Structural variation meets 454

Efforts to characterize human genomic structural variation have used a variety of microarray- and sequencing-based strategies, but these approaches have generally been limited in their ability to map breakpoints precisely or too laborious to apply on a genome-wide scale. Now, Michael Snyder and colleagues (*Science* published online 13 September 2007; doi:10.1126/science.1149504) describe a new method, based on pyrosequencing, that allows fine-scale mapping of genomic structural variation on an unprecedented scale. The approach, termed paired-end mapping, involves high-throughput analysis of end sequences from 3-kb genomic fragments followed by mapping of paired-end reads onto the reference genome. The method allows detection of most structural variants greater than 3 kb in size and achieves an average breakpoint resolution of less than 1 kb. The authors applied the method to analyze the genomes of two individuals, achieving up to fourfold genome coverage. At this depth of sequencing, the authors detected 1,297 structural variants, including 122 inversions. Breakpoint analyses revealed that most deletion events arose by nonhomologous end joining, whereas nonallelic homologous recombination accounted for a smaller proportion of the variation, providing insights into the mechanisms responsible for widespread structural variation. ***KV***

Variants of unknown significance

The sequences of the *BRCA1* and *BRCA2* genes have probably been determined in more individuals than those of any other genes in the genome. Numerous variants of unknown significance have been identified, for which clinical counseling is problematic. Now Doug Easton and colleagues present an analysis of disease causality for 1,433 sequence variants in *BRCA1* and *BRCA2* in over 70,000 individuals using data from the Myriad Genetic Laboratories database (*Am. J. Hum. Genet.* **81**, 873–883; 2007). The authors used a logistic regression model to combine likelihood ratios derived from three independent tests to assess each variant: co-occurrence in *trans* with a known deleterious mutation, assessment of personal and family history of cancer, and formal segregation analysis in pedigrees for variants in which pedigree data was available. The authors were able to classify 133 variants as neutral with odds greater than 100 to 1 and 43 variants as deleterious with odds of at least 20 to 1. The authors note that even though this method was not able to classify most variants, this analysis has the potential to provide useful information to 1,569 probands and their families, because the variants most likely to be classified were the ones with many independent observations. *EN*

JmjC-family proteins demethylate histone H3 lysine 27

Di- and trimethylation of lysine 27 (K27) on histone H3 is catalyzed by the polycomb repressive complex 2; these repressive chromatin modifications are thought to constitute a unique stem cell epigenetic state, and there is a rapid decrease of H3 di- and trimethylation during differentiation. Now, three groups have shown that members of the JmjC family of proteins are H3K27 demethylases. Ramin Shiekhattar and colleagues and Kristian Helin and colleagues reported that UTX demethylates di- and trimethyated H3K27 (*Science* published online 30 August 2007; doi:10.1126/science.1149042; and *Nature* advanced online publication 22 August 2007; doi:10.1038/nature06145), and Helin and colleagues and Charlie Degui Chen and colleagues reported that JMJD3 demethylates H3K27 trimethylation (same *Nature* paper, and *Cell Research* published online 9 October 2007; doi: 10.1038/cr.2007.83). In addition, Shiekhattar and colleagues and Helin and colleagues both showed that UTX associates with *HOX* gene promoters during retinoic acid–induced differentiation of embryonic carcinoma cells. UTX is required for loss of H3K27 methylation and transcriptional activation at these loci. These are the first reports of H3K27 demethylases. *EN*

Subcellular mRNA localization patterns

The subcellular localization of specific mRNAs during *Drosophila melanogaster* embryogenesis is a well-established mechanism for generating graded protein distributions, but it is unclear to what extent this mode of regulation occurs across diverse classes of cellular transcripts. To address this question, Henry Krause and colleagues (*Cell* **131**, 174–187; 2007) applied a sensitive *in situ* hybridization strategy to assess the subcellular distribution patterns of 2,314 transcripts across early stages of *Drosophila* embryogenesis. Unexpectedly, they found that 71% of the transcripts they analyzed were spatially distributed in the embryo. In addition to previously well-described patterns such as anterior or posterior localization, the authors observed other striking distribution patterns, including apical or basal localization, membrane association or localization with cytoskeletal networks. In many cases, the observed transcript distribution patterns are consistent with known functional roles. In other cases, where the functions of the encoded proteins are unknown, the patterns of distribution and co-localization can be used to infer possible functional relationships. Although it remains uncertain how broadly this mode of regulation occurs in other systems, these observations suggest that transcript localization may be more prominent in development than has generally been appreciated. *KV*

CNC evolution

Several recent studies have examined evidence for selection on a genome-wide scale, but it remains of interest to characterize evolution in noncoding regions, and in particular of conserved noncoding elements (CNC). A new study by Su Yeon Kim and Jonathan Pritchard now examines the adaptive evolution of CNCs (*PloS Genet.* 3, e147; 2007) and offers a new method to detect evidence of selection. The authors began by scanning for regions conserved in up to eight vertebrates, with a dataset of 98,910 CNCs. They then used a statistical method called a 'shared rates test' to identify CNCs that showed significant variation in substitution rates across branches of a phylogenetic tree, and they applied this method to alignments from human, chimpanzee, dog, mouse and rat genomes. They found that about 68% of CNCs evolve according to a model with a single constant evolution rate among phylogenies, whereas the rest show departures from this basic model. From these, they found that about 75% showed a rate change on a single lineage (in a two-parameter model), whereas the remainder showed two or more (up to a seven-parameter model) rate changes. They further found a subset of CNCs that showed evolution rates faster than the neutral model on a particular branch, suggesting adaptive evolution. ***OB***

Written by Orli Bahcall, Emily Niemitz and Kyle Vogan