bjh research paper

Haplotypes in the tumour necrosis factor region and myeloma

Gareth J. Morgan,[1] Peter J. Adamson,[2] Fiona K. Mensah,[2] Charlotte F. Spink,[3] Graham R. Law,[2] Leigh J. Keen,[3] Eve Roman,[2] Faith E. Davies,[1] Sara Rollinson,[4] J. A. Child[4] and Jeffrey L. Bidwell[3]

[1]Section of Haemato-Oncology, Institute for Cancer Research, Royal Marsden Hospital, London, [2]Epidemiology and Genetics Unit, Leukaemia Research Fund Epidemiology and Genetics Unit, University of York, York, [3]Department of Pathology and Microbiology, University of Bristol, Bristol, and [4]Department of Haematology and Oncology, University of Leeds, Leeds, UK

Received 16 December 2004; accepted for publication 24 February 2005

Correspondence: Gareth Morgan, Professor of Haematology, Haemato-Oncology Unit, Royal Marsden Hospital, Downs Road, Sutton, Surrey SM2 5PT, UK. E-mail: gareth.morgan@rmh.nhs.uk

Summary

This study described the haplotypic structure across a region of chromosome 6 including the tumour necrosis factor (TNF) gene, and investigated its influence on the aetiology of myeloma. A total of 181 myeloma cases from the Medical Research Council Myeloma VII trial and 233 controls from the Leukaemia Research Fund Case Control Study of Adult Acute Leukaemia were included in the analysis. Genotyping by induced heteroduplex generator analysis was carried out for single nucleotide polymorphisms (SNP) located at positions −1031, −863, −857, −308 and −238 of the 5′ promoter region of TNF-α gene, and 252 in the LT-α gene; and five microsatellites, TNF*a*, *b*, *c*, *d* and *e*. Haplotypes were inferred statistically using the PHASE algorithm. A limited diversity of haplotypes was observed, with the majority of variation described by 12 frequent haplotypes. Detailed characterization of the haplotype did not provide greater determination of disease risk beyond that described by the TNF-α −308 SNP. Some evidence was provided for a decreased risk of myeloma associated with the TNF-α −308 variant allele A, odds ratio, 0·57; 95% confidence interval, 0·38–0·86. The results of this study did not support our starting hypothesis; that high producer haplotypes at the TNF locus are associated with an increased risk of developing myeloma.

Keywords: genetic variation, tumour necrosis factor and lymphotoxin alpha haplotypes, microsatellites, chromosome 6p, class III human leucocyte antigen.

The tumour necrosis factor (TNF) cytokine is central to pro-inflammatory response, and is therefore a strong candidate for mediating the risk of developing immunologically related malignant disease. The TNF is involved in T-cell-dependent B-cell responses, T-cell proliferation, natural killer (NK)-cell activity and dendritic cell maturation. It is of particular relevance to myeloma, where it acts as a recognized growth and survival factor for myeloma cells. The TNF gene region is embedded within the human leucocyte antigen (HLA) class III region on chromosome 6. This highly polymorphic region, 6p21, contains a number of immune response genes including those of lymphotoxin alpha (LT-α), leucocyte-specific transcript-1 (LST1) and allograft inflammatory factor (AIF). Thus far, 15 TNF single nucleotide polymorphisms (SNP) have been identified, mainly in the 5′-promoter region. A number of microsatellite polymorphisms, existing together with these SNP in extended haplotypes, have also been described (Nedospasov *et al*, 1991); Over larger genomic distances TNF microsatellites are inherited in combination with class I and class II HLA types (Degli-Esposti *et al*, 1995); which may also influence TNF responses (Udalova *et al*, 1993).

Prior studies of disease risk associated with genetic variation and myeloma have targeted 'candidate genes' for investigation. Often, candidate polymorphisms have been chosen based on inferred functionality and analysed in isolation to the genetic structure of the rest of the chromosome. An alternative approach is to use haplotype structure, inferred from allelic variants at a number of loci, which may be used to give a representation of a chromosomal region (Stephens *et al*, 2001; Stephens & Donnelly, 2003). This genetic information is then used to study disease association and is not based on prior assumptions of function for the variants analysed. In this study, we explore the influence of genetic variation, on the aetiology of myeloma, across a chromosomal region that includes the TNF gene through consideration of haplotypes, inferred using seven SNP and five microsatellites. These were chosen for examination based on a reasonably high minor allele frequency and previous evidence of association in the literature, with the overall aim of having a combination of

microsatellites and SNP that would provide a manageable number of haplotypes for statistical analysis.

