Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 6240–6244, September 1985
Genetics

Variable numbers of pepsinogen genes are located in the centromeric region of human chromosome 11 and determine the high-frequency electrophoretic polymorphism

(gastric protease/gene duplication/protein polymorphism/DNA restriction fragment length polymorphism)

R. THOMAS TAGGART*, T. K. MOHANDAS†, THOMAS B. SHOWS‡, AND GRAEME I. BELL§

*Human Genetics Research (151M), Veterans Administration Medical Center, Sepulveda, CA 91343, and Department of Medicine, University of California School of Medicine, Los Angeles, CA 90073; †Division of Medical Genetics, University of California Los Angeles, UCLA-Harbor General Hospital, 1000 West Carson Street, Torrance, CA 90509; ‡Department of Human Genetics, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, NY 14263; and §Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608

Communicated by Harry Harris, June 7, 1985

ABSTRACT **A panel of 26 mouse–human somatic cell hybrids containing different human chromosome complements was analyzed with a cloned human pepsinogen cDNA probe to determine the chromosomal location and the number of genes encoding these proteins. A complex containing variable numbers of pepsinogen genes was localized to the centromeric region of human chromosome 11 (p11→q13). Examination of somatic cell hybrids containing single copies of chromosome 11 and the corresponding human parental cell lines revealed a restriction fragment length polymorphism determined by pepsinogen haplotypes that contained two or three genes, respectively. Concurrent studies of DNA from individuals exhibiting the most common pepsinogen electrophoretic phenotypes with exon-specific probes demonstrated that the absence of one gene among the different restriction fragment patterns correlated with the absence of one specific isozymogen (Pg 5). Thus, our studies demonstrate that this genetic polymorphism involving intensity variation of individual pepsinogen isozymogens results from chromosome haplotypes that contain different numbers of genes. The regional localization of this polymorphic gene complex will facilitate detailed linkage analysis of human chromosome 11.**

The human group I pepsinogens (PG I or PGA) are the inactive zymogen precursors of pepsin, which is the major acid protease activity found in the stomach (1). These electrophoretically distinguishable pepsin precursors exhibit extensive intensity variation of the individual isozymogens Pg 5, Pg 4, Pg 3, and Pg 2. Initial family studies of urinary pepsinogen phenotypes were unable to determine whether the polymorphic variation of the individual isozymogens resulted from multigenic or allelic variation (2–5). A major difficulty for the genetic and biochemical characterization of these closely related proteins has been the inability to detect a defined molecular difference among them beyond that observed by electrophoresis. However, recent studies utilized monoclonal antibodies to identify an antigenic difference among the isozymogens, to purify the three major isozymogens from gastric mucosa (Pg 3, Pg 4, and Pg 5) and to localize net ionic charge differences to the activation peptide and pepsin regions of the isozymogens. These studies suggest that the net charge amino acid differences among them result from three different *PGA* genes (6).

The genetic polymorphism of PGA, as demonstrated by electrophoretic techniques, is manifested by qualitative and quantitative differences in the pepsin activity derived from individual isozymogens. Samloff and Townes originally reported the variation of Pg 5 (2). They demonstrated that the presence of Pg 5 (phenotype A) could be explained by the dominant expression of one allele, Pg^a, of an autosomal diallelic locus. Individuals homozygous for the alternative allele, Pg^b, did not exhibit Pg 5 activity (designated phenotype B). In addition to the absence of Pg 5, variations in the intensity of the other PGA isozymogens were subsequently reported by several investigators (3–5). Recent studies of several rare PG I variants in the Dutch population have provided evidence for the existence of multiple genes for PG I (7). However, the complexity of the isozymogen patterns and the indirect nature of protein electrophoretic studies have precluded determination of the gene copy number and the organization of the PG I genes (*PGA*).

In the present report, we describe the isolation of a cDNA clone that contained a coding sequence corresponding to a portion of exon 3 through exon 8 of the nine-exon human *PGA* gene (8). We utilized this cDNA probe to determine the chromosomal location and the organization of the *PGA* genes present in human cells and human–mouse somatic cell hybrids. We present evidence that intensity variation of the individual PGAs Pg 3, Pg 4, and Pg 5 results from haplotypes containing different combinations of the genes (*PGA3*, *PGA4*, and *PGA5*) that encode each protein. Our studies of individuals exhibiting informative PGA phenotypes have identified further the gene fragments determining the Pg 5 isozymogen (*PGA5*).

