MOLECULAR AND CELLULAR BIOLOGY, Aug. 1985, p. 2114–2122
0270-7306/85/082114-09$02.00/0

Vol. 5, No. 8

A Genetic Locus Closely Linked to a Protease Inhibitor Gene Complex Controls the Level of Multiple RNA Transcripts

ROBERT E. HILL,[1]* PHILLIP H. SHAW,[2] RICHARD K. BARTH,[3] AND NICHOLAS D. HASTIE[1]

Molecular Genetics Section, Medical Research Council Clinical and Population Cytogenetics Unit, Western General Hospital, Edinburgh EH4 2XU, Scotland[1]; Swiss Institute for Experimental Cancer Research, CH-des Boveressez, CH-1066, Lausanne, Switzerland[2]; and Division of Biology, California Institute of Technology, Pasadena, California 91125[3]

Received 11 February 1985/Accepted 21 May 1985

The two major protease inhibitors in mouse plasma are α_1-protease inhibitor (α_1-PI), putative inhibitor of neutrophil elastase, and contrapsin, an inhibitor in vitro of trypsinlike proteases. We have shown by nucleotide sequence analysis that these two inhibitors are related (R. E. Hill, P. H. Shaw, P. A. Boyd, H. Baumann, and N. D. Hastie, Nature (London) 311:175–177, 1984). Here, we show that the contrapsin and α_1-PI genes are members of two different multigene families, each containing at least three genes in mice and rats. We established the chromosomal locations of these genes by analyzing the segregation of restriction fragment length polymorphisms in recombinant inbred mouse strains. These experiments show that the multiple genes in each family are clustered and that the two gene families are closely linked on chromosome 12. Thus the genes for contrapsin and α_1-PI are likely to have evolved by duplication of a common ancestral gene. The contrapsin multigene family codes for multiple mRNA transcripts in the liver. There is a genetic difference among inbred mouse strains in the regulation of two of these transcripts. In some inbred strains the transcripts are synthesized constitutively; in others they are induced by inflammation. We mapped in recombinant inbred strains the regulatory locus responsible for this genetic variation and found it is linked to the contrapsin multigene family, which suggests a *cis*-acting regulatory element. We also found that the contrapsin and the α_1-PI multigene families have acquired very different regulatory responses since the time of the gene duplication event.

The two major mouse plasma protease inhibitors are α_1-proteinase inhibitor (α_1-PI; also called α_1-antitrypsin) and contrapsin, which together are responsible for most of the trypsinlike inhibitory activity of mice (27, 28). These two inhibitors are synthesized by hepatocytes, and secretion from this cell type is primarily responsible for their high levels in plasma (3, 4, 21). Mouse α_1-PI is homologous to (14) and appears to be functionally equivalent to human α_1-PI; in contrast, contrapsin activity appears to be novel to rodents (27). At the nucleotide sequence level, however, the contrapsin gene is closely related to that encoding human plasma α_1-antichymotrypsin (α_1-achy) (14). In addition, contrapsin and α_1-PI are related to each other (59% nucleotide homology) but to a lesser extent than with their counterparts in humans (14).

These inhibitors have been placed into a larger superfamily of proteins (sharing 29 to 34% amino acid homology with each member) which includes other protease inhibitors such as antithrombin III (7) as well as proteins with no known inhibitory activity, including chicken ovalbumin (15) and rat angiotensinogen (10). It is not known whether these members evolved from a common ancestral gene or are related through a process of convergent evolution (19). Genetic analysis of the chromosomal location of related genes may provide insight into their mode of evolution. Because α_1-PI and contrapsin are the most closely related members of this superfamily, it is of particular interest to map their chromosomal locations.

We showed, first, that both these genes are members of small multigene families in rodents. By using recombinant inbred strains of mice, we showed that the genes in each multigene family are clustered and that the two gene families are closely linked on chromosome 12. Therefore, the gene families are likely to have evolved by duplication from a single common ancestral gene. These two families are designated the *Spi* (serine protease inhibitor) complex; the *Spi*-1 locus contains the α_1-PI gene, and the *Spi*-2 locus contains the contrapsin gene.

Secondly, we studied the expression of these multigene families in an attempt (i) to establish the number of genes expressed and the level and regulation of expression of each and (ii) to identify any expression differences that occur among inbred mouse strains. A genetic locus involved in controlling the level of a particular mRNA could be instrumental in identifying elements at the molecular level that are involved in gene expression. We found that in the liver the contrapsin gene family can encode multiple mRNA transcripts detected as three different molecular weight species. The level of two of these differs dramatically among inbred strains. The strains that produce low levels of these two RNA species can be stimulated to synthesize levels comparable to those of the high-producing strains by treatment with bacterial lipopolysaccharides (LPS) or dexamethasone. We mapped this regulatory variation to the structural locus *Spi*-2, which suggests it functions as a *cis*-acting element. This regulatory locus is designated *Spi*-2r.

MATERIALS AND METHODS

Treatment of animals. Hypophysectomized mice were purchased from Charles River Breeding Laboratories, Inc. Hormonal treatment of the mice was begun on day 11 after surgery. Dexamethasone (Sigma Chemical Co.) was injected subcutaneously (2.5 mg in 300 μl of olive oil) at 24-h intervals for 2 consecutive days, and the animals were killed on the third day. Mice that received no hormonal treatment were killed on day 13 after hypophysectomy. Inflammation was

* Corresponding author.

induced in intact mice by a single subcutaneous injection of *Escherichia coli* LPS (serotype no. 0127:B8; Sigma) at a dose of 25 μg in 30 μl of phosphate-buffered saline (pH 7.3), and the mice were killed 24 h later.

Isolation of DNA and restriction endonuclease analysis. Nuclei from the livers of C57BL/6J and BALB/cBy mice and Buffalo rats were isolated by homogenization of the tissue in 0.3 M sucrose–3 mM EDTA–A buffer (60 mM KCl, 0.5 mM spermine, 0.15 mM spermidine, 14 mM 2-mercaptoethanol, 15 mM Tris hydrochloride [pH 7.4]) (20) and centrifugation of the nuclei through a 1.37 M sucrose cushion in A buffer. DNA was isolated from the nuclear pellet, which was suspended in pronase buffer (50 mM Tris [pH 10], 150 mM NaCl, 100 mM EDTA), lysed in 0.1% sodium dodecyl sulfate, and digested with 100 μg of RNase A per ml for 30 min at 37°C. Protein was removed by further digestion with protease K (100 μg/ml) in 1% sodium dodecyl sulfate for 4 h at 37°C, followed by phenol-chloroform extractions. The DNA was spooled onto a glass rod after the addition of 2 volumes of ethanol. Restriction digestions were done as recommended by the commercial supplier. Digested DNA was fractionated electrophoretically in 0.8% agarose gels, and the DNA was transferred to nitrocellulose by the method of Southern (26).

