III.3. Evidence for horizontal gene transfer between eukaryotes and Myxobacteria

Evidence for horizontal gene transfer from Eukaryotes to Myxobacteria

Noureddine Loukili[*], Marta Fontes[†], Juan José Lozano[‡], Víctor M. Díaz[*], Bru Cormand[§], Johanna Scheper[*], Florenci Serras[§], Julio Rozas[§], Francisco Murillo[†], and Timothy M. Thomson[*]

[*] Instituto de Biología Molecular, Consejo Superior de Investigaciones Científicas, Barcelona

[†] Departamento de Microbiología y Genética, Facultad de Ciencias, Universidad de Murcia

[‡] Centro de Biología Molecular, Consejo Superior de Investigaciones Científicas, Madrid

[§] Departament de Genètica, Facultat de Biologia, Universitat de Barcelona

Corresponding author:

Dr. Timothy M. Thomson

Instituto de Biología Molecular, CSIC

c. Jordi Girona 18-26

08034 Barcelona, Spain

Phone: 34 93 489 4063

Fax: 34 93 489 4064

E-mail: tthomson@hg.vhebron.es

The occurrence of horizontal gene transfer (HGT) is generally inferred from significant discrepancies found between the observed phylogenetic topology for a given gene or protein, and organismal phylogenetic trees based on the comparison of single or multiple conserved genes. There are currently only a few well-supported cases of prokaryote-to-eukaryote transfer that do not involve mitochondria or plastids, and yet fewer cases suggest the transfer of genes from eukayotes to prokaryotes. Here we show that the recently described human Kua protein presents a striking homology to the CarF protein from the bacterium *Myxococcus xantus*, the only prokaryote in which this gene was found. *M. xanthus* CarF is highly homologous to Kua proteins from protists such as slime molds or protozoa, and the metazoans *Drosophila melanogaster*, *Caenorhabditis elegans* and *Homo sapiens*, and is more distantly related to plant Kua proteins. The unequivocal branching of *M. xanthus* CarF with non-plant Kua proteins, along with the absence of related proteins in other prokaryotes, suggests the occurrence of HGT between *M. xanthus* or related bacteria and an ancestor common to protists and metazoa. Contemplated in the context of a second potential case of HGT involving *M. xanthus*, that of the βγ-crystallin-related protein S, our analysis suggests the occurrence of a gene transfer event from eukaryotes related to Mycetozoa into prokaryotes related to Myxobacteria. In *M. xanthus*, CarF is a sensor for blue light that regulates the expression of genes in the carotenoid biosynthetic pathways. When expressed in mammalian cells, human Kua and *M. xanthus* CarF co-localized in cytosolic structures, together with *M. xanthus* CarR, a protein previously proposed to interact with CarF. Finally, we have analyzed the expression of Kua/CarF during the embryonal development of two distant organisms, *Drosophila melanogaster* and *Mus musculus*, evidencing a restricted pattern of expression in specific neuronal structures in

early stages of mouse development. Therefore, Kua/CarF represents a remarkable example of a protein highly conserved between a bacterium and protists/metazoans, potentially involved in eukaryote-to-prokaryote HGT.

Kua is a protein of unknown function whose gene was found in the course of a comparative analysis of the organization of the genes for the human polyubiquitination regulator UEV1 and its orthologs in other organisms[1]. In humans, the genes for Kua and UEV1 are physically and functionally associated, such that they can be transcribed either as separate units or as a chimeric transcript and protein, in a process that involves alternative splicing and the use of two independent promoters[1]. In organisms such as *Drosophila melanogaster* or *Caenorhabditis elegans*, the gene for Kua is unlinked to that for UEV1, and the two proteins do not appear to be related in any obvious genetic, biochemical or functional fashion.

The Gram-negative bacterium *Myxococcus xanthus* is a sulfate-reducing microorganism that dwells in the soil, feeds on other bacteria and has a complex developmental behaviour that involves cellular specialization processes requiring intercellular signaling systems that modulate specific regulatory pathways[2]. In addition to these features, *M. xantus* responds to blue light by producing carotenoids, which protect the cells against phytochemically generated oxidants[3,4]. Light-induced gene expression in *M. xanthus* involves three unrelated loci (carQ, crtI, carB and carA). Another gene, carD, unlinked to the others, is not regulated by light but is necessary for carQRS and crtI activation. Protein CarD contains a DNA-binding domain similar to that of eukaryotic high-motility proup I(Y) HMGI(Y) proteins, which bind a particular cis-acting site of the carQRS promoter region[5].

In a screen for Tn5 insertion mutants affecting carotenoid biosynthetic pathways in *M. xanthus*, a new gene, designated CarF, was recently isolated and shown to belong to a distinct complementation group (Fontes et al., upublished observations). *M. xanthus* CarF mutants do not express the enzymes of the carotenoid biosynthetic pathway in response to blue light. The CarF mutation does not revert the mutation in another locus, CarR, that encodes a negative regulator of the carotenoid biosynthetic pathway[6,7], while the wild-type carF gene restores the response to blue light in bacteria with wild-type but not mutant CarR (Fontes et al., unpublished observations). This suggests that CarF exerts opposite effects to CarR as an upstream regulator. Initial sequence analysis of CarF indicated a high degree of homology to the Kua proteins from *H. sapiens* and other organisms[1] (Fig. 1). The construction of phylogenetic distance trees with the aligned sequences showed unequivocal clustering of *M. xanthus* CarF with Kua proteins from protists such as *Dictyostelium discoideum*, *Leishmania major* or *Trypanosoma cruzi*, as well as with Kua from metazoans (Fig 2B). Kua proteins from the plants *Arabidhopsis thaliana*, *Glycine max* and *Mycebrotum crystallinum* formed a distintc phyletic group (Fig. 2B). A thorough search for genes or proteins related to CarF or Kua in prokaryotes other than *M. xanthus* or in the yeasts *Saccharomyces cerevisiae* or *Schizosaccharomyces pombe*, as well as other microorganisms for which complete genomes have been sequenced, failed to yield any sequences with significant similarities.

