# Polymorphism in TGFB1 is associated with worse non-relapse mortality and overall survival after stem cell transplantation with unrelated donors

Esteban Arrieta-Bolaños,[1,2,3] Neema P. Mayor,[1,2] Steven G.E. Marsh,[1,2] J. Alejandro Madrigal,[1,2] Jane F. Apperley,[4] Keiren Kirkland,[5] Stephen Mackinnon,[6] David I. Marks,[7] Grant McQuaker,[8] Julia Perry,[5] Michael N. Potter,[9] Nigel H. Russell,[10] Kirsty Thomson,[11] and Bronwen E. Shaw[1,2]

[1]Anthony Nolan Research Institute, London, UK; [2]Cancer Institute, University College London, UK; [3]Centro de Investigaciones en Hematología y Trastornos Afines (CIHATA), Universidad de Costa Rica, San José, Costa Rica; [4]Haematology Department, Hammersmith Hospital, London, UK; [5]BSBMT Data Registry, Guy's Hospital, London, UK; [6]Department of Haematology, University College London, Royal Free Campus, UK; [7]Adult BMT Unit, University Hospitals Bristol NHS Trust, Bristol, UK; [8]Bone Marrow Transplant Unit, Beatson West of Scotland Cancer Centre, Glasgow, UK; [9]Section of Haemato-oncology, Royal Marsden Hospital, London, UK; [10]Centre for Clinical Haematology, Nottingham University Hospital, and Academic Haematology, Nottingham University Hospitals, UK; and [11]Department of Haematology, University College Hospital, London, UK

**Haematologica** 2016
Volume 101(3):382-390

**Correspondence:**
a.madrigal@ucl.ac.uk

*Received:* August 10, 2015.
*Accepted:* November 20, 2015.
*Pre-published:* November 26, 2015.

doi:10.3324/haematol.2015.134999

*Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/101/3/382*



## ABSTRACT

Transforming growth factor β-1, encoded by the *TGFB1* gene, is a cytokine that plays a central role in many physiological and pathogenic processes. We have sequenced *TGFB1* regulatory region and assigned allelic genotypes in a large cohort of hematopoietic stem cell transplantation patients and donors. In this study, we analyzed 522 unrelated donor-patient pairs and examined the combined effect of all the common polymorphisms in this genomic region. In univariate analysis, we found that patients carrying a specific allele, 'p001', showed significantly reduced overall survival (5-year overall survival 30.7% for p001/p001 patients *vs.* 41.6% others; *P*=0.032) and increased non-relapse mortality (1-year non-relapse mortality: 39.0% *vs.* 25.4%; *P*=0.039) after transplantation. In multivariate analysis, the presence of a p001/p001 genotype in patients was confirmed as an independent factor for reduced overall survival [hazard ratio=1.53 (1.04-2.24); *P*=0.031], and increased non-relapse mortality [hazard ratio=1.73 (1.06-2.83); *P*=0.030]. In functional experiments we found a trend towards a higher percentage of surface transforming growth factor β-1-positive regulatory T cells after activation when the cells had a p001 allele (*P*=0.07). Higher or lower production of transforming growth factor β-1 in the inflammatory context of hematopoietic stem cell transplantation may influence the development of complications in these patients. Findings indicate that *TGFB1* genotype could potentially be of use as a prognostic factor in hematopoietic stem cell transplantation risk assessment algorithms.

## Introduction

Hematopoietic stem cell transplantation (HSCT) is a medical procedure used to treat malignant and non-malignant diseases of the blood, as well as solid tumors. The outcome of HSCT is influenced both by clinical and genetic factors. Compatibility between the recipient and the donor in terms of HLA is a well-known limiting factor for the success of allogeneic HSCT.[1] In addition, genes other than those of the HLA system, in particular those that are highly polymorphic, have been proposed as potential factors affecting the success of this therapy.[2]

One of the genes that are likely to play an important role in the outcome of allogeneic HSCT is *TGFB1*, which encodes transforming growth factor-β1 (TGF-β1). TGF-β1 is a cytokine that plays a central role in many physiological and pathogenic processes, having pleiotropic effects on cell proliferation, differentiation, migration and survival, as well as being a fundamental component of the immune system. TGF-β1 is likely to be relevant for both therapeutic and pathogenic immune processes associated with the different stages of HSCT.[3] Genetic variation resulting in differences in its production and/or function could play a role in the way that this cytokine modifies these immune processes.

Regulatory activity for this gene, located at chromosome 19q13.1-q13.3, has been mapped to approximately 3.0 kilobases (kb) from positions -2665 to +423 in its exon 1 (+1 being the translation start site). This region includes two promoter sites, two negative regulatory elements and two enhancers lying upstream of the first promoter.[4]

Several polymorphisms in *TGFB1* regulatory region have been identified, and these are known to cause alterations in cytokine secretion in several settings.[4] Previous work allowed for the definition of 17 *TGFB1* regulatory region and exon 1 alleles, which are formed by the combination of 18 SNPs and other kinds of variation (*Online Supplementary Table S1*).[4,5] We have recently expanded this inventory of *TGFB1* alleles with the discovery of other less common variant combinations.[6]

The role of polymorphism in *TGFB1* in the outcome of HSCT has been examined in some studies.[7] However, the results have not been consistent. In this study, we aimed at comprehensively analyzing the role of genetic variation in *TGFB1* regulatory region and exon 1 in a large cohort of UD-HSCT recipients and donors. In addition, since regulatory T cells (Treg) are major producers of TGF-β1 and have the unique ability of expressing its latent form on their surface upon stimulation,[8] as well as being likely effectors or targets during the immunological events taking place prior, during and after HSCT, we have performed functional assays to further understand the effect of this variation on the way that TGF-β1 is expressed by human regulatory Treg.

## Methods

### Patients, donors, and clinical data

Hematopoietic stem cell transplantation patient and donor samples are part of the Anthony Nolan Research Institute's stem cell transplantation sample repository (*www.myresearchproject.org.uk,* application number MREC 01/8/31).

Healthy volunteer donors were used to obtain mononuclear cells for functional experiments. Patients' clinical data were collected by the Anthony Nolan Research Institute in collaboration with the British Society for Blood and Marrow Transplantation.

All samples were collected according to the Anthony Nolan Research Institute's review board-approved guidelines and written informed consent was obtained from all participants.

