Complexity and Analogy in Science
Pontifical Academy of Sciences, *Acta* 22, Vatican City 2014
www.pas.va/content/dam/accademia/pdf/acta22/acta22-alberghina.pdf

Towards an Understanding of the Molecular Complexity of Cancer

Lilia Alberghina

Introduction

The molecular biology of cancer cells has a crucial role not only in promoting scientific knowledge, but also in developing better treatments for patients of a disease that is still in strong medical need. Since the eighties of the last century, the research in this field has been characterized by a genetic approach. The recombinant DNA technologies, developed in those years, allowed to isolate genes (called "oncogenes"), which were shown to be able to promote both *in vitro* and *in vivo* cancer growth. More than one hundred oncogenes have been described, together with a few "oncosuppressors", genes whose function is, on the contrary, to block cancer cells proliferation.

Extensive analysis of the molecular functions of both oncogenes and oncosuppressors has shown, to the surprise of the scientific community, that they encode for mutated or overexpressed proteins, which, in the wild type form and at normal level, are constituents of healthy cells. The large part of these proteins are involved in signaling, that is in the cascade of biochemical reactions that convey the cues coming from the environment (from nutrient availability and/or from the presence of growth factors) to the cell machinery.

Given that oncogenes are found to activate signaling, thereby allowing cancer cells to grow under conditions in which normal cells do not, a new drug discovery strategy was developed: to construct molecules able to block, in a specific way, the various oncogenes, whose activation underlays each given cancer type. Although these new drugs present, at first, very positive clinical responses, quite often, after several months of treatment, there is a recurrence of the disease, due to innate or acquired resistance (1, 2). New signaling pathways, able to sustain the growth and survival of cancer cells, are activated, mostly by mutations, generated by the genomic instability of cancer cells (3), so to bypass the oncogenic signaling pathway which was the initial target of the therapeutic treatment. It has become clear, therefore, that the genetic approach, which polarized the interest on specific oncogenes (and on the signaling pathways in which they are involved), while it considered the growth and survival of cancer cells as a "read out" of the genetic analysis, is no longer tenable, and instead investigations have to be directed to the molecular basis of the specific "read out" of cancer cells.

The molecular setup of a human cell is extremely complex: while we have about 50,000 billion cells in a human body, each cell contains, on average, about 1 billion proteins of about 10,000 different sequences, being the proteins the components that carry out the large majority of the activities of the cell. Furthermore, each of these proteins may be localized in different compartments of the cell (having a distinct role in each compartment) and may be modulated in its activity by a number of different post-translational modifications (i.e. phosphorylation, acetylation, etc.) that are under the control of the signaling apparatus.

Molecular complexity of cancer cells

The "omics" technologies (transcriptome, proteome and metabolome analyses), together with the classical approaches of cellular and molecular biology, have collected a wealth of information on molecular changes observed in cancer cells, when compared with normal ones. The various types of tumors and the many cancer cell lines that have been analyzed, show changes, for instance, in several thousand gene products, which are largely different between one type and another type of tumor or cell line (4). The necessity thus emerged to develop automatic, computational algorithms able to extract information from those sets of Big Data (5). Statistical analysis and graph theory have offered the first tools used to extract an order from this deluge of data (6), but they have obtained a very limited success in terms of a better understanding of the disease.

Following a line of thought first elaborated by René Descartes on how to treat complex problems to be solved: "to divide each difficulty which I examined into as many parts as possible and as might be necessary to resolve it better", Hanahan and Weiberg (7) recognized that, under the vast catalogue of different cancer phenotypes, there are a small number of essential alterations in cell biochemistry and physiology (that they called "hallmarks"), which, integrated in various ways, generate the features of cancer disease.

Some of these hallmarks are specifically expressed in the organism, for instance the ability to metastasize, others are properties that can be investigated, under controlled conditions, also in cancer cells growing "in vitro". A reorganization of cancer hallmarks, following a systems approach, identified a number of functional modules, each one characterized by a measurable input, a molecular network that elaborates the input, to produce a measurable output (8).

