

3D reconstruction of a nucleus. Each colour represents a different chromosome, and each dot is a gene location.

REVEALING CHROMOSOME SHAPE, ONE DOT AT A TIME

Methods that meld imaging with sequencing are revealing the organizational principles of the genome – and how it influences cell fates. **By Amber Dance**

Someone has drawn a mascot on the whiteboard in Mitchell Guttman's molecular-biology laboratory at the California Institute of Technology (Caltech) in Pasadena. It looks like a tangled ball of blue yarn a cat would chase, complete with eyes, jaunty grin, arms and legs.

Named SHARP-Y, after the gene-silencing protein SHARP the Guttman group studies, it could be the mascot for any of a handful of labs that are analysing similar tangled features – not balls of yarn, but the web of DNA in the nucleus. As these researchers are discovering, those tangles are anything but random. Chromosomes are precisely organized, as are the RNAs they make and the proteins that interact with them, and this organization seems to be crucial for gene expression to work as it should.

Efforts to trace chromatin – the complex of

DNA and protein that makes up a chromosome – drive a small but growing field that is concerned with the 3D spatial positioning and dynamics of the molecular components that comprise the 'nucleome'.

These researchers are tackling a seemingly straightforward question: how does the genetic material arrange itself, physically, inside the nucleus? Biologists typically think of DNA as a string, a linear sequence of the nucleotide letters A, T, G and C that make up the DNA double helix. But cells can't treat their genetic material in that way, says Guttman. For example, when a cell has to adjust to an environmental change, a protein called a transcription factor enters the nucleus, seeking specific genes to activate for the appropriate response. But a linear search would take hours, too long for a timely response. Organization solves the problem: each chromosome has its

own 'territory', where it is further subdivided into sections that are open for transcription or closed off. Those are then split into smaller domains, which unite sequences that tend to interact with each other. That way, genes and proteins can find their partners efficiently.

Thinking of DNA in 3D also solves a problem that genome sequencing has not, says Ana Pombo, a genome biologist at the Max Delbrück Center for Molecular Medicine in Berlin. Only 1–2% of the human genome encodes proteins directly. Much of the rest – where many disease-linked mutations can reside – performs regulatory roles, often influencing the expression of far-flung genes. But it isn't always easy to link these regulatory sequences to the genes they control. Chromosome structure can help to resolve those connections.

Disease links are already apparent. The

gene-imprinting conditions Prader–Willi syndrome and Angelman syndrome, which cause developmental delays and intellectual disabilities, have been associated with structural differences between sister chromosomes in a person's cells, says Guttman. And scientists reported in 2016 that a genetic mutation involved in brain cancer produces an abnormal metabolite that interferes with the normal boundaries between DNA domains in chromatin¹. Last year, in work that has not been peer reviewed, a team led by researchers at Columbia University in New York City suggested² that the coronavirus SARS-CoV-2 alters the architecture of chromosomes in olfactory cells, causing some people to lose their sense of smell.

Scientists have long had a well-stocked toolkit for studying these associations biochemically, for instance using the technique Hi-C to crosslink DNA regions that are found in close proximity to each other. But those tools offer only an average view of chromosome arrangement; things can look different at the single-cell level. Imaging offers a richer picture. Some approaches build on fluorescence *in situ* hybridization (FISH), a long-standing method used to 'paint' chromosomes or identify individual genes using fluorescent tags. Others use *in situ* sequencing to find the location of specific genetic targets or a random subset of the genome in chemically fixed cells or tissues. Researchers are also combining methods to gain a holistic view of the nucleus, creating 'multi-omic' data sets.

"You don't have to choose between imaging and sequencing," says Xiao Wang, a genomics researcher at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts. "You can do both in the same sample."

FISHing for loci

Caltech bioengineer Long Cai's approach to spatial genomics stemmed from a simple realization: "Fundamentally, a DNA sequencer is a microscope." Many modern sequencing machines decode DNA by incorporating fluorescently tagged nucleotide bases into the DNA as it is copied, reading those additions letter by letter. Cai figured: "Why take everything out of the cell, prepare it, and put it in the sequencer?" He wondered whether he could instead analyse nucleic acids right where they lie.

