The Human Cancer Genome Project—one more misstep in the war on cancer

George L Gabor Miklos

Strap yourself in and get ready for some serious 'more of the same.' A recent proposal to sequence cancer genomes holds out the promise of personalized cures for each of 50 different cancers. The cost? A mere $12 billion at today's prices[1]. This human cancer genome megaproject is the equivalent of 12,500 human genome projects and already has the backing of several prominent scientists. Harold Varmus believes that the project could "completely change how we view cancer"[1]; Eric Lander argues that "knowing the defects of the cancer cell points you to the Achilles' heel of tumors"[1]; and Francis Collins predicts that he "can confidently tell you that something will happen here"[1]. More pragmatically, Craig Venter points out that "...it's not clear what answer we'd get; there might be better ways to move cancer research forward"[1].

In a nutshell, the megaproject aims to catalog all somatic mutations from primary tumors as the basis for designer drugs to cure most cancers. Success is predicated on the assumption that drugs can be targeted to very specific mutated regions of gene products. However, most patients with a localized primary tumor are cured by surgery and local radiation. It is not the primary tumor, but the metastatic spread of a small population of deadly cells that ultimately compromises a normal tissue or organ, that kills in cancer[2] (for an excellent popular account, see ref. 3). Are primary tumors therefore the appropriate focus for such a massive project and are their bulk mutational spectra therapeutically useful?

George L. Gabor Miklos is at Secure Genetics Pty Limited, 19 Bungan Head Road, Newport Beach, Sydney, New South Wales, Australia 2106. He was an advisor to the Berkeley Drosophila Genome Project and to the human and mouse genome projects at Celera.
e-mail: gmiklos@securegenetics.com

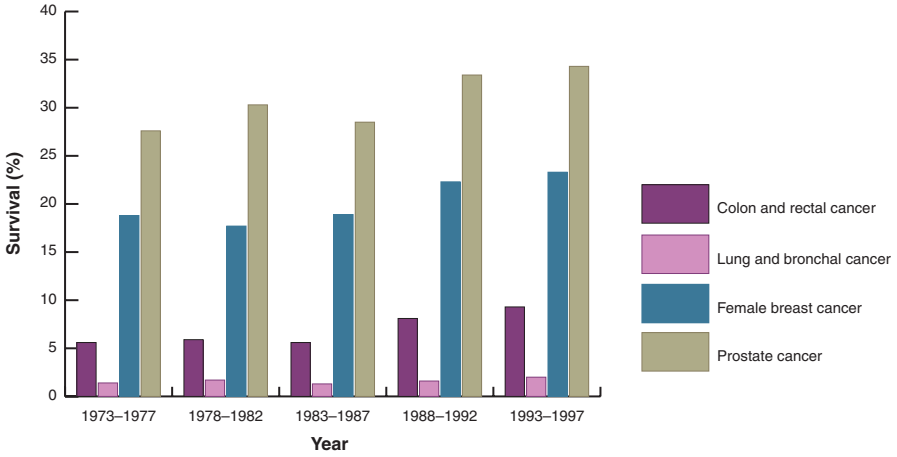

Figure 1 Glacial progress. National Cancer Institute data of 5 year relative survival for patients with distant metastases of colorectal, lung, breast and prostate cancer.

The clinical track record

Cancer research has consumed hundreds of billions of dollars to date[3] and yet the main killers—breast, prostate, lung and colorectal cancer—are essentially as deadly as ever[4] (see **Fig. 1**). Despite the glacial progress in treatment and the advent of 'molecularly targeted' therapy, cancer research continues to focus myopically on individual oncogenes, tumor suppressors and repair genes[5], with little effort devoted to alternative mechanisms and targets[6–8].

Although conventional chemotherapeutic agents remain the first-line treatment of choice, newer molecularly targeted therapies are now reaching the market. Thus far, however, these therapies have had very limited success against solid tumors, which after all make up 90% of all cancers. Success has been largely restricted to rare leukemias; for example, imatinib mesylate (Gleevec; Novartis, Basel) has initially proven effective in patients with chronic myelogenous leukemia (CML).

