Evidence for the derivation of the *Drosophila fushi tarazu* gene from a Hox gene orthologous to lophotrochozoan *Lox5*

Maximilian J. Telford

The DNA-binding homeobox motif was first identified in several *Drosophila* homeotic genes but also in *fushi tarazu*, a gene found in the Hox cluster yet involved in segmentation, not anteroposterior patterning [1]. Homeotic transformations are not seen in insect *ftz* mutants, and insect *ftz* genes do not have Hox-like expression except within the nervous system [2,3]. Insect *ftz* homeobox sequences link them to the *Antp*-class genes and *Tribolium* and *Schistocerca* orthologs have *Antp*-class YPWM motifs amino-terminal to the homeobox [2,3]. Orthologs of *ftz* cloned from a centipede and an onychophoran [4] show that it predates the emergence of the arthropods, but the inability to pinpoint non-arthropodan orthologs suggested that *ftz* is the product of a Hox gene duplication in the arthropod ancestor [4,5]. I have cloned *ftz* orthologs from a mite and a tardigrade, arthropod outgroups of the insects [6]. Mite *ftz* is expressed in a Hox-like pattern, confirming its ancestral role in anteroposterior patterning. Phylogenetic analyses indicate that arthropod *ftz* genes are orthologous to the *Lox5* genes of lophotrochozoans

deuterostomes and would therefore have been present in the triploblast ancestor.

Address: Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.

Correspondence: Maximilian J. Telford
E-mail: M.Telford@nhm.ac.uk

Received: **21 October 1999**
Revised: **12 January 2000**
Accepted: **26 January 2000**

Published: **10 March 2000**

Current Biology 2000, 10:349–352

0960-9822/00/$ – see front matter

Results and discussion

Alftz and *Mtftz* are orthologs of insect *ftz* genes

Phylogenetic analyses of the homeodomain securely unite the mite (*Archegozetes longisetosus*; *Alftz*) and tardigrade (*Milnesium tardigradum*, *Mtftz*) *ftz* genes with the *ftz* genes of insects as well as those of a crustacean, a centipede and an onychophoran (data not shown). All arthropod and onychophoran *ftz* sequences are grouped with bootstrap supports of 74% (neighbor joining; NJ) and 51% (maximum parsimony; MP) and 66% support from Puzzle Maximum Likelihood (PML). Striking amino acid identities are seen in the hexapeptide consensus F(F/Y)PWM(K/R)SYTD (in the single-letter amino acid code) in onychophora, tardigrades, mites and centipedes. In further support of this orthology relationship, *Alftz* was found adjacent to, upstream of and in the same orientation as the *Archegozetes Sex-combs-reduced* ortholog (*AlScr*) in genomic library lambda clones in precisely the relative position and orientation inferred for the ancestral insect *ftz* homolog [8] (data not shown).

The insect and crustacean *ftz* genes are particularly divergent, as can be inferred through outgroup comparison. Chelicerates and centipedes are phylogenetically closer to the insects than to the onychophoran, yet their *ftz* homeodomain is considerably closer in sequence to onychophoran and tardigrade *ftz* genes. This demonstrates that the onychophoran, tardigrade, chelicerate and centipede genes are closest to the ancestral arthropod *ftz* sequence and that the differences between them and the crustacean/insect lineage are due to changes in the latter. In subsequent phylogenetic analyses, the complete chelicerate *ftz* ortholog is used as representative of arthropod *ftz* genes in general.

The primitive anterior boundary of arthropod *ftz* expression is in the maxillary segment

Earliest *ftz* expression in insects is in a broad band throughout the blastoderm. The anterior-most early *ftz* expression in the beetle *Tribolium* lies at the front of the maxillary segment, just behind the mandibular *engrailed* stripe (in register with the segments rather than the parasegments), but later fades as expression resolves into a pair-rule pattern [2]. *Drosophila ftz* is also expressed early on in a broad domain in the syncytial blastoderm, again resolving into seven stripes marking the front of even-numbered parasegments [9]. Unlike in *Tribolium*, the early anterior boundary of *Drosophila ftz* is in the posterior of the maxillary segment (the parasegment 1/2 (PS1/2) boundary); it seems, however, that there would be an exact correspondence between *Tribolium* and *Drosophila* anterior boundaries of *ftz* expression at the front of the maxillary segment if *Drosophila ftz* were not repressed in the front of the maxilla by *evenskipped* (*eve*) expression [9]. *Drosophila ftz* expands anteriorly into this domain in *eve*$^-$ embryos. This repression by *eve* of *Drosophila ftz* is likely to be a derived feature, as *eve* and

ftz are co-expressed in *Schistocerca* and to some degree in *Tribolium*. There are no segmental markers to identify the earliest anterior boundary of expression of *Sgdax*, which has been suggested to be a *Schistocerca ftz* ortholog [3]. In all of the insects, the broad early domain disappears very early and long before there is any evidence of appendages.