Preparation of radioactive probes and analysis of Southern blots. Plasmid DNA and insert fragments were radiolabeled with [α-^{32}P]TTP (800 Ci/mmol; Amersham Corp.) by nick translation (23). The small insert fragments of the contrapsin cDNA clone (no. pLv54) (3) were isolated from the complete pLv54 insert fragment by digestion with *Pvu*II or *Hae*II and fractionated on 1.5% agarose gels. The appropriate fragments were electroeluted from the gel onto NA-45 filter paper (Schleicher & Schuell, Inc.), and the bound DNA was eluted by the method of Dretzen et al. (11). The probes were hybridized to Southern blots in 5× SSC at 68°C and washed in 2× SSC at the same temperature (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate).

RNA isolation and fractionation, Northern blot analysis, and in vitro translation. Total RNA was isolated from mouse livers by using guanidinium hydrochloride (3). Cytoplasmic and polysomal RNA were isolated by homogenizing four mouse livers in 40 ml of 0.35 M sucrose in the TKMS buffer described by Shore and Tata (25) which contains 200 mM Tris-acetate [pH 8.5; at 4°C], 50 mM KCl, 10 mM $Mg(CH_3COO)_2$, and 6 mM 2-mercaptoethanol. This was made 8 mM Vanadyl Ribonucleoside Complex (Bethesda Research Laboratories, Inc.) before use. The homogenate was spun down at 8,000 rpm in a Sorvall HB-4 swinging bucket rotor for 10 min. Heparin (grade 1 from Sigma) was added to a final concentration of 500 μg/ml; 5 ml of supernatant was then used to extract cytoplasmic RNA with guanidium hydrochloride. The remaining supernatant was made 1% Triton X-100 and 1% sodium deoxycholate and homogenized in a Dounce homogenizer. This supernatant was layered on a discontinuous sucrose gradient with the 3-ml bottom layer containing 2.0 M sucrose and the 4-ml top layer containing 1.6 M sucrose, each with TKM buffer–6 mM 2-mercaptoethanol–100 μg of heparin per ml. These were spun in the Sorvall T865.1 fixed-angle rotor for 16 h (2°C) at 40,000 rpm. RNA was isolated from the pelleted polysomes by guanidinium hydrochloride extraction.

Electrophoresis of RNA and Northern blot analysis were done as described by Meehan et al. (21). Hybrid selection of RNA complementary to the pLv54 DNA was done by the method of Barth et al. (3). Translation of poly(A) RNA and hybrid-selected RNA was done in a fractionated reticulocyte lysate cell-free translation system as described by Held et al. (13).

RESULTS

Both contrapsin and α_1-PI are represented by small multigene families in the mouse and rat genome. Southern blot analysis of rodent genomic DNA by using either the pLv54 DNA probe (contrapsin cDNA clone) or the pLv1796 DNA probe (α_1-PI cDNA clone) revealed a number of restriction fragments which indicated that within the genome there are several related genes homologous to each cDNA probe (data not shown). To understand this complexity, we analyzed the restriction fragment patterns for DNA from two inbred mouse strains, C57BL/6J and BALB/c, and from Buffalo rats by using as probes small DNA fragments from different regions of the cloned insert. Figure 1 shows the results of two identical Southern blots hybridized with different fragments from within the contrapsin-cloned insert. One insert fragment contained DNA from the internal *Hae*II site to the 5′ end of the insert (220 bases of cDNA), and the other contained DNA from the *Pvu*II site to the 3′ end of the insert (288 bases of cDNA) (designated in Fig. 1). The DNA restriction fragments that hybridized with both the 5′ and 3′ insert probes must represent separate and distinct genes or pseudogenes. Several restriction enzymes were used to reduce the chance of comigration of fragments. The fragments in common represent a minimum number of genes, and the data suggest that there are at least three genes in both inbred strains of mice and in the rats. The gel lanes containing the *Sst*I digestions indicated that *Sst*I restriction sites exist in the genes between the regions represented by the probes. In all digests, the more 5′ probe produced a higher degree of complexity than the 3′ probe. We do not yet know whether this represents more introns within the region of the 5′ probe or whether the 3′ ends of some genes have diverged and will not hybridize under our conditions (5× SSC at 68°C).

A similar experiment with different insert fragments of the mouse α_1-PI DNA clone suggested that there are at least four related genes in the BALB/c mouse, five in the C57BL/6 mouse, and three in the rat haploid genomes (data not shown). These data indicate that each of the two protease inhibitors is encoded by a gene (or genes) that is a member of a family of related genes.

The multigene families for α_1-PI and contrapsin are clustered and closely linked on chromosome 12 of the mouse. To determine the relative chromosomal location of the members in these multiple gene families, we used DNA isolated from recombinant inbred (RI) strains of mice (2, 29). We wanted to determine, first, whether the multiple genes of each family are clustered or dispersed throughout the genome and, secondly, the chromosomal location of each family or family member. Briefly, RI strains are the result of systematic inbreeding of the F_2 progeny of two different inbred strains. This essentially fixes within each resulting strain a specific pattern of recombination. Having established restriction fragment length polymorphisms between two inbred strain progenitors, one can compare the strain distribution pattern of the unmapped gene in a set of RI strains relative to well-characterized genomic markers. Restriction fragment length polymorphisms in RI strains have proved very useful in the mapping of a number of mouse genes and gene families (5, 6, 12, 16).

Figure 2 shows the pattern of restriction fragments obtained for α_1-PI-related genes generated by the restriction enzyme *Bgl*II in C57BL/6J and DBA/2 mice. These are the

