Characterization of four human YAC libraries for clone size, chimerism and X chromosome sequence representation

Ramaiah Nagaraja, Juha Kere, Sandra MacMillan, Mochubeloa W.J.Masisi, Douglas Johnson, Barbara J.Molini, George R.Halley, Kim Wein, Michelle Trusgnich, Brian Eble, Brian Railey, Bernard H.Brownstein and David Schlessinger*

Center for Genetics in Medicine, Washington University School of Medicine, St Louis, MO 63110, USA

Received April 22, 1994; Revised and Accepted July 11, 1994

ABSTRACT

Four collections of human X-specific YACs, derived from human cells containing supernumerary X chromosomes or from somatic cell hybrids containing only X human DNA were characterized. In each collection, 80–85% of YAC strains contained a single X YAC. Five thousand YACs from the various libraries were sized, and cocloning was assessed in subsets by the fraction of YAC insert-ends with non-X sequences. Cocloning was substantial, ranging up to 50% for different collections; and in agreement with previous indications, in all libraries the larger the YACs, the higher the level of cocloning. In libraries made from human–hamster hybrid cells, expected numbers of clones were recovered by STS-based screening; but unexpectedly, the two collections from cells with 4 or 5 X chromosomes yielded numbers of YACs corresponding to an apparent content of only about two X equivalents. Thus it is possible that the DNA of inactive X chromosomes is poorly cloned into YACs, speculatively perhaps because of its specialized chromatin structure.

INTRODUCTION

The quality of yeast artificial chromosomes (YACs) containing inserts of human genomic DNA has been sufficient to achieve maps with long-range continuity across the human genome at levels of resolution up to 100 kb (1–4); but the degree to which YACs provide fidelity as well as coverage of the genome has remained unclear. Even at low resolution, however, one problem is apparent: in every collection of YACs reported to date, a fraction of the clones are chimeric, bringing together in covalent linkage fragments of DNA from more than one chromosomal location. But the precise extent of cocloning in various collections of YACs, and its relation to parameters such as YAC size and genomic location has been estimated only for portions of one X

and one chromosome 11 library (5), and have otherwise generally been studied only anecdotally.

The analysis of YACs is simplified by studying collections or subcollections of clones targeted to a specific region of the genome. Such collections have been generated by several techniques, including the subfractionation of YAC libraries of total human DNA (3) and the formation of YACs from flow-sorted chromosomes (6); but the most generally useful method has been based on the use of rodent/human somatic hybrid cells containing a single human chromosome.

Among the first such libraries constructed were two from hybrid cells containing Xpter–q27.3 (7) or Xq24–qter (8–9) as their only detectable content of human DNA. Preliminary data on those libraries, based on the use of hybridization probes, has been reported (1,7). During extensive mapping of the X chromosome, we have studied these two collections in further detail, and report here the levels of cocloning observed, along with some comparative characterization of collections of X-containing clones derived from two libraries of total human DNA made from cells with 4 (10) or 5 X chromosomes.

MATERIALS AND METHODS

YAC libraries

Collections of 3109 clones specific for Xpter–q27.3 (from hybrid cell Micro-21D; ref. 7) and 800 clones specific for Xq24–qter (from hybrid cell X3000.11; ref. 8–9) were the resource for the production of most of the end-insert STSs studied; they are referred to as the F and X3000 libraries, respectively, in the text.

Representative clones were also characterized from two other YAC libraries, derived from lymphoblastoid cells. One set of 36 000 YAC clones, kindly provided by Zeneca Pharmaceuticals, has been reported by Anand *et al.* (10) from a cell containing 4 X chromosomes (GM1416 in the Camden Repository); it is called the I library in the text. A second set of 20 500 clones was prepared by F.Abidi, J.Y.Yoon and R.Mazzarella from a

*To whom correspondence should be addressed

cell containing 5 X chromosomes (GM06061B), using a protocol essentially the same as that of Anand *et al.*; it is called the E library in the text.

