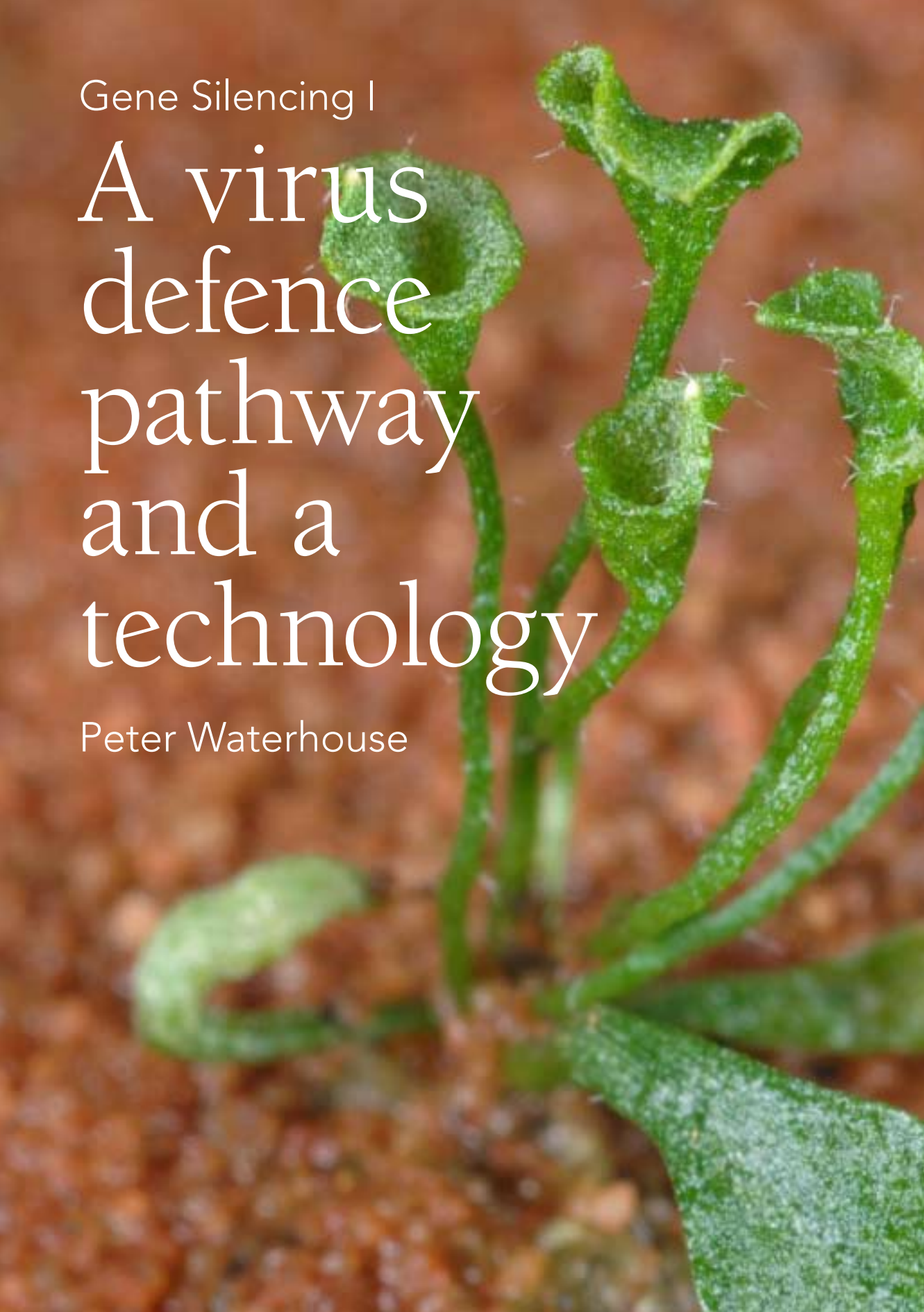
Gene Silencing I

# A virus defence pathway and a technology

Peter Waterhouse



The development and use of vaccines against viruses such as polio, smallpox, and measles have to be among the great accomplishments of medical science. The history of how it all started from Edward Jenner's discovery (that milkmaids and dairymen infected with the mild cowpox virus were protected against smallpox) is widely known. However, it is not so generally appreciated that plants can also be protected from a severe virus by prior infection with a mild strain of a closely related virus. This "cross protection" in plants was recognized as early as the 1920s, but plants do not possess an antibody-based immune system, so the mechanism underlying this defence remained a mystery for many decades. A few years prior to the turn of the millennium, our understanding began to dawn and after a flurry of research the existence of an unsuspected, but immensely powerful, mechanism, in both plants and animals, has been revealed. This mechanism can be triggered and directed to not only provide protection against viruses but also to silence any gene, and has led to a technology

called RNA interference (RNAi) which is being used for applications ranging from improved agricultural traits to fighting cancer.

In this first of my two chapters, I will describe how the gene silencing pathway was discovered, and how it works, and then give some examples of how it has been exploited. In the second chapter, I will describe how this mechanism turned out to be a sophisticated multidimensional pathway which not only protects cells against viruses but also tightly controls the regulation of genes required for normal development in almost all forms of multi-cellular life.

