Highlights

- Identification of a new member of the Rab-guanosine disassociation inhibitor family
- Only found in fish and anurans
- This gene is developmentally regulated and necessary for normal development

A newly identified *Rab-gdi* paralogue has a role in neural development in amphibia.

Liliya Nazlamova, Anna Noble, Frank R. Schubert, John McGeehan, Fiona Myers, Matt Guille, Garry Scarlett[1]

[1]Corresponding author. Biophysics Laboratories, School of Biological Sciences, Institute of Biomedical and Biomolecular Science, University of Portsmouth, St Michaels Building, White Swan Road, Portsmouth PO1 2DT, UK.

Abstract

Vesicle shuttling is critical for many cellular and organismal processes, including embryonic development. Gdi proteins contribute to vesicle shuttling by regulating the activity of Rab GTPases, controlling their cycling between the inactive cytosol and active membrane bound states. Whilst identifying genes controlled by A-form DNA sequences we discovered a previously unknown member of the *Gdi* family, *gdi3*. The *gdi3* gene is found only in amphibians and fish and is developmentally expressed in *Xenopus* from neurula stages onwards in the neural plate, and subsequently in both dorsal and anterior structures. Depletion or over-expression of the Gdi3 protein in *Xenopus* embryos gives rise to very similar phenotypes, suggesting that strict control of Gdi3 protein levels is required for correct embryonic development. Our analysis suggests the evolutionary origins of *gdi3* and that it is functionally distinct from *gdi1*. Predicted structural analysis of Gdi3 suggests that the key difference between Gdi1 and Gdi3 lies in their lipid binding pockets.

Key words

gdi1, gdi2, gdi3, *Xenopus,* vesicle trafficking, Rab

1. Introduction

The molecules that regulate vesicle shuttling are important for both cellular survival and correct embryological development. Central to vesicle shuttling are the Rab proteins, which are critical for the correct regulation of the movement of vesicles between the components of the endomembrane system such as the Golgi body, endoplasmic reticulum and the cell membrane. Rab proteins are small GTPases belonging to the Ras superfamily, with more than 60 distinct Rab proteins known in mammalian cells (Colicelli 2004; Stenmark & Olkkonen 2001). These Rabs cycle between a GTP-bound active form and a GDP-bound inactive form, the GDP-bound form joining a pool of available Rabs for the next round of vesicle shuttling (Soldati et al. 1993). When in the activated GTP-bound state they are anchored to the membrane through a lipid isoprenoid attachment at their C-terminus. The status of Rab proteins is regulated by a number of effectors; such as Rab escort protein (REP); GDP exchange factor (GEF); GDP activating protein (G-) and GDP dissociation inhibitor (GDI). Together REPs, GDIs, GEFs and GAPs contribute to the effective targeting of the appropriate Rabs to membranes and their subsequent recycling. In particular, Gdi proteins help maintain the free pool of Rab-GDP by solubilising the isoprenoid-modified Rab proteins, burying the lipid moiety and protecting it from the hydrophilic intracellular environment (Pylypenko et al. 2006). As a consequence, depletion of Gdi in yeasts has been shown to lead to a loss of the soluble pool of Rabs in the cytosol, which in turn leads to inhibition of vesicle trafficking (Garett, M. D. et al. 1994). The process of Rab recycling and the role of the various effectors are highlighted in figure 1.

There are only a small number of known Gdi proteins compared to the Rabs with which they interact. For instance, in humans there are only two Gdi paralogues (GDI1 and GDI2) but more than 60 Rab genes (Pereira-Leal & Seabra 2001). Human GDI1 has been the most extensively studied Gdi, although the crystal structure of Gdi1 was first solved for the bovine version that has a 98% identity to the human orthologue. The structure of bovine Gdi1 revealed that the three conserved primary sequence elements (SCR1, SCR2 and SCR3) fold to produce domains I and II (Luan et al. 2000; Rak et al. 2003). Domain I contains the Rab binding platform, while domain II contains a mobile effector loop as well as a hydrophobic pocket which interacts with the Rab's geranylgeranyl attachment (Luan et al. 2000). In general it is thought the Rab proteins are initially anchored to the membrane via a prenylated C-terminus and Gdis bind the Rab via Gdi domain I,

triggering a conformational change that enhances the binding of the Gdi domain II to the isoprenoid and subsequent Rab extraction from the membrane (Rak et al. 2003).

It has been proposed that different Rabs exhibit differential binding to either Gdi1 or Gdi2 irrespective of their cellular concentration (Erdman & Maltese 2001). Furthermore, although the mammalian *Gdi1* and *Gdi2* paralogues have 86% primary protein sequence identity, they display distinct expression patterns. For example, in rats *Gdi1* is predominantly expressed in the neural tissues while *Gdi2* is ubiquitously expressed (Nishimura et al. 1994), suggesting that there are functional differences between different Gdis and at least in the case of *Gdi1*, a tissue specific role (Erdman, R., Maltese, W. 2001). In particular, *Gdi1* appears involved in neural development. For instance in *Drosophila* lethal Gdi mutations affect the mitogenic signalling to the imaginal discs and brain, leading to small imaginal discs and small larval brains (Ricard et al. 2001). Further, GDI1 mutations in humans can cause X-linked, non-specific mental retardation and severe impairment of learning abilities (D'Adamo et al. 1998). Similar effects are also observed in Gdi1 deficient mice, which have cognitive impairments associated with altered synaptic vesicles (Bianchi et al. 2009). Although Gdi1 deficient mice are fertile and anatomically normal their ability to perform tasks requiring short-term memory is affected and they have altered social behaviour (D'Adamo et al. 2002).

Here we report the identification and characterisation of a new *Rab-gdi* gene *(gdi3)* that is developmentally regulated, tissue specific and found only in fish and frogs where it is necessary for normal development.

2. Materials and Methods

2.1 Cloning and sequencing

RNA was extracted from stage 18 *Xenopus tropicalis* embryos (Guille, 1999) and used to generate a cDNA library (Precision nanoScript™ 2 Reverse Transcription kit from Primerdesign). Gdi3 gene specific primers (forward GTACTAGACTCTAGAATGGAGGAGATGTATGATGTC and reverse CATTCTTGGAAATCTGAGTTGTC) were used to amplify a 1332-bp product by extension with Q5 high fidelity Taq (New England Biolabs). 3' A overhangs were added to the amplified product by incubation with standard Taq polymerase (New England Biolabs) and the amplicon cloned into a pGEM-TEasy® TA vector (pGEM-TEasy, Promega). Recombinants were

selected by restriction digest screening and the full sequence confirmed by Sanger dideoxy sequencing and primer walking (Source Bioscience).

2.2 Bioinformatic analysis

Gdi protein sequences from a number of species were obtained from Ensembl, NCBI and Xenbase cDNA and genomic databases. Sequences were aligned using the ClustalW multiple sequence alignment method (Larkin, et al. 2007), positions in which most sequences had a gap were removed. The aligned sequences were subsequently analysed with PhyML 3.0 (Phylogeny.fr platform, Dereeper, et al. 2008) using JTT+G model to produce a maximum likelihood phylogenetic tree. The tree topology was further tested for statistical significance by Bootstrap analysis with 100 replicates. Synteny was determined by comparing the genome loci of the respective Gdi genes obtained from the Ensembl, UCSC and Xenbase genome browsers. The tertiary structural analysis was based upon the X-ray crystallographic model of bovine Gdi1 (PDB ID: 1D5T). The crystallographic model of the yeast monoprenylated Ypt1:RabGDI complex (PDB ID 1UKV) was aligned with the bovine Gdi structure to provide contextual reference for the interaction surfaces and lipid cavity. The program PyMOL (Schrödinger, LLC) was used to display the structural models and highlight the relative positions of the key substitutions in Gdi1 versus Gdi3.

