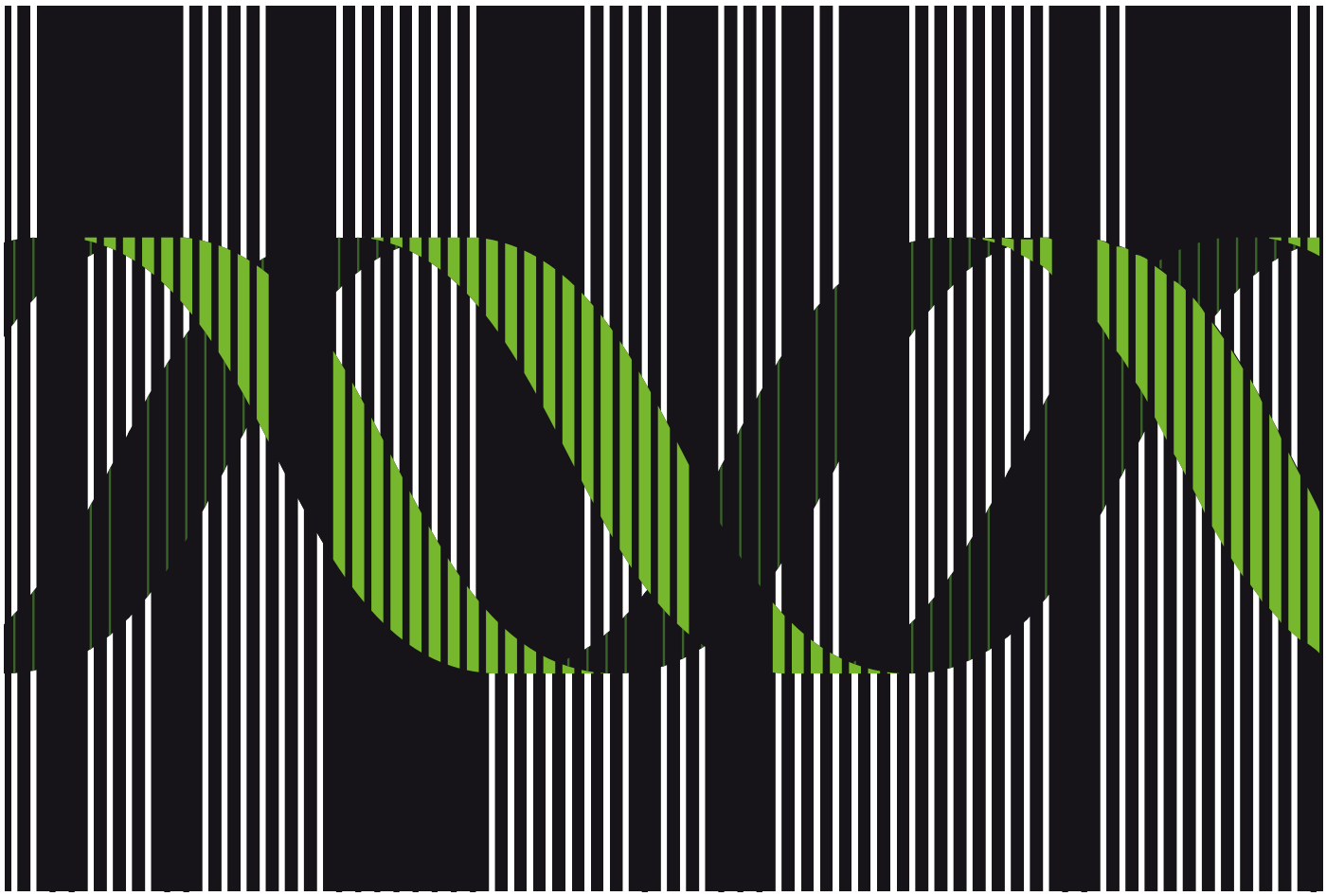


ILLUSTRATION BY *NATURE*

MOLECULAR BARCODES REVEAL TUMOUR LINEAGES

Researchers are blending tools from developmental biology with technologies such as cell sorting and CRISPR to gain fresh insight into cancer. **By Jyoti Madhusoodanan**

What Anna Obenauf needed was a time machine. In 2016, Obenauf, then at Memorial Sloan Kettering Cancer Center in New York City, read a case report about a person with melanoma who, in 2011, had been given a drug that blocks a cancer-promoting form of the protein BRAF. Before treatment, the person's torso was covered with tumours. Within weeks, the growths had disappeared. But six months later, they were back — and were resistant to the drug. The person died.

