Genes and Immunity (2003) 4, 441–449

www.nature.com/gene

Genomic analysis of Th1–Th2 cytokine genes in an AIDS cohort: identification of IL4 and IL10 haplotypes associated with the disease progression

A Vasilescu[1,2], SC Heath[2], R Ivanova[1,2], H Hendel[1], H Do[1,2], A Mazoyer[1,2], E Khadivpour[1], FX Goutalier[1], K Khalili[3], J Rappaport[3], GM Lathrop[2], F Matsuda[2] and J-F Zagury[1,3]

[1]*Centre de Recherche des Cordeliers, Université Pierre et Marie Curie, Paris, France;* [2]*Centre National de Génotypage, 2, rue Gaston Crémieux, Évry, France;* [3]*Center for Neurovirology and Cancer Biology, Temple University, Philadelphia, PA, USA*

Polymorphisms of Th1–Th2 cytokine genes have previously been implicated in the rate of progression to AIDS in seropositive patients. To evaluate further the impact of these genes in the development of AIDS, we have performed an extensive genetic analysis of IL2, IL4, IL6, IL10, IL12p35 and p40, IL13 and IFNγ. The coding regions and promoters of these genes were sequenced in a Caucasian cohort of 337 HIV-1 seropositive extreme patients (the GRIV cohort) consisting of patients with slow progression and rapid progression, and up to 470 healthy controls. In all, 64 single nucleotide polymorphisms (SNPs) and four deleterious polymorphisms with frequency $> 1\%$ were identified and evaluated for their association with disease. Statistically significant associations were observed with haplotypes of the IL4 and IL10 genes, but no relation was found with variants of other genes. The catalogue of SNP and haplotypes presented here will facilitate further genetic investigations of Th1–Th2 cytokines in AIDS and other immune-related disorders.
Genes and Immunity (2003) **4,** 441–449. doi:10.1038/sj.gene.6363983

Keywords: *Th1–Th2 cytokines; AIDS progression; genotyping; single nucleotide polymorphism; haplotype; GRIV cohort*

Introduction

Cytokines have a vital role within the immune system, and cytokine genes have been targeted for studies for association to susceptibility and development of several immune-related diseases.[1,2,4–9] Such genetic investigations can provide key information on both the function of such genes in the immune system and the molecular etiology of the diseases in question, and they are expected to lead to the development of new therapies and diagnostic tools.[3]

Th1–Th2 cytokines are of particular interest as targets for genetic studies since they are the fundamental messengers of adaptive immunity and, as such, are likely to be involved in pathogenic mechanisms. In the case of AIDS, although many hypotheses have been presented to explain the molecular mechanisms leading to CD4 cell depletion,[10–15] the most conclusive and reproducible observations implicate Th1–Th2 cytokines, for example through the so-called Th1/Th2 shift[15] and the great fragility of the cells leading to their apoptosis.[13] Thus, genetic studies of Th1–Th2 cytokines are of particular relevance in this disease. Although to date, most genetic studies in AIDS have focused on the chemokine/chemokine receptor system and HLA,[16,17] promoter polymorphisms in a limited number of Th1–Th2 cytokine genes have been examined for association with the disease,[18–20] some of which have been correlated to the level of production of cytokine.[21] In particular, IL10 and IL4 variants have been reported to be associated with disease progression to AIDS.[18,19] These studies underscore the possible involvement of Th1–Th2 cytokines in the development of the disease, and the interest of a complete and systematic exploration of these other Th1–Th2 cytokine genes in AIDS. In regard to the IL10 and IL4 variants that have been implicated in disease progression, an important next step is to seek confirmation in independent cohorts, as such confirmation is a key element in the validation of genetic association.

Studies of the relation between disease progression in AIDS and genetic variants have largely used cohorts of seropositive patients analyzed with respect to their progression to disease.[22] In the present study, we have taken an alternative approach based on the comparison between patient groups having extreme patterns of disease progression. We identified a group of 253 seropositive patients with slow progression (SP) of the disease, which represents a 1% subset of patients with extreme phenotypes among a cohort of 25 000 patients at all stages of disease and a group of 84 patients with rapid progression (RP). This cohort, called GRIV (genetics of resistance to immunodeficiency virus), is the largest of its kind in the world and its usefulness has been validated in studies of other genes including CCR5 and HLA.[23–25] We have performed a systematic genetic study of eight Th1–Th2 cytokine genes, IL2, IL4, IL6, IL10, IL12 p35 and p40, IL13 and IFNγ, by extensive nucleotide sequencing of a group of up to 807 Caucasian individuals, including the 337 members of the GRIV

Correspondence: Dr J-F Zagury, Centre de Recherche des Cordeliers, 15, rue de l'Ecole de Médecine, 75006 Paris, France. E-mail : jfz@ccr.jussieu.fr

cohort and 470 Caucasian controls of similar ethnic origin. Single nucleotide polymorphisms (SNPs) and other genetic variants identified in through sequencing were evaluated for their association with disease susceptibility and progression.

Results

SNP discovery

We sequenced the exons, flanking regions (up to 1000 base pairs) and promoter regions to detect polymorph-

