

Construction of a BAC Contig Map of Chromosome 16q by Two-Dimensional Overgo Hybridization

Cliff S. Han,¹ Robert D. Sutherland, Phillip B. Jewett, Mary L. Campbell, Linda J. Meincke, Judy G. Tesmer, Mark O. Mundt, John J. Fawcett, Ung-Jin Kim,² Larry L. Deaven, and Norman A. Doggett

DOE Joint Genome Institute, Bioscience Division and Center for Human Genome Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545 USA; ²Division of Biology, California Institute of Technology, Pasadena, California 91125 USA

We have used sequence-based markers from an integrated YAC STS-content/somatic cell hybrid breakpoint physical map and radiation hybrid maps of human chromosome 16 to construct a new sequence-ready BAC map of the long arm of this chromosome. The integrated physical map was generated previously in our laboratory and contains 1150 STSs, providing a marker on average every 78 kb on the euchromatic arms of chromosome 16. The other two maps used for this effort were the radiation hybrid maps of chromosome 16 from Whitehead Institute and Stanford University. To create large sequenceable targets of this chromosome, we used a systematic approach to screen high-density BAC filters with probes generated from overlapping oligonucleotides (overgos). We first identified all available sequences in the three maps. These include sequences from genes, ESTs, STSs, and cosmid end sequences. We then used BLAST to identify 36-bp unique fragments of DNA for overgo probes. A total of 906 overgos were selected from the long arm of chromosome 16. Hybridizations occurred in three stages: (1) superpool hybridizations against the 12× coverage human BAC library (RPC1-II); (2) two-dimensional hybridizations against rearranged positive BACs identified in the superpool hybridizations; and (3) pooled tertiary hybridizations for those overgos that had ambiguous positives remaining after the two-dimensional hybridization. For the superpool hybridizations, up to 236 overgos have been pooled in a single hybridization against the 12× BAC library. A total of 5187 positive BACs from chromosome 16q were identified as a result of five superpool hybridizations. These positive clones were rearranged on membranes and hybridized with 161 two-dimensional subpools of overgos to determine which BAC clones were positive for individual overgos. An additional 46 tertiary hybridizations were required to resolve ambiguous overgo-BAC relationships. Thus, after a total of 212 hybridizations, we have constructed an initial probe-content BAC map of chromosome 16q consisting of 828 overgo markers and 3363 BACs providing >85% coverage of the long arm of this chromosome. The map has been confirmed by the fingerprinting data and BAC end PCR screening.

The major initial objectives of the Human Genome Project included the construction of genetic and physical maps of each human chromosome (Botstein 1990). Such maps permit rapid access to any chromosomal region, facilitating the identification of disease genes and accelerating our understanding of many aspects of cellular and molecular biology. Furthermore, chromosomal physical maps constructed as cloned DNA segments supply the necessary materials to support large-scale DNA sequencing, the ultimate goal of the Human Genome Project. Yeast artificial chromosome (YAC)-based physical maps have been constructed for the euchromatic regions of chromosome 16 (Doggett et al. 1995) and as a framework covering significant stretches of the entire genome (Cohen et al. 1993; Chumakov et al. 1995). YACs have an advantage over other

cloning systems in that their large size (~350- to >1000-kb average size for most libraries) facilitates the construction of maps with long-range continuity (Burke et al. 1987). However, YACs are poor substrates for DNA sequencing, and most YAC libraries have high rates of chimerism and deletions that can further limit their utility (Green et al. 1991; Selleri et al., 1992). Bacterial artificial chromosome (BAC) clones (~120- to 200-kb average size for available libraries) offer high clonal stability and reduced cloning biases and are easily purified into DNA suitable for DNA sequencing (Shizuya et al. 1992). There are two major collections of BAC libraries (Kim et al. 1996; Osoegawa et al. 1998) that are currently used for most large-scale mapping and sequencing efforts.

Screening of BAC libraries for map construction is possible by either a PCR- or hybridization-based approach. PCR-based screening requires pooling of the

¹Corresponding author.
E-MAIL chan@telomere.lanl.gov; FAX (505) 665-3024.

