NEW POLYNUCLEOTIDES AND POLYPEPTIDES OF THE IFNα-21 GENE

RELATED APPLICATIONS:

The present invention claims priority to French Patent Application 0104404 filed on 30 March 2001 titled <<Nouveaux polynucléotides comportant des polymorphismes de type SNP fonctionnels dans la séquence nucléique du gène IFNalpha21 ainsi que de nouveaux polypeptides codés par ces polynucléotides et leurs utilisations thérapeutiques.>>

BACKGROUND OF THE INVENTION

Field of the Invention.

The present invention relates to new polynucleotides derived from the nucleotide sequence of the IFNα-21 gene comprising new SNPs, and new polypeptides derived from the natural wild-type IFNα-21 protein comprising mutations caused by these SNPs, as well as their therapeutic uses.

Related Art.

The interferon alpha 21 gene, hereinafter referred to as IFNα-21, is described in the publications:

- Goeddel, D. V., Leung, D. W; "The structure of eight distinct cloned human leukocyte interferon cDNAs"; Nature 290 (5801), 20-26 (1981).

- Olopade OI., Bohlander SK.; "Mapping of the shortest region of overlap of deletions of the short arm of chromosome 9 associated with human neoplasia"; Genomics 14 (2), 437-443 (1992).

The nucleotide sequence of this gene is accessible in the HTG section of the GenBank database under accession number AC009445.

The sequence of the messenger RNA of IFNα-21 is mentioned in the database of the NCBI, under accession code NM_002175.

IFNα-21 is a gene possessing a structural and functional homology close to that of human interferons alpha (IFNα), specifically IFNα-2.

The IFNα are known for their cellular antiproliferative effects and their involvements in antiviral and antiparasitic responses.

The IFNα are also known to inhibit the expression of several other cytokines at the level of the hematopoietic stem cells, as well as to inhibit the cellular proliferation of certain

tumors.

The IFNα are also known to reduce the expression of the receptors to the EGF in renal carcinomas, to inhibit the expression of certain mitochondrial genes, to inhibit the proliferation of fibroblasts, monocytes and B lymphocytes, especially in vitro, and to block the synthesis of antibodies by B lymphocytes.

The IFNα are also known to induce the expression of tumor specific antigens on the surface of tumor cells and also to induce the genes placed under the control of promoter regions of the ISRE type (Interferon-Stimulated Response Element) by acting on the specific transcription factors of these ISRE.

It is known that the IFNα are involved in different disorders and/or human diseases, such as the different cancers like for example, carcinomas, melanomas, lymphomas, leukemias and cancers of the liver, neck, head and kidneys, cardiovascular diseases, metabolic diseases such as those that are not connected with the immune system like, for example, obesity, infectious diseases such as hepatitis B and C and AIDS, pneumonias, ulcerative colitis, diseases of the central nervous system like, for example, Alzheimer's disease, schizophrenia and depression, the rejection of tissue or organ grafts, healing of wounds, anemia in dialyzed patients, allergies, asthma, multiple sclerosis, osteoporosis, psoriasis, rheumatoid arthritis, Crohn's disease, autoimmune diseases and disorders, gastrointestinal disorders or even disorders connected with chemotherapy treatments.

The IFNα are particularly used for the treatment of certain leukemias, metastasized renal carcinomas as well as tumors that appear following an immunodeficiency, such as Kaposi's sarcoma in the case of AIDS. The IFNα are also effective against other types of tumors and against certain viral infections. The IFNα are also recognized by the FDA (Food and Drug Administration) for the treatment of genital warts or venereal diseases.

More specifically, IFNα-21 was located by *in situ* hybridization in the brains of patients suffering from Parkinson's disease or Alzheimer's disease.

Compared to other cells, microglial cells express IFNα-21 in large quantities.

In patients suffering from Alzheimer's disease, the presence of IFNα-21 was shown in the neurons of the parietal lobes, suggesting that IFNα-21 may be involved in this pathology (See e.g. Kawaguchi N, Yamada T, Yoshiyama Y. No To Shinkei. 1997 Jan.; 49(1): 69-73).

However, the IFNα, and in particular IFNα-21, have numerous side effects when they are used in pharmaceutical compositions, such as reactions of acute hypersensitivity

(urticaria, bronchoconstriction, anaphylactic shock etc.), cardiac arrythmias, low blood pressure, epileptic seizures, problems with thyroid functions, flu-like syndromes (fevers, sweats, myalgias) etc.

Furthermore, the patients treated with IFNα can develop antibodies neutralizing these molecules, thus decreasing their effectiveness.

The inventors have found new polypeptide and new polynucleotide analogs to the IFNα-21 gene capable of having a different functionality from the natural wild-type IFNα-21 protein.

These new polypeptides and polynucleotides can notably be used to treat or prevent the disorders or diseases previously mentioned and avoid all or part of the disadvantages, which are tied to them.

BRIEF SUMMARY OF THE INVENTION

The invention has as its first object new polynucleotides that differ from the nucleotide sequence of the reference wild-type IFNα-21 gene, in that it comprises one or several SNPs (Single Nucleotide Polymorphism).

The nucleotide sequence SEQ ID NO. 1 of the human reference wild-type IFNα-21 gene is composed of 2001 nucleotides and comprises a coding sequence of 570 nucleotides, from nucleotide 670 (start codon) to nucleotide 1239 (stop codon).

The applicant has identified 8 SNPs in the nucleotide sequence of the reference wild-type IFNα-21 gene. These 8 SNPs are the following: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c.

It is understood, in the sense of the present invention, that the numbering corresponding to the positioning of the SNP previously defined is relative to the numbering of the nucleotide sequence SEQ ID NO. 1.

The letters a, t, c and g correspond respectively to the nitrogenous bases adenine, thymine, cytosine and guanine.

The first letter corresponds to the wild-type nucleotide, whereas the last letter corresponds to the mutated nucleotide.

Thus, for example, the SNP c794g corresponds to a mutation of the nucleotide cytosine (c) at position 794 of the nucleotide sequence SEQ ID NO. 1 of the reference wild-type IFNα-21 gene, into nucleotide guanine (g).

These SNPs were identified by the applicant using the determination process

described in applicant's patent application FR 00 22894, entitled "Process for the determination of one or several functional polymorphism(s) in the nucleotide sequence of a preselected functional candidate gene and its applications" and filed December 6, 2000, cited here by way of reference.

The process described in this patent application permits the identification of one (or several) preexisting SNP(s) in at least one individual from a random population of individuals.

In the scope of the present invention, a fragment of the nucleotide sequence of the IFNα-21 gene, comprising, for example, the coding sequence, was isolated from different individuals in a population of individuals chosen in a random manner.

Sequencing of these fragments was then carried out on certain of these samples having a heteroduplex profile (that is a profile different from that of the reference wild-type IFNα-21 gene sequence) after analysis by DHPLC ("Denaturing-High Performance Liquid Chromatography").

The fragment sequenced in this way was then compared to the nucleotide sequence of the fragment of the reference wild-type IFNα-21 gene and the SNPs in conformity with the invention identified.

Thus, the SNPs are natural and each of them is present in certain individuals of the world population.

The reference wild-type IFNα-21 gene codes for an immature protein of 189 amino acids, corresponding to the amino acid sequence SEQ ID NO. 2, that will be converted to a mature protein of 166 amino acids, by cleavage of the signal peptide that includes the first 23 amino acids.

Each of the coding SNPs of the invention, namely: c794g, c973a, g1011c, t1049a, t1155a, a1204g, causes modifications, at the level of the amino acid sequence, of the protein encoded by the nucleotide sequence of the IFNα-21 gene.

These modifications in the amino acid sequence are the following:

The SNP c794g causes a mutation of the amino acid alanine (A) at position 42 in the immature protein of the IFNα-21 gene, corresponding to the amino acid sequence SEQ ID NO. 2, in glycine (G) and at position 19 of the mature protein. In the description of the present invention, one will call the mutation encoded by this SNP either A19G or A42G according to whether one refers to the mature protein or to the immature protein respectively.

The SNP c973a causes a mutation of the amino acid glutamine (Q) at position 102 in

the immature protein of the IFNα-21 gene, corresponding to the amino acid sequence SEQ ID NO. 2, in lysine (K) and at position 79 of the mature protein. In the description of the present invention, one will call the mutation encoded by this SNP Q79K or Q102K according to whether one refers respectively to the mature protein or to the immature protein.

The SNP g1011c causes a mutation of the amino acid glutamine (Q) at position 114 in the immature protein of the IFNα-21 gene, corresponding to the amino acid sequence SEQ ID NO. 2, in histidine (H) and at position 91 of the mature protein. In the description of the present invention, one will call the mutation encoded by this SNP Q91H or Q114H according to whether one refers respectively to the mature protein or to the immature protein.

The SNP t1049a causes a mutation of the amino acid valine (V) at position 127 in the immature protein of the IFNα-21 gene, corresponding to the amino acid sequence SEQ ID NO. 2, in aspartic acid (D) and at position 104 of the mature protein. In the description of the present invention, one will call the mutation encoded by this SNP V104D or V127D according to whether one refers respectively to the mature protein or to the immature protein.

The SNP t1155a causes a mutation of the amino acid cysteine (C) at position 162 in the immature protein of the IFNα-21 gene, corresponding to the amino acid sequence SEQ ID NO. 2, in stop codon (stop) and at position 139 of the mature protein. In the description of the present invention, one will call the mutation encoded by this SNP C139stop or C162stop according to whether one refers respectively to the mature protein or to the immature protein.

The SNP a1204g causes a mutation of the amino acid lysine (K) at position 179 in the immature protein of the IFNα-21 gene, corresponding to the amino acid sequence SEQ ID NO. 2, in glutamic acid (E) and at position 156 of the mature protein. In the description of the present invention, one will call the mutation encoded by this SNP K156E or K179E according to whether one refers respectively to the mature protein or to the immature protein.

The SNPs c794g, c973a, g1011c, t1049a, t1155a, a1204g cause modifications of the spatial conformation of the polypeptides in conformity with the invention compared to the polypeptide encoded by the nucleotide sequence of the wild-type reference IFNα-21 gene.

These modifications can be observed by computational molecular modeling, according to methods that are well known to a person skilled in the art, making use of, for example, the modeling tools *de novo* (for example, SEQFOLD/MSI), homology (for example, MODELER/MSI), minimization of the force field (for example, DISCOVER, DELPHI/MSI) and/or molecular dynamics (for example, CFF/MSI).

Examples of such models are given hereinafter in the experimental section.

Computational molecular modeling shows that the mutation Q79K on the mature mutated protein involves the displacement of helix C N-end in the wild-type IFNα-21 protein due to hydrogen bonds disturbance as shown in Figures 1A and 1B.

Indeed, hydrogen bonds between the oxygen atom of Q79 residue's side chain, E83 residue's acidic group and helix C of the wild-type IFNα-21 protein disappear in the Q79K mutated IFNα-21 protein.

Thus, the Q79K mutated protein possesses a three-dimensional conformation different from the natural wild-type IFNα-21 protein involving a significant change in its structure and function.

Computational molecular modeling shows that the mutation Q91H on the mature mutated protein involves a displacement of helix C at the mutation position as shown in Figures 2A and 2B. Several hydrogen bonds and salt bridges appear, especially between H91 and D76 amino acids side chains, which make the helix more rigid.

Thus, the Q91H mutated protein possesses a three-dimensional conformation different from the natural wild-type IFNα-21 protein involving a significant change in its structure and function.

Computational molecular modeling shows that the mutation V104D on the mature mutated protein involves modifications in the structure of the loop between helices C and D at the mutation position as shown in Figures 3A and 3B. In the mutated structure several hydrogen bonds appear (between Q102 and G105 on the one hand, and Q52, E107 and T109 on the other hand) which make the loop between helices C and D more rigid. Moreover, a slight displacement of helix B N-end is also observed.

These spatial modifications affect the residues involved in IFNα-21 binding to its receptor.

Thus, the V104D mutated protein possesses a three-dimensional conformation different from the natural wild-type IFNα-21 protein involving a significant change in its structure and function.

Computational molecular modeling shows that the mutation C139stop on the mature mutated protein causes a premature arrest in protein translation leading to the disappearance of a polypeptidic fragment normally involved in helix E in the wild-type IFNα-21 protein, as shown in Figure 4.

Helix E is essential for IFNα-21 binding to its receptor. The absence of helix E causes an incorrect folding of the mutated protein and leads to a modification in the three-

dimensional conformation of the protein in which the hydrophobe core of the protein is in contact with the hydrophilic external medium. Thus, the mutated protein must modify its three-dimensional conformation so as its hydrophobic core is covered with hydrophilic residues in order to avoid contact with hydrophilic external medium.

Thus, the C139stop mutated protein possesses a three-dimensional conformation different from the natural wild-type IFNα-21 protein involving a significant change in its structure and function.

Computational molecular modeling shows that the mutation K156E on the mature mutated protein involves unfolding of helix E C-end and modification of the C-terminus loop shape as shown in Figures 5A and 5B.

This mutation increases the hydrogen bonds network and creates a salt bridge between E156 and R161 residues. These modifications render IFNα-21 protein structure more rigid in this area. This area in the protein is known to be involved in its antiviral activity. Thus, it is possible to predict that the K156E mutated IFNα-21 protein's antiviral activity is dramatically disturbed and that the glutamic acid at position 156 causes a modification in the structure and the function of mature IFNα-21.

Other SNPs in conformity with the invention, namely: t1265c, t1277c do not involve modification of the protein encoded by the nucleotide sequence of the IFNα-21 gene at the level of the amino acid sequence SEQ ID NO. 2. The SNPs t1265c, t1277c are non-coding.

Genotyping of the polynucleotides in conformity with the invention can be carried out in such a fashion as to determine the allelic frequency of these polynucleotides in a population. Examples of genotyping are given, hereinafter, in the experimental section.

The determination of the functionality of the polypeptides of the invention can equally be carried out by a test of their biological activity.

In this regard, it is possible to measure, for example, the anti-proliferative effect on a Daudi cell line of polypeptides in conformity with the invention in comparison with the natural wild-type IFNα-21 protein (Pielher et al. J. Biol. Chem.; Vol. 275, Issue 51, 40425-40433, December 22, 2000; "New structural and Functional Aspects of the type I Interferon-Receptor interaction revealed by comprehensible mutational analysis of the binding interface").

The invention also has for an object the use of polynucleotides and of polypeptides in conformity with the invention as well as of therapeutic molecules obtained and/or identified

starting from these polynucleotides and polypeptides, notably for the prevention and the treatment of certain human disorders and/or diseases.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B represent a model of the encoded protein according to the invention comprising the SNP c973a (Q79K) and the natural wild-type IFNα-21 protein. Figure 1B represents a close up of the model of the inferior part of each one of the proteins represented in Figure 1A.

