

ISOLATION OF HUMAN CHROMOSOME 21-SPECIFIC COSMIDS AND THEIR USES IN MAPPING OF COSMID CONTIGS ON CHROMOSOMAL SUBREGIONS

De-Xing HOU,* Haruo KISHIDA, Takanori SHIMOKAWA,
and Eiichi SOEDA

*Gene Bank, Tsukuba Life Science Center, The Institute of Physical and Chemical Research
(RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan*

Summary A cosmid library of 3×10^5 clones has been constructed from a human \times hamster hybrid cell line, 153E9a3, which contains human chromosome 21 (HC21) as the only human chromosome. From 56,500 clones of this library, 229 HC21-specific cosmids have been isolated by their hybridization to total human DNA and by their failure to hybridize to total Chinese hamster DNA. The cosmids isolated were then characterized, of these, 28 cosmids (12.2% of those tested) contained *NotI* site(s), and 41 cosmids were localized on the eight subregions of HC21 by differential hybridization with *Alu*-PCR products obtained from a hybrid mapping panel. The cosmids localized were further integrated into the existing contigs using the end-specific probes of the clone insert. Therefore, they provided useful anchor points for contig mapping and walking.

Key Words human chromosome 21, cosmid library, cosmid contig, regional mapping

INTRODUCTION

Human chromosome 21 (HC21) represents a model for physical mapping of the human genome because it is the smallest human autosome and associates with several genetic diseases, including Down's syndrome, some forms of Alzheimer's disease and other neurological diseases (Stewart *et al.*, 1989). The cloning and regional mapping of chromosomal DNA will speed the characterization of the disease-related genes. Recently, a continuous unit of yeast artificial chromosomes (YACs) covering the entire human chromosome 21q has been constructed using

Received August 24, 1994; Revised version accepted October 4, 1994.

* To whom correspondence should be addressed.

sequence-tagged sites (STS) as landmarks (Chumakov *et al.*, 1992a). YACs provide an efficient way to order large genomic regions but they tend to have high frequency of chimerism, deletions and rearrangements (Schlessinger, 1990; Christopher, 1993). Unlike YACs, cosmids tend to have more stable inserts with little chimerism. Such cosmid contig maps will provide easily accessible materials for determining genome organization and for obtaining the DNA sequence.

For these reasons, we have begun to construct the HC21 contig map with cosmid clones. Our project consisted of two logical steps: first, the cosmid clones from a primary library were assembled into sets of overlapping cosmids (contigs) based on the fingerprinting method (Carrano *et al.*, 1989); second, contigs were mapped onto HC21 with known markers, and the gaps between contigs were closed up by other cosmid clones. To date, we have fingerprinted approximately 9,300 cosmid clones from a HC21 flow-sorted library (constructed in Lawrence Livermore National Laboratory) by automated fluorescence-based fingerprinting method (Carrano *et al.*, 1989), and assembled 4,440 cosmids of them into 467 contigs by a computer algorithm on the basis of measured overlap probability (the log likelihood ratio) statistics that exceeded an acceptable threshold (Branscomb *et al.*, 1990) (submitted for publication). Now, we are mapping these contigs onto HC21 and trying to close the gaps between the contigs. However, we have noted that some gaps, where no junction clone exists in this library, must be closed with other libraries.

Here, we constructed a cosmid library from a human \times hamster hybrid cell, 153E9a3, which retains HC21 as the only human chromosome. The HC21-specific cosmids were isolated from this library and then localized on HC21 subregions. The cosmids localized were further used to map the existing contigs and to close the gaps between contigs.

MATERIALS AND METHODS

Library construction. The human \times hamster hybrid cell line 153E9a3, which retains human chromosome 21 as the only human chromosome (Patterson *et al.*, 1985), was used as the cloning source. The cell was cultured in F12 medium with 10% fetal bovine serum. The unique presence of human chromosome 21 was certified by fluorescence *in situ* hybridization (Hou, unpublished data). High-molecular-weight DNA was extracted and partially digested by *Sau3AI* according to standard method (Maniatis *et al.*, 1989). Fragments between 35 and 50 kb were collected by sedimentation through a sucrose density gradient and then dephosphorylated. A high efficiency cosmid vector, sCos-1 (Evans *et al.*, 1989), was prepared by *XbaI* digestion and dephosphorylation. After the cloning site was cleaved with *BamHI*, the vector DNA was ligated with dephosphorylated insert DNA and then packaged using the Gigapack Gold System (Stratagene). Finally, the assembled particles were transfected into *E. coli* 490A (Stelnmetz *et al.*, 1982)

and plated under ampicillin selection. This library has been designated as RL01.