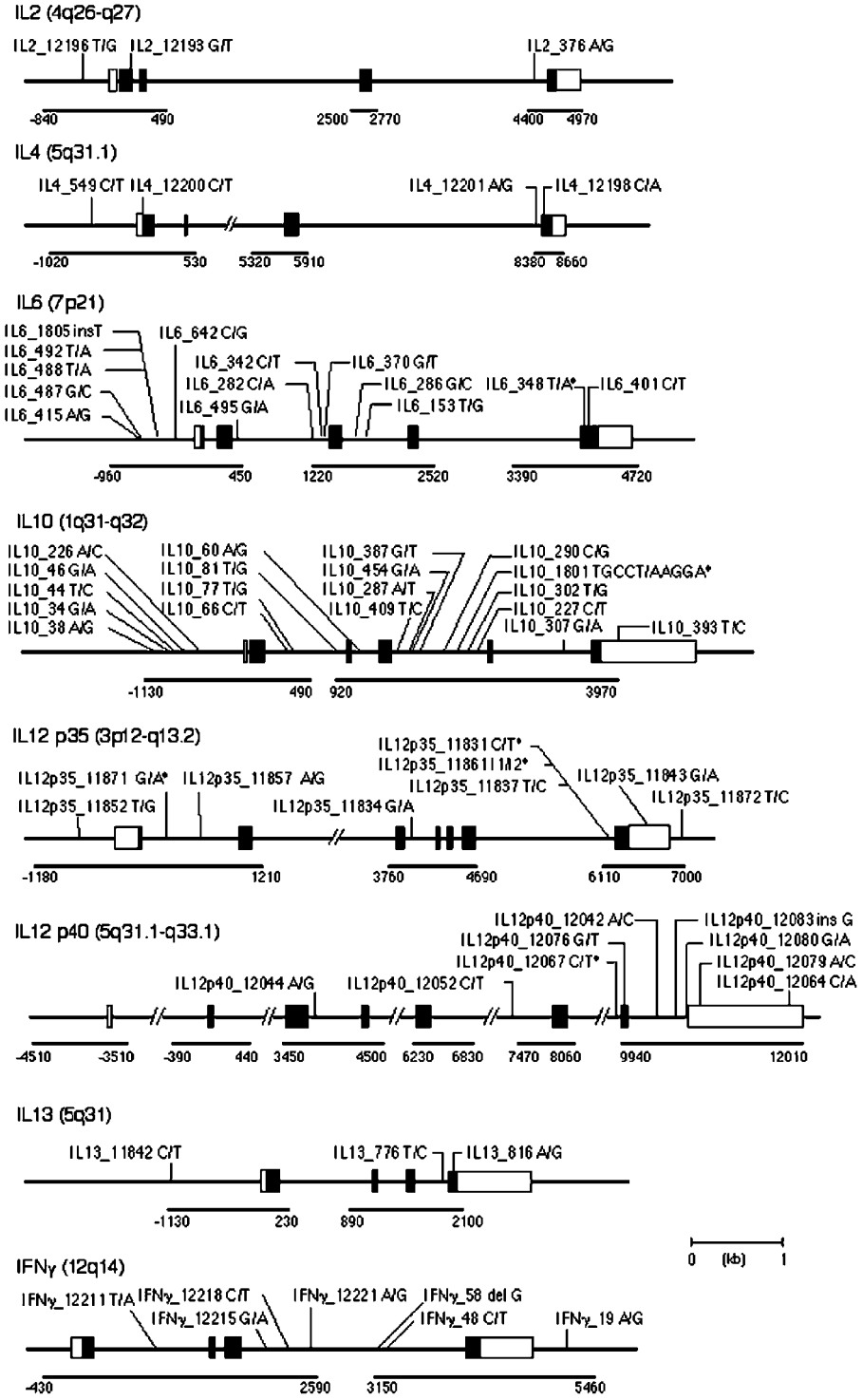


Figure 1 Organization of the eight cytokine genes investigated in the study. Coding and untranslated regions of each gene are indicated by solid and open rectangles, respectively. The position of 68 polymorphisms with frequencies of greater than 1% is indicated with the name and nucleotide changes. The asterisk corresponds to polymorphisms that are newly characterized in this study. The regions that have been sequenced are shown by a horizontal line below each gene with start and end positions according to the first nucleotide of the initiation codon as +1.

isms in IL10, IL4 and the six other cytokine genes included in the study in a group of up to 807 Caucasian subjects (Figure 1). Initially, we included DNA from 154 controls for sequencing, but this was extended to 470 control samples for gene regions that contained polymorphisms showing potential association with progression. Of 205 polymorphisms identified, 68 polymorphisms (64 SNPs and four insertion/deletion polymorphisms) have an allele frequency of 1% or greater in the combined cohorts (Table 1). These include promoter polymorphisms in IL10 and IL4 that were reported in other studies to be associated with progression to AIDS.[18,19] A cluster of five SNPs in intron 3 of the IL10 gene that exhibit complete linkage disequilibrium (LD) (TGCCT or AAGGA) in our data have been assigned a single SNP designation (IL10_1801) here (see Figure 1 and Table 2). The polymorphism adjacent to exon 7 designated as IL12p35_11861 is an insertion of multiple nucleotides of either 6 bp (I1: GTGGCA) or 14 bp (I2: TTGGATAGGGTTATGCCT) in length that are exclusive to each other.

Overall, we found one polymorphism in every 154 bp and one with a frequency higher than 1% in every 464 bp, in agreement with previous genomic studies (Table 1).[41,42] Compared to the other cytokines, IL10 appeared to have a greater number of polymorphisms (one polymorphism with more than 1% frequency in every 246 bp). Table 2 summarizes genotype counts and allele frequencies of each polymorphism in the SP, RP and control groups, their associations with AIDS progression (see below), and the relevant information known to date for each polymorphism. The genotype distributions were compatible with Hardy–Weinberg equilibrium in the SP, RP and control subgroups for all polymorphisms.

Ten of the 64 frequent SNPs were located in exons. Three of these introduce nonsynonymous changes, namely, IL6_348 (Asp to Glu at position 162), IL12p40_12076 (Phe to Val at position 298) and IL13_816 (Glu to Arg at position 130). Two other variants with a synonymous change are IL2_12193, IL6_401 and the other five are located either in the 5′-untranslated region (IL4_12200) or the 3′-untranslated region (IL10_393, IL12p40_12080, 12079 and 12064). For the 58 intronic polymorphisms, comparisons of human and mouse sequences around did not show any specific conservation indicating interspecies functional constraints.

Several SNPs located in promoter regions of the gene studies here have previously been reported to affect the production of cytokines. An extensive study of the IL10 promoter showed that IL10_46 and IL10_226 are in the negative regulatory region of transcription, suggesting their functional involvement in IL10 production.[43] IL10_38 is located in the ETS-like recognition site.[43] Measures of IL10 expression *in vitro*, however, are still contradictory, probably because of different protocols or the limited number of samples tested.[4,27] The T allele of IL4_549 was shown to be associated with higher production of this cytokine in the Japanese population.[21] For the other promoter polymorphisms, the degree of evolutionary conservation was higher than in the intron between mouse and human; however, computer searches failed to localize any to motifs that were recognized as related to transcription. These polymorphisms are candidates for experimental evaluation to elucidate their biological significance.

Linkage disequilibrium and haplotypes

Figure 2 shows pairwise LD measured between each pair of polymorphisms using D and D' statistics as described in Materials and methods. It should be noted that the maximum value of D depends on the allele frequencies of the polymorphisms involved (which is why the D' values are also shown). As expected, the majority of high LD values occur within genes. In some instance, a high LD measure was observed between polymorphisms in different genes. Mostly these values are related to a relatively rare variant that is detected on few chromosomes, as these may be associated with variants in other genes by chance. For example, IL10_302, which has a frequency of 0.02 (15 heterozygotes observed in the study cohorts), is in LD (as measured by D') with polymorphisms in most of the other genes examined. In the case of IL4 and IL13, however, the strong LD that is observed between several polymorphisms in these genes can be attributed to their physical proximity on chromosome 5 (12.8 kb apart).

