Identification of miRNAs and their targets

Introduction

As we discussed in the first lesson, miRNAs were only discovered two decades ago. One reason for their late discovery was certainly their small size. When the researchers around Victor Ambros tried to identify the transcript of the *lin-4* gene, the bands they detected of a size of 60 nucleotides (corresponding to the *lin-4* precursor miRNA) already seemed incredibly small for a "real" gene product. At first, they simply neglected the tiny band of 20 nucleotides appearing at the bottom of the gel as background degradation products of no biological relevance. However, they later realized that this tiny band of 20 nucleotides was actually the major *lin-4* gene product, produced by Dicer.

Due to their small size and their high degree of homology, identifying and quantifying miRNAs in the lab experimentally is challenging. Here, we will discuss the methods commonly used to identify and experimentally detect miRNAs themselves, but also the mRNA targets they regulate.

Quantifying miRNAs

After the identification of the miRNA *let-7*, genome database searches using the *let-7* sequence revealed perfect 22-nucleotide matches in the *Drosophila*- and human-genome sequence, confirming that there are *let-7* homologs in other organisms (this was not true for *lin-4*). Several labs started out to identify novel miRNAs by cloning and subsequent sequencing of endogenous small RNAs of 21-25 bp length in flies, worms and mammals. Here, small RNAs were isolated form organisms, these RNAs were transcribed into DNA (a technique called RT-PCR, see below) and sequenced. This led to the identification of several dozens of miRNAs that were deposited in the miRbase repository (http://www.mirbase.org), the main miRNA resource, which now includes thousands of miRNAs from all organisms. These discoveries represented the starting point for looking for more miRNA genes by intensive bioinformatics research. The miRNAS discovered so far served to define the rules that were applied to predict miRNA genes by computational approaches. These computer predictions are based on the hairpin-shaped secondary structures and the high evolutionary conservation of miRNAs from species to species.

Although bioinformatics plays an important part in identifying putative miRNAs based on homology searches, miRNAs also need to be experimentally verified in the lab before the sequence is finally designated to encode a miRNA. Traditional methods for detecting miRNAs include Northern blotting, RT-PCR, microarrays and sequencing, which we will each explain in more detail in the following.

Northern blotting

Northern blotting (NB) is a technique to study gene expression by the detection of RNA, e.g., miRNAs, in a sample isolated from cells, tissues, or whole organisms. Using this method, one can determine whether a certain RNA of interest is present in a sample, and at which level. That, of course, requires the researcher to know the sequence of the investigated RNA. Typically, RNA samples extracted from cells consist to a large extent of ribosomal RNAs and mRNAs, whereas miRNAs only make up a small part of the sample and have a much smaller size. miRNA can be better detected if the total RNA samples are separated by size using electrophoresis, usually on 15-18% polyacrylamide gels in which small RNAs (20-200 nt) RNAs can be separated from the larger one (although the small RNA fraction also contains very abundant transfer RNAs!). After electrophoresis, the RNA is transferred onto a membrane, a process called blotting (hence the name). This membrane is incubated with a hybridization probe, a nucleic acid that is complementary to the target sequence (the RNA to be detected). To detect the hybridization between the probe and the target RNA, the probe has to be labeled, e.g., with radioactive phosphorus (^{32}P) that is incorporated into the nucleic-acid probe. After

hybridization, an X-ray film is applied to the dried membrane to detect the radioactive signals in those samples where hybridization took place (see figure 2-1).