In Figures 1A and 1B, the black ribbon represents the structure of the natural wild-type IFNα-21 protein and the white ribbon represents the structure of the Q79K mutated IFNα-21 protein.

Figures 2A and 2B represent a model of the encoded protein according to the invention comprising the SNP g1011c (Q91H) and the natural wild-type IFNα-21 protein. Figure 2B represents a close up of the model of the left part of each of the proteins represented on Figure 2A.

In Figures 2A and 2B, the black ribbon represents the structure of the natural wild-type IFNα-21 protein and the white ribbon represents the structure of the Q91H mutated IFNα-21 protein.

Figures 3A and 3B represent a model of the encoded protein according to the invention comprising the SNP t1049a (V104D) and the natural wild-type IFNα-21 protein. Figure 3B represents a close up of the model of the superior part of each of the proteins represented on Figure 3A.

In Figures 3A and 3B, the black ribbon represents the structure of the natural wild-type IFNα-21 protein and the white ribbon represents the structure of the V104D mutated IFNα-21 protein.

Figure 4 represents a model of the encoded protein according to the invention comprising the SNP t1155a (C139stop) (Figure 4B) and the natural wild-type IFNα-21 protein (Figure 4A). Figure 4C represents the superposition of the two proteins of figures 4A and 4B. In Figure 4, the black ribbon represents the structure of the natural wild-type IFNα-21 protein and the white ribbon represents the structure of the C139stop mutated IFNα-21 protein.

Figures 5A and 5B represent a model of the encoded protein according to the invention comprising the SNP a1204g (K156E) and the natural wild-type IFNα-21 protein.

Figure 5B represents a close up of the model of the upper part of each of the proteins represented on Figure 5A. In Figure 5, the black ribbon represents the structure of the natural wild-type IFNα-21 protein and the white ribbon represents the structure of the K156E mutated IFNα-21 protein.

Figure 6 represents the survival rate of mice previously infected by VSV virus and treated with A42G, Q114H/V127D, or K179E mutated IFNα-21 protein, in comparison to those treated with wild-type IFNα-2, or those that have not been treated.

In this figure, the abscissas correspond to the time of survival (days) and the ordinates correspond to the relative survival rate of VSV infected mice. The black triangles, the crosses, the black diamonds represent the data for VSV infected mice treated with A42G mutated IFNα-21, Q114H/V127D mutated IFNα-21, and K179E mutated IFNα-21, respectively. The black squares represent the data for VSV infected mice treated with wild-type IFNα-2, and the open triangles represent the data for VSV infected mice that have not been treated.

DETAILED DESCRIPTION OF THE INVENTION

Definitions.

"Nucleotide sequence of the reference wild-type gene" is understood as the nucleotide sequence SEQ ID NO. 1 of the human IFNα-21 gene.

This sequence is accessible in GenBank under Accession number AC009445 and the sequence of the IFNα-21 messenger RNA is mentioned in the database of the NCBI under accession code NM_002175. Moreover, the human IFNα-21 gene is described in Goeddel, D. V., Leung, D. W "The structure of eight distinct cloned human leukocyte interferon cDNAs"; Nature 290 (5801), 20-26 (1981), and Olopade OI., Bohlander SK."Mapping of the shortest region of overlap of deletions of the short arm of chromosome 9 associated with human neoplasia"; Genomics 14 (2), 437-443 (1992).

"Natural wild-type IFNα-21 protein" is understood as the mature protein encoded by the nucleotide sequence of the reference wild-type IFNα-21 gene. The natural wild-type immature protein IFNα-21 corresponds to the peptide sequence shown in SEQ ID NO. 2.

"Polynucleotide" is understood as a polyribonucleotide or a polydeoxyribonucleotide that can be a modified or non-modified DNA or an RNA.

The term polynucleotide includes, for example, a single strand or double strand DNA, a DNA composed of a mixture of one or several single strand region(s) and of one or several double strand region(s), a single strand or double strand RNA and an RNA composed of a mixture of one or several single strand region(s) and of one or several double strand region(s). The term polynucleotide can also include an RNA and/or a DNA including one or several triple strand regions. By polynucleotide is equally understood the DNAs and RNAs containing one or several bases modified in such a fashion as to have a skeleton modified for reasons of stability or for other reasons. By modified base is understood, for example, the unusual bases such as inosine.

"Polypeptide" is understood as a peptide, an oligopeptide, an oligomer or a protein comprising at least two amino acids joined to each other by a normal or modified peptide bond, such as in the cases of the isosteric peptides, for example.

A polypeptide can be composed of amino acids other than the 20 amino acids defined by the genetic code. A polypeptide can equally be composed of amino acids modified by natural processes, such as post translational maturation processes or by chemical processes, which are well known to a person skilled in the art. Such modifications are fully detailed in the literature. These modifications can appear anywhere in the polypeptide: in the peptide skeleton, in the amino acid chain or even at the carboxy- or amino-terminal ends.

A polypeptide can be branched following an ubiquitination or be cyclic with or without branching. This type of modification can be the result of natural or synthetic post-translational processes that are well known to a person skilled in the art.

For example, polypeptide modifications is understood to include acetylation, acylation, ADP-ribosylation, amidation, covalent fixation of flavine, covalent fixation of heme, covalent fixation of a nucleotide or of a nucleotide derivative, covalent fixation of a lipid or of a lipidic derivative, the covalent fixation of a phosphatidylinositol, covalent or non-covalent cross-linking, cyclization, disulfide bond formation, demethylation, cysteine formation, pyroglutamate formation, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodization, methylation, myristoylation, oxidation, proteolytic processes, phosphorylation, prenylation, racemization, seneloylation, sulfatation, amino acid addition such as arginylation or ubiquitination. Such modifications are fully detailed in the literature: PROTEINS-STRUCTURE AND MOLECULAR PROPERTIES,

2nd Ed., T. E. Creighton, New York, 1993, POST-TRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983, Seifter et al. "Analysis for protein modifications and nonprotein cofactors", Meth. Enzymol. (1990) 182: 626-646, and Rattan et al. "Protein Synthesis: Post-translational Modifications and Aging", Ann NY Acad Sci (1992) 663: 48-62.

"Isolated polynucleotide" or "isolated polypeptide" are understood as a polynucleotide or a polypeptide respectively such as previously defined which is isolated from the human body or otherwise produced by a technical process.

"Identity" is understood as the measurement of nucleotide or polypeptide sequence identity.

Identity is a term well known to a person skilled in the art and well described in the literature. See COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., Ed., Oxford University Press, New York, 1998; BIOCOMPUTING INFORMATICS AND GENOME PROJECT, Smith, D.W., Ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M. and Griffin H.G., Ed, Humana Press, New Jersey, 1994; and SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987.

The methods commonly employed to determine the identity and the similarity between two sequences are equally well described in the literature. See GUIDE TO HUGE COMPUTER, Martin J. Bishop, Ed, Academic Press, San Diego, 1994, and Carillo H. and Lipton D., Siam J Applied Math (1988) 48: 1073.

A polynucleotide having, for example, an identity of at least 95 % with the nucleotide sequence SEQ ID NO. 1 is a polynucleotide which contains at most 5 points of mutation over 100 nucleotides, compared to said sequence.

These points of mutation can be one (or several) substitution(s), addition(s) and/or deletion(s) of one (or several) nucleotide(s).

In the same way, a polypeptide having, for example, an identity of at least 95 % with the amino acid sequence SEQ ID NO. 2 is a polypeptide that contains at most 5 points of mutation over 100 amino acids, compared to said sequence.

These points of mutation can be one (or several) substitution(s), addition(s) and/or deletion(s) of one (or several) amino acid(s).

The polynucleotides and the polypeptides according to the invention which are not totally identical with respectively the nucleotide sequence SEQ ID NO. 1 or the amino acid

sequence SEQ ID NO. 2, it being understood that these sequences contains at least one of the SNPs of the invention, are considered as variants of these sequences.

Usually a polynucleotide according to the invention possesses the same or practically the same biological activity as the nucleotide sequence SEQ ID NO. 1 comprising at least one of the SNPs of the invention.

In similar fashion, usually a polypeptide according to the invention possesses the same or practically the same biological activity as the amino acid sequence SEQ ID NO. 2 comprising at least one of the coding SNPs of the invention.

A variant, according to the invention, can be obtained, for example, by site-directed mutagenesis or by direct synthesis.

By "SNP" is understood any natural variation of a base in a nucleotide sequence. A SNP, on a nucleotide sequence, can be coding, silent or non-coding.

A coding SNP is a polymorphism included in the coding sequence of a nucleotide sequence that involves a modification of an amino acid in the sequence of amino acids encoded by this nucleotide sequence. In this case, the term SNP applies equally, by extension, to a mutation in an amino acid sequence.

A silent SNP is a polymorphism included in the coding sequence of a nucleotide sequence that does not involve a modification of an amino acid in the amino acid sequence encoded by this nucleotide sequence.

A non-coding SNP is a polymorphism included in the non-coding sequence of a nucleotide sequence. This polymorphism can notably be found in an intron, a splicing zone, a transcription promoter or a site enhancer sequence.

By "functional SNP" is understood a SNP, such as previously defined, which is included in a nucleotide sequence or an amino acid sequence, having a functionality.

By "functionality" is understood the biological activity of a polypeptide or of a polynucleotide.

The functionality of a polypeptide or of a polynucleotide according to the invention can consist in a conservation, an augmentation, a reduction or a suppression of the biological activity of the polypeptide encoded by the nucleotide sequence of the wild-type reference gene or of this latter nucleotide sequence.

The functionality of a polypeptide or of a polynucleotide according to the invention can equally consist in a change in the nature of the biological activity of the polypeptide

encoded by the nucleotide sequence of the reference wild-type gene or of this latter nucleotide sequence.

The biological activity can, notably, be linked to the affinity or to the absence of affinity of a polypeptide according to the invention with a receptor.

Polynucleotide

The present invention has for its first object an isolated polynucleotide comprising:

a) a nucleotide sequence having at least 80 % identity, preferably at least 90 % identity, more preferably at least 95 % identity and still more preferably at least 99 % identity with the sequence SEQ ID NO. 1 or its coding sequence (from nucleotide 670 to nucleotide 1239), it being understood that this nucleotide sequence comprises at least one of the following coding SNPs c794g, c973a, g1011c, t1049a, t1155a, a1204g or

b) a nucleotide sequence complementary to a nucleotide sequence under a).

It is understood, in the sense of the present invention, that the numbering corresponds to the positioning of the SNPs in the nucleotide sequence SEQ ID NO. 1.

The present invention relates equally to an isolated polynucleotide comprising:

a) a nucleotide sequence SEQ ID NO. 1 or its coding sequence, it being understood that each of these sequences comprises at least one of the following coding SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, or

b) a nucleotide sequence complementary to a nucleotide sequence under a).

Preferably, the polynucleotide of the invention consists of the sequence SEQ ID NO. 1 or its coding sequence, it being understood that each of these sequences comprises at least one of the following coding SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g.

According to the invention, the polynucleotide previously defined comprises a single coding SNP selected from the group consisting of: c794g, c973a, g1011c, t1049a, t1155a, and a1204g.

A polynucleotide such as previously defined can equally include at least one of the following non-coding SNPs: t1265c, t1277c.

The present invention equally has for its object an isolated polynucleotide comprising or consisting of:

a) a nucleotide sequence SEQ ID NO. 1 or its coding sequence, it being understood that each of these sequences comprises at least one of the following non coding SNPs: t1265c, t1277c, or

b) a nucleotide sequence complementary to a nucleotide sequence under a).

The present invention also concerns an isolated polynucleotide consisting of a part of:

a) a nucleotide sequence SEQ ID NO. 1 or its coding sequence, it being understood that each of these sequences comprises at least one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c, or

b) a nucleotide sequence complementary to a nucleotide sequence under a).

said isolated polynucleotide being composed of at least 10 nucleotides.

Preferably, the isolated polynucleotide as defined above is composed of 10 to 40 nucleotides.

The present invention also has for its object an isolated polynucleotide coding for a polypeptide comprising:

a) the amino acid sequence SEQ ID NO. 2, or

b) the amino acid sequence comprising the amino acids included between positions 24 and 189 in the sequence of amino acids SEQ ID NO. 2,

it being understood that each of the amino acid sequences under a) and b) comprises at least one of the following coding SNPs: A42G, Q102K, Q114H, V127D, C162stop, K179E.

It is understood, in the sense of the present invention, that the numbering corresponding to the positioning of the A42G, Q102K, Q114H, V127D, C162stop, K179E SNPs is relative to the numbering of the amino acid sequence SEQ ID NO. 2.

According to a preferred object of the invention, the previously defined polypeptide comprises a single coding SNP such as defined above.

Preferably a polynucleotide according to the invention is composed of a DNA or RNA molecule.

A polynucleotide according to the invention can be obtained by standard DNA or RNA synthetic methods.

A polynucleotide according to the invention can equally be obtained by site-directed mutagenesis starting from the nucleotide sequence of the IFNα-21 gene by modifying the wild-type nucleotide by the mutated nucleotide for each SNP on the nucleotide sequence SEQ ID NO. 1.

For example, a polynucleotide according to the invention, comprising a SNP c794g can be obtained by site-directed mutagenesis starting from the nucleotide sequence of the

IFNα-21 gene by modifying the nucleotide cytosine (c) by the nucleotide guanine (g) at position 794 on the nucleotide sequence SEQ ID NO. 1.

The processes of site-directed mutagenesis that can be implemented in this way are well known to a person skilled in the art. The publication of TA Kunkel in 1985 in "Proc. Natl. Acad. Sci. USA" 82:488 can notably be mentioned.

An isolated polynucleotide can equally include, for example, nucleotide sequences coding for pre-, pro- or pre-pro-protein amino acid sequences or marker amino acid sequences, such as hexa-histidine peptide.

A polynucleotide of the invention can equally be associated with nucleotide sequences coding for other proteins or protein fragments in order to obtain fusion proteins or other purification products.

A polynucleotide according to the invention can equally include nucleotide sequences such as the 5' and/or 3' non-coding sequences, such as, for example, transcribed or non-transcribed sequences, translated or non-translated sequences, splicing signal sequences, polyadenylated sequences, ribosome binding sequences or even sequences which stabilize mRNA.

A nucleotide sequence complementary to the nucleotide or polynucleotide sequence is defined as one that can hybridize with this nucleotide sequence, under stringent conditions.

"Stringent hybridization conditions" is generally but not necessarily understood as the chemical conditions that permit a hybridization when the nucleotide sequences have an identity of at least 80 %, preferably greater than or equal to 90 %, still more preferably greater than or equal to 95 % and most preferably greater than or equal to 97 %.

