

# An ordered *Arabidopsis thaliana* mitochondrial cDNA library on high-density filters allows rapid systematic analysis of plant gene expression: a pilot study

Philippe Giegé<sup>1,\*</sup>, Zoltán Konthur<sup>2</sup>, Gerald Walter<sup>2</sup> and Axel Brennicke<sup>1</sup>

<sup>1</sup>Universität Ulm, Allgemeine Botanik, Albert-Einstein-Allee, D-89069 Ulm, Germany, and

<sup>2</sup>MPI für Molekulare Genetik, Abteilung Lehrach, Ihnestr. 73, D-14195 Berlin, Germany

## Summary

The availability of the complete sequence of a genome allows a systematic analysis of its expression. Gene-specific variations of transcription levels and phenomena such as transcript processing and RNA editing require large numbers of clones to be examined. For the completely sequenced mitochondrial genome of *Arabidopsis thaliana* we adapted robot technology to identify and characterize expressed genes. A cDNA library of about 50 000 clones was constructed, robot-ordered into 384-well microtitre plates and spotted onto high-density filter membranes. These filters permit the isolation of large numbers of specific cDNA clones in a single hybridization step. The *cox1*, *cox2* and *cox3* genes were used to evaluate the feasibility and efficiency of this approach. A cluster of RNA editing sites observed outside the *cox3* coding region identifies a novel reading frame *orf95* in higher plants with significant similarity to a subunit of respiratory chain complex II.

## Introduction

The rapid progress in genome analysis produces exponentially growing amounts of DNA sequence data from a wide variety of organisms. To fully evaluate the realised information potential of a given genome, expression analyses, using for example cDNA libraries, are essential. To test feasibility and handling of large ordered cDNA libraries of, e.g. plant genomes, we investigated the potential of robot-assisted cDNA ordering for a comparatively small plant mitochondrial genome. In one of our laboratories, we have completely sequenced the mitochondrial genome of the model plant *Arabidopsis thaliana* (Unseld *et al.*,

1997). This genome is 366 924 nucleotides long and contains 57 known genes. Another 85 open reading frames (ORFs) larger than 100 codons and beginning with an AUG start codon have been identified (Unseld *et al.*, 1997). To evaluate their functional significance, their status of transcription and RNA processing needs to be investigated. The latter processes include RNA editing, commonly found in assigned plant mitochondrial mRNAs (Bonnard *et al.*, 1992; Maier *et al.*, 1996). It is therefore necessary to systematically isolate and sequence cDNA clones of mitochondrial ORFs to detect potential editing sites. A transcribed ORF processed by RNA editing is also very likely to be translated into a functional protein. At the same time, a whole-genome study of cDNA sequences corresponding to the identified mitochondrial genes in *Arabidopsis thaliana* will allow an overview of the entire extent of sequence variation created by RNA editing in this model plant. Such a full-scale analysis will potentially give some clues on how individual RNA editing sites are specified and determined *in vivo*. Considering the divergent intensities of individual gene expression levels in plant mitochondria (Finnegan and Brown, 1990; Mulligan *et al.*, 1988) and the variation in the editing status of individual mRNAs, any meaningful mitochondrial cDNA library of plants must have a large number of individual clones.

At this point, a fast and easy method to specifically and reproducibly isolate large numbers of different cDNA clones is required. Obviously, the classical library screening technique of repeated plating resulting in individual clone patterns is not suitable for systematic studies. As an alternative, automated technology which allows the generation of large, ordered libraries has been developed in one of our laboratories (Maier *et al.*, 1994; 1997). This technology was specifically designed for high-throughput genome work and enables the automated handling of molecular libraries in the order of up to several millions of clones. A robotic device is used to pick bacterial colonies from agar plates and to transfer them into 384-well microtitre plates. Another module of the same robot is then employed to make high-density arrays by spotting aliquots of the ordered cultures onto filter membranes.

We used the completely sequenced mitochondrial genome of the model plant *Arabidopsis thaliana* in a pilot study aimed at the adaptation and transfer of automated, high-throughput library technology from human systems to plants. A mitochondrial cDNA library was made and

Received 20 May 1998; accepted 25 June 1998.