2.3 RT-PCR

Total RNA was extracted from a series of embryonic developmental stages and adult *Xenopus tropicalis* tissues (Guille, 1999) and reverse transcribed (Precision nanoScript^{TM} 2 Reverse Transcription kit, Primerdesign). The quality and quantity of the cDNA prepared was estimated by amplification of a 190 bp region of the house-keeping gene ODC (ornithine decarboxylase). Gene specific primers (forward TCAGATGGGAAATACGTGGC and reverse CACGTTGTCTCAAAGTGCG) were used to amplify a 210 bp fragment of the *gdi3* gene. The PCR was conducted with 33 cycles, which had been previously experimentally determined to be within the linear range of this amplification. For more precise quantification Sybr Green qPCR was employed. For the qPCR all reactions were conducted as technical triplicates and relative fold changes calculated by the $\Delta\Delta C_t$ method.

2.4 Wholemount in situ hybridisation

A 494 bp 3' region of *gdi3* was PCR amplified from the full gene previously cloned into pGEM-T Easy vector. The forward 5'- GTACTAGACCTCGAGTTTGTGACCCCAGTTATGTG -3' and the reverse 5'- GAACTATCCTCTAGATCATTCTTGGAAATCTGAGTTG-3' primers encoded XhoI and XbaI restriction sites, in order to facilitate cloning into the pBluescriptKS+ vector. The new construct was linearized with either NotI or KpnI and *in vitro* transcribed with SP6 or T7 RNA Polymerase to produce sense and antisense mRNA probes respectively. Probes were used in a standard whole-mount *in situ* hybridization protocol (Guille, 1999). Colour change was observed periodically to avoid over staining.

2.5 Microinjection and western blotting

Xenopus tropicalis embryos were microinjected with either synthetic mRNA or morpholino (GeneTools, LLC) using a micromanipulator (Sutter-M33) connected to a microinjector (Medical Systems Corp. PLI-100). Needles were calibrated using a 30 mm capillary tube as a reference, in general a single injection was in the range of 10 nl. Embryos were injected at the one-cell stage and once injected were transferred to a Petri dish containing fresh 4 % Ficoll in 1 X MBS and allowed to recover for several hours at 14 °C. Embryos were then moved to 0.1 X MBS solution at 23 °C and monitored until the required stage. Expression of exogenously injected RNA was confirmed by Western blotting (Guille, 1999) and the blot visualised by chemiluminescence.

3. Results

3.1 Cloning, sequencing and phylogenetics

We have previously described a novel promoter type dependent on an A-form DNA structure and identified 86 such potential examples in the *Xenopus tropicalis* genome (Whitley et al. 2014). A-form DNA is a an 'overwound', non-canonical DNA helical structure favoured by dehydration. However, certain DNA sequences can adopt A-form conformations under physiological conditions. It has previously been implicated in transcriptional regulation of the *gata2* promoter (Llewelyn et. al. 2009; Scarlett et.al. 2004). One of these was the predicted but uncharacterised gene, designated in Ensembl (http://www.ensembl.org/) as ENSXETG00000009337 with a corresponding transcript ENSXETT00000020569. This gene also occurs in NCBI (XM_004914093.2) as a predicted protein. Comparison of the predicted open reading frame with the human genome showed ENSXETG00000009337 to have highest similarity to *GDI2* (identity 70.8%) followed by *GDI1* (identity 68.8%). Orthologues of both *GDI1* and *GDI2* have been

previously identified in *Xenopus tropicalis*, suggesting that ENSXETG00000009337 is a novel gene that is distinct from the *gdi1* and *gdi2* paralogues, and also that it is present in frogs but absent in humans. A single version of the gdi3 is found in the closely related pseudo-tetraploidy species *Xenopus laevis* on chromosome 4L.

We therefore preliminarily designated this new gene *gdi3*. The *gdi3* sequences available on Ensembl (genome assembly 4.2) and Xenbase (genome assembly 7.1 where it is designated Xetro.D0227.1) differ from each other in terms of both exon structure and sequence. To determine the true *gdi3* sequence, total mRNA was isolated from stage 18 *Xenopus tropicalis* embryos and the corresponding *gdi3* cDNA fully sequenced by primer walking. The observed sequence most closely corresponded to the later 7.1 assembly. However, a number of short nucleotide polymorphisms (SNPs) were identified between our sequence and that published in the 7.1 assembly (figure 2). Comparison on the protein sequence level showed that out of the identified 76 SNP variations at the DNA level only 9 were non-synonymous and would be translated as alternative amino acids.

To determine whether *Xenopus gdi3* was the result of a recent gene duplication in amphibians or part of the gnathostome Gdi repertoire, we retrieved Gdi protein sequences from a variety of vertebrate species. The sequences were aligned and a phylogenetic tree was generated using maximum likelihood. In the tree (figure 3A), the gnathostome Gdi proteins are separated into three distinct paralogue groups, corresponding to Gdi1, Gdi2 and Gdi3. Both the alignment and the branch lengths in the tree indicate that the Gdi1 sequences are most highly conserved in all taxa except reptiles. Interestingly, all bird genomes analysed lack a Gdi1 orthologue. Gdi2 was found in all gnathostome species, although the Gdi2 sequences generally showed more variability; the exceptions were reptiles and birds where the Gdi2 sequence is highly conserved. Gdi3 sequences were found, besides amphibians, in cartilaginous fish, actinopterygians and the coelacanth. While clearly grouping together, the Gdi3 sequences were the most divergent of all Gdi genes, particularly for the actinopterygians. No Gdi3 orthologue was found in amniotes. Lampreys have two Gdi genes, but neither groups clearly with the gnathostome paralogues and hence we identified these as gdiA and gdiB.

The phylogenetic tree suggested a monophyletic origin of the *Gdi3* genes in cartilaginous fish, actinopterygians and sarcopterygians. To further test this, we analysed the genomic environment of the *Gdi* paralogues in different species. This synteny analysis (figure 3B) revealed a number of shared neighbouring genes such as *Arih2*, *Slc25a20* and *Prkar2a* between the *Gdi3* orthologues,

despite some rearrangements affecting the gene order. The locus as such is also well conserved in the *Gdi3* lacking amniotes, but without the presence of *Gdi3*. The *Gdi1* and *Gdi2* paralogues are surrounded by their own, distinct sets of genes, for instance characteristic for *Gdi1* orthologues is the presence of *Atp6ap1*, while *Ankrd16* is linked with *Gdi2*. The similarity of the *Gdi3* locus between species, and its distinctness from the synteny of *Gdi1* and *Gdi2,* confirmed that the *Gdi3* genes identified in our analysis are indeed orthologues, and distinct from *Gdi1* and *Gdi2*.

To gain further insight into the evolution of the Gdi family in chordates, we analysed the protein sequences in more detail. Our analysis revealed a number of positions in which the three gnathostome (Gdi1, Gdi2, Gdi3) and two agnathan (gdi1A, gdiB) paralogue groups differ from each other (Table 1). Out of these 20 positions, Gdi1 and Gdi2 share the same amino acid in 9 positions, Gdi1 and Gdi3 in 5, and Gdi2 and Gdi3 in 4. This would support the grouping of Gdi1 and Gdi2 in the phylogenetic tree. The lamprey gdiA genes share about the same number of amino acid similarities with all three gnathostome Gdis, but gdiB has more positions in common with Gdi1 and Gdi2. We next mapped these variable residue sites on to the known three dimensional structure of bovine Gdi1 (PDB ID: 1D5T) which provided a useful platform for structural analysis due to the high amino acid sequence identity to *Xenopus* Gdi3. In order to interrogate protein-protein interactions in the Rab GDP-dissociation inhibitor complex, we employed the crystallographic model of the yeast monoprenylated Ypt1:RabGDI complex (PDB ID 1UKV). This structure has a resolution of 1.5A with good geometry and usefully contains the Guanosine-5'-diphosphate, geranylgeranyl ligand and a bound Magnesium ion. This structure was aligned with the bovine Gdi structure to provide contextual reference for the interaction surfaces and lipid cavity. The Rab binding surface was strikingly clear of variable sites between Gdi1, Gdi2 and Gdi3, with the bulk of the variable sites mapping distal to the Rab binding platform of the Gdi. Notably two of the variable sites between Gdi1 and Gdi3 (Y117H and N130D) mapped to the start of a short helix lining the lipid binding pocket. This helix is shown in an 'open' position in the yeast structure (which is with bound lipid) but 'closed' in the bovine structure (which is without the lipid). The two substitution sites in Gdi3 map directly to the flanking hinge regions for this helix (figure 3C) and are likely to alter the dynamics of the pocket.

