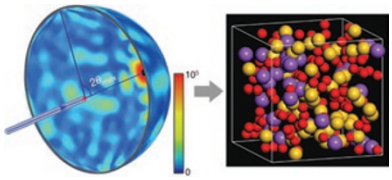
IMAGING AND VISUALIZATION

One-shot structure determination

By sampling a two-dimensional diffraction pattern on a spherical detector, three-dimensional structure determination of single molecules should be possible from a single measurement.

The three-dimensional (3D) structure of an object can reveal a lot about its functional purpose. This is especially true at the microscopic scale, which is why many different methods have been developed to image biological particles, all the way down to the atomic level. To obtain 3D structures, however, one must reconstruct them from 2D snapshots. For a technique such as confocal microscopy, this involves sectioning biological specimens into planes and then assembling them back together to reconstruct the 3D image. For the technique of X-ray crystallography, this requires taking shots of multiple, identical copies of a biological molecule at different orientations, and reconstructing the 3D image from the diffraction patterns.

Although X-ray crystallography can yield atomic-resolution 3D structures of proteins and other biological molecules, it has two limitations. "The problem is that it requires crystals and also it [yields] average structures," explains Jianwei Miao of the University of California, Los Angeles. Besides the challenge and time expense of crystallization, because X-ray crystallography is an 'averaging' technique, it cannot be used to look at single molecules. Thus the small structural differences between copies of a biomolecule, which are often crucial for understanding biological function, cannot be observed.



The concept of ankylography. Left, a coherent beam illuminates a particle and the diffraction pattern is collected on a spherical surface ($2\theta_{max}$ is the diffraction angle). Right, the 3D structure of a sodium silicate glass particle is encoded in the 2D spherical diffraction pattern. Reprinted from *Nature*.

But now, Miao and his colleagues show that obtaining enough information to construct a 3D structure from a single diffraction measurement of a single particle should in principle be possible, using a new concept called ankylography.

The key idea of ankylography, which comes from the Greek word *ankylos*, meaning 'curved', is that a spherical (rather than planar) detector is used to record the diffraction pattern. A coherent X-ray beam is fired at the particle, and the diffraction pattern of the resulting scattered waves is recorded on a curved surface by, for example, a charge-coupled device (CCD) camera. If the diffraction pattern is sampled at a very fine scale, the 3D structure of the object is encoded by the 2D spherical pattern. Thus ankylography should allow the determination of 3D structures of single particles in one shot.

Using simulated diffraction data for a sodium silicate glass particle and for poliovirus,

GENOMICS

TO EACH HIS OWN

***De novo* assembly of human genome sequences that are not currently included in the reference genome opens the possibility of a human pan-genome.**

For Jun Wang at the Beijing Genome Institute-Shenzhen and his international collaborators, the idea of a human pan-genome was triggered by the development of a new short sequence read assembler. Wang recalls that during their testing of SOAPdenovo—part of the short oligonucleotide analysis package (SOAP)—they assembled the genomes of an African and an Asian individual but discovered that 5 megabases (Mb) of the approximately 3 gigabases of total assembled sequence for each genome did not align to the current human reference.

The current human reference genome GRCh37, which stands for 'genome reference consortium human build 37', is the representation of a haploid human genome, derived from sequences of multiple individuals. The assembly is divided into a linear primary assembly and a series of alternate loci in regions with great diversity.

Instead of just trying to fit the 5 Mb of diverse sequence into alternate loci to a single linear reference assembly, Wang and colleagues borrowed a concept from microbiologists who are used to dealing with multiple different genomes in bacterial communities. In the microbial world, a pan-genome describes the diversity of genomes in a population, and accordingly, Wang and his collaborators defined a human pan-genome as a nonredundant set of human DNA sequences that includes all genetic information of human populations.

To ascertain that the diverse sequences were indeed authentic, the researchers confirmed that a substantial fraction of the 5 Mb aligned to other mammalian sequences stored at the National Center for Biotechnology information (NCBI). When they compared these sequences between the African and Asian individuals' genomes they saw that these sequences were polymorphic and not simply a rearrangement of repeat regions. "We started to guess," says Wang, "that they could be individual-specific or even population-specific sequences, and it would be interesting to see whether there are any potential functional elements in them."

Wang and his colleagues estimated that these individual-specific regions will encompass a total of up to 40 Mb, around 1.3% of the entire human genome. As these sequences vary greatly between individuals, Wang speculates that they can only be identified by *de novo* assembly of human genomes. To

Miao and his colleagues demonstrated that this is possible, at least theoretically. They also imaged an etched silicon nitride substrate using experimental soft X-ray diffraction data, but because their CCD was equipped with a planar detector, they needed to mathematically interpolate what the pattern on a spherical detector would be.