The stringent conditions can be obtained according to methods well known to a person skilled in the art and, for example, by an incubation of the polynucleotides, at 42° C, in a solution comprising 50 % formamide, 5xSSC (150 mM of NaCl, 15 mM of trisodium citrate), 50 mM of sodium phosphate (pH = 7.6), 5x Denhardt Solution, 10 % dextran sulfate and 20 µg denatured salmon sperm DNA, followed by washing the filters at 0.1x SSC, at 65° C.

Within the scope of the invention, when the stringent hybridization conditions permit hybridization of the nucleotide sequences having an identity equal to 100 %, the nucleotide sequence is considered to be strictly complementary to the nucleotide sequence such as described under a).

It is understood within the meaning of the present invention that the nucleotide sequence complementary to a nucleotide sequence comprises at least one anti-sense SNP according to the invention.

Thus, for example, if the nucleotide sequence comprises the SNP c794g, its complementary nucleotide sequence comprises the cytosine (c) nucleotide at the equivalent of position 794.

Identification, hybridization and/or amplification of a polynucleotide comprising a SNP.

The present invention also has for its object the use of all or part of:

a) a polynucleotide having 80 to 100 % identity (preferably at least 90 % identity, more preferably 95 % identity and particularly 100 % identity) with the nucleotide sequence SEQ ID NO. 1, and/or

b) a polynucleotide according to the invention comprising at least one SNP

in order to identify, hybridize and/or amplify all or part of a polynucleotide having 80 to 100 % identity (preferably at least 90 % identity, more preferably 95 % identity and particularly 100 % identity) with the nucleotide sequence SEQ ID NO. 1 or if necessary its coding sequence (of the nucleotide 670 to the nucleotide 1239),

it being understood that each one of these sequences comprises at least one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c.

Genotyping and determination of the frequency of a SNP

The present invention equally has for its object the use of all or part of:

a) a polynucleotide having 80 to 100 % identity (preferably at least 90 % identity, more preferably 95 % identity and particularly 100 % identity) with the nucleotide sequence SEQ ID NO. 1, and/or

b) a polynucleotide according to the invention comprising at least one SNP

for the genotyping of all or part of a polynucleotide having 80 to 100 % identity (preferably at least 90 % identity, more preferably 95 % identity and particularly 100 % identity) with the nucleotide sequence SEQ ID NO. 1 or if necessary its coding sequence (from nucleotide 670 to nucleotide 1239),

it being understood that each one of these sequences comprises at least one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c.

According to the invention, the genotyping may be carried out on an individual or a

population of individuals.

Within the meaning of the invention, genotyping is defined as a process for the determination of the genotype of an individual or of a population of individuals. Genotype consists of the alleles present at one or more specific loci.

"Population of individuals" is understood as a group of individuals selected in random or non-random fashion. These individuals can be humans, animals, microorganisms or plants.

Usually, the group of individuals comprises at least 10 persons, preferably from 100 to 300 persons.

The individuals can be selected according to their ethnicity or according to their phenotype, notably those who are affected by the following disorders and/or diseases: carcinomas, melanomas, lymphomas, leukemias and cancers of the liver, neck, head and kidneys, cardiovascular diseases, metabolic diseases such as those that are not connected with the immune system like, for example, obesity, infectious diseases in particular viral infections like hepatitis B and C and AIDS, pneumonias, ulcerative colitis, diseases of the central nervous system like, for example, Alzheimer's disease, schizophrenia and depression, the rejection of tissue or organ grafts, healing of wounds, anemia in dialyzed patients, allergies, asthma, multiple sclerosis, osteoporosis, psoriasis, rheumatoid arthritis, Crohn's disease, autoimmune diseases and disorders, gastrointestinal disorders or even disorders connected with chemotherapy treatments.

A functional SNP according to the invention is preferably genotyped in a population of individuals.

Multiple technologies exist which can be implemented in order to genotype SNPs (see notably Kwok Pharmacogenomics, 2000, vol 1, pp 95-100. "High–throughput genotyping assay approaches"). These technologies are based on one of the four following principles: allele specific oligonucleotide hybridization, oligonucleotide elongation by dideoxynucleotides optionally in the presence of deoxynucleotides, ligation of allele specific oligonucleotides or cleavage of allele specific oligonucleotides. Each one of these technologies can be coupled to a detection system such as measurement of direct or polarized fluorescence, or mass spectrometry.

Genotyping can notably be carried out by minisequencing with hot ddNTPs (2 different ddNTPs labeled by different fluorophores) and cold ddNTPs (2 different non labeled ddNTPs), in connection with a polarized fluorescence scanner. The minisequencing protocol with reading of polarized fluorescence (FP-TDI Technology or Fluorescence Polarization

Template-direct Dye-Terminator Incorporation) is well known to a person skilled in the art.

It can be carried out on a product obtained after amplification by polymerase chain reaction (PCR) of the DNA of each individual. This PCR product is selected to cover the polynucleotide genic region containing the studied SNP. After the last step in the PCR thermocycler, the plate is then placed on a polarized fluorescence scanner for a reading of the labeled bases by using fluorophore specific excitation and emission filters. The intensity values of the labeled bases are reported on a graph.

For the PCR amplification, in the case of a SNP of the invention, the sense and antisense primers, respectively, can easily be selected by a person skilled in the art according to the position of the SNPs of the invention.

For example, the sense and antisense nucleotide sequences for the PCR amplification primers can be, respectively:

SEQ ID NO. 3: Sense primer: GGTTCAAGGTTACCCATCTC

SEQ ID NO. 4: Antisense primer: TTTGAAATGGCAGAAGTCAT

The nucleotide sequences permit amplification of a fragment having a length of 696 nucleotides, from nucleotide 620 to nucleotide 1315 in the nucleotide sequence SEQ ID NO. 1.

A statistical analysis of the frequency of each allele (allelic frequency) encoded by the gene comprising the SNP in the population of individuals is then achieved, which permits determination of the importance of their impact and their distribution in the different sub-groups and notably, if necessary, the diverse ethnic groups that constitute this population of individuals.

The genotyping data are analyzed in order to estimate the distribution frequency of the different alleles observed in the studied populations. The calculations of the allelic frequencies can be carried out with the help of software such as SAS-suite® (SAS) or SPLUS® (MathSoft). The comparison of the allelic distributions of a SNP of the invention across different ethnic groups of the population of individuals can be carried out by means of the software ARLEQUIN® and SAS-suite®.

SNPs of the invention as genetic markers.

Whereas SNPs modifying functional sequences of genes (e.g. promoter, splicing sites, coding region) are likely to be directly related to disease susceptibility or resistance, all SNPs (functional or not) may provide valuable markers for the identification of one or several

genes involved in these disease states and, consequently, may be indirectly related to these disease states (See Cargill et al. (1999). Nature Genetics 22:231-238; Riley et al. (2000). Pharmacogenomics 1:39-47; Roberts L. (2000). Science 287: 1898-1899).

Thus, the present invention also concerns a databank comprising at least one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c, in a polynucleotide of the IFNα-21 gene.

It is well understood that said SNPs are numbered in accordance with the nucleotide sequence SEQ ID NO. 1.

This databank may be analyzed for determining statistically relevant associations between:

(i) at least one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c, in a polynucleotide of the IFNα-21 gene, and

(ii) a disease or a resistance to a disease.

The present invention also concerns the use of at least one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c, in a polynucleotide of the IFNα-21 gene, for developing diagnostic/prognostic kits for a disease or a resistance to a disease.

A SNP of the invention such as defined above may be directly or indirectly associated to a disease or a resistance to a disease.

Preferably, these diseases may be those which are defined as mentioned above.

Expression vector and host cells.

The present invention also has for its object a recombinant vector comprising at least one polynucleotide according to the invention.

Numerous expression systems can be used, including without limitation chromosomes, episomes, and derived viruses. More particularly, the recombinant vectors used can be derived from bacterial plasmids, transposons, yeast episomes, insertion elements, yeast chromosome elements, viruses such as baculovirus, papilloma viruses such as SV40, vaccinia viruses, adenoviruses, fox pox viruses, pseudorabies viruses, retroviruses.

These recombinant vectors can equally be cosmid or phagemid derivatives. The nucleotide sequence can be inserted in the recombinant expression vector by methods well known to a person skilled in the art such as, for example, those that are described in

MOLECULAR CLONING: A LABORATORY MANUAL, Sambrook *et al.*, 4th Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001.

The recombinant vector can include nucleotide sequences that control the regulation of the polynucleotide expression as well as nucleotide sequences permitting the expression and the transcription of a polynucleotide of the invention and the translation of a polypeptide of the invention, these sequences being selected according to the host cells that are used.

Thus, for example, an appropriate secretion signal can be integrated in the recombinant vector so that the polypeptide, encoded by the polynucleotide of the invention, will be directed towards the lumen of the endoplasmic reticulum, towards the periplasmic space, on the membrane or towards the extracellular environment.

The present invention also has for its object a host cell comprising a recombinant vector according to the invention.

The introduction of the recombinant vector in a host cell can be carried out according to methods that are well known to a person skilled in the art such as those described in BASIC METHODS IN MOLECULAR BIOLOGY, Davis *et al.*, 2nd ed., McGraw-Hill Professional Publishing, 1995, and MOLECULAR CLONING: A LABORATORY MANUAL, *supra*, such as transfection by calcium phosphate, transfection by DEAE dextran, transfection, microinjection, transfection by cationic lipids, electroporation, transduction or infection.

The host cell can be, for example, bacterial cells such as cells of streptococci, staphylococci, *E. coli* or *Bacillus subtilis*, cells of fungi such as yeast cells and cells of *Aspergillus, Streptomyces*, insect cells such as cells of *Drosophila* S2 and of *Spodoptera* Sf9, animal cells, such as CHO, COS, HeLa, C127, BHK, HEK 293 cells and human cells of the subject to treat or even plant cells.

The host cells can be used, for example, to express a polypeptide of the invention or as active product in pharmaceutical compositions, as will be seen hereinafter.

Polypeptide.

The present invention also has for its object an isolated polypeptide comprising an amino acid sequence having at least 80 % identity, preferably at least 90 % identity, more preferably at least 95 % identity and still more preferably at least 99 % identity with:

a) the amino acid sequence SEQ ID NO. 2 or with

b) the amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2,

it being understood that each of the amino acid sequences under a) and b) contains at least one of the following coding SNPs: A42G, Q102K, Q114H, V127D, C162stop, K179E.

The polypeptide of the invention can equally comprise:

a) the amino acid sequence SEQ ID NO. 2, or

b) the amino acid sequence containing the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2,

it being understood that each of the amino acid sequences under a) and b) contains at least one of the following coding SNPs: A42G, Q102K, Q114H, V127D, C162stop, K179E.

The polypeptide of the invention can more particularly consist of:

a) the amino acid sequence SEQ ID NO. 2, or

b) the amino acid sequence containing the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2,

it being understood that each one of the amino acid sequences under a) and b) contains at least one of the following coding SNPs: A42G, Q102K, Q114H, V127D, C162stop, K179E.

Preferably, a polypeptide according to the invention contains a single coding SNP selected from the group consisting of: A42G, Q102K, Q114H, V127D, C162stop, and K179E.

The present invention equally has for its object a process for the preparation of the above-described polypeptide, in which a previously defined host cell is cultivated in a culture medium and said polypeptide is isolated from the culture medium.

The polypeptide can be purified starting from the host cells' culture medium, according to methods well known to a person skilled in the art such as precipitation with the help of chaotropic agents such as salts, in particular ammonium sulfate, ethanol acetone or trichloroacetic acid, acid extraction; ion exchange chromatography; phosphocellulose chromatography; hydrophobic interaction chromatography; affinity chromatography; hydroxyapatite chromatography or exclusion chromatographies.

"Culture medium" is understood as the medium in which the polypeptide of the invention is isolated or purified. This medium can be composed of the extracellular medium and/or the cellular lysate. Techniques well known to a person skilled in the art equally permit the latter to give back an active conformation to the polypeptide, if the conformation of said polypeptide was altered during the isolation or the purification.

Antibodies.

The present invention also has for its object a process for obtaining an immunospecific antibody.

"Antibody" is understood as the monoclonal, polyclonal, chimeric, simple chain, humanized antibodies as well as the Fab fragments, including Fab or immunoglobulin expression library products.

An immunospecific antibody can be obtained by immunization of an animal with a polypeptide according to the invention.

The invention also relates to an immunospecific antibody for a polypeptide according to the invention, such as defined previously.

A polypeptide according to the invention, one of its fragments, an analog, one of its variants or a cell expressing this polypeptide can also be used to produce immunospecific antibodies.

The term "immunospecific" means that the antibody possesses a better affinity for the polypeptide of the invention than for other polypeptides known in the prior art.

The immunospecific antibodies can be obtained by administration of a polypeptide of the invention, of one of its fragments, of an analog or of an epitopic fragment or of a cell expressing this polynucleotide in a mammal, preferably non human, according to methods well known to a person skilled in the art.

For the preparation of monoclonal antibodies, typical methods for antibody production can be used, starting from cell lines, such as the hybridoma technique (Kohler et al., Nature (1975) 256: 495-497), the trioma technique, the human B cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4: 72) and the EBV hybridoma technique (Cole *et al.*, "The EBV-hybridoma technique and its application to human lung cancer," in Monoclonal Antibodies and Cancer Therapy (Vol. 27, UCLA Symposia on Molecular and Cellular Biology, New Series) (eds. R.A. Reisfeld and S.Sell), pp. 77-96, Alan R. Liss, Inc. N.Y., 1985, pp. 77-96).

The techniques of single chain antibody production such as described, for example, in US Patent No. 4,946, 778 can equally be used.

Transgenic animals such as mice, for example, can equally be used to produce humanized antibodies.

Agents interacting with the polypeptide of the invention.

The present invention equally has for its object a process for the identification of an agent activating or inhibiting a polypeptide according to the invention, comprising:

a) the preparation of a recombinant vector comprising a polynucleotide according to the invention containing at least one coding SNP,

b) the preparation of host cells comprising a recombinant vector according to a),

c) the contacting of host cells according to b) with an agent to be tested, and

d) the determination of the activating or inhibiting effect generated by the agent to test.

A polypeptide according to the invention can also be employed for a process for screening compounds that interact with it.

These compounds can be activating (agonists) or inhibiting (antagonists) agents of intrinsic activity of a polypeptide according to the invention. These compounds can equally be ligands or substrates of a polypeptide of the invention. See Coligan et al., Current Protocols in Immunology 1 (2), Chapter 5 (1991).

In general, in order to implement such a process, it is first desirable to produce appropriate host cells that express a polypeptide according to the invention. Such cells can be, for example, cells of mammals, yeasts, insects such as *Drosophila* or bacteria such as *E. coli*.