The study[1] identified the mutation that made the tumours resistant, but Obenauf, a cancer researcher, wanted to know where that mutation came from — a tricky task without a time machine. When did that first mutant cell arise, and how was it different from others in the tumour? Answering those questions could help researchers to spot and overcome treatment resistance, which causes many cancers to recur. But the only way to do that was to catch the first mutant cell — known as a founder clone — before it was exposed to drugs, and that clone was long gone. So Obenauf, who now leads her own laboratory at the Research Institute of Molecular Pathology in Vienna, began seeking out more practical options.

Efforts to observe how cells evolve or adapt over time typically rely on lineage tracing. This method, which follows a particular cell and its descendants as they divide and spread in an organism or tissue, traces its origins back to developmental biology, in which scientists track how a single fertilized cell differentiates and grows into a fish, plant or person. As early as 1973, researchers were injecting black ink into the cells of freshwater invertebrates called hydra; the ink, visible as a black dot, was passed to daughter cells on division, enabling scientists to see where these cellular offspring ended up in the hydra's body.

Today, barcodes built of nucleic acids or protein work similarly. Like the sticker on a piece of fruit at the grocery shop, these molecular barcodes mark cells so that researchers can identify, track and study the populations that arise from a cell. Cancer biologists use them to home in on a tumour's origins or understand how its mutations change over time. But reading molecular barcodes requires breaking cells open to sequence DNA. Although it's possible to track millions of cells through time in this way, the cells are destroyed in the process and

are thus unavailable for further experiments.

Now, researchers are expanding the tumour-barcoding toolkit to avoid destroying barcoded cells. New techniques instead enable the cells to be isolated and injected into animal models to understand drug resistance, or to be identified in tumours and then modified *in situ* to see how the tumour microenvironment changes. Some teams have coupled DNA-sequence-based identifiers with gene-editing circuits and fluorescent proteins to detect a cell's molecular and behavioural changes, using flow cytometry, fluorescence microscopy or other tools. Others have similarly adapted molecular barcodes so that they can be applied in high-throughput workflows, and then used to sort live cells.

"A lot of the excitement around these barcoding tools has been just how easily they slot into current workflows," says Amy Brock, a cancer researcher at the University of Texas, Austin. "They expand the capabilities of workflows that everybody is already doing."

Tracking drug resistance

Whatever the application, barcodes broadly fall into two categories, says Barbara Grüner, an oncology researcher at University Hospital Essen in Germany. The first are simple DNA sequences that allow researchers to follow cell states and fates. The second enables a certain transcript or cell to be tracked down for further analysis.

In both cases, it is theoretically possible to use the technology to individually label every cell in an organism or tumour, although the data would probably be too complex to be useful, says Brock. When her team started using barcodes, she says, "We were so excited by the combinatorial possibilities of being able to measure 10 million different labels. But it's not very practical, and it turns out you don't really need that many." Instead, researchers typically strive to balance experimental complexity with cost.

For Obenauf, there was another key consideration. Existing barcodes could help her to trace cellular lineages, but not to isolate them. So, she built her own system. In a study[2] published last year, Obenauf and her colleagues created a library of some 130 million barcodes, each of which was connected to a fluorescent protein, and used them to tag a melanoma cell line so that each cell contained a unique barcode. (The barcodes also contained a short piece of RNA, which the team could use later to guide components of the CRISPR–Cas9 gene-editing system.) They then injected those tagged cells into mice, causing tumours to grow, and treated half the animals with a compound that inhibits the cancer-promoting version of the BRAF protein. The remaining animals were untreated.

