



Cross-screening: a new method to assemble clones rapidly and unambiguously into contigs.

J Locke, G Rairdan, H McDermid, et al.

Genome Res. 1996 6: 155-165

Access the most recent version at doi:[10.1101/gr.6.2.155](https://doi.org/10.1101/gr.6.2.155)

References This article cites 19 articles, 4 of which can be accessed free at:
<http://genome.cshlp.org/content/6/2/155.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner with a teal background. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." in white. In the center, there is a white-bordered box containing the words "LEARN MORE" in blue. On the right, there is a photograph of a woman wearing a red and white superhero costume and a red visor. To the right of the photo is a green molecular structure icon and the word "CELLECTA" in white.

CRISPR and RNAi Genetic Screening.
Your new superpower.

LEARN MORE

CELLECTA

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © Cold Spring Harbor Laboratory Press

GENOME METHODS

Cross-screening: A New Method to Assemble Clones Rapidly and Unambiguously into Contigs

John Locke,¹ Greg Rairdan, Heather McDermid, David Nash,
David Pilgrim, John Bell, Kenneth Roy, and Ross Hodgetts

Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9

We have developed a new procedure that relies on an array of cross-hybridization tests to order a set of random clones into a contig. The method, called cross-screening, uses each clone as a target and its end sequences as probes, in a matrix of reciprocal cross-hybridization tests performed on a single blot. The relationships among the clones are determined rapidly from the pairwise tests, allowing clone order to be determined directly. We have applied this technique to DNAs from a set of overlapping λ clones from *Drosophila* chromosome 4. The location and orientation of each clone derived from the cross-screening data was that expected from the map assembled from overlapping restriction sites and chromosomal walking. The procedure provided additional information on a previously unknown, internally repeated DNA sequence. To demonstrate the general utility of the procedure, we have applied it to a previously described clone set within a contig in region 22q12 of human chromosome 22. The correct relative position and orientation of these clones were derived from the cross-screening data without knowledge of, or reference to, any nucleotide sequence or restriction site analysis of the DNA concerned. The cross-screening procedure is fast, economical, and robust and allows clone overlaps to be determined efficiently, with minimal interference from repeated DNA sequences. This new procedure is specifically designed for small groups of overlapping clones (tens to hundreds) and should facilitate the ordering of subclone libraries derived from small chromosomes or the large cloned inserts carried in YAC, BAC, and PI vectors.

The mapping of whole genomes now utilizes yeast artificial chromosomes (YACs) and other vectors [P1, bacterial artificial chromosomes (BACs)] capable of carrying very large segments of DNA. These clones have made it possible to assemble overlapping clone sets that span extended regions of the genome, such as whole human chromosomes (Chumakov et al. 1992). However, the size of these large clones makes restriction mapping, sequence analysis, and gene localization difficult. To analyze these regions, clones with smaller, more manageable fragments in cosmid or λ phage vectors are recovered from other libraries (e.g., a chromosome-specific cosmid library) to span these large regions (e.g., Whittaker et al. 1993; Fischer et al. 1994). DNA from these large clones may also be subcloned directly into such vectors. In either case, the resulting unorganized minilibraries of smaller

clones must be assembled into ordered overlapping sets, or contigs. To minimize the effort invested to contig a clone set, clone overlaps need to be determined quickly, efficiently, and accurately. Consequently, a number of methods have been developed that improve upon standard chromosomal walking techniques (e.g., Bender et al. 1979), which can be time consuming and significantly impeded by the presence of repetitive sequences.

The identification of clone overlaps has been facilitated by the use of a semiautomated robotic system to generate gridded arrays of random clones (Nizetic et al. 1991). These gridded array systems permit easy screening of large numbers of clones organized in high density on replica hybridization membranes. By screening replicate filters with probes, such as whole YAC or P1 clones, a set of cosmid clones can be recovered for a specific region (Baxendale et al. 1991). Then, through further rounds of hybridization with different probes (used in series or in parallel, where possible), the overlaps among a set of

¹Corresponding author.
E-MAIL john_locke@biology.ualberta.ca; FAX (403)492-9234.

LOCKE ET AL.

clones can be identified and a contig constructed. This method of contig assembly has been very useful where overlapping YAC clones have been available for a chromosome or subchromosomal region (e.g., Fischer et al. 1994).

Adaptations of chromosomal walking that initiate walks at multiple points have also been successful (Whittaker et al. 1993). They conduct simultaneously what are normally sequential rounds of probing, thereby expediting a standard chromosomal walk. In these cases, the linear methodology requires much time and redundant effort to generate data for a contig. Ideally, a procedure with the capacity to identify in parallel all clone overlaps in a set of clones should enhance the rate of contig assembly considerably.

Another widely used procedure for determining clone overlaps for higher resolution contig maps is based on restriction fragment fingerprinting analysis. This technically complex method depends on sophisticated computer technology for the identification of possible clone overlaps based only on several restriction fragments of similar size. This methodology has been applied successfully to the assembly of contigs of large genomes (Sulston et al. 1992).

In this paper we describe a new procedure that approaches this ideal by carrying chromosome walking to its extreme. By commencing the walk at all points, any theoretical requirement for sequential, repeated probing is eliminated. This procedure, which we call cross-screening, uses conventional "inward-facing" viral transcription promoters (e.g., SP6, T3, or T7) to generate terminal RNA probes for each clone in a set of overlapping clones (minilibrary). Every clone's probe is then concurrently hybridized to the entire set of clones. Because these cross-hybridization tests are accomplished on a single blot using a commercially available blotting/hybridization apparatus, savings in time, materials, and effort are substantial.

