GENOMICS

Asian genomes

GenomeAsia100K Consortium. *Nature* https://doi.org/10.1038/s41586-019-1793-z (2019).

In spite of its large number of populations, genomic data from Asian people are underrepresented in reference genome datasets. This poses challenges in medical genetic studies of Asian people, as well as hinders a detailed understanding of their demographic history. The GenomeAsia100K Consortium aims to "facilitate and accelerate genetic studies in Asian populations by coordinating sequencing efforts among its members." In their GenomeAsia Pilot (GAsP) project, 1,267 whole-genome sequences were generated with high coverage. By including additional human genome sequences from previous sequencing studies, they built a reference dataset with 1,739 individuals from 219 population groups and 64 countries across Asia. Analyzing this GAsP dataset revealed ancestries and founder effects for different Asian populations. The rich information about genetic variation in Asian people also helps improve accuracy in pinpointing disease-relevant variants in Asian populations. *LT**

https://doi.org/10.1038/s41592-020-0753-8

STRUCTURAL BIOLOGY

Reprogramming allosteric control in proteins

Chen et al. *Nat. Chem. Biol.* **16**, 77–86 (2020).

Allostery is mediated by a network of coupled residues that propagate small local structural changes into long-distance structural modulation. Although coupled allosteric residues can be identified from concerted motion in molecular dynamics simulations, designing allosteric control is not trivial. Chen et al. develop a framework for computational protein design that takes into account the protein conformational stability, motion, binding properties and known allosteric regulation to reprogram signaling activity. First molecular dynamics simulations are used to predict allosteric sites, followed by the use of normal-mode calculations and multi-state design for de novo design for reprogramming the long-range coupling. The researchers demonstrate on G-protein-coupled receptors (GPCRs), which involve receptor activation through highly conserved sites called microswitches. They use the method to engineer constitutive and ligand-induced G protein activation in a GPCR, the dopamine D2 receptor, for which the active-state structure is not known. *AS*

https://doi.org/10.1038/s41592-020-0755-6

MICROBIOLOGY

Oceans of genomes

Pachiadaki, M. G. et al. *Cell* **179**, 1623–1635 (2019).

Sequencing-based approaches have empowered us to decipher the largely unknown set of marine microorganisms and their roles in geochemical cycling and global ecosystems. However, the scarcity of reference genomes remains an obstacle for taxonomic and functional annotations. Pachiadaki et al. collected 28 water samples from the tropical and subtropical Atlantic and Pacific Oceans and used a single-cell genomics approach for surveying the complex microbiome in seawater samples. Unlike metagenomics analysis, single-cell genomics generates sequences of individual cells that do not make assumptions based on microbial abundances. The researchers generated untargeted libraries of single amplified genomes of prokaryoplankton, which recovered partial genomes of 12,715 sequenced cells. Surprisingly, in a single 0.4 mL seawater sample, they recovered a subpopulation of 6,236 genomes. They also offer a reference database, GORG-Tropics, for further prokaryoplankton taxonomy and function assignments. *LT*

https://doi.org/10.1038/s41592-020-0754-7

SENSORS AND PROBES

Encoding the glow

Gregor, C. et al. *Proc. Natl Acad. Sci. USA* **116**, 26491–26496 (2019).

Bioluminescence from luciferases and their substrate luciferins have found extensive use in biomedical research. In comparison to fluorescent proteins, they are beneficial because they do not need light to deliver signal. However, unlike fluorescent proteins, the most widely used systems are not fully genetically encoded. Recent work from Stefan Hell's lab revealed the full pathway for luciferin formation in a bacterial system of bioluminescence. Gregor et al. now show that this entire pathway can be expressed in mammalian cells to create cell lines that luminesce without the need for exogenous luciferin. For the best signal, the researchers carried out codon optimization of the relevant genes and tested multiple plasmids in mammalian cells. Their work creates expression systems with signals on par with that of firefly luciferase and has important implications for the future of bioluminescence imaging. *RS*

https://doi.org/10.1038/s41592-020-0756-5

Arunima Singh, Rita Strack, Lei Tang, Lin Tang* and Nina Vogt