Materials and methods

Study subjects

A total of 181 participants from the Medical Research Council (MRC) Myeloma VII trial (Child *et al*, 2003), were included in the analysis. Cases diagnosed under the age of 65, between 1993 and 2000, in the UK were included in the trial. This represents a young population of myeloma cases compared with the UK distribution of myeloma, because of trial entry criteria. Blood or bone marrow samples, taken at presentation or at follow-up, were used to prepare DNA from 233 UK population-based controls from the Leukaemia Research Fund Case Control Study of Adult Acute Leukaemia (Kane *et al*, 1999), covering a comparable age range to the myeloma cases, were used. DNA was prepared from blood samples provided by the controls.

Genotyping

Five SNP located in the 5′-promoter region of TNF-α were studied at positions −1031 (T–C substitution), −863 (C–A), −857 (C–T), −308 (G–A) and −238 (G–A); and SNP at position +489 (G–A) located in intron 1 of the TNF-α gene and +252 (A–G) in the LT-α gene. In addition, five TNF region dinucleotide repeat microsatellites were analysed termed TNF*a*, *b*, *c*, *d* and *e* (Udalova *et al*, 1993). TNF*a*, *b* and *c* are located in the LT-α gene, TNF*d* and *e* in intron IV of the LST-1 gene.

The SNP were genotyped using induced heteroduplex generator (IHG) analysis (Morse *et al*, 1999), a high-throughput polymerase chain reaction (PCR)-based genotyping tool. Briefly, an artificial DNA construct was designed that mimicked the sequence flanking the SNP of interest. Adjacent to the SNP, a polynucleotide insertion or deletion was added to the construct. Genomic DNA and construct were then amplified in separate PCR reactions using identical primers and PCR parameters. The two amplicons were mixed and cross-matched by heating at 95°C for 5 min followed by controlled ramping to 37°C over 30 min. This leads to the formation of artificial heteroduplexes between complementary strands of the genomic DNA and IHG construct. The polynucleotide insertion/deletion causes a distortion of the resulting heteroduplex and consequent retardation during gel electrophoresis. The banding patterns of retarded heteroduplexes can then be used to identify both homozygous and heterozygous individuals. Basic PCR parameters were as follows: PCR mixes (20 μl) contained 0·5 μmol/l each of forward and reverse primers, 1·5 mmol/l $MgCl_2$, 200 μmol/l of each dNTP, 1X *Taq* polymerase buffer (67 mmol/l Tris–HCl pH 8·8, 16 mmol/l $(NH_4)_2SO_4$, 0·01% v/v Tween), 0·2 units *Taq* polymerase (Advanced Biotechnologies, Columbia, MD, USA) and either 50 ng of genomic DNA or diluted IHG reagent. Primers and annealing temperatures are detailed in Appendix 1.

Samples were electrophoresed using 12–15% polyacrylamide gels (37·5:1 acrylamide:bisacrylamide, Protogel; National Diagnostics, Atlanta, GA, USA) for 50 (12%) or 90 (15%) min on a triple wide minigel format. Resulting gels were poststained with either ethidium bromide or SYBR green stain and digitally imaged on a Kodak imaging system. For the microsatellite genotyping, PCR was carried out as for the SNP with previously described primers (Udalova *et al*, 1993). Samples were electrophoresed on 12% polyacrylamide gels (19:1 acrylamide: bisacrylamide, Accugel; National Diagnostics) for 90–135 min and poststained with ethidium bromide.

Statistical analysis

The Hardy–Weinberg equilibrium was tested, for each locus, within the case and control groups. Distributions of alleles and genotypes, in case and control groups, were compared using permutation testing (Good, 2000). Disease association was described using odds ratios (OR) and 95% confidence intervals (95% CI) estimated by logistic regression. *P*-values of 0·01 or less were considered to provide evidence for statistical significance in the analysis.

Haplotypes were inferred using PHASE, version 2.0.2 (Stephens & Donnelly, 2003), using the model for recombination. All cases and controls were included, and haplotypes were defined by the 12 loci in sequence, where there was incomplete genotype data, potential haplotype reconstructions, given the genotype data observed at the other loci, were inferred. A sample of 100 haplotype reconstructions, were determined for all individuals within the study group, such that uncertainty in inference of haplotypes could be accounted for in subsequent analyses (Stephens & Donnelly, 2003).