The tree topology observed for Kua/CarF suggests three possible evolutionary scenarios: (1) An origin of Kua/CarF in prokaryotes, with evolution through common ancestry, along with selective losses in all prokaryotes except Myxobacteria and, in eukaryotes, in yeasts and fungi. Since *M. xanthus* CarF clearly groups with protist and animal Kua, while plant Kua proteins form a distinc phyletic group, this scenario would

also imply a significantly faster mutational rate for plant Kua proteins, as compared to non-plant Kua proteins including *M. xanthus* CarF. (2) An origin of Kua/CarF in prokaryotes, followed by horizontal transfer of Kua/CarF from *M. xanthus*, or an ancestor, into eukaryotes at the root of the branching between plants and protist/metazoa. As in the first scenario, this would also imply selective loss in all prokaryotes except Myxobacteria, and also in yeasts and fungi. It also implies faster mutational rates for plant Kua proteins as compared to all other Kua/CarF proteins, including *M. xanthus* CarF. (3) An origin of Kua in eukaryotes, at the root of the bifurcation between plant and protist/metazoan Kua, followed by horizontal transfer into *M. xanthus* from an ancestor organism at the root of the protist/metazoan branching. As in the two previous hypotheses, this scenario assumes the occurence of early divergence between plant and non-plant Kua, accompanied by duplication and faster mutation for plant Kua genes and proteins. However, there is no need to invoque losses in prokaryotes in this case. Therefore, the third scenario would represent the most parsimonious possibility.

Independent evidence for a HGT event would be the occurrence of discrepancies in distance trees for genes in *M. xanthus* other than CarF, or the observation of significant differences in GC content[8] or codon usage[9] for these genes relative to the entire gene complement of the recipient organism. *M. xanthus* is an organism long known to harbor unusually high numbers of genes and proteins with putative counterparts in eukaryotes, including proteins that regulate specialized processes in signal transduction and gene transcription[10,11]. In order to determine if *M. xanthus* proteins other than CarF presented similar phylogenetic topologies, we analyzed selected genes and proteins that showed the strongest degrees of similarity between *M. xanthus* and the eukaryotic organisms for which Kua-related sequences are available. Two informative cases were those for *M. xanthus*

protein phosphatase 1(ref. 12), and protein S, related to eukaryotic beta/gamma crystallins[13]. *M. xanthus* Pph1 has orthologues in plants and animals and, in contrast to CarF, it also has homologues in other prokaryotes (Fig. 2C). *M. xanthus* Pph1 clusters with its prokaryote counterparts, with which it forms a phyletic group unequivocally distinct from the plant or animal protein phosphatases (Fig. 2C). The tree for Pph1 closely follows that determined for *M. xanthus* on the basis of a number of genes and proteins[14], therefore suggesting evolution through common ancestry for this gene. This is the pattern followed by most *M. xanthus* genes and proteins with related sequences in eukaryotes (not shown).

M. xanthus protein S represents a different case. This protein is clearly related to animal β/γ crystallins, and does not appear to have homologues in any other prokaryotes or in such eukaryotes as yeasts, protozoa or plants (Fig. 2D). The lack of protein S-related sequences in these organisms could be interpreted either as vertical evolution with selective gene loss affecting one or more phyla[15], or as being a consequence of lateral transfer between *M. xanthus*, or a related ancestor, and organisms at or preceding the root of the protozoa/metazoa branching[16]. Thus, the tree topology for crystallin/protein S presents some similarities to that for Kua/CarF. A significant difference between these two cases resides in the absence of protein S or crystallin-related sequences in most non-animal eukaryotes, such as plants, which difficults evolutionary and phylogenetic analyses and inferences. In both instances, the proteins in eukaryotes with the closest similarity to the corresponding proteins in *M. xanthus* are those present in two closely related organisms, *Dictyostelium discoideum* for Kua and *M. xanthus* CarF (Fig. 2B), and *Physarium polycephalum* for the crystallin-related protein spherulin 3A (ref. 15), and *M. xanthus* protein S (Fig. 2D).

The most parsimonious explanation suggested by our phylogenetic analysis of Kua/CarF and βγ-crystallin/protein S is the occurrence of a gene transfer event from an organism at the root of the protist/metazoan branching, possibly related to Mycetozoa, into a bacterium related to *M. xanthus*. A strong argument for the direction of this transfer is provided by the absence in the available databases of genes related to Kua/CarF or to crystallin/protein S in prokaryotes other than *Myxococcus xanthus*. An origin in prokaryotes and posterior transfer into eukaryotes would require subsequent multiple losses, in both prokaryote and eukaryote lineages. Horizontal gene transfers are inferred from non-congruent phylogenies, that is, the observation of topologies that are in disagreement with organismal trees generated using multiple genes and proteins[17]. In the case of interkingdom or interclade transfers, this implies the occurrence of significantly larger number of sequences with similarities to the transferred gene in the kingdom of origin of the gene, and significantly fewer sequences with strong similarity in the kingdom or clade of the recipient organisms. This criterion has been used to infer the transfer of genes from eukaryotes into the bacterium *Mycobacterium tuberculosis*[18]. In the case of Kua/CarF, the application of this criterion leaves a eukaryote-to-bacterium transfer as the most plausible hypothesis, since *M. xanthus* CarF is the sole gene of its class in prokaryotes, while it is represented in most eukaryotes, a notable exception being fungi.

Other criteria often used to support cases of HGT between prokaryotes are the analysis of GC content[8], the search for biases in codon usage[9], or nucleotide frequencies in the first two positions of codons[19]. None of these three parameters differ significantly in the CarF gene as compared to the sum of the available *M. xanthus* gene sequences (data not shown). This suggests that the the transfer of Kua between eukaryotes and prokaryotes

must have occurred sufficiently back in time to allow the adaptation of the transferred gene to its recipient organism[9].