library into sets of clones (often involving combinations of plate, row, and column pools) such that each clone can be uniquely identified by its detection in the sets of clones in which it occurs (Green and Olsen 1990; Bruno et al. 1995). Hybridization-based screening can be performed against high-density gridded macroarrays of an entire library using radioactively labeled probes either singly or as groups. The most commonly used probes for hybridization have been subcloned DNA fragments, PCR amplicon products or DNA oligonucleotides. Subcloned fragments or PCR amplicons, which are usually several hundred base pairs, have been preferred over oligonucleotides because of their ease of labeling by nick translation or random priming to generate high specific activity probes but have the disadvantage of having a higher likelihood of containing repeat elements. With oligonucleotides it is easy to avoid repeats, but end-labeling does not generate very high specific activity probes, and hybridization kinetics are not as strong as for the larger probes. A novel approach for making oligonucleotide probes developed by John McPhearson (Ross et al. 1999) maintains the advantage that oligonucleotides have for control of sequence content but yields slightly larger probes with better hybridization kinetics and higher specificity labeling than can be achieved for conventional oligonucleotide probes. These probes, termed overgos are made by annealing two overlapping oligonucleotides and filling in the overhanging bases with Klenow enzyme and radiolabeled nucleotides. For the effort presented herein, we have slightly modified the labeling protocol and developed a stringent criteria for the selection of sequences used for overgo probes. Following this approach, we have been able to pool >200 overgo probes in a single hybridization for highly paralleled hybridization-based screening of a 12 \times coverage total genomic human BAC library.

We had constructed previously an integrated physical map of human chromosome 16 that provided essentially complete coverage in YACs, which were ordered by STS content to somatic cell hybrid breakpoint intervals (Doggett et al. 1995). Continued efforts advanced this map to include a total of 1150 STSs and ~500 cosmid contigs of 110-kb average size. For an ~20-Mb region extending from 16p13.11 to p11.2, we sup-

plied the available STSs and clones to Caltech to facilitate their efforts at constructing a BAC contig of this interval (Cao et al. 1999). To create long-range continuity in sequenceable targets of an 11-Mb region from 13p13.3 to 16p13.12 (distal and adjacent to the Caltech region) and of the 45-Mb euchromatic long arm of this chromosome, we developed a systematic approach for screening high-density BAC filters with evenly spaced overgo probes. We first identified all available sequences in the integrated map. These include sequences from genes, ESTs, STSs, and cosmid end sequences. We then used BLAST to identify 36-bp unique fragments of DNA for use as overgo probes. Pooled overgo hybridizations were performed against the 12 \times coverage human BAC library (RPCI-11, sections 1 and 2). Positive BACs identified from these hybridizations were rearranged on membranes and hybridized with two-dimensional subpools of overgos to determine which BAC clones were positive for individual overgos. In some instances a third hybridization was included to resolve ambiguous probe-BAC relationships that remained after the two-dimensional hybridizations. Probe-content BAC contigs were constructed in this manner. Restriction mapping was subsequently performed to select the optimal tiling sets for sequencing. Here, we report on the mapping of 828 overgos to generate ~40 Mb of probe-content BAC contigs covering ~85% of the long arm of chromosome 16.

RESULTS

Overgo Development

A total of ~1000 STSs were available for the long arm of chromosome 16 from three sources—the integrated chromosome 16 physical map generated at Los Alamos National Laboratory (Doggett et al. 1995), the RH and STS-content maps generated at Whitehead Institute (<http://www-genome.wi.mit.edu>), and the STSs markers and RH maps developed at the Stanford Human Genome Center (<http://shgc.stanford.edu>) (Table 1). The group of 402 overgos labeled STS-YAC were developed from sequences within the STS-content YAC map portion of the LANL integrated chromosome 16 physical map and are finely ordered. The 128 overgos within group STS-SCH were derived from STSs mapped to only the somatic cell hybrid breakpoint map compo-

Table 1. STS Sources Used for Overgo Development and BAC Library Screening

Source ^a	LANL		Whitehead	Stanford		Total
	STS-YAC	STS-SCH		RH map	unmapped	
No. of overgos	402	128	166	64	148	908

^aSee text for description.

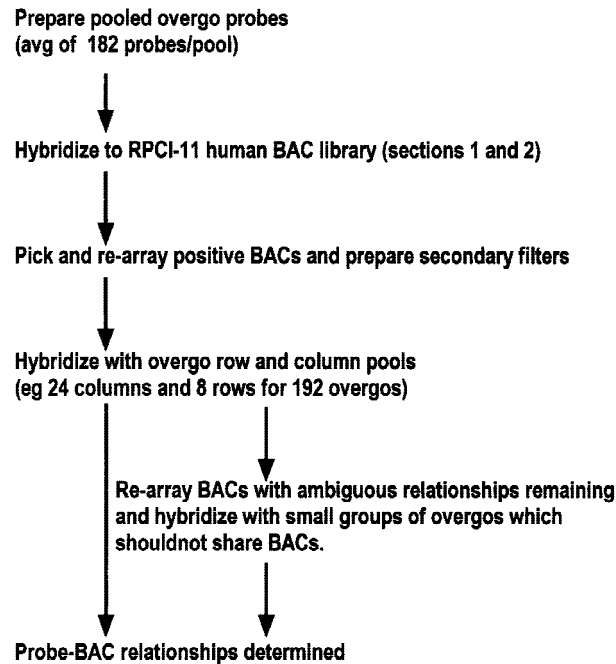


Figure 1 Overall two-dimensional overgo hybridization process.