To use ankylography in real-life experiments, additional technological developments are needed. For one, spherical detectors are needed. Also, robust computational algorithms to reconstruct 3D images from 2D spherical diffraction patterns will be necessary. Additionally, a new generation of ultrapowerful X-ray free electron lasers (XFELs), which are expected to transform the field of X-ray structure determination, are only just beginning to become available. These XFELs would be ideal to couple with ankylography; Miao predicts that 2-nanometer resolution from a single shot using an XFEL beamline would be possible, enabling 3D imaging of a broad range of biological specimens.

Technology development could potentially proceed rapidly. Hundreds of millions of dollars have already been invested for the development of XFELs and improved detectors in the US, Europe and Japan, says Miao, and many groups worldwide are working on the coherent imaging field using these advanced X-ray sources. "I think in a few years from now it should be much clearer what can be done with this technique," he says. "This is a completely new idea so the community is still debating it; we've created a lot of discussion."

Allison Doerr

RESEARCH PAPERS
Raines, K.S. *et al.* Three-dimensional structure determination from a single view. *Nature* advance online publication (16 December 2009).

get a firm handle on the function of these regions, genomes of many more individuals will need to be sequenced.

To complete the entire human pan-genome will most likely never be feasible because everyone would need to have their genome sequenced and assembled. Currently whole-genome sequencing is extremely expensive, if done by a commercial provider and, in addition, there are ethical and legal issues that need to be worked out, not to speak of a person's right not to know.

Realizing this, Wang presents an alternative solution, "We could start with over a hundred individuals to get the common alleles with a frequency higher than 1% in a population." Population-based pan-genomes would already be a great help in finding genomic locations associated with phenotypic traits in a given population.

But ultimately Wang thinks that only the individual-specific, rather than population-specific sequences, provide insight into biological functions and disease mechanisms, and, despite all challenges, he thus summarizes their work in a simple message: "Everyone should have their genome assembled to get complete information."

Nicole Rusk

RESEARCH PAPERS
Li, R. *et al.* Building the sequence map of the human pan-genome. *Nat. Biotechnol.* advance online publication (7 December 2009).

NEWS IN BRIEF

GENE REGULATION

Mutagenesis screens in human cells

The classical mutagenesis approach is a powerful method to determine the genes induced in various biological processes. However, the diploid nature of the human and other mammalian genomes has limited large-scale mutagenesis studies. Now, Carette *et al.* describe an insertional mutagenesis screening approach for human cells that uses a chronic myeloid leukemia cell line with a haploid karyotype, for all but chromosome 8. They used this approach to identify genes required for the lethal effects of bacterial toxins and for infection with influenza virus.

Carette, J.E. *et al. Science* **326**, 1231–1235 (2009).

PROTEIN BIOCHEMISTRY

Structures of transient complexes

Proteins carry out most of their functions by interacting with other proteins. Often these interactions are short-lived and are therefore difficult to isolate and study structurally. Schug *et al.* describe an approach to obtain structural models for these transient protein complexes. They combine existing structures for the individual proteins with an analysis of mutational patterns to identify interacting surface amino acids and then use molecular dynamics simulations to resolve structural models of the transient complexes.

Schug, A. *et al. Proc. Natl. Acad. Sci. USA* **106**, 22142–22129 (2009).

MICROARRAYS

DNA stiffness as a microarray readout

Gene expression analysis via microarrays usually involves fluorescence labeling. Husale *et al.* describe an alternative, label-free readout, based on the unique nanomechanical response of hybridized versus single-stranded DNA. They use an atomic force microscope to scan the surface of the array and create a stiffness map; hybridized DNA molecules are less stiff than flat-lying single-stranded DNA probes. This nanomechanical readout offers attomolar sensitivity and a large dynamic range.

Husale, S. *et al. Nature* **462**, 1075–1078 (2009).

IMAGING AND VISUALIZATION

Cell cycle imaging in fish

Cell cycle–dependent protein ubiquitination has been previously used by Miyawaki *et al.* as the basis for genetically encoded fluorescent reporters for the cell cycle in mammalian systems. Using the fish orthologs of these proteins, the researchers now extend this 'Fucci' approach to zebrafish. The transparency of the fish embryo permits the imaging of cell-cycle transitions during development.

Sugiyama, M. *et al. Proc. Natl. Acad. Sci. USA* **106**, 20812–20817 (2009).

BIOSENSORS

Multiplexed biomarker detection

Making the detection of clinical biomarkers routine requires inexpensive, highly sensitive and reliable methods that can be multiplexed. Washburn *et al.* demonstrate that the relatively new technology of silicon photonic microring resonators can fulfill these requirements. These optical sensors are very sensitive to changes in the refractive index upon binding of an antigen to an antibody, for example. Washburn *et al.* used this technology for the multiplexed detection of five protein biomarkers in complex solutions.

Washburn, A.L. *et al. Anal. Chem.* **82**, 69–72 (2010).