Identification of human cosmid clones. The cosmid clones were plated at a density of 1,000–2,000 clones per 150-mm plate and then transferred to nylon membranes. The human placental DNA (Sigma) was labeled with [α - 32 P]dCTP by the random primer labeling method (Feinberg and Vogelstein, 1983) and hybridized to the filters according to standard methods (Maniatis *et al.*, 1989). The positive clones were picked and transferred into individual wells of microtiter plates. After overnight culture, two replica filters were made using a 96-needle transferring device. One filter was hybridized with labeled human DNA and the another was hybridized with labeled hamster DNA. The clones that hybridized to human DNA with no detectable signal to hamster DNA were transferred into individual wells of microtiter plates, grown overnight, and then frozen at -80°C in LB media plus 20% glycerol and 100 $\mu\text{g}/\text{ml}$ ampicillin.

Hybrid mapping panel of chromosome 21. Figure 1 presents a diagram of the hybrid mapping panel. The following seven cell hybrids, containing partial chromosome 21 and dividing the chromosome 21 into eight regions, were used as template

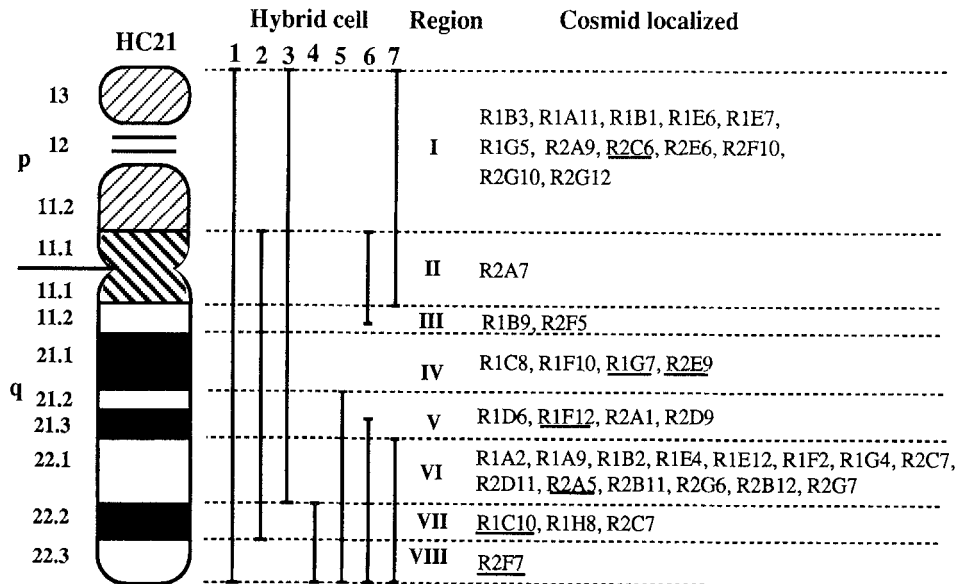


Fig. 1. Regional mapping of human chromosome 21-specific cosmids. (Left) Schematic representation of the human chromosome 21 contents of the cell hybrids. The vertical lines indicate the portion of chromosome 21 presents in each hybrid. The cell hybrids divide the chromosome 21 into eight subregions. 1, 153E9a3; 2, R2-10W; 3, 21q⁺; 4, 8q^{-a}; 5, 1881C-13b; 6, ACEM2-10d; 7, JC-6A. (Right) A total of 41 cosmids was localized in the eight subregions of chromosome 21 by differential hybridization of 229 cosmid clones with *Alu*-PCR products obtained from seven hybrid cells. The cosmid with an underline indicates that it contains *NotI* site.

