Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 5446–5449, August 1985
Genetics

A 78-kilobase region of mouse chromosome 3 contains salivary and pancreatic amylase genes and a pseudogene

(multigene family/cosmid cloning/subcloning strategy)

KARIN WIEBAUER*, DEBORAH L. GUMUCIO, JULIE M. JONES†, ROBERT M. CALDWELL, HEATHER T. HARTLE‡, AND MIRIAM H. MEISLER§

Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109-0010

Communicated by Salome G. Waelsch, April 30, 1985

ABSTRACT **Genetic studies have demonstrated that salivary and pancreatic amylase genes are closely linked in human and mouse. To analyze the arrangement of genes within the amylase cluster, a library of YBR mouse genomic DNA was cloned in the cosmid vector pJB8. Clones containing amylase genes were identified by hybridization with amylase cDNA probes. Salivary and pancreatic amylase genes were isolated on separate cosmid clones, but no overlapping clones were evident from the initial screening. A strategy for the rapid isolation of terminal noncoding fragments from the cosmid clones was developed. By using these terminal fragments for chromosome "walking," a map of 78 kilobases of the amylase gene region was constructed. The salivary and pancreatic amylase genes are present within this region in the same 5′-to-3′ orientation, separated by 22 kilobases of genomic DNA. A truncated amylase pseudogene is located 10 kilobases downstream from the pancreatic amylase gene.**

In the human and mouse genomes, distinct genes encode the α-amylases (EC 3.2.1.1) produced in salivary gland and in pancreas (1–4). The extensive sequence homology between salivary (Amy-1) and pancreatic (Amy-2) amylase cDNAs (>90%) indicates that the two genes, *Amy-1* and *Amy-2*, are derived from a common ancestral gene (1, 4). Close linkage of *Amy-1* and *Amy-2* has been established by genetic analysis of the segregation of electrophoretic variants (3, 5). The offspring of more than 1500 meioses have been studied in crosses between inbred mice without observation of recombination between these genes (6). This data suggested that the intergenic distance between *Amy-1* and *Amy-2* may be less than 150 kilobases (kb). Consistent with their close linkage and sequence homology, the amylase genes appear to have undergone "correction" events during mammalian evolution (4, 7).

The present study was undertaken to determine the precise organization of amylase gene copies within the mouse multigene cluster. To facilitate analysis of an extended chromosome region, cloning was carried out in a cosmid vector that accommodates 35- to 45-kb inserts of genomic DNA. We chose mouse strain YBR for these studies because of our interest in the independent regulation of its two active pancreatic amylase genes, *Amy-2.1* and *Amy-2.2* (3, 8–10). We report here the characterization of a 78-kb portion of the amylase gene region that includes one salivary amylase gene, the *Amy-2.1* pancreatic amylase gene, and an apparent amylase pseudogene. A subcloning strategy that facilitates chromosome "walking" in cosmid libraries is also described.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Genomic DNA libraries were constructed from partially *Mbo* I-digested DNA isolated from livers of YBR/Ki mice. Restriction fragments of 35–50 kb were size-selected by centrifugation through a neutral sucrose gradient and inserted into the *Bam*HI site of the cosmid vector pJB8 (11, 12). Recombinant molecules were packaged *in vitro* into λ phage heads and used to infect *Escherichia coli* strain 490A (13). The yield of infectious phage particles containing recombinant cosmids was greater than 10^5 per μg of chromosomal DNA. In order to screen a genomic library of 200,000–300,000 clones, approximately 10,000 colonies were plated per 150-mm Petri dish and replica-plated onto nitrocellulose filters (Millipore) (14). Filters were hybridized with a mixture of salivary and pancreatic amylase cDNAs from A/J mice (pMPa21 and pMSal04; ref. 1). These clones contain all of the coding regions of the mRNAs. Recombinant DNA was isolated in minipreparations (15, 16). Restriction site mapping was carried out by analyzing single, double, and partial digests of the cloned DNA.