## Genes and Transgenes

Genes are encoded in the nucleotide sequences of double stranded (ds) DNA molecules, which are folded up to form chromosomes in the nucleus of eukaryotic cells. Each gene is made up of three adjacent sections: the "promoter", the "coding region", and the " terminator", (Figure 1). The promoter sequence defines where an enzyme (called a polymerase) binds to the DNA and starts to copy the sequence of one strand of the DNA into molecules of single stranded (ss) RNA. This copying proceeds across the coding region and stops when it reaches the terminator sequence. The RNA production is in the nucleus but once made, each RNA molecule (called messenger RNA) is transported to the main compartment of the cell (the cytoplasm) where it is used as the template for another enzyme complex (the ribosome) to decipher the sequence for the production of a protein of the inferred amino-acid sequence. This DNA to RNA to Protein is called the central dogma. As I hope will become clear, it is also important to note that the DNA is in a double stranded form (with the two strands binding together like a zipped-up zip), but the messenger RNA is single stranded (like one side of an unzipped zip). And the zipping rule of dsDNA is very simple, for two strands to bind together they must have "complementary" sequences. Each nucleotide can be one of four types called (in abbreviated form) A, C, G or T, and A will only bind to T and C will only bind to G. So, for example, if a short strand of DNA has the nucleotide sequence =>ACGTAT it will only zip up nicely with a strand having the sequence TGCATA<= (the arrows are to show that strands have polarity and when two strands zip up they actually point in opposite directions).

The other take-home message I want to convey is that our understanding of how a gene is composed of three sections allows us to make transgenes. As you will see in the next section, we make a transgene that is inserted into a plant so that it makes the coat protein of a virus but not the whole virus. We do this simply by taking a dsDNA copy of a plant gene, replacing the coding region of the plant gene with the coding region of the virus coat protein gene, and then inserting this [plant promoter- virus coat protein coding region – plant terminator] piece of dsDNA into a chromosome of a plant (Figure 2).

