GENOMICS

Story of C's

Chromosome conformation capture (3C) graduates to a tool for mapping interaction networks, thanks to a clever way of making molecular 'carbon copies'.

Functional genomic annotation initiatives, such as the ENCODE project, are churning out an incredible number of regulatory elements in the human genome. The question that emerges now is how each of these elements can be connected to their target genes. Indeed, regulatory elements in mammalian systems are capable of acting on genes located far away on the same chromosome or even on a different chromosome.

A few years ago, Job Dekker, now at the University of Massachusetts Medical School, developed the chromosome conformation capture assay, known as 3C, precisely for the purpose of identifying physical associations between distant genomic regions. The principle of 3C is to capture these physical interactions by cross-linking associated regions of the genome (**Fig. 1**). By a remarkable molecular *haute couture* process, the cross-linked fragments are then cut with a restriction enzyme and stitched together to form a library of ligation products, each representative of a physical interaction and detectable by PCR. The problem with 3C is that pairs of interacting elements have to be tested one by one, therefore limiting the scope of analysis.

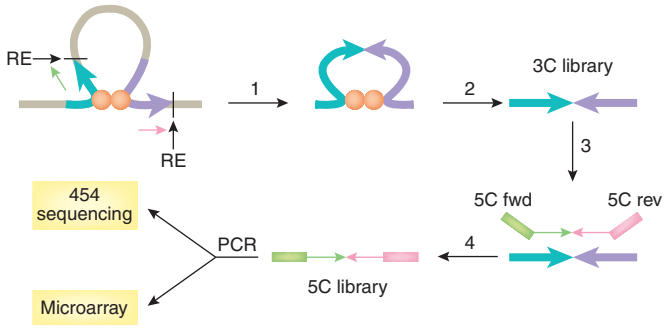

Figure 1 | The 3C method captures physical associations between distal genomic regions using cross-linking followed by restriction enzyme (RE) digestion and ligation (1), then cross-link reversal (2). A 3C library can now be copied and amplified into a 5C library, by multiplex hybridization of 5C primers (3) and ligation-mediated amplification (4). Adapted from Dotsie *et al.*, 2006.

To simultaneously interrogate a 3C library for many interactions why not make… carbon copies? This is the principle of 5C, which stands for 3C carbon copy, the new method from the Dekker laboratory, recently published in *Genome Research*.

The 'reproduction process', as you will have guessed, is more complicated than your usual trip to the Xerox machine. It relies on ligation-mediated amplification and careful design of primers at the cleavage sites of the restriction enzyme used to prepare the 3C library. A large number of these 5C primers are hybridized to the 3C library. When a 3C

IMAGING AND VISUALIZATION

THE INSIDE TAG

An uncharged CoA precursor that can enter the cell is used for covalent, site-specific labeling of proteins inside living cells.

Live-cell imaging allows an unprecedented window into the cell, but it requires more specific and efficient covalent labeling methods than the popular fluorescent protein fusions and antibody conjugates. Although these are powerful techniques, the large size of fluorescent protein tags and the membrane impermeability of antibodies make them less than ideal for *in vivo* labeling.

To get around these problems, Michael Burkart and his colleagues at the University of California, San Diego use the natural phosphopantetheinyltransferase (PPTase) enzyme to covalently label peptide carrier protein fusions with analogs of coenzyme A (CoA) that contain reporter molecules. "If you could somehow sneak your CoA analog inside the cell, your organism will label [the carrier protein fusion] with your analogs because most of these PPTases are very permissive of the identity of what is coming off the CoA molecule," explains Burkart.

The uncharged CoA precursor that his team had tested was pantetheine. When they added a fluorescently labeled pantetheine analog to *Escherichia coli* cells expressing a carrier protein fusion and PPTase, the fusion protein was labeled (Clarke *et al.*, 2005). This was accomplished through uptake of the pantetheine analog, conversion of this analog into a CoA analog through the constitutive CoA biosynthetic pathway, and PPTase transfer of the CoA analog onto the carrier protein.

Now, in an article in the *Journal of the American Chemical Society*, Burkart and coworkers report an optimized synthesis procedure and analysis of pantetheine analogs containing a fluorescent reporter and bioorthogonal tags, allowing for more versatile detection and isolation than is possible with antibody detection techniques (Meier *et al.*, 2006). The pantetheine analogs were all efficiently converted into CoA analogs and attached to the carrier protein. Pointing out the broad applicability of this method, Burkart says: "What's nice about these is that as long as you can make some CoA analog, you can covalently attach that to a fusion protein. And it's fairly limitless what you can attach to CoA and attach to a carrier protein [with these tools]."

Notably, for the first time, Burkart and colleagues demonstrated *in vivo* carrier protein labeling using these molecules inside the cell. Although other groups have been able to label cell-

template allows them to do so, two 5C primers bind in close proximity and are ligated to each other; then the product is amplified by PCR (**Fig. 1**). The resulting 5C library can be further amplified and analyzed using a high-throughput readout such as microarray or 454 sequencing.