RESULTS

The Cross-screening Technique

Overview

The cross-screening technique involves hybridization in a two-dimensional array much like the Southern cross method described previously for assembling clone restriction fragments into a

map (Potter and Dressler 1986). In both methods the target DNA is immobilized as a series of parallel stripes on a hybridization membrane and the probe is hybridized as a series of parallel stripes at right angles to the target. The stripes in the Southern cross originate from gel electrophoresis separation of restriction fragments of clone DNA, whereas in cross-screening a manifold apparatus with repeating, parallel channels is used to produce the stripes (see below). At the intersection of each target and probe stripe, a cross-hybridization test occurs and a hybridization signal appears if the probe can hybridize with the target. In the Southern cross this signal indicates shared sequences between two restriction fragments (cut with different enzymes) within a single clone. This information facilitates the construction of a restriction map of the clone. In cross-screening, a signal indicates that a sequence is shared between two different clones, as would be expected for overlapping clones. This information facilitates the assembly of clones into an overlapping series, or contig.

The cross-screening technique uses a Miniblotter 45 Channel System (Immunitics, Inc.), which was originally devised to perform multiple antibody tests on Western-blotted, electrophoretically fractionated protein preparations. A similar apparatus has also been used to detect DNA-DNA hybridizations to evaluate clinical samples against a large number of pathogen target DNAs (Socransky et al. 1994). Only minimal procedural adaptations were needed for the use described here. The cross-screening procedure uses DNA from each clone as both a target and a source of probe in a set of cross-hybridization tests (Fig. 1). Target DNAs were fixed on a membrane in a series of stripes, or columns (Fig. 1B). The RNA probes were hybridized as a series of stripes, or rows at right angles to the columns (Fig. 1C). At the intersection of each column and row a test of cross-hybridization occurs for each pair of clones. Cross-hybridization in these pairwise tests can be detected as the presence of a characteristically square signal on the blot (Fig. 1D). From the pattern of cross-hybridizations the relative position of each clone in a contig can be determined (Fig. 1A).

Organizing Target DNA on the Blot

To analyze a set of clones, two aliquots of DNA from each clone were digested separately with two restriction enzymes (e.g., *Hae*III and *Rsa*I; see

RAPID CONTIG ASSEMBLY USING CROSS-SCREENING

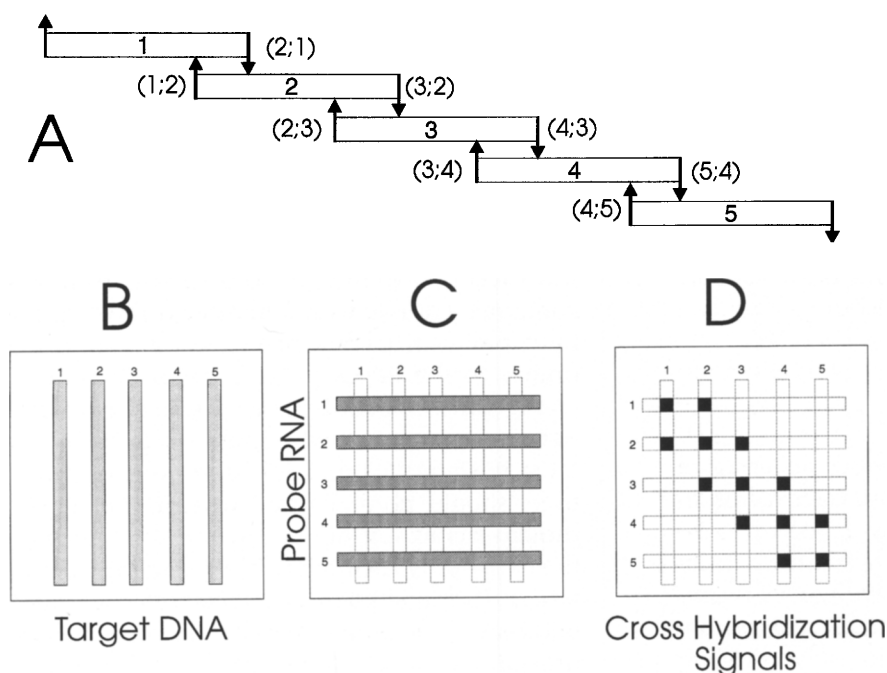


Figure 1 Diagram showing an overview of the major steps in the cross-screening procedure using clones from a hypothetical contig (A). First, the target DNA from each clone (1–5) is deposited as a series of vertical stripes (columns) on a hybridization membrane (B) using a Miniblotter 45 Channel System (Immunetics, Inc.) apparatus. Next, the membrane is repositioned in the apparatus with a 90° rotation. Then the RNA probes from each of the series of clones are added to the lanes of the Miniblotter to form a set of rows that intersect the target DNAs and produce an array of cross-hybridization tests (C). After removal of the hybridization solution and washing to remove the unbound probes, cross-hybridizations can be detected and will appear as squares of signal (D). This grid contains a set of all pairwise cross-hybridizations in both directions among the hypothetical set of clones shown in A. The terminal probes from each clone are shown as arrows in A. Their respective positions in the final array of cross-hybridization tests in D are shown in parentheses (column/target;row/probe). All probes hybridize with their own template DNA to give a series of diagonal squares from (1;1) to (5;5).

below). After heat inactivation of the restriction enzymes, aliquots from each digest were combined, and for each clone, a sample of the mixture containing 0.3 μg of DNA, diluted to 140 μl with water, was carefully deposited in one tunnel of the Miniblotter 45 Channel System. This apparatus has an upper and lower Plexiglas block between which the hybridization membrane is held. A series of 45 parallel channels on the bottom face of the upper block forms tunnels over the hybridization membrane and through which a solution can be applied in a manner similar to that of Socransky et al. (1994). Once in the tunnel, the DNA solution diffuses through the hybridization membrane (Gene Screen Plus,

NEN Dupont) and into three sheets of filter paper placed beneath the membrane. Thus, each DNA sample is deposited as a long stripe of DNA on the membrane, much as a slot blot apparatus deposits DNA in a series of slots. A set of clones (up to 45 at a time) results in a series of parallel, closely spaced columns, with one clone per column, on one membrane (Fig. 1B). The membrane was removed from the Miniblotter, denatured in 0.5 M NaOH (2 \times 3 min), neutralized in 1.0 M Tris-HCl (pH 7.5) (2 \times 3 min), and dried to bind the DNA to the membrane.