\*For correspondence (fax + 49 73150 22626; e-mail Philippe.Giege@biologie.uni-ulm.de).

about 50 000 clones were robot-picked and spotted onto sets of two high-density filters. Pilot screening analyses with the *cox1*, *cox2* and *cox3* genes revealed a high extent of RNA editing in *cox2* and *cox3* transcripts but no editing in *cox1* RNAs. A separate cluster of editing sites identifies a novel reading frame in higher plants. Using this technology, specific clones, including low copy transcripts, can now be identified in a single round of hybridization.

## Results and discussion

### High-throughput library handling and screening on high-density filters

An ordered *Arabidopsis thaliana* mitochondrial cDNA library was prepared in a high-density format using robot technology (Maier *et al.*, 1997), and clones were picked into 160 × 384-well microtitre plates. Subject to overnight growth, 144 × 384-well microtitre plates were replicated, and five copies of this library were stored at -80°C. A full set of 55 296 clones was spotted onto two high-density filters, with 12 copies produced initially. These filters were air dried and can be stored at room temperature for years. New copies can be made at any time, making the ordered cDNA library a very convenient resource for comparative hybridization studies (Lennon and Lehrach, 1991). Library screening on high-density filters can be done in only 2 days by a single round of hybridization. This speed marks a considerable advancement over classical library screening technology (Sambrook *et al.*, 1989). Formerly, phage library screening required three rounds of hybridisation. Each time, phage had to be plated and grown overnight, replica filters from agar plates had to be processed and hybridised for another night. Finally, when positive phage could be identified, phagemids were still to be excised. In contrast, the method described in this paper reduces the total time of the screening process from weeks to a few days.

### Library size

In contrast to smaller mitochondrial genomes (e.g. mammalian mitochondria) whose genetic information is very densely packed (Anderson *et al.*, 1981; Bibb *et al.*, 1981), plant mitochondrial genomes contain large intergenic spaces (Ward *et al.*, 1981). Only about 30% of the 366 924 nucleotides of the *Arabidopsis thaliana* mitochondrial genome are part of coding regions (genes/open reading frames or introns; Unseld *et al.*, 1997). The genome seems to contain approximately 100 genes in total, including 57 identified genes and other open reading frames of yet unknown functions. Therefore, in a cDNA library of about 50 000 clones, each gene should be represented by an average of 200 clones, a much higher coverage than normally intended with genomic libraries. However, in

**Table 1.** Evaluation of the *Arabidopsis thaliana* cDNA library

	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>
Positive clones (percentage of the library)	104 (0.19%)	132 (0.24%)	142 (0.26%)
Analysed clones	20	10	15
Different clones	10	7	10
False positives	2	0	1

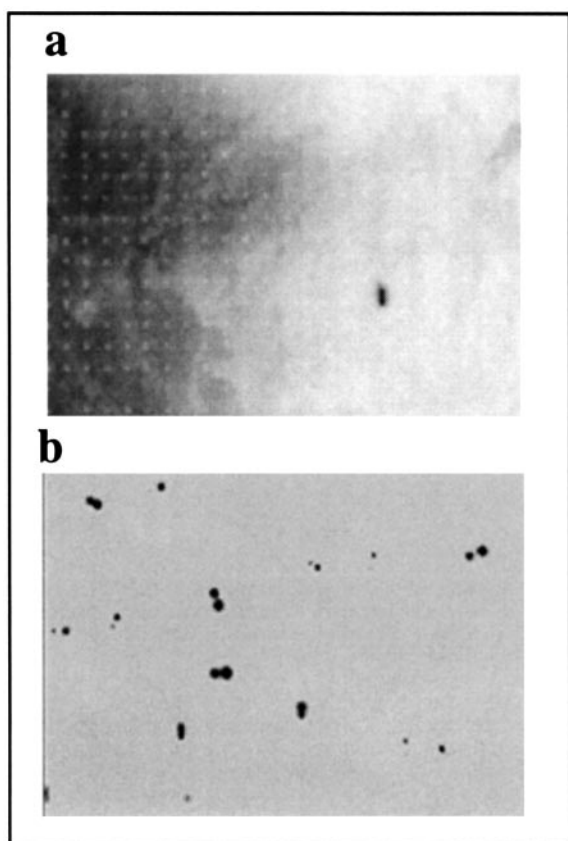
**Table 2.** Extent of RNA editing in the analysed open reading frames

