Vol. 138, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 16, 1986 Pages 289-296

CLOSE LINKAGE OF HUMAN CHROMOSOMAL PEPSINOGEN A GENES

Toshiya Hayano[1,2], Kazuhiro Sogawa[1], Yoshikazu Ichihara[1,2], Yoshiaki Fujii-Kuriyama[1], and Kenji Takahashi[2]

[1]Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan

[2]Department of Biophysics and Biochemistry, Faculty of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Received May 15, 1986

SUMMARY: We have obtained a clone containing two pepsinogen A genes in a single insert by screening a recombinant cosmid library for human genomic DNA. Restriction endonuclease mappings of this cloned DNA showed that these two genes are very similar, but distinct in structure, and that they are closely linked to one another in the human chromosome DNA. The close arrangement of the genes with very similar structures could facilitate the homologous recombination or the unequal crossing-over which accounts for high frequency of haplotype variation in copy number of pepsinogen A genes as reported by Taggart et al.

Pepsinogen, the inactive precursor of pepsin which belongs to the group of aspartic proteinase, is synthesized in chief cells of gastric glands and is secreted into gastric lumen. It is demonstrated that seven electrophoretically separate pepsinogens exist in human gastric mucosa and that they are classified into two groups (1). One group of the zymogens, designated pepsinogen A (PGA), consists of Pg1-Pg5 which are characterized by their electrophoretically faster migration and their localization in the fundus and the body of the stomach, while the other group, termed pepsinogen C (PGC), is composed of Pg6 and Pg7 which are localized in the whole stomach.

Previous studies suggested that the heterogeneity of pepsinogens was derived from their multiple genes as well as post-translational modifications (2) and that the polymorphism of

0006-291X/86 $1.50

the individual isozymogens is derived from multigenic or allelic variation (3). These suggestions have recently been verified in part by the study of Taggart et al.(4) indicating that at least three PGA genes are located in their different combination in the centromeric region of individual human chromosome 11. On the basis of these observations, they explained quantitative and qualitative differences of individual isozymogen patterns.

In the present study, we screened a cosmid human genomic library in order to investigate how these multiple PGA genes are arranged in chromosomal DNA. Our results revealed that at least two PGA genes, with very similar structures, are closely linked to each other. These data strongly suggest that variant PGA genes are clustering in the limited region of human chromosomal DNA and that different haplotypes of PGA genes in individuals are products of homologous recombination or unequal crossing-over facilitated by their close arrangement.

MATERIALS AND METHODS

Materials: Restriction enzymes, T4 DNA ligase, and alkaline phosphatase were purchased from Takara Shuzo Co. (Kyoto, Japan) or Bethesda Research Laborarory (Gaithersburg, MD). {α-^{32}P}dCTP ($\geq$ 3000 Ci/mmol) was from the Radiochemical Centre, Amersham, England.

Preparation and screening of the recombinant cosmid library: Human genomic fragments generated by partial Sau3AI digestion were cloned into the BamHI site of cosmid vector, pHC79, as described previously (5). The cosmid library was transfected into Escherichia coli strain, ED8767, by in vitro packaging method and screened for human PGA genes with the nick-translated rat PGC cDNA, pRPC1 (Ichihara et al., to be published), as a probe. The hybridization was performed overnight at 65°C in 50mM Tris-HCl buffer, pH 7.5, containing 1M NaCl, 10mM EDTA, 0.1% SDSa (sodium dodecyl sarcosinate), 0.2% polyvinylpyrrolidone, 0.2% Ficol and 0.2% bovine serum albumin. Following the hybridization, the filters were washed in 2XSSC (1XSSC is 0.15M NaCl, 0.015M sodium citrate, pH 7.0), 0.1% SDSa at 65°C. Autoradiography was carried out at -70°C in the presence of an intensifying screen.

Restriction endonuclease analysis and Southern blot analysis: High-molecular-weight nuclear DNA from human placenta was isolated as previously described (6). λDNA and cosmid DNA were purified as described previously (7-9). High-molecular-weight DNA was digested with restriction endonucleases according to the procedures recommended by the vendors. The digests were subjected to electrophoresis on 0.8% agarose gel and transferred to

nitrocellulose filters as described (10). The filters were probed with a nick-translated human PGA cDNA, pTM003g-35. The hybridization was performed as described above. After the hybridization, the filters were washed in 0.1XSSC, 0.1% SDSa at 65°C.

