Research Highlights

Avoiding TEN

Carbamazepine has a number of uses. It blocks epileptic seizures. It stabilises mood swings caused by bipolar disorder and relieves some of the suffering associated with trigeminal neuralgia. Other off-label uses are proliferating. Unfortunately, a few individuals who take this drug experience Stevens–Johnson syndrome (SJS) or the related disease, toxic epidermal necrolysis (TEN). The likelihood of developing one of these potentially deadly adverse effects has been linked to the *HLA B*1502* allele. Recently, a group of clinicians tested the hypothesis that by screening for *HLA B*1502* they could profoundly decrease the prevalence of SJS–TEN in the carbamazepine treatment population. Almost 8 per cent of 5,000 patients who were candidates for carbamazepine therapy tested positive for *HLA B*1502* and were prescribed something else. The remainder received carbamazepine, and not one case of SJS–TEN was observed. Score another point for pharmacogenetic screening.

Chen P. *et al.* (2011), 'Carbamazepine-induced toxic effects and *HLA-B*1502* screening in Taiwan', *N. Engl. J. Med.* Vol. 364, pp. 1126–1133.

Too much of a good thing

The immune response, if not terminated, may move from eradication of pathogens to destruction of host tissues (as is seen in autoimmune disease). The FAS receptor mediates this by initiating lymphocyte apoptosis. In studies of sub-Saharan malaria, elevated levels of FAS and FAS ligand have been measured in multiple studies. Rather than simply considering this a mark of elevated immune function, Schuldt and co-workers used this as the basis for a focused examination of the *FAS* gene. Interestingly, they identified a promoter polymorphism (−436C > A) which conferred a 29 per cent reduced risk of children developing severe malaria. Individuals homozygous for this substitution showed a three-fold increase in FAS expression in blood-derived lymphocytes, as compared with individuals with the wild-type sequence. These data strongly suggest that the *removal* of activated lymphocytes is an important step in the response to this parasite.

Schuldt, K. *et al.* (2011), 'A −436C > A polymorphism in the human *FAS* gene promoter associated with severe childhood malaria', *PLoS Genet.* Vol. 7, pp. 1–10.

Cyclin D1 does more

Cyclin D1 is overexpressed in many cancers. To understand its function better, Jirawatnotai *et al.* laboriously identified cyclin D1-interacting proteins via immunoprecipitation followed by repeated rounds of liquid chromatography–mass spectrometry (LCMS). The result: cyclin D1 was found to interact not only with the expected cell cycle machinery, but also with components of the DNA repair complex. The authors went on to show that cyclin D1 interacts directly with RAD51 and that this interaction is completely independent of its cyclin-dependent kinase (CDK) activity. A reduction in cyclin D1 expression in cancer cells increased their sensitivity to radiation, suggesting that it may serve as a useful new target. We need all the targets we can get.

Jirawatnotai, S. *et al.* (2011), 'A function for cyclin D1 I DNA repair uncovered by protein interactome analyses in human cancers', *Nature* Vol. 474, pp. 230–234.

Unravelling the shape of the genome

DNA is organised in a three-dimensional structure comprising great open loops interspersed among

tightly packed areas of chromatin. Proteins called insulators help to define these architectural regions; however, how they do so is still being worked out. It is clear that this shape plays a critical role in defining when and where genes will be expressed. CCCTC-binding factor (CTCF) is, at present, the best understood of these insulators and now we understand it even better. Using a combination of chromatin immunoprecipitation–sequencing (ChIP-Seq), chromosome conformation capture and chromosome analysis by paired-end tag sequencing (ChIA-PET), Handoko and colleagues have performed an exhaustive analysis of mouse embryonic stem cells. They found that CTCF organises the genome into distinct architectural domains and subnuclear compartments that exhibit distinct epigenetic states and transcriptional activities. CTCF binding sites were correlated with p300 and nuclear lamina contact points. Together, the data suggest four modes of regulation for CTCF: creating distinct chromatin hubs, harnessing clusters of genes via coordinated expression, providing a link between regulatory elements and promoters, and creating barriers that serve to demarcate boundaries between subnuclear compartments.

