

## Polymorphism in the Interferon- $\alpha$ Gene Family

Irina Golovleva,<sup>1</sup> Martyna Kandefer-Szerszen,<sup>3</sup> Lars Beckman,<sup>2</sup> and Erik Lundgren<sup>1</sup>

Departments of <sup>1</sup>Cell and Molecular Biology and <sup>2</sup>Medical Genetics, University of Umeå, Umeå, Sweden; and <sup>3</sup>Department of Microbiology, Maria Curie-Skłodowska University, Lublin

### Summary

A pronounced genetic polymorphism of the interferon type I gene family has been assumed on the basis of RFLP analysis of the genomic region as well as the large number of sequences published compared to the number of loci. However, *IFNA2* is the only locus that has been carefully analyzed concerning gene frequency, and only naturally occurring rare alleles have been found. We have extended the studies on a variation of expressed sequences by studying the *IFNA1*, *IFNA2*, *IFNA10*, *IFNA13*, *IFNA14*, and *IFNA17* genes. Genomic white-blood-cell DNA from a population sample of blood donors and from a family material were screened by single-nucleotide primer extension (allele-specific primer extension) of PCR fragments. Because of sequence similarities, in some cases “nested” PCR was used, and, when applicable, restriction analysis or control sequencing was performed. All individuals carried the interferon- $\alpha$  1 and interferon- $\alpha$  13 variants but not the LeIF D variant. At the *IFNA2* and *IFNA14* loci only one sequence variant was found, while in the *IFNA10* and *IFNA17* groups two alleles were detected in each group. The *IFNA10* and *IFNA17* alleles segregated in families and showed a close fit to the Hardy-Weinberg equilibrium. There was a significant linkage disequilibrium between *IFNA10* and *IFNA17* alleles. The fact that the extent of genetic polymorphism was lower than expected suggests that a majority of the previously described gene sequences represent nonpolymorphic rare mutants that may have arisen in tumor cell lines.

### Introduction

Interferon- $\alpha$  (IFN- $\alpha$ ) produced in various types of cells represents a family of proteins with pleiotropic biological functions, including antiviral, antiproliferative, and immunomodulatory abilities (Lengyel 1982; Pestka et al. 1987; Landolfo et al. 1995). Recently, a physical map of the type I interferon gene cluster was constructed

disclosing 26 *IFN* genes, of which several are pseudogenes clustered in a 400-kb region on chromosome 9p (Diaz et al. 1994b). Within the cluster, the single IFN- $\beta$  gene was found to be located at the most telomeric position, while at least 14 different true IFN- $\alpha$  genes were scattered in the region together with some pseudogenes. The biological significance for the existence of the multiple IFN- $\alpha$  genes is not understood. It could be because of the evolutionary accidents of neutral biological significance or to the development of specific functions for each subtype of the IFN- $\alpha$  genes. It has already been shown that several subtypes differ in antiviral activity (Weck et al. 1981) or in binding activity, as in the case of Hu-IFN- $\alpha$ A and Hu-IFN- $\alpha$ B2 (Cleary et al. 1994).

A considerable genetic polymorphism has been assumed to exist among the IFN- $\alpha$  genes. RFLP analysis has revealed polymorphism and a high degree of heterozygosity (Ohlsson et al. 1985). Furthermore, in the latest accepted nomenclature of interferon genes, at least 51 IFN- $\alpha$  sequence variants were listed that were isolated from genomic DNA or from cDNA libraries often derived from tumor cell lines (Diaz et al. 1994a). Conventionally, the variants that have close sequence similarities have been assumed to be allelic variants, although this assumption has been proved in only a few cases (Henco et al. 1985; Diaz et al. 1994a).

In the *IFNA2* locus, three variants have been described, designated as “IFN- $\alpha$ 2,” “IFN- $\alpha$ 2(Arg),” and “IFN-A.” IFN- $\alpha$ 2 was isolated as a cDNA from peripheral blood cells (Streuli et al. 1980), while IFN-A and IFN- $\alpha$ 2(Arg) were cloned from two different tumor cell lines (Goeddel et al. 1980; Maeda et al. 1980; Dworkin-Rastl et al. 1982; Lund et al. 1985). It was shown in several ethnic populations of Caucasian, Arabic, African, Asian, Chinese, and Japanese origin that only IFN- $\alpha$ 2 could be detected (Hosoi et al. 1992; Emanuel and Pestka 1993; Gewert et al. 1993; Crowe et al. 1994; Kaluz et al. 1994). In a study of genomic DNA from pooled white blood cells of >28,000 individuals, D. Testa and collaborators also found that IFN- $\alpha$ 2 was predominant, that IFN- $\alpha$ 2(Arg) was rare, and that IFN-A could not be detected (Dipaola et al. 1994; Liao et al. 1994). A previous suggestion that IFN-A and IFN- $\alpha$ 2(Arg) might represent distinct IFN- $\alpha$  genes was not confirmed in the physical map of the IFN locus (Hotta et al. 1988; Diaz et al. 1994b). The finding of IFN-A

Received December 7, 1995; accepted for publication June 5, 1996.

Address for correspondence and reprints: Dr. Erik Lundgren, Department of Cell and Molecular Biology, University of Umeå, S-901 87 Umeå, Sweden. E-mail: erik.lundgren@cmb.umu.se

© 1996 by The American Society of Human Genetics. All rights reserved.  
0002-9297/96/5903-0012\$02.00

and IFN- $\alpha$ 2(Arg) variants in individuals of African and Afro-Caribbean origin confirms that these variants represent alleles (Gewert et al. 1995).

The rareness of allelic variants of IFN- $\alpha$ 2 was somewhat surprising, in light of the reported degree of heterozygosity and the many known IFN- $\alpha$  sequence variants. Therefore, we extended the analysis of proposed allelic variants to IFNA genes other than *IFNA2*. We have analyzed six IFN- $\alpha$  genes by a sensitive single-nucleotide primer extension method, which allows us to detect single-nucleotide differences in PCR fragments. We have especially looked for the presence of previously described sequence variants assumed to represent allelic products. In our report we use the designation of the latest nomenclature of interferon genes that has been approved by Human Genome Mapping Workshop on Chromosome 9 and the Nomenclature Committee of the International Society of Interferon and Cytokine Research (Diaz et al. 1994a). IFNA designates a gene or a locus, while IFN- $\alpha$  usually refers to a sequence variant.