3.2 Temporal and spatial expression of gdi3

In order to determine the expression levels of the newly designated *gdi3* gene during early *Xenopus* development we performed an RT-PCR time course. Specific primers were designed to

span an intron and against the 3' end of the *gdi3* sequence, the most variable region when aligned to *gdi1* and *gdi2*. *Xenopus tropicalis* embryos were collected at stage 2 (prior to the midblastula transition); stage 8 (blastula); stage 10 (early gastrula); stage 12 (late gastrula); stage 16 (mid-neurula); stages 18, 19, 20, 21, (late neurula) and finally tailbud stages 22, 23, 25 30 and 40. *Gdi3* expression in *Xenopus tropicalis* is only weakly maternal but the gene is zygotically expressed from stage 12, where it rapidly increases to stable levels at stage 18, and stays constant across the remaining developmental stages examined (figure 4A). The expression profile of gdi1 and gdi2 can be obtained from the Wellcome trust *Xenopus* server in which transcriptome-wide analysis shows their expression to increase over 66 hours post-fertilizastion, compared to the peak of gdi3 expression we detect at 20 hours (stage 18).

We next studied the spatial distribution of *gdi3* mRNA in *Xenopus tropicalis* early development. Anti-sense and sense control RNA *in situ* hybridisation probes were *in vitro* transcribed and used to probe a range of *Xenopus* early developmental stages. Probe sequences were again designed to be complementary against the 3' end of the *gdi3* of the mRNA to provide specificity to this paralogue. The RT-PCR temporal analysis had shown *gdi3* to be mainly expressed from mid-neurula stages onwards; we therefore collected developmental stages corresponding to this developmental process at stages 16, 18, 22, 25 and 30. Analysis of the *in situ* hybridisation results of *Xenopus tropicalis* embryos revealed *gdi3* to be expressed throughout the neural plate at stages 16 and 18 but with higher levels at the border. At the later stages gdi3 expression was more clearly seen at the anterior regions of the embryo (figure 4B), particularly in the eye although some expression could still be observed in the neural tube. As *in situ* hybridisation is not possible in adult frogs we tested the spatial distribution of *gdi3* mRNA in adult frog tissues by dissection and RNA extraction from specific organs; these samples were then used in a RT-PCR reaction. We dissected and analysed brain, eyes, liver, leg muscle, intestine, heart and kidney tissue. Despite some degradation of the liver mRNA sample (as is commonly seen by our laboratory from RNA prepared from this tissue (unpublished data)), in general *gdi3* expression was observed at equal levels in the eye, heart and leg muscle but no transcript was detected in brain, liver or intestines (figure 4C). In the adult frog, gdi1 and gdi2 EST profiles (NCBI, UniGene) show that gdi1 is expressed in the head and thymus while gdi2 is expressed in all analysed tissues except bone, heart and kidney.

3.3 Gdi3 functional analysis

To investigate the function of *gdi3* we performed both a knock-down and an over-expression analysis of the gene. We firstly knocked-down levels of gdi3 protein through the use of anti-sense morpholino oligonucleotides (AMOs), an approach widely used to prevent translation of the endogenous target mRNA (Summerton & Weller 1997; Heasman et al, 2000). AMOs are designed to prevent translation of the target mRNA by hybridising to, and consequently blocking, the translation start site, thus hindering ribosome progression. However, AMOs have also been shown to be effective if they hybridise upstream of the start codon in the 5' UTR (Summerton 1999). Therefore, as a control for possible off-target effects we designed two non-overlapping AMOs, the first of these targeted the AUG and the second was just upstream of this region (figure 5A). We designated these AMOs MO1 and MO2 respectively; effects specific to gdi3 knock-down should be mirrored in phenotypes arising from either of the morpholinos. Preliminary morpholino injections of between 2 and 16 ng were conducted and 8ng was selected as optimal for both penetrance of the phenotype and the survival of the embryos, data not shown. Injected embryos were allowed to develop until stage 40 with survival rates of 81% (total of 97 surviving embryos) for MO1 and a survival rate of only 28% (total of 60 surviving embryos) for MO2 observed. Injection of either AMO gave identical phenotypes of a shortened A-P axis and severe anterior defects (a 100% penetrance for MO1 and 95% for MO2). We also injected 8 ng standard control morpholino (SCMO) available Genetools (http://www.gene-tools.com/), SCMO injected embryos displayed no effect (0% from 60 surviving embryos) on the A-P axis. Microcephaly was also visible in a subset of embryos from stage 26 onwards. These phenotypes were not exhibited by either the uninjected controls or embryos injected with a standard control morpholino which was designed to not hybridise to any *Xenopus tropicalis* transcript.

We also exogenously over-expressed a 6xmyc-tagged fusion form of the Gdi3 protein by microinjection of *in vitro* transcribed mRNA into one-cell *Xenopus* embryos. The *gdi3* open reading frame of this synthetic mRNA was flanked by the 5' and 3 'UTRs of globin, to improve both RNA stability and translational efficiency (Iebhaber 2003). We tested a preliminary range of 14 to 112 pg of *gdi3* mRNA amounts injected into embryos and found optimal penetrance to survival rates at 20 pg, data not shown. Therefore, 20 pg of *gdi3* synthetic mRNA was injected at the one-cell stage and the embryos photographed when the uninjected controls were at stage 40. The presence of the translated, exogenous protein was confirmed by Western blotting of total embryo extract from a sub-set of the injected embryos collected at stage 18, the extract was probed with anti-myc antibody and showed increasing protein levels with increasing amounts of

injected mRNA (figure 5B). Quantification of the band intensities by ImageQuantTL showed an increase of 1.4 fold in intensity in the 24 pg lane compared the 12 pg injection. Over-expression phenotypes (figure 5C) routinely showed loss of anterior structures, with the embryos exhibiting anterior-posterior axis malformations at later developmental stages. To explore the basis of the phenotype we analysed a range of anterior markers in Gdi3 over-expressing embryos by RT-PCR. As early marker genes we selected Noggin (early mesodermal induction) and sox2 (neural ectoderm). We also included a number of later anterior markers. Progressing from most anterior these were Xanf1 (forebrain), pax6 (eye and forebrain) otx2 (forebrain/hindbrain boundary), engrailed (hindbrain) and krox20 (hindbrain). Only a minor reduction of expression in the hindbrain marker *krox20* and smaller reduction in the forebrain marker *pax6* (figure 5D) were detected by semi-quantitative RT-PCR. However closer analysis of all these marker genes by more quantitative RT-QPCR showed no significant variation, suggesting the minor changes observed by semi-quantitative analysis were the consequence of minor loading variation.

4. Discussion

We have previously shown that the *Xenopus gata2* promoter is regulated, at least partially, by an A-form DNA structure (Llewellyn et al. 2009; Scarlett et al. 2004). A range of genes were identified in a screen to reveal other, similarly regulated targets and our data suggested that the promoter of one of these genes was probably, like *gata2*, a target for the Ilf3 transcription factor (Whitley et al. 2014), here we have described the identification and characterisation of this gene. After initial comparison of this uncharacterised gene to the published *Xenopus* and human genomes, it was identified as a novel member of the *Xenopus Rab-gdi* family that we designated, according to current gene designation criteria, *gdi3*. This name was subsequently approved by the *Xenopus* gene nomenclature committee after consultation with HUGO in 2014.

The phylogenetic analysis of all of the known *Rab-gdi* members showed a distinct grouping of the two known paralogues, *gdi1* and *gdi2,* away from *gdi3,* with g*di3* present in all gnathostome groups except amniotes. Our phylogenetic and synteny analyses suggest that the ancestral *Gdi* gene was duplicated twice in the course of the two vertebrate genome duplications (1R and 2R). Based on both the phylogenetic tree and the analysis of indicative amino acids it is likely that the first duplication gave rise to the Gdi1/2 and Gdi3/4 precursors. Of the four initial paralogues after the second genome duplication, three have been retained in most gnathostomes. *Gdi4* must have

been lost early after 2R, and *gdi3* was subsequently lost in the amniote lineage. Birds appear to have shed their *gdi1* gene, which is still found, though quite diverged, in reptiles.