Handoko, L. (2011), 'CTCF-mediated functional chromatin interactome in pluripotent cells', *Nat. Genet.* Vol. 43, pp. 630–638.

Just in case you felt the transcriptome was not complex enough . . .

RNA is not nearly the perfect copy of DNA that we had supposed. In a comparison of DNA with RNA sequences in 27 immortalised unrelated human B cell lines, over 28,000 events scattered over 10,000 exonic sites were observed. While a few were due to adenosine deamination, the majority were not. Using mass spectrometry, the authors demonstrated that these sequence changes were maintained through translation. All 12 possible alterations were observed; however, in any one location in any one individual, only one change occurred. The underlying reasons for this variation remain a mystery but it is clearly not 'noise' but rather a process with a physiological purpose.

Li, M. *et al.* (2011), 'Widespread RNA and DNA sequence differences in the human transcriptome', *Science* Vol. 333, pp. 53–58.

Autistic networks

Using a post-mortem tissue bank, Voineagu and colleagues sampled brain regions associated with autism and performed Illumina-based array analysis and RNA sequencing. Using a process called weighted-gene co-expression network analysis, they then integrated the differences into a set of gene networks. Interestingly, they found that significant differences which normally differentiate the frontal and temporal cortex were attenuated in tissue samples from autistic individuals. RNA sequencing identified splicing defects. Finally, using a published genome-wide association study (GWAS) dataset, they demonstrated that neuronal function genes within their identified network modules were enriched for variants associated with autism, while immune function genes were not. It will be interesting to see how this network model will be applied both to understand and treat autism, as well as other psychiatric disorders.

Voineagu, I. *et al.* (2011), 'Transcriptomic analysis of autistic brain reveals convergent molecular pathology', *Nature* Vol. 474, pp. 380–384.

No gene left undisrupted

Nothing beats a conditional knockout for the investigation of gene function within a tissue of interest. The past 20 years of effort has given us, perhaps, 5,000 knockouts, of which a minority have been conditionally targeted. The International Knockout Program, comprising organisations from both the USA and the European Union, aims to complete the list, making all genes within the mouse available as a conditional knockout. Recently, a high-throughput gene-targeting pipeline has been established which will form the basis

for this amazing resource. The pipeline consists of computational allele design, coupled with a high-efficiency recombineering gene-targeting strategy set upon an automated platform. Over 9,000 embryonic stem (ES) cell lines have been constructed and await you. Access detailed gene information at http://www.knockoutmouse.org. Targeting constructs and ES lines are available on request from either http://www.komp.org (US) or http://www.eummcr.org (EU).

Skarnes, W.C. *et al.* (2011), 'A conditional knockout resource for the genome-wide study of mouse gene function', *Nature* Vol. 474, pp. 337–342.

Mapping CLL mutations

Chronic lymphocytic leukaemia (CLL), a common blood cancer, has been divided into two subtypes characterised by the number of somatic hypermutations in the variable region of the immunoglobulin locus. Whole-genome sequencing was recently applied to this disease, with four patients volunteering their DNA for this purpose. The 46 somatic mutations found therein were then examined in a larger population of 363 patients. From this, four genes were found to be often mutated — the genes encoding notch 1 (*NOTCH1*), exportin 1 (*XPO1*), myeloid differentiation primary response gene 88 (*MYD88*) and kelch-like 6 (*KLHL6*). These were of particular interest to the authors, in that patients with unmutated hypervariable regions were more likely to have mutations in *NOTCH1* and *XPO1*, while those with hypervariable mutations were more likely to also have mutations in *MYD88* and *KLHL6*. Additionally, the authors claim that these mutations are likely to be cancer drivers, rather than simply markers. Indeed, they show data for the upregulation of *MYD88* and *NOTCH1* signalling which support this contention. The mapping of cancer's genetic landscape continues with exciting rapidity.

Puente, X.S. *et al.* (2011), 'Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia', *Nature* Vol. 475, pp. 101–105.