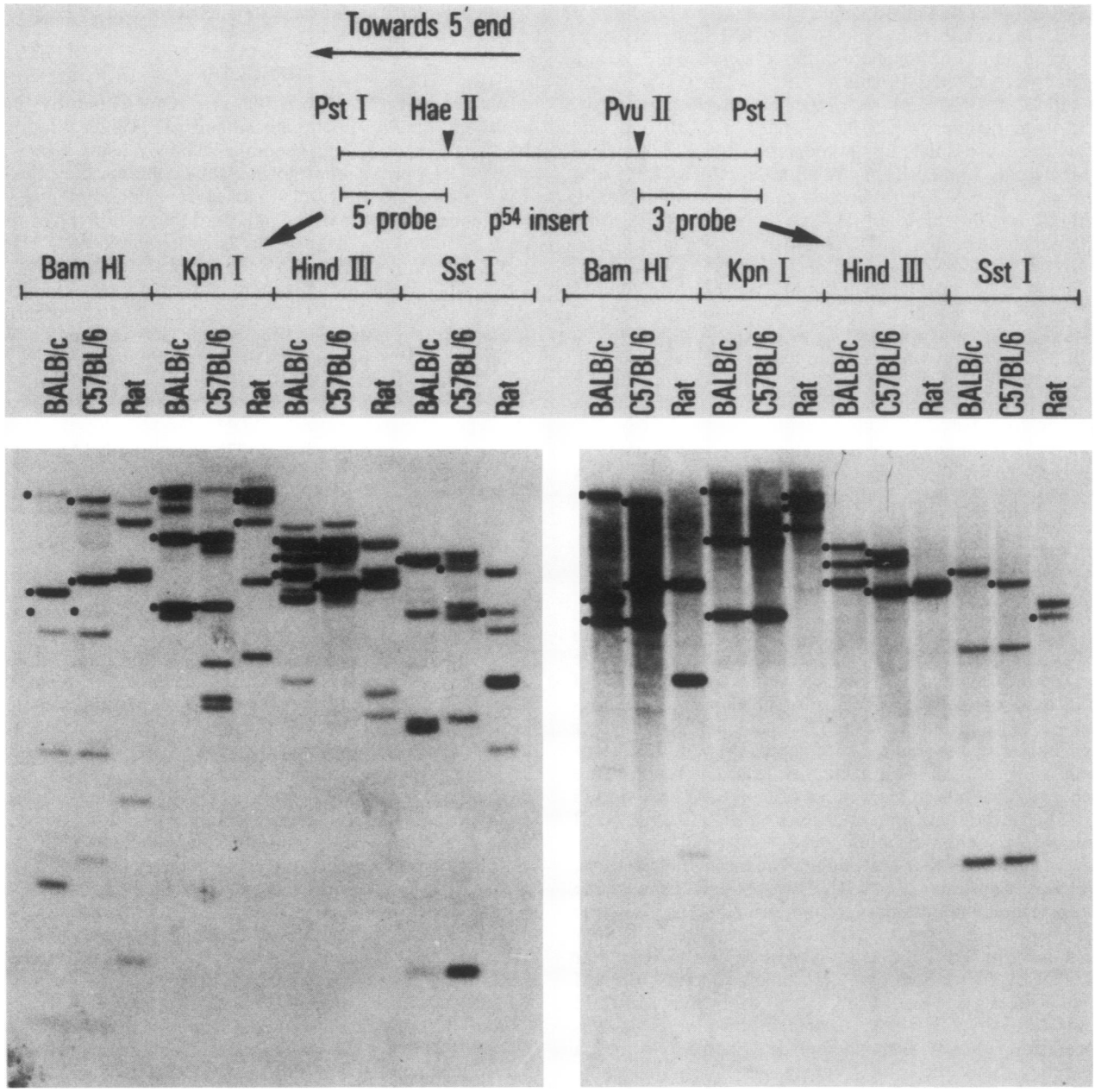

FIG. 1. Complexity analysis of the mouse multigene family of which contrapsin is a member. Two identical Southern transfers of mouse and rat DNAs were probed with ^{32}P-labeled fragments from the pLv54 cDNA insert. The fragments used as probes are identified at the top of the figure. The dots indicate the restriction fragments that are in common between the two transfers, representing distinct homologous genes. Unfortunately, the *Bam*HI restriction bands hybridized with the 5′ end probe in the lanes designated BALB/c and C57BL/6 did not reproduce well.

progenitor strains for the B×D series of RI strains. Of the restriction enzymes screened, *Bgl*II produced the most extensive variation of restriction fragments between these two inbred strains. The pattern of restriction fragments in each RI strain showed inheritance from either the C57BL/6 or the DBA/2 progenitor; no intermediate pattern representing recombination between the α_1-PI-related genes was generated. Therefore most if not all of the genes mapped to a single chromosome, and the lack of detectable recombination upon examination of 25 RI strains suggests that all are clustered within a region of approximately one centimorgan (cM). We designated this genetic locus *Spi*-1. To determine the chromosomal location of these genes, it was necessary to determine linkage of *Spi*-1 to known mouse chromosomal markers. Comparison of the strain distribution patterns of the above data with those of markers previously determined (computer analysis performed by Benjamin Taylor of Jackson Laboratory) revealed that *Spi*-1 is tightly linked to the prealbumin-1 (*Pre*-1) locus (Table 1). This places the *Spi*-1 locus on mouse chromosome 12. No recombination events occur between the *Pre*-1 and *Spi*-1 loci in this series of RI strains; however, both show seven recombination events with the immunoglobulin heavy chain-1 (*Igh*-1) locus in B×D RI strains.

We performed a similar analysis for the mouse contrapsin gene family. With the restriction enzyme *Hin*dIII, more than

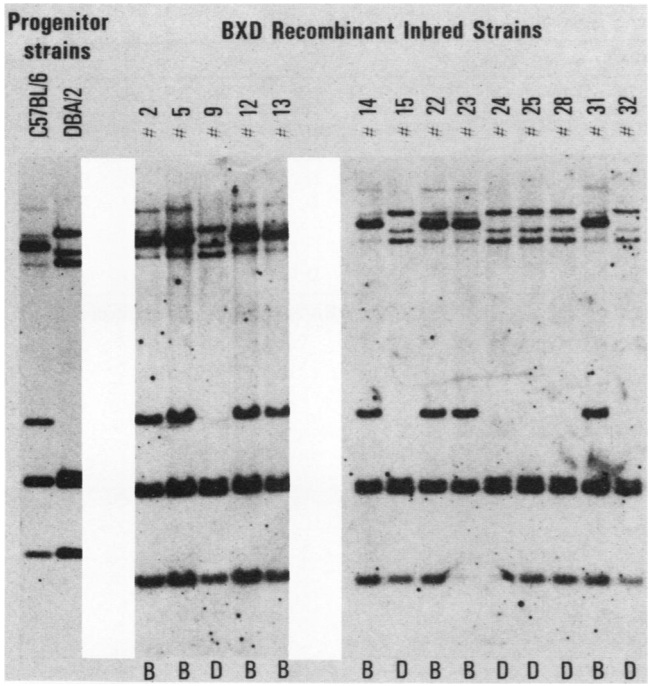

FIG. 2. Mapping restriction fragment length polymorphisms of the mouse α_1-PI multigene family in RI strains. *Bgl*II-digested DNA from the prongenitor strains and 13 RI strains is shown. Radiolabeled p1796 DNA was used as the probe. The progenitor strain restriction pattern of each RI strain analyzed is indicated below each gel lane as B (C57BL/6J) or D (DBA/2).

half the restriction fragments, including the three distinct genes shown in Fig. 1, were polymorphic (data not shown). All these polymorphic bands segregated together in RI strains, showing that, like the α_1-PI-related genes, they are clustered. We designated this gene cluster *Spi*-2. Furthermore, the strain distribution pattern for the restriction fragments corresponding to *Spi*-2 is the same as that for *Spi*-1 in the B×D RI strains, except for RI strain 13 (Table 1). DNA from strain 13 has the C57BL/6 restriction pattern for the *Spi*-1 locus and the DBA/2 pattern for the *Spi*-2 locus, representing a recombination event between the two gene clusters. Therefore, the two protease inhibitor gene families are linked on chromosome 12, with approximately one recombination unit between the two.