## Material and Methods

### Populations and DNA Samples

Blood samples were obtained from a random population sample of blood donors from the County of Västerbotten, Sweden, and from family material from the same area, comprising two generations of families with one to three children, originally collected in a study of heredity palmoplantar keratoderma (Lind et al. 1994).

Genomic DNA was prepared from human white blood cells according to the method of Sambrook et al. (1989). DNA concentration was measured by spectrophotometry and adjusted to a concentration of 100 ng/ $\mu$ l.

### Oligonucleotides

The primers were chosen according to known sequences reported by Henco et al. (1985), and the nucleotide positions were designated accordingly (Rychlik et al. 1993). In some cases *Bam*HI or *Pst*I restriction sites for subcloning were included into the primers. The sequences of the primers used for PCR, allele-specific primer extension, and direct sequencing are shown in table 1.

### PCR Amplification

PCR was carried out in a total volume of 50  $\mu$ l in 10 mM Tris-HCl, pH 8.3; 50 mM KCl; 1.5 mM MgCl<sub>2</sub>; 100 ng/ml gelatin; 10 pmol each primer; 8  $\mu$ l of working stock containing 1.25 mM each of dNTPs (dATP, dTTP, dCTP, dGTP); and 2.5 U of *Taq* DNA polymerase (Boehringer-Mannheim). PCR was performed in MicroAmp tubes (Perkin-Elmer) with 500 ng of genomic DNA. Thirty cycles were performed in the Programmable Thermal Controller model PTC-100 (SDS, MJ Re-

search). One cycle consisted of denaturation at 94°C for 1 min, annealing for 1 min at temperatures calculated individually for each separate pair of primers, and extension at 72°C for 1 min. Before the first cycle, the initial denaturation step was 3 min, and in the last cycle the extension was increased to 10 min. Annealing temperature was calculated according to the method of Thein and Wallace (1986). After the end of the last cycle the temperature was held at 4°C. A negative control without DNA was used in all experiments.

The size of the PCR fragments was controlled by electrophoresis in 1% agarose gel (SeaKem) in 1  $\times$  TBE and then visualized with ethidium bromide (0.5  $\mu$ g/ml) under UV light.

### Single-Nucleotide Primer Extension

The experiments were done according to the method of Kuppaswamy et al. (1991). In brief, the PCR products were purified by excision from 0.75% (SeaPlaque) agarose gels. After excision and agarose melting, DNA was purified by the Gene Clean kit (Bio 101), and then DNA was dissolved in 10  $\mu$ l of 1  $\times$  TE buffer.

Each allele-specific single-nucleotide primer extension reaction was carried out in 10  $\mu$ l, containing 1  $\mu$ l PCR product, 1  $\mu$ l of the PCR 10 $\times$  buffer (the same as for PCR amplification), 0.25  $\mu$ M allele-specific primer, 1 U *Taq* DNA polymerase, and 0.1  $\mu$ l of the appropriate  $\alpha$ -<sup>32</sup>P-labeled nucleotide (10 mCi/ml, 3,000 Ci/mmol).

The mixtures were denatured at 94°C for 3 min then subjected to 10 cycles, each consisting of 5 s of annealing (temperature was dependent on used primer), 5 s of extension at 72°C, and 30 s of denaturation at 94°C. After completion, 2  $\mu$ l from each reaction was electrophoresed in a 15% polyacrylamide gel with 7 M Urea in 1  $\times$  TBE, for 2.5 h. The gels were fixed in a solution containing 10:10 methanol:acetic acid and then dried. Autoradiography was done by overlaying Kodak X-AR 5 film and exposing at room temperature overnight. The sequences of the allele-specific extension primers are shown in table 1, and discriminatory positions are shown in table 2.

### Subcloning of the PCR Products and Sequencing

The *IFNA2* PCR fragments were subcloned into pBlue sk+ vector, after cleavage at *Bam*HI and *Pst*I sites introduced into the PCR primers. After transformation of *Escherichia coli* XL Blue white colonies were picked from L-agar plates containing IPTG (60  $\mu$ g/ml), ampicillin (100  $\mu$ g/ml) and X-Gal (40  $\mu$ g/ml). Small-scale plasmid preparations were sequenced using the Sequenase Version 2.0 kit as recommended by the manufacturer (U.S. Biochemical). Sequencing primers used were T3 and T7.

The *IFNA10* and *IFNA17* PCR fragments with single 3'-A-nucleotide overhangs were subcloned into the pT7Blue T-vector (Novagen). Sequencing was per-

**Table 1**  
Sequences of Oligonucleotides for PCR and Allele-Specific Primer Extension

Gene and Primer <sup>a</sup>	Sequence <sup>b</sup>	Sequence Region <sup>c</sup> (nt)	Size of the Amplified Fragment (bp)
<i>IFNA1</i> : <sup>d</sup>			
1P1	5'-GCGGGATCCGGCTCTAAACTCATGTAAAG	731–750	857
1P2	5'-GCGCTGCAGGACAGAATTCATGAAAGCGTG	1549–1569	
1-842	5'-GTATGTTCCCTATTTAAGGCAT	820–841	
1-855	5'-GCTCTAGGTTCTCTGAAGGC	856–876	
1-1341	5'-AGAACTCCCCTGATGAATG	1331–1340	
<i>IFNA2</i> :			
2P1	5'-GCGGGATCCAGTCAAGCTGCTCTGTGGGC	981–1000	589
2P2	5'-GCGCTGCAGGTGAGCTGGCATAACGAATCA	1532–1551	
2-68	5'-GCTCCTGGCACAGATGAGGA	1048–1067	
2-1068	5'-AGCAGGAGAAAAGAGAGATT	1088–1069	
2-1101	5'-CTCCTGCTTGAAGGACAGAC	1081–1110	
<i>IFNA10</i> :			
10P1	5'-GCGGGATCCCTAGCAATATTTGCACAT	908–926	612
10P2	5'-GCGCTGCAGTGCCATGTTGAACCAGTTTT	1501–1520	
10P3	5'-TGATTTCCGAATCCCCCAGG	1102–1121	418
10-991	5'-GGCAGATCACAGCCCAGAGA	992–1111	
10-1265	5'-GTCATTCAGTTGCTGGTAAA	1266–1285	
<i>IFNA14</i> :			
14P1	5'-GCGGGATCCAGTGTACCCTCAATCAACC	885–904	647
14P2	5'-GCGCTGCAGTAGTCAATGAGAATCATTTT	1518–1537	
14-922	5'-ACCAGCCCAGCAGCATCTTC	902–921	
14-1183	5'-TCTGTCTCCATGAGATGAT	1163–1182	
14-1454	5'-GAGCAGAAATCATGAGATCC	1434–1453	
<i>IFNA17</i> :			
17P1	5'-GCGGGATCCGATCTGCCTCAGACCCACAG	1004–1023	548
17P2	5'-GCGCTGCAGCAATCAGGATCATTGCCATG	1514–1533	
17P3	5'-TGACTTTGGACTTCCCCCAGG	1102–1121	432
17-101	5'-CTCCTGCCTGAAGGACA	1081–1100	
17-1101	5'-CCTGGGGAAGTCCAAAGTCA	1102–1121	
17-1453	5'-AAGTTTGTGAAAAAGAGAG	1454–1473	
17-1482	5'-TTCATCCTTCCTCCTAAT	1483–1502	