Probing the Blot in the Second Dimension

The membrane, containing columns of target DNA, was replaced in the Miniblotter at 90° to its previous orientation (Fig. 1C). For each clone in the set, a terminal probe was then inserted into a Miniblotter tunnel to form a series of rows. Being at 90° to the previous orientation, the probe tunnels intersect the columns of target DNA laid down previously. The single-stranded digoxige-

nin-labeled RNA probes (see Methods) were applied in 140 μl of hybridization solution consisting of 50% formamide, 5 \times SSC, 1% SDS, and 1% blocking agent (Boehringer Mannheim Canada, Inc.) and allowed to hybridize for ~2 hr in an oven at 55°C. Plastic pipette tips were temporarily wedged in each end of the tunnel openings to prevent accidental probe cross-contamination.

After hybridization, the Miniblotter washing manifold was used to remove most of the unhybridized probe. All channels were rinsed briefly with 2 \times SSC, 0.1% SDS, which the manifold draws through the channels of the apparatus. The membrane was then removed from the apparatus and washed further in a tray with

LOCKE ET AL.

2× SSC, 0.1% SDS, for 5 min at room temperature and then washed twice in 0.1% SSC, 0.1% SDS, for 30 min at 65°C. Hybridized digoxigenin-labeled probe was detected on the membrane according to the manufacturer's instructions (Boehringer Mannheim Canada, Inc.). Cross-screen blots were scanned with a Bio-Rad GS670 Scanner (reflectance mode) and imaged using the Bio-Rad Molecular Analyst Software V1.2 on a 486DX2 IBM-compatible computer.

Analysis of Cross-screening Blots

Each cross-screening blot provides an array of cross-hybridization tests. Because each region of potential cross-hybridization is expected to be characteristically square and in a known location, the identification of positive signals is simplified. From each blot two sets of data are obtained: (1) Reading down columns of target DNA, one ascertains those termini that are included in the sequence of a given clone; (2) reading across rows identifies those clones that hybridize to a given probe. At least one signal is expected in each row and each column, as each probe should hybridize to its own template. This signal acts as a positive control, and its absence most likely indicates unsatisfactory probe labeling. If the order of the probes is the same as the order of the targets from which they were derived, then the positive control signals are expected to fall on a diagonal of the cross-hybridization matrix (Fig. 1). The evaluation and subsequent interpretation of all the signals identify clones that share overlaps. Combining this information can order the clones into contigs.

Probing can be done simultaneously with pooled transcripts from both ends (T7 and T3/SP6) of each clone as diagramed in Figure 1, but additional information can be obtained by analyzing the two probes separately. Separate probing permits monitoring of both terminal probes from each clone and ascertainment of the orientation of each clone. It can be done as pairs of rows on one blot, but in this paper we have done them on separate blots, an arbitrary choice. Therefore a complete matrix of results for n clones consists of two $n \times n$ blots. Contig construction requires that information from both blots be integrated as in Figures 3C and 4C (below).

Because a contig is usually substantially longer than the average clone size, most ran-

domly chosen pairs of clones will show no cross-hybridization. For clone sets spanning regions that have no DNA sequences repeated within the cloned region (but not necessarily within the genome) hybridization will usually involve only one of the two probes from each clone. Such cross-hybridization most likely indicates that the clones overlap and therefore extends the contig. (An exception to this would be if one clone was contained entirely within another, in which case both probes from the smaller clone would hybridize to the larger.) The terminal clones of a contig can be identified, as they will normally have one probe that hybridizes only to itself (Fig. 1A at positions 1;1 and 5;5), and for each contig there should normally be two such clones. Although the extent of overlap is not revealed by the data—only their relative order and the orientation—the analysis does allow selection of a minimal set of clones necessary to cover the entire contig (minimal tiling set).

In the data shown here, the clones have been ordered to correspond with their sequence on the physical map. We recommend that any contig assembled from the cross-hybridization pattern be confirmed in this manner, using residual material made for the initial experiment. The confirmatory blot can be organized to reflect the physical map, and the two probes from each clone can be pooled without significant loss of information. The option also exists to simplify the matrix by eliminating redundant clones. However, sometimes apparently redundant clones prove valuable in another experimental context. Thus, we recommend confirming the analysis on as many clones as is reasonable, as the protocol allows for the preparation of sufficient quantities of probe and target DNA to do so with little extra expenditure of time or resources.

Although some of the more common cross-hybridization patterns have been mentioned above, a description of the much larger set of possible patterns for normal and aberrant clones is beyond this initial report.

