Modelling how initiating and transforming oncogenes cooperate to produce a leukaemic cell state

Simon E. Richardson[1]

Summary of and comment on a recent *Cell Stem Cell* paper entitled 'Gene sets identified with oncogene cooperativity analysis regulate in vivo growth and survival of leukemia stem cells' (Ashton et al., 2012).

Progress in cancer research in the decade since the human genome was published has provided more questions than answers. A theme that has emerged from this enormous body of work is that the molecular and functional heterogeneity of cancer makes the study of established cancers complex (Stratton et al., 2009). Although the degree and nature of variegation differs between cancer types, heterogeneity seems to be a fundamental characteristic of neoplasia that is acquired from the earliest stages of tumour development. At the molecular level, first-hit genetic (or epigenetic) alterations are thought to initiate a pre-malignant process by interacting with the existing machinery of the physiological cell state (Speck and Gilliland, 2002). Additional second-hit mutations are necessary to drive transformation (Knudson, 1971) and, as the malignancy progresses, further mutations lead to the development of sub-clonal variegation. The functional consequences of this molecular heterogeneity are modulated by factors that are both intrinsic and extrinsic to the cell of origin. The cell state into which mutations are introduced is of crucial importance; for example, the tissue and developmental origins of the cell (Li et al., 2005) and its position in the cellular hierarchy (e.g. somatic stem cell, committed progenitor, mature cell) (Bonnet and Dick, 1997) are among the many factors that affect the outcome of mutations. Thus, molecular and functional heterogeneity forms substrates for Darwinian natural selection at the cellular level (Anderson et al., 2011). Sub-clones 'compete' in a complex multicellular organism that provides many microenvironmental niches, an immune system that is hard-wired to prevent malignancy and the selective pressure of cancer therapies.

Against this background of heterogeneity, it is perhaps remarkable that drugs targeting a single part of this complex mechanism exert any clinically relevant effect, let alone afford a cure. This has been the experience clinically: drugs have been developed that can induce responses and even complete remissions, but rarely cure.

Thus, more accurate models of the cancer cell state are needed to elaborate on how early initiating and transforming oncogenes cooperate to affect a functional cellular hierarchy. A recent publication by Ashton et al. (Ashton et al., 2012) capitalised on an established mouse model of chronic myeloid leukaemia blast crisis (CML-BC) that arises from the interaction between two different fusion genes (*BCR-ABL* and *NUP98-HOXA9*) and results in a functional hierarchy that has been previously characterised (Neering et al., 2007). Ashton et al. combine this defined model of disease progression with bioinformatic microarray analysis and in silico connectivity mapping to identify candidate drugs. Could this study inform a new approach to modelling the cancer cell state?

Summary of paper

The study by Ashton et al. used a previously published model of CML-BC involving retroviral overexpression of the *BCR-ABL* and *NUP98-HOXA9* oncogenes in mouse haematopoietic stem cell (HSC)-enriched bone marrow. Using a combination of HSC markers, these authors had previously characterised a functional hierarchy, which includes 'leukaemia stem cells' (LSCs; Lin– Sca+), progenitors (Lin– Sca–) and bulk populations (Lin+). A controversial interpretation of functional hierarchies in cancer is that 'cancer stem cells' have the greatest potential for self-renewal and are more resistant to therapy, making them the likely source of relapse. According to this theory, these cells are the key therapeutic target.

It is known that driver oncogenes can cooperate synergistically to change gene expression (McMurray et al., 2008), so the team used microarrays to identify essential cooperation response genes (CRGs) in primitive (Lin–) leukaemia cells. CRGs were defined as those genes that are activated more than additively by *BCR-ABL* and *NUP98-HOXA9*, compared with the effects of each translocation product alone. A total of 72 CRGs were identified and were validated by quantitative RT-PCR in more than 95% of cases. Of validated targets, 50 genes were upregulated. Interestingly, most CRGs that were identified had not previously been associated with leukaemia, and ranged from genes involved in signal transduction to metabolism and cell adhesion, potentially providing a broad range of previously unknown therapeutic targets for CML-BC.

