Gene Silencing II

Gene regulation

Peter Waterhouse

I n my first chapter, I described how the existence of a virus protection mechanism in plants was discovered, how it operates, and how the mechanism has been exploited in biotechnology applications. In this second chapter, I continue the story about this RNA degradation mechanism - of which has some even more unexpected twists to come.

Plant and Animal Dicer Bioinformatics

The term "bioinformatics" has slightly different meanings to different people, but to me it means the computer-mediated analysis of nucleotide and amino acid sequences in genes and genomes, and this is becoming an increasingly important tool in biological research. One aspect of bioinformatics is to look at the amino acid (aa) sequences of proteins that are known to have related biological functions and to see if there are stretches of aa sequences that are conserved within them. These regions are called domains. One can then look at the aa sequences of biologically uncharacterised

proteins to see if they contain any known domains, and if they do, to use them to make predictions about the possible biological functions of the proteins under scrutiny. This is exactly the approach we took to search for plant and animal Dicer genes. Some elegant biochemical work by researchers in the USA, using a purified protein from *Drosophila* identified it to be responsible for cutting up dsRNA into ~21nt dsRNA fragments and to be the Dicer protein mediating RNAi. When the sequence of this protein was examined, it was found to contain a number of domains (Figure 1A). These were: two different helicase domains, two RNAseIII domains, one dsRNA binding domain, one PAZ domain and one DUF283 domain. The presence of the helicase, RNAseIII and dsRNA binding domains was totally understandable: a helicase is an enzyme that unwinds dsRNA or dsDNA, RNAseIII is an enzyme that cuts RNA, and a dsRNA binding domain (not surprisingly) is found in proteins that bind to dsRNA. These are properties quite likely to be possessed by an enzyme that is going to bind, unwind (to allow access), and cut dsRNA. The PAZ domain gets its name from a stretch of aa sequence found to be shared by three proteins called : Piwi, Argonaute and Zwille. DUF283 is the wonderfully named domain of unknown function number 283. Armed with these domains we searched plant, fungal and other animal genomes for genes that encoded single proteins containing all of these domains. We found (Figure 1B) only one such Dicer gene in each mammalian genome, two such genes in insects, but found plants to have taken it to the extreme and have at least four of these genes (in *Arabidopsis*), and up to six (in rice) .

Why does a plant need four Dicers?

Mice and men can survive quite happily with one Dicer, so why have plants got so many?

Figure 1A: Domains found in an Animal Dicer Protein
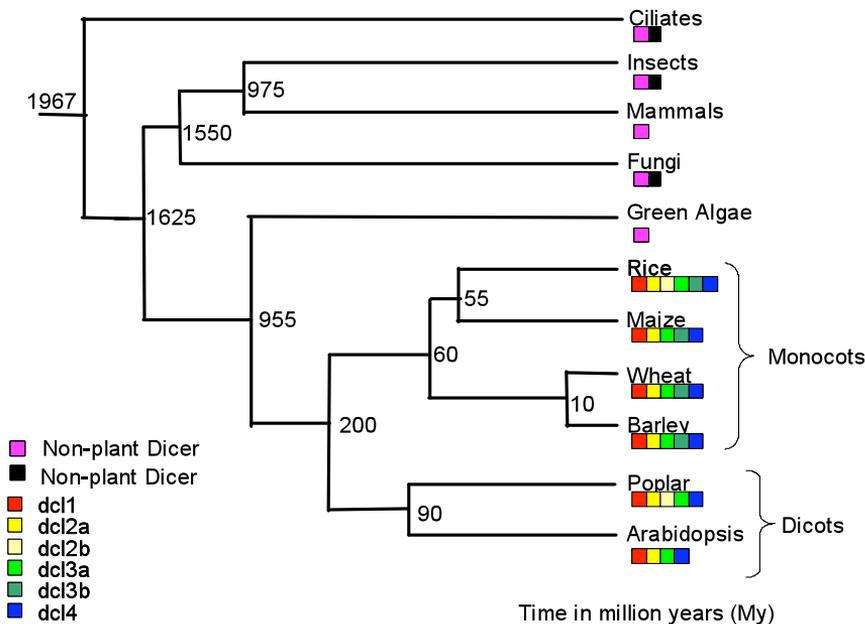