All eukaryotic Kua proteins are predicted to have two two-pass transmembrane domains, characteristic of many proteins that reside in the endoplasmic reticulum[1]. They also contain several juxtamembrane histidine-rich sequence motifs reminiscent of similar motifs found in fatty acid hydoxylases and dehydrogenases[1,20]. *M. xanthus* CarF maintains the same general domain topology and histidine-rich motifs (Fig 3A), and therefore it would be predicted to localize to endomembranes if expressed in eukaryotic cells. Indeed, a GFP-CarF fusion protein expressed in Cos-7 cells (Fig 3C) localized to cytoplasmic structures, with the exclusion of nuclei, and co-localized with a human Kua protein bearing a HA epitope in a subpopulation of cells (Fig 3B, top), although these two proteins to not appear to co-localize in a subpopulations of transfected cells (Fig 3B, bottom). *M. xanthus* CarF has a carboxyterminal tail that is absent in human Kua (Fig 1), a difference that could account for the partial differences in subcellular localization of both proteins in mammalian cells. What remains remarkable is the colocalization in most transfected cells of these proteins from organisms as distant as a bacteria and humans. Therefore, the conservation in sequence between *M. xanthus* CarF and human Kua translates into conservation in signals for subcellular localization, such as the putative two-pass transmembrane domains.

Based on these observations, we predict a conserved biochemical function for these two proteins. Epistasis analysis for CarF places it upstream of CarR[6]. The latter is a membrane-associated protein[6,7], that dissappears after exposure of bacteria to blue light, releasing the transcriptional co-activator CarQ[6]. The localization of CarF in membranes in mammalian cells and, presumably, also in bacteria, leaves open the possibility of a physical or tight functional interaction between CarF and CarR. For instance, blue light could induce

the CarF-mediated degradation of CarR. The juxtamembrane histidine-rich domains of CarF and Kua resemble similar motifs in other proteins that co-ordinate transition metal atoms, such as iron, zinc or nickel[21]. The electronic state of the coordinated metal atoms can shift with changes in the redox conditions in the immediate vicinity[22], and this is translated into subtle but significant variations in the strength and number of co-ordinations with the relevant histidine residues on the protein, which in turn results in detectable changes in the function of the protein[23]. In *M. xanthus*, blue light could be the inducer of changes in the intracellular redox status that would activate the anti-CarR function of CarF. If, as suggested by our observations, similar general functions are conserved for Kua, we predict that this protein could also serve as a redox sensor for a variety of stimuli. We have not found proteins with significant similarity to CarR in eukaryotes, and thus the potential targets of redox regulation by Kua in eukaryotes remain to be identified.

In the developing mouse, expression of the *Kua* gene follows a specific spatial and temporal pattern (Fig. 4). At day 9.5 of mouse development, the expression of *Kua* is mainly restricted to the central nervous system (Fig. 4A). The staining is strongest in the neuroepithelium lining the neuropore, especially at the ventricles and the telencephalic and optic vesicles (Fig. 4D). This pattern suggests an association of high levels of expression of this gene with proliferating neuronal precursors. By days 11.5 and 13.5 pc, expression of *Kua*, while maintaining a high level in the central nervous system, extends to non-neuronal structures, including liver, lung and gut (Fig. 4B, C). At these stages, expression in neuronal structures is not limited to subventricular and subvesicular layers in the CNS. Thus, dorsal root ganglia show a strong staining, and, in the neural tube, groups of dorsal and ventral neurons express *Kua* (not shown). Also, expression is seen in postreplicative and postmigratory compartments that form the neopallial cortex (Fig. 4B-D). Of particular

interest are the temporal and spacial patterns of expression that *Kua* follow in the developing eye: as mentioned above, by day 9.5 pc, the gene is expressed in the proliferating cell layer lining the optic vesicle and cup (Fig. 4D); by day 11.5 pc, *Kua* is strongly expressed in the entire neural retina and in primary lens fibers (Fig. 4E); by day 15.5 pc, expression of this gene in the neural retina appears to be limited to peripheral cells, as well as differentiated ganglion cells and photoreceptors (Fig. 4F). At this stage, expression of *Kua* in the lens is limited to the germinal epithelium and the equatorial zone (Fig. 4F). Thus, while expression of *Kua* in the early stages of eye development in the mouse is associated with proliferating cellular compartments, expression at later stages occurs also in non-proliferating, differentiated cells.

In *D. melanogaster*, *Kua* is expressed in the earliest stages of development, indicating a maternal origin of the transcripts (Fig. 5A). The expression is ubiquitous, without evidence for polarization or segmentation, and continues to be obvious throughout the gastrula stage (Fig. 5C). In third-instar larvae, the staining with the *Kua* probes is significantly weaker than that for embryos, indicating a generalized decrease in *Kua* transcript levels. At this stage, the expression of *Kua* adopts a more restricted pattern, with weak but detectable expression in imaginal discs (Fig. 5D, E), but not in brain (not shown).

Thus, despite the strongly conserved sequences and possible cellular or biochemical functions of Kua, the comparative analysis of the expression of the Kua genes in the mouse and in *Drosophila* shows that, in the early developmental stages of these two very different organisms, they are not likely to be involved in developmental decisions affecting orthologous structures. This appears to be particularly evident in the restricted expression of *Kua* in central nervous system structures in the early development of the mouse, which contrasts with a more generalized pattern of expression in *Drosophila* embryos and

imaginal discs, but not brain. The lack of polarization, compartmentalization or segmentation in its expression patterns suggests that its function is not related to the organization of the body plan of *Drosophila*, but rather to more basic or general cellular processes.

