

**DIABETES SUSCEPTIBILITY IN THE CANADIAN OJI-CREE
POPULATION IS MODERATED BY ABNORMAL mRNA
PROCESSING OF *HNF1A* G319S TRANSCRIPTS.**

Running title: *HNF1A* G319S and mRNA splicing

L.W.Harries PhD^{*1}, M.J. Sloman BSc^{*2}, E.A.C Sellers MD³, A.T.Hattersley
DM, FRCP¹ and S.Ellard PhD FRCPATH^{1,2}

**These authors contributed equally to this work.*

¹Institute of Biomedical and Clinical Sciences, Peninsula Medical School,
Exeter, UK

²Department of Molecular Genetics, Royal Devon and Exeter Foundation
Trust, Exeter, UK.

³Department of Paediatrics and Child Health, University of Manitoba,
Winnipeg, Canada.

Corresponding Author

L. W. Harries,
Institute of Biomedical and Clinical Sciences,
Peninsula Medical School,
Barrack Road,
Exeter,
UK

Phone: 44 (0)1392 406779

Fax: 44 (0)1392 406767

Email: L.W.Harries@exeter.ac.uk

Received for publication 19 November 2007 and accepted in revised form 2
April 2008.

Objective

The G319S *HNF1A* variant is associated with an increased risk of type 2 diabetes in the Canadian Oji-Cree population. We hypothesised that the variant site at the 3' end of exon 4 might influence splicing and characterised mRNA transcripts to investigate the mutational mechanism underlying this susceptibility to diabetes.

Research design and methods

We established lymphoblastoid cell lines from a G319S homozygote and controls. *HNF1A* transcripts were characterised in the cell lines and pancreatic tissue by sequence analysis of RT-PCR products and quantification using real-time PCR. Susceptibility to mRNA surveillance was investigated using cycloheximide.

Results

Full-length G319S mRNA accounted for 24% of mRNA transcripts in the homozygous G319S cell line. A novel isoform lacking the terminal 12 bases of exon 4 was up-regulated (55% of mRNA transcripts) compared to control cell lines (33%) and human pancreatic tissue (17%). Two abnormal transcripts present only in the G319S cell line included premature termination codons as a result of the inclusion of 7 nucleotides from intron 4 or the deletion of exon 8. Cycloheximide treatment increased the levels of both transcripts.

Conclusions

The G319S variant results in the production of two abnormal transcripts and an alteration in the relative balance of normal splicing products. This is predicted to lead to a reduction in total *HNF1A* transcript levels but residual HNF-1 α protein activity in G319S homozygotes may still reach up to 66% of normal levels. A combination of abnormal splicing and reduced activity of the G319S protein may explain the diabetes susceptibility.

The Oji-Cree are an isolated population from North Central Canada, who are amongst the most diabetes-prone sub-populations in the world; almost 40% of adults (1) and a high proportion of adolescents (2) have diabetes. Clinical heterogeneity in disease phenotype suggests that the diabetes may have both genetic and environmental components (2-4). The most common genetic susceptibility determinant is a G>A base substitution at nucleotide 955 (c.955G>A) in the HNF1 homeobox A (*HNF1A*) gene, rare mutations in which cause maturity-onset diabetes of the young (MODY) (5). This variant is found only in the Oji Cree and is predicted to result in a glycine to serine substitution at codon 319 (p.Gly319Ser; G319S). Almost 40% of diabetic subjects carry the G319S variant with an odds ratio of 1.97 for heterozygotes and 4.00 for G319S homozygotes, compared to subjects without the variant (6).

Oji-Cree subjects without the variant show the characteristic features of type diabetes with high levels of obesity, high plasma insulin and insulin resistance (2). The diabetic phenotype of G319S carriers is more consistent with a defect in insulin secretion, with an earlier onset of diabetes, less obesity and lower insulin levels than diabetic non-carriers (2). Although some of the insulin resistance features that have been reported in Oji-Cree diabetes may arise from their unique genetic background, it is more likely however that most of these features arise from the concurrent obesity in this population (2).

