Genes and Immunity (2004) 5, 76–79

www.nature.com/gene

BRIEF COMMUNICATION

# *Discrimination of suballeles present at the TNFd microsatellite locus using induced heteroduplex analysis*

CF Spink[1], LJ Keen[1], PG Middleton[2] and JL Bidwell[1]
[1]*Department of Pathology and Microbiology, University of Bristol, UK;* [2]*Department of Haematology, School of Clinical Laboratory Sciences, University of Newcastle, UK*

*Polymorphism at the TNF*d *locus has been implicated in a number of disease association studies. The TNF*d *locus consists of three regions of (GA)*n *repeats separated by an imperfect repeat of two guanine bases. TNF*d *alleles are genotyped by the number of repeats in the first (GA)*n *repeat region, and until now the second repeat region had been thought to be nonpolymorphic. We report the existence of suballeles present within the TNF*d *microsatellite locus, detected using induced heteroduplex generator (IHG) technology. These alleles cannot be detected using conventional typing strategies as they represent altered distribution of the (GA)*n *repeats or sequence variation within the repeat. The suballeles affect the frequencies of the conventional d3 and d4 alleles leading to significantly altered allele frequencies. Some studies have associated the* d3 *and d4 alleles with disease outcome. We re-analysed one such study cohort using IHG technology and demonstrated a high proportion of incorrectly assigned TNF*d3 *alleles.*
Genes and Immunity (2004) **5,** 76–79. doi:10.1038/sj.gene.6364039

**Keywords:** *IHG technology; TNFd suballeles; homoplasy*

## Introduction

The *LST-1* gene is located 15 kilobases (kb) upstream of the tumour necrosis factor alpha (TNFα) locus (*TNFSF2*) and encodes a small protein that modulates immune responses and cellular morphogenesis and is constitutively expressed in leucocytes and dendritic cells.[1] Its gene product has been shown to have an inhibitory effect on lymphocyte proliferation.[2] Polymorphism has been described in intron 3 of the *LST-1* gene, including two dinucleotide repeats termed TNF*d* and TNF*e*.[3] It has been reported that variability in TNFα production is associated with alleles at the TNF*d* locus. A strong correlation between inheritance of the TNF*d3* allele and high TNFα production by leucocytes *in vitro* was shown.[4] The experiment suggested strong linkage disequilibrium between an as yet undescribed functional polymorphism in the TNFα gene and the TNF*d3* allele, despite a distance of some 8 kb between the TNF*d* locus and the TNFα gene.

TNF region microsatellites have been typed in a number of disease association studies.[5–7] With the conventional size-based method of typing, there is an assumption that alleles of identical size are identical in sequence. Often this is not the case, due to the presence of variable base insertions/deletions either within or in close proximity to the microsatellite, termed homoplasy. Homoplasy has been reported within the TNF*a* microsatellite, which fundamentally affects the calculation of TNF*a* allelic size.[8,9] Thus a number of suballeles have now been described at the TNF*a* microsatellite.[10]

*Correspondence: Miss C Spink, Department of Pathology and Microbiology, University of Bristol, Homoeopathic Hospital Site, Cotham, Bristol BS6 6JU, UK. E-mail: C.F.Spink@bristol.ac.uk*
Received 07 August 2003; revised 02 October 2003; accepted 03 October 2003

The presence of homoplasy in the TNF*d* microsatellite would have strong implications in disease association studies that have analysed this locus. TNF region polymorphism has been strongly implicated in severe acute graft-versus-host disease (aGVHD) and other bone marrow transplant (BMT)-related complications. Middleton *et al*[11] demonstrated TNF*d3* homozygosity associated with grade III/IV aGVHD in HLA-matched sibling BMT. A second study also tested association of aGVHD with TNF*d* in a large cohort of sibling donor/recipient pairs.[12] Their genotype results correlated with acute and chronic GVHD and mortality, and in addition patients who were homozygous *d3* had higher transplant-related mortality rates.