Based on our analysis, we speculate that the donor eukaryote in the horizontal transfer of Kua to Myxobacteria was an organims related to slime molds. *D. dictyostelium* and *P. polycephalum* belong to the Mycetozoa and, together with the protostelid slime molds, form a monophyletic group that has been placed at the "crown" of the eukaryote tree[24,25]. Both *M. xanthus* and *D. discoideum* grow in overlapping ecological niches and present a "social" behaviour, with the development of fruiting bodies consisting of vegetative stalks and spores, under conditions of nutrient restriction[2]. In both cases, the morphological differentiation into two populations requires the use of highly regulated signaling mechanisms that modulate the expression of subpopulation-specific sets of genes[26]. The selective advantage for the acquisition and retention of this gene and protein in the recipient organism must be discussed in the light of the known and predicted functions of CarF and Kua. In *Myxococcus xanthus*, CarF functions as a light sensor that triggers the synthesis of carotenoids, which protect the bacterium from reactive oxygen species. To date, a sensor similar to CarF, or a regulatory pathway as clearly defined as that represented by *M. xanthus* CarQRS and CarD, has not been described in eukaryotes. In certain multicellular eukaryotes, such as plants, the biosynthesis of carotenoids is regulated by oxidative stress[22]. The striking conservation observed here between Kua and CarF suggests that this protein could be a general redox sensor in eukaryotes as well as in prokaryotes, which was acquired by Myxobacteria as a part of a tightly regulated survival mechanism against oxidative stress. Protein S, also possibly acquired by *Myxococcus* by horizontal

transfer from eukaryotes, is part of a protective mechanism against stress both in this bacterium and in *P. polycephalum*[15]. Horizontal gene transfer between prokaryotes has been shown to constitute a major contributor to the evolution and diversification of microorganisms[27]. Evidences for horizontal transfer from eukaryotes, such as those shown here, indicate that microorganisms can acquire virtually any genes from other organisms, provided that such genes confer competitive or survival advantages to the recipient organims[28], and also show that any organism can be a donor of genes in these transfers, without restrictions of kingdom, phylum, or clade. The driving force for the acquisition and retention of such genes and traits could reside either on symbiotic associations with mutually beneficial metabolic interchanges[29], or in a one-way take-over with obvious benefits only for the acquiring organism.

Methods

Sequence alignments and tree construction

General and organism-specific sequence data banks were searched for Kua orthologs with BLASTP, TBLASTN and PSI-BLAST. The retrieved sequences were aligned with ClustalW, using the longest available sequences for each entry. For subsequent analysis, the longest segments that showed sound alignments for all entries were selected. and used for tree construction by the neighbour joining method (Philip package). Computed trees were then subjected to bootstrapping.

Expression constructs and transient transfection

Full-length human *Kua* cDNA was cloned into pcDNA3.1 (Invitrogen) bearing sequences coding for the hemagluttinin epitope (HA), such that this sequence was placed in-frame at the carboxy terminus of cDNA. Full-length *M. xanthus* CarF was subcloned

into pEGFP (Clontech). For transient transfections, 1 μg of endotoxin-free plasmid was transfected with Lipofectamine Plus (Life Sciences) into Cos-7 cells grown on glass coverslips. As controls, the corresponding insert-less vectors were used. Twenty-four hours after transfection, cells were washed in cold PBS, and fixed in 4% paraformaldehyde/PBS. In the case of fluorescent chimeric proteins, cells were directly mounted in Immuno-fluor (ICN). For transfections with vectors expressing the HA epitope, cells were permeabilized with 1% saponin/2% BSA/PBS, incubated for 1h with rat monoclonal anti-HA antibody (Boehringer-Mannheim), washed, and further incubated for 1h with FITC-conjugated anti-rat Ig (Dako) for Kua-HA transfections, or with TRITC-conjugated anti-rat Ig for CarF and Kua-HA double transfections. Cells were observed under a Leica confocal microscope (Wetzlar, Germany).

For inmunoblotting, 4 μg of endotoxin-free plasmid were transfected per 100-mm plate with Cos-7 cells growing at 70-80% confluence. Twenty-four hours after transfection, cells were scraped, washed and directely lysed in Laemmli loading buffer, and 50 μg of total protein per sample separated by 12% SDS-PAGE. After electrophoresis, proteins were transferred to nitrocellulose membranes, and processed according to standard methods for Western blotting[30], using antibodies to HA (Boehringer-Mannheim) or to GFP (Clontech). Secondary antibodies were anti-mouse or anti-rat Ig conjugated to horseradish peroxidase (Sigma), and reactivities were revealed by chemoluminiscence (Amersham).

Whole mount *in situ* hybridization of *Drosophila* embryos

D. melanogaster Canton S strain embryos and larvae were processed according to previously described procedures[31,32]. Briefly, dechorionated embryos were rinsed in methanol and postfixed in 5% formaldehyde in PBS/0.1% Tween-20 for 20 min, and

washed 5 times, 5 min each, with PBS/Tween-20. Third-instar larvae were dissected in PBS and the imaginal discs and brain fixed in 4% paraformaldehyde in PBS for 20 min, washed in PBS and postfixed in 4% paraformaldehyde in PBS, 0.1% Triton X-100 and 0.1% deoxycholate for 20 min. Specimens were dehydrated and incubated in xylene for 3 h and rinsed in ethanol and kept in methanol at -20ºC overnight. Digoxigenin-labeled (Boehringer-Mannheim) sense and anti-sense riboprobes were generated by transcription with SP6 or T7 RNA polymerase from linearlized pGEM-T/DmKua. Hybridization was conducted at 65ºC in a 50% formamide solution. Samples were incubated with alkaline phosphatase-conjugated anti-digoxigenin antibody (Boehringer-Mannheim), washed and developed with NBT/BCIP as a substrate.

In situ hybridization on mouse embryos

Mouse embryos were fixed in RNAse-free 4% paraformaldehyde solution at 4ºC overnight, sequentially dehydrated through graded ethanols, embedded in paraffin, and cut at 5 μM thickness. Mouse testis RNA was used in RT-PCR reactions with primers K4 and K5 for *Kua*. Amplification products were cloned in pGEM-T (Promega) and sequenced. Digoxigenin-labeled sense and anti-sense RNA probes were synthesized from these templates, using a digoxinenin RNA labeling (SP6/T7) *in vitro* transcription kit (Boehringer-Mannheim). Hybridization was performed as previously described[33]. After hybridization and washing, samples were incubated with alkaline phosphatase-conjugated anti-digoxigenin antibody (Boehringer Mannheim) and developed with NBT/BCIP. All slides were mounted with Crystal/Mount (Biomeda Corp, Foster City, CA). Whole-mount *in situ* hybridization and posthybridization sectioning were performed as described[34].