MATERIALS AND METHODS

Somatic Cell Hybrids. Hybrid cell clones with the prefix 84 were isolated by the polyethylene glycol-mediated fusion of mouse B82 fibroblast cells (GM 0347A) and human fibroblast cells (IMR 91, Mutant Cell Repository, Camden, NJ) using standard procedures. The derivation and characterization of the remaining mouse–human somatic cell hybrid lines used in the study have been described (9–12). Hybrid cell clones were harvested for chromosome characterization, isozyme analysis, and DNA preparation at the same passage. Identification of the individual chromosome 11 homologs retained by the hybrid cell lines was achieved by analysis of the restriction fragment length polymorphism (RFLP) associated with the human insulin gene (*INS*) (13, 14).

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.

Abbreviations: PG I or PGA, group I pepsinogen = pepsinogen A, (E.C. 3.4.23.1, pepsin zymogen precursor); *PGA*, gene symbol for PG I or PGA; Pg 5, Pg 4, Pg 3, and Pg 2, the individual PG I isozymogens separated by electrophoresis; PG II or PGC, group II pepsinogen = pepsinogen C = gastricsinogen, (E.C. 3.4.23.3, pepsin II or gastricsin precursor); *PGC*, gene symbol for PG II or PGC; *INS*, gene symbol for insulin; bp, base pair; kb, kilobase pair; RFLP, restriction fragment length polymorphism.

cDNA Probes. A human gastric mucosa cDNA library was constructed from the poly(A)$^+$ RNA fraction purified by the procedure of Cathala *et al.* (15). Double-stranded cDNA was prepared by the procedure of Land *et al.* (16). The dC-tailed double-stranded cDNA was annealed with dG-tailed *Pst* I-digested pBR322 and then used to transform *Escherichia coli* HB101. Transformants containing plasmids encoding PGA were identified by hybridization with a synthetic 25-base oligonucleotide probe synthesized by P. J. Barr (Chiron Corporation, Emeryville, CA). The oligonucleotide (5′ ATC-CTCAGGGTTGAAGCGGTTGTGG) was complementary to residues 392–416 of the PGA mRNA molecule predicted by the genomic clone nucleotide sequence (8).

The nucleotide sequence of the *PGA* plasmid (phpep14–21) insert used in this report was determined by the dideoxy chain-termination method (17, 18). The phpep14–21 insert detected exons 3 through 8 of the nine-exon *PGA* gene. Probes specific for exons 3–5 and exons 6–8 were obtained by digestion of the phpep14–21 insert with *Hae* II. The human insulin gene (*INS*) probes, phins310 and phins96, detected the RFLP adjacent to the 5′ end of *INS* as described (13, 14).

Gene Marker Analysis. Ten micrograms of DNA isolated from cell lines or whole blood (13, 14) was digested in 250 μl for 16–20 hr with *Eco*RI or *Pvu* II under the reaction conditions suggested by the manufacturer (New England Biolabs and Bethesda Research Laboratories). Samples containing digested DNA were subjected to electrophoresis on 0.75% or 1.0% agarose gels, transferred to nitrocellulose, and hybridized with the PGA or INS cDNA probes under conditions previously described (13, 14). The *PGA* and *INS* plasmid inserts were labeled to high specific activity (1–3 × 10^8 cpm/μg) by nick-translation using standard procedures. We used high-stringency washing conditions (0.015 M NaCl/0.0015 M sodium citrate, containing 0.1% $NaDodSO_4$ at 60°C). Under these conditions, no cross-hybridization of the human PGA cDNA probe with the mouse gene(s) was detected, and only weak hybridization of the human INS cDNA probe with the mouse gene was detected.

PGA Electrophoresis. Urine samples were separated by polyacrylamide gel electrophoresis and stained for pepsinogen-derived pepsin activity (4). Pepsinogen phenotypes were classified as previously described, except the common *A*, *B*, and *C* alleles refer to specific haplotypes that contain three (*PGA3-PGA4-PGA5*), two (*PGA3-PGA4*), and one (*PGA4*) of the *PGA* genes, respectively.

