Am. J. Hum. Genet. 55:126–133, 1994

Physical and Genetic Mapping of the Serpin Gene Cluster at 14q32.1: Allelic Association and a Unique Haplotype Associated with α_1-Antitrypsin Deficiency

Barbara C. Byth,[1] Gail D. Billingsley,[1] and Diane W. Cox[1,2]

[1]The Research Institute, The Hospital for Sick Children, and [2]Departments of Paediatrics and of Molecular and Medical Genetics, University of Toronto, Ontario

Summary

The α_1-antitrypsin (PI) gene is part of a cluster of structurally related serine protease inhibitor genes localized at chromosome 14q32.1, a cluster that includes the α_1-antichymotrypsin (AACT), protein C inhibitor (PCI), and corticosteroid-binding globulin (CBG) genes and the α_1-antitrypsin–like pseudogene (PIL). The order of the genes is refined here by genetic mapping using simple tandem repeat polymorphisms (STRPs) and by physical mapping in YACs. The order of the genes is (centromere)–CBG–PIL–PI–PCI–AACT–(telomere). Analysis of DNA haplotypes comprising STRP and RFLP markers in the serpin genes reveals considerable allelic association throughout the cluster. Furthermore, the common α_1-antitrypsin deficiency allele, *PI*Z,* has a unique DNA haplotype at the CBG, PIL, and PI loci, which extends over 60 kb in 97% of cases and in 44% of cases includes the PCI and AACT loci. This unique haplotype will be of use in examining a number of other diseases, particularly those with an inflammatory component, thought to be associated with α_1-antitrypsin deficiency or partial deficiency.

Introduction

α_1-Antitrypsin (α1AT; also called "α_1-protease inhibitor") is the most abundant component of a family of serine protease inhibitors (serpins) in human plasma. Deficiency of α1AT, inherited as an autosomal recessive trait, is one of the most common genetic abnormalities of Caucasians, occurring at a frequency of 1 in 7,000 to 1 in 2,000. The deficiency is associated with an increased risk for obstructive lung disease in adults and with liver disease in children (reviewed in Cox, in press). The α1AT gene (PI locus) is located on chromosome 14, within a cluster of related serpin genes: α_1-antichymotrypsin (AACT), protein C inhibitor (PCI), corticosteroid-binding globulin (CBG), and the α1AT-like pseudogene sequence (PIL). These genes occupy a 280-kb region of chromosome 14q32.1 (Sefton et al. 1990; Billingsley et al. 1993). We report here (*a*) the development of simple tandem repeat polymorphisms (STRPs) for each of the serpin genes and (*b*) their use in physical mapping and genetic linkage studies.

α1AT deficiency, usually caused by homozygosity for the common deficiency allele *PI*Z,* is known to be associated with a number of other diseases, including rheumatoid arthritis, panniculitis, asthma, certain forms of cancer, and immune complex diseases (Byth and Cox, submitted). It is possible that the observed associations arise as a result of allelic association between the PI locus and a neighboring disease-causing or disease-susceptibility locus. Alternatively, the PI locus itself may act as a modifying locus in these diseases, which are likely to be of multifactorial etiology. Assessment of the degree of allelic association present throughout the serpin gene cluster allows the development of a strategy for studying these potential disease associations by using haplotype analysis.

Material and Methods

Development and Analysis of STRPs

The development and analysis of the AACT, PI, and PCI STRPs has been described elsewhere (Byth and Cox 1993*a*, 1993*b*; Byth et al. 1993); their characteristics are described in table 1. Bacteriophage clones containing the CBG gene were subcloned into the *Bam*HI site of pBluescript SK(+) vector (Stratagene) by using the restriction enzyme *Sau*3AI. The subclones were screened with a selection of five trimer and five tetramer repeat oligonucleotides; positively hybridizing clones were selected for DNA sequencing. The oligonucleotide sequences used were as follows: $(AAAT)_5$, $(AAAC)_5$, $(AAAG)_5$, $(AATG)_5$, $(AAGG)_5$, $(CCG)_7$, $(AGC)_7$, $(AAC)_7$, $(AAT)_7$, and $(AGG)_7$. A tetramer repeat motif of sequence (AAAC) was detected (table 1). Oligonucleotide primers designed to flank the repeat were used in the PCR, under the following conditions: 50 Mm KCl; 10 mM Tris-Cl pH 8.4; 2 mM $MgCl_2$; 200 M each dCTP, dATP, dGTP, and dTTP; 50 ng genomic DNA, and

Received November 10, 1993; accepted for publication March 17, 1994.

Address for correspondence and reprints: Dr. D. W. Cox, Department of Genetics, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada.

