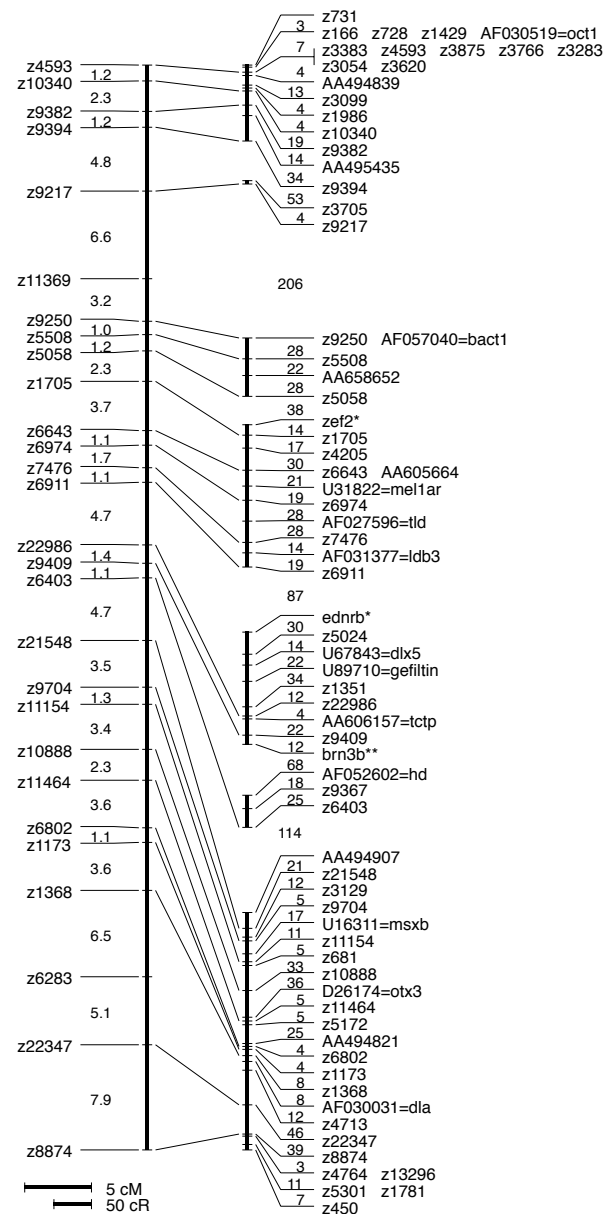


# A radiation hybrid map of the zebrafish genome

Robert Geisler<sup>1</sup>, Gerd-Jörg Rauch<sup>1</sup>, Herwig Baier<sup>1</sup>, Frauke van Bebber<sup>1</sup>, Linda Broß<sup>1</sup>, Marcus P.S. Dekens<sup>1</sup>, Karin Finger<sup>1</sup>, Cornelia Fricke<sup>1</sup>, Michael A. Gates<sup>2</sup>, Horst Geiger<sup>1</sup>, Silke Geiger-Rudolph<sup>1</sup>, Darren Gilmour<sup>1</sup>, Stefanie Glaser<sup>1</sup>, Lara Gnügge<sup>1</sup>, Hinrich Habeck<sup>1</sup>, Katy Hingst<sup>1</sup>, Scott Holley<sup>1</sup>, Jeremy Keenan<sup>1</sup>, Anette Kirn<sup>1</sup>, Holger Knaut<sup>1</sup>, Deval Lashkari<sup>3</sup>, Florian Maderspacher<sup>1</sup>, Ulrike Martyn<sup>1</sup>, Stephan Neuhaus<sup>1</sup>, Carl Neumann<sup>1</sup>, Teresa Nicolson<sup>1</sup>, Francisco Pelegri<sup>1</sup>, Russell Ray<sup>1</sup>, Jens M. Rick<sup>1</sup>, Henry Roehl<sup>1</sup>, Tobias Roeser<sup>1</sup>, Heike E. Schauerte<sup>1</sup>, Alexander F. Schier<sup>2</sup>, Ulrike Schönberger<sup>1</sup>, Helia-Berit Schönthaler<sup>1</sup>, Stefan Schulte-Merker<sup>1</sup>, Catrin Seydler<sup>1</sup>, William S. Talbot<sup>2</sup>, Christian Weiler<sup>1</sup>, Christiane Nüsslein-Volhard<sup>1</sup> & Pascal Haffter<sup>1</sup>