It seems clear that insect *ftz* is derived from a homeotic gene and I suggest that the earliest expression of *ftz* in insects indicates its primitive expression domain with a boundary at the anterior of the maxillary segment. This suggestion is supported by *in situ* hybridizations using *Alftz* probes on *Archegozetes* embryos, which show that *Alftz* is expressed in a typically Hox-like pattern in all stages examined, with a sharp anterior boundary at the front of the second leg segment (Figure 1). This segment has been shown to be homologous with the insect maxilla [10,11]. The posterior boundary is at the rear of the fourth leg bud (Figure 1). Unlike the early broad insect *ftz* domain, expression in the insect nervous system is slightly more anterior and in parasegmental register (its anterior border is at the PS0/1 boundary) in the posterior of the mandibular segment. This discrete parasegmental anterior boundary is similar to the nervous system expression of the canonical Hox genes [12] and is conserved in all insects studied. Nervous system expression has not been studied in *Archegozetes*.

A possible explanation for the change in *ftz* function in the insects

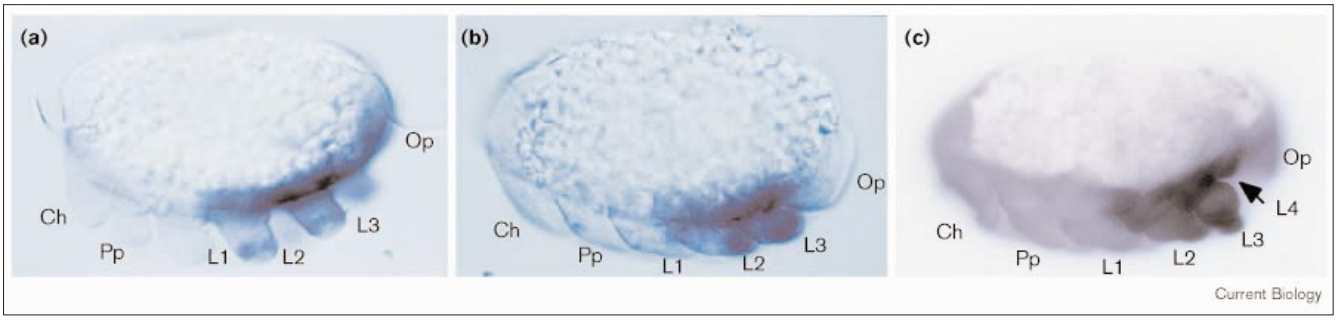
Alftz expression overlaps almost exactly with *AlScr* expression [10] in the fourth appendage-bearing segment (the mite second leg, which is equivalent to the insect/crustacean second maxilla). This anterior boundary of *Scr* expression seems to be primitive for the arthropods, as it is seen almost identically in the isopod crustacean *Porcellio* [13], although it is expressed only in the posterior of the maxilla in insects. It seems plausible that *ftz* has lost its role as a homeotic gene in the ancestor of the insects as a result of redundancy of function following overlap of its expression domain with that of *Scr*. Coincidence of anterior boundaries also correlates with a loss of homeotic function of the *Hox3/zen* homolog in arthropods [14,15]. Identical with *Hox3/zen* in insects, the loss of AP patterning function might have released the homeodomain from stabilizing selection and led to the rapid sequence divergence seen in the insect *ftz* genes; indeed in its new role as a pair-rule gene, *Drosophila ftz* can function even with its homeodomain almost entirely deleted [16].

Sequence analyses suggest that *ftz* orthologs are present in all protostomes and possibly all triploblasts

In further phylogenetic analyses, I compared arthropod Ftz homeodomain amino acid sequences with others of the closely related central-group Hox genes (those related to *Antp*, *Ubx* and *abdA*) from a range of other animals, as well as with *Hox4/Dfd* and *Hox5/Scr* orthologs to root the trees. Recent phylogenetic analyses of Hox genes have implied that arthropod *ftz* and *Antp* genes arose from one duplication and *Lox5*-like genes and the genes related to nemertean *LsHox7* arose from another separate duplication such that there is no ortholog of *ftz* outside the Arthropoda and Onychophora [5].