Systems Biology is emerging as the necessary approach to investigate how the interactions of a large number of molecular players generate the functional properties of living cells (9). The integration of molecular analysis

with mathematical modeling and simulation, in an iterative process, characterizes the systems biology approach, since it allows to investigate the role of non-linear dynamic steps in the network under investigation (10). Mathematical models may be constructed at different levels of resolution (11) and based on different rationales (12, 13).

Due to the non-linearity of biological complex processes, their functions emerge as a system-level property, making the relation between genotype and phenotype not straightforward (as it may happen for simple functions, like the transport of respiratory gases by hemoglobin), but very difficult to predict, if one knows only the properties of the individual components and not the map and the strength of their interactions, organized in a dynamic mathematical model. This is the reason why a purely reductionist approach is not, and will not, be able to explain complex functions, as cancer, and systems biology is needed to move towards the understanding of complex functions as emergent properties.

Going back to the cancer hallmarks, the more basic property, necessary for the development of the disease, is the uncontrolled, enhanced cell growth, that is well characterized under "in vitro" conditions. Extensive transcriptome analysis has shown that almost 3000 gene products are differentially expressed in cancer cells as compared to their normal counterparts (14). Clustering analysis has identified, among various changes, alterations in metabolism, both glycolytic and mitochondrial (15).

Cancer metabolism rewiring

In 1956 Otto Warburg reported that most cancer cells present a high rate of glycolysis, even in the presence of oxygen. Pyruvate, instead of being oxidized by mitochondria, as normal cells do, is converted to lactate. He postulated that this change of metabolism be the fundamental cause of malignant transformation (16).

The Warburg hypothesis was scoffed at, and later forgotten in the period in which the genetic approach to oncogenes was popular. Things started to change when evidence accumulated to show that different oncogenes have as a common "read out": the stimulation of glycolysis (17).

The turnaround occurred a few years ago, when it was shown that, besides having stimulated glycolysis, several cancer cells are "glutamine addicted", that is, they require glutamine (a non essential aminoacid in humans) for proliferation (18, 19). Glutamine is utilized by reductive carboxylation, an unusual pathway which may lead to lipid synthesis (20).

In K-Ras transformed cells, utilization of glucose and glutamine has been shown to be decoupled: glucose goes to lactate, producing ATP on

the way, while glutamine is utilized both as a carbon and nitrogen source for biomass production (21). One has to recall that normal cells utilize glutamine only as a source of organic nitrogen and excrete the remaining glutamic acid (the carbon moiety). It should be added that in the same transformed cells, mitochondria are quite inefficient, having the Complex I dysfunctional (22).

Is it possible to connect all the information that we have by now on the ample cancer metabolic rewiring (CMR): stimulated glycolysis and reductive carboxylation of glutamine, mitochondrial dysfunction, enhanced lipid synthesis, sustained protein and nucleic acid production, etc.? This is going to be a great task for systems biology.

First of all, we need to reconstruct the large metabolic network that underlays CMR. A redox control has been identified from the analysis of the map, linking the various areas of the metabolism and explaining their interlocking (23). Advanced computational approaches are required to unravel the complexity of metabolism and of its regulation (24). Of particular interest are constraint-based models (12, 13). This ingenious approach, that requires to know the stoichiometry of the reactions involved without asking for their kinetics parameters, which are very difficult to obtain by experimental analysis, allows to estimate the flux, that, in steady state, traverse the various pathways of the metabolism, thereby offering an informative description of the functionality of the entire network. Flux Balance Analysis has been developed both for small networks and for genome-scale networks (25, 26). Given that experimental findings indicate that various types of tumors may have different metabolic pathways (some are reported to be glycolytic while others are oxidative), an ensemble evolutionary constraint-based approach was developed to better understand how various phenotypes of CMR metabolism may be generated (27).

Studies, now underway in my laboratory, based on the various computational approaches indicated above, are recognizing the design principles that underlay CMR in various cancer cells.

