

Characterisation of multiple regulatory domains spanning the major transcriptional start site of the FUS gene, a candidate gene for motor neurone disease

Khursheed, K*., Wilm, T.P*., Cashman, C., Quinn, J.P., Bubb, V.J. and Moss, D.J.

Institute of Translational Medicine, Sherrington Buildings, Ashton St, Liverpool University
Liverpool L69 3GE, UK

*joint 1st authors

Dr D J Moss

Department of Cellular & Molecular Physiology

Institute of Translational Medicine

Sherrington Buildings, Crown St

Liverpool University

Liverpool L69 3BX

Tel: 44 (0)151 794 5521

Fax: 44 (0)151 794 5517

Email: d.moss@liv.ac.uk

Running title: Regulatory domains of FUS, candidate motor neurone disease gene

Key words: evolutionary conserved domains, reporter gene assays, chick embryo model, ALS, Parkinson's disease, FTD

Abstract

Fused-In-Sarcoma (FUS) is a candidate gene for neurological disorders including motor neurone disease and Parkinson's disease in addition to various types of cancer. Recently it has been reported that over expression of FUS causes motor neurone disease in mouse models hence mutations leading to changes in gene expression may contribute to the development of neurodegenerative disease. Genome evolutionary conservation was used to predict important cis-acting DNA regulators of the FUS gene promoter that direct transcription. The putative regulators identified were analysed in reporter gene assays in cells and in chick embryos. Our analysis indicated in addition to regulatory domains 5' of the transcriptional start site an important regulatory domain resides in intron 1 of the gene itself. This intronic domain functioned both in cell lines and *in vivo* in the neural tube of the chick embryo including developing motor neurones. Our data suggests the interaction of multiple domains including intronic domains are involved in expression of FUS. A better understanding of the regulation of expression of FUS may give insight into how its stimulus inducible expression may be associated with neurological disorders.

1. Introduction

FUS (fused in sarcoma) is an RNA binding protein, associated with the spliceosomal complex, which was identified as a candidate gene for genetic association with amyotrophic lateral sclerosis (ALS) susceptibility or survival [1-2]. The number of ALS cases attributed to mutations in the FUS gene is small; FUS mutations are present but rare in Sporadic ALS at around 1% [3-6] and found in only 3-5% of Familial ALS [1-2]. However recently, further mutations were identified located in the 3'UTR of the FUS gene in some individuals with ALS (1.2 %)[7]. It has been proposed that levels of FUS protein may be critical for cell phenotype and initiation or progression of specific neurological disorders including ALS [8-10]. For example wild type FUS when over expressed in a transgenic mouse model resulted in progressive degeneration of motor neurons in a dose-dependent manner [9]. Furthermore the 3'UTR mutations described above, were associated with increased levels of Fus protein and empirical observations of the limited number of cases identified, suggested that the level of protein over-abundance was mirrored by the severity of the disease. *In vitro*, FUS knockdown leads to preferential inclusion of tau exons 3 and 10 which in turn leads to disruption of tau function. Similarly tau exon 10 inclusion is also associated with frontotemporal dementia (FTD) and Parkinson' disease. Thus tau splicing is a potential target of FUS function in neurons [11]. Taken together this suggests that mechanisms that control FUS expression, either the absolute levels of expression, the isoforms expressed or indeed post-transcriptional or translational control mechanisms may be critical for the involvement of this protein in neurological disorders. This would be consistent with a model where either mutation of the protein or differential regulation of the gene transcription or translation could lead to dysregulation of gene networks or pathway. For example, both the mutant forms of FUS or excessive amounts of normal FUS protein lead to perturbed localization of FUS in the cytoplasm rather than in the nucleus of the cell [9,12-14]. Mutant FUS is located in stress granules while wild type FUS forms globular and skein-like inclusions. In both cases the pathological accumulation of cytoplasmic FUS is thought to contribute directly to the death of the motor neurones.

Other than the recent findings describing mutations in the 3'UTR, little is known about the mechanisms regulating FUS expression. Here, we define important regulatory domains spanning the predicted major transcriptional start site using *in vitro* reporter gene assays. These domains were predicted using comparative genomics, selecting evolutionary conserved regions (ECRs) and putative promoter for the *in vitro* analysis. We next established an *in vivo* model to address their potential

function in an appropriate anatomical location, specifically in motor neurones of the chick embryo [15].

2. Results

2.1 An upstream (ECR) and intronic (INT) region are conserved during evolution

Comparative genome analysis can highlight which regions of the genome have been conserved during evolution using genome browsers such as UCSC genome browser and the ECR browser [16-17]. This analysis of course identifies the exons, but can also pinpoint regions in the non coding genome that may have a function. When this analysis was applied to the FUS gene we observed that a high degree of conservation was observed both 5' and 3' of the predicted major transcriptional start site (MTSS), located at Chr16: 31,191,429 (Figure 1A). In this analysis the shaded area indicated more than 70% conservation and the peak height corresponded to the percentage conservation over 100bp regions. The width of the peak corresponds to number of bases in the domains. This analysis identified, as might be predicted, a domain continuous with the MTSS, we termed the proximal promoter (PP). This region included 137 bp upstream of the MTSS and the untranslated part of exon 1 (106bp). A further upstream region we termed evolutionary conserved region (ECR), was also conserved. It was 311bp in length and located 29bp 5' to PP; co-ordinates were Chr 16:31,190,954 – 31,191,264. However we found that a 5' portion of the first intron continuous with exon 1, a domain we termed INT (704bp) was the most highly conserved region amongst the species analysed and could therefore contain an important regulatory domain. No other significant conserved homology was observed in a region encompassing 4kb on either side of the MTSS in the species analysed, other than the exons. These three domains were amplified by PCR and cloned into reporter gene plasmids as shown in Figure 1B & C, to test their ability to support reporter gene expression.

2.2 ECR and INT enhance expression with both FUS and SV40 promoters *in vitro*

Regulatory domains often work better with their own proximal promoter [18], therefore the initial set of constructs analysed were contiguous with the predicted MTSS of FUS and tested in the neuroblastoma cell line, SK-N-AS in which we had confirmed expression of the endogenous FUS gene (data not shown). PP (243bp) was a robust and strong promoter which drove expression of luciferase marker gene when tested in the promoter-less pGL3b vector ($p < 0.002$) (Figure 2A). Constructs that contained either the ECR or INT domain in combination with PP, both gave small but significant increases in expression, consistent with an ability to act as regulatory domains (Figure

2A). Thus ECR and INT are proposed to be activators of transcription rather than repressors of expression in the SK-N-AS cell line. In these experiments the ECR and INT domains were maintained in the same order, relative to the MTSS, as found in the genome, specifically the INT domain was maintained after the MTSS, therefore it could still modulate via post transcriptional activity. We then tested the ability of both the ECR and INT domains to act as transcriptional regulators by cloning them upstream of the heterologous SV40 promoter in the pGL3p reporter gene vector (Fig 1B). In this model both domains acted as positive regulators of expression with INT having significantly greater effect (Figure 2B).