<sup>a</sup> 1P1, 1P2, 2P1, 2P2, 10P1, 10P2, 10P3, 14P1, 14P2, 17P1, 17P2, and 17P3 are primers for PCR; the rest of the primers are for primer extension. Pairs of primers 10P1-10P2 and 10P3-10P2 allow us to amplify fragments 612 bp and 418 bp, respectively. In the *IFNA17* group, primer pair 17P1-17P2 allows us to amplify a fragment 548 bp, and primer pair 17P2-17P3 allows us to amplify a fragment 432 bp.

<sup>b</sup> Underlined sequences represent extensions of the primers for PCR to incorporate *Bam*HI (1P1, 2P1, 10P1, 14P1, 17P1) or 1 *Pst*I (1P2, 2P2, 10P2, 14P2, 17P2) restriction sites for subcloning.

<sup>c</sup> See Henco et al. (1985).

<sup>d</sup> Primers for PCR amplification and allele-specific primer extension reactions were identical for *IFNA1* and *IFNA13* groups.

formed with Sequenase Version 2.0 using T7 promoter sequencing primer.

Direct sequencing (Meltzer et al. 1993) was applied as indicated in the Results section. The fragments were purified by agarose gel electrophoresis, excision, and treatment by the Gene Clean kit. All sequencing steps were done according to manufacturer's descriptions (Stratagene). In sequencing reaction  $\alpha$ -<sup>35</sup>S-dATP was used.

Cycling of the sequencing reaction was run with the following temperature profile. The initial denaturation at 94°C was 5 min. Then followed 30 cycles each with

denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 1 min. The electrophoresis was run at 60 W. The gel was fixed in a solution containing 10% metanol:10% acetic acid, dried for 2 h and exposed on Kodak X-AR 5 film for 15–36 h.

#### Statistical Analysis

Segregation of the *IFNA17* and *IFNA10* alleles was demonstrated in pedigrees. Frequencies of the *IFNA17* and *IFNA10* alleles were calculated by simple gene count, and agreement with Hardy-Weinberg equilibrium

**Table 2**  
Discriminatory Positions for Variants in the IFN- $\alpha$  Gene Family

Group of Genes	Discriminatory Positions <sup>a</sup>			
<i>IFNA1</i> :	842	1341		
IFN- $\alpha$ 1	T	C		
LeIF-D	T	T		
<i>IFNA13</i> :	842	1341		
IFN- $\alpha$ 13	C	C		
<i>IFNA2</i> :	1068	1101		
IFN- $\alpha$ 2	G	A		
IFN- $\alpha$ 2(ARG)	G	G		
IFN-A	A	A		
<i>IFNA10</i> :	991	1265		
$\Psi$ IFN- $\alpha$ 10	C	A		
$\Psi$ LeIF-L	A	A		
LeIF C	T	C		
<i>IFNA14</i> :	992	1183	1454	
LeIF H	G	G	T	
$\lambda$ 2h	A	G	C	
IFN- $\alpha$ 14	G	G	C	
GX-1	G	C	T	
<i>IFNA17</i> :	1101	1453	1482	
IFN- $\alpha$ 88	A	C	G	
$\lambda$ 2C <sub>1</sub>	C	T	T	
IFN- $\alpha$ T	A	T	T	

<sup>a</sup> Numbers for discriminatory position were taken from Henco et al. (1985).

was tested by the  $\chi^2$  test. Haplotype frequencies were estimated from population data according to the principles outlined by Hill (1974). The genotype frequency of the double heterozygotes was computed from the haplotype frequencies obtained from the unequivocal combinations.

## Results

Six groups of human IFN- $\alpha$  genes were selected for this study, and specific primers were designed to detect 16 different published sequence variants, assigned to these groups and assumed to represent allelic variants. In the first step, PCR primers were designed specifically for the six groups as shown in table 1. Only in the case of *IFNA17* was contamination from another locus observed (see below). The obtained PCR fragments were then tested with the respect to the individual sequence variants assigned to this group by single-nucleotide primer extension, using the primers shown in the table 1. Although the variants are very similar, this method was highly discriminatory, as demonstrated in figure 1. As described below, the results could always be confirmed either by sequencing or by restriction analysis.

### The IFNA1 and IFNA13 Subgroups

The IFN- $\alpha$ 1 gene was originally isolated from chromosomal DNA (Nagata et al. 1980) or as a cDNA clone

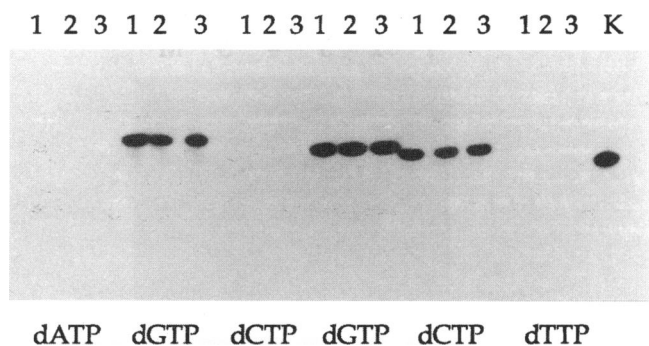
derived from virus-induced leukocyte mRNA (Mantei et al. 1980). The LeIF-D gene was isolated from a cDNA bank derived from the tumor cell line KG1 (Goeddel et al. 1981; Goeddel and Pestka 1982, 1989) and differs from IFN- $\alpha$ 1 in two positions (nucleotide position 1341 and 1502, according to the numbering used by Henco et al. [1985]).

