© 1995 Oxford University Press Nucleic Acids Research, 1995, Vol. 23, No. 19 **4013–4014**

A rapid procedure for the compositional analysis of yeast artificial chromosomes

Albertina De Sario+, Eva-Maria Geigl and Giorgio Bernardi*

Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France

Received May 10, 1995; Revised and Accepted August 18, 1995

The genomes of warm-blooded vertebrates are made up of large DNA segments, the isochores, which are compositionally uniform, but belong to a small number of families that cover a wide GC range (1,2). Hybridization of gene probes on compositional fractions led to the discovery of a strikingly non-uniform gene distribution in the human genome (1). Indeed, the concentration of genes is low in GC-poor isochores and increases in GC-rich isochores to reach the highest values in the GC-richest isochores (3,4). Because of the precise correlation between GC level and gene concentration in the human genome (see ref. 5), compositional mapping (2) is equivalent to mapping of gene concentration. So far, compositional mapping has been performed by hybridizing probes from single-copy sequences located in physically mapped regions of the human genome on compositional human DNA fractions. This approach has been used to assess the GC levels of isochores located on the long arm of chromosome 21 (6), and on the q26–q28 region of the X chromosome (7), as well as of all the isochores forming the dystrophin locus (8). Although very satisfactory, the procedure is time-consuming, and requires a commercially unavailable sequence-specific ligand, BAMD (3,6-bis-acetato mercurimethyl)-1,4-dioxane), as well as an analytical ultracentrifuge in order to determine the CsCl profile and the GC level of DNA fractions. Another procedure (9), based on the enzymatic digestion of radiolabeled YAC DNA, followed by thin layer chromatography separation of the nucleotides, requires YAC purification by pulsed field gel electrophoresis (PFGE), which is inconvenient when analyzing a large number of YACs. Here we present an alternative approach which has none of the problems just mentioned and which can be applied to obtain the compositional map of any human genome region covered by yeast artificial chromosomes (YACs).

Cells from the diploid strain *Saccharomyces cerevisiae* AB1380, containing a human DNA insert cloned in vector pYAC4 (10), were grown on a selective medium [5 g/l $(NH_4)_2SO_4$; 10 g/l acid-hydrolysed casein; 20 mg/l adenine hemisulfate and 2% glucose], transferred into 10 ml YPD (1% yeast extract, 2% bactopeptone, 2% glucose) and grown to stationary phase. Cells ($\sim 2 \times 10^{10}$) were harvested, washed twice with 1 M sorbitol and resuspended in 1 ml spheroplasting buffer (0.9 M sorbitol; 50 mM sodium phosphate, pH 7.5; 14 mM β-mercaptoethanol; 1 mM EDTA). Cells were incubated with 20 U zymolyase 20T (Seikagaku Corporation, Tokyo) for 1 h at 37°C to produce spheroplasts that were pelleted, resuspended in 180 µl lysis buffer (10 mM Tris–HCl, pH 8.0; 50 mM EDTA; 100 mM NaCl; 1 mg/ml proteinase K), and transferred to a Quick-seal 13 × 51 mm centrifuge tube using a Pasteur pipette. Twenty µl of 25% *N*-lauroyl sarcosine sodium salt (Sigma, St Louis) and 4 drops of paraffin oil (Merck, Darmstadt) were added to the centrifuge tube which was incubated at 50°C for 3 h. After lysis, 5 ml of a solution of CsCl in TE (10 mM Tris; 1 mM EDTA), having a density of 1.699 g/cm^3, were added together with 250 ng each of DNAs from T4 (a gift of Gilbert Orsini, Gif-sur-Yvette, France) and λ (Boehringer, Mannheim) bacteriophages, in <10 µl TE. The buoyant densities of the marker DNAs were found by analytical centrifugation to be 1.709 g/cm^3 for λ DMA and 1.700 g/cm^3 for T4 DNA. These density markers were chosen because their base compositions largely cover the buoyant density range (1.698 g/cm^3 to 1.714 g/cm^3) of human DNA. The buoyant density of bacteriophage P1 DNA (a gift of Werner Arber, Basel), used to test the procedure, was found to be 1.707 g/cm^3.