Figure 1B: Evolution of Dicer Genes

When we compared the sequences of the Dicer-like (DCL) genes in plants we found that although rice has six, it has four different types (two are duplicated), which match the four different genes in *Arabidopsis*. So, if we call the *Arabidopsis* DCL genes 1, 2, 3 and 4, then rice contains one DCL1, two DCL2s, two DCL3s and one DCL4. To see what functions these genes might have, we obtained *Arabidopsis* plants which are singly mutant for each of these genes. We were expecting these mutant plants to look fairly normal but to be more susceptible to virus infection. We also thought that having four DCLs may simply be a way of

boosting a plant's defence capacity. Mammals have an immune system to combat viruses and may therefore not need this increased multiple Dicer-mediated capacity. Uninfected *Arabidopsis* plants that are mutant for DCL2, DCL3 or DCL4 all look much like wild-type plants (Figure 2), but plants that are mutant for DCL1 are very peculiar – dwarfed and twisted – and I will come back to what DCL1 is doing later in the chapter. So we challenged the DCL2, DCL3 and DCL4 mutant plants with a virus and looked to see if they developed extreme virus symptoms and whether the virus RNA genome was being chopped up into ~21nt fragments. We, also, stacked the mutations so that we had double and triple mutants. The triple DCL2/DCL3/DCL4 mutant plant looked very similar to wild type - when not infected by a virus (Figure 3), but when infected had much higher levels of virus and more severe symptoms than the wild-type plants, or indeed the single or double mutants. This confirmed our hypothesis that these three Dicers were providing defence against viruses. The double mutants also gave us a nice insight into the processing of the dsRNAs. The DCL2/DCL3 mutant produced virus dsRNA fragments of 21nts, the DCL3/DCL4 mutant produced 22nt virus fragments, and the DCL2/DCL4 mutant gave 24nt virus fragments. This shows that each of the three DCLs is chopping up the virus dsRNA and that the DCL4 enzyme produces 21nt fragments, the DCL2 enzyme produces 22nt fragments and DCL3 produces 24nt fragments (Figure 4). When we transformed plants with hpRNA transgenes, their RNAs were also processed by DCL2, DCL3 and DCL4 into 22nt, 24nt
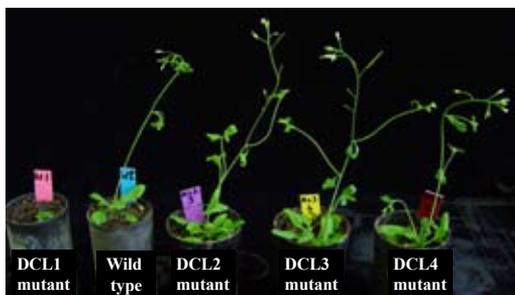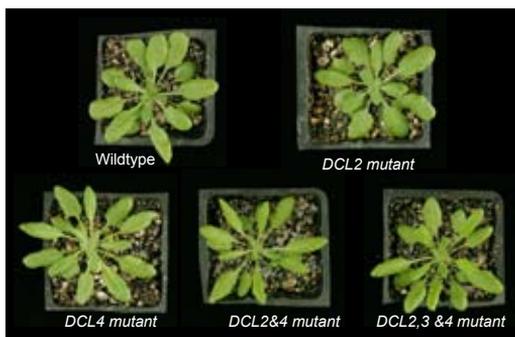

Figure 2: Dicer mutants in Arabidopsis

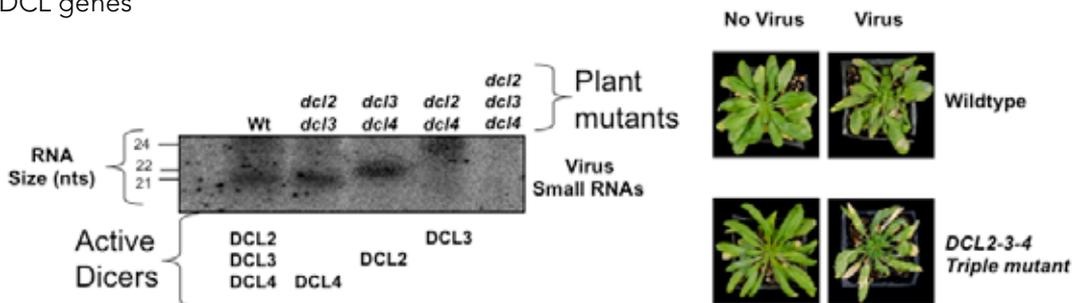**Figure 3:** Arabidopsis plants with mutant DCL genes

Figure 4: DCL2, DCL3 and DCL4 in viral defence

and 21nt fragments, showing that RNAi, as proposed in the first chapter, is operating by the virus defence mechanism. Interestingly, we could detect almost no small dsRNA-derived fragments in the DCL2/DCL3/DCL4 triple mutant, suggesting that DCL1 is doing something other than virus protection.