To estimate the sizes of YACs, DNA was prepared from spheroplasts of yeast clones embedded in agarose blocks (11), and a small bit of each block was applied to a lane of a gel for pulsed-field electrophoresis (11) in a Biorad CHEF apparatus (model DRII) with a typical running time of 18 hrs at 6volts/cms in 0.5×TBE, starting with switch times of 30 sec forward and reverse, and ending with a switching time of 60 sec forward and reverse. After electrophoresis, the DNA was transferred by Southern blotting to a Sureblot nylon filter (Oncor) and hybridized to a probe of total human DNA labeled by the random priming method (12) with an Amersham Enhanced Chemiluminescence kit to detect the mobility of corresponding YACs in relation to control yeast chromosomes. DNA from yeast strain AB972 was included in each gel to provide size standards (13). In many experiments, additional hybridization probes of pBR322 and/or hamster DNA were used to confirm the presence of additional YACs or the presence of clones containing hamster DNA.

Screening for cognate YAC clones with STSs

Screening for YACs was entirely PCR-based. Robot-assisted screening was carried out using the uniform temperature regimen and buffers as described in ref. 14.

Recovery and testing for chimerism with YAC insert-ends

YAC ends were isolated using ligation-mediated PCR as described (15). Sequences were determined by the dideoxy chain termination method (16), using vector-arm specific labeled primers in an ABI 373A sequencer. The sequences of the primers tagged with fluorescent labels for left end and right end respectively are: 5' TCTCGGTAGCCAAGTTGGTTTAAGG and 5' TCGAACGCCCGATCTCAAGATTAC.

Sequences were edited and compared to the GenBank database using the FASTA program of the GCG package (17). Sequences that showed significant homology to Alu (18) or L1 (19) repeats were not used. The remaining sequences were analysed by the OSP program (20) to predict and select oligonucleotides suitable for PCR. In general, more than 97% of primer pairs tested were effective in PCR assays with parental clones as templates, and were further studied.

To test for cocloning in a YAC, PCR was carried out with the end-insert primer pairs (STSs; 21) on a series of templates. Since all of the YACs studied here were recovered with STSs made from X-chromosome specific probes, the testing panel included total human DNA (from CGM-1 cells; ref.22) and two hamster/human somatic cell hybrids containing either a single X chromosome (GM06318B) or the Xq24-qter portion of the X chromosome (23) as their only content of human DNA. A further test that the PCR primer pair indeed defined a unique STS was carried out with DNA from a third hybrid cell containing human chromosome 7 (GM10791; see Results, Figure 1; cf. the testing panels in refs. 15,24).

RESULTS

As reported here, the size distribution and fraction of clones containing single YACs is very comparable in the four libraries tested. Unexpectedly, X-specific DNA from cells with multiple X chromosomes is cloned relatively inefficiently; but in all cases, efficient recovery and characterization of sequences from YAC

Table 1. Frequency of cocloning in YAC insert-ends

ELIBRARY		
NUMBER OF SEQUENCES ANALYZED	715	% OF TOTAL
Xpter-Xq24	159	22.2
Xq24-qter	73	10.2
Human Repetitive Sequences*	285	39.9
Hamster	198	27.7
ELIBRARY		
NUMBER OF SEQUENCES ANALYZED	174	
Xpter-Xq24	17	9.8
Xq24-qter	51	29.3
Human repetitive sequences*	69	39.7
Non X Sequences	37	21.2
ILIBRARY		
NUMBER OF SEQUENCES ANALYZED	73	
Xpter-Xq24	24	32.8
Xq24-qter	2	2.7
Human repetitive sequences*	36	49.3
Non X Sequences	11	15.0

*Determined by comparison to GenBank database.