Modifier loci for cystic fibrosis

While cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, the clinical variation observed cannot be explained by this gene alone. Studies to identify modifier loci face technical challenges, including the need for a sufficiently large sample size and consistent diagnoses across different clinics. To this end, the authors created the North American Cystic Fibrosis Gene Modifier Consortium and agreed upon a means of quantifying disease. With these tools in hand, the authors embarked upon a full GWAS and linkage study. A number of suggestive loci were identified, of which two were deemed significant. One mapped to chromosome 11p13 in an intergenic region 3′ of the genes encoding apoptotic protease activating factor-1-interacting protein (*APIP*) and Ets homologous factor (*EHF*). APIP, an apoptosis regulator expressed both in lung tissues and monocytes, fits well with the emerging concept that cystic fibrosis may be worsened by decreased neutrophil clearance via decreased apoptosis. EHF, in turn, is a transcription factor expressed in bronchial epithelial cells. New pathways, when validated, may provide new targets against which sorely needed therapeutics may be developed.

Wright, F.A. *et al.* (2011), 'Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2', *Nat. Genet.* Vol. 43, pp. 539–548.

A catalogue of molecular aberrations

The application of genomic analysis to cancer marches onward with a large study cataloguing the changes associated with ovarian cancers. The authors examined mRNA and microRNA (miRNA) expression, promoter methylation and DNA copy number in almost 500 ovarian adenocarcinomas, along with exome sequencing in over 300 of these tumours. Tumour protein 53 (TP53) was mutated in 96 per cent of tumours. Other recurrent mutations were seen in genes such as that encoding breast cancer gene 1 (*BRCA1*), *BRCA2*,

the retinoblastoma gene (*RB*), and *CDK12*. Copy number mutations and promoter methylation events, as well as subtypes of miRNA, transcriptional patterns and promoter methylation patterns, were enumerated. Pathway analysis fingered NOTCH and forkhead box M1 (FOXM1) signalling as being implicated in disease progression. This catalogue should prove quite useful.

Cancer Genome Atlas Research Network. (2011), 'Integrated genomic analyses of ovarian carcinoma', *Nature* Vol. 474, pp. 609–615.

Optogenetic transcription

Using light to drive transgene expression is attractive for numerous reasons. One can implant cells expressing a therapeutic transgene and then drive expression of that gene when desired using light pulses delivered externally. This concept was put to the test recently. Melanopsin (sensitive to blue light) induces the release of intracellular calcium which is sufficient to drive the transcription factor, nuclear factor of activated T cell (*NFAT*). The authors created NFAT-expressing cell lines containing melanopsin along with an NFAT-driven transgene, and demonstrated in culture that expression was light sensitive. They then implanted a cell line that produced a glucagon-like peptide into a diabetic mouse and demonstrated that blood glucose levels could be regulated by pulses of light. This has applications to the eventual construction of the closed-loop glucose-sensing/glucose homeostasis system that would ease the lives of millions of individuals with diabetes, as well as having many other creative applications.

Ye, H. *et al.* (2011), 'A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice', *Science* Vol. 332, pp. 1565–1568.

Pure stemness

Purified haematopoietic stem cells (HSCs) have many uses. Bone marrow transplants would become mostly survivable, for example. The problem was always the purification ... until now. Notta and colleagues hit upon the bright idea that, since integrins often programme cells to remain in specific environments, it might be possible to identify an integrin which is specifically expressed in HSCs. To this end, they had previously created a robust HSC xenograft assay involving femoral injection of single cells using female NOD-*scid* $IL2Rgc^{-/-}$ mice. Using this assay, they indeed found that integrin α6 (CD49f) greatly improved the yield of $Thy1^{+}$ cells capable of HSC behaviour. To further purify cells capable of full repopulation, they sorted on the ability of these cells to efflux the mitochondrial dye, rhodamine-123. Single $thy1^{+}$ $Rho^{lo}CD49f^{+}$ cells were capable of complete self-renewal for 14–28 per cent of the time — a remarkable achievement.

Notta, F. *et al.* (2011), 'Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment', *Science* Vol. 333, pp. 218–221.