RESULTS

Approximately 200,000 recombinant cosmid clones for human genomic DNA fragments were screened at low stringency as described in MATERIALS AND METHODS, probed with the radiolabeled rat PGC cDNA, pRPC1. Thus we obtained eight positive clones. Judging from the strength of their hybridization to human PGA cDNA or rat PGC cDNA, we classified six of them into PGA clones and the rest into PGC ones.

Figure 1 shows restriction and hybridization analysis for one of the PGA clones, designated cosHPGA12-1, and λHPGA1 isolated previously by Sogawa et al. (11). As can be seen, there are common restriction fragments between them ; that is, 3.4 kb EcoRI fragment, 2.9 kb and 1.9 kb BamHI ones, 12.4 kb HindIII one and 7.1 kb XhoI one. Further, for each digestion, cosHPGA12-1

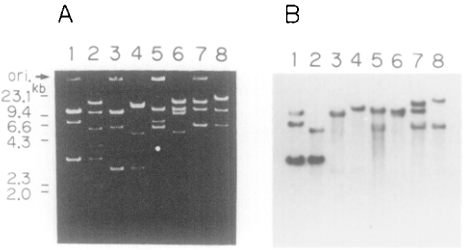

Fig.1 Restriction endonuclease and Southern blot analysis of cosHPGA12-1 and λHPGA1

Cosmid DNA (cosHPGA12-1) and λDNA (λHPGA1) digested with restriction endonucleases were subjected to electrophoresis on a 0.8% agarose gel and they were transferred to a nitrocellulose filter, followed by hybridization to the ^{32}P-labeled human PGA cDNA as described in MATERIALS AND METHODS. A. A photograph of the ethidium bromide-stained agarose gel. Lanes 1, 3, 5, and 7, cosHPGA12-1 DNA digested with EcoRI, BamHI, HindIII and XhoI, respectively. Lanes 2, 4, 6, and 8, λHPGA1 DNA digested with EcoRI, BamHI, HindIII and XhoI, respectively. B. An autoradiogram of the filter hybridized to the probe. Lanes 1, 3, 5, and 7, cosHPGA12-1 DNA digested with restriction endonucleases described above, respectively. Lanes 2, 4, 6, and 8, λHPGA1 DNA digested with restriction endonucleases above, respectively. The origin and the size markers are indicated.

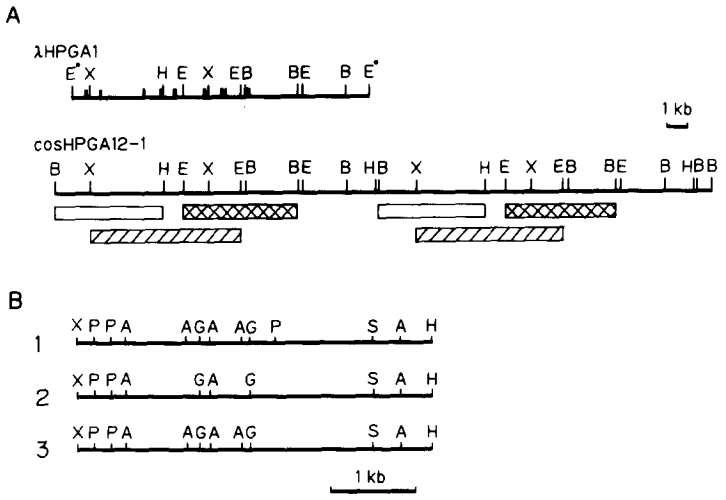

Fig.2 Restriction endonuclease maps of λHPGA1 and cosHPGA12-1

A, Restriction endonuclease maps of λHPGA1 (upper) and cosHPGA12-1 (lower). Closed blocks show the exons of human PGA gene and artificial EcoRI sites (E*) are indicated in the map of λHPGA1. Open boxes, shaded ones and cross-hatched ones indicate the fragments hybridized to 5'-, middle, and 3'-portions of the rat PGC cDNA, respectively, in the map of cosHPGA12-1. B, Restriction endonuclease maps of 4.2 kb XhoI/HindIII fragments which were independently isolated from 5'-portion of the three PGA genes (the one isolated by Sogawa et al. and two putative ones contained in cosHPGA12-1). 1, 2, and 3 show the restriction maps of the 4.2 kb XhoI/HindIII fragments isolated from λHPGA1, the upstream gene, and the downstream one in the cosHPGA12-1, respectively. E, EcoRI; B, BamHI; H, HindIII; X, XhoI; P, PvuII; A, AvaII; G, BglII; S, SstI.