	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>
Editing sites observed	0	15	8
Calculated number of mRNAs differing in individual editing sites	1	2 <sup>15</sup> =32768	2 <sup>8</sup> =256
Observed number of different editing patterns	1	6	6
Completely unedited clones	10	1	1

a non-preselected cDNA library, some transcripts (e.g. ribosomal RNA clones) are expected to be massively over-represented at the expense of other, mainly low copy transcripts. To be able to screen for the latter, we chose to make our ordered cDNA library as large as conveniently possible. In addition, large numbers of clones were considered necessary for the detection of transcripts of different processing and editing states. The medium-sized *Arabidopsis thaliana* mitochondrial genome provides a suitable test case for high-throughput screening of cDNA libraries. One hundred and forty-four microtitre plates or two filters, respectively, contain all clones necessary for a 200-fold coverage of the expected gene content. For the detection of low copy transcripts and different processing/editing states in larger libraries (e.g. nuclear cDNA libraries), demands on high-throughput library handling and screening will of course be greater. Considering a library of a million different clones, more than 2600 × 384-well microtitre plates would have to be handled and stored. It is expected that further development in automation will be needed to reduce storage space and allow higher clone densities on filters or other surfaces (e.g. chip arrays, Schena, 1996).

### Evaluation of the library

The *Arabidopsis thaliana* mitochondrial cDNA library was preliminarily screened with specific probes of bona fide genes. The two high-density filters were test-hybridised with PCR-amplified *cox1* (Isaac *et al.*, 1985), *cox2* (Fox & Leaver, 1981) and *cox3* (Hiesel *et al.*, 1987) fragments, yielding more than 100 positive clones each (Figure 1, Table 1). To evaluate the reliability of double spotting for distinguishing positive signals from artefacts, about 10% randomly selected positives (Table 1) were analysed by DNA sequencing. It was

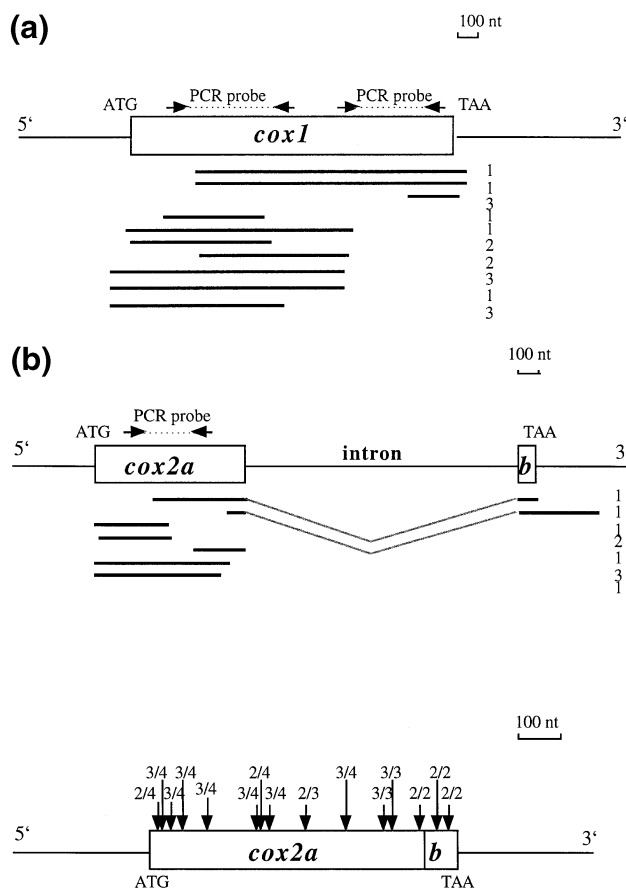


**Figure 1.** Example of a high density filter screen hybridization. The same sixth of a filter with 4608 clones is represented in both pictures. Each clone is spotted twice for identification. (a) shows the 24/16 guide dot frame visible above the hybridization background. (b) is a representative filter section hybridised with the *cox3* coding region. The location of a positive clone in a microtitre plate is identified by position and pattern of the two radioactive spots relative to the guide dot frame.

