European Journal of Human Genetics (2009) 17, 1454–1462

www.nature.com/ejhg

ARTICLE

A critical assessment of the factors affecting reporter gene assays for promoter SNP function: a reassessment of −308 *TNF* polymorphism function using a novel integrated reporter system

Mahdad Karimi[1], Lauren C Goldie[1,3], Mark N Cruickshank[1], Eric K Moses[2] and Lawrence J Abraham*[,1]

[1]*Department of Biochemistry and Molecular Biology, School of Biomedical, Biomolecular and Chemical Sciences, University of Western Australia, Crawley, WA, Australia and* [2]*Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX USA*

One of the greatest challenges facing genetics is the development of strategies to identify functionally relevant genetic variation. The most common test of function is the reporter gene assay, in which allelic regulatory regions are used to drive the expression of a reporter gene, and differences in expression in a cell line after transient transfection are taken to be a reflection of the polymorphism. Many studies have reported small differences in single nucleotide polymorphism (SNP)-specific reporter activity, including the tumor necrosis factor (*TNF*) G−308A polymorphism. However, we have established that many variables inherent in the reporter gene approach can account for the reported allelic differences. Variables, such as the amount of DNA used in transfection, the amount of transfection control vector used, the method of transfection, the growth history of the host cells, and the quality and purity of DNA used, all influence *TNF* −308 SNP-specific transient reporter gene assays and serve as a caution for those researchers who apply this method to the functional assessment of polymorphic promoter sequences. We have developed an integrated reporter system that obviates some of these problems and shows that the *TNF* G−308A polymorphism is functionally relevant in this improved assay, thus confirming that the −308A allele expresses at a higher level compared with the −308G allele.
European Journal of Human Genetics (2009) **17,** 1454–1462; doi:10.1038/ejhg.2009.80; published online 27 May 2009

Keywords: functional assay; TNF; genetic variant; integrated reporter

Introduction

One of the greatest challenges in genetics is the development of strategies to distinguish genetic variation that is functionally significant from variation that is not. Approximately 14.7 million single nucleotide polymorphisms (SNPs) have been found in the human genome to date, over half of them being located in non-coding regions (see NCBI dbSNP build 129). Genome-wide association studies are currently being used to identify polymorphisms that underlie disease susceptibility in the population at large.[1] The consensus now emerging from these studies is that, either a large number of low penetrance polymorphisms or a small number of high penetrance polymorphisms contribute to the genetic predisposition to common diseases.[2] These studies test only a small subset of

*Correspondence: Professor LJ Abraham, Department of Biochemistry and Molecular Biology, School of Biomedical, Biomolecular and Chemical Sciences, University of Western Australia, 35 Stirling Highway, Crawley, WA, 6009, Australia.
Tel: +61 8 648 830 41; Fax: +61 8 648 811 48;
E-mail: labraham@cyllene.uwa.edu.au
[3]Present address: Therapeutic Target Discovery Division, Regeneron Pharmaceuticals Inc. New York, NY, USA.
Received 13 January 2009; revised 1 April 2009; accepted 15 April 2009; published online 27 May 2009

the full SNP complement, relying on (fortuitously) either typing the functional SNP(s) or on the possible existence of linkage disequilibrium between genotyped SNPs and the functional polymorphism(s). The unpredictability of linkage disequilibrium means that this approach does not necessarily have the power, for instance in the case of rare variants, to identify individual causative SNPs. One approach to overcome problems in identifying individual genetic variants that cause a functional change in the associated gene and hence, the disease susceptibility locus, is to statistically prioritize candidate SNPs for their likely effect on phenotype. This prioritization can be done either before or after carrying out an association study. One such post-analysis approach involves defining all genetic variations in a defined genomic region and then using Bayesian statistics (BQTN analysis) to objectively prioritize the potentially functional SNPs within the particular region.[3] As an example, we have used such an approach to identify the −137 *Vanin 1* promoter SNP as being most likely to affect the *Vanin 1* expression.[4] Such prioritization is usually followed by the validation of function using a range of biological tests.[5–7] The most common test of function is the reporter gene assay, in which allelic regulatory regions are used to drive the expression of a reporter gene, and differences in expression in a cell line after transient transfection are taken to be a reflection of the polymorphism in the regulatory sequences. However, the reporter gene assay was originally designed to define the major transcriptional elements that have an effect on basal and/or inducible gene expression. Differences caused by a regulatory SNP contributing to a quantitative trait are likely to be at least an order of smaller magnitude (usually <two-fold) than those measured when deleting or mutating transcriptional elements essential for transcription. Small differences in reporter activity can be caused by a range of confounding factors rather than by a true reflection of sequence differences. For instance, the concentration and quality of the reporter DNA used[8] or the use of a co-transfected control reporter can have a major effect on reporter gene activity. Another major determinant of the veracity of the assay is the amount of reporter that gets into a particular cell.[9] Hundreds or even thousands of copies of the reporter gene construct can enter the cell, depending partly on the method of transfection. As most transcription factors are present at low concentration, only a small proportion of reporters entering the cell receive the full complement of proteins needed for the proper function of the control region,[10] leading to an aberrant function of specific control elements.[11–13]

Many gene regulatory regions have been shown to function inappropriately in transient transfection assays. In addition to the concentration effects outlined above, aberrant activity may also be caused by the lack of an appropriate chromatin conformation.[12] In many cases, long-range enhancer and silencer elements are not functional unless the reporter gene is integrated into the genome by either stable transfection or transgenic means. As the chromatin context in which a particular transcriptional element functions is likely to affect the interaction of the transcription factors with that element, the transient reporter gene assay, although a useful indicator of important regulatory regions, is not always an appropriate assay to assess the influence of small differences due to SNP variation, on the transcriptional activity of a particular gene.