ment of the LANL integrated physical map. One hundred sixty-six overgos were derived from Whitehead Institute maps, 66 overgos were developed from Stanford RH mapped STSs, and another 148 were developed from only chromosome 16 assigned but not further mapped STSs (unmapped). STSs occurring in multiple maps are listed only once. Overgos from the LANL STS-content YAC map served as the backbone markers for construction of the BAC map. The criteria applied to select overgos from STSs sequences are described in Methods. Of 964 STS sequences, 908 (94%) met these criteria for overgo probe generation. All of the overgo sequences are at our web site, http://www-ls.lanl.gov/Chr_16_mapping_Home/chr16map.html.

Top Pool Hybridizations

The overall hybridization strategy is shown in Figure 1. The 908 overgos were divided into five pools for the top pool hybridizations of the RPCI-11 human BAC library. Pool size ranged from 96 to 236 overgos for each hybridization. The results of these screenings are

summarized in Table 2, and representative hybridizations are shown in Figure 2. A total of 5177 BACs were positive as a result of these hybridizations. These BACs, plus 413 BAC clones mapped by Caltech to chromosome 16q, were rearrayed in 96-well microtiter plates and used to produce membrane sets (8- × 12-cm filters, eight filters/set) for subsequent two-dimensional hybridizations. On average, 61.5% of BACs that were identified in top pool hybridizations were confirmed by subsequent hybridizations, whereas 38% (157) of the BACs from Caltec were confirmed.

Two-Dimensional Hybridizations

Two-dimensional overgo hybridizations were performed against the rearrayed positive BAC clones to determine probe-BAC relationships. The results of these hybridizations are summarized in Table 3 and a sample row and column hybridization is shown in Figure 3. These subpool hybridizations were performed in five groups, according to the five different sources of overgos. The overgos in each group were pooled by row and column. All row pools included a fixed number of overgos (24). The column pools contained a varying number of overgos (3–17) depending on the number of overgos in the group. One consequence of the two-dimensional pooling strategy is that ambiguous positives can be inferred when a BAC clone is positive with multiple row pools and column pools. We attempted to reduce the likelihood of such occurrences by ordering the overgos along each row following their known map order. This method of ordering increased the likelihood that all positive overgos for a single BAC were placed in a single row.

After completing the two-dimensional hybridizations, a total of 458 overgos and 371 BACs were involved in 2343 ambiguous probe-BAC relationships that occurred when a BAC was positive with multiple row and column pools. These 458 overgos were divided in 46 groups for tertiary hybridizations against the 371 BAC clones. Each of the 46 groups contained overgos that did not share any of the possible positive BAC clones. Six hundred eighty-seven of the probe-BAC relationships were confirmed by these 46 tertiary hybridizations. After completion of all hybridizations, 828 overgos (91%) were located to at least one BAC clone.

Table 2. Top Pool Hybridization Results

Pool name	1	2-9	3-4	5-6	7-8-10	Total	Caltec
No. of overgos	96	192	192	192	236	908	—
No. of positive BACs	620	1941	1937	1955	2688	5177 ^b	413
Percent of BACs confirmed ^a	85.2%	75.8%	58%	69.5%	73.6%	61.5%	38%

^aConfirmation by two-dimensional and tertiary hybridizations.

^bCommon positives between groups are the reason that the number of the total positives is less than the combination of the positives of the five groups.