Regulatory domains can demonstrate cell/ tissue specific and stimulus-inducible parameters when tested *in vitro* and *in vivo*. To address this, we tested the pGL3p constructs in the hybrid neuroblastoma/motor neuron cell line NSC-34 cells (Figure 2C). In this line we again generated evidence that these elements could act as positive regulators of expression furthermore the INT domain showed an increased level of activity to that observed in the SK-N-AS cell line whilst the level of activity of the ECR was the same. This analysis indicated that in the cell models tested, the INT domain had a major transcriptional regulatory activity.

2.3 ECR and INT enhance expression *in vivo* in developing motor neurons

To further investigate the function of these domains we extended our analysis to include a chick embryo *in vivo* model [15]. Electroporation of chick embryos is a rapid and inexpensive method of examining the *in vivo* behaviour of DNA regulatory domains in the appropriate tissues, in this case motor neurons. To establish the validity of analysing reporter gene expression in embryonic day 5 (E5) embryos, endogenous chick FUS expression was analysed by RT-PCR and immunohistochemistry (Figure 3). By about E5 (HH25) the motor column is clearly apparent within the neural tube and markers of developing motor neurons are expressed [19]. FUS expression was detected in the neural tube at E5 (Fig 3C) and immunohistochemistry demonstrated that at this stage of development FUS was expressed ubiquitously in cross sections of the trunk of the embryo (Fig 3A) and more specifically cell nuclei in all regions of the neural tube including the motor column were stained. The region of the developing motor column was identified by Hox gene family member, HB9 staining (S1) [20].

For *in vivo* analysis, the putative regulatory domains were cloned upstream of a hrGFP reporter gene (Figure 1C). TdTomato fluorescent protein expressed under a constitutive β actin promoter was used as a marker to identify embryos which had undergone successful injection and electroporation of the neural tube. Preliminary experiments were carried out to ensure there was no overflow of

fluorescence between the green and red channels when visualising hrGFP and Tomato proteins (data not shown). When the ECR PP construct was electroporated into the neural tube, hrGFP was detected along the length of the neural tube corresponding to the area of successful electroporation as delineated by the tdTomato marker gene activity (Figure 4A), this demonstration corroborated the finding that ECR PP had activity in SK-N-AS cells (Figure 2A). Cross sections of the neural tube confirmed that hrGFP fluorescence was seen in all regions that had been successfully electroporated including the developing motor column (Figure 4 D, E). To test whether ECR was acting as a regulatory domain, enhancing expression, the PP domain alone construct was electroporated. It was found that PP alone was not a sufficiently strong promoter to drive detectable levels of hrGFP fluorescence in this model (Figure 4 G, H). Since hrGFP fluorescence could not be detected *in vivo* driven by PP we confirmed that this plasmid was functional by transfection into SK-N-AS cells. We were able to show the PP construct drove expression of hrGFP in transfected SK-N-AS (Figure S2). These results demonstrated the ability of the ECR domain to enhance the activity of PP both *in vivo* and *in vitro* as well as enhancing the activity of the heterologous SV40 promoter in pGL3p.

INT was cloned upstream of PP and ECR PP (Figure 1C) to test whether this region could also act as a regulatory domain for transcription *in vivo*. INT PP drove hrGFP expression in the neural tube, thus confirming the ability of this region to act as a positive regulator of transcription in the *in vivo* model as well as in cell lines (Figure 5). As could be predicted, INT ECR PP also drove expression in the neural tube (Figure 5). Taken together these results show that both the ECR upstream of the FUS promoter and the evolutionary conserved region in intron 1 act as positive regulators of FUS expression in both cell lines and in the chick neural tube including the region of the developing motor neurons.

3. Discussion

The use of comparative genome analysis highlighted the potential for an important regulatory domain of FUS in the 1st intron of FUS. Conventional analysis would more likely have used a large 5' promoter fragment reporter gene construct with subsequent deletions from the 5' end of the fragment to define elements within the promoter. Transcriptional regulatory domains that can be as distant as 100kb+ from the gene they are predicted to regulate [18,21]. The comparative genome approach immediately identifies some of the important domains with no bias as to their location. The 704bp fragment of intron 1 and the 311bp fragment upstream of the PP both enhanced

transcription when linked with either the SV40 promoter or the FUS PP. In tissue culture experiments with SK-N-AS it was possible that INT might have acted as a post transcription activator however in subsequent experiments where it was clone upstream of SV40 or PP in the phrGFP plasmid INT must have been acting on transcriptional activity. It was also of interest that INT supported the greatest activity in the hybrid neuroblastoma/motor neuron cell line. Surprisingly PP alone worked only in cultured cells but not in the chick embryo neural tube suggesting either the cell environment and cell phenotype play an important role in governing the response of regulatory sequences or there was a substantial difference in the sensitivity of the two systems. Chick embryos have been used as a model for understanding the tissue-specific regulation of genes that are important during development [15] but so far this approach has not extended to genes primarily of interest due to roles in the adult and neurodegenerative diseases. Our results suggest that chick embryos will be a useful addition to the reporter gene assays currently used provided the candidate gene of interest is expressed in the differentiating cells in the embryo. Use of chick embryos synergises particularly well with the comparative genome approach where human regulatory sequences are identified based on their conservation with chick or even fish genomes. Unsurprisingly the human regulatory sequences clearly respond appropriately to the chick environment and previous work has shown that human sequences respond to a mouse environment even when the sequences are not conserved in mouse [22-23]. Most recently a primate specific retrotransposon of the SINE-VNTR-Alu (SVA) family, located upstream of the FUS gene has also been shown to drive expression in the chick embryo neural tube further demonstrating the versatility of this model [24].

Polymorphisms or mutations leading to dysregulation of gene expression may lie at the heart of understanding the causes of neurodegenerative diseases like ALS and frontotemporal lobar degeneration, rather than mutations within the coding region of proteins. Identification of the DNA sequences that regulate FUS expression in motor neurons provides an important step in our understanding of the genetic features that promote these diseases. Mutations within the cis-acting regulatory domains or 3' UTR mutations of FUS may have a profound effect on its spatial and temporal expression levels and these in turn can lead to changes in the cellular location of FUS and the formation of globular and granular cytoplasmic inclusions and eventual loss of motor neurons [7,12-14]. In keeping with this proposal SNPs occur within PP and INT although none so far are have been identified within the upstream ECR. Further analysis will be required to demonstrate whether these or other genetic changes modify the expression of the FUS protein.