The G319S variant is located within the proline-rich domain II of the

transactivation domain (7), but outside the regions important for nuclear localisation or maximum transactivation capacity (8). *In-vitro* transient transactivation studies demonstrated reduced transactivation activity of the HNF-1 α protein by 54% compared to the wild type protein. No dominant negative effects or influences on protein stability or DNA binding were observed (4). The G319S missense variant results from a base substitution at the most 3' nucleotide of exon 4. This base is located within the conserved intron 4 splice donor site and we hypothesized that the c.955G>A variant might affect splicing. In this study, we determined the sequence and levels of *HNF1A* mRNA transcripts expressed in EBV-transformed lymphoblastoid cells derived from a G319S homozygote and two controls. We demonstrate that the G319S variant leads to the production of two abnormal transcripts and disrupts the normal balance of *HNF1A* expression.

RESEARCH DESIGN AND METHODS

Subjects, materials and methods.

The G319S homozygous subject was diagnosed with diabetes at 13 years of age. He was asymptomatic at presentation, with no history of polyuria, polydipsia or weight loss. He had no acanthosis nigricans, dyslipidaemia, hypertension or evidence of non-alcoholic fatty liver disease. There was a family history of diabetes affecting two maternal aunts, the maternal grandmother, the father and two paternal uncles. Informed consent was obtained from all participants. The study was approved by the Health Research Ethics Board, University of Manitoba and was carried out in accordance

with the Helsinki declaration. Commercially available human pancreatic RNA was obtained from a pool of five donors aged from 24 to 77 yrs (Clontech, Oxford, UK).

EBV transformation. Cell lines were established from the G319S homozygous subject and 2 unrelated controls by Epstein Barr virus (EBV) transformation of peripheral blood lymphocytes (ECACC, Porton Down, Salisbury, UK). Cell lines were maintained in 1X RPMI-1640 (Gibco Life Technologies, Paisley, UK), supplemented with 10% foetal calf serum (Gibco Life Technologies, Paisley, UK).

RNA extraction and reverse transcription. Total RNA was extracted from approximately 1×10^6 lymphoblastoid cells using the Perfect RNA Mini RNA kit (Eppendorf, Hamburg, Germany). Complementary DNA (cDNA) was synthesized from 4.5 μ g of the total RNA using the Thermoscript RT-PCR system (Life Technologies, Paisley, UK) using 50°C as the incubation temperature.

Characterisation of *HNF1A* transcripts. *HNF1A* transcripts were amplified in triplicate from G319S and control mRNAs as previously described (9). Individual splice products were then isolated by bandstab PCR and sequenced (10). To investigate the occurrence of variants *in cis/trans*, long range PCR was employed with primers EX4FP, EX4FS, EX9RP and EX9RS (see table 1).

Real time PCR quantification of G319S and normal transcripts. Ectopic mRNA transcripts were amplified from lymphoblastoid cells using a single tube TaqMan™ approach on the ABI Prism 7900 platform (Applied Biosystems, Warrington, UK). Primer and probe sequences are given in table 1.

Triplicate single round reactions were carried out for each cell line using 2 μ l cDNA. PCR products were detected by the use of mutation-specific probes that were identical except for the site of the mutation (see table 1) as previously described (9; 10). Final profiles were pooled from at least two separate reverse transcriptions to ensure accuracy.

Relative quantification of normal and mutant transcripts. Relative quantification of the two transcripts was carried out using the equation $2^{-\Delta\Delta C_t}$ described by Applied Biosystems, Foster City, USA (11). The ubiquitously expressed beta-2 Microglobulin ($\beta 2M$) gene was selected as an endogenous control. An assay for this gene (Assay Hs00187842) was purchased by Assays-on-Demand from Applied Biosystems (Foster City, USA).

Inhibition of NMD. In order to determine whether any observed reductions in the amount of mutant transcripts were due to NMD, cell cultures were split and treated for 4 hours with either solvent alone (1% DMSO), or cycloheximide (100 μ g/ml in DMSO) (Sigma, Poole, UK). Cells were then washed once with PBS (Gibco Life Technologies, Paisley, UK) and harvested by centrifugation.

In silico prediction of splicing patterns and mRNA folding. The software package employed for splice site prediction is available from www.fruitfly.org and the secondary structure of mRNAs was predicted using the MFOLD software (<http://www.bioinfo.rpi.edu/applications/mfold>) (12; 13).