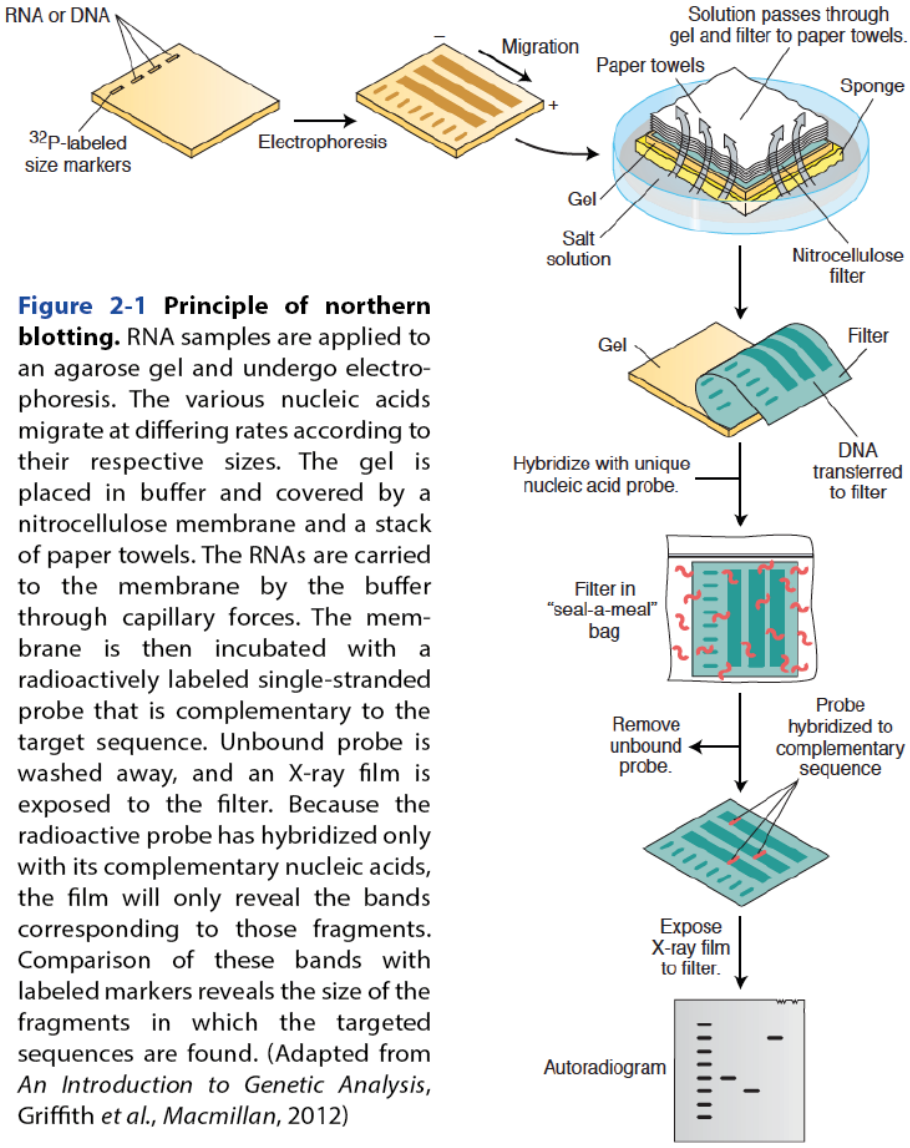

Figure 2-1 Principle of northern blotting. RNA samples are applied to an agarose gel and undergo electrophoresis. The various nucleic acids migrate at differing rates according to their respective sizes. The gel is placed in buffer and covered by a nitrocellulose membrane and a stack of paper towels. The RNAs are carried to the membrane by the buffer through capillary forces. The membrane is then incubated with a radioactively labeled single-stranded probe that is complementary to the target sequence. Unbound probe is washed away, and an X-ray film is exposed to the filter. Because the radioactive probe has hybridized only with its complementary nucleic acids, the film will only reveal the bands corresponding to those fragments. Comparison of these bands with labeled markers reveals the size of the fragments in which the targeted sequences are found. (Adapted from *An Introduction to Genetic Analysis*, Griffith *et al.*, *Macmillan*, 2012)

Northern blotting is the only technique that allows for the (semi-) quantitative visualization of miRNAs. Because gel electrophoresis separates RNA sizes, mature as well as pre-miRNAs can be analyzed. However, compared to other methods, e.g., qRT-PCR (described below), northern blotting suffers from low sensitivity (RNAs with low abundance cannot be detected), low throughput, and high input RNA requirements (typically in the order of 5?50 ?g total RNA per sample have to be loaded onto the gel). Radioisotopes (^{32}P) are still the most commonly used labeling system for northern-blot detection; however, they pose several safety concerns for the researchers using them and the environment where they are disposed. In addition to these safety concerns, the use of radioisotopes greatly increases the amount of time required to conduct the experiment, and in some cases, ^{32}P labels must be exposed for days to detect weak signals. The probes used for miRNA detection by northern blotting are usually miRNA-complementary DNA oligonucleotides. An oligonucleotide is a synthesized, short RNA or single-stranded DNA molecule used in many applications in molecular biology.

The biggest problem of northern blotting is the low sensitivity of this method, such that low abundant RNAs, such as some miRNAs, are difficult to detect. The sensitivity of detection can be increased by using locked nucleic acids (LNA) in the probes. LNAs are RNA nucleotides, in which the ribose moiety is modified with an extra bridge connecting the 2' oxygen and 4' carbon. The bridge "locks" the ribose in the 3'-endo conformation. LNAs hybridize with DNA or RNA according to Watson-Crick base-pairing rules, where the locked ribose conformation increases the hybridization properties of the probe. LNA oligomers are synthesized and commercially available and are now frequently used for miRNA detection. miRNA detection via radioactive labelling and hybridization remains the number-one method of miRNA analysis in plants and many other species, because it is comparatively cheap, easy to perform in every laboratory, and does not rely on high-tech equipment, such as sequencing machines.

