IN BRIEF

GENE REGULATION

A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements

Mittler, G. *et al. Genome Res.* 17 Nov 2008 (doi:10.1101/gr.081711.108)

A new study describes the use of SILAC (stable isotope labelling by amino acids in cell culture) for high-throughput screening of sequence-specific DNA–protein interactions. Cellular proteins are labelled *in vivo* using amino acids substituted with non-radioactive isotopes. Labelled proteins that bind to specific DNA elements can then be affinity purified and analysed by mass spectrometry. The authors identified proteins that preferentially bind to methylated CpG sites, showing that the approach was also applicable to DNA modification-dependent interactions.

GENOME VARIATION

Extensive genomic copy number variation in embryonic stem cells

Liang, Q. *et al. Proc. Natl Acad. Sci. USA* **105**, 17453–17456 (2008)

Extensive copy number variation (CNV) in mouse and human genomes arises during meiosis and contributes to disease traits. CNV is highly variable between inbred mouse strains, but is fixed within strains. However, by analyzing the genomes of 50 embryonic stem-cell clones derived from three parental lines the authors reveal extensive mitotically generated CNV in these cells (affecting over 1,000 genes) and show that variation is transmitted to the germ line. It is therefore plausible that our somatic tissues are CNV mosaics of the zygotic genome.

TRANSCRIPTOMICS

Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome

Cheung, V. *et al. PLoS Biol.* **6**, e277 (2008)

This study suggests that cryptic transcripts, which are expressed from within coding regions in certain mutant backgrounds, provide a source of alternative genetic information. Cheung and colleagues show that many cryptic transcripts that are expressed in *Saccharomyces cerevisiae* mutants have ORFs of at least 100 codons, and that several cryptic transcripts are translated into protein. Furthermore, when wild-type cells were subjected to a nutritional shift, several cryptic transcripts were expressed, suggesting that the transcripts have functions that are induced under altered conditions.

HUMAN DISEASE

Genome-wide analysis of human disease alleles reveals that their locations are correlated in paralogous proteins

Yandell, M. *et al. PLoS Comp. Biol.* **4**, e1000218 (2008)

The authors develop a computational method to identify sequence variants that might have phenotypic consequences. Rather than look for sequence conservation across homologous proteins in different species, the method is applied to paralogous sequences in the same genome. Variants from the SNP database dbSNP mapped to homologous positions on the 17,000 paralogous human proteins studied, and disease-causing mutations in paralogues tended to align with one another. Alignment of paralogous proteins might therefore be a resource for identifying disease-associated variants.