A rare variant in the osteoarthritis-associated locus GDF5 is functional and reveals a site that can be manipulated to modulate GDF5 expression

Andrew W Dodd, Catherine M Syddall and John Loughlin

Osteoarthritis (OA) is a polygenic disease characterized by cartilage loss, with the single-nucleotide polymorphism (SNP) rs143383 (C/T) influencing OA susceptibility across a range of ethnic groups. The SNP resides within the 5′-UTR of the growth and differentiation factor 5 gene (GDF5), with the OA-associated T-allele mediating reduced GDF5 expression. As GDF5 codes for a cartilage anabolic protein, this reduced expression may explain why the T-allele of rs143383 is an OA risk factor. Our deep sequencing of GDF5 identified a CA transversion located −41 bp relative to the gene’s transcription start site. This promoter variant is predicted to affect transcription factor binding and it may therefore highlight a regulatory site that could be exploited to manipulate GDF5 expression and alleviate the detrimental effect mediated by the T-allele of rs143383. Here, we describe our functional assessment of the −41 bp variant. Using reporter constructs we demonstrated that the transversion leads to increased gene expression to such a degree that the A-allele is able to compensate for the reduced expression mediated by the T-allele of rs143383. Using electrophoretic mobility shift assays we identified YY1 as a trans-acting factor that differentially binds to the alleles of the −41 bp variant, with more avid binding to allele A. Knockdown of YY1 led to a significant reduction in GDF5 expression, supporting YY1 as a GDF5 activator. In conclusion, we demonstrated that the −41 bp variant is functional and we have identified a regulatory region of GDF5 that can be exploited to overcome the OA genetic deficit mediated by the T-allele of rs143383.

European Journal of Human Genetics (2013) 21, 517–521; doi:10.1038/ejhg.2012.197; published online 29 August 2012

Keywords: osteoarthritis; GDF5; association; expression; YY1; EMSA

INTRODUCTION

Growth and differentiation factor 5 (GDF5), also known as cartilage-derived morphogenetic protein 1, is a signaling molecule that acts extracellularly to promote the development, maintenance and repair of synovial joint tissues, particularly bone and cartilage.1 It achieves this by regulating the expression of a number of anabolic genes, including COL2A1, which codes for type II collagen, and ACAN, which codes for aggrecan.1 Homozygous mutations within the GDF5 gene can lead to severe skeletal defects such as those observed in Grebe type and Hunter–Thompson chondrodysplasias.2,3 Furthermore, brachydactyly type C, which is characterized by the shortening or loss of the phalanges, is due to haploinsufficiency of GDF5 and resides in patients heterozygous for the Grebe type mutation CDMPC400Y.4 These discoveries led to the hypothesis that skeletal elements can vary in their sensitivity to the loss of GDF5 function.4 This hypothesis is supported by studies of the brachypodism mutation in mice, which involves a frame-shift mutation of gdf5 that produces a truncated and inactive GDF5. Mice homozygous for the mutation present with overt skeletal defects, whereas heterozygous mice do not manifest an obvious abnormal phenotype but are predisposed to develop an osteoarthritis (OA)-like condition when challenged.5,6

Based on the known role of GDF5 in skeletogenesis, Miyamoto et al.7 selected GDF5 as a plausible candidate for harboring OA susceptibility in humans. OA is an age-related common polygenic disease characterized by the thinning and loss of the articular cartilage in synovial joints such as the hips and the knees.8 This study by Miyamoto et al.7 revealed that a single-nucleotide polymorphism (SNP) in the 5′-untranslated region (5′-UTR) of GDF5, rs143383, was associated at genome-wide significance with OA in a Japanese population. This result was quickly replicated in a number of other OA populations.9 rs143383 is a C/T transition and the OA-associated T-allele of rs143383 has been shown to produce less mRNA transcript compared with the ancestral C-allele, indicating that reduced GDF5 expression is likely to be the mechanism through which this OA susceptibility is operating.7,10 The effect that rs143383 has on GDF5 expression is modulated by a second C/T transition SNP, rs143384, which is also located within the 5′-UTR of the gene; luciferase reporter assays have revealed that the reduction in GDF5 expression occurs only when the T-allele of rs143383 is present on the same chromosome with the T-allele of rs143384.11