0002-9297/94/5501-0018$02.00

Table 1

Characteristics of the Serpin STRP Markers

Gene	Motif	No. of Alleles	Size Range (bp)	PIC	Primer Sequences[a]
CBG	$(AAAC)_n$	2	186–190	.30	ACTCCAGCTTGGGCAACAAGA CTGTGCATTTTTATATGGCTGGG
PI	$(CA)_n$	18	155–189	.85	TTGCAGGGAGTCAGGTGTATG GCATCACACAGAGACACGGAT
PCI	$(CA)_n$	15	128–156	.79	CATCTGAGGCCTCCATATTCTC CCTCTGCACCTGGTAATTAGCC
AACT	$(TA)_n(GA)_n$	17	141–173	.68	ATACATATATACGTATACACAT GTAAGCAATGAGCAATGATT

[a] Sequences are written in the 5′ to 3′ direction.

0.5 unit Amplitaq (Perkin Elmer), per 20-μl reaction. Thirty cycles were performed, each of 30 s at 94°C (denaturing), 60°C (annealing), and 72°C (extension). The PCR products were analyzed by electrophoresis on 0.75-mm thick, 20-cm × 20-cm, 12% nondenaturing polyacrylamide gels and were visualized by silver staining (using the BioRad Silver Stain Kit according to the manufacturer's recommended conditions; fig. 1). There are two alleles—one of 90 bp (CBG-1) and one of 86 bp (CBG-2)—corresponding to $(AAAC)_6$ and $(AAAC)_5$, respectively; the allele frequencies (calculated from 392 unrelated Caucasians) are .2 and .8, respectively. Forty reference pedigrees from the CEPH database were typed for the PCI and CBG STRPs.

Markers from the 14q32.1 Region Used in the Linkage Analysis

In addition to the STRPs described above, a number of two-allele RFLPs were used in the linkage analysis. Both a *Bst*EII RFLP in exon III of the PI gene and a *Bgl*II RFLP 5′ of exon III of the PIL gene were analyzed by hybridizing Southern blots of genomic DNA, double-digested with *Bst*EII and *Bgl*II, with a PI genomic probe, as described elsewhere (Cox et al. 1987). Typings for a number of other STRP markers (D14S48, D14S51, D14S55, and D14S45) and for a *Taq*I RFLP in exon III of the AACT gene were obtained from the CEPH database (version 6.0).

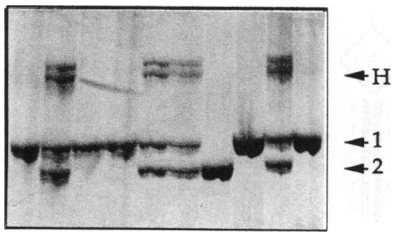

Figure 1 Silver-stained polyacrylamide gel showing the CBG $(AAAC)_n$ STRP. Alleles 1 (90 bp) and 2 (86 bp) are marked. H = heteroduplexes seen in heterozygous DNA samples.

Statistical Analyses

Linkage analysis was performed using the LINKAGE programs MLINK and CLODSCORE (Lathrop and Lalouel 1984). χ^2 and P values were calculated using the Contingency Tables option of the Utility Programs for the Analysis of Genetic Linkage, provided by J. Ott (Columbia University, New York). For the large tables generated using the multiallelic markers, data for rare alleles were combined so that no cell had an expected value <5 (Sherrington et al. 1991; Lerner et al. 1994). D' was calculated using the formula $D' = D/D_{max}$ (Lewontin 1964), where $D = h(ab) - P(a)P(a)$, h is the frequency of the haplotype containing the two rare alleles, and $P(a)$ and $P(b)$ are the frequencies of these rare alleles. If $D < 0$, then D_{max} is the lesser of $P(a)P(b)$ and $P(A)P(B)$; if $D > 0$, then D_{max} is the lesser of $P(a)P(B)$ and $P(A)P(b)$. $P(A)$ and $P(B)$ are the frequencies of the common alleles. D' varies between +1 and −1; a positive value is taken to indicate the association of the rare alleles at the two loci, and a negative value is taken to indicate the association of the rare allele at one locus with the common allele at the other.

Results

Linkage Analysis Using STRPs

The maximum two-point LOD scores and recombination values obtained between the serpin STRPs and flanking markers are shown in table 2. The confidence intervals (Conneally et al. 1985) for the genetic distances between the serpin gene pairs are also shown in table 2. The close linkage of the four serpin genes is as expected from their physical proximity.

The PI and AACT genes had previously been oriented relative to the centromere, by linkage analysis (Billingsley et al. 1993, and submitted), with PI centromeric to AACT. Physical mapping by pulsed-field gel electrophoresis ordered the CBG, PIL, and PCI genes relative to the PI and AACT genes, to give the following order: (centromere)–CBG–PIL–PI–(PCI/AACT)–(telomere) (Billingsley et al.