The watchmaker at work

While we strongly suspect that evolution makes its greatest leaps through mutations in developmental regulators, we have not yet had a good example of this. Frankel and colleagues have now provided us with an excellent dissection of the contributions of single nucleotide substitutions to the function of the *Drosophila* promoter for the transcription factor known as shavenbaby (svb), which occurred when strains diverged approximately 50,000 years ago. This factor was so named for its control of the formation of the larval cuticle. The authors examined the effect of each substitution on both the timing and level of *svb* expression using an assay developed especially for this purpose). The contribution of each change was subtle, although in aggregate, the effect was significant. Importantly, the effects were non-additive (ie most of the changes needed to be present before the effect was observed). This mode of selection was in contrast to other studies demonstrating large deletions and insertions within transcriptional enhancers. These sorts of changes generally cause a wholesale loss of tissue specificity and seldom create a viable modification. Thus, the authors conclude, single nucleotide substitutions are probably the favoured tool of evolutionary change.

Frankel, N. *et al.* (2011), 'Morphological evolution caused by many subtle-effect substitutions in regulatory DNA', *Nature* Vol. 474, pp. 598–603.

Your gut microbes are what you eat

The ecosystem of the gut must certainly be sensitive to diet. Recently, Faith and co-workers began the process of figuring out just how human gut bacteria react to a defined change in diet. They created gnotobiotic mice colonised with ten bacterial species found in humans and fed the mice a simple diet of protein (casein), fat (corn oil), polysaccharide (starch) and simple sugar (sucrose). By sequencing DNA found in faeces, they were able to determine the abundance of each species. This system allowed them to create a mathematical model capable of predicting how dietary changes would affect the gut ecosystem. The model allowed prediction of changes in bacterial growth in the presence of more complex human-like diets. This work will become especially important as links between gut flora and disease are generated.

Faith, J. *et al.* (2011), 'Predicting a human gut microbiota's response to diet in gnotobiotic mice', *Science* Vol. 333, pp. 101–104.

America's most wanted

Displaying an image of the brigand for all to see is sometimes all that is necessary. Kottke and co-workers have devised a molecular version of this strategy to allow the immune system to identify tumour cells. They packaged cDNA into the highly immunogenic vesicular stomatitis virus (VSV) to create a series of virally expressed epitope libraries. Intravenous injection of these libraries into mice with established tumours promoted anti-tumour activity without evidence of inflammation or autoimmunity. Tumours that escaped immune attack were treated effectively with repeat injections of the immunotherapy. What is remarkable about this process is that it does not require identification of specific tumour antigens or injection anywhere near the site of the tumour. Some years ago, yeast was used as the immunogenic stimulator (using single antigens): this technology is currently successfully moving forward in clinical trials. The future for cancer immunotherapy engenders both excitement and hope.

Kottke, T. *et al.* (2011), 'Broad antigenic coverage induced by vaccination with virus-based cDNA libraries cures established tumors', *Nat. Med.* Vol. 17, pp. 854–860.

RNA polymerase thinks small

The C terminal domain of RNA polymerase II (pol II) is well known for its extensive post-translational modifications. The Reinberg group has found yet another important node of regulation within this domain. It seems that the arginine at position 1810 is methylated by co-activator associated arginine methyltransferase (CARM1). When this amino acid is mutated, normal gene expression appears to be unaltered, except for the loss of a specific subset of small nuclear and nucleolar RNAs. Methylation at this position occurs in opposition to ser2 and ser5 phosphorylation events associated with classical gene expression, thus presenting us with an elegant switch. The ramifications of how this switch is regulated should prove to be anything but small.

Sims, R.J. *et al.* (2011), 'The C-terminal domain of RNA polymerase II is modified by site-specific methylation', *Science* Vol. 332, pp. 99–103.

Autism subtypes allow SNP identification

Despite a massive undertaking, few loci have been reproducibly associated with autism spectrum disorders (ASDs). This may be due to the heterogeneity of these diseases and the genetics that underlie them. The authors recently published evidence supporting the concept of dividing ASD into at least four separate groups. Here, using this differentiating technique, they were able to identify 18 novel single nucleotide polymorphisms (SNPs) upon reanalysis of published data. Interestingly, none of the SNPs were located in exonic regions, suggesting regulatory rather than structural defects.