Oligonucleotide pairs were selected from sequenced YAC insert-ends that were found to be unique sequence by the FASTA program (15). STSs were tested as described in Methods (see also Figure 1). Results are shown for the number of end-sequences analyzed from the F, E and I library. Note that the % of total cocloned ends is accurately based on the number of hamster sequences for the F library, but is underestimated for E and I, because X-specific and autosome-specific cocloned repetitive sequences are indistinguishable.

insert-ends permits the estimation of the fraction of cocloned ends; and a concordant qualitative estimate can be independently inferred for the X-only libraries from the representation of X sequences.

Content and size of single YACs in four libraries

In pulsed-field gel analyses of 750 X3000 library clones, 2811 F clones, 432 I clones, and 717 E clones, the fraction of clones containing single YACs ranged from 80 to 88%, comparable to that originally found in a sample of the first library of total DNA assembled in YACs (22).

The average YAC sizes for the X3000, F, I and E collections were, respectively, 257, 225, 323 and 257 kb, as determined by comparison to the independently-sized chromosomes of yeast strain AB972 (11). The distribution of sizes is shown in Figure 2 (filled bars). Of the two collections made from somatic hybrid cells, the F library shows a more restricted size distribution than the X3000 library, and the clones are somewhat smaller on average (225 vs 280 kb) than that reported for a subset examined for ref. 7; this may reflect a systematic difference in pulsed-field gel electrophoretic sizing gels or some bias in sampling. The other two libraries, E and I, were constructed with a size-fractionation step on gels to increase average size.

Chimerism inferred from tests of insert end-clones

In sets of chromosome-specific YACs, cocloning can be scored easily by several methods. When the YACs were made from somatic cell hybrids (X3000 and F libraries), those containing human DNA were directly recognized among a more than hundred-fold excess of hamster YACs by their hybridization to radiolabeled probes of total human DNA. Those YACs containing a substantial content of rodent as well as human sequences can

Figure 1. STS testing for X chromosome localization. Each STS developed from a YAC insert-end is tested by PCR against yeast DNA (YY); a random pool of human YACs (PO); the random pool to which a trace of total human DNA has been added (SP); total human DNA (HU); and DNAs from hamster-human somatic cell hybrids containing only X (XO), X3000.11 (X3) or chromosome 7 (7O). Results are shown that localize an STS to Xpter-q24 (panel A), Xq24-qter (panel B), or to hamster DNA (C).

then be recognized by an additional hybridization with total hamster DNA as a probe. Such experiments yielded estimates of 12–20% cocloning in the two libraries. However, these experiments are not applicable to sets of clones made from total human DNA, and like *in situ* measurements, are relatively insensitive and less definitive than the analysis of end-inserts applied here.

Ligation-mediated PCR as reported by Kere *et al.* (15) has proven to be highly efficient. Carried out on YACs in groups of 96 in microtiter plate format, the procedure yielded 68% of insert-ends when a single enzyme (*RsaI*) was used. Yields of left or right insert-ends reached 99% when the protocol was repeated with *PvuII*, *AluI* or *ScaI* on clones which had yielded no product with *RsaI*.

From insert-ends that provided enough good sequence in a single run (about 80% in typical cases), cocloning was assessed, and is summarized in Table 1.

An initial comparison detected most of the Alu and L1-like elements. From the remaining sequences, primer pairs for STSs were predicted by the OSP program, synthesized, and partially characterized by PCR testing against a panel of DNAs.

Typical testing profiles for clones from the X-specific YACs of the F library are shown in Figure 1.

In panel A, an STS (sWXD1185, entered in GDB) showed a unique signal of the expected size from total human DNA, from human DNA diluted into a pool of indifferent YACs, and from DNA of a hybrid cell containing total X DNA, but not from a hybrid containing Xq24-qter DNA. These results imply that the end-clone is human-specific and falls in the Xpter-q24 region. As a further indication of its X-specificity, the primer pair amplified no product from hybrid cell DNA containing human chromosome 7. Comparable tests in Figure 1B shows results for an STS (sWXD1184) in the Xq24-qter region.