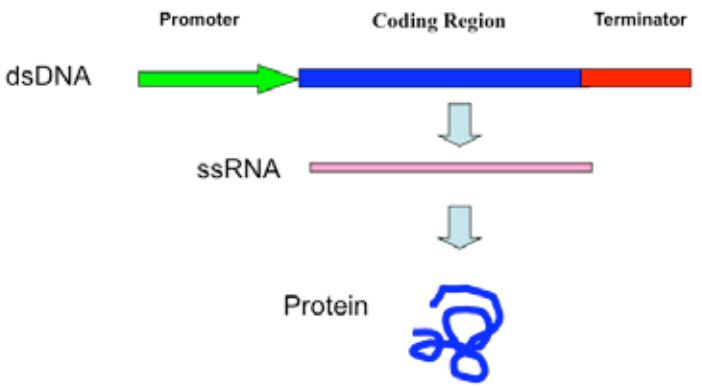



**Figure 1:** Layout of a gene and the central dogma

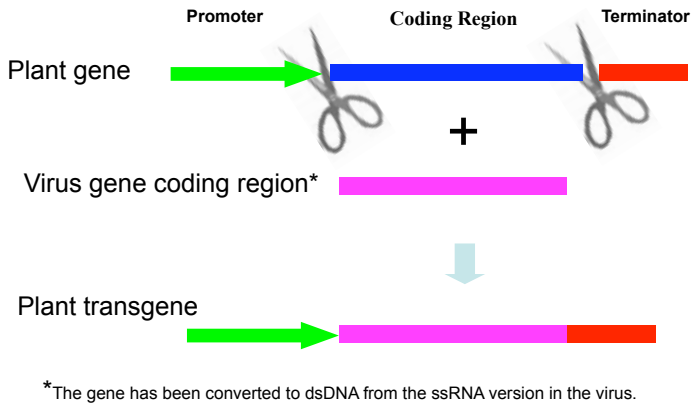



**Figure 2:** Making a plant transgene

## Virus Protection.

Viruses can cause serious losses to almost all of our food crops and plant breeders have spent considerable effort to find and incorporate natural virus-resistance genes into them. Nevertheless, for many virus/crop combinations there are no known natural resistance genes or there are single resistance genes which are under threat of being overcome by evolving virus strains. However, in the early 1980's researchers started determining the nucleotide sequences of plant viruses and this led the way to the production of pathogen-derived resistance genes. Most plant viruses have genomes made of single stranded RNA which encode at least three genes: a replicase gene – to replicate the genomic RNA, a movement protein gene – to help the virus genome spread from cell to cell, and a coat protein gene to wrap up the genome into a protective particle for movement, often by an insect vector, from plant to plant. (Figure 3). In 1986, a team of pioneering virologists showed that when a transgene made from the coat protein gene of tobacco mosaic virus (TMV) was inserted into the chromosomes of a plant, the "transformed" plant became TMV resistant. This stimulated plant virologists, including my group at CSIRO, to make transgenes from many different viruses and put them into lots of different crop plants – and

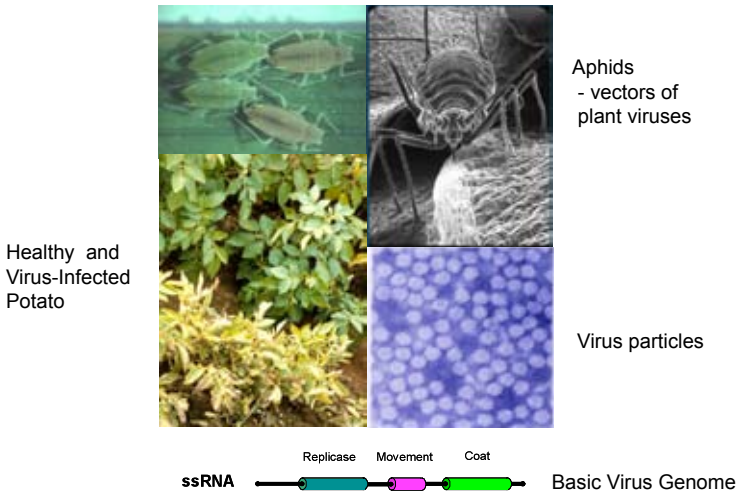



**Figure 3:** Plant viruses

with some success. The curious thing was that, with the exception of TMV, only a small proportion of plants containing these transgenes had virus resistance and the ones that produced the highest levels of transgene protein tended to be the ones that had no protection. Also, completely counter to expectation, the plants in which the transgene seemed to be producing little or no protein were the ones with the virus resistance. So what was going on?