FISH provided the starting point. With this method, scientists design fluorescent nucleic-acid probes that are specific to the sequences they want to light up, and use microscopy to pinpoint the probes' location in the cell. However, the method can look at only a handful of sites in the same sample, because microscopes can distinguish between only a few colours.

The Cai group's innovation was to label a single sample repeatedly with

different-coloured probes for several genetic loci, then decode the images later. They call the technique seqFISH, or sequential fluorescence *in situ* hybridization (see 'Mapping a chromosome'). In their first demonstration, the researchers assigned each of 12 RNAs a unique, sequential barcode such as blue–yellow, green–purple, yellow–blue or purple–green, using four colours in total. Then they designed FISH probes using those colours for each RNA, and performed two rounds of labelling and imaging of yeast cells. Each spot on the image indicated an RNA, and the colours it flashed in the two rounds indicated its identity³.

The maximum number of targets this approach can label is 16 (or 4²: 4 dyes and 2 rounds of labelling). But when graduate student Yodai Takei joined the Cai lab in 2015, he wanted to see thousands of target sequences – and not just RNA, but nuclear DNA as well. Last year, he and his colleagues reported doing just that⁴.

Takei labelled 3,660 DNA loci in slices of mouse cerebral cortex, imaging them over

"You're really looking at spatial organization of the nucleus. This is really the future of genomics."

125 rounds of data collection. By spacing those sites one million bases apart, Takei obtained a pattern of dots that, when joined up as in a connect-the-dots puzzle, provided a low-resolution approximation of the chromosome's conformation. The data revealed that chromosomes in the same types of cell were arranged and interacted in similar patterns. The approach could be used to explore how the nucleus is organized in many other cell types.

But 125 rounds of imaging? Working manually, each round of probe binding, imaging and stripping takes at least 50 minutes, Takei says; 125 rounds would have required, at a minimum, 7 consecutive 15-hour days. Fortunately for Takei, an automated microscope did the work for him. A typical experiment still takes about a week, but Takei – now a postdoc at Caltech – can do other things while it runs.

Cai employs two mechanical engineers to build automated microscopes such as these. In the lab's microscopy room sits a handful of machines, each occupying its own small space shrouded in black curtains to block out ambient light. Takei's set-up is built on a Leica microscope, but decking it out with an automated sampler, custom fluidics and a computer script to control it took two years. But the finishing touch is decidedly low-tech: the sample is protected from light by an upside-down cardboard box.

That's not the kind of microscope you

can buy off the shelf – at least, not yet. Cai co-founded the California-based firm Spatial Genomics to commercialize seqFISH technology, and a product is expected later this year, according to Brian Fritz, vice-president of marketing for the company.

Another firm, Acuity Spatial Genomics, which has offices in Newton, Massachusetts, and San Jose, California, is commercializing a different spatial-imaging technology. Called OligoFISSEQ, it was developed in the laboratory of Ting Wu, a chromosome biologist at Harvard Medical School in Boston, Massachusetts.

OligoFISSEQ combines fluorescence *in situ* sequencing (FISSEQ) – a technique that sequences nucleic acids in their tissue context – with barcoded versions of Oligopaints, which are FISH probes invented by the Wu group. The team engineered the probes so they can reveal chromosome topology in three ways: sequencing by hybridization (as for FISH); sequencing by synthesis; and sequencing by ligation. Sequencing by synthesis is the technology that many next-generation sequencers use, except in this case, the sequences are read in the tissue rather than being extracted first. Sequencing by ligation uses short, fluorescently labelled strands of DNA called oligonucleotides that are repeatedly attached to the Oligopaints barcode, imaged and then removed⁵.

Wu's team used that technology to trace the shape of the X chromosome through 46 loci spaced about 3 million bases apart. Using the specific barcodes and four rounds of imaging in the study⁵, the hybridization approach could, in theory, detect up to 1,296 loci. The other two sequencing strategies could yield as many as 65,536 loci after 8 rounds of sequencing. Wu co-founded Acuity to commercialize the approach, and the company is currently working on a product.

