Molecular Evolution of the Mouse Proline-rich Protein Multigene Family

INSERTION OF A LONG INTERSPERSED REPEATED DNA ELEMENT

(Received for publication, March 31, 1988)

David K. Ann, M. Kathleen Smith, and Don M. Carlson
From the Department of Biochemistry and Biophysics, University of California, Davis, California 95616

Proline-rich proteins (PRPs) in the salivary glands of mice, rats, and hamsters are encoded by tissue-specific inducible multigene families. Mouse PRP genes are located on chromosome 8, and transcription is dramatically induced (about 70-fold) by isoproterenol treatment. Clones containing two nonallelic PRP genes (MP2 and M14) were isolated from cosmids and phage libraries of CD-1 mouse genomic DNA. The cloned regions comprise a contiguous block of 77-kilobase pairs of the mouse genome. Restriction mapping established the physical lineage of PRP genes MP2 and M14, and they are tandemly arrayed. The DNA sequence analysis presented in this report suggests that genes M14 and MP2 (Ann, D. K., and Carlson, D. M. (1985) J. Biol. Chem. 260, 15863–15872) arose via a gene duplication of a common ancestor. Two major differences between M14 and MP2 were observed. PRP gene MP2 has 13 tandemly arrayed 42-nucleotide repeats in exon II, whereas M14 has 17 repeats, and PRP gene M14 has an insertion by transposition of a 2-kilobase pair member of the long interspersed repeated DNA (LINE) family (LIMd) into intron I. The evolution of this PRP multigene family has been dominated by intra-exonic amplification of repeating nucleotide units coding for these and other proline-rich repeated peptides and by gene duplication. The LIMd element gives rise to heterogenous EcoRI, BamHI, and HindIII restriction enzyme patterns, and this insertion is also present in BALB/c, C57BL/6J, and DBA/2J mice.

Mammalian proline-rich proteins (PRPs) are encoded by tissue-specific multigene families whose members have diverged with respect to structure and regulation (1–9). The nucleotide sequences of several PRP mRNAs from rat (4, 5), mouse (6), and human (7) and the structure and organization of complete PRP genes from the mouse (7), hamster (8), and human (9) have been reported. The common evolutionary origin of these genes is evident from the extensive conservation of 5′-untranslated regions, coding sequences, and exon/exon structures. It has been proposed (7) that the 42-nucleotide repeat unit CCA CCA CCA GGA GGC CCA CAG CCG AGA CCC CCT CAA GGC is the ancestral unit. During gene duplication multiplets of three bases were likely recruited into or deleted from this ancestral unit, and gene conversion homogenized the divergence between the internal repeats (7).

The mouse PRP genes are clustered on chromosome 8 (2). The unusual strain differences of PRP mRNAs in parotid glands of isoproterenol-treated mice have been reported (3). Whatever the mechanism is, interspersed repeated DNA elements may have an effect on gene expression and evolution. One major class of long interspersed repeated DNA (LINE) (>10^4 copies/genome) in the mouse genome is LIMd formerly known as BamH1 (10) and MIF-I (11–14). The size of LIMd has been estimated to be as large as 7 kb, but most members are truncated at apparently random distances from a common 3′ end (14). The 3′ end of individual LIMd elements contains an adenine-rich tail. This, coupled with the observation that individual LIMd elements are flanked by small, less than 15-base pair direct repeats, suggests that each element is generated via an RNA intermediate which is subsequently dispersed to distant locations (13, 14).

In this report, we demonstrate that the EcoRI restriction enzyme site differences in two members of the mouse PRP gene family (7) are due to the presence of a highly repeated DNA element LIMd. We have isolated PRP genes from cosmids and phage libraries of CD-1 mouse genomic DNA. Restriction site mapping of a 77-kb region revealed the tandem alignment of two nonallelic PRP genes (MP2 and M14) in the same 5′ to 3′ orientation. Nucleotide sequence analysis of M14 suggests that MP2 and M14 arose from a common ancestor, with the following major differences; an insertion into intron I of a 2 kb LIMd element in the opposite orientation of the M14 gene (3′ → 5′) and M14 contains 17 tandemly arrayed 42-nucleotide repeats versus 13 in MP2. The location and the nature of the insertion site for the LINE element and the implications for genomic evolution are discussed.