Here we have used a combination of models to demonstrate robustly the activity of the cis-acting elements identified in the 5' region of the FUS gene. Crucially INT, a 704bp conserved region within

intron 1, increases expression of reporter genes in two cell lines, with two promoters (FUS and SV40) and in the neural tube *in vivo* making it a region of specific interest in view of recent results linking FUS over expression with motor neuron degeneration [9]. The approach used here will be applicable to the analysis of several genes whose over expression is implicated in the development and progression of ALS and other neurodegenerative diseases [25].

4. Materials and Methods

4.1 Comparative Genomics

Evolutionary conserved regions (ECRs) potentially containing regulatory elements were identified by genomic comparison using the ECR Browser (<http://ecrbrowser.dcode.org>; [17]) and the UCSC Genome Browser (<http://genome.ucsc.edu>; [16]). Human FUS was compared with chimpanzee, dog, rat and mouse using 70% identity and a length of 100bp as the parameters to identify conserved sequences. A region 4kB upstream and downstream was analysed and coordinates of conserved regions were obtained from Hg19 build.

4.2 Reporter Gene Construction for *in vitro* analysis

Regions of interest identified above, were extended by 100-200 bp on either side to facilitate primer design. All regions were amplified by high fidelity PCR from pooled mixed gender human genomic DNA preparations (Promega, USA) using *Pfu* DNA polymerase (Promega, USA). Restriction enzyme sites were included at the 5' end of primers to facilitate directional cloning (forward: *Nhe*I, reverse: *Bgl*II underlined below,) and the first two PCR-cycles were performed at annealing temperatures to match template-specific sequences exclusively. The following primers were designed: INT-forward 5'-GGCTAGCCCATGGCCTCAAACGGTAGGTAAGG-3', INT-reverse 5'-GAGATCTCGAAAGAAATTTAGGCGGGAAAACTCTCGGGC-3', PP-forward 5'-GGCTAGCCCATGATTCTAGTAACTTGTTCCCTTCTGCCTGCTCGGACCC-3', PP-reverse 5'-GAGATCTCGTCCGCGCACGCGCGCAC-3', ECR-forward 5'-GGCTAGCCCAGAGGGAGGGCCAGTGTGCTGCC-3', ECR-reverse 5'-GAGATCTCGAGGAAGAACCCTCTCTCCTCGGCAG-3'. PCR products INT (704bp), PP (243bp), ECR (311bp), PP INT (948bp) and ECR PP (582bp) were cloned into firefly luciferase reporter gene expressing vectors pGL3p, containing a SV40 minimal promoter element, or pGL3b, containing no promoter (both Promega, USA). Correct cloning and sequence were verified by bi-directional sequencing using standardized primers.

4.3 Cell Culture and Cell Transfection

Human neuroblastoma SK-N-AS cells (American Type Culture Collection Resource Centre stock number CRL-2137) were maintained in Dulbecco's Modified Eagle's high glucose medium (Sigma, D5672), 10% fetal bovine serum (ThermoScientific/Hyclone), 100 U/ml penicillin, 100 µg/ml streptomycin, (Sigma P0781), 1% Non-Essential Amino Acids (Sigma, M7145) and 200 mM L-glutamine (Sigma, D7513), in 5% CO₂ at 37°C. Murine motor neuron-like neuroblastoma hybrid NSC-34 cells were obtained from Professor Pamela Shaw, Department of Neuroscience, University of Sheffield and maintained in Dulbecco's Modified Eagle's high glucose medium (Sigma, D5672), 10% fetal bovine serum (ThermoScientific/Hyclone,) and 100 U/ml penicillin, 100 µg/ml, streptomycin (Sigma P0781), in 5% CO₂ at 37°C [26]. SK-N-AS and NSC-34 cells were co-transfected with 1 µg of test reporter gene construct and 20 ng of internal control DNA (pMLuc2; *renilla* luciferase reporter gene; Merck Biosciences, Nottingham, UK) using TurboFect Transfection Reagent (ThermoScientific/Fermentas, R0531) according to manufacturer's protocol in the 24-well plate format. Each transfection was performed in six wells and experiments were repeated at least three times. Similarly SK-N-AS cells were transfected with 1 µg/well of GFP reporter plasmids. Transfectant was removed after 4 hours of incubation and exchanged with fresh medium and analysed using a Leica inverted microscope 48 hours later (n=4).

4.4 Luciferase-Assays.

Luciferase activity was measured using dual luciferase assay kits (Promega). Briefly, cell cultures were lysed and the proteins stabilized using the passive lysis buffer as per the manufacturer's instructions. Luciferase activity was measured using a Glomax 96-microplate luminometer (Promega). Background from untransfected cells was subtracted from all luciferase and *renilla* values. Luciferase values were normalised to *renilla* values and expressed as mean fold increase from untreated cells. Measurements were averaged from 6-fold replicates to minimize pipetting errors and repeated at least three times to confirm results. Statistical analyses were performed using MSExcel software and a one tailed t-test to measure the significance of fold activity.

4.5 RT – PCR

Total RNA was prepared from embryonic day 5 (E5) chick neural tube and E17 cerebellum and the human-derived cell line SK-N-AS using the RNeasy mini kit (Qiagen). First strand cDNA was produced

using either Reverse Transcription (Promega) or First- Strand cDNA Synthesis (Invitrogen) kits in accordance with manufacturer's protocols. RT-PCR was performed using chicken FUS primers designed to bridge adjacent exons: forward 5'-CCAACGATTACAGCCAGACC-3' and reverse 5'-GAAACCTTGATGGGGTTGC-3' and platinum Taq DNA polymerase (Promega) following manufacturer's protocol. Amplicon size was 994bp when derived from cDNA. Endogenous FUS expression in SK-N-AS cell line was confirmed by RT-PCR from purified total RNA preparations using the following primers: (forward) 5'- AGGTGACTGTTTAGTGGGTAGGTC-3' and (reverse) 5'-ATAGCCGGACACAGTATCTCA CAC-3'.

4.6 Reporter gene construction for *in vivo* analysis in chick embryo model.

Tomato reporter gene sequence was PCR-amplified from pG-TdTomato (a kind gift from Marco Marcello, University of Liverpool) using primers forward 5'-ATAGGAATTCGGTGTACGGTGGGAGGTCTA-3' and reverse 5'-GGCCGTCGACATCATTTTACGTTTCTCGTTC-3'. These introduced *Eco RI* and *Sal I* restriction sites as indicated, for directional cloning into the plasmid pIRES EGFP (kind gift from John Gilthorpe King's College London). The pIRES EGFP cassette was removed from the vector using *EcoRI* and *XhoI* restriction sites and replaced by the Tomato reporter gene, such that the Tomato reporter was located downstream of the chick beta actin promoter.