Haplotype frequencies were estimated for each gene based on an EM algorithm. The results are summarized in Figure 2 for those haplotypes with frequency estimates of 1% or more. The average number of haplotypes per gene with frequencies of >1 or $>5\%$ in the control population is 6.3 (range: 3–12) and 4.0 (range: 2–5), respectively. Cumulatively, haplotypes with frequencies

Table 1 Incidence of SNPs identified in Th1–Th2 cytokine genes

		Number of polymorphisms (number of insertions/deletions)			*Frequency (bp/polymorphism)*	
Gene	*bp sequenced*	<1%	>1%	Total	>1%	All
IL2	2170	8 (2)	3 (0)	11 (2)	723	197
IL4	2420	8 (0)	4 (0)	12 (0)	605	202
IL6	4040	13 (0)	14 (1)	27 (1)	289	150
IL10	4670	21 (0)	19 (0)	37 (0)	246	126
IL12p35	4210	21 (1)	9 (1)	30 (2)	468	140
IL12p40	6140	26 (3)	9 (1)	35 (4)	682	175
IL13	2570	16 (2)	3 (0)	20 (2)	643	129
IFNg	5330	24 (3)	7 (1)	31 (4)	761	172
Total	31 550	137 (11)	68 (4)	205 (15)	464	154

Table 2 Summary of polymorphisms of Th1–Th2 cytokine genes

							P-values for test statistics				*References to previous study of the variant in AIDS or for evaluation of its biological effect*
SNP				*Freq A1*			*CvSvR*		*SvR*		
Gene	*ID*	*A1*	*A2*	*CTR*	*SP*	*RP*	*Nominal*	*Adjusted*	*Nominal*	*Adjusted*	
IL2	12196	T	G	0.69	0.68	0.69	0.902	0.983	0.847	0.957	(26), IL2 production (27)
	12193	G	T	0.69	0.68	0.66	0.867	0.975	0.684	0.887	(26)
	376	A	G	0.74	0.73	0.74	1.000	1.000	1.000	1.000	NCBI (rs2069772)
IL4	549	C	T	0.81	0.82	0.87	0.206	0.400	0.146	0.267	(28), progression (19), IL4 production (21)
	12200	C	T	0.84	0.83	0.90	0.098	0.216	**0.034**	0.072	(29), IL4 production (21)
	12201	A	G	0.81	0.84	0.88	0.176	0.352	0.259	0.437	NCBI (rs2243289)
	12198	C	A	0.81	0.84	0.88	0.225	0.428	0.259	0.438	NCBI (rs2243290)
IL6	415	A	G	0.64	0.58	0.64	0.167	0.131	0.308	0.841	NCBI (rs1800797)
	487	G	C	0.96	0.96	0.91	0.064	0.423	**0.043**	0.251	(30)
	488	T	A	0.63	0.59	0.64	0.406	0.356	0.384	0.901	(30)
	492	T	A	0.61	0.59	0.57	0.599	0.396	0.668	0.990	(30)
	1805	-	T	0.76	0.72	0.73	0.368	0.183	0.905	1.000	(30)
	642	C	G	0.63	0.58	0.61	0.336	0.182	0.628	0.990	(31)
	495	G	A	1.00	0.97	0.99	0.229	0.194	0.667	0.990	NCBI (rs2069831 & 2069858)
	282	C	A	0.37	0.39	0.43	0.486	0.371	0.480	0.955	NCBI (rs1474347 & 3087231)
	342	C	T	0.95	0.95	0.97	0.633	0.885	0.420	0.927	NCBI (rs1524107 & 3087233)
	370	G	T	0.95	0.95	0.96	0.846	0.887	0.609	0.986	NCBI (rs2066992)
	286	C	G	0.65	0.66	0.63	0.740	0.802	0.486	0.958	NCBI (rs2069840)
	153	T	G	0.59	0.59	0.56	0.829	0.715	0.654	0.990	NCBI (rs1554606)
	348	T	A	0.98	0.99	0.98	0.683	0.667	0.638	0.990	**NEW**
	401	C	T	0.97	0.98	0.98	0.425	0.316	0.489	0.960	NCBI (rs2069849)
IL10	38	A	G	0.61	0.55	0.51	**0.042**	0.347	0.403	0.955	(32), IL10 production (27)
	34	G	A	0.97	0.95	0.99	0.102	0.584	0.057	0.412	(33)
	44	T	C	0.28	0.25	0.24	0.247	0.841	0.916	0.999	(32)
	46	G	A	0.98	1.00	0.98	**0.008**	0.104	0.100	0.590	(33)
	226	A	C	0.28	0.25	0.24	0.277	0.869	0.916	0.999	(32), progression (18), IL10 production (4)
	66	C	T	0.41	0.44	0.50	0.270	0.862	0.241	0.856	NCBI (rs2222202)
	77	T	G	0.70	0.74	0.75	0.705	0.994	0.880	0.999	NCBI (rs3024490)
	81	T	G	0.29	0.27	0.17	0.189	0.768	0.112	0.625	NCBI (rs1518110)
	60	A	G	0.25	0.23	0.19	0.246	0.840	0.483	0.975	NCBI (rs1518111 & rs3024506)
	409	T	C	0.22	0.17	0.14	**0.020**	0.209	0.440	0.965	(34)
	287	A	T	0.78	0.75	0.75	0.321	0.901	1.000	1.000	NCBI (rs3024492)
	454	G	A	0.98	0.98	0.99	0.585	0.983	0.467	0.973	NCBI (rs3024507)
	387	G	T	0.88	0.88	0.80	0.064	0.450	**0.040**	0.313	NCBI (rs3024493)
	290	C	G	0.60	0.53	0.53	0.168	0.734	0.852	0.998	NCBI (rs1878672)
	1801	T	A	0.98	0.98	0.98	0.828	0.998	0.719	0.995	**NEW**
	302	T	G	1.00	0.98	0.97	0.154	0.707	0.486	0.975	NCBI (rs3024508)
	227	T	C	0.95	0.94	0.97	0.295	0.883	0.139	0.694	NCBI (rs3024509)
	307	G	A	0.88	0.86	0.82	0.090	0.547	0.233	0.850	NCBI (rs3024495)
	393	T	C	0.61	0.55	0.52	**0.028**	0.260	0.523	0.982	NCBI (rs3024496)
IL12p35	11852	T	G	0.90	0.89	0.92	0.517	0.892	0.296	0.296	NCBI (rs2243115)
	11871	G	A	0.76	0.74	0.71	0.389	0.820	0.360	0.360	**NEW**
	11857	G	A	0.51	0.51	0.52	0.934	0.993	0.787	0.787	NCBI (rs647801)
	11834	G	A	0.96	0.98	0.97	0.435	0.850	0.713	0.713	NCBI (rs2243130 & rs 2243129)
	11831	C	T	0.96	0.99	1.00	0.074	0.340	1.000	1.000	**NEW**
	11861	I2	I1	0.85	0.82	0.85	0.429	0.846	0.402	0.402	**NEW**
	11837	T	C	0.87	0.90	0.91	0.333	0.775	0.765	0.765	NCBI (rs2243136)
	11843	A	G	0.87	0.89	0.91	0.287	0.732	0.643	0.643	NCBI (rs568408)
	11872	C	T	0.89	0.91	0.93	0.518	0.893	0.661	0.661	NCBI (rs640039)
					I1=TTGGATAGGGTTATGCCT I2=GTGGCA						
IL12p40	12044	A	G	0.81	0.77	0.79	0.495	0.916	0.661	0.661	(35)
	12052	C	T	0.81	0.76	0.79	0.326	0.793	0.490	0.490	NCBI (rs2421047)
	12067	C	T	0.88	0.88	0.79	0.416	0.870	0.320	0.320	**NEW**
	12076	G	T	0.97	0.97	0.99	0.417	0.870	0.262	0.262	NCBI (rs3213119).
	12042	A	C	0.85	0.83	0.77	**0.043**	0.208	0.111	0.111	NCBI (rs2853697)
	12083	-	G	0.77	0.77	0.79	0.885	0.994	0.740	0.740	NCBI (rs3213120)
	12080	G	A	0.95	0.96	0.98	0.367	0.830	0.374	0.374	NCBI (rs3213120)
	12079	A	C	0.82	0.82	0.84	0.899	0.995	0.690	0.690	(36)
	12064	C	A	0.82	0.84	0.82	0.690	0.975	0.607	0.607	NCBI (rs1368439)

