

TECHNICAL ADVANCE

Analysis of clones carrying repeated DNA sequences in two YAC libraries of *Arabidopsis thaliana* DNA

Renate Schmidt, Joanna Putterill, Joanne West, Gerda Cnops[†], Frances Robson, George Coupland and Caroline Dean*

Department of Molecular Genetics, Cambridge Laboratory, AFRC Institute of Plant Science Research, John Innes Centre, Colney, Norwich NR4 7UJ, UK

Summary

YAC clones carrying repeated DNA sequences from the *Arabidopsis thaliana* genome have been characterized in two widely used *Arabidopsis* YAC libraries, the EG library and the EW library. Ribosomal, chloroplast and the paracentromeric repeat sequences are differentially represented in the two libraries. The coordinates of YAC clones hybridizing to these sequences are given. A high proportion of EG YAC clones were classified as containing chimaeric inserts because individual clones carried unique sequences and repetitive sequences originating from different locations in the genome. None of the EW YAC clones analysed were chimaeric in this way. YAC clones carrying tandemly repeated sequences, such as the paracentromeric or rDNA sequences, exhibited a high degree of instability. These observations need to be taken into account when using these libraries in the development of a physical map of the *Arabidopsis* genome and in chromosome walking experiments.

Introduction

Arabidopsis thaliana is a plant species suited to gene cloning strategies which rely on knowledge of the map position of the gene (Arondel *et al.*, 1992; Giraudat *et al.*, 1992). Loci involved in many aspects of plant growth and development have been identified in *Arabidopsis* and their cloning using chromosome walking techniques initiated (reviewed by Schmidt and Dean, 1992). Map-based gene cloning would be more efficient if a complete physical map of the *Arabidopsis* genome were available.

Considerable progress has been made towards this goal. First, a set of overlapping cosmid clones has been assembled into 750 contigs covering 90–95% of the genome (Hauge *et al.*, 1991). Second, efforts are underway to construct overlapping YAC contigs for the five *Arabidopsis* chromosomes. The three YAC libraries commonly in use are the EG (Grill and Somerville, 1991), EW (Ward and Jen, 1990) and Yup (Ecker, 1990) libraries. These are all derived from the Columbia ecotype. The average insert size in the EG and the EW libraries is 160 kb and the libraries represent three to four genome equivalents each, while the Yup library contains YACs with an average insert size of 250 kb covering more than five genome equivalents. YAC clones containing 30% of the *Arabidopsis* genome have already been identified by hybridization to 125 mapped markers (Hwang *et al.*, 1991). The ultimate aim is to integrate the *Arabidopsis* phenotypic, RFLP and physical maps.

As part of the effort to generate an overlapping YAC library, we have identified YAC clones carrying some classes of repeated DNA sequences in the EG and EW libraries. The main types of repeated DNA sequences within total *Arabidopsis* DNA are chloroplast and ribosomal DNA (rDNA) (Pruitt and Meyerowitz, 1986). A diploid *Arabidopsis* cell contains about 660 copies (Pruitt and Meyerowitz, 1986) of the 155 kb chloroplast genome (Brennicke and Schuster, personal communication). Furthermore, in a number of plant species, nuclear-derived sequences with homology to the chloroplast genome (Timmis and Steele Scott, 1985) have been described. Ribosomal DNA is present in 570 tandem copies of the 25S–18S rDNA repeat unit in the haploid *Arabidopsis* nuclear genome (Pruitt and Meyerowitz, 1986). *In situ* hybridization experiments have identified two rDNA loci, one each on chromosomes 2 and 4 (Maluszynska and Heslop-Harrison, 1991).

In addition, a highly repeated DNA sequence of 180 bp has been described. This *Hind*III repeat sequence is arranged in tandem arrays (Martinez-Zapater *et al.*, 1986; Simoens *et al.*, 1988) and is closely related to a 500 bp repetitive sequence (Simoens *et al.*, 1988). These repeated sequences account for 1–1.6% of the *Arabidopsis* nuclear genome (Simoens *et al.*, 1988) and are part of the paracentromeric heterochromatin of all five chromosomes (Maluszynska and Heslop-Harrison, 1991).

Received 2 August 1993; revised 20 December 1993; accepted 11 January 1994.

*For correspondence (fax +44 603 505725).

[†]Present address: Laboratorium Genetika, Universiteit Gent, Ledeganckstraat 35, 9000 Gent, Belgium.

The purpose of the experiments described here was to identify and characterize YAC clones containing these main categories of repeated DNA sequences. This would enable chromosome walks into these highly repeated regions to be quickly recognized and terminated. YAC clones containing chimaeric inserts between repeated and unique DNA sequences could also be more readily identified. The stability of YAC clones carrying the repeated DNA sequences has been investigated.

Results

Identification of YAC clones containing chloroplast DNA sequences

Two cosmid clones, representing more than 80% of the *Arabidopsis* chloroplast genome were used as probes in colony hybridization experiments to identify clones

containing chloroplast DNA. One-hundred-and-forty-six or 6.7% of the YAC clones in the EW library hybridized to these probes (Figure 1). The coordinates of these clones are given in Table 1. The majority of the YAC clones hybridized to both of the chloroplast-specific probes, but 27% hybridized to only one.

The results of the colony hybridization experiments were confirmed by Southern blot analysis. In all but one case, the restriction fragments that hybridized to the chloroplast probes were the same size as fragments from chloroplast DNA. This supports the conclusion that the YAC clones did contain chloroplast DNA. One of the 146 YAC clones (EW10H10, data not shown) showed different sized fragments, suggesting it may contain chloroplast-DNA-related sequences encoded by the nuclear genome.

The two chloroplast DNA probes were also used in yeast-colony hybridization experiments with the EG

Table 1. YAC clones hybridizing to repeated DNA sequences

EW YAC clones containing chloroplast DNA sequences

1B4, 1F6, 1F7, 2B8, 2C11, 2D4, 2E4, 2G11, 3B4, 3C9, 3E4, 3F4, 3F11, 3G5, 3G10, 3H3, 4A3, 4A4, 4B9, 4C2, 4C6, 4F1, 4F9, 4G5, 4G10, 4H5, 4H8, 5A5, 5A11, 5B3, 5C3, 5C11, 5E6, 5G8, 5H6, 5H11, 6A4, 6A9, 6C4, 6D5, 6E8, 6G7, 7A11, 7B12, 7D2, 7D3, 7D5, 7D6, 7F3, 7G1, 7G8, 7G10, 7H6, 8A4, 8A5, 8E1, 8F1, 8G3, 8G12, 8H6, 9D10, 9E1, 9G11, 9G12, 9H8, 10B7, 10C11, 10E5, 10H5, 10H10, 10H11, 11D11, 11F3, 11G1, 12B10, 13A7, 13B2, 13B8, 13D1, 13D6, 13F7, 13H3, 13H6, 14E10, 14F3, 14G12, 15A5, 15A7, 15B10, 15D9, 15E4, 15E5, 15E7, 15E8, 15F3, 15G2, 15H1, 15H4, 16A3, 16A8, 16B6, 16D9, 16F6, 17C5, 17C9, 17C12, 17F7, 17H4, 18C2, 18C6, 18C12, 18D5, 18E5, 18G2, 18H8, 18H9, 18H10, 18H11, 19A7, 19A8, 19D1, 19G4, 19H8, 20A6, 20C5, 20D5, 20G4, 20G9, 21A3, 21B3, 21E6, 22A7, 22B2, 22B12, 22C4, 22D5, 22E12, 22G3, 22G10, 22G12, 22H9, 23B3, 23B8, 23C5, 23C10, 23F7