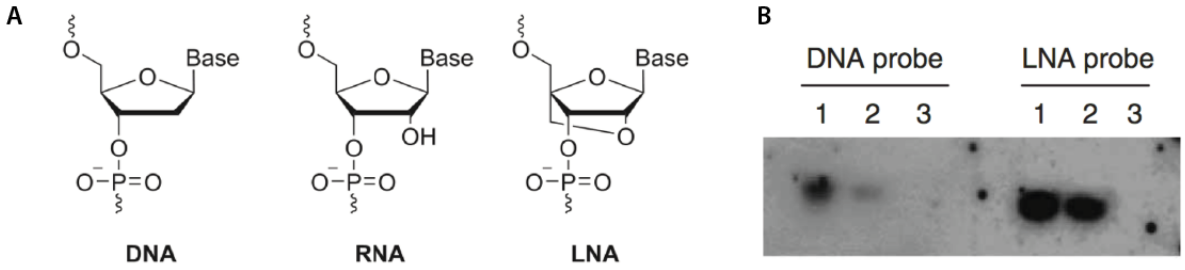

Figure 2-2 **(A)** Comparison of the structures of DNA, RNA, and LNA (locked nucleic acids) nucleotides. (Adapted from K. Astakhova, *Chemosensors*, 2014) **(B)** Northern blot from RNAs isolated from *A. thaliana* leaves using a conventional DNA oligonucleotide probe and an LNA (locked nucleic acids) probe to detect a low abundant miRNA. The detection sensitivity of a certain, plant-specific miRNA improved using LNA probes compared to a conventional DNA oligonucleotide probe. Samples 1 and 2 are from *A. thaliana* flowers and leaves, respectively, and sample 3 is from mouse liver. (Adapted from E. Vàrallylay *et al.*, *Nature Protocols*, 2008)

Quantitative reverse-transcriptase PCR

Another very popular technique for validating and accurately quantifying miRNAs is quantitative reverse-transcriptase PCR (qRT-PCR). Here, you also aim at identifying whether a specific transcript is present in your sample (the RT-PCR part) and additionally determine to which amount this RNA is present (the quantitative part). Let's first look at the RT-PCR part of this technique. It starts with a population of RNA molecules isolated from a tissue or cell culture, which you would like to amplify. Here, the first obstacle is that in order to amplify the RNA using PCR, the RNA must somehow serve as a template. However, PCR uses DNA polymerase, and most DNA polymerases cannot use RNA as a template. Thus, RNA has to be converted into DNA first. How is this achieved? Here, a unique type of polymerase known as reverse transcriptase (RT), which is able to synthesize DNA using an RNA template, is used (RTs are encoded by retroviruses to convert their RNA into DNA that can be inserted into the host genome). This method is therefore called RT-PCR.

In a typical RT-PCR reaction, RNA is first reverse-transcribed into DNA using a primer that binds to the RNA of interest. The resulting DNA is called copy DNA (cDNA). In the next step, the second strand is synthesized by PCR. This double-stranded DNA molecule now serves as a template for amplification in the subsequent PCR cycles.

When setting up the RT-PCR reaction, the first question is how to choose the primers used for reverse transcription and for the PCR reaction. There are several strategies, depending on the question you want to answer.

1. If you want to analyze the expression of a single gene, you can use primers that will anneal specifically with the RNA of interest. Thus, you use gene-specific primers.

2. In many cases, however, you would like to analyze the expression of several genes within the same sample. To be able to reverse-transcribe different RNAs at the same time, there are two options: First, you can use primers that randomly bind to RNAs. These are usually short primers of six nucleotides in a random arrangement (also called random hexamers), e.g., ACGGCA. Second, you can use a primer that relies on the fact that mRNA molecules carry a poly(A) tail at their 3' end, to which primers containing multiple T nucleotides can bind. These primers are called poly dT primers or oligonucleotides.

For the second part of the RT-PCR, the PCR reaction, the primers are chosen specifically for each RNA to be amplified. Thus, for each analyzed RNA, a PCR reaction is set up with a forward and a reverse primer that specifically bind to the target (see figure 2-3).