## It is not the protein that does it!

When a transgene was made by placing the coding region for the virus gene in a back-to-front orientation between the promoter and the terminator and then transformed into plants, surprisingly, some of them were resistant to the virus. This transgene could not be producing the virus protein. It would be like trying to read a sentence of English from right to left - we call this an antisense gene. This showed us that it was not the protein, itself, but something else about the transgene that was conferring the resistance. When we looked at the plants that showed resistance and those that didn't, the striking feature was that those with resistance had multiple adjacent copies of the transgene (and there were always two adjacent copies in opposite orientations); those with no protection had only one or a few copies (Figure 4A). One possible explanation was that the terminators were not working with 100% efficiency in these adjacent and opposite transgene copies, so the promoter of one sometimes directed the synthesis of RNA molecules that continued into the coding region of the adjacent gene, and vice versa. In this situation the plant would be making two opposite strands of RNA that could bind together in a similar way to dsDNA. (Figure 4B).

## The double stranded RNA experiment

From our earlier experiments, we had plants with a single copy of a virus transgene (let's call this a sense transgene) and plants with a single copy of the antisense version of the transgene. None of these plants showed any resistance to the virus. We postulated that if the formation of virus-derived dsRNA from the transgenes was the key to generating virus resistance we could test this by a simple crossing experiment. We took pollen from a "sense" transgene plant and crossed it onto emasculated flowers of an "antisense" transgene plant, collected the seed that was set, germinated them, then inoculated the seedlings with the virus - and got a beautiful result. One quarter of the plants showed resistance to the virus and three quarters were susceptible. And when we analysed the genetic make-up of the plants, the quarter with resistance had inherited both the sense and antisense transgenes, whereas the susceptible plants had inherited either the sense transgene,

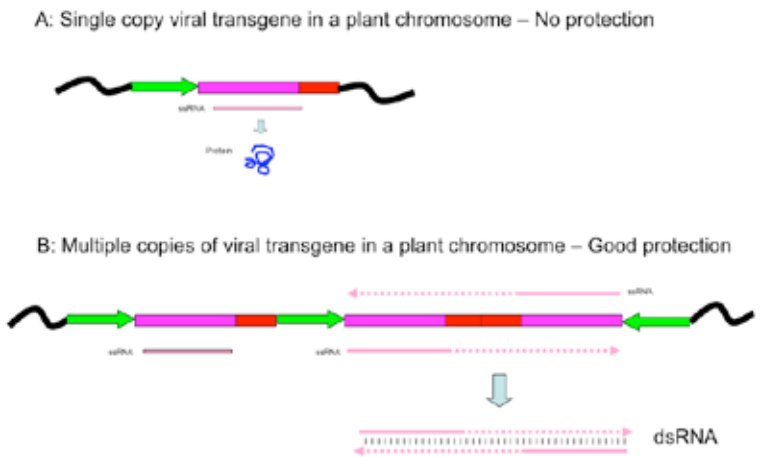



**Figure 4:** Integration of virus transgene

or the antisense transgene, or neither (Figure 5). This convinced us that dsRNA was the trigger that was somehow protecting the plants against the virus.

## How does it work? part 1

As mentioned in the first section of this chapter, a healthy cell contains lots of dsDNA and ssRNA, but it does not contain dsRNA. The only time a cell contains dsRNA is when it is infected with a replicating RNA virus. Therefore, the model we proposed was that a virus defence pathway already exists in plants and is triggered by dsRNA; it operates by using the sequence of the dsRNA to direct enzymes to destroy ssRNA molecules of the same or complimentary sequence (Figure 6). So, what we had been doing by using transgenes to express a piece of the virus genome as dsRNA was to forewarn the cell of the virus's sequence so that it was primed and ready to destroy it even before infection. This has obvious parallels with the vaccination strategy we use to protect ourselves from viruses such as polio and measles.