Whereas the initial clinical success of imatinib in CML was spectacular, this has not been repeated in most of the succeeding cancer therapies against solid tumors. Gefitinib (Iressa; AstraZeneca, London)-based treatment shrinks tumors in only about 10% of advanced non-small cell lung cancer patients[9]. A recent study in one very small (35) patient group indicated that trastuzumab (Herceptin; Genentech, S. San Francisco) induces a partial response in only 23% of individuals with advanced *HER-2/neu*-overexpressing breast cancers[10]; early indications are that bevacizumab (Avastin; Genentech, S. San Francisco) is not much better in colon cancer. All of these agents have serious associated toxicities[11–13], most extend patient survival only by a matter of months and there is a variable period of remission before a resistant form of the cancer returns, even in the case of imatinib[14].

In the light of these findings, the concept of intervening in cancer networks at a single 'oncoprotein' or 'tumor suppressor protein'

has thus far had decidedly mixed results when translated to the clinic. This should not be surprising as the majority of solid tumors are characterized not by single gene-based events, but by multiple genomic alterations, which are specific to each tumor in an individual.

Correlations and mutational data

Analysis of somatic mutations in disseminated single tumor cells from the bone marrow of breast cancer patients reveals that contrary to dogma, mutation of *TP53* (also known as *p53*) is not an early event in systemic breast cancer, but may occur in some individuals later during metastatic progression[2,15,16]. Likewise, whereas mutations in the phosphatidylinositol 3-kinase gene have been documented in 74 out of 199 colorectal tumors and 4 out of 15 glioblastomas, they are found in only 1 out of 12 breast cancers, 1 out of 24 lung cancers, none out of 11 pancreatic cancers and none out of 12 medulloblastomas[17]. In the case of the epidermal growth factor receptor (EGFR), somatic mutations are found in 1 out of 61 non-small cell lung cancer tumors from US patients, but in 15 out of 58 Japanese patients with the same cancer[18]. Furthermore, although somatic mutations in *CDC4* are claimed to be a chief cause of chromosomal instability in cancers generally[19], only 22 out of 190 colon tumor samples have these mutations. It is not known which of the above mutations existed in a homo-, hetero- or hemizygous condition within particular cells in these samples—a prerequisite for functional and causal interpretations. Thus, correlations remain weak and the above data are equally consistent with the majority of these mutations being innocent bystanders in the processes of tumor progression.

There is enormous phenotypic variation in the extent of human cancer phenotypes, even among family members inheriting the same mutation in the adenomatous polyposis coli (*APC*) gene believed to be causal for colon cancer. In the experimental mouse knockout of the catalytic gamma subunit of the phosphatidyl-3-OH kinase, there can be a high incidence of colorectal carcinomas or no cancers at all, depending on the mouse strain in which the knockout is created, or into which the knockout is crossed[20]. Finally, the experimental centerpiece on which theories of human oncogenesis are based, namely that overexpression of only two oncogenes and a telomerase catalytic subunit is sufficient to create a malignant human tumor[21], simply does not hold up experimentally for a diploid human cell[22].

Massive genomic imbalances and methylation

The Human Genome Project used diploid genomes in which the principles of mutation, reversion, suppression and methylation are reasonably understood, and in which the transcriptomes, proteomes and higher-order systems are characteristic of particular cell types at equilibrium. In contrast, all the genomes of the thousands of solid tumors that have been examined microscopically or molecularly depart from diploidy[23–25], as do multiple myelomas[26]. Some regions of a cancer genome are differentially amplified, others are deleted and many are rearranged.