DNAs for *Alu*-PCR amplification: 153E9a3, R2-10W (containing 21p11.1-q22.2), 21q⁺ (containing 21pter-q22.1), 8q^{-a} (containing 21q22.2-qter), 1881C-13b (containing 21q21.2-qter), ACEM (containing 21cen-q11.1 and q21.3-qter), and JC-6A (containing 21pter-q11.1 and 21q21.3-qter). Detailed analyses of these hybrids have been described elsewhere (Patterson *et al.*, 1985; Gardiner *et al.*, 1990).

Alu polymerase chain reaction (Alu-PCR). An *Alu*-primer 3144 (Monaco *et al.*, 1991) was used to amplify human DNA sequences from seven cell hybrids (Fig. 1). It was synthesized on an Applied Biosystems 380B DNA synthesizer and subsequently purified by OPC column. The PCR reaction was performed on a GeneAmp PCR System 9600 (Perkin Elmer Cetus). The PCR reaction mixture contained 1 µg of each cell hybrid DNA and 1.25 µM of the *Alu* primer in 100 µl of 0.01 M Tris-HCl (pH 8.3), 0.0015 M MgCl₂, 0.05 M KCl, 250 µM of each dNTPs, and 2.5 units of Cetus *Taq* polymerase. The template DNAs were initially denatured at 94°C for 2 min, 45 cycles of PCR were then performed, each cycle consisting of a 2-min denaturation at 94°C, a 2-min annealing at 57°C, and 4-min extension at 74°C. The final extension was carried out at 74°C for 10 min. The PCR products were analyzed by electrophoresis on a 1.2% agarose gel. The remainder was used as a probe after it was extracted with phenol-chloroform.

Differential hybridization. A method based on the differential hybridization of cosmid clones with amplified inter-*Alu* sequences from a set of somatic cell hybrids was used to localize the cosmids onto subchromosomal regions (Zucman *et al.*, 1992). The HC21-specific cosmid clones were spotted onto nylon membrane using a robotic device (Biomech 1000, Beckman), and grown at 37°C for overnight. The filters were baked and prehybridized for 2 h at 72°C in the hybridization solution of 1 M NaCl, 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, 1% SDS, 10% dextran sulphate, and 200 µg/ml sonicated human placental DNA (Sigma). One hundred nanogram PCR product from each cell hybrid was labeled by the random primer labeling method, and purified through Sephadex G-50. One hundred nanogram of the labeled DNA and 40 µg of human placental DNA in 400 µl of 0.5× hybridization solution (final) were denatured at 95°C for 7 min, and then incubated at 72°C for 2 h. After suppression, the probe DNA was added to the hybridization bag, and hybridized for 14 h at 72°C (Chumakov *et al.*, 1992b). The filters were washed with 2× SSC at room temperature for 15 min, and with 0.1× SSC, 0.1% SDS at 65°C for 30 min. The autoradiography was done for 24–36 h at –80°C with one intensifying screen.

Linear PCR. Linear PCR for creating end-specific probes of the clone insert was prepared by two independent reactions, using T3 and T7 primers, which correspond to the T3 and T7 promoter sequences adjacent to the cloning site of the vector (Caerano-Anollies *et al.*, 1991). The *Eco*RI-digested cosmid DNA was purified through phenol-chloroform extraction and then used as a template in linear PCR with [α -³²P]dCTP (Mizukami *et al.*, 1993). The PCR was done with 30 cycles, each cycle consisting of 1-min denaturation at 94°C, 1.5-min annealing at 53°C,

and 1-min extension at 74°C. The linear PCR products were directly hybridized to the arrayed cosmid filters according to the method as described previously (Mizukami *et al.*, 1993).

RESULTS

Library composition

A cosmid library containing 3×10^5 clones was constructed from the cell hybrid, 153E9a3, which contains human chromosome 21 as the only human chromosome. Of the library, 56,500 clones have been screened and the library composition was shown in Table 1. The result of the first screening with labeled human DNA indicate that 0.4% clones of the library showed positive signals to human DNA probe. The 282 clones obtained from the first screening were further hybridized to labeled human DNA and hamster DNA, respectively. Of which, 229 clones (81.2%) hybridized to human DNA, 24 clones (8.5%) to hamster DNA, 15 clones (5.3%) to both, and 14 clones (5.0%) to neither. The insert size of the cosmid clones was estimated as 38 ± 8 kb by digesting 10 cosmids with *EcoRI* endonuclease.

