The Latest Surge in the War on Cancer; Tour de Force or Tour de Farce?

George L Gabor Miklos PhD
Secure Genetics
www.securegenetics.com

Phillip J Baird MD PhD
Integrated Diagnostic Pathology

The March issue of Scientific American features an article by Francis Collins, head of the National Human Genome Research Institute and Anna Barker, deputy director of the US National Cancer Institute, setting out how to wage a smarter War on Cancer (1). This 'new' initiative, packaged as The Cancer Genome Atlas (or TCGA for cute), aims to catalogue all the somatic mutations in selected lung, brain and ovarian cancers as a $100 million pilot project and then anticipates moving on to catalogue mutations in another 50 or so cancers, provided further billions of dollars of taxpayers money is forthcoming. This initiative was originally highlighted as the Cancer Genome Project in the *New York Times* (2) and in *Nature Biotechnology* (3), and is the brainchild of Eric Lander of the Broad Institute (4). When in full production, TCGA anticipates sequencing the DNA of over 12,000 tumor samples to reveal all the DNA mutations within them, the underlying assumption being that the generation of tera bytes of cancer DNA sequence data will seamlessly equate with drug-based therapeutic empowerment. The reasoning goes as follows. By knowing all the mutations in a primary tumor sample from a patient, a physician will be able to prescribe a cocktail of drugs together with a chemotherapeutic regimen which *is specific for that individual*. Any new drugs will be intelligently designed by pharmaceutical companies to match the alterations in the defective proteins that arise from these DNA mutations in primary and metastatic tumors. In a nutshell, this is the Holy Grail of personalized cancer medicine.

In their article, Collins and Barker chart the challenges that lie ahead by using the analogy of the original 1804-1806 Lewis and Clark expedition to explore the unknown Northwest territories on orders from President Jefferson. Collins and Barker seek gravitas and the high ground by highlighting the input of Nobel Laureates who look favorably upon their endeavor. However, reputations count for little in any war and the War on Cancer is no exception. This one is long running, shows little sign of being won and has continually raised and then dashed the hopes of cancer patients. By contrasting the practicalities of the bedside with the pure research efforts, we illustrate how TCGA could benefit from more cerebral input prior to such profligacy of DNA sequencing. Readers can judge for themselves whether the strategy put forward in the Scientific American article is indeed as momentous as claimed, or whether this is another Roadmap within the Land of Unfulfilled Promise.

The TCGA Roadmap

Some truly astonishing statements emanate from senior members of the US cancer community and continue to achieve prominence in *Science* and the tabloids. The outgoing director of the National Cancer Institute, Andrew von Eschenbach put forth a plan by which **suffering and death from cancer will be eliminated by 2015** (5). Harold Varmus, a previous Director of the NCI, has further upped the stakes by stating that The Cancer Genome Atlas aims to provide the basis for **all future studies of cancer in the laboratory and the clinic** (6). If the foregoing was not breathtakingly expansive enough, then Anna Barker's statements as the acting deputy director of the NCI are truly astounding. She describes TCGA as not only **a turning point in biomedical research** ….but…a **turning point for medicine.** She defends the goal of eliminating suffering and death from cancer by 2015 arguing that it is achievable if scientists **think outside the box**. Political support derives from the 372 members of Congress who have written to President Bush supporting the 2015 timeline. Thus, there is now a plan, a timetable *and* an exit strategy for the War on Cancer. What more could one hope for over the next eight years? Well for one thing, some clinical and therapeutic realism.

Fortunately a few cooler heads still remain and these stratospheric marketing statements can be more calmly evaluated at ground level. On the basis of signal to noise problems with data dimensionality, Lee Hood has described TCGA proposals as *naïve*. David Baltimore has pointed out that it is *dangerous to bet on future discoveries. Particularly in the area of cancer, it could be very dangerous*. Craig Venter, whose first hand knowledge of bringing diverse genome projects to successful conclusions is unrivalled, muses that *it's unclear what answers one may get*.

Spin doctors versus real doctors

Masses of mutational data are not criteria of immediate use to a physician; it is understanding them in a therapeutic context that would be momentous.

