Cancer Genome Sequencing in Wonderland Ad Infinitum

The most extensive results to date concerning the repertoire of point mutations in cancer samples revealed by gene sequencing is out in the 29-author treatise published by one of the science industry's Holy Trinity (CSN--Cell, Science, Nature) (*The Consensus Coding Sequences of Human Breast and Colorectal Cancers,* Science 314:268, 2006). The study costing about $5 million is an analysis that shows that the number of mutations in products of about half of the estimated genes of the human genome from eleven cell lines or xenografts from breast or colorectal cancers exceeds a mind boggling 800,000+ compared to a normal sample. Mutations are errors in the four letter alphabet that spells out in words of three the sequence of the 20 amino acid units that make a functional protein. Thousands of functional proteins working together at wide ranges of activity due to thousands of forces that modify them make up a healthy cell. Xenografts are chunks of cancer tissue removed from patients and selected for their ability to survive and grow in the artificial environment under the skin of a mouse without being attacked by the mouse's immune system so as to provide enough material to analyze. Cell lines are mixtures of cancer cells selected one step further for their ability to survive and grow in a Petri dish like bacteria independent of their existence in tissue and even the artificial, but more close to physiological environment in the mouse.

Mutations thought to be methodological artifact, sequence differences between the two normal tissue samples, and "silent mutations" that don't change the amino acid sequence of protein products were subtracted. [Although each amino acid of a protein is spelled out by the sequence of three letters of a four letter alphabet, the third letter of each triplet for an amino acid can be variable referred to as codon degeneracy. A mutation to one of the variants that can code for the same amino acid is termed by the authors a "silent mutation"). It was concluded that a total of 1307 mutations that could affect a protein amino acid sequence occurs in 1149 genes in the samples of the two cancers.

When the sequence of the 1149 mutated genes from the cell lines and xenografts was examined in 24 additional tissue samples of tumors taken directly from patients without xenograft or cell culture, an additional 365 mutations in 236 genes was detected showing that different tumor samples have different numbers and types of mutations. After all this, 921 and 751 mutations of potential interest (MOI's) in the breast and colorectal tumor samples, respectively, for a total of 1672 cancer MOI's were noted. Recall the above numbers need to be doubled since only about half of the suspected coding sequences in the genome were examined. This is an enormous number of MOI's to begin to sort through concerning questions of whether any one or combination have anything to do with the individual cancers sampled much less other samples of the same types of tumors.

To grasp the sheer magnitude of observed mutations detected in a cancer sample one should add back a substantial portion (about 30% of the original 800,000 plus) of the "silent mutations". "Silent mutations" in this study more properly should be referred to as "synonymous mutations," "silent mutations" in coding sequences of genes (exons). Moreover "synonymous mutations" can

have impact on the final product of genes even though they are predicted not to change amino acid sequence (http://en.wikipedia.org/wiki/Silent_mutations).

The large number of these types of mutations in which conservation of amino acid sequence and protein function is preserved speaks well for the evolution of a single cancer according to Darwinian principles. Cancer cells can only tolerate so many mutations; they have to survive to be observed. Perhaps it is these gene products that should be examined as the "Achilles heel" of cancer cells if they are so important as to be intolerant of a mutation that alters the amino acid sequence of the protein product. On this note one wonders how many of the apparently "normal" sequences were mutated at one time and reverted to normal during progression to the stage sampled.

Moreover, current studies this one being of greatest magnitude so far only consider mutations in areas of the genome that code for protein products. The great majority of the rest of the genome sequence in which a proportional number of mutations are predicted to occur per amount of genome has not even begun to be approached. Mutations in noncoding areas may have equal or greater impact on consequence to cancer cells as those in coding sequences.

To attempt the impossible task of reducing this mind boggling number of mutations to some sort of meaning in terms of causality, the authors went on to try to rank mutations in terms of genes that might be involved in a tumor property described in the literature in single gene functional studies (CAN-genes for cancer candidate genes) in various tumor models or analyses relative to those that are along for the ride ("passenger mutations"). Mutations were found in genes that have been previously implicated in diverse cancers and cancer models and others that have not been mentioned before. Most important was the author's conclusion that among the 24 breast and colorectal cancer samples examined, the two cancer types exhibit their own signature of mutations and each cancer sample of the same cancer type also exhibits its own signature of CAN-gene mutations and no two individual sample had more than 6 CAN-genes in common. The grapevine of unpublished results from diverse laboratories like the Sanger Institute in England examining diverse subsets of genes across diverse cancers and samples of the same cancers are reporting in with the same results, the diversity and number of mutations among different cancers is astounding.

The current Science report is being hailed as a landmark study, a "tour de force by other cancer scientists" states Science's News section (*First Pass at Cancer Genome Reveals Complex Landscape*, Science 313, 1370, 2006). Rightly so, this report confirms with the most extensive analysis to date (half of the coding gene sequences in the human genome) what the data from hundreds, thousands of prior studies studying lesser numbers of single genes in diverse cancer-related samples varying from one to a few to hundreds predict despite that the authors interpreted their data as a breakthrough in cause of at least some cancers. The genome of cancers is a tangled mess awash with mutations, the number and types of mutations vary with the individual type of cancer, an individual cancer within a single type of cancer, the time in

progression of an individual cancer within a single type of cancer, how an individual cancer at a specific time in progression within a single type of cancer was sampled (big chunk, little chunk) or manipulated prior to analysis (primary tissue, xenograft or cell line) and location of the individual cancer at a specific time in progression (location within the organ of origin or location outside the organ). Mutations may be happening at a rate in individual cells within a given cancer faster than the cancer can be sampled and sequenced. And the mind boggling number and diversity of mutations revealed in this first study among two types of cancers and a few samples of them focus on parts of the total genome that code for proteins. A proportional number likely occurs all up and down the genome that do not directly have to do with appearance of a protein. These may be equally important to the properties of cancers as what is seen so far.