These cells or membrane extracts of these cells are then put in the presence of compounds to be tested.

The binding capacity of the compounds to be tested with the polypeptide of the invention can then be observed, as well as the inhibition or the activation of the functional response.

Step d) of the above process can be implemented by using an agent to be tested that is directly or indirectly labeled. It can also include a competition test, by using a labeled or non-labeled agent and a labeled competitor agent.

It can equally be determined if an agent to be tested generates an activation or inhibition signal on cells expressing the polypeptide of the invention by using detection means appropriately chosen according to the signal to be detected.

Such activating or inhibiting agents can be polynucleotides, and in certain cases oligonucleotides or polypeptides, such as proteins or antibodies, for example.

The present invention also has for its object a process for the identification of an agent activated or inhibited by a polypeptide according to the invention, comprising:

a) the preparation of a recombinant vector comprising a polynucleotide according to the

invention containing at least one coding SNP,

b) the preparation of host cells comprising a recombinant vector according to a),

c) placing host cells according to b) in the presence of an agent to be tested, and

d) the determination of the activating or inhibiting effect generated by the polypeptide on the agent to be tested.

An agent activated or inhibited by the polypeptide of the invention is an agent that responds, respectively, by an activation or an inhibition in the presence of this polypeptide. The agents, activated or inhibited directly or indirectly by the polypeptide of the invention, can consist of polypeptides such as, for example, membranal or nuclear receptors, kinases and more preferably tyrosine kinases, transcription factor or polynucleotides.

Detection of diseases.

The present invention also has for object a process for analyzing the biological characteristics of a polynucleotide according to the invention and/or of a polypeptide according to the invention in a subject, comprising at least one of the following:

a) Determining the presence or the absence of a polynucleotide according to the invention in the genome of a subject;

b) Determining the level of expression of a polynucleotide according to the invention in a subject;

c) Determining the presence or the absence of a polypeptide according to the invention in a subject;

d) Determining the concentration of a polypeptide according to the invention in a subject; and/or

e) Determining the functionality of a polypeptide according to the invention in a subject.

These biological characteristics may be analyzed in a subject or in a sample from a subject.

These biological characteristics may permit to carry out a genetic diagnosis and to determine whether a subject is affected or at risk of being affected or, to the contrary, presents a partial resistance to the development of a disease, an indisposition or a disorder linked to the presence of a polynucleotide according to the invention and/or a polypeptide according to the invention.

These diseases can be disorders and/or human diseases, such as cancers and tumors, infectious diseases, venereal diseases, immunologically related diseases and/or autoimmune

diseases and disorders, cardiovascular diseases, metabolic diseases, central nervous system diseases, and disorders connected with chemotherapy treatments.

Said cancers and tumors include carcinomas comprising metastasizing renal carcinomas, melanomas, lymphomas comprising follicular lymphomas and cutaneous T cell lymphoma, leukemias comprising hairy-cell leukemia, chronic lymphocytic leukemia and chronic myeloid leukemia, cancers of the liver, neck, head and kidneys, multiple myelomas, carcinoid tumors and tumors that appear following an immune deficiency comprising Kaposi's sarcoma in the case of AIDS.

Said infectious diseases include viral infections comprising chronic hepatitis B and C and HIV/AIDS, infectious pneumonias, and venereal diseases, such as genital warts.

Said immunologically and auto-immunologically related diseases may include the rejection of tissue or organ grafts, allergies, asthma, psoriasis, rheumatoid arthritis, multiple sclerosis, Crohn's disease and ulcerative colitis.

Said metabolic diseases may include such non-immune associated diseases as obesity.

Said diseases of the central nervous system may include Alzheimer's disease, Parkinson's disease, schizophrenia and depression.

Said diseases and disorders may also include healing of wounds, anemia in dialyzed patient, and osteoporosis.

This process also permits genetic diagnosis of a disease or of a resistance to a disease linked to the presence, in a subject, of the mutant allele encoded by a SNP according to the invention.

Preferably, in step a), the presence or absence of a polynucleotide, containing at least one coding SNP such as previously defined, is going to be detected.

The detection of the polynucleotide may be carried out starting from biological samples from the subject to be studied, such as cells, blood, urine, saliva, or starting from a biopsy or an autopsy of the subject to be studied. The genomic DNA may be used for the detection directly or after a PCR amplification, for example. RNA or cDNA can equally be used in a similar fashion.

It is then possible to compare the nucleotide sequence of a polynucleotide according to the invention with the nucleotide sequence detected in the genome of the subject.

The comparison of the nucleotide sequences can be carried out by sequencing, by DNA hybridization methods, by mobility difference of the DNA fragments on an electrophoresis gel with or without denaturing agents or by melting temperature difference. See Myers et al., Science (1985) 230: 1242. Such modifications in the structure of the

nucleotide sequence at a precise point can equally be revealed by nuclease protection tests, such as RNase and the S1 nuclease or also by chemical cleaving agents. See Cotton et al., Proc. Nat. Acad. Sci. USA (1985) 85: 4397-4401. Oligonucleotide probes comprising a polynucleotide fragment of the invention can equally be used to conduct the screening.

Many methods well known to a person skilled in the art can be used to determine the expression of a polynucleotide of the invention and to identify the genetic variability of this polynucleotide (See Chee et al., Science (1996), Vol 274, pp 610-613).

In step b), the level of expression of the polynucleotide may be measured by quantifying the level of RNA encoded by this polynucleotide (and coding for a polypeptide) according to methods well known to a person skilled in the art as, for example, by PCR, RT-PCR, RNase protection, Northern blot, and other hybridization methods.

In step c) and d) the presence or the absence as well as the concentration of a polypeptide according to the invention in a subject or a sample form a subject may be carried out by well known methods such as, for example, by radioimmunoassay, competitive binding tests, Western blot and ELISA tests.

Consecutively to step d), the determined concentration of the polypeptide according to the invention can be compared with the natural wild-type protein concentration usually found in a subject.

A person skilled in the art can identify the threshold above or below which appears the sensitivity or, to the contrary, the resistance to the disease, the indisposition or the disorder evoked above, with the help of prior art publications or by conventional tests or assays, such as those that are previously mentioned.

In step e), the determination of the functionality of a polypeptide according to the invention may be carried out by methods well known to a person skilled in the art as, for example, by in vitro tests such as above mentioned or by an use of host cells expressing said polypeptide.

Therapeutic compounds and treatments of diseases.

The present invention also has for its object a therapeutic compound containing, by way of active agent, a polypeptide according to the invention.

The invention also relates to the use of a polypeptide according to the invention, for the manufacture of a therapeutic compound intended for the prevention or the treatment of different human disorders and/or diseases. These diseases can be disorders and/or human diseases, such

as cancers and tumors, infectious diseases, venereal diseases, immunologically related diseases and/or autoimmune diseases and disorders, cardiovascular diseases, metabolic diseases, central nervous system diseases, and disorders connected with chemotherapy treatments.

Said cancers and tumors include carcinomas comprising metastasizing renal carcinomas, melanomas, lymphomas comprising follicular lymphomas and cutaneous T cell lymphoma, leukemias comprising hairy-cell leukemia, chronic lymphocytic leukemia and chronic myeloid leukemia, cancers of the liver, neck, head and kidneys, multiple myelomas, carcinoid tumors and tumors that appear following an immune deficiency comprising Kaposi's sarcoma in the case of AIDS.

Said infectious diseases include viral infections comprising chronic hepatitis B and C and HIV/AIDS, infectious pneumonias, and venereal diseases, such as genital warts.

Said immunologically and auto-immunologically related diseases may include the rejection of tissue or organ grafts, allergies, asthma, psoriasis, rheumatoid arthritis, multiple sclerosis, Crohn's disease and ulcerative colitis.

Said metabolic diseases may include such non-immune associated diseases as obesity.

Said diseases of the central nervous system may include Alzheimer's disease, Parkinson's disease, schizophrenia and depression.

Said diseases and disorders may also include healing of wounds, anemia in dialyzed patient, and osteoporosis.

Preferably, a polypeptide according to the invention can also be used for the manufacture of a therapeutic compound intended for the prevention or the treatment of different human disorders and/or diseases, such as certain viral infections such as chronic hepatitis B and C, leukemias such as hairy-cell leukemia and chronic myeloid leukemia, multiple myelomas, follicular lymphomas, carcinoid tumors, malignant melanomas, metastasized renal carcinomas, Alzheimer's disease, Parkinson's disease, as well as tumors that appear following an immune deficiency, such as Kaposi's sarcoma in the case of AIDS, and genital warts or venereal diseases.

Certain of the compounds permitting to obtain the polypeptide according to the invention as well as the compounds obtained or identified by or from this polypeptide can likewise be used for the therapeutic treatment of the human body, i.e. as a therapeutic compound.

This is why the present invention also has for an object a medicament containing, by way of active agent, a polynucleotide according to the invention containing at least one

previously defined coding SNP, a previously defined recombinant vector, a previously defined host cell, and/or a previously defined antibody.

The invention also relates to the use of a polynucleotide according to the invention containing at least one previously defined coding SNP, a previously defined recombinant vector, a previously defined host cell, and/or a previously defined antibody for the manufacture of a medicament intended for the prevention or the treatment of different human disorders and/or diseases. These diseases can be disorders and/or human diseases, such as cancers and tumors, infectious diseases, venereal diseases, immunologically related diseases and/or autoimmune diseases and disorders, cardiovascular diseases, metabolic diseases, central nervous system diseases, and disorders connected with chemotherapy treatments.

Said cancers and tumors include carcinomas comprising metastasizing renal carcinomas, melanomas, lymphomas comprising follicular lymphomas and cutaneous T cell lymphoma, leukemias comprising hairy-cell leukemia, chronic lymphocytic leukemia and chronic myeloid leukemia, cancers of the liver, neck, head and kidneys, multiple myelomas, carcinoid tumors and tumors that appear following an immune deficiency comprising Kaposi's sarcoma in the case of AIDS.

Said infectious diseases include viral infections comprising chronic hepatitis B and C and HIV/AIDS, infectious pneumonias, and venereal diseases, such as genital warts.

Said immunologically and auto-immunologically related diseases may include the rejection of tissue or organ grafts, allergies, asthma, psoriasis, rheumatoid arthritis, multiple sclerosis, Crohn's disease and ulcerative colitis.

Said metabolic diseases may include such non-immune associated diseases as obesity.

Said diseases of the central nervous system may include Alzheimer's disease, Parkinson's disease, schizophrenia and depression.

Said diseases and disorders may also include healing of wounds, anemia in dialyzed patient, and osteoporosis.

Preferably, the invention concerns the use of a polynucleotide according to the invention containing at least one previously defined SNP, a previously defined recombinant vector, a previously defined host cell, and/or a previously defined antibody, for the manufacture of a medicament intended for the prevention or the treatment of different human disorders and/or diseases, such as certain viral infections such as chronic hepatitis B and C, leukemias such as hairy-cell leukemia and chronic myeloid leukemia, multiple myelomas, follicular lymphomas, carcinoid tumors, malignant melanomas, metastasized renal

carcinomas, Alzheimer's disease, Parkinson's disease, as well as tumors that appear following an immune deficiency, such as Kaposi's sarcoma in the case of AIDS, and genital warts or venereal diseases.

The dosage of a polypeptide and of the other compounds of the invention, useful as active agent, depends on the choice of the compound, the therapeutic indication, the mode of administration, the nature of the formulation, the nature of the subject and the judgment of the doctor.

When it is used as active agent, a polypeptide according to the invention is generally administered at doses ranging between 1 and 100 μg/kg of the subject.

The invention also has as an object a pharmaceutical composition that contains, as active agent, at least one above-mentioned compound such as a polypeptide according to the invention, a polynucleotide according to the invention containing at least one previously defined SNP, a previously defined recombinant vector, a previously defined host cell, and/or a previously defined antibody, as well as a pharmaceutically acceptable excipient.

In these pharmaceutical compositions, the active agent is advantageously present at physiologically effective doses.

These pharmaceutical compositions can be, for example, solids or liquids and be present in pharmaceutical forms currently used in human medicine such as, for example, simple or coated tablets, gelcaps, granules, caramels, suppositories and preferably injectable preparations and powders for injectables. These pharmaceutical forms can be prepared according to usual methods.

The active agent(s) can be incorporated into excipients usually employed in pharmaceutical compositions such as talc, Arabic gum, lactose, starch, dextrose, glycerol, ethanol, magnesium stearate, cocoa butter, aqueous or non-aqueous vehicles, fatty substances of animal or vegetable origin, paraffinic derivatives, glycols, various wetting agents, dispersants or emulsifiers, preservatives.

The active agent(s) according to the invention can be employed alone or in combination with other compounds such as therapeutic compounds such as other cytokines such as interleukins or interferons, for example.

The different formulations of the pharmaceutical compositions are adapted according to the mode of administration.

The pharmaceutical compositions can be administered by different routes of administration known to a person skilled in the art.

The invention equally has for an object a diagnostic composition that contains, as active agent, at least one above-mentioned compound such as a polypeptide according to the invention, all or part of a polynucleotide according to the invention, a previously defined recombinant vector, a previously defined host cell, and/or a previously defined antibody, as well as a suitable pharmaceutically acceptable excipient.

This diagnostic composition may contain, for example, an appropriate excipient like those generally used in the diagnostic composition such as buffers and preservatives.

The present invention equally has as an object the use:

a) of a therapeutically effective quantity of a polypeptide according to the invention, and/or

b) of a polynucleotide according to the invention, and/or

c) of a host cell from the subject to be treated, previously defined,

to prepare a therapeutic compound intended to increase the expression or the activity, in a subject, of a polypeptide according to the invention.

Thus, to treat a subject who needs an increase in the expression or in the activity of a polypeptide of the invention, several methods are possible.

It is possible to administer to the subject a therapeutically effective quantity of a polypeptide of the invention, with a pharmaceutically acceptable excipient.

It is likewise possible to increase the endogenous production of a polypeptide of the invention by administration to the subject of a polynucleotide according to the invention. For example, this polynucleotide can be inserted in a retroviral expression vector. Such a vector can be isolated starting from cells having been infected by a retroviral plasmid vector containing RNA encoding for the polypeptide of the invention, in such a fashion that the transduced cells produce infectious viral particles containing the gene of interest. See Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, Chapter 20, in Human Molecular Genetics, Strachan and Read, BIOS Scientifics Publishers Ltd (1996).

In accordance with the invention, a polynucleotide containing at least one coding SNP such as previously defined will be preferably used.

It is equally possible to administer to the subject host cells belonging to him, these host cells having been preliminarily taken and modified so as to express the polypeptide of the invention, as previously described.