Confirming the Order of an Overlapping Set of λ Clones on *Drosophila* Chromosome 4

The cross-screening procedure was developed to facilitate the identification of overlapping clones in the gene-rich euchromatic banded region of *Drosophila melanogaster* chromosome 4, where we are currently constructing a physical map. This region is atypically rich (at least for *D. melano-*

RAPID CONTIG ASSEMBLY USING CROSS-SCREENING

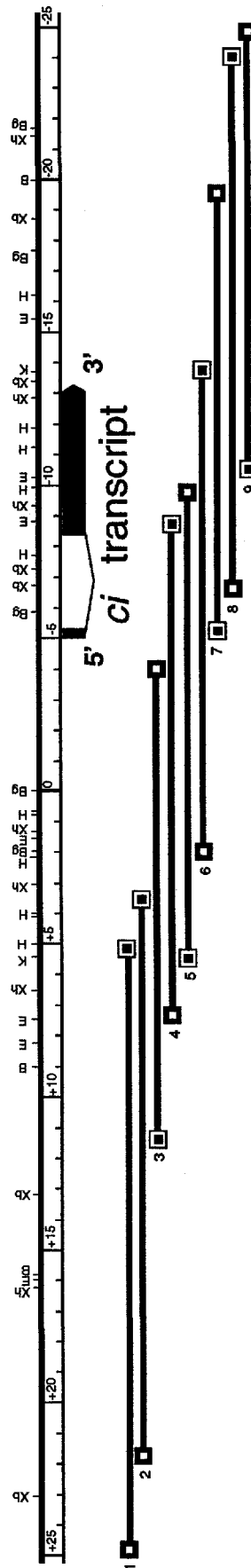


Figure 2 Restriction map and clones from the *ci* region on *D. melanogaster* chromosome 4 used in the testing of the cross-screening technique. This composite map is derived from λ clones 1–9 shown below the map (and others not shown) by standard chromosome walking techniques. The position of the *ci* gene is shown below the map. The positions of the T7, T3, and SP6 promoters are shown for each λ clone. The restriction enzyme sites shown are *Xba*I (Xb), *Xho*I (Xh), *Bam*HI (B), *Eco*RI (E), *Kpn*I (K), and *Bgl*II (Bg). The scale is in kb and extends from +25 to –25 relative to the *Bgl*II site, which is arbitrarily set at the 0.0 position.

LOCKE ET AL.

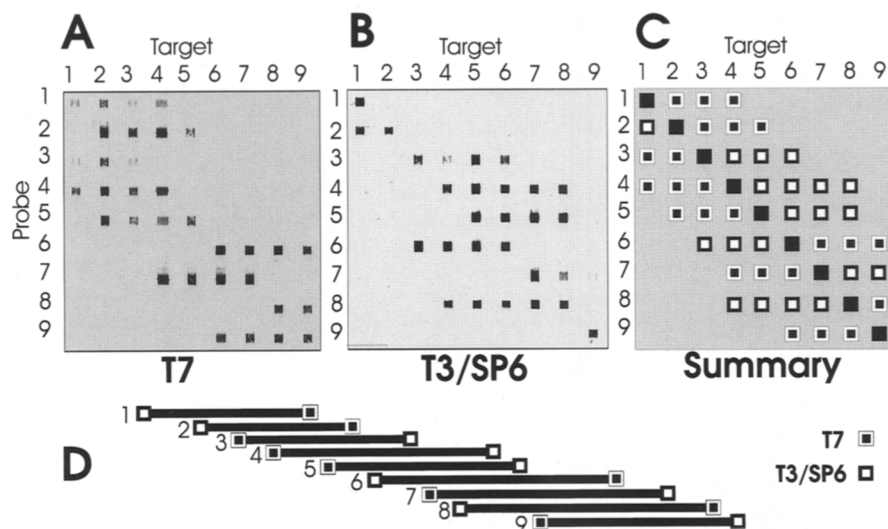


Figure 3 Cross-screening blots of nine λ clones (1–9) spanning the *ci* locus on *D. melanogaster* chromosome 4. In each blot, target λ clone DNA (0.3 μ g/lane) was placed in vertical columns while probe from T7 and T3 or SP6 transcription reactions was placed in horizontal rows. Cross-hybridizations appear as dark squares of signal on the blot as revealed by the alkaline phosphatase detection system (A,B). The diagonal (top left to bottom right) series of squares are positive controls where each clone's probe hybridizes to itself. The summary diagram (C) combines the cross-hybridization data derived from the T7 (dark-centered box, \blacksquare) and T3/SP6 (light-centered box, \square) probeings in A and B. The T7 and T3/SP6 symbols denote the blot where the signal originated. The solid squares (\blacksquare) positioned along the diagonal represent both T7 and T3/SP6 signal. (D) The clone overlap relationships as derived from these summary data. This organization is consistent with the relative positions of the clones derived from the chromosome walk shown in Fig. 2.

gaster) in interspersed moderately repeated sequences (Miklos and Cotsell 1990). To test the reliability of the cross-screening technique, DNAs from a set of λ clones on *Drosophila* chromosome 4 (Locke and Tartof 1994) were analyzed. These clones have been numbered sequentially according to their locations on the physical map (Fig. 2; clones 1–9).

The T7 and T3(SP6) probes from each clone were used in separate cross-screening blots (Fig. 3, A and B, respectively). Because the order of the targets and the probes has been organized to reflect the sequence of the clones within the contig, the cross-hybridization signals cluster along the diagonals in Figure 3, A and B. The hybridization of each clone to its own template (a positive control), confirms that target DNA is present in the appropriate columns and probe in the appropriate rows. The ordering of this clone set is unambiguous and leads to the situation shown, where each terminal probe hybridizes to targets

either on the left or the right side of the diagonal, but not on both sides. This pattern is not inevitable; circumstances where clones are wholly included within others in the set or sequences repeated within the region spanned by the clone set could give signals on both sides in an ordered clone set.