To isolate the fragment used for chromosome "walking," approximately 60 ng of *Hin*dIII-digested DNA from cosmid cSam47 was self-ligated in 10 μl of 66 mM Tris·HCl buffer (pH 7.5) containing 5 mM $MgCl_2$, 5 mM dithiothreitol, and 1 mM ATP, in the presence of 1–3 Weiss units of T4 DNA ligase (New England Biolabs) at 16°C overnight. After *E. coli* strain DH1 was transformed with 1 μl of the ligation mix, colonies were selected on Luria broth (LB) agar containing ampicillin at 50 μg/ml. The plasmid DNA isolated from transformants contained the vector fragment (with the *ori* and Amp^R sequences) plus the 3′-terminal insert fragment (see the legend to Fig. 2). This insert fragment was isolated from the vector by digestion with *Eco*RI and *Hin*dIII, followed by gel purification. The fragment was tested for the presence of repetitive sequences by probing a Southern blot of the subcloned DNA with nick-translated genomic DNA and also by using the labeled fragment to probe a Southern blot of genomic DNA.

RESULTS

In the initial screening of two independently constructed cosmid gene libraries, several recombinant cosmids were isolated by hybridization with a mixture of salivary and pancreatic amylase cDNA clones. The cloned genes were readily classified as *Amy-1* or *Amy-2* by comparison of their

Abbreviation: kb, kilobase(s).

*Present address: Friedrich Miescher Institute, CH-4002, Basel, Switzerland.
†Present address: Department of Microbiology and Molecular Genetics, University of California, Irvine, CA 92717.
‡Present address: Department of Internal Medicine, Medical College of Virginia, Richmond, VA 23298.
§To whom reprint requests should be addressed.

restriction maps with those for the amylase genes from strain A/J (18), as shown in Fig. 1 for a representative clone of each type. Of the 10 mapped restriction sites in *Amy-1* from A/J, 9 were found in clone cSam47. The eight restriction sites in *Amy-2*^{a} from A/J were all present in the pancreatic gene in clone cPamE1, which is designated *Amy-2.1*. The identification of the salivary and pancreatic amylase genes has been confirmed by sequence analysis (data not shown). In contrast to the coding regions, the restriction sites in the flanking regions of the A/J and YBR genes are not conserved.

None of the cosmid clones contained more than a single amylase gene copy, with the exception of the cosmid cPamE1, which contains an apparent pseudogene located approximately 10 kb downstream of the *Amy-2.1* gene. The pseudogene hybridizes with a full-length amylase cDNA but does not hybridize with a fragment containing only exons a and b. There were no obvious overlaps between the flanking regions of different clones.

To obtain overlapping clones that would link the salivary and pancreatic amylase genes, the cosmid libraries were screened with terminal insert fragments from the amylase clones. Terminal fragments were isolated by recircularization of cosmid DNA after digestion with a restriction enzyme that cleaves the vector DNA at a single site close to the cloning site. The method is outlined in Fig. 2. This approach provides a rapid procedure for subcloning one end fragment of a cloned insert. Depending upon the orientation of the gene within the insert, either a 5′ or a 3′ terminal fragment will be isolated. The advantage of this strategy is that chromosome walking can be carried out without extensive mapping of the insert. Comparison of the clones obtained by chromosome walking with the original insert also provides confirmation of the integrity of the original clone.

Using this procedure, we subcloned from cosmid cSam47 a terminal fragment located 14 kb downstream from the *Amy-1* gene. Since the subcloned fragment did not contain *Eco*RI sites, the intact 0.75-kb insert could be recovered after digestion with *Eco*RI and *Hin*dIII. After ensuring that the gel-purified fragment did not contain repetitive elements, we used it to screen a new cosmid library consisting of approximately 200,000 colonies. Two of the recovered clones, cSamJ4 and cSamJ8, proved to include segments that are present in the *Amy-1* clones (cSam47 and cSamD4) and in the *Amy-2.1* clone (cPamE1). The restriction maps of these clones provide the overlap between the *Amy-1* and *Amy-2.1* genes (Fig. 3). The two genes are present in the same orientation, 5′ to 3′, along the chromosome and are separated by approximately 22 kb of intergenic DNA.