The IFN- $\alpha$ 13 gene was isolated from chromosomal DNA (Todokoro et al. 1984). The entire coding sequence is identical to the sequence of the IFN- $\alpha$ 1 gene, while 22 nucleotide differences have been found in the 5' and 3' noncoding flanking regions. It has been shown that IFN- $\alpha$ 1 and IFN- $\alpha$ 13 represent different genes in spite of having the same coding sequence (Todokoro et al. 1984; Diaz et al. 1994b).

The PCR primers for isolation of both IFN- $\alpha$ 1 and IFN- $\alpha$ 13 were identical, as well as primers for detection of discriminatory positions using the primer extension protocol (tables 1 and 2). In the population sample and in the family material, 857-bp PCR fragments were obtained with the specific primers. It was found that all samples tested ( $n = 69$ ) had T and C in position 842 (according to the numbering of Henco et al. [1985]) and that all had C in position 1341. Thus, we could not detect LeIF D in our study population, while all individuals had both IFN- $\alpha$ 1 and IFN- $\alpha$ 13. This effectively excludes allelism of IFN- $\alpha$ 1 and IFN- $\alpha$ 13, since all 69 individuals could not be heterozygotes. Analysis of the family material also showed that IFN- $\alpha$ 1 and IFN- $\alpha$ 13 were not alleles (data not shown).

### The IFNA17 Group

At least three sequence variants have been assigned to this gene.  $\lambda$ 2c<sub>1</sub> was isolated from chromosomal DNA



**Figure 1** PAGE analysis of the allele-specific primer extension products for the *IFNA14* group. The number of the individuals is given as 1, 2, or 3. K is the control size marker (kinased oligonucleotide 14-922). Primer extension reaction has been done as described in Material and Methods. To determine variants of IFN- $\alpha$  within the *IFNA14* group, the following primers for discriminatory positions were used: primer 14-922 was used with radioactive dATP and dGTP; primer 14-1183 was used with radioactive dCTP and dGTP; and primer 14-1454 was used with dCTP and dTTP. According to these results, individuals 1-3 have G in position 922, G in position 1183, and C in position 1454. It means that the tested genomic DNA contains the IFN- $\alpha$ 14 variant.

(Lawn et al. 1981), and the coding region is identical to IFN- $\alpha$ 17 (Henco et al. 1985). It differs from IFN- $\alpha$ T isolated from chromosomal DNA (von Gabain et al. 1986) in one position (1101, see table 2) and in three positions from IFN- $\alpha$ 88 (Lund et al. 1985) isolated from a cDNA library derived from Namalwa cells.

The sequence variants in this group have a high degree of homology with IFN- $\alpha$  variants of the *IFNA10* group. Thus, control sequencing of the PCR products (548 bp) obtained using the 17P1 and 17P2 primers showed *IFNA10* sequences. Therefore, "nested" primers were used, and, when advantage was taken of discriminatory differences in the 1101-1120 region, the 17P3 primer (1102 to 1121) was used both for the "nested" PCR and as a sequencing primer.

The fragments (432 bp) obtained with 17P3 and 17P2 primers were digested with the restriction enzyme *HinfI*. The 48-bp, 64-bp, and 110-bp fragments diagnostic of the *IFNA17* group are seen in material from five individuals (fig. 2A). For comparison the cleavage of the *IFNA10* PCR fragments are shown (fig. 2B) with a characteristic 110/112 bp doublet. None of the *IFNA17* "nested" fragments were contaminated with *IFNA10* DNA. This finding was confirmed by sequencing.

We could exclude the presence of  $\lambda$ 2c<sub>1</sub> by primer extension of the first round of amplification (primers 17P1 and 17P2). It was shown by primer extension using the primer 17-1101 that all tested DNA had an A in the position 1101 (see table 2). Discrimination between IFN- $\alpha$ 88 and IFN- $\alpha$ 17 was done probing positions 1453

and 1482 by allele-specific primers. Thus, 126 individuals in the population sample carried IFN- $\alpha$ T, 29 had IFN- $\alpha$ 88 and IFN- $\alpha$ T, and 2 had only IFN- $\alpha$ 88.

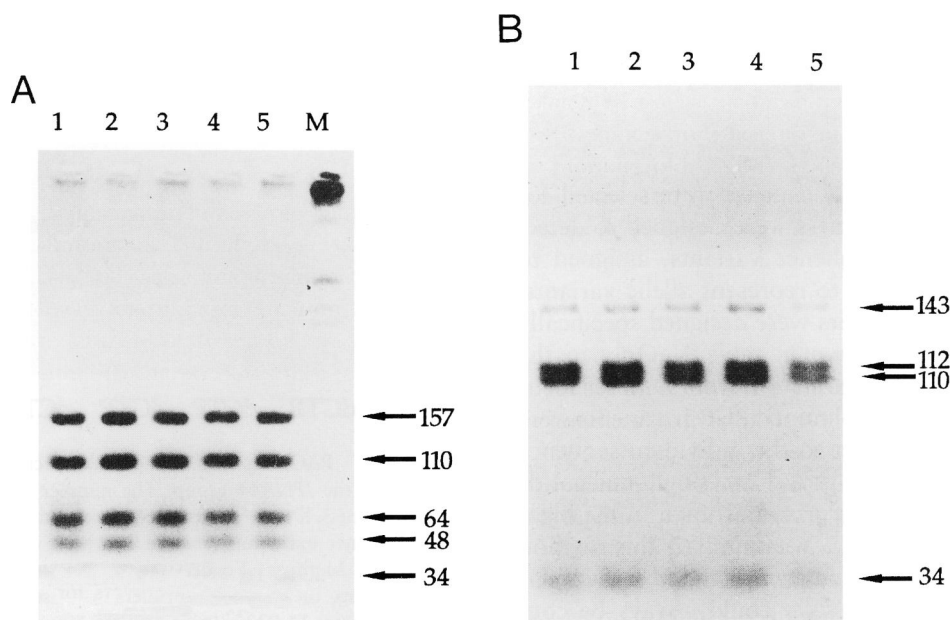
The results were confirmed by restriction analysis with *SspI*. Figure 3 shows an example of one individual with a 380-bp fragment (homozygous for IFN- $\alpha$ T) and one individual with two fragments of 432 bp and 380 bp (heterozygous for IFN- $\alpha$ T and IFN- $\alpha$ 88).