EG YAC clones containing chloroplast DNA sequences

1A6, 1A9, 1A10, 1A11, 1A12, 1B3, 1B7, 1B10, 1B11, 1C2, 1C8, 1C11, 1D4, 1E1, 1E2, 1E4, 1F1, 1F2, 1F4, 1F7, 1G5, 1G6, 1H3, 1H8, 1H9, 2A2, 2A8, 2B4, 2B8, 2B10, 2C4, 2C8, 2D5, 2D6, 2D7, 2E5, 2E7, 2E8, 2F1, 2F11, 2G1, 2G2, 2G4, 2G6, 2G7, 2H4, 2H5, 2H6, 3A1, 3A3, 3A4, 3A5, 3A6, 3A7, 3A10, 3A11, 3A12, 3B1, 3B5, 3B10, 3C1, 3C3, 3C7, 3C9, 3C11, 3D4, 3D5, 3D6, 3D8, 3D9, 3D10, 3E2, 3E3, 3E4, 3E5, 3E6, 3E7, 3E9, 3E10, 3E12, 3F1, 3F3, 3F4, 3F5, 3F6, 3F7, 3F11, 3G1, 3G8, 3G12, 3H1, 3H2, 3H5, 3H10, 3H12, 4A5, 4A6, 4A8, 4A9, 4A12, 4B1, 4B2, 4B7, 4C2, 4C12, 4D1, 4D6, 4D7, 4D12, 4E2, 4E3, 4E6, 4E7, 4E8, 4F5, 4F6, 4F7, 4F9, 4G4, 4H3, 4H4, 4H9, 5A1, 5A10, 5B3, 5B6, 5B9, 5C1, 5D1, 5D5, 5D8, 5E1, 5E4, 5E5, 5E12, 5F1, 5F3, 5F12, 5G1, 5G2, 5G3, 5G4, 5G5, 5G7, 5G10, 5H3, 5H5, 5H9, 5H12, 6A10, 6A11, 6D1, 6D2, 6D3, 6D6, 6D12, 6E8, 6F6, 6G2, 6G3, 6G4, 6H2, 6H12, 7A9, 7A11, 7B1, 7B6, 7B8, 7B9, 7B10, 7B11, 7C4, 7C6, 7C9, 7C10, 7D3, 7D5, 7D7, 7D12, 7E1, 7E3, 7F7, 7F12, 7G2, 7G12, 7H2, 7H4, 7H11, 7H12, 8A2, 8B3, 8B5, 8B6, 8B9, 8C6, 8C7, 8D2, 8D7, 8F2, 8F3, 8F6, 8F7, 8F9, 8G1, 8G2, 8G4, 8G7, 8G8, 8G9, 8H1, 9A1, 9A5, 9A6, 9A10, 9A12, 9B4, 9C11, 9D5, 9D8, 9D11, 9E3, 9E4, 9E5, 9E8, 9E12, 9F1, 9F2, 9F4, 9F10, 9F11, 9G1, 9G2, 9H5, 9H10, 9H11, 10A7, 10A8, 10C7, 10C8, 10C9, 10C10, 10C12, 10D1, 10D2, 10D4, 10D6, 10D8, 10E1, 10E6, 10E8, 10E9, 10E11, 10F1, 10F2, 10F3, 10F4, 10F8, 10F9, 10F12, 10G6, 10G8, 10G9, 10G11, 10H6, 10H9, 11A2, 11A3, 11A4, 11A7, 11A10, 11A11, 11A12, 11B3, 11B8, 11B9, 11C1, 11C10, 11D3, 11D5, 11D6, 11D10, 11E2, 11E4, 11E5, 11E7, 11E10, 11F2, 11F4, 11F8, 11F11, 11G11, 11H1, 11H2, 11H4, 11H5, 11H6, 11H8, 12A4, 12A6, 12A9, 12A11, 12B4, 12B7, 12C1, 12C3, 12D7, 12D11, 12E1, 12E5, 12F1, 12F2, 12F3, 12F5, 12F6, 12F7, 12F8, 12G2, 12G4, 12G5, 12G6, 12G11, 12H1, 12H3, 12H6, 12H12, 13A1, 13A11, 13B1, 13B3, 13B10, 13B11, 13C7, 13C11, 13D4, 13D7, 13D9, 13D10, 13D11, 13E3, 13E4, 13E8, 13F9, 13F10, 13G6, 13G8, 13H6, 14A2, 14A10, 14A11, 14B2, 14B4, 14B10, 14C3, 14C7, 14C11, 14D2, 14D3, 14D4, 14D7, 14D9, 14D10, 14E1, 14E8, 14F1, 14F3, 14F9, 14G3, 14G6, 14G7, 14H5, 14H7, 14H8, 14H10, 15A1, 15A2, 15A12, 15B4, 15B8, 15C2, 15C4, 15C7, 15C10, 15D1, 15D6, 15D10, 15D12, 15E1, 15E3, 15E9, 15E12, 15F1, 15F2, 15F3, 15F8, 15G6, 15G9, 15G11, 15H2, 15H3, 15H6, 15H7, 15H10, 16A1, 16A2, 16A6, 16A12, 16B4, 16B7, 16B12, 16C2, 16C11, 16D1, 16D7, 16E1, 16E2, 16E5, 16E6, 16E7, 16F3, 16F4, 16F5, 16F7, 16F9, 16F10, 16G1, 16G3, 16G4, 16G7, 16G9, 16H2, 16H4, 16H11, 17A5, 17A11, 17B6, 17B7, 17B9, 17C7, 17D5, 17D9, 17E3, 17E5, 17E9, 17E10, 17F5, 17F6, 17G10, 17H6, 18A2, 18A4, 18A7, 18A12, 18B1, 18B2, 18B7, 18B9, 18B10, 18B12, 18C2, 18C6, 18D1, 18D2, 18E6, 18E7, 18E12, 18F1, 18F11, 18F12, 18G1, 18G3, 18G6, 18G8, 18G10, 18H4, 18H5, 18H6, 18H7, 18H8, 18H10, 19B4, 19C2, 19C3, 19C4, 19C9, 19D12, 19E5, 19E7, 19E8, 19E9, 19E12, 19F8, 19F9, 19G2, 19G5, 19G8, 19H2, 19H3, 19H8, 19H9, 20A1, 20A8, 20A11, 20B1, 20B2, 20B12, 20C1, 20C6, 20C10, 20D2, 20D4, 20D5, 20D7, 20D10, 20D11, 20D12, 20E4, 20E9, 20F6, 20F9, 20G3, 20G6, 20H12, 21B1, 21B6, 21B10, 21C3, 21C7, 21C12, 21D4, 21D6, 21E4, 21F4, 21F5, 21G7, 21H1, 22A4, 22A6, 22A9, 22B5, 22B8, 22B11, 22B12, 22C4, 22D4, 22D8, 22D9, 22D10, 22E1, 22E2, 22E10, 22E11, 22E12, 22F2, 22F3, 22F10, 22F11, 22G1, 22G2, 22G3, 22G7, 22G12, 23A3, 23A4, 23A6, 23B2, 23B3, 23B4, 23B7, 23C7, 23C9, 23D1, 23D4, 23E3, 23E6, 23E9, 23E11, 23E12, 23F12, 23G1, 23G3, 23G12, 23H2, 23H4, 23H8, 24A8, 24B5, 24B6, 24B7, 24B8, 24B10, 24C1, 24C5, 24C11, 24D1, 24D2, 24D3, 24D4, 24E1, 24E7, 24E10, 24E11, 24F6, 24G4, 24G5, 24H3, 24H4