DISCUSSION

Chromosome walking usually requires generation of a restriction map for the entire cloned insert, followed by isolation and subcloning of fragments that have been mapped to the terminal positions. This is particularly difficult in the case of chromosome walking with cosmids because of the complexity of mapping the large cosmid inserts. The procedure outlined in Fig. 2 provides considerable advantage because subclones of a terminal fragment are obtained directly without extensive mapping or subcloning.

The results of our studies demonstrate that a mouse salivary amylase gene is located approximately 22 kb upstream from a pancreatic amylase gene. This small intergenic distance can account for the lack of recombination (6), since the predicted recombination rate would be approximately 1 per 10,000 meioses if recombination events were randomly distributed in the genome. Our result agrees closely with the intergenic distance recently observed for *Amy-1* and *Amy-2* in strain A/J (24) and, thus, is likely to be a general result for *Mus domesticus*.

The YBR genome contains an additional amylase gene, *Amy-2.2* (25). We have mapped 10 kb of the 5′ flanking region and 7 kb of the 3′ flanking region of *Amy-2.2* without encountering overlap with the 78-kb gene cluster described here. The minimal size of the YBR amylase gene region is thus 100 kb. In other inbred strains of mice, which contain extra copies of *Amy-2* (19, 26), the amylase gene region must be considerably larger.

Approximately 10 kb downstream from the YBR *Amy-2.1* gene is a 5-kb region that hybridizes with the full-length amylase cDNAs but not with a fragment containing only exons a and b. The restriction map of the 5-kb region is not obviously related to the *Amy-1* or *Amy-2* genes. It presumably represents an amylase pseudogene or, less likely, an

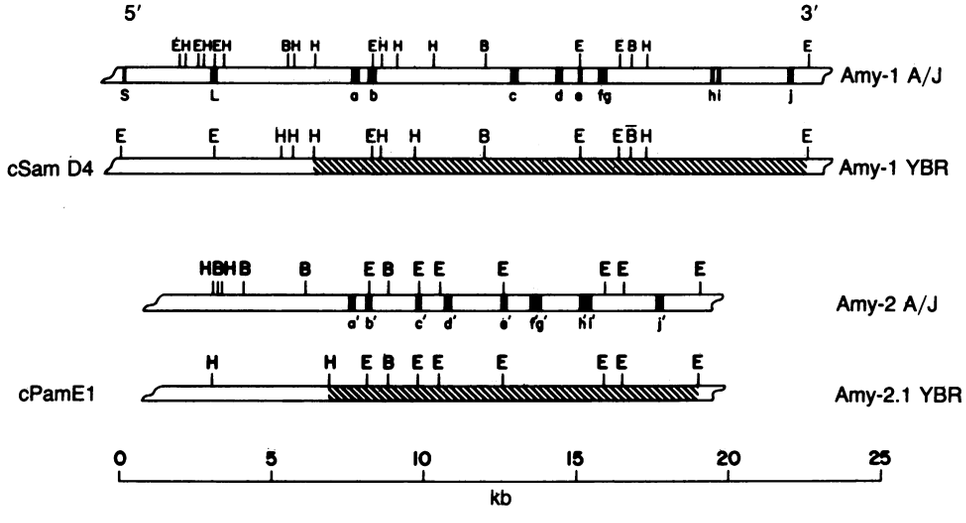

FIG. 1. Restriction maps of the coding regions of mouse *Amy-1* and *Amy-2* genes. Cosmids cPamE1 and cSam47 were isolated from a library of YBR genomic DNA. Coding regions were mapped by partial digestion and by Southern blot analysis of the products of single and double restriction endonuclease digestion using subclones of the cDNAs as probes. The extent of the hybridizing regions is indicated by the hatched boxes. The restriction maps of the coding regions of the two YBR genes are compared with those of the *Amy-1* and *Amy-2* genes from strain A/J, which have been redrawn from Schibler *et al.* (18). The exons are indicated by lower case letters. E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; $\overline{B}$ represents two *Bam*HI sites separated by 0.2 kb.