RESULTS

Bioinformatic tools predicted a high score (0.99) for the native splice donor site of intron 4 which was reduced to 0.94 by the G319S

variant. An alternative cryptic splice site 12bp 5' of the native splice donor site was also indicated with a score of 0.94 in both normal and G319S sequences.

Analysis of cDNA from the homozygous G319S subject revealed four different sequences (Figures 1a-1d). Transcript 1 corresponds to the full length *HNF1A* product including the G319S substitution (figure 1a). Transcript 2 corresponds to use of the alternative splice donor site and the generation of an mRNA transcript lacking 12 base pairs at the 3' end of exon 4 which includes nucleotide 955 (the site of the G319S substitution), (figure 1b). These mRNAs will result in the translation of an in frame protein of either 631 or 627 amino acids, respectively. Transcript 3 shows a small insertion of 7bp of intron 4 immediately following the G319S variant (figure 1c). This results from the use of a further intron 4 splice donor site 3' of the native site. The 7bp insertion disrupts the reading frame and a premature termination codon (PTC) is introduced at codon 332 in exon 5. The fourth transcript shows complete skipping of exon 8 in addition to the 7bp insertion from intron 4 predicted to result in premature termination at codon 332 (figure 1d). Long range PCR analysis confirmed that the IVS4nt+7 insertion and the exon 8 deletion are present within the same transcript (*in cis*).

The four mRNA transcripts identified in the G319S homozygous cell line were quantified by real-time PCR (see figure 2a). The most abundant transcript was transcript 2, which comprised 55% of *HNF1A* mRNA transcripts. This transcript results from the use of a cryptic splice site

12bp proximal to the native splice donor site of intron 4. The full-length G319S transcript (transcript 1) represented 24% of total *HNF1A* expression, with transcripts 3 and 4 at approximately 15 and 6% respectively. The levels of transcripts 3 and 4 were increased 1.6 and 1.7 fold respectively by the addition of the translation blocker cycloheximide, which abolishes the nonsense-mediated decay mRNA surveillance pathway (14). Transcript 2 was also present in both normal cell lines and in human pancreatic tissue (28, 38 and 17% of total *HNF1A* expression respectively), which suggests that it is a normal splice variant and may not compromise HNF-1 α activity (figures 2b and 2c).

The c.955G>A base substitution is predicted to alter the secondary structure of the *HNF1A* pre-RNA (figures 3a and 3b). When guanine is present at nucleotide 955, the sequences flanking the splice donor site of exon 4 are held in a relatively open conformation, but when adenine is substituted, they are incorporated into a double stranded region (figures 3c and 3d). These changes in mRNA secondary structure predict a global change in conformation, which may obscure binding sites for splicing factors.

DISCUSSION

The *HNF1A* G319S variant may moderate diabetes susceptibility through the disruption of normal mRNA splicing. We detected two abnormal transcripts (transcripts 3 and 4) in the G319S cell line that result in premature termination codons. We also identified a novel isoform, transcript 2, which arises from the use of a cryptic splice donor

site 12 bases upstream of the native intron 4 splice donor site. This is the major *HNF1A* isoform (55%) in G319S lymphoblastoid cells, but was also present in normal control cell lines and human pancreas at lower levels (approximately 33% and 17% respectively). The G319S variant does not completely abolish the use of the native splice donor site, since full-length G319S transcripts were also present (24% of transcripts).

The identification of an mRNA transcript lacking exon 8 was unexpected since this exon is located 4.6 kb downstream from the c.955G>A substitution. Mutations located close to one splice site causing deletion of remote exons have been reported before (15) but they are rare. One possible explanation may be that alterations in the conformation of the nascent hnRNA (pre mRNA) transcripts may influence both splice site choice and the stability of the resultant mRNA (16-18). The inclusion of sequences from intron 4 may therefore disrupt the secondary structure of the transcript and prevent binding of proteins to additional splicing factors such as exon splicing enhancers. The predicted changes in RNA secondary structure resulting from the intron 4 insertion (see figure 3b) are consistent with this hypothesis.