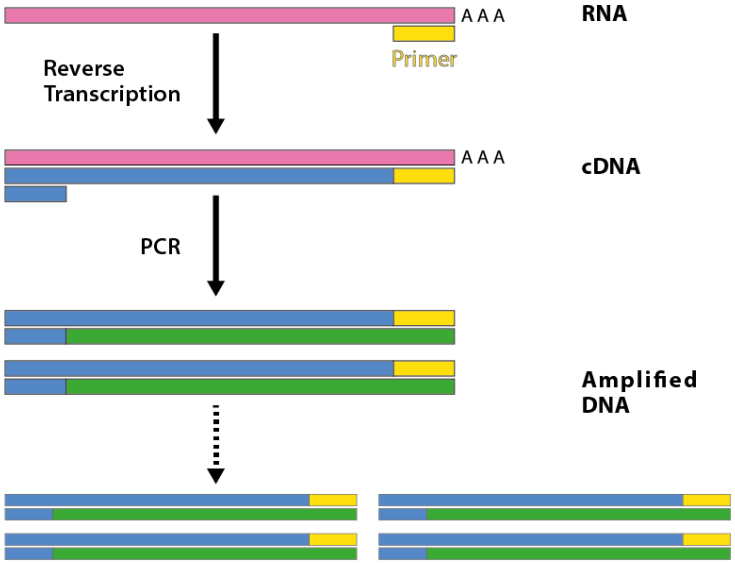

Figure 2-3 The reverse-transcriptase polymerase-chain-reaction (RT-PCR) technique for determining whether a particular type of mRNA is present. First, the mRNA from a small sample is converted into double-stranded cDNA using the enzyme reverse transcriptase and a primer binding to the RNA (yellow; either a random or specific primer, depending on the research question). The second strand is completed using Taq polymerase. After denaturing the DNA, the primers hybridize to opposite ends of the target sequence, the complementary strand is synthesized, and the target DNA is amplified like in a conventional PCR.

The workflow presented above applies to mRNAs that contain a poly(A) tail. However, if we wish to analyze miRNAs, we face a problem: miRNAs are tiny, with a length comparable to that of a typical DNA primer. Thus, to reverse-transcribe miRNAs into DNA, the size of the miRNAs have to be increased. This can be achieved by adding a poly(A) tail using the poly(A) polymerase. This added poly(A) tail can then be used for universal priming of miRNAs with oligo dTs. These oligo dTs contain a universal adapter, a nucleotide sequence that can later on be recognized by a primer in the PCR reaction. Once the miRNA has been converted to cDNA, it can be assayed using the same approach as a conventional PCR experiment. In the PCR reaction, a miRNA-specific forward primer together with a universal reverse primer (that binds to the adapter at the 3' end and was introduced in the RT reaction) is used (see figure 2-4).

1. **Poly(A) Tailing**
Poly(A) Polymerase, a template-independent enzyme, catalyzes the transfer of adenosine residues from ATP to the 3′-end of all RNA, including miRNA.

2. **cDNA Synthesis**
miRNA reverse transcription using an oligo-dT primer. The oligo-dT primer includes an adapter sequence at its 5′-end.

3. **Quantitative miRNA PCR**
qPCR with a forward primer complementary to a specific miRNA and a reverse primer complementary to the adapter sequence.

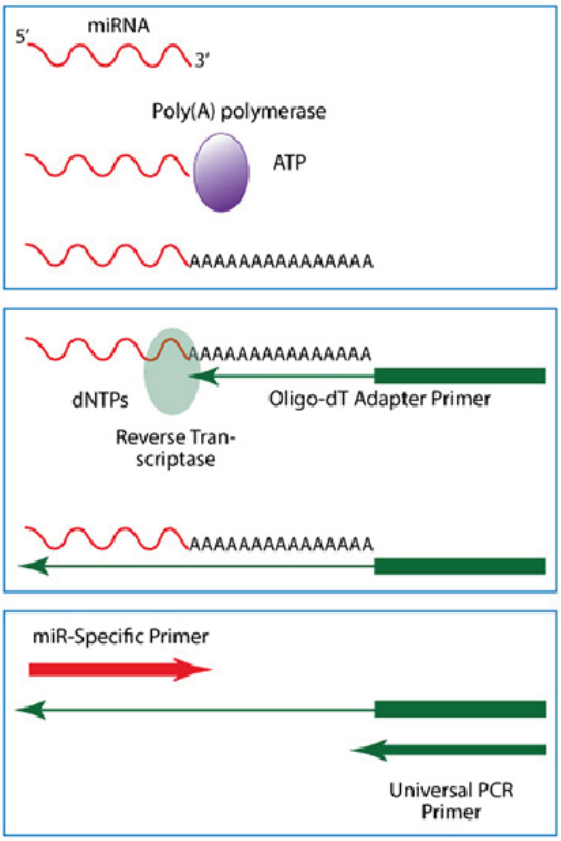

Figure 2-4 qRT-PCR method to detect miRNAs. (1) Addition of a poly(A) tail to the non-adenylated miRNAs in a sample of total RNAs. **(2)** Reverse transcription with an adapter primer (oligo dT with unique sequence at the 5′ end for universal reverse primer binding) synthesizes cDNA. **(3)** PCR with miRNA-specific forward primer and universal reverse primer. (Adapted from sigmaaldrich.com)