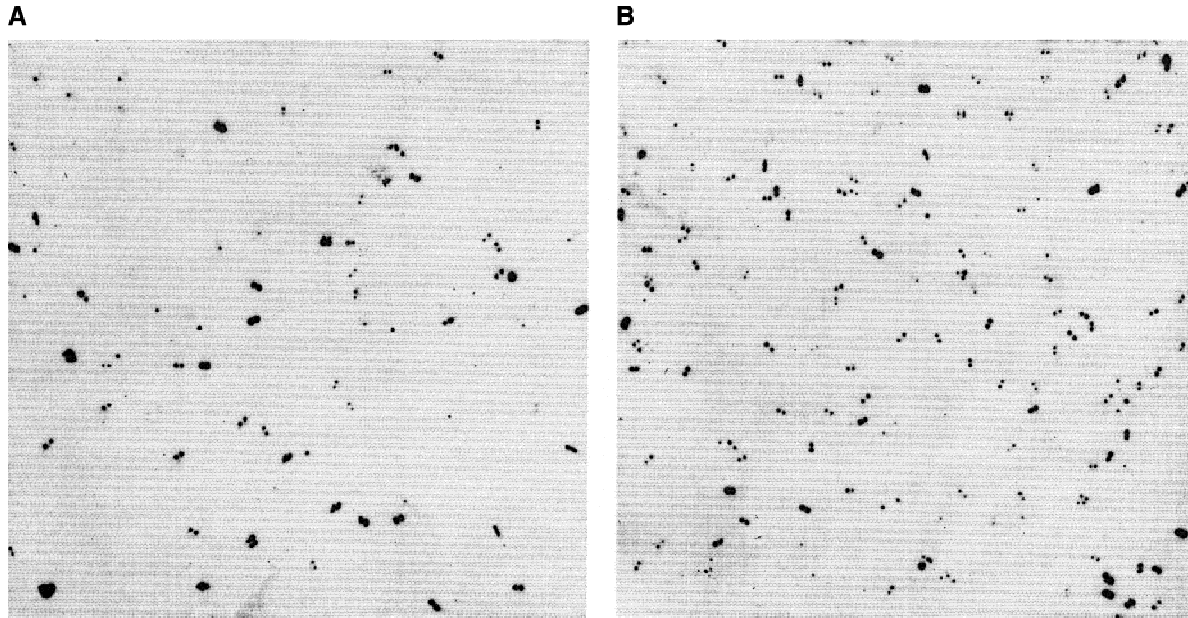


Figure 2 Top pool hybridization. (A) Pool 1 with 96 overgos hybridized with the seventh high-density filter of the 12 RPCI-11 human BAC library. (B) Pool 3–4 with 192 overgos hybridized with the fourth high-density filter of the same BAC library. There are six panels on each filter and $4 \times 4 \times 384$ plates gridded on each panel. Each plate is gridded in duplicate in different orientations on the membranes. The signals are graded according to their strength into four groups: strong, media, weak, and very weak. The signal in the first three group are picked up as positives.

The average number of BACs that were positive for each overgo probe was 10.6 (Table 3). The individual overgo screening results are available at http://www-ls.lanl.gov/Chr_16_mapping_Home/chr16map.html.

Contig Construction and Validation

Overgo-content BAC contigs were constructed according to the hybridization results, using the LANL STS–YAC-derived overgos as the framework for ordering. The overgos from other groups were integrated into the map according to the common BACs hybridized. Finally, 53 contigs were constructed with 6 of these >3 Mb. These contigs cover >85% of the long arm of chromosome 16.

To confirm the overgo–BAC contigs and to select

the minimal tiling set of BACs for high throughput sequencing, 3363 clones were fingerprinted by *EcoRI* digestion. One hundred six of these clones failed in fingerprinting. We used a software tool GRAM (Soderlund and Burks 1994) to assemble the restriction contigs prior to selection of clones for sequencing. A summary restriction map covering 16q12.1–q12.2 is shown in Figure 4. There were 561 BACs identified with the 161 overgos from this region. These clones were evaluated using GRAM in restriction map assembly. Four hundred thirty-seven (75%) of the clones are fit well into the restriction mapped contig. The majority of the remaining clones partially fit with the restriction map, indicating to us that these most likely were derived from duplicated regions of the genome and regions

Table 3. Results of Two-Dimensional Hybridization

Source ^a	LANL		Whitehead	Stanford		Total
	STS–YAC	STS–SCH		RH map	unmapped	
No. of overgos	402	128	166	64	148	908
Column pools	24	24	24	24	24	120
Row pools	17	6	8	3	7	41
Overgos with at least one positive BAC	366 (91%)	120 (94%)	147 (89%)	60 (95%)	135 (92%)	828 (91%)
Overgos requiring 3 hybridizations ^a	234 (64%)	61 (51%)	55 (37%)	12 (20%)	96 (71%)	458 (50%)
Average hits per overgo	12.6	11.4	10.3	10.2	9	10.6

^aTo resolve ambiguous BAC–probe relationships (see text for details).