Lewontin's D' coefficient (Lewontin, 1964; Hedrick, 1987) was estimated for each pair of loci. D' coefficients were estimated for each of the sampled haplotype reconstructions, with the mean value taken as an overall estimate. Calculations of D' included all cases and controls, and were made using the *haploxt* programme within graphical overview of linkage disequilibrium (GOLD)(Abecasis & Cookson, 1974). The GOLD software was used to produce a graphical representation of linkage disequilibrium for the region.

Estimates of association between disease and haplotype were made using logistic regression, defining the most commonly observed haplotype as the baseline. Haplotypes with a frequency of <1% in the combined sample of cases and controls were grouped as 'other haplotypes'. The OR were repeatedly estimated for each sampled haplotype reconstruction, with mean OR and associated 95% CI derived following the principle of inference using multiple imputation (Schafer, 1999; Little & Rubin, 2002). Haplotypes categorized by LT-α +252 and TNF-α −308, by all SNP, and by all SNP and microsatellite loci were considered. Models were compared using the Akaike Information Criteria (AIC)

(Akaike, 1974) using the mean AIC over the sampled reconstructions. Distributions of haplotypes in case and control groups, were compared using permutation testing, in which permutations within each haplotype reconstruction were made, and the total number of permutations attaining a chi-squared value above the realised value for each reconstruction, were summed to determine statistical significance.

Table I. Genotype and allelic association of SNP loci with myeloma

Locus	Genotype/allele	Cases (n = 181/362) [n (%)]	Controls (n = 233/466) [n (%)]	OR	95% CI	P*
LT-α	AA	85 (50)	91 (42)	1·00		
	AG	75 (44)	100 (46)	0·80	0·53–1·22	
	GG	10 (6)	25 (12)	0·43	0·19–0·94	0·09
	SNA	11	17			
	G vs. (A)	95 (28)	150 (35)	0·73	0·54–0·99	0·05
	HWE: P†	(0·21)	(0·75)			
−1031	TT	98 (55)	147 (64)	1·00		
	TC	68 (38)	75 (32)	1·36	0·90–2·06	
	CC	12 (7)	9 (4)	2·00	0·81–4·93	0·15
	SNA	3	2			
	C vs. (T)	92 (26)	93 (20)	1·38	1·00–1·92	0·06
	HWE: P†	(0·96)	(0·88)			
−863	CC	122 (68)	165 (71)	1·00		
	CA	52 (29)	61 (26)	1·15	0·74–1·79	
	AA	6 (3)	5 (2)	1·62	0·40–6·88	0·63
	SNA	1	2			
	A vs. (C)	64 (18)	71 (15)	1·19	0·82–1·72	0·37
	HWE: P†	(0·87)	(0·82)			
−857	CC	160 (89)	200 (87)	1·00		
	CT	18 (10)	30 (13)	0·75	0·40–1·39	
	TT	2 (1)	1 (0)	2·50	0·13–148·20	0·49
	SNA	1	2			
	T vs. (C)	22 (6)	32 (7)	0·87	0·50–1·53	0·62
	HWE: P†	(0·08)	(0·91)			
−308	GG	141 (78)	158 (68)	1·00		
	GA	36 (20)	64 (27)	0·63	0·40–1·01	
	AA	3 (2)	11 (5)	0·31	0·06–1·19	0·04
	SNA	1	0			
	A vs. (G)	42 (12)	86 (18)	0·58	0·39–0·87	0·01
	HWE: P†	(0·69)	(0·18)			
−238	GG	161 (89)	217 (93)	1·00		
	GA	18 (10)	15 (6)	1·62	0·79–3·31	
	AA	1 (1)	1 (0)	1·35	0·02–106·28	0·49
	SNA	1	0			
	A vs. (G)	20 (6)	17 (4)	1·54	0·80–3·01	0·20
	HWE: P†	(0·53)	(0·20)			
+489	GG	163 (91)	201 (87)	1·00		
	GA	15 (8)	29 (13)	0·64	0·33–1·23	
	AA	2 (1)	1 (0)	2·47	0·13–146·20	0·33
	SNA	1	2			
	A vs. (G)	19 (5)	31 (7)	0·77	0·43–1·40	0·42
	HWE: P†	(0·03)	(0·97)			

Odds ratio (OR) and 95% confidence interval (95% CI) are determined by exact methods where five or less of the genotype, in either cases or controls.

SNA, sample not amplified.

*Significance probability determined by permutation test, comparing genotype or allele frequency in cases and controls.

†Significance probability determined by the Hardy–Weinberg equilibrium (HWE) test.