We analysed the TNF*d* microsatellite locus using DNA conformational approaches in order to determine the presence or absence of further suballeles. This paper demonstrates that induced heteroduplex generator (IHG) technology can successfully identify and genotype the TNF*d* microsatellite, detecting sequence variations, leading to the identification of TNF*d1* and TNF*d4* suballeles, a clear advance over existing typing methods. The TNF*d1* suballele would otherwise have been typed as a TNF*d3* allele using conventional typing approaches. Using the Arlequin software package we have also demonstrated strong linkage disequilibrium between the suballeles and other polymorphic loci in the TNF region. In addition we have retyped the same DNA sibling transplant panel, previously typed as TNF*d3* homozygotes by Middleton *et al*,[11] for the TNF*d1* suballele using IHG analysis.

## Results and discussion

### Identification of TNF*d* locus suballeles

Figure 2a shows a conventional size-based typing approach for the TNF*d* microsatellite locus using

```
                                        <------ repeat region 1------>  < repeat region 2>
Allele d5   CATAGTGGACTCTGTCTCCAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGGAGAGAGAGAGAGAGAGA
Allele d4   CATAGTGGACTCTGTCTCCAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGA--GGGAGAGAGAGAGAGAGA
Allele d3   CATAGTGGACTCTGTCTCCAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGA----GGGAGAGAGAGAGAGAGA
Allele d1   CATAGTGGACTCTGTCTCCAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGA--------GGGAGAGAGAGAGAGAGA
IHG 2       CATAGTGGACTCTGTCTCCAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGA----------GGGAGAGAGAGAGAGAGA
IHG 1       CATAGTGGACTCTGTCTCCAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGA------------GGGAGAGAGAGAGAGAGA

                 <Repeat3>
Allele d5   GGGAGAGAAGTAAGAAAGGCTGGAGGTGGGAGCAGAACTCACAGGGAAGGATCT
Allele d4   GGGAGAGAAGTAAGAAAGGCTGGAGGTGGGAGCAGAACTCACAGGGAAGGATCT
Allele d3   GGGAGAGAAGTAAGAAAGGCTGGAGGTGGGAGCAGAACTCACAGGGAAGGATCT
Allele d1   GGGAGAGAAGTAAGAAAGGCTGGAGGTGGGAGCAGAACTCACAGGGAAGGATCT
IHG 2       GGGAGAGAAGTAAGAAAGGCTGGAGGTGGGAGCAGAACTCACAGGGAAGGATCT
IHG 1       GGGAGAGAAGTAAGAAAGGCTGGAGGTGGGAGCAGAACTCACAGGGAAGGATCT
```

**Figure 1** Nucleotide sequence alignment of TNF*d5, d4, d3* and *d1* alleles with IHG1 and IHG2. Forward and reverse primer annealing sites are shown underlined. Repeat regions 1, 2 and 3 are highlighted in bold typeface. The IHGs were synthesized as single oligonucleotides, using 0.2 μmol membrane columns on a PerSeptive 8900 Nucleic Acid Synthesis System. After synthesis, deprotection and precipitation, the oligonucleotides were amplified by PCR using the primers as shown. Prior to analysis, a stepwise dilution of each IHG was carried out. The optimal range was between $10^{-4}$ and $10^{-6}$ dilution.

nondenaturing polyacrylamide gel electrophoresis (PAGE). Natural heteroduplexes are formed during each reannealing stage of the PCR reaction when two microsatellite alleles from a heterozygous individual cross-hybridize, forming a heteroduplex. In using the conventional typing procedure it was noticeable that samples with the same genotype gave different natural heteroduplex patterns (Figure 2a, arrows). This could only occur where there is a further level of polymorphism within or adjacent to the repeat sequence, a phenomenon termed homoplasy. It was noted that the variation appeared to be restricted to TNF*d3- and d4*-positive genotypes and we therefore decided to attempt to identify unequivocally the potential variant allele using DNA conformational approaches. An IHG reagent, designated IHG1, was synthesized which mimicked an undescribed TNF*d* allele (Figure 1). Using this reagent in combination with PAGE we were able to identify all existing TNF*d* alleles and genotypes. Using IHG1 we identified a novel *d4*-related allele that was not always