Figure legends

Figure 1. Alignment of *Myxococcus xanthus* CarF with Kua proteins from organisms for which probable full-length sequences were retrieved from databases. MxCarF, *Myxococcus xanthus* carF (acc. g13397952), HsKua, *Homo sapiens* Kua (acc. g6448867), DmKua, *Drosophila melanogaster* Kua (acc. g6002805), CeKua, *Caenorhabditis elegans* Kua (acc. g7510148), LmKua, *Leishmania major* Kua (acc. g11061657), AtKua1, At Kua2, AtKua3, *Arabidopsis thaliana* Kua (acc. g3445206, g7485344, and g2160143, respectively).

Figure 2. (A) Alignment for a segment conserved between Kua proteins from all organisms for which homologous sequences were retrieved from the databases. The alignment includes all the sequences in Figure 1, with the addition of sequences from *Dictyostelium discoideum* (acc. JC1c85g09), *Xenopus laevis* (acc. AW643048), *Trypanosoma cruzi* (acc. AI652134), *Glycine max* (acc. BF424782) and *Mycebrotum crystallinum* (TIGR TC364). (B) Neighbor-joining tree with bootstrap test generated from the alignment in (A). (D) Neighbor-joining tree for *M. xanthus* Pph1 and selected orthologues in other prokaryotes and eukaryotes[12]. Only alignable segments from all sequences were selected to generate the tree. (D) Neighbor-joining tree with bootstrap test for *M. xanthus* protein S1 and protein S2 and selected crystallins with highest similarities to the *M. xanthus* proteins[13,15].

Figure 3. (A) Predicted transmembrane topology for human Kua and Myxococcus xanthus CarF, generated with TMpred (http://www.ch.embnet.org/software/TMPRED_form.html). (B) Flourescent detection of Kua-HA and GFP-CarF co-transfected into Cos-7 cells, as indicated. For Kua-HA, an indirect immunofluorescence procedure was followed, using a

TRITC-conjugated secondary antibody for detection through the red light emission channel. GFP-CarF emission was detected through the green light emission channel. (C) Western blotting of Cos-7 cells transfected with Kua-HA (lane 1), GFP-CarF (lane 2) or GFP vector alone (lane 3), and probed with antibodies to HA (lane 1) or to GFP (lanes 2 and 3).

Figure 4. *In situ* hybridization of mouse embryos with *M. musculus* Kua-specific riboprobes. (A) Day 9.5 pc embryos. (B) Day 13.5 pc embryos. (C) Day 15.5 pc embryos. (D) Coronal section at the forebrain of a day 9.5 pc embryo. (E) Section through the developing eye of a day 13.5 pc embryo. (F) Section through the developing eye of a day 15.5 pc embryo.

Figure 5. *In situ* hybridization of *Drosophila* embryos (A-C) and wing (D) and eye (E) imaginal discs from third-instar larvae, with *D. melanogaster* Kua-specific riboprobes. (A) Syncitial blastoderm. (B) Cellular blastoderm. (C) Gastrula.

References

1. Thomson, T.M. et al. Fusion of the human gene for the polyubiquitination coeffector UEV1 with Kua, a newly identified gene. *Genome Res.* **10,** 1743-1756 (2000).
2. Kaiser, D. Control of multicellular development: *Dictyostelium* and *Myxococcus*. *Annu. Rev. Genet.* **20,** 539-566 (1986).
3. Burchard, R.P., & Dworkin, M. Light-induced lysis and carotenogenesis in *Myxococcus xanthus*. *J. Bacteriol* **91,** 535-545 (1966).
4. Burchard, R.P., & Hendricks, S.B Action spectrum for carotenogenesis in *Myxococcus xanthus*. *J. Bacteriol* **97,** 1165-1168 (1969).
5. Nicolas, F.J., Ruiz-Vazquez, R.M., & Murillo, F.J A genetic link between light response and multicellular development in the bacterium *Myxococcus xanthus*. *Genes Dev.* **8,** 2375-2387 (1994).
6. Gorham, H.C., McGowan, S.J., Robson, P.R. & Hodgson D.A Light-induced carotenogenesis in *Myxococcus xanthus*: light-dependent membrane sequestration of ECF sigma factor CarQ by anti-sigma factor CarR. *Mol. Microbiol* **19,** 171-186 (1996).
7. Hodgson, D.A. Light-induced carotenogenesis in *Myxococcus xanthus*: genetic analysis of the carR region. *Mol. Microbiol* **7,** 471-488 (1993).
8. Lawrence, J.G., & Ochman, H. Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. USA*. **95,** 9413-9417(1998).
9. Lawrence, J.G., & Ochman, H. Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* **44,** 383-397 (1997).

10. Muñoz-Dorado, J., Inouye, S. & Inouye, M. Eukaryotic-like protein serine/threonine kinases in *Myxococcus xanthus*, a developmental bacterium exhibiting social behaviour. *J. Cell. Biochem.* **551,** 29-33 (1993).

11. Inouye, S., *et al*. A large family of eukaryotic-like protein Ser/Thr kinases of *Myxococcus xanthus*, a developmental bacterium. *Microb. Comp. Genomics* **5,** 103-120 (2000).

12. Treuner-Lange, A., Ward, M.J, & Zusman D.R. Pph1 from *Myxococcus xanthus* is a protein phosphatase involved in vegetative growth and development. *Mol. Microbiol.* **40,** 126-140 (2001).

13. Bagby, S., Havery, T.S., Eagle, S.G., Inouye, S., & Ikura, M. Structural similarity of a developmentally regulated bacterial spore coat protein to βγ-crystallins of the vertebrate eye lens. *Proc. Natl. Acad. Sci. USA* **91,** 4308-4312 (1994).

14. Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., & Doolittle, W.F. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* **290,** 972-977 (2000).

15. Wistow, G. Evolution of a protein superfamily: relationships between vertebrate lens crystallins and microorganism dormancy proteins. *J. Mol. Evol.* **30,** 140-145 (1990).

