BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **220,** 233–237 (1996)
ARTICLE NO. 0386

Mermaid, a Family of Short Interspersed Repetitive Elements, Is Useful for Zebrafish Genome Mapping

Nobuyoshi Shimoda,* Mario Chevrette,† Marc Ekker,‡ Yutaka Kikuchi,*,§ Yoshiki Hotta,*,¶ and Hitoshi Okamoto*,§,[1]

**Division of Cellular Communication, National Institute for Basic Biology, Nishigonaka 38, Myodaiji-cho, Okazaki, Aichi 444, Japan; †Department of Surgery, Urology Division, McGill University and Montreal General Hospital Research Institute, 1650 Avenue Cedar, Montreal, Quebec, Canada, H3G 1A4; ‡Loeb Institute for Medical Research, Ottawa Civic Hospital, University of Ottawa, 725 Parkdale Avenue, Ottawa, Ontario, Canada, K1Y 4E9; ¶Department of Physics, Graduate School of Science, University of Tokyo, Hongo7-3-1, Bunkyo, Tokyo, 113, Japan; and §Department of Physiology, Keio University, School of Medicine, 35 Shinanomachi, Shinjuku, Tokyo 160, Japan*

Received February 1, 1996

A family of short interspersed repetitive elements (SINEs), designated *mermaid,* is present in the genomes of fish, amphibian and primates, but absent in the mouse genome. We have demonstrated that the sequences of the *mermaid* family are highly polymorphic in the zebrafish genome as in the human genome. We have also shown that the *mermaid* sequence can be used to recover zebrafish specific DNA from zebrafish–mouse cell hybrids by using *mermaid*-specific oligonucleotides as PCR primers. Thus, the *mermaid* family serves as a valuable genetic tool for the zebrafish genome mapping.

The zebrafish is an attractive organism to analyze the early development of vertebrates by a systematic approach, based on large-scale mutagenesis followed by characterization of mutants (1,2). The methodology to induce mutations in the zebrafish has already been established using a chemical mutagen, *N*-ethyl-*N*-nitrosourea, which introduces point mutations in the zebrafish genome at high rate (3,4,5). To clone genes carrying point mutations, both a high-resolution genetic linkage map and appropriate genomic libraries are indispensable. The first genetic linkage map of the zebrafish has been constructed with random amplified polymorphic DNA (RAPD) markers (6), and efforts are under way to generate another linkage map with $(CA)_n$ microsatellite markers (7), which are abundant and highly polymorphic in zebrafish and other vertebrate genomes (8). However, as it has been shown in the construction of the human linkage map (9), various types of informative markers will be required to achieve a high-density genetic map of the zebrafish. Indeed, once tightly linked markers are found near a gene of interest, large genomic fragments need to be isolated from this region. This process is greatly facilitated if chromosome-specific libraries are available since they reduce the numbers of clones to be screened. Stable somatic hybrids between zebrafish and mouse that may allow the construction of such libraries of the zebrafish were produced (10).

Since short interspersed repetitive elements (SINEs) appear to be inserted irreversibly, they have been successfully used as genetic markers for population studies, evolutionary analysis, and the mapping of vertebrate genomes (11,12,13,14). For example, *Alu* repeats were utilized to construct human chromosome-specific libraries in combination with somatic hybrid cells that retain one human chromosome or subchromosome in a rodent cell (15,16). Clones containing human DNA can be isolated from genomic libraries of hybrid cells with the use of *Alu*-PCR products as probes. These clones can be aligned according to the hybridization patterns of restriction digested fragments with *Alu* probes (*Alu* fingerprinting) (17).

We have found a novel family of short interspersed repetitive elements (SINEs), named *mer-*

[1] To whom correspondence should be addressed. Fax: 81-3-3357-5445; e-mail: QZB00077@nifty serve.or.jp.

0006-291X/96 $18.00

maid, that are present in the genomes of fish, frog, and primates (unpublished). In this study, we show that the sequences of the *mermaid* family members present in fish and human are highly polymorphic. We also demonstrate that *mermaid* sequences can be used for recovery of zebrafish specific DNA from total DNA of zebrafish-mouse hybrid cells. The *mermaid* family should offer valuable information on vertebrate evolution and facilitate the genome mapping of the zebrafish.