We have chosen the well-studied tumor necrosis factor (*TNF*) G−308A promoter SNP to assess parameters that influence allele-specific reporter gene activity. The −308A allele is of interest as it has been linked to a wide variety of both MHC- and non-MHC-associated diseases with a dominant autoimmune and/or inflammatory component. A number of studies on the possible functional relevance of the −308 polymorphism have been carried out using transient transfection assays of reporter genes controlled by allelic variants of the *TNF* promoter. Some studies claim to show significant differences in the transcriptional activity between *TNF* G−308A allelic reporters,[12,14–17] whereas others have not been able to detect any significant differences in activity.[18–21] Numerous factors may be involved in the divergent results, including cell type and culture conditions, transfection methodology, and type of reporter construct.[22,23] In this study, we assess some of the variables that may affect the results of the reporter gene assay, and validate an alternative reporter gene system that provides compelling evidence for a differential functional role of the −308 SNP.[24]

Materials and methods

Cell lines

The Jurkat E6-1 cell line (ATTC) and the derivative Flp-In -Jurkat (Invitrogen, Carlsbad, CA, USA) were cultured in RPMI-1640 (Trace Biosciences, Castle Hill, NSW, Australia) supplemented with 100 μg/ml each of penicillin and streptomycin (Trace Biosciences) and 10% FBS (Thermo Trace Ltd, Noble Park, VIC, Australia) at 37°C in 5% CO_2.

TNF/GL3Luc and TNF/GL3Luc/3′UTR reporter constructs

Allelic −308G and −308A *TNF* promoter reporters were constructed as described,[15] except that the original pGL2 reporter backbone was replaced with pGL3 vector sequences. Each of the reporters contained the −992 to +110 region of the *TNF* promoter. When required, linearized −308G/GL3Luc and −308A/GL3Luc reporter DNA was prepared by digestion at the *BamH* I site 2004 nucleotides immediately downstream of the poly(**A**) signal (from pGL3-Basic, Promega Inc, Madison, WI, USA) and ethanol precipitated.

Transient transfection assays

Mycoplasma-free Jurkat E6-1 cell lines were transiently transfected using either electroporation or when required, the lipid-mediated transfection protocol (Lipofectamine 2000, Invitrogen, Carlsbad, CA, USA). In general, transfections were performed using endotoxin-free supercoiled plasmid DNA from aliquots diluted to 1 μg/μl stored at −20°C prepared using the QIAFilter Endofree Plasmid Maxi Kit (Qiagen, Valencia, CA, USA). Unless stated otherwise, transfections were performed using electroporation; the cells were grown to a late-log phase of growth (1×10^6 cells per ml) and harvested in a total volume of 400 μl composed of complete growth media (2.5×10^7 cells per ml). For each transfection, the total volume was transferred into a sterile Bio-Rad GenePulser electroporation cuvette (0.4 cm; Bio-Rad Laboratories, Regent Park, NSW, Australia) along with 10 μg of TNF/GL3Luc and/or 3′UTR plasmid DNA, gently mixed, then pulsed with 240 V, 960 μFD capacitance using a Bio-Rad GenePulser electroporation apparatus (Bio-Rad Laboratories). Freshly electroporated cells were then transferred into 25 ml of fresh complete growth media and distributed as 1-ml aliquots (4×10^5 cells per ml) in 24-well flat-bottom culture dishes. The transfected cells were then incubated at 37°C and 5% CO_2 for 24 h unless stated otherwise. Cell lysates were prepared and assayed for both firefly and Renilla luciferase activity using the Dual-Luciferase Assay Kit (Promega) according to the manufacturer's instructions.

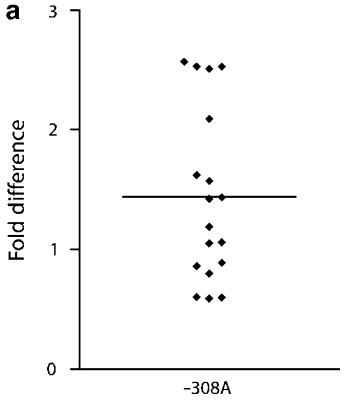
Cell cycle analysis and Flow Cytometry A method for the quantitation of DNA content without loss of GFP was modified from published methods.[25] A gentle fixation step of 0.25% paraformaldehyde for 5 min was followed by three washes in 1 × phosphate-buffered saline (PBS; Gibco BRL, Melbourne, VIC, Australia). Cell membranes were then permeabilized in 70% ethanol on ice for 30 min, washed thrice in PBS, and resuspended in 0.5 ml of Propidium Iodide staining solution (5 μg/ml PI (Sigma Chemical, Castle Hill, NSW, Australia), 5 mM EDTA (Sigma), 5 mM EGTA (Sigma), 0.1% NP-40 (Roche Diagnotic, Sydney, NSW, Australia), and 100 μg/ml RNase A (Roche) in PBS) left on ice until flow cytometric analysis.

Flow cytometric analysis was performed on an EPICS XL flow cytometer (Beckman Coulter, Mannheim, Germany). For each sample, a minimum of 10 000 events was collected and multivariate data analysis was performed using FlowJo (TreeStar Inc., Ashland, OR, USA).