An embryonic developmental time-course revealed that *gdi3* transcripts were detectable at all stages analysed but only weakly present prior to the mid-blastula (stage 9) that corresponds to the onset of zygotic transcription. As we were able to detect *gdi3* mRNA at stage 2 i.e. before the onset of zygotic transcription, *gdi3* mRNA must be maternally deposited. Equilibrium *gdi3* mRNA levels rose very markedly at the beginning of the neurula stages and remained high until the last stage recorded (tadpole stage), although levels did decrease slightly at later stages. By contrast both *gdi1* and *gdi2* in *Xenopus* are strongly expressed maternally (xenbase.org). Spatial expression patterns of *gdi3* as observed by *in situ* hybridization agreed with the RT-PCR data i.e. that *gdi3* mRNA is first zygotically activated at neurula stages of development. The observed neural expression of *gdi3* in early *Xenopus* development is similar to the expression observed for *gdi1* in mouse and rat (Bächner et al. 1995; Nishimura et al. 1994). Indeed, *gdi3* mRNA overlaps with the reported expression of *gdi1* on AxelDB at stage 30, although the anterior predominance of *gdi3* is much more marked. Although *gdi3* is expressed in neurula tissues in tadpoles it is absent in adult frog brain, suggesting that it is a developmentally important gene involved in setting up neurula and anterior structures but not necessarily required for their maintenance. Possibly Most likely the role of gdi3 is restricted to early development and the gdi1 and 2 family members contribute to adult brain function.

The striking similarity between both the over-expression and the knock-down resultant phenotypes suggests that Gdi3 protein levels require precise regulation for normal activity. Loss and gain of function producing the same phenotypic change is by no means unknown. A 'squelching' effect, where the titration out of a critical factor by the exogenously over-expressed protein mirrors the effect of the absence of the protein, has been reported for the Pacman gene in *Drosophila* (Waldron et al. 2015). Such a squelching mechanism may operate through the Gdi3 - Rab-GTPase interaction; since either an excess or reduction in Gdi3 would alter the pool of available Rabs. As these effects are specifically in A-P axis formation and the anterior structures of *Xenopus* embryo it may be that Gdi3 is interacting with a particular subset of Rabs. Altered Gdi levels may affect Rab3d Rab4, Rab5, Rab7, Rab8, or Rab40 activity, all of which are implicated in early development (Kim & Han 2011). Perhaps the most likely key interaction with Gdi3 is that of Rab40, which is known to be involved in early *Xenopus* A-P axis formation through the non-canonical Wnt/PCP signalling pathway (Lee et al. 2007). Knock-down and overexpression of Rab40 show gastrulation defects and inhibition of the A-P axis elongation, phenotypes similar to

those obtained from our experiments with Gdi3. Furthermore, both GDI1 and GDI2 have previously been shown to have very low affinity for Rab40 (Kirsten M.L. 2011), suggesting an interaction through an alternative Gdi. Another possible Gdi3 target may be Rab3d which has also been shown to be required for *Xenopus* anterior regulation where it regulates the secretion of the BMP antagonist noggin in the neural plate (Kim & Han 2011). The expression of *gdi3* in the neural plate may also suggest important Rab interactions; Rab4 and Rab5 have been shown to regulate axon elongation via endosome recruitment in the growth cone (Falk et al. 2014).

As both *gdi1* and *gdi3* are expressed in the neural system and share considerable identity a number of questions arise: do *gdi1* and *gdi3* represent redundancy; do they have different targets; are they controlled similarly? Protein structure analysis may be used to begin to answer such questions. Our analysis of the *Gdi* gene family identified a number of amino acid variations between the paralogues. When mapped against the tertiary structure of Gdi proteins, nearly all of these variable sites map to the distal face of the Gdi, away from the Rab binding platform. In particular, the Y117H and N130D variation between Gdi1 and Gdi3 map to a critical region of the lipid binding pocket. These data suggest that the regulation or kinetics of release of the Gdi-Rab complexes from membranes may be distinct for Gdi1 and Gdi3 but that they are likely to bind the same Rabs.

Overall we have identified a new member of the Gdi family, Gdi3, which has been lost during evolution in amniotes. Gdi3 levels need to be carefully regulated during normal development or axial patterning of the embryo fails, possibly due to mis-regulation of Rab40 or Rab3b, which control wnt and BMP signalling respectively. Gdi3 is therefore a critical component of the embryo in early development. However, its relationship with Gdi1, which has an overlapping expression pattern but is unable to substitute for it, will need further clarification in the future as will the biochemistry of Gdi3-Rab and Gdi3-lipid interactions.

Acknowledgements

We would like to thank Dr Colin Sharpe and Julian Mitchell for discussions and technical advice. We would also like to thank the European *Xenopus* Resource Center (EXRC) for providing expertise concerning embryo manipulation.

References

Bächner, D., Sedlacek Z., Korn B., Hameister H., Poustka A. 1995. Expression patterns of two human genes coding for different rab GDP-dissociation inhibitors (GDIs), extremely conserved proteins involved in cellular transport. *Human molecular genetics*, 4(4), pp.701–

708.

Bianchi, V., Farisello P., Baldelli P., Meskenaite V., Milanese M., Vecellio M., Muhlemann S., Lipp H.P., Bonanno G., Benfenati F., Toniolo D., D`Adamo P. 2009. Cognitive impairment in Gdi1-deficient mice is associated with altered synaptic vesicle pools and short-term synaptic plasticity, and can be corrected by appropriate learning training. *Human molecular genetics*, 18(1), pp.105–17.

Colicelli, J. 2004. Human RAS superfamily proteins and related GTPases, *Sci. STKE* (250), pp.1–53.

D'Adamo, P., Welzi H., Papadimitriou S., Raffaele di Barletta M., Tiveron C., Tatangelo, L., Pozzi, L., Chapman P.F., Knevett, S.G., Ramsay M.F., Valtorta F., Leoni C., Menegon, A., Wolfer D.P., Lipp H., Toniolo. 2002. Deletion of the mental retardation gene Gdi1 impairs associative memory and alters social behavior in mice. *Human molecular genetics*, 11(21), pp.2567–80.

D'Adamo, P., Menegon A., Lo Nigro, C., Grasso, M., Gulisano, M., Tamanini, F., Bienvenu T., Gedeon, A., Oostra B., Wu, S., Tandon, A., Valtorta, F., Balch, W.E., Chelly, J., Toniolo, D. 1998. Mutations in GDI1 are responsible for X-linked non-specific mental retardation. *Nature genetics*, 19(2), pp.134–9.

Dereeper A1, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research.* 36 (Web Server issue):W465-9

Erdman, R., Maltese, W. 2001. Different Rab GTPases associate preferentially with alpha or beta GDP-dissociation inhibitors. *Biochemical and biophysical research communications*, 282(1), pp.4–9.

Falk, J., Konopacki F. A., Zivraj, K. H., Holt, C.E. 2014. Rab5 and Rab4 Regulate Axon Elongation in the Xenopus Visual System. *Journal of Neuroscience*, 34(2), pp.373–391.

Guille, M. 1999. *Molecular Methods in Developmental Biology* (Guille, M., ed) Vol. 127, Humana Press

Heasman J, Kofron M, Wylie C., 2000. Beta-catenin signaling activity dissected in the early *Xenopus* embryo: a novel antisense approach. *Developmental. Biology.* Jun 1;222(1), pp.124-34

Kim, H., Han, J.-K. 2011. Rab3d is required for Xenopus anterior neurulation by regulating Noggin secretion. *Developmental dynamics : an official publication of the American Association of Anatomists*, 240(6), pp.1430–9.