### Sequencing of the regulatory region of TGFB1

The 3.0 kb upstream regulatory region of TGFB1 was analyzed for polymorphism by Sanger sequencing, as explained elsewhere.[6] Briefly, based on the studies by Shah *et al.*,[4,5] the region extending from -2,664 to +423 according to this gene's translation start site was sequenced and the sequenced fragments were then analyzed,

**Table 1. Patients', donor and transplant characteristics in the hematopoietic stem cell transplantation cohort (n=504*).**

| | N | % |
|---|---|---|
| Patient age (years) | | |
| Median 27.9 (range 0.4-63.8) | | |
| 0-20 | 176 | 34.9 |
| 20-40 | 201 | 39.9 |
| 40-60 | 126 | 25.0 |
| >60 | 1 | 0.2 |
| Donor age (years) | | |
| Median 35.0 (range 19.2-60.4) | | |
| 0-20 | 2 | 0.4 |
| 20-40 | 369 | 73.2 |
| 40-60 | 132 | 26.2 |
| >60 | 1 | 0.2 |
| Sex (male) | | |
| Patients | 322 | 63.9 |
| Donors | 382 | 75.8 |
| Sex matching | | |
| Patient-donor | | |
| Male-male | 256 | 50.8 |
| Male-female | 66 | 13.1 |
| Female-female | 56 | 11.1 |
| Female-male | 126 | 25.0 |
| HLA-matching | | |
| 10/10 matched | 358 | 71.0 |
| 1 mismatch | 100 | 19.8 |
| >1 mismatch | 46 | 9.1 |
| Disease | | |
| AML | 136 | 27.0 |
| ALL | 155 | 30.8 |
| CML | 111 | 22.0 |
| MDS | 45 | 8.9 |
| Other[1] | 57 | 11.3 |
| Disease status | | |
| Complete remission/chronic phase | 417 | 82.7 |
| Other | 77 | 15.3 |
| Unknown | 10 | 2.0 |
| CMV status | | |
| Patient(+)-donor(+) | 64 | 12.7 |
| Patient(+)-donor(-) | 88 | 17.4 |
| Patient(-)-donor(+) | 51 | 10.1 |
| Patient(-)-donor(-) | 282 | 56.0 |
| Unknown | 19 | 3.8 |
| TBI | | |
| Yes | 434 | 86.1 |
| No | 55 | 10.9 |
| Unknown | 15 | 3.0 |
| SC source | | |
| BM | 340 | 67.5 |
| PB | 159 | 31.5 |
| Both | 2 | 0.4 |
| Unknown | 3 | 0.6 |
| T-cell depletion | | |
| Yes | 427 | 84.7 |
| No | 22 | 4.4 |
| Unknown | 55 | 10.9 |
| GvHD prophylaxis | | |
| None | 4 | 0.8 |
| Cyclosporin | 131 | 26.0 |
| Cyclosporin+MTX | 329 | 65.3 |
| Other | 28 | 5.6 |
| Unknown | 12 | 2.4 |
| Transplant year | | |
| 1996-2001 | 253 | 50.2 |
| 2002-2009 | 251 | 49.8 |

*[1]Includes secondary acute leukemia, non-Hodgkin lymphoma, primary immune deficiency, bone marrow failure, multiple myeloma, metabolic disease, myeloproliferative neoplasia, biphenotypic acute leukemia, Hodgkin disease, undifferentiated acute leukemia. *18 pairs out of 522 eligible lacked clinical data. AML: acute myeloid leukemia; ALL: acute lymphoid leukemia; CML: chronic myeloid leukemia; BM: bone marrow; CMV: cytomegalovirus; GvHD: graft-* versus-*host disease; HLA: human leukocyte antigen; MDS: myelodysplastic syndrome; MTX: methotrexate; PB: peripheral blood; SC: stem cell; TBI: total body irradiation.*

**Table 2.** ***TGFB1*** **regulatory region and exon 1 allele frequencies found in the UD-HSCT patient-donor cohort.**

| Allele | Patients | | | | Donors | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Frequency | Copies | Carriers | Carrier frequency | Frequency | Copies | Carriers | Carrier frequency | Z | *P* |
| p001 | 0.3004 | 307 | 264 | 0.5166 | 0.2865 | 294 | 256 | 0.4990 | 0.7 | 0.49 |
| p003 | 0.5411 | 553 | 415 | 0.8121 | 0.5331 | 547 | 405 | 0.7895 | 0.4 | 0.72 |
| p006 | 0.0753 | 77 | 73 | 0.1429 | 0.0867 | 89 | 88 | 0.1715 | 0.9 | 0.34 |
| p009 | 0.0000 | 0 | 0 | 0.0000 | 0.0010 | 1 | 1 | 0.0019 | N/A | N/A |
| p013 | 0.0020 | 2 | 2 | 0.0039 | 0.0000 | 0 | 0 | 0.0000 | N/A | N/A |
| p014 | 0.0773 | 79 | 75 | 0.1468 | 0.0877 | 90 | 87 | 0.1696 | 0.9 | 0.39 |
| Total | 0.9961 | 1018 | | | 0.9951 | 1021 | | | | |

and used to assign a *TGFB1* regulatory region and exon 1 allelic genotype[4,5] based on the genotypes for 18 known polymorphic positions. In cases where there were theoretical ambiguities, the phase of the relevant polymorphic positions was defined by allele-specific amplification strategies using different primer combinations.[6]

### Cellular assays

$CD4^+CD25^-$ and $CD4^+CD25^+$ cells were isolated from peripheral blood mononuclear cells (PBMC) with a human $CD4^+CD25^+$ Regulatory T-Cell Isolation Kit (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany).

Isolated cell fractions were stained with antibodies against CD4 (PerCP, clone SK3, BD Biosciences, Oxford, UK; APC, clone RPA-T4, eBioscience, San Diego, USA), CD127 (FITC, clone eBioRDR5, eBioscience; PerCP, clone eBioRDR5, eBioscience, San Diego, USA), and CD25 (APC, clone 2A3, BD Biosciences, Oxford, UK; PerCP-Cy5.5, clone BC96, BioLegend, San Diego, USA). Surface TGF-β1 expression on isolated Treg was assessed by staining of its latency-associated peptide (αLAP-PE, clone 27232, R&D Systems, Abingdon, UK) on resting and activated CD4+CD25+CD127lo cells.

The cells were activated with antibodies against CD3 and CD28 (NA/LE mouse, clones HIT3a and CD28.2, respectively, BD Biosciences, Oxford, UK) at 10 μg/mL. Non-stimulated and plate-bound antibody-stimulated cells were used as controls.

### Statistical analysis

The Z-test was used to compare *TGFB1* regulatory region and exon 1 allele variant and genotype frequencies between the HSCT patient and donor cohorts (*Online Supplementary Table S2*). Deviation from Hardy-Weinberg equilibrium was assessed with Fisher's exact test or $\chi^2$ test (*Online Supplementary Tables S3 and S4*). Detailed information on SNP and allele frequencies and the assignment of *TGFB1* regulatory region and exon 1 genotypes is available in the *Online Supplementary Appendix*.