This workflow is also applied if the aim is to simply detect a specific RNA. How can this system be turned into a technique to measure RNA quantities? The trick is to use the rate of the PCR reaction as a measure of the RNA amount present in the sample, because the rate at which the PCR product is generated is directly related to the amount of RNA present in the sample. Thus, at each cycle of the PCR, the amount of PCR product is determined. This is achieved by adding chemical dyes to the PCR that fluoresce only when bound to double-stranded DNA. Thus, by simply measuring the fluorescence after each cycle, the progress of the reaction can be monitored (see figure 2-5). The intensity of the fluorescence signal reflects the amount of RNA amplified. If a reaction reaches a high fluorescence intensity already after a few cycles, this indicates that the RNA amount in this sample was high. To be able to actually quantify the amount of RNA present, the RNA of an internal reference gene is also amplified by RT-PCR. Here, genes are chosen that are not supposed to change their expression, so called "housekeeping genes", such as ribosomal genes. Then, the relative amount of each RNA to this rRNA is calculated.

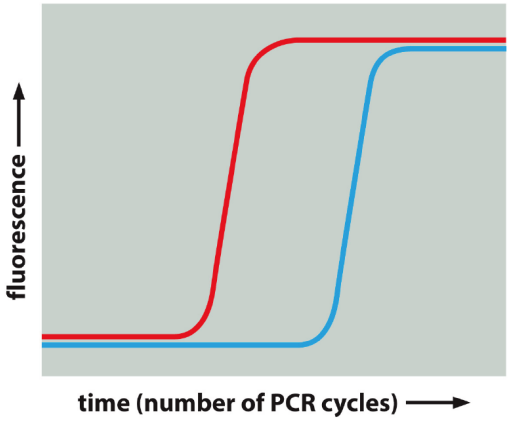

Figure 2-5 RNA levels can be measured by quantitative RT-PCR. The fluorescence measured is generated by a dye that fluoresces only when bound to the double-stranded DNA products of the RT-PCR. The red sample here has a higher concentration of the mRNA being measured than the blue sample, since it requires fewer PCR cycles to reach the same half-maximal concentration of double-stranded DNA. (Adapted from Figure 8-63, *Molecular Biology of the Cell*, 6th edition, Alberts *et al.*, *Garland Science*)

qRT-PCR is relatively fast to perform in the lab, and allows direct quantification of the RNA present in a sample. As well as being sensitive and quantitative, qRT-PCR is also relatively inexpensive and flexible, making it the preferred choice for validating novel miRNAs. However, qPCR has its limitations: unlike in conventional qRT-PCR, to detect miRNAs, only one flanking primer can be specific to the miRNA, so care must be taken to ensure only one product is being amplified. The short template length can prove a particularly problematic issue when trying to distinguish miRNA isoforms that may only differ by a few nucleotides, because the primers may anneal to, and amplify several distinct miRNA species at once.

Ultra-deep sequencing (RNA-Seq)

RNA can, just as DNA, be sequenced, with a method called RNA-Seq. RNA-Seq uses next-generation-sequencing techniques to reveal the presence and quantify the amounts of RNA present in a sample. Thus, it is also applicable for identifying de novo small RNAs, such as microRNAs. The workflow for this technique includes the isolation of RNA, the conversion form RNA into cDNA, and then sequencing of the cDNA using a next-generation-sequencing platform, e.g., the Illumina technique that we have discussed in lesson 1 of the reverse genetics part.

Why can't we sequence RNA directly, but have to convert it into DNA first? First, RNA is considered highly unstable, because it is a lot more prone to hydrolysis due to the nucleophilic 2'-hydroxyl groups. In contrast, DNA is more biologically stable, so converting RNA to DNA ensures the stability of the sample's information content. Second, in most cases, you will have small amounts of RNA, and you will have to amplify it before you can use it for sequencing. PCR amplification only works on DNA, so unless you can obtain enough RNA to feed directly into your sequencing protocol, you need to amplify it using PCR, and you therefore have to reverse-transcribe it to cDNA. Third, many sequencing techniques still rely on DNA polymerase for sequencing.

When sequencing small RNAs, the RNAs to be analyzed are isolated through size selection to enrich for those small RNA molecules (e.g., with a size-exclusion gel, through size-selection magnetic beads,

or with a commercially developed kit). Then, the RNA must be converted to double-stranded complementary DNA. Currently available sequencing technologies require a DNA template with platform-specific "adaptor" sequences at either end of each molecule. Generating the complementary DNA, adding the adaptors, and amplifying the DNA for sequencing are part of a process called "library preparation". RNA-Seq procedures differ in the library preparation, and the protocols highly depend upon which strategy and sequencing platform is used.

