Research Highlights

Till death do us part

Malaria, the great killer, is carried by the mosquito, *Plasmodium falciparum*. It is found throughout the tropics but, while much is known of the biology of this insect, relatively little is known about its genetic diversity. Tanabe and colleagues examined 519 isolates from around the world and sequenced two housekeeping genes, identifying a series of single nucleotide polymorphisms (SNPs). Interestingly, the diversity of various geographical populations of *P. falciparum* almost exactly matches that of humans indigenous to that locale. In both cases, this corresponds to a decrease in diversity as the region analysed moves further from sub-Saharan Africa. This strongly supports the notion of a joint origin between *P. falciparum* and our own species and a co-evolutionary relationship with this parasite for most of our existence.

Tanabe, K. *et al.* (2010), '*Plasmodium falciparum* accompanied the human expansion out of Africa', *Curr. Biol.* Vol. 20, pp. 1283–1289.

A coagulating prediction

Anticoagulation therapies, such as coumarin, provide protection from stroke but getting the dose right is tricky. The new patient undergoes a time-consuming iterative process of dosage and clotting tests. Unfortunately, the variation in optimum dosage for different individuals is great, necessitating a great deal of testing. For this reason, the work of Verde and colleagues is of both genetic and clinical interest. They examined the two major genes associated with variation in the response to anticoagulation therapy, *CYP2C9* and *VKORC1*. They assessed the combined influence of the major variant alleles of these genes using a 'total genotype score' methodology and generated an algorithm. The results for this small test population were quite satisfying; genotype scores correlated well with the appropriate dosages. Now we must await a study in a larger population to see if the algorithm continues to predict optimal dosages.

Verde, Z. *et al.* (2010), 'A novel, single algorithm approach to predict acenocoumarol dose based on *CYP2C9* and *VKORC1* allele variants', *PLoS One* Vol. 5, p. e11210.

Beyond Mendel

How can we be as complex as we are with only 20,000 genes? Two partial answers to this question include alternative splicing and RNA editing. RNA editing involves the conversion of adenosine to inosine via the enzyme adenosine deaminase acting on RNA (ADAR). It occurs at thousands of sites and in all tissues examined. Most occur in non-coding Alu elements but there are still quite a few that can be found in coding regions. RNA editing is increased in brain tissues and ADAR appears to be essential for organismal viability, at least in mice. In this report, Osenberg and co-workers have extended these mouse studies to human embryonic stem cells (hESCs) and watched how RNA editing rates change as these cells differentiate. They found that as hESCs differentiate into neurones, the high rate of editing in Alu elements declines and that modulation of ADAR expression in these cells has marked effects on gene expression patterns. These data suggest that RNA editing has a role to play in the 'decisions' that hESCs make as they differentiate and contribute to our complexity. Unfortunately, for those who would use DNA sequence variations as predictors of function, this complexity now makes their job that much harder.

Osenberg, S. *et al.* (2010), 'Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing', *PLoS One* Vol. 5, p. e11173.

TERT extends its influence

Approximately 14,000 women in the USA will die from ovarian cancer this year. While mutations in *BRCA1/2* increase the risk of developing either breast or ovarian cancer (1.4 per cent versus 40 per cent), the number of ovarian cancer patients with a known mutation is small (3–5 per cent). With this in mind, Johnatty and colleagues screened thousands of patients and control subjects to search for new SNPs that might associate with ovarian cancer. Of 173 genes examined, only TERT (the active subunit of telomerase) showed a consistent association. TERT has been associated with several cancers and now we may add ovarian cancer to the list.

Johanatty, S.E. *et al.* (2010), 'Evaluation of candidate stromal epithelial cross-talk genes identifies association between risk of serous ovarian cancer and TERT, a cancer susceptibility "hot-spot"', *PLoS Genet.* Vol. 6, p. e1001016.

Living the high life

The people who ultimately became indigenous Tibetans are believed to have colonised their home between 25,000 and 11,000 years ago. They faced a formidable barrier as the elevation of their home often is of the order of 4,000 metres (approximately 13,000 feet), an altitude that normally causes chronic mountain sickness in long-term residents. Three groups, which originally worked independently but have now joined forces, now report on the first gene to be identified that begins to explain the ability of these people to live at high altitude. The gene, *EPAS1*, encodes a transcription factor named HIF2α. This transcription factor is known to play a role in stimulating red blood cell production. Alleles of *EPAS1* found in the Tibetan population are associated with much lower haemoglobin levels, thus protecting them from chronic mountain sickness. How these alleles actually work to protect high-altitude residents and whether or not protection from chronic mountain sickness was the driving selective pressure are questions that remain to be answered.

Beall, C.M. *et al.* (2010), 'Natural selection on *EPAS1* (HIF2α) associated with low hemoglobin concentration in Tibetan highlanders', *Proc. Natl. Acad. Sci. USA* Vol. 107, pp. 11459–11464.