Results

Reproducibility of the −308 *TNF* reporter gene assay

To assess the veracity of our earlier published data on the functional effects of the G−308A *TNF* promoter polymorphism, [15,23] 18 independent transient transfection experiments were carried out in Jurkat cells. Cells were transfected with freshly prepared DNA using electroporation without co-transfection with the normalization vector. Instead, cells were separately transfected with normalization control vector paired for each test transfection. Transfection efficiency as assessed by Renilla luciferase activity, did not vary by >10% within any experiment. Statistical analysis of the test results shows that a modest but significant increase in the transcriptional activity of the −308A allele was observed (Figure 1a). However, an inspection of the data suggested that the results fell into two distinct groups. Cluster analysis was carried out to confirm that there were multiple, statistically distinct groups using the hierarchical TwoStep Cluster analysis procedure (SPSS TwoStep Cluster Component SPSS Inc., Chicago Ill, USA), which assumes



b

Cluster	*N*	% of combined	Mean
1	5	27.8%	2.4460
2	13	72.2%	1.0516
Total	18	100%	1.4389

Figure 1 Distribution of Luciferase expression from 18 independent experiments of −308A normalized to −308G and TwoStep cluster analysis. (**a**) Graphical representation of the fold difference in the transcriptional activity of −308A over −308G. Jurkat cells were transiently transfected using either −308G/GL3Luc or −308A/GL3Luc DNA and the cell lysates were harvested after 24 h. Luciferase activity was determined using the Dual-Luciferase reporter assay system (Promega). Each value ($n = 18$) represents the mean fold difference in luciferase activity of −308A/GL3Luc over −308G/GL3Luc, for which each experiment was performed in triplicate. (**b**) The percentage and mean of the data values from panel a clustered in two groups. Initially, the TwoStep Clustering Component of the SPSS Statistical Analysis software package was used to automatically determine whether the ideal number of clusters was two, using both the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC).

that variables in the cluster model are independent. To determine which number of clusters was the best, each of the cluster solutions were compared using both Schwarz's Bayesian Criterion and Akaike Information Criterion. Both analyses predicted that there were, optimally, two clusters. Figure 1b shows the percentage and mean of the data for each cluster, whereby either a 2.45-fold increase (cluster 1) or a 1.05-fold increase (cluster 2) in the level of −308A activity compared with −308G is predicted.

Given that the only difference between −308G/GL3Luc and −308A/GL3Luc reporter constructs is the G/A substitution at −308, these results strongly suggest that the −308 SNP is regulated differentially and results in up to a 2.5-fold increase in expression of the −308A allele compared with the −308G allele only under particular conditions.

DNA concentration affects differential expression of the −308G/A SNP in the *TNF* promoter

The optimal concentration of DNA in any transfection is influenced by several factors, including promoter strength, transfection efficiency of the cell line being used, and cellular toxicity. Too much DNA may result in growth retardation of the cells or in poor complexation with the transfection agent. Conversely, if too little is transfected, the reporter activity may be undetectable or may make the analysis of reporter differences difficult. To determine whether the concentration of DNA used in the reporter assay was able to explain the clustering of results observed earlier, a series of −308 reporter assays was carried out using different concentrations of reporter plasmid DNA. The results (Figure 2a) using −308 reporters that contained only the −993 to +110 region of the *TNF* gene,[15] showed that the use of 5 μg of DNA in a standard transfection led to a 1.53-fold difference in the expression of the two −308 allelic constructs. Using 3 μg of DNA resulted in a greater difference between reporter constructs (1.97-fold). At 1 μg per transfection (most commonly used in our transient transfection experiments), a 2.05-fold difference was observed. Notably, reduction of the DNA concentration to 0.5 μg resulted in the complete loss of −308 expression difference (1.0-fold).

Constructs containing both the *TNF* promoter region and from +1957 to +2792 of the *TNF* 3′UTR[15] were also tested. In this case, no significant differences were observed between the −308A and −308G constructs at any of the tested DNA concentrations (Figure 2b). These results contrasted with those generated using the promoter-only constructs and suggested that there may be a complex interaction between multiple variables as we have reported >two-fold differences in the past when the 3′UTR was present.[15,23] In those studies, co-transfection with a different transfection normalization control vector (pCAT) to that used in this study (pRL-TK) or different cell culture conditions may have affected the results (see results presented below and in Figure 3).

Transfection control vectors may negatively affect −308G/A *TNF* expression differences

To further investigate the nature of the variables that might influence expression of the polymorphism, the effect of the transfection normalization control vector, pRL-TK, which we now use routinely in our experiments, was assessed. A titration of the amount of control vector used in co-transfection, with constant amounts of the allelic *TNF* constructs, was carried out. The results showed a 1.57-fold greater luciferase activity of the −308A allele in the absence of pRL-TK (Figure 3). Co-transfection with increasing amounts of pRL-TK lead to a decrease in the difference and at higher amounts (>30 ng) of the control vector, the difference in activity between alleles is reversed, with the −308G allelic reporter showing higher levels of activity.