16. Doolittle, W.F. Phylogenetic classification and the universal tree. *Science* **284,** 2124-2128 (1999).

17. Stanhope, M.J., et al. Phylogenetic analyses do not support horizontal gene transfers from bacteris to vertebrates. *Nature* **411,** 940-944 (2001).

18. Gamieldien, J., Ptitsyn, A., & Hide, W. Eukaryotic genes in *Mycobacterium tuberculosis* could have a role in pathogenesis and immunomodulation. *Trends Genet.* **18,** 5-8 (2002).

19. Gautier, C. Compositional bias in DNA. *Curr. Opin. Genet. Dev.* **10,** 656-661 (2000).

20. Shanklin, J., Whittle, E., & Fox, B.G. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. *Biochemistry* **33**, 12686-12694 (1994).

21. Fox, B. G., Shanklin, J., Somerville, C. R., & Munck, E. Stearoyl-acyl carrier protein Δ^9 desaturase from *Ricinus communis* is a diiron-oxo protein. *Proc. Natl. Acad. Sci. USA* **90**, 2486-2490 (1993).

22. Bouvier, F., Backhaus, R.A., & Camara, B. Induction and control of chromoplast-specific carotenoid genes by oxidative stress. *J. Biol. Chem.* **273,** 30651-30659 (1998).

23. Shanklin, J., Achim, C., Schmidt, H., Fox, B.G., & Munck, E. Mossbauer studies of alkane omega-hydroxylase: evidence for a diiron cluster in an integral-membrane enzyme. *Proc. Natl. Acad. Sci. USA* **94**, 2981-2986 (1997).

24. Baldauf, S.F., & Doolittle, W.F. Origin and evolution of the slime molds (Mycetozoa). *Proc. Natl. Acad. Sci. USA* **94,** 12007-12012 (1997).

25. Baldauf, S.L., & Palmer, J.D. Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. *Proc. Natl. Acad. Sci. USA*. **90,** 11558-11562 (1993).

26. Kaiser, D. Cell fate and organogenesis in bacteria. *Trends Genet*. **15,** 273-277(1999).

27. Ochman, H., Lawrence, J.G.,& Groisman, E.A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405,** 299-304 (2000).

28. Moreira, D., López-García, P. Symbiosis between methanogenic archaea and δ-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J. Mol. Evol.* **47,** 517-530 (1998).

29. Doolittle, W.F. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* **14,** 307-311 (1998).

30. Harlow, E., & Lane, D. Antibodies. A laboratory manual. Cold Spring Harbor Laboratory (1988).

31. Tautz, D., & Pfeifle, C. A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* **98,** 81-85 (1989).

32. Lehner, C.F., & O'Farrell, P.H. *Drosophila* cdc2 homologs: a functional homolog is coexpressed with a cognate variant. *EMBO J.* **9,** 3573-3581 (1990).

33. Carrió, et al., A1 demonstrates restricted tissue distribution during embryonic development and functions to protect against cell death. *Am. J. Pathol.* **149,** 2133-2142 (1996).

34. Nieto, M.A., Patel, K., & Wilkinson, D.G. In situ hybridization analysis of chick embryos in whole mount and tissue sections. *Methods Cell. Biol.* **51,** 219-235 (1996).