The team also investigated whether the CRG expression pattern was maintained throughout the functional hierarchy of primitive LSCs, early progenitors and mature bulk cells. In total, 85% of the CRGs were common to all three fractions, suggesting that many of the core properties of oncogene cooperativity are conserved throughout the leukaemic hierarchy.

Subsequent experiments investigated whether knocking down the expression of upregulated CRGs affected leukaemia growth in vivo. The team used an in vivo short hairpin RNA (shRNA) screen to simultaneously knock down the 50 upregulated CRGs in isolated leukaemia cells in vitro, then transplanted the cells into mice. After 8 days, shRNAs against 35 of the 50 targeted CRGs were depleted in bone

[1]University College London Cancer Institute, Paul O'Gorman Building, 72 Huntley Street, London, WC1E 6DD, UK
e-mail: simon.richardson@ucl.ac.uk

marrow and spleen, indicating a functional role for the corresponding CRGs in maintaining the leukaemic population. Six CRGs with varying degrees of depletion were then investigated further. Individually knocking down each of the six candidates led to reduced leukaemia engraftment, but had little effect on long-term engraftment of normal HSCs, validating the roles of these CRGs in maintaining the abnormal self-renewal characteristics of CML-BC.

The greatest degree of depletion post-transplant was induced by shRNAs against *SerpinB2*. To further examine the role of the encoded protein in CML-BC, *SerpinB2*$^{-/-}$ LSCs expressing *BCR-ABL* and *NUP98-HOXA9* fusions were generated. Compared with SerpinB2-sufficient LSCs, *SerpinB2*$^{-/-}$ LSCs showed impaired engraftment. Interestingly, *SerpinB2*$^{-/-}$ LSCs expressing *BCR-ABL* alone had similar engraftment to SerpinB2-sufficient LSCs, supporting the idea that SerpinB2 can contribute specifically to blastic transformation of CML. Ex vivo, *SerpinB2*$^{-/-}$ LSCs showed defects in colony formation in methylcellulose, and increased apoptosis, but no change in cell cycle parameters, suggesting that SerpinB2 is necessary for the survival of LSCs prior to or during engraftment.

Previous work by Lamb et al. measured the effects of a large number of well-characterised drugs on the gene expression signatures of a range of commonly used cell lines to create a 'connectivity map' (Lamb et al., 2006). This group developed an Open Access online database to help researchers match disease-related gene expression profiles to candidate drugs. Using this database, Ashton et al. identified compounds that were predicted to reverse the CRG changes associated with the expression of *BCR-ABL* and *NUP98-HOXA9*. Two compounds exerted dose-dependent toxicity against leukaemia cells, particularly LSCs, with minimal toxicity to wild-type HSCs.

Finally, the team validated their findings using primary samples from patients with CML-BC (all of which carried *BCR-ABL* translocations, but a variety of other cooperating mutations). The authors carried out gene expression profiling for a subset of the CRGs identified in their mouse experiments, and showed that 13 of 33 evaluable genes were also dysregulated in the patient samples. Notably, in contrast to what was found in mice, *SerpinB2* was not identified as a CRG in the patient samples. When the human CML-BC CRG profile was analysed on the online connectivity map, small-molecule inhibitors were suggested that were similar to those identified based on the analysis of mouse CRGs; again, these compounds showed preferential toxicity to human leukaemia cells (particularly progenitors) compared with control cells.

In summary, the study by Ashton et al. combines previously published techniques to address three of the most vexing problems in modelling cancer: the unpredictable cooperativity of early driver mutations on cell state, functional heterogeneity and the identification of candidate drugs that can reverse this state.