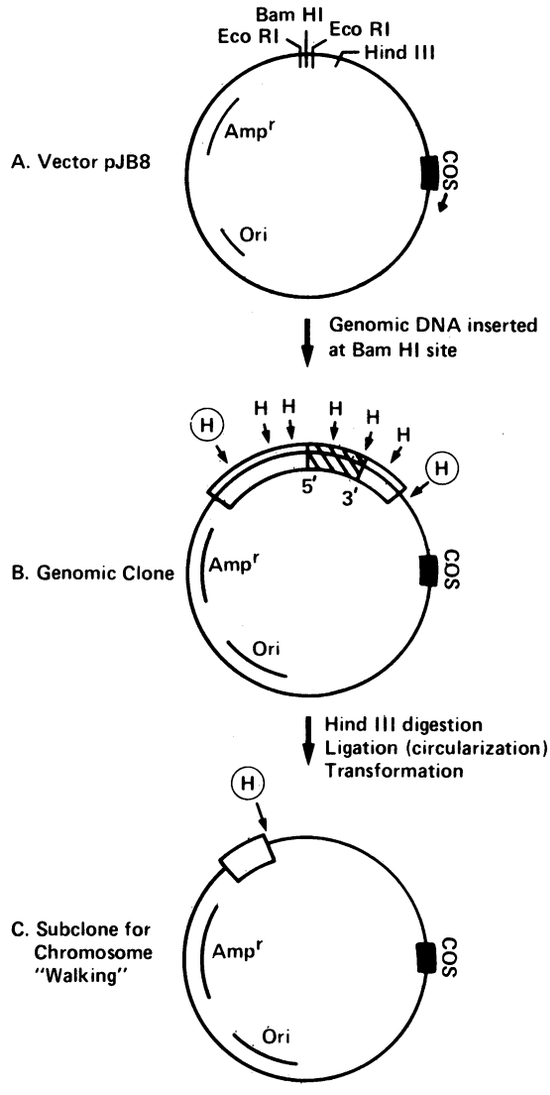

FIG. 2. Strategy for isolation of terminal subclones from inserts in pJB8. (*A*) The *Bam*HI cloning site and flanking *Eco*RI sites of the cosmid vector pJB8. (*B*) A recombinant molecule with a cloned insert containing a gene (hatched region) in one possible orientation. To isolate terminal fragments from the insert, the recombinant DNA is digested to completion with the restriction endonuclease *Hin*dIII, which cleaves the inserts at multiple sites but cleaves the vector at a single site close to the cloning site. The restriction fragments are allowed to circularize in the presence of DNA ligase. *E. coli* strain DH1 is transformed with the ligation mix, and transformants are selected on LB plates containing ampicillin. Only cells transformed by the vector fragment containing the *ori* and Amp^R sequences will grow. The subclone also will contain one of the terminal insert fragments; in this example, the 5'-terminal fragment of the insert is recovered. Note: The figure is not drawn to scale; the actual size of the insert (35–45 kb) is greater than that of the vector (5.4 kb). H, *Hin*dIII.

amylase gene with an extremely divergent 5' region. Additional characterization of the pseudogene is in progress.

This report is an initial step in the characterization of the chromosomal organization of the amylase gene family, which has been less well studied than other gene clusters. Duplication of the amylase gene has been followed by divergence of regulatory signals controlling tissue specificity. The time of divergence of salivary and pancreatic amylase genes can be estimated from the time of origin of salivary glands, first seen in amphibia approximately 400 million years ago. Since several extant amphibian species produce salivary amylase, it is likely that both *Amy-1* and *Amy-2* are present in the amphibian genome. Thus, the linkage of *Amy-1* and *Amy-2* has survived for an extended period of evolution. During the same period, the α-globin and β-globin gene families, which are linked in the amphibian *Xenopus*, have become dispersed on different mammalian chromosomes (21). One consequence of retained linkage of duplicated genes is the continuing potential for correction of sequences by gene conversions. Such events clearly have occurred between *Amy-1* and *Amy-2* in the mouse and human genomes, since in each species the amylase genes are more similar to each other than to the homologous gene of the other species (4). Correction over an intergenic distance of 22 kb is not without precedent, as it apparently has occurred between the mouse immunoglobin γ2b and γ2a chain genes which are separated by the comparable distance of 15 kb (22).