MATERIALS AND METHODS

Cell lines. The somatic cell hybrids were derived by fusing a ZF4 zebrafish cell line (18), which contained the APH gene, with the mouse melanoma B78 cell line, followed by selection in the presence of G418 and ouabain. The details are described elsewhere (10).

Primers and PCR. The PCR was carried out in a total volume of 50 μl with 50 ng of genomic DNA, primers at 1 μM, in 50 mM KCl, 10 mM TRIS-HCl pH 8.0, 1.5 mM $MgCl_2$, 0.01% gelatin, 200 μM dNTPs, and 2.5 units of Taq polymerase (Promega) for 30 cycles of 94°C denaturation (1 min), 55°C annealing or 63°C for *mermaid* PCR (45 sec), and 72°C extension (1 min) in an automated thermal cycler (ASTEC, model PC-700). Amplification of the *mermaid* family-containing loci was performed using primer pairs flanking the *mermaid* insertions:
5′-CCTTCCTTTCTACAAGATGC-3′ and 5′-CCTAAATTTCTCTCCAACATG-3′ (Z. mer1);
5′-CAGACATGACGTGACACTCC-3′ and 5′-TTGCCAACTTCACAGAGGC-3′ (Z. mer2);
5′-TGTCCTTATTAAGCCCAGCG-3′ and 5′-TATCAAATGACCCAGCCGAG-3′ (Z. mer3);
5′-TTGTTATGTGTCGCCACCG-3′ and 5′-CGCATGTGACCTCATTCTCC-3′ (Z epd).
The sequences of the degenerated nucleotides, oligoA and oligoB, are
5′-AGAA(C/T)(A/G)TGCAAACTCCACACAGA-3′ and
5′-CCTGGAG(G/A)AAACCCAC(G/A)CA(G/A)ACA-3′, respectively.

Southern blot analysis. The *mermaid* PCR probe derived from the somatic cell hybrids, ZFB 42-2, was labeled with [^{32}P]dCTP by the BcaBest labeling kit (Takara). 5 μg of zebrafish DNA and 10 μg of mouse DNA were digested with *Dra*I and loaded on a 1% agarose gel and transferred to a nylon membrane (Hybond-N plus, Amersham). Hybridization was carried out at 65°C in 5XSSC, 0.1%SDS (w/v), 200 μg/ml sonicated zebrafish DNA, and 5×Denhardt. Posthybridization washes were done in 2×SSC, 0.1% (w/v) SDS at room temperature for 15 min, 2×SSC, 0.1% SDS (w/v) at 65°C for 20 min, and 0.1×SSC/0.1% SDS (w/v) at 65°C for 20 min.

RESULTS

The Mermaid Family Serves as Informative Polymorphic Markers

We have first found a sequence that belongs to a novel family of short interspersed repetitive elements (SINEs), designated *mermaid*, in the genome of the Japanese medaka fish (18, GenBank accession number: D78164). In addition to the genomes of other fish such as zebrafish, ray, and shark, sequences related to the medaka *mermaid* were found in the human genome (19). Interestingly, most of the human *mermaid* sequences were originally published as sequence tagged sites, and used as markers to map the human genome because of the highly polymorphic nature of $(GT)_n$ repeats (L30507, M98545, Z17075, G02317, L00843). In fact, these human *mermaid* sequences contained distinctive $(GT)_n$ repeats at the middle of their sequences, indicating that the human *mermaid* sequences are highly polymorphic in the genome and can be used as polymorphic markers. The $(GT)_n$ dinucleotides were not clustered but frequently occurred within *mermaid* sequences in fish and frog. We thus tested the possibility that *mermaid* sequences could be used as polymorphic markers. Using primers flanking four different *mermaid* loci, we amplified genomic DNAs from two distantly related zebrafish strains C32 and SJD (20). Fragments containing Z. mer1, Z. mer2, and Z. epd sequence could be produced from both the C32 and SJD strains, while the fragment that span Z. mer3 was not generated from SJD DNA with our primers. Since fragment size polymorphisms of the amplified products could not be observed on an agarose gel, these products were then subjected to digestion with restriction enzymes. As shown in Fig. 1, we detected RFLPs in the PCR products of Z. mer2 and Z. epd, but not in Z. mer1. These results indicated that three of the four *mermaid* sequences could be used to detect polymorphism between C32 and SJD strains.