In contrast, Figure 1C shows an STS that yielded no PCR product from human DNA but was positive with all three hamster/human hybrid cells. It is scored presumptively as hamster in origin. In further tests, 20/20 putative hamster STSs were

Figure 2. Sizes of YACs and cocloned YACs in four libraries. YACs were sized by pulsed-field gel electrophoresis compared to natural chromosomes of yeast strain AB972. Filled bars: the percent of YACs in various size ranges are summed for various libraries; 100% represents, respectively, 717 E library, 2811 F library, 432 I library, or 750 X3000 library clones. Hatched bars: For comparison, the sizes of a subset of clones known to be cocloned is plotted as a percent of that group; 100% represents, respectively, 39, 178, 14 and 187 clones for the E, F, I and X3000 libraries.

positive, as expected, when tested directly against purified hamster DNA.

Table 1 summarizes data for groups of YAC end-inserts. Of 719 F library sequences, four proved to be yeast sequences and were excluded from further analysis. Non-repetitive sequences were recovered from both ends of only a fraction of the YACs studied, but a measure of overall cocloning could be inferred from the fraction of ends which tested as hamster-specific: 198 of 715, or 27.7%. Assuming that only one end of a chimeric clone is hamster DNA, the frequency of cocloning is double the occurrence of hamster-specific ends; and thus 27.7% cocloned ends imply that 55% of YACs in the collection were cocloned.

Table 1 indicates that the 715 end-sequences also included 285 repetitive human sequences, or about 39.9% of the ends scored. The other primer pairs became STSs, distributed by the PCR panel tests into Xpter-q24 (22.2%) and Xq24-qter (10.2%).

It is of interest that 31% (10.2/32.4) of the STSs derived from end-sequences of anonymous X-specific YACs fell into Xq24-qter. For the X chromosome, containing about 160 Mb, 31% would represent 50.3 Mb. This is very close to current YAC clone coverage of Xq24-qter (CCM93), with greater than 98% estimated coverage in 49.5 Mb of mapped DNA.

Figure 3. Number of cognate YACs recovered by STSs from the E, F and I libraries. The ordinate plots the number of STSs for which 0, 1,...8 YACs were recovered from a library (see text).

Comparable tests with the X3000 library gave corresponding estimates for the degree of cocloning and the distribution of end-sequences into the categories of hamster, repetitive sequences, and Xpter-q24 and Xq24-qter categories. Because that library is largely organized into contigs across cytogenetic bands (1,8,25,26, and mss. in preparation), the degree of cocloning can be estimated by region. For example, for Xq28, 79 of 92 end-probes were human DNA when tested by hybridization against hamster and human DNA; and 36 of 41 inserts were found to be human DNA after sequencing and STS development were pursued (Palmieri *et al.*, ms. in preparation). Both the hybridization and PCR-based tests indicated that about 13% of YAC ends in the region contained hamster sequences, implying a cocloning rate of 26%. In other subregions, the number of cocloned ends varied from 12 to 25%.

Comparable testing was adapted for X-specific clones isolated from the E and I collections. Clones were recovered with X-specific STSs and end-inserts isolated as above; but the analysis of cocloning was limited by the inability to determine whether repetitive sequences arose from X-specific portions of clones or from cocloned autosomal material. As indicated in Table 1, the fraction of unique sequences that amplified products from total human DNA but were demonstrably non-X was 21.2% for the E clones. The minimum estimate of chimerism for the E library would thus be 42%. This would be augmented by the inclusion of a fraction of the repetitive sequences, leading to an overall rate that is likely to be comparable to the figure for the F library. For the I library, the non-X fraction of 15% (Table 1) would yield a tentative value of 30% chimerism, but the numbers are low, and more extensive studies indicate much lower levels of chimerism for the range of genomic clones in the collection (see Discussion).

Cocloning, and its relation to YAC size and STS representation

Since the size data for all clones was available, the sizes of YACs containing non-X end-sequences could be compared to the total

distribution of sizes. The results are incorporated into Figure 2. For all collections, the probability of cocloning increases sharply with clone size, as already reported for the X3000 YAC collection (1) and a library of chromosome 11-specific clones (5), and as suggested for the F library (7).