RESULTS

The phpep14–21 plasmid with 705-base-pair (bp) insert was isolated from the human gastric mucosa cDNA library and shown to contain a 686-bp portion of the exact nucleotide sequence of a human genomic *PGA* clone (8). The cDNA insert encoded a portion of exon 3 through exon 8 of the nine-exon *PGA* gene corresponding to amino acids 80 through 307 (residues 332 through 1017 of the predicted 1374-residue mRNA sequence). It was used to analyze the panel of human–mouse somatic cell hybrids that were informative for each human chromosome (Table 1). The presence of human *PGA* fragments detected by Southern blot hybridization (Fig. 1 *Upper*) correlated exactly with the segregation

Table 1. Mouse–human somatic cell hybrid panel

	Human chromosomes																								Genes	
Clone	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	*INS*	*PGA*
84-2/8	82	73		82	64	100	91	91			9	73		100	100	73	100	91	73	100		100			+	+
84-4/8	83	83	92	92		100	83	83		75			100	100	92		92	92	92	100					−	−
84-5/8		90	80		40	20		100					20	90	100		100		50		90	80	60	90	−	−
84-7/8			80	67		67	7	20			53	27		87			87			47					+	+
84-13/8			19			19				19		19	19		19	19	100	19		19	19				−	−
84-20/8		50	64	57	14		93	43			36	50	93	93	7	43	100	93		36	93	43	57		+	+
84-25/8				62	92	85		62		54	85	77					100		15		100				+	+
84-27/8		91			91	94		88			100	91			97		88	91		91	97				+	+
84-34/9				33		100	100	58		83	92			100			92		25	58		42			+	+
84-3/5	91	81	94	97		100	94	91		91			97	81	97		97	97	97	100					−	−
84-21/5							82			85		82		97			100			97	100			18	−	−
84-26/5	79		100	82	72	93	97	97		50		46	82	79	71	39	96	93	86	61	96	82		61	−	−
84-30/5			71	43	26	49				49		20	77	83			100	69		20	91	57			−	−
84-37/4				97	70	97	93	13			63	100		80			93			90	83	80			+	+
84-39/5			94				91	97							85		100								−	−
37-6/12		53	89		81	86	94			67	83*			67	67	67	61	78	67	72	47		83*		−	−
DUA5BSAgA			41		88						81			38			88	81			75				+	+
DUA3BSAgA		53					94	44					69	28			72								−	−
TSL-1			37	43					3	47	110		20	10		10	77	73		77	7				+	+
REX11BSAgB		53					94	44					69	28			72								−	−
SIR-8	73	30	30	77	73		53	70	37	60	63	120	30	40	13	97	53	53			93	33	73		+	+
WIL-14	47		41		3		68	35		38		41		12	24		59						65		−	−
REW-11				10							107	7	27			23				33	80	3	77		+	+
REW-8D				10				7						7			67			40	43	57	53		−	−
EXR-4CsAz	27		10	23	57	30	30	17	10	47	13†	27	20	33	13		10	33	17	33	53	10	13†		−	−
XER-7	47	20	27	30	97	40	90	60	23	113	67‡	43	23	47	17			40	47				67‡		−	+
Discordancy	15	12	16	8	9	13	14	12	11	17	0	10	15	15	17	9	15	13	14	13	11	12	11	15		
	26	26	26	26	26	26	26	26	26	26	23	26	26	26	26	26	26	26	26	26	26	26	23	26		

The frequency of each chromosome is indicated. The total number of each particular chromosome homolog observed was divided by the total number of metaphase preparations examined and, thus, potentially ranged from 0 to >100%. The chromosome was ascertained as present if >5% of the metaphase preparations examined contained the intact chromosome; >25 metaphase preparations were examined for each cell line.

*Contained an X/11 translocation chromosome Xqter → Xp2105::11q13 → 11qter.

†Contained X/11 translocation chromosome Xpter → Xq22::11q13 → 11qter; also contained 3% intact X.

‡Contained 11/X translocation chromosome 11qter → 11p11::Xq11 → Xqter; also contained 57% intact X.