In the case of contrapsin, we were able to show that the particular gene(s) encoding contrapsin is within the cluster predicted by the DNA analysis. This was done by using a size polymorphism in the in vitro-synthesized contrapsin polypeptide detected in different inbred mouse strains. Because of the abundance of contrapsin mRNA (3), contrapsin is detected as a major protein band in the liver mRNA translation product analyzed by sodium dodecyl sulfate-urea polyacrylamide gel electrophoresis. The relative position of the contrapsin polypeptide band was shown by isolating contrapsin mRNA from liver poly(A) RNA of C57BL/6 or BALB/c mice by hybrid selection (3) with the pLv54 cDNA clone, translation in vitro, and analysis by electrophoresis (Fig. 3a). The mRNA selected from BALB/c liver poly(A) RNA directed the synthesis of an apparently smaller polypeptide than that from C57BL/6 (Fig. 3a). These polypeptides can be detected as two abundant polypeptides in the total translation product of liver poly(A) RNA from (C57BL/6 × BALB/c)F_1 mice (Fig. 3a). When total mRNA from 17 different inbred strains was translated in vitro and the product was analyzed, a polypeptide showing the abundance and size of one of the polymorphic forms of contrapsin was found in each strain. A representative sample is shown in Fig. 3b. Because C57BL/6 and DBA/2 express this polymorphism, it was possible to map the contrapsin gene, analyzing the translation products of liver mRNA from the B×D RI strains (Fig. 3c). Analysis of the strain distribution pattern of the in vitro-synthesized polypeptides (Table 1) places the contrapsin gene(s) within the multigene cluster, showing the recombination in RI strain 13 between the contrapsin gene and the *Spi*-1 locus.

A summary of the segregation analysis in the B×D RI strains is presented in Table 1. The results suggest the following arrangement of the loci on mouse chromosome 12. (i) The *Spi*-1 locus made up of at least four tightly clustered members is near or perhaps includes the *Pre*-1 locus showing seven recombinations with the *Igh*-1 locus. (ii) The *Spi*-2 locus (at least three tightly linked genes) is 1 ± 1 cM from the *Spi*-1 and *Pre*-1 loci (one recombination in 23 RI strains) to the distal side of the *Igh*-1 locus with which it has undergone eight recombination events.

Expression of multiple contrapsin mRNA transcripts is genetically variant in inbred mouse strains. We screened liver mRNA from several different inbred mouse strains to study the level of contrapsin and α_1-PI mRNA in an attempt to establish any genetic variation that may exist between inbred strains. Figure 4 shows a Northern blot analysis of liver RNA from 22 different inbred mouse strains with cloned contrapsin cDNA as probe. This analysis clearly showed that among these species there is a difference in the mRNA molecular weight species expressed in the liver. In some strains, represented by the C57BL/6 and C3H/He strains, only one contrapsin-related mRNA, which migrated at 18S (1.9 kilobases), was detected; however, in other strains, represented by BALB/c and DBA/2 strains, three RNA species migrating at 18S, 20S (approx. 2.4 kilobases), and 28S (approx. 5.0 kilobases) predominate. All 22 strains had one RNA pattern or the other; however, noticeable differences did exist in the relative level of RNA in some of these bands, e.g., C57Br/Cd.

Long exposure of Northern blots of C57BL/6 liver mRNA revealed no detectable contrapsin-related 20S mRNA, and the level was determined to be at least 20-fold lower in this strain than in DBA/2. However, the 28S contrapsin-related RNA was detected in C57BL/6 at very low levels (data not shown). The difference between DBA/2 and C57BL/6 in the total steady-state level of the contrapsin-related mRNAs (which is approximately 1% of total liver mRNA [3]) was less than twofold; the most easily discernible difference was in the expression of the upper-molecular-weight mRNAs.

Comparison of the distribution of protein polymorphism in the 17 inbred strains discussed above to that of these RNA patterns in the corresponding strains shows a distinct relationship between mobility of the in vitro-synthesized protein and mRNA pattern (Table 2). Strains that contain the three contrapsin-related liver RNA species show the less mobile polypeptide, and strains with one mRNA band show the more mobile polypeptide. This relationship of RNA pattern to protein mobility suggests that the RNA variation is a true genetic difference.

Liver mRNA from these inbred strains was also screened with the cloned α_1-PI cDNA probe. The α_1-PI RNA species is relatively invariant in inbred mice and is composed of a single molecular weight species migrating at 16-17S (approx. 1,600 bases) (data not shown).

TABLE 1. Segregation analysis of chromosome 12 genetic loci in B×D RI strains[a]

Genetic loci	Parameter	Patterns for strain no.																							
		1	2	5	6	8	9	11	12	13	14	15	16	18	19	21	22	23	24	25	27	28	29	30	31	32
Igh-1		B	B	B	D	B	B	D	D	B	B	D	D	D	B	B	B	B	D	B	D	D	B	D	D	D
						X	X		X					X	X					X					X	
Pre-1		B	B	B	D	D	D	D	B	B	B	D	D	B	D	B	B	B	D	D	D	D	B	D	B	D
α_1-PI	DNA	B	B	B	D	D	D	D	B	B	B	D	D	B	D	B	B	B	D	D	D	D	B	D	B	D
										X																
Contrapsin	Protein	B	B	B	D	D	D	D	B	D	B		D		D	B	B	B	D			D	B	D	B	D
	DNA	B	B	B		D	D	D	B	D	B	D	D	B	D	B	B	B	D	D	D	D	B	D	B	D

[a] The letters represent the progenitor strain patterns to which RI strains correspond for each locus. B, C57BL/6J; D, DBA/2. The Xs represent recombination events that occurred between two loci in an RI strain.

