



editorial



Paul Ko Ferrigno

Will systems biology translate into ever higher healthcare costs, or are there savings to be made?

It has been estimated that as much as \$300 billion may have been spent on cancer research since the war on cancer was announced by President Nixon in 1971 [1]. As well as bringing about huge improvements in cancer care – an average of four extra years in life expectancy per cancer patient, totaling 23 million years for the US alone – this has had a huge impact on our understanding of biology. The impact has been both direct (increased understanding of cell cycle control, DNA replication and repair and so on) but also, and perhaps more importantly, indirect impact through an explosion in the availability of technologies that are being used to study life from the level of single molecules to the level of the entire genome.

But did we really need the expensive, labour-intensive, open-ended use of genome-wide interrogation to develop four new multi-gene prognostic tests for breast cancer? This has led to yet more expense as clinical trials are performed to evaluate and compare each of the tests to the others. Is this worth it when a simple appraisal of breast cancer biology allowed the elaboration of a classical four protein immuno-histochemical test that appears to perform just as well. This question is particularly acute in the light of the huge debate that is raging on the cost of healthcare (e.g. [2]). Will systems biology add to these costs, or can it reduce them?

The greatest transformation in cancer biology in the last two decades has been in our genome-wide approach to systems biology, where we now conceive of gene and protein networks, as opposed to individual genes or even pathways, specific to different cancers. Research on breast cancer, in particular, has been a trailblazer in terms of impact and innovation, with a terrifically positive outcome for many women who now are experiencing an 80% survival rate, as opposed to 80% mortality, five years from diagnosis. And yet the means by which this has been achieved – the identification of gene signatures that can be used to guide treatment for patients with breast cancer – highlights one of the great problems that still faces biologists and clinical scientists: How can we efficiently translate the information we get in the test tube, from a tissue culture or from a whole genome microarray into a meaningful understanding of the underlying biology? The key word here is 'efficiently'. There are currently five distinct tests, which each analyse an almost completely distinct sets of genes and proteins, yet all of which are used to the same end: determining which patients with breast cancer belong to the 15% who need adjuvant chemotherapy to prevent early recurrence, and which belong to the 85% who can be spared the burden of a toxic and traumatic treatment that they do not need. Four of the five prognostic tests that are currently used to guide decision-making in breast cancer (Oncotype Dx, PAM50, Mammostrat and Mammaprint) have been developed through open-ended approaches, using empirical algorithms to detect patterns that correlate with disease. The other, IHC4, was developed on the basis of a knowledge of the biology of breast cancer: that it is growth factor (Her2) and hormone (estrogen and/or progesterone) dependent, and that it involves an increase in cell proliferation that can be measured by staining for the Ki67 antigen.

Oncotype Dx was developed from a candidate set of 250 genes that could be measured by PCR from formalin-fixed clinical biopsy samples; the goal was to identify genes whose expression changes correlated to some degree with breast cancer recurrence [3]. The PAM50 gene set was identified using a microarray-based approach to identify genes whose expression was affected by chemotherapy [4]. Mammprint was perhaps the series pathfinder, using DNA microarrays to identify a signature of early recurrence of disease to identify a 70 gene signature [5]. Finally, Mammostrat captures both the beauty of the open-ended approach, and the ease of a guided approach [6]: it used gene expression studies in multiple cancers to identify 700 proteins of interest. They raised and validated antibodies against those proteins, ending with a set of 20 validated antibodies, and were able to build a model using just 5 of them to identify patients at high risk of recurrence using immuno-histochemistry. Altogether, then, four very similar approaches were taken to reach the same goal, a signature that predicts whether a patient is at high risk of cancer recurrence, which nonetheless end up producing surprisingly different results. If they are reporting on the underlying biology, one would predict that of the 130-odd genes or genetic elements on which these signatures report, many would be shared; but this is not the case. There are 5 genes shared between Pam50 and Oncotype Dx, one between Oncotype Dx and Mammprint, and one between Mammprint and Pam50. By extending from genes to related genes of gene ontologies, these numbers can be increased, but in all cases there are many more genes unique to a single test than shared between tests. Perhaps most telling is that, although the Oncotype Dx test does include a look at the genes encoding all the proteins currently measured by the gold standard immuno-histochemical IHC4 test (ER, HER2, PR and Ki67), none of the other gene-based tests do. As biologists, we wonder why this is. As biomedical scientists, we must ask ourselves 'Does this matter?'

A great deal of time and money is now being spent to determine which of these tests should be used. Three of the five tests (IHC4, PAM50 and Oncotype Dx) perform similarly in determining whether patients with ER positive, HER2 negative, node negative breast cancer should be given chemotherapy, or spared its side effects [7]. In the light of these data, we must ask ourselves whether we should be devoting so much effort and resource into

open-ended, 'omic'-style approaches to biomarker discovery, or whether we have been distracted by the glamour of new technologies away from the tried and effective path of turning biological insights into diagnostic tests? When comparing the relatively complex 70-gene Mammprint test to the relatively simple 21-gene Oncotype Dx test, Ross and colleagues speculated that the benefit of broader tests derives from the ability to assess additional pathways and potentially provide additional information [8].

This then must be the challenge when translating system biology data into clinical tests: the resulting assays must be able to report on the underlying biology and pathology (in the sense of malfunctioning biology that is causing the disease), predict the course of the disease *and* predict the best possible course of treatment. Our challenge is the development of tests that combine diagnosis, prediction and prognosis - areas that are traditionally held to be distinct. Ideally these would grow from tests developed as companion diagnostics throughout the drug discovery process. Ultimately, however, such progress can only come from rational application of that shiny new toy: systems biology-wide, genomic and/or proteomic data.

References

- 1 Sun, E.C. *et al.* (2009) An economic evaluation of the war on cancer. In *National Bureau of Economic Research Working Paper 15574*. <http://www.nber.org/papers/w15574>
- 2 Konski, A. (2011) The war on cancer: progress at what price? *J. Clin. Oncol.* 29, 1503–1504
- 3 Paik, S. *et al.* (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* 351, 2817–2826
- 4 Korde, L.A. *et al.* (2010) Gene expression pathway analysis to predict response to neoadjuvant docetaxel and capecitabine for breast cancer. *Breast Cancer Res. Treat.* 119, 685–699
- 5 van't Veer, L.J. *et al.* (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530–536
- 6 Ring, B.Z. *et al.* (2006) Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer. *J. Clin. Oncol.* 24, 3039–3047
- 7 Dowsett, M. *et al.* (2013) Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J. Clin. Oncol.* 31, 2783–2790
- 8 Ross, J.S. *et al.* (2008) Commercialized multigene predictors of clinical outcome for breast cancer. *Oncologist* 13, 477–493

Paul Ko Ferrigno

Avacta Group plc, Unit 651, Street 5, Thorp Arch Estate, Wetherby LS23 7GA, United Kingdom