Results

Allele and genotype frequencies

No evidence of deviation from Hardy–Weinberg equilibrium, at $P = 0{\cdot}01$ level, was observed for any SNP genotype in either the case or control population (Table I). Allelic and genotype frequencies for cases and controls are presented in Table I. A decreased frequency of the TNF-α −308A allele was observed in cases (12%) compared with the controls (18%), corresponding to a decreased risk of myeloma, OR 0·58, 95% CI 0·39–0·87. The AG and AA genotypes for TNF-α −308 were associated with decreased risks, OR 0·63, 95% CI 0·40–1·01 and OR 0·31, 95% CI 0·06–1·19 respectively. The frequency of LT-α +252G allele was slightly decreased in cases (28%) compared with the controls (35%), corresponding to a decreased risk of myeloma, OR 0·73, 95% CI 0·54–0·99. The AG and GG genotypes for LT-α +252 were associated with decreased risks, OR 0·80, 95% CI 0·53–1·22 and OR 0·43, 95% CI 0·19–0·94 respectively. No significant differences in the allelic or genotype distributions, between cases and controls, were observed for the other SNP considered. The distribution of alleles, described by numbers of repeats, for each of the TNF microsatellite loci are presented in Table II. No significant difference in the allelic distribution, between cases and controls, were observed for any of the microsatellite loci.

Linkage disequilibrium and haplotype structure

Strong linkage disequilibrium in the region was determined (see Table III and Fig 1), particularly between the SNP in the TNF gene and its promoter region, extending to the SNP within the LT-α gene. Linkage disequilibrium was less strong for pair-wise comparisons of microsatellite loci. Limited haplotype diversity was described where, for the control population, six frequent (>1%) haplotypes described 99% of the observed haplotypic variation over the seven SNP loci; and 11 frequent haplotypes described 82% of the observed haplotypic variation over the seven SNP and five microsatellite loci. Similar coverage was observed for the case population, whereby 97% and 80% respectively, of the variation was described by haplotypes corresponding to those inferred frequently in the control population.

Haplotypic association with myeloma

Comparison of haplotype association models using the AIC demonstrated that the most parsimonious model was a simple allelic relative risk for the TNF-α −308 variant allele. The minimum AIC was observed for this model, and thus increasing the complexity of the haplotype analysis did not significantly improve the fit of the models, given the number of parameters required. In all association analyses, a decreased risk was observed for all frequent (>1%) haplotypes, of which the TNF-α −308 variant allele A was constituent. Of particular interest was the model including haplotypes defined by LT-α +252 and TNF-α −308, in which the GA haplotype showed association with myeloma, OR 0·57, 95% CI 0·38–0·86, in contrast to the GG haplotype, OR 0·92, 95% CI 0·63–1·35, demonstrating that the univariate association observable between LT-α +252 and risk of myeloma, OR 0·73, 95% CI 0·54–0·99 for the G allele, (Table I), may be attributable to the strong linkage this SNP has with TNF-α −308.

Table II. Allelic association of microsatellite loci with myeloma.

Locus	Allele repeats	Cases ($n = 362$), n (%)	Controls ($n = 466$), n (%)	P^*
TNF*b*	1	56 (17)	63 (14)	
	3	36 (11)	67 (15)	
	4	151 (45)	176 (40)	
	5	92 (27)	129 (29)	
	7	3 (1)	3 (1)	0·29
	SNA	24	28	
TNF*a*	1	1 (0)	5 (1)	
	2	105 (30)	133 (29)	
	3	2 (1)	3 (1)	
	4	24 (7)	43 (9)	
	5	11 (3)	22 (5)	
	6	48 (14)	63 (14)	
	7	33 (9)	31 (7)	
	8	3 (1)	3 (1)	
	9	7 (2)	7 (2)	
	10	32 (9)	36 (8)	
	11	78 (22)	96 (21)	
	12	1 (0)	1 (0)	
	13	7 (2)	10 (2)	
	14	0 (0)	1 (0)	0·86
	SNA	10	12	
TNF*c*	1	255 (70)	342 (74)	
	2	107 (30)	122 (26)	0·30
	SNA	0	2	
TNF*e*	1	59 (16)	66 (15)	
	2	5 (1)	15 (3)	
	3	294 (82)	365 (82)	0·19
	SNA	4	20	
TNF*d*	1	31 (9)	55 (12)	
	2	9 (3)	7 (2)	
	3	184 (53)	251 (56)	
	4	105 (31)	114 (25)	
	5	14 (4)	20 (4)	
	6	1 (0)	3 (1)	0·35
	SNA	18	16	

*Significance probability determined by permutation test, comparing allele frequency in cases and controls.
SNA, sample not amplified.