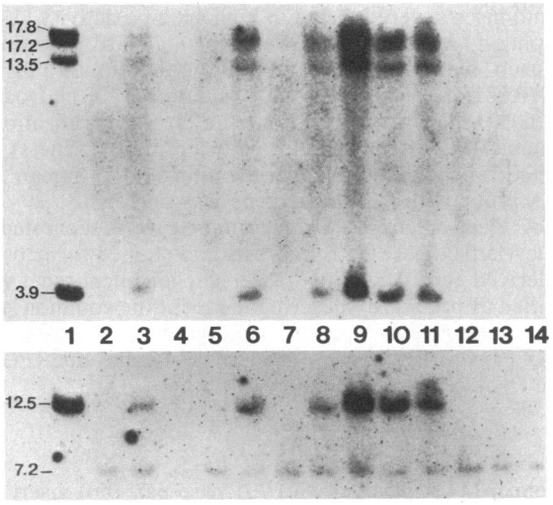

FIG. 1. Southern blot analysis of human–mouse somatic cell hybrids for *PGA* and *INS*. (*Upper*) The *Eco*RI DNA restriction fragments hybridizing with the PGA cDNA probe are indicated (17.8, 17.2, 13.5 and 3.9 kb). (*Lower*) Human and mouse DNA restriction fragments hybridizing with the INS cDNA probe (12.5 and 7.2 kb, respectively) on the same nitrocellulose blot indicated (lower portion). Lanes: 1, the human fibroblast cell line IMR-91; 2, the mouse cell line B82 (GM 347); 4, 5, 7, and 12–14, mouse–human hybrids negative for *PGA* and *INS*; lanes 3, 6, and 8–11, mouse–human hybrids positive for *PGA* and *INS*.

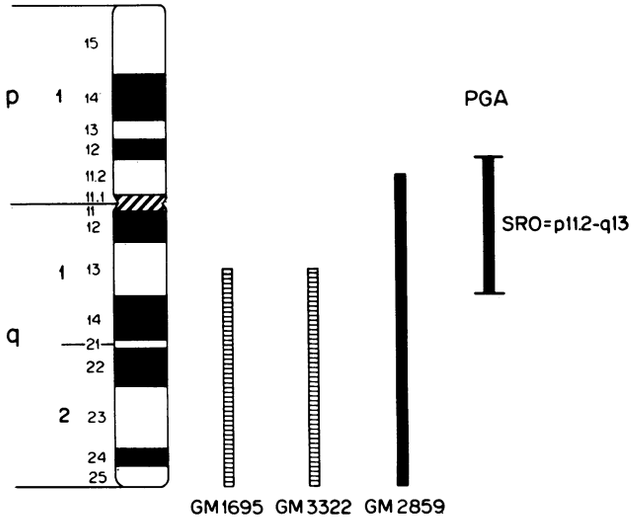

FIG. 2. Regional localization of the human *PGA* locus on human chromosome 11. The banding pattern of human chromosome 11 is indicated according to the standard nomenclature. The regions of chromosome 11 contained in hybrids constructed from three human cell lines are indicated with their General Medical identification designation (GM 1695, GM 3322, and GM 2859). The chromosomal subregion that contained the *PGA* genes is indicated with the solid line, whereas the subregions that did not contain the *PGA* genes are indicated with the hatched line. The shortest region of overlap (SRO) containing the *PGA* complex, 11p11→q13, is indicated.

pattern of human chromosome 11. No other chromosome exhibited a segregation pattern that was concordant with the pattern observed with the *PGA* genes as the degree of discordancy varied from 31% to 65%. Concurrent analysis of the same nitrocellulose blots with an INS cDNA probe as a marker of human chromosome 11 (Fig. 1 *Lower*) demonstrated that the *INS* and *PGA* genes exhibited similar autoradiographic intensities and the identical segregation pattern among the somatic cell hybrids containing an intact human chromosome 11 (Table 1). The regional assignment of the PGA genes to 11p11→q13 was accomplished by analysis of hybrids that contained three different X/11 translocations (Fig. 2).

Two distinct patterns of hybridization of the PGA cDNA probe with the human and somatic hybrid cell lines were observed. One pattern contained four *Eco*RI restriction fragments of 17.8, 17.2, 13.5, and 3.9 kilobases (kb), and the other pattern contained only three fragments of 17.8, 17.2, and 3.9 kb. In the latter pattern, the 13.5-kb fragment was missing, and the 3.9-kb fragment also exhibited less autoradiographic intensity (Figs. 3 and 4). To identify the exon composition of the fragments, we performed hybridization with probes specific for exons 3–5 and exons 6–8. The 17.8-, 17.2- and 13.5-kb fragments each hybridized with the probe specific for exons 3–5 whereas only the 17.8- and 3.9-kb fragments hybridized with the probe specific for exons 6–8. These results suggested that the two different restriction fragment patterns observed contained three and two *PGA* genes, respectively (Fig. 3 *Upper*). Alternatively, they could have resulted from an allelic RFLP involving only two genes.