As in the 2011 case report, tumours in the treated animals initially disappeared, but then reappeared as cancer cells became resistant to the inhibitor. By reading the barcodes from both sets of animals, the researchers found that the cell populations in the two groups had diverged: there were 17-fold fewer barcodes in resistant tumours relative to untreated ones, with only 2–6 barcodes in each tumour. This showed that the treatment had killed off most cell types, leaving only a few drug-resistant clones.

That's a dramatic illustration of the selective pressure anticancer therapies can exert, but no more than conventional barcodes could reveal. "This is where most traditional barcoding tools stop," Obenauf says. "You perform this next-generation sequencing and identify the barcodes, then do statistical analysis. But we go a step further" — in this case, isolating individual cells for further analysis.

"A lot of the excitement around these barcoding tools has been how easily they slot into current workflows."

Using a strategy called CaTCH (CRISPRa tracing of clones in heterogeneous cell populations), the team targeted the barcodes using CRISPR–Cas9 gene editing. The researchers used a short piece of RNA to guide a modified Cas protein towards the barcode, inducing expression of a fluorescent protein gene alongside it. The resulting fluorescence provided a handle that Obenauf's team could use to isolate those particular barcoded cells using flow cytometry. When they injected the isolated cells into new animals and repeated the experiment, Obenauf got her answer: the tumours didn't start out resistant to BRAF inhibition, but acquired it in response to the treatment.

"Targeted therapy is not usually thought to induce mutations," Obenauf says. That is more typically associated with harsher treatments such as DNA-damaging chemotherapy or radiation. "I was actually surprised because I wasn't expecting a mutation to cause this resistance."

A powerful approach

Coupling CRISPR with DNA barcodes is a powerful approach, says Brock, who has also developed a barcoding system based on guide RNAs[3]. Developers have turned the CRISPR system into a kind of molecular Swiss army knife, and researchers can exploit those different tools to achieve multiple goals. Brock's team has developed a technique called ClonMapper, and used it to induce selected tumour-cell populations to die. "The barcodes are not only these very powerful and quantitative labels of some population, but we can also use them to activate gene circuits in clone-specific ways," she says. "There's a lot of possibilities."

In exploring those possibilities, barcode length is a crucial factor, Brock says. A longer tag means more cells can receive unique labels. But sequencing costs and data interpretation can become impractical. In their study, Brock and her team decided a 20-nucleotide sequence was sufficient to create strong binding sites for guide RNAs. But shorter tags can yield insight into cancer, too.

Complex tumour microenvironments are tough to mimic in a plastic dish, but have crucial roles in driving drug resistance and cancer spread. They can also pose hurdles in drug development: molecules that seem promising in a dish often fail to work as well in animals. But researchers cannot screen hundreds of molecules in animal models. "It's not justifiable on an ethical level, of course, but also not in terms of time or money," Grüner says.

To bridge that gap, Grüner and her colleagues generated distinctly labelled copies of a pancreatic cancer cell line and grew them in 96-well plates. Each well was treated with one of 712 compounds, and to differentiate the various groups, the team used a six-base barcode coupled to a fluorescent protein — enough to resolve 4,096 treatments (or 4^6 — 4 base possibilities at each position of the 6-base sequence). "You can't make [the barcode] too much shorter, because if single-base mutations arise, you can't distinguish them in a sequencer," Grüner says. "But it's not adding anything to make them longer."

The team pooled the treated cells and injected them into animals. After about two days — long enough for cells to begin spreading, but not so long they would divide and lose the effects of drug treatment — the team retrieved fluorescence-tagged cancer cells from the mouse lung, where pancreatic cancer cells often spread. The researchers then sequenced them to count the barcodes. A decrease in the frequency of a given barcode indicated that the associated treatment had worked: it had inhibited cells with that barcode from spreading.

One treatment that worked particularly well, the team found, targeted not cell division or cell-death pathways, but a lipase, an enzyme that digests fats. The team's subsequent testing suggests that the compound prevents cells from squeezing through blood-vessel walls to infiltrate tissues[4]. "It influences the physical properties of the cell — how sticky or squishy it is — rather than more conventional aspects, like killing the cell," Grüner says.