EXPERIMENTAL PROCEDURES

Materials—The following substances were purchased from the respective companies and were used according to the specifications of the suppliers: restriction enzymes, DNA polymerase I (Klenow fragment), T4 DNA ligase, T4 polynucleotide kinase, and calf intestine phosphatase were from Bethesda Research Laboratories, Du Pont-New England Biolabs, and Boehringer Mannheim; [α-32P]dXTP (X = A, G, C), [α-32P]dATP, [α-32P]dGTP were from Amersham Corp.; oligo-labeling kit was from Pharmacia LKB Biotechnology Inc., NYTRAN membrane, and Elutip™-d columns were from Schleicher and Schuell; agarose, low-melting agarose, 5-bromo-4-chloro-3-indoly-β-D-galactoside, isopropyl-1-thio-D-galactoside, and Bal31 nuclease were from Bethesda Research Laboratories; CD-1 mice, Charles River Breeding Labs; lambda (λ) in vitro packaging kits were from Amersham Corp. and Stratagene. Specific DNA probes

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

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J03891.

1 The abbreviations used are: PRP, proline-rich protein; MP2, M14, mouse proline-rich protein genes; bp, base pair; kb, kilobase pairs, CRPs, contiguous identical repeat peptides.
from PRP gene MP2 (5', exon II, 3', etc.) were prepared as described earlier (7).

**Screening of Mouse Genomic Libraries**—Liver DNA was isolated from CD-1 mice and was used to construct cosmids libraries and a λ EMBL-3 library according to established procedures (15). For the construction of the λ DNA, the genomic DNA was partially digested with Sau3AI to a length of 30-50 kb and size-fractionated by centrifugation through a 1.25-5 M NaCl gradient. The preparation of vector (pT7CF) arms, ligation, and packaging were carried out according to Grossveld et al. (16). The infectious bacteriophage particles were used to transduce *Escherichia coli* ED8767, with an efficiency of about 320,000 transformants/µg of size-fractionated genomic DNA. Approximately 320,000 recombinant colonies were grown on nitrocellulose filters (10,000 colonies/82-mm filter) and were hybridized with 32P-labeled 5' and exon II probes of mouse PRP gene MP2 (7) according to the screening procedure of Hanahan and Meselson (17). DNA was isolated from positive colonies by alkaline lysis (18). Approximately 200,000 recombinant phages were screened and 13 positive clones were purified and characterized as we reported previously (7, 8).

**Characterization of Cosmid Clones**—Complete restriction maps were derived by a five-step procedure. The vector (pT7CF) contains two SalI sites flanking the BamHI cloning site, which facilitates mapping of the recombinant clones. First, the coding regions of the genes were mapped by Southern blotting (19) of restriction digests and hybridization with fragments of the 5' and exon II probes of MP2. Second, the terminal fragments were subcloned by recircularization of cosmid DNA after digestion with restriction enzymes HindIII or BamHI. The plasmid DNA from transformants contained the vector fragment with the origin of replication and AmpR sequences. Aliquots of the digest were removed at several intervals between 1 and 30 min and placed in tubes containing 0.5 M EDTA. Fragments contained in the aliquots were combined and subjected to electrophoresis, and the separated fragments were blotted to a NYTRAN membrane. Two specific vector probes (NruI/SalI and SalI/NruI) from the left and right side of the BamHI cloning site, respectively, were labeled and hybridized to the blots. Fourth, the complete spectrum of restriction fragments was visualized by autoradiography of gels of single and double digestion with several restriction enzymes (six-base cutters), and the digests were labeled by filling in with appropriate [-32P]-deoxynucleotides. Fifth, the exact maps were ascertained by subcloning either SalI or BamHI fragments into pUC19 for detailed restriction enzyme analysis.