For the *in vivo* analysis of FUS promoter in chick embryos, human FUS gene promoter fragments ECR PP (582bp) and PP (243bp) were cut out from their pGL3b plasmids using *SacI* and *BglII* restriction enzymes. These fragments were cloned into the *BamHI/XbaI* sites of the multicloning site of promoter-less vector phrGFP (Stratagene, UK) upstream of the GFP reporter gene. The clones were named ECR PP GFP and PP GFP.

In order to clone the putative regulatory element from FUS intron 1 into a location upstream of the FUS promoter sequence in ECR PP GFP and PP GFP a 704bp fragment was PCR-amplified from human genomic DNA (Promega) using proofreading Phusion DNA polymerase (NEB, UK) and primers 5'-GATGAGATCTATGGCCTCAAACGGTAGGTAAGG and 5'-AGGTGCTAGCGAAAGAAATTTAGGCGGGAAAACTCTCGGGC introducing *NheI* and *BglII* restriction sites for directional cloning. The resulting constructs were named INT ECR PP GFP and INT PP GFP, respectively.

4.7 Electroporation of chick embryos

Fertile chicken eggs were incubated at 37.8 °C for about 60-65 hours until they were approximately developmental stage 14 HH. 2-3 ml of albumen was removed and a window was cut in the egg.

Embryos were staged according to Hamburger and Hamilton (1951). In those at stage 11-14 the vitelline membrane was removed to aid manipulation of the embryo. The lumen of the neural tube was injected with a solution containing 2-5 $\mu\text{g}/\mu\text{l}$ of test DNA reporter plasmid, 1 $\mu\text{g}/\mu\text{l}$ of TdTomato plasmid (control for successful injection) in PBS and 1mM MgCl_2 containing 0.2% fast green to help visualization. Injections were undertaken with a micropipette pulled from borosilicate capillaries (Warner Instruments). Post-injection, DNA was immediately electroporated into the cells of the neural tube; gold plated electrodes of 3mm length (Harvard Apparatus) were placed either side of the embryo with an internal gap of 5mm and five 50ms square wave pulses with 100ms gaps were delivered [27]. Electroporated embryos were incubated at 37.8°C for at least 48 hours until they were approximately HH stage 25 or E5 and then assessed for expression of hrGFP and TdTomato. Electroporated embryos were dissected free of membranes and yolk and photographed using epifluorescent microscopy. Red fluorescence was photographed on the same settings (previously determined to be appropriate) to determine the consistency of electroporation. Green fluorescence was photographed on consistent settings and where necessary on longer exposures such that faint fluorescence could be documented. The GFP fluorescence was recorded as either present or absent without further quantification.

4.8 Immunofluorescence

Dissected embryos were washed with PBS and fixed using 4% paraformaldehyde (w/v) in 0.12M phosphate buffer for 1 h. The trunks of the embryos were dissected out and placed sequentially in 6%, 12 % and 18% (w/v) sucrose solution until the tissues sank. Embryonic tissues were then mounted in Cryo-M-Bed embedding compound (VWR International Ltd.), frozen in isopentane held over liquid nitrogen and 10 μm cryostat sections cut [28].

For immunofluorescence, the following primary antibodies were used: Rabbit anti Fus (1:500, Abcam ab23439), Mouse Hb9, (1:5, Developmental Studies Hybridoma Bank, Department of Biology, University of Iowa). Secondary antibodies were goat anti mouse alexa488 1:250 and goat anti rabbit alexa488 1:250 (Invitrogen) [28]. Cell nuclei were detected with DAPI.

Acknowledgements

This work was supported by the Motor Neurone Disease Association

- [1] Kwiatkowski, T.J., Jr. et al. (2009). Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323, 1205-8.
- [2] Vance, C. et al. (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323, 1208-11.
- [3] Chio, A. et al. (2011). A de novo missense mutation of the FUS gene in a "true" sporadic ALS case. *Neurobiol Aging* 32, 553 e23-6.
- [4] Corrado, L. et al. (2010). Mutations of FUS gene in sporadic amyotrophic lateral sclerosis. *J Med Genet* 47, 190-4.
- [5] Lai, S.L. et al. (2011). FUS mutations in sporadic amyotrophic lateral sclerosis. *Neurobiol Aging* 32, 550 e1-4.
- [6] Sproviero, W. et al. (2012). FUS mutations in sporadic amyotrophic lateral sclerosis: clinical and genetic analysis. *Neurobiol Aging* 33, 837 e1-5.
- [7] Sabatelli, M. et al. (2013). Mutations in the 3' untranslated region of FUS causing FUS overexpression are associated with amyotrophic lateral sclerosis. *Hum Mol Genet*
- [8] Mackenzie, I.R., Rademakers, R. and Neumann, M. (2010). TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol* 9, 995-1007.
- [9] Mitchell, J.C. et al. (2013). Overexpression of human wild-type FUS causes progressive motor neuron degeneration in an age- and dose-dependent fashion. *Acta Neuropathol* 125, 273-88.
- [10] Neumann, M., Rademakers, R., Roeber, S., Baker, M., Kretzschmar, H.A. and Mackenzie, I.R. (2009). A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain* 132, 2922-31.
- [11] Orozco, D., Tahirovic, S., Rentzsch, K., Schwenk, B.M., Haass, C. and Edbauer, D. (2012). Loss of fused in sarcoma (FUS) promotes pathological Tau splicing. *EMBO Rep* 13, 759-64.
- [12] Gal, J., Zhang, J., Kwinter, D.M., Zhai, J., Jia, H., Jia, J. and Zhu, H. (2011). Nuclear localization sequence of FUS and induction of stress granules by ALS mutants. *Neurobiol Aging* 32, 2323 e27-40.
- [13] Ito, D., Seki, M., Tsunoda, Y., Uchiyama, H. and Suzuki, N. (2011). Nuclear transport impairment of amyotrophic lateral sclerosis-linked mutations in FUS/TLS. *Ann Neurol* 69, 152-62.
- [14] Kryndushkin, D., Wickner, R.B. and Shewmaker, F. (2011). FUS/TLS forms cytoplasmic aggregates, inhibits cell growth and interacts with TDP-43 in a yeast model of amyotrophic lateral sclerosis. *Protein Cell* 2, 223-36.
- [15] Uchikawa, M. (2008). Enhancer analysis by chicken embryo electroporation with aid of genome comparison. *Development Growth & Differentiation* 50, 467-74.
- [16] Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res* 12, 996-1006.
- [17] Ovcharenko, I., Nobrega, M.A., Loots, G.G. and Stubbs, L. (2004). ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res* 32, W280-6.
- [18] Shanley, L., Davidson, S., Lear, M., Thotakura, A.K., McEwan, I.J., Ross, R.A. and MacKenzie, A. (2010). Long-range regulatory synergy is required to allow control of the TAC1 locus by MEK/ERK signalling in sensory neurones. *Neurosignals* 18, 173-85.
- [19] Ericson, J., Thor, S., Edlund, T., Jessell, T.M. and Yamada, T. (1992). Early stages of motor neuron differentiation revealed by expression of homeobox gene *Islet-1*. *Science* 256, 1555-60.
- [20] Nakano, T., Windrem, M., Zappavigna, V. and Goldman, S.A. (2005). Identification of a conserved 125 base-pair Hb9 enhancer that specifies gene expression to spinal motor neurons. *Dev Biol* 283, 474-85.
- [21] MacKenzie, A. and Quinn, J.P. (2004). Post-genomic approaches to exploring neuropeptide gene mis-expression in disease. *Neuropeptides* 38, 1-15.