In order to convert RNA to DNA, a reverse transcription reaction as described in the qPCR assay is performed. All of the current protocols for library preparation use reverse transcriptase to synthesize a DNA strand using RNA as a template. RT, like other polymerases, requires a primer annealed to either DNA or RNA to initiate polymerization. Therefore, oligonucleotides are ligated to the 3' and 5' ends of the miRNA (see figure 2-6). The ligation step adds RNA adaptors to both ends of the small RNAs, which act as primer binding sites during reverse transcription and PCR amplification. A single-stranded, RNA 3' adaptor is ligated to the small RNAs using a ligating enzyme such as T4 RNA ligase. This primer is designed to specifically capture miRNAs that have a 3' hydroxyl group resulting from the enzymatic cleavage by Dicer-like enzymes. Similarly, a 5' adaptor is added that captures the 5' phosphate group in a second ligation step. The 5' adaptor is also designed to capture small RNAs with a 5' phosphate group, characteristic for all Dicer products, including microRNAs, rather than RNA degradation products with a 5' hydroxyl group. The next step is the reverse transcription that converts the small adaptor-ligated RNAs into cDNA clones used in the sequencing reaction. PCR is then carried out to amplify the pool of cDNA sequences (see figure 2-6). This creates cDNA constructs that contain the 5' and 3' adapter sequences. The PCR is performed using primers that anneal to the end of the adapters. Finally, massive parallel sequencing is done using a next-generation-sequencing platform, such as Illumina.

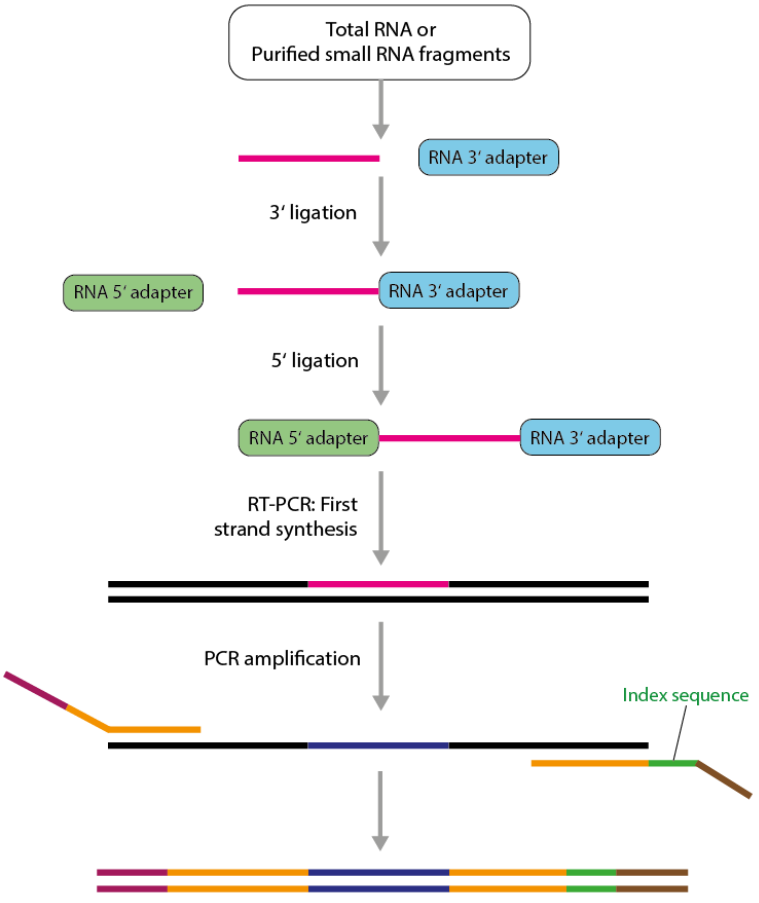

Figure 2-6 Library preparation workflow for miRNA-Seq using a high-throughput, next generation sequencing platform. From a total RNA isolate, miRNAs are enriched for by size fractionation. Then, RNA adapters are ligated to each end of the miRNA molecule (pink) and an RT reaction is used to create single-stranded cDNA. The cDNA is then PCR amplified using a primer binding the 5′ adapter and a primer binding the 3′ adapter (yellow). The primers also contain adapters that are specific for the sequencing platform to be used (red and brown). The reverse primer contains one of 48 index sequences, a short sequence of six nucleotides that serve as a barcode (green). The index sequence can be used to identify a certain cDNA within a pool of cDNAs. This is important because the massive parallel sequencing step following the library preparation is carried out with a pool of up to 48 samples in one sequencing reaction. The resulting sequences contain the index sequence and therefore indicate where this sequence is derived from (e.g., from a cDNA library derived from a certain tissue). The workflow shown here is for Illumina sequencing.