**DNA Sequence Analysis**—Fragments containing the coding regions were subcloned from the phage clones. Progressive and nonrandom shortening of the inserts, from either end separately, was carried out according to the method of Maxam and Gilbert (21). Computer analyses were performed on an IBM PC/XT computer using Pustell sequence analysis programs (22).

**RESULTS**

**Isolation and Characterization of the Mouse PRP Gene M14**—Southern blot hybridization analysis of mouse genomic DNA had shown that two EcoRI fragments (3.1 and 8.6 kb) hybridized to the 5'-specific probe of MP2 (7). Based on restriction enzyme mapping analysis, mouse PRP gene MP2 was represented by the 8.6-kb EcoRI fragment (Fig. 1). The initial purpose of this study was to identify the mouse PRP clone(s) representing the 3.1-kb EcoRI fragment. Two plasmid clones, M14 and M56, were isolated from an unamplified phage (EMBL3) genomic library. Both clones hybridized to 5' and exon II probes, whereas only clone M14 hybridized to the 3'-specific probe (data not shown). The 3.1-kb EcoRI fragment that hybridized to the 5'-specific probe was detected in both clones. Restriction enzyme analysis showed that M14 contained an 18.4-kb insert and M56 contained a 10.5-kb insert. The 5' part of M14 is identical to M56 with a 2.1-kb extension on the 5' region (Fig. 1). The 3.1-kb EcoRI fragment (and the 3.8-kb EcoRI/SalI fragment of clone M14 (Fig. 1, fragment d and part of fragment e) which contained the entire PRP gene M14 was each subcloned into pUC19 for sequencing. By using the nonrandom subcloning strategy (20), we determined 7306 base pairs of the M14 gene and flanking regions 6914 base pairs starting from the 5'-end of the EcoRI site of fragment d to the SalI site in fragment e plus 392 base pairs downstream from the SalI site of fragment e (Figs. 1 and 2).

**FIG. 1. Linkage of mouse PRP genes MP2 and M14.** Linkage of genes MP2 and M14 was demonstrated by mapping a series of overlapping genomic clones designated M2, M14, and M56 for phage clones and MC16 and MC22 for cosmids clones. The organization of genes MP2 and M14 is shown by the expanded scales and the relative lengths by the 1-kb bar. Open boxes (on the GENES line) indicate the phage clones used for sequence analysis and detailed comparisons. Solid bars show the three exonic regions, and cross-hatched boxes are the simple repetitive sequences in intron I. The large open arrow in intron I of M14 represents the 2-kb LINE insert. The B1 probe (28) (mouse Alu-equivalent) hybridized to EcoRI fragments g (MP2) and f (M14) as indicated by the boxes below the GENES line. Probes A-E correspond to the following gene regions: A, 5'-upstream of MP2; B, 5' of MP2; C, exon II of MP2; D, simple repetitive sequence of M14; E, LINE element of M14 (bp 933-2994) (see Fig. 5). Arrows show direction of transcription.

**The Structure of M14**—The location of M14 coding and noncoding sequences in Fig. 2 indicates that M14 is arranged in the exon/intron configuration characteristic of mammalian PRP genes (7-9), similar to MP2 (7), and that no intron is located between repeats (7-9). M14 has three exons (amino acids -15 to 6 and 7 to 312, and the 3'-untranslated region) (Figs. 1 and 2). Both introns are flanked by consensus splice junctions of RNA transcripts synthesized by RNA polymerase II (23). However, relative to intron I of MP2, which is 1433 base pairs in length, intron I of the M14 gene is 3784 base pairs long. Actually, intron I of M14 had two separate insertions of 223 and 2005 base pairs into intron I of MP2.

