Exploring melanoma’s massively parallel universe

Keith S. Hoek

For molecular biologists the latest ‘big thing’ in high throughput technology has summarily dispensed with an old workhorse, the DNA microarray platform. With RNA-seq, researchers are nearer to the dream of sequencing and quantifying every molecule of a sample’s mRNA than ever before. Because the method acquires sequence data and quantifies transcripts without using probe hybridization, the probe affinity and cross-hybridization issues which plague microarrays are absent, and so the sensitivity and accuracy of detection is greatly enhanced. Furthermore, the technology enables researchers to characterize aspects of sequence structure beyond ordinary gene expression, including detection of alternative splicing events, allele-specific gene expression and sequence mutations. Key to all of this is massively parallel sequencing (MPS) which has accelerated the rate of large scale DNA sequencing efforts, er, massively (Shendure and Ji, 2008).

There was a time when it took years for whole institutions to obtain the full genomic sequence of a human by standard Sanger sequencing, but by using MPS one lab recently traced out James Watson’s complete sequence over a couple of months (Wheeler et al., 2008). Combined with the capacity to align sequences against comprehensive genomic databases, RNA-seq promises to be a powerful strategy for the detection of novel disease-specific events. Indeed, it seems set to revive hopes that were awakened, but never fully realised, with the rise of the microarray.

A team led by Levi Garraway of the Dana-Farber Cancer Institute has recently published a study which turned the focus of MPS to melanoma. By combining RNA-seq and high-resolution chromosomal copy number analysis, Berger and co-workers subjected the complete melanoma transcriptome to minute dissection. The principal finding is that of novel gene fusion events, the first described for melanoma. Their work represents an impressive technological feat, but has it contributed meaningfully to our understanding of melanoma biology?

After validating their strategy by positively identifying the BCR-ABL1 gene fusion in a chronic myelogenous leukaemia line (K-562), Berger and co-workers proceeded to analyze eight short term melanoma cultures and two cell lines (MeWo and 501 Mel). In six of the ten samples, the authors identified ten gene fusion events which they could confirm with Sanger sequencing. Gene fusions are exciting things to find in cancer because they are sometimes, like BCR-ABL1, found to be oncogenic. Gene fusions that are disease-specific are also useful for establishing precise diagnoses. However, each of the identified fusion events was pursued by RT-PCR in a further 90 melanoma samples and not found. With none of the events being identified more than once, the authors acknowledge that these may therefore be ‘passenger’ mutations with a diminished role in driving the disease. Perhaps more interesting was the identification of 12 novel ‘readthrough’ or chimeric transcripts in which neighbouring genes are concatenated during transcription. Seven of these were detected in at least two study samples each, meaning that they are probably feature events in melanoma. The most significant was the CDK2-RAB5B readthrough transcript detected in seven of the ten samples. Importantly, this readthrough was not detected in non-melanoma samples, indicating that it is specific to melanoma and may be useful in diagnosis. The biological significance of the CDK2-RAB5B readthrough is, however, unclear. Although CDK2 is a known MITF target important for cell growth and RAB5B is a RAS-related protein expressed in melanoma (Du et al., 2004; Meije et al., 2002), whether or not frequent readthrough of their sequences has a driving effect on the disease remains unknown. The authors also looked for novel somatic mutations, differentiating them from inherited single nucleotide polymorphisms by reference to patient-matched samples. Because they restricted themselves to considering mutations which were confirmed by at least ten separate reads, the search was limited to the top ~12% expressing transcripts (and so did not look at e.g. BRAF or NRAS). Again, while Berger et al. did find a range of novel somatic mutations (27 of which they validated by genotyping) none of these were found to occur more than once across the samples. Clearly the search should be expanded to cover as much of the remaining transcript pool as is possible, and the authors indicate that since their data was generated the ‘searchable fraction’ has grown to about 40% of expressed transcripts. Interestingly, Berger et al. report (as others have done before) mutation rates for their samples and confirm that melanoma’s mutation rate is significantly higher than it is in other cancers. The MeWo line had such a high mutation rate that the authors considered that it may harbour a ‘mutator phenotype’. The calculation of mutation rates in this paper needs some qualification, because the authors did not actually examine the rate of mutations within samples over time but rather compared the number of found mutations between samples. Therefore, one can’t help but think that because MeWo has been floating around various laboratories since the 1970s it may have had ample time to accumulate in vitro an excess of somatic changes.

Despite the likelihood that the gene fusion and readthrough events described are ‘passengers’ and not ‘drivers’ in melanoma, genomic rearrangements may still play a role in melanoma biology. For example, it has to be kept in mind that the gene fusions identified are only those which were transcribed into mRNA at a level high enough to be detected. This limitation restricted the current RNA-seq findings to a fraction of the transcriptome and even as the technology improves such that more and more transcripts become detectable, the transcriptome itself still represents a very small fraction of the full genomic sequence. It is highly probable that different regions of the genome are being rearranged without necessarily disrupting the structure or activity of genes located within them. Therefore, there may yet be specific DNA fusion events which will prove to be of general significance to the disease. But such events in non-coding regions will only be precisely


identified by full genome (i.e. DNA) sequencing efforts.

Berger et al. have successfully demonstrated just how powerful a technique RNA-seq really is for cancer research. They found, alas, that melanoma is an even more heterogeneous disease than was previously imagined. Large numbers of mutations spread across the samples speaks of a highly unstable genome subject to frequent alterations. The association of up-regulated CDK2 expression with its frequent read-through to the RAB5B gene indicates that at least some genomic changes affect genes that are actively being transcribed. One reflects on MITF amplification in melanoma, an observation also made by Garraway’s lab (Garraway et al., 2005), and wonders if that event too is somehow encouraged by the gene’s transcription. In any case, with each sample yielding its own galaxy of mutations, rearrangements, duplications and deletions, the described findings should give pause to any laboratory considering to fully sequence melanoma. The search for consensus alterations looks to be a long and expensive road to follow.

Technology advancement continues at a pace which is sometimes hard for ordinary laboratories to keep up with. Spiralling equipment and consumable expenses, particularly under the presently dire financial climate, remain major hurdles. Per sample it costs at least an order of magnitude more to analyse gene expression by MPS than with microarrays. The computing infrastructure required to assemble millions of sequence reads is similarly prohibitive. However, with time costs tend to come down and an increasing number of labs are bound to take it up. When they do, we should brace ourselves for a likely repeat of recent history. It was not so long ago that researchers were agog with excitement about the potential uses for DNA microarrays. As tickets to ride that band-wagon steadily cheapened, more and more laboratories jumped aboard, grabbed their instruments and feverishly played along. Subsequently, an astounding amount of data was generated and duly published. However, in retrospect it is difficult not to conclude that the main use of high throughput analysis appears to be for the production of papers thick with data but thin on content. Relative to output, few results have spawned new and scientifically profitable lines of study. This is especially true for melanoma research, which was present at the rise of the DNA microarray and has since been deluged by scores of (mostly poor) expression profiling studies. Many researchers, perhaps after being ‘burned’ at the bench while following up ultimately spurious conclusions, have subsequently learned to consider such works with a somewhat jaundiced eye. Indeed, anyone choosing to pursue the role of candidate genes raised by this paper should proceed carefully and pay close attention to the authors’ qualifying remarks. Nevertheless, because their RNA-seq study demonstrates in unprecedented detail the high degree of molecular heterogeneity with which melanoma biologists must contend, Berger and co-worker’s paper is a remarkable piece of work. What it is not is yet another high throughput exercise in ‘yellow press’ publication. Perish the thought.

References