rs143383 is common in Asians and Europeans, with a minor allele frequency (MAF) > 20% for the C-allele. Genome-wide association scans of polygenic traits have revealed that common alleles individually contribute only modestly to trait variance and there are claims that rare variants of greater individual impact could account for some of the missing heritability.12,13 As such we were keen to assess whether GDF5 harbored rare variants that could influence OA susceptibility by...
either acting as independent risk factors or by modulating the influence of rs143383, akin to rs143384. We therefore performed a deep sequence analysis of GDF5, on >1900 European individuals, and focused our sequencing on the proximal regulatory elements of the gene (promoter, UTRs, intron/exon boundaries) and on the protein coding sequence. This revealed an absence of rare GDF5 variants, with the gene harboring either known common variants, with MAFs ≥ 2.5%, or very rare novel variants, with MAFs ≤ 0.006%. In total, the study uncovered six novel variants; five within the protein coding sequence of GDF5 and one in the proximal promoter, 41 bp upstream of the transcription start site. Clearly, the extremely low frequency of these variants means that none can have a genetic impact at the population level on OA risk. We were, however, intrigued by the −41 bp variant, since this C/A transversion is located in a conserved sequence of GDF5 and a search of databases revealed that the novel A-allele is likely to impact on the binding of trans-acting regulators of gene transcription. If the −41 bp variant does highlight a genuine regulatory domain of GDF5, then this may offer us with a DNA site that could be exploited to manipulate the expression of the gene and therefore potentially alleviate the detrimental effect that the T-allele of rs143383 has on GDF5 expression and therefore on OA risk. In this report we describe the experiments that were performed to assess whether the −41 bp variant is functional.

MATERIALS AND METHODS

Construction of GDF5-pGL3-basic luciferase reporter plasmids

We investigated the −41 bp variant in combination with rs143384 and rs143384. The physical order of the three variants is −41bp-rs143383-rs143384. To generate constructs for use in the GDF5 promoter/5′-UTR luciferase reporter gene assays, a 403-bp fragment of genomic DNA corresponding to the GDF5 proximal promoter and part of the 5′-UTR (−97 to +305 with respect to the transcriptional start site of GDF5) and encompassing the three variants was PCR amplified using gene-specific primers containing either a MluI (forward primer 5′-ACTCAAGCGGCCTG GATTCAAAACTAGGGG-3′) or a BglII (reverse primer 5′-GCAAAGATGC TAGGGCGTGAATAAGCACAAAAAG-3′) restriction site at the 5′ end of the primer. The PCR fragment was then cloned into the MultiBglII site of the purified pGL3-basic vector’s multiple cloning site (Promega, Southampton, UK). Plasmid DNA was transformed into MACH 1 bacterial cells (Invitrogen, Paisley, UK) and following growth on agar plates, clones were picked and sequenced by Sanger chemistry using capillary electrophoresis. The starting input DNA for this cloning strategy came from the individual in whom the −41 bp variant had been originally detected. This individual was heterozygous for the −41 bp variant and for rs143384 and rs143384, and carried a −41 bp (A)−rs143383 (C)−rs143384 (C) haplotype (in future known as the A-C-C haplotype) as well as a −41 bp (C)−rs143384 (T)−rs143384 (T) haplotype (in future known as the C-T-T haplotype). To generate A-T-T and C-C-C haplotypes we used the Agilent QuickChange II site-directed mutagenesis kit (Agilent, Wokingham, UK). To generate A-T-T we used the C-T-T clone and the forward primer 5′-CAGGATACGGCATTCTCCTTT GGAATA-3′ and the reverse primer 5′-TCTTCAAGGAAATTGCTTGGTT ATGTGTT-3′. To generate C-C-C we used the A-C-C clone and the forward primer 5′-CAGGATTACGGCATTCTCCTTTGGAATA-3′ and the reverse primer 5′-TCTTCAAGGAAATTGCTTGGTT ATGTGTT-3′. Purified DNA for each clone was prepared using the Maxiprep kit (Qiagen, Crawley, UK) before transfection.

