low recombination. Nucleotide and haplotype diversity will also probably parallel recombination rates. Although our baseline long-range recombination rates will be useful, they should be recalculated when the human genomic sequences are finished and as higher resolution genetic maps become available. In the more distant future, genotyping greater numbers of reference families at much higher polymorphism densities will lead to short-range maps of recombination hot spots.

Methods

Connection of genetic and physical maps

We used short, single-pass genomic sequences and/or PCR primer sequences for STRPs to identify draft or finished bacterial artificial chromosome (BAC) or cosmids sequences within GenBank that span the STRPs using BLAST24 and ePCR. Blast criteria were score (bits) > 200, expect (E) value < e−50, and ratio of matched bases to marker sequence length > 85%. ePCR criteria were no more than one base mismatch in each primer and size of PCR product within allele size range for the STRP. About 75% of the STRPs were connected to the long genomic sequences. The reasons for failure of the remaining 25% are not fully understood, but include absence of the corresponding sequence in GenBank and poor quality of the STRP sequences. As the genetic maps are marker rich, the absence of 25% was not a serious limitation. Tables of STRPs with GenBank sequence accession numbers for encompassing BACs, genetic map positions and recombination rates are available from the Marshfield web site.

Determination of recombination rates

For each sequence assembly we built new female, male and sex-average genetic maps, using the marker order provided by the assemblies and using the genotyping data from the eight CEPH reference families.5 We fitted cubic splines to plots of genetic versus physical distance, and from these curves we obtained recombination rates as first derivatives.15 The statistical significance of the recombination rates was estimated by computer simulation of STRP densities were obtained from programs written and tested at Marshfield29. STR sequences were 12, 11 or 19 ((AC)n, (A)(GT)n, (AGAT)n, (AAN)n) sequences were 12, 11 or 19 ((AC)n, (A)(GT)n, (AGAT)n, (AAN)n), respectively. We obtained interspersed repetitive element densities using the program Repeat Masker (http://ftp.genome.washington.edu/RM/RepeatMasker.html). SINES and LINEs were defined by Repeat Masker and consist primarily of Alu and L1 elements, respectively. We computed all DNA sequence parameters over 250-kb windows centred about each STRP. For markers ≤ 125 kb from the ends of the sequence assemblies, we defined the window as the 125 kb of proximal sequence plus all available distal sequence. Unknown bases in the sequence assemblies were excluded from analysis. All parameters were corrected for reduced window size owing to unknown bases or proximity to ends.

Measurement of linkage disequilibrium

Recombination deserts and jungles were selected as those chromosomal regions with sex-average recombination rates of <0.3 or >3.0, respectively. We measured linkage disequilibrium for all pairs of STRPs within the deserts (449 pairs) and jungles (467 pairs) using Fisher’s exact test.27 Only disequilibrium results that were significant at P ≤ 0.01 were plotted in Fig. 2. An overall P-value was obtained by a permutation test treating the regions as units in order to account for the dependence between marker pairs within a region.

Received 22 October; accepted 8 December 2000.