This study has identified a novel *HNF1A* isoform (transcript 2) expressed in lymphoblastoid cell lines and adult human pancreas. This isoform lacks the 3' terminal four amino acid residues of exon 4 which are not located within any of the regions required for maximal transactivation capacity or nuclear localisation (8). It seems unlikely that this isoform would have reduced functional activity. Previous in vitro

functional studies of full length G319S protein showed a reduction in HNF-1 α activity by approximately 54% (4). This transcript represented 24% of total mRNA and is estimated to account for around 11% of active HNF-1 α protein in G319S homozygotes. This, together with the predicted 55% activity attributable to transcript 2 suggests that G319S homozygotes might retain up to 66% of HNF-1 α activity. The levels of active HNF-1 α protein in a G319S heterozygote are predicted to be intermediate between these and normal levels.

The Oji-Cree diabetic phenotype associated with the G319S variant is consistent with reduced insulin secretion and is likely to be a consequence of reduced HNF-1 α activity. Our results suggest that the residual level of transactivation activity is greater than that observed in *HNF1A*-MODY where the mutational mechanism of haploinsufficiency was confirmed by the identification of heterozygous gene deletions (20-22). The reduction in HNF-1 α activity may result from a combination of NMD-mediated reduction in mRNA levels due to the PTC-containing transcripts and the reduced function of the full length G319S protein. The aberrant splicing resulting from the base substitution at the 3' end of exon 4 moderates the diabetic phenotype through the upregulation of a novel isoform that is predicted to have normal transactivation activity.

We considered the possibility that other coding polymorphisms associated with type 2 diabetes might also affect splicing, but the *HNF1A* coding SNPs A98V and I27L are not located close to conserved splice sites or exonic splice

enhancers. Our study demonstrates that susceptibility to type 2 diabetes in the Oji-Cree may be mediated in part by aberrant splicing, highlighting the potential utility of mRNA analysis in the investigation of disease susceptibility alleles.

ACKNOWLEDGEMENTS

We are grateful to the Research and Development Directorate at the Royal Devon and Exeter NHS Foundation Trust for funding this

study. We are grateful to Dr Heather Dean for clinical collaboration and constructive comment on the manuscript. We thank Andrew Parrish for assistance in preparing the figures. We would also like to acknowledge the European Collection of Cell Cultures (ECACC) for providing the transformed cell lines. ATH is a Wellcome Trust Research Leave Fellow and LWH is an RCUK fellow.