The main clinical end point was overall survival (OS). Secondary end points were event-free survival (EFS), non-relapse mortality (NRM), acute graft-*versus*-host disease (aGvHD) (grades I-IV, II-IV or III-IV), and relapse. EFS was defined as survival without relapse (an event was death of any cause or relapse). For univariate analysis of time-to-event data (OS, EFS), the Kaplan-Meier method was used. Log rank statistics were used to compare OS and EFS probabilities between groups of interest. The probabilities of NRM and relapse were estimated by the cumulative incidence method, and compared using Gray's test, with relapse and death without relapse as competing events, respectively. aGvHD frequencies were compared by means of the $\chi^2$ test, or by Fisher's exact test.

Multivariate analyses were performed using Cox's regression (OS, EFS), the Fine-Gray method (NRM, relapse), or logistic regression (aGvHD) as appropriate. Clinical variables with $P \leq 0.2$ in univariate models for association with transplant outcome were selected for multivariate analyses.

Kruskal-Wallis and Mann-Whitney tests were used to compare Treg LAP expression levels between *TGFB1* regulatory region and exon 1 genotype groups.

## Results

### Cohort

The cohort was composed of 522 unrelated myeloablative transplants performed between 1996 and 2009. Typing was possible for only patient or donor DNA for 9 and 11 pairs, respectively. Although permission for genetic testing was granted, permission for use of clinical data was not granted in 18 cases [patient *TGFB1* genotypes: p001/p003 (n=7), p003/p003 (n=6), p001/p001 (n=2), p006/p014 (n=2), and p014/p014 (n=1)] and these were thus excluded. Consequently, when clinical data were analyzed, the final number of pairs included for patient and donor genotypes were 493 and 495, respectively ('whole cohort'). The characteristics of the patients, their donors and the transplants are presented in Table 1. T-cell depletion with alemtuzumab was used in 85% of the patients.

### Descriptive results for the typing of TGFB1 regulatory region and exon 1 alleles in the patient-donor cohort

Only six of the previously reported alleles were seen in the cohort: p001, p003, p006, p009, p013, and p014, four of which were predominant (Table 2). *Online Supplementary Table S5* shows the genotype frequencies observed. Neither the allele nor the genotype frequencies differed significantly between patients and donors (Z test; $P>0.050$).

Nine samples (5 donors and 4 patients) showed genotypes that did not correspond with any allelic combination based on the previously known 17 *TGFB1* regulatory region and exon 1 alleles. These samples were shown to carry a combination of a known allele and a novel allele.[6]

### Survival analysis

Median follow up in the cohort was 20.5 months (range 0.2-178.9 months). Five-year OS and EFS in the whole cohort were 40.9% (95%CI: 36.6%-45.2%) and 30.4% (95%CI: 26.3%-34.5%), respectively. Median OS was

21.6 months (95%CI: 11.5-31.6 months). Median EFS was 9.9 months (95%CI: 7.6-12.2 months). One-year cumulative incidence for NRM was 26.8% in the whole cohort. Five-year relapse cumulative incidence was 39.0%. Median time to relapse was 51.6 months (95%CI: 9.5-93.8 months).

The effect of *TGFB1* polymorphism on survival was assessed in three models: recessive allelic, dominant allelic, and SNP (-1347C>T)-associated effects. Both the effect of donor and patient-borne polymorphism was independently assessed.

### Recessive models

Variation in patient *TGFB1* regulatory region and exon 1 had a significant effect on the OS of the whole cohort. When homozygosity for alleles p001 and p003 and heterozygosity were compared (Figure 1A), significant differences were found (n=486 when excluding patients homozygous for p006 and p014 due to low numbers; *P*=0.041). Patients homozygous for p003 (n=132) had the highest median OS (43.8 months), while patients homozygous for the p001 allele (n=41) had the lowest (7.9 months). When pairwise comparisons were considered, there was a significant difference between patients homozygous for p001 and p003 (*P*=0.014), and a trend between p001 and the heterozygous group (n=313; *P*=0.071). Patients with a p001/p001 genotype (n=41) show significantly lower OS than the rest of the patients of any other genotype (n=452; 5-year OS 30.7% for p001/p001 patients *vs.* 41.6% others; *P*=0.032) (Figure 1B).

No differences in OS according to donor allele were found (n=491; *P*=0.47). Other *TGFB1* alleles could not be tested due to low numbers of homozygotes. Among all patient genotypes with n>20, only p001/p001 shows a significant effect on OS in the whole cohort when compared to the rest of the genotypes (*data not shown*).

### Dominant models

No effect of patient alleles was seen using this model (Figure 1C for p001). However, patients whose donors carried at least one copy of p001 had worse OS than patients whose donor lacked this allele (median OS 13.7 *vs.* 39.5 months, respectively; *P*=0.043). Alleles p003, p006 and p014 did not have statistically significant dominant donor effects on OS in this cohort.

### *TGFB1* -1347C>T (rs1800469)

The -1347T variant was close to a marker for allele p001 in this cohort (41/43 *TGFB1* -1347TT patients were p001/p001). There was no statistical evidence for a recessive effect of either patient (n=493; *P*=0.11) or donor (n=495; *P*=0.11) genotype on OS (Figure 2A). When the dominant model for -1347C was tested (i.e. TT *vs.* CT+CC) (Figure 2B), patients that had the -1347TT genotype (n=43) had significantly lower OS than that of patients bearing at least one C variant (median OS 7.9 *vs.* 25.1 months; *P*=0.036). No evidence of a donor genotype effect on OS was found in this model (*P*=0.82) (Figure 2C).

### Analysis of EFS, NRM, relapse and aGvHD

There was a significant increase in the incidence of NRM among patients that bear the p001 allele (1-year NRM: 39.0%; *P*=0.039) or the -1347T (1-year NRM: 39.5%; *P*=0.029) in a homozygous manner when compared to other genotypes (1-year NRM: 25.4% and 25.3%,