Direct sequencing of a 548-bp fragment covered the discriminatory positions 1453 and 1482 for the three *IFNA17* gene variants  $\alpha$ 88,  $\alpha$ T, and  $\lambda$ 2c<sub>1</sub>. All the tested sequences had T in both positions. In the cases classified as heterozygotes, T and C were found in position 1453 and G and T in position 1482 as expected. The *IFNA17* allele and genotype frequencies are shown in table 3.

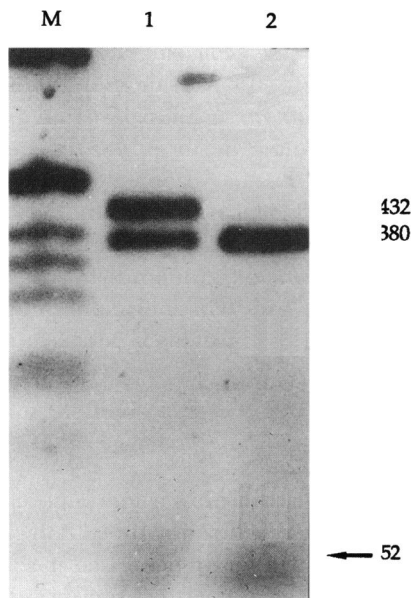
### The *IFNA10* Subgroup

LeIF C was first isolated as cDNA from the myeloid cell line KG1 (Goeddel et al. 1981; Goeddel and Pestka 1982, 1989).  $\Psi$ LeIF-L and  $\Psi$ IFN- $\alpha$ 10 represent DNA sequences in clones  $\lambda$ chr-10 (Brack et al. 1981; Henco et al. 1985). LeIFC is very similar to  $\Psi$ IFN- $\alpha$ 10 and  $\Psi$ LeIF-L. It differs only in four positions, suggesting that it can be allelic to the pseudogenes (Henco et al. 1985).

The same approach was used as for the analysis of the *IFNA17* subgroup. PCR fragments were amplified with primers 10P1 and 10P2 (612 bp, table 1) and then with "nested" primer 10P3 and 10P2 (418-bp fragment). By primer extension it was found that 52 individuals had C in the position 1265 (homozygous for



**Figure 2** Restriction endonuclease analysis of the PCR fragments for the *IFNA10* and *IFNA17* groups. PCR fragments after amplification with "nested" primers for the *IFNA10* group (10P3 and 10P2) and the *IFNA17* group (17P3 and 17P2) were digested with *HinfI* and run on a 4% NuSieve gel. A, Five individuals, showing the 48-bp, 64-bp, and 110-bp fragments diagnostic of the *IFNA17*. M represents DNA molecular-weight marker X (Boehringer Mannheim). B, Five individuals, showing a 110/112-bp doublet characteristic of the *IFNA10*.



**Figure 3** Restriction endonuclease analysis of the PCR fragments in the *IFNA17* group. PCR fragments after amplification with primers for the *IFNA17* group (17P3 and 17P2) were digested with *SspI* and run on a 2% Nusieve gel. Lanes 1 and 2 represent the DNA for individuals 1 and 2. M represents DNA molecular-weight marker X (Boehringer Mannheim). Individual 1 has two variants of *IFNA17*, IFN- $\alpha$ T and IFN- $\alpha$ 88. Individual 2 has one variant, IFN- $\alpha$ -T.

LeIFC), and 11 had both C and A. This finding means that they are heterozygous and carry either  $\Psi$ IFN- $\alpha$ 10 or  $\Psi$ LeIF-L together with LeIF C. We did not further analyze which of these two pseudogenes that were present and designated them  $\Psi$ 10 in this paper.

After purification, the fragments were subjected to direct sequencing with primer 10P3, which confirmed the results from primer extension. Table 4 shows the *IFNA10* genotype and allele frequencies.

#### The *IFNA14* Subgroup

IFN- $\alpha$ 14 and  $\lambda$ 2h are chromosomal clones from Maniatis gene bank and are closely related. They differ in seven positions in the 5' flanking region, two positions in the 5' noncoding region, and in one in the 3' noncoding

**Table 3**

#### *IFNA17* Genotypes and $\alpha$ 88 Allele Frequency in Northern Sweden

	<i>IFNA17</i> GENOTYPES		
	$\alpha$ T/ $\alpha$ T	$\alpha$ T/ $\alpha$ 88	$\alpha$ 88/ $\alpha$ 88
Observed	126	29	2
Expected	125.76	29.51	1.73

NOTE.—The table shows the no. of observed and expected no. of genotypes of 157 individuals (*n*). The frequencies of the  $\alpha$ 88 and  $\alpha$ T alleles were .105 and .895, respectively ( $\chi^2_{HW} = 0.05$ ).

**Table 4**

#### *IFNA10* Genotypes and $\Psi$ 10 Allele Frequency in Northern Sweden

	<i>IFNA10</i> GENOTYPES		
	LeIFC/LeIFC	LeIFC/ $\Psi$ 10	$\Psi$ 10
Observed	52	11	0
Expected	52.52	10.01	.48

NOTE.—The table shows the no. of observed and expected no. of genotypes of 63 individuals. The frequencies of the LeIFC and  $\Psi$ 10 alleles were .087 and .913, respectively ( $\chi^2_{HW} = 0.58$ ).

region. LeIF H is a cDNA clone derived from KG-1 cells (Goeddel et al. 1981; Henco et al. 1985; Lund et al. 1984). GX-1 is probably a product of a recombination event between *IFNA14* and *IFNA5* (Diaz et al. 1994b). The discriminatory positions selected are listed in table 2.