Kirsten, M.L. 2011. Biophysical Studies of Rab GTPase Membrane Binding. Unpublished PhD thesis. Imperial College London

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG., 2007. ClustalW and ClustalX version 2. *Bioinformatics* 23(21), pp *2947-2948.*

Lee, R.H.K., Iioka, H. Ohashi, M., Lemura, S., Natsume, T., Kinoshita, N. 2007. XRab40 and XCullin5 form a ubiquitin ligase complex essential for the noncanonical Wnt pathway. *The EMBO journal*, 26(15), pp.3592–606.

Llewellyn, K.J.,Cary, P., McClellan J.A., Guille, M. J., Scarlett, G.P. 2009. A-form DNA structure is a determinant of transcript levels from the Xenopus gata2 promoter in embryos. *Biochim Biophys Acta*, 1789, pp.675–680.

Luan, P., Heine, A., Zeng, K., Moeyer, B., Greasely S.E., Kuhn, P., Balch, W.E., Wilson, I. A. 2000. A new functional domain of guanine nucleotide dissociation inhibitor (alpha-GDI) involved in Rab recycling. *Traffic*, 1(3), pp.270–81.

Nishimura, N., Nakamura, H., Takai, Y., Sano, K. 1994. Molecular cloning and characterization of two rab GDI species from rat brain: brain-specific and ubiquitous types. *The Journal of Biological Chemistry*, 269(19), pp.14191–8.

Pereira-Leal, J.B. & Seabra, M.C. 2001. Evolution of the Rab family of small GTP-binding proteins. *Journal of Molecular Biology*, 313(4), pp.889–901.

Pylypenko, O., Rak, A., Durek, T., Kushnir, S., Dursina, B.E., Thomae, N. H., Constantinescu, A.T., Brunsveld, L., Watzke, A., Waldmann, H., Goody, R.S., Alexandrov, K. 2006. Structure of doubly prenylated Ypt1:GDI complex and the mechanism of GDI-mediated Rab recycling. *The EMBO journal*, 25(1), pp.13–23.

Rak, A., Pylypenko, O., Durek, T., Watzke, A., Kushnir, S., Brunsveld, L., Waldmann, H., Goody, R.S., Alexandrov, K. 2003. Structure of Rab GDP-dissociation inhibitor in complex with prenylated YPT1 GTPase. *Science*, 302(5645), pp.646–50.

Ricard, C.S., Jakubowski, J.M., Verbsky, J.W., Barbieri, M.A., Lewis, W.M., Fernandez, G.E., Vogel, M., Tsou, C., Prasad, V., Stahl, P.D., Waksman, G., Cheney, C.M. 2001. Drosophila Rab GDI Mutants Disrupt Development but Have Normal Rab Membrane Extraction. *Genesis*, 29, pp.17–29.

Scarlett, G.P., Elgar, S.J., Cary, P.D., Noble, A.M., Orford, R. L., Kneale, G.G., Guille, M.J. 2004. Intact RNA-binding Domains Are Necessary for Structure-specific DNA Binding and Transcription Control by CBTF 122 during Xenopus Development. *Journal of Biological Chemistry*, 279(50), pp.52447–52455.

Soldati, T., Riederer, M. a & Pfeffer, S.R. 1993. Rab GDI: a solubilizing and recycling factor for rab9 protein. *Molecular biology of the cell*, 4(4), pp.425–34.

Stenmark, H., Olkkonen, V.M., 2001. The Rab GTPase family. *Genome biology*, 2(5), p.3007.

Summerton, J. 1999. Morpholino antisense oligomers: the case for an RNase H-independent structural type. *Biochimica et Biophysica Acta*, 1489(1), pp.141–158.

Summerton, J.,Weller, D.1997. Morpholino antisense oligomers: design, preparation, and properties. *Antisense & nucleic acid drug development*, 7, pp.187–195.

Waldron, J.A., Jones, C.I., Towler, B.P., Pashler, A.L., Grima, D. P, Hebbes, S., Crossman, S.H, Zabolotskaya M.V., Newbury, S. F. 2015. Xrn1/Pacman affects apoptosis and regulates expression of hid and reaper. *Biology Open*, pp.1–12.

Waggoner, S.A. and Liebhaber, S. A. 2003. Regulation of α-Globin mRNA Stability. *Experimental Biology and Medicine*, pp.387–395.

Whitley, D.C, Runfola, V., Cary, P., Nazlamova, L., Guille, M., Scarlett, G.P. 2014. APTE: identification of indirect read-out A-DNA promoter elements in genomes. *BMC Bioinformatics*, 15(1), p.288.

Figure legends

Figure 1. The Rab GTPase recycling mechanism is controlled by a variety of effectors to ensure a constant Rab soluble pool in the cell cytosol. GDP bound Rab proteins are escorted by Rab escort protein (REP) in order to be prenylated by geranyl-geranyl tranferase (GGT). Once the isoprenoid group is attached, Rab-GDP is solubilised by guanine dissociation inhibitor (GDI) and is ready to regulate vesicle shuttling. Subsequently Rab-GDP is activated by exchange to Rab-GTP, a process assisted by guanine exchange factor (GEF). Once activated Rab-GTP can bind vesicles and export them to the correct destination where Rab-GTP is converted back to Rab-GDP. The conversion from Rab-GTP to Rab-GDP is catalysed by GTPase-activating protein (GAP).

Figure 2. Alignment of the *Xenopus tropicalis* gdi3 protein sequence against the JGI 7.1 gdi3 predicted protein sequence. The upper row corresponds to the observed *gdi3* DNA sequence and the lower row the Gdi3 predicted protein sequence obtained from Xenbase genome assembly JGI 7.1 of *Xenopus tropicalis.* Bases highlighted indicate SNPs between the two sequences while the predicted exonic arrangement is indicated by grey arrows.

Figure 3. Phylogenetic and syntenic analysis of the gdi genes. (A) Phylogenetic tree of chordate Gdi proteins. The tree shows the phylogenetic relationship of Gdi protein sequences from representatives of all main vertebrate classes. *Nanorana parkeri* (XP_018431581 (NCBI)), *Latimeria chalumnae* (XP_005992613 (NCBI)), *Lepisosteus oculatus* (XP_015203569 (NCBI)), *Callorhinchus milii* (XP_007888513 (NCBI)), *Scyliorhynchus canicula* (SSC-transcript-ctg25260 (SkateBase)), *Danio rerio* (NP_001307001 (NCBI)), *Gasterosteus aculeatus* (ENSGACP00000001552 (Ensembl)), *Xenopus tropicalis* (XP_004914150.1 (NCBI)). The tree was produced using the maximum likelihood method, using the Gdi sequence from the basal chordate amphioxus as outgroup. The gnathostome sequences are divided into three distinct groups (Gdi1 in blue, Gdi2 in green, Gdi3 in red), with Gdi1 and Gdi2 grouped together. The lamprey gdiA and gdiB sequences are placed on their own, unique branches. (B) Organisation of the *Gdi3* loci in different vertebrates. The *gdi3* gene with its four upstream and downstream neighbours is shown for the Western clawed frog, coelacanth, spotted gar, stickleback and elephant shark. Teleost fish have undergone a third genome duplication. Both loci are shown, since most genes, including gdi3, have only been retained in one of the two loci. The corresponding locus is also shown for human, chicken and anole lizard, but these species lack the