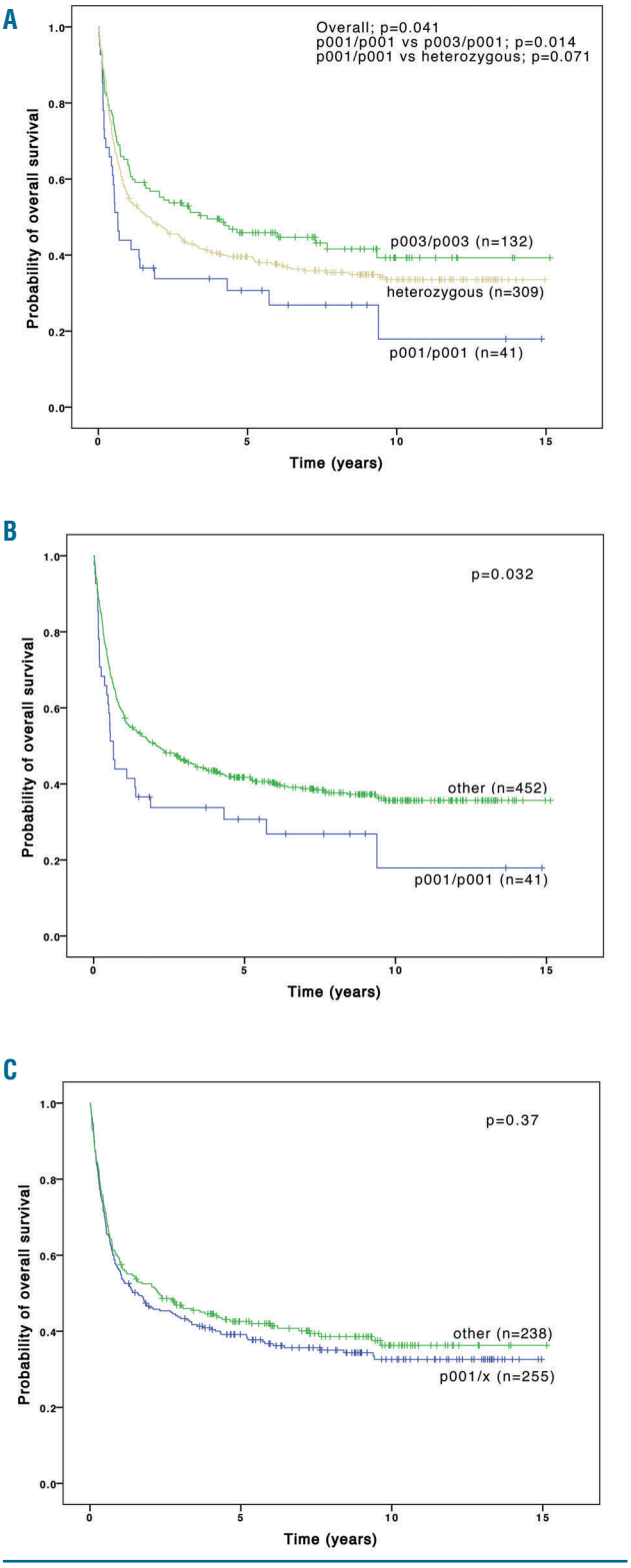


**Figure 1. Survival analysis according to the effects of patient *TGFB1* regulatory region and exon 1 allele for the whole cohort.** (A) Patients homozygous for allele p001 (n=41) show significantly worse overall survival (OS) when compared to patients homozygous for p003 (n=132; *P*=0.014). A trend for lower OS in patients homozygous for p001 was also found when compared to heterozygous patients (n=309; *P*=0.071). (B) Patients homozygous for *TGFB1* allele p001 (n=41) show reduced OS when compared with all other genotypes (n=452; *P*=0.032). (C) Patients with at least one copy of *TGFB1* allele p001 (n=255) do not show significantly different OS from patients with other genotypes (n=238; *P*=0.37).

respectively) (Figure 3A-C). There was no effect of the dominant presence of p001 among donors (Figure 3D).

None of the models tested impacted on EFS, disease relapse or aGvHD (grades I-IV, II-IV or III-IV).

### Multivariate analyses

Based on the univariate analyses for the clinical factors (Table 3), patient age, donor age, patient sex, HLA matching, disease status, cytomegalovirus (CMV) matching, and use of total body irradiation (TBI) were selected for inclusion in the multivariate model for OS. Likewise, patient age, HLA matching, CMV matching, use of TBI, and use of T-cell depletion were selected for the NRM model.

For OS, disease status at transplant and patient age together with the recessive allelic model were significant factors associated with this outcome (Table 4). When the -1347C dominant and the 'p001/p001 *versus* other genotype' models were examined, both were found to be significantly associated with OS together with patient age, HLA matching, and disease status. Overall, patients older than 40 years of age, not transplanted in complete remission/chronic phase nor from 10/10 HLA-matched donors, and being homozygous for *TGFB1* p001 (or -1347T) were associated with decreased OS.

**Table 3.** Analysis of the univariate association between clinical factors and overall survival and non-relapse mortality in the whole cohort (n=504).

| Risk factor | Median OS (months) | *P* | NRM (cumulative incidence)[3] | *P* |
|---|---|---|---|---|
| Patient age | | | | |
| > 40 years | 7.6 | <0.001 | 36.2 | 0.002 |
| < 40 years | 40.4 | | 23.6 | |
| Donor age | | | | |
| > 30 years | 17.0 | 0.081 | 28.8 | 0.38 |
| < 30 years | 45.0 | | 21.3 | |
| Sex | | | | |
| Female patients | 40.5 | 0.15 | 24.7 | 0.37 |
| Male patients | 17.0 | | 28.0 | |
| Female donors | 34.0 | 0.32 | 27.9 | 0.98 |
| Male donors | 18.7 | | 26.4 | |
| Sex matching | | | | |
| Overall matched | 22.4 | 0.84 | 26.9 | 0.72 |
| Overall mismatched | 21.3 | | 26.6 | |
| Female to male | 12.5 | 0.59 | 30.3 | 0.69 |
| Other | 22.4 | | 26.2 | |
| HLA-matching | | | | |
| 0 mismatches | 28.6 | 0.051 | 23.7 | 0.048 |
| 1 mismatch | 11.3 | | 31.0 | |
| >1 mismatch | 8.6 | | 41.3 | |
| Disease status | | | | |
| Low risk[1] | 26.4 | 0.004 | 26.4 | 0.44 |
| High risk | 7.6 | | 32.5 | |
| CMV status | | | | |
| Matched[2] | 27.2 | 0.035 | 24.0 | 0.012 |
| Not matched | 12.2 | | 35.5 | |
| TBI presence | | | | |
| Yes | 18.7 | 0.089 | 27.9 | 0.080 |
| No | N/A | | 18.1 | |
| SC source | | | | |
| BM | 19.5 | 0.66 | 29.4 | 0.24 |
| PB | 27.8 | | 21.4 | |
| T-cell depletion | | | | |
| Yes | 22.1 | 0.95 | 27.4 | 0.18 |
| No | 14.4 | | 13.6 | |
| GvHD prophylaxis | | | | |
| None | 7.8 | 0.23 | N/A | 0.74 |
| Cyclosporin | 19.1 | | 26.7 | |
| Cyclosporin+MTX | 25.7 | | 26.4 | |
| Other | 8.9 | | 28.6 | |
| Unknown | 7.9 | | 33.3 | |
| Transplant year | | | | |
| 1996-2001 | 14.4 | 0.23 | 29.6 | 0.36 |
| 2002-2009 | 32.1 | | 23.9 | |

*[1]Complete remission/chronic phase; [2]cytomegalovirus (CMV) matching: CMV positive-CMV positive; CMV negative-CMV negative; [3]1-year NRM cumulative incidence. BM: bone marrow; GvHD: graft-*versus*-host disease; HLA: human leukocyte antigen; MTX: methotrexate; NRM: non-relapse mortality; OS: overall survival; PB: peripheral blood; SC: stem cell; TBI: total body irradiation.*

For NRM, patient homozygosity for *TGFB1* p001 (or -1347T), patient age older than 40 years, and the presence of one or more allelic HLA mismatches (i.e. ≤9/10) were associated with increased probability of death (Table 4).

