

People & Ideas

Marian Walhout: Transcriptional mapmaker

Walhout uses the genome as a base camp for exploring transcriptional regulation.

Trying to pick your way through all the factors that govern differential gene expression patterns in multicellular organisms can be a tricky business. A quick glance at Marian Walhout's maps of the interactions between transcription factors and DNA elements could easily leave one feeling more lost than ever. But Walhout, who constructs these maps using high-throughput screens and computational methods, successfully navigates these complex networks by following the trail of specific biological questions.

Walhout's interest in gene expression began as a graduate student at Utrecht University in the Netherlands, where she studied the DNA-binding properties of the transcription factors c-myc and max (1, 2). As a postdoc with Marc Vidal at Harvard Medical School, Walhout branched out into the developing field of systems biology, using large-scale yeast two-hybrid screens to map protein interaction networks in *C. elegans* (3, 4). In 2003, Walhout began her own laboratory at the University of Massachusetts Medical School, where she turned her mapmaking skills back toward understanding transcriptional regulation. Her group used yeast one-hybrid screens (Walhout calls this a gene-centered approach) to chart the protein–DNA interactions that control gene transcription in worms, revealing some fundamental design principles of this regulatory network (5, 6). More recently, Walhout has expanded her focus to include the post-transcriptional regulation of gene expression, exploring how micro-RNAs (miRNAs) function on a genome-wide scale (7).

In a recent interview, Walhout mapped out the course of her own scientific career.

A GLOBAL VIEW

How do you define systems biology?

Systems biology is still a relatively new field, so I think that there is no clear-cut definition. The way I would phrase it is

that rather than looking at a single gene or molecule, or even at a few, you really want to understand the relationships between different molecules, how these are organized into networks, and relate them to global phenomena that you see in the living system.

That sounds a little vague and abstract, but really, in my opinion, that's what it is: asking questions that go beyond a single gene or protein.

How did you become a systems biologist?

My PhD was challenging. I worked on the transcription factors myc and max. Myc is one of the most insoluble proteins in the world—it took me about two and a half years to make enough soluble protein using vaccinia virus for DNA binding studies. It was frustrating to be in the cold room when things didn't work. I was close to calling it quits, but I still loved science and wanted to give it another shot.

I wanted to learn more about genetics, and to work with yeast. And I was vaguely infatuated with the human genome project—this was in 1996–97—although

I didn't quite get how it was going to help us. I contacted Marc Vidal, looking for a postdoc. He had these very ambitious, “crazy” ideas that I really liked, and when we first met and started talking science, I realized we thought in very similar ways. He's

been the biggest influence on my career on many different levels, and he's become one of my best friends; it's great to talk with him about science and other things. People think we've been married for 50 years. We'll say the same thing and people start laughing.

The whole systems biology approach suited me much better, but my love for transcription, which I learned in my PhD laboratory, never really went away. So when I started my own laboratory, I was fortunate enough to combine the two loves into one.

“The greater complexity of humans has to come from somewhere.”



Marian Walhout

Why does gene expression fascinate you?

Relatively simple multicellular organisms like *C. elegans* have roughly the same number of genes as humans. We have 25,000 genes, worms have 20,000. The greater complexity of humans has to come from somewhere—one hypothesis is that transcriptional and post-transcriptional regulation is much more complex in humans. My goal is to understand how turning on and off, and fine-tuning the expression of all the genes in the genome makes a functioning organism. Right now we're trying to understand this at a global level in *C. elegans*, and it's fascinating; there's still very little known.

GENE-CENTERED

You describe your approach to mapping transcription regulatory networks as gene-centered. What does that mean?

Most people trying to understand transcription at a global scale are doing ChIP—chromatin-immunoprecipitation. That's a transcription factor-centered approach, or protein-to-gene: you pull down a factor, and identify a bunch of DNA fragments by PCR sequencing or microarray (ChIP-chip). It's beautiful, but it does have its limitations. If you have a transcription factor with a restricted expression pattern—maybe it's expressed at very low levels or only in one or two cells of the worm—then ChIP isn't technically feasible yet. And there aren't enough good antibodies to do this on a genome-wide scale, which partly explains the paucity of data for 99% of transcription factors in the human genome.

So we go the other way around. We start with a piece of DNA, and identify the transcription factors that can bind to it—what we call a transcription factor binding profile—using the yeast one-hybrid system. That goes from gene to protein.