Phylogenetic analyses using the 60 amino acids of the homeodomain show that, in fact, arthropod *ftz* genes cluster with the *Lox5* genes of lophotrochozoa (molluscs, annelids, and relatives; Figures 2,3). Furthermore the ecdysozoan *Antp* genes (except *AlAntp* and the priapulid *Antp* orthologs) cluster with the lophotrochozoan *LsHox7* and *LaHB1* genes (Figure 3) There is no bootstrap support for this latter

Figure 1



In situ hybridizations of *Alftz* riboprobes to *Archegozetes* embryos. **(a)** Hybridization with homeobox-containing probe on a young embryo before leg extension. The anterior boundary of expression is at the front of the second leg. There is some staining in the opisthosoma. The fourth leg buds are stained but are out of the plane of focus. **(b)** Hybridization with the same probe on a later embryo after leg extension. The anterior boundary is the same but the opisthosomal expression has disappeared. The fourth leg buds are stained but are out of the plane of focus. **(c)** Hybridization with a non-homeobox-containing probe with relatively high non-specific background. The anterior boundary is identical. The fourth leg buds are visible in this image and are stained (arrow). Ch, chelicerae; Pp, pedipalps; L1–L4, legs 1–4; Op, opisthosoma. Embryos are approximately 170 μm long.

Figure 2

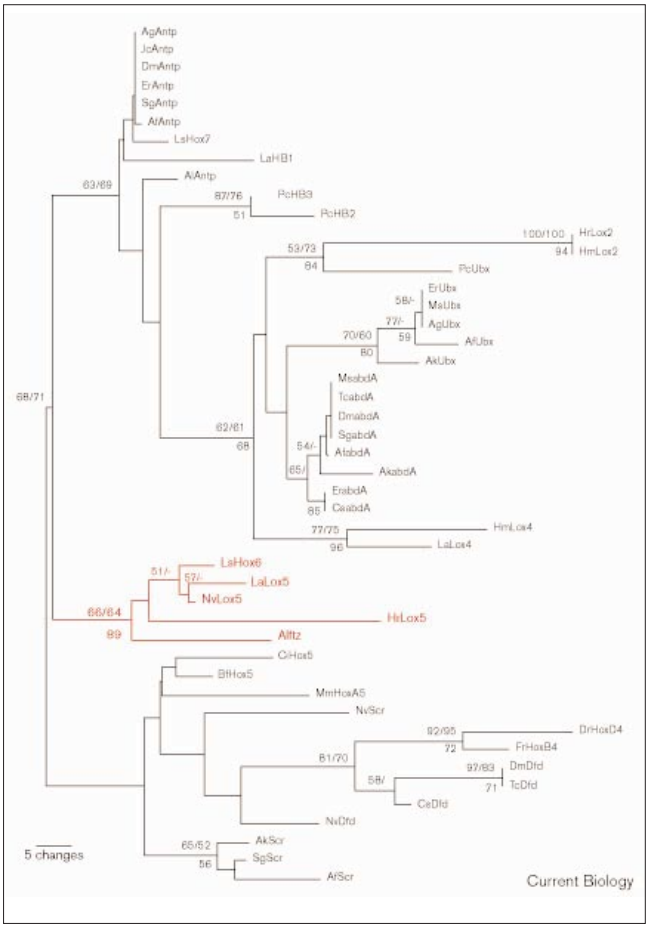
Alignment of Hox genes showing diagnostic residues. Sequences are grouped as follows. 1, *Dfd*; 2, *Scr*; 3, arthropod *ftz*; 4, lophotrochozoan *Lox5*; 5, *Antp*; 6, lophotrochozoan *Lox2*, *Lox4*; 7, ecdysozoan *Ubx*/*abdA*. Dark blue residues are inferred ancestral (plesiomorphic) states. Red or green residues are shared derived features (synapomorphies) linking two or more of the above groups. Light blue residues are specific to the *Lox5* genes or the *ftz* genes (autapomorphies). I equate the following amino acids, in the single-letter code: Y/F, aromatic; I/L/V/M, hydrophobic; K/R, positive; D/E, carboxylate/negative; S/T, hydroxyl/small. Noteworthy positions are in bold and are indicated by a lower case letter above: j and k support monophyly of *ftz*, *Lox5, Antp*-like and *UbdA*-like genes (groups 3–7). Although I equate S + T, the S in group 2 and the T in groups 3–7 at position k have been shown experimentally to be functionally distinct [17]. Positions c–f, i and l support monophyly of *ftz* plus *Lox5*. Further similarities at the carboxyl terminus are unreliable because of alignment uncertainty (positions q–u), although I have attempted to align *ftz*, *Lox5* and *Antp*-like genes. A uniquely positioned intron four residues amino-terminal to the homeodomain in HrLox5, LzLox5, NvLox5 and ErFtz and AkFtz gives further support to the orthology of *ftz* and *Lox5*. The ancestral state at position i may be T (some *Scr* and *Dfd* genes), making this state plesiomorphic. In this case, however, along with h, *Antp*-like and *UbdA*-like genes must be more derived than *ftz*/*Lox5*. It is clear that *UbdA*/*Lox2–4* genes were present in the ancestral protostome, so *ftz* and *Lox5* are yet more primitive and must therefore have existed in the ancestral protostome. To assume *ftz* and *Lox5* are not orthologs requires the unparsimonious assumption of additional duplications and gene losses (see Supplementary material). Genes are named with first two letters representing genus and species. Ecdysozoa: Ag, *Anopheles gambiae*; Er, *Ethmostigmus rubripes*; Ak, *Acanthokara kaputensis*; Tc, *Tribolium castaneum*; Af, *Artemia franciscana*; Al, *Archegozetes longisetosus*; Dm, *Drosophila melanogaster;* Sg, *Schistocerca gregaria*; Cs, *Cuppienius salei*; Pc, *Priapulus caudatus;* Mt, *Milnesium tardigradum*;. Sc, *Sacculina carcini*. Lophotrochozoa: Cv, *Chaetopterus variopedatus;* La, *Lingula anatina*; Nv, *Nereis virens*; Hm, *Hirudo medicinalis*; Hr, *Helobdella robusta*; Ls, *Lineus sanguineus*; Dj, *Dugesia japonica*; Dt, *Dugesia tigrina.* Deuterostomia: Dr, *Danio rerio*. GenBank accession numbers are *Alftz*: AF237818 and *Mtftz*: AF237819.