The present invention equally relates to the use:

a) of a therapeutically effective quantity of a previously defined immunospecific antibody, and/or

b) of a polynucleotide permitting inhibition of the expression of a polynucleotide according to the invention,

in order to prepare a therapeutic compound intended to reduce the expression or the activity, in a subject, of a polypeptide according to the invention.

Thus, it is possible to administer to the subject a therapeutically effective quantity of an inhibiting agent and/or of an antibody such as previously defined, possibly in combination, with a pharmaceutically acceptable excipient.

It is equally possible to reduce the endogenous production of a polypeptide of the invention by administration to the subject of a complementary polynucleotide according to the invention permitting inhibition of the expression of a polynucleotide of the invention.

Preferably, a complementary polynucleotide containing at least one coding SNP such as previously defined can be used.

The present invention concerns also the use of a IFNα-21 protein for the preparation of a medicament for the prevention or the treatment of a patient having a disorder or a disease caused by a IFNα-21 variant linked to the presence in the genome of said patient of a nucleotide sequence having at least 95% identity (preferably, 97% identity, more preferably 99% identity and particularly 100% identity) with the nucleotide sequence SEQ ID NO. 1, provided that said nucleotide sequence comprises one of the following SNPs: c794g, c973a, g1011c, t1049a, t1155a, a1204g, t1265c, t1277c.

Preferably, said medicament is used for the prevention or the treatment of one of the diseases selected from the group consisting of cancers and tumors, infectious diseases, venereal diseases, immunologically related diseases and/or autoimmune diseases and disorders, cardiovascular diseases, metabolic diseases, central nervous system diseases, and disorders connected with chemotherapy treatments.

Said cancers and tumors include carcinomas comprising metastasizing renal carcinomas, melanomas, lymphomas comprising follicular lymphomas and cutaneous T cell lymphoma, leukemias comprising hairy-cell leukemia, chronic lymphocytic leukemia and chronic myeloid leukemia, cancers of the liver, neck, head and kidneys, multiple myelomas, carcinoid tumors and tumors that appear following an immune deficiency comprising Kaposi's sarcoma in the case of AIDS.

Said infectious diseases include viral infections comprising chronic hepatitis B and C and HIV/AIDS, infectious pneumonias, and venereal diseases, such as genital warts.

Said immunologically and auto-immunologically related diseases may include the

rejection of tissue or organ grafts, allergies, asthma, psoriasis, rheumatoid arthritis, multiple sclerosis, Crohn's disease and ulcerative colitis.

Said metabolic diseases may include such non-immune associated diseases as obesity.

Said diseases of the central nervous system may include Alzheimer's disease, Parkinson's disease, schizophrenia and depression.

Said diseases and disorders may also include healing of wounds, anemia in dialyzed patient, and osteoporosis.

Mimetic compounds of an IFNα-21 polypeptide comprising the SNPs of the invention.

The present invention also concerns a new compound having a biological activity substantially similar to that of the polypeptide of:

a) amino acid sequence SEQ ID NO. 2, or

b) amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2;

provided that said amino acid sequences under a) and b) comprise the SNP K179E.

Said biological activity may be evaluated, for example, by measuring the capacity to stimulate dendritic cell maturation, cytokine release by CD4+ or CD8+ T-lymphocytes, cytokine release by monocytes, in vitro or in vivo antiviral activity, cellular antiproliferative activity on Daudi Burkitt's cell line, cellular antiproliferative activity on TF-1 cell line as described in the experimental part.

As mentioned in the experimental section, the K179E mutated IFNα-21 differs from the wild-type IFNα-2 in the following ways:

- K179E mutated IFNα-21 possesses higher capacity to stimulate dendritic cell maturation;
- K179E mutated IFNα-21 possesses a higher capacity to stimulate IFN-gamma release by CD4+ or CD8+ T-lymphocytes;
- K179E mutated IFNα-21 possesses an antiviral activity in cell culture infected with VSV which is lower than that of wild-type IFNα-2.

As mentioned in the experimental part, K179E mutated IFNα-21 possesses a cellular antiproliferative activity on Daudi Burkitt's cell line which is slightly lower than that of the natural wild- type IFNα-21.

Also as mentioned in the experimental part, the K179E mutated IFNα-21 possesses a cellular antiproliferative activity on TF-1 cell line which is similar to that of wild-type IFNα-

2, and an antiviral activity in EMCV mouse model which is similar to that of wild-type IFNα-2.

A new compound of the invention, such as previously defined, may possess a biological activity substantially similar to that of the K179E mutated IFNα-21.

Said compound may also have a biological activity which is even lower or higher, according to the kind of biological activity considered, than that of the K179E mutated IFNα-21.

Said compound may be a biochemical compound, such as a polypeptide or a peptide for example, or an organic chemical compound, such as a synthetic peptide-mimetic for example.

The present invention also concerns the use of a polypeptide of the invention containing the K179E SNP, for the identification of a compound such as defined above.

The present invention also concerns a process for the identification of a compound of the invention, comprising the following steps:

a) Determining the biological activity of the compound to be tested, such as dendritic cell maturation, cytokine release by CD4+ or CD8+ T-lymphocytes, cytokine release by monocytes, in vitro or in vivo antiviral activity, cellular antiproliferative activity on Daudi Burkitt's cell line, for example;

b) Comparing:

i) the activity determined in step a) of the compound to be tested, with

ii) the activity of the polypeptide of amino acid sequence SEQ ID NO. 2, or of amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2; provided that said amino acid sequences comprise the K179E SNP; and

c) Determining on the basis of the comparison carried out in step b) whether the compound to be tested has a substantially similar, or lower or higher, activity compared to that of the polypeptide of amino acid sequence SEQ ID NO. 2, or of amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2; provided that said amino acid sequences comprise the K179E SNP.

Preferably, the compound to be tested may be previously identified from synthetic peptide combinatorial libraries, high-throughput screening, or designed by computer-aided drug design so as to have the same three-dimensional structure as that of the polypeptide of

amino acid sequence SEQ ID NO. 2, or of amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2; provided that said amino acid sequences comprise the K179E SNP.

The present invention also concerns a new compound having a biological activity substantially similar to that of the polypeptide of:

a) amino acid sequence SEQ ID NO. 2, or

b) amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2;

provided that said amino acid sequences under a) and b) comprise the SNPs Q114H and V127D.

Said biological activity may be evaluated, for example, by measuring the capacity to stimulate dendritic cell maturation, cytokine release by CD4+ or CD8+ T-lymphocytes, cytokine release by monocytes, in vitro or in vivo antiviral activity, cellular antiproliferative activity on Daudi Burkitt's cell line, cellular antiproliferative activity on TF-1 cell line as described in the experimental part.

As mentioned in the experimental section, IFNα-21 containing both SNPs Q114H and V127D (such double mutation shall be referred to hereafter as "Q114H/V127D") differs from the wild-type IFNα-2 in the following ways:

- the Q114H/V127D mutated IFNα-21 possesses a higher capacity to stimulate IFN-gamma release by CD4+ or CD8+ T-lymphocytes;
- the Q114H/V127D mutated IFNα-21 possesses a lower antiviral activity in cell culture infected with VSV;

As mentioned in the experimental part, the Q114H/V127D mutated IFNα-21 possesses a cellular antiproliferative activity on Daudi Burkitt's cell line which is lower than that of the natural wild- type IFNα-21.

Also as mentioned in the experimental part, the Q114H/V127D mutated IFNα-21 possesses an antiviral activity in EMCV mouse model which is similar to that of wild-type IFNα-2, a capacity to stimulate IL-10, IL-12 and TNF-α release by monocytes which is similar to that of wild-type IFNα-2, and cellular antiproliferative activity on TF-1 cell line which is similar to that of wild-type IFNα-2.

A new compound of the invention, such as previously defined, may possess a biological activity substantially similar to that of the Q114H/V127D mutated IFNα-21.

Said compound may also have a biological activity which is even lower or higher, according to the kind of biological activity considered, than that of the Q114H/V127D mutated IFNα-21.

Said compound may be a biochemical compound, such as a polypeptide or a peptide for example, or an organic chemical compound, such as a synthetic peptide-mimetic for example.

The present invention also concerns the use of a polypeptide of the invention containing the Q114H/V127D SNP, for the identification of a compound such as defined above.

The present invention also concerns a process for the identification of a compound of the invention, comprising the following steps:

a) Determining the biological activity, such as dendritic cell maturation, cytokine release by CD4+ or CD8+ T-lymphocytes, cytokine release by monocytes, in vitro or in vivo antiviral activity, cellular antiproliferative activity on Daudi Burkitt's cell line, for example;

b) Comparing:

i) the activity determined in step a) of the compound to be tested, with

ii) the activity of the polypeptide of amino acid sequence SEQ ID NO. 2, or of amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2;

provided that said amino acid sequences comprise the Q114H/V127D SNP; and

c) Determining on the basis of the comparison carried out in step b) whether the compound to be tested has a substantially similar, or lower or higher, activity compared to that of the polypeptide of amino acid sequence SEQ ID NO. 2, or of amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2; provided that said amino acid sequences comprise the Q114H/V127D SNP.

Preferably, the compound to be tested may be previously identified from synthetic peptide combinatorial libraries, high-throughput screening, or designed by computer-aided drug design so as to have the same three-dimensional structure as that of the polypeptide of amino acid sequence SEQ ID NO. 2, or of amino acid sequence comprising the amino acids included between positions 24 and 189 of the amino acid sequence SEQ ID NO. 2; provided that said amino acid sequences comprise the Q114H/V127D SNPs.

The methods to identify and design compounds are well known by a person skilled in the art.

Publications referring to these methods may be, for example:

- Silverman R.B. (1992). "Organic Chemistry of Drug Design and Drug Action". Academic Press, 1st edition (January 15, 1992).

- Anderson S and Chiplin J. (2002). "Structural genomics; shaping the future of drug design" Drug Discov. Today. 7(2):105-107.

- Selick HE, Beresford AP, Tarbit MH. (2002). "The emerging importance of predictive ADME simulation in drug discovery". Drug Discov. Today. 7(2):109-116.

- Burbidge R, Trotter M, Buxton B, Holden S. (2001). "Drug design by machine learning: support vector machines for pharmaceutical data analysis". Comput. Chem. 26(1): 5-14.

- Kauvar L.M. (1996). "Peptide mimetic drugs: a comment on progress and prospects" 14(6): 709.

The compounds of the invention may be used for the preparation of a medicament intended for the prevention or the treatment of one of the diseases selected from the group consisting of cancers and tumors, infectious diseases, venereal diseases, immunologically related diseases and/or autoimmune diseases and disorders, cardiovascular diseases, metabolic diseases, central nervous system diseases, and disorders connected with chemotherapy treatments.

Said cancers and tumors include carcinomas comprising metastasizing renal carcinomas, melanomas, lymphomas comprising follicular lymphomas and cutaneous T cell lymphoma, leukemias comprising hairy-cell leukemia, chronic lymphocytic leukemia and chronic myeloid leukemia, cancers of the liver, neck, head and kidneys, multiple myelomas, carcinoid tumors and tumors that appear following an immune deficiency comprising Kaposi's sarcoma in the case of AIDS.

Said infectious diseases include viral infections comprising chronic hepatitis B and C and HIV/AIDS, infectious pneumonias, and venereal diseases, such as genital warts.

Said immunologically and auto-immunologically related diseases may include the rejection of tissue or organ grafts, allergies, asthma, psoriasis, rheumatoid arthritis, multiple sclerosis, Crohn's disease and ulcerative colitis.

Said metabolic diseases may include such non-immune associated diseases as obesity.

Said diseases of the central nervous system may include Alzheimer's disease, Parkinson's disease, schizophrenia and depression.

Said diseases and disorders may also include healing of wounds, anemia in dialyzed patient, and osteoporosis.

The compounds of the invention may be used for the preparation of a medicament intended for the prevention or the treatment of one of the diseases selected from the group consisting of certain viral infections such as chronic hepatitis B and C, leukemias such as hairy-cell leukemia and chronic myeloid leukemia, multiple myelomas, follicular lymphomas, carcinoid tumors, malignant melanomas, metastasized renal carcinomas, Alzheimer's disease, Parkinson's disease, as well as tumors that appear following an immune deficiency, such as Kaposi's sarcoma in the case of AIDS, and genital warts or venereal diseases.

EXPERIMENTAL SECTION

Example 1: Modeling of a protein encoded by a polynucleotide of nucleotide sequence containing SNP c794g, c973a, g1011c, t1049a, t1155a, a1204g and of the protein encoded by the nucleotide sequence of the wild-type reference gene

In a first step the three-dimensional structure of IFNα-21 was constructed starting from that of IFNα-2 whose structure is available in the PDB database (code 1ITF) and by using the software Modeler (MSI, San Diego, CA).

The mature polypeptide fragment was then modified in such a fashion as to reproduce the mutation A19G, Q79K, Q91H, V104D, C139stop, and K156E.

A thousand molecular minimization steps were conducted on this mutated fragment by using the programs AMBER and DISCOVER (MSI: Molecular Simulations Inc.).

Two molecular dynamic calculation runs were then carried out with the same program and the same force fields.

In each case, 50,000 steps were calculated at 300°K, terminated by 300 equilibration steps.

The result of this modeling is visualized on Figures 1, 2, 3, 4, and 5.

Example 2: Genotyping of the SNPs c794g, c973a, g1011c, t1049a, t1155a, a1204g in a population of individuals

The genotyping of SNPs is based on the principle of the minisequencing wherein the product is detected by a reading of polarized fluorescence. The technique consists of a fluorescent minisequencing (FP-TDI Technology or Fluorescence Polarization Template-

direct Dye-terminator Incorporation).

The minisequencing is performed on a product amplified by PCR from genomic DNA of each individual of the population. This PCR product is chosen in such a manner that it covers the genic region containing the SNP to be genotyped. After elimination of the PCR primers that have not been used and the dNTPs that have not been incorporated, the minisequencing is carried out.

The minisequencing consists of lengthening an oligonucleotide primer, placed just upstream of the site of the SNP, by using a polymerase enzyme and fluorolabeled dideoxynucleotides. The product resulting from this lengthening process is directly analyzed by a reading of polarized fluorescence.

All these steps, as well as the reading, are carried out in the same PCR plate.

Thus, the genotyping requires 5 steps:

1) Amplification by PCR

2) Purification of the PCR product by enzymatic digestion

3) Elongation of the oligonucleotide primer

4) Reading

5) Interpretation of the reading

The genotyping steps 1 and 2 are carried out in the same conditions for each of the SNPs c794g, c973a, g1011c, t1049a, t1155a, a1204g. The steps 3, 4 and 5 are specific to each one of these polymorphisms.

1) The PCR amplification of the nucleotide sequence of the IFNα-21 gene is carried out starting from genomic DNA coming from 268 individuals of ethnically diverse origins.