contains additional other restriction fragments hybridized to the probe than those found with λHPGA1 (Fig.1). We constructed precise restriction cleavage maps of this insert fragment by single and double digestions with the restriction enzymes described above, followed by hybridization to the 5'-, middle and 3'- portions of the rat PGC cDNA, pRPC1. These results are summarized in Fig.2-A, together with the structure of one of the PGA genes previously determined by our group (11). From comparison with the previously reported structure of a PGA gene, it is clearly demonstrated that cosHPGA12-1 contains two PGA genes which are tandemly arranged in a short stretch of chromosomal DNA. As far as cleavage maps of the four restriction enzymes are concerned, no difference was observed between these

two tandemly arranged genes. This close homology was also extended to the structure of a PGA gene reported by Sogawa et al.

In order to clarify the identity among these three PGA genes, 4.2 kb XhoI/HindIII fragments which are commonly present in these genes were separately isolated and further analyzed with four other restriction endonucleases, AvaII, PvuII, BglII, and SstI. As can be seen in Fig.2-B, any one of the three DNA fragments can be distinguished from the two others by digestion with AvaII and PvuII, although BglII and SstI yielded the same digests from the three DNA fragments. Taken together, these results indicate that these three PGA genes are very similar in structure but not identical with one another.

To detect genomic DNA fragments carrying PGA gene sequence, high-molecular-weight DNA from human placenta digested with some restriction enzymes was subjected to Southern blot analysis with radiolabeled human PGA cDNA, pTM003g-35, as a probe (Fig.3).

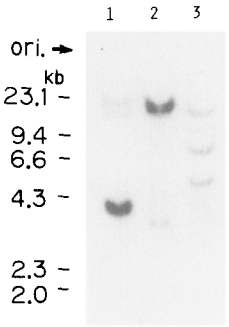

Fig.3 Southern blot analysis of restriction endonuclease digested high-molecular-weight DNA

10 μg of high-molecular-weight DNA isolated from human placenta was digested with EcoRI, BamHI, and HindIII, respectively, electrophoresed, blotted and hybridized to ^{32}P-labeled human PGA cDNA as described in MATERIALS AND METHODS. The origin and the size markers of the electrophoresis are indicated. Lane 1, EcoRI digested DNA; 2, BamHI one; 3, HindIII one.

The following genomic DNA fragments which hybridized to the probe, were included in cosHPGA12-1. (Fig.1-A and 2); 12.2 kb and 3.4 kb EcoRI fragments, 11.5 kb and 2.9 kb BamHI ones and 12.4 kb and 6.6 kb HindIII ones. Other fragments detected, namely 16 kb EcoRI one, 8.6 kb and 4.6 kb HindIII ones allow us to suspect the existence of other PGA-related DNA sequence which failed to be obtained by the screening. It could be that all the PGA genes or PGA gene-related sequences are clustered in the limited region of the chromosomal DNA. Precise chromosomal arrangement of all these DNA sequences will be a next target to be elucidated.

DISCUSSION

This study presents direct evidence for close linkage of the two human PGA genes within a single haplotype. Of three or more PGA genes that could be detected by genomic Southern blot analysis, two were arranged in a head-to-tail orientation, separated by approximately 16 kb of intergenic DNA. Detailed comparison of the restriction maps of these genes and the one isolated by Sogawa et al. reveals that they are structurally very similar but distinct from one another.

Taggart et al. (4) and Banke et al. (12) reported that human PGA genes are located in the centromeric region of human chromosome 11. Taggart et al. (4) also reported that differences in the copy number of three PGA genes are the primary determinant of PGA polymorphism that occurs at high frequency in the United States. The close association of the genes with very similar structures could facilitate the homologous recombination or the unequal crossing-over, resulting in the high frequency of haplotype variation of PGA genes.