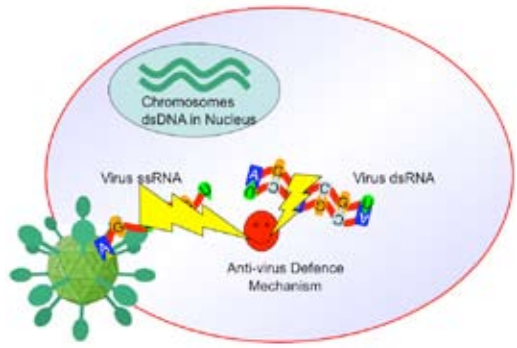



**Figure 6:** The intrinsic virus defence mechanism in plants

## Can we use this pathway to silence genes?

There are lots of interesting and useful things that can be done if one can silence specific genes in a plant or animal, as we will see later. Therefore, we wondered if we could use this viral defence to specifically silence some of a plant's own genes. On the basis of our dsRNA induction model, all we need to do to silence a specific gene is make the cell perceive that the messenger RNA of that gene is from a virus, and we can do this by expressing transgenic

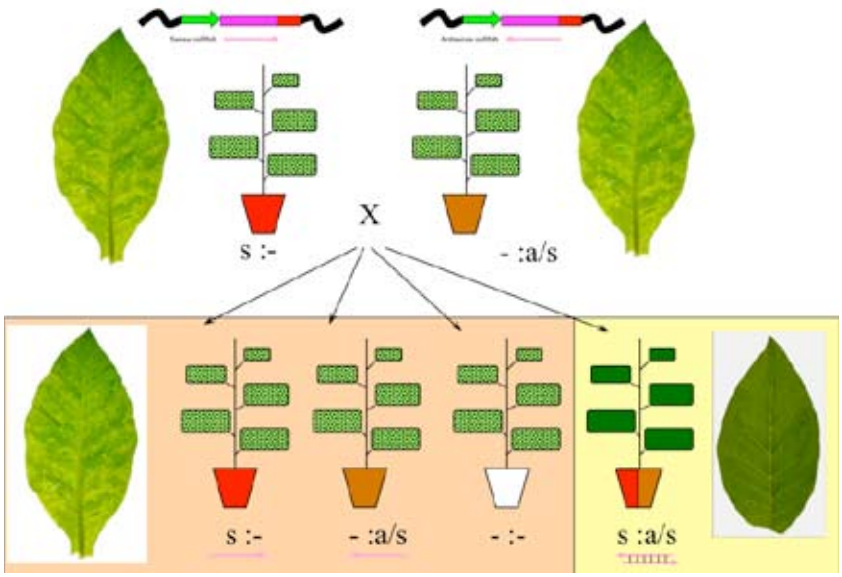



**Figure 5:** Sense x antisense experiment

dsRNA containing the same sequence as the target messenger RNA. The defence mechanism will then be directed to destroy the messenger RNA before it can be deciphered into the protein it encodes.

## Hairpin RNAs

Making a plant with both a sense and an antisense transgene, either by the crossing strategy already described or by transforming a plant simultaneously with two transgenes, is hard work. Also, when two complementary RNA molecules are made from two different genes the two molecules have to find each other in the cell before they can zip up to form dsRNA. We thought of a solution: make a transgene that makes an RNA which looks like a hairpin. This is a single strand of RNA, coded for by one transgene, but the last section of the molecule is complementary to the first section so it folds back and zips up into a hairpin shape (Figure 7). One of our first tests, in whole plants, was to try to silence the chalcone synthase gene in our favourite model plant - Arabidopsis. This plant produces dark brown seeds and the enzyme that makes this brown pigment is chalcone synthase. So, we took a section of the coding region of the chalcone synthase gene, made it into a hairpin (hp) RNA transgene and transformed Arabidopsis with it.