The termini of the contig are easily identified as clones 1 and 9, because each produces a probe that hybridizes only to its own template (see Fig. 3B). In this particular set of data, many probes hybridize to more than a single additional target site, a phenomenon that reflects the fact that there are more clones in the set than the minimum needed to define the contig. The minimum clone set needed in this case would consist of four clones, and there are four combinations possible (1,3,6,9; 1,4,6,9; 1,4,7,9; and 1,4,8,9). In these sets,

only the middle members of each set act as prototypical cross-hybridizing pairs in which no probe hybridizes with more than its template and one other target.

From the complete array of cross-hybridization information in Figure 3C, the relative order of the clones can be determined and is shown in Figure 3D. The organization is equivalent to that determined from the restriction maps shown in Figure 2. The orientation of each clone can also be derived from the cross-hybridization data, and for each clone the orientation was that predicted from the previous map.

Repeated Sequences

The presence of sequences repeated within a contig can complicate the analysis of a clone set. The situations generated by each instance where hybridization signals result from repeated sequences within a cross-screen will differ, but

RAPID CONTIG ASSEMBLY USING CROSS-SCREENING

most will result in a nonlinear contig map. There was one instance of a repeat sequence anomaly present in our *Drosophila* contig. Clone 7 had a sequence at its right end, within the initial SP6 probe, that not only hybridized to clones 7–9 as expected (Fig. 2) but also to clones 4–6 (data not shown). Consequently, this pattern did not fit that of a standard series of overlapping clones. With the cross-screening technique, anomalies such as this can be removed entirely from the overall data set. The residual contig map would still be coextensive with the one shown in Figure 3D because the probing of other clones to clone 7 DNA would have positioned clone 7 correctly within the contig.

DNA from clone 7 had already been restriction mapped (Locke and Tartof 1994), and supplementary work (not shown) identified several fragments that contained sequences repeated in the *Drosophila* genome but not necessarily repeated within the contig. This information suggested the possibility that the clone 7 anomaly might stem from repeat sequence DNA. Knowledge of the map and repeat location allowed us to eliminate the cross-hybridization problem, by digesting the clone 7 DNA with *Xba*I, which cuts the insert DNA at coordinate –19 kb (Fig. 2) between the repeat sequence and the right end of the clone. Thus, the right-hand probe (SP6) for

clone 7 in Figure 3B was produced after an *Xba*I digest of the clone. With this modification of the usual protocol (see Methods) to remove a repeated sequence from the probe, the hybridization fits clone 7 into the contig map in precisely the position predicted from the previous mapping. The discovery that the initial right end SP6 probe (made with *Rsa*I or *Hae*III digests) of clone 7 contained sequence homologies close to its left end, where it overlaps with clones 4, 5, and 6, is typical of serendipitous information the technique can reveal about repeated sequences within the contig.

Ordering an Overlapping Set of Cosmid Clones on Human Chromosome 22

To demonstrate the general utility of the procedure, we have applied it to clones spanning a segment of human DNA. In human genomic DNA the presence of dispersed highly repeated elements (e.g., Alu elements) could be located within a probe template and thereby confound the hybridization tests. Therefore, as a precaution, all human probes were preannealed (to $C_0t = 27$) with denatured total human genomic DNA to remove labeled repeat sequence from the single-stranded component of the probe preparations.

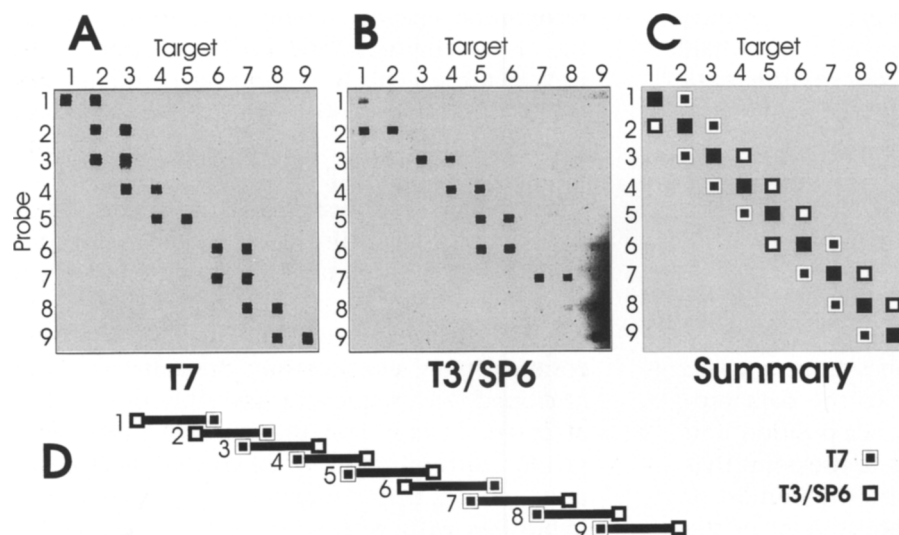


Figure 4 Cross-screening blots of nine cosmid clones (1–9) that represent a tiling path of clones from the 22q12 region of human chromosome 22. T7 end probes (A), T3/SP6 end probes (B), and a summary diagram (C) are marked as in Fig. 3, with an overlapping contig derived from the data (D). This contig is consistent with that determined by Xie et al. (1993) and also provides the orientation of these clones.