The examples of the linkage have already been reported with other eukaryotic genes (13-15). They have been characterized by

their developmental or tissue specific regulations of gene expression. In the case of the pepsinogen family, prochymosin is known as a neonatal form. It is rarely observed except for in bovine. However, recent studies reported its existence in some other mammals (16,17), allowing us to presume the developmentally regulated expression of pepsinogen and its related genes. Furthermore, it is demonstrated that PGCs are found in prostate and seminal fluid while PGAs are not (18,19), suggesting some differences with tissue specificity between them. Therefore, it would be of interest to clarify how the PGA and its related genes are localized in the chromosome, in view of comparative biochemistry of multigene families such as globin, α-amylase and so on. This paper presents the first step toward the characterization of the chromosomal organization of human pepsinogen gene family.

ACKNOWLEDGMENT: We thank Dr. N. Nobuhara (Mochida Pharmaceutical Co., Tokyo) for a generous gift of the human pepsinogen A cDNA, pTM003g-35.

REFERENCES

1. Samloff,I.M. (1969) Gastroenterology 57, 659-669
2. Taggart,R.T., Karn,R.C., Merritt,A.D., Yu,P.l., and Conneally, P.M. (1979) Hum. Gent. 52, 227-238
3. Samloff,I.M., and Townes,P.L. (1975) Science 168, 144-1454
4. Taggart,R.T., Mohandas,T.K., Shows,T.B., and Bell,G.I. (1985) Proc. Natl. Acad. Sci. USA 82, 6240-6244
5. Hohn,B., and Collins,J. (1980) Gene 11, 291-298
6. Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982) "Molecular Cloning", Cold Spring Harbor Laboratory
7. Maniatis,T., Hardison,R.C., Lacy,E., Lauer,J., O'Connell,C., Quon,D., Sim,G.K., and Efstratiadis,A. (1978) Cell 15, 681-701
8. Birnboim,H.C., and Doly,J. (1979) Nucl. Acid Res. 7, 1513-1523
9. Ish-Horowicz,D., and Burke,J.F. (1981) Nucl. Acid Res. 9, 2989-2998
10. Southern,E.M. (1975) J. Mol. Biol. 98, 503-517
11. Sogawa,K., Fujii-Kuriyama,Y., Mizukami,Y., Ichihara,Y., and Takahashi,K. (1983) J. Biol. Chem. 258, 5306-5311
12. Bauke,Z., van Kessel,A.G., de Wit,J., Evers,P., Arwert,F., Pronk,J.C., Mager,W.H., Planta,R.J., Eriksson,A.W., Frants,R.R. (1985) Hum. Genet. 70, 337-340

13. Efstratiadis,A., Posakony,J.W., Maniatis,T., Lacon,R.M., O'Connell,C., Spritz,R.A., DeRiel,J.K., Forget,B.G., Weissman,S.M., Slightom,J.L., Blechl,A.E., Smithies,O., Baralle,F.E., Soulders,C.C., and Proudfoot,N.J. (1980) Cell 21, 653-668
14. Ingram,R.S., Scott,R.W., and Tilghman,S.M. (1981) Proc. Natl. Acad. Sci. USA 78, 4694-4698
15. Wiebauer,K., Gumucio,D.L., Jones,J.M., Caldwell,R.M., Hartle,H.T., and Meisler,M.H. (1985) Proc. Natl. Acad. Sci. USA 82, 5446-5449
16. Foltmann,A.L., Jensen,A.L., Lϕnblad,P., Smidt,E., and Axelsen, N.H. (1981) Comp. Biochem. Physiol. 68B, 9-13
17. Jensen,T., Axelsen,N.H., and Foltmann,B. (1982) Biochim. Biophys. Acta. 705, 249-256
18. Samloff,I.M., and Liebman,W.M. (1972) Clin. Exp. Immunol. 11, 405-414
19. Chiang,L., Contreras,L., Chiang,J., and Ward,P.H. (1981) Arch. Biochem. Biophys. 210, 14-20