In searching for a molecular explanation for the results, an inspection of the pRL-TK sequence was carried out. Earlier findings established that the *TNF* −308G transcriptional elements contained binding sites for nuclear proteins.[15] The −308A element lacked the binding site because of the G/A change at −308. An inspection of the

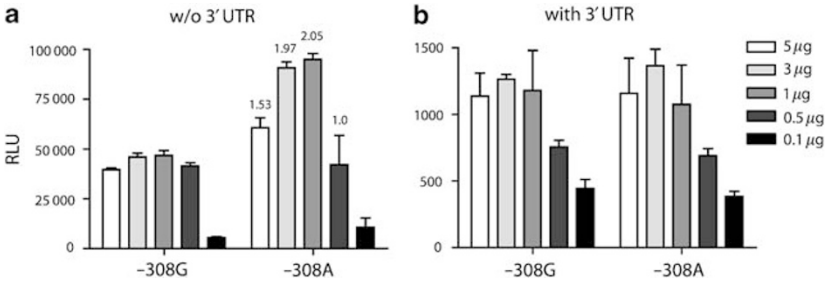


Figure 2 Titration of −308G/GL3Luc and −308A/GL3Luc DNA with or without the human *TNF* untranslated region (3′UTR). Jurkat cells were transiently transfected with varying amounts of reporter DNA (5, 3, 1, 0.5, and 0.1 μg) and harvested after 24 h to assay for luciferase activity using the Dual-Lucifrase reporter assay system (Promega). Reporter plasmids that were used include (**a**) −308G/GL3Luc and −308A/GL3Luc without 3′UTR or (**b**) −308G/GL3Luc and −308A/GL3Luc in context of the 3′UTR. Experiments were performed in triplicate. Values represent fold difference of raw luciferase relative light units (RLU) from −308A/GL3Luc over corresponding −308G/GL3Luc.

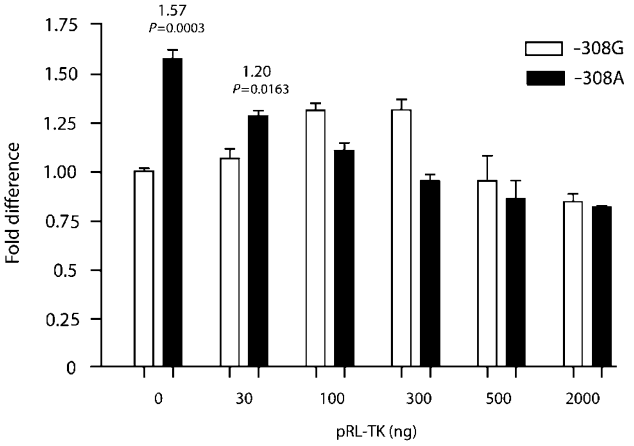


Figure 3 Effects of co-transfection with pRL-TK control plasmid. Represents the fold difference in transcriptional activity of −308A/GL3Luc and −308G/GL3Luc after normalizing RLU/s to −308G/GL3Luc only, which was set at 1. Jurkat cells were transiently co-transfected with biallelic −308 *TNF* reporter constructs along with different amounts of pRL-TK (30, 100, 300, 500 ng, and 2 μg) and harvested after 24 h to assay for luciferase activity using the Dual-Luciferase reporter assay system (Promega). The experiment was performed twice in triplicate. Values represent fold difference of raw luciferase RLU/s from −308A/GL3Luc over corresponding −308G/GL3Luc. Statistical significance was determined by Student's unpaired *t*-test. $P<0.05$ was considered significant.

HSV-TK promoter of pRL-TK established that the same nine base pair elements were also present in this promoter and may explain why the pRL-TK vector influences the −308 *TNF* polymorphism expression.

Effect of transfection method on expression of the *TNF* −308 SNP

To determine whether the method of transfection affects −308 *TNF* allele-specific regulation, two transfection methods in use within our laboratory were investigated. Figure 4a represents the results of the transfection of −308G/GL3Luc and −308A/GL3Luc in Jurkat cells using either electroporation or a lipid-mediated protocol as a means of introducing the DNA into the cells. A significantly greater activity of the −308A promoter (1.49-fold) over −308G was observed when the DNA was introduced into the cell using a lipid-formulated transfection reagent. When cells of the same culture grown to the same late-log phase of growth before transfection were used for electroporation, the difference between the polymorphic *TNF* reporter constructs was not significant (0.9-fold). The data suggested that expression of the *TNF* −308 polymorphism may be influenced by the nature or the amount of DNA entering the cell or by changing the intracellular milieu as a result of the transfection process.

Growth history of cells affects differential expression of −308G/A SNP in the *TNF* promoter

Preliminary observations suggested that cells that had been grown for a prolonged period did not show an allele-specific expression of the *TNF* −308 polymorphism. To test this formally, cells freshly revived from cryopreservation were expanded to a mid-log phase and cryopreserved in identical aliquots suitable to carry out replicate experiments. Figure 4b represents a time course. Cells were taken from cryopreservation, grown for 1 week, and then transfected using lipid-mediated transfection and assayed 24 h later. The results show a significant 2.47-fold increase in luciferase expression of the −308A reporter construct. Cells were further grown, expanded, and transfected at 2 weeks post cryopreservation recovery. In this case, the results showed that the difference in expression between the reporter constructs was reduced to 1.55-fold of *TNF* −308A over −308G. After 3 weeks in culture, the difference in activity was insignificant.

DNA conformation affects expression of the *TNF* G−308A promoter SNP

An investigation of the possible effects of DNA quality and integrity was carried out. Freshly prepared DNA was compared with DNA that had been stored frozen at −20°C in transfections. Figure 4c represents three consecutive transfections performed in Jurkat cells. Using stored DNA, a significant 2.29-fold increase in the expression of −308A over −308G was obtained. In contrast, no significance difference (0.92-fold) was seen when freshly prepared DNA was used. An inspection of the DNA by agarose gel electrophoresis indicated an increase in the proportion of open circular (nicked) plasmid DNA (>40%) on storage (data not shown). To further investigate whether the conformation of the DNA might have an effect on the results, the reporter vectors were linearized at the single *BamH* I restriction endonuclease cleavage site and used to transfect Jurkat cells by electroporation. The experiment was performed twice in duplicate. The results (Table 1) show that neither of the construct was expressed, suggesting that conformation of the DNA was important in the expression of either reporter and that the relative proportion of open circular *vs* supercoiled DNA may influence the results obtained when comparing allelic reporters.