A number of publications will assist readers to view the claims of TCGA in a more realistic light (7, 8, 9, 10). All of them clarify why this War on Cancer is stuck in a quagmire of its own making. Despite the media hype from the spin doctors of the research establishment, as against the real doctors treating patients with radiation, chemotherapy and drugs, the generals and politicians of TCGA persist with the same restricted mindset as their military and civilian counterparts and pay lip service to the important therapeutic issues while pursuing pure research avenues largely tangential to the direct therapeutic benefit of the patient. The knub of the problem is that tumors are heterogeneous collections of cells which together contain millions of genomic perturbations all the way from aberrantly methylated bases to gross genomic imbalances at the chromosomal level. The central issue of clinical genetics and of cancer has always been the following. How do we sort the mutations of very different types which *may* produce a clinical outcome *in the unique genetic and epigenetic background of that particular individual*, from those flotsam and jetsam mutations which are clinically asymptomatic. This problem is alluded to by Collins and Barker, but then glossed over, in the headlong rush for further data accumulation.

The first step towards any therapeutic goal is to ensure that the search is conducted in the correct area. Pivotal to this is understanding that 90% of all deaths from cancer are not due to the primary tumor *per se* (11). It is the departure of certain cells from the primary tumor, their invasion of distant organs and the formation of multiple secondary cancerous growths that is the killer. These invading cells form metastatic growths which destroy the surrounding tissue, compromise organ function and result in the death of the patient (10). Most importantly, however, only 1 in 50,000 or so of the cells in a primary tumor ever develops the requisite genomic alterations to emigrate and embark upon a journey which will leave its descendants in far off anatomical locations. It is these rare rogue cells that pose the danger, not the cells that make up the bulk of the primary tumor itself. These so-called cancer stem cells have now been described in breast, brain, colorectum, pancreas, prostate, skin and in tissues of the head and neck. Alan Bernstein, the President of the Canadian Institute of Health Research, Canada's major medical funding agency, has contrasted the cells in the bulk of the primary tumor on which most cancer research has been carried out, with the rare cancer stem cells in such a tumor; ….*all this time, the 30 or 40 years that chemotherapy and radiation have been around, we've been going after the wrong cells* (12).

Second, except for aggressive brain cancers where special problems arise in an organ strongly fortified by a blood brain barrier and further constrained within a braincase, most of the cells in a primary tumor never leave the primary site. Thus if a tumor is detected before any of its cells have left, then surgical removal will leave a patient completely cured of that cancerous growth.

Third, pathological analyses, particularly at autopsy, invariably confirm that tumors are heterogeneous and can contain a mixture of cancerous as well as benign foci (10). In prostate and cervical cancer, for example, there are often many foci, ***only one of which may ultimately turn out to be invasive*** (13). As all practicing pathologists will attest, the practicalities of finding and then sampling each and every different focus in a single prostate gland or in various cervical lesions, and prescribing a personalized drug cocktail on the basis of what may, or may not, turn out to be the most dangerous focus, is clinically unrealistic.

Thus analyzing an entire solid tumor for all its mutations, as codified by the NCI and the NHGRI, is close to an exercise in futility. At the genomic level, most of the millions of coding, noncoding and epigenetic alterations found in the heterogeneous cell population will be irrelevant to metastasis, as most of these mutations do not originate from the rare subpopulation whose descendants disseminate to establish metastatic foci in distant organs. The deadly *combination* of genomic alterations that *do* occur in the rare subpopulation will be comprehensively obliterated by the avalanche of phenotypically benign mutations from the bulk of the tumor. Since there is no bioinformatic way of determining which *combinations* of mutations came from which cells (since the whole tumor is used for DNA sequencing), attempts to extract meaningful clinical information from a system where the noise overwhelms the signal by a factor of between 10,000 and 100,000, or higher, borders on legerdemain. What is the point of sequencing every piece of altered genomic flotsam and jetsam in the overwhelming majority of tumor cells whose impact on the individual patient is largely inconsequential? Yet flotsam and jetsam mutational cataloging is precisely what the NCI and the NHGRI have so ardently and assiduously embarked upon.