The cross-screening procedure was applied to a set of cosmid clones from a contig in region 22q12 of human chromosome 22 (see Methods). The results of the cross-screening are shown in Figure 4. As in the previous example, the T7 and T3/SP6 probes were used separately (Fig. 4A,B), but in this case, the clones already comprised a minimal tiling path. The terminal probes (clones 1 and 9), coincidentally also in the T3/SP6 set, again hybridize only to themselves. Otherwise, each probe hybridizes only to itself, as seen as signals on the descending diagonals, and to one other clone. Because there are no clones wholly included in other clones, the orienta-

LOCKE ET AL.

tion of the clones in the contig relative to their individual orientation in the vector is fully defined.

The aggregate picture derived from the data is shown in Figure 4D. Note that this diagram was constructed without knowledge of the length of individual clones (hence, the use of standardized bars in Fig. 4D) and without reference to, or need for, any sequence or restriction analysis of the DNA concerned. The order of the clones correlates exactly with that given in Xie et al. (1993); furthermore, the data provide the orientation of each cloned insert within its vector.

DISCUSSION

Practical Advantages of the Cross-screening Procedure

The main advantage of the technique is the speed at which a large set of cohesive data can be produced. The multiple probings on one cross-screening blot obviate the need for multiple rounds of probings that are required in most other methods. For example, to find overlaps among 120 λ clones derived from a 400-kb YAC, Whittaker et al. (1993) had to perform >100 blot hybridizations. With each round, less and less new information was obtained. Using the cross-screening procedure, this clone set could have been ordered with a maximum of just over fourteen 45×45 cross-hybridization grids, and half that number if clone/vector orientation data were sacrificed. Furthermore, if the grids were constructed on successive days, it would be possible to progressively eliminate redundant clones from further study (as connected segments of the contig emerged) and thereby further reduce the workload.

A less obvious advantage of cross-screening is its resilience to failure in probe production. If only one of the two end probes from a given clone is effective, the totality of the data produced may still indicate the exact position and orientation of the clone, at least to the point that it can be decided whether further definition of the clone is critical to the establishment of the contig. If both probes fail, the cross-hybridizations to the probes of other clones may still permit it to be positioned in the contig, but without orientation.

Another advantage of this technique is the reciprocal probing of each clone pair. If two

clones overlap then positive signals are expected from both termini included within the overlap. Thus, true clone overlaps produce a reciprocal pair of signals, in addition to the internal positive control of the template DNA itself. Absence of such reciprocal signals indicates that the clones' relationship may not be a simple overlap. Lack of reciprocal signals is expected in the event of chimeric clones, terminal probes with repeated sequences, and internally deleted clones. Thus, this aspect of the cross-screening may not only confirm legitimate overlaps of authentic cloned sequences but also identify aberrant clones in the set.

Problems Encountered

Initial minor problems, such as lane leakage, were remedied easily by appropriate gasket pressure. We have found that during the hybridization step it is essential to seal the tunnel exits and fill all lanes with fluid to prevent seepage between adjacent tunnels, regardless of whether or not they contain experimental samples.

Variation in signal intensity was also noted. Potentially, some clones could yield no effective probe because restriction sites are positioned poorly. The use of two separate 4-nucleotide restriction enzymes reduces this possibility but does not eliminate it altogether. The use of 6-bp recognition enzymes could be used to increase the probe length and, hence, signal strength (as well as for selectively reducing probe length to avoid confounding repeated sequences, as we did with *Drosophila* clone 7). Severe differences in signal strength from probe to probe can also be accommodated by adjusting the probe concentration. The signal can also be optimized by controlling the development time in the detection system.

We have found one unusual artifact with cosmid clones. On occasion, the rearrangement of cloned DNA sequences (probably by deletion of insert during DNA preparation) produces a probe contaminated with labeled vector sequences. In this situation, the probe cross-hybridizes with vector sequences in every column, an easily identified result. We have eliminated this artifact by preannealing with unlabeled vector sequences, just as was done for highly repeated sequences in the human clones. The alternative, clone reisolation and preparation of probe without the contaminant, also

eliminated the spurious cross-hybridization signals.

Repeated Sequences

Cross-screening can handle clones with internally repeated sequences. The standard technique of competitive hybridization (Litt and White 1985; Evans et al. 1989) can be applied to clones with highly repeated interspersed DNA (e.g., Alu elements). However, the competitive hybridization technique is of questionable value for clones with some dispersed repeated sequences that are not distinguished from unique sequences by dramatic differences in C_0t value (e.g., transposable elements in *Drosophila*; moderate to low repeats in humans). If the interspersed repeats are widely separated and only moderately frequent (as in *Drosophila*), the problems generated by repeat sequences should be relatively rare. Our use of restriction enzymes with 4-bp recognition sequences to limit the probe to relatively short templates at the end of each clone also reduces the frequency of repeated elements becoming a terminal probe.

Where a repeat is part of a terminal probe, it will hybridize to all other clones that contain the repeat, but the reciprocal hybridization will not occur. This nonreciprocity flags the clones' relationship as not a simple, authentic overlap and therefore deserving of further consideration. Such a repeat sequence probe, and its template clone, may be eliminated from the overall analysis. The important point is that such problems will arise only in a minority of instances where repeats are present, and only in a small minority of these cases will they be undetected by the analysis.

General Application of Cross-screening

This procedure can be used easily with any set of clones with a vector that has inwardly facing viral transcription start sites. Alternatively, the knowledge of the sequences adjacent to the cloning sites should also permit the use of linear PCR to generate probes specific for the insert termini in any vector system. The procedure also permits different vectors to be used in the same cross-screen blot. For example, one could use plasmids containing cDNA sequences as probes (rows) to map them within a λ , cosmid, or YAC tiling set (columns). Furthermore, the basic procedure is

RAPID CONTIG ASSEMBLY USING CROSS-SCREENING

simple and can be performed with minimal materials in about one working day. The results reveal clone relationships quickly and efficiently with minimal interference from repeated DNA sequences and allow for the efficient selection of a minimal tiling set of the region for further analysis (restriction mapping, DNA sequencing, gene localization).