Table 2

Maximum LOD Scores (below the Diagonal) and Recombination Fractions (above the Diagonal) between the Serpin STRPs and Flanking 14q32.1 Markers

	D14S55	D14S48	CBG	PI	PCI	AACT	D14S45	D14S51
D14S55		.010	.200	.217	.161	.110	.205	.219
D14S48	52.13		.119	.153	.124	.132	.170	.172
CBG	3.21	9.49		.013	.045	.039	.023	.051
PI	7.00	15.41	20.24		.017	.035	.099	.114
PCI	20.11	44.26	17.60	40.84		.000	.075	.058
AACT	8.77	13.13	5.68	12.18	39.56		.067	.049
D14S45	8.01	20.33	10.58	16.03	45.59	13.77		.023
D14S51	13.14	33.00	17.69	24.47	68.24	22.63	51.25	

NOTE.—Confidence intervals for recombination fractions for selected marker pairs are as follows: PI-AACT, .008–.106; PI-CBG, .000–.058; AACT-CBG, .005–.135; PI-PCI, .003–.044; AACT-PCI, .000–.014; and CBG-PCI, .012–.118.

1993). The PCI and AACT genes could not be oriented, because of their physical proximity. Linkage analysis in CEPH reference pedigrees, using the serpin STRPs, identifies recombination events that can be used to order the genes (fig. 2*a–c*). Three recombination events, between the PI and PCI-AACT genes in individuals 1345-03 (fig 2*a*), 1345-05 (fig. 2*a*), and 1416-03 (fig. 2*b*), position PCI and AACT distal to PI. In pedigree 1416, individual 1416-05 also has a recombinant, either between D14S48 and PI/PIL or between PI/PIL and the PI STRP; the former possibility is more likely, as there is ~15% recombination between D14S48 and PI (*Bst*EII) (see table 2). CBG was uninformative in this family. A single recombination event was identified in CEPH individual 1424-08, between CBG and PI/PIL (fig. 2*c*). This recombination event confirms the position of CBG proximal to PIL and PI. This individual also has a recombinant between the PCI and AACT STRPs, positioning AACT distal to PCI and the other serpin genes. Individual 1424-07 has a D14S48–CBG recombinant on the other chromosome, placing D14S48 proximal to the serpin cluster.

Physical Mapping of PCI and AACT

Physical mapping of the PCI and AACT genes was used to confirm the order given by analysis of recombinant individuals. The analysis of cosmids containing the two genes positions them 35 kb apart (data not shown). Pooled YAC DNA from the CEPH YAC library (Albertsen et al. 1990) was screened by the PCR using the primer pair for the PI STRP. Secondary screening was performed by D. Le

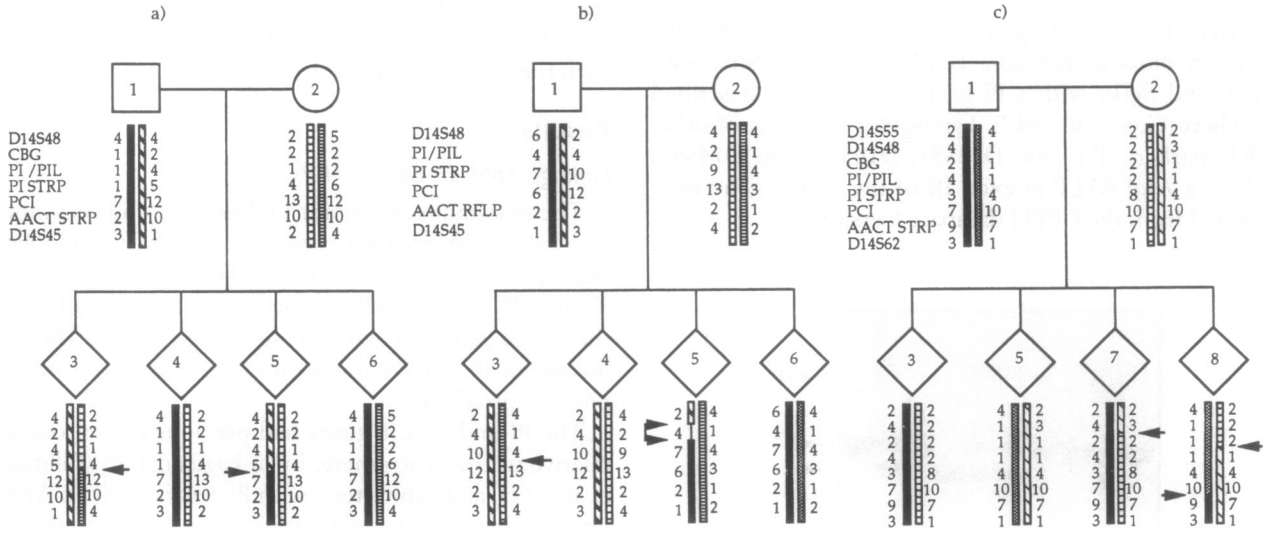

Figure 2 Recombination events within the serpin gene cluster in CEPH families. In each case only a portion of the sibship is shown. Markers are as in the text, except for PI/PIL, which is a haplotype of the PI *Bst*EII RFLP and the PIL *Bgl*II RFLP (Cox et al. 1987). *a*, CEPH 1345. *b*, CEPH 1416. *c*, CEPH 1424.