Recent large-scale mutagenesis screens have made the zebrafish the first vertebrate organism to allow a forward genetic approach to the discovery of developmental control genes<sup>1-3</sup>. Mutations can be cloned positionally, or placed on a simple sequence length polymorphism (SSLP) map<sup>4-6</sup> to match them with mapped candidate genes and expressed sequence tags<sup>7,8</sup> (ESTs). To facilitate the mapping of candidate genes and to increase the density of markers available for positional cloning, we have created a radiation hybrid (RH) map of the zebrafish genome. This technique is based on somatic cell hybrid lines produced by fusion of lethally irradiated cells of the species of interest with a rodent cell line. Random fragments of the donor chromosomes are integrated into recipient chromosomes or retained as separate minichromosomes<sup>9,10</sup>. The radiation-induced breakpoints can be used for mapping in a manner analogous to genetic mapping, but at higher resolution and without a need for polymorphism. Genome-wide maps exist for the human, based on three RH panels of different resolutions<sup>11-13</sup>, as well as for the dog<sup>14</sup>, rat<sup>15</sup> and mouse<sup>16,17</sup>. For our map of the zebrafish genome, we used an existing RH panel<sup>18,19</sup> and 1,451 sequence tagged site (STS) markers, including SSLPs, cloned candidate genes and ESTs. Of these, 1,275 (87.9%) have significant linkage to at least one other marker. The fraction of ESTs with significant linkage, which can be used as an estimate of map coverage, is 81.9%. We found the average marker retention frequency to be 18.4%. One cR<sub>3000</sub> is equivalent to 61 kb, resulting in a potential resolution of approximately 350 kb.

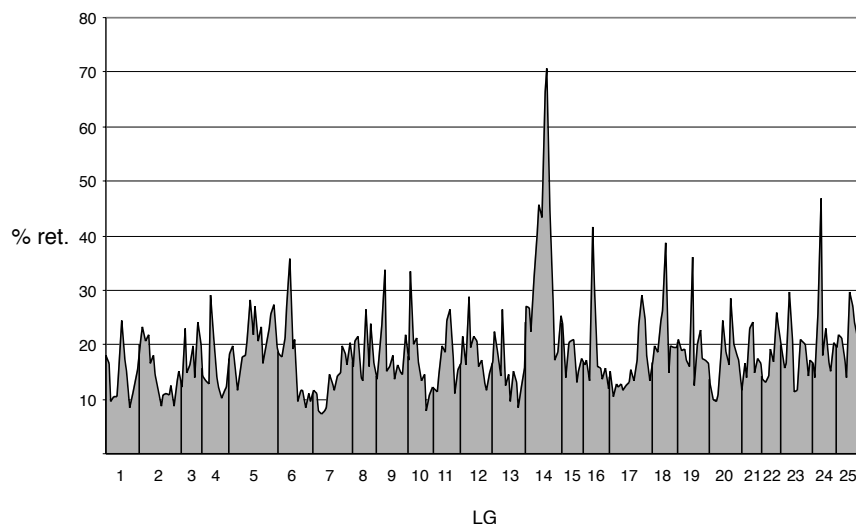
Our map is based on a panel of 94 radiation hybrids<sup>18,19</sup>. This panel, designated T51, was created by irradiating the diploid fibroblast-like cell line AB9, derived from a fin amputation of zebrafish, with a dose of 3,000 rad and fusing it with Wg3H, a hypoxanthine-guanine phosphoribosyltransferase-deficient (HPRT<sup>-</sup>) hamster cell line. Before mapping, we tested 2,125