Virus-encoded silencing suppressors

If plants have a virus defence mechanism that destroys the replicating and single stranded forms of their RNA genome, as suggested above, how is it that virtually every species of plant is known to be susceptible to infection by at least one – and commonly by many – different viruses? I previously described plant viruses as having a basic set of three genes encoding a replicase, a movement protein, and a coat protein. However, I did not describe the function of one other type of viral gene. When we were first determining the sequences of virus genomes, it was relatively easy to identify the replicase, movement and coat protein genes but there was often a gene in the viral genome with a function that was difficult to predict. It turns out that these genes are silencing suppressor genes (Figure 5). The protein from such a gene suppresses the plant's RNAi mechanism and by doing so protects the virus from being destroyed. This is a classic example of the war that is waged between a pathogen and its host. The virus infects the plant, the plant responds by evolving a defence mechanism, the virus

counters this by evolving a way of inactivating the plant's defence mechanism. Different types of plant viruses have different silencing suppressor genes. Some bind to the small RNAs made by the DCLs preventing them from being loaded into Argonautes, others inactivate DCLs - preventing them from dicing the viral dsRNAs. But my favourite group of viruses, the luteoviruses, have an even more elegant way of inactivating the RNAi mechanism. Plants have an intrinsic system that degrades their own proteins ensuring that each protein has a specific working life. This operates by a system called the proteosome which degrades proteins when they have been labelled with a tag, and this tag is placed onto proteins by a mechanism that is guided by "F-box" proteins. The first protein that a luteovirus produces when it infects a plant cell is a silencing suppressor protein that mimics a plant F-box protein. This protein guides the tagging of Argonaute proteins. So, by making one small protein the virus guides the plant's protein turn-over system against a key component of the plant's viral defence system!

What is DCL1 doing?

Returning to the mutant DCL experiment, the plants that were mutant for DCL1 showed a really weird phenotype, even when they were not infected by a virus. Also, we found in the virus-infected DCL2/DCL3/DCL4 triple mutant that the virus was not diced into small RNAs. Taken all together this suggested DCL1 was doing something other than virus defence. From the domains contained in the DCL1 protein, one would predict that it would have a similar enzymatic activity to DCL2, DCL3 and DCL4 - and it does. We thought we had been really clever when we designed single stranded hairpin (hp) RNAs as a way of guiding RNAi against our target genes. In fact, nature was hundreds of millions of years ahead of us. It turns out that almost all multi-cellular organisms (plant and animal) produce their own hpRNAs that are processed by DCL1 to produce 21nt RNAs, which are used by Argonaute to guide the regulation of endogenous messenger RNAs, and that DCL1 is essential for normal development. For normal development to
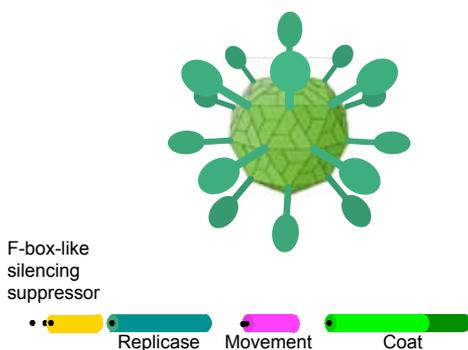

F-box-like silencing suppressor

Replicase   Movement   Coat

Figure 5: Luteovirus particle and its RNA genome

occur certain genes must be switched off, especially at the time of developmental transitions, such as developing from a juvenile to an adult form of a nematode or insect, or from vegetative to floral growth in plants. This switching off is performed by the small RNAs produced by DCL1 from special hpRNAs produced by the plant at the right time and in the right tissue. DCL1 differs from DCLs 2,3 and 4 in that it specifically cuts out only one 21nt small RNA from its hpRNA precursor (Figure 6), and this is called a micro (mi)RNA. In Arabidopsis there are at least 187 different miRNAs and there are many more than this in animals (see Table 1).

If DCL1 is not produced, the miRNAs cannot be cut out of their precursor hpRNA molecules and cannot be used by Argonaute to regulate the target genes. This explains why our DCL1 mutant plant has such a weird phenotype – it is a plant with incorrect developmental regulation. Also, if we look at plants mutant for Argonaute, they too look bizarre (Figure 7). This is not surprising because the regulation system cannot operate without both of these proteins. In fact, both the DCL1 mutant and Argonaute mutant shown in the photographs are not completely inactive mutants but rather a mutation that produces a truncated version of

Table 1

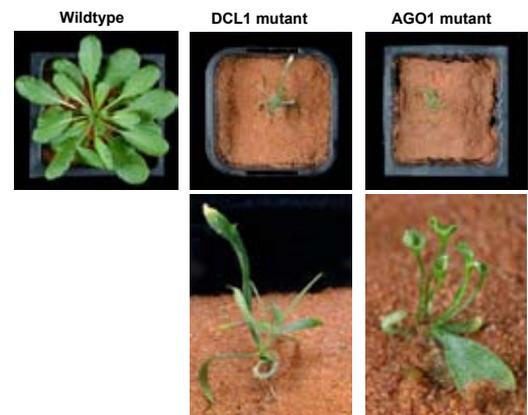| Species	Number of different microRNAs
Arabidopsis	187
Nematode	156
Drosophila	152
Mouse	547
Human	706