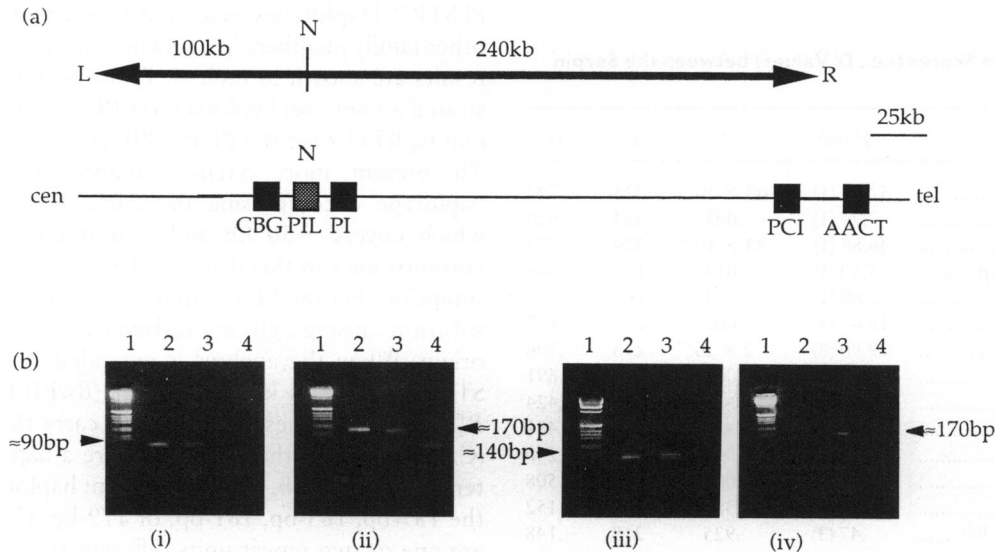

Figure 3 Physical ordering of the PCI and AACT genes. *a*, Simplified map of the YAC CEPH 350A11. N = *Not*I site in the PIL gene; and L and R = left and right arms, respectively, of the YAC. *b*, PCR analysis of YAC CEPH 350A11, by using the STRP primers for CBG (i), PI (ii), PCI (iii), and AACT (iv). In each panel, lane 1 is the size marker (kilobase ladder), lane 2 is YAC DNA, lane 3 is the genomic DNA control, and lane 4 is the negative (i.e., no-template) control. The size of the amplimer is indicated in each case.

Paslier (CEPH, Paris). Two YAC clones containing the PI, PIL, and CBG genes were identified. One of these YACs, CEPH 230C2, was shown to be chimeric by hybridization of end probes to DNA from a somatic cell hybrid (WEGROTH-B3) containing chromosome 14 as its only human component. The other YAC, CEPH 350A11, was mapped using restriction enzymes that cut infrequently in the human genome and was oriented relative to the serpin gene cluster by utilizing a *Not*I site in the PIL gene (a partial map is shown in fig. 3*a*). The genes present on this YAC were characterized by using the STRP primers: PCR analysis showed that the YAC contains the CBG, PI, and PCI genes but not the AACT gene (fig. 3*b*). Although it is possible that the YAC contains a small deletion encompassing the AACT gene, this result together with the genetic data presented above positions AACT distal to PCI. No other YACs containing both the PI and PCI genes have been identified.

Analysis of Allelic Association

In order to assess the extent of the allelic association within the serpin gene cluster, haplotypes were constructed by typing parents and grandparents from the CEPH reference pedigrees and from our collection of families with α1AT deficiency, providing a maximum of 325 haplotypes. The haplotype consisted of the four serpin STRPs, the PI *Bst*EII RFLP, the PIL *Bgl*II RFLP, and the AACT *Taq*I RFLP (fig. 4). The families used are of Caucasian, northern European origin; therefore the effect of racial admixture on the allelic association values can be considered to be negligible.

The D' measure (Lewontin 1964) was used to calculate allelic association, as this statistic maximizes the detectable association and is largely independent of allele frequency (Hedrick 1988; Thompson et al. 1988). Multiallelic data were combined so that no fewer than five observations were expected in any cell. These manipulations affected the magnitude of the probability values but did not affect either their overall significance or the D' values. For calculations of D' from the multiallelic data, a computer program was used that performed pairwise comparisons of each allele with the sum of the remaining alleles, thereby reducing the large tables in each case to a two-by-two table. The two-by-two table that gave the highest χ^2 value was then used to calculate the corresponding D' value. These D' values are presented in table 3.