Table 1. YAC clones hybridizing to repeated DNA sequences (*continued*)*EW YAC clones containing rDNA sequences*

1E2, 1E3, 1B8, 3C12, 3D9, 3G6, 4E4, 5D10, 6C9, 6D2, 6F1, 6F11, 6H2, 8A12, 8F6, 10A3, 10B1, 10C1, 10G7, 10G10, 11B7, 11C8, 11H11, 12A6, 12A8, 12F9, 13D8, 14C8, 14H11, 15G4, 16A12, 16F3, 17D3, 17E8, 18C4, 18D6, 18G1, 19H4, 19H12, 20B2, 20G7, 20H9, 21A11, 21D5, 21F7, 22B5, 22H3, 23A2, 23B11, 23F8, 23H8

EG YAC clones containing rDNA sequences

1A8, 1B3, 1C2, 1C3, 1C12, 1D4, 1E4, 1E7, 1E8, 1E9, 1E10, 1F1, 1F2, 1F8, 1F9, 1G1, 1G2, 1G6, 1H5, 1H6, 1H7, 2A3, 2A7, 2B4, 2B7, 2C3, 2D2, 2D5, 2D6, 2E7, 2E8, 2E9, 2E11, 2F5, 2F6, 2F7, 2F9, 2F11, 2G5, 2G8, 2G9, 2G10, 2H5, 2H12, 3A1, 3A3, 3A6, 3A8, 3B9, 3B10, 3C9, 3D1, 3E1, 3F4, 3F6, 3F8, 3F12, 3G1, 3G4, 3G9, 3H9, 4A11, 4B1, 4B2, 4B3, 4B4, 4B5, 4B12, 4C2, 4C6, 4C11, 4D10, 4D11, 4D12, 4F9, 4G2, 4G9, 4H8, 5B2, 5B5, 5C11, 5D8, 5D11, 5E9, 5F9, 5H12, 6A4, 6A5, 6A7, 6B8, 6B9, 6B11, 6B12, 6C2, 6D1, 6D3, 6D10, 6E4, 6E6, 6F9, 6G5, 6H7, 6H11, 7B1, 7B8, 7C9, 7D1, 7D10, 7G2, 7G6, 7G12, 7H2, 7H4, 7H11, 7H12, 8B4, 8B12, 8C7, 8C9, 8D3, 8E6, 8E7, 8E8, 8F2, 8F7, 8F8, 8F10, 8F12, 8G2, 8H6, 8H7, 8H9, 8H10, 8H11, 9A7, 9A8, 9B5, 9C12, 9D4, 9E2, 9E4, 9F1, 9G6, 9H6, 9H11, 10A2, 10A7, 10A8, 10B2, 10C1, 10C2, 10D4, 10D6, 10D7, 10D8, 10E1, 10E10, 10F2, 10G3, 10G8, 10G9, 10H4, 10H12, 11B3, 11B4, 11B5, 11B12, 11C2, 11D1, 11D2, 11D4, 11D5, 11D6, 11D7, 11D9, 11E1, 11E5, 11E6, 11G8, 11H9, 12A4, 12A7, 12A10, 12B4, 12B9, 12C3, 12C7, 12E1, 12E3, 12F1, 12F4, 12F7, 12G8, 12G12, 13B10, 13D5, 13D6, 13F1, 13F11, 13H9, 13H11, 13H12, 14A10, 14A11, 14B6, 14B10, 14C3, 14C4, 14C8, 14F9, 14G2, 14G5, 14G12, 14H8, 15A3, 15A11, 15B1, 15B3, 15B5, 15B11, 15C2, 15C6, 15C7, 15C10, 15D1, 15E1, 15E3, 15F3, 15F8, 15G6, 15G7, 15G11, 15H3, 15H7, 16A1, 16B7, 16C2, 16C10, 16D10, 16E1, 16E7, 16E8, 16E9, 16E12, 16F11, 16F12, 16G2, 16G7, 16H4, 16H11, 17A3, 17A11, 17B3, 17D3, 17E9, 17G2, 17G10, 17H7, 18A2, 18A12, 18B2, 18B4, 18B8, 18C2, 18D2, 18D4, 18D5, 18D7, 18D12, 18E2, 18G6, 18G8, 18H5, 18H6, 18H7, 19A1, 19A5, 19B4, 19C1, 19D12, 19E10, 19E11, 19G3, 19G5, 19G8, 19H1, 19H3, 20A4, 20B2, 20B4, 20B6, 20B10, 20C8, 20C9, 20E4, 20F2, 20G6, 20G9, 20G12, 20H9, 20H10, 20H12, 21A7, 21C2, 21C10, 21C11, 21D6, 21D12, 21E1, 21E9, 21F7, 21G10, 21H3, 22A1, 22A7, 22B2, 22B11, 22D8, 22E3, 22E11, 22F3, 22F5, 22F6, 22F10, 22F12, 22H4, 22H5, 22H12, 23A3, 23C4, 23C8, 23D5, 23E6, 23F4, 23F5, 23F11, 23F12, 23G4, 23G12, 23H10, 24A3, 24B12, 24C1, 24C12, 24D1, 24D4, 24D8, 24E7, 24E8, 24E10, 24E11, 24F6, 24G1, 24H6

EW YAC clones containing the paracentromeric repeat sequences

1D5, 1D8, 1H3, 2A3, 2A6, 2D9, 2E12, 2F1, 2F12, 3A6, 3B1, 3B5, 3B9, 3C1, 3D8, 3E2, 3F1, 3F5, 3F6, 3F12, 3G4, 3G8, 3G11, 3H11, 4B2, 4B12, 4E11, 4H2, 5C2, 5D3, 5F3, 5F4, 5F9, 5H1, 6B4, 6B10, 6C1, 6C6, 6C10, 6C12, 6E6, 6G1, 6G6, 6G12, 6H12, 7A4, 7B10, 7B4, 7B11, 7D8, 7E11, 7F6, 8A7, 8B3, 8B11, 8B12, 8C1, 8F3, 8F11, 8G1, 8G4, 8H9, 9A5, 9A7, 9D1, 9D8, 9D11, 9E6, 9E11, 9G1, 9G9, 9H11, 10B11, 10D12, 10E2, 10E8, 10H8, 11A3, 11B10, 11D4, 11E11, 11F5, 11H6, 11H7, 12A3, 12A5, 12A10, 12B6, 12C2, 12D7, 12F7, 12F12, 12G3, 12G8, 12H3, 12H7, 13A10, 13B5, 13B7, 13B11, 13F10, 13G2, 13G12, 13H1, 13H2, 14A10, 14B1, 14C1, 14C5, 14C7, 14C10, 14D9, 14E2, 14F7, 15B8, 15C2, 15C5, 15C9, 16B5, 16D7, 16D9, 16H5, 17A6, 17C4, 17D2, 17D10, 17D11, 17E4, 17E10, 17F12, 17G1, 17G2, 17H1, 17H10, 18A2, 18B4, 18B9, 18C5, 18C11, 18D1, 18D9, 18F4, 18F9, 18G4, 18H6, 19A2, 19B4, 19E6, 19F1, 19G2, 19G8, 19G12, 20A9, 20B2, 20B5, 20C1, 20D10, 20D11, 20E1, 20E2, 20F5, 21A10, 21B6, 21B8, 21C10, 21D2, 21D3, 21D6, 21E12, 21F3, 21F6, 21F9, 21G1, 21H5, 22A5, 22B8, 22C1, 22E1, 22H7, 22H10, 23E3, 23E7, 23G1, 23G5, 23H9, 23H12