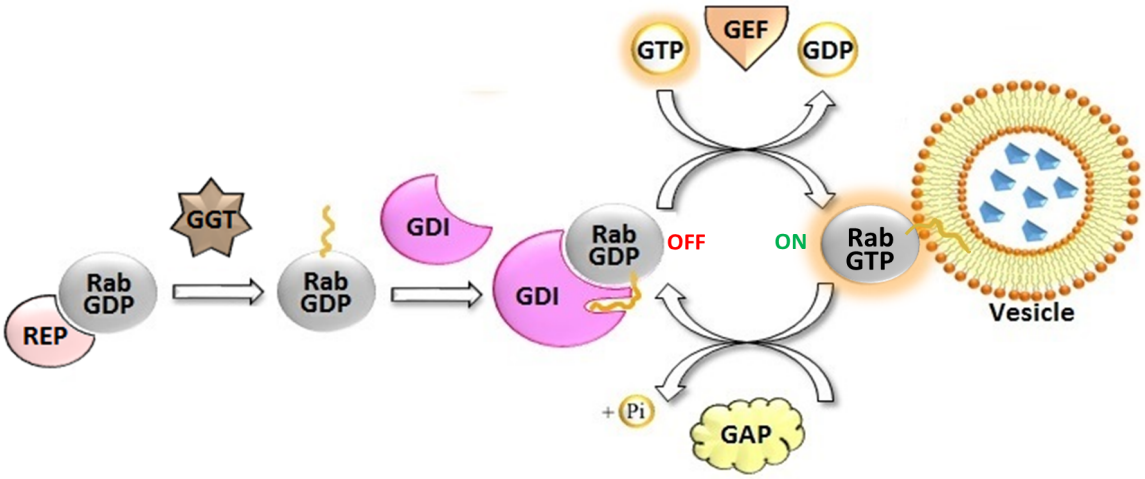
Gdi3 gene. For comparison, the *Gdi1* and *Gdi2* genes with their neighbours are shown for the Western clawed frog. The *Gdi* genes are marked in yellow, while for the neighbouring genes orthologues are shown in the same colour. The *Gdi3* loci are well conserved between species, notably with *Arih2*, *Slc25A20* and *Prkar2a* always present close to *Gdi3*. (C). Two substitution sites in Gdi3 compared to Gdi1 (Y117H and N130D, shown in red) when plotted against the tertiary structure map to the flanking hinge regions of a helix lining the lipid binding pocket. Yeast gdi is shown in blue and bovine gdi1 in yellow and the lipid moiety is highlighted in green. Gdi1 helices D and E are part of the lipid binding site of domain II and beta-sheet b2 is contained in domain I.

Figure 4. *Gdi3* expression in *X. tropicalis*. (A) *Xenopus tropicalis* embryos were harvested at the Nieuwkoop-Faber stages shown and RNA extracted prior to RT-PCR analysis by odc (as an internal reference) and *gdi3* specific primers. *Gdi3* expression is observed throughout all the sampled developmental stages. (B) *In situ* hybridisation analysis of *Xenopus tropicalis* embryos for *gdi3*. The upper row of each stage set corresponds to probing with an antisense probe and the bottom row is probed with a control sense probe. The red arrow shows the anterior region of the embryos. In Nieukoop and Faber (NF) stages 16, 18 and 22 embryos are shown from lateral, dorsal and anterior views left to right. Embryos at stages 25 and 30 are shown in the lateral and dorsal view only. (C) Analysis of *gdi3* expression in adult *Xenopus tropicalis* tissues show presence of mRNA in the eyes, leg muscle and heart but not in the brain, liver, intestine and kidney.

Figure 5. Functional analysis of gdi3 in *Xenopus tropicalis.* (A) The binding sites of MO1 and MO2 to the *Xenopus gdi3* 5' UTR. (B) Embryos injected at one cell stage with either 12 or 24 pg of gdi3 synthetic mRNA were grown to stage 18 and harvested and tested for translation of the exogenous gdi3. Extraction equivalence was confirmed by running a parallel gel in the same electrophoresis tank that was stained with Coomassie Brilliant Blue. (C) Embryos injected with either morpholino for protein level knock-down or gdi3 mRNA for protein over-expression demonstrate loss of anterior-posterior axis. In each panel pictures of three different embryos are shown. (D) A variety of anterior marker genes were analysed for their changes as a consequence of gdi3 over-expression by RT-PCR at stage 18.

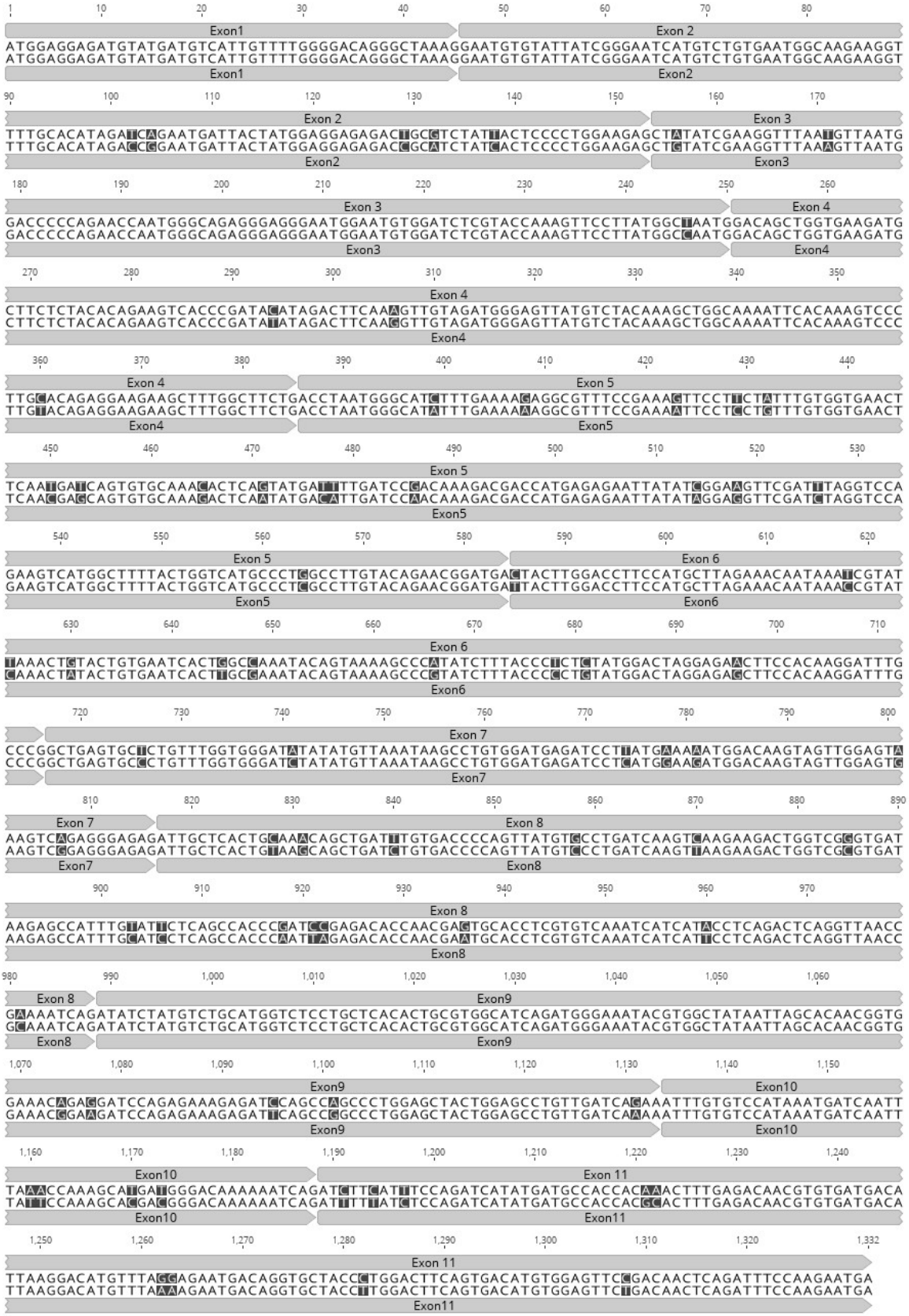
Table 1. Gdi family's indicative amino acids. The numbering of the amino acids (aa) is in relation to the bovine Gdi1 primary sequence.