To assess further the quality and X-specific DNA content of the YAC libraries, clones of each library were screened to recover YACs cognate for a series of STSs. From clone sizes and STS determinations, a second estimate of the total X-specific DNA content could be determined.

For the X3000 and F libraries. Since the X3000 and F libraries are made from hybrid cells in which the human DNA is exclusively X-specific, the number of YACs multiplied by the average YAC size gives the total content of putative X-specific DNA equivalents, about 3.5 in both cases.

If all of the YAC insert DNA in these two libraries is truly X-specific, then the number of equivalents of DNA should equal the average number of clones recovered for each of a series of probes ('hitness'). For the X3000 library, with an average clone size of 210 kb and a calculated content of 3.5 equivalents of a 50 Mb region, hybridization data reported earlier (1) and ongoing PCR-based screening both recover 3.0 cognate clones per probe. The difference between 3.5 and 3 is consistent with estimates that cocloning in this library accounts for no more than 25% of the total DNA.

For the F library, the difference is more marked. With 2800 clones and an average size of 220 kb, the collection should contain about 3.8 equivalents of a chromosome of 160 Mb. Yet the average number of clones recovered per STS is slightly above 2.3; the spread of values for 600 STSs is shown in Figure 3, and is roughly in accord with Poisson expectation (Schlessinger *et al.*, 1991). This suggests that the order of 40% of the clone content may be hamster DNA, consistent with the estimate above that more than half of the clones are chimeras containing hamster DNA.

For the E and I libraries. Since these collections contain YACs made from all the human chromosomes, only YACs recovered by screening with X-specific probes have been compared with clones from other libraries.

The results of screening with STSs were somewhat unexpected. From the average size and numbers of clones (38 000 for I and 20 500 for E), the two libraries would be expected to contain roughly 4 and 2 equivalents of total genomic DNA. But the I and E libraries were made, respectively, from cells containing 4 and 5 X chromosomes, and would therefore be expected to contain 8 and 4 equivalents of X-specific clones—and therefore to yield 8 and 5 clones per probe. Instead, screenings of the two libraries have given about 4 and 2.3 clones per probe; again, the distributions are given in Figure 3. [Note that the distribution for the I library is bimodal. This is because the STSs started with 100 which had found 0 or 1 clone in the E and F libraries combined. About 30 of these have found no YACs in libraries containing 10 genomic equivalents of DNA, and are therefore essentially unrecoverable. Apart from these, the average number of YACs recovered per STS is >3.6.] These results imply that the collections contain no more X YACs than are expected for autosomes (see Discussion).

In general, if all DNA were equally cloned into YACs, the numbers of YACs recovered by probes can be described by the Poisson distribution, and 'hitness' can also be estimated from

the numbers of probes which recover no YACs (e^{-n} , where n is the average yield of clones/probe; see, e.g., ref. 1). Apart from a fraction of probes (20 of 600) which appear to be essentially unrecoverable in any YAC library tested, 'hitness' assessed in this way was consistent with the overall screening results; for example, about 60 probes, or 10% (roughly e^{-2}), recovered no clones in the F library.

No independent estimate of cocloning can be obtained by comparing total X DNA content with 'hitness' for these libraries, since only some of the X-specific YACs have been recovered and the libraries are made from total human genomic DNA.

DISCUSSION

Detection and levels of cocloning

One reason for the slow characterization of cocloning early on was the dependence on *in situ* hybridization tests to detect clones which mapped to more than one genomic location. Such methods are too labor-intensive to be readily adapted to systematic determinations and are relatively insensitive to small extents of chimerism; in our experience (28 and Featherstone *et al.*, work in progress), cocloned regions less than about 25% the length of a typical YAC are often not detected in comparison to the stronger signal from the rest of the clone.