Testing cosmids for NotI restriction site

To investigate the presence of *NotI* site of the cosmid clones, 229 of HC21-specific cosmids were digested with five units of *NotI* endonuclease per microgram of DNA under the conditions recommended by the manufacturer, and then fractionated by electrophoresis in a 1% agarose gel. Since the vector, sCos-1, contains two *NotI* sites, more than two bands would be detected if the insert DNA contains *NotI* site. We found a total of 28 cosmids (12.2% of those tested) contained one or two *NotI* sites.

Chromosomal mapping of the cosmids

Human DNA sequences from somatic cell hybrids were amplified with the *Alu* primer 3144. The total complexity of the *Alu*-PCR products appears to correlate with that of the human contents of the respective hybrids, and no amplification product was detected in hamster DNA (data not shown). Each *Alu*-PCR product obtained from seven cell hybrids was labeled and hybridized to the replica filters of the arrayed HC21 cosmid clones with total human DNA suppression. Figure

Table 1. Assessment of library composition by species-specific colony hybridization.

Screening	Clones	Positive clones		Percentage
First screening	70,500	human	282	0.40%
Second screening	282	human	229	81.2%
		hamster	24	8.5%
		human+hamster	15	5.3%
		no signal	14	5.0%

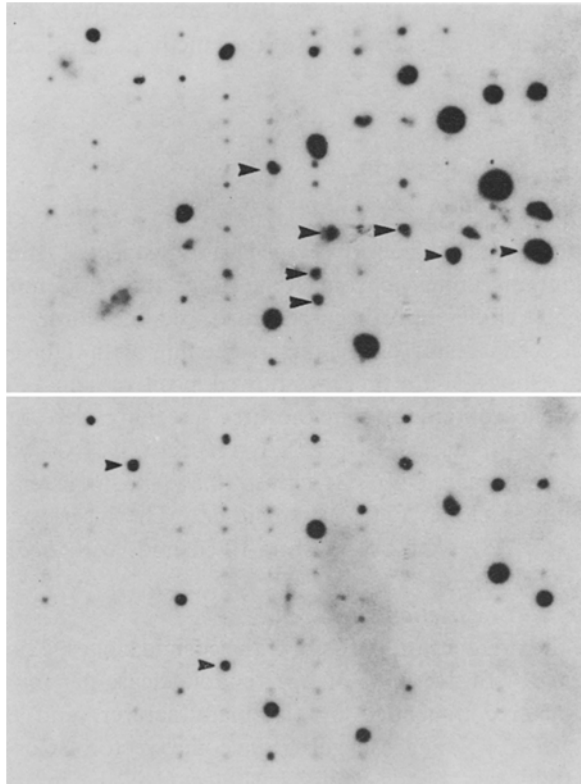


Fig. 2. Differential hybridization signals obtained when the *Alu*-PCR products from the cell hybrids R2-10W (upper) and JC-6A (lower) were hybridized to the same filter of cosmid clones. The positive clones are observed on the autoradiogram. Arrows indicate the cosmid clones showing obvious differential hybridization signals for the two probes.

2 shows the representative differential hybridization signals of cosmid clones to the *Alu*-PCR products obtained from cell hybrids, R2-10W and JC-6A. Using clear-cut stringent criteria for analysis of differential hybridization (Zucman *et al.*, 1992), essentially the positive signal cosmid could be unambiguously grouped in one of eight different regions according to its pattern of hybridization. This analysis identified 41 cosmids potentially located in the eight different regions of chromosome 21 (Fig. 1). Ten cosmids did not meet the clear-cut stringent criteria, and they were not further analyzed. To estimate the targeting efficiency of this method, we used a YAC clone, SOD1 that has been localized on 21q22.1 (region VI) (Imai *et al.*, 1990), as probe to screen the arrayed cosmid filters. Four positive clones (R1A9, R1B2, R1E4, and R1G4) that were potentially localized on region VI (Fig. 1) were obtained. Furthermore, DNA from these four cosmids was used for *in situ* hybridization to normal human metaphase, and a pair of strong signals