Unresolved issues and future directions

Suitability of mouse models of malignancy

An obvious concern with this study is how closely the transcriptional profile identified in the mouse model mirrors that of human CML-BC. Retroviral overexpression of just two oncogenes in HSCs from inbred laboratory mice cannot model the physiological cell state of transformed human HSCs, nor the molecular and functional heterogeneity established during the human disease. Untreated CML in humans runs a chronic course, longer than the entire life expectancy of a mouse. This time course provides the opportunity to build up a range of different mutations and sub-clones, with blast crisis usually preceded by an accelerated phase of variable duration, suggesting a multi-step pathogenesis. Nevertheless, Ashton et al. used a well-characterised mouse model, and it should be pointed out that it is surprisingly difficult to make accurate mouse models of the various morphological (let alone myriad molecular) leukaemic sub-types (Tsuzuki et al., 2004). Despite the caveats of using mouse models, the authors identified significant overlap in both the gene expression profile and drug sensitivity between their mouse model and the molecularly heterogeneous human disease samples.

CRGs in molecular and functional heterogeneity

The identification of CRGs could be a strategy to sort through heterogeneity. Excitingly, this study seems to show that a small number of CRGs are preserved throughout the functional leukaemic hierarchy. This suggests that, at least in CML-BC, the CRGs drive a cancer cell state through which a physiological cellular hierarchy can be expressed. However, this study does not seek to validate a CRG approach to identifying drugs that circumvent molecular heterogeneity. First, it is believed that many leukaemias are preceded by the presence of an ancestrally initiated, but non-transformed, 'pre-leukaemic clone' harbouring the first-hit mutation (Wiemels et al., 1999; Hong et al., 2008). Work in the paradigmatic TEL-AML1 childhood acute lymphoblastic leukaemia (ALL) has clearly shown that such 'pre-leukaemic' clones commonly initiate relapse after chemotherapy (Mullighan et al., 2008). Because drugs targeting CRG profiles can only reverse the transformed state, they would not be expected to eradicate the pre-leukaemic compartment from which relapse can be initiated. Second, even in relatively tractable childhood leukaemias, such as TEL-AML1 ALL, there are several possible transforming second hits (Bateman et al., 2010). It is not known whether these second hits cooperate with the first hit in a convergent manner, to produce a common gene-expression profile, or whether they provide a diversity of potentially overlapping cell states. It is clear that transformation is more than a two-hit process in most cancers, and certainly several mutations are present in most clones in most patients at diagnosis.

So, does a core CRG profile persist through this molecular heterogeneity? The hope is that, despite the molecular complexity, oncogenes and transcription factors interact in a predictable way, permitting only a limited number of cell state 'solutions' (Enver et al., 2009). This moves us away from targeting individual molecular lesions to an approach in which drugs are used to make the cancer cell state solution thermodynamically 'unfavourable'.

Targeting CRGs for drug development

Using model systems with defined genetic events to identify synergistic changes to gene expression that occur during disease progression provides a manageable list of candidate genes to examine. The proportion of CRGs validated as having functional relevance during the shRNA screen is impressive, and provides an array of unexpected novel drug targets. However, beyond the identification of individual drug

targets, connectivity mapping can identify drugs that broadly antagonise the CRG profile. This not only provides a short-cut in drug screening, but can also identify compounds with efficacy in both a mouse model of CML-BC and on a heterogeneous selection of human CML-BC samples.

As Lamb et al. noted in their earlier paper (Lamb et al., 2006), the design of the reference connectivity map also used by Ashton et al. seeks to simplify biological complexity. Variables that are standardised in these maps include: cell types used; the dose, timing and combination of drugs assayed; and the measured output (mRNA, protein, epigenetic or functional changes). It remains difficult to predict whether a drug-induced change in the gene expression profile of a specific cell line will translate into functional changes to cancer cells in vivo, but this work shows that in silico connectivity mapping is a promising approach.