The analysis of genomic DNA from 50 individuals showed that G was present at position 922, G was present at position 1183, and C at position 1454 (fig. 1). Thus, all screened genomic DNA samples contained IFN- $\alpha$ 14, and no other variants were detected by the primer extension method.

To confirm these data, subcloning into the pBlue sk+ vector and sequencing were done. All clones that tested positive had the IFN- $\alpha$  14 sequence.

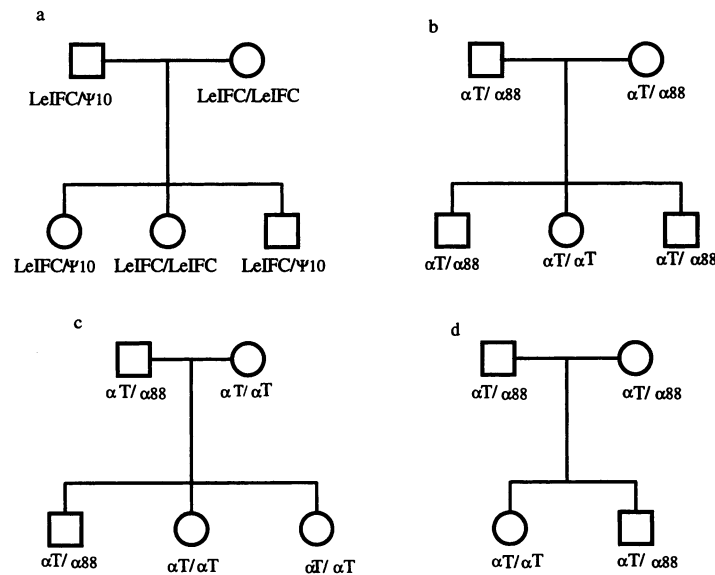
#### The *IFNA2* Subgroup

As expected from results reported elsewhere (Emanuel and Pestka 1993; Gewert et al. 1993; Dipaola et al. 1994; Liao et al. 1994), we found only IFN- $\alpha$ 2 in 45 examined individuals. Control DNA from Namalwa cells could be confirmed to contain both IFN- $\alpha$ 2 and IFN- $\alpha$ 2(Arg).

#### Allele and Haplotype Frequencies

Under the assumption that IFN- $\alpha$ T and IFN- $\alpha$ 88 are alleles, the *IFNA17* genotypes and the frequency of the less frequent allele 88 were calculated (table 3). There was a close agreement between the observed numbers and numbers expected assuming a Hardy-Weinberg equilibrium ( $\chi^2_{HW} = 0.05$ ;  $P = .88$ ). Furthermore, IFN- $\alpha$ T and IFN- $\alpha$ 88 were found to segregate in the families (fig. 4).

Table 4 shows *IFNA10* genotypes and the frequency of the  $\Psi$ 10, calculated on the assumption that LeIFC and  $\Psi$ 10 are alleles. Also, in this case there was a close agreement between observed and expected numbers assuming a Hardy-Weinberg equilibrium ( $\chi^2_{HW} = 0.58$ ;  $P = .45$ ). One back-cross family segregating LeIF C and  $\Psi$ 10 was found (fig. 4). The families were few and contained a small amount of informative haplotypes. Therefore, haplotypes and linkage disequilibrium were calcu-



**Figure 4** Pedigrees showing segregation of *IFNA10* alleles (a) and segregation of *IFNA17* alleles (b–d)

lated from population data. Data on both *IFNA17* and *IFNA10* genotypes were available in 59 individuals. Haplotypes frequencies calculated from this population sample were as follows: LeIFC/*IFN*- $\alpha$ T = .856; *IFN*- $\alpha$ T/ $\Psi$ 10 = .034; *IFN*- $\alpha$ 88/LeIFC = .059; and *IFN*- $\alpha$ 88/ $\Psi$ 10 = .051. Thus, the  $\alpha$ 88 allele was associated with the  $\Psi$ 10 allele and the  $\alpha$ T allele with the LeIFC allele. The linkage disequilibrium was significant ( $\chi^2 = 21.6$ , Yates correction,  $P = 3 \times 10^{-6}$ ) and 55% of  $D_{\max}$ , the maximum possible disequilibrium.

## Discussion

In the human genome, only one *IFN*- $\beta$  gene and one *IFN*- $\gamma$  gene have been identified, and they have been mapped to chromosomes 9 and 12, respectively (Taniguchi et al. 1980; Trent et al. 1982). In contrast, *IFN*- $\alpha$  is a multigene family—14 genes have been found with 80%–95% homology at a nucleotide level (Diaz et al. 1994b). Moreover, *IFN*- $\alpha$ -related sequences have been found, which are reported as pseudogenes, and which hybridize weakly to an *IFNA1* probe (Henco et al. 1985). The diversity in the *IFN*- $\alpha$  gene family appears to be due to both gene duplications and allelic variation (von Gabain et al. 1986). Elsewhere, it was shown that allelic forms of pseudogenes can be functional genes (Bartholomew and Windass 1989). In contrast, only silent nucleotide changes have been found in *IFN*- $\beta$  genes sequenced from different individuals (Taniguchi et al. 1980; Lund et al. 1985). The reason for the evolutionary multiplication of the *IFN*- $\alpha$  genes remains unknown. The latest observations on binding activities of different subtypes of *IFN*- $\alpha$  support the idea on a specific function for each subtypes (Cleary et al. 1994).

Analysis of RFLP in different individuals indicates a pronounced polymorphism and high degree of heterozygosity at the *IFN*- $\alpha$  and - $\beta$  loci (Ohlsson et al. 1985; von Gabain et al. 1986). So far, at least 51 *IFN*- $\alpha$  and *IFN*- $\alpha$ -related pseudogene sequences have been reported from cDNA libraries from both normal cells and tumor cells and from genomic libraries (Goeddel et al. 1980, 1981; Nagata et al. 1980; Goeddel and Pestka 1982; Langer et al. 1984; Lund et al. 1985), while the physical map suggests that there are only 14 gene loci. But so far there has been only limited information as to whether the many sequence variants represent allelic forms present in polymorphic frequencies in populations. Most authors agree that similar sequences represent alleles, and they have therefore been assigned to different genes following these criteria (Henco et al. 1985; Diaz et al. 1994a).