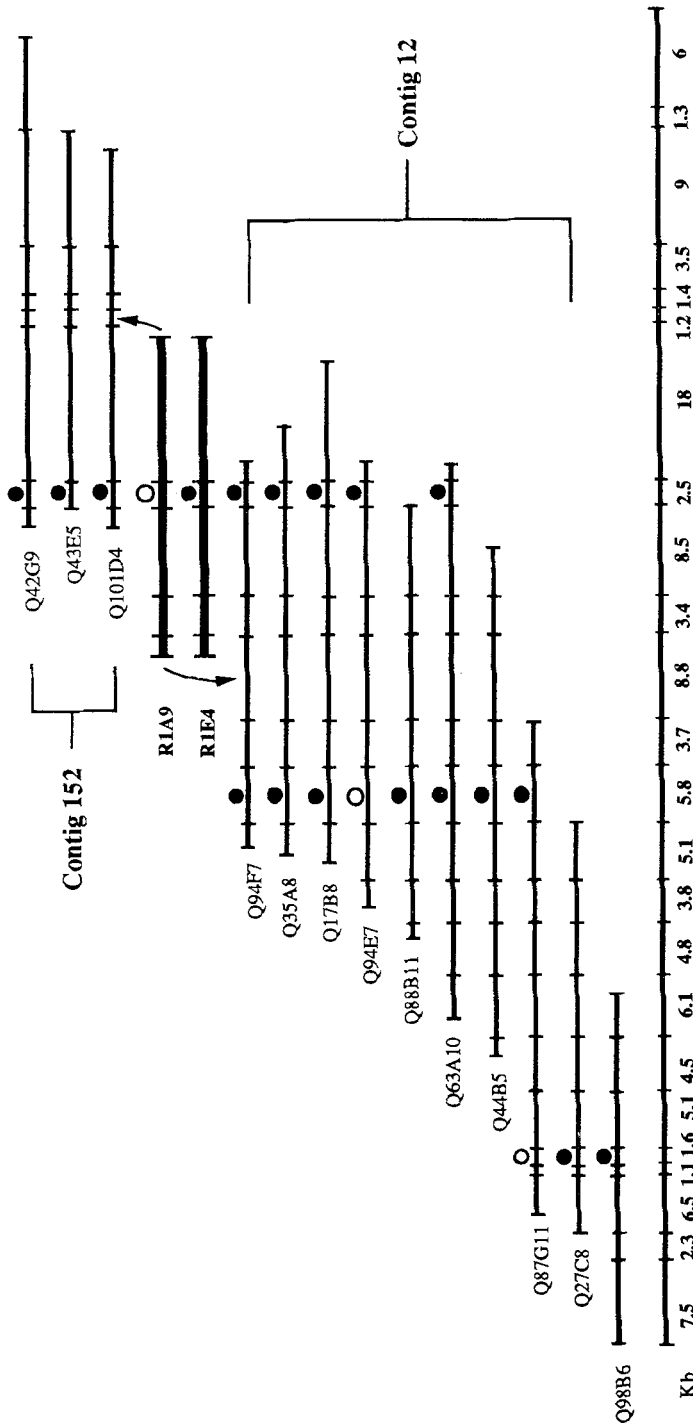


Fig. 3. Contig mapping and gap closure with the clone R1A9. The original contigs 12 and 152 assembled from a flow-sorted HC21 cosmid library by fingerprinted strategy were indicated by bars, while the endprobe walk of the clone R1A9 between two contigs were indicated by arrows. The contigs were further confirmed by *Eco*RI restriction enzyme mapping, and three junction fragments (circle symbols) were detected by Southern blot hybridization with the defined probes (open circles). The number at left indicates the cosmid number that was given by the originator. The contig connected with the clone R1A9 can span approximately 121.5 kb of the genome.

without background were observed in the region VI (data not shown).

