Department
of Plant Systems Biology

Interview *with* Yves Van de Peer

Yves Van de Peer has always been interested in evolution. Whether or not this is pure coincidence, his amazement regarding life and its origin brought him to Ghent 3 years ago. Today, he is leading the bioinformatics group in the Department of Plant Systems Biology. When **Ann Van Gysel** visited the group, she found an ambitious, but contented, Group Leader.

Yves has always been occupied with nature, so no one was surprised when he began to study Biology. For his Master's thesis – about evolution (how could it have been otherwise?) – he chose **Rupert De Wachter**'s research group at the University of Antwerp. In 1987, computers in a lab were still more the exception than the rule, and so Yves made his first evolution tree on the one and only computer that the group had. Despite his initial lack of interest in computers, computer fever seized him and he began his doctoral studies. With the help of rRNA, relationships were sought between the most divergent life forms: plants, mammals, fungi and bacteria. Yves is still intrigued by how life has evolved over millions of years and the fact that humans are only one of the many millions of species that have ever populated this planet. And all of it based on pure coincidence…

Yves: 'In fact, I took too long to get my doctorate. But I learned a lot and was able to publish well. Because of the many publications, by the way, I was able to get an FWO post-doc grant. But when it became clear that my options for the future were limited in Antwerp – people didn't really believe in bioinformatics there, nor in molecular evolution – it was time to look for new horizons. An interview with **Axel Meyer**, an authority in the field of evolutionary research, accelerated everything and before I knew it I was in Germany. I deferred my FWO grant, twice actually, because I was learning so much in that period. Before that, I had had relatively few interactions with other scientists, and suddenly I found myself among post-docs from at least 13 different countries. With Meyer, you learned to think big – everyone was very ambitious – and although he led his group (too) severely, I'm very glad to have had that experience. It was an indispensable eye-opener on the level of ambitious scientific thinking. That post-doc really prepared me for my position today as Group Leader.'

What was the focus of your research during your post-doc in Germany?

'With Meyer, I began my research on duplicated genes. People assumed that evolution from a 'gene set' like a bacterial genome can lead to various kinds of bacteria but never to a more complex organism. For that, you need more genetic material, on which evolution can then operate through natural selection. And how do you create more genetic material? By duplicating genes. The first copy fulfills the necessary functions, but evolution gets free play on the second set, so that slightly other functions develop – the beginning of biological complexity.

Fish are very interesting in this regard. There are over 25,000 different fish species! A lot more than mammals or birds. How have they originated? You find the answer in their genome. When we go looking for genes in fish that we also find in humans, we usually find them in 2 copies. Our research has shown that this is a recurring pattern. We came to the conclusion that an extra genome duplication had taken place in fish that had not occurred in the vertebrates on land. That's how we find correlations between our genome studies and the evolutionary findings.'

After 2 years in Germany, Belgium lured you home?

Not exactly. But at that moment they started looking for professors in bioinformatics in Germany. Bioinformatics was really on the rise. It was an opportunity that I wanted to grab. I had also become convinced that I wanted to build an academic career, and a professorship seemed to me to be the ideal way to gain academic freedom. However, I was also contacted at that moment by the universities of Antwerp, Ghent, and Leuven. And after my conversations with **Marc Zabeau**, at that time Head of the Department of Plant

Systems Biology, I decided to try my luck in Ghent. The proposal from VIB was strong enough for me to resist a very tempting offer from Münster. An essential factor was that I got the opportunity to develop a group that would have enough critical mass to make it in the quite competitive field of bioinformatics. And look how far we've come: today, I'm leading a group of 24 people. Marc Zabeau strongly supported the formation of the bioinformatics group, and together with Group Leader **Pierre Rouzé** and the current Scientific Director **Dirk Inzé** we have been able to build a dynamic, young group. We are a multi-disciplinary group with biologists, IT specialists, mathematicians, and physicists. Meanwhile, I was appointed to Ghent University two years ago as well.'

Is your research now more geared toward plants?

'Actually, yes. The main theme of my group is evolutionary research, but this can also be done with a focus on plants. Plants have always trailed behind in this, by the way. For a long time, there was only the *Arabidopsis* genome, then rice. But in order to carry out detailed genome comparisons, these sequences have already diverged from each other a bit too long: some 200,000 years. But in the meantime, we have poplar – situated somewhere in between – and soon *Medicago* as well. Thus, we are able to increase the resolution and contract the evolutionary distance between the various species. In this way, you can identify functional domains and discover new relevant sequence motifs. Comparative genomics is really growing in importance. Through Pierre Rouzé, the lab already had a lot of expertise in the domain of genome annotation, and so we're looking to work together on several genome projects. This will give us quicker access to new data, an important time advantage in a competitive field.

