

## Clinical genome interpretation

Isaac Kohane, David Margulies and colleagues report on the CLARITY Challenge, which sought to develop standards for the analysis, interpretation and reporting of clinical sequencing for the diagnosis of genetic disorders (*Genome Biol.* 15, R53, 2014). The study challenge included information on 12 individuals from 3 families with different heritable disorders, with clinical information provided from medical records, exome sequencing using the SOLiD platform and whole-genome sequencing by Complete Genomics. Thirty international groups participated in the challenge and were assessed on the basis of the methods they used for analysis and interpretation, whether these methods were efficient, scalable and replicable, and the clinical usefulness of their case reports. Only two groups identified the consensus candidate variants in all of the cases, although there was significant overlap in the candidates reported across the teams and some consensus emerging on the methods used for alignment, variant calling and pathogenicity prediction. For example, most of the groups used GATK and SAMtools, either alone or in combination, for variant calling and used both SIFT and PolyPhen for pathogenicity prediction. The authors highlight the need for broader adoption of standard data formats, consideration of coverage along with an estimation of false negative rates for candidate genes and further development of publicly available genomic databases. **OB**

## Oncogenic chimeric transcript

Hepatitis B virus (HBV) is thought to drive hepatocellular carcinoma (HCC) development through viral insertion into the host genome, but the mechanism is not fully understood. Now, Nathalie Wong, Ting-Fung Chan and colleagues report the identification of viral-human gene fusions and a recurrent chimeric transcript with oncogenic functions (*Cancer Cell* 25, 335–349, 2014). The authors performed transcriptome sequencing of six HBV-positive HCC cell lines and identified chimeric viral-human transcripts in four of the lines. They found virus-driven expression of a long interspersed nuclear element (LINE1) repetitive sequence generating a chimeric HBx-LINE1 transcript, which was expressed in about 23% of HBV-associated HCC tumors. Knockdown of HBx-LINE1b reduced the migratory and invasive properties of HCC cells, decreased expression of markers of the epithelial-to-mesenchymal transition and reduced the transactivation activity of  $\beta$ -catenin, whereas transgene-driven expression of HBx-LINE1b promoted cell migration and invasion. Only the full-length HBx-LINE1 transcript had functional effects, suggesting that the transcript functions as a long noncoding RNA. Finally, the authors generated HBx-LINE1 transgenic mice, which were found to be more susceptible to diethylnitrosamine-induced HCC formation. **EN**

## Epigenetic reprogramming in treatment-resistant breast cancer

Steffi Oesterreich and colleagues report that the homeobox gene *HOXC10* is hypermethylated at a CpG island shore in aromatase inhibitor (AI)-resistant breast tumors (*Sci. Transl. Med.* 6, 229ra41, 2014). The authors profiled the methylomes of two cell line models of AI-resistant breast

cancer and identified 72 hypermethylated regions enriched for developmental genes, including *HOXC10*. DNA and histone methylation at an estrogen-responsive element in the distal promoter region of *HOXC10*, near an extensive CpG island, was associated with transcriptional repression. In cell lines, *HOXC10* silencing led to an increase in cell growth, proliferation and motility and to a decrease in apoptosis. In mouse xenograft models, cell line-derived tumors that developed resistance to estrogen deprivation showed hypermethylation at *HOXC10* compared to estrogen-stimulated controls. Finally, the authors showed that *HOXC10* expression levels were lower in a small set of patient-derived tumors that recurred after AI therapy than in matched primary tumors. **BL**

## MRSA toxicity and virulence

Ruth Massey and colleagues report an analysis used to predict the toxicity of methicillin-resistant *Staphylococcus aureus* (MRSA) isolates based on a genetic signature (*Genome Res.* doi:10.1101/gr.165415.113, 9 April 2014). They characterized 90 independent MRSA ST239 isolates for adhesiveness and toxicity, which are critical virulence factors. They found limited interstrain variation in binding to fibrinogen and fibronectin, whereas there was an 18-fold difference between the most and least toxic strains. To examine the relationship between toxicity and disease severity *in vivo*, they infected mice with the isolates with the highest and lowest toxicity, demonstrating that the more toxic isolates caused the most severe disease. They conducted a genome-wide association analysis for toxicity, identifying 100 SNPs and 22 indels associated at  $P < 0.05$ . Next, the authors screened for evidence of repeated independent evolution and segregation in clusters of highly toxic isolates, narrowing the candidate loci to four. They also tested for epistatic interactions between toxicity-associated loci. Finally, they built a predictive model for toxicity classification that uses a random forest machine learning algorithm and is based on data for a set of 31 SNPs and 21 indels. Using this model, they were able to correctly predict all 27 strains classified as having either high or low toxicity in this study, with lower accuracy shown in predicting strains of medium toxicity. **OB**

## Organizing the body axis

In many developmental systems, the patterning of an undifferentiated pool of cells into an organized structure is achieved through the production of secreted morphogens that establish gradients of signaling activity that coordinate cell movements and specify distinct cell fates. Christine Thisse and colleagues now show that opposing gradients of two such morphogens, Nodal and bone morphogenetic protein (BMP), are sufficient to organize uncommitted zebrafish blastula cells into a well-patterned embryo (*Science* 344, 87–89, 2014). To establish the role of these morphogens, the authors generated opposing gradients of Nodal and BMP by injecting their respective mRNAs into different blastomeres at the undifferentiated animal pole of the zebrafish blastula. Whereas Nodal alone induced a radially symmetric structure with dorsal identity, the introduction of a gradient of BMP opposite to the Nodal gradient induced a bilaterally symmetric secondary body axis that was independent of the orientation of the primary axis. The authors repeated these experiments using explants of animal pole cells cultured *in vitro* and again found that opposing gradients of Nodal and BMP could produce a complete embryonic axis, showing that the activity of these two morphogens is sufficient to organize the basic vertebrate body plan. **KV**

Written by Orli Bahcall, Brooke LaFlamme, Emily Niemitz & Kyle Vogan