Hu, V. *et al.* (2011), 'Novel autism subtype-dependent genetic variants are revealed by quantitative trait and subphenotype association analyses of published GWAS data', *PLoS One* Vol. 6, p. e19067.

Cis-eQTLs get perturbed

To what extent do environmental perturbations modulate gene expression differently due to genetic variation? Grundberg *et al.* addressed this question by first examining gene expression patterns upon stimulation by a panel of substances. A subset of the stimuli that elicited the most robust responses was then used for expression profiling and genome-wide SNP analysis. The result: the identification of two genes encoding myocin VI (*MYO6*) and tenascin C (*TNC*) within cis regulatory variants (cis expression quantitative trait loci [cis-eQTLs]) located within the promoter region. These rare variants, which respond differently to environmental stimuli, are likely to be clinically important biomarkers.

Grundberg, E. *et al.* (2011), 'Global analysis of the impact of environmental perturbation on cis-regulation of gene expression', *PLoS Genet.* Vol. 7, p. e1001279.

Help from the other side

The list of microRNAs (miRNAs) that are associated with cancer is ever growing. Thus, it is heartening to hear of one that protects. The miR-17-92 cluster is overexpressed in a number of cancers, and numerous members of this cluster have oncogenic properties — particularly those derived from the 5′ arm. Interestingly, one miRNA derived from the 3′ arm, miR-17*, does something completely different. It acts to suppress mitochondrial antioxidant enzymes, making prostate cancer cells more susceptible to oxidative stress.

Xu, Y. *et al.* (2010), 'miR-17* suppresses tumorigenicity of prostate cancer by inhibiting mitochondrial antioxidant enzymes', *PLoS One* Vol. 5, p. e14356.

Impulsive founders

If one examines a genetically homogeneous founder population (the theory goes), one has a better chance of detecting the effects of rare disease alleles. This idea was recently tested by comparing Finns with extraordinarily poor impulse control (so poor that they were incarcerated) with matched controls. Exon-centric sequencing of serotonin and dopamine genes identified a stop codon in serotonin receptor 2B (*HTR2B*). Knockout of *Htr2b* in mice did the same thing. At the time of writing, the mice remain incarcerated!

Bevilacqua, L. *et al.* (2010), 'A population-specific *HTR2B* stop codon predisposes to severe impulsivity', *Nature* Vol. 468, pp. 23–30.

New loci for the heart

Coronary artery disease (CAD) represents one of the great killers within the developed nations. It has been estimated that as much as 30–60 per cent of the risk is inherited. Yet, as with many other heritable risks, only a tiny fraction of this risk has been mapped to specific loci. The CARDIoGRAM consortium pooled data from 14 GWASs, representing 22,233 patients and 64,762 controls, and performed a meta-analysis. The most promising SNPs were then genotyped in 56,682 new subjects. The results: 13 new susceptibility loci, with a plethora of new genes not previously associated with CAD, as well as a collection of genes that appear to contribute to multiple diseases. A rich harvest, which, interestingly, only accounts for 10 per cent of the risk. Thus, there remains much to do in this area.

Schunkert, H. *et al.* (2011), 'Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease', *Nat. Genet.* Vol. 43, pp. 333–338.

PhD basket weaving

Our advances in understanding are, in part, due to advances in the manipulation of materials and to the construction of clever devices that can do things we could not easily do before. In this regard,

Dongran Han and colleagues have worked out how to generate complex synthetic three-dimensional DNA structures using a technique developed by others, termed 'DNA origami'. The authors have advanced the field by developing design principles that allow a greater freedom of curvature, using strand crossover events to stabilise complex shapes. They demonstrate this with the construction of a series of DNA objects, including baskets (hemisphere) and other pleasing nanostructures. Something useful should come of this eventually. One might imagine a DNA 'cage' that delivers a therapeutic molecule to a specific region of the genome, for example.

Han, D. *et al.* (2011), 'DNA origami with complex curvatures in three-dimensional space', *Science* Vol. 332, pp. 342–435.