To distinguish between these alternative explanations of the *PGA* hybridization patterns, we analyzed the hybrid cell lines for the *INS* polymorphism. All of the hybrid cell lines with one exception were found to contain only one human chromosome 11 homolog by analysis for the *INS* RFLP and by quantitative cytogenetic analysis. Since only single chromosome 11 homologs (with a single *INS* allele) were present in the hybrids, we demonstrated that the two different hybridization patterns resulted from chromosome haplotypes that contained 3 or 2 *PGA* genes. Comparison of the *PGA* fragment patterns observed in the somatic cell hybrids and the human parental cell lines revealed the same fragment pattern with one exception. In this instance, a hybrid containing the three-*PGA*-fragment pattern was derived from a parental human cell line containing the four-*PGA*-fragment pattern. This result was explained by the presence of the two different haplotypes in the original human parental cell line of which only one was present in the hybrid cell analyzed (Fig. 3 *Lower*).

We suspected that the two different restriction fragment patterns observed in the hybrid and human cell lines reflected the common genetic polymorphism involving the absence of the Pg 5 isozymogen. Comparison of the gene fragment patterns obtained from individuals who were missing the Pg 5 isozymogen (phenotype BB) with those containing the isozymogen (AA and AB) revealed that the absence of Pg 5 isozymogen correlated with the absence of the 13.5-kb fragment and a reduced 3.9-kb fragment signal (Fig. 4). We identified 12 BB individuals in five families, in which all were missing the 13.5- and 3.9-kb fragments. We performed repeated Southern blot analysis of the RFLP patterns obtained with AA and AB individuals; however, it was difficult to reliably distinguish between the DNA fragment patterns obtained from individuals with these two urinary phenotypes. Examination of phenotypes involving the C phenotype indicated that it is determined by a haplotype that contains a single *PGA4* gene (Fig. 3 *Lower*). Thus, the existence of at least three haplotypes, each containing different combinations of the *PGA5*, *PGA4*, and *PGA3* genes, were indicated by these studies (Fig. 3 *Lower*).

DISCUSSION

A central problem for defining the basis of the PGA (or PG I) heterogeneity has been the inability to detect molecular differences among the isozymogens beyond that due to differences in net ionic charge. In this regard, recent studies with monoclonal antibodies indicate that Pg 3, Pg 4, and Pg 5 are the primary PGA gene products and that the antigenic