A realistic therapeutic test of TCGA strategy

As proof of principle of the impending power of The Cancer Genome Atlas, Collins and Barker cite the largest cancer DNA mutation data set which has just been published by a consortium of prominent cancer laboratories, those of Vogelstein, Kinzler and Velculescu (14) and hailed by some as a Tour de Force. In this study, samples from patients with colorectal cancer having metastases to the liver, lung and lymph nodes and patients with primary breast tumors, as well as cell lines and xenografts, were exhaustively sequenced for somatic mutations. Hundreds of thousands of mutations from these diverse cell populations were catalogued, bioinformatically processed and winnowed down to yield 189 genes that were mutated at significant frequency. It is these genes that are held to provide new targets for diagnostic and therapeutic intervention, yet none of them have a causal connection to clinical phenotype in *an individual patient*. These data on selected coding regions represent only a few percent of the genome and are still to be supplemented by mutations from the remaining 97% of the genome as well as from the yet to be added epigenetic "mutations" in which the DNA base sequence *per se* is not changed in the tumor. Such methylation changes in coding, regulatory and noncoding regions are pervasive in all tumors and their already demonstrated clinical effects are considerable. In total, there will be millions of somatic mutations in a single primary tumor of which the important, but totally unknown fraction, will be in the subpopulation that comprises only 1 cell in 50,000 or so of the tumor sample. Mutations in this rare subpopulation, which will provide the engine rooms of dissemination and metastasis, can only be triaged by prior fractionation of tumor cell populations. In such a situation, bioinformatic prioritizations from a whole tumor have little predictive therapeutic value as regards *the individual patient*.

A far more important finding from these data, which is not emphasized either by the above authors or by Collins and Barker, is that the mutational signature of each patient is unique (14), a finding which has enormous implications for drug treatment. The promise of TCGA is personalized therapeutics and as such, predictions of appropriate therapeutic regimens should follow smoothly from such data. Velculescu describes a future where a cancer patient comes into a clinic, the tumor is analyzed and the treatment *is based on a spectrum of mutations with a cocktail of drugs. It doesn't mean a new drug for each person, just a different combination of drugs* (15). If TCGA does indeed have therapeutic utility, it should be a straight forward matter to match the mutational spectra generated by the Vogelstein, Kinzler and Velculescu laboratories with an already approved FDA drug combination for each patient. However, the FDAs Orange Book, largely for small molecule drugs, and the Center for Biologics Evaluation and Research, for biological drugs, reveal that there are only 324 drug targets for *all* classes of therapeutic drugs not just those used in oncology (16). Since there will be billions of unique mutational signatures from tumor samples of millions of cancer patients, the 1 to 1 matching of a given mutational spectrum of a patient to a specific therapeutic drug regimen is a pipedream.

It is hardly a surprise therefore, that no therapeutic predictions for the individual patient have been forthcoming from this extensive data set; the authors remain strangely silent on the *only* issue of relevance to their data. Such reticence is understandable. At the only place where it really matters, namely a physicians' treatment of the individual cancer patient, these much vaunted data do not pass the most elementary scientific requirements of testable predictions. Having amassed the data, what are the therapeutic predictions flowing from them? What *dosages* of FDA approved drug combinations and chemotherapeutic regimens would a practicing physician prescribe for any of the colorectal and breast cancer patients whose mutational signatures derived from the bulk of tumors, rather than the dangerous so-called stem cell population? Each of us metabolize different drugs at very different rates in different tissues owing to our different genetic and epigenetic backgrounds. We also carry different combinations of Single Nucleotide Polymorphisms and other larger genomic imbalances that interact in nonlinear ways to produce responses to drug dosages. How then does a physician begin to determine dosages and drug interactions at this level of *Systems* complexity? Whatever else, it is the uniqueness of each patient that is the stumbling block to personalized therapy. The fact that robust therapeutic predictions have not emerged from the mutational data is a telling indictment of TCGA strategy. TCGA is certainly keeping sequencing centers busy, but offering little in the way of therapeutic clarity. The inconvenient truth is that without precise therapeutic predictions that *alter the treatment of the individual patient*, TCGA is running on vapor. Its blinkered strategic vision is one of congealed inflexibility; it conforms to the dictum, *when your favorite tool is a hammer, every problem looks like a nail.*