Table 2 (*continued*)

							P-values for test statistics				*References to previous study of the variant in AIDS or for evaluation of its biological effect*
	SNP			*Freq A1*			*CvSvR*		*SvR*		
Gene	*ID*	*A1*	*A2*	*CTR*	*SP*	*RP*	*Nominal*	*Adjusted*	*Nominal*	*Adjusted*	
IL13	11842	C	T	0.81	0.78	0.81	0.382	0.698	0.566	0.848	(37), IL13 production (38)
	776	T	C	0.79	0.76	0.82	0.316	0.621	0.189	0.451	(37)
	816	A	G	0.81	0.76	0.83	0.116	0.336	0.122	0.347	(37), IgE production (37)
IFNG	12211	T	A	0.46	0.46	0.43	0.892	0.991	0.714	0.957	(39)
	12215	A	G	0.74	0.73	0.71	0.845	0.985	0.760	0.957	(20)
	12218	T	C	0.73	0.72	0.72	0.939	0.995	0.837	0.970	(20)
	12221	G	A	N.A.	0.90	0.93	0.694	0.953	0.694	0.943	NCBI (rs2069716)
	58	G	-	0.78	0.77	0.78	0.991	0.999	1.000	1.000	NCBI (rs2069718)
	48	C	T	0.63	0.64	0.56	0.145	0.422	0.069	0.226	NCBI (rs2069718)
	19	G	A	0.47	0.45	0.39	0.161	0.459	0.159	0.440	(40)

Nature and frequencies of the polymorphisms studied in the control, SP and RP groups, *P*-values for statistical tests comparing the frequencies in the three groups and in SP *vs* RP, and references to polymorphisms that have previously been investigated in AIDS or for biological effects. dbSNP IDs are also given for those that are already reported and registered in dbSNP. Nominal *P*-values of <0.05 are shown in bold. The adjusted *P*-values have been corrected for the number of polymorphisms within each gene taking account of LD. CvSvR means CTR versus SP versus RP, SvR means SP versus RP.

of greater than 1% (5%) account for 97% (92%) of the total number of chromosomes tested on average per gene.

SNP and haplotype associations with HIV-1

Allele and haplotype frequencies were compared between SP and RP groups, and between SP, RP and controls using contingency table analysis. Nominal *P*-values (not adjusted for multiple testing) <0.05 were found for eight SNPs (Table 2). While none of the results were significant after adjustment for testing multiple polymorphisms within the same gene using a permutation test (see Methods), it is of interest that five of these SNPs were found in IL10 and one in IL4. Polymorphisms in these genes have previously been implicated in progression to AIDS.[18,19]

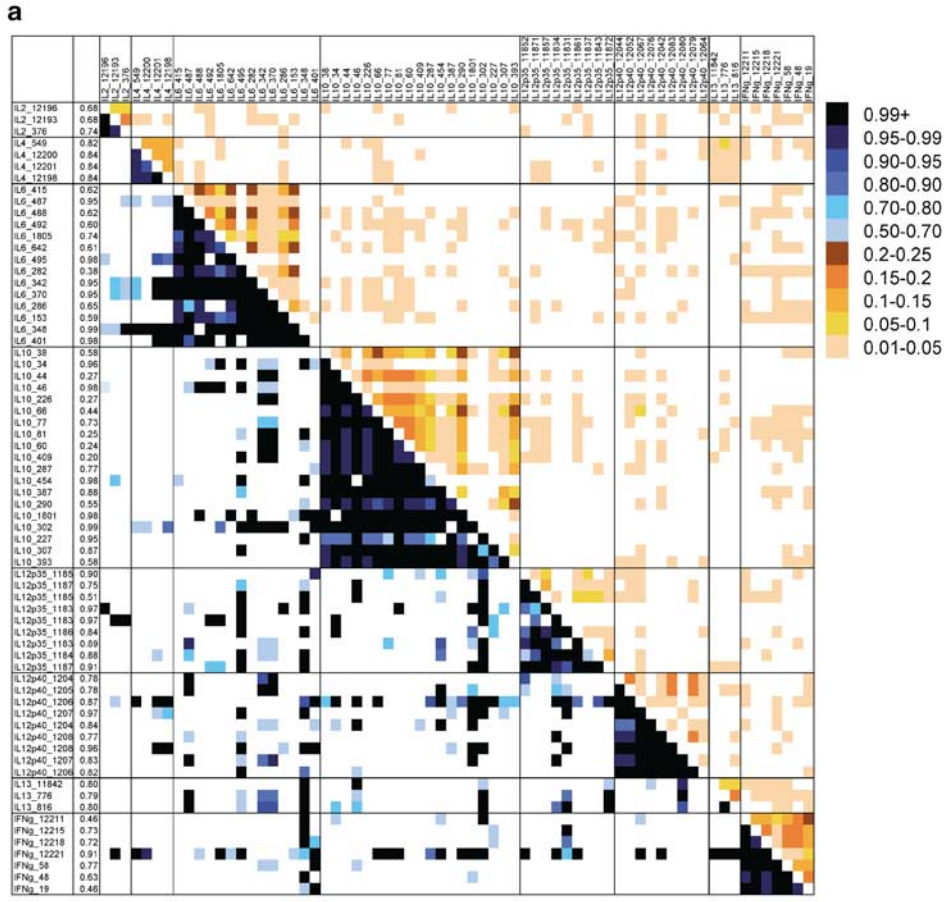
Haplotype analysis provided further insight into the possible relation between SNPs in these genes and disease (Table 3). Two of the 10 IL10 haplotypes (IL10-H4 and IL10-H10) with estimated frequencies >0.01 were found to have nominal *P*-values <0.05 in the association tests, but only one of these was significant after the *P*-values were adjusted for the number of tests (IL10-H10). IL10-H4 is a relatively common haplotype with a higher frequency in the RP group compared to others. The frequency of this haplotype is 0.188 in the RP group, compared to 0.126 in the SP group (relative risk (RR) $=1.5$; nominal $P=0.042$) and 0.115 in controls (RR $=1.7$; nominal $P=0.011$). As shown in Figure 2, IL10-H4 is the only IL10 haplotype that carries a T allele at the IL10_387 site or an A allele at the IL10_307 site. The second of these haplotypes (IL10-H10) is rare overall, but with a lower frequency in the SP group compared to others: 0.002 in the SP group compared to 0.018 in the RP group (RR $=9.0$; nominal $P=0.05$) and 0.023 in controls (RR $=11.5$; nominal $P=0.0013$). IL10-H10 is distinguished from other IL10 haplotypes by the presence of an A allele at the IL10_46 site in the promoter region.