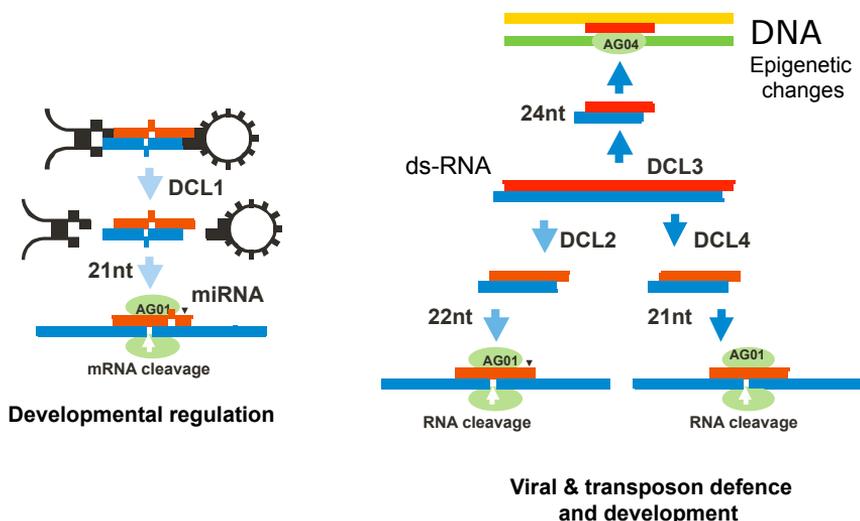**Figure 7:** DCL1 and Argonaute (AGO1) mutant plants

Figure 6: DCL-mediated production of small RNAs and actions of guided Argonautes

the DCL1 protein (missing the last 1/20th piece of the protein) and a mutation that has only one amino acid different from the wild-type Argonaute protein. Indeed, a totally inactive DCL1 or Argonaute gene in multi-cellular organisms is almost always lethal.

Future Research and Technology

In an earlier section, I described how DCL3 chops up viral dsRNA into 24nt small RNAs as part of an RNA virus defence pathway. It has recently been found that DCL3's major role is not so much in the control of RNA viruses, but more in the control of DNA viruses. DCL3 still cuts up dsRNA that is somehow produced from the DNA virus, but the 24nt RNAs that are produced are not loaded into the Argonaute that I have been discussing (AGO1), but rather into a related protein called Argonaute 4. This protein does not cleave single stranded RNA like its cousin, but rather directs the compression of dsDNA with sequences that are complementary to the 24nt RNA fragments. This prevents the genes in the viral dsDNA from being copied into messenger RNA. Significantly, this also happens for the protection of plants from transposons – which are like retroviruses – which are in very large numbers in a normal plant genome and if not repressed by DCL3 would move around causing all kinds of mutations. It is also just beginning to emerge that this form of RNA-directed DNA compression is a very important process, not only for plants in their control of viruses and transposons, but also in both plants and animals in terms of normal gene regulation – for example – in the inactivation of one of the X chromosomes (an essential process) of human females. This control of gene expression by DNA compression is called epigenetics and this is a research field that, I think, will be at the forefront of medical and molecular plant and animal research over the next decade.

Another area which will probably become a major new technology is the use of artificial miRNAs. We understand how to make transgenes that encode hpRNA templates which DCL1 will process into miRNAs. Therefore, we can use them to silence genes in much the same way as RNAi but with more precision.

Indeed, plants and animals altered by artificial miRNAs are now beginning to be produced.

Further Reading

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001, 409:295-6.

Ding SW, Voinnet O. Antiviral immunity directed by small RNAs. Cell. 2007, 130: 3-26.

Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM. The evolution and diversification of Dicers in plants. FEBS Lett. 2006, 58: 2442-50.

Mallory AC, Vaucheret H. Functions of micro-RNAs and related small RNAs in plants. Nat Genet. 2006, 38:850.

Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, Ellacott GA, Watson JM, Wang MB, Brosnan C, Carroll BJ, Waterhouse PM. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. EMBO Rep. 2006, 11: 168-75.

Pfeffer, P. Dunoyer, F. Heim, K. E. Richards, G. Jonard, and V. Ziegler-Graff . P0 of Beet Western Yellows Virus Is a Suppressor of Posttranscriptional Gene Silencing Journal of Virology, 2002, 76: 6815-6824.