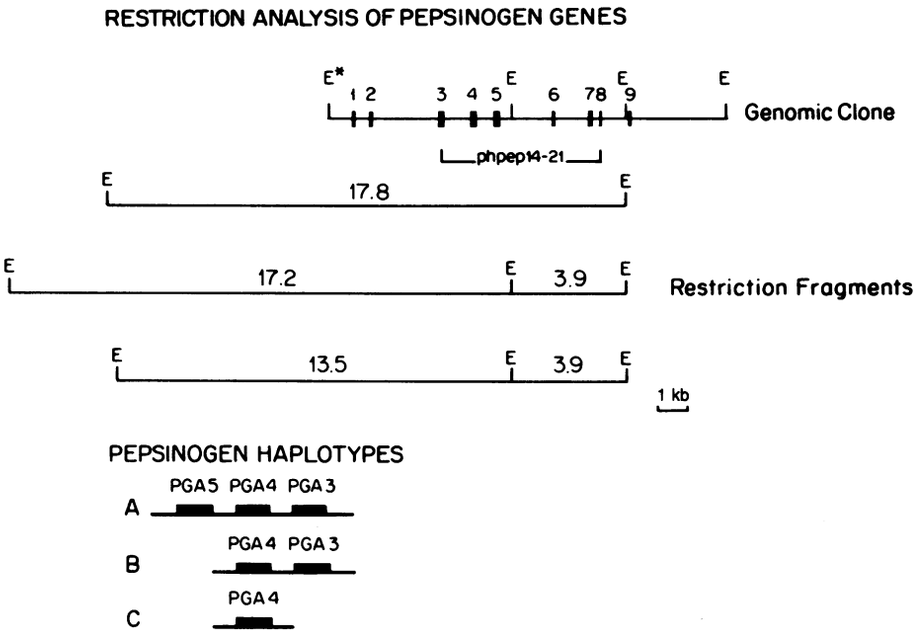

FIG. 3. Organization of the human *PGA* genes. (*Upper*) Restriction analysis of *PGA* genes. The exons contained in the *Eco*RI fragments (E) of a *PGA* gene isolated by Sogawa *et al.* (8) are indicated. An artificial *Eco*RI site (E*) at the 5′ end of the genomic clone is indicated. The phpep14-21 cDNA probe encoded sequences corresponding to a portion of exon 3 through exon 8 of the nine-exon *PGA*. The exon composition of the 17.8-, 17.2-, 13.5- and 3.9-kb (doublet) fragments was determined by hybridization as described. The three- and four-restriction fragment patterns (17.8, 17.2, 13.5, and 3.9 kb vs. 17.8, 17.2, and 3.9 kb) resulted from the absence of the 13.5- and 3.9-kb fragments contained in the composite gene shown at the bottom. (*Lower*) A model for the *PGA5*, *PGA4*, and *PGA3* gene composition of the *A*, *B*, and *C* haplotypes is presented. The haplotypes differ in that the *PGA5* or *PGA5* and *PGA3* genes are not present in the *B* and *C* haplotypes. The *PGA5* gene corresponds to the 13.5-kb and 3.9-kb restriction fragments. Additional haplotypes suggested by genetic studies of the electrophoretic polymorphism, including those missing the *PGA3* or *PGA4* and *PGA3* genes are not presented.

difference between them resided on the pepsin region of the molecules (6, 19). Sogawa *et al.* (8) isolated and characterized a pepsinogen gene from a human genomic library by using a porcine cDNA pepsinogen probe. The clone was identified to contain the 9.4-kb pepsinogen gene by comparison with partial amino acid sequences of human pepsinogen and nucleotide sequences of the homologous porcine pepsinogen. Comparison of the predicted amino acid sequence of the pepsinogen gene described by Sogawa *et al.* (8) with our preliminary amino acid sequencing studies of Pg 3, Pg 4, and Pg 5 confirmed that it was encoded by one of the *PGA* genes. Since the structure of only a single genomic clone was reported, it was not known if other nonallelic *PGA* genes existed as predicted by genetic studies of the protein polymorphism. To assess the relationship between the molecular structure and organization of the *PGA* genes with the previously described genetic variation, we isolated a human cDNA clone and used it to determine the chromosomal location of the *PGA* complex and to investigate its organization.

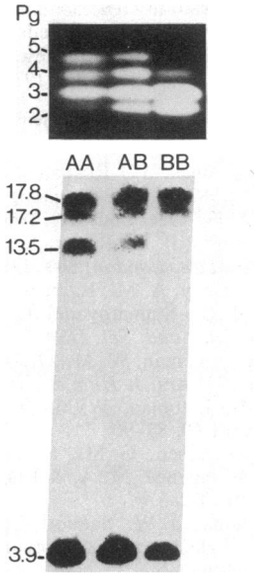

FIG. 4. The urinary pepsinogen phenotypes AA, AB, and BB (*Upper*) and the corresponding *Eco*RI restriction fragment patterns detected with the PGA cDNA probe (*Lower*). Urine from AA, AB, and BB individuals was subjected to polyacrylamide gel electrophoresis and stained for pepsin-derived activity. The location of the individual PGA isozymogens, Pg 5, Pg 4, Pg 3, and Pg 2, are indicated. DNA from the individuals was isolated from whole blood, digested with *Eco*RI, separated on 0.75% agarose, transferred to nitrocellulose, and hybridized with the PGA cDNA probe. The sizes of the gene fragments hybridizing with the PGA probe are indicated.