**Fig. 1** Radiation hybrid map of zebrafish chromosome LG1 anchored to the published genetic map<sup>5,6</sup>. Only markers selected for anchoring the radiation hybrid map are indicated. The radiation hybrid map on the right shows all markers linked to at least one other marker with a lod score of 6 or more. Distances on the genetic map are shown in cM, distances on the radiation hybrid map in cR. Labels in lowercase are marker names (the prefix 'z' indicates SSLPs). Labels in uppercase are GenBank accession numbers (the prefix 'AA' indicates ESTs). Gaps indicate two-point lod scores of less than 6. \* Ref. 7. \*\* T.R. and H.B., in preparation. The other linkage groups (LG2-LG25) are published electronically only ([http://genetics.nature.com/supplementary\\_info](http://genetics.nature.com/supplementary_info)).

<sup>1</sup>Max-Planck-Institut für Entwicklungsbiologie, Spemannstr. 35, 72076 Tübingen, Germany. <sup>2</sup>Developmental Genetics Program, Skirball Institute of Biomolecular Medicine, New York University Medical Center, 540 1<sup>st</sup> Ave, New York, New York 10016, USA. <sup>3</sup>Beckman Center B407, Stanford University School of Medicine, Stanford, California 94305, USA. Correspondence should be addressed to R.G. (e-mail: [robert.geisler@tuebingen.mpg.de](mailto:robert.geisler@tuebingen.mpg.de)).

**Fig. 2** Marker retention of the zebrafish radiation hybrid panel. The width of the x axis corresponds to the map length in cR. The y axis indicates retention frequency in per cent, averaged over 100-cR intervals.



zebrafish STS markers for their PCR performance with zebrafish and Wg3H DNA. We rejected 467 markers (21.9%) at this stage, mostly because they gave no PCR product, or the zebrafish and hamster bands were not clearly distinguishable. We used the remaining markers and an additional 145 untested markers for PCR with the 94 radiation hybrids and zebrafish and Wg3H controls. We carried out reactions in duplicate using standard conditions derived from our mutant mapping project. Of these markers, 352 (19.5%) failed to give bands that could be scored.

We eventually used 1451 markers for construction of the radiation hybrid map. These markers include 1,029 SSLP markers that had been previously generated and in part genetically mapped<sup>5,6</sup>. We used 551 of these to anchor our radiation map to the genetic map. If possible, we selected one anchor for each genetic map position. Additional anchors were selected where necessary. We also used 229 cloned genes, representing most zebrafish genes published to date. We designed new primers for 150 of them, mostly from sequences available in GenBank. Primer sequences for *acvr2* and *twist* were obtained from ref. 7, and the remainder from ref. 8. Finally, we used 193 ESTs, most of them generated by the Washington University Zebrafish EST Project (M. Clark *et al.*, unpublished data). We made no attempt to select ESTs by similarity to known genes. We designed primers for 149 of the ESTs from sequences available in GenBank, the remaining primer sequences are published<sup>8</sup>.

To construct the map, we chose the computer program SAMapper (ref. 13; K. McKusick and D.R. Cox, unpublished data). SAMapper was developed for constructing the human Stanford G3 map and is particularly suited to automatic analysis of large data sets. Rather than identifying a high-confidence set of framework markers, SAMapper considers all markers

simultaneously. The markers are preordered based on two-point distances and on the genetic framework order, and the order is subsequently optimized by simulated annealing<sup>20,21</sup>. As the simulated annealing step uses even insignificant linkages across gaps, entire radiation hybrid linkage groups may be erroneously inverted or transposed. We corrected such errors to restore the framework order, but made no changes within the linkage groups.