Genetics 101

Conventional interpretations of cancer are oblivious to a number of fundamental genetic constraints. Genetics 101 places limits on the number of genes that can be mutated in a diploid genome and still leave a viable somatic cell. In a diploid genome, mutations, deletions, amplifications, rearrangements, methylation, genomic imprinting, position effect variegation and transvection, are all perturbations which conform to reasonably understood molecular rules. However, neither the cells which leave the primary tumor, nor the metastatic

cancerous growths in distant organs, any longer possess pristine diploid genomes and therein lies the rub. This very dangerous minority of tumor cells are genomically very different from the bulk of the stay-at-home cells. This is illustrated by the single cell analyses of Riethmuller and Klein and their colleagues on breast cancer patients whose cells have disseminated to the bone marrow or lymph nodes, as well as their analyses of patients with cancer of the unknown primary (17, 18). Additionally, descendants of the biochemically isolated so-called stem cells that have been isolated from solid tumors have all been found to have grossly distorted genomic contents (19, 20). The take home message is that the cells which leave a primary tumor consist of a heterogeneous population of cells with varying combinations of translocations, inversions, chromosomal as well as extrachromosomal amplifications, segmental aneuploidy and assuredly, aberrant genomic methylation. The rules by which somatic mutations can exert their influences on phenotype, when they are in imbalanced genomes, are radically different to the rules which apply in diploid genomes. The only genome wide dissection of a complex multicellular organism at the subchromosomal dosage level clearly illustrates the systematic phenotypic consequences of segmental aneuploidy in a defined genetic background (21). Conventional dominance and recessiveness have little meaning in imbalanced genomes, as it is the perturbed *flux* through reorganized networks at different dosages that is now the major driving force of phenotype. So also in a therapeutic context, a diploid cell population will respond quite differently to drug treatment compared to a genomically heterogeneous cell population consisting of imbalanced genomes (22, 23).

The reality is that Systems-based approaches predicated on massively imbalanced genomes, at genetic as well as epigenetic levels, need to be embraced in order to predict therapeutically relevant outcomes (3, 24). Such approaches involve accepting that cancer is first and foremost a disease of differentiation, not cell multiplication, the compelling data for which have been carefully summarized by Harris (9). Underpinning these issues is the unique nature of disseminated and metastatic cancer genomes and their consequences for drug resistance. These involve an understanding and emphasis of the effects of drugs on *network fluxes*, and less so on their individual components. Drug resistance as commonly interpreted in the context of diploid genomes and somatic mutations is simply obsolete in a clinical context (22, 23).

Dream come true or continuing nightmare?

The existing data pertaining to TCGA have already demonstrated that therapeutic solutions will not easily be forthcoming by pouring billions of taxpayers dollars into the mind numbing analysis of individual mutational components stemming from tumors that are a heterogeneous mixture of imbalanced and differentially methylated genomes. No matter what fashionable halo is attached to TCGA, or by how many institutions and high profile individuals the TCGA is anointed, the only therapeutic yardstick of progress in the solid tumor area is a reduction in metastasis-based body count. These data are the bottom line and nothing else cuts the mustard. If TCGA stays the present course, Collins and Barker estimate that it will take the equivalent of 10,000 Human Genome Projects to elucidate the standard mutational spectra of sundry tumors. We assume that TCGA will probably embark upon another surge of 10,000 Human Epigenome Projects to elucidate the standard epigenetic mutational spectrum. These figures should give serious pause for a more reflective analysis of whether the current strategy of TCGA is therapeutically redeemable as it now stands, or is an expensively misplaced moon shot. This excruciatingly painful examination of all the mutations in a tumor is akin to examining each tapestry in the Louvre, thread by thread, in an increasingly desperate search for the essence of art. The proponents are reluctant to actively engage with the therapeutic issues of the elimination of the so-called stem cell population and the issues of drug resistance which after all, are the only major issues of relevance to patients. Cancer genomes harboring multiple deficiencies, duplications and variously methylated regions are not amenable to facile high throughput sequencing and yet they may contribute as much or more to metastasis as the single base pair changes that are the predominant target for TCGA.

The fact that most metastatic cancers are almost as deadly today as they were 50 years ago, is information that is publically available from the NCIs own databases and examples are illustrated in *Fortune* and *Nature Biotechnology* (3,7). We reiterate that it is when those rare genomically heterogeneous cells in a solid tumor begin to emigrate that the serious problems really begin. Dissemination can occur very early when tumors are only millimeters in diameter and probably even before this size is reached (18). The clinical data of Tarin, Riethmuller and Klein, as well as the follow up data from more than 300,000 cancer patients in the Munich Cancer Registry gathered over a 20 year period, unambiguously demonstrate that the majority of cells in a tumor never acquire the ability to either disseminate or to metastasize; this property is only achieved by the rare so-called stem cells now being described in a plethora of cancers. From the vantage point of the surgeon and the pathologist, it is indeed axiomatic that *if* the majority of cells in a primary tumor could metastasize, as inferred from microarray–based studies (25), there would be no residual tumor in its original location (10).