Single copy isogenic A−308G reporter cell lines confirm differential expression of the −308 *TNF* polymorphism

To obviate the potential problems associated with transient transfection, we have recently developed novel alternative reporter gene vectors, which permit the comparison of transcriptional activities, using GFP as a reporter, after site-specific genomic integration.[24] Using the Flp recombinase-mediated integration, the system allows the integration and expression of allelic reporter gene

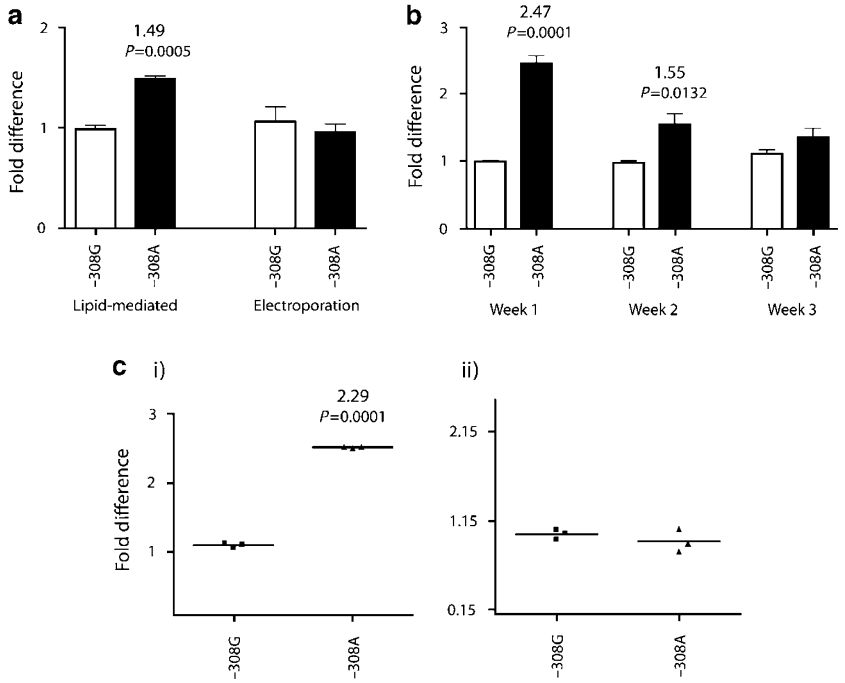


Figure 4 Differential expression is abrogated with transfection, time in culture, and age of DNA. Represents the fold difference in transcriptional activity after normalizing luciferase RLUs/ to −308G/GL3Luc, which was set at 1. (**a**) Jurkat cells were transiently transfected with either −308G/GL3Luc or −308A/GL3Luc DNA and harvested after 24 h to assay for luciferase activity using the Dual-Luciferase reporter assay system (Promega). Transfections were performed using Lipofectamine 2000 (Lipid-mediated; Invitrogen) or by electroporation as indicated. The experiment was performed thrice in triplicate. (**b**) Jurkat cells were cultured for 1, 2, or 3 weeks before transfection. The experiment was performed in triplicate. (**c**) (i) Represents three consecutive experiments performed in triplicate using DNA prepared from the QIAfilter Endofree Plasmid Maxi Kit (Qiagen) earlier and stored at −20°C. (ii) Represents the next three consecutive experiments performed in triplicate using fresh unstored DNAs prepared from the QIAfilter Endofree Plasmid Maxi Kit. The average fold differences of −308A/GL3Luc over −308G/GL3Luc from all experiments after normalization of raw luciferase RLU/s values to −308G/GL3Luc are shown. Statistical significance was determined by Student's unpaired *t*-test. $P<0.05$ was considered significant.

Table 1 Linearized −308 biallelic *TNF* reporter constructs do not express

Sample	*Relative light units*
−308G/GL3Luc (undigested) positive control	18498 ± 1396
−308A/GL3Luc (undigested) positive control	17457 ± 1230
−308G/GL3Luc (digested)	842 ± 284
−308A/GL3Luc (digested)	600 ± 64
Negative control	289 ± 14

constructs at exactly the same genomic location and orientation in all cells of any one culture.[26] The resulting reporter gene lines have a single copy of the reporter incorporated within an appropriately chromatinized region of the genome. The allelic reporters, although not a true *in vivo* system but rather an improved *in vitro* system, operate in an environment that more closely reflects that of the endogenous gene compared with an extra-chromosomal plasmid reporter.[24]

We have validated the use of this system by comparing G−308A allele-specific variation in the expression from the *TNF* promoter. Each −308 allelic promoter sequence was cloned into the GFP reporter, pFRT-GFP, and integrated into the FRT-Jurkat T-cell line.[24] Each allelic reporter cell line was analyzed by flow cytometry for GFP expression. The cell line carrying the −308G reporter showed a mean fluorescence (MF) of 18.17, in which only 39.9% of cells expressed GFP (Figure 5a). In contrast, the line carrying the −308A reporter showed a MF of 38.0 and a >two-fold increase (91.7%) in the number of GFP-expressing cells compared with −308G (Figure 5b) confirming that the transcriptional element within which the −308A SNP sits, serves to increase expression from the *TNF* promoter, not by increasing the overall level of expression of the gene but rather by increasing the number of cells that do express. This is an attribute of a classical enhancer.