The currently available next-generation-sequencing platforms are able to perform massive parallel sequencing, which allows sequencing a genomic region or a miRNA molecule multiple times, sometimes hundreds or even thousands of times in a short time frame. This procedure to aim for a high number of replicate reads for a DNA or RNA molecule is referred to as "deep sequencing". This approach allows researchers to detect rare molecule species comprising as little as 1% of the original sample. Deep sequencing of RNAs thus not only provides the sequence, but also the frequency of RNA molecules that are present at any particular time in a specific cell type, tissue, or organ. Another advantage of RNA-Seq is that it can identify both known and novel miRNAs (unlike northern blots, for example, where knowledge of the sequence of the miRNA is required). However, a limitation of this technique is the medium-high input quantity required (500 ng - 5 μg total RNA). Furthermore, the library preparation and PCR amplification can potentially induce errors in the sequence and some bias toward preferential sequencing of particular sRNA species to the detriment of others may occur due to preferential ligation. Finally, the library preparation is time-consuming and expensive, and the costs for deep sequencing are still relatively high although prices continue to be reduced steadily.

Identification of miRNA Targets: Target predictions

To be able to understand the function of miRNAs, the identification of their target genes is essential. The major part of animal miRNA targets has been identified through computer-based predictions. Generally, computational predictions using miRNA characteristics are difficult, because miRNAs are only 21-25 nucleotides long. However, in plants, bioinformatics has served to identify many miRNA targets, mainly because in plants, most miRNAs are nearly perfectly complementary to their targets. Interestingly, many predicted plant miRNA targets are transcription factors, highlighting possible miRNA functions in regulating diverse developmental processes in plants.

The computational prediction of animal miRNA targets, on the other hand, is less straightforward, because the miRNA-target pairings are not entirely complementary and mostly rely on the seed-complementary region. As a result, computer-assisted predictions initially had to rely on rules that were built on a few known miRNA-target interactions, and eventually generalized for genome-wide searches; as targets became experimentally confirmed and new miRNA:target pairs discovered, these rules were considerably refined throughout time. For example, starting from three animal miRNAs, *lin-4*, *let-7*, and *Bantam*, the targets of which have been experimentally validated, the fly genome was searched for miRNA targets on the basis of the following three criteria: perfect complementarity between the target 3'-untranslated region (UTR) and the seed sequence of the miRNA; favorable structural and thermodynamic heteroduplex formation between miRNA and its putative targets; and evolutionary conservation of miRNA target sites and of the seed in particular between closely related species, such as *Drosophila melanogaster* and *Drosophila pseudoobscura*.

To experimentally validate potential targets, the researchers fused a GFP reporter open reading frame upstream of the predicted target's 3' UTR (this generated a construct known as a "miRNA sensor") and examined GFP expression in *Drosophila* tissues with and without overexpression of the corresponding miRNA or in animals carrying loss-of-function mutations in the miRNA of interest. Using this approach, they gained experimental evidence for the computationally predicted targets of three miRNAs, including *Notch* genes, pro-apoptotic genes, and genes that encode metabolic enzymes. There exist several software packages, the algorithm of which take into account many parameters including conservation across species, multiplicity of target sites within a predicted target, thermodynamic data for the free energy of the miRNA-target hybrid, and mRNA accumulation data if available in the cell type, tissue, or organ under consideration. Although the prediction success rate is difficult to determine on a genome-wide scale, these studies are important steps towards understanding miRNA function in animals as well as the scale of miRNA-guided regulation in genetic network.

Identification of miRNA Targets: Validations

After targets have been computationally predicted, they have to be experimentally validated to verify that they actually constitute a real target. The GFP-sensor approach described above is one way to monitor how the expression of the target gene depends on the activity of a certain miRNA in vivo both in animals and plants (see course 4). Another type of reporter frequently used for miRNA target validation in cultured cells uses luciferase, an enzyme that produces luminescence by converting the substrate luciferin. In the so called dual-luciferase assay for miRNA-target validation, the 3' UTR of the gene suspected to be a miRNA target is cloned downstream of the translational stop codon of a *luciferase* gene. For this purpose, commercially available vectors carrying the *luciferase* gene are used; and they usually also contain a second type of *luciferase* gene (producing a protein that emits in a different light spectrum) that is devoid of the artificial miRNA target site. This second luciferase is used for normalisation for the light emission quantification. By transfecting the vector construct into cells, a fusion mRNA between the *luciferase* gene and the 3' UTR of the potential miRNA target is transcribed. If the predicted target is the actual miRNA target, then the endogenous miRNA will cause a reduction in activity of the miRNA-reporter luciferase compared to that of the control luciferase, indicating that the miRNA can bind to the 3? UTR target gene under investigation (see figure 2-7). The light emission by the luciferase reaction can then be measured in a luminometer, providing a convenient way of monitoring and also precisely quantifying the effect of the miRNA on its target.