I'm not saying that one hybrid is better than ChIP. I cannot stress enough that they're complementary. We have limitations too: not everything we find is necessarily meaningful in vivo and we can't detect transcription factor heterodimers. I think you need to combine approaches like these with computational biology to get to the answers.

How did you use this approach to study the regulation of miRNA expression?

Uri Alon has shown that there's little feedback in purely transcriptional networks even though, as biologists, we all know that feedback mechanisms are everywhere. So a talented graduate student in my laboratory, Natalia Martinez, said "Isn't that strange? I wonder if there's feedback in regulatory networks once we incorporate miRNAs." She set out to delineate an initial miRNA regulatory network using yeast one-hybrid. It's a great example of the gene-centered approach—if Natalia had addressed this question by ChIP she would have had to immunoprecipitate all 940 transcription factors in all conditions in all cells, which of course is not feasible.

We combined the miRNA regulatory network with a computationally derived network of miRNA targets. And yes—there are many feedback loops in which a transcription factor regulates a miRNA, and is then itself regulated by that miRNA. It's something that you see more often than you would expect by chance.

Now, this is all very cute, but then you think why is that? The system has evolved to use these loops over and over again. This brings you to the design principles of networks and I think the field of synthetic biology is going to be very important for answering those kinds of questions.

What's harder: obtaining high-throughput data, or extracting meaning from it?

Well, it's challenging to generate data because we always want to do more. But, for me, the

bigger challenge is that, although I'm 100% an experimentalist by training, we do quite a bit of computation, math, and statistics in my laboratory. So I'm learning every day from the people in my laboratory, and I love it.

But, as I say to my laboratory, it's important to constantly think what the question is that you want to ask. We always try to relate our observations back to biological principles and biological questions.

What advice would you give a cell biologist who wants to take a more systems-based approach to their question?

If you believe in something, then go for it. When we started out, people said it was never going to work, and I wanted to prove at least to myself whether it would or wouldn't. And don't be scared of the scale of the project. It's a mindset. Working on a single protein can be repetitive too, if you have to immunoprecipitate it for months in a row to get your data. Only the scale is different. But once you get into it, it's not such a big deal at all.

A RECIPE FOR SUCCESS

Were you tempted to return to Europe after your postdoc?

I came to UMass with my husband, Job Dekker—who is also a scientist—and it's the best decision we could've made. On my very first day, I remember standing in an empty laboratory with my first ice bucket and thinking, "Wow, this system has enough belief in us and what we want to do that it gives us a bunch of money and space, and then says: go do it." It's very unusual to get such an opportunity in Europe, and I think that really makes the United States special.

And UMass is a special place too. The science is fantastic and we have a wonderful cohort of collaborative, friendly colleagues. When Job and I started here, I used to say about once a week, "I think UMass is the best-kept secret in science," but I think the secret is out now.

What are your hobbies?

I love cooking for people and trying out new dishes. My friends enjoy my food.

Are you an experimenter in the kitchen, or do you strictly follow the recipe?



Walhout researches ideas for her next culinary experiment.

Oh, no, I'm an experimenter, just like in science. I say to my students that you first have to follow the protocol and make sure it works, and then you can try to improve it. Not the other way around; that's usually not a good idea!

What's next for you?

We want to do a much larger study in *C. elegans*, to interrogate parts of the genome other than promoters. We want to understand the relative importance of promoters versus introns and other sequences, and, with our approach, we can address those questions because we start from the genome.

And another thing that I really want to start doing is human studies. I didn't want to begin with human because people can be nervous about these types of global approaches, and we really wanted to leverage the power of a genetically tractable metazoan like *C. elegans* to follow up on the data that we got.

But now we know how the yeast one-hybrid system behaves, so we're ready to start the much more daunting task of interrogating the human genome. The human genome is 30 times bigger, with large introns and intergenic regions, so it's much less clear where you need to look.

1. Walhout, A.J., et al. 1997. *Nucleic Acids Res.* 25:1493–1501.
2. Walhout, A.J., et al. 1998. *Biochim. Biophys. Acta.* 1397:189–201.
3. Walhout, A.J., et al. 2000. *Science.* 287:116–122.
4. Walhout, A.J., et al. 2002. *Curr. Biol.* 12:1952–1958.
5. Deplancke, B., et al. 2006. *Cell.* 125:1193–1205.
6. Vermeirssen, V., et al. 2007. *Genome Res.* 17:1061–1071.
7. Martinez, N.J., et al. 2008. *Genes Dev.* 22:2535–2549.