REFERENCES

1. Hegele RA, Cao H, Harris SB, Hanley AJ, Zinman B: Hepatocyte nuclear factor 1 alpha G319S: A private mutation in Oji-Cree associated with diabetes. *Diabetes Care* 22:524, 1999
2. Sellers EAC, Triggs-Raine BL, Rockman-Greenberg C, Dean HJ: The prevalence of the HNF-1alpha G319S mutation in Canadian aboriginal youth with type 2 diabetes. *Diabetes Care* 25:2202-2206, 2002
3. Hegele RA, Cao H, Harris SB, Zinman B, Hanley AJ, Anderson CM: Gender, obesity, hepatic nuclear factor-1alpha G319S and the age-of-onset of type 2 diabetes in Canadian Oji-Cree. *International journal of obesity* 24:1062-1064, 2000
4. Triggs-Raine BL, Kirkpatrick RD, Kelly SL, Norquay LD, Cattini PA, Yamagata K, Hanley AJ, Zinman B, Harris SB, Barrett PH, Hegele RA: HNF-1alpha G319S, a transactivation-deficient mutant, is associated with altered dynamics of diabetes onset in an Oji-Cree community. *Proc Natl Acad Sci U S A* 99:4614-4619, 2002
5. Owen K, Hattersley A: The young patient with type 2 diabetes: a diagnostic and therapeutic challenge. *Modern Diabetes Management* 2:10-14, 2001
6. Hegele RA, Cao H, Harris SB, Hanley AJ, Zinman B: The hepatic nuclear factor-1alpha G319S variant is associated with early-onset type 2 diabetes in Canadian Oji-Cree. *Journal of Clinical Endocrinology & Metabolism* 84:1077-1082, 1999
7. Vaxillaire M, Abderrahmani A, Boutin P, Bailleul B, Froguel P, Yaniv M, Pontoglio M: Anatomy of a homeoprotein revealed by the analysis of human MODY3 mutations. *Journal of Biological Chemistry* 274:35639-35646, 1999
8. Bjorkhaug L, Bratland A, Njolstad PR, Molven A: Functional Dissection of the HNF-1alpha Transcription Factor: A Study on Nuclear Localization and Transcriptional Activation. *DNA and Cell Biology* 24:661-669, 2005
9. Bulman MP, Harries LW, Hansen T, Shepherd M, Kelly WF, Hattersley AT, Ellard S: Abnormal splicing of hepatocyte nuclear factor 1 alpha in maturity-onset diabetes of the young. *Diabetologia* 45:1463-1467, 2002
10. Harries LW, Ellard S, Jones RW, Hattersley AT, Bingham C: Abnormal splicing of hepatocyte nuclear factor-1 beta in the renal cysts and diabetes syndrome. *Diabetologia* 47:937-942, 2004
11. Applied-Biosystems: Relative quantitation of gene expression. *User bulletin #2*:11-15, 2001
12. Zuker M: Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406-3415, 2003
13. Matthews DH, Sabina J, Zuker M, Turner DH: Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *Journal of Molecular Biology* 288:911-940, 1999
14. Frischmeyer PA, Dietz HC: Nonsense-mediated mRNA decay in health and disease. *Human Molecular Genetics* 8:1893 - 1900, 1999
15. Lo Ten Foe JR, Kryut F, Zweekhorst M, Pals G, Gibson R, Mathew CG, Joenje H, Arwert F: Exon 6 skipping in the Fanconi anemia C gene associated with a nonsense/missense mutation (775C-->T) in exon 5: the first example of a nonsense mutation in one exon causing skipping of another downstream. *Human Mutation* suppl 1:S25-S27, 1998

16. Chamary JV, Hurst L: Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome Biology* 6:R75, 2005
17. Chamary JV, Parmley JL, Hurst L: Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nature Reviews Genetics* 7:98 - 108, 2006
18. Shabalina S, Oqurtsov A, Spiridonov N: A periodic pattern of mRNA secondary structure created by the genetic code. *Nucleic Acids Res* 34:2428 - 2437, 2006
19. Winckler W, Burt NP, Holmkvist J, Cervin C, de Bakker PI, Sun M, Almgren P, Tuomi T, Gaudet D, Hudson TJ, Ardlie KG, Daly MJ, Hirschhorn JN, Altshuler D, Groop L: Association of common variation in the HNF1alpha gene region with risk of type 2 diabetes. *Diabetes* 54:2336-2342, 2005
20. Thomas H, Badenberg, B., Bulman, M., Lemm, I., Lausen, J., Kind, S., Roosen, S., Ellard S., Hattersley, A. T. and Ryffel, G.: Evidence for haploinsufficiency of the HNF1[alpha] gene revealed by functional characterisation of MODY3-associated mutations. *Biological Chemistry* 383:1691 - 1700, 2002
21. Harries LW, Hattersley AT, Ellard S: Messenger RNA transcripts of the hepatocyte nuclear factor-1alpha gene containing premature termination codons are subject to nonsense-mediated decay. *Diabetes* 53:500-504, 2004
22. Ellard S, Thomas K, Edghill EL, Owens M, Ambye L, Cropper J, Little J, Strachan M, Stride A, Ersoy B, Eiberg H, Pedersen O, Shepherd MH, Hansen T, Harries LW, Hattersley AT: Partial and whole gene deletion mutations of the GCK and HNF1A genes in maturity-onset diabetes of the young. *Diabetologia* 50:2313-2317, 2007

TABLES

Table 1: Nested and real-time PCR probe and primer sequences. The sequences of nested PCR primers and oligonucleotides used in the real-time quantification are given here. The label 6-FAM refers to 6-fluorosceine and MGB refers to the minor groove binding non-fluorescent quencher. Forward primers are given by F suffix, reverse primers by R suffix and probes by P suffix.