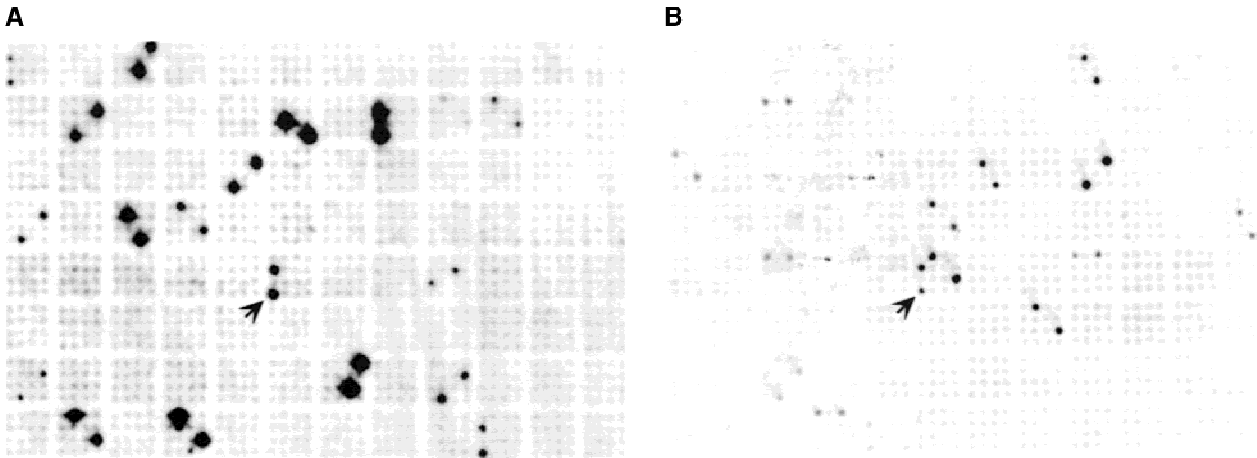


Figure 3 Two-dimensional hybridization. (A) Probe pool, column 17 of the Whitehead group; filter number 5. (B) Probe pool, row 6 of the Whitehead group; filter number 5. Arrows point to the same positive clone hybridized with the two pools. The overgo probe in common at the intersection with the two pools is WI-14879.

containing low abundance repeats. Less than 10% of the clones showed no significant match to the restriction map. These false positives could have arisen from overgo sequences that were duplicated in the genome, but whose surrounding restriction map was not conserved, or from a low level of cross-contamination in the BAC library.

Portions of the overgo-BAC map have also been independently confirmed by PCR. For gap closing efforts we have pooled the same RPCI-11 BAC library following an eight-dimensional pooling strategy. PCR primers were generated from BAC end sequences adjacent to gaps in the overgo-BAC map (Fig. 4). To date, 213 PCR screenings with BAC end sequences have been completed. We found that over two-thirds of the BACs identified were previously located on the overgo-based map. The new clones identified by STS screening could readily be incorporated in overgo-content map in a manner that was consistent with the existing data, including regions that were analyzed at the restriction fragment level. The STS-based map data are also available at our web site, http://www-ls.lanl.gov/Chr_16_mapping_Home/chr16map.html.

DISCUSSION

Overgo Hybridization

The advantages of the overgo hybridization method include low background of hybridization, low rate of false positives, high throughput, and ease of handling compared with conventional cloned probes. The foundation of these advantages is the short sequence (36 bp) that is needed for an overgo probe, which increases the likelihood of finding suitable single copy regions from STS sequences. The importance of uniqueness of each overgo in highly pooled hybridizations is critical,

and thus, the selection criteria for uniqueness should be quite stringent. We have found that screening against repeat databases is not sufficient for overgo selection. An early problem with our first overgo hybridizations demonstrated this fact. A pool of eight overgos hybridized to >1000 clones on a single $1 \times$ BAC library membrane (data not shown). The eight overgos were picked using a popular primer program and checked against a repeat sequence database. The overgos were subsequently screened against GenBank, and two of them were found to be >90% similar with >10 sequence entries. BLAST searches against GenBank are now an integral part of our selection criteria for overgos (see Methods).

The overgo labeling method used in this paper was modified slightly from the original protocol kindly provided by Dr. McPherson (Washington University, St. Louis, MO). Cold dATP and dCTP were added to the reaction after labeling for 1 hr. This made it possible for the overgos to be fully filled in at both ends. This was necessary because the concentration of radiolabeled dATP and dCTP was only $0.165 \mu\text{M}$ in the labeling reaction compared with a $1 \mu\text{M}$ concentration for the overgo. We also found that better results were obtained by labeling the overgos individually and then pooling them for hybridization. For this effort we pooled 96–236 overgos together for the top pool hybridizations. The rate of false-positive hybridization ranged from 15% for a pool of 96 overgos to a maximum of 42% for one of the pools of 192 overgos. We found, however, that even this higher rate of false positives for the top-level pool results is tolerable for large-scale mapping because these are eliminated in subsequent hybridizations.

After the two-dimensional hybridization, 9% of the overgos failed to pick up any BAC clones. We checked these overgos and the original STS sequences.

We believe that a majority of these failures may be attributable to base-calling errors in the original STS sequence within the region selected for the overgo. Remaining failures may be attributable to secondary structures occurring within the overgos or failures of the labeling reaction.

Two-Dimensional Hybridizations

A two-dimensional pooling strategy can result in ambiguous positive results even when all overgos in the pools are unique. These ambiguous positives occur when a BAC is positive with multiple rows and columns of overgos. In the case of two positive overgos occurring in different rows and columns, there are two positive rows and two positive columns, resulting in four intersection points (overgo positions) half of which are true positives. Several conditions contribute to the rate of ambiguous positives in a two-dimensional hybridization. The first is the density (D) of overgos in the region covered by the two-dimensional hybridization. The higher the density of markers the greater the possibility of producing ambiguous positives. The useful limit for the overgo density (D_{MAX}) is the maximum number of overgos in a row or column of the array (N) divided by the average length of BACs (B), $D_{MAX} = N / B$.