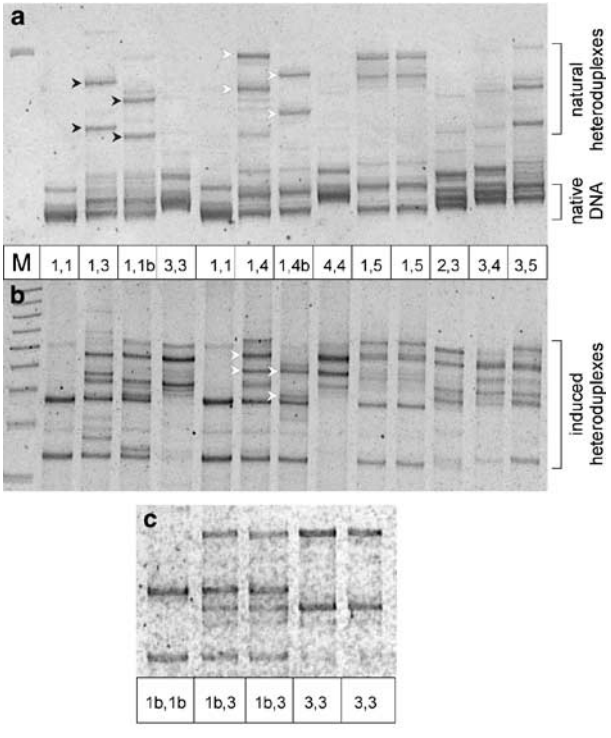


**Figure 2** (**a**) Conventional nondenaturing PAGE TNF*d* microsatellite typing. Black and white arrows indicate natural heteroduplexes that indicate sequence variation not identified by conventional analysis of PCR product length. PCR primers from a previous study[13] were used for amplification of the TNF*d* microsatellite and IHG1 and IHG2 (IHG and primer sequences shown in Figure 1). PCR amplifications were carried out in a total reaction volume of 20 μl, containing 100 ng of genomic DNA or IHG, 0.5 U *Taq* polymerase (Advanced Biotechnologies), 0.5 μM of each primer (forward and reverse), 1.5 mM $MgCl_2$, 200 μm of each dNTP, 67 mM Tris-HCL (pH 8.8), 16 mM $(NH_4)_2SO_4$ and 0.01% Tween. The PCR was optimized on an MJ PTC-100 thermal cycler (GRI, Braintree, UK). The PCR parameters for both genomic DNA and IHGs were as follows: initial denaturation for 5 min at 95°C, followed by 32 cycles of 95°C for 1 min, 62°C for 30 s, 72°C for 1 min with a final extension at 72°C for 5 min. (**b**) TNF*d* microsatellite typing using induced heteroduplex analysis with IHG2. White arrows indicate induced heteroduplexes, which clearly discriminate between alleles *4* and *4b*. The method of IHG analysis has been described previously:[14,15] cross-matching parameters were 95°C for 2 min, followed by an initial controlled cooling step to 65°C (ramping rate=1.0°C/s) and holding for 1 min. A second controlled cooling step was inserted to 45°C (ramping rate 0.03°C/s) and held for 1 min, with a final holding step at 4°C. Heteroduplexes were resolved on nondenaturing 15% Protogel® polyacrylamide gels (National Diagnostics, Hull, UK) using a 'triple wide' minigel system (30 cm × 8 cm; CBS Scientific Company, Del Mar, USA): gel constitution 37.5:1 (w/v acrylamide:bisacrylamide). Electrophoresis was carried out at 300 V for 90 min. Gels were post-stained in 1 × Tris borate EDTA (TBE) buffer containing 0.5 μg/ml ethidium bromide and examined using a 302 nm UV transilluminator. Images for analysis were acquired using an EDAS 120 system (Kodak). M, molecular weight marker (100 bp ladder). (**c**) TNF*d* microsatellite typing using induced heteroduplex analysis with IHG1. Induced heteroduplexes clearly discriminate between alleles *3* and *1b*. Genotypes are shown for each lane.