Solid tumors consist of a heterogeneous population of aneuploid and/or segmentally aneuploid cells, where spontaneous mitotic nondisjunction at a single cell division can inexorably change the dosages of thousands of genes, microRNAs as well as other noncoding entities and the dosage-mediated interactions of thousands of noncoding single nucleotide polymorphisms, all without any mutational input whatsoever. Equally important, cancer genomes undergo massive changes in hyper- and hypomethylation[27] leading to large alterations in gene activity. These clinically profound genome-wide methylation changes in genes and regulatory regions can occur completely independently of mutation.

This huge departure from diploidy combined with a changing methylome introduces novel properties to cellular networks, such that conventional interpretations of phenotypic change through mutations have little traction. Transcriptomes emanating from variably aneuploid genomes are subjected to perturbations that far exceed anything that single gene effects can muster. The descriptions of homo-, hetero- and hemizygosity, dominance and recessiveness, and neomorphic, antimorphic, hypermorphic and hypomorphic alleles lose their conventional meaning because allelic and methylomic dosages are mere parameters in perturbed networks where the key is network flux.

It is in such aneuploid contexts that multidrug resistance rapidly develops, even when major multidrug resistance genes are experimentally deleted from a genome[28]. As deficiencies involved in loss of heterozygosity cannot revert, and because back-mutation frequencies are low, conventional mutational interpretations of the rapidity of multidrug resistance and reversion do not hold[29].

The real killer—metastasis

Primary tumors are heterogeneous at all levels[30,31] and many remain dormant[32]. The heterogeneity is illustrated by microarray data showing that the correlation among the expression profiles for three different parts of the same kidney tumor is poor[33]. The heterogeneity among tumors from the same individual is also extensive. In prostate cancer, there are often many distinct foci in the same prostate, only one of which may be invasive and have a deleterious effect on the patient[34]. Thus sequence data derived from a primary tumor are problematic because the important functional variation between cells is obliterated. Single-cell data are far more informative.

Analysis of single disseminated tumor cells after curative resection of the primary breast cancer reveals that disseminated cells can exhibit changes completely different from those observed in the primary tumor[15]. Similarly, when bone marrow micrometastases are compared with the primary colorectal tumors from the same patient, cells disseminated to the bone marrow do not always carry the same K-*ras* mutations as the primary tumor[35].

Money well spent?

No one doubts that primary tumors accumulate somatic mutations over time. However, the Achilles' heel of cancer is not the mutational baggage train of the primary tumor, but the genomic imbalances and methylation changes of the deadly cohort of cells that metastasize in different genetic backgrounds. As a megaproject in advancing cancer research and ultimate cures, the human cancer genome project thus is fundamentally flawed.

First, as the mutational spectra of the primary tumor and its metastatic derivatives may only partially overlap, therapeutic strategies developed for specifically targeting mutations in primary tumors are unlikely to eradicate cells that have already left the primary tumor and are evolving along different genomic trajectories. The clinically significant entity is not the primary tumor *per se* but the rare cells within it that give rise to metastases and the particular genetic background within which they occur[36–40].

Second, there is growing evidence for the profound clinical effects of genome-wide methylation changes in genes and regulatory regions[41]. These changes can take place completely independently of oncogenic or tumor suppressor or mismatch repair mutations and would not be detected by a human cancer genome sequencing effort.

Thus, although a mutation-cataloging research megaproject may be a diverting occupation for sequencing centers and gene hunters, leading scientists should think carefully before they tout its therapeutic promise to patients and politicians. The simple truth is that the money would be much better spent

if research priorities were reevaluated. A good place to start would be to dismiss the fallacious notion that single mutations in primary tumors are the optimal starting point for research that would lead to the discovery of new, more effective cancer drugs. The clinical reality is that it is not single genes, but rather the properties of aneuploid-based methylated networks that allow metastatic cancer cells to explore novel niches in different genetic backgrounds and to rapidly become resistant to drug-based therapies.