Cell-cycle effects on the differential expression of the −308 *TNF* polymorphism

The isogenic stable reporter system allows more detailed investigations into the effects of regulatory polymorphisms on expression. Given the above results indicating that the cellular growth condition may influence the allele-specific

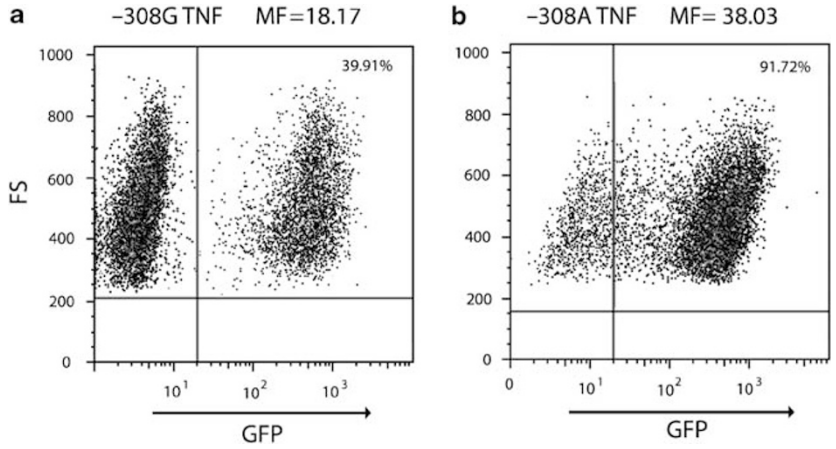


Figure 5 Comparative analysis of polymorphic −308 *TNF* promoter sequences. (**a**) Representation of both the −308G *TNF* and −308A *TNF* allelic promoter/GFP reporter constructs. (**b**) Both reporters were integrated into the FRT site of the FRT-Jurkat line and GFP expression assessed by flow cytometry. For cells carrying the −308G reporter, 34.5% of the cell population expresses GFP, whereas for the −308A reporter, 79.3% of the cell population expresses GFP. MF, mean fluorescence.

readout from the *TNF* reporters, the influence of the cell cycle on *TNF* −308 allelic expression was investigated using the stable FRT-Jurkat reporter lines. Both −308 allelic reporters were grown to mid-log phase analyzed by flow cytometry. GFP reporter-expressing cells were analyzed for DNA content by PI staining and gated as being in the G1, S, or G2M phase of the cell cycle (Figure 6). Comparison of the allelic reporters indicated that a lower proportion of GFP-expressing cells were present in the G1 phases in the −308A reporter cell population and this was reversed in the S phase. GFP-positive cells in the G2M phase showed a similar proportion for both alleles. This result indicated that expression of G−308A alleles is differentially regulated during the cell cycle and may explain why the growth history of the host cell effects expression of the polymorphism.

Discussion

A number of methods are available for assessing the effect of regulatory genetic variation on gene expression. One approach is to compare the total amount of RNA or protein in tissue from individuals of different genotypes. An alternative approach involves comparing the RNA levels of the two alleles of a given gene in heterozygous individuals. In this case, each allele acts as an internal control for the other, allowing assessment of the overall *cis*-acting influences on the expression of a gene.[27] These *in vivo* assays are very useful for examining the influence of specific haplotypes on expression of the gene under study. Haplotypes that are associated with differences in transcription also can be identified using the 'haploChIP' method, which assesses RNA polymerase loading at each allele.[28] However, for directly testing the effect of individual SNPs (rather than haplotypes) on gene expression, the only available method is the reporter gene assay. Despite being an *in vitro* assay, its popularity is because of the ease with which results can be obtained. However, as documented in a recent study, reporter gene assays do not always predict function *in vivo*.[29] Our study highlights a number of confounding factors that may affect significantly the veracity of reporter gene data.

Along with the gel shift assay, reporter gene analysis is used to establish a *prima facie* case for further, more detailed investigation of a particular SNP or indel that shows an association with a particular phenotype.[5,6] As outlined earlier, the contribution of any individual regulatory SNP to phenotypic variation is likely to be small. The transient reporter gene assay is not designed to reliably measure statistically meaningful differences of these magnitudes. Our results indicate that subtle variations in the manner by which the reporter gene assay is conducted can confound the measurement. This is exemplified by the number of conflicting reports regarding the functional attributes of the G−308A SNP of the *TNF* locus, which we have chosen to study further (for review see Bayley *et al*[30]).

Interest in the −308 *TNF* promoter polymorphism is related to its association with MHC-linked autoimmune and inflammatory diseases (see recent review Fernando *et al*[31]). Given the high degree of linkage disequilibrium within the MHC,[32,33] it remains controversial whether the −308 *TNF* polymorphism is causal. By using reporter gene constructs to compare the transcriptional activity of *TNF* −308 promoter alleles *in vitro*, many studies have sought to examine the consequences of this SNP on TNF expression independently of HLA associations. Some of these studies claim to show differences in transcriptional activity between the G−308A allelic promoters,[14–17] whereas others could not detect any differences.[18–20] However, all