- [22] MacKenzie, A., Payne, C., Boyle, S., Clarke, A.R. and Quinn, J.P. (2000). The human preprotachykinin-A gene promoter has been highly conserved and can drive human-like marker gene expression in the adult mouse CNS. *Mol Cell Neurosci* 16, 620-30.
- [23] MacKenzie, A. and Quinn, J. (1999). A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. *Proc Natl Acad Sci U S A* 96, 15251-5.
- [24] Savage, A.L. et al. (2014). An evaluation of a SVA retrotransposon in the FUS promoter as a transcriptional regulator and its association to ALS. *PLoS One* 9, e90833.
- [25] Yamada, M., Iwatsubo, T., Mizuno, Y. and Mochizuki, H. (2004). Overexpression of alpha-synuclein in rat substantia nigra results in loss of dopaminergic neurons, phosphorylation of alpha-synuclein and activation of caspase-9: resemblance to pathogenetic changes in Parkinson's disease. *J Neurochem* 91, 451-61.
- [26] Gomes, C., Palma, A.S., Almeida, R., Regalla, M., McCluskey, L.F., Trojanowski, J.Q. and Costa, J. (2008). Establishment of a cell model of ALS disease: Golgi apparatus disruption occurs independently from apoptosis. *Biotechnol Lett* 30, 603-10.
- [27] Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y. and Kondoh, H. (2003). Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell* 4, 509-19.
- [28] Carter, R., Mullassery, D., See, V., Theocharatos, S., Pizer, B., Losty, P.D., Jesudason, E. and Moss, D.J. (2012). Exploitation of chick embryo environments to reprogram MYCN-amplified neuroblastoma cells to a benign phenotype, lacking detectable MYCN expression. *Oncogenesis* 1, e24.
- [29] Ovcharenko, I., Stubbs, L. and Loots, G.G. (2004). Interpreting mammalian evolution using Fugu genome comparisons. *Genomics* 84, 890-5.

Figure Legends

Figure 1 *Schematic representation of FUS gene locus organization and reporter gene construct design.* (A) Human FUS is located in p11.2 on chromosome 16. A human region upstream of the prospective promoter to exon 3 is expanded and aligned with corresponding FUS genomic regions of chimp, dog, mouse and rat highlighting evolutionary conservation in coding and non-coding sequences of the FUS locus (ECR browser; [29]). Bent arrow indicates transcriptional start (MTSS) and ATG indicates the first translated codon. Three regions, non-coding ECR (classical evolutionary conserved region), PP (proximal promoter) and INT (intron 1) were cloned. (B) For SK-N-AS and NSC-34 cell culture experiments regions of interest were cloned into either pGL3-Promoter containing minimal SV40 promoter (mP) or pGL3-Basic (without minimal SV40 promoter) vectors using firefly luciferase (Luc) as the reporter gene. (C) For *in vivo* studies in chick embryos, regions of interest were cloned into vector phrGFP using GFP as the reporter gene.

Figure 2 *Dual luciferase assays of FUS reporter gene constructs in SK-N-AS and NSC-34 cells.* (A) FUS promoter constructs cloned into promoter-less pGL3b expressed in SK-N-AS cells. All three constructs were significantly different when compared with pGL3b ($p < 0.002$ PP and PP-INT, $p < 0.002$ ECR-PP). ECR-PP and PP-INT had significantly higher transcriptional activity ($p < 0.04$, $p < 0.03$ respectively) than PP alone. (B, C) Upstream (ECR) and intronic (INT) FUS putative regulatory

domains cloned into minimal SV40 promoter-controlled pGL3p and expressed in SK-N-AS cells (B) and NSC-34 (C). Both ECR and INT increased luciferase expression compared to the SV40 promoter in SK-N-AS cells ($p < 0.02$, $p < 0.0002$) and NSC-34 cells ($p < 0.02$, $p < 0.004$). Bars represent the fold change in firefly luciferase activity normalized against *renilla* luciferase activity and mean calculated from n experiments. Transfection experiments were performed in double triplicates.

Figure 3 *FUS is expressed in the neural tube of E5 chick embryos*

Frozen section of E5 chick embryo stained with (A). FUS antisera, (B). DAPI. FUS is ubiquitously expressed in all cells in the section including cells in the developing motor column. Scale bar 100 μ m (C). RT-PCR of E5 neural tube and E17 cerebellum. The arrow indicates 994bp the expected size of the FUS PCR product.

Figure 4 *ECR PP enhances expression compared to PP alone*

(A,B) and (C) show embryos that have been electroporated with 1 μ g/ μ l dTomato, 2.2 μ g/ μ l ECR PP GFP. (A) dTomato fluorescence indicating successful electroporation. (B) GFP fluorescence indicating the activity of ECR PP in the neural tube. (C) Bright field showing the location of the four successfully electroporated chick embryos. The arrow indicates the electroporated trunk region of one of the embryos. Scale bar 2mm. (D, E and F) are frozen sections of one of the embryos from (A, B and C). (D) shows the extent of successful electroporation in the neural tube. It is restricted to the left hand side as the DNA within the neural tube is directed to the cathode. (E) shows that ECR PP drives GFP expression in all the electroporated cells including the developing motor neurons. Scale bar 100 μ m (G, H and I) show embryos that have been electroporated 1 μ g/ μ l dTomato, 2.0 μ g/ μ l PP GFP. (G) is TdTomato, (H) PP GFP and (I) bright field. Scale bar 2mm