1. The BAC Resource Consortium. Integration of cytogenetic landmarks into the draft sequence of the human genome. Nature 409, 533–541 (2000).
2. The International Human Genome Mapping Consortium. A physical map of the human genome. Nature 409, 934–941 (2001).
3. Deloukas, P. et al. A physical map of 30,000 human genes. Science 282, 744–746 (1998).
4. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409, 860–921 (2001).
5. Brosn, K. W., Murray, C. J., Sheffield, V. C., White, R. L. & Weber, J. L. Comprehensive human genetic maps: Individual and sex-specific variation in recombination. Am. J. Hum. Genet. 63, 861–869 (1998).
6. Dunham, I. et al. The DNA sequence of human chromosome 22. Nature 402, 489–495 (1999).
7. Hartman, M. et al. The DNA sequence of chromosome 21. Nature 405, 311–319 (2000).
8. Fain, P. R., Kort, E. N., Yousef, C., James, M. R. & Litt, M. A high resolution CEFP crossover mapping panel and integrated map of chromosome 11. Hum. Genet. Mol. Genet. 5, 1631–1636 (1999).
9. Boulflard, G. G. et al. A physical map of human chromosome 7: An integrated YAC contig map with average STS spacing of 79 kb. Genome Res. 7, 673–692 (1997).
10. Nagaraja, R. et al. X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. Genome Res. 7, 210–222 (1997).
11. Mohrenweiser, H. W., Tsuchimoto, S., Gordon, L. B. & Olsen, A. S. Regions of sex-specific hypo- and hyper-recombination identified through integration of 180 genetic markers into the metric physical map of human chromosome 19. Genomics 47, 153–162 (1998).
12. Nicolas, A. Relationship between transcription and initiation of meiotic recombination: toward chromatin accessibility. Proc. Natl Acad. Sci. USA 95, 87–89 (1998).
13. Wahls, W. P. Meiotic recombination hotspots: shaping the genome and insights into hypervariable minisatellite DNA change.Curr. Top. Dev. Biol. 37, 37–75, (1998).
14. Fari, J. D., Haen, K. M. & Gill, B. S. Saturation mapping of a gene-rich recombination hot spot region in wheat. Genetics 154, 823–835 (2000).
15. Klama, R. M. & Hey, J. Reduced natural selection associated with low recombination in Drosophila melanogaster. Mol. Biol. Evol. 10, 1239–1258 (1993).
16. Chen, T. L. & Manuelidis, L. SINES and LINE clusters in distinct DNA fragments of Gmiens band size. Chromosoma 98, 309–316 (1989).
17. Payser, B. A. & Nachman, M. W. Microsatellite variation and recombination rate in the human genome. Genetics 156, 1285–1298 (2000).
18. Nachman, M. W., Bauer, L. V., Crowell, S. L., and Aquadro, C. F. DNA variability and recombination rates at X-linked loci in humans. Genetics 150, 1133–1141 (1998).
19. Nachman, M. W. & Crowell, S. L. Contrasting evolutionary histories of two introns of the Duchen muscular dystrophy gene in humans. Genetics 155, 1855–1864 (2000).
20. Majewski, J. R. et al. GT repeats are associated with recombination on human chromosome 22. Genome Res. 10, 1108–1114 (2000).
21. Bernardi, G. Isochores and the evolutionary genomics of vertebrates. Gene 241, 1–17 (2000).
22. Eisenbarth, J., Vogel, K., Krone, W., Vogel, W. & Assmann, G. An isochore transition in the NF1 gene region coincides with a switch in the extent of linkage disequilibrium. Am. J. Hum. Genet. 67, 873–880 (2000).
23. Yu, J. et al. Individual variation in recombination among human males. Am. J. Hum. Genet. 59, 1186–1192 (1996).
24. Lien, S., Syla, J., Schlochinger, R., Bappert, G. & Arheim, N. Evidence for heterogeneity in recombination in the human pseudoautosomal region: High resolution analysis by sperm typing and radiation-hybrid mapping. Am. J. Hum. Genet. 66, 557–566 (2000).
25. Carrington, M. Recombination within the human MHC. Immunol. Rev. 167, 243–256 (1999).
26. Jeffreys, A. J., Ritchie, A. & Neumann, R. High resolution analysis of haplotype diversity and meiotic crossovers in the human TP2 recombination hot spot. Hum. Mol. Genet. 9, 725–733 (2000).
27. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389 (1997).
28. Schuler, G. D. Sequence mapping by electronic PCR. Genome Res. 7, 541–550 (1997).
29. Zhao, C., Heil, J. & Weber, J. L. A genome wide portrait of short tandem repeats. Am. J. Hum. Genet. 65, (Suppl.) A102 (1999).
30. Huntley, G., Smith, M. W., Carrington, M. & O’Brien, S. J. A scan for linkage disequilibrium across the human genome. Genetics 152, 1711–1722 (1999).

Supplementary information is available from Nature’s World Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of Nature, and from the Marshfield Web site (http://research.marshfieldclinic.org/genetics).

Acknowledgements

This work was supported by contracts from the US National Institutes of Health and Department of Energy. Assistance was also provided by the chromosome 6 and 20 project groups at the Cancer Centre, supported by the Wellcome Trust.

Correspondence and requests for materials should be addressed to J.L.W. (e-mail: weberj@cmg.mldclinic.edu).
human genome; provides an independent validation of the sequence map and framework for contig order and orientation; surveys the genome for large-scale duplications, which are likely to require special attention during sequence assembly; and allows a stringent assessment of sequence differences between the dark and light bands of chromosomes. It also provides insight into large-scale chromatin structure and the evolution of chromosomes and gene families and will accelerate our understanding of the molecular bases of human disease and cancer.