In the present investigation we report results of the analysis of *IFN*- $\alpha$  genes in a northern Swedish population and family material from the same area. We choose six groups of *IFN*- $\alpha$  that comprise 16 sequence variants according to the latest accepted nomenclature of the *IFN*- $\alpha$  genes (Diaz et al. 1994a). By using group specific primers, PCR fragments were obtained, which were probed for the published sequence variants by using a single-nucleotide primer extension method (Kuppussamy et al. 1991).

In four groups, *IFNA1*, *IFNA2*, *IFNA13*, and *IFNA14*, only 4 variants, 1 from each group, were detected, instead of the reported 10. Since the number of alleles examined in those four groups was 138, 90, 138, and 160, respectively, the results indicate that these four loci are nonpolymorphic. The physical map based on a contig of YAC clones shows the nonallelic character of

the *IFNA1* and *IFNA13* genes (Diaz et al. 1994b). *IFNA1* and *IFNA13* are also distinct from the *IFNA2*, *IFNA10*, and *IFNA14* genes in the population studied. Two variants of each of the genes *IFNA10* and *IFNA17* were found. Because of sequence similarities, we had to use “nested” PCR primers to resolve these two groups. The good agreement with Hardy-Weinberg equilibrium and the segregation in families show rather convincingly that IFN- $\alpha$ T and IFN- $\alpha$ 88 are alleles at the *IFNA17* locus and that LeIFC and  $\Psi$ 10 are alleles at the *IFNA10* locus. The linkage disequilibrium between *IFNA10* and *IFNA17* alleles was expected and was in agreement with previous findings of disequilibria between RFLP alleles at the IFN- $\alpha$  and - $\beta$  loci (Ohlsson et al. 1985).

When the previously reported numerous sequence variants are taken into account, the finding of polymorphism at only two of six loci was clearly less than expected. On the other hand, the two polymorphisms reported here together with the five reported before (Ohlsson et al. 1985) can be said to represent a normal level of genetic polymorphism. The major part of the sequence variants reported—for example, the six variants in the *IFNA1*, *IFNA2*, *IFNA13*, and *IFNA14* groups that were not found in the population study—may represent rare (nonpolymorphic) mutants, some of which may have arisen as somatic mutations in cell lines.

Diaz et al. (1994b) pointed out that the whole IFN- $\alpha$  gene cluster probably is a result of repeated gene duplications. Two major duplication units were proposed, of which the *IFNA10* and *IFNA17* loci belonged to the more telomeric group. Some genes in the region joining the proposed duplications units have signs of gene conversions or recombinations.

The genetic polymorphism in the region detected by RFLP markers may indicate a high recombination activity. Our data indicate rather few sequence variants, which suggests selection for sequence conservation. The finding that IFN- $\beta$ , which is functionally homologous and binds to the same receptor as IFN- $\alpha$ , only has silent nucleotide changes supports a notion that there exists a mechanism for sequence conservation in the IFN- $\alpha$  gene locus, in spite of the high recombination frequency.

## Acknowledgments

This study was performed by grants from Lion's Foundation, Umeå University, and BioNative AB, Umeå. J. Wahlquist is acknowledged for skillful introduction of the method used for primer allele-specific primer extension. We are grateful to Dr. A. Lundström, Department of Dermatology, Umeå University, for providing us with blood DNA samples for the population study.

## References

Bartholomew G, Windass JD (1989) Identification of a functional allele of a human interferon- $\alpha$  gene previously characterized as a pseudogene. *J Interferon Res* 9:407–417

- Brack C, Nagata S, Mantei N, Weissmann C (1981) Molecular analysis of the human interferon-alpha gene family. *Gene* 15:379–394
- Cleary CM, Donnelly RJ, Soh J, Mariano TM, Pestka S (1994) Knockout and reconstitution of a functional human type I interferon receptor complex. *J Biol Chem* 269:18747–18749
- Crowe JS, Gewert DR, Barber KA, Lewis AP, Sims MJ, Davies SL, Salom CL, et al (1994) Interferon- $\alpha$ 2 genotype analysis of Chinese chronic hepatitis B patients undergoing recombinant interferon- $\alpha$ 2 therapy. *J Infect Dis* 169:875–878
- Diaz MO, Bohlander S, Allen G (1994a) Nomenclature of the human interferon genes. *J Interferon Res* 14:221–222
- Diaz MO, Pomykala HM, Bohlander SK, Maltepe E, Malik K, Browntein B, Olopade OI (1994b) Structure of the human type-I interferon gene cluster determined from a YAC clone contig. *Genomics* 22:540–552
- Dipaola M, Smith T, Ferencz-Brio K, Liao MJ, Testa D (1994) Interferon-alpha 2 produced by normal human leukocytes is predominantly interferon-alpha 2b. *J Interferon Res* 14:325–332
- Dworkin-Rastl E, Dworkin MB, Swelty P (1982) Molecular cloning of human alpha and beta interferon genes from Namalwa cells. *J Interferon Res* 2:575–585
- Emanuel SL, Pestka S (1993) Human interferon-alpha A, -alpha2, and -alpha 2 (Arg) genes in genomic DNA. *J Biol Chem* 268:12565–12569
- Gewert D, Salom C, Barber K, Macbride S, Cooper H, Lewis A, Wood J, et al (1993) Analysis of interferon-alpha 2 sequences in human genomic DNA. *J Interferon Res* 13:227–231
- Gewert D, Sharp NA, Barber KA, Cooper H, Tucker D, Lewis AP, Thursz M, et al (1995) Detection of rare allelic variants of the interferon-2 gene in human genomic DNA. *J Interferon Res* 15:403–406
- Goeddel DV, Leung DW, Dull TJ, Gross M, Lawn RM, McCandliss R, Seeburg PH, et al (1981) The structure of eight distinct cloned human leukocyte interferon cDNAs. *Nature* 290:20–26
- Goeddel DV, Pestka S (1982) European patent application 81105067.3. Polypeptides, process for their microbial production, intermediates therefore and compositions containing them (microbial production of mature human leukocyte interferons)
- (1989) Microbial production of mature human leukocyte interferon K and L. U.S. Patent 4,801,685 (application 56,623), filed June 1, 1987, issued January 31, 1989; European patent application 82107337.6
- Goeddel DV, Yelverton E, Ullrich A, Heyneker HL, Miozzari G, Holmes W, Seeburg PH, et al (1980) Human leukocyte interferon produced by *E. coli* is biologically active. *Nature* 287:411–416
- Henco K, Brosius J, Fujisawa A, Fujisawa JI, Haynes JR, Hochstadt J, Kovacic T, et al (1985) Structural relationship of human interferon alpha genes and pseudogenes. *J Mol Biol* 185:227–260
- Hill WG (1974) Estimation of linkage disequilibrium in randomly mating populations. *Heredity* 33:229–239
- Hosoi K, Utsumi JL, Kitagawa T, Shimizu H, Kobayashi S (1988) Structural characterization of fibroblast human interferon. *J Interferon Res* 8:375–384