Anchoring cosmid contigs onto chromosome 21 using the mapped cosmids

The end-specific probe of clone localized was generated by the inward-facing T3 or T7 promoter located at the clone site of the vector, and then hybridized to the arrayed cosmid filters. A typical example is shown in Fig. 3, the T3 and T7 end probes from the clone R1A9, that has been localized on region VI (21q22.1) by *Alu*-PCR method, was used to screen the arrayed cosmid clones from Livermore's library and our library. The T3-end probe hybridized to seven clones (Q94F7, Q35A8, Q17B8, Q94E7, Q88B11, Q63A10, and Q44B5) in existing contig 12, the T7-end probe hybridized to all of three clones in existing contig 152, and both of probes hybridized to the clone R1E4 that was a similar clone to R1A9. To define the overlap relationship between the clone R1A9 and two contigs, all of clones were digested with *EcoRI*, and then analyzed by Southern hybridization. The junction fragments were also detected on autoradiograms as unique fragments by using overlapping fragments as probes (Fig. 3). The results indicate that the gap between contigs 12 and 152, where no junction clone exists in the original library, can be closed up with the clone R1A9 and both contigs were mapped onto the region VI (21q22.1) of HC21. The localization of the contig 152 was also confirmed by PCR-specific detection of these clones with the primer 93B1012 that is a STS landmark localized at 21q22.1 (Hou *et al.*, unpublished data).

DISCUSSION

We have begun to construct the HC21 contig maps with cosmid clones. As a complementary step of this project, we have constructed a cosmid library from a human × hamster hybrid cell, 153E9a3, which contains HC21 as the only human chromosome. To map the cosmid clones onto chromosomal subregions rapidly, we carried out differential hybridization of the HC21-specific cosmid clones with amplified inter-*Alu* sequences obtained from somatic cell hybrids. Of these, 41 cosmids have been localized in eight subregions of HC21. The targeting efficiency of this method was evaluated by YAC-cosmid hybridization and *in situ* hybridization. The map location of cosmids derived by these methods showed excellent correspondence.

We have further identified 28 cosmids (12.2% of those tested) that contained one or two *NotI* sites. *NotI* site is present, as rare restriction site, in vertebrate genome and its distribution is related to the distributions of genes (Bird, 1986). It has been reported that 9.4% of chromosome 3-specific cosmids carried one or two *NotI* sites and they contained gene sequences (Golembieski *et al.*, 1991). Our cosmids also carried *NotI* site with similar ratio. These cosmids should be a valuable resource for the identification of chromosome 21-specific genes.

Finally, we used these localized clones to anchor our existing contigs on HC21

and to close the gaps between the contigs. Of course, we can also do these works by the fingerprinting method. But, we have noted that the fingerprinting method tends to assemble the additional, redundant clones into existing contigs, instead of covering gaps. To discover the small overlaps between clones, we used end-specific probes of a clone insert, made available by the inward-facing T3 and T7 promoters located at the cloning site of the cosmid vector. Our results indicate that some contigs can be anchored onto HC21 and some gaps between contigs can be effectively closed up with the end-specific probes of these cosmids localized. Thus, these cosmids provided useful anchor points for contig mapping and walking. Of course, we also recognized that all of gaps between 4,440 contigs are difficult to be closed up with our 229 cosmids. The final closure of these gaps still needs other cosmids or YAC clones. The works of contig anchoring and gap closure are in progress, the final results will be summarized in another paper.

Acknowledgments We thank Dr. David Patterson for kindly providing the hybrid cell lines, and Dr. Katy Walton for shipping all the cell lines. We also thank Mr. Kazuo Ozawa for technical help in the initial stage of the cosmid library construction. This work was supported by the human genome research program of RIKEN and the Science and Technology Agency of Japan.