With the draft of the human genome available, scientists can conduct global analyses of its gene content, structure, function and variation. One important challenge is to define the genetic contribution to human diseases. For many developmental disorders, inherited conditions and cancers, gross chromosomal aberrations provide clues to the locations of the causative molecular defects. These aberrations are visible as alterations in chromosomal banding patterns or in the number or relative positions of DNA sequences labelled by fluorescence in situ hybridization (FISH). Although tracing gross abnormalities to the DNA sequence has revealed the genetic causes of many diseases, molecular characterization of chromosomal aberrations has lagged far behind their discovery. To proceed from cytogenetic observation to gene discovery and mechanistic explanation, scientists will need access to a resource of experimental reagents that effectively integrates the cytogenetic and sequence maps of the human genome.

We describe here the results of a concerted effort to assemble such a genome-wide resource of well mapped, large-insert DNA clones. Each clone has been localized directly to chromosomal band(s) by FISH (Fig. 1a) and assigned one or more unique sequence tags, which can anchor the clone to the emerging draft sequence. We used complementary strategies to amass the current set of 8,877 clones.

![Figure 1](image-url)
duplicated at more than one location in the genome (to flag regions of the genome that might complicate sequence assembly\textsuperscript{7}). The molecular signatures are STSs (many corresponding to genes or expressed sequence tags (ESTs)), BAC end sequences, or the actual draft or final sequence of the clone (Table 1). Earlier publications have described genome-wide and chromosome-specific subsets of this collection\textsuperscript{8–11}.

Each clone is publicly available as single-colony-purified bacterial stocks and is ready for distribution. Each clone can each be obtained from one of three stock centres by e-mail: mapped-clones@mail.cho.org, libraries@resgen.com and clonerequest@sanger.ac.uk. The website http://www.ncbi.nlm.nih.gov/genome/cyto provides information about all clones in this collection, including how to obtain each clone. (Additional information can be obtained at the websites listed in Supplementary Information 1).

The 8,877 clones provide excellent coverage of the human genome (Table 1), with at least one clone on average per megabase (Mb) for 23 of the 24 chromosomes. Clone density ranges from greater than 5 clones per Mb for chromosomes 1, 6, 20, 22 and X to about 0.3 clones per Mb for chromosome Y.

Our study provides an assessment of the representation of the human genome in the RPCI-11 BAC library\textsuperscript{13}, which serves as the intermediate template for most sequencing efforts and the foundation of genome-wide contig assembly by fingerprint analyses.\textsuperscript{12} We randomly selected 1,243 clones from this library for FISH analysis. The number of clones assigned to each chromosome correlated well with chromosome size, with no significant bias in the distribution of clones between Giemsa (G)-dark and G-light bands of chromosomes (see Supplementary Information 2 and 3).

Cytogenetic mapping is one of several methods that can produce a framework of ordered clones upon which the human sequence can be assembled. The resource provides an opportunity to cross-check these critical framework maps, because over 3,300 FISH-mapped clones have STSs that reference the radiation hybrid\textsuperscript{14} or linkage maps\textsuperscript{15,16}. Overall, the concordance between cytogenetic map order and marker order established by radiation hybrid and linkage mapping is very high for clones with single cytogenetic locations (94–98%, depending on the map; Table 2). Significant discrepancies were observed for only around 140 of these clones and are probably due to errors in clone tracking. Integration of cytogenetic and linkage maps also aids efforts to map disease genes. The location of the cytogenetic abnormality in one patient can guide the choice of polymorphic markers to assess linkage in other families that have similar phenotypes, but no visible chromosomal aberrations.