Discussion

This study described the haplotypic structure across a region of chromosome 6 focussing on the 'TNF gene', Table IV. Strong linkage disequilibrium observed across the region was

Table III. Linkage disequilibrium between loci*.

D'*	TNF*b*	TNF*a*	LT-α	TNF*c*	−1031	−863	−857	−308	−238	489	TNF*e*
TNF*d*	0·60	0·72	0·34	0·82	0·92	0·93	0·82	0·65	0·96	0·98	0·93
TNF*e*	0·80	0·78	0·98	0·94	0·81	0·89	0·69	1·00	1·00	1·00	
489	1·00	0·98	1·00	0·93	1·00	1·00	1·00	1·00	1·00		
−238	1·00	0·84	1·00	1·00	1·00	1·00	1·00	1·00			
−308	0·69	0·77	1·00	0·94	0·94	1·00	1·00				
−857	1·00	0·98	1·00	0·93	1·00	1·00					
−863	0·87	0·77	0·98	0·95	0·99						
−1031	0·67	0·76	0·89	0·91							
TNF*c*	0·58	0·66	0·96								
LTα	0·81	0·65									
TNF*a*	0·81										

*Mean of pair-wise D' coefficients calculated for each of 100 haplotype reconstructions.

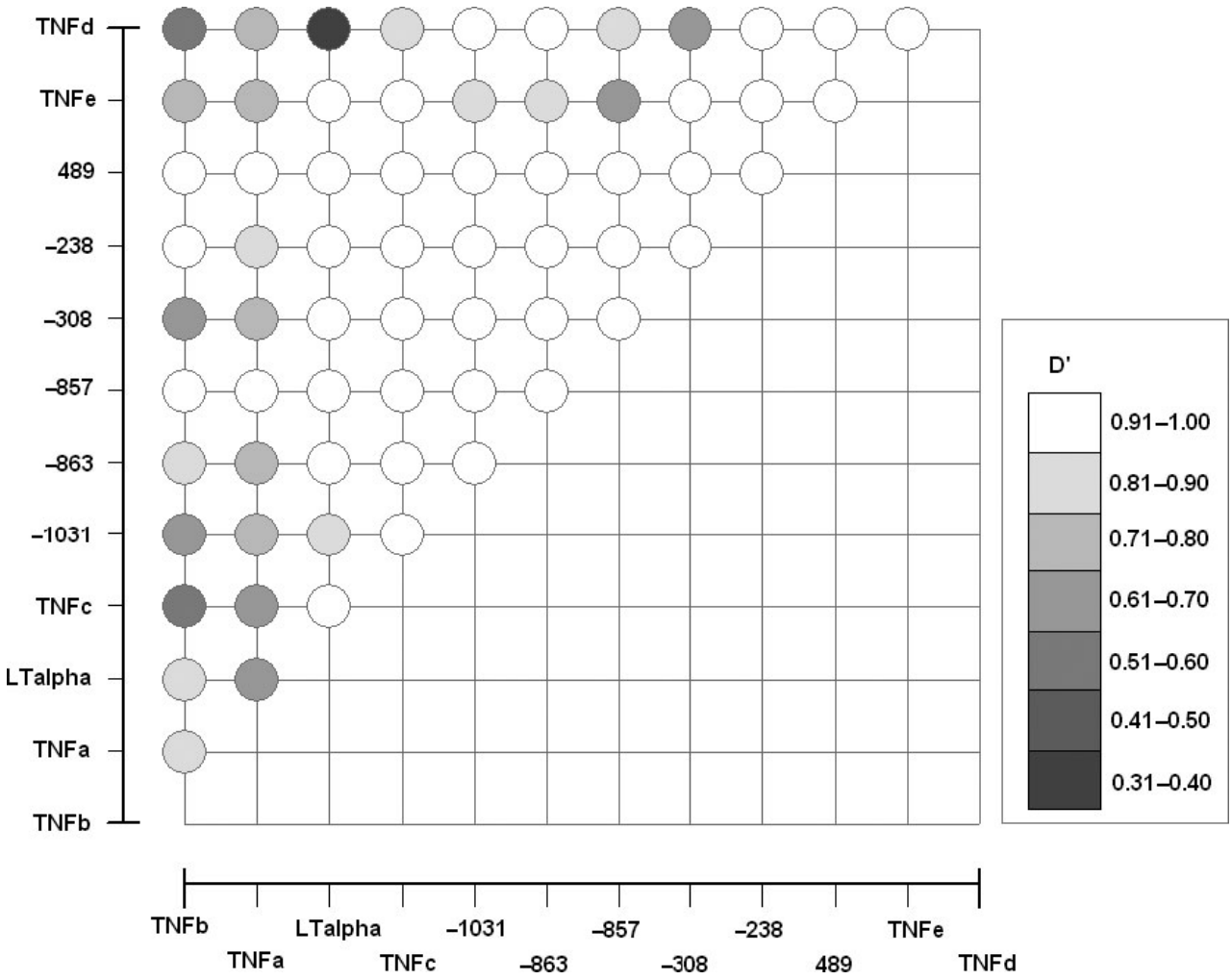

Fig 1. Graphical representation of linkage disequilibrium between loci (mean of pair-wise D' coefficients calculated for each of 100 haplotype reconstructions).

associated with a limited diversity of haplotypes, with the majority of haplotypic variation in the case and control populations being described by 12 frequent haplotypes. These haplotypes were used to study myeloma disease association, investigating whether the grouping of SNP within the extended haplotype provided any greater determination of disease risk than the consideration of individual SNP variation. The main finding of the study was an apparent reduced risk of myeloma

Table IV. Haplotypic associations with myeloma.

Haplotype	Cases (n = 362) [n(%)]*	Controls (n = 466) [n (%)]*	OR†	95% CI†	P‡
−308					
G	320·0 (88)	380·0 (82)	1·00		
A	42·0 (12)	86·0 (18)	0·58	0·39–0·87	0·01