found that all but three of the 45 clones sequenced were correctly identified and indeed contained parts of the *cox1*, 2 or 3 sequences. False positive clones are cDNA sequences found unrelated to the probe and are due to either hybridization artefacts or to erroneously selected clones. Nearly all the *cox2* and *cox3* clones were found to be partially or fully edited (Table 2), only one cDNA in each case being unedited and identical to the genomic sequence. All *cox2* cDNA clones covering the respective area are bona fide spliced cDNAs and have lost the single intron (Figure 2). No editing activity was observed in *cox1* clones. Provided that the number of positive clones identified in the library reflects the abundance of the respective steady state transcripts, the three genes investigated (*cox1*, *cox2* and *cox3*) are transcribed with comparable abundance, each represented by 0.19–0.26% of the clones (Table 1).

#### Use of the library for gene identification

The use of automated technology for the rapid isolation of defined clones from a mitochondrial cDNA library of

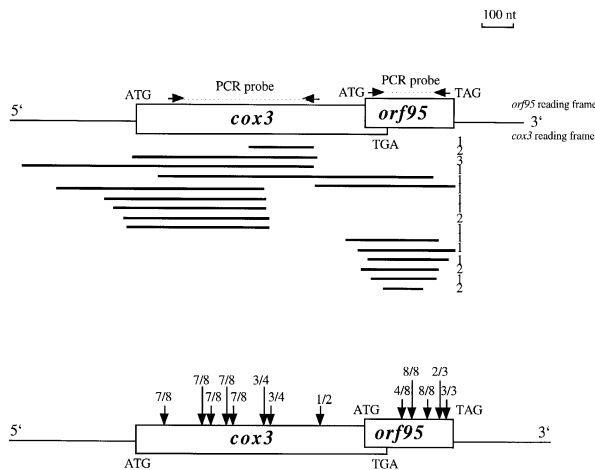


**Figure 2.** Analysis of the cDNA clones identified for the *cox1* (a) and *cox2* (b) genes.

Arrows show the position of the primers used to amplify the specific PCR probes. Lines underneath the schematic reading frames indicate position and extent of the different cDNA clones analysed by DNA sequencing. The respective number of clones identified is given on the right. The lower part in (b) shows the positions of the editing sites detected in the *cox2* open reading frame (vertical arrows). The number of edited clones per total number of clones analysed is given for each site. For *cox2* the thin lines indicate the absence of intron sequence in the analysed clones. No editing sites are observed in any of the *cox1* cDNA clones. In the EMBL database, RNA editing will be inserted as annotations to the genomic sequence with accession numbers Y08501 and Y08502 as soon as formats have been coordinated with the European MITBASE project.

*Arabidopsis thaliana* enables high-throughput gene expression, i.e. transcription studies of this genome. One of the main goals is the detection of actively expressed open reading frames (ORFs) beyond the already identified genes. The identification of cDNA clones representing a given ORF is evidence for its transcription, and suggests that this specific ORF is an actively used gene. Furthermore, RNA editing is about 100 times more frequent in genes than in non-coding regions (Bonnard *et al.*, 1992; Maier *et al.*, 1996). Therefore, evidence for RNA editing activity modifying a given ORF will further pinpoint candidate genes.

Indeed, when investigating cDNA sequences of the *cox3* gene, a cluster of RNA editing events is seen downstream of the *cox3* reading frame (Figure 3). Close examination



**Figure 3.** A cluster of RNA editing sites downstream of the *cox3* reading frame identifies a novel orf in the *Arabidopsis* mitochondrial genome. Arrows identify the primers of the specific PCR probes. Frequency, position and extent of the different cDNA clones are indicated. The lower part shows the positions of the editing sites detected in the depicted open reading frames (vertical arrows). The number of edited clones per total number of clones analysed is given for each site. The novel open reading frame downstream of, and partially overlapping with, the *cox3* gene shows highest similarity to the fourth subunit of the respiratory chain complex II, the succinate-ubiquinone oxidoreductase (*sdh4*; P. Giegé, V. Knoop and A. Brennicke, to be detailed elsewhere).

of this sequence identifies a novel open reading frame potentially coding for a protein of 95 amino acids. Database searches with the deduced protein sequence reveal significant similarity to a respiratory chain complex II subunit (*sdh4*) encoded in the mitochondrial genomes of some algae and the liverwort *Marchantia* (Oda *et al.*, 1992; P. Giegé, V. Knoop and A. Brennicke, unpublished results). This first identification of a potential complex II orf in higher plants by the cDNA analysis described here thus illustrates the potential and power of this approach.