The epigenetic hypothesis

The nature/nurture battle has left us with nature and nurture entwined. While there is beauty in this, it does little to satisfy the researcher who wishes to unravel the basis for complex phenotypes. Arturas Petronis has recently proposed a grand hypothesis, elevating epigenetic imprinting as a unifying principle. His reasoning is that DNA methylation patterns can be malleable enough to change in response to environmental stimuli and stable enough to be passed from mother to offspring. One can speak of a 'mutation rate' for methylation patterns as they are transmitted from mother cell to daughter cell. There is a certain degree of germline reprogramming and repair, which may serve as a source of both sporadic and familial diseases. In the words of the author: 'Taking an epigenetic perspective allows handling the same bundle of data as before, but placing them in a new system of relations with one another by giving them a different framework.'

Petronis, A. (2010), 'Epigenetics as a unifying principle in the aetiology of complex traits and diseases', *Nature* Vol. 465, pp. 721–727.

Deciphering the human foetal globin gene-silencing process

Hereditary persistence of foetal haemoglobin (HPFH) is characterised by persistently high levels of foetal haemoglobin (HbF) in adult life, and is mainly due to genetic factors. Genome-wide scans followed by linkage analysis in a large Maltese family identified a candidate region on chromosome 19p13.12–13. This harboured a nonsense mutation in the erythroid-specific factor, KLF1 (p.K288X), which ablated the DNA-binding domain of this key erythroid transcriptional regulator. Functional assays suggested that, in addition to its established role in regulating adult globin

expression, KLF1 also has a fundamental role in foetal globin gene silencing by activating *BCL11A*, a suppressor of HbF expression. The results of this multicentre study, coordinated by researchers from the University of Malta (Alex Felice's group), the University of Patras, Greece (George P. Patrinos' group) and the Erasmus University Medical Center, Rotterdam, the Netherlands (Sjaak Philipsen's group), provide a sound model for human globin gene switching and suggest that attenuation of KLF1 activity may be a fruitful approach for increasing HbF levels in individuals with β-type haemoglobinopathies, hence providing a novel approach to therapies for these conditions.

Borg, J. *et al.* (2010), 'Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin', *Nat. Genet.* Vol. 42, pp. 801–805.

Regeneration in mammals requires turning off ARF

It is widely accepted that re-entry of a post-mitotic cell to the cell cycle is a major prerequisite for cell repair and regeneration. In urodeles, which possess impressive regenerative capabilities, it has been shown that inactivation of retinoblastoma (Rb) allows cell cycle re-entry and DNA synthesis of muscle cells. Such an inactivation, however, is not enough in mammalian primary cells. The authors of this paper show that inactivation of Rb and also of the tumour suppressor alternative reading frame (ARF) allows cell cycle re-entry of C2C12 myotubes and consequently the production of cells that on transplantation can differentiate into myofibres. The authors state that because ARF is not present in urodeles, it was an invention of mammals to suppress regeneration. The absence of ARF from newts is inferred by its absence in lower organisms (the newt genome has not been sequenced). Thus, the authors speculate that ARF might be the answer to regeneration in newts. While this is highly interesting, such a claim can only be substantiated if removal of both Rb and ARF from, say, mouse limb would result in similar regeneration to the newts. Until then, the newts remain an indispensable animal model for vertebrate regeneration.

Pajcini, K.V. *et al.* (2010), 'Transient inactivation of Rb and ARF yields regenerative cells from post-mitotic mammalian muscle', *Cell Stem Cell* Vol. 7, pp. 198–213.

Comparing RNA sequencing methods

Strand-specific massively parallel cDNA sequencing (RNA-seq) is currently the best way to assess all transcribed sequences. By knowing which sequence corresponds to the original RNA, one can identify the presence of regulatory anti-sense RNA transcripts, establish the sequence of non-coding RNAs and demarcate the exact boundaries of adjacent genes which happen to be transcribed on opposite strands. There are several ways to do this, but little has been written in the way of a relative comparison of these methodologies. Levin and co-workers have rectified this situation in their recent report. Seven different strand-specific RNA-seq protocols were compared using *Saccharomyces cerevisiae* poly(A)+RNA libraries. The methods varied extensively in their ability to identify strands accurately, provide even coverage of the library and measure expression levels accurately. Anyone planning to perform RNA-seq should take a careful look at these results.

Levin, J.Z. *et al.* (2010), 'Comprehensive comparative analysis of strand-specific RNA sequencing methods', *Nat. Methods* Vol. 7, pp. 709–715.

Robert I. Scheinman
Associate Research Professor
University of Colorado Anschutz Medical Campus
Aurora, CO, USA
E-mail: robert.scheinman@ucdenver.edu