Although the title of the current subject paper is misleading, e.g. "*The Consensus Coding Sequences of Breast and Colorectal Cancers*," the report is a landmark because the data in it question whether anything can be learned about cancer, its prevention and treatment from mutational analysis. It questions the whole single or collective gene mutation theory of cancer causation at least for what is currently observable as a cancer sample. It should be the basis for calls for new ideas and re-evaluation of where the whole basic science, clinical science and pharmaceutical industry is headed in respect to cancer, its understanding and management based on the collective mutation theory of cancer.

BUT ALAS, observe the completely opposite spin on the data, even by the lead author of the study who privately knows the consequence of the results, both from this study and previous ones and expresses his opinion privately to colleagues. Vogelstein prophetically remarks in the above Science News piece "It will take a long time to unravel all of this, but this is what cancer *is*. It's a much more complex picture than we had anticipated." But instead of denouncing the current juggernaut of "pilot funding" in the billions of dollars and calling for a new direction rather than more of the same sequencing ad infinitum, the Science paper concludes "these data provide new targets for diagnostic and therapeutic intervention, and open fertile avenues for basic research in tumor biology." A conclusion that will likely be repeated in hundreds, thousands of papers reporting the same sort of sequencing data in different cancer samples ad infinitum as the sequencing juggernaut kicks in and expands.

The Science News article states "the results appear to bolster The Human Cancer Genome Atlas, an ambitious $1.5 billion federal project (Science, 29 July 2005, p. 693)." The big project figureheads chime in with "I see this as a big shot in the arm for the argument that this strategy is going to work"—Francis Collins (National Human Genome Research Institute) who is soon to announce a $100 million 3-year pilot project, a pilot project mind you of which the true project could be unlimited. (Well put, Dr. Collins, this endless and mindless sequencing is indeed a futile addiction). Eric Lander another mega-sequencer chimes in with "This is a beautiful demonstration that if you turn over every rock, there is a lot more to be found." (Well put, Dr. Lander, indeed behind every rock

are endless variants in grains of sand, no two of which will likely tell much about the landscape).

Lone voice in the wilderness Harvard geneticist Steve Elledge continues to appeal in behalf of more rationale individual-investigator based projects on the biology of cancers that are much more cost effective in terms of long term conceptual information that is needed to ever hope to understand and attack cancer (Science, 21 October 2005, p.439). But lone voices like Dr. Elledge are quickly neutralized by the powerful bishops of the industry like director of the Sloan-Kettering Instutute and former director of the National Institutes of Health Harold Varmus and the Cold Spring Harbor Laboratory Bruce Stillman (Science, 9 December 2005, p. 1615).

Meanwhile, as mounting billions of dollars gradually feed out across the land for institutional cancer genome sequencing ad infinitum, soon to follow global analyses of the proteome (the far more diverse protein products of genes) and about every other “-ome” one can come up with, and the manipulation of every mutated gene in every way imaginable in a mouse, support of new and fresh ideas on the cause and origin of cancer, both specific ones and in general, gradually decreases. Peer reviewers carefully chosen who are dependent on and guaranteed to chime the dogma increasingly populate review panels to insure that only the repetitive documentation of trivia concerning the POI, personal gene of interest, gets funded. This insures that any young investigator with a fresh idea cannot survive as well as the established ones who dare to propose a fresh idea.

And lastly, what is Pharma to do with the thousands, maybe millions of altered genes and their altered products that are emerging, the “treasure trove” referred to by Dana-Farber Cancer Institute Ronald DePinho (co-chair of the advisory committee for the federal Cancer Genome Atlas project? Flip a coin?

****

Just at this moment, somehow or other, they began to run.

Alice never could quite make out, in thinking it over afterwards, how it was that they began: all she remembers is, that they were running hand in hand, and the Queen went so fast that it was all she could do to keep up with her: and still the Queen kept crying `Faster! Faster!'.

The most curious part of the thing was, that the trees and the other things round them never changed their places at all: however fast they went, they never seemed to pass anything. `I wonder if all the things move along with us?' thought poor puzzled Alice. And the Queen seemed to guess her thoughts, for she cried, `Faster! Don't try to talk!'

Not that Alice had any idea of doing that. She felt as if she would never be able to talk again, she was getting so much out of breath: and still the Queen cried `Faster! Faster!' and dragged her along. `Are we nearly there?' Alice managed to pant out at last.

`Now! Now!' cried the Queen. `Faster! Faster!' And they went so fast that at last they seemed to skim through the air, hardly touching the ground with their feet, till suddenly, just as Alice was getting quite exhausted, they stopped, and she found herself sitting on the ground, breathless and giddy.

The Queen propped her up against a tree, and said kindly, `You may rest a little now.'

Alice looked round her in great surprise. `Why, I do believe we've been under this tree the whole time! Everything's just as it was!'

`Of course it is,' said the Queen, `what would you have it?'

`Well, in our country,' said Alice, still panting a little, `you'd generally get to somewhere else -- if you ran very fast for a long time, as we've been doing.'

`A slow sort of country!' said the Queen. `Now, here, you see, it takes all the running you can do, to keep in the same place.

If you want to get somewhere else, you must run at least twice as fast as that!

From Lewis Carroll's Alice in Wonderland, The Garden of Live Flowers