The entire *PGA* gene complex containing variable numbers of genes was localized to the centromere of human chromosome 11 in the region p11→q13. This finding was not unexpected, since the only undisputed positive *PGA* linkage data obtained by previous studies was with the β-globin gene that is located on the short arm of chromosome 11 (4). We identified a very unusual RFLP characterized by variable numbers of *PGA* genes. It is unlikely that our PGA cDNA probe detects the PG II gene(s) (*PGC*). The immunochemical and biochemical differences between PG I and PG II are dramatic and they are encoded by different gene loci (1, 20–22). In addition, the high-stringency hybridization conditions used revealed no cross-hybridization of the human PGA cDNA probe with the mouse *Pga* genes, and blot hybridizations of the original poly(A)$^+$ RNA preparation revealed a single hybridizing species of 1350 bp (data not shown). We used exon-specific probes to identify the exons contained in the restriction fragment patterns and demonstrated the existence of two haplotypes that contained two and three *PGA*

genes, respectively, by detailed analysis of cell hybrids containing individual chromosome 11 homologs. The presence or absence of a specific isozymogen, Pg 5, was correlated with the gene copy differences observed in the two restriction fragment patterns detected with our cDNA probe. Therefore, this analysis demonstrated that the 13.5- and 3.9-kb fragments contained a portion of the *PGA5* gene.

The molecular characterization studies reported here verify the predictions of previous genetic and biochemical studies that differences in the copy number of three PG I genes are the primary determinant of PG I polymorphism as illustrated in Fig. 3 (4, 6, 7, 19). A number of individual factors and combinations thereof can result in protein heterogeneity. These include multiple gene loci, multiple alleles at a single gene locus, transcriptional processing, and secondary or posttranslational modifications of primary gene products (4, 20, 23, 24). This type of multigene protein polymorphism is extraordinarily rare, having only been described for low-frequency haplotypes of α-globin that contain null, single, and triplicate copies of the gene (25). Normally, two copies of the α-globin genes are found in man, and the consequence of the deficient haplotypes is α-thalessemia. Therefore, human PGA represents a unique protein polymorphism in that its heterogeneity is determined by three factors: variation in the number of structural genes, rare allelic variation at the individual *PGA* gene loci, and posttranslational modification of the primary gene products. It is interesting to speculate about the evolutionary significance of the PGA haplotypes. This haplotype polymorphism occurs at a high frequency in the United States, and >55% of the population is heterozygous. The incidence of the different phenotypes also varies greatly among other populations, since previous studies have reported the absence of the B phenotype in Oriental populations (26). We speculate that this gene complex is undergoing a period of evolutionary change by the process of gene duplication and selection. Because of the high frequency of haplotype variation, PGA will be an invaluable gene marker for linkage analysis of human chromosome 11.

The potential for relationships between the pepsinogen haplotype variation and specific clinical disease is speculative at this point. Samloff and co-workers have found that approximately half of the duodenal ulcer population has hyperpepsinogenemia I; in family and sibship studies, they have found that this trait is inherited in an autosomal dominant manner and is a subclinical marker of the disease (27–29). Genetic factors have long been known to contribute to the pathogenesis of several gastroduodenal disorders, including atrophic gastritis and pernicious anemia, gastric ulcer, and duodenal ulcer. In addition, associations have been reported between pepsinogen phenotypes and several gastroduodenal disorders (30–33). It should be noted that these studies have a low power of discrimination, since the associations described are based on a simple polymorphism—namely, the relative intensity or absence of Pg 5. The molecular characterization of the pepsinogens, including the development of probes specific for each gene, will provide for additional sensitive and specific studies of these proteins in health and in disease.

The technical assistance of A. Bell, M. Byers, C. Cass, R. Eddy, N. Fong, A. Graham, L. Hanley, and M. Stempien is gratefully acknowledged. We wish to thank Drs. P. J. Barr, L. S. Cousens, L. B. Rall, J. I. Rotter, I. M. Samloff, and R. Sanchez-Pescador for their invaluable suggestions. This work was supported by Merit Review from the Veterans Administration and by Research Grants AM 32015, GM 20454, and HD 05196 from the National Institutes of Health.