We used a homology matrix, or “dot plot” method, to compare these two sequences for investigating the relationship between MP2 and M14. A clear homology between the two sequences is revealed using a minimum stringency of 28 out of 40 nucleotides (Fig. 3). This plot shows virtually no spurious background. The displacememt of the diagonal line indicates (i) insertions of 223 and 2005 nucleotides in the M14 structure, (ii) fractional sequence differences on the simple repetitive sequences (CA, TA, TAGA) between M14 and MP2 genes, and (iii) four more 42-nucleotide internal repeats in M14 than in MP2. The dot plot of Fig. 3 serves as a basis for aligning the two sequences (data not shown). Excluding the insertions, the sequences and the four

---

2 The 223-nucleotide segment in intron I of M14 may be either the result of an insertion into MP2 or a deletion from M14 to form MP2. The authors are assuming that this 223-nucleotide segment is an insertion, and it is treated accordingly in the paper.
Fig. 2. Nucleotide sequence of mouse PRP gene M14 and flanking sequences. Sequences for three exons, the adjoining introns, and the flanking regions are shown. The proposed TA-TAA box, the CAAT box, and two polyadenylation signals are underlined. The putative mRNA cap site, based mainly on the sequence analysis and relationship to MP2, is denoted +1. The derived amino acids are numbered below starting with the proposed amino-terminal Ala (+1). Residues 65-78 are underlined to represent a typical 14-amino acid repeat. Imperfect repeats generated by the LIMd-PRP insertion (10 bp) are boxed. The simple repetitive sequences are in parentheses.

extra internal repeats in the exon II sequences of the two genes including the introns are nearly identical. Percentage divergences for the various gene regions are summarized in Table I. The remarkable identity indicates that M14 and MP2 have a common ancestor. There are two typical poly(A) addition signals, AATAAA (24), separated by 662 base pairs (Fig. 2). Whether the second signal is used or whether any differential regulation is involved is not known.

 Interruption of Intron I by Insertion of a LIMd-PRP Element—To investigate the nature of the large inserted element (2005 bp), a SpeII/AuII fragment of M14 (bp 933-2994) was used to probe a CD-1 mouse genomic total DNA blot (Fig. 1, probe E). The characteristics of the repetitive DNA and the pattern of hybridization to various restriction enzyme digests of mouse total DNA suggested that this inserted DNA was a member of the mouse LINE family (LIMd). Comparison of this inserted sequence and the LIMd-A2 of BALB/c mouse (25) clearly showed that the inserted sequence was the 3' portion of the mouse LINE element and that it had been transposed into intron I in the opposite direction of M14. The LIMd-PRP, like most mouse LINE elements, is truncated at the 5' end (12, 13). It contains the typical polyadenylation signal AATAAA (24) and an adenine-rich element, and it is flanked by a pair of 10-bp imperfect direct repeats (TGTCTTTTTT) (Figs. 2 and 4). We subsequently aligned the nucleotide sequences bordering LIMd-PRP with their homologous sequences in MP2 (Fig. 4). The 5' boundary of LIMd-PRP was confirmed by the existence of 10-bp imperfect direct repeats at both the putative 5' boundary and adjacent to the 3'-poly(A) tract. This 10-bp sequence is present only once in the MP2 target region (Fig. 4). The presence of this direct repeat, apparently generated by duplication of the single target sequence, is consistent with the hypothesis that LIMd-PRP entered the PRP locus via transposition. The 10-bp direct repeat is imperfect due to the presence of one nucleotide substitution (T/C), which was presumably introduced during or after the transposition event. However, each of the direct repeats differs from the MP2 target sequence by one base substitution. Another sequence feature is the presence of 17-18 bases outside either border of the putative target site of
transposition that show dyad symmetry (Fig. 4).
Comparison of the other insertion sequence of 223 nucleotides with respect to either the LIMd consensus sequence or to the mouse Alu-equivalent B1 and B2 elements (26, 27) showed no evidence of sequence homology. Although this segment appears to be an insertion, it does not contain features typical of insertion elements. Inverted repeats are not present, and there are no target site-derived direct repeats.