The intergenic distance and orientation of the salivary and pancreatic amylase genes resembles certain other gene families that have been studied. For example, the mouse albumin and α-fetoprotein genes are separated by 13.5 kb and also are transcribed from the same DNA strand (17). These genes have diverged with respect to developmental regulation while retaining the same tissue specificity. Two copies of the chicken δ-crystallin gene are also present in the same 5'-to-3' orientation, separated by 4.2 kb (23). Both crystallin genes are expressed in lens tissue, but one gene is considerably more active. As these examples demonstrate, duplicated gene copies that remain in close physical proximity may diverge with respect to tissue specificity, developmental timing, and quantitative expression.

We thank P. K. Wellauer, U. Schibler, O. Hagenbüchle, M. Tosi, and R. A. Young (Swiss Institute for Cancer Research, Lausanne) for generously providing the amylase cDNA clones used in the project; Richard W. Padgett (University of North Carolina) for stocks of pJB8 and *E. coli* strain 490A; and David Hewett-Emmett, Linda Samuelson, and Laurie Treisman for helpful discussions. K.W. was the recipient of an Arthur F. Thurnau Postdoctoral Fellowship from the University of Michigan. This work was supported by Grant GM 24872 from the National Institutes of Health.

1. Hagenbüchle, O., Bovey, R. & Young, R. A. (1980) *Cell* **21**, 179–187.
2. Meisler, M., Strahler, J., Wiebauer, K. & Thomsen, K. K. (1983) in *Isozymes: Current Topics in Biological and Medical Research*, eds. Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), Vol. 7, pp. 39–57.
3. Meisler, M. H. & Gumucio, D. L. (1985) in *Molecular and Cellular Biology of Digestion*, eds. Desnuelle, P., Sjöström, H. & Noren, O. (Elsevier/North-Holland, Amsterdam), in press.
4. Nakamura, Y., Ogawa, M., Nishide, T., Emi, M., Kosaki, G., Himeno, S. & Matsubara, K. (1984) *Gene* **28**, 263–270.
5. Merritt, A. D. & Karn, R. C. (1977) in *Advances in Human Genetics*, eds. Harris, H. & Hirschhorn, K. (Plenum, New York), Vol. 8, pp. 135–234.
6. Bloor, J. & Meisler, M. (1980) *J. Hered.* **71**, 449–451.
7. Milner-White, E. J. (1984) *Trends. Biochem. Sci.* **9**, 517–519.
8. Strahler, J. R. & Meisler, M. (1982) *Genetics* **99**, 91–102.
9. Dranginis, A., Morley, M., Nesbitt, M., Rosenblum, B. B. & Meisler, M. H. (1984) *J. Biol. Chem.* **259**, 12216–12219.
10. Bloor, J., Meisler, M. & Nielsen, J. T. (1981) *J. Biol. Chem.* **256**, 373–377.
11. Chia, W., Scott, R. D. & Rigby, P. W. J. (1982) *Nucleic Acids Res.* **10**, 2503–2520.
12. Ish-Horowicz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* **9**, 2989–2993.
13. Enquist, L. & Sternberg, N. (1979) *Methods Enzymol.* **68**, 281–298.
14. Hanahan, D. & Meselson, M. (1980) *Gene* **10**, 63–67.
15. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1515.