Significant allelic association (defined as a $P<.05$) was observed within and between the members of the CBG–PIL–PI cluster and the PCI–AACT cluster. The rare CBG allele (CBG-1) is strongly associated with the common PI *Bst*EII allele (the "+" allele; $D' = -.620$) and with the rare PIL allele (the "−" allele; $D' = .774$) and shows a weaker association with the PI STRP ($D' = .177$). The rare PI *Bst*EII allele and the common PIL allele are strongly associated with each other ($D' = -.792$). The PI STRP is associated with the common PI *Bst*EII allele ($D' = -.691$) and with the rare PIL allele ($D' = .706$). Within the PCI-AACT cluster, there is significant association between the PCI STRP and both the rare AACT *Taq*I allele (the "−" allele; $D' = .424$) and the AACT STRP ($D' = .481$). There is also significant association between the two clusters: the rare

Table 3

Allelic Association Scores (i.e., *D'* Values) between the Serpin Genes

Locus Pair	χ^2 (df)	*P*	*n*[a]	*D'*
PI *Bst*EII-PIL	52.98 (1)	$\ll 1 \times 10^{-6}$	325	−.792
CBG-PI *Bst*EII	7.94 (1)	.005	324	−.620
CBG-PIL	46.88 (1)	$\ll 1 \times 10^{-6}$	324	.774
PI*Bst*EII-AACT *Taq*I	5.50 (1)	.014	155	−.456
PIL-AACT *Taq*I	2.80 (1)	.094	167	.177
CBG-PI STRP	16.61 (5)	.005	215	.177
PIL-PI STRP	29.90 (8)	2×10^{-4}	218	.706
PI*Bst*EII-PI STRP	20.24 (6)	.003	226	−.691
PCI-AACT *Taq*I	34.27 (5)	2×10^{-6}	168	.424
PCI-AACT STRP	66.38 (12)	$\ll 1 \times 10^{-6}$	219	.481
PI *Bst*EII-PCI	16.39 (5)	.006	238	.240
PIL-PCI	19.90 (7)	.006	236	−.508
PIL-AACT STRP	4.19 (4)	.381	212	−.152
PI *Bst*EII-AACT STRP	.47 (3)	.925	218	−.148
CBG-PCI	2.84 (4)	.584	186	−.600
CBG-AACT STRP	2.79 (4)	.424	206	.090

[a] No. of haplotypes analyzed.

PI *Bst*EII allele is associated with the common AACT *Taq*I allele ($D' = -.456$) and with the PCI STRP ($D' = .240$). The common PIL allele is also associated with the PCI STRP ($D' = -.508$).

To estimate the age of the CBG polymorphism (which might affect the observed allelic association), 18 individuals of Asian origin were typed for the STRP. The rarer of the two CBG alleles, CBG-1, was not observed in these 36 chromosomes, implying that the CBG-1 allele either is absent in Asians or is present at a frequency significantly lower than that in Caucasians (allele frequency <.03 in Asians, compared with .2 in Caucasians; $P<.01$). The CBG-1 allele may have arisen later than the *Bst*EII and *Bgl*II RFLPs, both of which are present in Asians.

*Construction of Haplotypes Associated with the PI*Z Allele*

The series of families with α1AT deficiency contains 90 independent *PI*Z* chromosomes. These were excluded from the data set used for the calculation of allelic association. The PI type ZZ individuals were all ascertained through the presence of obstructive lung disease or liver disease. The allele frequencies of the CBG and PI STRPs on *PI*Z* chromosomes were compared with those present in the non-*PI*Z* chromosomes (taken from the α1AT-deficiency families and the CEPH reference pedigrees). The results are shown in figure *5a* and *b;* in both cases the differences in allele frequencies between the non-*PI*Z* and the *PI*Z* chromosomes are highly significant ($P \ll 1 \times 10^{-6}$).

All 90 *PI*Z* chromosomes were typed for the PI and PIL RFLPs and for the CBG STRP; 62 were also typed for the PI STRP. Haplotypes were derived using information from other family members (parents or children). The haplotype results are shown in table 4. Previous work had demonstrated a common haplotype on *PI*Z* chromosomes containing RFLPs for the PI and PIL genes (Cox et al. 1985). The present, more extensive analysis reveals one major haplotype encompassing the CBG, PIL, and PI genes, which covers ~50 kb and is found on 97% of *PI*Z* chromosomes in this data set. These data support the assumption that the *PI*Z* mutation (a single nucleotide substitution causing a glycine-to-lysine transition) had a single origin. When the analysis is extended to include the PI STRP (located 11 kb distal to the *Bst*EII RFLP), 84% of *PI*Z* chromosomes are shown to carry the same 183-bp $(CA)_n$ allele and, therefore, to share a haplotype that extends over ~60 kb. Of the 10 variant haplotypes, 8 include the 187-bp, 185-bp, 181-bp, or 179-bp $(CA)_n$ alleles; they are one or two repeat units different from the "founder" PI 183-bp allele. Eight of the 10 share the same flanking markers (CBG-1, PI *Bst*EII "−," and PIL *Bgl*II "−"). These observations suggest that the *PI*Z* mutation arose on a PI 183-bp background and that in the majority of cases replication slippage, rather than recombination, has generated these variant haplotypes. These observations are similar to those found in an analysis of STRPs associated with the fragile X mutation (Oudet et al. 1993) or with the spinal muscular atrophy gene (Brzustowicz et al. 1993).