#### Use of the library for gene expression studies

In addition, the characterisation of all editing sites in the genome will possibly give clues to the specificity of the editing mechanism. As a positive selective pressure, which is presumably needed to maintain such a complex and specific activity as RNA editing, this mechanism seems to be indispensable for the creation of most functional mitochondrial proteins. Besides finding potentially new genes, screening of the ordered cDNA library will also yield information on expression levels of already identified genes. This can be complemented by studies at the protein level, i.e. the screening of cDNA expression libraries with sets of antibodies. Alternatively, it seems feasible to screen a cDNA library, to identify edited ORFs and then to create specific antibodies against candidate proteins in order to test for their translation.

#### Comparative expression study strategies

An *Arabidopsis* expression study consisting of the specific isolation of cDNA clones can complement the EST (expressed sequence tags) strategy of random cDNA sequencing in *Arabidopsis* and many other organisms. The EST method has allowed the rapid detection of thousands of unknown genes (Cook *et al.*, 1997) and is particularly effective in expression studies of new genomes. The fact that EST sequences are selected randomly implies that genes expressed at a low level, such as regulatory genes, will be difficult to detect. In an ordered library of cDNA clones, unique gene sequences randomly found by the EST approach can be specifically used to isolate further cDNA clones very rapidly and economically. An ordered cDNA library can thus be used to complement EST and genomic sequencing projects.

#### Experimental procedures

##### RNA preparation from mitochondria of *Arabidopsis thaliana*

Total mitochondrial RNA (mtRNA) was prepared from an *Arabidopsis thaliana* var. Columbia cell suspension culture. Cells were disrupted with a Waring blender by three strokes of 5 sec each at high speed. Mitochondria were isolated by differential centrifugation using a discontinuous sucrose gradient with 20, 35, 47 and 60% sucrose steps. Two mitochondrial fractions were taken at the 35/47% and 47/60% sucrose interphases. Mitochondria were lysed 5 min on ice in a 0.2 M Tris-HCl pH 7.5, 0.1 M EDTA and 0.1 mg ml<sup>-1</sup> sodium sarcosyl solution. Mitochondrial RNAs were purified by phenol/chloroform extractions. Two successive 2 M LiCl precipitations were done to fractionate high molecular weight RNA from DNA and small RNA molecules. The final mtRNA pellet was precipitated with ethanol and dried. Further purification of mitochondrial RNA by polyA<sup>+</sup> selection is not feasible for plant mitochondrial transcripts, since transcripts are generally not polyadenylated in these organelles. Therefore, the abundant mitochondrial rRNAs cannot be completely removed, and about 40% of the cDNA clones are derived from fragments of rRNA in addition to mRNA sequences. The smaller the cloned mRNA derived sequence, the more frequently connected rRNA sequences are co-inserted.

##### Construction of the mitochondrial cDNA library

A Lambda ZapII (Stratagene) phage library was constructed from 20 µg of *Arabidopsis* mtRNA. First-strand cDNA synthesis was initiated with 10 µg of random hexamer primers. Further construction procedures followed the manufacturer's instructions. Double stranded *EcoRI/EcoRI* cDNA fragments were fractionated by size on an S400 sephacryl gel filtration column. cDNAs larger than 200 nucleotides were ligated with Lambda ZapII vector and packaged.