Shin *et al*[18] have implicated the less frequent allele at the IL10_226 (A allele) promoter site in acceleration of progression to AIDS. The IL10-H4 haplotype, which is more frequent in the RP compared to SP groups (ie compatible with acceleration of progression), carries the alternative C allele at this site. Overall, we found no correlation in our data between the presence of the A allele at this site and rapid progression. Similarly, the less frequent allele at the IL10_38 site (G allele) is not associated with rapid progression in our data as found by Shin *et al*[18] in their cohorts.

In the IL4 gene, a single haplotype (IL4-H2) exhibited nominally significant different frequencies in the RP compared to the SP cohorts (RR $=1.9$; nominal $P=0.016$); the association remained significant after adjustment for the number of tests. The frequency of IL4-H2 in the SP group was similar to that in controls (0.16 *vs* 0.15), and lower in the RP group (0.09). The comparison of the RP cohort and control cohort frequency gave an RR $=1.7$ (nominal $P=0.05$). IL4-H2 is distinguished from the most common haplotype (IL4-H1; frequency $>80\%$ in all the groups we studied) at all four polymorphic sites identified in the gene. IL4-H2 carries the T allele at the IL4_549 site consistent with the results of Nakayama *et al*[19], who reported slower progression to AIDS associated with the IL4_549 T allele.

Discussion

We have undertaken a systematic investigation of the association of genetic variation in eight Th1–Th2 cytokines to susceptibility and disease progression in AIDS. In all, 68 variants of the eight genes that occurred with frequencies of 1% or greater were detected. Although the majority of these reside outside of exons, promoters or other regions that are known to have biological function, all were considered as possible candidates for association



b

IL2				
	Freq	12196	12193	376
H1	0.374	T	G	A
H2	0.312	G	G	A
H3	0.250	T	T	G
H4	0.056	T	T	A

IL4				
	Freq	549	12200	12201	12198
H1	0.821	C	C	A	C
H2	0.149	T	T	G	A
H3	0.011	T	C	A	C

IL10																				
	Freq	38	34	44	46	226	66	77	81	60	409	287	454	387	290	1801	302	227	307	393
H1	0.312	T	G	C	G	C	C	G	G	G	C	A	G	G	C	TGCCT	T	T	G	T
H2	0.184	C	G	C	G	C	T	G	G	G	C	T	G	G	G	TGCCT	T	T	G	C
H3	0.168	T	G	T	G	A	C	T	T	A	T	A	G	G	C	TGCCT	T	T	G	T
H4	0.115	C	G	C	G	C	T	G	G	G	C	A	G	T	G	TGCCT	T	T	A	C
H5	0.036	T	G	T	G	A	C	T	T	A	C	A	G	G	C	TGCCT	T	T	G	T
H6	0.025	C	A	C	G	C	T	G	G	G	C	A	G	G	G	TGCCT	T	C	G	C
H7	0.019	T	G	T	G	A	C	T	G	G	C	A	G	G	C	TGCCT	T	T	G	T
H8	0.021	T	G	T	G	A	C	T	T	A	T	A	A	G	C	TGCCT	T	T	G	T
H9	0.020	C	G	C	G	C	T	G	G	G	C	A	G	G	G	TGCCT	T	C	G	C
H10	0.023	T	G	T	A	A	C	T	T	A	T	A	G	G	C	TGCCT	T	T	G	T
H11	0.016	C	G	C	G	C	T	G	G	G	C	T	G	G	G	4AGGA	T	T	G	C
H12	0.010	C	G	C	G	C	T	G	G	G	C	T	G	G	G	TGCCT	G	T	G	C

IL12p40									
	Freq	12044	12052	12067	12076	12042	12083	12080	12079	12064
H1	0.283	A	C	C	G	A	-	G	A	A
H2	0.224	G	T	C	G	A	G	G	C	A
H3	0.18	A	C	C	G	A	-	G	A	C
H4	0.151	A	C	C	G	C	-	G	A	A
H5	0.117	A	C	T	G	A	-	G	A	A
H6	0.032	A	C	T	T	A	-	A	A	A

IFNg							
	Freq	12211	12215	12218	12221	58	48	19
H1	0.353	T	A	T	A	G	C	G
H2	0.224	A	G	C	A	-	T	A
H3	0.164	A	A	T	A	G	C	A
H4	0.101	T	A	T	G	G	C	G
H5	0.100	A	A	T	A	G	T	A
H6	0.041	A	G	C	A	G	T	A

IL13				
	Freq	11842	776	816
H1	0.719	C	C	G
H2	0.108	T	T	A
H3	0.081	C	T	A
H4	0.074	T	C	G
H5	0.011	C	T	G

IL12p35									
	Freq	11852	11871	11857	11834	11831	11861	11837	11843	11872
H1	0.323	T	G	A	G	C	I2	T	G	T
H2	0.236	T	A	G	G	C	I2	T	G	T
H3	0.152	T	G	A	G	C	I1	T	G	T
H4	0.116	T	G	G	G	C	I2	T	A	C
H5	0.095	G	G	G	G	C	I2	C	G	T
H6	0.030	T	G	G	A	C	I2	T	G	T
H7	0.016	T	G	A	G	T	I2	T	G	T