### *Functional consequences of* TGFB1 *regulatory region and exon 1 alleles in Treg*

When Treg from healthy donors were stimulated with antibodies against CD3 and CD28, an upregulation of surface LAP, which peaked at 24 h of culture, was detected. This upregulation was observed only on the $CD4^+$, CD127lo cells and $CD25^+$ cells, as previously described.[8]

As shown in Figure 4, *TGFB1* genotype appears to influence the levels of LAP expressed by Treg upon TCR stimulation. A trend towards higher LAP+ levels was seen when the sample expressed a p001 allele (Mann-Whitney test; *P*=0.07). An analysis of p001/p001 individuals on their own was not possible because of the reduced frequency of this genotype among available healthy volunteer donors.

## Discussion

The present study revealed that homozygosity for a *TGFB1* p001 allele in UD-HSCT patients was associated with significantly worse OS and NRM. Cellular experiments suggest a potential functional effect of TGFB1 p001, as there was a trend toward higher expression of surface TGF-β1 on *in vitro* stimulated Treg that bore this allele. This study is the largest performed so far on the role of *TGFB1* polymorphisms in HSCT. In contrast to previous studies, the analysis encompassed the combined effect of various polymorphisms organized in defined alleles in a genomic region of approximately 3 kb.

A few studies have analyzed *TGFB1* polymorphism in HSCT, but their heterogeneity makes comparisons difficult.[7] Most previous studies are small (54% included less than 100 pairs) or have investigated rare alleles. Moreover, most of the studies have focused their analysis on one or two SNPs and only on their impact on GvHD. Two early studies also used pre-existing classifications of the genotypes in "high producer" and "low producer" groups, potentially introducing a bias in their analyses.[9-11]

Studies in mostly related donor cohorts have found no association for *TGFB1* +29T>C and +74G>C or their combined +29~+74 genotypes with OS, GvHD, engraftment or infections.[9,12] In a larger study with mismatched UD-HSCT, there was no consistent association of *TGFB1* -1347C>T with OS, engraftment or GvHD, despite initial findings in a discovery cohort.[13] Finally, in two recent reports analyzing relatively large cohorts of mostly related donor transplants, *TGFB1* -1347 TT and CT patients showed increased incidence of aGvHD, but no effect on OS, EFS, or NRM.[14,15] *TGFB1* -1347T was found to be a risk factor for skin aGvHD but protective against lung cGvHD.[15]

In our analyses, there was no statistical confirmation of a role of +29T>C16 (*data not shown*). An explanation for this difference could be the fact that +29CC genotypes could include both p001 and p014.[4] The presence of p014 could not be analyzed for a recessive effect in our cohort. However, since OS in +29CC patients was not statistically different from +29TT and TC individuals, this could suggest the lack of effect from p014. Interestingly, a study performed in Chinese HSCT patients (n=240) found lower

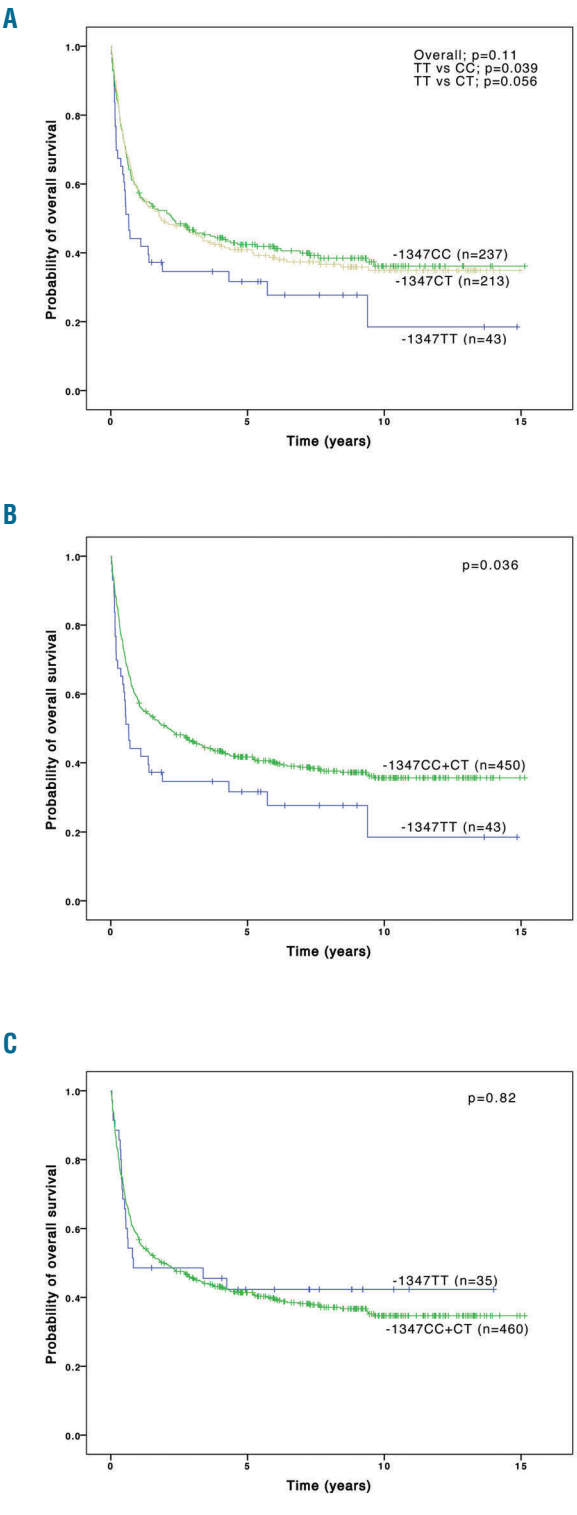


**Figure 2. Survival analysis according to the effects of patient TGFB1 -1347C>T for the whole cohort.** (A) Among the whole cohort (n=493), patients bearing a -1347TT genotype (n=43) show a trend toward lower OS (*P*=0.11) when compared to other genotypes. When pairwise comparisons are made, patients with a -1347TT genotype (n=43) have significantly lower OS than patients with -1347CC (n=237; *P*=0.039). In a dominant model for the C variant, (B) the presence of a TGFB1 -1347TT in patients results in significantly lower overall survival (OS) [(median OS TT (n=43) *vs.* CC+CT (n=450) 7.9 *vs.* 25.1 months; *P*=0.036)]. (C) No effect of the donor genotype (TT, n=35; CC+CT, n=460) was suggested by this dominant model (*P*=0.82).

incidence of aGvHD in patients whose donors were *TGFB1* -1347TT individuals, and also in patients who bore at least one copy of the T variant, but with no effect on OS, NRM or relapse.[17] However, it is uncertain if *TGFB1* -1347T correlates with allele p001 in the Chinese population.

