NOVEL ALLELES OF THE TRANSFORMING GROWTH FACTOR β-1 PROMOTER AND EXON 1

E Arrieta-Bolaños[1,2,3], JA Madrigal[1,2], BE Shaw[1,4]

1 Anthony Nolan Research Institute, Royal Free Hospital, Pond Street, London, UK, 2 Cancer Institute, University College London, Royal Free Campus, London, UK, 3 Centro de Investigaciones en Hematología y Trastornos Afines (CIHATA), Universidad de Costa Rica, San José, Costa Rica, 4 Haemato-Oncology Research Unit, Division of Molecular Pathology, The Institute of Cancer Research, London, UK

Introduction

Transforming growth factor β-1, encoded by the *TGFB1* gene, is a cytokine that plays a central role in many physiologic and pathogenic processes, having pleiotropic effects on cell proliferation, differentiation, migration and survival. Regulatory activity for this gene has been demonstrated for approximately 3.0 kb between positions -2,665 and +423 from its translational start site. This region includes two promoter sites, two negative regulatory elements and two enhancers, in addition to part of the protein's signal peptide. An important amount of polymorphism has been reported for this gene, including 17 promoter and exon 1 alleles defined on the basis or 18 polymorphisms (**Table 1**). Polymorphism in *TGFB1* has been associated to differential levels of expression of this cytokine and to genetic risk in cancer and hematopoietic stem cell (HSCT) and organ transplantation. In this report, we extend this list by presenting novel alleles formed by new polymorphic positions and a new combination of the previously described polymorphisms. These novel alleles have been found during the typing of a cohort of unrelated haematopoietic stem cell transplantation patient-donor pairs, and of healthy volunteer donors.

Table 1 Known *TGFB1* regulatory region alleles. Alleles as defined by Shah R, et al. 2006 and 2009 are formed by combinations of genotypes found in 18 polymorphic positions within *TGFB1*'s exon 1 and upstream regulatory region.

Allele[b]	Position[a]																	
	-2410	-2391	-1985	-1638	-1347	-1287	-1169[c]	-1154	-827	-778	-768	-469	-387	-229	-14	+29	+74	+91
p001	*A*	-	*C*	*G*	*T*	*G*	*TT*	*C*	*G*	*G*	-	*C*	*C*	*C*	*G*	*C*	*G*	*T*
p002	-	-	-	*A*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p003	*G*	*AGG*	-	-	*C*	-	-	-	-	-	-	-	-	-	-	*T*	-	-
p004	*G*	*AGG*	*G*	-	*C*	-	-	-	-	-	-	-	-	-	-	*T*	-	-
p005	-	*AGG*	-	-	*C*	*A*	-	*T*	*C*	-	-	-	*T*	-	*A*	-	-	-
p006	*G*	*AGG*	-	*A*	*C*	-	-	-	-	-	-	-	-	-	-	*T*	-	-
p007	*G*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
p008	*G*	*AGG*	-	-	*C*	-	-	-	-	-	-	*A*	-	-	-	*T*	-	-
p009	-	-	-	-	-	-	-	-	-	*A*	-	-	-	-	-	-	-	-
p010	-	-	-	-	*C*	-	-	-	-	-	-	-	-	-	-	-	-	-
p011	-	-	-	-	*C*	-	-	-	-	-	-	-	-	-	-	*T*	-	-
p012	-	-	-	-	-	-	*del*	-	-	-	-	-	-	-	-	-	-	-
p013	*G*	*AGG*	-	-	*C*	-	*del*	-	-	-	-	-	-	-	-	*T*	-	-
p014	-	*AGG*	-	-	*C*	-	-	-	-	-	*C*	-	-	-	-	-	*C*	-
p015	*G*	*AGG*	-	-	-	-	-	-	-	-	-	-	-	*G*	-	*T*	-	-
p016	*G*	*AGG*	-	-	-	-	-	-	-	-	-	-	-	-	-	*T*	-	-
p017	*G*	-	-	-	*C*	-	-	-	-	-	-	-	-	-	-	*T*	-	*G*

a. Nucleotide position relative to the major translation start site (+1) as described in Shah R, *et al*. Human Genet (2006) 119: 61-74 and Shah R, *et al*. Tissue Antigens (2009) 74: 50-6. Only positions that are polymorphic are included.
b. Promoter allele defined by the sequence of the regulatory region and exon 1. Sequences are compared to allele p001 (GenBank accession no. AY871232).
c. Indicates deletion of nucleotide(s) (del).