Figure 5 *INT and INT ECR enhance expression compared to PP alone*

(A, B and C) show embryos that have been electroporated with 1 μ g/ μ l dTomato, 3 μ g/ μ l INT PP GFP. (A) dTomato fluorescence indicating successful electroporation. (B) GFP fluorescence indicating the activity of ECR PP in the neural tube. Scale bar 1mm (C) Bright field. (D, E and F) show embryos that have been electroporated with 1 μ g/ μ l TdTomato, 3 μ g/ μ l INT ECR PP. D. dTomato fluorescence. E. GFP fluorescence indicating the activity of INT ECR PP in the neural tube. F. Bright field. Scale bar 500 μ m

Supplementary figure 1 *Hb9 staining indicates the motor column in E5 neural tube*

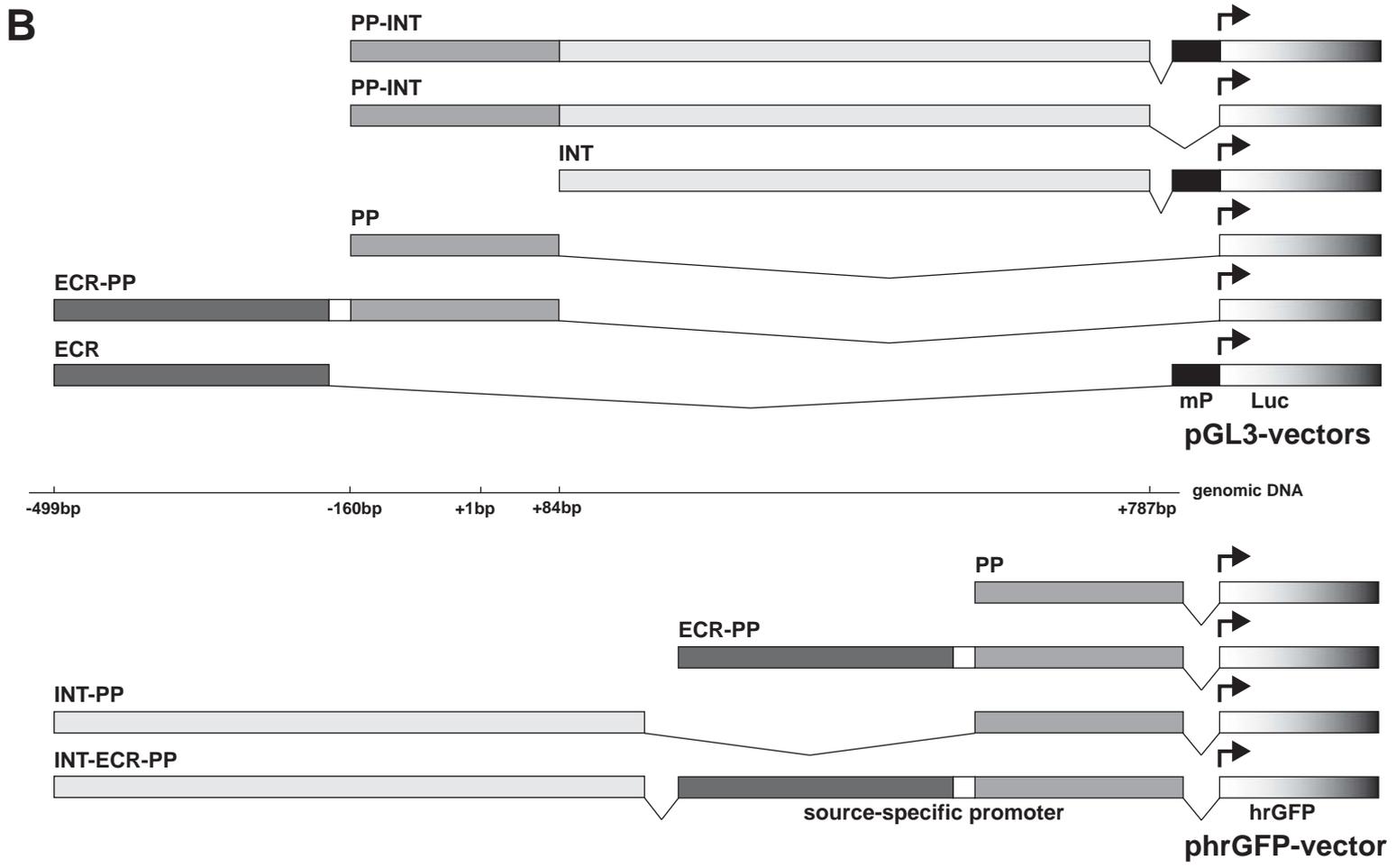
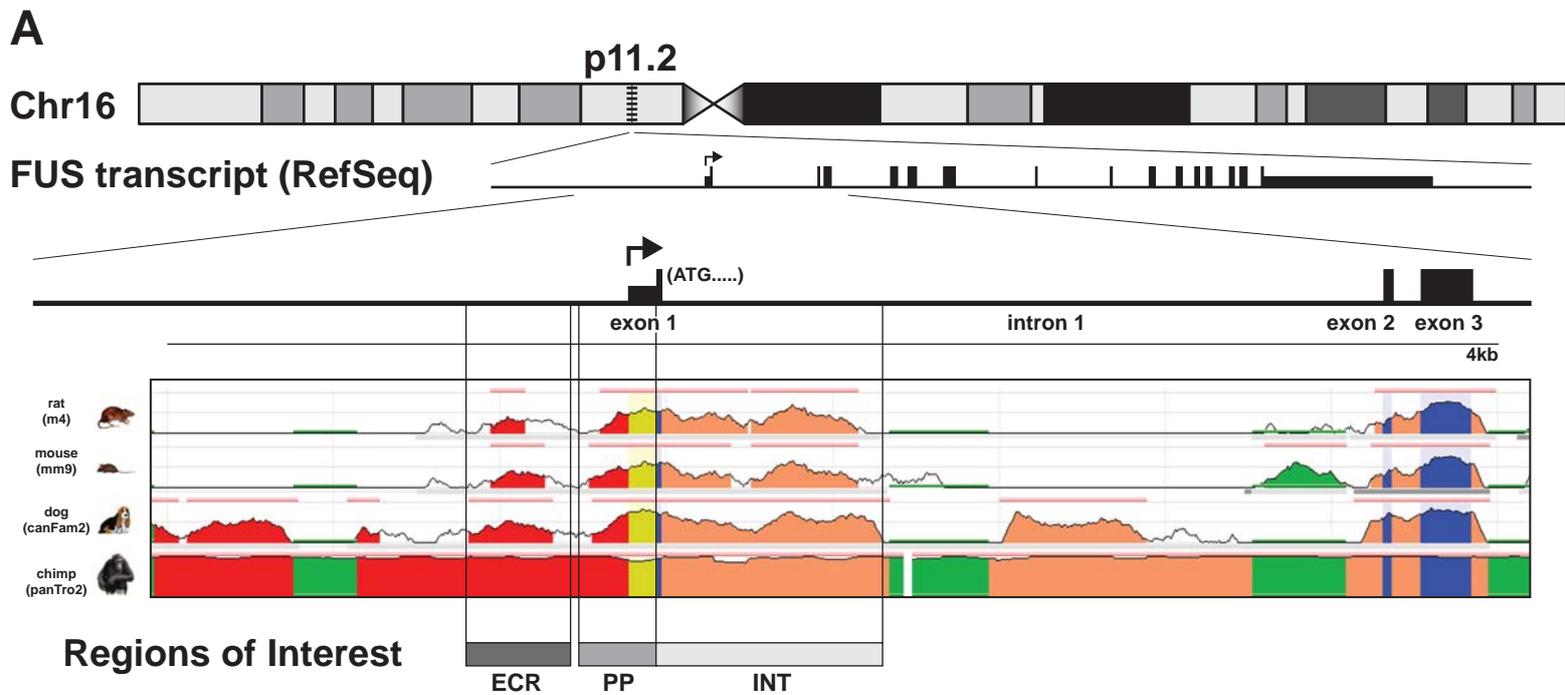
Frozen section of E5 neural tube was stained with anti Hb9. Hb9 is expressed specifically in the nuclei of motor neurons. Scale bar 100µm