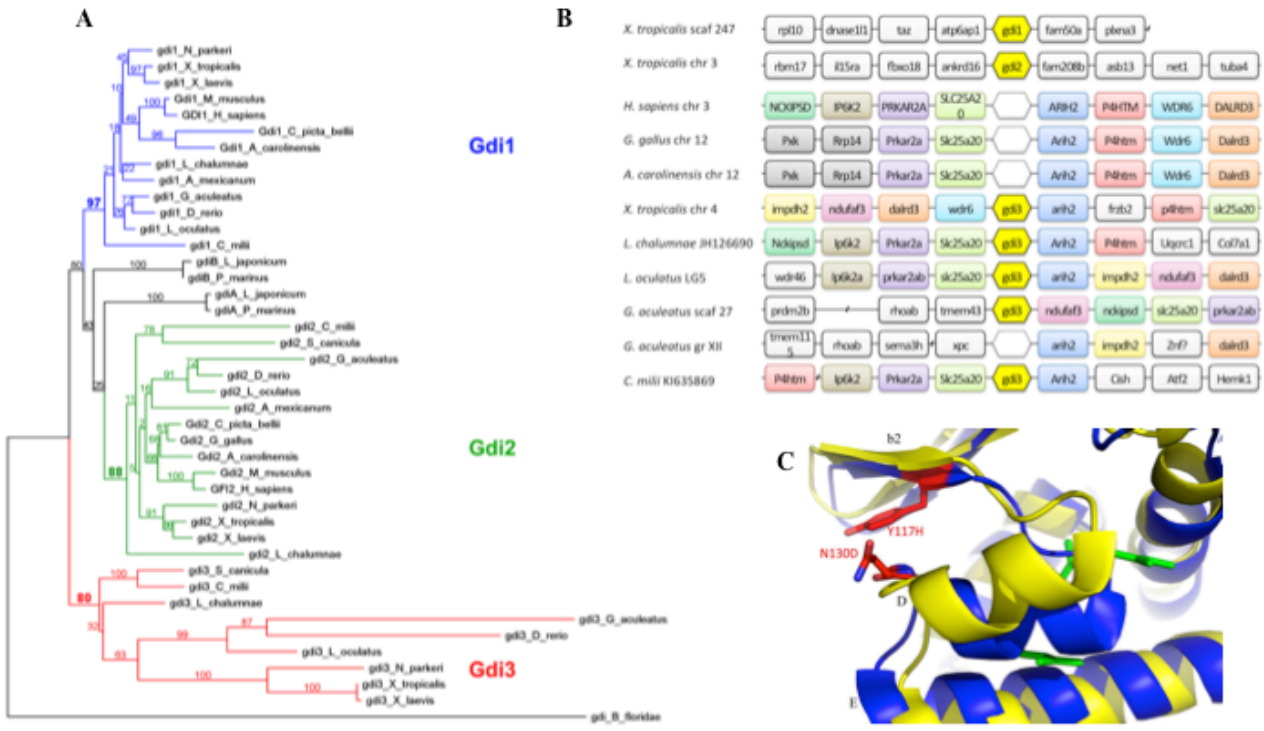
aa	Gdi1	Gdi2	Gdi3	gdiA	gdiB	B_floridae	Exceptions
2	D	N	D/E	N	D	D	C_milii gdi2 D, L_oculatus gdi3 Q
15	T	T	K	K	T	K	
44	S	A	A	S	A	A	
51	E	D	E/D	D	E	E	
105	V	I	V	V	V	C	
117	Y	Y	H	H	Y	H	C_picta beli gdi1 H, G_aculeatus gdi2 H
130	N	S	D	G	N	S	
220	G	G	S/N	G	G	G	L_chalumnae gdi3 K
255	V	I	V	V	V	E	
294	A	V	V/T	V	S	V	L_chalumnae gdi1, D_rerio gdi1 V
296	Q	Q	R/K	Q	Q	Q	A_carolinesis gdi1 K, C_picta beli gdi1 R, A_carolinesis gdi2 K, C_picta beli gdi2 L, N_parkeri gdi2 K, N_parkeri gdi3 Q
324	N	N	T	N	T	N	D_rerio gdi3 A, S_canicula gdi3 M
337	I	I	V/M	V	I	V	
346	Q	Q	D/E	T	Q	K	L_chalumnae gdi3 G
361	D/E	D/E	N/S	N	D	N	X_laevis gdi3 A, X_tropicalis gdi3 D
367	E	K/R	Q	Q	L	K	A_carolinesis gdi1 Q, C_milii gdi1 Q, S_ca gdi2 Q, G_aculeatus gdi3 E
382	A	S	S	S	T	S	S_canicula gdi3 C
392	D	L/M/V	D	L	L	D	G_aculeatus gdi3 G, C_milii gdi3 E
395	D/E	D/E	R/K	E	E	N	C_milli gdi3 A, S_canicula gdi3 T
433	N	E	N/D	E	E	-	A_carolinesis gdi1 S, C_picta beli gdi1 S, A_mexicanum gdi2 Q, G_aculeatus gdi3 G