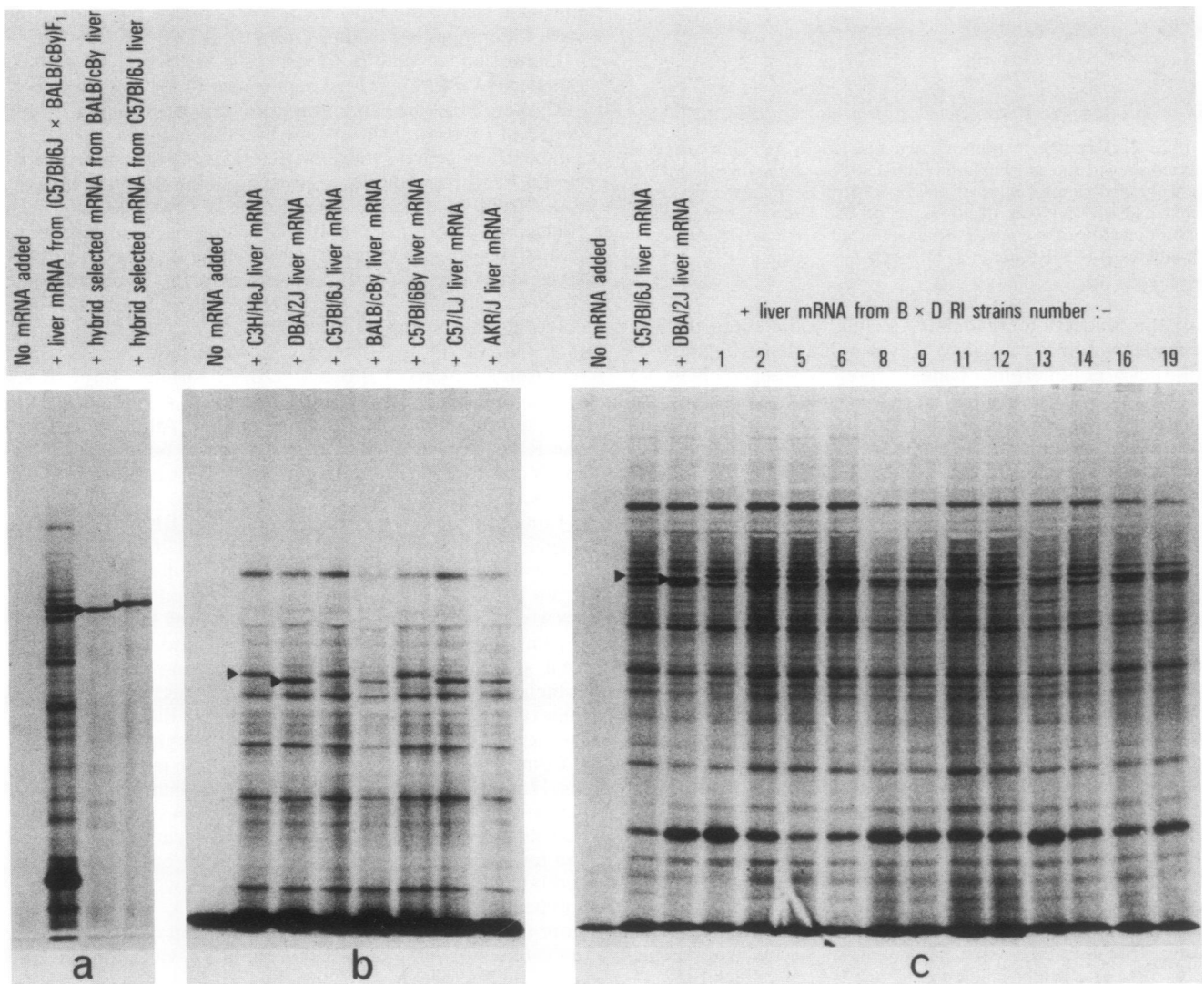

FIG. 3. Mapping the contrapsin gene by using size polymorphisms of the in vitro-synthesized protein. (a) Analysis of polypeptide polymorphism of [^{35}S]methionine-labeled translation product directed by total liver poly(A) RNA from (C57BL/6J × BALB/cBy)F_1 mice and from the RNA complementary to the pLv54 cDNA from the F_1 progenitors. The arrowheads in all panels point to contrapsin variants. (b) Analysis of the translation products from liver poly(A) RNA of RI progenitor strains. (c) Analysis of the translation products from liver poly(A) RNA from the B×D RI strains. Panels (a) and (c) are autoradiographs of 10% polyacrylamide gels, and panel (b) is a 7% gel.

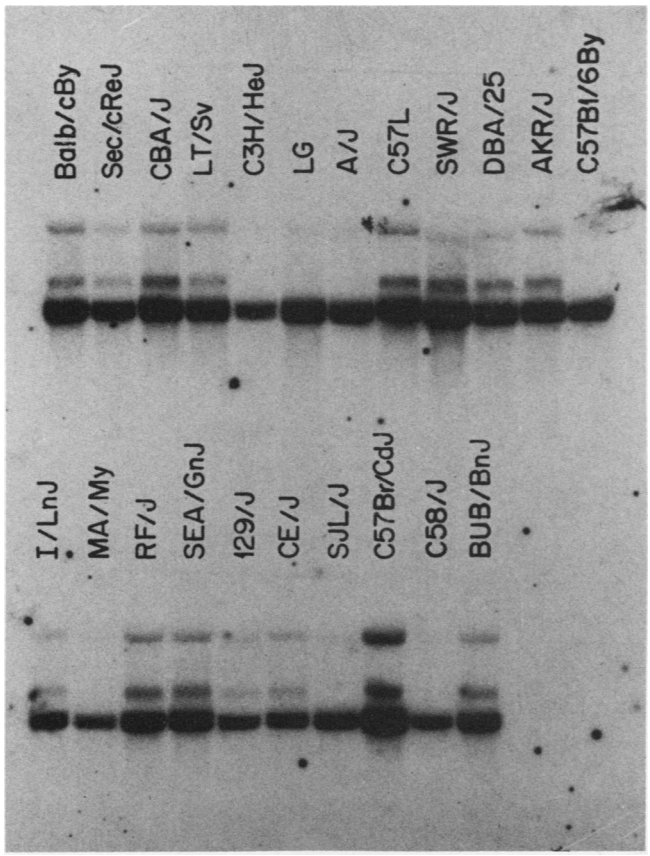

FIG. 4. Analysis of liver RNA from different inbred strains of mice for the expression of contrapsin mRNA. Radiolabeled pLv54 cDNA was used as the probe to screen the Northern blot. The name of each inbred strain used is indicated above the respective gel lane.