These genomic DNAs were provided by the Coriell Institute in the United States.

The 268 individuals are distributed as follows:

Phylogenic Population	Specific Ethnic Population	Total	%
African American	African American	*50*	*100.0*
	Subtotal	***50***	***18.7***
Amerind	South American Andes	*10*	*66.7*
	South West American Indians	*5*	*33.3*
	Subtotal	***15***	***5.6***
Caribbean	Caribbean	*10*	*100.0*
	Subtotal	***10***	***3.7***
European Caucasoid	North American Caucasian	*79*	*79.8*
	Iberian	*10*	*10.1*
	Italian	*10*	*10.1*
	Subtotal	***99***	***36.9***
Mexican	Mexican	*10*	*100.0*
	Subtotal	***10***	***3.7***
Northeast Asian	Chinese	*10*	*50.0*
	Japanese	*10*	*50.0*
	Subtotal	***20***	***7.5***
Non-European Caucasoid	Greek	*8*	*21.6*
	Indo-Pakistani	*9*	*24.3*
	Middle-Eastern	*20*	*54.1*
	Subtotal	***37***	***13.8***
Southeast Asian	Pacific Islander	*7*	*41.2*
	South Asian	*10*	*58.8*
	Subtotal	***17***	***6.3***
South American	South American	*10*	*100.0*
	Subtotal	***10***	***3.7***
	Total	***268***	***100***

The genomic DNA coming from each one of these individuals constitutes a sample.

For all the SNPs, the PCR amplification is carried out starting from the following primers:

SEQ ID NO. 5: Sense primer: GGTTCAAGGTTACCCATCTC

SEQ ID NO. 6: Antisense primer: TTTGAAATGGCAGAAGTCAT

These nucleotide sequences permit amplification of a fragment of a length of 696 nucleotides, from nucleotide 620 to nucleotide 1315 in the nucleotide sequence SEQ ID NO. 1.

For each SNP, the PCR product will serve as a template for the minisequencing

The total reaction volume of the PCR reaction is 5 µl per sample.

This reaction volume is composed of the reagents indicated in the following table:

Supplier	Reference	Reactant	Initial Conc.	Vol. per tube (µl)	Final Conc.
Life Technology	Delivered with Taq	Buffer (X)	10	0.5	1
Life Technology	Delivered with Taq	$MgSO_4$ (mM)	50	0.2	2
AP Biotech	27-2035-03	dNTPs (mM)	10	0.1	0.2
	On request	Sense Primer (µM)	10	0.1	0.2
	On request	Antisense Primer (µM)	10	0.1	0.2
Life Technology	11304-029	Taq platinum	5U/ µl	0.02	0.1 U/ reaction
		H_2O	Qsp 5 µl	1.98	
		DNA (sample)	2.5 ng/ µl	2	5 ng/ reaction
		Total volume		5 µl	

These reagents are distributed in a black PCR plate having 384 wells provided by ABGene (ref: TF-0384-k). The plate is sealed, centrifuged, then placed in a thermocycler for 384-well plates (Tetrad of MJ Research) and undergoes the following incubation: PCR Cycles: 1 min at 94° C, followed by 36 cycles composed of 3 steps (15 sec. at 94° C, 30 sec. at 56° C, 1 min at 68° C).

2) The PCR amplified product is then purified using two enzymes: Shrimp Alkaline Phosphatase (SAP) and exonuclease I (Exo I). The first of these enzymes permits the dephosphorylation of the dNTPs which have not been incorporated during the PCR amplification, whereas the second eliminates the single stranded DNA residues, in particular the primers which have not been used during the PCR.

This digestion is done by addition, in each well of the PCR plate, of a reaction mixture of 5 µl per sample. This reaction mixture is composed of the following reagents:

Supplier	Reference	Reactant	Initial Conc.	Vol. per tube (µl)	Final conc.
AP Biotech	E70092X	SAP	1 U/ µl	0.5	0.5/reaction
AP Biotech	070073Z	Exo I	10 U/ µl	0.1	1/reaction
AP Biotech	Supplied with SAP	Buffer SAP (X)	10	0.5	1
		H_2O	Qsp 5 µl	3.9	
		PCR product		5 µl	
		Total vol.		10 µl	

Once filled, the plate is sealed, centrifuged, then placed in a thermocycler for 384 well plates (Tetrad of MJ Research) and undergoes the following incubation: Digestion SAP-EXO: 45 min at 37° C, 15 min at 80° C.

The elongation or minisequencing step is then carried out on the product of PCR digested by addition of a reaction mixture of 5 µl per prepared sample.

The minisequencing 3) and the reading steps 4) and interpretation of reading 5) are specific to each SNP c794g, c973a, g1011c, t1049a, t1155a, a1204g.

All these steps are described hereinafter precising the specific conditions used for each one of these polymorphisms.

3) Minisequencing

The sequences of the two minisequencing primers necessary for the genotyping were determined in a way to correspond to the sequence of the nucleotides located upstream of the site of a SNP according to the invention. The PCR product that contains the SNP being a double stranded DNA product, the genotyping can therefore be done either on the sense strand or on the antisense strand. The selected primers are manufactured by Life Technologies Inc.

The following table indicates, for each SNP, the sequence of the minisequencing primers that have been tested and the optimal condition retained for the genotyping:

SNP	Primers tested	Optimal condition for the genotyping
c794g	SEQ ID NO. 7: Sense: gagggccttgatactcctgg SEQ ID NO. 8: Antisense: gagagattcttcccatttgt	antisense primer + ddGTP-R110 + ddCTP-Tamra
c973a	SEQ ID NO. 9: Sense: actcatctgctacttgggaa SEQ ID NO. 10: Antisense: aaatttttctaggaggctct	antisense primer + dGTP-R110 + ddTTP-Tamra
g1011c	SEQ ID NO. 11: Sense: ttttccactgaacttaacca SEQ ID NO. 12: Antisense: gcttccaggtcattcagctg	antisense primer + ddGTP-R110 + ddCTP-Tamra
t1049a	SEQ ID NO. 13: Sense: agcctgcgtgatacaggagg SEQ ID NO. 14: Antisense: ggggagtctcttccacccca	antisense primer + ddATP-R110 + ddTTP-Tamra
t1155a	SEQ ID NO. 15: Sense: gagaagaaatacagcccttg SEQ ID NO. 16: Antisense: gctctgacaacctcccaggc	antisense primer + ddATP-R110 + ddTTP-Tamra
a1204g	SEQ ID NO. 17: Sense: tgagatccttctctttatca SEQ ID NO. 18: Antisense: taatctttcttgaaaaattt	sense primer + ddGTP-R110 + ddATP-Tamra

The minisequencing of the SNPs was first validated over 16 samples, then genotyped over the set of the population of individuals composed of 268 individuals and 10 controls.

The elongation or minisequencing step is then carried out as indicated in the following table:

Supplier	Reference	Reactant	Initial conc.	Vol. per tube (µl)	Final conc.
Own preparation		Elongation Buffer[1] (X)	5	1	1
Life Technologies	On request	Miniseq Primer (µM) A or B	10	0.5	1
AP Biotech	27-2051 (61,71,81)-01	ddNTPs[2] (µM) 2 are non labeled	2.5 of each	0.25	0.125 of each
NEN	Nel 472/5 and Nel 492/5	ddNTPs[2] (µM) 2 are labeled with Tamra and R110	2.5 of each	0.25	0.125 of each
AP Biotech	E79000Z	Thermo-sequenase	3.2 U/ µl	0.125	0.4 U/ reaction
		H_2O	Qsp 5 µl	3.125	
		digested PCR product		10	
		Total volume		15	

[1] The 5X elongation buffer is composed of 250 mM Tris-HCl pH 9, 250 mM KCl, 25 mM NaCl, 10 mM $MgCl_2$ and 40 % glycerol.

[2] For the ddNTPs, a mixture of the 4 bases is carried out according to the polymorphism studied. Only the 2 bases of interest (wild-type nucleotide/mutated nucleotide) composing the functional SNP are labeled, either in Tamra, or in R110. For example:

For SNP c973a, the mixture of ddNTPs is composed of:

- 2.5 µM of ddCTP non labeled,
- 2.5 µM of ddATP non-labeled,
- 2.5 µM of ddTTP (1.875 µM of ddTTP non labeled and 0.625 µM of ddTTP Tamra labeled),
- 2.5 µM of ddGTP (1.875 µM of ddGTP non labeled and 0.625 µM of ddGTP R110 labeled).

Once filled, the plate is sealed, centrifuged, then placed in a thermocycler for 384-well plates (Tetrad of MJ Research) and undergoes the following incubation: Elongation cycles: 1 min. at 93° C, followed by 35 cycles composed of 2 steps (10 sec. at 93° C, 30 sec. at 55° C).

After the last step in the thermocycler, the plate is directly placed on a polarized fluorescence reader of type Analyst® HT of LJL Biosystems Inc. The plate is read using Criterion Host® software by using two methods. The first permits reading the Tamra labeled base by using emission and excitation filters specific for this fluorophore (excitation 550-10 nm, emission 580-10 nm) and the second permits reading the R110 labeled base by using the excitation and emission filters specific for this fluorophore (excitation 490-10 nm, emission 520-10 nm). In the two cases, a dichroic double mirror (R110/Tamra) is used and the other

reading parameters are:

Z-height: 1.5 mm

Attenuator: out

Integration time: 100,000 µsec.

Raw data units: counts/sec

Switch polarization: by well

Plate settling time: 0 msec

PMT setup: Smart Read (+), sensitivity 2

Dynamic polarizer: emission

Static polarizer: S

A file result is thus obtained containing the calculated values of mP (milliPolarization) for the Tamra filter and that for the R110 filter. These mP values are calculated starting from intensity values obtained on the parallel plane (//) and on the perpendicular plane (⊥) according to the following formula:

$$MP = 1000(// - g\perp)/(// + g\perp).$$

In this calculation, the value ⊥ is weighted by a factor g. It is a machine parameter that must be determined experimentally beforehand.

4) and 5) Interpretation of the reading and determination of the genotypes.

The mP values are reported on a graph using Microsoft Inc. Excel software, and/or Allele Caller® software developed by LJL Biosystems Inc.

On the abscissa is indicated the mP value of the Tamra labeled base, on the ordinate is indicated the mP value of the R110 labeled base. A strong mP value indicates that the base labeled with this fluorophore is incorporated and, conversely, a weak mP value reveals the absence of incorporation of this base.

Up to three homogenous groups of nucleotide sequences having different genotypes may be obtained.

The use of the Allele Caller® software permits, once the identification of the different groups is carried out, to directly extract the genotype defined for each individual in table form.

Results of the minisequencing for the SNPs c794g, c973a, g1011c, t1049a, t1155a, a1204g

After the completion of the genotyping process, the determination of the genotypes of the individuals of the population of individuals for the SNPs studied here was carried out using the graphs described above.

For SNP c794g, the genotype is in theory either homozygote CC, or heterozygote CG, or homozygote GG in the tested individuals. In reality, and as shown below, the homozygote genotype GG is not detected in the population of individuals.

For SNP c973a, the genotype is in theory either homozygote CC, or heterozygote CA, or homozygote AA in the tested individuals. In reality, and as shown below, the homozygote genotype AA is not detected in the population of individuals.

For SNP g1011c, the genotype is in theory either homozygote GG, or heterozygote GC, or homozygote CC in the tested individuals. In reality, and as shown below, the homozygote genotype CC is not detected in the population of individuals.

For SNP t1049a, the genotype is in theory either homozygote TT, or heterozygote TA, or homozygote AA in the tested individuals. In reality, and as shown below, the homozygote genotype AA is not detected in the population of individuals.

For SNP t1155a, the genotype is in theory either homozygote TT, or heterozygote TA, or homozygote AA in the tested individuals. In reality, and as shown below, the homozygote genotype AA is not detected in the population of individuals.

For SNP a1204g, the genotype is in theory either homozygote AA, or heterozygote AG, or homozygote GG in the tested individuals. In reality, and as shown below, the homozygote genotype GG is not detected in the population of individuals.

The results of the distribution of the determined genotypes in the population of individuals and the calculation of the different allelic frequencies for the 6 SNPs studied are presented in the following tables:

Phylogenic Population	Total	c794g (A42G)								
		f	(95% CI)	CC	%	CG	%	GG	%	Total
African American	50	5,0	(0.7, 9.3)	45	90,0	5	10			50
Amerind	15			15	100					15
Caribbean	10			10	100					10
European Caucasoid	99			99	100					99
Mexican	10			10	100					10
Non-European Caucasoid	37			37	100					37
Northeast Asian	20			20	100					20
South American	10			10	100					10
Southeast Asian	17	2,9	(0, 8.6)	16	94,1	1	5,9			17
Total	268	1,1	(0.2, 2.0)	262	97,8	6	2,2			268

Phylogenic Population	Total	c973a (Q102K)								
		f	(95% CI)	CC	%	CA	%	AA	%	Total
African American	50			49	100					49
Amerind	15			15	100					15
Caribbean	10			10	100					10
European Caucasoid	99	1,5	(0, 3.2)	96	97,0	3	3,0			99
Mexican	10			9	100					9
Non-European Caucasoid	37	1,4	(0, 4.0)	36	97,3	1	2,7			37
Northeast Asian	20			20	100					20
South American	10			10	100					10
Southeast Asian	17			17	100					17
Total	268	0,8	(0, 1.5)	262	98,5	4	1,5			266

Phylogenic Population	Total	g1011c (Q114H)								
		f	(95% CI)	GG	%	GC	%	CC	%	Total
African American	50			49	100					49
Amerind	15			15	100					15
Caribbean	10			10	100					10
European Caucasoid	99	0,5	(0, 1.6)	92	98,9	1	1,1			93
Mexican	10			10	100					10
Non-European Caucasoid	37			37	100					37
Northeast Asian	20			20	100					20
South American	10			10	100					10
Southeast Asian	17			17	100					17
Total	268	0,2	(0, 0.6)	260	99,6	1	0,4			261

Phylogenic Population	Total	t1049a (V127D)								
		f	(95% CI)	TT	%	TA	%	AA	%	Total
African American	50	1,0	(0, 3.0)	48	98,0	1	2,0			49
Amerind	15			14	100					14
Caribbean	10			10	100					10
European Caucasoid	99			98	100					98
Mexican	10			10	100					10
Non-European Caucasoid	37			37	100					37
Northeast Asian	20			20	100					20
South American	10			10	100					10
Southeast Asian	17			17	100					17
Total	**268**	**0,2**	**(0, 0.6)**	**264**	**99,6**	**1**	**0,4**			**265**

Phylogenic Population	Total	t1155a (C162STOP)								
		f	(95% CI)	TT	%	TA	%	AA	%	Total
African American	50	1,0	(0, 3.0)	48	98,0	1	2,0			49
Amerind	15			15	100					15
Caribbean	10			9	100					9
European Caucasoid	99	3,6	(1.0, 6.2)	91	92,9	7	7,1			98
Mexican	10			10	100					10
Non-European Caucasoid	37			37	100					37
Northeast Asian	20			20	100					20
South American	10			10	100					10
Southeast Asian	17			17	100					17
Total	**268**	**1,5**	**(0.5, 2.5)**	**257**	**97,0**	**8**	**3,0**			**264**

Phylogenic Population	Total	a1204g (K179E)								
		f	(95% CI)	AA	%	AG	%	GG	%	Total
African American	50			50	100					50
Amerind	15	3,6	(0, 10.4)	13	92,9	1	7,1			14
Caribbean	10			10	100					10
European Caucasoid	99	1,0	(0, 2.4)	97	98,0	2	2,0			99
Mexican	10	10,0	(0, 23.1)	8	80,0	2	20,0			10
Non-European Caucasoid	37	2,7	(0, 6.4)	35	94,6	2	5,4			37
Northeast Asian	20	17,5	(5.7, 29.3)	13	65,0	7	35,0			20
South American	10	5,0	(0, 14.6)	9	90,0	1	10,0			10
Southeast Asian	17			17	100					17
Total	**268**	**2,8**	**(1.4, 4.2)**	**252**	**94,4**	**15**	**5,6**			**267**

In the above table,

- N represents the number of individuals,
- % represents the percentage of individuals in the specific sub-population,
- the allelic frequency represents the percentage of the mutated allele in the specific sub-population,
- 95 % IC represents the minimal and maximal interval of confidence at 95 %.