This procedure should be ideally applicable to the mapping of small genomes or genome regions, such as the construction of cosmid or λ contigs of YAC clones. For example, the cross-screening method has been used to organize 20 fosmid clones derived from a YAC clone containing 500 kb of human chromosome 22q11.2 DNA (C. Bell, pers. comm.). These fosmids were organized into a contig of 14 overlapping clones, another contig of 3 clones, and 2 singletons that hybridize only with themselves and not with any other clone in the set; the remaining clone failed to give end probes. In general, 45 cosmid clones should be able to provide a complete clone set (about fivefold coverage) for a 300-kb YAC and can be analyzed and ordered by using a single cross-screening blot. If more clones need to be ordered, as with a larger region or clone, then a larger apparatus with more lanes can be used. Alternatively larger clone numbers covering larger regions (e.g., ~1000-kb YACs) can be handled by pooling clone target DNAs (several clones per lane). With this modification, little information would be lost if only the target DNAs are pooled but the probes are still done individually for each clone (nonpooled probes). With target DNA pooling, an increase in clone number may increase the number of cross-hybridizations in a set only linearly, rather than exponentially as in the standard cross-screening with an $n \times n$ grid.

The cross-hybridization data are normally comprehensible enough that advanced specialized computers and programs are not required for analysis and interpretation. However, the use of computer programs that could manipulate such cross-hybridization information on common desktop computers (e.g., see Cuticchia et al. 1992) should facilitate the routine handling of results from repeated cross-screening blots.

METHODS

Vectors and Clones

We have used an overlapping set of λ clones that span ~50 kb of the proximal portion of *D. melanogaster* chromosome

LOCKE ET AL.

4 (Locke and Tartof 1994). These clones represent ~5% of the banded euchromatic region of polytene chromosome 4. The clones have been recovered from several different libraries using typical chromosome walking methods and include regions surrounding the *cubitus interruptus (ci)* locus (Orenic et al. 1990; Locke and Tartof 1994). The vectors used in these libraries have a T7 promoter adjacent to one cloning site and either a T3 (λ DASH II) or SP6 (λ KT7) promoter adjacent to the other. This provided an established contig upon which the cross-screening procedure was tested.

Xie et al. (1993) described a set of overlapping human cosmid clones that span ~1600 kb in region 22q12 near the neurofibromatosis type 2 (NF2) gene and include the Ewings' sarcoma t(11;22) translocation break point gene. We have used some of these clones for testing the cross-screening procedure on an established contig of human DNA sequences. They were taken from a gridded set of human chromosome 22-specific clones constructed by deJong et al. (1989) and Zucman et al. (1992). The cosmids used were E107G9, N87G7, N63H9, N57B5, E98C4, E113B1, N112B12, E115F7, and N49G8 [clones designated E are from deJong et al. (1989); clones N are from Zucman et al. (1992)]. For simplicity, they were renumbered 1–9 in the order given above. The cosmid vectors used in these libraries were LAWRIST 16 or LAWRIST 5, which have a λ origin of replication, kanamycin resistance (as the selectable marker), and T3 (or SP6) and T7 viral promoters facing the BamHI cloning site where the human DNA is inserted.

Isolation of Clone DNA

Lambda clone DNA was obtained via standard methods from either large-scale or minipreparations (Sambrook et al. 1989). Cosmid clones were usually grown in 10 ml of Superbroth (Tartof and Hobbs 1988) with 10 μ g/ml of kanamycin. DNA was isolated by a small-scale SDS–NaOH lysis method. DNA (2–10 μ g) was usually collected for each clone.

Transcription Labeling of Probes from Clone Termini

The inward-facing bacteriophage promoters that flank each cloned DNA segment were used to generate probes specific for the termini of each sequence. To generate only terminal transcripts, restricted DNA was used as the template. Two different enzymes were used in case one enzyme cut too close to a promoter site to provide a good template. We usually added 0.1 μ g each of a *HaeIII*-digested aliquot and an *RsaI*-digested aliquot to the transcription reactions (total, 0.2 μ g). In the event that this strategy fails to generate an effective probe, an appropriate pair of 6-nucleotide recognition site restriction enzymes could be used instead. A "six-cutter" was used, but for a different reason, with clone 7 in the *D. melanogaster* set, as described in Results.

T7 and T3 (or SP6) transcripts were generated in separate reactions to provide a probe specific for each end of each cloned sequence. The *in vitro* transcription reactions with SP6, T3, and T7 were carried out in 20 μ l of transcription buffer according to the manufacturer's instructions (Boehringer Mannheim Canada, Inc.) with 0.5 μ l of digoxigenin-labeled UTP (RNA labeling mix; Boehringer Mann-

heim Canada, Inc.). This reaction generates probes of ~200–2000 bases in length. Usually 2, 3, or 5 μ l of the T7, T3, or SP6 reactions, respectively, was added to the hybridization solution, without the removal of unincorporated nucleotides. Thus each 20- μ l transcription reaction could provide several probings (e.g., a test to determine whether the reaction worked). In certain cross-screen blots the amount of probe was adjusted to optimize the signal.

Interspersed highly repeated sequences located near the termini of the clone sequences would be expected to contribute to the probe sequences and confound the hybridization pattern. This potential problem is significant in the analysis of human DNA. These highly repeated sequences were selectively depleted as double-stranded reactants by preannealing the probe with denatured total genomic DNA (Litt and White 1985; Evans et al. 1989). Typically 2–5 μ l of probe was added to 18–15 μ l of 2.5 mg/ml sonicated human placental DNA, boiled, and annealed at 65°C for >2 hr. Presumably preannealing with commercially available C_{θ} t 1 DNA would also be effective.