At present, 7,303 clones that map to single cytogenetic locations are positioned by their sequence tags on the draft sequence assembly of 7 October 2000 (Table 1). The fraction of clones located on the draft sequence ranges from 76% to 91% across different

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Figure2.png}
\caption{The correspondence between cytogenetic location and position on the 7 October 2000 draft sequence for chromosome 12. The band location of each clone is indicated by a range on the y-axis. Clones mapping to chromosomes other than 12 are indicated at the bottom. Colours differentiate assignments made in different laboratories. Each clone is anchored on the draft sequence by one or more sequence tags. Plots for the other chromosomes and the 5 September, 2000 assembly can be found at http://genome.ucsc.edu/goldenPath/mapPlots/. Genome browsers that assist researchers in navigating from cytogenetic location to other maps and detailed, annotated sequence information are available at http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch (NCBI Mapviewer, which includes chromosomal aberrations associated with cancer and inherited disorders), http://www.ensembl.org/ and http://genome.ucsc.edu.}
\end{figure}
regions encompassing sequence tags of multi-site clones (either the sequence of the FISH-mapped clone or a surrogate clone from the assembly) contain blocks of homology found at an average of around 3.9 chromosomal locations (compared to around 1.3 for the regions underlying clones with single FISH signals). The regions observed by FISH and revealed through homology searches are not fully congruent, however (not shown). These findings indicate that both FISH and sequence analyses may underestimate large-scale duplications and that these complex, inter-related regions of the genome will require special attention during the finishing stages of genome sequencing.

The extensive integration of cytogenetic and primary sequence data gives investigators access to fine-structure information—including details on predicted genes—for cytogenetic locations of interest. Tools such as NCBI’s MapViewer and the UCSC and ENSEMBL genome browsers (see Fig. 2 for URLs) allow researchers to navigate readily between chromosomal location and annotated sequence.

This integration provides insight into the sequence differences underlying cytogenetic banding patterns. Sequence analyses of 200-kilobase (kb) regions surrounding the sequence tags of 338 clones mapped with the finest band resolution reveal more striking differences in the base-pair composition between Giemsa-positive and -negative bands than were predicted from earlier studies. These clones were mapped with high precision to 850-level bands of varying staining intensity on seven chromosomes. The AT content of 58 of the 59 clones in the darkest G-bands exceeds the genome-wide average of 0.59 (mean 0.63), whereas the AT content of only 22 of the 143 clones in G-negative bands is higher than average (mean 0.55; \( \chi^2 = 43, P < 0.005 \)). These data confirm that dark G-bands are more AT-rich than G-negative bands.

The utility of a sequence-integrated cytogenetic resource is illustrated by two examples. In the first, alloys are applied in conventional FISH assays to rapidly narrow the search for candidate genes disrupted or deregulated by translocations causing developmental disorders. The process is expedited by selection of clones assigned to the regions implicated by banding analyses. In a patient with multiple congenital malformations and mental retardation (DGAP012, http://dgap.harvard.edu), a breakpoint-spanning clone was identified (Fig. 1b). This clone spans a 170-kb interval containing the gene for MKK7, a human mitogen-activated protein kinase, and a novel sequence with homology to the tre-2 oncogene, both plausible candidate genes. More typically, breakpoints will be mapped to an interval between neighbouring clones. For example, a translocation implicated in mental retardation in another patient maps to an interval containing at least 12 genes, including protocadherin 8, a promising candidate given its exclusive expression in fetal and adult brain.

In the second example, an array of around 2,000 BAC clones from the collection is used to perform a genome-wide scan for segmental aneuploidy by comparative genomic hybridization (CGH) (Fig. 3 and A. Snijders et al., in preparation). The array format offers better sensitivity and resolution than metaphase chromosomes, the traditional target for CGH, and, because the arrayed clones are integrated into the draft, copy-number abnormalities can be related directly to sequence information. To illustrate the power of array CGH, the ML-2 cancer cell line was ‘karyotyped’ using the array. Array CGH revealed relative copy-number losses on 1p, 6q, 11q and 20p and gains of 12, 13 and 20q (Fig. 3). Copy-number abnormalities on chromosomes 6, 11 and 20 were subsequently confirmed by FISH using clones predicted by array CGH to be included in the region of loss. Several of these alterations were noted in previous banding analyses (2p, 6q, 11q, 12p, +12, +13q+) in a human mitogen-activated protein kinase, and a novel sequence with homology to the tre-2 oncogene, both plausible candidate genes. More typically, breakpoints will be mapped to an interval between neighbouring clones. For example, a translocation implicated in mental retardation in another patient maps to an interval containing at least 12 genes, including protocadherin 8, a promising candidate given its exclusive expression in fetal and adult brain.