A selection of the *PI*Z* chromosomes were further typed for the PCI and AACT STRPs: of the 44 chromosomes fully typed, 18 carried the PCI 132-bp allele and the AACT 155-bp allele: 41% of *PI*Z* chromosomes therefore share an identical haplotype, which extends over 280 kb. These haplotypes were unique to *PI*Z* chromosomes and were not observed on any of the non-*PI*Z* chromosomes studied. The allele frequencies of the PCI and AACT STRP alleles on the *PI*Z* and non-*PI*Z* chromosomes were compared. The differences in allele frequency at the PCI locus were significant ($P = 3.4 \times 10^{-5}$; fig. 6), whereas those at the AACT locus were not significant ($P=.13$; data not

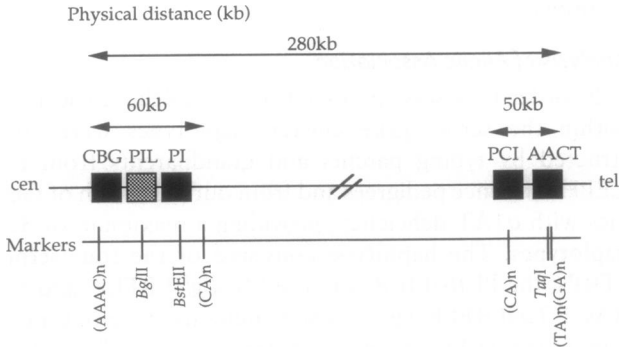

Figure 4 Map of the serpin gene cluster, showing the relative positions of the serpin gene markers used in the association studies.

Table 4

Haplotypes Associated with *PI*Z* Chromosomes

Locus			Size of STRP (bp)				
CBG	PIL[b]	PI*Bst*EII[b]	PI	PCI	AACT	*PI*Z* Frequency[a]	Distance (kb)
Major haplotypes:							
1	−	−				87/90 (97%)	50
1	−	−	183			52/62 (84%)	60
1[c]	−	−	183	132	155	18/44 (41%)	280
Minor haplotypes:							
1	−	−	187			1/62 (1.6%)	60
1	−	−	185	132	155	2/62 (3.2%)	280
1	−	−	181			3/62 (4.8%)	60
1	−	−	179			2/62 (3.2%)	60
2	+	−	169			1/62 (1.6%)	60
2	−	−	181	132	155	1/62 (1.6%)	280

[a] No. of *PI*Z* chromosomes with each haplotype/total no. of *PI*Z* chromosomes.
[b] + = Presence of restriction site; and − = absence of restriction site.
[c] All CBG-PIL-PI haplotypes are shown, whereas only the major CBG-PIL-PI-PCI-AACT haplotype is shown.

shown). This discrepancy is due, in part, to the small numbers studied (48 *PI*Z* chromosomes) but, in part, also occurs because the AACT STRP allele found on the majority of *PI*Z* chromosomes (34 of 48, the 155-bp allele), is also the most common in the total population (97 of 188 chromosomes).

Discussion

The genetic and physical data on the organization of the serpin gene cluster presented here indicate that the order of the genes is as follows: (centromere)–CBG–PIL–PI–PCI–AACT–(telomere). The genetic distance between the genes at the ends of the cluster, CBG and AACT, is ~4 cM (q=.045), within a confidence interval of .005–.135.

There is significant allelic association measurable within the serpin gene cluster. The use of the D' measure of allelic association is preferable to the use of other measures (such as D and r), as it is less dependent on allele frequency (Hedrick 1988; Thompson et al. 1988). This is particularly im-