Supplementary figure 2 *PP drives GFP expression in SKNAS cells*

SKNAS cells were transfected with (A, B) TdTomato, (C, D) ECR PP GFP, (E, F) PP GFP and (G, H) phr GFP. (A, C, E and G) show transfected cells and (B, D, F and H) show phase contrast micrographs of SKNAS cells. PP GFP was successfully transfected SKNAS cells along with TdTomato and ECR PP GFP.

Contributions

KK and TW carried out all the experiments; CC was responsible for some of the cloning and preparation of key reagents, JQ initiated the project and supervised the *in vitro* studies, VB supervised cloning and analysis of some of the experiments, DM initiated and supervised the *in vivo* model and JQ, VB and DM drafted the manuscript.



Figure

[Click here to download Figure: Figure 2.pdf](#)

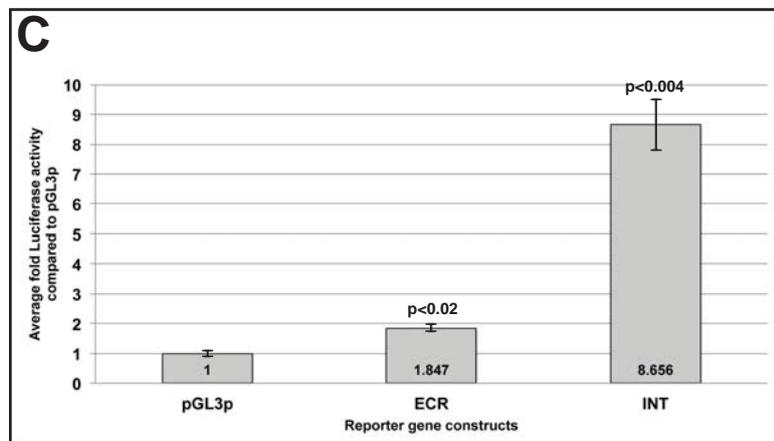
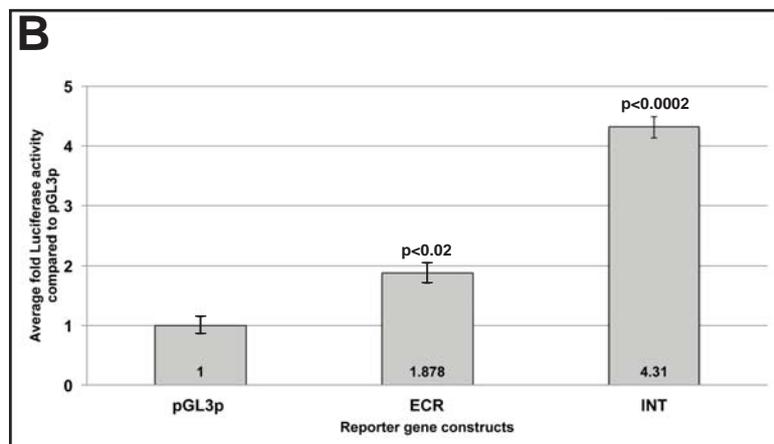
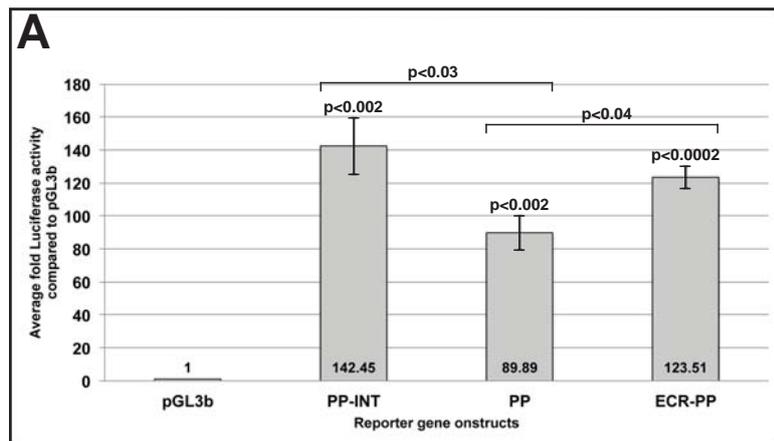


Figure
[Click here to download high resolution image](#)