Transfection of cell lines

Two human cell lines were tested; the SW1353 chondrosarcoma cell line and the MG63 osteosarcoma cell line. SW1353 cells were seeded in Dulbecco’s modified Eagle’s medium/nutrient mixture F-12 (DMEM/F-12; Gibco, Paisley, UK), containing 5% fetal bovine serum (FBS), penicillin, streptomycin (Sigma Aldrich, Gillingham, UK) and glutamine at a density of 17,500 cells per well in a 48-well cell culture plate (Costar, Gillingham, UK). MG63 cells were seeded as above but with Dulbecco’s modified Eagle’s medium (DMEM; Gibco). SW1353 cells were cultured for 48 h and MG63 cells for 24 h. The cells were then transfected using 500 ng of pGL3 plasmid DNA and 15 ng of pTK-RL Renilla plasmid in combination with Exgen 500 Transfection reagent (Fermentas, San Diego, CA, USA). Twenty-four hours after transfection the cells were lysed, using passive lysis buffer (Promega), and the protein extracted and stored at −20°C. Lysate sample was mixed with luciferase activating reagent II (Promega) and a 1-s luciferase activity reading was measured using a MicroLumat Plus LB96V luminometer (EG&G Berthold, Bad Wildbad, Germany). Stop and Glo (Promega) was added to each sample to measure renilla activity, using the same exposure time. An empty vector transfection was performed as a control. Each experiment contained six replicates, and was repeated three times, producing a total of 18 data points. A Student’s t-test was performed to assess any significant differences in luciferase activity between the different haplotypes.

Nuclear protein extraction

SW1353 and MG63 cells were plated on 500 cm² plates at a density of 20 × 10⁴ cells per plate. After 24 h the culture medium was removed and the cells were washed in ice-cold PBS. The PBS was then removed and the cells were scraped into 5 ml of fresh PBS and centrifuged for 30 s at 10000 × g at 4°C. The cells were then re-suspended in 1 ml of hypotonic buffer (Roche, Burgess Hill, UK) and incubated on ice for 15 min. The cells were then re-pelleted and the pellet re-suspended in ice-cold hypotonic buffer supplemented with 0.25% sucrose. The cells were centrifuged and the pellet was re-suspended in 1.5 ml high salt buffer (20 mM HEPES, pH 7.9, 420 mM NaCl, 20% glycerol (v/v), 1 mM DTT, 10 mM NaF, 1 mM NaVO₄ and one complete protease inhibitor cocktail tablet per 50 ml of buffer). Following a 30-min incubation on ice the cells were centrifuged and the supernatant containing nuclear protein was collected and stored at −80°C.

Electrophoretic mobility shift assay (EMSA)

Fluorescently labeled (5′ DY682) oligonucleotides (Eurofins MWG Operon, London, UK) were re-suspended in water (Sigma Aldrich) to a concentration of 100 pmol/μl. The oligonucleotides were annealed to create double-stranded probes at 95°C for 5 min in a solution containing EMSA annealing buffer (Odyssey Infrared EMSA kit, Li-Cor Biosciences, Cambridge, UK). Following this incubation, the reaction was allowed to cool to room temperature for 2–3 h. Five μg of nuclear protein extract was then incubated with 200 fmol of double-stranded probe and incubated for 20 min at room temperature and then in the dark. The samples were then electrophoresed through a native 2.3% (weight/volume) polyacrylamide gel at 100 V in the dark for ~4 h, or until the dye front had reached the gel end. Visualization was carried out using an Odyssey Infrared Imaging (Li-Cor Biosciences). When carrying out super-shift EMSAs, 2 μg of antibody (Transcruz, Santa Cruz, CA, USA) was added in place of the unlabeled competitor oligonucleotides. Supplementary Table S1 lists the nucleotide sequences of the two −41 bp probes and of the competitor oligonucleotides used.

YY1 knockdown

Cells were seeded at a density of 4000 cells per well in a 96-well cell culture plate (Costar) for PCR or 125 000 cells per well in a six-well plate (Costar) for transfection. Transfection of siRNA was then performed using ON TARGET Plus smartpool human siRNA (Dharmacon, Chicago, IL, USA). The cells were lysed using Cells to CT lysis buffer (Invitrogen) and the RNA was reverse transcribed using MMLV (Invitrogen). Gene expression was measured by quantitative real-time PCR using PrimeTime Mini qPCR Assays for YY1 and GAPDH (Integrated DNA Technologies, Coralville, IA, USA), an ABI gene expression assay for GDF5 (Applied Biosystems, Foster City, CA, USA) and an ABI PRISM 7900HT Sequence Detection System. Expression of YY1 and of GDF5 relative to GAPDH was determined using the 2−ΔΔCT method. Twelve replicates were performed. Differences in expression were assessed using a Student’s t-test. Cells from the six-well plate were harvested using lysis buffer (50 μl Tris, 10% glycerol, 50 μl NaF, 1 μl EGTA, 1 μl EDTA, 10 μl glycerol phosphate, 1% Triton X-100, 1x complete inhibitor cocktail, 1 μl).
microcin-LR and 1 mM Na3VO4). Following quantification using Bradford reagent (Expeodon, Cambridge, UK), 10 μg of protein was used for western blotting with a monoclonal β actin antibody (Sigma Aldrich) as a loading control and the same YY1 antibody as used in the EMSAs.