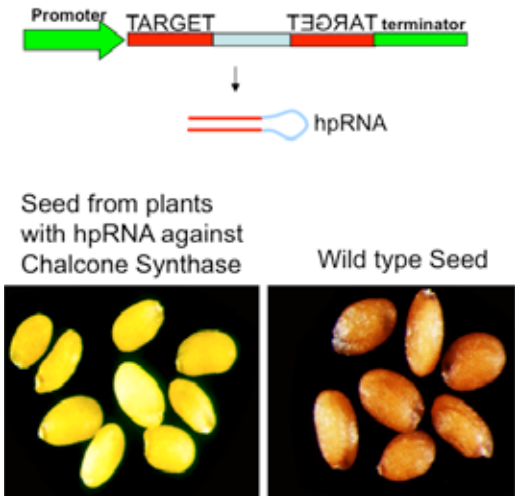



**Figure 7:** One of the first hpRNA transgenes silencing a plant gene

This gave us an unequivocal result, the plants produced bright yellow seed (Figure 7). When we got this result, we knew we were really on to something. Technology with widespread applications – and this technology of introducing a dsRNA or hpRNA into plant or animal cells for the silencing of genes has become known as RNA interference or RNAi.

## How does it work? part 2

While we were finding out that dsRNA can direct silencing in plants, Andy Fire and Craig Mello, in the United States, were making similar discoveries in nematodes. In fact, they were to win the Nobel Prize in 2007 for their discovery. And it turns out that this pathway exists in almost all eukaryotic multicellular organisms ranging from mosses to mammals. By looking at how the gene silencing pathway could be disrupted by different mutants in plants, insects and nematodes has led to a deep understanding of the enzymes and processes involved, with perhaps the best understanding coming from work using *Drosophila* (the fruit fly).

All of these eukaryotes share two key proteins (Dicer and Argonaute - some call this latter one, Slicer) which, with a number of accessory proteins, make the silencing process work (Figure 8). Dicer recognises dsRNA and cuts it up into fragments, about 21 nucleotides long, and transfers each fragment to an Argonaute molecule. The Argonaute protein cuts and discards one of the strands but retains the other to use as a guide. Using the retained strand, the Argonaute examines all of the ssRNA molecules in the cell and if an RNA is found that has a


**Figure 8:** The RNA interference mechanism

stretch of 21 nucleotides that is exactly complementary to the guide sequence, the Argonaute acts like a pair of scissors and cuts the "target" RNA in the middle of the recognised sequence. Because the dsRNA being introduced into the cell is usually several hundred nucleotides long, it is Diced up into many different 21nt fragments, and each one is loaded into a different Argonaute molecule. This means that a single target ssRNA molecule may be cleaved in several different places by the loaded Argonaute population. Chopping the target RNA into pieces prevents it from being translated into protein, thus silencing the gene.

## A few examples

RNAi has been used or has shown potential for many purposes including human therapeutics (Table 1) and functional genomics. The entire nucleotide sequences of the genomes of a number of species (including human, fruit fly, nematode, and *Arabidopsis*) have been determined. From these sequences we can predict the coding regions of all of the genes within each genome. However for a significant proportion of these genes, we have little or no idea about their functions. So, large scale projects are currently underway using RNAi to silence each of these genes, one by one, in nematodes and in *Arabidopsis* so that the changes in form or behaviour that result from silencing the genes can give clues about their roles.

**Table 1**

| Therapeutic potential of RNAi in Humans | |
|---|---|
| Neurological disorder | Obesity |
| Various forms of Cancer | High blood Cholesterol |
| Spinal muscular atrophy | Depression |
| Growth hormone deficiency | HIV |
| Diabetes | Hepatitis B |
| Malaria | Hepatitis C |

There have been some amazing applications of RNAi in plants, such as coffee plants that produce decaffeinated beans and opium poppies that produce desirable pharmaceutical compounds, and I would like to finish this chapter with three further examples that hopefully give an idea of how useful and versatile the technology can be.

One of the important aspects of crop production is flowering time. For instance, if a cereal crop flowers too early, it may have not yet made sufficient energy stores to fuel its maximum grain production. Similarly, if it flowers too late in the season there may be insufficient time to produce a good yield. So, being able to control flowering time in plants could be a very useful tool in horticulture and agriculture. In *Arabidopsis*, there is a gene called FLC which represses flowering and we have used RNAi to switch it off and bring on flowering (Figure 9). This clearly shows that the technology has the potential to regulate flowering time in crops.