Probe/Primer name	Sequence	Labels
EX4FP	5' TGCCTGTCTACAACCTGGTTT 3'	none
EX9RP	5' AGGGTGGTGGCCTGAGAT 3'	none
EX4FS	5' TGGCCATGGACACGTACAG 3'	none
EX9RS	5' GTGAAGCCCGGACTCACTG 3'	none
HNF1A-WT-F	5' CCCTCTCCCCCAGTAAGGT 3'	none
HNF1A-WT-R	5' AGGGTACTTCTGCAGTCTCACT 3'	none
HNF1A-WT-P	5' CACGGTGTGCGCTATG 3'	5' 6-FAM, 3'-MGB
G319S-FL-F	5' TGCCCTCTCCCCCAGTAAG 3'	none
G319S-FL-R	5' AGGGTACTTCTGCAGTCTCACT 3'	none
G319S-FL-P	5' TCCACAGTGTGCGCTATG 3'	5' 6-FAM, 3'-MGB
Ex412bpdel-F	5' CCTCCACCTGCCCTCTC 3'	none
Ex412bpdel-R	5' AGGGTACTTCTGCAGTCTCACT 3'	none
Ex412bpdel-P	5' CCCAGTGTGCGCTATG 3'	5' 6-FAM, 3'-MGB
IVS4ins7-F	5' CTCCACCTGCCCTCTCC 3'	none
IVS4ins7-R	5' TCGCAGGCTGTCCATAGC 3'	none
IVS4ins7-P	5' ACCACTTACTGTGGACCTTA 3'	5' 6-FAM, 3'-MGB
Ex8del-F	5' CCCCTTCATGGCCACCAT 3'	none
Ex8del-R	5' GTGTGAAGCCCGGACTCA 3'	none
Ex8del-P	5' CCCACGGTCTTCACC 3'	5' 6-FAM, 3'-MGB

FIGURE LEGENDS

Figure 1. Schematic of G319S variant transcripts

The mRNA sequences of a) transcript 1, b) transcript 2, c) transcript 3 and d) transcript 4 are given below. Lower case letters in italics refer to the conserved splice sites present in the introns. The position of G319S at nt955 is marked by an arrow. Cryptic splice sites are marked with stars. The junction sequences of the abnormal transcripts are provided. The sequence deriving from the 5' exon are in normal type and the sequence deriving from the 3' spliced exon are underlined and given in italics. Sequences derived from introns are in lower case. The position of the 12bp deletion is given by a hatched box and the position of the 7bp insertion is given by a black box. The position of the G319S variant on the sequence is marked in capitals where appropriate.

Figure 2. Real-time PCR quantification of G319S and wild type transcripts in lymphoblastoid cells and normal human pancreas

Real-time quantitative data for G319S and normal transcripts are given for a) normal b) homozygous G319S cells and c) normal human pancreas. The percentages given are calculated relative to β -2-microglobulin (*B2M*) expression. The X-axis refers to the transcript identity and the Y-axis refers to the percentage of total *HNF1A* expression represented by each transcript. Error bars represent the upper and lower limits of quantification based on three independent measurements. Differences in expression profile were analysed statistically by comparison of the delta Ct values (ie the crossing point of the test transcript relative to the endogenous control) for full length and 12bp deletion-bearing transcripts in lymphoblastoid cells by Mann-Whitney U analysis.

Figure 3. Predicted secondary structure of G319S hnRNAs and mRNAs

The predicted secondary structures of a) G319 pre-spliced hnRNA, b) S319 pre-spliced hnRNAs, c) G319S full length processed mRNA and d) G319S IVS4ins7 processed mRNA transcripts are given. The positions of residue 319 within the transcript are marked by stars. Note the alteration of secondary structure