Conclusions and perspectives

Analytical technologies, based on mass spectrometry, allow to obtain experimental findings not only on metabolic profiling and on metabolic pathway analysis, but more importantly also on metabolic flux analysis (25, 26), which can be used to ascertain the validity of the indications coming from computational Flux Balance Analysis, for various tumors and cell lines. The details of CMR in various types of cancer are thus going to be ascertained. Given that inhibition of CMR has been shown to block cancer cell growth,

the knowledge of the biochemical pathways involved in CMR is going to be the stepping-stone for a new strategy of anticancer drug discovery.

The molecular complexity of cancer is a very clear example of the "organized complexity" anticipated by Warren Weaver (28). While classical physics has been devoted to few-variable problems and many areas of scientific investigation are analyzing "disorganized complexity", in which numerous variables can only be treated by probability analysis, biology offers problems of "organized complexity".

A moderate number of variables, often linked by non-linear relationships, generate functions as emergent properties. As discussed above, only the integration of extensive molecular analysis and of computational algorithms and models may allow to understand how biological functions are generated by organized molecular complexity.

As shown in this note, science has become able to tackle with success this new frontier, having developed appropriate experimental and computational approaches. It is a real paradigm change: from the consolidated molecular and reductionist approach, biology is moving towards a new integrated, multidisciplinary approach, that is going to be "Big Science" oriented. As predicted by Weaver, large communities of scientists are going to collaborate to reach macro-objectives: from a better understanding of major diseases to unraveling the relations brain/mind.

The flourishing of biological organized complexity studies is expected to elucidate new concepts that are going to impact also on the analysis and management of man-made organized complexity, such as that found in social or financial organizations.

References

1. Eberlein, C.A., Stetson, D., Markovets, A.A., Al-Kadhimi, K.J., Lai, Z., Fisher, P.R., Meador, C.B., Spitzler, P., Ichihara, E., Ross, S.J., *et al.* (2015) Acquired Resistance to the Mutant-Selective EGFR Inhibitor AZD9291 Is Associated with Increased Dependence on RAS Signaling in Preclinical Models. *Cancer Res.*, 10.1158/0008-5472.CAN-14-3167.
2. Holohan, C., Van Schaeybroeck, S., Longley, D.B. and Johnston, P.G. (2013) Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer*, 13, 714-726.
3. Stoler, DL, Chen, N, Basik, M, Kahlenberg, MS, Rodriguez-Bigas, MA, Petrelli, NJ, Anderson, GR (1999). The onset and extent of genomic instability in sporadic colorectal tumor progression. *Proc. Natl. Acad. Sci. U.S.A.*, 96, 15121-6.
4. Shaw, P.G., Chaerkady, R., Wang, T., Vasilatos, S., Huang, Y., Van Houten, B., Pandey, A. and Davidson, N.E. (2013) Integrated proteomic and metabolic analysis of breast cancer progression. *PloS One*, 8, e76220.
5. Berger, B., Peng, J. and Singh, M. (2013) Computational solutions for omics data. *Nat. Rev. Genet.*, 14, 333-346.