**Linkage of MP2 and M14**—To link MP2 and the duplicated gene M14, two cosmids libraries were screened directly after construction. Positive clones were mapped with several restriction enzymes and were compared with the restriction maps of existing clones. Two specific cosmids clones, MC16 and MC22, were selected for further characterization. A compiled map of these PRP genes is shown in Fig. 1. In total, 77 kb were cloned from mouse genomic libraries. The positions of the exons and the orientations of the two genes have been determined by blot analysis of cosmids clones (MC16, MC22) and phage clone (M56) using various probes (Figs. 1 and 5). MP2 is located 29 kb upstream from M14, and they are tandemly arrayed.

Fig. 5 shows Southern blots of two overlapping PRP cDNA clones and one phage clone digested with EcoRI and hybridized with 5' (probe B) and exon II (probe C) sequences of MP2. Both the 8.6- and 3.1-kb fragments (Fig. 1, fragments b and d), which were detected in genomic DNA by using the 5' probe, were observed in MC22 (lane 2, probe B). MC16 contained only the 8.6-kb EcoRI fragment (lane 3, probe B). This suggested that MC22 contained portions of both PRP genes MP2 and M14. When the blots were hybridized to exon II, MC22 contained the EcoRI fragment of 8.6 kb (Fig. 1, fragment b) as predicted by the total genomic blot and a 7.5-kb fragment (Fig. 5, lane 2, probe C). The 7.5-kb fragment represented a truncated EcoRI sequence (Fig. 1, fragment e) plus part of the pTCE vector (SalI-EcoRI). This analysis suggested that MC22 as well as M56 contained a 5'-truncated M14 gene.

To show that these two clones were overlapping and that no recombination had occurred during the cloning, the blot was screened with a terminal insert fragment from MC22 (probe A). Terminal fragments were isolated by recircularization of the cosmid DNA after digestion with BamHI which does not cleave the vector (see "Experimental Procedures"). An 8.5-kb EcoRI fragment from MC16 hybridized to this terminal probe (Fig. 1, fragment a, and Fig. 5, lane 2, probe A) as well as truncated EcoRI fragment a of MC22 plus part of the vector (Fig. 5, lane 3, probe A). Interestingly, a 5.5-kb EcoRI fragment from MC22 and M56 also showed hybridization (Fig. 5, lanes 1 and 2, probe A, and fragment c in Fig. 1). This terminal fragment did not contain repetitive elements which suggested that the 5' upstream regions of M14 and MP2 shared sequence homology. Both MP2 and M14 also had two HindIII sites separated by 0.5 kb and about 3.5 kb upstream from the EcoRI sites (Fig. 1). This is in contrast to other multigene families, such as the mouse amylase gene families (28), where restriction enzyme sites in the flanking regions are not conserved. Blots were screened with either LIMd-PRP, alternating purine and pyrimidine sequences (simple repetitive sequences) (Fig. 1), lanes 1 and 2, probe A) and the Alu-equivalent B1 probe of mouse (29) (Fig. 5). Only one copy of LIMd was observed in this 77-kb DNA fragment (Fig. 5, probe E), and no other stretches of simple repetitive se-

**FIG. 3.** Homology matrix comparison of PRP genes MP2 and M14. Strangency of comparison as the ratio of required matches to the length of comparison window is 28/40. A scale diagram showing the length of each gene is included for reference. For M14, the inserted LIMd-PRP was not plotted. Exons and introns 1 are located in the 3'-untranslated region of each gene, is not labeled.