**YY1 expression in human articular chondrocytes**

We collected articular cartilage specimens from OA patients who had undergone elective joint replacement surgery of the hip or the knee, using informed consent and local ethical committee approval. RNA was extracted from the cartilage and cDNA was synthesized as described previously. The expression of YY1 was then assessed by quantitative real-time PCR, as described above. For protein analysis the chondrocytes were extracted from the cartilage extracellular matrix by enzymatic digestion using hyaluronidase, trypsin and collagenase. Following extraction, chondrocytes were cultured in DMEM culture medium containing 10% FBS, 2 mM l-glutamine, 200 IU/ml penicillin, 200 μg/ml streptomycin and 40 IU/ml Nystatin. These cells were then seeded at a density of 500,000 cells/well in a six-well plate. When the cells were 80–90% confluent the nuclear protein was extracted, as described above.

**RESULTS**

The −41 bp variant was discovered in an individual who was heterozygous at this site and also for the rs143383 and rs143384 SNPs. Cloning and sequencing of this person’s DNA revealed the two haplotypes to be A-C-C and the C-T-T. We therefore investigated the effect of the −41 bp variant within the context of these natural haplotypes and in the context of the alternative A-T-T and C-C-C haplotypes.

In the comparison of A-C-C with C-T-T it was apparent that A-C-C drove greater luciferase expression than C-T-T, with a 30% difference (P < 0.001) seen in MG63 and a 20% difference (P < 0.001) seen in SW1353 (Figure 1 and Supplementary Table S2). As noted earlier, it is already known that the T-allele of rs143383, when in combination with the T-allele of rs143384, results in reduced expression. As such, the greater expression seen here for the A-C-C haplotype could reflect the functionality of rs143383 and rs143384 rather than any functionality of −41 bp. To resolve this dilemma we studied the alternative haplotypes A-T-T and C-C-C. When the A-T-T haplotype was compared with the C-T-T haplotype, A-T-T drove greater expression, with a 21% difference (P < 0.001) in MG63 and a 10% difference (P = 0.044) in SW1353. Furthermore, when the C-C-C haplotype was compared with the A-C-C haplotype, C-C-C resulted in reduced expression, with a 17% difference (P = 0.0072) in MG63 and a 25% difference (P < 0.001) in SW1353. These data clearly demonstrate that the −41 bp variant is itself functional and that the A-allele discovered in our deep-sequencing experiments mediates increased gene expression.

What was particularly exciting about our luciferase data were the discovery that introducing a −41 bp A-allele on to a rs143383-rs143384 T-T haplotype neutralizes or even reverses the capacity of this T-T haplotype to mediate reduced expression. This is apparent when one compares the relative expression seen between the C-C-C and the A-T-T haplotypes in Figure 1. In MG63 the A-T-T expression is now greater than that of C-C-C (P = 0.0025). Having established that the −41 bp variant is functional in respect of gene expression we then went on to use EMSAs to characterize trans-acting factors that bind to this site. We investigated both alleles of −41 bp and nuclear extracts from MG63 and SW1353 cells. We observed a differential pattern of protein:DNA complexes on addition of labeled DNA probes corresponding to each allele, implying a difference in nuclear protein binding between the alleles (lanes C and A in Supplementary Figure S1). Based on intensity there were four major protein:DNA complexes for the A-allele, two of which were also observed for the C-allele. The four protein:DNA complexes were outcompeted with increasing concentrations of unlabeled competitor, indicated with arrows in Supplementary Figure S1. The unlabeled A-allele competitor was then used to compete binding to the C-allele and vice versa in order to assess the affinity of binding to each allele. The A-allele competitor was able to compete binding of the protein:DNA complexes using a lower amount of competitor in comparison with the C-allele competitor, implying that the protein complexes are binding more strongly to the A-allele.