AIC§: 1131·3					
LT-α, −308					
A**G**	259·8 (72)	303·5 (65)	1·00		
G**A**	42·0 (12)	86·0 (18)	0·57	0·38–0·86	
G**G**	60·2 (17)	76·5 (16)	0·92	0·62–1·35	0·02
AIC§: 1133·1					
LT-α, −1031, −863, −857, −308, −238, 489					
ATCC**G**GG	147·9 (41)	180·1 (39)	1·00		
ACAC**G**GG	63·9 (18)	72·0 (16)	1·08	0·72–1·63	
ATCT**G**GA	19·0 (5)	31·0 (7)	0·75	0·40–1·39	
ACCC**G**AG	21·0 (6)	17·0 (4)	1·50	0·76–2·99	
GTCC**A**GG	41·3 (11)	85·0 (18)	0·59	0·38–0·92	
GTCC**G**GG	56·7 (16)	74·8 (16)	0·92	0·61–1·41	
Other¶	12·1 (3)	6·1 (1)	2·42	0·88–6·71	0·03
AIC§: 1134·9					
TNF***b***, TNF***a***, LT-α, TNFc, −1031, −863, −857, −308, −238, 489, TNF***e***, TNF***d***					
4 11 A 1 TCC**G**GG 3 3	77·1 (21)	95·6 (21)	1·00		
4 7 A 1 TCC**G**GG 3 3	29·7 (8)	25·5 (5)	1·44	0·77–2·72	
5 5 A 2 TCC**G**GG 2 4	5·0 (1)	11·9 (3)	0·52	0·17–1·59	
4 13 A 1 TCC**G**GG 3 5	6·0 (2)	8·0 (2)	0·94	0·30–2·91	
4 9 A 2 TCC**G**GG 3 2	6·0 (2)	3·0 (1)	2·48	0·59–10·47	
1 2 A 2 CAC**G**GG 1 4	47·9 (13)	58·9 (13)	1·01	0·61–1·66	
4 10 A 1 TCT**G**GA 3 3	18·9 (5)	29·2 (6)	0·80	0·41–1·56	
5 2 A 2 CCC**G**AG 3 4	20·0 (6)	11·9 (3)	2·09	0·94–4·62	
3 2 G 1 TCC**A**GG 3 1	28·1 (8)	52·4 (11)	0·66	0·38–1·16	
5 4 G 1 TCC**A**GG 3 3	8·1 (2)	24·5 (5)	0·41	0·17–0·99	
5 6 G 1 TCC**G**GG 3 3	42·0 (12)	57·9 (12)	0·90	0·54–1·50	
5 4 G 1 TCC**G**GG 3 4	5·4 (1)	6·1 (1)	1·08	0·30–3·90	
Other**	68·1 (19)	81·2 (17)	1·04	0·66–1·64	0·10
AIC§: 1141·8					

*Mean frequency over 100 haplotype reconstructions.

†Combined odds ratio and 95% confidence interval for each of 100 haplotype reconstructions.

‡Significance probability determined by permutation test, comparing distribution of haplotype in cases and controls in each of 100 haplotype reconstructions.