An application of RNAi in plants that is much closer to agricultural use is the silencing of genes involved with seed-oil production. Some seed-oils are much better for human health than others, and some oils are more stable at high temperatures than others. It all depends on the fatty acid composition of the oil. For example palm oil is very high in palmitic acid which makes it stable at high temperatures but also unhealthy for human consumption, as it raises LDL cholesterol levels. Olive oil, on the

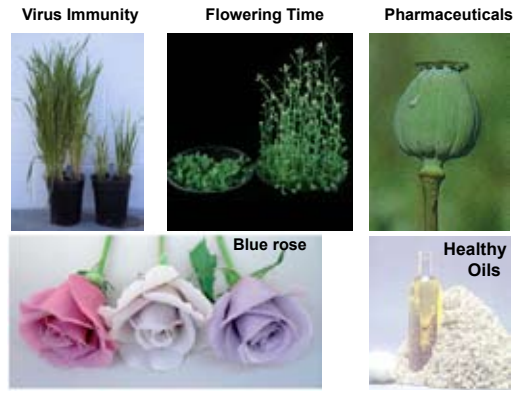



**Figure 9:** Custom-made changes using RNAi

other hand, is high in linoleic acid which is much healthier for human consumption, but it is not stable at high temperatures and therefore not good for frying. Almost everyone knows about growing cotton plants for their fibre, but it is less well known that the seeds contain high levels of oil. Unfortunately, the oil composition is similar to palm oil. The best oil for heat stability and with no negative effects on cholesterol levels is one which is high in oleic acid. We have used RNAi to silence the gene in cotton which codes for the enzyme that converts oleic acid into a different fatty acid. This has altered the seed-oil from being around 10% to an impressive 75% oleic acid. If these plants were used in agriculture it would produce two crops, fibre and seed-oil, for the price of one. The last example is very close to being a commercial reality. It will not feed our stomachs but it may soothe our souls. A biotech company with its origins in Melbourne has produced a blue rose – a colour that has not been achieved during centuries of rose breeding. The first step was to introduce a transgene into roses that produced the blue pigment from a different plant species. Unfortunately, the rose kept on making a red pigment, resulting in a purple flower. However, by adding an RNAi transgene to silence the red pigment gene, the world now has a beautiful blue rose.

## Further Reading


Abel PP, Nelson RS, De B, Hoffmann N, Rogers SG, Fraley RT, Beachy RN.Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene.Science. 1986; 232:738-43.

Hamilton AJ, Baulcombe DC.A species of small antisense RNA in posttranscriptional gene silencing in plants.Science. 1999; 286: 950-2.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC.Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.Nature. 1998; 391: 806-11.

Waterhouse PM, Graham MW, Wang MB.Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA.Proc Natl Acad Sci U S A. 1998; 95: 13959-64.

Zamore PD, Tuschl T, Sharp PA, Bartel DP.RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals.Cell. 2000; 101: 25-33.

Wesley SV, Helliwell CA, Smith NA, Wang MB, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG, Waterhouse PM. Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J. 2001, 27: 581-90.

Liu Q, Singh SP, Green AG. High-stearic and High-oleic cottonseed oils produced by hairpin RNA-mediated post-transcriptional gene silencing.Plant Physiol. 2002; 129: 1732-43.

Allen RS, Millgate AG, Chitty JA, Thisleton J, Miller JA, Fist AJ, Gerlach WL, Larkin PJ. RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. Nature Biotechnol. 2004; 22: 1559-66

Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, Brugliera F, Holton TA, Karan M, Nakamura N, Yonekura-Sakakibara K, Togami J, Pigeaire A, Tao GQ, Nehra NS, Lu CY, Dyson BK, Tsuda S, Ashikari T, Kusumi T, Mason JG, Tanaka Y. Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating delphinidin. Plant Cell Physiol. 2007; 48: 1589-600.