Conclusion

The past decade of cancer research has witnessed a revolution in our capacity to measure and analyse biological data, but, when deployed on human cancer samples, such analyses have revealed unexpected complexity. This study suggests that the earliest lesions cooperate synergistically to define the cancer cell state and that we can use these data to screen candidate drugs in silico. Perhaps the biggest challenge to overcome before this approach can be fully exploited is the lack of a cellular 'toolbox' that can accurately model the earliest cell states of multiple different cancers. An ideal system would compare wild-type human cells in their physiological cell state to those harbouring authentic initiating first-hit and transforming second-hit mutations. Directed differentiation of human pluripotent stem cells might provide the ability to model physiological cell states in vitro, whereas the ability to precisely genome engineer these cells with targeted nucleases could be used to model the earliest stages of cancer. An exciting possibility is that disease-specific model systems will identify disease-defining CRG profiles that persist through clonal diversification, allowing the identification of drugs that can more broadly antagonise the cancer cell state (Stratton, 2009).

REFERENCES

Anderson, K., Lutz, C., van Delft, F. W., Bateman, C. M., Guo, Y., Colman, S. M., Kempski, H., Moorman, A. V., Titley, I., Swansbury, J. et al. (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469**, 356-361.

Ashton, J. M., Balys, M., Neering, S. J., Hassane, D. C., Cowley, G., Root, D. E., Miller, P. G., Ebert, B. L., McMurray, H. R., Land, H. et al. (2012). Gene sets identified with oncogene cooperativity analysis regulate in vivo growth and survival of leukemia stem cells. *Cell Stem Cell* **11**, 359-372.

Bateman, C. M., Colman, S. M., Chaplin, T., Young, B. D., Eden, T. O., Bhakta, M., Gratias, E. J., van Wering, E. R., Cazzaniga, G., Harrison, C. J. et al. (2010). Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* **115**, 3553-3558.

Bonnet, D. and Dick, J. E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730-737.

Enver, T., Pera, M., Peterson, C. and Andrews, P. W. (2009). Stem cell states, fates, and the rules of attraction. *Cell Stem Cell* **4**, 387-397.

Hong, D., Gupta, R., Ancliff, P., Atzberger, A., Brown, J., Soneji, S., Green, J., Colman, S., Piacibello, W., Buckle, V. et al. (2008). Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* **319**, 336-339.

Knudson, A. G., Jr (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. USA* **68**, 820-823.

Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., Lerner, J., Brunet, J. P., Subramanian, A., Ross, K. N. et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929-1935.

Li, Z., Godinho, F. J., Klusmann, J. H., Garriga-Canut, M., Yu, C. and Orkin, S. H. (2005). Developmental stage-selective effect of somatically mutated leukemogenic transcription factor GATA1. *Nat. Genet.* **37**, 613-619.

McMurray, H. R., Sampson, E. R., Compitello, G., Kinsey, C., Newman, L., Smith, B., Chen, S. R., Klebanov, L., Salzman, P., Yakovlev, A. et al. (2008). Synergistic response to oncogenic mutations defines gene class critical to cancer phenotype. *Nature* **453**, 1112-1116.

Mullighan, C. G., Phillips, L. A., Su, X., Ma, J., Miller, C. B., Shurtleff, S. A. and Downing, J. R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377-1380.

Neering, S. J., Bushnell, T., Sozer, S., Ashton, J., Rossi, R. M., Wang, P. Y., Bell, D. R., Heinrich, D., Bottaro, A. and Jordan, C.T. (2007). Leukemia stem cells in a genetically defined murine model of blast-crisis CML. *Blood* **110**, 2578-2585.

Speck, N. A. and Gilliland, D. G. (2002). Core-binding factors in haematopoiesis and leukaemia. *Nat. Rev. Cancer* **2**, 502-513.

Stratton, M. R., Campbell, P. J. and Futreal, P. A. (2009). The cancer genome. *Nature* **458**, 719-724.

Tsuzuki, S., Seto, M., Greaves, M. and Enver, T. (2004). Modeling first-hit functions of the t(12;21) TEL-AML1 translocation in mice. *Proc. Natl. Acad. Sci. USA* **101**, 8443-8448.

Wiemels, J. L., Cazzaniga, G., Daniotti, M., Eden, O. B., Addison, G. M., Masera, G., Saha, V., Biondi, A. and Greaves, M. F. (1999). Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet* **354**, 1499-1503.