Our map contains 1,275 markers linked to at least one other marker at a lod score of 6 or more, equivalent to 87.9% of the successfully scored markers, occupying 1,080 separate map positions (Fig. 1). These markers include 355 cloned genes and ESTs, 207 of which had not been mapped previously. Conversely, 27 genetically mapped genes and ESTs have no linkage on our map. The number of markers per chromosome varies between 29 (LG4) and 81 (LG5). The markers constitute 243 high-confidence 1,000:1 bins, that is, groups of markers that do not overlap each other in any order up to 1,000 times less likely than the best one (data not shown). The map has 190 linkage groups, equivalent to 7.6 linkage groups per chromosome, separated by gaps with two-point lod scores of less than 6. To assess the coverage of the map, it is useful to consider the set of ESTs, because they represent a less biased collection than genes cloned as candidates. The fraction of scored ESTs that were mapped was 81.9%.

**Table 1 • Properties of the zebrafish radiation hybrid panel compared with other panels**

| Organism and panel                             | Zebrafish Goodfellow T51               | Human Stanford G3 <sup>13</sup> | Human GeneBridge 4 <sup>11,12</sup> | Dog RHDF <sub>5000</sub> <sup>14</sup> | Rat Goodfellow T55v3 <sup>15</sup> | Mouse Goodfellow T31 <sup>16,17</sup> |
|--|--|---------------------------------|-------------------------------------|--|------------------------------------|---------------------------------------|
| Radiation dose                                 | 3,000 rad                              | 10,000 rad                      | 3,000 rad                           | 5,000 rad                              | 3,000 rad                          | 3,000 rad                             |
| Number of hybrids                              | 94                                     | 83                              | 91                                  | 126                                    | 96                                 | 93                                    |
| Average retention                              | 18.4%                                  | 16%                             | 30%                                 | 21%                                    | 27%                                | 30.0%                                 |
| Map length                                     | 27,729 cR <sub>3000</sub> <sup>a</sup> | 103,303 cR <sub>10000</sub>     | 11,042 cR <sub>3000</sub>           | 7,995 cR <sub>5000</sub>               | 28,234 cR <sub>3000</sub>          | 29,559 cR <sub>3000</sub>             |
| Potential resolution <sup>b</sup>              | 0.35 Mb                                | 0.24 Mb                         | 1.1 MB                              | 0.63 Mb                                | 0.41 Mb                            | 0.35 Mb                               |
| STSs mapped                                    | 1,275 <sup>c</sup>                     | 6,490 <sup>c</sup>              | 6,193                               | 347                                    | 5,255                              | 2,486                                 |
| Random STSs mapped at lod score $\geq$ 6.0     | 82%                                    | 74%                             | 100%                                | n.d.                                   | n.d.                               | 100%                                  |
| 1,000:1 bins per chromosome                    | 9.7                                    | 73.6                            | n.a.                                | n.a.                                   | n.a.                               | n.a.                                  |
| Non-anchor markers that violate bin boundaries | 2.1%                                   | 8.0%                            | n.a.                                | n.a.                                   | n.a.                               | n.a.                                  |

<sup>a</sup>Estimated distances across gaps are included. <sup>b</sup>Calculated as explained in the main text. <sup>c</sup>Expanders included.

Of 241 SSLPs that appear on the genetic as well as the radiation hybrid map, but were not used for anchoring, 14 (5.8%) are not in the same order. Most of those are inversions between neighbouring markers, which can be explained by statistical error of the genetic or radiation hybrid map or by a small number of genotyping errors in either data set. Only 5 of these SSLPs (2.1%) violate 1,000:1 bin boundaries. All of them were also assigned to different chromosomes (*z3054* to LG1, *z13289* to LG5, *z1322* to LG7, *z13* to LG8 and *gof2* to LG17). Because the genetic framework order was used only for preordering the markers, an additional 40 (8.1%) of the 492 anchor SSLPs that appear on the radiation hybrid map are in a different order, and 2 (0.4%) violate bin boundaries. To further test the accuracy of our map, we compared the radiation hybrid and genetic chromosome assignments of 148 cloned genes and ESTs. Three of the genes (2.0%) were assigned to different chromosomes (*brn1.2* to LG4, *tbx6* to LG5 and *p110a* to LG24).