EG YAC clones containing the paracentromeric repeat sequences

2A11, 2B6, 2C6, 2D5, 2E6, 5F4, 6A8, 7E10, 8C9, 8F11, 9A4, 9B3, 10D12, 10F5, 10F10, 10H10, 11B11, 12C9, 12E5, 12G1, 12G6, 12H2, 14A8, 14G10, 16E6, 17C8, 17C9, 17G2, 18C12, 18D5, 19E6, 20D7, 20E12, 23A12, 23B2, 23H5, 23H6

library. Six-hundred-and-six (26%) YAC clones hybridized to these probes (Figure 1; Table 1). Fifty-one per cent of the clones hybridized to one or other of the two chloroplast probes, 49% to both. The positive score in the colony hybridization experiments was confirmed for a subset of 95 of the EG YAC clones using Southern blot analysis.

Identification of YAC clones containing rDNA sequences

Ribosomal DNA gene sequences are highly conserved between species. Unlike the rDNA genes, rDNA intergenic spacers are highly variable and not conserved between species (Rogers and Bendich, 1987). A probe derived from the intergenic spacer DNA of *Arabidopsis* was used to identify YAC clones containing *Arabidopsis* rDNA sequences in yeast-colony hybridizations of the *Arabidopsis* YAC libraries.

The rDNA intergenic spacer probe hybridized to 51 EW YACs (2.3%) and 354 EG YACs (15%) in yeast-colony hybridizations (Figure 1; Table 1). In all 13 cases tested, YAC clones initially selected by use of the intergenic spacer probe also hybridized to rDNA gene sequences. This result indicated that both rDNA gene and spacer sequences were present in the YAC clones (data not shown). The colony hybridization results were confirmed by Southern blot analysis for all the coordinates of the EW library and for a subset of 71 EG YAC clones.

Identification of YAC clones containing the paracentromeric heterochromatin HindIII repeat sequences

The YAC libraries were screened by colony hybridization with the 180 bp, and the closely related 500 bp, *HindIII* paracentromeric heterochromatin repeat sequences

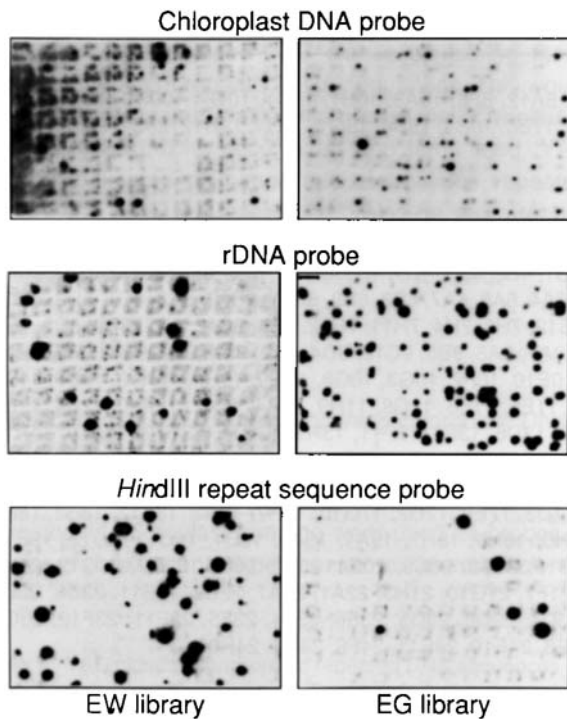


Figure 1. Uneven representation of repeated sequences in YAC libraries. Colonies (768) from the EW and the EG libraries were arrayed on single filters. The results of colony screenings using chloroplast, rDNA and *Hind*III repeat probes are shown. The repeated sequences are unevenly represented in the two YAC libraries. Chloroplast DNA and rDNA sequences are much more frequent in the EG library than in the EW library, the contrary is true for the *Hind*III repeat sequence probe.

(Simoens *et al.*, 1988). Both probes revealed a very similar hybridization pattern (Figure 1; and data not shown). In total these probes hybridized to 186 (8.5%) of the EW YAC clones and 37 (1.6%) of the EG YAC clones (Table 1). Positively hybridizing YAC clones were subjected to Southern blot hybridization experiments which confirmed that the majority of the YAC clones contained sequences homologous to both the 180 bp and the closely related 500 bp sequence repeat (data not shown).

Identification of YAC clones containing chimaeric inserts

Extensive chromosome walking experiments (Putterill *et al.*, 1993; Schmidt *et al.*, 1992) indicated the existence of EG YAC clones with chimaeric inserts of unique DNA combined with either chloroplast or rDNA sequences. One (EG1C12RE) of the 75 end probes derived from 53 YAC clones hybridized to rDNA sequences and four (EG1B10RE, EG4E2RE, EG15C10LE, EG19B1RE) hybridized to YAC clones containing chloroplast DNA. Clones EG1B10, EG4E2 and EG15C10 hybridized to the cosmids which were used to identify YAC clones containing chloroplast DNA sequences (Table 1). Clone

EG19B1, however, did not. It is an example of a YAC clone containing a part of chloroplast DNA not spanned by the two cosmids. All five YAC clones had already been shown to contain single-copy nuclear DNA. EG15C10 hybridized to RFLP marker m326, EG1B10 to RFLP marker g10086 and EG4E2 to RFLP marker m217 (Hwang *et al.*, 1991). YAC clones EG19B1 and EG1C12 are part of a large YAC contig around the chalcone synthase gene (Putterill *et al.*, 1993). Therefore, at least five (9%) of the 53 YAC clones analyzed in these contigs are not collinear with chromosomal DNA, but are chimaeric between single-copy nuclear DNA and repetitive DNA sequences.

The same kind of analysis of YAC end probes was performed on EW YAC clones. None of the 47 EW YAC end probes derived from 35 YAC clones hybridized to rDNA or to chloroplast sequences. This suggests that EW YAC chimaeras between single-copy nuclear DNA sequences and the analysed repeated DNA sequences are uncommon.