Scattered sequencing

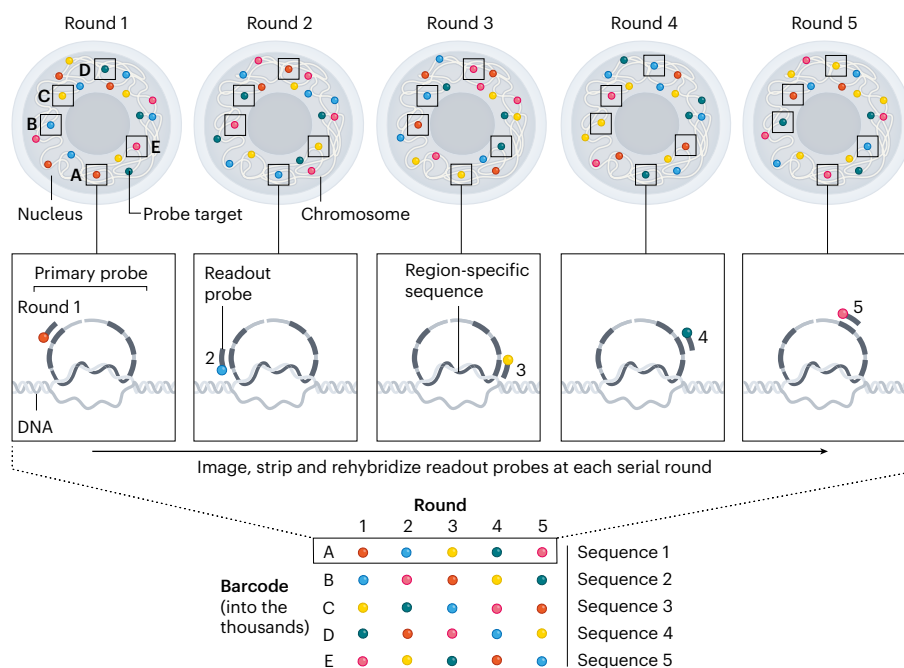
FISH's strength is its signal: researchers can tile multiple probes next to one another at each genomic locus, creating a strong, bright, fluorescent output. But researchers usually design probes only for the genes they care about. "It's not a very good discovery tool," says Guttman.

His team uses a biochemical technique called SPRITE to crosslink sequences in chromosomes, then barcode them at random to label any loci, without bias, that tend to be found near each other⁶. Sequencing of the barcodes and what they're attached to reveals the physical associations. With collaborators, Guttman's team has applied SPRITE in tissues from mouse brains and beetles to the plant *Arabidopsis*.

Image-based techniques also support untargeted searches through *in situ* sequencing of genomic DNA on a microscope slide. But because a single sequence wouldn't be very bright, researchers first amplify the signal by

MAPPING A CHROMOSOME

Researchers can trace chromosome shape using a technique called seqFISH (sequential fluorescence *in situ* hybridization). Repeated rounds of probe binding, imaging and removal creates a 'sequential barcode' that maps a sequence of colours onto a physical position in the genome. Thousands of gene locations can be barcoded.



repeatedly copying the sequences.

If that sounds simple, trust genomic scientist Fei Chen when he says it wasn't. His team at the Broad Institute spent several years developing *in situ* genome sequencing⁷, which they reported in 2020.

The process unfolds in three steps. First, the scientists take fixed cells or embryos and sprinkle sequencing adapters into the genome at random, creating an unbiased sample that preserves the fragments' spatial positions. Each adapter contains a unique, 20-base barcode to help the scientists read out the sequence later. Then they use a technique called rolling circle amplification to produce a 'DNA nanoball', measuring 400–500 nanometres across, which contains multiple copies of the barcoded DNA.

Next, the researchers decode those nanoballs using sequencing by ligation. But that method can read only about 20 bases: too few to conclusively identify a genetic region. This is where the barcodes come in. On the slide, the researchers sequence only the barcodes. Then they break up the cells and extract their DNA to sequence them again using standard sequencing by synthesis. Most next-generation sequencers can easily read the unique barcode together with 100 or more bases from the genomic locus where that barcode landed, allowing the scientists to match barcodes to loci on the linear sequence.