The team might not have seen this type of change using other methods, says Grüner. And, thanks to the barcoding strategy, she adds, the researchers were able to make that observation using just 36 mice, rather than the hundreds that would typically have been needed.

Spatial specifics

Other researchers are using barcodes to explore the spatial organization of tumours. At Stanford University in California, for instance, a team led by Garry Nolan and Julien Sage is exploring

the tumour microenvironment using barcodes built not of nucleic acids, but of proteins.

The scientists started with a set of six peptides and chose a combination of three peptides per barcode to create a library — an approach they call six-choose-three. To detect the barcodes, the team uses a technique called multiplexed ion beam imaging (MIBI). Each peptide has a corresponding antibody attached to a unique heavy-metal atom, which produces a distinct MIBI signal. By treating barcoded cells with all six antibodies and then probing them in a mass spectrometer, the team can capture the tumours' spatial organization. Like mixing colours on an artist's palette, each unique three-peptide barcode (called an EpicTag) produces a different combination of signals, and thus, a different (false-colour) hue.

The researchers applied this approach to small-cell lung cancer, labelling 20 human lung cancer cell lines with unique barcodes and pooling the samples to generate tumours in mice.

Small-cell lung cancers contain cells with different genetics, and often a mix of two functionally distinct cell types, known as neuroendocrine and non-neuroendocrine cells. Using their barcodes, the researchers scanned tissue sections and watched how distinct subgroups of cells arranged themselves in the tumours. They found that cells with the same barcode, although genetically identical, could form both functional subgroups. Within tumours, non-neuroendocrine cells tended to form larger clusters, or 'patches', suggesting that they created their own suitable environments, or that neuroendocrine cells somehow inhibited their growth.

The team then knocked out expression of *PTEN*, a tumour-suppressor gene that is often mutated in lung cancer, expecting to see the mutant cells dominate the resulting tumours. Instead, they found that both mutant and non-mutant cells formed patches, hinting at some interaction between the two cellular groups[5]. "All the clonal populations grew in larger patches, indicating that the *PTEN* mutants had some effect on unmutated cells as well, changing their growth," says cancer researcher Alexandros Drainas, who is joint first author on the study.

Although studies of the tumour microenvironment have often focused on the interactions between immune and cancer cells, interactions between subtypes of tumour cells matter, too. With barcoding tools such as EpicTags, it's possible to go from observing spatial patterns to modifying them, says Xavier Rovira-Clavé, a cancer researcher at Stanford and joint first author of the work with Drainas. "We're trying to infer causation and function from these patterns."

Jean Berthelet, a postdoctoral researcher at the Olivia Newton-John Cancer Research Institute in Heidelberg, Australia, is also probing the spatial organization of cancer. Working in the lab of researcher Delphine Merino, Berthelet developed an optical barcoding strategy called LeGO that relies on fluorescence to visualize and capture clones of tumours that have spread. Berthelet and his colleagues used a lentiviral vector carrying the sequences for different fluorescent proteins to barcode a patient-derived breast-cancer cell line so that each barcoded cell received either one or two proteins in different combinations, producing 31 unique colours[6].

When the barcode was injected into mice, the researchers could see how different clones spread (or metastasized) and grew in different tissues. "The first thing we spotted was this striking difference between lung and liver metastases — the lung metastases each contained more colours than in the liver," Merino says.

That difference hinted at varying trajectories of tumour growth in each tissue. When the researchers isolated the different clones, they discovered that cells that colonized lung tissue consistently activated a genetic pathway that spurred metastasis, but did not maintain it. Unlike the cells in Obenauf's study, which had acquired drug resistance through mutation, these cells had changed their gene-expression patterns in a reversible manner.

"It's a very big difference in the biology of metastasis between two organs by the exact same cells," Merino says. "Without barcoding, we might conclude that the cells that migrate to each tissue are different and thus behave differently, but it's actually the same cells behaving differently."