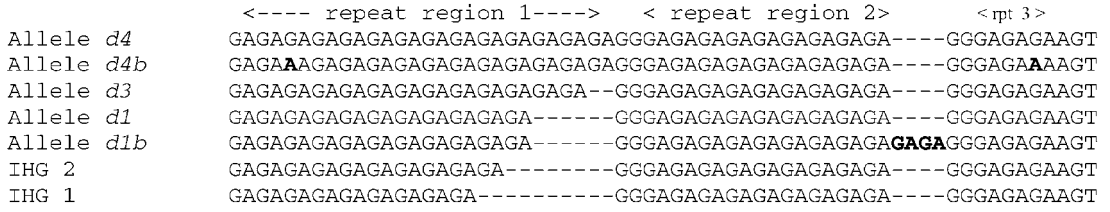


**Figure 3** Nucleotide sequence alignment of TNF*d4*, *d4b*, *d3*, *d1* and *d1b* with IHG 1 and IHG 2. Alleles *d1* and *d1b* differ by two GA repeats in repeat region 2. Alleles *d4* and *d4b* differ by two imperfect repeats (GA >AA) in repeat region 1 and repeat region 3.

**Table 1** Allele frequencies in 100 Caucasian adult volunteer bone marrow donors from southwest England, showing a comparison of alleles detected by nondenaturing PAGE and IHG analysis

| *Conventional nondenaturing PAGE* | | *IHG-based analysis* | |
|---|---|---|---|
| *TNFd allele* | *Allele frequency* | *TNFd allele* | *Allele frequency* |
| 1 | 0.08 | 1 | 0.08 |
| 1b | Not detectable | 1b | 0.155 |
| 2 | 0.015 | 2 | 0.015 |
| 3 | 0.545 | 3 | 0.39 |
| 4 | 0.295 | 4 | 0.255 |
| 4b | Not detectable | 4b | 0.04 |
| 5 | 0.06 | 5 | 0.06 |
| 7 | 0.005 | 7 | 0.005 |

identified using nondenaturing PAGE (Figure 2b, arrows). However, using IHG1, we were unable to discriminate clearly between the TNF*d3* allele and its potential variant. Therefore, a second IHG was constructed (designated IHG2) with one extra GA repeat than present in IHG1 (see Figure 1). Using this reagent it was possible to identify clearly the presence of a previously undescribed variant allele designated TNF*d1b* (Figure 2c) in individuals previously genotyped as *d3* homozygotes. IHG1 and IHG2 may therefore be used in combination to type accurately these two novel alleles, and all previously described alleles in the TNF*d* microsatellite locus.

Figure 3 shows the nucleotide sequences of the two novel TNF*d* alleles. We have classified the variant allele previously typed erroneously as *d3*, as a suballele of TNF*d1* and have designated it TNF*d1b*. This designation is based on the principle of TNF*d* allelic discrimination by sequence variation in the first of three GA repeat blocks. The TNF*d1b* allele has the same number of repeats as the TNF*d1* allele in the first repeat block but has two extra GA repeats in the second repeat block. The situation with the TNF*d4* suballele is more complex. Sequence variation involves the creation of two imperfect repeats via guanine to adenosine substitutions, one in the first GA repeat region and one in the third. We have designated this new suballele TNF*d4b*.

### Allele frequencies of the two novel TNF*d* alleles

Using these new IHG-based typing methods, we genotyped a random cohort of 100 individuals from a BMT donor registry using both size-based and conformational approaches. Conventional TNF*d* allele frequencies did not differ significantly from previous studies.[16] However, when IHG typing was incorporated, allele frequencies changed significantly (Table 1). The TNF*d1b* allele had an allele frequency of 0.155. This figure represents 28.4% of previously classified 'TNF*d3*' alleles. In our study, the TNF*d1b* allele is the third most frequent allele. The TNF*d4b* allele had an allele frequency of 0.04, 13.5% of previously classified 'TNF*d4*' alleles.