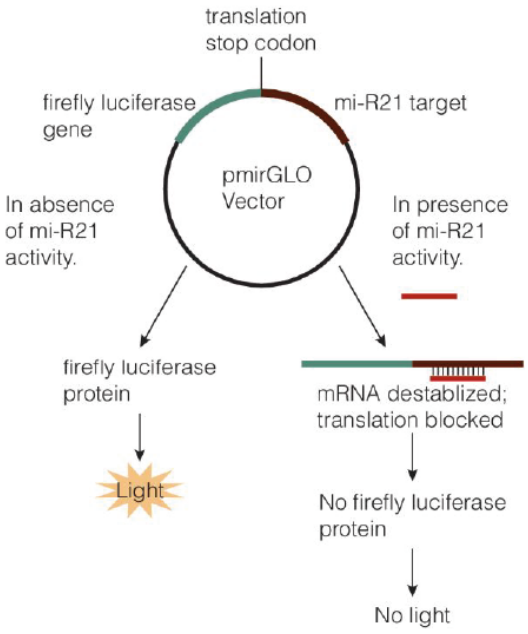

Figure 2-7 The luciferase reporter assay to validate an miRNA target. The 3′ UTR sequence of a suspected target gene is cloned downstream of a *luciferase* gene. After transfection of this construct in cells, a *luciferase*-3′ UTR fusion mRNA is made from this construct. If the target represents a real target, the presence of the respective miRNA will silence expression of this fusion mRNA, thus, luciferase activity is reduced. Note that the second *luciferase* gene used for normalization is not represented in this scheme. (Adapted from promega.com)

Thousands of miRNA targets have been validated in a variety of organisms. These miRNA-target interactions are deposited in a widely accessible database, the miRTarBase (mirtarbase.mbc.nctu.edu.tw). miRTarBase has accumulated more than 360'000 miRNA-target interactions derived from extensive literature search (see table 2-1). Generally, the collected interactions are validated experimentally by reporter assays, western blot, microarray, and next-generation-sequencing experiments. The database also contains information about expression levels and integrates data from other databases, e.g., the cancer genome atlas (TCGA).

Species	No. of miRNA-target interactions (MTIs)	No. of miRNAs	No. of target genes	No. of articles	No. of MTIs are validated by				
					Western blot	Luciferase assay	pSILAC	Microarray	NGS
Human	324,219	2,619	12,738	4,264	3,627	5,081	500	12,592	304,988
Mouse	41,588	1,885	7,146	669	744	1,038	2	753	39,818
Rat	488	150	247	198	212	220	0	204	2
Chicken	55	31	34	27	11	45	0	3	0
Cattle	242	21	111	10	4	8	0	1	0
Sheep	3	1	3	4	0	1	0	0	0
Japanese medaka	1	1	1	1	0	1	0	0	0
Zebrafish	135	39	98	52	52	103	0	7	0
Fruit fly	147	45	86	47	19	141	0	13	0
Silkworm	3	3	2	2	0	3	0	0	0
African clawed frog	4	2	4	5	0	1	0	0	0
Nematode	3,221	154	454	32	2	41	0	1	3,178
Western clawed frog	2	1	2	2	0	1	0	0	0
Rice	1	1	1	7	0	0	0	0	0
Thale cress	94	40	68	19	7	3	0	13	21
Epstein-Barr virus	1	11	1	7	1	1	0	0	0
HHV-8	1	8	1	3	0	1	0	0	0
Human cytomegalovirus	2	3	2	6	1	2	0	0	0
VSV	4	0	2	1	0	4	0	0	0
Total	370,211	5,015	21,001	5,356	4,580	6,695	502	13,587	348,007

Table 2-1 Current statistics of miRNA-target interactions in the miRTarBase.

Summary

In this lesson, we have discussed northern blotting, quantitative RT-PCR and deep RNA sequencing as the main techniques to identify and quantify miRNAs. Despite these techniques, bioinformatic predictions based on the characteristics of the so-far known miRNAs can identify novel miRNA genes in different organisms. Once a novel miRNA is identified, it is important to find the target of this miRNA to be able to understand the processes it regulates. Here, bioinformatics has served to identify various miRNA targets. Predicted miRNA have to be experimentally validated, which can be done using luciferase reporter assays. Once discovered, miRNAs as well as their targets are deposited in openly accessible databases that represent an important resource for the small RNA research community.