To obtain more accurate frequency estimates for the presence of chimaeric YAC clones, we compared the coordinates of the YAC clones which hybridized to chloroplast DNA with the coordinates of clones which had been mapped previously on to chromosome 4 and the top half of chromosome 5 using RFLP markers as probes (Hwang *et al.*, 1991; Schmidt *et al.*, unpublished results). One-hundred-and-ninety-five clones of the EG library were previously identified as hybridizing to RFLP markers mapping to chromosome 4 and the top half of chromosome 5 (data not shown). These represent 155 independent clones; the remaining 40 were identical to clones in adjacent well positions and are likely to be duplicate clones caused by spill-over from the neighbouring well. Southern blot analysis of YACs separated by pulsed field gel electrophoresis established that 26 (17%) of these 155 clones are chimaeric between chloroplast DNA and single-copy nuclear DNA (Table 2, Table 3a). Clones EG1F4 and EG20G6 also hybridized to chloroplast DNA as well as to RFLP marker DNA, however, YAC bands of different sizes hybridized to the RFLP marker DNA and the chloroplast DNA, respectively. Thus, multiple YACs rather than chimaeric clones explain the hybridization results in these two cases.

The 155 clones from chromosomes 4 or 5 were then compared with the coordinates of clones corresponding to rDNA or the *Hind*III repeat sequences. This analysis identified 19 clones corresponding to rDNA as well as to single-copy nuclear DNA, and two clones hybridizing to the *Hind*III repeat sequences as well as to single-copy nuclear DNA (Table 2, Table 3a). Although some of these clones are clearly chimaeric (e.g. EG7G2, EG15C10), others could represent the cloning of single-copy

Table 2. YAC clones carrying single-copy nuclear sequences as well as repetitive DNA sequences

YAC coordinate	Chromosome ^a	Hybridization to repetitive DNA sequences
EG1B10	4	chloroplast DNA
EG2G4 / H4 / H6	4	chloroplast DNA
EG3A10	5	chloroplast DNA
EG4D1 / E2	5	chloroplast DNA
EG5C1	4	chloroplast DNA
EG7B11	4	chloroplast DNA
EG7G2 / H2	5	chloroplast DNA / rDNA
EG8F7		chloroplast DNA / rDNA
EG8G8	4	chloroplast DNA
EG9F2	5	chloroplast DNA
EG10C12	4	chloroplast DNA
EG10F2	4	chloroplast DNA / rDNA
EG10F12	5	chloroplast DNA
EG10G6 / H6	4	chloroplast DNA
EG12F8	5	chloroplast DNA
EG12H12	5	chloroplast DNA
EG15C10	4	chloroplast DNA / rDNA
EG15E1	4	chloroplast DNA / rDNA
EG16C2	5	chloroplast DNA / rDNA
EG17A11	4	chloroplast DNA / rDNA
EG18A7	5	chloroplast DNA
EG18G6 / H5 / H6 / H7	5	chloroplast DNA / rDNA
EG19H3	4	chloroplast DNA / rDNA
EG21B6	5	chloroplast DNA
EG23A3	4	chloroplast DNA / rDNA
EG24C1 / D1	4	chloroplast DNA / rDNA
EG1C12	5	rDNA
EG3F12	4	rDNA
EG11B5	5	rDNA ^b
EG11C2	4	rDNA
EG11G8	4	rDNA
EG14G5	4	rDNA
EG15A11	5	rDNA
EG18B4	4	rDNA
EG17C8 / C9	4	<i>Hind</i> III repeat sequences
EG23H5	5	<i>Hind</i> III repeat sequences

^a Refers to chromosome to which the RFLP marker, used to identify the YAC clone, has previously been mapped.

^b Not confirmed in Southern blot experiments.

sequences which are genuinely located adjacent to repetitive DNA sequences in the *Arabidopsis* genome.

YAC clones hybridizing to chloroplast sequences were compared with those hybridizing to rDNA sequences and the *Hind*III repeat sequences. One-hundred-and-forty clones hybridize to more than one of the repetitive DNA sequences (Table 3b). One clone (EG2D5) showed hybridization to all three repeated DNA sequences. One-hundred-and-thirty-seven clones hybridized to chloroplast DNA and rDNA or the paracentromeric repeat sequences. Thus, up to 23% of the EG YAC clones carrying chloroplast DNA could be chimaeric with other classes of repeated DNA sequences.

A Southern blot analysis on YACs separated by pulsed field gel electrophoresis was carried out on a subset of these 137 clones to distinguish if the clones were chimaeric or if the colony hybridization data resulted from multiple YACs. Repetitive sequences as probes confirmed that YACs EG12E5, EG12G6, EG16E6, EG20D7 and EG23B2 are chimaeric between chloroplast DNA sequences and the *Hind*III repeat sequences. A similar analysis revealed that 29 of the 30 clones tested are chimaeric between rDNA and chloroplast DNA sequences (EG10A7, EG10E1, EG10G8, EG11B3, EG11D5, EG12A4, EG12F1, EG13B10, EG14F9, EG14H8, EG15C2, EG15C7, EG15E3, EG15F8, EG15G6, EG15G11, EG15H7, EG20G6 and YAC coordinates in Table 2). Clone EG12E1, however, contained two YACs. The band of approximately 145 kb in size hybridized to chloroplast DNA whereas the 90 kb band hybridized to rDNA, hence this clone is not chimaeric. Therefore, most of the EG YAC clones hybridizing to chloroplast DNA sequences and rDNA sequences or the *Hind*III repeat sequences are genuine chimaeric clones.

A comparison of the 146 EW YAC clones containing chloroplast DNA with 144 EW YAC clones which have been mapped on to chromosome 4 and the top half of chromosome 5 using RFLP markers (Schmidt *et al.*, unpublished results) revealed one YAC clone in common (Table 3a). This was the clone, however, previously identified as possibly containing sequences cross-hybridizing to chloroplast DNA (EW10H10, see above) and is not classified as a chimaeric clone. A comparison of all EW clones hybridizing to chloroplast DNA with the clones containing rDNA and the paracentromeric repeat showed one (0.7%) chimaeric clone (EW16D9 hybridizes to the *Hind*III repeat sequence as well as to chloroplast DNA, Table 3b). Therefore it can be concluded, that chimaeric EW clones containing the repetitive DNA sequences that we have so far analysed are uncommon.

Instability of YAC clones carrying tandemly repeated DNA

All the EW YACs containing rDNA sequences were fractionated by PFGE and were found to be uniformly small in size at ~ 50 kb. However, 52 EG YAC clones containing these sequences were tested and ranged in size from 80 to 250 kb. Scrutiny of the gels stained with ethidium bromide, showed that subsets of both groups of clones gave two or three YAC bands. Multiple YACs can also be observed in the autoradiogram of EG YACs in Figure 2.

PFGE Southern blot analysis of YAC clones containing the tandem *Hind*III repeat sequences also showed that there was a high incidence of multiple YAC bands in each colony. The multiple bands hybridized to YAC vector

Table 3. Comparative analysis of the EG and the EW library

	YAC clones hybridizing in the	
	EG library	EW library
<i>(a) YAC clones hybridizing to single-copy nuclear sequences as well as repeated DNA sequences</i>		
Chromosome 4 / 5 ^a	155	144
Chromosome 4 / 5 ^a as well as chloroplast DNA	26 (17%)	1 (0.7%)
Chromosome 4 / 5 ^a as well as rDNA	19 (12%)	0
Chromosome 4 / 5 ^a as well as <i>Hind</i> III repeats	2 (1.3%)	0
Chromosome 4 / 5 ^a as well as repeated DNA sequences ^b	36 (23%)	0
<i>(b) YAC clones hybridizing to repeated DNA sequences</i>		
Chloroplast DNA	606 (26%)	146 (6.7%)
rDNA	354 (15%)	51 (2.3%)
<i>Hind</i> III repeats	37 (1.6%)	186 (8.5%)
Chloroplast DNA as well as rDNA	132 (5.7%)	0
Chloroplast DNA as well as <i>Hind</i> III repeats	6 (0.3%)	1 (0.05%)
rDNA as well as <i>Hind</i> III repeats	4 (0.2%)	1 (0.05%)
Repeated DNA sequences ^c	857 (37%)	381 (17%)

^aYAC clones corresponding to RFLP markers which have previously been mapped to chromosome 4 or the top half of chromosome 5 (Schmidt *et al.*, unpublished results).