It is necessary to specify that for SNP c973a, for example, the allele g read in antisense corresponds to the allele c read in sense, and is related to the presence of a glutamine (Q) at position 102 of the immature IFNα-21 protein sequence and therefore that the allele t read in antisense corresponds to the allele a read in sense corresponding to a lysine (K) for this position in the sequence of the corresponding protein.

By examining these results by phylogenic population, and by SNP, it is observed that:

- for SNP c794g, the 6 heterozygote individuals CG come from the sub-populations African American and Southeast Asian.
- for SNP c973a, the 4 heterozygote individual CA come from the sub-populations European and non-European Caucasoid.
- for SNP g1011c, the unique heterozygote individual GC comes from the sub-population European Caucasoid.
- for SNP t1049a, the unique heterozygote individual TA comes from the sub-population African American.
- for SNP t1155a, the 8 heterozygote individual TA come from the sub-populations African American and European Caucasoid.
- for SNP a1204g, the 15 heterozygote individual AG come from the sub-populations Amerind, European and non-European Caucasoid, Mexican, Northeast Asian and South American.

Example 3. Expression of natural wild-type IFNα-21 and A42G, Q102K, Q114H/V127D, K179E mutated IFNα-21 in yeast

a) Cloning of the natural wild-type IFNα-21 and mutated IFNα-21 in the eukaryote expression vector pPicZα-topo

The nucleotide sequences coding for the mature part of the natural wild-type IFNα-21 and A42G, Q102K, Q114H/V127D, K179E mutated IFNα-21 are amplified by PCR using as template genomic DNA from an individual who is heterozygote for the corresponding SNP(s).

The PCR primers permitting such an amplification are:

SEQ ID NO. 19: Sense primer: TGTGATCTGCCTCAGACCCAC

SEQ ID NO. 20: Antisense primer: TCATTCCTTCCTCCTTAATCTTTCTTG

The PCR products are inserted in the eukaryote expression vector pPicZα-TOPO under the control of the hybrid promoter AOX1 inducible by methanol (TOPO™-*cloning*; Invitrogen Corp.).

This vector permits the heterologous expression of eukaryote proteins in the yeast *Pichia pastoris*.

After checking of the nucleotide sequence of the region of the vector coding for the recombinant proteins, the vector is linearized by the Pme1 restriction enzyme, and the *P. pastoris* yeast strain (Invitrogen) is transformed with these recombinant expression vectors.

b) Heterologous expression in *P. pastoris* and purification of the natural wild-type IFNα-21 and mutated IFNα-21 proteins

Two saturated pre-cultures of 50 mL of BMGY medium (2% Peptone, 1% yeast extract, 1.34% YNB, 1% Glycerol, 100 mM potassium phosphate, 0.4 mg/Liter biotin pH 6.0) containing a clone coding for natural wild-type IFNα-21 or that coding for A42G, Q102K, Q114H/V127D, or K179E mutated IFNα-21, were carried out for 24-48 hours at 30°C at an agitation of 200 rotations per minute (rpm).

When the culture reaches a saturating cellular density (corresponding to an optical density of 12 measured at a wavelength of 600 nm), it is used to inoculate, at 5 OD/mL, 250 mL of BMMY medium (2% Peptone, 1% yeast extract, 1.34% YNB, 0.5%

Methanol, 100 mM potassium phosphate, 0.4 mg/Liter biotin pH 6.0).

The expression of the protein is then induced by methanol at a final concentration of 1%, for 24 hours at 30 °C, with an agitation of the culture flask at 180 rpm.

Due to the presence of the signal peptide sequence of the "alpha factor", upstream of the coding sequence, the proteins are secreted by the yeasts in the culture medium. The alpha factor is naturally cleaved during the processing.

The suspension is centrifuged and the protein is purified by HPLC starting from the obtained supernatant.

In a pre-started step, an ultrafiltration (Labscale, cut-off 5000Da, Millipore) followed by a dialysis permits a ten times concentration of the yeast supernatant in a buffer of 50 mM Tris-Cl pH 9.0, 25 mM NaCl.

The first chromatographic step permits protein recovery by affinity on a blue sepharose column (Amersham Pharmacia). The presence of the protein in the collected fractions is verified, on the one hand by electrophoresis of SDS PAGE type and on the other hand by immuno-detection by a specific antibody directed against the IFNα-21 protein. At this step, the purity of the protein of interest is higher than 75%.

In a second purification step, a gel filtration permits buffer exchange of the collected fractions corresponding to IFNα-21 proteins against 50 mM Tris pH 9.0, 25 mM NaCl.

The last step of the purification consists of a separation of the proteins on an ion exchange chromatography column.

The fractions containing the recombinant protein are injected on an anion exchange column (ResourceQ 6.0 mL, Pharmacia) equilibrated beforehand in Tris 50 mM pH 9, NaCl 25 mM buffer. The elution of the proteins is carried out by the migration of a gradient between 0,025 and 1 M NaCl in the Tris 50 mM pH 9 buffer.

The purity of the protein of interest is estimated on SDS/PAGE gel and the protein concentrations were measured by densitometry (Quantity one, Biorad) and BCA assay (bicinchoninic acid and copper sulfate, Sigma).

Purified natural wild-type IFNα-21 and A42G, Q102K, Q114H/V127D, or K179E mutated IFNα-21 proteins obtained according to this protocols, eventually

scaled-up to produce higher amount of proteins, are used for the functional tests described below.

Example 4. Evaluation of immunomodulatory activity of natural wild-type IFNα-21 and A42G, Q114H/V127D, or K179E mutated IFNα-21

IFNs type I (IFN alpha and IFN beta) are able to modulate certain functions of the immune system. They have been demonstrated to increase the dendritic cells (DC) maturation: increase in the expression of MHC class I (HLA- ABC) and II (HLA-DR) molecules, increase in the expression of the molecules involved in the co-stimulation of the T-lymphocytes, CD80, CD86 and CD83 molecules and increase in the stimulating function of T-lymphocyte.

a) Effect of A42G, Q114H/V127D, or K179E mutated IFNα-21 on dendritic cell maturation.

Immunomodulatory activity of A42G, Q114H/V127D, or K179E mutated IFNα-21 was first investigated on dendritic cells maturation and compared to that of wild-type IFNα-2 chosen as representative of commercial Intron A product.

To do so, dendritic cells were first generated from adult peripheral blood monocytes cultivated in the presence of GM-CSF and IL-4 cytokines. After purification using a CD14+ cells purification kit, these dendritic cells were placed in presence of 100 ng/mL of wild-type IFNα-2 or A42G, Q114H/V127D, or K179E mutated IFNα-21 and their phenotype was determined by FACS analysis aiming at looking for the expression of the MHC class I and II molecules and the CD40, CD80, CD86, CD83 and CD1a markers. The maturation state of these dendritic cells has also been compared to that obtained without IFNα treatment, to provide a control with non-stimulated dendritic cells.

The median value of the measures of fluorescence intensity for each marker and for the three experimental conditions, expressed as arbitrary unit, are presented in the following table:

	HLA ABC	HLA DR	CD40	CD80	CD86	CD83	CD1a
No stimulation	64	133	24	25	14	15	26
Wild-type IFNα-2	87	281	331	76	45	15	155
A42G mutated IFNα-21	76	127	117	37	5	8	209
Q114H/V127D mutated IFNα-21	68	122	163	46	8	8	191
K179E mutated IFNα-21	181	322	491	87	44	14	127

The results of this test demonstrate that the K179E mutated IFNα-21 protein possesses a high capacity to stimulate dendritic cell maturation, this stimulatory effect being higher than that of the wild-type IFNα-2.

b) Effect of Q114H/V127D, or K179E mutated IFNα-21 on cytokine release by T lymphocytes

Immunomodulatory activity of Q114H/V127D, or K179E mutated IFNα-21 was also investigated by measuring cytokine release by T lymphocytes placed in presence of the corresponding mutated IFNα-21 proteins and with or without a strong antigen (SEB) in order to mimic an immune response against an aggression. This test was also performed in presence of wild-type IFNα-2 used as control and chosen as representative of the Intron A commercial product.

To do so, peripheral blood mononuclear cells (PBMC) were isolated from healthy donors and stimulated for 16 hours in an appropriate medium containing anti-CD3 and anti-CD28 antibodies or SEB. In each culture was added 4 µg/mL of wild-type IFNα-2 or Q114H/V127D, or K179E mutated IFNα-21. After stimulation, T lymphocytes were extracellularly labelled with anti-CD3, anti-CD4 and anti-CD69 antibodies or anti-CD3, anti-CD8 and anti-CD69 antibodies, and intracellularly labelled with specific antibodies directed against Th1-type cytokines (IFN-gamma) or Th2-type cytokines (IL-10). Fluorescent cells were analysed using FACScalibur and CellQuest software.

The results obtained indicate that mutated IFNα-21 proteins and wild-type

IFNα-2 do not stimulate IL-10 and IFN-gamma release and, thus, do not activate T lymphocytes in absence of SEB. In contrast, mutated IFNα-21 proteins and wild-type IFNα-2 stimulate cytokines (IL-10 and IFN-gamma) release by SEB-activated T-lymphocytes as shown in the table below. This table represents the cytokine release by T-lymphocytes in presence of SEB, expressed as percentage of the CD4+ CD69+ cells or CD8+ CD69+ cells for the CD4+ T-lymphocytes and CD8+ T-lymphocytes, respectively, and the percentage of CD69+ cells on total cells.

T-lymphocyte		IFN gamma	IL-10	CD69+ cells/total
CD4+ CD69+	Negative control	11.9	7.5	1.26
	Wild-type IFNα-2	19.6	24.68	2.7
	Q114H/V127D IFNα-21	38.9	14.6	4.67
	K179E IFNα-21	29.5	15.1	3.84
CD8+ CD69+	Negative control	8.73	0.65	4.69
	Wild-type IFNα-2	16.37	4.26	10.02
	Q114H/V127D IFNα-21	32.24	4.91	14.98
	K179E IFNα-21	28.28	3.8	13.48

These results clearly demonstrate that Q114H/V127D mutated IFNα-21 and K179E mutated IFNα-21 strongly stimulate cytokine release (IFN gamma and IL-10) by CD4+ and CD8+ T-lymphocytes previously activated by SEB antigen. In this test, the interferon gamma production by T-lymphocytes is higher in presence of Q114H/V127D mutated IFNα-21 or K179E mutated IFNα-21 than in presence of wild-type IFNα-2.

c) Effect of Q114H/V127D or K179E mutated IFNα-21 on cytokine release by monocytes

Finally, immunomodulatory activity of Q114H/V127D or K179E mutated IFNα-21 was investigated by measuring cytokine release by monocytes in absence or in presence of a bacterial toxic agent (LPS). This test was also performed in presence of

wild-type IFNα-2 used as control and chosen as representative of the Intron A commercial product.

To do so, human peripheral blood mononuclear cells (PBMC) were isolated from healthy donors and their phenotype was analyzed to determine the relative amount of CD64+ CD4dim cells (CD64 and CD4dim are markers for blood monocytes). After an over-night culture, these PBMC were incubated in the culture medium alone (not stimulated cells) or in presence of LPS (stimulated cells). In each culture, 4 µg/mL of wild-type IFNα-2 or mutated IFNα-21 was added. After culture, cells were extracellularly labelled with anti-CD64 and anti-CD4dim, and intracellularly labelled with specific antibodies directed against Th1-type cytokines (TNF-alpha), IL-12 and IL-10.

Fluorescent cells were analyzed using FACScalibur and CellQuest software.

The results obtained indicate that mutated IFNα-21 proteins and wild-type IFNα-2 do not stimulate cytokines (IL-10, IL-12 and TNF-alpha) release in absence of LPS. In contrast, in presence of LPS, Q114H/V127D and K179E mutated IFNα-21 proteins and wild-type IFNα-2 stimulate cytokines (IL-10, IL-12 and TNF-alpha) release by monocytes as shown in the table below. This table represents cytokine release by monocytes in presence of LPS, expressed as percentage of the CD64+ CD4dim cells, and the percentage of CD4dim CD64+ cells on total cells.

	IL-10	**IL-12**	**TNF-α**	**CD4dim CD64+ cells/total**
No stimulation	16.21	8.52	13.88	3.1
Wild-type IFNα-2	49.34	34.48	50.87	2.71
Q114H/V127D IFNα-21	50.63	31.81	56.5	2.31
K179E IFNα-21	60.14	36.42	60.16	4.43

Example 5. Evaluation of *in vitro* antiproliferative activity of A42G, Q102K, Q114H/V127D, and K179E mutated IFNα-21

a) on the human lymphoblasts of Daudi Burkitt's cell line proliferation

These tests are carried out on A42G, Q102K, Q114H/V127D, and K179E

mutated IFNα-21 proteins and wild-type IFNα-21 protein. Cells (human Daudi Burkitt's lymphoma cell line, hereinafter called "Daudi cells") cultivated beforehand in a RPMI 1640 medium (supplemented with 10% fetal calf serum and 2 mM of L-Glutamine) are inoculated in 96-well plates at the cellular density of 4.10^4 cells/ well.