6. Wang, J., Zuo, Y., Man, Y.-G., Avital, I., Stojadinovic, A., Liu, M., Yang, X., Varghese, R.S., Tadesse, M.G. and Ressom, H.W. (2015) Pathway and network approaches for identification of cancer signature markers from omics data. *J. Cancer*, 6, 54-65.
7. Hanahan, D. and Weinberg, R.A. (2000) The hallmarks of cancer. *Cell*, 100, 57-70.
8. Alberghina, L., Gaglio, D., Gelfi, C., Moresco, R.M., Mauri, G., Bertolazzi, P., Messa, C., Gilardi, M.C., Chiaradonna, F. and Vanoni, M. (2012) Cancer cell growth and survival as a system-level property sustained by enhanced glycolysis and mitochondrial metabolic remodeling. *Front. Physiol.*, 3, 362.
9. Alberghina, L. and Westerhoff, H. Systems Biology: Definitions and Perspectives. *Springer-Verlag Berlin Heidelberg. (2005)*
10. Kitano, H. (2002) Systems biology: a brief overview. *Science*, 295, 1662-1664.
11. Noble, D. (2002) The rise of computational biology. *Nat. Rev. Mol. Cell Biol.*, 3, 459-463.
12. Stelling, J. (2004) Mathematical models in microbial systems biology. *Curr. Opin. Microbiol.*, 7, 513-518.
13. Stelling, J. and Gilles, E.D. (2004) Mathematical modeling of complex regulatory networks. *IEEE Trans. Nanobioscience*, 3, 172-179.
14. Akan, P., Alexeyenko, A., Costea, P.I., Hedberg, L., Solnestam, B.W., Lundin, S., Hällman, J., Lundberg, E., Uhlén, M. and Lundeberg, J. (2012) Comprehensive analysis of the genome transcriptome and proteome landscapes of three tumor cell lines. *Genome Med.*, 4, 86.
15. Balestrieri, C, Vanoni, M, Hautaniemi, S, Alberghina, L, Chiaradonna, F (2012). Integrative transcriptional analysis between human and mouse cancer cells provides a common set of transformation associated genes. *Biotechnol. Adv.*, 30, 16-29.
16. Warburg, O. (1956) On the origin of cancer cells. *Science*, 123, 309-314.
17. Reitzer, LJ, Wice, BM, Kennell, D (1979). Evidence that glutamine, not sugar, is the major energy source for cultured HeLa cells. *J. Biol. Chem.*, 254, 2669-76.
18. DeBerardinis, R.J., Mancuso, A., Daikhin, E., Nissim, I., Yudkoff, M., Wehrli, S. and Thompson, C.B. (2007) Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, 104, 19345-19350.
19. Wise, D.R., DeBerardinis, R.J., Mancuso, A., Sayed, N., Zhang, X.-Y., Pfeiffer, H.K., Nissim, I., Daikhin, E., Yudkoff, M., McMahon, S.B., *et al.* (2008) Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl. Acad. Sci. U.S.A.*, 105, 18782–18787.
20. Metallo, C.M., Gameiro, P.A., Bell, E.L., Mattaini, K.R., Yang, J., Hiller, K., Jewell, C.M., Johnson, Z.R., Irvine, D.J., Guarente, L., *et al.* (2012) Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature*, 481, 380-384.
21. Gaglio, D., Metallo, C.M., Gameiro, P.A., Hiller, K., Danna, L.S., Balestrieri, C., Alberghina, L., Stephanopoulos, G. and Chiaradonna, F. (2011) Oncogenic K-Ras decouples glucose and glutamine metabolism to support cancer cell growth. *Mol. Syst. Biol.*, 7, 523.
22. Baracca, A., Chiaradonna, F., Sgarbi, G., Solaini, G., Alberghina, L. and Lenaz, G. (2010) Mitochondrial Complex I decrease is responsible for bioenergetic dysfunction in K-ras transformed cells. *Biochim. Biophys. Acta*, 1797, 314-323.
23. Alberghina, L. and Gaglio, D. (2014) Redox control of glutamine utilization

in cancer. *Cell Death Dis.*, 5, e1561.
24. Cazzaniga, P., Damiani, C., Besozzi, D., Colombo, R., Nobile, M.S., Gaglio, D., Pescini, D., Molinari, S., Mauri, G., Alberghina, L., *et al.* (2014) Computational strategies for a system-level understanding of metabolism. *Metabolites*, 4, 1034-1087.
25. Gatto, F., Miess, H., Schulze, A. and Nielsen, J. (2015) Flux balance analysis predicts essential genes in clear cell renal cell carcinoma metabolism. *Sci. Rep.*, 5, 10738.
26. Orth, J.D., Thiele, I. and Palsson, B.Ø. (2010) What is flux balance analysis? *Nat. Biotechnol.*, 28, 245-248.
27. Damiani, C., Pescini, D., Colombo, R., Molinari, S., Alberghina, L., Vanoni, M. and Mauri, G. (2014) An ensemble evolutionary constraint-based approach to understand the emergence of metabolic phenotypes. *Nat. Comput.*, 13, 321-331.
28. Weaver, W. (1948) Science and Complexity. *Am. Sci.*, 36, 536-544.