##### Robot-assisted picking of the phagemid library

The phage library was mass-excised with Ex Assist helper phage, and the phagemids obtained were transformed into SOLR *E. coli*

cells. The resulting bacterial library was plated onto 222 × 222 mm Genetix square dishes containing 2YT agar with 2% (w/v) glucose and 100 µg ml<sup>-1</sup> ampicillin and grown at 37°C overnight. The bacterial density was adjusted to 3000–6000 colonies per plate. Colonies were picked into 384-well microtitre plates containing 2YT medium with 2% (w/v) glucose, 100 µg ml<sup>-1</sup> ampicillin and 10% Hogness Modified Freezing Medium (HMFM) by a robot as described previously (Maier *et al.*, 1997). Briefly, the robot visualises colonies using a CCD camera and chooses individual colonies for picking, according to shape and size. Evaluation parameters were adjusted to the colony appearance of the SOLR strain. Target colonies were individually picked using a 96-pin picking device mounted to the xyz-drive of the robot. Inoculated 384-well microtitre plates were incubated at 37°C overnight, replicated using Genetix 384-pin replicators and stored at -80°C.

#### Production and processing of high-density filters

Using the same robotic system, colonies freshly grown in 384-well microtitre plates were spotted by transferring 50 nL of bacterial culture onto Hybond-N+ 222 mm × 222 mm square filters (Amersham). High-density filters were produced in a 5 × 5 format with the smallest spotting unit consisting of a square of 25 spots. The central position of the square was used as a guide dot marked with black ink. A 48 × 48 guide dot frame was used with 12 different clones spotted in a specific duplicate pattern around each guide dot. Duplicate spotting is crucial for distinguishing positive signals from hybridization artefacts (Figure 1). Two filters displaying 27 648 clones each were produced of the *Arabidopsis thaliana* cDNA library making a total of 55 296 clones available for screening. The filters were transferred onto 222 × 222 mm Genetix square dishes containing 2YT agar with 2% (w/v) glucose and 100 µg ml<sup>-1</sup> ampicillin, and bacterial colonies were grown at 37°C overnight. Subsequently, filters were processed by denaturation and neutralisation steps using a modification of the method described by Hoheisel *et al.* (1991). Briefly, colonies were denatured by two 4 min incubations on Whatman paper soaked in denaturing solution (0.5 M NaOH, 1.5 M NaCl). After the first incubation at room temperature, filters were transferred to a glass plate floating on a 95°C waterbath. Filters were neutralised for 4 min at room temperature on Whatman paper soaked in neutralising solution (1 M Tris-HCl pH 7.4, 1.5 M NaCl). To digest the proteins, filters were submerged for 30 min in prewarmed Proteinase K solution (50 mM Tris-HCl pH 8.5, 50 mM EDTA, 100 mM NaCl, 1% (w/v) N-Lauryl sarcosine, 150 mg Proteinase K/600 ml solution) at 37°C. The filters were air-dried for 2 days at room temperature, and finally, the DNA was auto-crosslinked by short-wave UV using a Stratalinker (Stratagene).

#### Hybridization with specific PCR probes

PCR products specific for *cox1* (Isaac *et al.*, 1985), *cox2* (Fox and Leaver, 1981) and *cox3* (Hiesel *et al.*, 1987; 417 and 437 nucleotides long for *cox1*, 374 and 462 nucleotides long for *cox2* and *cox3*, respectively) were used to test-hybridize the high-density filters. Oligos Acox1 5'-TTCATTCTGGTTGTTG CCAC-3', Bcox1 5'-CGAAGAACAAAAGAGATGC-3', A2cox1 5'-TTCC ATTATGACTTTCTATGGG-3', B2cox1 5'-GCTTTTCGCTCCTTGATAG C, Acox2 5'-CAATTAGGATCTCAAGACGC-3', Bcox2 5'-TGTCAAAAGTG AGT-GACTGC-3', Acox3 5'-ATCCATGTTCTCGTATGGTGG-3' and Bcox3 5'-AAACCATGAAAGCCTGTTGC-3' were derived from genomic sequence data (Unseld *et al.*, 1997). Further *orf95* specific library screening was undertaken using a PCR probe amplified between

oligonucleotides A95 5'-ATGAAGGAACGAAAGAGTGGATTAC-3' and B95 5'-GGTCATTCTTGG TGAACATGATC-3'. 50 ng of each PCR probe were internally labeled with 50 µCi of α-<sup>32</sup>P dATP. Prehybridization was performed for 1 h at 65°C in a 1% SDS, 1 M NaCl and 50 mM Tris-HCl pH 7.5 solution, containing 200 mg ml<sup>-1</sup> calf thymus DNA denatured for 5 min at 100°C. Per filter, 15 ml of the hybridization solution were used in a rotating hybridization cylinder. 1 µCi ml<sup>-1</sup> of labeled probe was added to the solution and hybridized overnight at 65°C. Blots were washed with increasing stringency at 65°C for 15 min with 1 × SSC, 0.1% SDS and for 15 min with 0.1 × SSC, 0.1% SDS.