Finally, researchers use the barcodes to match up the thousands of dots seen in the microscope image, like nuclear confetti, with the linear sequence. Doing so allowed Chen and his colleagues to observe how cells with

shared lineages have more similar chromosome architecture than do cells without common ancestry.

Multi-omics

Chromosome models in papers look like highly articulated puzzles, with coloured balls and rods approximating the shape of a chromosome in the cell. But DNA by itself provides an incomplete picture of genetic activity, Guttman says. RNAs present near a DNA locus indicate that transcription is under way. And DNA can interact with or be anchored by nuclear structures, such as the nucleolus that generates ribosome components and the nuclear speckles that contain RNA splicing factors. To get a more comprehensive view of nuclear architecture, researchers need to image the whole set of DNAs, RNAs and proteins in the same sample.

During his 125 imaging rounds, Takei included labels for 76 cellular RNAs and 8 nuclear structures and epigenetic markers. As a result, he could see that chromatin architecture, as well as a gene's proximity to nuclear speckles and chromatin modifications, correlated with gene expression. Yet at the single-cell level, cells of the same type showed differences in nucleome structure. The significance of this variation is still uncertain; one possibility Takei suggests is that the organization could reflect different external stimuli.

Xiaowei Zhuang, a biophysicist at Harvard University in Cambridge, Massachusetts, has also collected images of DNA, RNA and proteins together using a technique called multiplexed error-robust FISH (MERFISH), which her group developed for imaging RNA. In the

team's latest work⁸, MERFISH allowed imaging of around 2,200 DNA loci and RNA species in single cells. Antibody stains for nuclear structures completed the picture, helping her team to visualize not just chromatin interactions and other nuclear structures, but also how that arrangement influenced the production of RNAs.

With Zhuang's and Cai's approaches, "you're really looking at spatial organization of the nucleus", says Bing Ren, a molecular biologist at the University of California, San Diego, who wasn't involved in either project. "This is really the future of genomics and epigenomics."

And that future is becoming more widely accessible. Vizgen, a genomics company in Cambridge, Massachusetts, now sells a custom system for MERFISH studies, called MERSCOPE. (Zhuang is a co-founder of and consultant for the company.) 10x Genomics, based in Pleasanton, California, is also commercializing multiplex and other spatial technologies.

Meanwhile, researchers continue to innovate, for instance by combining imaging techniques with enhanced resolution methods, such as STORM, which maps chromosome domains in fine detail, and expansion microscopy, which physically expands the volume of specimens to make *in situ* RNA sequencing more visible. They are also devising ways to make chromosome structure data easily available, for example through the 4D Nucleome Data Portal, where scientists can search and visualize data on nuclear components. "It's almost like having a genome browser," says Ren, "but now, in the 3D form."

Wang says she sees two main applications for such data. One is to study subcellular biology, including genome organization and cellular distribution of RNAs. The other is to delineate different cell types in a complex tissue on the basis of their nucleome arrangements. With her own imaging-sequencing technique, called StarMAP, Wang is mapping chromatin, RNAs and proteins in the nuclei of several organs from mice and humans. Those data form the early stages of a new kind of cell atlas, which she hopes to share in the next couple of years.

The pace of innovation is frenetic, but invigorating, says Wu. "Inventions are happening left and right. I think everyone's extremely excited to see what the next year's going to bring."

Amber Dance is a science writer in Los Angeles, California.

1. Flavahan, W. A. et al. *Nature* **529**, 110–114 (2016).
2. Zazhytska, M. et al. Preprint at bioRxiv <https://doi.org/10.1101/2021.02.09.430314> (2021).
3. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. *Nature Methods* **11**, 360–361 (2014).
4. Takei, Y. et al. *Science* **374**, 586–594 (2021).
5. Nguyen, H. Q. et al. *Nature Methods* **17**, 822–832 (2020).
6. Quinodoz, S. A. et al. *Cell* **174**, 744–757 (2018).
7. Payne, A. C. et al. *Science* **371**, eaay3446 (2020).
8. Su, J.-H., Zheng, P., Kinrot, S. S., Bintu, B. & Zhuang, X. *Cell* **182**, 1641–1659 (2020).