The new method can be used in several configurations. One can ask how a single or a handful of 'fixed' elements interact with the neighboring chromatin, which the authors demonstrated on a 400-kb region containing the well-characterized β-globin locus. Alternatively, 5C is amenable to mapping a network of interactions between two large sets of elements. Dekker's group performed a proof-of-principle experiment examining a 100-kb conserved gene desert region, and they now plan to establish such maps for the ENCODE regions.

This is where the readout by sequencing comes in handy. Even if considering a single small chromosome, the number of possible interactions between the genes and regulatory elements present on this chromosome can easily reach several million. "By sequencing," explains Dekker, "you will only detect those interactions that are actually occurring, but with a microarray you would have to represent each possible combination." Capitalizing on the new sequencing technologies such as 454, which works optimally on short reads, Dekker's group specifically designed the 5C method so that short reads would be sufficient.

"With sequencing methods allowing you to read 100,000 or even 1 million sequences," explains Dekker, "this really opens up the possibility to map this network of interactions that we believe occur in the genome... linking things that are far apart on a chromosome, but actually are functionally related and physically associated."

Veronique Kiermer

RESEARCH PAPERS

Dostie, J. *et al.* Chromosome conformation capture carbon copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **16**, 1299–1309 (2006).

surface proteins using a CoA-based method, the inability to get the CoA analog inside the cell limited this technique to use only for membrane-protein modification. In addition to bringing this methodology inside the cell, the researchers plan to extend this technology to species other than *E. coli*. They also would like to improve the detection limit to allow work with native protein levels instead of relying on overexpression of the carrier protein fusion and PPTase, as they did in this study.

As they continue to develop this CoA-directing technology, Burkart notes that other researchers can now use their own fusions to study protein function inside the cell. "I think what will be exciting in this area now is to see others use their carrier protein fusions inside the cell with our CoA delivery technique to figure out what happens to proteins in the cellular context," he adds.

Irene Kaganman

RESEARCH PAPERS

Clarke, K.M., *et al. In vivo* reporter labeling of proteins via metabolic delivery of coenzyme A analogues. *J. Am. Chem. Soc.* **127**, 11234–11235 (2005).

Meier, J.L. *et al.* Synthesis and evaluation of bioorthogonal pantetheine analogues for *in vivo* protein modification. *J. Am. Chem. Soc.* **128**, 12174–12184 (2006).

NEWS IN BRIEF

PROTEOMICS

Automating phosphoproteomics

One of the most powerful applications of mass spectrometry–based proteomics is the ability to identify and map post-translational modifications. As data sets grow larger, however, the manual validation of such sites becomes nearly impossible. Beausoleil *et al.* have devised an automated scoring tool they call 'Ascore' that, using the intensities of site-determining ions from tandem mass spectrometry data, measures the probability of correct phosphorylation site localization.

Beausoleil, S.A. *et al. Nat. Biotechnol.* **24**, 1285–1292 (2006).

CELL BIOLOGY

Visualizing disulfide reduction

Yang *et al.* present a FRET-based strategy to image disulfide-bond cleavage in live cells. Using the folate receptor as an example to probe the endocytic pathway, they created a FRET reporter consisting of folate, rhodamine and BODIPY. When the disulfide bond between folate and rhodamine is cleaved, the fluorescence switches from red to green, providing a visual readout of cellular location.

Yang, J. *et al. Proc. Natl. Acad. Sci. USA* **103**, 13872–13877 (2006).

CHEMICAL BIOLOGY

Reversing reactions to make new sugars

Sugar-based natural products with therapeutic activities are notoriously difficult to make using organic synthesis alone. With the discovery that several glycosyltransferase enzymes can catalyze reversible reactions, Zhang *et al.* describe a powerful tool to synthesize exotic natural product variants and to incorporate chemical handles onto sugar scaffolds.

Zhang, C. *et al. Science* **313**, 1291–1294 (2006).

MICROBIOLOGY

Identification of parasite genes silenced to evade host immunity

A *Vibrio cholerae* strain expressing a Tet repressor–sensitive GFP and a transposon carrying the repressor permits the identification of genes that are silenced as the parasite adapts from the surface-water environment to its mammalian host. Transposon-containing bacteria clones turn green only when adapting from *in vitro* culture to the infection of infant mice, allowing identification of genes that are silenced to escape host immunity.

Hsiao, A. *et al. Proc. Natl. Acad. Sci. USA* **103**, 14542–14547 (2006).

IMAGING AND VISUALIZATION

Improving FlAsH

The fluorogenic, membrane-permeable biarsenical dye called FlAsH has been indispensable for the chemical labeling of proteins in cells, requiring only that the protein be engineered with a tetracysteine motif. Spagnuolo *et al.* have developed a new biarsenical dye they name F2FlAsH, which exhibits higher absorbance, quantum yield and photostability, with a reduced dependence on pH as compared to the original FlAsH.

Spagnuolo, C.C. *et al. J. Am. Chem. Soc.* **128**, 12040–12041 (2006).