IN BRIEF

CANCER GENOMICS

Cancer classification within tissues and beyond

In a combined analysis of multiple omics platforms (5 genomic and 1 proteomic) on 3,527 specimens from 12 cancer types, researchers have determined shared genomic signatures to yield a refined classification into 11 major cancer subtypes that goes beyond tissue of origin. For example, one subtype characterized by alterations in *TP53* and *TP63*, as well as by increased expression of immune and proliferation pathway genes, comprised lung squamous, head and neck, and a subset of bladder cancers. As this study shows, next-generation sequencing technologies yield omics data sets that can be successfully integrated to facilitate large-scale studies of complex systems, thus providing important, multilayered insights into biology and disease.

ORIGINAL RESEARCH PAPER Hoadley, K. A. *et al.* Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **158**, 929–944 (2014)

HUMAN EVOLUTION

Selection — a balancing act

In humans, the pervasiveness of hard sweeps (that is, for strong selection to drive the rapid fixation of new mutations in particular populations) has been strongly argued against. As an alternative, various selective processes categorized as balancing selection could maintain genetic variation in populations. However, the actual impact of balancing selection on the human genome was unknown. Now, DeGiorgio *et al.* have developed two composite likelihood ratio tests that model the spatial distribution of polymorphism expected near a site under long-term balancing selection. Applied to whole-genome data from unrelated Africans and Europeans, these methods detected several genes at the *HLA* region, which is a locus well-known to be under balancing selection in humans. In addition to many previously identified loci, the tests detected new sites displaying signatures of balancing selection.

ORIGINAL RESEARCH PAPER DeGiorgio, M., Lohmueller, K. E. & Nielsen, R. A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genet.* **10**, e1004561 (2014)

GENE EXPRESSION

A global assessment of RNA-seq performance

To characterize how technical variables affect the reliability of high-throughput RNA sequencing (RNA-seq), the SEQC/MAQC-III Consortium carried out RNA-seq using different sequencing platforms and protocols across different geographical laboratory sites. From the >100 billion sequence reads generated, they compared the ability of RNA-seq procedures to identify known features of the RNA samples relative to each other and relative to gene expression microarrays and quantitative PCR. Overall, good accuracy and reproducibility were seen across approaches, platforms and locations when identifying differentially expressed genes (if appropriate bioinformatic filters were used). However, RNA-seq and microarrays showed inaccuracies for the quantification of absolute RNA levels, and all approaches had gene-specific biases. In parallel, Li *et al.* show that RNA-seq platforms differ in their ability to detect RNA variants, such as splice site isoforms.

ORIGINAL RESEARCH PAPERS SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nature Biotech.* http://dx.doi.org/10.1038/nbt.2957 (2014) | Li, S. *et al.* Multi-platform assessment of transcriptome profiling using RNA-seq in the ABRF next-generation sequencing study. *Nature Biotech.* http://dx.doi.org/10.1038/nbt.2972 (2014)