Acknowledgments

We thank Jan Dumanski for clones in the cosmid contig on chromosome 22 and Callum Bell for making his cross-screening results available prior to publication. We thank Barbara Hepperle, Hilary Kemp, Lynn Podemski, and Eugene Chomey for technical assistance. Financial support was provided by the Canadian Genome Analysis and Technology (CGAT) Program.

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

- Baxendale, S., G.P. Bates, M.E. MacDonald, J. F. Gusella, and H. Lehrach. 1991. The direct screening of cosmid libraries with YAC clones. *Nucleic Acids Res.* **19**: 6651.
- Bender, W., P. Spierer, and D. Hogness. 1979. Gene isolation by chromosomal walking. *J. Supra. Mol. Struct.* (Suppl.) **10**: 32.
- Chumakov, I., P. Rigault, S. Guillou, P. Ougen, A. Billaut, G. Guasconi, P. Gervy, I. LeGall, P. Soularue, L. Grinas et al. 1992. Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* **359**: 380–387.
- Cuticchia, A.J., J. Arnold, and W.E. Timberlake. 1992. The use of simulated annealing in chromosome reconstruction experiments based on binary scoring. *Genetics* **132**: 591–601.
- deJong, P.J., K. Yokobata, C. Chen, F. Lohman, L. Pederson, J. McNinch, and M. van Dilla. 1989. Human chromosome-specific partial digest libraries in human lambda and cosmid vectors (A2333). *Cytogenet. Cell Genet.* **51**: 985.
- Evans, G.A., K.A. Lewis, and B.E. Rothenberg. 1989. High

RAPID CONTIG ASSEMBLY USING CROSS-SCREENING

efficiency vectors for cosmid micro-cloning and genome analysis. *Gene* **79**: 9–20.

Fischer, S.G., E. Cayanis, J.J. Russo, I. Sunjevaric, B. Boukhgalter, P. Zhang, M.-T. Yu, R. Rothstein, D. Warburton, I.S. Edelman, and A. Efstatidis. 1994. Assembly of ordered contigs of cosmids selected with YACs of human chromosome 13. *Genomics* **21**: 525–537.

Litt, M. and R. White. 1985. A highly polymorphic locus in human DNA revealed by cosmid derived probes. *Proc. Natl. Acad. Sci.* **82**: 6206–6210.

Locke, J. and K.D. Tartof. 1994. Molecular analysis of *cubitus interruptus* (*ci*) mutations suggests an explanation for the unusual *ci* position effects. *Mol. & Gen. Genet.* **243**: 234–243.

Miklos, G.L.G. and J.N. Cotseff. 1990. Chromosome structure at interfaces between major chromatin types: *alpha*- and *beta*-heterochromatin. *BioEssays* **12**: 1–6.

Nizetic, D., G. Zehetner, A.P. Monaco, L. Gellen, B.D. Young, and H. Lehrach. 1991. Construction, arraying, and high-density screening of large insert libraries of human chromosomes X and 21: Their potential use as reference libraries. *Proc. Natl. Acad. Sci.* **88**: 3233–3237.

Orenic, T.V., D.C. Slusarski, K.L. Kroll, and R.A. Holmgren. 1990. Cloning and characterization of the segment polarity gene *cubitus interruptus* *Dominant of Drosophila*. *Genes & Dev.* **4**: 1053–1067.

Potter, H. and D. Dressler. 1986. A “Southern Cross” method for the analysis of genome organization and the localization of transcription units. *Gene* **48**: 229–239.

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.

Socransky, S.S., C. Smith, L. Martin, B.J. Paster, F.E. Dewhirst, and A.E. Levin. 1994. “Checkerboard” DNA-DNA hybridization. *BioTechniques* **17**: 788–792.

Sulston, J., Z. Du, K. Thomas, R. Wilson, L. Hillier, R. Staden, N. Halloran, P. Green, J. Thierry-Mieg, L. Qiu, S. Dear, A. Coulson, M. Craxton, R. Durbin, M. Berks, M. Metzstein, T. Hawkins, R. Ainscough, and R. Waterston. 1992. The *C. elegans* genome sequencing project: A beginning. *Nature* **356**: 37–41.

Tartof, K.D. and C.A. Hobbs. 1988. New cloning vectors and techniques for easy and rapid restriction mapping. *Gene* **67**: 169–182.

Whittaker, P.A., L. Wood, M. Mathrubutham, and R. Anand. 1993. Generation of ordered phage sublibraries of YAC clones: Construction of a 400-kb phage contig in the human dystrophin gene. *Genomics* **15**: 453–456.

Xie, Y.-G., F.-Y. Han, M. Peyrard, M.H. Rutledge, I. Fransson, P. DeJong, J. Collins, I. Dunham, M. Nordenskjold, and J. P. Dumanski. 1993. Cloning of a

novel, anonymous gene from a megabase-range YAC and cosmid contig in the neurofibromatosis type 2/meningioma region of human chromosome 22q12. *Hum. Mol. Genet.* **2**: 1361–1368.

Zucman, J., O. Delattre, C. Desmaze, B. Plougastel, I. Joubert, T. Melot, M. Peter, P. DeJong, G. Rouleau, A. Aurias, and G. Thomas. 1992. Cloning and characterization of the Ewing’s sarcoma and peripheral neuroepithelioma t(11;22) translocation breakpoints. *Genes Chromosomes Cancer* **5**: 271–277.

Received October 25, 1995; accepted in revised form January 5, 1996.