Commentary

Protein Complex, Gene, and Regulatory Modules in Cancer Heterogeneity

Nikolaos A Papanikolaou[1] and Athanasios G Papavassiliou[2]

[1]Ludwig Institute for Cancer Research, New York, New York, United States of America and [2]Department of Biological Chemistry, University of Athens Medical School, Athens, Greece

Online address: http://www.molmed.org
doi: 10.2119/2008-00083.Papanikolaou

INTRODUCTION

The recent excitement generated by the analysis of human cancer genomes is tempered by the uncovered regulatory complexity. Interactions between heterogeneous multiprotein complex and genetic and regulatory modules form networks unique not only to different tumor types, but also, surprisingly, to similar type tumors generating molecular/cellular heterogeneity that ultimately control cancer cell proliferation and invasion, endowing tumors with robustness. It is, however, unknown how cancer-specific modules and their networks determine tumor behavior, or how they directly or indirectly reorganize signaling complexes into novel networks even within the same tumor, making tumor treatment problematic. Experimental and theoretical data from high-throughput studies suggest that the key lies in characterizing cancer-specific protein complex modules, identifying common nodes, that is, the key proteins (oncogenes or other molecules) shared by different complex modules and networks, and in dissecting adaptive, inter-module functional communication in the control of tumor growth by rearrangement of protein complex modules.

Altered genetic and epigenetic networks in cancer cells are organized in functional modules (1) and result in permanent changes in the types, numbers, and interactions of protein complex genes that endow tumors with heterogeneity and robustness, and largely determine disease progression. Robustness as used above is defined as the ability of genes to resist changes in function from perturbations in individual components (2,3). They originate from mutated/overexpressed oncogenic or tumor suppressor molecules (Figure 1) in key pathways that control cell and tissue growth, such as the cell cycle (4), signal transduction pathways (5), or chromatin remodeling pathways (6,7). Genomic, proteomic, and other high-throughput sequencing efforts are revealing an unprecedented complexity and heterogeneity of human cancer genomes (8,9).

No general principles have emerged that describe heterogeneity (10), and most methods currently in use, such as polymerase chain reaction (PCR), derive their utility from their clonal nature, which, although permitting genome-wide, high-throughput analyses of many samples, are limited by sensitivity, sample size, and noise. Recent computational and experimental results on yeast protein network architecture (11) open the possibility to formulate general experimental hypotheses that bridge theory and experiment in identifying cancer pathways that drive overall tumor behavior. Central to such hypotheses will be the identification of key node proteins (nodes) in different protein modules that permit inter-module communication between different cell populations and their networks (11). In this commentary, we discuss sources of human solid tumor heterogeneity, robustness, and their dependence on the existence of differently reorganized complexes and networks, using the example of epidermal growth factor receptor (EGFR)/MET complex reorganization in gefinitib resistance.

The Nature of Tumor Heterogeneity and Robustness

High-throughput gene (12,13) and protein expression data, as well as massively parallel signature sequencing (MPSS) which reveals the mutation landscape of tumors, generate a complex picture that can be analyzed in terms of genetic and transcriptional regulatory modules, that is, sets of transcription factors that regulate these gene modules (14,15) in a temporal and signal-dependent manner. Most solid human tumors are communities of several different types of cells with varied gene and protein expression profiles that generate hundreds of multimolecular complexes in cells, organized in functional signaling pathways (16–18)

Address correspondence and reprint requests to *Athanasios G Papavassiliou, Department of Biological Chemistry, University of Athens Medical School, 75 M. Asias Street, 11527 Athens, Greece. Phone: + 30-210-746-2508/9; Fax: + 30-210-779-1207; E-mail: papavas@med.uoa.gr.*
Submitted July 18, 2008; Accepted for publication July 18, 2008; Epub (www.molmed.org) ahead of print July 18, 2008.

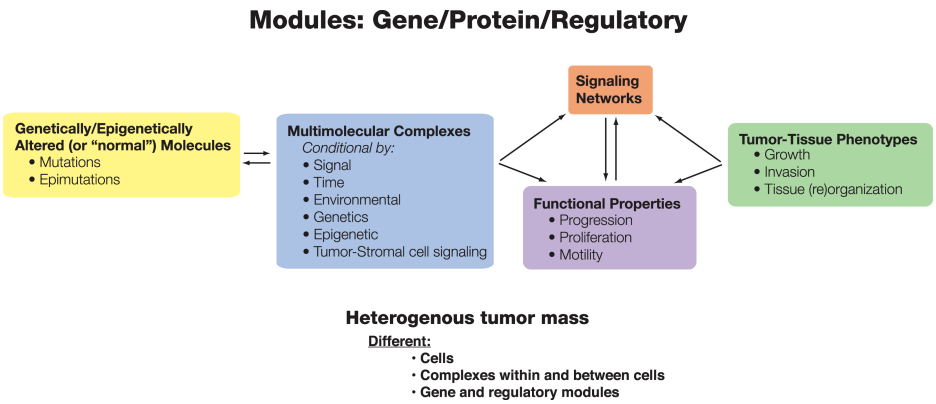

Figure 1. Model of hierarchical protein complex, gene, and regulatory module organization. Mutations and epimutations in "cancer genes" give rise to molecules (proteins and so on) that form altered multi-molecular complex functional modules. These are, in turn, reorganized into novel, tumor, stage, or even sample-specific, autonomous, physical, as well as temporal growth-promoting networks that control tumor tissue properties, such as proliferation and invasion.