In the second example, an array of around 2,000 BAC clones from the collection is used to perform a genome-wide scan for segmental aneuploidy by comparative genomic hybridization (CGH) (Fig. 3 and A. Snijders et al., in preparation). The array format offers better sensitivity and resolution than metaphase chromosomes, the traditional target for CGH, and, because the arrayed clones are integrated into the draft, copy-number abnormalities can be related directly to sequence information. To illustrate the power of array CGH, the ML-2 cancer cell line was ‘karyotyped’ using the array. Array CGH revealed relative copy-number losses on 1p, 6q, 11q and 20p and gains of 12, 13 and 20q (Fig. 3). Copy-number abnormalities on chromosomes 6, 11 and 20 were subsequently confirmed by FISH using clones predicted by array CGH to be included in the region of loss. Several of these alterations were noted in previous banding analyses (1p, 6q, 11q, +12, +13q) in a human mitogen-activated protein kinase, and a novel sequence with homology to the tre-2 oncogene, both plausible candidate genes. More typically, breakpoints will be mapped to an interval between neighbouring clones.

Many of these duplications are functionally significant, as some have generated multigene families, and some are potential sites of recombination events, which can result in chromosome abnormalities. The cytogenetic data should greatly facilitate analyses of these regions, which are likely to pose challenges to sequence assembly. The sequence tags of 84% of the clones that hybridize to more than one site were placed in the 7 October 2000 draft assembly, and the location(s) were roughly consistent with at least one FISH observation for 88% of these clones. Collectively, the multiscite clones highlight regions that are more likely to become entangled with other regions of the genome during sequence assembly than clones with single FISH locations. Indeed, global BLAST analyses show that
combination with increasingly detailed knowledge of genes and other functional motifs in the human sequence will transform the process of identifying genes that are altered in cancer and other diseases. Ultimately, this resource will contribute to a better understanding of the organization of the cell nucleus, the compacting of DNA into mitotic chromosomes, and the basis of the chromosomal banding patterns that have been so valuable in uncovering the etiology of human diseases.

Methods

GenBank was screened for draft, finished or end sequences derived from clones in this collection. BACs were screened for STS content by a combination of hybridization and polymerase chain reaction (see refs 8, 25 and Supplementary Information for details). Sequence tags were located on the draft sequence by a combination of methods (see Supplementary Information and refs 26, 27). Sequence at these locations was compiled with the results of a genome-wide BLAST analysis (ref. 2 and J. A. Bailey and E. E. Eichler, in preparation) to identify paralogous regions of the genome (regions in the draft sequence containing ~20 kb of sequence that match sequence of the FISH-mapped clone or that of a surrogate clone from the assembly at ~90% identity in non-repeat-masked bases over each 1-kb segment), and these locations were translated into estimated band positions using a dynamic programming algorithm (T. S. Furey et al., in preparation; and see Supplementary Information).

Details of FISH procedures are provided elsewhere28. Only locations of unique or low-copy portions of the clone are identified, because high-copy interspersed repetitive sequences were suppressed by addition of unlabelled copy portions of the clone are identified, because high-copy interspersed repetitive sequences were suppressed by addition of unlabelled DNA. FISH analyses were performed using DNA from the bacterial stock used for STS typing. Fluorescence ratios of Cy3 to Cy5 fluorescence for each BAC normalized to the median ratio for all 2,000 clones on the array, ordered from 1pter to Xpter. Arrows, chromosomal regions showing significant copy number variations. The lower ratio on the X indicates expected ratio for mismatched sex of test and reference DNAs. Fluorescence ratios of clones on chromosomes 11 (b) and 20 (c) are shown with clones ordered according to position of their STSs on the G3 radiation hybrid or Genethon linkage maps, respectively.

Received 7 November 1999; accepted 20 December 2000.

1. The International Human Genome Mapping Consortium. A physical map of the human genome. Nature 409, 934–941 (2001).
2. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409, 860–921 (2001).
3. Caspersson, T. et al. Chemical differentiation along metaphase chromosomes. Exp. Cell Res. 49, 219–222 (1968).
4. Trask, B. J. in Genome Analysis: A Laboratory Manual Vol. 4, 303–413 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1999).
5. Collins, F. S. Positional cloning moves from perditional to traditional. Genome Res. 9, 347–350 (1995).
6. Mittelman, F. Catalog of Chromosome Aberrations in Cancer (Wiley, New York, 1998).
7. Eichler, E. E. Masquerading repeats: paralogous pitfalls of the human genome. Genome Res. 8, 758–762 (1998).