The −41 bp variant resides within a sequence of GDF5 that is conserved amongst mammals (Supplementary Figure S2). Using the online prediction databases TESS and Promo 3, we identified several transcription factors whose consensus sequences indicated a potential binding with the sequence encompassing the variant and in which this binding is predicted to be influenced by the C/A substitution. We tested the ability of YY1, SOX9, abaA, SEF4, VDR and GRLF consensus sequences to out-compete the binding to the −41 bp C-allele and A-allele probes, using the competitor oligonucleotides.

![Figure 1](image-url)
We finally demonstrated expression of YY1 protein and a subsequent significant reduction in expression of GDF5 binding more avidly to the rare A-allele. YY1 assessed the effect this had on GDF5 confirming binding of YY1 protein. The supershift occurs for both the C and A alleles. We did, however, detect competition for YY1 in both MG63 and SW1353 nuclear extracts (Figure 3). The same band seen in Figure 2, which has now supershifted following binding to the YY1 antibody (upper arrows). A supershift occurs for both the C and A alleles.

We then used siRNA to knockdown YY1 in MG63 cells and assessed the effect this had on GDF5 expression (Figure 4). Our successful knockdown of YY1 (P<0.001) resulted in a reduction of YY1 protein and a subsequent significant reduction in expression of GDF5 (P = 0.0097). Similar data was obtained for SW1353 (data not shown). This result supports YY1 as an activator of GDF5 expression. We finally demonstrated expression of YY1 in OA articular chondrocytes by quantitative PCR and the presence of YY1 protein in the same cells by western blotting (Supplementary Figure S3).

**DISCUSSION**

Using a range of techniques we have demonstrated that the −41bp variant discovered by our deep-sequencing analysis of GDF5 is functional, with the novel A-allele mediating increased gene expression. One of the particularly exciting outcomes of our luciferase assay was the discovery that this A-allele is able to neutralize the reduced expression that is mediated by the OA-associated T-allele of rs143383. This result clearly demonstrates that the activity of a susceptibility allele is highly context specific and that negative effects on gene expression mediated by cis-acting regulatory sites are potentially modifiable by targeting other cis sites within the gene.

We have previously demonstrated that DNA methylation of CpG dinucleotides can modulate the expression of GDF5. However, the −41bp variant neither destroys nor creates a CpG site nor is it part of a CpG island. Furthermore, the luciferase reporter assays and EMSA experiments that we performed do not use native DNA but instead use amplified DNA and synthesized oligonucleotide primers, respectively. As such they do not recapitulate any methylation profile that may exist in the native DNA. The differences that we identified between the expression of the C and A alleles of the −41bp variant are not therefore accounted for by differential methylation of GDF5.

We identified YY1 as a trans-acting factor that differentially binds to the C and A alleles of the −41bp variant, with more avid binding to the A-allele. Knocking down YY1 led to a reduction in GDF5 expression, implying that YY1 is an activator of GDF5 expression. YY1 (Yin Yang 1) is a widely expressed zinc-finger nuclear protein that is conserved between species and which can act as a transcriptional repressor or activator, depending on cellular context. It has roles in cellular proliferation, differentiation and apoptosis, with yy1 knockout mice having in utero lethality. As far as we are aware the protein has not been implicated in OA pathogenesis. Its discovery as a functional, with the novel A-allele mediating increased gene expression, implying that YY1 is an activator of GDF5 expression now opens up the possibility that it can be exploited to modulate the expression of this important OA susceptibility locus. This may be initiated by screening compound libraries to identify proteins or other factors that can upregulate the expression of YY1 in cartilage chondrocytes in vitro. Such approaches are routine in industry and are becoming established in academia.

When we carried out our deep sequencing of GDF5 our primary goal was to assess whether this gene harbored other, novel, polymorphisms that could directly contribute to OA genetic susceptibility at a population level; by definition, such polymorphisms would need
to be common. Our experiment, however, only identified extremely rare variants, with MAFs ≤ 0.006%.14 It is becoming increasingly apparent that the human genome has an abundance of such rare, often unique, variants with their existence accounted for by the relatively recent explosion in the growth of the population.20 The −41 bp variant would appear to fit into this category and its extremely low frequency means that it cannot impact at the population level on OA risk.