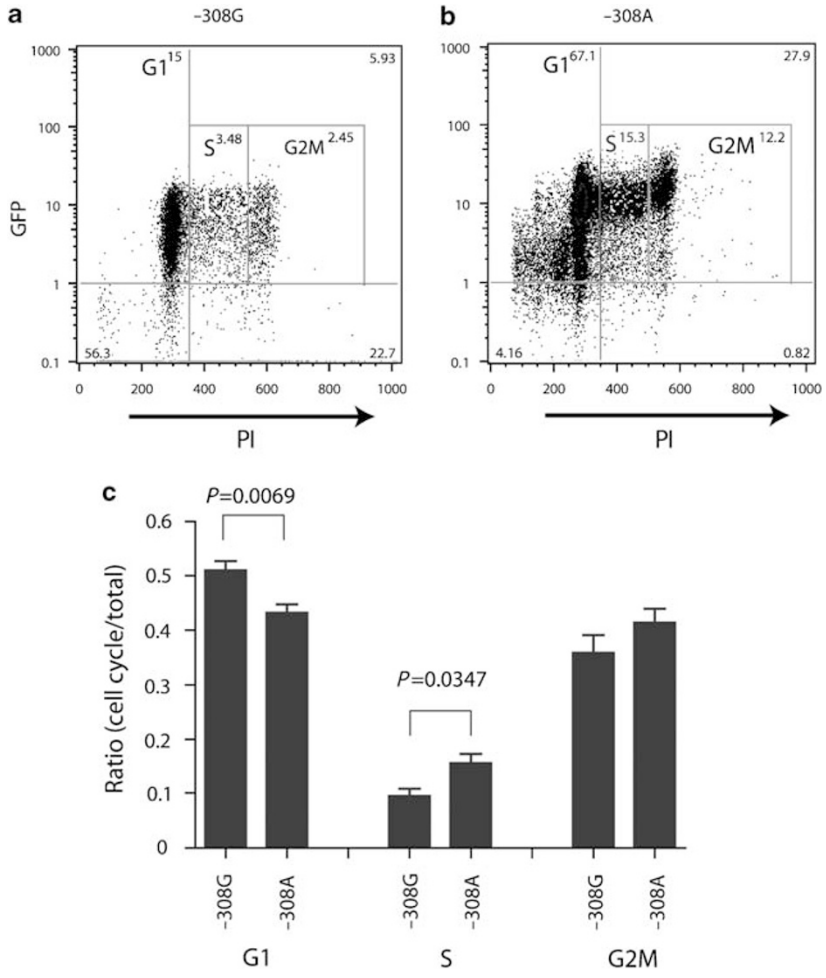


Figure 6 Cell cycle contribution to −308 allelic expression differences in the −308G and −308A FRT-Jurkat reporter cell lines (**a** and **b**). Dot plots of GFP fluorescence on *y* axis *vs* Propidium Iodide (PI) staining on the *x* axis to show sub-populations of cells in each phase of the cell cycle. G1, S, and G2M DNA contents are as labeled for each gate, with the percentage of GFP + cells in each population shown in superscript. Panel a represents cell cycle analysis of the −308G reporter cell line, whereas panel b represents the −308A reporter line. (**c**) Graphical representation of seven independent cell cycle experiments of the ratio of the mean fluorescence (percentage of GFP + cells multiplied by the mean channel fluorescence) in each cell cycle phase. Statistical significance was determined by Student's unpaired *t*-test. $P<0.05$ was considered significant.

of these studies used different reporter constructs, cells lines, and/or transfection techniques, which made any claim of *in vivo* function unsustainable. We have now established that factors, such as the amount of DNA used in transfection, the amount of transfection control vector used, the method of transfection, the growth history of the host cells, and the quality and purity of DNA used in transfections, all influence *TNF* −308 SNP-specific regulation in transient reporter gene assays and serve as a caution for those researchers who apply transient transfection assays to the functional assessment of polymorphic promoter sequences.

To obviate many of the potential problems associated with the transient reporter gene approach, including quality of reporter DNA, interference from normalization vectors, and errors associated with normalizing reporter activity data for transfection efficiency, we have developed an isogenic orientation-specific integrated reporter system that allows the assessment of promoter variation in a chromatinized setting that may more closely reflect the endogenous gene environment. This approach avoids many of the variables we have documented. Using our novel approach, we have provided compelling evidence in support of the functional relevance of the G−308A SNP of the *TNF* gene. In addition, the use of the integrated reporter system, although only an improved *in vitro* assay, has allowed us an insight into the mechanism of action of the promoter element within which the −308 SNP is

embedded, both with respect to the cell cycle and also the mode of transcriptional control, and increases our confidence in ascribing a function to the −308 *TNF* SNP.

Acknowledgements

The authors thank Dr Maria Franchina and Ms Meridith MacIntyre for their skillful technical assistance and helpful discussions. This work was supported by the National Health and Medical Research Council, Australia (Grant #139015 to LJA).