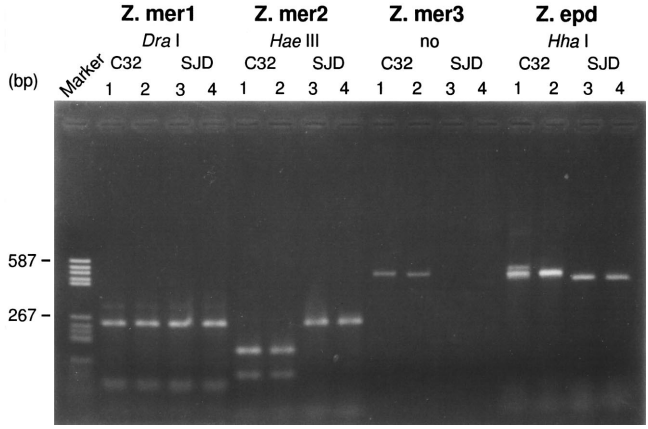

FIG. 1. RFLPs of the *mermaid* containing fragments derived from C32 and SJD zebrafish strains. Two individuals of each strain were used. PCR-amplified products containing the *mermaid* family derived from three loci were digested with the restriction enzymes presented and were fractionated on a 3% agrose gel by using *Hae*III-digested pBR322 DNA as a size standard. One locus containing Z. mer3 sequence could not be amplified from SJD strain.

Mermaid PCR on Hybrid Cell Line DNA

In order to provide rapid assignment of genetic markers to specific chromosomes, a panel of zebrafish-mouse somatic cell hybrids were produced (10). To assess whether the isolation of zebrafish chromosome-specific sequences from these hybrids was possible, we attempted PCR with three such hybrid cells, ZFB14, 33, 42-2, by using one of the two oligonucleotides, termed oligoA and oligoB, complementary to the zebrafish *mermaid* as a primer (*mermaid* PCR). Fig. 2 shows that each primer generated fragments amplified from the three DNA sources, and that the banding patterns are distinct for different hybrids. Both primers amplified a smear of fragments from ZFB33. OligoB generated a large number of fragments from ZFB14 and ZFB42-2, while oligoA could only amplify few bands from these two hybrids. These results suggested that each hybrid cells used in this study retained different zebrafish chromosomal elements, a result consistent with

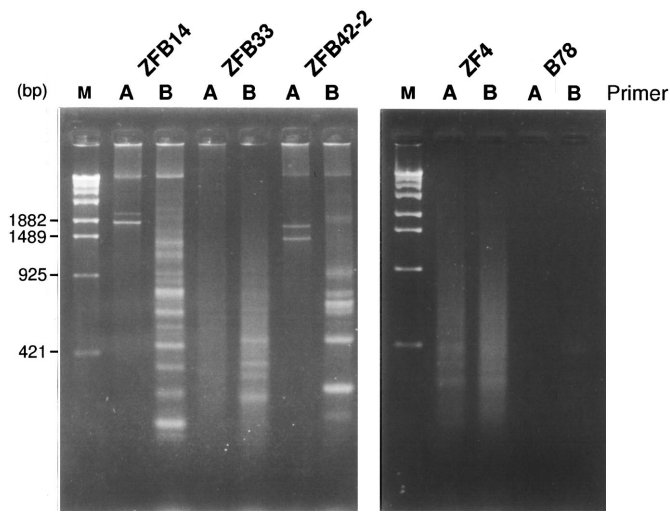

FIG. 2. Amplification of DNAs from three types of zebrafish/mouse hybrid cell, zebrafish ZF4 cell, and mouse B78 cell with oligoA and oligoB primers. PCR-amplified products were fractionated on a 3% agarose gel using *Sty*I-digested lambda DNA as a size standard (M).

the assignment of these hybrids to different linkage groups using specific markers (10). To ascertain that the amplified sequences were derived from zebrafish DNA, we hybridized one of these sequences to zebrafish and mouse genomic DNA. A PCR fragment of 1.4 kb that was generated from hybrid ZFB42-2 with the oligoA primer was used as a probe. The hybridization was also carried out in an excess of the zebrafish genomic DNA to block hybridization of the probe to the numerous *mermaid* sequences present in the PCR products. Fig. 3 shows that the probe hybridized to a fragment of DNA present in the zebrafish (Fig. 3, lane3) but not in the mouse (Fig. 3, lane2). The fragment was masked by the smear of hybridization signal attributable to the multiple *mermaid* sequence when zebrafish carrier DNA was omitted from the hybridization mixture (Fig. 3, lane1). This result demonstrated that *mermaid* PCR could produce zebrafish-specific DNA from zebrafish-mouse cell hybrids.