8. Cheung, V. G. et al. A resource of mapped human bacterial artificial chromosome clones. Genome Res. 9, 989–993 (1999).
9. Korenberg, J. J. et al. Human genome anatomy: BACs integrating the genetic and cytogenetic maps for bridging genome and biomedicine. Genome Res. 9, 994–1001 (1999).
10. Leversha, M. A., Dunham, I. & Carter, N. A. A molecular cytogenetic clone resource for chromosome 22. Chromosome Res. 7, 571–573 (1999).
11. Kirsch, I. R. et al. A systematic, high-resolution linkage of the cytogenetic and physical maps of the human genome. Nature Genet. 24, 339–340 (2000).
12. Kirsch, I. R. & Ried, T. Integration of cytogenetic data with genome maps and available probes: Present status and future promise. Semin. Hematol. 37, 420–428 (2000).
13. Osogawa, K. et al. Bacterial artificial chromosome library for sequencing the human genome. Genome Res. (in the press).
14. Olivieri, M. et al. A high resolution radiation hybrid map of the human genome draft sequence. Science (in press).
15. Yu, J. et al. Comparison of human genetic and sequence-based physical maps. Nature 409, 951–953 (2001).
16. Dib, C. et al. A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 388, 90–100 (1997).
17. Trask, B. J. et al. Large multi-chromosomal duplications encompass many members of the olfactory receptor gene family in the human genome. Hum. Mol. Genet. 7, 2007–2020 (1998).
18. Sancome, S. et al. Correlations between isochores and chromosomal bands in the human genome. Proc. Natl. Acad. Sci. USA 90, 11929–11933 (1993).
19. Mittelman, F. (Ed.) Genome Analysis: A Laboratory Manual (Macmillan Press, Cold Spring Harbor, New York, 1999).
20. Strehl, S. et al. Characterization of two novel protocadherins (PCDH8 and PCDH9) localized on chromosome 13 and 16. Chromosome Res. 7, 399–407 (1999).
21. Fair, D. et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nature Genet. 28, 207–211 (1998).
22. Solinas-Toldo, S. et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosom. Cancer 26, 399–407 (1997).
23. Kallioniemi, A. O. et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258, 818–821 (1992).
24. Ohyashiki, K., Ohyashiki, J. H. & Sandberg, A. V. Cytogenetic characterization of putative human myeloblastic leukemia cell lines (ML-1, -2, and -3): origin of the cells. Cancer Res. 46, 3642–3647 (1986).
25. Morley, M. GenMapDB: A database of mapped human BAC clones. Nucleic Acids Res. 29, 144–147 (2001).
26. Schuler, G. D. Electronic PCR: bridging the gap between genome mapping and genome sequencing. Trends Biotechnol. 16, 456–459 (1998).
27. Zhan, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7, 203–214 (2000).
28. Korenberg, J. J., Chen, X. N. Human cDNA mapping using a high-resolution R-banding technique and fluorescence in situ hybridization. Cytogenet. Cell Genet. 69, 196–200 (1995).
29. Albertson, D. G. et al. Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. Nature Genet. 29, 144–146 (2000).
30. Trask, B. J., van den Engh, G., Marziali, B. & Gray, J. W. Chromosome heteromorphism quantified by high resolution bivariate flow karyotyping. Am. J. Hum. Genet. 45, 738–752 (1989).

Supplementary information is available from Nature’s World-Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of Nature.
Acknowledgements

We thank M. Arcaro, M. Bakis, J. Burdick, J. Chang, H.-C. Chen, S. Chiu, Y. Fan, C. Harris, L. Haley, R. Hosseini, J. Kent, M. A. Leversha, J. Martin, L.-T. Nguyen, P. Quinn, Y. H. Ramsey, T. Reppert, L. J. Rogers, J. Shreve, J. Stalica, M. Wang, T. Weber, A. M. Yavor, I. Young, K. Zatloukal, and members of the TIGR BAC Ends Team for assistance. This work was supported by grants from NIH (NCI, NHGRI, NIDCD and NICHD), US DOE, NSF, HHMI, PPG, Merck Genome Research Institute, Vysis, Inc., and start-up funds provided by Obstetrics and Gynecology at Brigham and Women’s Hospital.