In theory, if $D \geq D_{MAX}$, the two-dimensional hybridization will fail without locating any BACs to individual overgos. If, on the other hand, $D \leq 1/BMAX_{TMAX}$ and overgos are evenly distributed, then no ambiguous positives will occur in the two-dimensional hybridization. A second factor is the average length of the BAC insert. The longer the average length of BAC inserts, the greater the possibility of producing ambiguous positives, which can be seen in the equation above. The third factor is the ordering of overgos within the two-dimensional array. With random order of overgos in the two-dimensional array, there is a greater possibility of producing ambiguous positives than occurs in two-dimensional arrays that are ordered according to the order of overgos on an accurate map. The three factors that contribute to the frequency of ambiguous positives in the two-dimensional hybridization are interrelated. For example, if the density of the overgos is so low that no two overgos share a common BAC, then ordering of overgos in the two-dimensional array is not an issue. Also, as BAC size decreases, a greater density of markers is tolerated.

The hybridization results in this paper testify to the theory exactly (Table 2). The three groups of overgos from the LANL STS–YAC, Whitehead and Stanford RH maps are all well ordered on chromosome 16q. For these groups, the ambiguous positive rate (percent of overgos requiring third hybridization) decreases with the number of overgos in the groups. Overgos within the group LANL STS–SCH are not as finely ordered and

have a higher ambiguous positive rate than the Whitehead group that has fewer overgos. Overgos derived from the Stanford unmapped STSs had the highest rate of ambiguous positives in the two-dimensional hybridizations. The following suggestions could be made to the experimental design to reduce the ambiguous positive rate for the two-dimensional hybridizations: (1) decrease the number of overgos in each dimensional hybridization; (2) increase the number of overgos in one dimension; and (3) expand the target region covered in each dimensional hybridization. Greater efficiency is obtained by increasing the target for mapping. For example, the whole mouse genome could be mapped with 20 $16 \times 12 \times 12$ three-dimensional hybridizations involving 46,080 overgos. The first two alternatives require more work and reduce the efficiency for mapping by multidimensional hybridizations.

Overgo–BAC Map

The overgo–BAC map, which contains 828 overgos and 3363 BACs, was constructed according to the overgo content for all BACs. The results of overgos from the LANL STS–YAC group were used as a backbone. The results of the overgo–BAC map largely confirmed those of the STS–YAC map. The greatest rearrangement occurred within an 800-kb region of 16q13. The results of the other overgo groups were integrated into the backbone map.

There are ~38 gaps in the overgo–BAC map as determined by probe content. The actual number of physical gaps may be substantially less because of undetected overlaps. The average contig size is ~750 kb. The average number of BACs per overgo is 10.6. There are 61 overgos that hybridized to >20 BACs each. Three of these overgos, which hybridized to 79, 69, and 47 BACs, respectively, were rejected. The remaining 58 overgos hybridized to an average of 24 BACs and were kept in the map for the following reasons: (1) The pattern that these BACs produce in the probe content map suggests that these clones occur in duplicated regions of chromosome 16; and (2) FISH performed for some of these BACs have confirmed that they are duplicated on chromosome 16 or contain low abundance repeats.

The overgo–BAC map was also confirmed by restriction mapping of the contigs covering 16q12.1–q12.2. Subsequent to the two-dimensional hybridization, all of the positive BACs were fingerprinted. Of the first grouping of 561 BAC clones that were analyzed by restriction mapping, 437 fit well into contigs that were consistent with their ordering in the probe-content map. The remaining clones did not fit well within the restriction maps. We note that for most of these cases the BACs were only positive for a single overgo in the contig. We believe that most of these were false posi-

tives arising from chromosome 16 duplications or low abundance repetitive sequences.

We have also compared our map with the contigs generated at Washington University by fingerprinting of the RPCI-11 BAC library. We had identified 3530 BACs from chromosome 16q (3363 by overgo hybridization plus 167 by either PCR screening of the RPCI-11 library or by BAC-end sequence searching from finished clones). Of these, 1826 (51.7%) were members of 658 Washington University BAC contigs. The remainder were not found in the contig sets that were available from the Washington University ftp site (http://genome.wustl.edu/gsc/human/human_ftp.pl). There were a total of 10,882 BACs in the Washington University contigs that contained at least one BAC on our map. Thus, only 16.8% (1826 out of 10,882) of BACs in Washington University contigs linked to our chromosome 16q map were included in our map. If we eliminated very large contigs with >40 BACs, then 22.7% (1826 out of 8046) of BACs in 616 contigs with at least one link to the chromosome 16q map were found in this map. Finally, if we required that at least two BACs from our map be contained in a fingerprint contig and also eliminated contigs with >40 members, then 66% (1826 out of 2766) of BACs in 236 contigs were included in the 16q map. These data suggest that confirmation of overlap between the query clone and at least one clone in contig retrieved is recommended when using the Washington University databases to extend clone coverage in the region of the query clone.