In the present study, the effect seen for allele p001 on OS and NRM could not be explained by increases in the incidence of aGvHD. This might be due to the fact that most of the transplants included in our cohort were T-cell depleted, and the incidence of aGvHD was low. Despite this, there remains the possibility that this cytokine could modify this complication, for example, by affecting the generation of Th17 cells.[18] Alternatively, the genetic association with NRM could potentially be explained by

**Table 4. Analysis of the multivariate association between clinical factors and TGFB1 regulatory region and exon 1 polymorphism and their effect on overall survival and non-relapse mortality in the whole cohort.**

| Outcome | Model | Factor1 | HR [95%CI] | *P* |
|---|---|---|---|---|
| OS | Recessive allelic (n=476) | Patient p001/p001 *vs.* heterozygous | 1.38 [0.94-2.04] | 0.103 |
| | | Patient p003/p003 *vs.* heterozygous | 0.78 [0.59-1.02] | 0.065 |
| | | Patient age < 40 *vs.* > 40 | 0.61 [0.47-0.78] | <0.001 |
| | | Disease status (low risk *vs.* high risk) | 0.72 [0.53-0.97] | 0.032 |
| | Dominant -1347C (n=483) | Patient -1347 TT *vs.* CC+CT | 1.52 [1.04-2.21] | 0.031 |
| | | Patient age < 40 *vs.* > 40 | 0.59 [0.46-0.76] | <0.001 |
| | | Disease status (low risk *vs.* high risk) | 0.73 [0.54-0.98] | 0.038 |
| | | HLA matched 10/10 *vs.* ≤ 9/10 | 0.77 [0.60-0.98] | 0.034 |
| | p001/p001 *vs.* other genotypes (n=483) | Patient p001/p001 *vs.* other | 1.53 [1.04-2.24] | 0.031 |
| | | Patient age < 40 *vs.* > 40 | 0.59 [0.46-0.76] | <0.001 |
| | | Disease status (low risk *vs.* high risk) | 0.73 [0.54-0.98] | 0.039 |
| | | HLA matched 10/10 *vs.* ≤ 9/10 | 0.77 [0.61-0.98] | 0.035 |
| NRM | Recessive allelic (n=486) | Patient p001/p001 *vs.* heterozygous | 1.80 [1.08-3.00] | 0.024 |
| | | Patient p003/p003 *vs.* heterozygous | 1.11 [0.76-1.60] | 0.580 |
| | | HLA matched 10/10 *vs.* ≤ 9/10 | 0.73 [0.52-1.01] | 0.065 |
| | | Patient age < 40 *vs.* > 40 | 0.60 [0.43-0.84] | 0.003 |
| | Dominant -1347C (n=493) | Patient -1347 TT *vs.* CC+CT | 1.79 [1.09-2.92] | 0.020 |
| | | HLA matched 10/10 *vs.* ≤ 9/10 | 0.70 [0.50-0.97] | 0.031 |
| | | Patient age < 40 *vs.* > 40 | 0.58 [0.42-0.81] | 0.001 |
| | p001/p001 *vs.* other genotypes (n=493) | Patient p001/p001 *vs.* other | 1.73 [1.06-2.83] | 0.030 |
| | | HLA matched 10/10 *vs.* ≤ 9/10 | 0.70 [0.50-0.97] | 0.032 |
| | | Patient age < 40 *vs.* > 40 | 0.58 [0.42-0.81] | 0.001 |

[1]*Factors are compared to the last one listed for their hazard ratio (HR). NRM: non-relapse mortality; OS: overall survival.*

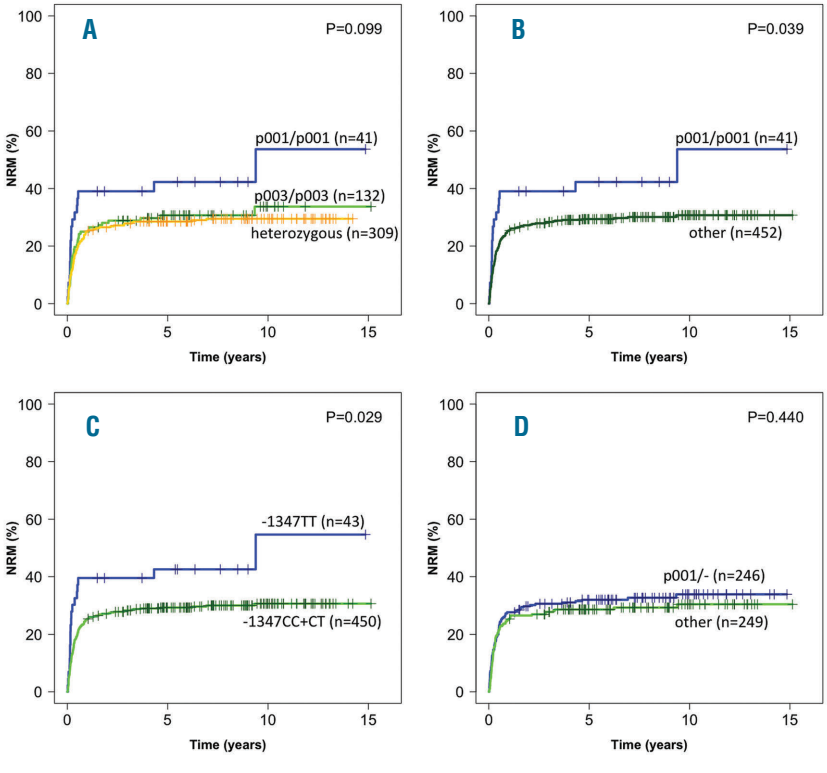


**Figure 3. Survival analysis according to the effects of patient and donor TGFB1 regulatory region and exon 1 genotype on non-relapse mortality (NRM) for the whole cohort. Allele p001 and -1347T homozygosity in patients showed a significant increase in NRM in the whole cohort.** No effect on NRM was observed for the donor p001 dominant model. (A) Recessive patient allelic model [(p001/p001 (n=41) *versus* p003/p003 (n=132) *versus* heterozygous (n=309); overall *P*=0.099)]; (B) patient p001/p001 (n=41) *versus* other (n=452; *P*=0.039); (C) dominant patient -1347C (-1347 TT (n=43) *versus* CT+CC (n=450); *P*=0.029); (D) dominant donor p001 (p001/- (n=246) *versus* other (n=249); *P*=0.440).