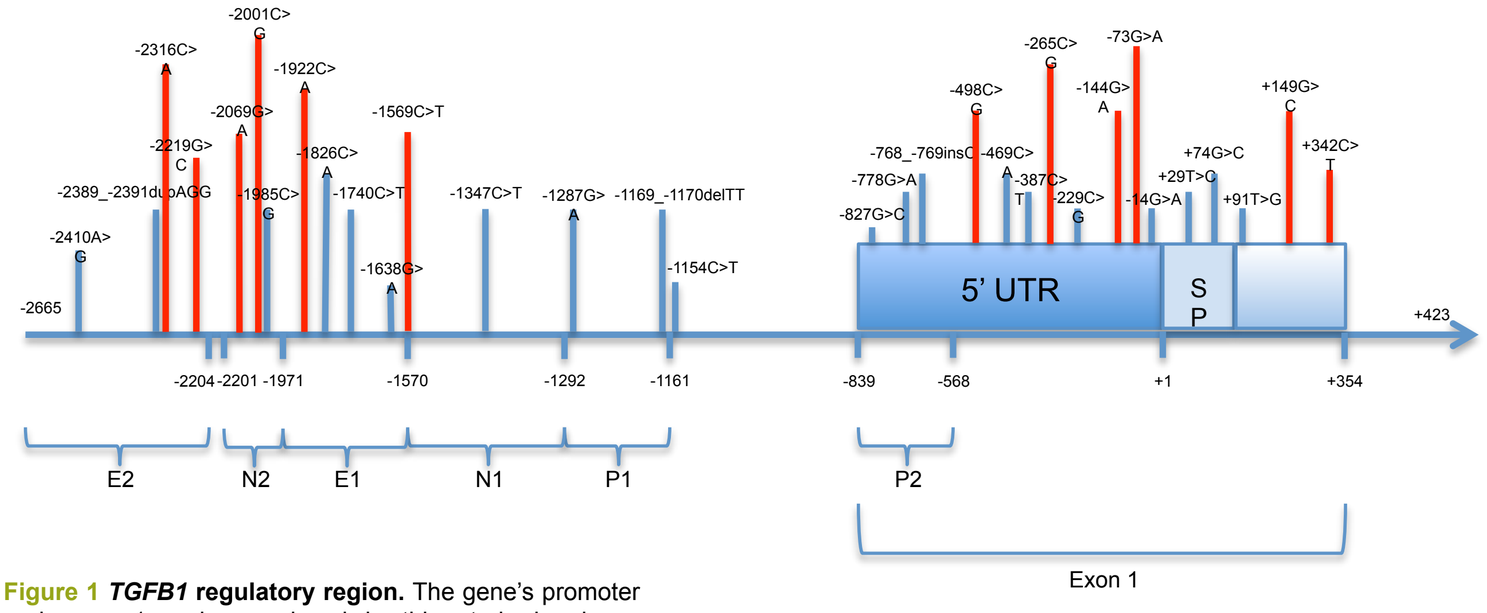

Figure 1 *TGFB1* regulatory region. The gene's promoter and exon 1 region analysed in this study is shown schematically and not to scale. Previously known variants that define known alleles are shown in blue, while novel variant positions discovered in this study are shown in red.

Materials and Methods

DNA from the Anthony Nolan Research Institute's sample repository has been used for the typing of a 3.0 kb region encompassing *TGFB1*'s promoter and exon 1. Peripheral blood from healthy volunteer donors was obtained by venipuncture. DNA was extracted by an in-house salting-out method. Molecular typing was carried out by an in-house amplification of 4 regions (~750 kb each) in *TGFB1*'s promoter and exon 1, followed by sequencing with M13 primers. Samples found to have previously unreported polymorphisms were confirmed by either (1) whole-region amplification followed by molecular cloning (Zero Blunt TOPO PCR Cloning Kit, Invitrogen) and/or (2) phase-specific whole-region amplification using primers specific for the -2391AGG duplication, both followed by sequencing in isolation.
In the case of the healthy volunteer donor, a previously unreported combination of the basic 18 polymorphic positions was suspected because the sample's genotype could not be inferred. A family study was designed and samples of the subject's parents and of one sibling were analysed.

Results

During the typing of 627 samples from the patient-donor cohort, a total of 12 new polymorphisms have been identified. Approximately 3% of the samples showed the presence of these new single nucleotide polymorphisms (SNP). Six of these SNP have been isolated and shown to be carried in alleles otherwise identical to the 4 frequent *TGFB1* promoter and exon 1 alleles found to occur in the patient-donor cohort (i.e. p001, p003, p006, p014). SNP at -498, -1569 and -2069 have been found in the context of p003, while those at -1743, -1878 and -1922 are carried by a p001-like allele. We are continuing the phase definition for the remainder of the new SNP.

The family study of the unsolvable genotype in a healthy donor revealed that the novel allele was carried by the subject's father, while the mother was homozygous for p001. The subject and the sibling had the same genotype. As shown in **Table 2** and confirmed by sequencing in isolation, the new allele is formed by a new combination of variants at positions -2410(A), -2391(AGG+), -1347(C) and +29(C), without polymorphism at positions -768(-) and +74(G).

Table 2 Results of family study of a new allele of *TGFB1*'s regulatory region.

	-2410	-2391	-1638	-1347	-768	+29	+74	
Subject	AA	AGG/-	GG	C/T	-/-	CC	GG	**p001/?**
Mother	AA	-/-	GG	TT	-/-	CC	GG	**p001/p001**
Father	G/A	AGG/AGG	GG	CC	-/-	T/C	GG	**p003/?**

Conclusions and future work

Our results show a considerable extent of unreported variation within *TGFB1*'s promoter and exon 1. Most of the new SNP found lie in either known regulatory regions such as enhancers, promoters or negative regulatory elements, or in the coding translated region of exon 1 causing amino acid changes in at least one case. All these variation is bound to have an effect in the regulation of this important gene. Knowledge of the extent and consequences of this variation are fundamental for understanding the immunogenetics of Transforming growth factor β-1 and its possible applications in Medicine.

We are currently continuing our analysis of the effect of polymorphisms within the regulatory region of *TGFB1* on the outcome of unrelated donor HSCT and on the function of regulatory T cells. These data will hopefully help to identify the risk factors associated with *TGFB1* expression in patients and donors and to generate clinically useful algorithms of genetic risk assessment in HSCT.

Acknowledgements

This work was supported by grants from the University of Costa Rica and the Costa Rican National Council for Scientific and Technologic Research (CONICIT).