Summary

Using overgo two-dimensional hybridization, we have constructed an overgo-content BAC clone map, which covers ~85%–90% of the long arm of human chromosome 16. BACs in the map have been fingerprinted to construct restriction-mapped contigs to select optimal tiling sets for genomic sequencing. The improved overgo hybridization methods described here are well suited for genome-scale physical mapping efforts.

METHODS

Overgo Hybridization

The overgo hybridization method was adapted from a protocol we originally received from John McPhearson. Briefly, overgos were selected following a BLASTsearch against GenBank to screen out repeated sequences and then purchased from Genosys. Overgos were labeled with [³²P]dATP and

[³²P]dCTP in overgo labeling buffer (Ross et al. 1999) at 37°C for 30 min. One microliter of cold dATP and dCTP solution (200 μM) was then added to the labeling tube, and the reaction was incubated at 37°C for another 30 min. Background overgo (BKG-F, aacgggagtgatgtaaaata; BKG-R, tgatgggatcgggctattttac), which was selected from K12 region of *Escherichia coli*, was labeled in the same manner, but only 0.1 μl of the reaction was used for every 20 ml of hybridization solution in the following steps. After removal of unincorporated nucleotides using Sephadex G50 spin columns, probes were denatured at 95°C for 4 min and added to the hybridization tubes. The hybridizations were performed at 58°C for 18 hr in hybridization solution (50 ml of 1% BSA, 1 mM EDTA, 0.5 mM EDTA at pH 8.0, 7% SDS, 0.5 mM sodium phosphate at pH 7.2). Filters were washed and exposed to film at 70°C overnight. A more detailed protocol is available at our Web site

BAC-End Walking

PCR Screening

The RPCI-11 library (sections 1 and 2) was pooled following an eight-dimensional pooling strategy designed by David Torney (Los Alamos National Laboratory, Los Alamos, NM) in increments of a quarter library at a time, each a 3× coverage of human genome containing 55,296 clones. Ninety-four pools were made from each quarter of the library including two sets of plate pools (11 pools each), two sets of row pools (12 pools each), two sets of column pools (12 pools each), and one set of left and one set of right diagonal pools (12 pools each). Each of these eight sets of pools was made from a plate-rearranged 24×24 array of 96-well microtitre dishes. Thus, most pools contain 4608 clones. The plate rearrangements were selected to minimize the numbers of coincidences of pairs of clones in pools. For the row, column, and diagonal pools, the “lines” of clones from the array were combined to make the final 12 pools, with no pool containing more than one line incident on any plate. We screened the library pools with PCR primers designed from BAC-end sequences. PCR screening results were decoded with a decoding program. Clones identified from the PCR screening were integrated into the existing map.

Sequence-Based Searching

In instances where we had generated either draft or finished sequence for BAC clones that were adjacent to gaps, we pursued a walking strategy that was dependent on searching the BAC End Sequence Database at TIGR (http://www.tigr.org/tdb/human/bac_end_search/bac_end_intro.html). BAC genomic sequences were masked with the RepeatMasker (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>) prior to searching for identities in the BAC-end sequence database at TIGR. A minimal 95% identity was required to be considered as sequence match. The clones identified by sequence-based searching were integrated into the existing map.

Figure 4 Detailed contig map of 16q12.1–q12.2. The consensus restriction map is shown at *top* in black and cyan. The bold regions are covered by clones that are sequenced or being sequenced. The size and order of the fragments in these regions are confirmed or partially confirmed at the genomic sequence level. Only fragments that are >2 kb are listed. The markers are listed vertically under the restriction map. The lines connecting markers with the restriction map indicate that these markers have been found within the draft or finished sequence of the BACs. The markers in black are overgos that were screened by hybridization, those in magenta are STSs that were screened by PCR, and those in green are sequence identities identified by searching the BAC-end database with the sequences of the sequenced clones. The arrows in bold indicate the backbone markers from the previous LANL STS–YAC map. The BACs with names in red have been confirmed to map to 16q12.1–q12.2 by FISH. The BACs whose names sbegin with “R” are from the RPCI-11 library, and the rest are from the BAC library of the California Institute of Technology. (*FRAGMENT) Gap closed by restriction data only.