- Hotta K, Monahan J, Collier KJ, Pestka S (1988) Detection of human leukocyte interferon- $\alpha$  A and - $\alpha$  2 genes in genomic DNAs by the use of deoxyoctadecyloligonucleotide probes. *J Interferon Res* 8:51-60
- Kaluz S, Kabat P, Gibadunilova A, Vojtassak J, Fuchsberger N, Kontsek P (1994) Interferon  $\alpha$  2b is the predominant subvariant detected in human genomic DNA. *Acta Virol (Praha)* 38:101-104
- Kuppuswamy MN, Hoffmann JW, Kasper CK, Spitzer S, Groce SL, Bajaj SP (1991) Single nucleotide primer extension to detect genetic diseases: experimental application to hemophilia B (factor IX) and cystic fibrosis genes. *Proc Natl Acad Sci USA* 88:1143-1147
- Landolfo S, Gribaudo G, Angeretti A, Gariglio M (1995) Mechanism of viral inhibition by interferons. *Pharmacol Ther* 65:415-442
- Langer JA, Pestka S (1984) Purification, bacterial expression, and biological activities of the human interferons. *J Invest Dermatol* 83:128s-136s
- Lawn RM, Adelman J, Dull TJ, Gross M, Goeddel D, Ullrich A (1981) DNA sequence of two closely linked human leukocyte interferon genes. *Science* 212:1159-1162
- Lengyel P (1982) Biochemistry of interferons and their actions. *Annu Rev Biochem* 51:251-282
- Liao MJ, Lee N, Dipaola M, Hussain M, Brissette R, Ni D, Smith T, et al (1994) Distribution of interferon- $\alpha$  2 genes in humans. *J Interferon Res* 14:183-185
- Lind L, Lundstrom A, Hofer PA, Holmgren G (1994) The gene for diffuse palmoplantar keratoderma of the type found in northern Sweden localizes to chromosome 12q11-q13. *Hum Mol Genet* 3:1789-1793
- Lund B, Edlund T, Lindenmaier W, Ny T, Collins J, Lundgren E, von Gabain A (1984) Novel cluster of  $\alpha$ -interferon gene sequences in a placental cosmid DNA library. *Proc Natl Acad Sci USA* 81:2435-2439
- Lund B, von Gabain A, Edlund T, Ny T, Lundgren E (1985) Differential expression of interferon genes in a substrain of Namalwa cells. *J Interferon Res* 5:229-238
- Maeda S, McCandliss R, Cross M, Sloma A, Familetti PC, Tabor JM, Evinger M, Levy WP, Pestka S (1980) Construction and Identification of bacterial plasmids containing nucleotide sequence for human leukocyte interferon. *Proc Natl Acad Sci USA* 77:7010-7013
- Mantei N, Schwarzstein M, Streuli M, Panem S, Nagata S, Weissmann C (1980) The nucleotide sequence of a cloned human leukocyte interferon cDNA. *Gene* 10:1-10
- Meltzer SJ (1993) Direct sequencing of polymerase chain reaction products. In: White BA (ed) *PCR protocols. Vol 15 in: Methods in molecular biology.* Humana Press, Totowa, NJ, pp 137-143
- Nagata S, Taira H, Hall A, Johnsrud L, Streuli M, Escodi J, Boll W, et al (1980) Synthesis in *E. coli* of a polypeptide with human leukocyte interferon activity. *Nature* 284:316-320
- Ohlsson M, Feder J, Cavalli-Sforza LL, von Gabain A (1985) Close linkage of  $\alpha$  and  $\beta$  interferons and infrequent duplication of  $\beta$  interferon in humans. *Proc Natl Acad Sci USA* 82:4473-4476
- Pestka S, Langer JA, Zoon KC, Samuel CE (1987) Interferons and their actions. *Annu Rev Biochem* 56:727-777
- Rychlik W (1993) Selection of primers for polymerase chain reaction. In: White BA (ed) *PCR protocols. Vol 15 in: Methods in molecular biology.* Humana Press, Totowa, NJ, pp 31-40
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Streuli M, Nagata S, Weissmann C (1980) At least three human type  $\alpha$  interferons: structure of  $\alpha$  2. *Science* 209:1343-1347
- Taniguchi T, Mantei N, Schwarzstein M, Nagata S, Muramatsu M, Weissmann C (1980) Human leukocyte and fibroblast interferons are structurally related. *Nature* 285:547-549
- Thein SL, Wallace RB (1986) The use of synthetic oligonucleotides as specific probes in the diagnosis of genetic disorders. In: Davis KE (ed) *Human genetic diseases: a practical approach.* IRL Press, Herndon, VA
- Trent JM, Olson S, Lawn RM (1982) Chromosomal localization of human leukocyte, fibroblast and immune interferon genes by means of in situ hybridization. *Proc Natl Acad Sci USA* 79:7809-7813
- Todokoro K, Kioussis D, Weissmann C (1984) Two non-allelic human interferon  $\alpha$  genes with identical coding regions. *EMBO J* 3:1809-1812
- von Gabain A, Ohlsson M, Lindstrom E, Lindstrom M, Lundgren E (1986) Polymorphism and gene duplication in the human IFN- $\alpha$  and - $\beta$  gene family. *Chemica Scripta* 26B:357-362
- Weck PK, Appererson S, May L, Stebbing N (1981) Comparison of the antiviral activities of various cloned human interferon-subtypes in mammalian cell cultures. *J Gen Virol* 57:233-237