Recovering the ends of YAC inserts and testing directly for their chromosomal origin, for example with a panel of somatic cell hybrids containing single human chromosomes, provides a more precise and straightforward measure of cocloning. This approach has become increasingly attractive for several reasons: end-inserts of YACs provide an immediate measure of the quality of individual YAC clones in mapping and in the selection of material for gene hunts; end-inserts of YACs provide optimal probes for contig expansion, reaching out into neighboring DNA to the maximum extent possible; and from the technical point of view, primer ligation-mediated methods (15,27) have facilitated the easy recovery of the vast majority of ends.

From the unequivocal estimates of cocloning from insert-ends, it is clear that the YAC libraries partially characterized here are comparable in cocloning levels, with about 1 YAC in two containing an appreciable amount of DNA from at least two different chromosomal loci. A similar result was obtained for a smaller collection of YACs from another library by Bates *et al.* (29). By discarding those clones which hybridize strongly to hamster DNA probes, the level of cocloning can be lowered in collections made from rodent/human hybrids. Even in those cases, however, many cocloned YACs are included in current libraries.

As one caveat, it should be noted that the estimates of cocloning have all been inferred here from X chromosome YACs. Extensive studies of end-clones from more than 200 YACs across the genome have found much less cocloning ($\sim 10\%$) in clones from the I library than has thus far been observed for X-specific clones in these studies (R.Anand, personal communication). Also, as mentioned earlier, cocloning rates were apparently higher in some regions of the X, so that overall percentages may not apply to subgenomic regions.

In these libraries, as in the earlier studies of the X3000 collection and the comparable examination of a set of clones specific for chromosome 11 (1,5), larger clones show increasing levels of cocloning; and by the time clones larger than 700 kb are examined, the vast majority show at least one chimeric end.

Two additional complications in the determination of cocloning levels should be noted. One comes from the observation that about 10–20% of all YAC strains contain at least two YACs. The additional YACs will in general tend to yield extraneous ends. To estimate rates of cocloning here, such strains have been excluded; but they can produce extraneous STSs in mapping strategies based on walking with YAC ends. A second and especially treacherous type of cocloning is that generated from two segments on the same chromosome. The data are not complete enough to analyze whether such clones occur in 1/20 cocloned YACs, as expected if intrachromosomal and interchromosomal chimeras arise by similar mechanisms; but for example, such cocloning has been detected in YAC yWDX843 (8) and in several others in ongoing mapping efforts, revealed by their inconsistent probe contents or lack of overlap with neighboring YACs.

These results underline a difficulty in the current practice of YAC-based mapping. On the one hand, long-range coverage is achieved more rapidly by the use of large clones, notably in the use of larger YACs to produce a first overall map of all human chromosomes (4). On the other hand, those clones are likely to be heavily cocloned, with estimates that may well go up from earlier determinations by *in situ* hybridization as end-clones are studied. This makes it more difficult to achieve high resolution maps with accurately determined intermarker distances unless large numbers of clones, including smaller clones that show less cocloning, are employed. One route to reduce the effort required to achieve highly resolved maps may be to avoid difficulties by developing mapping clones with less chimerism. These results do not imply any particular mechanistic origin for chimeras. As discussed previously (1), both standard models for the origin of chimeras—based on coligation of or recombination between two DNA fragments—would anticipate the higher levels of chimeras in larger clones. On the other hand, several studies show that at least some clone instability and chimerism arises from recombination events (30–32), and that clones with less chimerism can be recovered in yeast host strains disarmed in recombination pathways (32).

Unexpectedly low X-specific clone content in E and I libraries

If all their constituent DNA were comparably clonable, YAC collections from 4X and 5X-containing cells would yield about four times as many YACs per genome equivalent of DNA for X-specific probes as do libraries made from cells containing a single X chromosome (like the CGM library; ref. 22). Instead, the yield of YACs per genome equivalent of DNA was no greater. The apparent shortfall in yield probably reflects the true content of X-specific clones in the YAC collections—rather than a very high false-negative rate of screening—since 1) PCR-based results and hybridization-based results have agreed whenever compared directly, and have been capable of organizing essentially all the clones of the X3000 library; and 2) the recovery of YACs/probe from the X3000 and CGM libraries (and, somewhat less decisively, from the F clones) is in agreement with their content of X-specific DNA.