Nevertheless, our functional analysis of the variant has demonstrated that despite its uniqueness it offers insight into the regulation of expression of GDF5 and has provided us with a protein, YY1, that modulates GDF5 expression and which can now be exploited to potentially alleviate the OA risk coded for by this gene. Deep sequencing therefore offers not only a means to identify novel susceptibility alleles but also a means to identify a mechanism to counteract the effects of the susceptibility alleles that are already known.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This work was funded by Arthritis Research UK and by the UK NIHR Biomedical Research Center for ageing and age-related disease award to the Newcastle upon Tyne Hospitals NHS Foundation Trust. AWD is the recipient of a Dr William Harker Foundation studentship. CMS is the recipient of an Newcastle upon Tyne Hospitals NHS Foundation Trust. A WD is the recipient of a Dr William Harker Foundation studentship. CMS is the recipient of an Newcastle upon Tyne Hospitals NHS Foundation Trust. A WD is the recipient of a Dr William Harker Foundation studentship.

1 Mikic B: Multiple effects of GDF-5 deficiency on skeletal tissues: implications for therapeutic bioengineering. Ann Biomed Eng 2004; 32: 466–476.
2 Thomas JT, Lin K, Nandedkar M, Camargo M, Cervenka J, Luyten FP: A human chondrodysplasia due to a mutation in a TGF-β superfamily member. Nat Genet 1996; 12: 315–317.
3 Thomas JT, Kilpatrick MW, Lin K et al. Disruption of human limb morphogenesis by a dominant negative mutation in CDMP1. Nat Genet 1997; 17: 58–64.
4 Polinkovsky A, Robin NH, Thomas JT et al. Mutations in CDMP1 cause autosomal dominant brachydactyly type C. Nat Genet 1997; 17: 18–19.
5 Settle SH, Routtree RB, Sinha A, Thacker A, Higgins K, Kingsley DM: Multiple joint and skeletal patterning defects caused by single and double mutations in the Gdf6 and Gdf5 genes. Dev Biol 2003; 254: 116–130.
6 Daains M, Luyten FP, Lories RJ: GDF5 deficiency in mice is associated with instability-driven joint damage, gait and subchondral bone changes. Ann Rheum Dis 2011; 70: 208–213.
7 Miyamoto Y, Mabuchi A, Shi D et al. A functional polymorphism in the 5′ UTR of GDF5 is associated with susceptibility to osteoarthritis. Nat Genet 2007; 39: 529–533.
8 Brandt KD, Dieppe P, Radin E: Etiopathogenesis of osteoarthritis. Rheum Dis Clin North Am 2008; 34: 531–559.
9 Chapman K, Takashahi A, Meulebelt I et al. A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5′ UTR of GDF5 with osteoarthritis susceptibility. Hum Mol Genet 2008; 17: 1497–1504.
10 Southam L, Rodriguez-Lopez J, Wilkins JM et al. An SNP in the 5′UUTR of GDF5 is associated with osteoarthritis susceptibility in Europeans and with in vivo differences in allelic expression in articular cartilage. Hum Mol Genet 2007; 16: 2226–2232.
11 Egli R, Southam L, Wilkins JM et al. Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. Arthritis Rheum 2009; 60: 2055–2064.
12 Lango Allen H, Estrada K, Lettre G et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 2010; 467: 832–838.
13 Bodner W, Bonilla C: Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 2008; 40: 695–701.
14 Dodd AW, Rodriguez-Fontenla C, Calaza M et al. Deep sequencing of GDF5 reveals the absence of rare variants at this important osteoarthritis susceptibility locus. Osteoarthrits Cartilage 2011; 19: 430–434.
15 Ratnayake M, Reynard LN, Raine EVA, Santibanez-Koref M, Loughlin J: Allelic expression analysis of the osteoarthritis susceptibility loci that maps to MIGAL3. BMC Med Genet 2012; 13: 12.
16 Reynard LN, Bui C, Canty-Laid EG, Young DA, Loughlin J: Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. Hum Mol Genet 2011; 20: 3450–3460.
17 He Y, Casacca-Bonnefille P, The Yin and Yang of YY1 in the nervous system. J Neurochem 2008; 106: 1493–1502.
18 Donohoe ME, Zhang X, McGinnis L, Biggers J, Li E, Shi Y: Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. Mol Cell Biol 1999; 19: 7237–7244.
19 Roy A, McDonald PR, Sittampalam S, Chaguturu R: Open access high throughput drug discovery in the public domain: a Mount Everest in the making. Curr Pharm Biotechnol 2010; 11: 764–778.
20 Tennesse JA, Bigham AW, O’Connor TD et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 2012; 337: 64–69.

This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)