### Contribution of the two novel TNF*d* alleles to extended TNF haplotypes

In order to identify extended TNF region haplotypes that include the new TNF*d* alleles, we genotyped the same random cohort at a number of microsatellite and SNP loci. Maximum likelihood analysis (Arlequin 2.000[17]) was used to identify conserved haplotypes, and linkage disequilibrium analysis was used to analyse the association between loci. The TNF*d1b* allele was found to be significantly associated with a haplotype that includes the TNF*a6* allele. There was significant linkage disequilibrium ($P=4\times10^{-14}$) between the TNF*d1b* and TNF*a6* alleles. The TNF*d4b* allele was strongly associated with a haplotype including the TNF*a2*, TNF*d4b*, TNF$\alpha$ $-1031$C and $-238$A alleles (unpublished data). Linkage disequilibrium between the TNF*d4b* and TNF$\alpha$ $-238$A allele was highly significant ($P<0.000001$).

### Reanalysis of TNF*d3* homozygotes in an identical sibling bone marrow transplant cohort

A previous study on the influence of cytokine polymorphism on HLA-identical sibling stem cell transplantation[11] reported an association with homozygosity of the *d3* allele and increased risk of grade III/IV aGVHD. We revisited this study and genotyped the *d3*-homozygous samples using the TNF*d* IHG2 reagent, in order to identify any TNF*d1b* alleles. Of the 28 sib-transplant panel retyped for presence of the *d1* suballele, two individuals were homozygous and nine individuals were heterozygous for this suballele. In all, 17 individuals were negative for this subtype.

## Conclusions

In the current study, the TNF*d1b* allele has been shown to be the third most frequent allele in the control cohort. It comprises 28% of alleles that have been typed as TNF*d3* in previous disease association studies. As the TNF*d3* allele was by far the most frequent allele, this has obvious implications for those studies that have associated this allele with given conditions.[11,18–21] We have demonstrated this by reanalysing TNF*d3* homozygotes in an identical sibling BMT cohort.

The *d3* allele has been associated with a number of extended TNF haplotypes. If the TNF*a*/TNF*d* microsatellite association is analysed, the *d3* allele is associated with the *a6*, *a7*, *a10* and *a11* alleles. These are relatively common alleles in Western populations. We have shown

that the TNF*d1b* and not the TNF*d3* allele is strongly associated with the TNF*a6*-associated haplotype. This finding is important as a number of studies have associated this haplotype with disease association,[19,20] leading to the conclusion that the TNF*d1b* allele alone would be a strong marker of disease association in these studies. Further to this we have identified strong linkage disequilibrium between the TNF*d4b* allele and the TNF −238A allele. A number of studies have associated this allele with immune disease.[22]

Our IHG-based approach for the analysis of the TNF*d* microsatellite permits identification of homoplasy at the TNF*d* locus. The identification of two new suballeles (TNF*d1b* and TNF*d4b*) has implications for disease association and expression studies that have up to now been unable to identify them. We are currently applying IHG technology to all TNF region microsatellites in order to identify further suballeles.

## Acknowledgements

We would like to thank Dr. Nigel Wood for synthesis of the IHG reagents, and Doris Culpan for her help with the nucleotide sequencing.