Centrifugation was carried out in a vertical rotor VTi90 (Beckman) at 20°C and 35 000 r.p.m. for 16.5 h using a Kontron preparative ultracentrifuge with the brake off. Thirty-two fractions of 80 µl each were collected using a needle whose tip was positioned 1 cm above the bottom of the tube. DNA fractions were denatured in 0.4 M NaOH (final volume 600 µl) and transferred to a positively charged nylon membrane (Appligene, Pleasantown, CA), using a slot-blot apparatus (BioRad, Richmond, CA). Three series of membrane transfers were made. Two membranes (containing 100 µl of each fraction) were hybridized with λ and T4 bacteriophage DNA, respectively. One membrane (containing 400 µl of each fraction) was hybridized with a (1:1) mixture of plasmids containing an *Alu*I sequence (Blur8; 11) and a 3′ L1 sequence (CD11B; a gift of Ron Thayer, NIH, Bethesda, MD). The probes were radiolabeled using the random oligo primer method (12) and [α-^{32}P]dCTP, as one of the nucleotide precursors, to specific activities of $1–2 \times 10^9$ c.p.m./µg. Filters were hybridized in 0.5 M Na_2HPO_4 (pH 7.2) and 7% SDS in a hybridization oven (Appligene) at 65°C overnight, and washed in 40 mM Na_2HPO_4 (pH 6.5) and 1% SDS, twice at room temperature for 20 min and twice at 65°C for 10 min (13). Filters were analysed with a PhosphorImager (Molecular Dynamics) using the Image Quantification program. Each signal was analysed as an integrated volume, and values were recorded as percentages of the total amount of hybridization signals present in the analysed area of each filter. The percentages of hybridization signals were plotted against the corresponding fractions and a Gaussian curve was fitted to them using the MacCurveFit

* To whom correspondence should be addressed

+Present address: Laboratoire de Structure, Fonction et Evolution du Génome Eucaryote, Institut de Biologie, 4 bld Henri IV, 34060 Montpellier, France

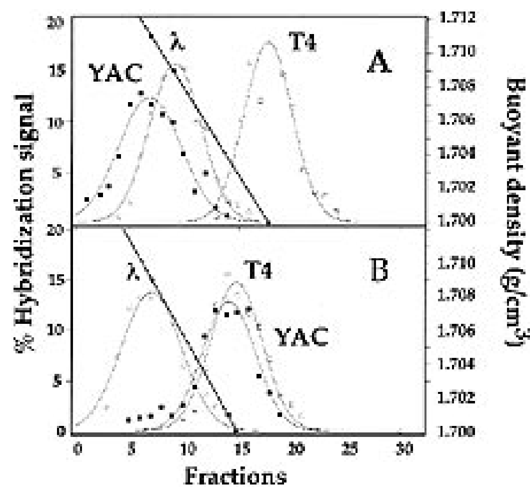

Figure 1. DNA distribution in a shallow CsCl gradient of a GC-rich YAC 917 (**A**) and a GC-poor YAC 813 (**B**) and of the density markers, λ and T4 bacteriophage DNAs. The intensities of the hybridization signals (left side) and the buoyant densities (right side) are plotted against the number of fraction collected from the gradient.

program (14). Three normal curves were thus obtained, which corresponded to the distributions in the gradient of λ, T4 and YAC DNAs, respectively. The peaks of the Gaussian curves of the markers were then plotted against their buoyant densities and used to calculate the buoyant density of the YAC. Finally the GC level of the YAC was calculated on the basis of its buoyant density using the equation $\rho = (GC \times 0.098/100) + 1.66$ (ref. 15).

The lysis of YAC clones in the centrifuge tube using *N*-lauroyl sarcosine (a detergent soluble in concentrated CsCl) was patterned on a previously developed procedure (16) to study very high molecular weight DNA from mouse thymocytes. In spite of the fact that the lysis of spheroplasts was done directly in the centrifuge tube, some yeast DNA fragmentation occurred, apparently as a result of some premature lysis following centrifugation and pipetting of spheroplasts. Indeed, when the lysates were loaded on a gel without any DNA pipetting and subjected to pulsed field gel electrophoresis, both yeast DNA, as observed in ethidium bromide stained gels, and YACs, as detected after Southern hybridization, produced smears. The majority of yeast DNA and of large YACs were broken into fragments of ~50–200 kb. This was not due to a contamination of the enzymes used (zymolyase and proteinase K), because the same enzyme batches were simultaneously used to prepare agarose plugs which were shown to carry intact high molecular weight DNA.