another cause of death in which TGF-β1 might play a role, such as impaired early immune responses to infectious agents,[19] organ damage,[20] or complications such as hepatic veno-occlusive disease.[21] However, this remains unclear. Interestingly, a recent report[22] has shown evidence of a role for TGF-β1 in limiting both the growth and function of the thymic medulla, another potential niche for its influence on the outcome of HSCT.

Our typing results revealed that four *TGFB1* regulatory region and exon 1 alleles predominated. Even though this was not a population study, our results provide insight into *TGFB1* regulatory region allelic diversity and frequencies and are a potential reference for future studies. Overall, the frequencies for variant polymorphic positions and for *TGFB1* regulatory region and exon 1 alleles agree widely with data previously reported by other studies.[23]

We speculated that the strong detrimental effect of patient p001 observed in this study was related to differences in functionality between *TGFB1* regulatory region and exon 1 alleles. Our study showed that the level of surface TGF-β1 on Treg after TCR stimulation appears to be modified by the presence of the p001 allele in *TGFB1*.

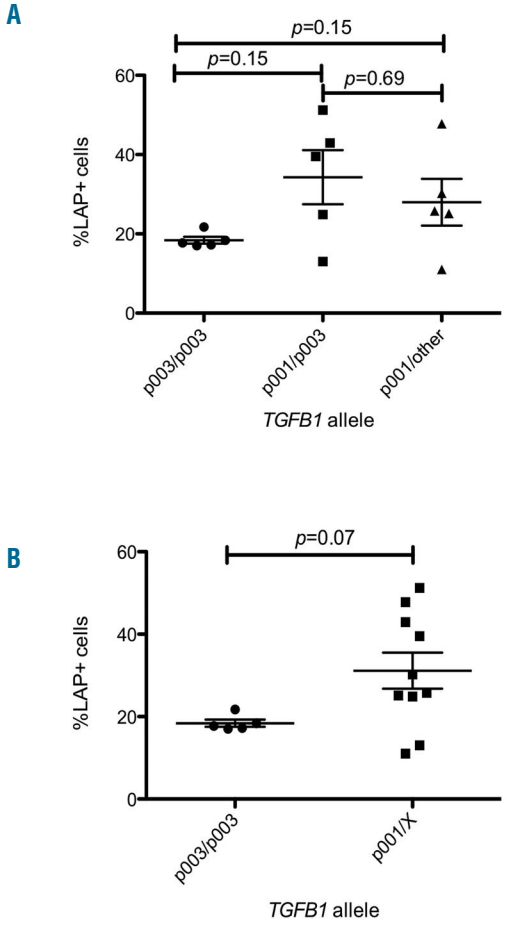


**Figure 4. Surface TGF-β1 (LAP) expression on Treg upon TCR stimulation differs according to *TGFB1* regulatory region and exon 1 genotype in healthy donors.** (A) Percentage of LAP$^+$ cells between different TGFB1 regulatory region and exon 1 genotypes. (B) The percentage of LAP+ cells within the CD4$^+$CD25$^+$CD127$^{lo}$ gate shows a trend toward higher expression when a p001 allele at TGFB1 regulatory region is present. p001/other: p001/p001 (2), p001/p014 (2), p001/new (1). Shown are mean and standard error.

Even though it did not reach statistical significance, *TGFB1* p001/x genotypes showed results that suggested higher generation of LAP+ cells when compared to TT individuals, following previous observations in other cell types and experimental systems.

The -1347T variant has been previously associated with higher TGF-β1 plasma levels,[24] as well as with a significant increase in *in vitro* TGF-β1 expression[25] *via* alteration of promoter interactions with transcription factors Yin Yang 126 and AP1.[27] Combining both the observations made for TGFB1 -1347C>T and those made for +29T>C,[28] Shah *et al.* proposed that *TGFB1* alleles that share a -1347T and +29C would represent a high production phenotype.[4] Allele p001 would be the sole representative of this cluster seen with significant frequencies in our cohort. Interestingly, a couple of studies have found opposite results and associated -1347T and +29C with lower plasma concentrations of this cytokine and lower reporter gene activities,[29] and a *TGFB1* upstream haplotype congruent with allele p001 with weaker promoter activity than another haplotype fitting with allele p003.[30] However, the genomic region examined in the latter study only partially spanned the one studied here and included different SNPs not characterized in this study. Finally, one study associated allelic variants carrying a proline either in codon 10 (+29C) or 25 (+74C) with reduced expression, but only included *TGFB1* coding region in *in vitro* constructs.[31]

In addition to a -1347T variant, a feature that is unique to p001 is the absence of the -2389AGG duplication (rs11466313), and this has been associated with the gain of allele DNA-protein complexes, potentially leading to novel transcription factor binding site motifs.[30]

Low frequency of homozygotes for some of the *TGFB1* alleles precluded thorough analysis of their effects. A much larger study would be needed for it to be possible to assess homozygous individuals. In addition, since our cohort was comprised mainly of alemtuzumab-T-cell- depleted transplants, changes in its dosage or dosage schedule could have taken place over the 13-year observation period, potentially having an impact on our results. Unfortunately, this information is not available for assessment. Finally, we do not have data on replication of these results in an independent cohort. Hence, these analyses should be confirmed in other settings, such as non-myeloablative transplants or transplants performed with alternative donors.

In conclusion, the fact that patients having a p001/p001 genotype have significantly higher probabilities of dying early after the transplant could potentially allow for better pre-emptive measures to improve the prognosis for these patients. However, further research is needed to understand the mechanism of this effect and the cause(s) of death associated with it.