Upper-molecular-weight contrapsin-related mRNA species are inducible in inbred strains that normally express them at low levels. The levels of several plasma proteins synthesized in the liver are known to increase during the initial phase of an inflammatory response and are known as acute-phase reactants (for a review see reference 17). In humans, two such proteins are α_1-PI and, especially, α_1-achy (1), the counterparts of the two major mouse protease inhibitors. An inflammatory response is induced in rodents with a single subcutaneous injection of bacterial LPS. Attempts to affect the levels of contrapsin-related mRNAs in DBA/2 mice were not successful. Twenty-four hours after treatment with LPS there was no change in the relative amount of these mRNA species (Fig. 5a). Unexpectedly, when C57BL/6 or C3H mice were treated, although there was little change in the 18S RNA species, there was a rapid increase in the 20S and 28S contrapsin-related mRNAs to a level comparable to that of DBA/2 mice. The inbred strains tested have the capacity to express at high levels; the 20S and 28S RNA species expressed either constitutively (DBA/2 mice) or in response to inflammation (C57BL/6 and C3H). It therefore appears that an undefined regulatory element is reponsible for the genetic difference in contrapsin-related mRNA expression.

We also showed another level at which the genes at the *Spi*-2 locus can be regulated. Hypophysectomy (removal of the hypothalmus and pituitary responsible for producing many of the tropic and stimulating hormones, most importantly here adrenocorticotropin) of BALB/c and C57BL/6 mice resulted in a dramatic decrease (approx. 50-fold) in the levels of all the contrapsin-related mRNAs (Fig. 5a). Treatment of these animals with the glucocorticoid analog dexamethasone resulted in a rapid return (within 24 h) of mRNA levels to normal. However, in C57BL/6 mice the 20S and to a lesser extent the 28S mRNA species also appeared in the liver. Therefore, it appears that the normal levels in the liver of these mRNA species are at least in part maintained by the circulating levels of glucocorticoids. Apparently, high pharmacological doses can induce the production of the upper-molecular-weight mRNA species in C57BL/6 mice.

In an initial approach to determine whether these mRNA species can be translated, we prepared RNA from the cytoplasmic and polysomal cell fractions isolated from BALB/c mouse livers (Fig. 5b). Cytoplasmic RNA clearly contained all three molecular weight species, suggesting that these RNAs are processed and potentially translatable. We had difficulty isolating completely intact polysomal RNA; however, although degradation slightly obscured the resolution of the RNA species, the polysomal fraction contained all three mRNAs, again suggesting that all these RNAs are translatable. We have not yet determined the nature of the protein encoded by these transcripts.

In contrast, a similar analysis with the α_1-PI cDNA clone showed that LPS (data not shown), hypophysectomy, and dexamethasome had little effect on the levels of α_1-PI mRNA in the liver (Fig. 5c).

Regulatory variation is linked to structural locus. The strain variation that exists in the regulation of the upper-molecular-weight contrapsin RNA species can be explained in several ways. LPS has little effect on the levels of contrapsin-related RNA species in strains constitutively expressing these RNAs, which may suggest that *trans*-acting factors involved in the induction of the RNAs are present at a high level in these animals. Alternatively, the strain variation may be owing to a *cis*-acting regulatory element that lies close to the contrapsin-related gene(s) involved.

To characterize the regulatory difference observed, we assayed liver RNAs from the 21 B×D RI strains of mice. We compared the pattern of RNA expression in the liver of each RI strain with those of the progenitor strains DBA/2 (all

TABLE 2. Distribution of mRNA polymorphism in inbred mouse strains[a]

Parameter	Pattern for the following strains:																
	BALB/cBy	Sec/cReJ	CBA/J	LT/SV	C3H/HeJ	LG	A/J	C57L/J	SWR	DBA/2J	AKR/J	C57BL/6	I/LnJ	MA/My	RF/J	C57Br/CdJ	C58/J
RNA	+	+	+	+	−	−	−	+	+	+	+	−	+	−	+	+	−
Protein	H	H	H	H	L	L	L	H	H	H	H	L	H	L	H	H	L

[a] + Designates strains in which the three contrapsin-related mRNA species are predominant, − designates strains in which only one mRNA species is predominant, and H and L designate high and low mobility, respectively, of the in vitro-synthesized contrapsin polypeptide in sodium dodecyl sulfate-urea polyacrylamide gels.

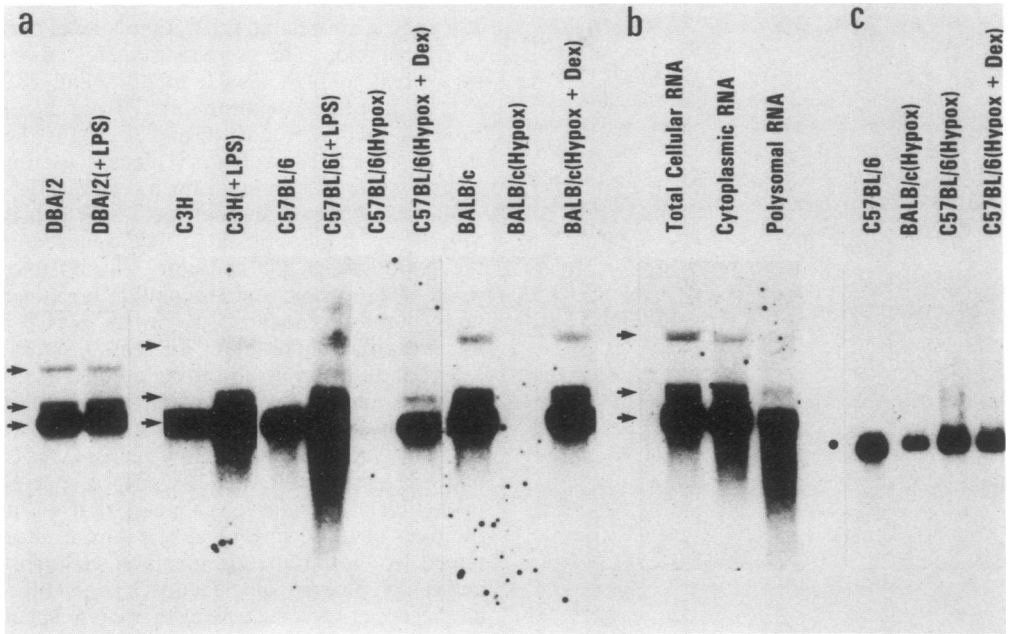

FIG. 5. Analysis of the levels of contrapsin and α_1-PI mRNA species in mouse livers. (a) Levels of contrapsin-related mRNAs in normal and LPS-treated DBA/2, C3H, and C57BL/6 mice and C57BL/6 and BALB/c mice that have been hypophysectomized (hypox) and left untreated or treated with dexamethasone (hypox + dex). (b) Fractionation of liver mRNA from BALB/c mice into total cellular RNA, total cytoplasmic RNA, or polysomal RNA. The arrows indicate the 18S, 20S, and 28S contrapsin-related RNA species. (c) Levels of α_1-PI mRNA in the livers of C57BL/6 mice that were normal, hypophysectomized, or hypophysectomized and subsequently treated with dexamethasone. The dot represents the position of the 17S RNA species.

three RNA species) and C57BL/6 (only the 18S RNA species). Each RI strain had either the DBA/2 or the C57BL/6 pattern. The strain distribution pattern of these RNA species (Table 3) is exactly that of the DNA polymorphism. Therefore, the regulatory element is located near or at the structural locus *Spi*-2. We designated this regulatory locus *Spi*-2r and suggest that this genetic element is *cis*-acting.