^bRefers to the three classes of repeated DNA sequences analysed: chloroplast DNA, rDNA and *Hind*III repeat sequences.

^c Refers to the three classes of repeated DNA sequences analysed: chloroplast DNA, rDNA and *Hind*III repeat sequences.

sequences as well as the repeated sequences (Figure 3). Both the EG and the EW YACs ranged in size from 50 to 250 kb. From 146 YAC clones analysed, 46 (32%) were shown to contain multiple YACs. Six of the YAC clones which contained multiple YAC bands were subjected to colony purification and intact chromosomal DNA was subsequently prepared from two to three single colonies. YACs of different sizes were still present in these preparations indicating that the original preparation was not derived from a mixture of clones. The relative intensity of the bands as well as the pattern varied between single colonies derived from the same clone (data not shown). This result suggests that a high percentage of YACs carrying the paracentromeric repeat sequences are unstable. However, despite this, all the clones do maintain some copies of the paracentromeric *Hind*III repeat sequences.

Discussion

The widespread adoption of *Arabidopsis* to study many plant processes using molecular genetic techniques has resulted in the extensive use of two of the published *Arabidopsis* YAC libraries (Grill and Somerville, 1991; Ward and Jen, 1990) for chromosome walking. These are also being used to construct YAC contigs around

numerous markers for the five *Arabidopsis* chromosomes (Hwang *et al.*, 1991). We have analysed the distribution and stability of three classes of repeated DNA sequences in YAC clones derived from these libraries. The coordinates of YAC clones carrying the paracentromeric *Hind*III repeat, ribosomal DNA and chloroplast DNA sequences are presented (Table 1).

The repeated DNA sequences analysed are differentially represented in the two YAC libraries. The three classes of repeated sequences, *Hind*III repeat, ribosomal and chloroplast sequences comprise 8.5%, 2.3% and 6.7%, respectively, of the EW library, totalling approximately 17% of clones, and 1.6%, 15% and 26%, respectively, of the EG library, totalling 37% of the clones (Table 3b). The paracentromeric repeat sequences account for 1–1.6% and the rDNA 6–8% of the haploid nuclear genome. A comparison of these figures with the representation of these sequences in the YAC libraries clearly indicates a bias in the cloning and/or the stability of these sequences within the libraries.

A much higher percentage of the EW YAC clones than the EG YAC clones carry the *Hind*III repeat. The tandemly repeated 180 bp *Hind*III sequences are arranged in arrays longer than 50 kb that lack *Bam*HI sites (Martinez-Zapater *et al.*, 1986). Most repeat arrays were therefore probably cloned as single units in the EG

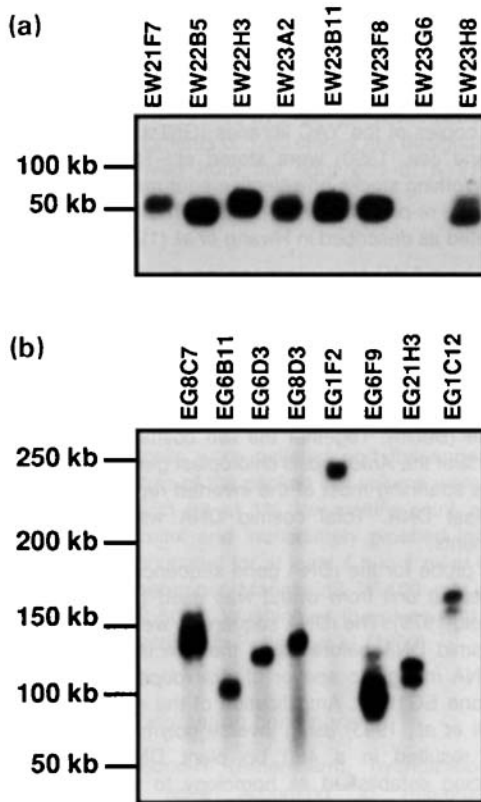


Figure 2. YAC clones containing rDNA show marked size differences between libraries.

YAC clones hybridizing to the *Arabidopsis* rDNA probe in colony hybridization experiments were fractionated by PFGE, gel blotted and hybridized to the rDNA probe. (a) and (b) Show the results for the EW and EG YAC clones, respectively. EW23G6 is a negative control. All EW YACs containing rDNA sequences are uniformly small in size, while the sizes of the EG YACs range from 80 to 250 kb. Multiple YAC bands are observed indicating instability of the clones (for example, EG21H3 and EG1C12).

library which was constructed using *Bam*HI partially digested DNA. The random shearing process which was used to fractionate the DNA during the construction of the EW library, could, however, have broken up the *Hind*III repeat arrays. This probably explains why these sequences are present in a higher percentage of clones in the EW than in the EG library. Differential representation of clones within the libraries was also found with single-copy probe hybridizations (Hwang *et al.*, 1991). These observations emphasize the desirability of searching for sequences of interest in multiple YAC libraries, preferably constructed using DNA cleaved in different ways.

The frequency of YAC clones carrying chimaeric inserts varies greatly in different YAC libraries and was 40–60% in one of the human YAC libraries (Green *et al.*, 1991). The occurrence of chimaeric YAC clones greatly reduces the efficiency of chromosome walking

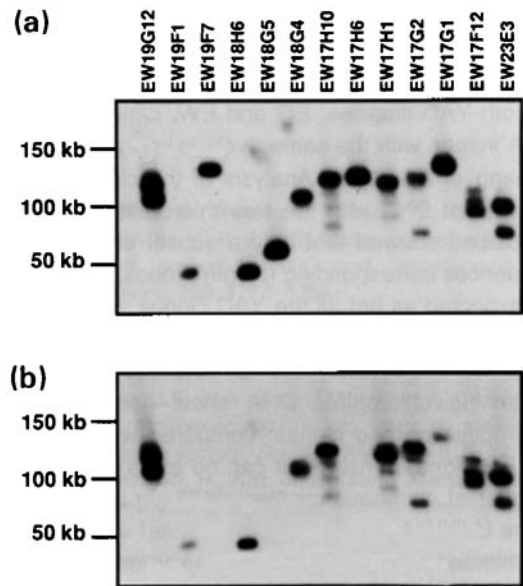


Figure 3. YAC clones carrying a paracentromeric repeat sequence show a high degree of instability.

