



A Conversation with Job Dekker

INTERVIEWER: RHIANNON MACRAE

Editor, Trends in Genetics

Job Dekker is a Professor in the Department of Biochemistry and Molecular Pharmacology and the Program in Systems Biology at the University of Massachusetts Medical School and an Investigator of the Howard Hughes Medical Institute.

Rhiannon Macrae: Your work is largely about the genome in three dimensions. I heard that you got your start doing NMR [nuclear magnetic resonance] spectroscopy, so I'm curious how you moved from structural biochemistry to looking at genome organization.

Dr. Dekker: I didn't actually work on NMR, but I worked with the people who did the NMR; I was the biochemist purifying the proteins. When I joined Nancy Kleckner's lab I had the notion that we could use biochemical approaches similar to what people were doing with NMR for proteins to study chromosome structure. If you look along a protein chain to see which atoms lie near each other for every amino acid or every atom, you can learn about the protein's structure. I thought that if you could do the same for chromosomes, to see what each locus on the chromosome is close to, we could learn something about folding. It was a crazy idea at the time, but it actually worked out quite well. That concept is what inspired me to develop a method that's now widely used as chromosome conformation capture or 3C.

Rhiannon Macrae: Could you describe how that method works and how it has since been adapted?

Dr. Dekker: There's a whole family of "C" methods now—4C (circularized chromosome conformation capture), 5C (chromosome conformation capture carbon-copy), Hi-C—but the basic concept is the same for all of them. The three-dimensional (3D) structure of a chromosome is made up of loci that are very near or even physically touching other loci. The method fixes those interactions in place, and the DNA can then be cut into small pieces. Things that were close to each other in three dimensions will continue to be linked to each other and we can religate them to make a ligation product that tells us that these two loci were in close proximity in 3D space. Then we can interrogate that sample by PCR [polymerase chain reaction], as in 3C, or by deep sequencing, which is ultimately what led to the Hi-C and 3C-seq methods. You end up with a contact map for the whole genome. You can see for every locus what it is close to in 3D space.

Rhiannon Macrae: Anytime someone talks about chromatin they describe this huge expanse of linear DNA jammed into a tiny space. Obviously, everything is going to be touching or nearly touching. How do you sort out what has a function versus what is simply a result of packing?

Dr. Dekker: That's the problem in this field. When we talk about close spatial proximity what we really mean is that they're not just packed, but that they have to be close enough to each other to cross-link. The biggest challenge for the last few years has been, now that such maps are available, how to interpret them.

First, we have to ask what we mean by "structure," and there are all kinds of structures that we can deduce from the maps. The next question is what you just asked: How can we find whether any of these structures are functional? This is a harder problem, but one that the field is on the edge of making tremendous breakthroughs. To assign function to a structure, you have to be able to perturb it. We can now do that using genome-editing tools. We can alter the 3D structure and then use other genomic methods like RNA-seq to see how it affects genome functions like gene expression. With 3C and Hi-C methods we can make 3D maps of the genome. There are imaging techniques where we can see the 3D genome in live cells in real time. By combining those techniques with genome-editing tools, we can really start to relate structure to function.

Rhiannon Macrae: Can you describe what a TAD is and the insights that have come out of discovering TADs?

Dr. Dekker: TADs—topologically associating domains—were defined in 5C and Hi-C interaction maps as contiguous chromosomal regions where all the loci in that region have an elevated interaction probability, as if they're all clustered in space. In mammalian cells they range from several hundred kilobases to very large ones several megabases in size. In flies and yeast and other organisms where people have found them that have a much smaller genome, they're only around 60 kb.

Rhiannon Macrae: Do they scale with genome size? If you have a really large genome are you more likely to have a large TAD?

Dr. Dekker: Not always. In collaboration with Barbara Meyer's lab at Berkeley we recently found similar types of domains in *C. elegans*, specifically on their X chromosome. The *C. elegans* genome is rather small, just like flies, but the TADs are the size of mammalian TADs. This goes back to your question: "What is a TAD?" We really don't know. They are defined as structural units. They might be loops. Some people have found that the ends of the domains interact to form loops, and that could result in the formation of a compact chromatin domain where everything inside that loop would interact frequently and they would live inside the TAD. But we now have these similar domains in *C. elegans* and they are different.