I1 = TTGGATAGGGTTATGCCT
I2 = GTGGCA

IL6															
	Freq	415	487	488	492	1805	642	495	282	342	370	286	153	348	401
H1	0.352	A	G	T	T	-	C	G	C	C	G	C	T	T	C
H2	0.168	G	G	A	T	-	G	G	A	C	G	G	G	T	C
H3	0.148	G	G	A	A	T	G	G	A	C	G	G	G	T	C
H4	0.099	G	G	A	A	-	G	G	A	C	G	C	G	T	C
H5	0.078	G	G	A	A	T	G	G	A	C	T	G	C	T	C
H6	0.047	G	C	A	A	-	G	G	A	T	T	C	G	T	C
H7	0.016	G	G	A	T	-	G	G	A	C	G	C	T	T	T
H8	0.014	G	G	A	T	-	G	G	A	C	G	G	G	A	C
H9	0.010	G	G	A	A	T	G	A	A	C	G	G	G	T	C

Figure 2 (**a**) LD plot of eight cytokine genes. The pattern of pairwise LD is shown with the *D* values above the diagonal and the *D*′ values below as described in Materials and Methods, and estimated allele frequencies for each polymorphism. Different colors are used to represent ranges of positive *D* and *D*′ values. (**b**) Summary of polymorphism combinations and estimated frequencies for haplotypes in the eight cytokine genes.

with AIDS, as genetic studies have previously identified polymorphisms with no predicted biological effects. Significant associations were observed only with haplotypes of the IL4 and IL10 genes. Although these associations were weak, we considered them to be of particular interest because polymorphisms of these genes have previously been implicated in progression to AIDS.

Our data were compatible with the reported association of IL4_549 T allele with slower progression to AIDS. The two IL10 promoter variants, IL10_226 and IL10_38, which are reported to be correlated with rate of disease

Table 3 Frequencies and *P*-values from test statistics comparing haplotype frequencies in the control, SP and RP groups for IL4 and IL10

Gene	*Hap*	*CTR*		*SP*		*RP*		*Nominal p-values*		*Adjusted p-values*	
		Counts	*Freq*	*Counts*	*Freq*	*Counts*	*Freq*	*CvSvR*	*SvR*	*CvSvR*	*SvR*
IL4	H1	718/874	0.821	407/500	0.815	145/168	0.861	0.349	0.159	0.620	0.323
	H2	130/874	0.149	82/500	0.164	15/168	0.088	0.053	0.016	0.113	0.035
	H3	10/874	0.011	3/500	0.006	2/168	0.012	0.636	0.604	0.903	0.877
IL10	H1	294/942	0.312	148/506	0.293	45/168	0.267	0.455	0.556	1.000	1.000
	H2	174/942	0.184	106/506	0.210	32/168	0.193	0.522	0.659	1.000	1.000
	H3	158/942	0.168	78/506	0.153	20/168	0.120	0.278	0.312	0.989	0.994
	H4	108/942	0.115	64/506	0.126	32/168	0.188	0.029	0.042	0.294	0.412
	H5	34/942	0.036	18/506	0.036	7/168	0.040	0.886	0.814	1.000	1.000
	H6	24/942	0.025	20/506	0.040	2/168	0.012	0.141	0.129	0.870	0.843
	H7	18/942	0.019	12/506	0.023	7/168	0.039	0.173	0.279	0.923	0.990
	H8	19/942	0.021	11/506	0.022	1/168	0.008	0.472	0.312	1.000	0.995
	H9	19/942	0.020	8/506	0.015	2/168	0.012	0.817	1.000	1.000	1.000
	H10	22/942	0.023	1/506	0.002	3/168	0.018	0.003	0.050	0.024	0.481
	H11	14/942	0.015	7/506	0.015	3/168	0.019	0.907	0.716	1.000	1.000
	H12	10/942	0.010	7/506	0.014	4/168	0.026	0.366	0.480	0.998	1.000

The allele combinations for the haplotypes are shown in Figure 2b.

progression elsewhere,[18] had statistically similar genotype frequencies in controls and the two disease groups in our study. However, we did find some evidence of possible association of disease progression with other IL10 sites. One of the IL10 haplotypes that we identified as possibly associated with progression (nominally significant P-value < 0.05) is the unique haplotype carrying the IL10_387 T and the IL10_307 A variants. Although neither of these sites is predicted to have a functional effect, it is possible that one or both variants have biological consequences that are as yet unknown. The second IL10 haplotype, for which the association remained significant after adjustment for the number of tests, is the unique haplotype carrying the IL10_46 A variant. This site is in the IL10 5′ promoter region.

The experimental GRIV groups have been voluntarily chosen with extreme patterns of progression in order to unravel differences in their genetic distributions, correlating with their phenotypes. Hence, the study design applied here involves the comparison of two extreme cohorts of patients and controls, whereas previous investigations implicating IL4 and IL10 were based on longitudinal follow-up of patients. In general, variants that affect risk or speed of developing AIDS should be detectable with either study design. For example, the CCR5-delta32 variant, which has a major protective effect on progression to AIDS in longitudinal studies,[44] gives an RR of 6.3 and a nominal P-value 0.0001 for comparison of the allele frequencies in the SP and RP groups from GRIV based on data presented elsewhere.[23] For a marker such as the IL10_226 A variant, previously reported to be associated with decreased risk of progression as described above, the GRIV study provides a power of 75% to detect an RR of approximately 0.6 in SP compared to controls at a significance level of 0.05, assuming an allele frequency of 0.28 in controls as observed here (implying an allele frequency of 0.19 in SP).

Our data provide some additional evidence for the possible involvement of two Th2 cytokines, IL10 and IL4. IL4 is reported to regulate differentially the expression of CCR5 and CXCR4 and to stimulate the expression of HIV-1, which leads to accelerated disease progression.[45] *In vitro* studies demonstrated the suppressive role of IL10 against HIV production either by direct inhibition of viral replication[46] or through suppression of the autocrine production of TNFα and IL6 that enhance HIV replication.[47] More importantly, an *in vivo* study showed the inhibitory effect of IL10 on acute HIV infection.[48] In the light of our observations and the previously published reports implicating these genes in progression to AIDS,[18,19] further studies in other cohorts to clarify the possible pathogenic role of IL10 and IL4 variants (IL4_H2 and IL10_H10) are now seen to be imperative. The data that we have catalogued on polymorphisms, LD and haplotype patterns in the Th1–Th2 cytokine genes studied here can serve for these investigations in AIDS, and also in other immune-related disorders.

Materials and methods

GRIV patients and control subjects

The GRIV cohort was established in 1995 in France to generate a large collection of DNAs for studies on genetic variations associated with rapid and slow progression of AIDS.[25] To avoid the confounding effects associated with racial/ethnic differences in the genetic analyses, only Caucasians of European descent living in France were recruited. This criterion limits the influence of the virogenetic and environmental factors (subjects are all infected by B strains and live in a similar environment) and puts emphasis on the genetic make-up of each individual to determine the various patterns of progression. Patients with SP were defined as asymptomatic individuals seropositive for 8 or more years with a CD4 cell count above 500/mm^3 in the absence of antiretroviral therapy. A seropositive test dating more than 8 years was necessary for inclusion in the study. Patients with RP show a drop in their CD4 cell count below 300/mm^3 in less than 3 years after the last seronegative test. All patients enrolled in the study signed an informed

consent. The DNA was obtained from fresh peripheral blood mononuclear cells or from EBV-transformed cell lines. The control series consisted of 470 French Caucasians from the EGEA study, ascertained without respect to disease status or seropositivity.