1. Pollack, A. *New York Times*, March 28**,** A1 (2005).
2. Klein, C.A. *Adv. Cancer Res.* **89**, 35–67 (2003).
3. Leaf, C. *Fortune*, **149**, 76–97 (2004).
4. SEER Program (www.seer.cancer.gov) SEER* Stat Database; SEER 18 Regs, Nov 2004 Sub (1973-2002 varying), National Cancer Institute, DCCPS, released December 2004.
5. Vogelstein, B. & Kinzler, K.W. *Nat. Med.* **10**, 789–799 (2004).
6. SEER Program (http://www.seer.cancer.gov) SEER* Stat Database; SEER 18 Regs, Nov 2004 Sub (1973-2002 varying), National Cancer Institute, DCCPS, released December 2004.
7. Sonnenschein, C. & Soto, A.M. *Mol. Carcinogen.* **29**, 205–211 (2000).
8. Harris, H. *Nature* **427**, 201 (2004).
9. Marx, J. *Science* **304**, 658–659 (2004).
10. Mohsin, S.K. *et al. J. Clin Oncol.* **23**, 2460–2468 (2005).
11. Hurwitz, H. *N. Engl J. Med.* **350**, 2335–2342 (2004).
12. Ozcelik, C. *et al. Proc. Natl. Acad. Sci. USA* **99**, 8880–8885 (2002).
13. Bendell, J.C. *et al. Cancer* **97**, 2972–2977 (2003).
14. Hofmann, W.-K. *et al. Lancet* **359**, 481–486 (2002).
15. Schmidt-Kittler, O. *et al. Proc. Natl. Acad. Sci. USA* **100**, 7737–7742 (2003).
16. Klein, C.A. *Cell Cycle* **3**, 29–31 (2004).
17. Samuels, Y. *et al. Science* **304**, 554 (2004).
18. Paez, J.G. *et al. Science* **304**, 1497–1500 (2004).
19. Rajagopolan, H. *et al. Nature* **428**, 77–81 (2004).
20. Barbier, M. *et al. Nature* **413**, 796 (2001).
21. Hahn, W.C. *et al. Nature* **400**, 464–468 (1999).
22. Akagi, T. *et al. Proc. Natl. Acad. Sci. USA* **100**, 13567–13572 (2003).
23. Pollack, J.R. *et al. Proc. Natl. Acad. Sci. USA* **99**, 12963–12968 (2002).
24. Aneuploidy Conference Abstracts. *Cell. Oncol.* **26**, 171–269 (2004).
25. Duesberg, P. *et al. Cell Cycle* **3**, 823–828 (2004).
26. Fonseca, R. *Blood* **102**, 2562–2567 (2003).
27. Rush, L.J. *Blood* **97**, 3226–3233 (2001).
28. Duesberg, P. *et al. Proc. Natl. Acad. Sci. USA* **98**, 11283–11288 (2001).
29. Duesberg, P. *et al. Proc. Natl. Acad. Sci. USA* **97**, 14295–14300 (2000).
30. Al-Hajj, M. *et al. Proc. Natl. Acad. Sci. USA* **100**, 3983–3988 (2003).
31. Al-Hajj, M. *et al. Curr. Opin. Genet. Dev.* **14**, 43–47 (2004).
32. Folkman, J. & Kalluri, R. *Nature* **427**, 787 (2004).
33. Vasselli, J.R. *et al. Proc. Natl. Acad. Sci. USA* **100**, 6958–6963 (2003).
34. Masters, J.R.W. & Lakhani, S.R. *Nature* **404**, 921 (2000).
35. Tortola, S. *et al. J. Clin. Oncol.* **19**, 2837–2843 (2001).
36. Fidler, I.J. & Kripke, M.L. *Nat. Genet.* **34**, 23 (2003).
37. Hunter, K. *et al. Nat. Genet.* **34**, 23–24 (2003).
38. Ramaswamy, S. *et al. Nat. Genet.* **34**, 25 (2003).
39. Dick, J.E. *Proc. Natl. Acad. Sci. USA* **100**, 3547–3549 (2003).
40. Kondo, T. *et al. Proc. Natl. Acad. Sci. USA* **101**, 781–786 (2004).
41. Egger, G. *Nature* **429**, 457–463 (2004).