***Acknowledgments***

*The authors would like to thank all of the UK transplant teams who have contributed patients' data and samples to this study. We thank Dr. Richard Szydlo for statistical advice and Prof. Katharina Fleischhauer for critically reviewing this manuscript.*

***Funding***

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

## References


1. Lee SJ, Klein J, Haagenson M, et al. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. Blood. 2007;110(13):4576-4583.
2. Dickinson AM, Norden J. Non-HLA genomics: does it have a role in predicting haematopoietic stem cell transplantation outcome? Int J Immunogenet. 2015;42(4):229-238.
3. Coomes SM, Moore BB. Pleiotropic effects of transforming growth factor-beta in hematopoietic stem-cell transplantation. Transplantation. 2010;90(11):1139-1144.
4. Shah R, Rahaman B, Hurley CK, Posch PE. Allelic diversity in the TGFB1 regulatory region: characterization of novel functional single nucleotide polymorphisms. Hum Genet. 2006;119(1-2):61-74.
5. Shah R, Selby ST, Yokley B, Slack RS, Hurley CK, Posch PE. TNF, LTA and TGFB1 genotype distributions among acute graft-vs-host disease subsets after HLA-matched unrelated hematopoietic stem cell transplantation: a pilot study. Tissue Antigens. 2009;74(1):50-56.
6. Arrieta-Bolaños E, Madrigal JA, Shaw BE. Novel alleles of the transforming growth factor -1 regulatory region and exon 1. Tissue Antigens. 2015;85(6):484-491.
7. Arrieta-Bolanos E, Alejandro Madrigal J, Shaw BE. Transforming growth factor-beta1 polymorphisms and the outcome of hematopoietic stem cell transplantation. Int J Immunogenet. 2012;39(3):192-202.
8. Chen W, Wahl SM. TGF-beta: the missing link in CD4+CD25+ regulatory T cell-mediated immunosuppression. Cytokine Growth Factor Rev. 2003;14(2):85-89.
9. Laguila Visentainer JE, Lieber SR, Lopes Persoli LB, et al. Relationship between cytokine gene polymorphisms and graft-versus-host disease after allogeneic stem cell transplantation in a Brazilian population. Cytokine. 2005;32(3-4):171-177.
10. Leffell MS, Vogelsang GB, Lucas DP, Delaney NL, Zachary AA. Association between TGF-beta expression and severe GVHD in allogeneic bone marrow transplantation. Transplant Proc. 2001;33(1-2):485-486.
11. Tambur AR, Yaniv I, Stein J, et al. Cytokine gene polymorphism in patients with graft-versus-host disease. Transplant Proc. 2001;33(1-2):502-503.
12. Sellathamby S, Lakshmi KM, Busson M, et al. Polymorphisms in the immunoregulatory genes are associated with hematopoietic recovery and increased susceptibility to bacterial infections in patients with thalassaemia major undergoing matched related hematopoietic stem cell transplantation. Biol Blood Marrow Transplant. 2012;18(8):1219-1226.
13. Harkensee C, Oka A, Onizuka M, et al. Single nucleotide polymorphisms and outcome risk in unrelated mismatched hematopoietic stem cell transplantation: an exploration study. Blood. 2012;119(26):6365-6372.
14. Kim DD, Yun J, Won HH, et al. Multiple single-nucleotide polymorphism-based risk model for clinical outcomes after allogeneic stem-cell transplantation, especially for acute graft-versus-host disease. Transplantation. 2012;94(12):1250-1257.
15. Kim D, Won HH, Su S, et al. Risk stratification of organ-specific GVHD can be improved by single-nucleotide polymorphism-based risk models. Bone Marrow Transplant. 2014;49(5):649-656.
16. Berro M, Mayor NP, Maldonado-Torres H, et al. Association of functional polymorphisms of the transforming growth factor B1 gene with survival and graft-versus-host disease after unrelated donor hematopoietic stem cell transplantation. Haematologica. 2010;95(2):276-283.
17. Xiao H, Cao W, Lai X, et al. Immunosuppressive cytokine gene polymorphisms and outcome after related and unrelated hematopoietic cell transplantation in a chinese population. Biol Blood Marrow Transplant. 2011;17(4):542-549.
18. Veldhoen M, Hocking RJ, Atkins CJ, Locksley RM, Stockinger B. TGFbeta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. Immunity. 2006;24(2):179-189.
19. Coomes SM, Wilke CA, Moore TA, Moore BB. Induction of TGF-beta 1, not regulatory T cells, impairs antiviral immunity in the lung following bone marrow transplant. J Immunol. 2010;184(9):5130-5140.
20. Anscher MS, Kong FM, Jirtle RL. The relevance of transforming growth factor beta 1 in pulmonary injury after radiation therapy. Lung Cancer. 1998;19(2):109-120.
21. Reimer J, Bien S, Ameling S, et al. Antineoplastic agent busulfan regulates a network of genes related to coagulation and fibrinolysis. Eur J Clin Pharmacol. 2012;68(6):923-935.
22. Hauri-Hohl M, Zuklys S, Hollander GA, Ziegler SF. A regulatory role for TGF-beta signaling in the establishment and function of the thymic medulla. Nat Immunol. 2014;15(6):554-561.
23. Hoffmann SC, Stanley EM, Cox ED, et al. Ethnicity greatly influences cytokine gene polymorphism distribution. Am J Transplant. 2002;2(6):560-567.
24. Grainger DJ, Heathcote K, Chiano M, et al. Genetic control of the circulating concentration of transforming growth factor type beta1. Hum Mol Genet. 1999;8(1):93-97.
25. Luedecking EK, DeKosky ST, Mehdi H, Ganguli M, Kamboh MI. Analysis of genetic polymorphisms in the transforming growth factor-beta1 gene and the risk of Alzheimer's disease. Hum Genet. 2000;106(5):565-569.
26. Silverman ES, Palmer LJ, Subramaniam V, et al. Transforming growth factor-beta1 promoter polymorphism C-509T is associated with asthma. Am J Respir Crit Care Med. 2004;169(2):214-219.
27. Shah R, Hurley CK, Posch PE. A molecular mechanism for the differential regulation of TGF-beta1 expression due to the common SNP -509C-T (c. -1347C > T). Hum Genet. 2006;120(4):461-469.
28. Dunning AM, Ellis PD, McBride S, et al. A transforming growth factorbeta1 signal peptide variant increases secretion in vitro and is associated with increased incidence of invasive breast cancer. Cancer Res. 2003;63(10):2610-2615.
29. Wang H, Zhao YP, Gao CF, et al. Transforming growth factor beta 1 gene variants increase transcription and are associated with liver cirrhosis in Chinese. Cytokine. 2008;43(1):20-25.
30. Healy J, Dionne J, Belanger H, et al. Functional impact of sequence variation in the promoter region of TGFB1. Int J Cancer. 2009;125(6):1483-1489.
31. Mohren S, Weiskirchen R. Non-synonymous gene polymorphisms in the secretory signal peptide of human TGF-beta1 affect cellular synthesis but not secretion of TGF-beta1. Biochem Biophys Res Commun. 2009;379(4):1015-1020.