DNA genotyping

For each gene, primers were established in order to amplify by PCR the exon-containing DNA fragments and the promoters. The PCRs were performed in a 15 μl reaction mixture containing 25 ng of DNA. The list of the primers for each gene is available on the web site of the CNG (www.cng.fr). Sequencing reactions were performed according to the dye terminator method using an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA). Alignment of experimental results, SNP discovery and genotyping were performed with the software Genalys, developed by the CNG.[49] Detailed information of all the SNPs including those with a frequency of less than 1% is also available from the CNG web site. The genomic sequences used for the alignment are IL2 (NT_022960.5), IL4 (NT_007072.7), IL6 (NT_025766.6), IL10 (NT_029217.3), IL12p35 (NT_005818.7), IL12p40 (NT_007006.7), IL13 (NT_007072.9) and IFNγ (NT_029419.2).

Statistical analysis

LD estimates between pairs of polymorphisms within genes were obtained by estimating the two polymorphism haplotype frequencies using the EM algorithm.[50] Haplotype frequencies using all polymorphisms for each gene were also estimated with the EM algorithm. Differences in the allele frequencies of individual polymorphisms between the three groups and between the SP and RP groups were examined using a Fisher's exact test on the resulting 2×3 and 2×2 table of counts. The nominal *P*-values were adjusted for the number of polymorphisms within each gene by a permutation test preserving the observed LD between the markers. For this, 100 000 replicates were produced by randomly permuting the group labels. Fisher's exact test was performed on all polymorphisms as in the original analysis, and the corrected significance was estimated from the proportion of replicates giving a minimum *P*-value at one site equal to or lower than the observed *P*-value. Differences between haplotype frequencies were examined in an analogous way to that used for the individual polymorphisms. For each haplotype (with frequency $>1\%$) in turn, the expected numbers of individuals in each group with and without that haplotype were computed from the estimated marginal haplotype probability distribution for each individual. These numbers were rounded to the nearest integer, and a nominal *P*-value was computed using Fisher's exact test. Corrected *P*-values were obtained using a permutation test in the same way as for the individual polymorphism analysis. Computation of *D* and *D*′ in Figure 2: let f, p, and q be the frequency of the ab haplotype, the a allele at locus 1, and the b allele at locus 2 respectively. The standard disequilibrium measure *D* is given by f-q *q. This measure is highly dependent on the allele frequencies, so we also calculated the standardized disequilibrium measure, *D*′, which ranges from 0 (no disequilibrium) to 1 (the highest amount of disequilibrium given the alleles frequencies).

Acknowledgements

We are grateful to all the patients and all the medical staff who have generously collaborated with the GRIV project. We also thank the EGEA cooperative group for having given access to data on the EGEA (Epidemiological study on the Genetics and Environment of Asthma) study, which was partly supported by an INSERM/MSD convention. The GRIV project is supported by grants from ANRS (Agence Nationale de Recherche sur le SIDA) and ACV (AIDS-Cancer Vaccine Development Foundation). The CNG is supported by the Ministere de la Recherche et des Nouvelles Technologies.

References

1 O'Shea JJ, Ma A, Lipsky P. Cytokines and autoimmunity. *Nat Rev Immunol* 2002; **2**: 37–45.
2 Bidwell J, Keen L, Gallagher G *et al*. Cytokine gene polymorphism in human disease: on-line databases, supplement 1. *Genes Immun* 2001; **2**: 61–70.
3 Hill AV. Immunogenetics and genomics. *Lancet* 2001; **357**: 2037–2041.
4 Gibson AW, Edberg JC, Wu J *et al*. Novel single nucleotide polymorphisms in the distal IL-10 promoter affect IL-10 production and enhances the risk for systemic lupus erythematosus. *J Immunol* 2001; **166**: 3915–3922.
5 Lee JY, Goldman D, Piliero LM, Petri M, Sullivan KE. IFN-gamma polymorphisms in systemic lupus erythematosus. *Genes Immun* 2001; **2**: 254–257.
6 Khani-Hanjani A, Lacaille D, Hoar D *et al*. Association between dinucleotide repeat in non-coding region of interferon-gamma gene and susceptibility to, and severity of rheumatoid arthritis. *Lancet* 2000; **356**: 820–825.
7 Constantin A, Navaux F, Lauwers-Cances V *et al*. Interferon-gamma gene polymorphism and susceptibility to, and severity of rheumatoid arthritis. *Lancet* 2001; **358**: 2051–2052.
8 Morahan G, Huang D, Ymer SI *et al*. Linkage disequilibrium of a type 1 diabetes susceptibility locus with a regulatory IL12B allele. *Nat Genet* 2001; **27**: 218–221.
9 Dahlman I, Eaves IA, Kosoy R *et al*. Parameters for reliable results in genetic association studies in common diseases. *Nat Genet* 2002; **30**: 149–150.
10 Fisher AG, Collalti E, Ratner L, Gallo RC, Wong-Staal F. A molecular clone of HTLV-III with biological activity. *Nature* 1985; **316**: 262–265.
11 Sodroski J, Goh WC, Rosen C, Campbell K, Haseltine WA. Role of the HTLV-III/LAV envelope in syncitium formation and cytopathicity. *Nature* 1986; **322**: 470–474.
12 Golding H, Robey FA, Gates III FT *et al*. Identification of homologous regions in human immunodeficiency virus 1 gp41 and human MHC class II beta 1 domain 1. Monoclonal antibodies against gp41-derived peptide and patients' sera react with native HLA class II antigens, suggesting a role for autoimmunity in the pathogenesis of AIDS. *J Exp Med* 1988; **167**: 914–923.
13 Ameisen JC, Capron A. Cell dysfunction and depletion in AIDS: the programmed cell death hypothesis. *Immunol Today* 1991; **12**: 102–105.
14 Imberti L, Sottini A, Bettinardi A, Puoti M, Primi D. Selective depletion in HIV infection of T cells that bear specific T cell receptor V beta sequences. *Science* 1991; **254**, 860–862.
15 Clerici M, Shearer GM. A Th1 to Th2 switch is a critical step in the etiology of HIV infection. *Immunol Today* 1993; **14**: 107–111.
16 O'Brien SJ, Moore JP. The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. *Immunol Rev* 2000; **177**: 99–111.
17 Al Jabri AA. HLA and *in vitro* susceptibility to HIV infection. *Mol Immunol* 2002; **38**: 959–967.