Correspondence should be addressed to B.J.T. (e-mail: btrask@fhcrc.org).

The BAC Resource Consortium

V. G. Cheung1, N. Nowak2, W. Jang3, I. R. Kirsch4, S. Zhao5, X.-N. Chen6, T. S. Furey7, U.-J. Kim8, W.-L. Kuo9, M. Olivier10, J. Conroy2, A. Kasprzyk11, H. Massa12, R. Yonescu4, S. Sait2, C. Thoreen13, A. Snijders9, E. Lemyre14, J. A. Bailey15, A. Bruzel1, W. D. Burrill11, S. M. Clegg11, S. Collins10, P. Dhami11, C. Friedman12, C. S. Han16, S. Herrick14, J. Lee9, A. H. Ligon14, S. Lowry17, M. Morley1, S. Narasimhan1, K. Osoegawa2,18, Z. Peng17, I. Plajzer-Frick17, B. J. Quade14, D. Scott17, K. Sirotkin2, A. A. Thorpe11, J. W. Gray9, J. Hudson19, D. Pinkel9, T. Ried4, L. Rowen20, G. L. Shen-Ong4, R. L. Strausberg4, E. Birney11, D. F. Callen17, J.-F. Cheng17, D. R. Cox10, N. A. Doggett16, N. P. Carter11, E. Eichler15, D. Haussler22, J. R. Korenberg14, C. C. Morton13, D. Albertson16, G. Schuler12, P. J. de Jong10,18 & B. J. Trask12

* These authors contributed equally to this work.

1, Department of Pediatrics, University of Pennsylvania, The Children’s Hospital of Philadelphia, 3516 Civic Center Boulevard, ARC 516, Philadelphia, Pennsylvania 19104, USA; 2, Roswell Park Cancer Institute, Elm and Carleton Street, Buffalo, New York 14263, USA; 3, National Center for Biotechnology Information, National Library of Medicine, Building 38A/Room 8N805, Bethesda, Maryland 20894, USA; 4, National Cancer Institute, NIH, Building 10/Room 12N214, Bethesda, Maryland 20889-5105, USA; 5, The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA; 6, Departments of Pediatrics and Human Genetics, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, California 90048, USA; 7, Computer Science Department, University of California Santa Cruz, 1156 High Street, Santa Cruz, California 95064-1077, USA; 8, Department of Biology, California Institute of Technology, Mail Code 147-75, Pasadena, California 91125, USA; 9, University of California San Francisco Cancer Center, Box 0808, San Francisco, California 94143-0808, USA; 10, Stanford University, Genome Lab, Mail Code 5120, Stanford, California 94305-5120, USA; 11, Sanger Center, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK; 12, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North C3-168, P.O. Box 19024, Seattle, Washington 98109-1024, USA; 13, Department of Molecular Biotechnology, University of Washington, Box 357730, Seattle, Washington 98195-7730, USA; 14, Departments of Obstetrics and Gynecology and Pathology, Brigham and Women’s Hospital, Amory Lab Building 3rd floor, Boston, Massachusetts 02115, USA; 15, Department of Human Genetics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106, USA; 16, Joint Genome Institute-Los Alamos National Laboratory, MS M888 B-N1, P.O. Box 1663, Los Alamos, New Mexico 87545, USA; 17, Joint Genome Institute-Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Mail Stop 84-171, Berkeley, California 94720, USA; 18, Children’s Hospital Oakland Research Institute, 747 52nd Street, Oakland, California 94609, USA; 19, Research Genetics, 2130 Memorial Parkway, Huntsville, Alabama 35801, USA; 20, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, Washington 98105-6099, USA; 21, Department of Cytogenetics and Molecular Genetics, Women’s and Children’s Hospital, 72 King William Road, North Adelaide, South Australia 5006, Australia; 22, Howard Hughes Medical Institute, Computer Science Department, University of California Santa Cruz, 1156 High Street, Santa Cruz, California 95064–1077, USA

† Present addresses: PanGenomics, 6401 Foothill Boulevard, Tujunga, California 91024, USA (U.-J.K.); Harvard Medical School, 240 Longwood Avenue, Cell Biology, Cambridge, Massachusetts 02115, USA (C.T.); Gene Logic, Inc., 708 Quince Orchard Road, Gaithersburg, Maryland 20878, USA (G.L.S.-O.).