The average marker retention frequency of the panel is 18.4% (Fig. 2). One of the chromosomes (LG14) shows a peak in marker retention, with an average retention frequency of 38.8%. This peak presumably indicates the position of the selective HPRT marker that was used for construction of the panel and must be retained in every hybrid. The average retention frequencies of the remaining chromosomes vary between 13.0% (LG7) and 21.5% (LG5). No clear trend is observable in the relative retention of centromeric and telomeric regions within a chromosome.

To obtain a conservative estimate of the length of the map, we eliminated markers that would expand the distance between the neighbouring markers by more than 20 cR. This expansion can be attributed to experimental error or to genuine breaks on both sides of the marker<sup>13</sup>. The sum of the distances between the remaining 1,126 markers, including estimated distances across gaps within a chromosome, is 27,729 cR. Assuming a physical size of the zebrafish genome of 1.7 Gb (refs 22,23), we estimate that 1 cR<sub>3000</sub> is approximately equivalent to 61 kb, and the average fragment size is 6.1 Mb. As each marker is, on average, represented in 17.3 radiation hybrid lines, the potential resolution of the panel is 6.1 Mb divided by 17.3, or approximately 350 kb. In comparison to the genetic map, central regions of the chromosomes are systematically better resolved by the radiation hybrid map than telomeric regions, possibly reflecting suppression of genetic recombination near the centromeres.

Our work has established that the Goodfellow T51 radiation hybrid panel is a useful tool for mapping zebrafish genes. The average retention is in between that of the human G3 radiation hybrid panel and the GeneBridge 4 panel (Table 1), and the potential resolution is close to that of the G3 panel despite the lower irradiation (3,000 versus 10,000 rad). This could be explained by a higher sensitivity of zebrafish cells to irradiation. Comparison with the G3 map suggests that it may not be possible to close all gaps and reach complete coverage by mapping additional markers. We expect, however, that the remaining gaps can be bridged by combining the T51 panel with another panel of lower resolution that has recently been characterized<sup>24</sup>.

In contrast to the human mapping effort, no prior chromosomal localization of markers is possible in the zebrafish, which would help to eliminate potential false linkages. Our map therefore contains only markers mapped at a lod score level of 6 or higher, whereas markers can be confidently positioned on the human G3 map at a lod score of 3 (ref. 13). Nevertheless, we could rapidly position 82% of the ESTs that we assayed with the RH panel, and this fraction will increase as more markers are added to the map. The zebrafish map reported here is the first radiation hybrid map of a non-mammalian genome and in-

cludes most of the available SSLPs, providing a framework for the EST mapping projects currently under way. It will also facilitate matching zebrafish mutations with candidate genes, analysis of conserved synteny between zebrafish and humans, and positional cloning of mapped mutations.

## Methods

**PCR.** We designed primers for cloned genes and ESTs with the program Primer3 (S. Rozen and H.J. Skaletsky, unpublished data available at [http://www-genome.wi.mit.edu/genome\\_software/other/primer3.html](http://www-genome.wi.mit.edu/genome_software/other/primer3.html)). Primer sequences are provided in Table 2, which is provided electronically only ([http://genetics.nature.com/supplementary\\_info/](http://genetics.nature.com/supplementary_info/)). All reactions were set up by a Biomek 2000 robotic workstation (Beckman-Coulter) on 96-well microtitre plates which were kept refrigerated below 10 °C. DNA samples of 94 zebrafish radiation hybrid lines and the zebrafish and hamster (Wg3H) controls (Research Genetics) were prediluted on a deep-well plate. Each 10- $\mu$ l reaction contained 7.14  $\mu$ l reaction mix (1  $\mu$ l of 10 $\times$ PCR buffer, 0.02  $\mu$ l each of 100 mM dATP, dCTP, dGTP and dTTP, 6.06  $\mu$ l water), 0.08  $\mu$ l each of 20  $\mu$ M forward and reverse primer, 0.2  $\mu$ l of 5 U/ $\mu$ l Taq polymerase (added just before distributing to the plates) and 2.5  $\mu$ l of 10 ng/ $\mu$ l radiation hybrid or control DNA. 10 $\times$ PCR buffer contained 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 15 mM MgCl<sub>2</sub> and 0.1% (w/v) gelatin. We doubled the radiation hybrid DNA concentration for duplicates of markers that had initially given only faint bands.