> " I am convinced that systems biology is the future of Biology. And, along with the other groups in the department, bioinformatics can play a major role in this. "

In addition to this, we are focusing on the biological questions that the department is tackling: the cell cycle, for example. When we verify the number of genes involved in the cell cycle in green algae – along with red algae the direct predecessors of plants – we find a single copy of these genes in each case. These marine creatures possess the 'basic genetic machinery'. By contrast, with *Arabidopsis* this is much more complex due to gene duplications. We're trying to find out why this is and whether this influences biological complexity. Thus, we see that in *Arabidopsis* – or better, its ancestors – certainly 3 genome duplications have taken place. We want to verify how important this has become for the origin of new plants and the appearance of new leaf forms, flower forms, and the increase in biological complexity.'

Can your research prove that genome duplications underlie biological complexity?

'For the moment, this is still controversial. Although, today we have clear indications that this is so. Then we're talking about large-scale duplications: thousands of genes at the same time. It's events like these that we believe signify a great leap forward for evolution – the sudden appearance of lots of new species, for example. Flowering plants originated rather suddenly some 250 to 300 million years ago, and then they very rapidly spread over the whole world. We are verifying whether genome duplications played a role in this.

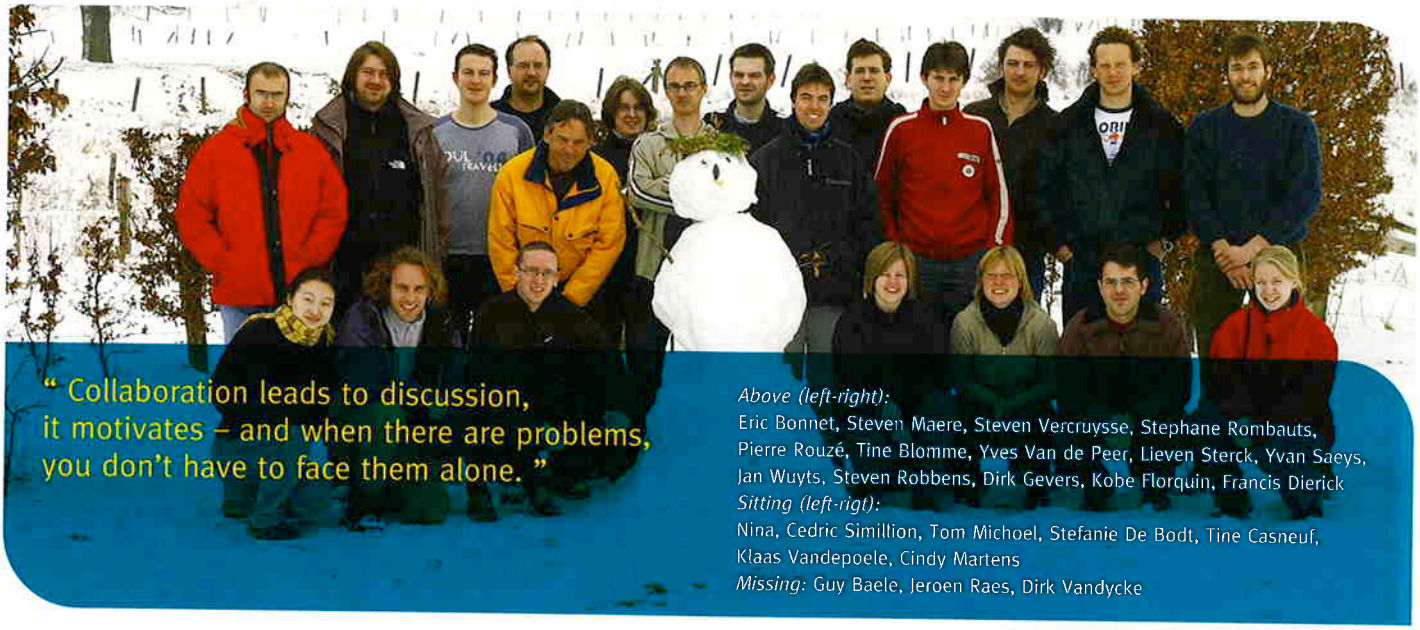
Our mathematical models indicate that a whole lot of genes remain in the genome when they are duplicated in a large-scale event. So, there is really a correlation between these large genome duplications and the retention of genes in the genome. At some moments, such large gene duplication events – which is, of course, a random and accidental process – can give that particular species a selective advantage, through which it is better able to survive. Then, with the extra genetic material, anything can happen. And when we assign a date to these duplications, we can deduce from fossils that they correlate with the emergence of a whole series of new leaf forms and flower structures.