apparent	in	c)	and	d).
----------	----	----	-----	-----

Figure 1

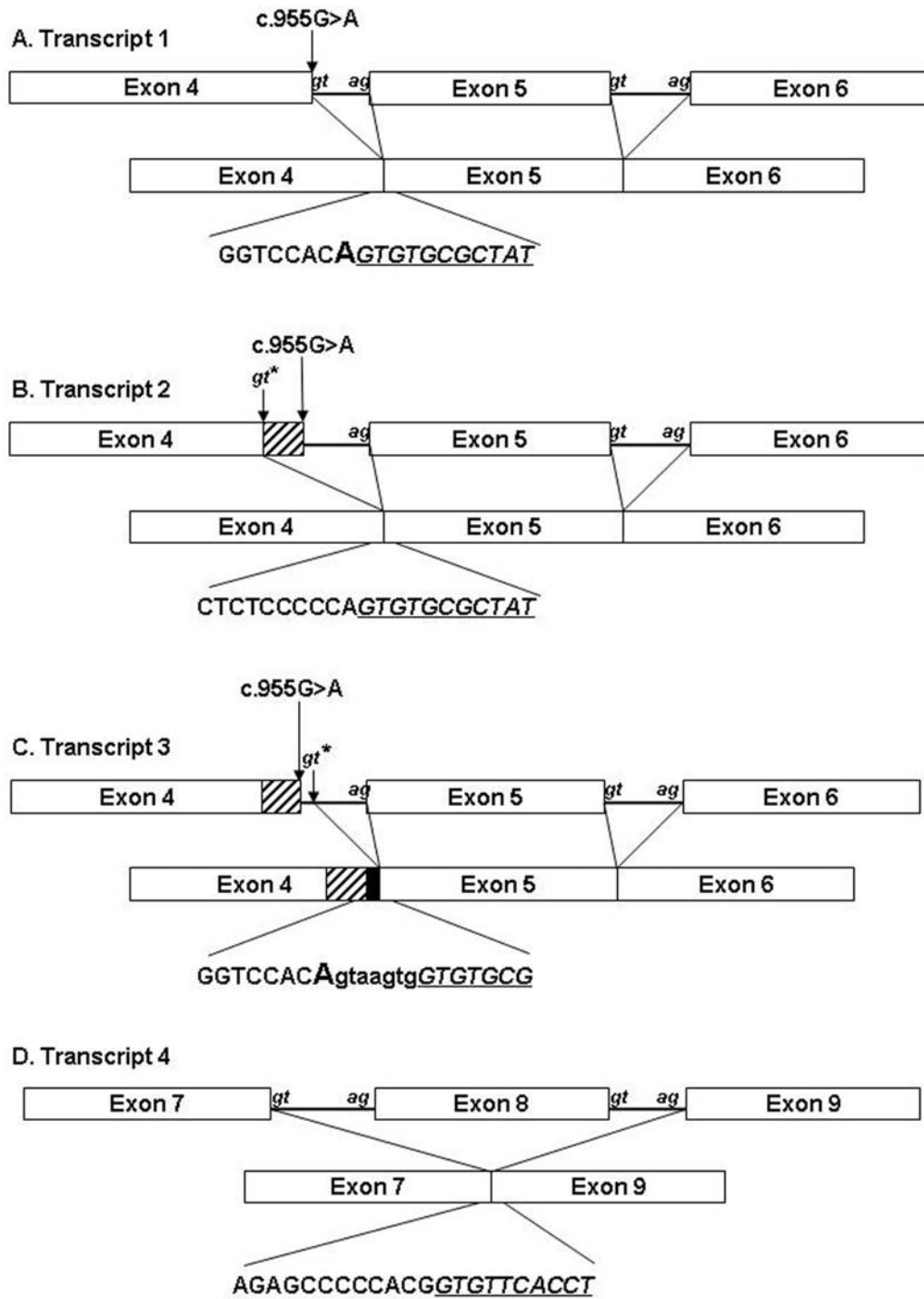