DISCUSSION

We have shown that there is a complex of clustered genes, designated the *Spi* complex, that code for serine protease inhibitors. The two inhibitors now characterized are α_1-PI and contrapsin. This complex can be separated into two multigene families: *Spi*-1, which contains the gene for α_1-PI, and *Spi*-2, which contains the gene for contrapsin, because (i) the cDNAs for these proteins, although related (59% nucleotide homology), do not crosshybridize under normal stringency conditions; and (ii) one recombination event has been shown between and not within the multigene families. However, these families are closely linked (1 ± 1 cM) and are located on chromosome 12.

Our finding that α_1-PI and contrapsin gene families map close together supports the idea that the precursor to these particular genes arose by a process of duplication from a common ancestral gene. We have predicted from the extent of nucleotide homology between contrapsin and α_1-PI that this occurred some 200 million to 300 million years ago (14). After these genes began to diversify, amplification events may have occurred to produce the two distinct gene families. Alternatively, one series of amplication events from a single primordial gene may have occurred, and the genes located most distantly diverged at different rates but locally remained more closely related by processes such as gene conversion.

In mice, α_1-PI is located approximately 8 to 12 cm from the *Igh*-1 locus on chromosome 12. The human α_1-PI locus has been mapped and assigned to chromosome 14, which is also linked to the immunoglobulin heavy chain locus (9). Thus, these loci on chromosome 12 in mice and chromosome 14 in humans represent conserved linkage groups. We have evidence that the human α_1-achy gene is also located on chromosome 14 (unpublished data). At least two different related α_1-PI genes map to chromosome 14 in humans (18), whereas the number of α_1-achy genes in humans has yet to be determined. It does appear that the number of genes in the two multigene families varies among mice, rats, and humans.

TABLE 3. Distribution of mRNA polymorphism in B×D RI strains[a]

Parameter	Pattern for the following strains:																				
	1	2	5	6	8	9	11	12	13	14	16	19	21	22	23	24	28	29	30	31	32
RNA	B	B	B	D	D	D	D	B	D	B	D	D	B	B	B	D	D	B	D	B	D
DNA	B	B	B		D	D	D	B	D	B	D	D	B	B	B	D	D	B	D	B	D

[a] B and D designate the progenitor strains (see the footnote to Table 1).

There are several possible consequences of the formation of a multigene family. (i) After amplification, one or a few genes remain active and the others eventually acquire mutations to form pseudogenes. (ii) An increase in gene number results in an increase in the level of the gene product. (iii) It allows for diversification in the function and regulation of the original primordial gene. Although we do not yet know the number of active or inactive genes within each family, we have shown that the regulation of expression between, as well as within, these families is quite different. We showed previously that the expression of the α_1-PI- and contrapsin-related mRNAs commences at different times in liver development (3, 21). We show here that these genes also have evolved very different regulatory mechanisms. Whereas the expression of contrapsin-related mRNAs is responsive to inflammation, glucocorticoid steroid hormone analogs, and hypophysectomy, the levels of α_1-PI mRNA(s) are not appreciably affected. The contrapsin gene family has acquired regulatory elements independent of the α_1-PI gene family; this must have occurred recently in evolutionary time. Of particular interest is the difference in regulation within the contrapsin gene family. The inbred strains we studied have acquired different mechanisms for regulating the level of 20S and 28S mRNA species. Inbred strains, represented by C57BL/6 mice, normally express very little of these mRNAs, but the level can be induced rapidly after stimulation of inflammation. However, other inbred strains, represented by DBA mice, constitutively synthesize a high level of 20S an 28S mRNA, comparable to that of the stimulated C57BL/6. Genetic analysis of this regulatory variation in the B×D RI strains allowed us to show that a single genetic locus, which we designated *Spi*-2r, is tightly linked to the structural locus. The data suggest that the regulatory element is *cis*-acting and that the difference is at the level of the primary structure of the gene(s) affected. Although it seems unlikely, we cannot dismiss the possibility that a very tightly linked *trans*-acting element is involved. Analysis of this genetic variation will be important in understanding the response to the inflammatory process of these genes and perhaps other liver genes.

We do not now know the nature of the different contrapsin-related mRNAs. All the transcripts are localized in the cytoplasm, and this suggests that they are not precursor RNAs and are capable of being translated. We do not believe that the higher-molecular-weight species are a result of nuclear leakage during the isolation of cytoplasmic RNA, because the relative ratio of the contrapsin RNA species is not altered when compared with that of the total RNA preparation. We also found multiple contrapsin-related mRNAs in rats. Here, we have evidence that these mRNAs are products of different genes (manuscript in preparation). We are now investigating whether the mRNAs code for different proteins. Interestingly, only a single substitution is required in the reactive center region (at the P_1 amino acid site) of the protein to change the specificity of the inhibitor (8, 22, 24). Therefore, it is possible that the genes of the family encode a set of very closely related protease inhibitors with diverse specificities and functions. We previously showed that mouse contrapsin is homologous to human α_1-achy, but these proteins have diverged significantly within their reactive center regions, resulting in inhibitors with different specificities (14). Contrapsin activity is not detected in human plasma (27), suggesting that the appearance of contrapsin in mice must have occurred at a time in evolutionary history after rodents and primates diverged. We are now trying to determine whether one gene in the mouse multigene family codes for a protein that more closely resembles α_1-achy at the reactive center region.

ACKNOWLEDGMENTS

We thank Benjamin Taylor for the computer analysis of the RI strains, Bill Held for the use of the fractionated reticulocyte lysate translation system, and Frank Berger for many of the RNA preparations. We also thank Mary Bodis for expert technical assistance, Richard Meehan for critical advice, Katie Rae for excellent assistance in the preparation of the manuscript, and Norman Davidson, Douglas Stuart, and Sandy Bruce for expert photographic work.