## References

1 de Baey A, Fellerhoff B, Maier S, Martinozzi S, Weidle U, Weiss EH. Complex expression pattern of the TNF region gene LST1 through differential regulation, initiation, and alternative splicing. *Genomics* 1997; **45**: 591–600.
2 Rollinger-Holzinger I, Eibl B, Pauly M *et al*. LST1: a gene with extensive alternative splicing and immunomodulatory function. *J Immunol* 2000; **15**: 169–176.
3 Holzinger I, de Baey A, Messer G, Kick G, Zwierzina H, Weiss EH. Cloning and genomic characterisation of LST1: a new gene in the human TNF region. *Immunogenetics* 1995; **42**: 315–322.
4 Turner DM, Grant SCD, Lamb WR *et al*. A genetic marker of high TNF-α production in heart transplant recipients. *Transplantation* 1995; **60**: 1113–1117.
5 Hajeer AH, Lear JT, Ollier WE *et al*. Preliminary evidence of an association of tumour necrosis factor microsatellites with increased risk of multiple basal carcinomas. *Br J Dermatol* 2000; **142**: 441–445.
6 Kunstmann E, Epplen C, Elitok E *et al*. *Helicobacter pylori* infection and polymorphisms in the tumor necrosis factor region. *Electrophoresis* 1999; **20**: 1756–1761.
7 Mu H, Chen JJ, Jiang Y, King MC, Thomson G, Criswell LA. Tumor necrosis factor a microsatellite polymorphism is associated with rheumatoid arthritis susceptibility through an interaction with the HLA-DRB1 shared epitope. *Arthritis Rheum* 1999; **42**: 438–442.
8 Honchel R, McDonell S, Schaid DJ, Thibodeau SN. Tumor necrosis factor-α allelic frequency and chromosome 6 allelic imbalance in patients with colorectal cancer. *Cancer Res* 1996; **56**: 145–149.
9 Greenberg SJ, Fujihara K, Selkirk SM *et al*. Novel compound tetra-, dinucleotide microsatellite polymorphism in the tumor necrosis factor/lymphotoxin locus. *Clin Diagn Lab Immunol* 1997; **4**: 79–84.
10 Van der Silk AR, Shing DC, Eerligh P, Giphart MJ. Subtyping for TNFa microsatellite sequence variation. *Immunogenetics* 2000; **52**: 29–34.
11 Middleton PG, Taylor PR, Jackson G, Proctor SJ, Dickinson AM. Cytokine gene polymorphisms associating with severe acute graft-versus-host disease in HLA-identical sibling transplants. *Blood* 1998; **92**: 3943–3948.
12 Cavet J, Middleton PG, Segall M, Noreen H, Savies SM, Dickinson AM. Recipient tumor necrosis factor-alpha and interleukin-10 gene polymorphisms associate with early mortality and graft-versus-host disease severity in HLA-matched sibling bone marrow transplants. *Blood* 1999; **94**: 3941–3946.
13 Udalova IA, Nedospasov SA, Webb GC, Chaplin DD, Turetskaya RL. Highly informative typing of the human TNF locus using six adjacent polymorphic markers. *Genomics* 1993; **16**: 180–186.
14 Wood N, Bidwell J. Genetic screening and testing by induced heteroduplex formation. *Electrophoresis* 1996; **17**: 247–254.
15 Morse HR, Olomolaiye OO, Wood NA *et al*. Induced heteroduplex genotyping of TNF-alpha, IL-1 beta, IL-6 and IL-10 polymorphisms associated with transcriptional regulation. *Cytokine* 1999; **11**: 789–795.
16 McDonnell GM, Kirk CW, Middleton D *et al*. Genetic association studies of tumour necrosis factor a and b and tumour necrosis factor receptor 1 and 2 polymorphisms across the clinical spectrum of multiple sclerosis. *J Neurol* 1999; **246**: 1051–1058.
17 Schneider S, Roessli D, Excoffier L. Arlequin ver 2.000: software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
18 Niizeki H, Naruse T, Hecker KH *et al*. Polymorphisms in the tumour necrosis factor (TNF) genes are associated with susceptibility to effects of ultraviolet-B radiation on induction of contact hypersensitivity. *Tissue Antigens* 2001; **58**: 369–378.
19 Mulcahy B, Waldron-Lynch F, McDermott MF *et al*. Genetic variability in the tumor necrosis factor-lymphotoxin region influences susceptibility to rheumatoid arthritis. *Am J Hum Genet* 1996; **59**: 676–683.
20 Hohler T, Grossmann S, Stradmann-Bellinghausan B *et al*. Differential association of polymorphisms in the TNFalpha region with psoriatic arthritis but not psoriasis. *Ann Rheum Dis* 2002; **61**: 213–218.
21 Turner DM, Pravica V, Sinnott PJ, Hutchinson IV. Polymorphism in the promoter region of the gene encoding human allograft inflammatory factor-1. *Eur J Immunogenet* 2001; **28**: 449–450.
22 Bidwell JL, Keen LJ, Gallagher G *et al*. Cytokine gene polymorphism in human disease: on-line databases. *Genes Immun* 1999; **1**: 3–19.