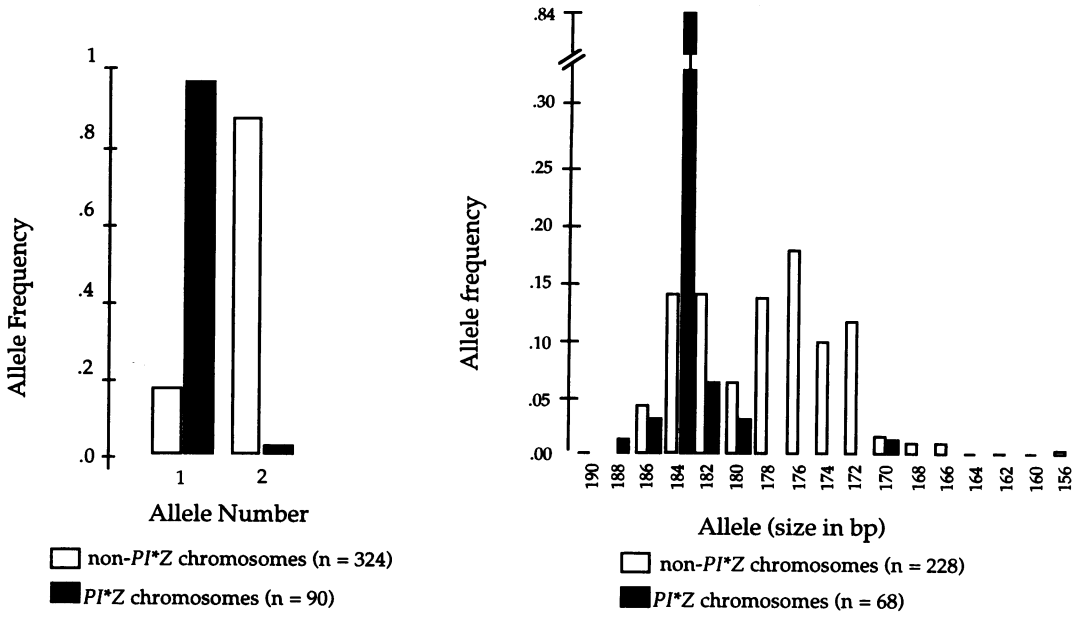

Figure 5 Comparison of PI and CBG STRP allele frequencies on *PI*Z* and non-*PI*Z* chromosomes from the CEPH reference pedigrees and the α1AT families in the present study. *Left,* Frequencies of the CBG STRP alleles. *Right,* Frequencies of the PI STRP alleles. The 160-bp allele does not appear in this data set but is presumed to exist, on the basis of size.

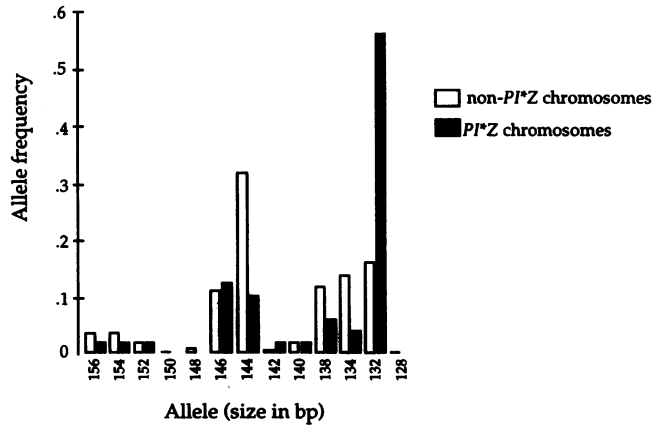

Figure 6 Comparison of PCI STRP allele frequencies on *PI*Z* and non-*PI*Z* chromosomes from the CEPH reference pedigrees and the α1AT families in the present study. The 136-bp and 130-bp alleles are not found in this data set but are presumed to exist, on the basis of size.

portant for the analysis of STRPs, whose individual alleles tend to have frequencies lower than those of typical RFLPs. The technique of deriving *D'* scores from multiallelic data by generating two-by-two tables is similar to those employed in the analysis of allelic associations with disease genes such as those for spinal muscular atrophy (Brzustowicz et al. 1993). The observation of significant allelic association between the two clusters, between the PI and AACT RFLPs, between the PI RFLP and the PCI STRP, and between the PIL RFLP and the PCI STRP differ from those presented by Kelsey et al. (1988), who did not detect association between the PI and AACT genes. However, those investigators used α1AT protein polymorphisms representing six different polymorphic sites within the PI gene, rather than the single *Bst*EII RFLP used here, which probably accounts for the differing results.

The α_1-antitrypsin deficiency allele *PI*Z* is associated with a unique haplotype of 60 kb encompassing the CBG, PIL, and PI genes, supporting the theory of a single origin for the *PI*Z* mutation. In 41% of 44 cases analyzed, the haplotype also extends over 280 kb to include the PCI and AACT genes. There are also significant differences between the allele frequency distribution on *PI*Z* chromosomes and that on non-*PI*Z* chromosomes.

The haplotype analysis also permits an estimate to be made of the age of the *PI*Z* mutation, an estimate based on the assumption of random recombination in a given area. Given that there is 1 crossover event/10^8 bp/generation (Kurnit 1979), the observation of 3.3% recombination (3 recombinant haplotypes, of 90 total) in the ~50-kb region containing the CBG, PI, and PIL RFLPs dates the mutation as having arisen 66 generations, or ~2,000 years, ago. This is slightly less than our previous estimate (Cox et al. 1985) but is in agreement with our suggestion that the value of 216 generations may have been an overestimate.

The identification of the unique and extensive haplotype associated with the *PI*Z* mutation will be of use in the analysis of diseases thought to be associated with α1AT deficiency. A number of diseases are associated with homozygosity or heterozygosity for the *PI*Z* allele; these diseases include panniculitis and immune complex diseases associated with PI type ZZ, as well as rheumatoid arthritis, asthma, and certain forms of cancer with the heterozygous PI types MZ and MS (reviewed in Byth and Cox, submitted). These associations may be a direct result of the protease-protease inhibitor imbalance or may reflect the influences of neighboring disease-causing or susceptibility genes. Of special interest is the proximity of the PI and CBG genes, as the corticosteroid-binding globulin is important in inflammation (Siiteri et al. 1983). The observed associations with PI alleles may be due, in part, to differences in corticosteroid function controlled by alleles at the CBG locus. Comparison of DNA haplotypes in individuals suffering from different manifestations of α1AT deficiency and associated diseases may resolve the involvement of these loci in disease etiology.

