Human microbiome genomes

The initial phase of the US National Institutes of Health Human Microbiome Project (HMP), which seeks to characterize the human microbiome, includes plans to produce reference genome sequences for 900 bacteria from the human microbiome. The Human Microbiome Jumpstart Reference Strains Consortium, which includes partners from four genome centers, now report their progress in sequencing 178 of these microbial reference genomes (*Science* 328, 994, 2010). They highlight the importance of strain selection, setting standards for sequencing and annotations, and data release strategies as guidelines for future studies. These microbial genomes were distributed amongst the five major body sites surveyed by HMP, including the gastrointestinal tract, oral cavity, urogenital tract, skin and respiratory tract. Pangenome analysis of four species, for which there were five or more annotated genomes each, provided initial estimates of their core genome sizes and suggested additional sequencing is needed to reach saturation. Overall, 547,968 predicted polypeptides were identified (94% of them unique), of which ~5% were defined as candidate novel peptides. Analysis of metagenomic shotgun datasets showed that ~33% of the human metagenome remains not well covered by any of the current reference genomes. ***OB***

Fusion gene function

Chromosomal rearrangements creating fusions of *TMPRSS2* with *ERG* occur in half of all prostate cancers, but the function of the TMPRSS2-ERG fusion protein is unclear. Now, Arul Chinnaiyan, Jindan Yu and colleagues report an investigation of the function of this fusion protein (*Cancer Cell* **17**, 443–454, 2010). The authors performed chromatin immunoprecipitation coupled with massively parallel sequencing (ChIP-seq) to map the genome-wide binding patterns of the androgen receptor (AR) and ERG and found that ERG and AR co-occupy target loci in prostate cancer cell lines and prostate tumor tissue. They further found by coimmunoprecipitation that ERG and AR proteins interact in prostate cancer cell lines and prostate tumor tissue. ERG overexpression had an inhibitory effect on expression of *AR*, and knockdown of ERG expression in a cell line with the *TMPRSS2-ERG* fusion resulted in derepression of AR target genes. Notably, the authors also found that ERG overexpression positively induced expression of the *EZH2* polycomb gene, and expression profiling in prostate tumors revealed a correlation between a polycomb expression signature and *TMPRSS2-ERG* fusion status. This work reveals interesting potential functions for the TMPRSS2-ERG fusion in disruption of prostate differentiation programs. *EN*

miRNA-33 and cholesterol

miR-33a and miR-33b are found within introns of *SREBP-1* and *SREBP-2*, respectively, which encode the sterol regulatory element–binding protein (SREBP) transcription factors. Now, Anders Naar and colleagues test whether miR-33a and miR-33b are functionally associated with their *SREBP* host genes (*Science* published online, doi:10.1126/science.1189123, 13 May 2010). Bioinformatic analysis predicted that a target of miR-33a and miR-33b is the ATP-binding cassette A1 (ABCA1) cholesterol transporter, a major regulator of high-density lipoprotein (HDL) homeostasis. Transfection of cell lines with synthetic miR-33 precursor oligonucleotides boosted miR-33a and miR-33b levels and led to decreased ABCA1 expression. In contrast, transfection with antisense oligonucleotides targeting miR-33a and miR-33b led to increased levels of ABCA1, showing that ABCA1 is regulated by miR-33a and miR-33b. The authors found that depleting cholesterol leads to increased expression of miR-33a and its *SREBP-2* host gene accompanied by a concomitant decrease in ABCA1 protein levels. The decrease in ABCA1 was partially reversed by antisense oligonucleotides that decrease miR-33a levels. Finally, the authors tested whether locked nucleic acid (LNA) antisense oligonucleotides that target miR-33a could modulate cholesterol levels in mice. The authors observed that plasma HDL cholesterol levels were substantially increased in LNA–miR-33a antisense-treated animals, raising the possibility of antisense therapeutic targeting of miR-33a and miR-33b as a treatment for cardiometabolic diseases. *PC*

Written by Orli Bahcall, Pamela Colosimo, Emily Niemitz & Kyle Vogan

CNVs and autism

Autism-spectrum disorder (ASD) is a heterogeneous group of conditions characterized by deficits in social interaction, impaired ability to communicate and stereotypic behavior patterns. Stephen Scherer and colleagues report a genome-wide study of 1,000 individuals with ASD (*Nature* published online, doi:10.1038/nature09146, 9 June 2010). The individuals with ASD (cases), their parents and 1,287 controls were genotyped using the Illumina 1M SNP microarray. Comparing ASD cases to controls, cases carry a higher burden of rare, genic copy number variants (CNVs). Among cases, 5.7% had one or more *de novo* CNV and >0.6% had two or more *de novo* CNVs. Although previous work has suggested a higher rate of *de novo* CNVs in families with one individual affected with ASD, this analysis showed a nearly equal rate of *de novo* CNVs in families with one affected individual and in those with more than one affected individual. Two hundred and twenty-six *de novo* and 219 inherited CNVs affecting single genes were observed in cases and not found in controls, implicating the new candidate ASD genes *SYNGAP1*, *DLGAP2* and *DDX53-PTCHD1*. To identify biological processes involved in ASD, the authors analyzed the function of genes affected by CNVs. The authors found that gene sets involved in GTPase-Ras signaling and cellular proliferation were enriched in genes disrupted by CNVs. *PC*

Turning T cells into natural killers

Natural killer (NK) cells are components of the innate immune system with important roles in tumor surveillance and antiviral defense. Pentao Liu and colleagues (*Science*, published online, doi:10.1126/science.1188063, 10 June 2010) now report that genetic ablation of the transcription factor gene *Bcl11b* in mice is sufficient to reprogram T cells into cells with NK-like properties. The authors used a tamoxifen-regulated Cre allele to conditionally delete *Bcl11b* in purified lymphocyte populations under various culture conditions. They found that cell populations from all developmental stages within the T-cell lineage could be converted to NK-like cells by eliminating *Bcl11b*. These induced T-to-natural-killer cells, or ITNKs, expressed markers characteristic of NK cells and had the ability to kill stromal cells in culture. ITNKs were also produced efficiently when *Bcl11b* was deleted *in vivo*. Notably, ITNKs showed the ability to kill tumor cells *in vitro* and to restrict tumor progression and metastasis *in vivo*. The potential for ITNKs to selectively eliminate malignant or virally infected cells without attacking normal cells could lead to the development of new cell-based approaches for treating cancer and viral infections. ***KV***