In each well, Daudi cells are placed in contact of increasing concentrations of either mutated or wild-type IFNα-21 proteins. For each IFNα-21 to be characterized, final concentrations of 0.003 pM to 600 nM are tested.

The Daudi cells are then incubated for 66 h at 37 °C under 5% CO_2 after which the Uptiblue reagent (Uptima) is added to the cultures. The rate of cell proliferation is quantified by measuring the fluorescence emitted at 590nm (excitation 560nm) after an additional period of incubation of 4 hours.

The antiproliferative activity of mutated or wild-type IFNα-21 is based on the measurements of the IC50 corresponding to the concentration of IFNα-21 inhibiting 50% of the cell growth.

For each experimental condition, at least three experiments have been carried out in triplicates, which permits the determination of the average IC50 value for each IFNα-21. The ratio corresponding to the value of the IC50 of the mutated protein over the value of the wild-type protein permits the comparison. The results are collected in the following table (in brakets is noted the standard deviation):

	Wild-type IFNα-21	**A42G IFNα-21**	**Q102K IFNα-21**	**Q114H/V127D IFNα-21**	**K179E IFNα-21**
IC50 (pM)	1.02	2.55	1.05	13.92	3.73
Ratio wild-type/ mutated	-	2.15 (0.78)	1.30 (0.24)	13.10 (3.06)	3.72 (0.83)

This test demonstrates that the cellular antiproliferative activity on Daudi cells of the A42G, Q114H/V127D, and K179E mutated IFNα-21 proteins is lower than that of wild-type IFNα-21. In particular, the cellular antiproliferative activity on Daudi cells is approximately 10 to 16-fold lower in presence of Q114H/V127D IFNα-21 by comparison with wild-type IFNα-21.

b) on the TF-1 erythroleukemia cell line

The effect of A42G, Q114H/V127D, and K179E mutated IFNα-21 was also evaluated on TF-1 erythroleukemia cell line. This test was also performed in presence of wild-type IFNα-2 used as control and chosen as representative of the Intron A commercial product.

To do so, TF-1 cells were placed in contact of increasing concentrations of mutated IFNα-21 or wild-type IFNα-2 (0.001 to 1000 ng/mL) and the cell proliferation measured.

The results are expressed as the IC30 corresponding to the IFNα concentration inhibiting proliferation of 30% of cells and collected in the following table:

	Wild-type IFNα-2	**A42G IFNα-21**	**Q114H/V127D IFNα-21**	**K179E IFNα-21**
IC30 (ng/mL)	0.66	2.33	1.97	1.62

These data indicate that the three mutated IFNα-21 have a weak antiproliferative effect on TF-1 cells, and this effect is similar to that of wild-type IFNα-2, suggesting that the A42G, Q114H/V127D, and K179E mutated IFNα-21's hematologic toxicity is not superior than that of wild-type IFNα-2.

Example 6. Evaluation of the antiviral activity of A42G, Q114H/V127D, and K179E mutated IFNα-21

The IFNs play an important role in the antiviral defence. The IFN antiviral activity is partly due to IFNs induced enzymatic systems, such as:

- The 2'5' oligoadenylate synthetase, an enzyme which catalyzes the adenosine oligomere synthesis. These oligomeres activate the RNase L, an endoribonuclease which destroy the viral RNA once activated.
- The Mx proteins (GTPases) which inhibit the synthesis and/or the maturation of viral transcripts. This activity is mainly exerted on the influenza virus.
- The PKR protein (or p68 kinase) which is activated by the double-stranded RNA. The activated PKR inhibits protein synthesis.

The IFNs antiviral activity is also induced by other mechanisms such as, in the case of retroviruses, the inhibition of viral particles entry into the cells, the replication, the binding, the exit of the particles and the infective power of viral particles.

Finally, the IFNs exert an indirect antiviral activity by modulating certain functions of the immune system, in particular by favoring the response to cellular mediation (including an increase of the MHC class I and II molecules, increase of IL-12 and IFN-gamma production, increase of the CTL activities, among others).

The antiviral activity of A42G, Q114H/V127D, and K179E mutated IFNα-21 has been evaluated both *in vitro* in cell culture and *in vivo* in mouse model. Both tests have been carried out in parallel with wild-type IFNα-2 used as control and chosen as representative of the Intron A commercial product.

a) Antiviral activity *in vitro* in cell culture

This assay permits evaluation of the antiviral activity of A42G, Q114H/V127D, and K179E mutated IFNα-21 and wild-type IFNα-2 in cell culture using the vesicular stomatitis virus (VSV).

To do so, WISH human epithelial cells are cultivated for 24 hours in the presence of decreasing concentrations of mutated IFNα-21 and wild-type IFNα-2. Then, the cells are infected by the virus of vesicular stomatitis (VSV) during 24 to 48 additional hours and the cell lysis is measured.

The antiviral effect of the different IFNα tested is determined by comparing the IC50 value corresponding to the IFN concentration inhibiting 50% of cell lysis induced by the VSV.

A similar experiment has been carried out two times, and the average IC50 values measured are presented in the following table:

	Wild-type IFNα-2	**A42G IFNα-21**	**Q114H/V127D IFNα-21**	**K179E IFNα-21**
IC50 (ng/mL)	4	14	22	25

Thus, in cell culture infected with VSV, the A42G, Q114H/V127D, and K179E mutated IFNα-21 have a lower antiviral activity than the wild-type IFNα-2.

b) Antiviral activity *in vivo* in mouse model

This test *in vivo* is performed in EMCV (Encephalomyocarditis virus) mouse model.

Human IFNs exhibit dose-dependent antiviral activity in the mouse which is in general 100 to 1,000 fold less than that exhibited by the same amount of mouse IFN (Meister et al. (1986). J. Gen. Virol. 67, 1633-1644).

Intraperitoneal injection of mice with Encephalomyocarditis virus (EMCV) gives rise to a rapidly progressive fatal disease characterized by central nervous system involvement and encephalitis (Finter NB (1973). Front Biol. 2: 295-360). Mouse and human interferon-alpha have both been shown to be effective in protecting mice against lethal EMCV infection (Tovey and Maury (1999). J. IFN Cytokine Res. 19: 145-155).

Groups of 20, six-week old Swiss mice were infected intraperitoneally (ip) with 100 x LD_{50} EMCV and treated one hour later, and then once daily for 3 days thereafter with 2 μg of A42G, Q114H/V127D, K179E mutated IFNα-21 or wild-type IFNα-2 preparations. A control group was performed with animals having been treated with excipient only. The animals were followed daily for survival for 21 days.

Results are presented in Figure 6 and indicate that the relative survival rate of the mice which have been treated with A42G, Q114H/V127D, or K179E mutated IFNα-21 is much higher than the survival rate of the non-treated mice but remains similar to that observed for the mice which have been treated with wild-type IFNα-2.

All of these results demonstrate that A42G, Q114H/V127D, and K179E mutated IFNα-21 possess unique biological properties.

30 SEP 2005

CLAIMS

1. An isolated polynucleotide comprising:
   a) a nucleotide sequence having at least 80% identity with the nucleotide sequence SEQ ID NO. 1, or its coding sequence, provided that such nucleotide sequence comprises at least one SNP selected from the group consisting of c973a, g1011c, t1049a, t1155a, a1204g, and t1265c; or
   b) a nucleotide sequence complementary to a nucleotide sequence under a).
2. An isolated polynucleotide according to claim 1, characterized in that the nucleotide sequence under a) which comprises at least one SNP selected from c973a, g1011c, t1049a, t1155a, a1204g and t1265c, is at least 90% identical with SEQ ID NO. 1, or its coding sequence.
3. An isolated polynucleotide according to claim 1, characterized in that the nucleotide sequence under a) which comprises at least one SNP selected from c973a, g1011c, t1049a, t1155a, a1204g and t1265c, is at least 95% identical with SEQ ID NO. 1, or its coding sequence.
4. An isolated polynucleotide according to claim 1, characterized in that the nucleotide sequence under a) which comprises at least one SNP selected from c973a, g1011c, t1049a, t1155a, a1204g and t1265c, is at least 99% identical with SEQ ID NO. 1, or its coding sequence.
5. An isolated polynucleotide characterized in that it encodes a polypeptide comprising the amino acid sequence SEQ ID NO. 2, and having at least one coding SNP selected from Q102K, Q114H, V127D, C162stop and K179E.
6. An isolated polynucleotide according to any one of claims 1 to 5, characterized in that it comprises c973g and/or encodes the coding SNP K179E.
7. Use of all or part of an isolated polynucleotide according to any one of claims 1 to 6 in order to identify, hybridize and/or amplify all or part of a polynucleotide having 80 to 100 % identity with the nucleotide sequence SEQ ID NO. 1, or its coding sequence, provided that each one of these sequences comprises at least one of the following SNPs: c973a, g1011c, t1049a, t1155a, a1204g, and t1265c.
8. Use of all or part of an isolated polynucleotide according to any one of claims 1 to 6 for the genotyping of all or part of a polynucleotide having 80 to 100 % identity with the nucleotide sequence SEQ ID NO. 1, or its coding sequence, provided that

each one of these sequences comprises at least one of the following SNPs: c973a, g1011c, t1049a, t1155a, a1204g, and t1265c.

9. A recombinant vector comprising a polynucleotide according to any one of claims 1 to 6.

10. A host cell comprising a recombinant vector according to claim 9.

11. A method for separating a polypeptide, comprising cultivating a host cell according to claim 10 in a culture medium and separating said polypeptide from the culture medium.

12. An isolated polypeptide characterized in that it is encoded by the isolated polynucleotide according to any one of claims 1 to 4.

13. An isolated polypeptide comprising an amino acid sequence having at least 80% identity with:

a) the amino acid sequence SEQ ID NO. 2; or

b) the amino acid 24 through 189 of the amino acid sequence SEQ ID NO. 2;

provided that the amino acid sequence under a) or b) contains at least one coding SNP selected from the group consisting of Q102K, Q114H, V127D, C162stop, and K179E.

14. An isolated polypeptide according to claim 13, characterized in that the amino acid sequence identity with said sequences under a) or b) is 90%.

15. An isolated polypeptide according to claim 13, characterized in that the amino acid sequence identity with said sequences under a) or b) is 95%.

16. An isolated polypeptide according to claim 13, characterized in that the amino acid sequence identity with said sequences under a) or b) is 99%.

17. An isolated polypeptide comprising amino acids 24 through 189 of the amino acid sequence SEQ ID NO. 2 and having at least one coding SNP selected from the group consisting of A42G, Q102K, Q114H, V127D, C162stop, and K179E.

18. An isolated polypeptide according to claim 17, and having the two coding SNPs Q114H and V127D.

19. An isolated polypeptide according to claim 17, and having the coding SNP K179E.

20. An antibody immunospecific for a polypeptide according to any one of claims 12 to 19.

21. A method for identifying an agent among one or more compounds to be tested which activates or inhibits the activity of an isolated polypeptide comprising all or

part of amino acid sequence SEQ ID NO. 2 and having at least one coding SNP selected from the group consisting of A42G, Q102K, Q114H, V127D, C162stop, and K179E, said method comprising:

a) providing host cells comprising the recombinant vector according to claim 9 containing at least one coding SNP selected from the group consisting of c794g, c973a, g1011c, t1049a, t1155a, and a1204g;

b) contacting said host cells with said compounds to be tested,

c) determining the activating or inhibiting effect upon the activity of said polypeptide whereby said activating or inhibiting agent is identified.

22. A method for analyzing the biological characteristics of a subject, comprising performing at least one of the following steps:

a) Determining the presence or the absence of the polynucleotide according to any one of claims 1 to 6 in the genome of a subject;

b) Determining the level of expression of the polynucleotide according to any one of claims 1 to 6 in a subject;

c) Determining the presence or the absence of the polypeptide according to any one of claims 12 to 20 in a subject;

d) Determining the concentration of the polypeptide according to any one of claims 12 to 20 in a subject; or

e) Determining the functionality of the polypeptide according to any one of claims 12 to 20 in a subject.

23. A therapeutic agent comprising one or more compounds selected from the group consisting of an isolated polynucleotide according to any one of claims 1 to 6; a recombinant vector according to claim 9; a host cell according to claim 10, an isolated polypeptide according to any one of claims 12 to 20; an antibody according to claim 21.

24. Use of one or more compounds selected from the group consisting of an isolated polynucleotide according to any one of claims 1 to 6; a recombinant vector according to claim 9, a host cell according to claim 10, an isolated polypeptide according to any one of claims 12 to 20, or an antibody according to claim 21for the preparation of a medicament for preventing or treating in a subject a disease selected from the group consisting of cancers and tumors, infectious diseases, immunologically and auto-immunologically related diseases, cardiovascular diseases, metabolic diseases, central nervous system diseases, disorders connected with chemotherapy treatments,

anemia in dialyzed patients, wound healing, and/or osteoporosis.

25. Use according to claim 24, characterized in that said cancers and tumors comprise metastasizing renal carcinomas, melanomas, lymphomas comprising follicular lymphomas, and cutaneous T cell lymphoma, leukemias comprising hairy-cell leukemia, chronic lymphocytic leukemia and chronic myeloid leukemia, cancers of the liver, neck, head and kidneys, multiple myelomas, carcinoid tumors and tumors that appear following an immune deficiency comprising Kaposi's sarcoma in the case of AIDS.

26. Use of one or more compounds selected from the group consisting of an isolated polynucleotide according to any one of claims 1 to 6; a recombinant vector according to claim 9, a host cell according to claim 10, an isolated polypeptide according to any one of claims 12 to 20, or an antibody according to claim 21 for preparing a medicament for increasing or decreasing the activity of the polypeptide according to claim 12.

27. Use of one or more compounds selected from the group consisting of an isolated polynucleotide according to any one of claims 1 to 6; a recombinant vector according to claim 9, a host cell according to claim 10, an isolated polypeptide according to any one of claims 12 to 20, or an antibody according to claim 21 for preparing a medicament for treating a disorder or a disease linked to the presence of the polynucleotide according to claim 1 in the genome.

28. A method for determining statistically relevant associations between at least one SNP selected from the group consisting of c794g, c973a, g1011c, t1049a, t1155a, a1204g, and t1265c in the IFNα-21 gene, and a disease or resistance to disease comprising:

a) Genotyping a group of individuals;

b) Determining the distribution of said disease or resistance to disease within said group of individuals;

c) Comparing the genotype data with the distribution of said disease or resistance to disease; and

d) Analyzing said comparison for statistically relevant associations.

29. A method for diagnosing or determining a prognosis of a disease or a resistance to a disease comprising detecting at least one SNP selected from the group consisting of c973a, g1011c, t1049a, t1155a, a1204g, and t1265c in the IFNα-21 gene.