**TABLE I**

| Region             | Mutations | % divergence |
|-------------------|-----------|--------------|
| 5'-Flanking region| 39/703    | 5.5          |
| 5'-Untranslated region | 1/33     | 3.0          |
| Exon I            | 1/97      | 1.0          |
| Intron I          | 126/1320  | 9.5          |
| Exon II           | 33/740    | 4.4          |
| Intron II         | 32/389    | 8.2          |
| Exon III          | 8/152     | 5.2          |
| 3'-Untranslated region | 3/115    | 2.6          |
| All exons         | 42/989    | 4.3          |
| All introns       | 138/1709  | 9.2          |
| All genic regions | 259/3401  | 7.0          |

**Fig. 4.** Comparisons of the 5' and 3' sequences of LIMd-PRP in M14 with the putative target site in MP2. The 10-bp imperfect repeats which flank LIMd-PRP (M14) and the 16-bp putative target site for insertion (MP2) are boxed. A dyad symmetry (18 bp) which flanks the target site of MP2 is overlined (L1-PRP(L)) and underlined (L1-PRP(R)). Identical sequences are indicated by dots. L and R denote the direction of transcription of the PRP gene (L — R), which is opposite to that of the LINE element.
hybridized to the 5′-upstream region subcloned from MC14 probe hybridized to two distinct EcoRI fragments, MC22 in Fig. 1 and mouse B1 probe. Protein-coding exons. Exon I corresponds generally to the introns of PRP genes in different species are considerably varied in size, and the peptide sequences differ in various family members. Some observations and speculations are given with respect to the transition and carboxyl terminus segments flanking the highly conserved repeats. These regions are variable in size, and the peptide sequences differ in various family members of the same species and of different species (1). However, these sequences are still relatively rich in the amino acids Pro, Glu + Gin, Gly, and Asp + Asn, which is a characteristic of the PRPs. The entire mature protein encoded by the PRP gene could have originated from the same simple oligonucleotide series which encodes the repeat, as proposed before (7, 8). Since there are no introns in this coding sequence, numerous duplications likely gave rise to both flanking (transition and carboxyl-terminal regions) and repeating segments over a rather short length of DNA. It is remarkable that the mechanism of duplication has preserved, with one exception, the exactness of the 42-nucleotide repeat in the highly conserved repeat region (7). Mutations likely occurred in the flanking regions which promoted variations in the transition region and the carboxyl terminus. A proposed role for introns is to limit amplification in genes whose protein products cannot tolerate variation in size (32). If so, the presence of introns between each repeat could have inhibited the development of new PRP genes. Ohno (33) has proposed that oligonucleotide repeats are the primordial source of all genes. It seems plausible that genes created recently would be the most likely to retain evidence of primordial repeats.

Recently, Heinrich and Habener (34) and Mirels et al. (35) reported on a multigene family encoding proteins containing glutamine/glutamic acid-rich secretory proteins or contiguous identical repeat polypeptides (CRPs) from rat submandibular gland. The CRPs are related to the PRPs, but they contain high amounts of Gln plus Glu (35%) and only about 10% proline. Other interesting similarities between the PRPs and the CRPs are that there are only two introns and three exons in the CRPs with essentially the same gene organization as shown for the mouse PRP genes (Fig. 6), and exons I of these two families (the signal peptide plus 5′-untranslated region) are highly conserved (7, 8, 34). These observations suggest two possibilities; the signal sequences encoded by exon I may be critical for secretion of these salivary gland products or there were recombinational events in the genesis of these multigene families. In other words, the signal peptide domains and 5′-untranslated regions were added to members of these gene families after duplication and evolution of primordial repeats.

One model for the evolution of the PRP genes is that 42-bp primordial oligonucleotides were duplicated and then they diverged in the flanking regions to form transcriptional units. These transcriptional units evolved independently and were subject to differential rates of change. The signal peptide domain may have been added to members of the gene family at this stage or later. Gene conversion could account for the homologies. The observation that gene conversion events in higher eukaryotic organisms often appears to occur within multigene families on the same chromosome (36) supports this contention. Unequal crossing over and gene conversion can result in the concerted evolution of gene families. In situations such as this, standard numerical techniques for

---

**Fig. 5. Southern blot analysis of overlapping cosmid and phage clones.** PRP phage clone M56 (lane 1) and cosmid clones MC22 (lane 2) and MC16 (lane 3) were digested with EcoRI and hybridized to the 5′-upstream region (probe A), the 5′-probe of MP2 (probe B), exon II of MP2 (probe C), the simple repetitive sequence subcloned from MC14 (probe D), L1M-PRP (probe E) as indicated in Fig. 1 and mouse B1 probe.