1. Samloff, I. M. (1982) *Gastroenterology* **82,** 26–33.
2. Samloff, I. M. & Townes, P. L. (1970) *Science* **168,** 144–145.
3. Weitkamp, L. R. & Townes, P. L. (1975) in *Isozymes: Genetics and Evolution*, ed. Markert, C. L. (Academic, New York), Vol. 4, pp. 829–838.
4. Taggart, R. T., Karn, R. C., Merritt, A. D., Yu, P. L. & Conneally, P. M. (1979) *Hum. Genet.* **52,** 227–238.
5. Korsnes, L. & Gedde-Dahl, T. (1980) *Ann. Hum. Genet. (London)* **43,** 199–212.
6. Taggart, R. T. & Samloff, I. M. (1984) *Gastroenterology* **86,** 1110 (abstr.).
7. Frants, R. R., Pronk, J. C., Pals, G., Defize, J., Westerveld, B. D., Meuwissen, S. G. M., Kreuning, J. & Eriksson, A. W. (1984) *Hum. Genet.* **65,** 385–390.
8. Sogawa, K., Fujii-Kuriyama, Y., Mizukami, Y., Ichihara, Y. & Takahashi, K. (1983) *J. Biol. Chem.* **258,** 5306–5311.
9. Owerbach, D., Bell, G. I., Rutter, W. J. & Shows, T. B. (1980) *Nature (London)* **286,** 82–84.
10. Owerbach, D., Bell, G. I., Rutter, W. J., Brown, J. A. & Shows, T. B. (1981) *Diabetes* **30,** 267–270.
11. Champion, M. J., Brown, J. A. & Shows, T. B. (1978) *Cytogenet. Cell Genet.* **22,** 498–502.
12. Mohandas, T., Sparks, R. S., Hellkuhl, B., Grzeschik, K. S. & Shapiro, L. J. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 6759–6763.
13. Bell, G. I., Karam, J. H. & Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 5759–5763.
14. Bell, G. I., Horita, S. & Karam, J. H. (1984) *Diabetes* **33,** 176–183.
15. Cathala, G., Savouret, J., Mendez, B., West, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983) *DNA* **2,** 329–335.
16. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. (1981) *Nucleic Acids Res.* **9,** 2251–2226.
17. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143,** 161–178.
18. Sanchez-Pescador, R. & Urdea, M. S. (1984) *DNA* **3,** 339–343.
19. Taggart, R. T. (1985) in *Pepsinogens in Man: Clinical and Genetic Advances*, eds. Kreuning, J., Samloff, I. M., Rotter, J. I. & Eriksson, A. W. (Liss, New York), pp. 15–22.
20. Taggart, R. T. (1980) in *Genetics and Heterogeneity of Common Gastrointestinal Disorders*, eds. Rotter, J. I., Samloff, I. M. & Rimoin, D. (Academic, New York), pp. 51–65.
21. Samloff, I. M. & Townes, P. L. (1970) *Gastroenterology* **58,** 462–469.
22. Foltman, B. & Jensen, A. L. (1982) *Eur. J. Biochem.* **128,** 63–70.
23. Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. (1982) *Nature (London)* **298,** 240–244.
24. Rosenfeld, M. G., Mermod, J., Amara, S. G., Swanson, L. W., Sawchenko, P. E., Rivier, J., Vale, W. W. & Evans, R. M. (1983) *Nature (London)* **304,** 129–135.
25. Goossens, M., Dozy, A. M., Embury, S. H., Zachariades, Z., Hadjiminas, M. G., Stamatoyannopoulos, G. & Kan, Y. W. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 518–521.
26. Samloff, I. M., Liebman, W. M., Glober, G. A., Moore, J. O. & Indra, D. (1973) *Am. J. Hum. Genet.* **25,** 178–180.
27. Samloff, I. M., Liebman, W. M. & Panitch, N. M. (1975) *Gastroenterology* **69,** 83–90.
28. Rotter, J. I., Petersen, G. M., Samloff, I. M., McConnell, R. B., Ellis, A., Spence, M. A. & Rimoin, D. L. (1979) *Ann. Intern. Med.* **91,** 372–377.
29. Rotter, J. I., Sones, J. W., Samloff, I. M., Richardson, C. T., Gursky, J. M., Walsh, J. H. & Rimoin, D. L. (1979) *N. Engl. J. Med.* **300,** 63–65.
30. Samloff, I. M. & Cole, D. (1975) *Gastroenterology* **68,** 144 (abstr.).
31. Ellis, A. & McConnell, R. B. (1982) *Gastroenterology* **83,** 1261–1263.
32. Ellis, A., Hughes, S. & McConnell, R. B. (1982) *Br. J. Cancer* **46,** 289–290.
33. Whitecross, D. P., Armstrong, C., Clarke, A. D. & Piper, D. W. (1973) *Gut* **14,** 850–859.