§Mean Akaike Information Criteria over 100 haplotype reconstructions.

¶10 haplotypes each of inferred frequency <1% in the combined case and control sample.

**211 haplotypes each of inferred frequency <1% in the combined case and control sample.

in persons with the A variant at position −308 of the TNF promoter region. Detailed characterization of the haplotype using seven SNP and five microsatellites did not provide greater determination of disease risk beyond that described by the TNF-α −308 SNP.

A single base substitution at position −308 of the TNF-α gene results in two allelic forms in which the presence of guanine defines the common variant TNF1; the presence of adenine defines the less common variant TNF2. Functional studies of the TNF2 allele have shown both higher constitutional and inducible expression, which may be explained by its situation in an AP-2 binding site within the promoter. In a prior less extensive and simpler analysis, we carried out an association study looking at the risk of developing myeloma in association with the −308 TNF and the LT-α +252 variants (Davies *et al*, 2000). 'The 181 cases included in this study were derived from the MRC VII study and includes our previously reported cases (Davies *et al*, 2000)'. In that study the approach taken was different to that currently used in that we followed the approach suggested by Warzocha *et al* (1998). In that approach, high producers were defined as individuals carrying more than one high producer haplotype at either of these loci. The study suggested that the higher producer haplotypes were associated with myeloma risk. There was an increased frequency of double heterozygotes (TNF1/2 LT10·5/5·5) in the cases compared with the controls,

with an associated twofold increased risk of developing myeloma. However, the evidence supporting this approach to analysis is limited and consequently we carried out the current study with the aim of examining the association further using a more genomic-based approach. In the current approach, rather than arbitrarily assigning individuals as high or low producers, we analysed association based on genetic variants distributed across the locus either singly or as haplotypes. By taking this approach we hoped to define at-risk haplotypes as well as to understand the potential biological function of any associated variants. For the TNF −308 variant, the distribution in the control group was identical between the studies, however, a strongly increased frequency of heterozygotes was observed for the case series in the previous result. In that series there was a strong deviation from Hardy–Weinberg equilibrium for the case groups, particularly for LT-α, with strong excess heterozygosity over that, which would be expected. One explanation for this may be that the genotyping for the original study was carried out using restriction fragment length polymorphism analysis rather than the more robust IHG analysis used on this occasion. This difference may well have been exacerbated by differences in the quality of DNA between the case and control groups.

Our starting hypothesis was that high producer haplotypes at the TNF locus would be associated with an increased risk of developing myeloma. This was based on a biological rationale as well as the results of the first analysis. The results of the current study do not support this hypothesis, with an apparent excess of low producer alleles of TNF −308 being seen in the case group. Thus, on the basis of this study, it must be considered that this association is not proven and that additional studies should address this question further. However, it would seem, that taking a haplotype-based approach does not improve the precision of the analysis over the use of the TNF −308 allele alone.

The current study also has important consequences for the interpretation of previous TNF locus association studies. Many of these studies have been small, often lacking power to define critical associations and may not have used optimum methodologies for detecting SNP variants. Future studies in this area need to incorporate robust allele discrimination and be large enough to provide sufficient power. Combination of smaller studies in meta-analysis may provide more definitive conclusions. Further, this study provides an example where detailed haplotype definition did not predict disease susceptibility beyond that predicted by a single potentially functional SNP locus. However, in such a region it is difficult to attribute functionality to any single SNP, as there is such strong association between the loci through linkage disequilibrium.

Acknowledgement

This work was supported by the Leukaemia Research Fund.

References

Abecasis, G.R. & Cookson, W.O. (1974) GOLD – graphical overview of linkage disequilibrium. *Bioinformatics*, **16,** 182–183.

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transaction and Automatic Control*, **AC-19,** 716–723.

Child, J.A., Morgan, G.J., Davies, F.E., Owen, R.G., Bell, S.E., Hawkins, K., Brown, J., Drayson, M.T. & Selby, P.J. (2003) High-dose chemotherapy with hematopoietic stem-cell rescue for multiple myeloma. *New English Journal of Medicine*, **348,** 1875–1883.

Davies, F.E., Rollinson, S.J., Rawstron, A.C., Roman, E., Richards, S., Drayson, M., Child, J.A. & Morgan, G.J. (2000) High-producer haplotypes of tumor necrosis factor alpha and lymphotoxin alpha are associated with an increased risk of myeloma and have an improved progression-free survival after treatment. *Journal of Clinical Oncology*, **18,** 2843–2851.