Southern blot analysis of a PFGE gel of 13 YAC clones derived from the EW YAC library, 10 of which were previously identified to contain the *Hind*III repeat sequence in colony hybridization experiments. EW19F7, EW18G5 and EW17H6 are negative controls. YAC clones were probed with pYAC vector sequences in (a) and with the 180 bp *Hind*III repeat in (b). All YAC clones hybridize to the pYAC vector sequence. The positive hybridization of the 10 clones in colony blot experiments with the *Hind*III repeat as a probe is confirmed in the Southern blot experiment. Six of the clones containing the *Hind*III repeat sequence (EW19G12, EW17H10, EW17H1, EW17G2, EW17F12 and EW23E3) show multiple bands indicating that these YAC clones are unstable.

experiments and hampers the construction of physical maps. Chloroplast DNA sequences and the rDNA sequences are much more prevalent in the EG library than in the EW library. This could reflect the higher number of YAC clones containing chimaeric inserts between single-copy nuclear and repeated sequences in the EG library. Different analyses were carried out to identify the frequency of chimaeric YACs among the clones carrying repeated DNA sequences. Chromosome walking experiments showed that at least 9% of the EG clones analysed contain unique and repeated DNA sequences from different locations in the *Arabidopsis* genome. Likewise, the comparison of EG YAC clones identified by RFLP markers with clones hybridizing to chloroplast DNA revealed a high percentage of chimaeric clones (17%, Table 3a). Finally, a comparison of coordinates corresponding to chloroplast DNA with clones hybridizing to rDNA or the *Hind*III repeat sequences revealed that 23% of the chloroplast-containing EG YAC clones (this figure corresponds to 6% of all EG YAC clones) are chimaeric with unlinked repetitive sequences (Table 3b). Similar analyses with

EW YAC clones showed that the proportion of chimaeric YAC clones containing the analysed repeated DNA sequences is well below 1% (Table 3).

Both YAC libraries, EG and EW, contain *Arabidopsis* DNA inserts with the same average insert size of 160 kb (Hwang *et al.*, 1991). Analysis of the clones containing chloroplast DNA with the two chloroplast DNA cosmids as probes showed that only a subset of clones contain sequences corresponding to both probes. This result is to be expected as not all the YAC clones will contain a full complement of the 155 kb *Arabidopsis* chloroplast genome. However, 73% of the EW YAC clones containing chloroplast DNA show homology to both chloroplast cosmid clones; compared with only 49% of the EG clones. This result can be explained by the high number of chimaeric clones containing chloroplast DNA in the EG library.

Chimaeric YAC clones carrying unique DNA derived from different areas of the genome have been detected in the EG as well as in the EW library. In contrast to this result, a high frequency of chimaeric clones containing repetitive DNA sequences is observed only in the EG library (Table 3). The identification of the YAC coordinates hybridizing to the repeated DNA sequences should be valuable for determining when chromosome walks reach problematic YAC clones. As all libraries tested to date contain unreliable YAC clones, it is an advantage to work with multiple libraries.

A high proportion of YAC clones containing the rDNA or *Hind*III paracentromeric repeats were observed to be unstable. All EW YAC clones harbouring ribosomal DNA were uniformly small in size; 50 kb as opposed to an average clone size of 160 kb. We attribute the small size of the EW YACs containing rDNA sequences to the profound instability of these cloned tandemly repeated sequences in YACs. We presume that the EG library also contains small YACs with rDNA inserts, but that our test sample size of 52 EG YAC clones was not large enough to see this subgroup. The normal range of sizes shown in the EG YACs could be explained if the great majority of EG YAC clones that hybridize to rDNA probes are chimaeric between rDNA and single-copy DNA. These chimaeric YAC clones may only contain a low number of copies of the ~9.9 kb rDNA repeat sequence (Pruitt and Meyerowitz, 1986) and thus not exhibit as much instability as the EW YAC clones containing only tandemly repeated rDNA. Approximately 32% of the clones containing the *Hind*III repeat sequences were shown to contain multiple YACs. Similar instability due to the presence of repeated sequences was previously observed in YAC clones (Dunford *et al.*, 1993; Neil *et al.*, 1990). Regions of highly repeated sequences will therefore need special attention during the completion of a physical map of *Arabidopsis*.

Experimental procedures

YAC libraries

Master copies of the YAC libraries (Grill and Somerville, 1991; Ward and Jen, 1990) were stored at -70°C (Schmidt *et al.*, 1992). Working stocks on selective agar medium (Putterill *et al.*, 1993) were re-plated every 3 months. The YAC coordinates are designated as described in Hwang *et al.* (1991).

DNA probes

Two non-overlapping cosmids containing *Arabidopsis* chloroplast DNA were obtained from A. Brennicke and W. Schuster (Berlin). Together the two cosmids account for more than 80% of the *Arabidopsis* chloroplast genome, with one of the cosmids spanning most of the inverted repeat sequence of the chloroplast DNA. Total cosmid DNA was used in labelling experiments.

As a probe for the rDNA gene sequences, a 9 kb, 25S–18S rDNA repeat unit from wheat was used (pTA71, Gerlach and Bedbrook, 1979). The rDNA sequences were purified away from the plasmid DNA before using them in labelling experiments. The rDNA intergenic spacer of *Arabidopsis* was derived from YAC clone EG1C12. Amplification of the right end of the YAC (Putterill *et al.*, 1993) using inverse polymerase chain reaction (IPCR) resulted in a 480 bp plant DNA fragment. Direct sequencing established its homology to the rDNA intergenic spacer DNA of *Arabidopsis* (Gruendler *et al.*, 1989). Labelling of the PCR product was carried out after removal of all pYAC vector sequences.

The 180 and 500 bp *Hind*III repeat sequences located in the paracentromeric heterochromatin were obtained as pGEM2 subclones (Simoens *et al.*, 1988). The plant DNA fragments were separated from the vector sequences before using them in labelling experiments.

The pYAC vector sequences (Burke *et al.*, 1987) were used as probes to detect YACs separated from the yeast chromosomes by PFGE.

Yeast colony hybridization

Preparation of the yeast colony filters was as outlined by Coulson *et al.* (1988) with the modifications detailed in Schmidt *et al.* (1992). Probe labelling, hybridization and washing conditions were the same as described in Schmidt *et al.* (1992).

Yeast genomic DNA for restriction enzyme digestion and YAC end fragment isolation

Yeast colonies were removed from agar plates and resuspended in 300 μl TE/SDS (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, 0.1% SDS). The cells were incubated for 15–30 mins at 65°C before they were cooled on ice. Precooled 3 M Acetate (150 μl , pH 4.8) was added and the preparations were incubated for 15–30 min. After centrifugation, the supernatant was transferred, extracted with phenol/chloroform/isoamylalcohol and precipitated with ethanol. The DNA pellet was resuspended in TE (10 mM Tris–HCl, pH 8.0, 1 mM EDTA) and used for restriction enzyme digestion.

Isolation of YAC end fragments

Isolation of YAC end fragments by IPCR was carried out as described in Schmidt *et al.* (1992) and Putterill *et al.* (1993). YAC left end fragments were isolated by plasmid rescue (Schmidt *et al.*, 1992). Left end fragments of YAC clones are designated LE, while sequences derived from the right end of a YAC are designated RE.