Figure 1

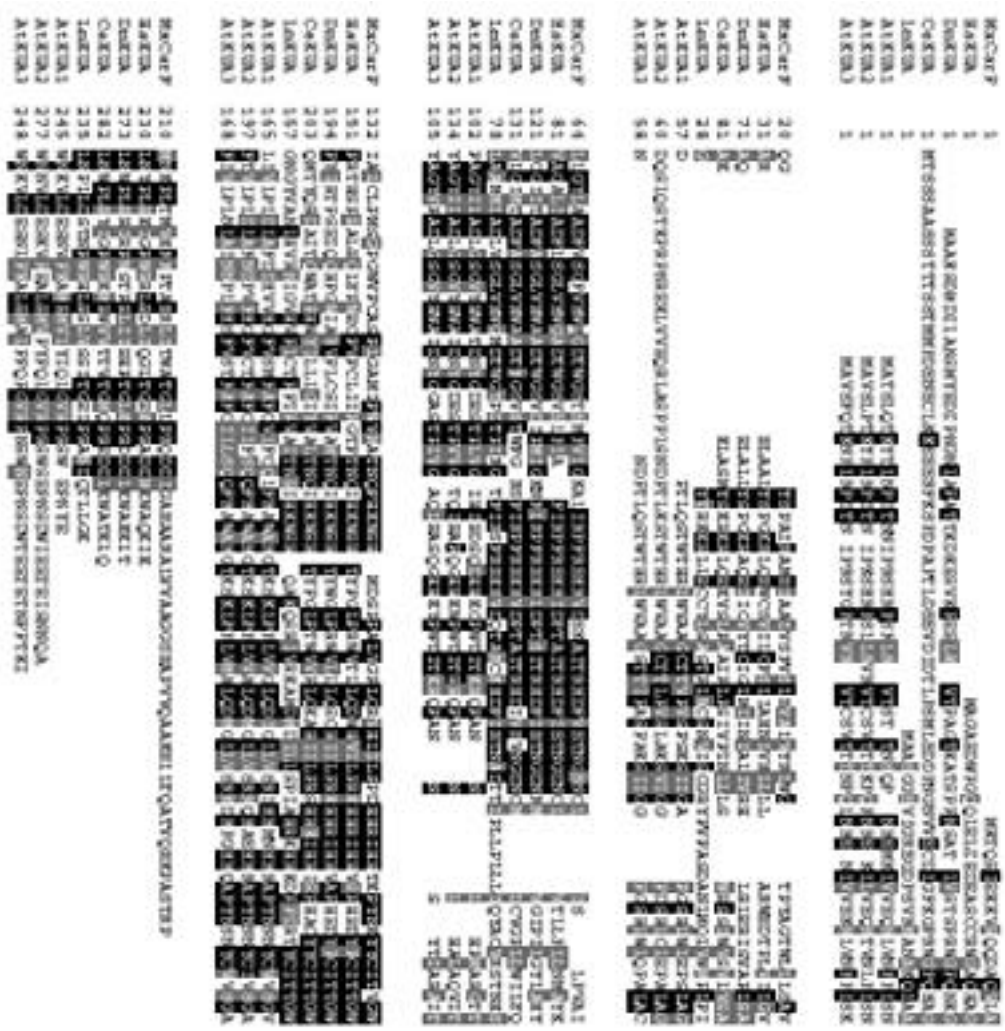

Figure 2

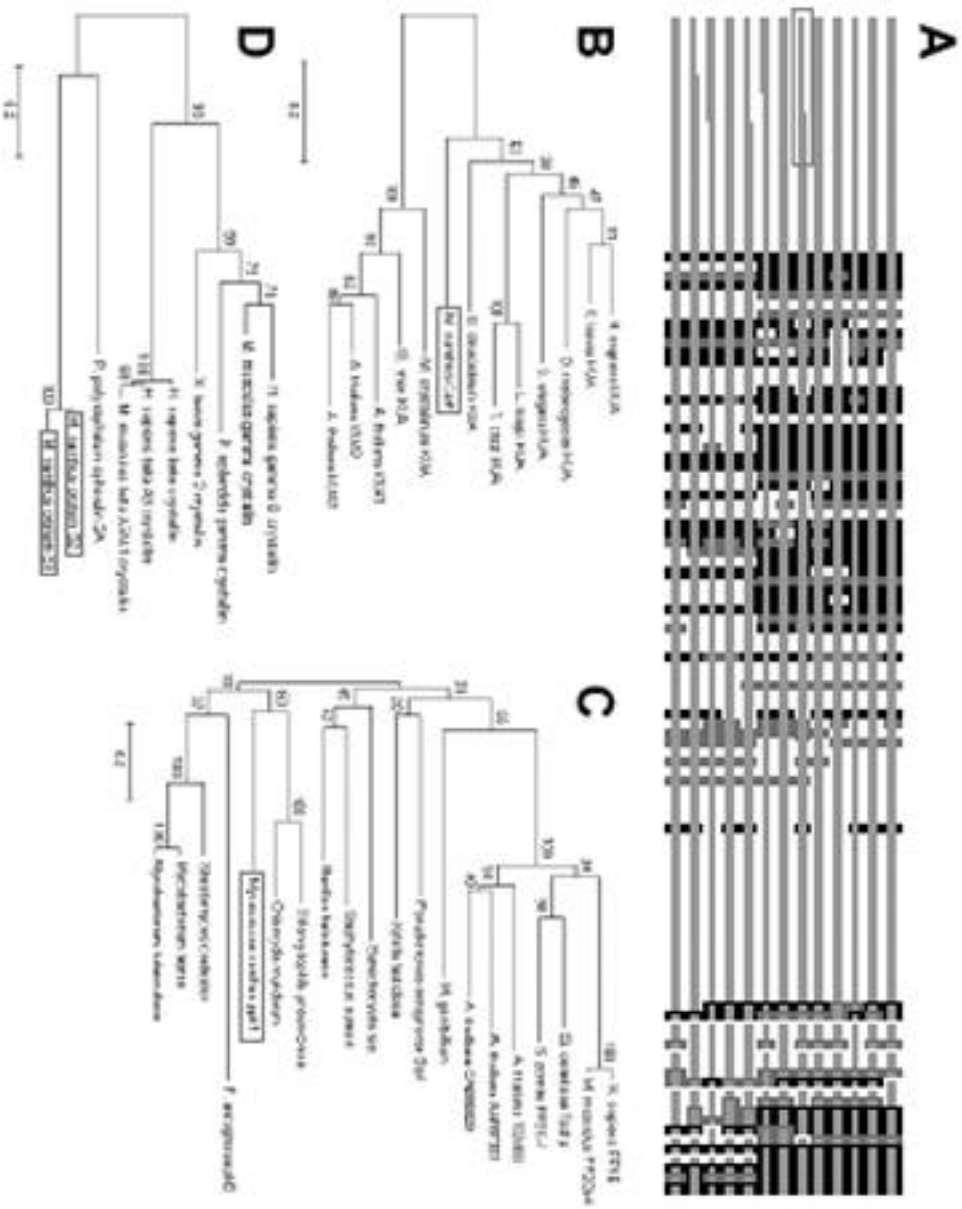

Figure 3

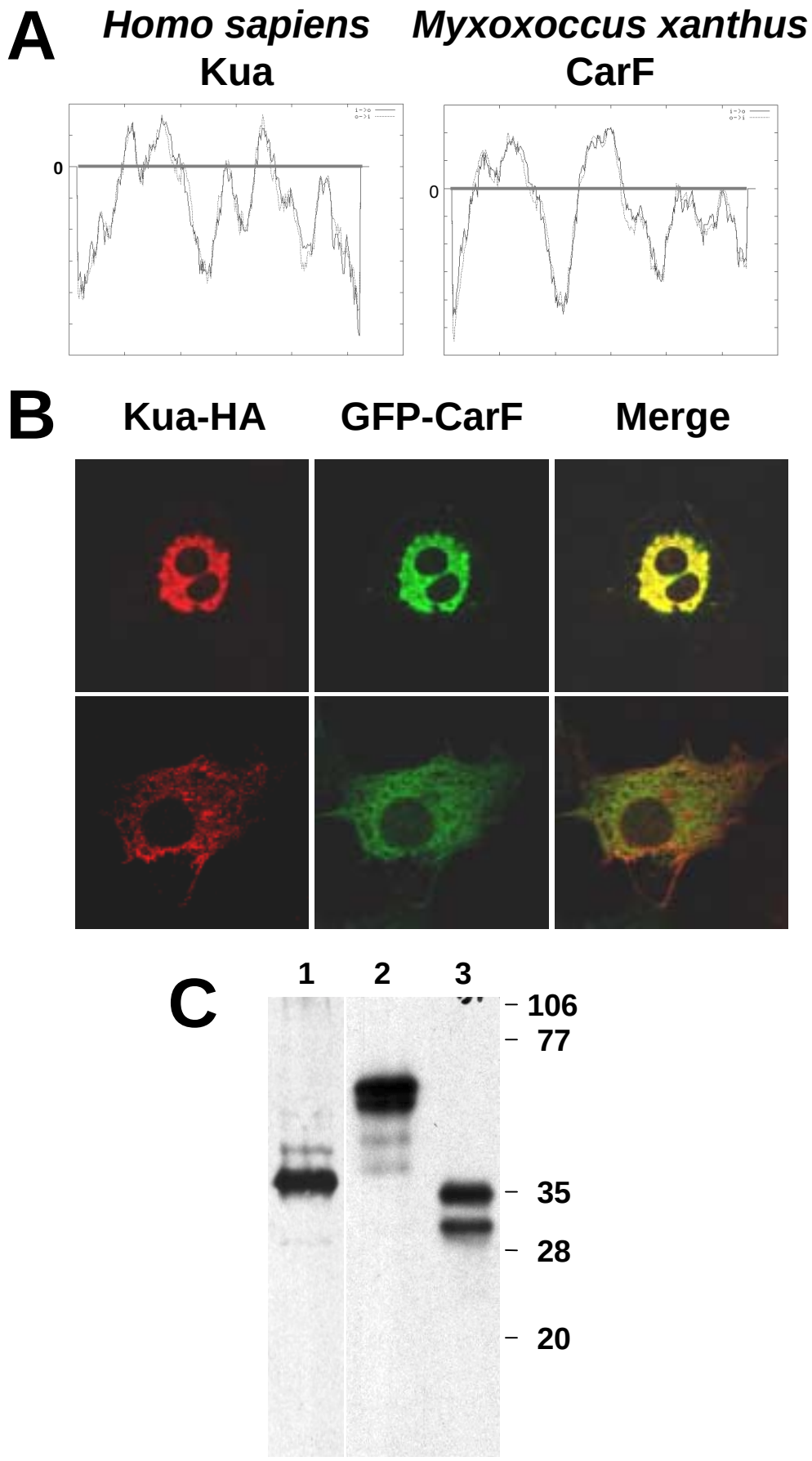