REFERENCES

- Bird AP (1986): CpG-rich islands and the function of DNA methylation. *Nature* **321**: 209–213
- Branscomb E, Slezak T, Pae R, Galas D, Carrano AV, Waterman M (1990): Optimizing restriction fragment fingerprinting methods for ordering large genomic libraries. *Genomics* **8**, 351–366
- Caerano-Anollis G, Bassam BJ, Gresshoff PM (1991): DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology* **9**: 553–557
- Carrano AV, Lamerdin J, Ashworth LK, Watkins B, Branscomb E, Slezak T, Raff M, deJong PJ, Keith D, McBride L, Meister S, Kronick M (1989): A high-resolution fluorescence-based, semiautomated method for DNA fingerprinting. *Genomics* **8**: 129–136
- Christopher A (1993): Genome shortcut leads to problems. *Science* **259**: 1684–1687
- Chumakov I, LeGall I, Billaut A, Ougen P, Soularue P, Guillou S, Regault P, Bui H, De Tend MK, Barillot E, Abderrahim H, Cherif D, Berger R, LePaslier D, Cohen D (1992a): Isolation of chromosome 21-specific yeast artificial chromosomes from a total human genome library. *Nature Genet* **1**: 222–225
- Chumakov I, Regault P, Guillou S, Ougen P, Billaut A, Guasconi G, Gervy P, LeGall I, Soularue P, Grinas L, Bougueleret L, Bellanné-Chantelot C, Lacroix B, Barillot E, Gesnouin P, Pook S, Vaysseix G, Frelat G, Schmitz A, Sambucy JL, Bosch A, Estivill X, Weissenbach J, Vignal A, Riethman H, Cox D, Patterson D, Gardiner K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Paslier DL, Heilig R, Antonarakis S, Cohen D (1992b): Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* **359**: 380–387
- Evans GA, Lewis K, Rothenberg BE (1989): High efficiency vectors for cosmid microcloning and genomic analysis. *Gene* **79**: 9–20
- Feinberg AP, Vogelstein B (1983): A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6–13
- Gardiner K, Horisberger M, Kraus J, Tantravahi U, Korenberg J, Rao V, Reddy S (1990): Analysis of human chromosome 21: correlation of physical and cytogenetic maps: gene and CpG

- island distribution. *EMBO J* **9**: 25–34
- Golembieski WA, Smith SE, Recchia F, Judge A, Shridhar V, Miller OJ, Drabkin H, Smith DI (1991): Isolation of large numbers of chromosome 3-specific cosmids containing clusters of rare restriction-endonuclease sites. *Am J Hum Genet* **49**: 581–589
- Imai T, Iida A, Miwa T, Tashiro H, Song T-C, Yokoyama K, Soeda E (1990): Analysis of YAC clones by Pulsed-field gel electrophoresis: Physical mapping of Cu/Zn superoxide dismutase gene locus. *Methods: A Companion to Methods in Enzymology* **1**: 180–185
- Maniatis T, Fritsch EF, Sambrook J (1989): *Molecular cloning. A Laboratory Manual* (2nd ed), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
- Mizukami T, Chang WI, Kaplan N, Lombardi D, Matsumoto T, Niwa O, Kounosu A, Yanagida M, Marr TG, Beach D (1993): A 13 kb restriction cosmid map of the 14 Mb fission yeast genome by nonrandom sequence-tagged site mapping. *Cell* **73**: 121–132
- Monaco AP, Lam VMS, Zehetner G, Lennon GG, Douglas C, Nizetic D, Goodfellow PN, Lehrach H (1991): Mapping irradiation hybrids to cosmid and yeast artificial chromosome libraries by direct hybridization of *Alu*-PCR products. *Nucleic Acids Res* **19**: 3315–3318
- Patterson D, Van Keuren M, Drabkin H, Watkins P, Gusella JF, Scoggin C (1985): Molecular analysis of chromosome 21 using somatic cell hybrids. *Ann NY Acad Sci* **450**: 109–120
- Schlessinger D (1990): Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet* **6**: 248–258
- Stelnmetz M, Winoto A, Minard K, Hood L (1982): Clusters of genes encoding mouse transplantation antigens. *Cell* **28**: 489–498
- Stewart GD, Van Keuren ML, Galt J, Kurachi S, Buraczynska ML (1989): Molecular structure of human chromosome 21. *Annu Rev Genet* **23**: 409–423
- Zucman J, Delatter O, Desmaze C, Azambuja C, Rouleau G, Jong PD, Aurias A, Thomas G (1992): Rapid isolation of cosmids from defined subregions by differential *Alu*-PCR hybridization on chromosome 22-specific library. *Genomics* **13**: 395–401