Preparation of intact chromosomal yeast DNA and fractionation by pulsed field gel electrophoresis (PFGE)

Selective medium (10 ml) (Putterill *et al.*, 1993) was inoculated with a single yeast colony and cultured at 30°C for 18–24 h. Yeast cells were harvested by centrifugation and then resuspended in 1 ml 1 M sorbitol. After transfer to a micro-centrifuge tube, the cells were pelleted and the supernatant removed. Fifty microlitres of the packed cell volume were mixed briefly by vortexing with 80 µl 1% low-melting-point agarose made up in 1 M sorbitol and immediately pipetted into plug moulds. Plugs were incubated for at least 4 h in 1 ml of 0.5 mg ml⁻¹ ProteinaseK (Boehringer Mannheim) in 100 mM EDTA, pH 8, 1% Sarkosyl at 50°C. Washing of the plugs and PFGE was carried out as described in Putterill *et al.* (1993).

Gel blotting and hybridization conditions

Gel transfer to Hybond-N (Amersham), hybridization and washing conditions were according to manufacturer's instructions except that DNA was UV nicked prior to transfer (UV Stratalinker, 4000 µJ × 100; Stratagene) and the DNA was fixed to the membranes by UV Stratalinker treatment (2400 µJ × 100) and/or baked at 80°C for 2 h. DNA fragments and cosmid DNA were labelled using random primer extension (Feinberg and Vogelstein, 1983).

Southern blot analysis of YAC clones using repeated sequences as probes

The analysis of YAC clones containing chloroplast DNA was performed using Southern blots of yeast genomic DNA digested with *EcoRI* and *BamHI*, while the DNA of YAC clones carrying the paracentromeric repeat sequences was digested with *HindIII*. To analyse YAC clones containing rDNA sequences, the YACs were separated from yeast chromosomes by PFGE. Blots of these gels were then hybridized to rDNA gene or intergenic spacer sequences.

Sizing of YACs and analysis of their stability was performed on intact chromosomes separated by PFGE using concatemers of λ DNA as a size standard (Bancroft and Wolk, 1988). Ethidium bromide staining was used to visualize the YACs and Southern blot analysis was carried out using pYAC vector (Burke *et al.*, 1987) and repeated DNA sequences as probes.

Acknowledgements

We gratefully acknowledge A. Brennicke and W. Schuster (Berlin) for kindly providing cosmid clones and restriction mapping data of the *Arabidopsis* chloroplast genome. The plasmid pTA71 was a gift from T. Schwarzacher (Norwich). We wish to thank David Flanders and Anuj Bhatt for critically reading the manuscript. This work was supported by grants to C. Dean

and G. Coupland from the AFRC PMB programme, EC training fellowships to R. Schmidt and G. Cnops and an EC Bridge grant to C. Dean.

References

- Aronel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H.M. and Somerville, C.R. (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science*, **258**, 1353–1354.
- Bancroft, I. and Wolk, C.P. (1988) Pulsed homogeneous orthogonal field gel electrophoresis (PHOGE). *Nucl. Acids Res.* **16**, 7405–7418.
- Burke, D.T., Carle, G.F. and Olson, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*, **236**, 806–812.
- Coulson, A., Waterston, R., Kiff, J., Sulston, J. and Kohara, Y. (1988) Genome linking with yeast artificial chromosomes. *Nature*, **335**, 184–186.
- Dunford, R., Vilageliu, L.I. and Moore, G. (1993) Stabilisation of a yeast artificial chromosome containing plant DNA using a recombination-deficient host. *Plant Mol. Biol.* **21**, 1187–1189.
- Ecker, J.R. (1990) PFGE and YAC analysis of the *Arabidopsis* genome. *Methods*, **1**, 186–194.
- Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6–13; Addendum (1984) **137**, 266–267.
- Gerlach, W.L. and Bedbrook, J.R. (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucl. Acids Res.* **7**, 1869–1885.
- Giraudat, J., Hauge, B.M., Valon, C., Smalle, J., Parcy, F. and Goodman, H.M. (1992) Isolation of the *Arabidopsis* ABI3 gene by positional cloning. *Plant Cell*, **4**, 1251–1261.
- Green, E.D., Riethman, H.C., Dutchik, J.E. and Olson, M.V. (1991) Detection and characterization of chimeric yeast artificial-chromosome clones. *Genomics*, **11**, 658–669.
- Grill, E. and Somerville, C. (1991) Construction and characterization of a yeast artificial chromosome library of *Arabidopsis* which is suitable for chromosome walking. *Mol. Gen. Genet.* **226**, 484–490.
- Gruendler, P., Unfried, I., Pointner, R. and Schweizer, D. (1989) Nucleotide sequence of the 25S–18S ribosomal gene spacer from *Arabidopsis thaliana*. *Nucl. Acids Res.* **17**, 6395–6396.
- Hauge, B.M., Hanley, S., Giraudat, J. and Goodman, H.M. (1991) Mapping the *Arabidopsis* genome. In *Molecular Biology of Plant Development* (Jenkins, G.I. and Schuch, W., eds). Cambridge: The Company of Biologists, pp. 45–56.
- Hwang, I., Kohchi, T., Hauge, B.M., *et al.* (1991) Identification and map position of YAC clones comprising one-third of the *Arabidopsis* genome. *Plant J.* **1**, 367–374.
- Maluszynska, J. and Heslop-Harrison, J.S. (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. *Plant J.* **1**, 159–166.
- Martinez-Zapater, J.M., Estelle, M.A. and Somerville, C.R. (1986) A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **204**, 417–423.
- Neil, D.L., Villasante, A., Fisher, R.B., Vetrie, D., Cox, B. and Tyler-Smith, C. (1990) Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucl. Acids Res.* **18**, 1421–1428.
- Pruitt, R.E. and Meyerowitz, E.M. (1986) Characterization of

- the genome of *Arabidopsis thaliana*. *J. Mol. Biol.* **187**, 169–184.
- Putterill, J., Robson, F., Lee, K. and Coupland, G.** (1993) Chromosome walking with YAC clones in *Arabidopsis*: isolation of 1700 kb of contiguous DNA on chromosome 5, including a 300 kb region containing the flowering-time gene *CO*. *Mol. Gen. Genet.* **239**, 145–157.
- Rogers, S.O. and Bendich, A.J.** (1987) Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Mol. Biol.* **9**, 509–520.
- Schmidt, R. and Dean, C.** (1992) Physical mapping of the *Arabidopsis thaliana* genome. In *Genome Analysis*, Volume 4, *Strategies for Physical Mapping* (Davies, K.E. and Tilghman, S.M., eds). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 71–98.
- Schmidt, R., Cnops, G., Bancroft, I. and Dean, C.** (1992) Construction of an overlapping YAC library of the *Arabidopsis thaliana* genome. *Aust. J. Plant Physiol.* **19**, 341–351.
- Simoens, C.R., Gielen, J., Van Montagu, M. and Inzé, D.** (1988) Characterization of highly repetitive sequences of *Arabidopsis thaliana*. *Nucl. Acids Res.* **16**, 6753–6766.
- Timmis, J.N. and Steele Scott, N.** (1985) Movement of genetic information between the chloroplast and the nucleus. In *Genetic Flux in Plants* (Hohn, B. and Dennis, E.S., eds). New York: Springer-Verlag, pp. 61–78.
- Ward, E.R. and Jen, G.C.** (1990) Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly-sheared *Arabidopsis thaliana* DNA. *Plant Mol. Biol.* **14**, 561–568.