18 Shin HD, Winkler C, Stephens JC *et al*. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci USA* 2000; **97**: 14467–14472.
19 Nakayama EE, Meyer L, Iwamoto A *et al*. Protective effect of IL-4–589T polymorphism on HIV-1 disease progression: relationship with viral load. *J Infect Dis* 2002; **185**: 1183–1186.
20 Bream JH, Carrington M, O'Toole S *et al*. Polymorphisms of the human IFNG noncoding regions. *Immunogenetics* 2000; **51**: 50–58.
21 Nakashima H, Miyake K, Inoue Y *et al*. Association between IL-4 genotype and IL-4 production in the Japanese population. *Genes Immun* 2002; **3**: 107–109.
22 O'Brien SJ, Nelson GW, Winkler CA, Smith MW. Polygenic and multifactorial disease gene association in man: lessons from AIDS. *Annu Rev Genet* 2000; **34**: 563–591.
23 Rappaport J, Cho YY, Hendel H *et al*. 32 bp CCR-5 gene deletion and resistance to fast progression in HIV-1 infected heterozygous. *Lancet* 1997; **349**: 922–923.
24 Hendel H, Henon N, Lebuanec H *et al*. Distinctive effects of CCR5, CCR2 and SDF1 genetic polymorphisms on AIDS progression. *J AIDS* 1998; **19**: 381–386.
25 Hendel H, Caillat-Zucman S, Lebuanec H *et al*. New class I and II HLA alleles strongly associated with opposite patterns of progression to AIDS. *J Immunol* 1999; **162**: 6942–6946.
26 John S, Turner D, Donn R *et al*. Two novel biallelic polymorphisms in the IL-2 gene. *Eur J Immunogenet* 1998; **25**: 419–420.
27 Hoffmann SC, Stanley EM, Darrin Cox E *et al*. Association of cytokine polymorphic inheritance and *in vitro* cytokine production in anti-CD3/CD28-stimulated peripheral blood lymphocytes. *Transplant* 2001; **72**: 1444–1450.
28 Rosenwasser LJ, Klemm DJ, Dresback JK *et al*. Promoter polymorphisms in the chromosome 5 gene cluster in asthma and atopy. *Clin Exp Allergy* 1995; **2**: 74–78.
29 Takabayashi A, Ihara K, Sasaki Y *et al*. Novel polymorphism in the 5′-untranslated region of the interleukin-4 gene. *J Hum Genet* 1999; **44**: 352–353.
30 Osiri M, McNicholl J, Moreland LW, Bridges Jr SL. A novel single nucleotide polymorphism and five probable haplotypes in the 5′ flanking region of the IL-6 gene in African Americans. *Genes Immun* 1999; **1**: 166–167.
31 Fishman D, Faulds G, Jeffery R *et al*. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. *J Clin Invest* 1998; **102**: 1369–1376.
32 Turner DM, Williams DM, Sankaran D *et al*. An investigation of polymorphism in the interleukin-10 gene promoter. *Eur J Immunogenet* 1997; **24**: 1–8.
33 D'Alfonso S, Rampi M, Rolando V, Giordano M, Momigliano-Richiardi P. New polymorphisms in the IL-10 promoter region. *Genes Immun* 2000; **1**: 231–233.
34 Donger C, Georges JL, Nicaud V *et al*. New polymorphisms in the IL-10 gene — relationships to myocardial infarction. *Eur J Clin Invest* 2001; **31**: 9–14.
35 Huang D, Cancilla MR, Morahan G. Complete primary structure, chromosomal localisation, and definition of polymorphisms of the gene encoding the human interleukin-12 p40 subunit. *Genes Immun* 2000; **1**: 515–520.
36 Hall MA, McGlinn E, Coakley G *et al*. Genetic polymorphism of IL-12 p40 gene in immune-mediated disease. *Genes Immun* 2000; **1**: 219–224.
37 Graves PE, Kabesch M, Halonen M *et al*. A cluster of seven tightly linked polymorphisms in the IL-13 gene is associated with total serum IgE levels in three populations of white children. *J Allergy Clin Immunol* 2000; **105**: 506–513.
38 van der Pouw Kraan TC, van Veen A, Boeije LC *et al*. An IL-13 promoter polymorphism associated with increased risk of allergic asthma. *Genes Immun* 1999; **1**: 61–65.
39 Awata T, Matsumoto C, Urakami T *et al*. Association of polymorphism in the interferon gamma gene with IDDM. *Diabetologia* 1994; **37**: 1159–62.
40 Wu S, Muhleman D, Comings DE. G5644A polymorphism in the interferon-gamma (INFG) gene. *Psychiatr Genet* 1998; **8**: 57.
41 Wang DG, Fan JB, Siao CJ *et al*. Large-scale identification, mapping, and genotyping of SNPs in the human genome. *Science* 1998; **280**: 1077–1082.
42 Cargill M, Altshuler D, Ireland J *et al*. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* 1999; **22**: 231–238.
43 Kube D, Platzer C, von Knethen A *et al*. Isolation of the human IL-10 promoter. Characterization of the promoter activity in Burkitt's lymphoma cell lines. *Cytokine* 1995; **7**: 1–7.
44 Dean M, Carrington M, Winkler C *et al*. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 1996; **273**:1856–1862.
45 Valentin A, Lu W, Rosati M *et al*. Dual effect of interleukin-4 on HIV-1 expression: implications for phenotypic switch and disease progression. *Proc Natl Acad Sci USA* 1998; **95**: 8886–8891.
46 Montaner LJ, Griffin P, Gordon S. Interleukin-10 inhibits initial reverse transcription of human immunodeficiency virus type 1 and mediates a virostatic latent state in primary blood-derived human macrophages in vitro. *J Gen Virol* 1994; **75**: 3393–3400.
47 Weissman D, Poli G, Fauci AS. Interleukin 10 blocks HIV replication in macrophages by inhibiting the autocrine loop of tumor necrosis factor alpha and interleukin 6 induction of virus. *AIDS Res Hum Retroviruses* 1994; **10**: 1199–1206.
48 Kollmann TR, Pettoello-Mantovani M, Katopodis NF *et al*. Inhibition of acute *in vivo* human immunodeficiency virus infection by human interleukin 10 treatment of SCID mice implanted with human fetal thymus and liver. *Proc Natl Acad Sci USA* 1996; **93**: 3126–3131.
49 Takahashi M, Matsuda F, Margetic N, Lathrop M. Automated identification of single nucleotide polymorphisms from sequencing data. *J Bioinf Computat Biol Proc. IEEE CS Bioinformatics Conference* 2002, pp 87–96.
50 Laird N. *Handbook of Statistics Vol. 9, Computational Statistic: the EM algorithm*. Elsevier Science Publishers BV: Amsterdam, 1993, pp 509–520.