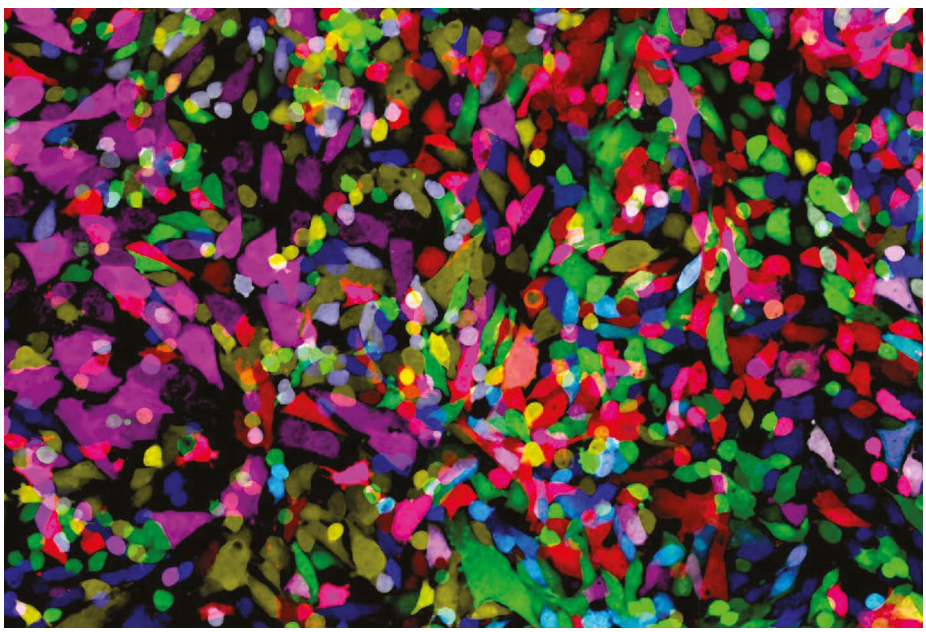
Experimental variables

Whatever the strategy, using barcodes in cancer studies requires careful planning. Researchers need to consider how many unique barcodes are needed to capture the expected populations, Merino says. With optical barcodes, for instance, "it helps to think about the ratio of cells and colours when setting up experiments, so two different cells don't receive the same colours", she says. With genetic barcodes, the aim is for balance: they must be long enough to generate distinct identifiers, but not so long as to make sequencing prohibitive, Brock says.

Start with a smaller library of colours so you can spot meaningful signals in the data, Merino suggests. Although there is no accepted rule of thumb, the minimum is typically three colours (red, green and blue), and the maximum that her team has used so far, *in vitro*, is six.

However unique the barcodes are, be aware that the ability to recover rare cells — although vastly improved by molecular barcoding strategies — is limited by other tools, especially flow cytometry. "If your endpoint is trying to capture a rare cell alive, sorting is always going to be the challenging step," says Brock. "The barcode is specific enough, but the physical step of pulling it out in flow cytometry is hard."

Still, the experimental versatility of molecular barcodes makes them worthy tools, she says. "You can get all this additional information from experiments you're already doing," Brock says. "I tell people that everyone should be barcoding their cells."



Breast cancer cells barcoded using an optical approach called LeGO.

JEAN BERTHELET AND DELPHINE MERINO

Jyoti Madhusoodanan is a science writer in Portland, Oregon.

1. Wagle, N. *et al. J. Clin. Oncol.* **29**, 3085–3096 (2011).
2. Umkehrer, C. *et al. Nature Biotechnol.* **39**, 174–178 (2021).
3. Gutierrez, C. *et al. Nature Cancer* **2**, 758–772 (2021).
4. Grüner, B. M. *et al. Nature Methods* **13**, 883–889 (2016).
5. Rovira-Clavé, X. *et al*. Preprint at bioRxiv https://doi.org/10.1101/2021.06.29.449991 (2021).
6. Berthelet, J. *et al. Sci. Adv.* **7**, eabf4408 (2021).