(see Figure 1). Tumor sample heterogeneity also is reflected by the presence of different putative tumor driver mutations in different specimens in the same tumor samples (8). Heterogeneity within the same tumor is more clearly apparent in data from high-throughput, targeted sequencing of specific kinase families in different tumor types. For example, genome-wide sequencing of samples from 72 breast tumors and nine cell lines reveal that only 6 out of 518 kinases bear non-synonymous, putative driver mutations (19). Similar results were obtained from lung cancer data (19–21) and testicular tumors (22).

Reducing Heterogeneity of Pathways and Genes to Genetic and Regulatory Modules

Reconstructing cancer pathways (23) addressed by genome-wide approaches is a formidable task (24,25). However, as mentioned in the previous section, these studies also reveal that, with few exceptions, profiles and distribution of mutations are variable, depending on sample, tumor type, and, within the same tumor type, on whether single cells or the gross tumor has been analyzed (8). Data also is confounded by noise and limited reproducibility, the non-concordance between mRNA, encoded protein, or small molecule (miRNA and so on) expression and processing. Genetic and regulatory modules are the units of cell- and tissue-level biological functions (26,27) and, in the long term, into targetable molecular mechanisms (28). Therefore, pattern identification relies on extracting coherent gene expression changes and on inferring the operative transcriptional and other cellular modules (26,27). More recent approaches attempt to integrate multiple data sources, including patient survival, with gene profile signatures in wound repair (29), to compile molecular concepts maps (MCMs) from gene/protein annotations, and to derive regulatory networks and expression profiles computationally (30), or to compare coherent sets of genes that already are mapped onto cancer pathways for association with either a single gene profile in microarrays or "phenotypes" using Kolmogorov-Smirnov sum statistics (31,32). Another approach uses "cancer gene and module compendium" from different studies to uncover shared genes and gene modules, and uses them to map onto and to predict clinical cancer properties (33,34). In many cases, use of gross tumor mass tissue results in an average profile that obfuscates the contribution of different cells and signals. Reduction of expression and other data to functional modules is more informative in simple model organisms, both because they are equipped with simpler genomes and because they consist of genetically similar or identical cells (less heterogeneous).

Identifying Key Protein Nodes in Complex and Functional Modules within Heterogeneous Tumors: Bridging Theory with Experiment

Although identification of gene and regulatory modules is revolutionizing our understanding of tumor growth, we are far from reconstructing shared molecular mechanisms in similar, let alone different, human cancers because there is a gap between experimental and theoretical approaches that define protein complex modules into realistic cancer cell networks. It is unknown how oncogene protein and other cancer protein complex modules are reorganized when tumors evolve, however, in spite of the complexity imposed by tumor cell genetic and biochemical heterogeneity, theoretical model building based on data from yeast interactomes and proteomes suggests that it is possible to deal with this issue by finding shared signatures from pure, isolated tumor cell populations. For example, analysis of the yeast protein interactome by Spirin and Mirny (11) revealed that two main modules exist: protein complex and functional modules with overlapping protein members in several growth-regulating pathways such as the cyclin-dependent kinase (CDK)/Cyclin; and mitogen-activated protein (MAP) kinase. This proposal remains purely hypothetical, however recent results on EGFR mutation-driven resistance to gefinitib in human lung cancers clearly support the idea that complex rearrangements are a major mechanism of robustness-driven resistance and growth within cells. EGFR mutations not affecting drug binding allow it to complex with the MET oncogene product, a receptor tyrosine kinase, thus permitting uncharacter-

ized MET-driven growth pathways to sustain cancer cell survival.

Stable or transient multiprotein complex modules of known oncogene proteins, such as MET/ErbB3, form novel protein complex modules that cannot be predicted by any current high-throughput strategy, however, their central role in tumor robustness (35) makes it imperative that their topology and interactions in pure tumor cell populations be defined.

CONCLUSION

Recent experimental and theoretical results on biological networks, their topology, and characterization of oncogenic protein complex modules suggest that mapping oncogene products and other cancer-involved genes to actual pathways in simple, homogeneous cell populations determining tumor behavior is a desired approach to unravel tumor heterogeneity and vulnerability. How oncoprotein complex modules change in response to intratumor or tumor-stromal signaling, and how gene, protein, and regulatory modules change by rearranging polypeptide composition in response to disease progression remain unanswered questions, and are vital to successful, combination therapeutic interventions.

ACKNOWLEDGMENTS

Much worthwhile work could not be included due to space constraints. We apologize to colleagues whose work had to be omitted.