We performed PCR on 12 plates in parallel on TouchDown (Hybaid) and Primus 96 plus (MWG-Biotech) cyclers, with initial denaturing at 94 °C for 2 min, followed by 35 cycles of denaturing at 94 °C for 30 s, annealing at 60 °C for 30 s and extension at 73 °C for 1 min, and a final extension at 73 °C for 5 min. We used 2 min extension time for markers that gave a PCR product of more than 800 bp.

**Gel electrophoresis and scoring.** We added 5  $\mu$ l of 6 $\times$ loading buffer (with approximately 10 ng/ $\mu$ l of undigested plasmid DNA as a loading control) to each sample. We carried out electrophoresis at 200 V for 45 min in 1 $\times$ TBE buffer, on 2% Qualex Gold agarose gels (AGS) with 8 $\times$ 30 lanes each. PCR products from two plate rows were interspersed on the gel. For each marker, we ran one lane with 80 ng of 100-bp Ladder (Pharmacia) as a size standard. We ran each marker in duplicate. We scored even very weak bands as positive if they were clearly distinguishable from the background. In case of any discrepancy (a hybrid scored as positive on one plate and as negative on the other), we re-scored the bands. If the discrepancy persisted, we ran a third plate and the RH types supported by two of the plates were used for map construction. We also re-scored markers if any RH type differed from that of both neighbouring markers, and in case of any uncertainty, we ran another plate. Because some PCR reactions failed to give bands that could be scored, this procedure left 0.2% of the RH types undecoded. The average fraction of inconsistent RH types was 3.5%.

We captured images with NIH Image 1.61 (developed at the National Institutes of Health, available at <http://rsb.info.nih.gov/nih-image/>) on a Power Macintosh 8500/120 computer equipped with a Cohu video camera and a Scion LG digitizer board, using on-chip integration to increase sensitivity as described in the NIH Image manual. We created NIH Image macros for capturing and semi-automatic scoring of the images. We used FileMaker Pro 3.0 databases to schedule PCR runs, catalogue images, store RH types and browse mapping data.

**Map construction.** We constructed the map with SAMapper 1.0 on a DEC-station 3000-600 following standard procedures described in the SAMapper manual. We set the lod score limit for linkage groups in the preorder phase and the maximum number of markers to be considered at once for simulated annealing to the recommended values of 6.0 and 35, respectively. Deviations from the framework order were corrected automatically by a program designed by us, essentially following the editing procedure described in the SAMapper manual. First, groups of markers separated by gaps (which we define as two-point intervals with a lod score of less than 6) were identified. If the first and last anchor marker of a group were not in the correct order, the group was inverted. The groups were then sorted according to anchor marker positions, and a map reflecting the changed marker order was prepared with the SAMapper subprograms makelinear and gencompr. Because this editing procedure was applied only to entire groups of markers separat-

ed by gaps with insignificant linkage and no changes were made inside the groups, no additional simulated annealing was required. We calculated the map for the entire genome in one SAMapper run, and then divided it into chromosomes at the intervals with the lowest two-point lod score between the last anchor marker on a chromosome and the first anchor marker on the next one. We performed five rounds of expander removal with SAContinuer, with subsequent editing, for calculation of the minimal map length only. Graphical maps were created with a custom Macintosh program. The full map and raw dataset are available on our web site (<http://wwwmap.tuebingen.mpg.de>) and the ZFIN database (<http://zfish.uoregon.edu/ZFIN/>).