DISCUSSION

We found that sequences of the *mermaid* family are highly polymorphic in the zebrafish and human genomes. In contrast to dinucleotide repeats whose polymorphism is detected by using denaturing polyacrylamide gel and radiolabeled probes, the polymorphism of the zebrafish *mermaid* were readily detected by PCR alone, or by PCR followed by restriction enzyme digestion. Polymorphism due to differences in the numbers of the $(GT)_n$ dinucleotide repeats may also be detected within the zebrafish *mermaid* sequences as they are in humans. Besides *mermaid*, a number of SINEs have been found to be associated with microsatellite repeats (21,22) and *Alu* has been postulated to be a source for the genesis of primate microsatellites (23).

In addition to the degree of polymorphism, the availability of a genetic marker for a linkage map largely depends on its abundance and distribution in genome. Although the extent to which the zebrafish *mermaid* sequence are dispersed in the genome is unclear, the average distance between the sequences was estimated to be 140 kb (18), corresponding to about 0.2 cM in the genome (6). Assuming that they are randomly distributed in the genome, the distance is close enough for high-resolution linkage mapping of the genome and one can expect one or a few *mermaid* sequences in every insert of P1 and YAC vectors, respectively (24). Since such moderately repetitive sequences in *C. elegans* and human have been proven to help to assemble contigs of those vectors using them as hybridization probes (17,25), a linkage map generated by the zebrafish *mermaid* sequence may be efficiently integrated with a physical map.

Finally, we used the absence of *mermaid* sequences from mouse DNA to directly amplify

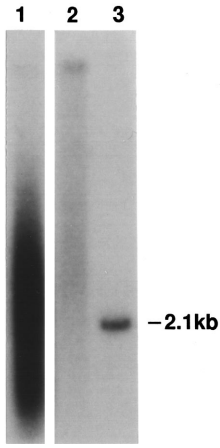

FIG. 3. Southern blot hybridization of *mermaid* PCR products from ZFB42-2 hybrids to genomic DNAs from zebrafish cell line, ZF4 (lane 1, 3), and mouse cell line, B78 (lane 2). Hybridization was performed in the absence (lane 1) or presence (lane 2, 3) of zebrafish genomic DNA added as carrier DNA to eliminate hybridization of the *mermaid* sequence.

zebrafish sequences from zebrafish/mouse hybrid cells by *mermaid* PCR. This technique provides a simple method for the isolation and analysis of specific chromosomal regions of the zebrafish. Thus, the zebrafish *mermaid* sequences will facilitate the construction of linkage and physical maps of zebrafish, and eventually the positional cloning of genes affected in zebrafish mutants. The *mermaid* family will also be useful for the genome analysis of other organisms where this family is found.

ACKNOWLEDGMENTS

We thank Stephen Johnson for providing C32 and SJD strains of the zebrafish. We are grateful to S. Yasumasu for providing the medaka genomic library, and S. Matsubayashi and K. Takamatsu for their assistance in maintaining the fish colony. We thank Cristofre Martin for critical reading of the manuscript. This work was supported by a Grant-in-Aid (07262217, 07279107, 07np0701) from Ministry of Education, Science and Culture of Japan, Special Coordination Funds for Promoting Science and Technology from Science and Technology Agency of Japan, and Uehara Memorial Foundation. N.S. and Y.K. are supported by National Institute for Basic Biology and Y.K. is supported by a fellowship from the Japan Society for the Promotion of Science for Japanese Junior Scientists.