Degli-Esposti, M.A., Leelayuwat, C., Daly, L.N., Carcassi, C., Contu, L., Versluis, L.F., Tilanus, M,G. & Dawkins, R.L. (1995) Updated characterization of ancestral haplotypes using the Fourth Asia-Oceania Histocompatibility Workshop panel. *Human Immunology*, **44,** 12–18.

Good, P.I. (2000) *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*, 2nd edn. Springer, New York, NY.

Hedrick, P.W. (1987) Gametic disequilibrium measures: proceed with caution. *Genetics*, **117,** 331–341.

Kane, E.V., Roman, E., Cartwright, R., Parker, J. & Morgan, G. (1999) Tobacco and the risk of acute leukaemia in adults. *British Journal of Cancer*, **81,** 1228–1233.

Lewontin, R.C. (1964) The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics*, **49,** 49–67.

Little, R.J.A. & Rubin, D.B. (2002) *Statistical Analysis with Missing Data*, 2nd edn. Wiley, Chichester.

Morse, H.R., Olomolaiye, O.O., Wood, N.A.P., Keen, L.J. & Bidwell, J.L. (1999) Induced heteroduplex genotyping of TNF-alpha, IL-1 beta, IL-6 and IL-10 polymorphisms associated with transcriptional regulation. *Cytokine*, **11,** 789–795.

Nedospasov, S.A., Udalova, I.A., Kuprash, D.V. & Turetskaya, R.L. (1991) DNA sequence polymorphism at the human tumor necrosis factor (TNF) locus. Numerous TNF/lymphotoxin alleles tagged by two closely linked microsatellites in the upstream region of the lymphotoxin (TNF-beta) gene. *Journal Immunology*, **147,** 1053–1059.

Schafer, J.L. (1999) Multiple imputation: a primer. *Statistical Methods in Medical Research*, **8,** 3–15.

Stephens, M. & Donnelly, P. (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73,** 1162–1169.

Stephens, M., Smith, N.J. & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68,** 978–989.

Udalova, I.A., Nedospasov, S.A., Webb, G.C., Chaplin, D.D. & Turetskaya, R.L. (1993) Highly informative typing of the human TNF locus using six adjacent polymorphic markers. *Genomics*, **16,** 180–186.

Warzocha, K., Ribeiro, P., Bienvenu, J., Roy, P., Charlot, C., Rigal, D., Coiffier, B. & Salles, G. (1998) Geneticpolymorphisms in the tumor necrosis factor locus influence non-Hodgkin's lymphoma outcome. *Blood*, **91,** 3574–3581.

Appendix 1

TNF polymorphism	Primer sequence 5′–3′	Oligo length (bp)	G and C content (bp)	Annealing temperature of primers for PCR (°C)	No. of PCR cycles
LT-α +252	F: CCATCTGTCAGTCTCATTGTC	21	10	58	32
	R: TGTCTCCCTCTGCTCACCTTGG	22	13		
TNF −1031	F: TGGACTCACCAGGTGAGGCCG	21	14	63	32
	R: GAGCTCCTGGGAGATATGGCC	21	13		
TNF −863 and −857	F: GAGATGTGACCACAGCAATGG	21	11	59	32
	R: CAGGACCTCCAGGTATGGAAT	21	11		
TNF −308 and −238	F: GTCCCCAAAAGAAATGGAGGC	21	11	59	32
	R: GGAGTGTGAGGGGTATCCTT	20			
TNF +489	F: GATGGGATGGGTGAAAGATGT	21	10	57	32
	R: GGGAAGAGAGAGAGAGAAAG	20	10		

TNF microsatellite	Primer sequence 5′–3′	No. of bases	Calculated temperature (°C)	Temperature used in PCR (°C)
TNF*a*	F: GCCTCTAGATTTCATCCAGCCACA	24	62·7	56
	R: CCTCTCTCTCCCCTGCAACACACA	22	64	
TNF*b*	F: GTGTGTGTTGCAGGGGAGAGAG	22	64	64
	R: GCACTCCAGCCTAGGCCACAGA	22	65·8	
TNF*c*	F: ACGCACCCTTGGGAGGAAGA	20	61·4	60
	R: TCTGTCTTCCGCCGCGTG	18	60·5	
TNF*d*	F: CATAGTGGGACTCTGTCTCCAAG	24	62·7	62
	R: AGATCCTTCCCTGTGAGTTCTGCT	24	62·7	
TNF*e*	F: TGAGACAGAGGATAGGAGAGACAG	24	62·7	62
	R: GTGCCTGGTTCTGGAGCCTCTC	22	65·8	