Figure 2

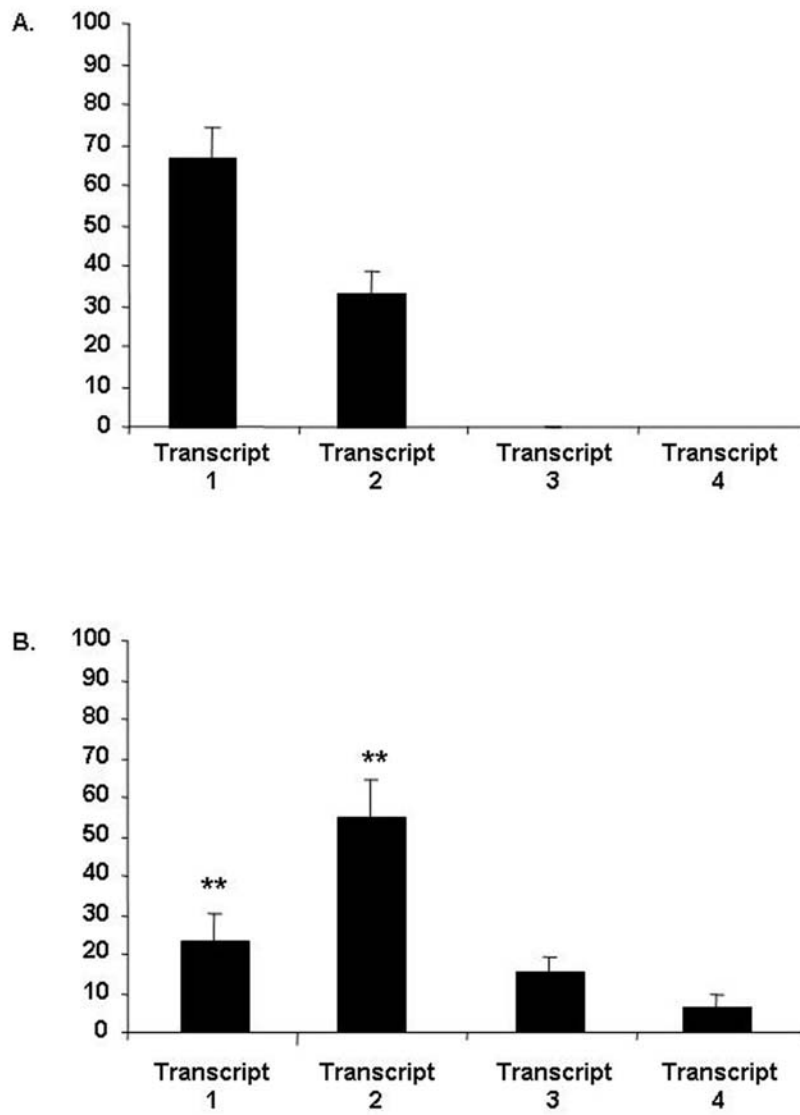


Figure 2

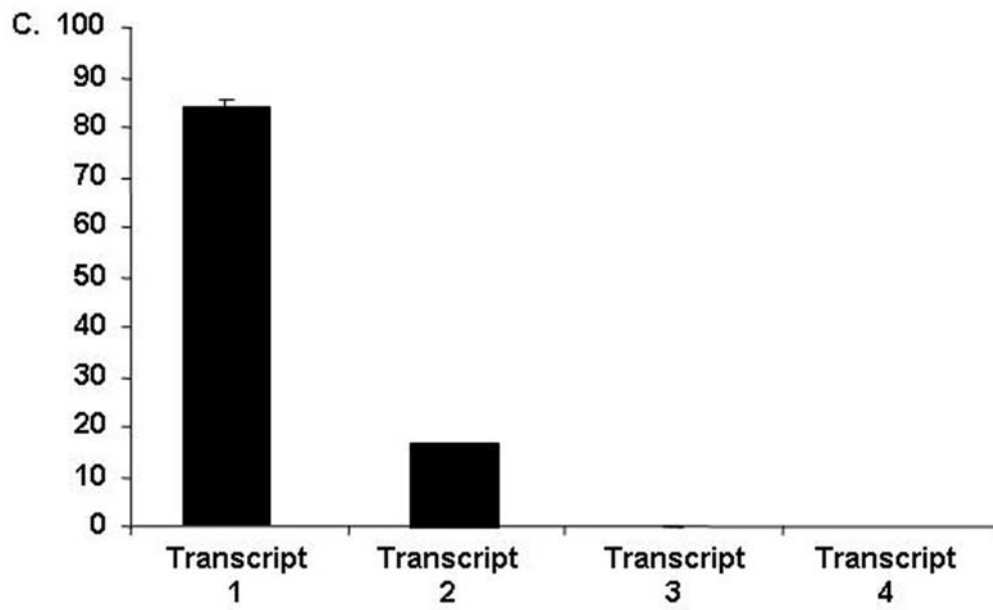


Figure 3