---

**Fig. 6. Comparison of PRP gene organizations in mouse, hamster, and human.** Related exonic regions (bars) are connected by dashed lines. The gene organizations of hamster H29 (8) and human PRH1 (9) show an extra 36-bp exonic sequence (EIIa).
inferring molecular phylogenies or estimating evolutionary distances cannot be directly applied. Further insight into the evolution of the mammalian PRP genes requires characterization of homologs in other species.

Restriction site differences at the mouse PRP gene loci MP2 and M14 are due to the presence of a 2.0-kb element. DNA sequence determination showed that the 2.0-kb element is a member of the highly repeated mouse LINE family. Four different strains of mice (DBA/2J, CD-1, C57BL/6J, and MP2 and M14) are due to the presence of a 2.0-kb element. The lower portion of the figure shows a diagrammatic representation of the relative organization of elements of intron I of PRP genes MP2 and M14. A and B denote the simple repetitive sequences. Inserts of a 223-bp (unknown element) and of the 2005-bp inserted elements are indicated by the horizontally striped boxes are present in M14, but not MP2. Exons are indicated by cross-hatched boxes, introns are indicated by two lines, and they are not drawn to scale.

In this report, the LINE element is inserted into the PRP gene in the opposite orientation. This insertion may affect isoproterenol inducibility of PRP gene M14. Hutchinson et al. (41) have proposed that gene duplications are the major source of genomic growth and that genome growth is a consequence of a higher rate of fixation than deletion during evolution. The altered locus could then be subject to a number of genetic effects, such as recombination of the repeated DNA sequences, to produce the plasticity of the genome. The mouse albumin and α-fetoprotein genes (42) and two chicken γ-crystallin genes (43) are present in the same 5′ to 3′ orientation. Both these genes are expressed in the same tissue, but they have diverged with respect to developmental regulation, and one gene of each pair is considerably more active. As these examples indicate, tandem gene duplication followed by divergence of various aspects can produce independently regulated genes in close physical proximity. This report represents a first step in characterizing the chromosomal organization of the mouse PRP multigene family.

Acknowledgments—We thank Drs. Jerry B. Lingrell and David D. Chaplin for providing mouse Alu-equivalent B1 probe (p158EH1E10) and cosmID pICF vector and Donna Gadbois for technical assistance.

REFERENCES

1. Carlson, D. M., Ann, D. K., and Mehansho, H. (1986) Microbiology (Wash., D.C.) 303–306
2. Azen, E. A., Carlson, D. M., Clements, S., Lalley, P. A., and Vanin, E. (1984) Science 226, 967–968
3. Ann, D. K., Clements, S., Johnstone, E., and Carlson, D. M. (1985) J. Biol. Chem. 260, 899–904
4. Ziemer, M. A., Swain, W. F., Rutter, W. J., Clements, S., Ann, D. K., and Carlson, D. M. (1984) J. Biol. Chem. 259, 10475–10480
5. Clements, S., Mehansho, H., and Carlson, D. M. (1985) J. Biol. Chem. 260, 13471–13477
6. Maeda, N., Kim, H-S., Azen, E. A., and Smithies, O. (1985) J. Biol. Chem. 260, 11123–11130
7. Ann, D. K., and Carlson, D. M. (1985) J. Biol. Chem. 262, 15863–15872
8. Ann, D. K., Gadbois, D., and Carlson, D. M. (1987) J. Biol. Chem. 262, 3958–3963
9. Kim, H-S., and Maeda, N. (1986) J. Biol. Chem. 261, 6712–6718
10. Soriano, P., Meunter-Rotival, M., and Bernardi, G. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1816–1820
11. Brown, S. D. M., and Dover, G. (1981) J. Mol. Biol. 150, 441–456
Proline-rich Protein Multigene Families