These unexpected results suggest that most of the X DNA in the hybrid cells is poorly cloned into YACs. One possibility is that only the single active X chromosome is well cloned. It may be relevant that DNA for YAC cloning, in contrast to preparations for small-insert cloning, is customarily only partially purified. Such minimal handling maintains high molecular weight DNA, but very likely retains some of the features of chromatin. The

DNA from inactive X chromosomes might then be disfavored for cloning because of special features like relatively tight coiling.

ACKNOWLEDGEMENTS

We are indebted to Dr Rakesh Anand and Zeneca Pharmaceuticals, and to Dr R. Nussbaum, for providing the I and F libraries, respectively; to our colleagues, Fatima Abidi and Richard Mazzarella for the E library; to Mary Blanchard and Volker Nowotny for YAC pools and screening protocols and to Sheila Kesterson and Frank Burroughs for data management. We thank E. Green for critical reading of the manuscript. Work was supported by NCHGR Center grant HG00201.

REFERENCES

- Schlessinger, D., Little, R.D., Freije, D., Abidi, F., Zucchi, I., Porta, G., Pilia, G., Nagaraja, R., Johnson, S.K., Yoon, J.Y., Srivastava, A., Kere, J., Palmieri, G., Ciccodicola, A., Montanaro, V., Romano, G., Casamassimi, A. and D'Urso, M. (1991) *Genomics* 11, 783–793.
- Foot, S., Vollrath, D. and Hilton, A. (1992) *Science* 258, 60–66.
- Chumakov, I., Rigault, P., Guilou, S., Ougen, P., Billaut, A., Guasconi, G., Gervy, P., LeGall, I., Soularue, P., Grinas, L., Bougueleret, L., Bellané-Chantelot, C., Lacroix, B., Barillot, E., Gesnoui, P., Pook, S., Vaysseix, G., Frelat, G., Schmitz, A., Sambucy, J.-L., Bosch, A., Estivill, X., Weissenbach, J., Vignal, A., Riethman, H., Cox, D., Patterson, D., Gardiner, K., Hattori, M., Sakaki, Y., Ichikawa, H., Ohki, M., Le Paslier, D., Heilig, R., Antonarakis, S. and Cohen, D. (1992) *Nature* 359, 380–386.
- Cohen, D., Chumakov, I. and Weissenbach, J. (1993) *Nature* 266, 698–700.
- Selleri, L., Eubanks, J.H., Biovannini, M., Hermanson, G.G., Romo, A., Djabali, M., Maurer, S., McElligott, D.L., Smith, M.W. and Evans, G.A. (1992) *Genomics* 14, 536–541.
- McCormick, M.K., Campbell, E., Deaven, L. and Moyzis, R. (1993) *Proc. Natl. Acad. Sci. USA* 90, 1063–1067.
- Lee, J.T., Murgia, A., Sosnoski, D.M., Olivos, I.M. and Nussbaum, R.L. (1992) *Genomics* 12, 526–533.
- Little, R.D., Pilia, G., Johnson, S., Zucchi, I., D'Urso, M. and Schlessinger, D. (1992) *Proc. Natl. Acad. Sci. USA* 89, 177–181.
- Abidi, F.E., Wada, M., Little, R.D. and Schlessinger, D. (1990) *Genomics* 7, 363–376.
- Anand, R., Riley, J.H., Butler, R., Smith, J.C. and Markham, A.F. (1990) *Nucl. Acids Res.* 18, 1951–1956.
- Chu, G., Vollrath, D. and Davis, R.W. (1986) *Science* 224, 419–426.
- Feinberg, A.P. and Vogelstein, B. (1983) *Analytical Biochem.* 132, 6–13.