References

1 Kraft P, Cox DG: Study designs for genome-wide association studies. *Adv Genet* 2008; **60**: 465–504.
2 Wray NR, Goddard ME, Visscher PM: Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res* 2007; **17**: 1520–1528.
3 Blangero J, Goring HH, Kent Jr JW *et al*: Quantitative trait nucleotide analysis using Bayesian model selection. *Hum Biol* 2005; **77**: 541–559.
4 Goring HH, Curran JE, Johnson MP *et al*: Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet* 2007; **39**: 1208–1216.
5 Buckland PR: The importance and identification of regulatory polymorphisms and their mechanisms of action. *Biochim Biophys Acta* 2006; **1762**: 17–28.
6 Chorley BN, Wang X, Campbell MR, Pittman GS, Noureddine MA, Bell DA: Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. *Mutat Res* 2008; **659**: 147–157.
7 Knight JC: Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. *Clin Sci* 2003; **104**: 493–501.
8 Butash KA, Natarajan P, Young A, Fox DK: Reexamination of the effect of endotoxin on cell proliferation and transfection efficiency. *Biotechniques* 2000; **29**: 610–614.
9 Bergeron D, Barbeau B, Leger C, Rassart E: Experimental bias in the evaluation of the cellular transient expression in DNA co-transfection experiments. *Cell Mol Biol Res* 1995; **41**: 155–159.
10 Carey M, Smale ST: *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2000.
11 Mercola M, Goverman J, Mirell C, Calame K: Immunoglobulin heavy-chain enhancer requires one or more tissue-specific factors. *Science* 1985; **227**: 266–270.
12 Smith CL, Hager GL: Transcriptional regulation of mammalian genes in vivo. A tale of two templates. *J Biol Chem* 1997; **272**: 27493–27496.
13 Agarwal ML, Agarwal A, Taylor WR, Chernova O, Sharma Y, Stark GR: A p53-dependent S-phase checkpoint helps to protect cells from DNA damage in response to starvation for pyrimidine nucleotides. *Proc Natl Acad Sci USA* 1998; **95**: 14775–14780.
14 Braun N MU, Ernst BP, Metzner R, Bitsch A, Weber F, Rieckmann P: Gene polymorphism at position −308 of the tumor-necrosis-factor-alpha (TNF-alpha) in multiple sclerosis and it′s influence on the regulation of TNF-alpha production. *Neurosci Lett* 1996; **215**: 75–78.
15 Kroeger KM, Carville KS, Abraham LJ: The −308 tumor necrosis factor-alpha promoter polymorphism effects transcription. *Mol Immunol* 1997; **34**: 391–399.
16 Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW: Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. *Proc Natl Acad Sci USA* 1997; **94**: 3195–3199.
17 Wu WS, McClain KL: DNA polymorphisms and mutations of the tumor necrosis factor-alpha (TNF-alpha) promoter in Langerhans cell histiocytosis (LCH). *J Interferon Cytokine Res* 1997; **17**: 631–635.
18 Brinkman BM ZD, Kaijzel EL, Breedveld FC, Verweij CL: Relevance of the tumor necrosis factor alpha (TNF alpha) −308 promoter polymorphism in TNF alpha gene regulation. *J Inflamm* 1995; **46**: 32–41.
19 Stuber F, Udalova IA, Book M *et al*: −308 tumor necrosis factor (TNF) polymorphism is not associated with survival in severe sepsis and is unrelated to lipopolysaccharide inducibility of the human TNF promoter. *J Inflamm* 1995; **46**: 42–50.
20 Uglialoro AM, Turbay D, Pesavento PA *et al*: Identification of three new single nucleotide polymorphisms in the human tumor necrosis factor-alpha gene promoter. *Tissue Antigens* 1998; **52**: 359–367.
21 Bayley JP, de Rooij H, van den Elsen PJ, Huizinga TW, Verweij CL: Functional analysis of linker-scan mutants spanning the −376, −308, −244, and −238 polymorphic sites of the TNF-alpha promoter. *Cytokine* 2001; **14**: 316–323.
22 Abraham LJ, Kroeger KM: Impact of the −308 TNF promoter polymorphism on the transcriptional regulation of the TNF gene: relevance to disease. *J Leukoc Biol* 1999; **66**: 562–566.
23 Kroeger KM, Steer JH, Joyce DA, Abraham LJ: Effects of stimulus and cell type on the expression of the −308 tumour necrosis factor promoter polymorphism. *Cytokine* 2000; **12**: 110–119.
24 Karimi M, Goldie LC, Ulgiati D, Abraham LJ: Integration site-specific transcriptional reporter gene analysis using Flp recombinase targeted cell lines. *Biotechniques* 2007; **42**: 217–224.
25 Schmid I, Uittenbogaart CH, Giorgi JV: A gentle fixation and permeabilization method for combined cell surface and intracellular staining with improved precision in DNA quantification. *Cytometry* 1991; **12**: 279–285.
26 Yan C, Wang H, Aggarwal B, Boyd DD: A novel homologous recombination system to study 92 kDa type IV collagenase transcription demonstrates that the NF-kappaB motif drives the transition from a repressed to an activated state of gene expression. *FASEB J* 2004; **18**: 540–541.
27 Bray NJ, Buckland PR, Owen MJ, O'Donovan MC: Cis-acting variation in the expression of a high proportion of genes in human brain. *Hum Genet* 2003; **113**: 149–153.
28 Knight JC, Keating BJ, Rockett KA, Kwiatkowski DP: In vivo characterization of regulatory polymorphisms by allele-specific quantification of RNA polymerase loading [see comment]. *Nat Genet* 2003; **33**: 469–475.
29 Cirulli ET, Goldstein DB: In vitro assays fail to predict in vivo effects of regulatory polymorphisms. *Hum Mol Genet* 2007; **16**: 1931–1939.
30 Bayley JP, Ottenhoff TH, Verweij CL: Is there a future for TNF promoter polymorphisms? *Genes Immun* 2004; **5**: 315–329.
31 Fernando MM, Stevens CR, Walsh EC *et al*: Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet* 2008; **4**: e1000024.
32 Awdeh ZL, Raum D, Yunis EJ, Alper CA: Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. *Proc Natl Acad SciUSA* 1983; **80**: 259–263.
33 Christiansen FT, Witt CS, Ciccone E *et al*: Human natural killer (NK) alloreactivity its association with the major histocompatibility complex: ancestral haplotypes encode particular NK-defined haplotypes. *J Exp Med* 1993; **178**: 1033–1039.