Contig Construction and Map Building Up

Restriction Mapping

BAC DNAs prepared in 96-well format were digested with *EcoRI* (New England Biolab, Beverly, MA). RNase was added before loading the samples onto the 0.7% LE agarose gels that were run at 35 V for 16 hr at room temperature. The gels were stained for 30 min with SybrGreen. Gel images were acquired using a Molecular Dynamics FluorImager 595. Images were analyzed with the BiImage Advanced Quantifier software. The fragment data of clones from the same region of the map were collected as input for GRAM to build restriction contigs.

Overgo-BAC Map Construction

The overgo-BAC map was assembled according to the probe content for all BACs. The results from overgos derived from the LANL STS-YAC map were used as a backbone for initial ordering. The map was maintained in Microsoft Excel and translated into a graphical output with the software PRTMP developed by C. Han, unpubl.

ACKNOWLEDGMENTS

This project was supported by the U.S. Department of Energy, OBER, under contract W-7405-ENG-36. We thank Dr. Paul S. White for help in the preparation of the manuscript.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Botstein, D., C. Cantor, J. Carbonell, A. Carrano, T. Caskey, F. Collins, U. Francke, E. Lander, L. Lerman, R. Moyzis et al. 1990. In *The Human Genome Project: The first five years FY 1991-1995*. DOE/ER-0452P, NIH 90-1590.
- Bruno, W.J., E. Knill, D.J. Balding, D.C. Bruce, N.A. Doggett, W.W. Sawhill, R.L. Stallings, C.C. Whittaker, and D.C. Torney. 1995. Efficient pooling designs for library screening. *Genomics* **26**: 21-30.
- Buckler, A.J., D.D. Chang, S.L. Graw, J.D. Brook, D.A. Haber, P.A. Sharp, and D.E. Housman. 1991. Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci.* **88**: 4005-4009.
- Burke, D.T., G.F. Carle, and M.V. Olson. 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806-812.
- Cao, Y., H.L. Kang, X. Xu, M. Wang, S.H. Dho, J. RyulHuh, B.-J. Lee, F. Kalush, D. Bocska, Y. Ding et al. 1999. A 12 Mbp complete coverage BAC contig map in human chromosome 16p13.1-11.2. *Genome Res.* **9**: 763-774.
- Chumakov, I.M., P. Rigault, I. Le Gall, C. Bellanné-Chantelot, A. Billault, S. Guillou, P. Soularue, G. Guasconi, E. Poullier, I. Gros et al. 1995. A YAC contig map of the human genome. *Nature (Suppl.)* **377**: 175-297.
- Cohen, D., I. Chumakov, and J. Weissenbach. 1993. A first-generation physical map of the human genome. *Nature* **366**: 698-701.
- Doggett, N.A., L.A. Goodwin, J.G. Tesmer, L.J. Meincke, D.C. Bruce, L.M. Clark, M.R. Altherr, A.A. Ford, H.C. Chi, B.L. Marrone. 1995. An integrated physical map of human chromosome 16. *Nature (Suppl.)* **37**: 335-365.
- Green, E.D. and M.V. Olson. 1990. Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain-reaction. *Proc. Natl. Acad. Sci.* **87**: 1213-1217.
- Green, E.D., H.C. Riethman, J.E. Dutchik, and M.V. Olson. 1991. Detection and characterization of chimeric yeast artificial-chromosome clones. *Genomics* **11**: 658-669.
- Kim, U.J., B.W. Birren, T. Slepak, V. Mancino, C. Boysen, H.L. Kang, M.I. Simon, and H. Shizuya. 1996. Construction and characterization of a human bacterial artificial chromosome library. *Genomics* **34**: 213-218.
- Osoegawa, K., P.Y. Woon, B. Zhao, E. Frengen, M. Tateno, J.J. Catanese, and P.J. de Jong. 1998. An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* **52**: 1-8.
- Ross, M.T., S. LaBrie, J. McPherson, and V.P. Stanton, Jr. 1999. Screening large-insert libraries by hybridization. In *Current protocols in human genetics* (ed. Ann Boyl), pp. 5.6.1-5.6.52. Wiley, New York, NY.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Selleri, L., J.H. Eubanks, M. Giovannini, G.G. Hermanson, A. Romo, M. Djabali, S. Maurer, D.L. Mcelligott, M.W. Smith, and G.A. Evans. 1992. Detection and characterization of "chimeric" yeast artificial chromosome clones by fluorescent in situ suppression hybridization. *Genomics* **14**: 536-541.
- Shizuya, H., B. Birren, U.J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, and M. Simon. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci.* **89**: 8794-8797.
- Soderlund, C., and C. Burks. 1994. GRAM and GENFRAGII: Solving and testing the single-digest, partially ordered restriction map problem. *Comput. Applic. Biosci.* **10**: 349-358.

Received July 29, 1999; accepted in revised form March 9, 2000.