- Link, A.J. and Olson, M.V. (1991) *Genetics* 127, 681–698.
- Blanchard, M.M., Sloan, D. and Nowotny, V. (1991) In Strimaitis, J.R. and Little, J.N. (eds), *Automation of the mapping of the human genome—Advances in Lab. Automat. Robotics*. Zymark Corp., Hopkinton, MA, Vol. 8, pp. 333–345.
- Kere, J., Nagaraja, R., Mumm, S., Ciccodicola, A., D'Urso, M. and Schlessinger, D. (1992) *Genomics* 14, 241–248.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Genetics Computer Group. Program Manual for the GCG package, Version 7, April 1991, 575 Science Drive, Madison, Wisconsin, USA 53711.
- Deininger, P.L. (1989) In Berg, D.E. and Howe, M.M., (eds.) *SINEs: short interspersed repeated DNA element in higher eukaryotes. Mobile DNA*, American Society for Microbiology, Washington, pp. 619–637.
- Hutchison, C.A.III, Hardies, S.C., Loeb, D.D., Shehee, W.R. and Edgell, M.H. (1989) In Berg, D.E. and Howe, M.M., (eds.) *LINEs and related retroposons: long interspersed repeated sequences in the eukaryotic genome. Mobile DNA*, American Society for Microbiology, Washington pp. 593–619.
- Hillier, L. and Green, P. (1991) *PCR Methods Applic.* 1, 124–128.
- Olson, M., Hood, L., Cantor, C. and Botstein, D. (1989) *Science* 245, 1434–1435.
- Brownstein, B.H., Silverman, G.A., Little, R.D., Burke, D.T., Korsmeyer, S.J., Schlessinger, D. and Olson, M.V. (1989) *Science* 244, 1348–1351.
- Nussbaum, R.L., Airhart, S.D. and Ledbetter, D.H. (1986) *Amer. J. Med. Genet.* 23, 457–466.
- Green, E.D., Mohr, R.M., Idol, J., Jones, M., Buckingham, J.M., Deaven, L.L., Moyzis, R.K. and Olson, M.V. (1991) *Genomics* 11, 548–564.
- Pilia, G., Little, R.D., Bernardi, G., Aissani, B. and Schlessinger, D. (1993) *Genomics* 17, 456–462.
- Schlessinger, D., Mandel, J.L., Monaco, A.P., Nelson, D.L. and Willard, H.F. (1993) *Cytogenet. Cell Genet.* 64, 148–194.
- Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C. and Markham, A.F. (1990) *Nucl. Acids Res.* 18, 2887–2890.
- Montanaro, V., Casamassimi, A., D'Urso, M., Yoon, J.-Y., Freije, W., Schlessinger, D., Muenke, M., Nussbaum, R.L., Saccone, S., Maugeri, S., Santoro, A.M., Motta, S. and della Valle, G. (1991) *Amer. J. Human Genet.* 48, 183–194.
- Bates, G.P., Valdes, J., Hummerich, H., Baxendale, S., de Paslier, D.L., Monaco, A.P., Tagle, D., MacDonald, M.E., Altherr, M., Ross, M., Brownstein, B.H., Bentley, D., Wasmuth, J.J., Gusella, J.F., Cohen, D., Collins, F. and Lehrach, H. (1992) *Nature Genetics* 1, 280–287.
- Green, E.D., Riethman, H.C., Dutchik, J.E. and Olson, M.V. (1991) *Genomics* 11, 658–669.
- Neil, D.L., Vaillasante, A., Fisher, R.B., Vetrie, D., Cox, B. and Tyler-Smith, C. (1990) *Nucl. Acids Res.* 18, 1421–1428.
- Ling, L.L., Ma, N. S.F., Smith, D.R., Miller, D.D. and Moir, D.T. (1993) *Nucl. Acids Res.* 21, 6045–6046.