12. Fanning, T. G. (1983) *Nucleic Acids Res.* 11, 5073-5091
13. Voliva, C. F., Jahn, C. L., Comer, M. B., Edgell, M. H., and Hutchison, C. A., III (1983) *Nucleic Acids Res.* 11, 8847-8859
14. Singer, M. F., and Skowronski, J. (1985) *Trends Biochem. Sci.* 10, 119-122
15. Manisatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Grosfeld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M., and Flavell, R. A. (1982) *Nucleic Acids Res.* 10, 6715-6732
17. Hanahan, D., and Meselson, M. (1980) *Methods Enzymol.* 100, 333-342
18. Birnboim, H. C. (1983) *Methods Enzymol.* 100, 243-255
19. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503-517
20. Guo, L., and Wu, R. (1983) *Methods Enzymol.* 100, 60-96
21. Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560
22. Pustell, J., and Kafatos, F. C. (1984) *Nucleic Acids Res.* 12, 643-655
23. Mount, S. M. (1982) *Nucleic Acids Res.* 10, 459-472
24. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Biochem.* 50, 349-383
25. Loeb, D. D., Padgett, R. W., Hardies, S. C., Steheee, W. R., Comer, M. B., Edgell, M. H., and Hutchison, C. A., III (1986) *Mol. Cell. Biol.* 6, 168-182
26. Coggins, L. W., Vass, J. K., Stinson, M. A., Lanyon, W. G., and Paul, J. (1982) *Gene (Amst.*) 17, 113-116
27. Kravt, A. S., Markushova, T. V., Kramerov, D. A., Ryskov, A. P., Skryabin, K. G., Baye, A. A., and Georgiev, G. P. (1982) *Nucleic Acids Res.* 10, 7461-7475
28. Wiebauer, K., Gumucio, D. L., Jones, J. M., Caldwell, R. M., Hartle, H. T., and Meisler, M. H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* 82, 5446-5449
29. King, D., Snider, L. D., and Lingrel, J. B. (1986) *Mol. Cell. Biol.* 6, 209-217
30. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987) *Science* 235, 1516-1522
31. Jeffreys, A. J., Wilson, V., and Thein, S. L. (1985) *Nature* 314, 67-73
32. Alexander, F., Yound, P. R., and Tilghman, S. M. (1984) *J. Mol. Biol.* 173, 159-176
33. Ohno, S. (1984) *J. Mol. Evol.* 20, 313-321
34. Heinrich, G., and Habener, J. F. (1987) *J. Biol. Chem.* 262, 5262-5270
35. Mirels, L., Bedi, G. S., Dickinson, D. P., Gross, K. W., and Tabak, L. A. (1987) *J. Biol. Chem.* 262, 7289-7297
36. Fink, G. R., and Pets, T. D. (1984) *Nature* 310, 728-729
37. Purano, A. V., Somerville, C. C., Tsichlis, P. N., and D’Ambrosio, E. (1986) *Nucleic Acids Res.* 14, 3717-3727
38. Tautz, D., and Renz, M. (1984) *Nucleic Acids Res.* 12, 4127-4138
39. Hanford, D. B., and Pulleyblank, D. E. (1985) *Nucleic Acids Res.* 13, 4343-4363
40. Kmiec, E. B., Angelides, K. J., and Holloman, W. K. (1985) *Cell* 40, 139-145
41. Hutchison, C. A., III, Hardies, S. C., Padgett, R. W., Weaver, S., and Edgell, M. H. (1984) *J. Biol. Chem.* 259, 12881-12889
42. Ingram, R. S., Scott, R. W., and Tilghman, S. M. (1981) *Proc. Natl. Acad. Sci. U. S. A.* 78, 4694-4698
43. Hawkins, J. W., Nickerson, J. M., Sullivan, M. A., and Pintigor-sk, J. (1984) *J. Biol. Chem.* 259, 9821-9825