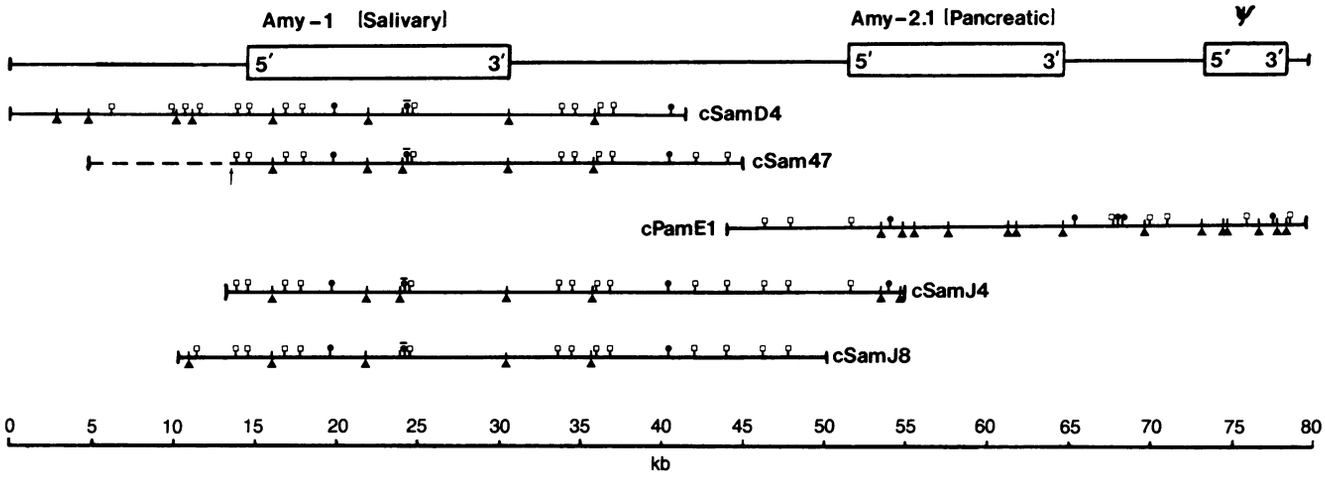

FIG. 3. Restriction maps of overlapping cosmid inserts from a 78-kb region of mouse chromosome 3. The schematic drawing (top line) summarizes the data from restriction site mapping and hybridization analysis of the overlapping cosmid clones obtained in the initial screening with amylase cDNA (cSam47, cSamD4, and cPamE1) and those obtained from the subsequent chromosome walk (cSamJ4 and cSamJ8). The boxed regions indicate the extent of restriction fragments that hybridize with amylase cDNA probes. The arrow indicates the position of a possible rearrangement within cSam47. △, *Eco*RI; □, *Hin*dIII; ●, *Bam*HI; $\overline{\bullet}$, the two *Bam*HI sites in *Amy-1* that are separated by 0.2 kb.

16. Holmes, D. S. & Quigley, M. (1981) *Anal. Biochem.* **114,** 193–196.
17. Ingram, R. S., Scott, R. W. & Tilghman, S. M. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 4694–4698.
18. Schibler, U., Pittet, A. C., Young, R. A., Hagenbüchle, O., Tosi, M., Gellman, S. & Wellauer, P. K. (1982) *J. Mol. Biol.* **155,** 247–266.
19. Hagenbüchle, O., Wellauer, P. K., Cribbs, D. L. & Schibler, U. (1984) *Cell* **38,** 737–744.
20. Tosi, M., Bovey, R., Astolfi, S., Bodary, S., Meisler, M. H. & Wellauer, P. K. (1984) *EMBO J.* **3,** 2809–2816.
21. Jeffreys, A. J., Wilson, V., Wood, D. & Simons, J. P. (1980) *Cell* **21,** 555–564.
22. Ollo, R. & Rougeon, F. (1983) *Cell* **32,** 515–523.
23. Hawkins, J. W., Nickerson, J. M., Sullivan, M. A. & Piatigorsky, J. (1984) *J. Biol. Chem.* **259,** 9821–9825.
24. Pittet, A.-C. & Schibler, U. (1985) *J. Mol. Biol.* **182,** 359–365.
25. Gumucio, D. L., Wiebauer, K., Dranginis, A., Samuelson, L. C., Treisman, L. O., Caldwell, R. M., Antonucci, T. K. & Meisler, M. H. (1985) *J. Biol. Chem.*, in press.
26. Bodary, S., Grossi, G., Hagenbüchle, O. & Wellauer, P. K. (1985) *J. Mol. Biol.* **182,** 1–10.