REFERENCES

1. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. (1999) From molecular to modular cell biology. *Nature* 402:C47–52.
2. Kitano H. (2004) Cancer as a robust system: implications for anticancer therapy. *Nat. Rev. Cancer* 4:227–35.
3. Kitano H. (2007) The theory of biological robustness and its implication in cancer. *Ernst. Schering. Res. Found. Workshop* 69–88.
4. Sherr CJ. (2000) Cell cycle control and cancer. *Harvey Lect.* 96:73–92.
5. Papatsoris AG, Karamouzis MV, Papavassiliou AG. (2007) The power and promise of "rewiring" the mitogen-activated protein kinase network in prostate cancer therapeutics. *Mol. Cancer Ther.* 6:811–9.
6. Davie JR, *et al.* (1999) Organization of chromatin in cancer cells: role of signalling pathways. *Biochem. Cell Biol.* 77:265–75.
7. Nikolaev AY, Papanikolaou NA, Li M, Qin J, Gu W. (2004) Identification of a novel BRMS1-homologue protein p40 as a component of the mSin3A/p33(ING1b)/HDAC1 deacetylase complex. *Biochem. Biophys. Res. Commun.* 323:1216–22.
8. Sjoblom T, *et al.* (2006) The consensus coding sequences of human breast and colorectal cancers. *Science* 314:268–74.
9. Vogelstein B, Kinzler KW. (2004) Cancer genes and the pathways they control. *Nat. Med.* 10:789–99.
10. Mesarovic MD, Sreenath SN, Keene JD. (2004) Search for organising principles: understanding in systems biology. *Syst. Biol. (Stevenage)* 1:19–27.
11. Spirin V, Mirny LA. (2003) Protein complexes and functional modules in molecular networks. *Proc. Natl. Acad. Sci. U. S. A.* 100:12123–8.
12. Balmain A, Gray J, Ponder B. (2003) The genetics and genomics of cancer. *Nat. Genet.* 33(Suppl): 238–44.
13. Clarke PA, te Poele R, Wooster R, Workman P. (2001) Gene expression microarray analysis in cancer biology, pharmacology, and drug development: progress and potential. *Biochem. Pharmacol.* 62:1311–36.
14. Lerman G, *et al.* (2007) Functional genomics via multiscale analysis: application to gene expression and ChIP-on-chip data. *Bioinformatics* 23:314–20.
15. Ren B, Dynlacht BD. (2004) Use of chromatin immunoprecipitation assays in genome-wide location analysis of mammalian transcription factors. *Methods Enzymol.* 376:304–15.
16. Lu P, Nakorchevskiy A, Marcotte EM. (2003) Expression deconvolution: a reinterpretation of DNA microarray data reveals dynamic changes in cell populations. *Proc. Natl. Acad. Sci. U. S. A.* 100:10370–5.
17. Nesbit CE, Tersak JM, Prochownik EV. (1999) MYC oncogenes and human neoplastic disease. *Oncogene* 18:3004–16.
18. von Mering C, *et al.* (2002) Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* 417:399–403.
19. Stephens P, *et al.* (2005) A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. *Nat. Genet.* 37:590–2.
20. Davies H, *et al.* (2002) Mutations of the BRAF gene in human cancer. *Nature* 417:949–54.
21. Davies H, *et al.* (2005) Somatic mutations of the protein kinase gene family in human lung cancer. *Cancer Res.* 65:7591–5.
22. Bignell G, *et al.* (2006) Sequence analysis of the protein kinase gene family in human testicular germ-cell tumors of adolescents and adults. *Genes Chromosomes Cancer* 45:42–6.
23. Fambrough D, McClure K, Kazlauskas A, Lander ES. (1999) Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell* 97:727–41.
24. Holstege FC, *et al.* (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95:717–28.
25. Tomlins SA, Chinnaiyan AM. (2006) Of mice and men: cancer gene discovery using comparative oncogenomics. *Cancer Cell* 10:2–4.
26. Segal E, *et al.* (2003) Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat. Genet.* 34:166–76.
27. Segal E, Friedman N, Koller D, Regev A. (2004) A module map showing conditional activity of expression modules in cancer. *Nat. Genet.* 36:1090–8.
28. Liu ET, Kuznetsov VA, Miller LD. (2006) In the pursuit of complexity: systems medicine in cancer biology. *Cancer Cell* 9:245–7.
29. Chang HY, *et al.* (2005) Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc. Natl. Acad. Sci. U. S. A.* 102:3738–43.
30. Tomlins SA, Rubin MA, Chinnaiyan AM. (2006) Integrative biology of prostate cancer progression. *Annu. Rev. Pathol.* 1:243–71.
31. Mootha VK, *et al.* (2003) Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. *Proc. Natl. Acad. Sci. U. S. A.* 100:605–10.
32. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102:15545–50.
33. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM. (2002) Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res.* 62:4427–33.
34. Rhodes DR, *et al.* (2004) Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc. Natl. Acad. Sci. U. S. A.* 101:9309–14.
35. Pereira-Leal JB, Levy ED, Teichmann SA. (2006) The origins and evolution of functional modules: lessons from protein complexes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361:507–17.