We predict that all the genes inside such domains, if they're functional, would be related. They probably share the same enhancers or some regulatory programs. They don't see most of the rest of the genome. That would predict that all the genes in the TAD would be somehow correlated or coordinated in their expression. That's the case in many examples that we looked at, which suggests that TADs are not only units of chromosome structure but also functional domains.

TADs can also be used to examine human disease genetics. There have been large efforts around the world to identify elements that carry genetic variations linked to disease. Often, these regions are linearly far removed from the regions they regulate, making it difficult to identify the genes involved in the disease. Using TAD information helps define target genes that could be affected by such variants, and many groups have already done this. It's been very satisfying to see my early interest in the structure of the chromosome having a direct impact on our understanding of the causes of human disease.

Rhiannon Macrae: You mentioned human disease. Does that suggest that we have the same TADs? Are TADs conserved between individuals and between species?

Dr. Dekker: They're very conserved between individuals, even between different cells in one individual, and between different species. It's a universal architecture but there are interesting variations to it. It would not be surprising if cells somehow use TADs to regulate genes. You could imagine having two TADs next to each other, with a gene present in one of them. It normally wouldn't be exposed to the regulatory elements in the other TAD, but the cell could lift the boundary between them and suddenly that gene is exposed to a new stretch of the genome with new regulatory elements, and sometimes these might be used during development to turn genes on or off. There are several examples in the Hox clusters where this could happen.

Rhiannon Macrae: Is anything known about what happens to TAD structure during genome reprogramming or genome-wide demethylation?

Dr. Dekker: We don't yet know the answer to that. We do know that every time cells go through mitosis TADs are disassembled, then they all reform when the 3D genome forms in early G₁. The cell has to constantly build and fold and unfold these TADs. I imagine that in reprogramming, the very pluripotent state is also very fluid. The TADs are there but they seem to be weaker than in differentiated cells. We don't know why.

Rhiannon Macrae: What do you mean by weaker?

Dr. Dekker: They look more diffuse. They look less separate from each other, as if there's more chances for mixing, as if the whole genome is still a little bit undecided as to how it's going to regulate its genes. Several people are looking into this process of pluripotency versus differentiation and then reprogramming of cells. I'm not familiar enough with the current data to know what the answer really is, but I wouldn't be surprised if there are some changes happening during that process.

Rhiannon Macrae: People have been working for a long time on proteins and whether or not you can computationally predict their 3D structure just from the amino acid sequence. Can we predict TAD structure from the linear DNA? Do you think that what's governing TAD formation is partly specified by the sequence?

Dr. Dekker: It would be great to be able to predict the 3D structure of a genome because it would suggest we understand how it works. We're still not able to do that, but I think we now understand some of the rules. The problem is that many of these rules were deduced from correlation studies like when the gene is "on" it is at a different location physically in the 3D genome than when it's "off," so transcription must be important. And yes, some things will be hardwired that we'll be able to deduce from the DNA sequence; it may be a little like protein structures, where you can place a hydrophobic amino acid somewhere and it will end up in a hydrophobic core of the protein. But I don't think that's the right way to do it. We have to use approaches such as genome engineering to directly manipulate it and see how we can alter things. Wendy Bickmore had a paper in *Science* where she showed that the local chromatin state can reposition a locus in the nucleus. This was a wonderful example of the type of information that is used to build 3D chromosome structures. Ultimately I think that, despite TADs being universal, a lot of 3D genome structure is very cell type-specific, so that's probably dependent on epigenetics. It's more likely we'll be able to predict 3D structure not from the DNA sequence but rather from the genome's linear epigenome: histone patterns, DNA methylation features, transcription patterns, nucleosomes. We still have to learn these rules but I think the field is poised now to start to ask these questions.

Rhiannon Macrae: We're honoring 150 years of Gregor Mendel's discoveries. Do you remember learning about Gregor Mendel's laws of inheritance or doing a Punnett square?

Dr. Dekker: I was probably in high school. I was intrigued by genetics from the beginning. What I like about it is that it's so simple once you know it. As Ken Zaret pointed out, if you just do that first

cross and don't know genetics, it's a mystery, it's magic. I think it's wonderful to know we've learned some of the details of the mechanisms and I still think it's magic.