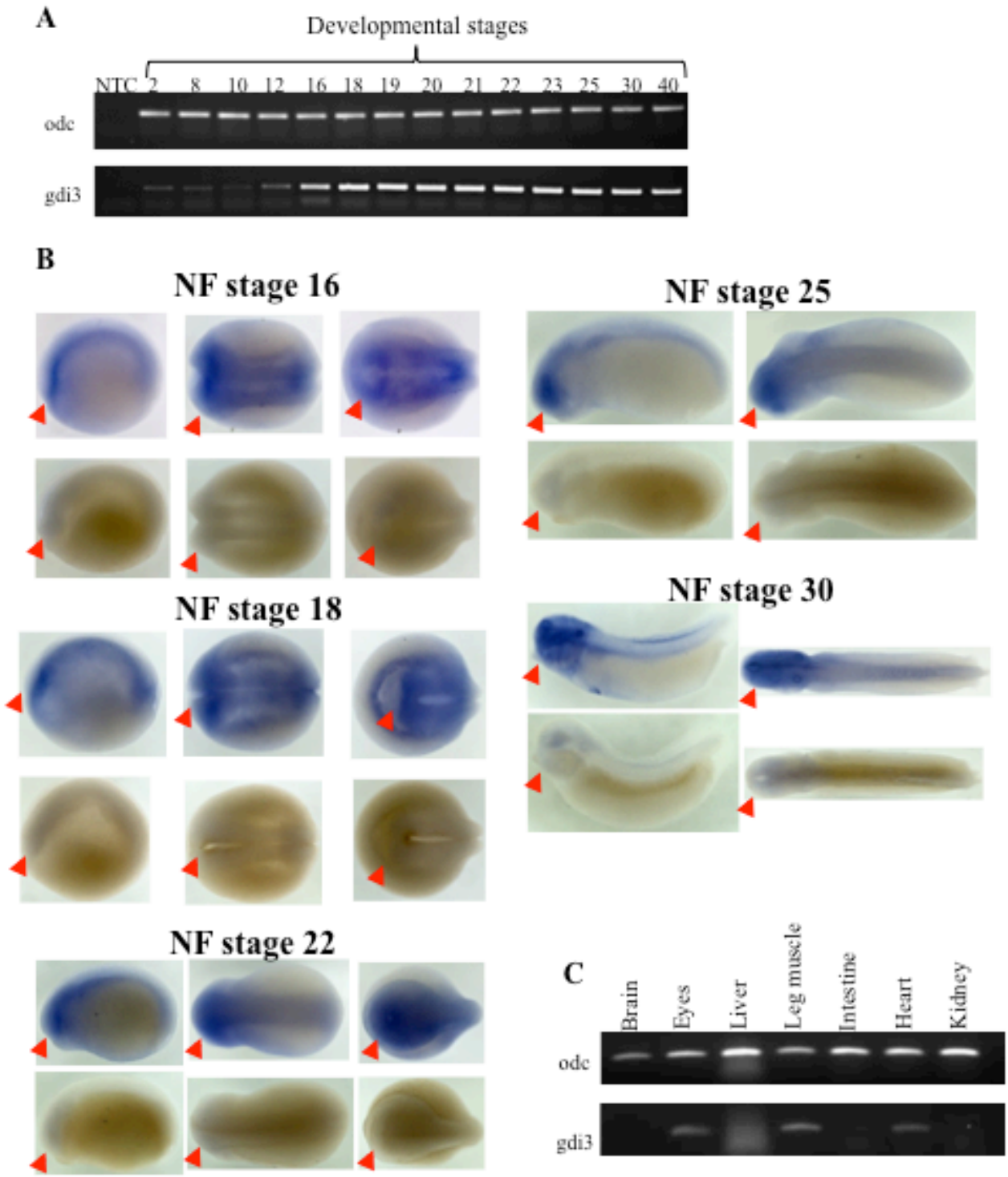

GTP
GEF
GDP
GGT
GDI
Rab
GDP
OFF
ON
Rab
GTP
REP
Rab
GDP
GDI
Vesicle
+ Pi
GAP

Gdi3 observed

Gdi3 JGI 7.1

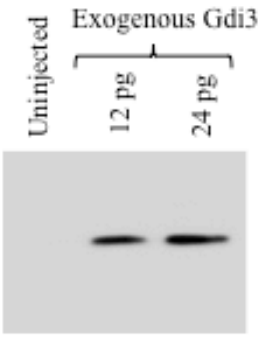



A
gdi1_N_parkeri
gdi1_X_tropicalis
gdi1_X_laevis
Gdi1_M_musculus
GDI1_H_sapiens
Gdi1_C_picta_bellii
Gdi1_A_carolinensis
gdi1_L_chalumnae
gdi1_A_mexicanum
gdi1_G_aculeatus
gdi1_D_rerio
gdi1_L_oculatus
gdi1_C_milii
gdiB_L_japonicum
gdiB_P_marinus
gdiA_L_japonicum
gdiA_P_marinus
gdi2_C_milii
gdi2_S_canicula
gdi2_G_aculeatus
gdi2_D_rerio
gdi2_L_oculatus
gdi2_A_mexicanum
Gdi2_C_picta_bellii
Gdi2_G_gallus
Gdi2_A_carolinensis
Gdi2_M_musculus
GFI2_H_sapiens
gdi2_N_parkeri
gdi2_X_tropicalis
gdi2_X_laevis
gdi2_L_chalumnae
gdi3_S_canicula
gdi3_C_milii
gdi3_L_chalumnae
gdi3_G_aculeatus
gdi3_D_rerio
gdi3_L_oculatus
gdi3_N_parkeri
gdi3_X_tropicalis
gdi3_X_laevis
gdi_B_floridae
Gdi1
Gdi2
Gdi3
B
X. tropicalis scaf 247
rpl10
dnase1l1
taz
atp6ap1
gdi1
farr50a
pbrm3
X. tropicalis chr 3
rbm17
il15ra
fbxo18
ankrd16
gdi2
farr208b
asb13
net1
tuba4
H. sapiens chr 3
NOXPSD
IP6K2
PRKAR2A
SLC25A20
ARIH2
PHHTM
WDR6
DALRD3
G. gallus chr 12
Pak
Rrp14
Prkar2a
Slc25a20
Arih2
Phhtm
Wdr6
Dalrd3
A. carolinensis chr 12
Pak
Rrp14
Prkar2a
Slc25a20
Arih2
Phhtm
Wdr6
Dalrd3
X. tropicalis chr 4
impdh2
ndufaf3
dalrd3
wdr6
gdi3
arih2
frzb2
pRhtm
slc25a20
L. chalumnae JH126690
Ndkipsd
Ip6k2
Prkar2a
Slc25a20
gdi3
Arih2
Phhtm
Ugcrc1
Col7a1
L. oculatus LG5
wdr46
Ip6k2a
prkar2ab
slc25a20
gdi3
arih2
impdh2
ndufaf3
dalrd3
G. aculeatus scaf 27
prdm2b
rhoab
tmem43
gdi3
ndufaf3
ndkipsd
slc25a20
prkar2ab
G. aculeatus gr XII
tmem115
rhoab
sema3h
xpc
arih2
impdh2
Znf7
dalrd3
C. milii KI635869
Phhtm
Ip6k2
Prkar2a
Slc25a20
gdi3
Arih2
Cish
Atf2
Hemk1
C
b2
Y117H
N130D
D

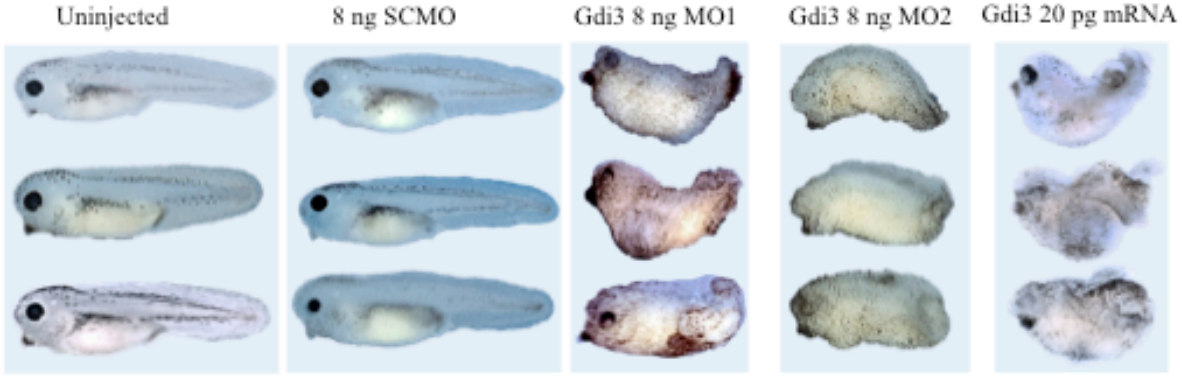

A
Developmental stages
NTC 2 8 10 12 16 18 19 20 21 22 23 25 30 40
odc
gdi3
B
NF stage 16
NF stage 25
NF stage 18
NF stage 30
NF stage 22
C
Brain
Eyes
Liver
Leg muscle
Intestine
Heart
Kidney
odc
gdi3

A

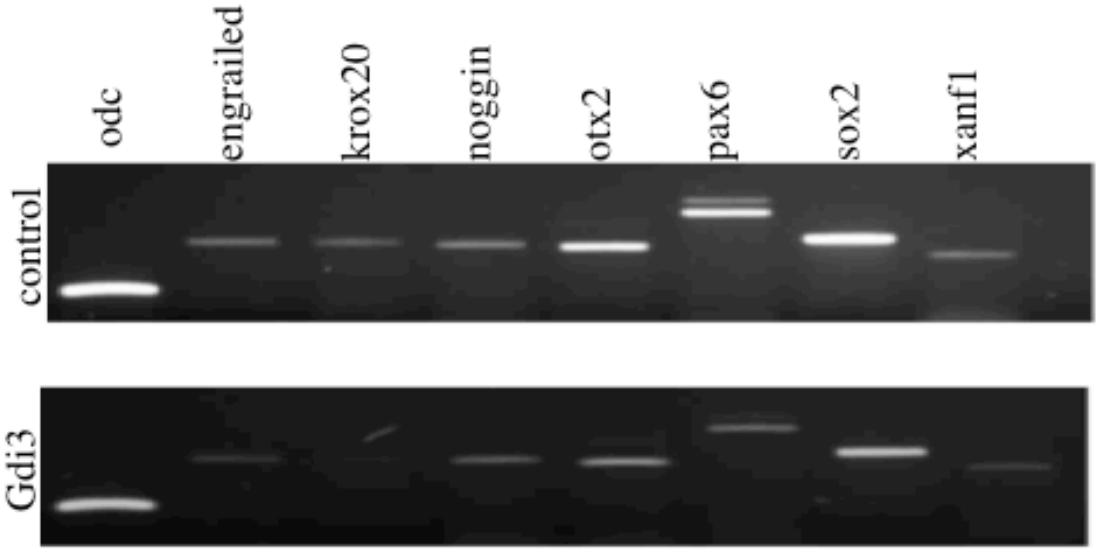
B


Uninjected
Exogenous Gdi3
12 pg
24 pg

C


Uninjected
8 ng SCMO
Gdi3 8 ng MO1
Gdi3 8 ng MO2
Gdi3 20 pg mRNA

D


odc
engrailed
krox20
noggin
otx2
pax6
sox2
xanf1
control
Gdi3