REFERENCES

1. Driever, W., Stemple, D., Schier, A., and Solnica-Krezel, L. (1994) *Trends Genet.* **10,** 152–159.
2. Nüsslein-Volhard, C. (1994) *Science* **266,** 572–574.
3. Riley, B. B., and Grunwald, D. J. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5997–6001.
4. Solnica-Krezel, L., Schier, A. F., and Driever, W. (1994) *Genetics* **136,** 1401–1420.
5. Mullins, M. C., Hammerschmidt, M., Haffter, P., and Nüsslein-Volhard, C. (1994) *Curr. Biol.* **4,** 189–202.
6. Postlethwait, J. H., Johnson, S. L., Midson, C. N., Talbot, W. S., Gates, M., Ballinger, E. W., Africa, D., Andrews, R., Carl, T., Eisen, J. S., Horne, S., Kimmel, C. B., Hutchinson, M., Johnson, M., and Rodriguez, A. (1994) *Science* **264,** 699–703.
7. Knapik, E. W., Goodman, A., Atkinson, O. S., Roberts, C. T., Shiozawa, M., Sim, C. U., Weksler-Zangen, S., Trolliet, M. R., Futrell, C., Innes, B. A., Koike, G., McLaughlin, M. G., Pierre, L., Simon, J. S., Vilallonga, E., Roy, M., Chiang, P-W., Fishman, M. C., Driever, W., and Jacob, H. J. (1995) *Development* in press.
8. Goff, D. J., Galvin, K., Katz, H., Westerfield, M., Lander, E. S., and Tabin, C. J. (1992) *Genomics* **14,** 200–202.
9. Murray, J. C., Buetow, K. H., Weber, J. L., Ludwigsen, S., Scherpbier-Heddema, T., Manion, F., Quillen, J., Sheffield, V. C., Sunden, S., Duyk, G. M., Weissenbach, J., Gyapay, G., Dib, C., Morrissette, J., Lathrop, G. M., Vignal, A., White, R., Matsunami, N., Gerken, S., Melis, R., Albertsen, H., Plaetke, R., Odelberg, S., Ward, D., Dausset, J., Cohen, D., and Cann, H. (1994) *Science* **265,** 2049–2054.
10. Ekker, M., Speevak, M. D., Martin, C. C., Joly, L., Giroux, G., and Chevrette, M. (1996) *Genomics,* in press.
11. Batzer, M. A., Stoneking, M., Alegria-Hartman, M., Bazan, H., Kass, D. H., Shaikh, T. H., Novick, G. E., Ioannou, P. A., Scheer, W. D., Herrera, R. J., and Deininger, P. L. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 12288–12292.
12. Murata, S., Takasaki, N., Saitoh, M., and Okada, N. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6995–6999.
13. Schalkwyk, L. C., Francis, F., and Lehrach, H. (1995) *Curr. Biol.* **6,** 37–43.
14. McCarthy, L., Hunter, K., Schalkwyk, L., Riba, L., Anson, S., Mott, R., Newell, W., Bruley, C., Bar, I., Ramu, E., Housman, D., Cox, R., and Lehrach, H. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5302–5306.
15. Nelson, D. L., Ledbetter, S. A., Corbo, L., Victoria, M. F., Ramírez-Solis, R., Webster, T. D., Ledbetter, D. H., and Caskey, C. T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6686–6690.
16. Wang, D., Zhu, Y., and Smith, C. L. (1995) *Genomics* **26,** 318–326.
17. Zucchi, I., and Schlessinger, D. (1992) *Gemomics* **12,** 264–275.
18. Driever, W., and Rangini, Z. (1993) *In Vitro Cell Dev. Biol.* **29A,** 749–754.
19. Shimoda, N., Chevrette, M., Ekker, M., Kikuchi, Y., Hotta, Y., and Okamoto, H. (1995) *Biochem. Biophys. Res. Commun.,* accompanying paper.
20. Johnson, S. L., Africa, D., Horne, S., and Postlethwait, J. H. (1995) *Genetics* **139,** 1727–1735.
21. Alexander, L. J., Rohrer, G. A., Stone, R. T., and Beattie, C. W. (1995) *Mammal. Genome* **6,** 464–468.
22. Pearlman, R. E., Tsao, N., and Moens, P. B. (1992) *Genetics* **130,** 865–872.
23. Arcot, S. S., Wang, Z., Weber, J. L., Deininger, P. L., and Batzer, M. A. (1995) *Genomics* **29,** 136–144.
24. Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A., and De Jong, P. J. (1994) *Nature Genet.* **6,** 84–89.
25. Cangiano, G., Ameer, H., Waterston, R., and La Volpe, A. (1990) *Nucleic Acids Res.* **18,** 5077–5081.