#### Acknowledgements

We thank F. Bonhoeffer for support of our project; N. Shimoda, D. Jackson and M. Fishman for genetic map data and primer sequences; and N. Hukriede for helpful discussions. W.S.T. is supported by NIH grant R01RR12349. P.H. is supported by a grant from the German Human Genome Project.

Received 9 March; accepted 5 August 1999.

1. Driever, W. *et al.* A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**, 37–46 (1996).
2. Haffter, P. *et al.* The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**, 1–36 (1996).
3. Haffter, P. & Nüsslein-Volhard, C. Large scale genetics in a small vertebrate, the zebrafish. *Int. J. Dev. Biol.* **40**, 221–227 (1996).
4. Knapik, E.W. *et al.* A reference cross DNA panel for zebrafish (*Danio rerio*) anchored with simple sequence length polymorphisms. *Development* **123**, 451–460 (1996).
5. Knapik, E.W. *et al.* A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genet.* **18**, 338–343 (1998).
6. Shimoda, N. *et al.* Zebrafish genetic map with 2000 microsatellite markers. *Genomics* **58**, 219–232 (1999).
7. Postlethwait, J.H. *et al.* Vertebrate genome evolution and the zebrafish gene map. *Nature Genet.* **18**, 345–349 (1998).
8. Gates, M.A. *et al.* A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. *Genome Res.* **9**, 334–347 (1999).
9. Goss, S.J. & Harris, H. New method for mapping genes in human chromosomes. *Nature* **255**, 680–684 (1975).
10. Walter, M.A., Spillett, D.J., Thomas, P., Weissenbach, J. & Goodfellow, P.N. A method for constructing radiation hybrid maps of whole genomes. *Nature Genet.* **7**, 22–28 (1994).
11. Hudson, T.J. *et al.* An STS-based map of the human genome. *Science* **270**, 1945–1954 (1995).
12. Gyapay, G. *et al.* A radiation hybrid map of the human genome. *Hum. Mol. Genet.* **5**, 339–346 (1996).
13. Stewart, E.A. *et al.* An STS-based radiation hybrid map of the human genome. *Genome Res.* **7**, 422–433 (1997).
14. Priat, C. *et al.* A whole-genome radiation hybrid map of the dog genome. *Genomics* **54**, 361–378 (1998).
15. Watanabe, T.K. *et al.* A radiation hybrid map of the rat genome containing 5,255 markers. *Nature Genet.* **22**, 27–36 (1999).
16. McCarthy, L.C. *et al.* A first-generation whole genome-radiation hybrid map spanning the mouse genome. *Genome Res.* **7**, 1153–1161 (1997).
17. Van Etten, W.J. *et al.* Radiation hybrid map of the mouse genome. *Nature Genet.* **22**, 384–387 (1999).
18. Kwok, C. *et al.* Characterization of whole genome radiation hybrid mapping resources for non-mammalian vertebrates. *Nucleic Acids Res.* **26**, 3562–3566 (1998).
19. Kwok, C., Critcher, R. & Schmitt, K. Construction and characterization of zebrafish whole genome radiation hybrids. *Methods Cell Biol.* **60**, 287–302 (1999).
20. Cox, D.R., Burmeister, M., Price, E.R., Kim, S. & Myers, R.M. Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250**, 245–250 (1990).
21. Boehnke, M., Lange, K. & Cox, D.R. Statistical methods for multipoint radiation hybrid mapping. *Am. J. Hum. Genet.* **49**, 1174–1188 (1991).
22. Hinegardner, R. & Rosen, D.E. Cellular DNA content and the evolution of teleostean fishes. *Am. Natur.* **166**, 621–644 (1972).
23. Bennet, M.D. & Smith, J.B. Nuclear DNA amounts in angiosperms. *Philos. Trans. R. Soc. Lond. B* **274**, 227–273 (1976).
24. Hukriede, N. *et al.* Radiation hybrid mapping of the zebrafish genome. *Proc. Natl Acad. Sci. USA* (in press).