A
Homo sapiens
Kua
Myxoxoccus xanthus
CarF
0
0
B
Kua-HA
GFP-CarF
Merge
C
1
2
3
106
77
35
28
20

Figure 4

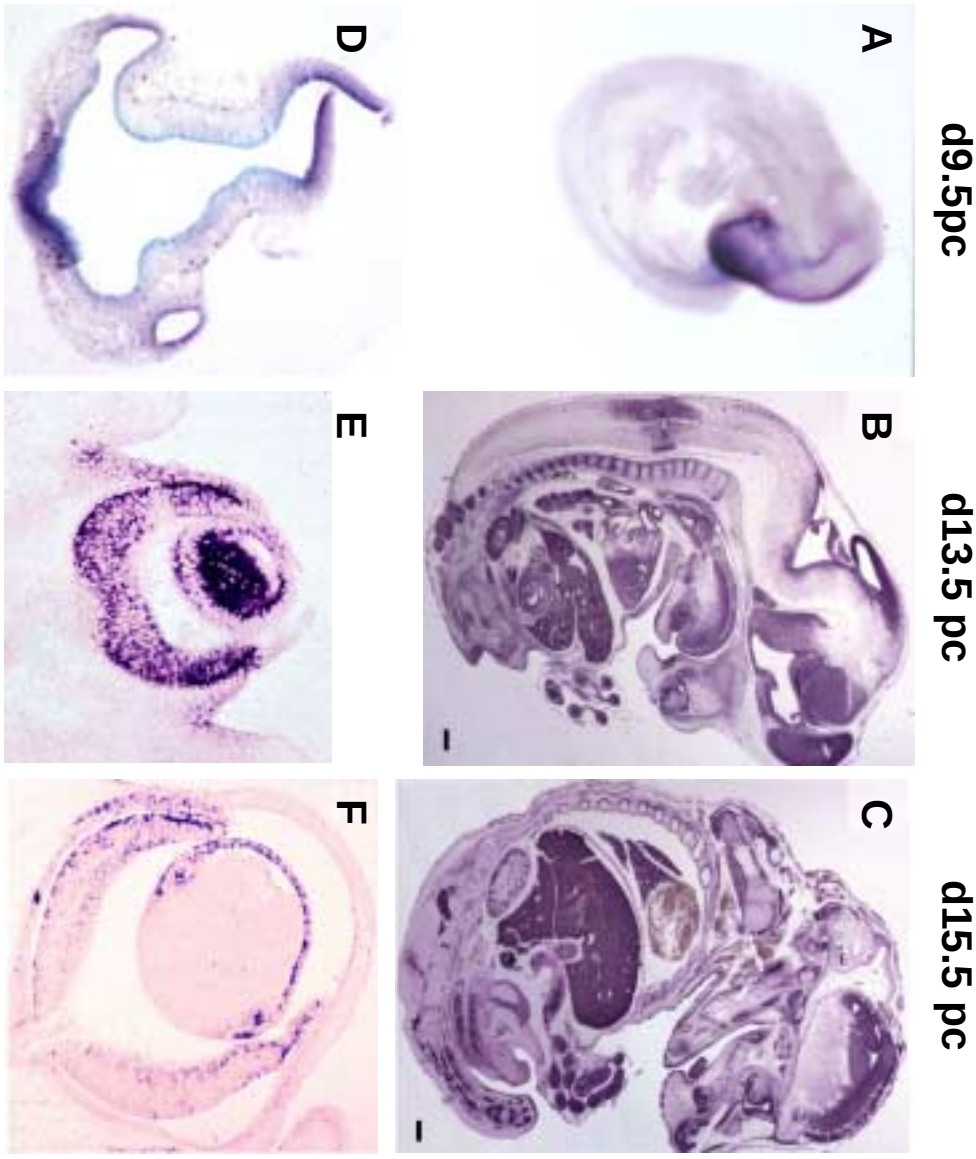

Figure 5