The resolution of standard CsCl gradients was improved by using shallow gradients obtained with a vertical rotor. Plots of the refractive index against fraction number showed that gradients were linear in the range of interest. Regression coefficients ranged from 0.97 to 1.00. On the average, the slope of the gradients was 1.1 mg cm^{-3}/fraction (which corresponds to ~1.1% GC per fraction) with a standard deviation of 0.13. Equilibrium was reached after an overnight run of 16.5 h. Longer runs did not affect the slope of the gradient, nor the width of the DNA distribution. The reproducibility of the results obtained with the method was evaluated with a series of 11 independent experiments in which the DNAs of the bacteriophages λ, P1 and T4 were used. The buoyant density of the DNA of bacteriophage P1, using λ and T4 DNAs as density markers, averaged 1.7068 ± 0.0004 g/cm^3, a value identical, within experimental error, to that obtained by analytical centrifugation, 1.7071 g/cm^3.

The *Alu* and L1 sequences used to detect the human insert cloned in the YACs are interspersed middle repetitive sequences that are widely, but not evenly, distributed in the human genome. Indeed, the former are more numerous in GC-rich isochores, while the latter are concentrated in GC-poor isochores (17). The use of both types of repeated sequences allowed a fair representation of YACs of all GC levels. *Alu* and L1 sequences were shown not to hybridize on the fractions from a gradient of a cell lysate of the recipient yeast strain. Hybridization of filter-immobilized aliquots of DNA, ranging in amount from 1 to 500 ng, produced a linear increase of the intensity of the hybridization signals. The intensities of the hybridization signals observed in each fraction were plotted against the fractions and fitted by a Gaussian curve. The means of the regression coefficients obtained from 50 independent experiments were 0.97 (SD = 0.024) for λ DNA and 0.97 (SD = 0.019) for T4 DNA. The effect of the vector sequence, which is 9.9 kb long and has a base composition of 43.7%, was negligible in most cases. Indeed, for a relatively small 100 kb YAC having the highest possible GC level, 54% GC, the correction would only be +1% GC.

Two examples of human YACs analysed by the procedure described above are shown in Figure 1. YAC 917 is GC-rich (ρ = 1.711 g/cm^3, 52.0% GC). YAC 813 is GC-poor (ρ = 1.701 g/cm^3, 41.8% of GC). These results are just two examples from an analysis of a collection of YACs ranging in size from 60 to 490 kb (De Sario *et al*., in preparation), which have been mapped to the human chromosomal band Xq28 (18) and span the whole GC range of human isochores. In conclusion, the method presented here seems to be satisfactory for the compositional mapping of YAC contigs, and to present advantages over alternative methods. Moreover, it can be applied to a variety of other problems, including the localization of genes to compositional fractions of animal and plant DNA.

ACKNOWLEDGEMENTS

We are grateful to Dr Gabriel Macaya for his advice in setting up the methodology as well as for discussions. EMG was found by an individual Human Capital and Mobility fellowship from the European Community (ERB4001GT930863).

REFERENCES

1 Bernardi, G. *et al.* (1985) *Science* **228**, 953–958.
2 Bernardi, G. (1989) *Annu. Rev. Genet.* **23**, 637–661.
3 Mouchiroud, D. *et al.* (1991) *Gene* **100**, 181–187.
4 Bernardi, G. (1993) *Gene* **135**, 57–66.
5 Bernardi, G. (1995) *Ann. Rev. Genet.* (in press)
6 Gardiner, K. *et al.* (1990) *EMBO J.* **9**, 1853–1858.
7 Pilia, G. *et al.* (1993) *Genomics* **17**, 456–462.
8 Bettecken, T. *et al.* (1992) *Gene* **122**, 329–335.
9 Krane D.E. *et al.* (1991) *Nucleic Acid Res.* **19**, 5181–5185.
10 Burke, D.T. *et al.* (1987) *Science* **236**, 806–812.
11 Deininger, P.L. *et al.* (1981) *J. Mol. Biol.* **151**, 17–33.
12 Feinberg, A. and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
13 Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
14 Raner, K. (1992) *MacTutor Magazine* **8**(3), 24.
15 Schildkraut, CL. *et al.* (1962) *J. Mol. Biol.* **4**, 430–443.
16 Macaya, G. *et al.* (1976) *J. Mol. Biol.* **108**, 237–254.
17 Soriano, P. *et al.* (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1816–1820.
18 Palmieri, G. *et al.* (1994) *Genomics* **24**, 149–158.