Acknowledgments

The authors are grateful to Dr. G. Hammond, University of Western Ontario, Canada, for the gift of the CBG l phage clones; to Dr. D. Le Paslier for supplying the CEPH YAC; and to Richard Wintle and Gordon Thomas, Department of Molecular and Medical Genetics, University of Toronto, for critical reading of the manuscript. The latter also provided the computer program used to analyze the multiallelic data. This work was supported by grant MT 5426 from the Medical Research Council of Canada.

References

Albertsen HM, Abderrahim H, Cann HM, Dausset J, Le Paslier D, Cohen D (1990) Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. Proc Natl Acad Sci USA 87:5019–5113

Billingsley GD, Nguyen V, Cox DW. A contiguous linkage map of chromosome 14 (submitted)

Billingsley GD, Walter MA, Hammond GL, Cox DW (1993) Physical mapping of four serpin genes: α_1-antitrypsin, α_1-antichymotrypsin, corticosteroid-binding globulin, and protein C inhibitor, within a 280-kb region on chromosome 14q32.1. Am J Hum Genet 52:343–353

Brzustowicz LM, Matseoane D, Wang CH, Kleyn PW, Vitale E, Penchaszadeh GK, Hausmanowa-Petrusewicz I, et al (1993) Linkage disequilibrium and haplotype analysis among Polish families with spinal muscular atrophy. Am J Hum Genet Suppl 53:982

Byth BC, Cox DW (1993*a*) A $(CA)_n$ repeat at the 5′ end of the α_1-antitrypsin gene. Hum Mol Genet 2:1752

——— (1993*b*) Two consecutive dinucleotide repeats constitute an informative marker at the α_1-antichymotrypsin (AACT) locus. Hum Mol Genet 2:1085

———. Alpha$_1$-antitrypsin deficiency: associations with inflammatory disease and implications for disease etiology (submitted)

Byth BC, Meijers JCM, Cox DW (1993) A $(CA)_n$ repeat within the protein C inhibitor gene. Hum Mol Genet 2:1752

Conneally PM, Edwards JH, Kidd KK, Lalouel J-L, Morton NE, Ott J, White R (1985) Report of the Committee on Methods of Linkage Analysis and Reporting. Human Gene Mapping 8. Cytogenet Cell Genet 40:356–359

Cox DW. α_1-Antitrypsin deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic basis of inherited disease, 7th ed. McGraw-Hill, New York (in press)

Cox DW, Coulson SE, Billingsley GD (1987) Unique DNA polymorphisms associated with the α_1-antitrypsin (PI) Z deficiency allele: application to prenatal diagnosis. Protides Biol Fluids 35:123–126

Cox DW, Woo SLC, Mansfield T (1985) DNA restriction fragments associated with α_1-antitrypsin indicate a single origin for the deficiency allele PI Z. Nature 316:79–81

Hedrick PW (1988) Inference of recombinational hotspots using gametic linkage disequilibrium values. Heredity 60:435–438

Kelsey GD, Abeliovich D, McMahon CJ, Whitehouse D, Corney G, Povey S, Hopkinson DA, et al (1988) Cloning of the human α_1-antichymotrypsin gene and genetic analysis of the gene in relation to α_1-antitrypsin deficiency. J Med Genet 25:361–368

Kurnit DM (1979) Evolution of sickle variant gene. Lancet 1:104

Lathrop GM, Lalouel JM (1984) Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36: 460–465

Lerner TJ, Boustany R-MN, MacCormack K, Gleitsman J, Schlumpf K, Breakfield XO, Gusella JF, et al (1994) Linkage disequilibrium between the juvenile neuronal ceroid lipofuscinosis gene and marker loci in chromosome 16p12.1. Am J Hum Genet 54:88–94

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49:49–67

Oudet C, Mornet E, Serre JL, Thomas F, Lentes-Zengerling S, Kretz C, Deluchat C, et al (1993) Linkage disequilibrium between the fragile X mutation and two closely linked CA repeats suggests that fragile X chromosomes are derived from a small number of founder chromosomes. Am J Hum Genet 52: 297–304

Sefton L, Kelsey G, Kearney P, Povey S, Wolfe J (1990) A physical map of the human PI and AACT genes. Genomics 7:382–388

Sherrington R, Melmer G, Dixon M, Curtis D, Mankoo B, Kalsi G, Gurling H (1991) Linkage disequilibrium between two highly polymorphic microsatellites. Am J Hum Genet 49:966–971

Siiteri PK, Murai JT, Hammond GL, Nisker JA, Raymoure WJ, Kuhn RW (1982) The serum transport of steroid hormones. Recent Prog Horm Res 38:457–510

Thompson EA, Deeb S, Walker D, Motulsky AG (1988) The detection of linkage disequilibrium between closely linked markers: RFLPs at the AI-CIII apolipoprotein genes. Am J Hum Genet 42:113–124