Double trouble

Ankylosing spondylitis (AS) is a somewhat common form of inflammatory arthritis, with significant heritability. About 10 per cent of these patients also have inflammatory bowel disease (IBD). Similarly, the incidence of AS in the population of IBD patients is also high. A GWAS screen was recently performed using susceptibility loci from a form of IBD known as Crohn's disease, applied to AS patients and controls. What popped out were loci near signal transducer and activator of transcription 3 (STAT3) — a transcription factor which plays a role in Th17 activity — and KIF21B — a member of the kinesin motor family. Th17 T cells have long been branded the perpetrators in several autoimmune conditions. The role of a kinesin motor is less clear.

Danoy, P. *et al.* (2010), 'Association of variants at 1q32 and STAT3 with ankylosing spondylitis suggests genetic overlap with Crohn's disease', *PLoS Genet.* Vol. 6, p. e1001195.

Targeting the seed

Over 50 per cent of human genes are regulated to some extent by miRNAs. Their dysregulation underlies numerous disease states, so inhibiting them is a current goal for numerous disease intervention strategies. Obad and colleagues have made an important advance in this regard. They make use of a recent advance in backbone modification, known as locked nucleic acids (LNAs). By employing an oxymethylene bridge between carbon atoms in ribose, the LNA is conformationally constrained and has much higher duplex stability than normal antisense sequences. This increase in duplex stability has allowed these antagomirs to target only the seed sequence of entire miRNA families, to great effect.

Obad, S. *et al.* (2011), 'Silencing of microRNA families by seed-targeting tiny LNAs', *Nat. Genet.* Vol. 43, pp. 371–378.

A new role for DICER

DICER is the ribonuclease that processes small interfering RNAs (siRNAs) and miRNAs such that they can enter into the RNA-induced silencing complex (RISC), where they can act as guides to target and destroy mRNAs. In this study, a loss of DICER1 function was associated with geographic atrophy, a form of age-related macular degeneration (AMD) and one of the major causes of blindness in the Western world. *DICER1* knockdown in both mice and in human retinal pigment epithelial cells (major players in AMD) resulted in an increase in Alu RNA. This retrotransposon-derived RNA appears, in this context, to play a pathogenic role. Thus, we have two surprises: DICER1 appears to regulate Alu RNA levels and Alu RNA can play a pathogenic role in AMD.

Kaneko, H. *et al.* (2011), 'DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration', *Nature* Vol. 471, pp. 325–332.

The genes of multiple myeloma

Multiple myeloma is a B cell cancer with a pathogenesis that remains unclear. Using whole-genome or whole-exome sequencing, the authors of this recent report have examined 38 patients comparing tumour tissue with normal tissue. While some of the usual suspects were rounded up, a number of new genes were identified. Nearly half the patients

had mutations in genes involving RNA processing, protein translation and the unfolded protein response. Others had mutations in genes involving histone methylation and blood coagulation. The nuclear factor κB pathway was broadly represented, with 11 activating mutations identified. Mutations in the oncogenic kinase, BRAF, were also identified in a small number of patients. This is of great interest, given the presence of BRAF inhibitors in the clinic. Once again, in the hunt for cancer genes, large-scale sequencing brings home the catch.

Chapman, M.A. *et al.* (2011), 'Initial genome sequencing and analysis of multiple myeloma', *Nature* Vol. 471, pp. 467–472.

Them's the breaks

Genomic rearrangements play a major role in establishing the course of development for many cancers. Indeed, different cancers have characteristically different sets of breakpoints. Visual assessment methodologies are giving way to powerful genomic sequencing approaches. DNA-paired-end-tag (DNA-PET) sequencing was used recently to generate maps of structural variations in epithelial cancers — and, in a companion paper, in breast cancers. Large duplications were found to be among the initial rearrangements that trigger genome instability in epithelial cancers. A novel segmental tandem duplication was also identified in many breast cancer genomes.

Hillmer, A. *et al.* (2011), 'Comprehensive long-span paired-end-tag mapping reveals characteristic patterns of structural variations in epithelial cancer genomes', *Genome Res.* Vol. 21, pp. 665–675.

Inaki, K. *et al.* (2011), 'Transcription consequence of genomic structural aberrations in breast cancer', *Genome Res.* Vol. 21, pp. 676–687.

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