioinformatics* 2012 ; 18(6) : 773-6.
19. Barillot E, Calzone L, Hupe P, Vert JP, Zinovyev A. *Computational Systems Biology of Cancer*. Boca Raton : CRC Press, 2012.
20. Hofree M, Shen JP, Carter H, Gross A, Ideker T. Network-based stratification of tumor mutations. *Nat Methods* 2013 ; 10 : 1108-15.
21. Zinovyev A, Fourquet S, Tournier L, Calzone L, Barillot E. Cell death and life in cancer: mathematical modeling of cell fate decisions. In : Goryanin I, Goryachev A, eds. *Advances in Experimental Medicine and Biology*, Vol. 736. New York : Springer, 2012.
22. Sahin O, Fröhlich H, Löbke C, et al. Modeling ERBB receptor-regulated G1/S transition to find novel targets for de novo trastuzumab resistance. *BMC Syst Biol* 2009 ; 3 : 1.
23. Fumiã HF, Martins ML. Boolean network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes. *PLoS One* 2013 ; 8 : e69008.
24. Saez-Rodriguez J, Alexopoulos LG, Zhang M, Morris MK, Lauffenburger DA, Sorger PK. Comparing signaling networks between normal and transformed hepatocytes using discrete logical models. *Cancer Res* 2011 ; 71 : 5400-11.