Of course, evolution is often controversial; we're trying as best we can to reconstruct what has happened, and then attempting to correlate this with what there is to find in the 'fossil record', for example. An interesting challenge, with scores of gaps. But the funny thing is that, with bioinformatics, we can verify what has occurred roughly 600 million years ago. That's not possible with experimental work, and I find that this makes our research very fascinating.

Now take human beings. How is it that we've been able to come into existence? In all probability, two genome duplications took place very early in the evolution of the vertebrates. If this would not have occurred then, there is a good chance that vertebrates never would have existed. This is very interesting! The origin of the human being is ultimately nothing more than coincidence…'

Your primary tool is not the lab but a computer?

'And genomes, naturally. The more data we have, the more reliable the conclusions are: it's a matter of statistics. We also write a lot of programs ourselves – and with relative success, because our competitors use them, too. A major principle in bioinformatics is to verify to what extent observations you've made could have also occurred randomly. We re-shuffle the entire genome and see whether duplicated areas can also be found by chance. But these are only small fragments in each



" Collaboration leads to discussion, it motivates – and when there are problems, you don't have to face them alone. "

Above (left-right):
Eric Bonnet, Steven Maere, Steven Vercruysse, Stephane Rombauts, Pierre Rouzé, Tine Blomme, Yves Van de Peer, Lieven Sterck, Yvan Saeys, Jan Wuyts, Steven Robbens, Dirk Gevers, Kobe Florquin, Francis Dierick
Sitting (left-rigt):
Nina, Cedric Simillion, Tom Michoel, Stefanie De Bodt, Tine Casneuf, Klaas Vandepoele, Cindy Martens
Missing: Guy Baele, Jeroen Raes, Dirk Vandycke

case, whereas in reality very large parts of the genome are often found in duplo, which then indicates a large-scale event, like a genome duplication. Inexplicable by chance, and thus biologically significant. Everything stands or falls with statistics: is there a significant difference with what you would expect through chance?'

How does bioinformatics compare with experimental research?
'Bioinformatics is very competitive. The genomes become available – and all the bioinformatics groups have their tools ready. So, you have to be faster or be able to extract more data if you want to be able to publish well. With bioinformatics, you can allow yourself to try out a topic and then decide whether it will yield something or not. We can change direction more quickly. That's frequently the frustration in a wet lab: several years can go by before you know whether it was worth the effort or not. With bioinformatics, you have more flexibility in directing the research. But this automatically brings diversification along with it, and that's something you must constantly be on the lookout for.

This all becomes clear quickly when you program software. You will certainly get bogged down, but you always manage in the end – because you yourself are the only limiting factor. In a lab, it's quite often otherwise. One less frustration in our job!'

Do you collaborate with your colleagues in the wet labs?
'Yes – more and more. It's very interesting to go into the lab to investigate the genes and motifs that we theoretically indicate as functional. What happens when such a motif is eliminated? As a matter of fact, such validation is often required by the scientific journals.

But the collaboration also works in the other direction. Bioinformatics can be very supportive for research, and so the expertise in our group is very useful. We try to find a good mix between service for the wet lab and our own research.

What are the important decisions that you still have to make as Group Leader?
'There's still so much to discover in bioinformatics, but I'm aware that too many topics can be disadvantageous. In that situation, it's impossible to follow all the literature, and so you run the risk of not being up-to-date, and you lose the 'cutting edge'. So, *focus* – which Dirk Inzé is always hammering home – is a key concept. It comes down to finding the right balance between enough divergence and still being goal-oriented in your work. This is perhaps every Group Leader's dilemma. My group is always growing, and a number of doctoral students want to start here. A big group is advantageous: there is a lot of complementary expertise. I'm all for working with 3 to 4 people on the same topic. Collaboration leads to discussion, it motivates – and when there are problems, you don't have to face them alone. But with a lot of young people, a good overview is still essential. That's why I would like to bring some more experienced post-docs into the group, to give the group some more maturity and experience. Preferably people from other countries, too: international input is enormously important for good science.

But it's still vitally important to make good choices. Sometimes there are opportunities that can quickly lead to good publications, and you should not miss them. Like our research on micro-RNAs, which led to a very nice publication in *PNAS*. Are we going further with this or not? It's a very competitive field, but then again we now have the poplar genome at our disposal, and this could well provide new information. So it would probably be regrettable not to follow this up further.'