The clinical and so-called cancer stem cell data clearly expose the danger of relying on bioinformatic analyses, as Lander and colleagues have done, to extrapolate from microarray data to clinical importance (25). Even though their primary and metastatic tumor samples did not even come from the same patients, these authors dismissed the notion that important signatures arose from rare cells within the primary tumor and instead placed their faith in signatures from whole tumors. However, cancer signatures from whole tumors are notoriously fragile; a fact which has been well documented by leading statisticians and mathematicians such as Tibshirani and Domany (26, 27). By contrast, clinical and pathological observations on patients and insights from the so-called cancer stem cell compartment, are far more informative than the artefact-prone results from expression-based analyses of whole tumors. Readers will now better appreciate that the therapeutic foundation of The Cancer Genome Atlas was predicated on a false premise; that the mutational signature of the bulk of the primary tumor would be congruent with its metastatic derivatives. The perplexing issue is that primary and metastatic tumors were already known to differ markedly at the gross genomic level years before the TCGA was proposed. Therefore its proponents are either still oblivious to a large amount of classical genetic, cell biological and clinical data which impinge on somatic processes, or choose to ignore them. The patient data which show that primary tumors and disseminated cells have different signatures at the single cell level (17, 18), reduce the TCGA to a weak force. Prescriptions of *personalized* drug combinations based on signatures from the bulk of a primary tumor are an illusion, whereas those from the so-called stem cell population still offer some hope.

In a therapeutic context, the description by Collins of The Cancer Genome Atlas as a *dream come true* and the call to arms by Collins and Barker of imploring scientists to *think outside the box*, is fatuous. We shall be much more enthusiastic about encouraging thinking outside the box, when there is concrete evidence of any thinking going on *inside* it.

References

1. Collins, F. S. & Barker, A. D. *Sci. Am*. March 50-57 (2007).
2. Pollack, A. *New York Times* March 28 A1 (2005).
3. Miklos, G.L.G. *Nat. Biotech*. **23**, 535-536 (2004).
4. Kaiser, J. *Science* **309**, 693 (2005).
5. Kaiser, J. *Science* **310**, 29 (2005).
6. Varmus, H. & Stillman, B. *Science* **310**,1615 (2006).
7. Leaf, C. *Fortune* **149**, 76-97 (2004).
8. Bialy, H. *Oncogenes, Aneuploidy and AIDS* (2004).
9. Harris, H. *BioEssays* **27**, 833-838 (2005).
10. Tarin, D. *Breast Disease* **26**, in press (2007).
11. Christofori, G. *Nature* **441**, 444-450 (2006).
12. Abraham, C. *globeandmail.com*; november 20 (2006).
13. Masters, J. R. W. & Lakhani, S. R. *Nature* **404**, 921 (2000).
14. Sjoblom, T. *et al*. *Science* **314**, 268-274 (2006).
15. Beckman, M. *J. National Cancer Institute* **98**,1758-1759 (2006).
16. Overington, J. P. *Nature Reviews Drug Discovery* **5**, 993-996 (2006).
17. Riethmuller, G. & Klein, C. A. *Cancer Biology* **11**, 307-311 (2001).
18. Klein, C. A. & Holzel, D. *Cell Cycle* **5**, 1788-1798 (2006).
19. Singh, S. K. *et al*. *Cancer Research* **63**, 5821-5828 (2003).
20. Galli, R. *et al*. *Cancer Research* **64**, 7011-7021 (2004).
21. Lindsley, D. *et al*. *Genetics* **71**, 157-184 (1972).
22. Duesberg, P. *Proc. Natl. Acad. Sci*. **97**, 14295-14300 (2000).
23. Duesberg, P. *Proc. Natl. Acad. Sci*. **98**, 11283-11288 2001).
24. Stock, R. & Bialy, H. *Nat. Biotech*. **21**, 13-14 (2003).
25. Ramaswamy, S. *et al*. *Nature Genetics* **33**, 49-54 (2003).
26. Ein-Dor, L. *et al*. *Proc. Natl. Acad. Sci*. **103**, 5923-5928 (2006).
27. Tibshirani, R. *N. Engl. J Med*. **352**, 1496 (2005).