LITERATURE CITED

1. **Aronsen, K. F., G. Ekelund, C. O. Kindmark, and C. B. Laurell.** 1972. Sequential changes of plasma proteins after surgical trauma. Scand. J. Clin. Lab. Invest. **29**(Suppl. 124):127–136.
2. **Bailey, D. W.** 1971. Recombinant inbred strains, an aid to finding identity, linkage and function of histocompatibility and other genes. Transplantation **11**:325–327.
3. **Barth, R. K., K. W. Gross, L. C. Gremke, and N. D. Hastie.** 1982. Developmentally regulated mRNAs in mouse liver. Proc. Natl. Acad. Sci. U.S.A. **79**:500–504.
4. **Baumann, H., G. P. Jahreis, and K. C. Gaines.** 1983. Synthesis and regulation of acute phase plasma proteins in primary culture of mouse hepatocytes. J. Cell Biol. **97**:866–876.
5. **Bennett, K. L., P. A. Lalley, R. K. Barth, and N. D. Hastie.** 1982. Mapping the structural genes coding for the major urinary proteins in the mouse: combined use of recombinant inbred strains and somatic cell hybrids. Proc. Natl. Acad. Sci. U.S.A. **79**:1220–1224.
6. **Blatt, C., K. Mileham, M. Haas, M. Nesbitt, M. E. Harper, and M. I. Simon.** 1983. Chromosomal mapping of the mink cell focus-inducing and xenotropic *env* gene family in the mouse. Proc. Natl. Acad. Sci. U.S.A. **80**:6288–6302.
7. **Carrell, R., M. Owen, S. Brennan, and L. Vaughn.** 1979. Carboxy terminal fragment of human α-1-antitrypsin from hydroxylamine cleavage: homology with antithrombin III. Biochem. Biophys. Res. Commun. **91**:1032–1037.
8. **Courtney, M., S. Jallat, L. H. Tessier, A. Benavente, R. G. Crystal, and J. P. Lecocq.** 1984. Synthesis in *E. coli* of α-antitrypsin variants of therapeutic potential for emphysema and thrombosis. Nature (London) **313**:149–151.
9. **Cox, D. W., V. D. Markovic, and I. E. Teshima.** 1982. Genes for immunoglobulin heavy chains and for α_1-antitrypsin are localized to specific regions of chromosome 14q. Nature (London) **297**:428–430.
10. **Doolittle, R. F.** 1983. Angiotensinogen is related to the antitrypsin-antithrombin-ovalbumin family. Science **222**:417–419.
11. **Dretzen, G., M. Bellard, P. Sassone-Corsi, and P. Chambon.** 1982. A reliable method for the recovery of DNA fragments for agarose and acrylamide gels. Anal. Biochem. **112**:295–302.
12. **Elliott, R. W., and F. G. Berger.** 1983. DNA sequence polymorphism in an androgen-regulated gene is associated with alteration in the encoded RNAs. Proc. Natl. Acad. Sci. U.S.A. **80**:501–504.
13. **Held, W. A., K. West, and J. F. Gallagher.** 1977. Importance of initiation factor preparations in the translation of reovirus and globin mRNA lacking a 5′-terminal 7-methylguanosine. J. Biol. Chem. **252**:8489–8497.
14. **Hill, R. E., P. H. Shaw, P. A. Boyd, H. Baumann, and N. D. Hastie.** 1984. Plasma protease inhibitors in mouse and man: divergence within the reactive center region. Nature (London) **311**:175–177.
15. **Hunt, L. T., and M. O. Dayhoff.** 1980. A surprising new protein superfamily containing ovalbumin, antithrombin III, and alpha$_1$-proteinase inhibitor. Biochem. Biophys. Res. Commun. **95**: 864–871.
16. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee.** 1981. Dilute (*d*) coat colour mutation of DBA/2J mice is associ-

ated with the site of integration of an ecotropic MuLV genome. Nature (London) **293**:370–374.
17. **Koj, A.** 1984. Acute phase reactants, p. 73–125. *In* A. C. Allison (ed.), Structure and function of plasma proteins, vol. 1. Plenum Publishing Corp., New York.
18. **Lai, E. C., F. T. Kao, M. L. Law, and S. L. Woo.** 1983. Assignment of the α_1-antitrypsin gene and a sequence-related gene to human chromosome 14 by molecular hybridisation. Am. J. Hum. Genet. **35**:385–392.
19. **Leicht, M., G. L. Long, T. Chandra, K. Kurachi, V. J. Kidd, M. Mace, E. W. Davie, and S. L. Woo.** 1982. Sequence homology and structural comparison between the chromosomal human α_1-antitrypsin and chicken ovalbumin gene. Nature (London) **297**:655–659.
20. **Marshall, A. J., and L. A. Burgoyne.** 1976. Interpretation of the properties of chromatin extracts from mammalian nuclei. Nucleic Acids Res. **3**:1101–1110.
21. **Meehan, R. R., D. P. Barlow, R. E. Hill, B. L. M. Hogan, and N. D. Hastie.** 1984. Pattern of serum protein gene expression in mouse visceral yolk sac and fetal liver. EMBO J. **3**:1881–1885.
22. **Owen, M. C., S. O. Brennan, J. H. Lewis, and R. W. Carrell.** 1983. Mutation of antitrypsin to antithrombin: α_1-antitrypsin Pittsburgh (358Met→Arg), a fetal bleeding disorder. N. Engl. J. Med. **309**:694–698.
23. **Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
24. **Rosenburg, S., P. J. Barr, R. Najarian, and R. A. Hallewell.** 1984. Synthesis in yeast of a functional oxidation-resistant mutant of human α_1-antitrypsin. Nature (London) **312**:77–80.
25. **Shore, G. C., and J. R. Tata.** 1977. Two fractions of rough endoplasmic reticulum from rat liver. J. Cell Biol. **72**:726–743.
26. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
27. **Takahara, H., and H. Sinohara.** 1982. Mouse plasma trypsin inhibitors: isolation and characterisation of α_1-antitrypsin and contrapsin, a novel trypsin inhibitor. J. Biol. Chem. **257**: 2438–2446.
28. **Takahara, H., and H. Sinohara.** 1983. Inhibitory spectrum of mouse contrapsin and α_1-antitrypsin against mouse serine protease. J. Biochem. (Tokyo) **93**:1411–1419.
29. **Taylor, B. A.** 1982. Recombinant inbred strains, p. 423–438. *In* H. C. Morse III (ed.), Origins of inbred mice. Academic Press, Inc., New York.