High-density filters of the described *Arabidopsis thaliana* mitochondrial cDNA library are available from the Ressourcenzentrum im Deutschen Humangenomprojekt (RZPD), MPI für Molekulare Genetik, Heubnerweg 6, D-14059 Berlin, Germany, under ARAMt-cDNA.

#### Acknowledgements

This work is supported by a grant from the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.

#### References

- Anderson, S., Bankier, A.T., Barrell, B.G. *et al.* (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.
- Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell*, **26**, 167–180.
- Bonnard, G., Gualberto, J.M., Lamattina, L. and Grienenberger, J.M. (1992) RNA editing in plant mitochondria. *Crit. Rev. Plant Sci.* **10**, 503–524.
- Cooke, R., Raynal, M., Laudié, M. and Delseny, M. (1997) Identification of members of gene families by contig construction from partial cDNA sequences: 106 genes encoding 50 cytoplasmic ribosomal proteins. *Plant J.* **11**, 1127–1140.
- Finnegan, P.M. and Brown, G.G. (1990) Transcriptional and post-transcriptional regulation of RNA levels in maize mitochondria. *Plant Cell*, **2**, 71–83.
- Fox, T.D. and Leaver, C.J. (1981) The *Zea mays* mitochondrial gene coding cytochrome oxidase subunit II has an intervening sequence and does not contain TGA codons. *Cell*, **26**, 315–323.
- Hiesel, R., Schobel, W., Schuster, W. and Brennicke, A. (1987) The cytochrome oxidase subunit I and III genes in *Oenothera* mitochondria are transcribed from identical promoter sequences. *EMBO J.* **6**, 29–34.
- Hoheisel, J.D., Zehetner, G., Lehrach, H. and Lennon, G.G. (1991) Use of high coverage reference libraries of *Drosophila melanogaster* for relational data analysis. A step towards mapping and sequencing of the genome. *J. Mol. Biol.* **220**, 903–914.
- Isaac, P., Jones, V.P. and Leaver, C.J. (1985) The maize cytochrome c oxidase subunit I gene: sequence, expression and rearrangement in cytoplasmic male sterile plants. *EMBO J.* **4**, 1617–1623.
- Lennon, G.G. and Lehrach, H. (1991) Hybridization analyses of arrayed cDNA libraries. *Trends Genet.* **7**, 314–317.
- Maier, E., Bancroft, D.R. and Lehrach, H. (1997) Large-scale library characterisation. In *Automation Technologies for Genome Characterization* (Beugelsdijk, T.J., ed.). New York: John Wiley & Sons Inc., pp. 65–88.

- Maier, E., Meier, E.S., Ahmadi, A.R., Curtis, J. and Lehrach, H.** (1994) Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridisation. *J. Biotech.* **35**, 191–203.
- Maier, R.M., Zeltz, P., Kössel, H., Bonnard, G., Gualberto, J.M. and Grienberger, J.M.** (1996) RNA editing in plant mitochondria and chloroplasts. *Plant Mol. Biol.* **32**, 343–365.
- Mulligan, R.M., Leon, P. and Walbot, V.** (1988) Transcriptional and post-transcriptional regulation of maize mitochondrial gene expression. *Mol. Cell. Biol.* **11**, 533–543.
- Oda, K., Yamato, K., Ohta, E. et al.** (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. *J. Mol. Biol.* **223**, 1–7.
- Sambrook, J., Fritsch, E.F. and Maniatis, T.** (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Schena, M.** (1996). Genome analysis with gene expression microarrays. *Bioessays*, **18**, 427–431.
- Unseld, M., Marienfeld, J.R., Brandt, P. and Brennicke, A.** (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet.* **15**, 57–61.
- Ward, B.L., Anderson, R.S. and Bendich, A.J.** (1981) The mitochondrial genome is large and variable in a family of plants. *Cell*, **25**, 793–803.