


LINNAEUS MEETS THE INTERNET

Publication of new species names breaks with more than 200 years of history. go.nature.com/peKLQz

A. CANO, USM / S. KNAJPP

Genomics goes beyond DNA sequence

A technology that simultaneously reads a DNA sequence and its crucial modifications makes its debut.

What makes two individuals different? Biologists now know that the genome sequence holds only a small part of the answer, and that key elements of development and disease are controlled by the epigenome — a set of chemical modifications, not encoded in DNA, that orchestrate how and when genes are expressed. But whereas faster, cheaper and more accurate sequencing technologies have developed rapidly, techniques to map the epigenome have lagged behind.

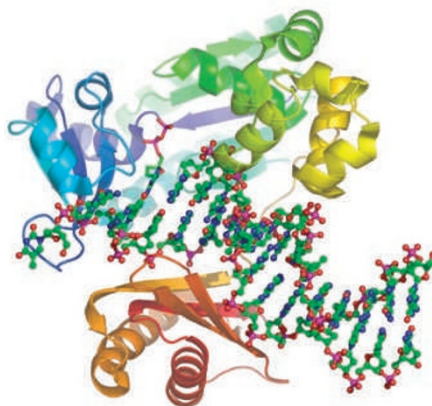
Sequencing company Pacific Biosciences, based in Menlo Park, California, has now developed an integrated system that simultaneously reads a genome sequence and detects an important epigenetic marker called DNA methylation. “I think it’s an important step forward, although I think it is a baby step,” says Joseph Ecker, a plant geneticist at the Salk Institute for Biological Studies in La Jolla, California, who was not involved in the work.

DNA methylation — the addition of methyl groups to individual bases — is just one of many epigenetic markers of DNA and its associated proteins. Others include modification of the histone proteins that DNA winds around to form chromatin — the tightly packed cluster that makes up chromosomes — and the activation of small non-coding RNA molecules.

DNA methylation, which reduces gene expression, is linked to key developmental events, as well as many types of cancer. It is the best-studied epigenetic modification, mainly because tools have existed to study it, says Susan Clark, an epigeneticist at the Garvan Institute of Medical Research in Sydney, Australia.

The gold-standard method for detecting DNA methylation, which Clark’s group developed more than 15 years ago, is bisulphite sequencing, in which unmethylated versions of the base cytosine are chemically converted into another base, uracil. Sequencing the converted DNA allows scientists to reconstruct a genome-wide methylation map. But the technique has several drawbacks. Not only is it expensive and time consuming, it also damages DNA, reducing the map’s accuracy. And it doesn’t detect methylation at adenine bases, which are very prevalent in organisms such as bacteria.

Pacific Biosciences’ approach for detecting DNA methylation, published this month in *Nature Methods*¹, builds on the company’s



DNA polymerase (shown flanking the double helix) can reveal genomic and epigenomic detail.

sequencing technology. The system uses an enzyme called DNA polymerase to read a strand of DNA and build a complementary strand out of nucleotides labelled with fluorescent molecules. As each component is added to the growing strand, it produces a flash of light — the colour of the light corresponds to the identity of the base, and thus reveals the sequence of the template DNA.

Analysing the pulses of light, and the time between them, can also show whether methylation is affecting polymerase activity. This has now been exploited to detect methyladenine, methylcytosine and a poorly understood modification called 5-hydroxymethylcytosine. “We foresee with this technology that in the future there will be a unification of the fields of epigenomics and genomics,” says Stephen Turner, the company’s founder and chief technology officer.

Game changer?

Although the data are promising, obstacles remain. “There are distinct advantages, but we’re not rushing out tomorrow to apply this because it’s not prime time for human methylation mapping,” says Ecker.

One problem is that although the technique is great at distinguishing adenine from methyladenine, it doesn’t quite reach single-base resolution for cytosine and methylcytosine. It also lacks one of the key promised benefits of Pacific Biosciences’ sequencing technology: its ability to read long sequences of DNA, up to 8,000–10,000 base pairs, which makes it easier

to assemble the data into complete genomes. Instead, the reported methylation read-length is only about 1,000 base pairs.

Turner says that the company is working to solve these problems. It will ship the first sequencers that use fluorescent labelling this year, and plans to add the methylation mapping capability next year.

“What needs to be done now is to make it robust and accurate,” says Clark, a steering-committee member of the International Human Epigenome Consortium, a bid launched in January to map the epigenome in multiple cell types². “There’s a lot of troubleshooting that needs to be done to get it to be accurate enough to be able to compete with bisulphite sequencing.”

Several companies are working on similar technologies. UK-based Oxford Nanopore Technologies published a report last year showing that it could detect methylated DNA at a single-molecule level³. But that system and others are still at an earlier stage of development.

Some say that the promise for such a technique is huge. Bisulphite sequencing for a single human genome can cost up to US\$100,000, says Robert Martienssen, a geneticist at Cold Spring Harbor Laboratory in New York. With the latest technique, the cost of a full-genome methylation map would drop to \$100–1,000, he says. “That will change everything.”

There is no shortage of epigenetic questions ripe for probing. One example is in tumour biology, where different cancer cells are likely to have different methylation patterns. Another is how cells in a single organism take on different functions despite having identical genomes. “This is exactly the technology you could use to look for epigenomic changes in specific cell types,” says Martienssen, who is also on the International Human Epigenome Consortium’s steering committee.

Ecker says researchers still haven’t pinned down the significance of, say, having a methylation mark in one position and not another, and what’s really needed is more studies that unify genomic and epigenomic information. “As you get more genomes to compare, then of course the differences take on some meaning,” he says. “We’re just lacking numbers at this point.” ■

Alla Katsnelson

1. Flusberg, B. A. *et al. Nature Meth.* doi:10.1038/nmeth.1459 (2010).
2. Abbott, A. *Nature* **463**, 596–597 (2010).
3. Clark, J. *et al. Nature Nanotechnol.* **4**, 265–270 (2009).

LAGUNA DESIGN/SPL