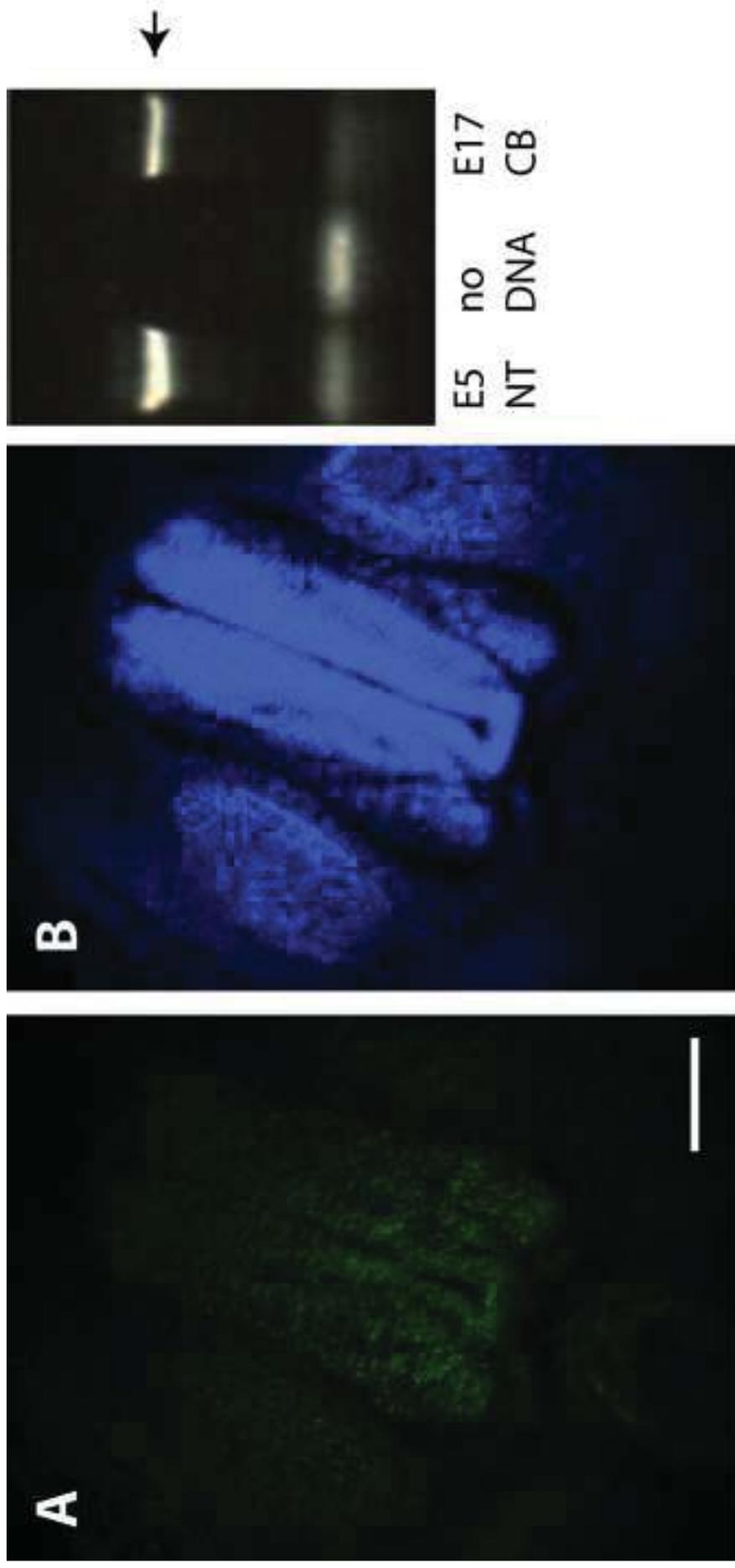


Figure
[Click here to download high resolution image](#)

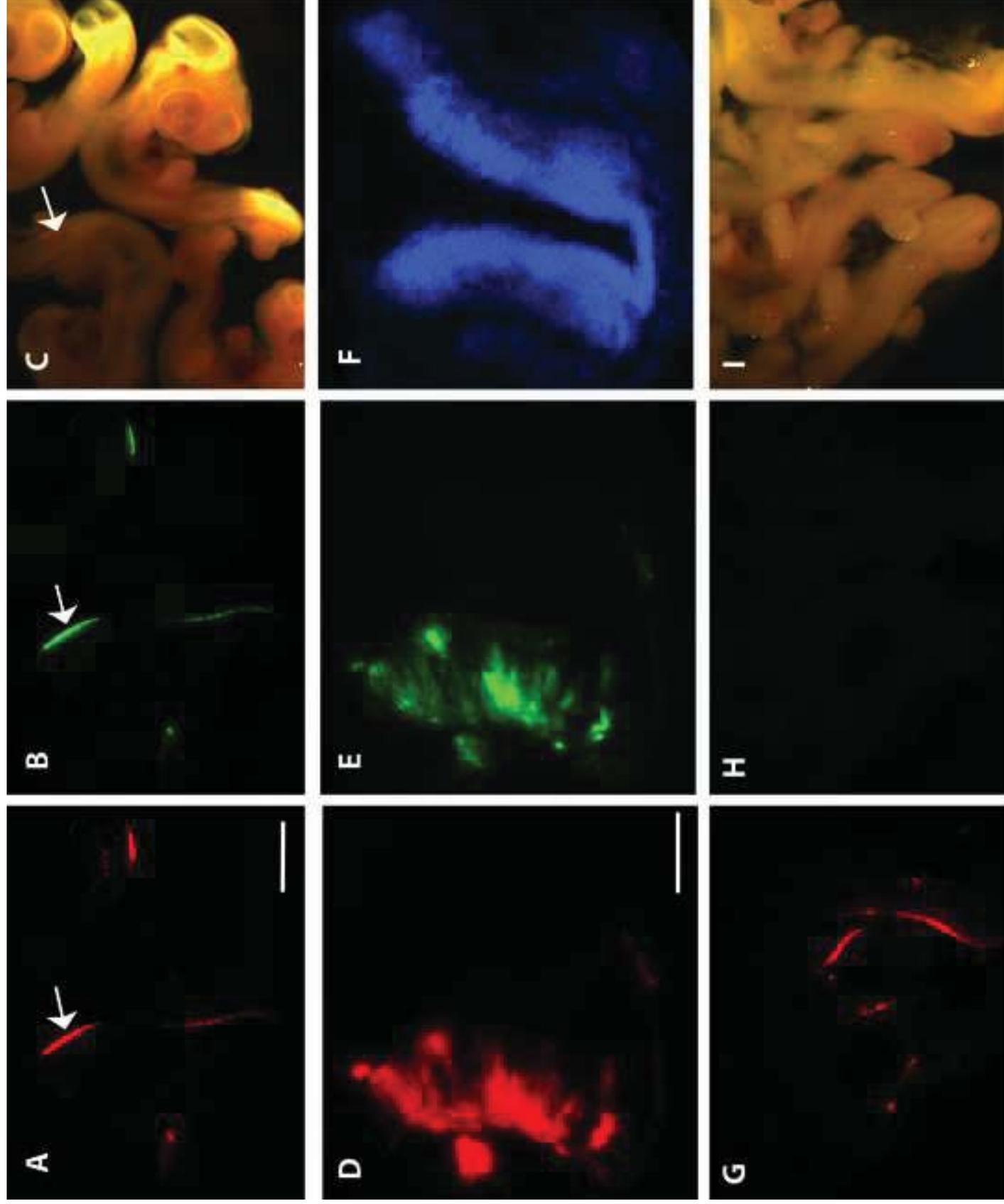


Figure
[Click here to download high resolution image](#)