clade but there is support separating *Antp+LsHox7+LaHB1* plus *UbdA+Lox2+Lox4* from *ftz+Lox5* and the outgroup as well as support separating *UbdA+Lox2+Lox4* from *Antp+LsHox7+LaHB1* (*UbdA* refers to *Ubx* and *abdA* genes). A number of further similarities outside the homeobox also support the orthology of *ftz* and *Lox5* (Figure 2). This scenario requires the minimum number of gene duplications (two) and losses (none) to produce *UbdA+Lox2+Lox4*, *Antp+Lshox7* and *ftz+Lox5* representatives in the ecdysozoa and lophotrochozoa, and as such is a parsimonious explanation of the data. All alternative trees were rejected either as significantly less parsimonious (using the Templeton and Kishino-Hasegawa tests) or requiring four or more gene duplication/loss events (see Supplementary material).

In further support of the proposed orthology relationship between *ftz* and *Lox5*, the anterior expression boundary of leech *Lox5* is immediately posterior to that of the leech *Scr* ortholog (*Lox20*), as would be expected of a *ftz* ortholog according to the rule of colinearity between chromosomal position and anterior boundary of expression domain. Genes from two other ecdysozoans (*Mab-5* genes from two nematodes) and from a deuterostome (amphioxus *Hox-6*) also group with *ftz*/*Lox5* in further analyses (data not shown), and these share with *ftz*/*Lox5* one or other of the two amino acids typical of *ftz*/*Lox5* (Figure 2). These genes also have chromosomal positions (amphioxus) or anterior boundaries of expression (nematodes) adjacent to the respective *Scr* orthologs lending further credence to these orthology assignments.

In conclusion, the demonstration that *ftz* is derived from a Hox gene and has lost its anteroposterior patterning role in the insects means that the number of HOM/Hox genes involved in anteroposterior patterning in the arthropod/onychophoran ancestor must have been ten (eight Hox genes plus *zen*/*Hox3* and *ftz*). Furthermore, the most parsimonious conclusion from my sequence analyses shows that a *ftz* ortholog was present in the protostome ancestor and possibly in all bilaterians. As there is no convincing evidence for grouping *Ubx* and *abdA* with each other rather than each being an ortholog of either *Lox2* or *Lox4* of lophotrochozoans

Figure 3



Phylogenetic analysis of *Antp*-class Hox genes showing that *ftz* genes group with *Lox5* genes. NJ tree rooted by *Deformed* (*Dfd*) and *Sex combs reduced* (*Scr*). *Lox5* and *ftz* genes are grouped (in red). This branch was also found in MP and PML analyses. NJ and MP bootstrap values of >50% are shown above branches and PML percentage reliability values are shown below branches. Separate NJ analyses grouped all divergent insect *ftz* genes, the nematode *Mab-5* genes and Amphioxus *Hox-6* with *ftz*/*Lox5* but with bootstrap support <50% (data not shown). Gene names are as in Figure 2 with the addition of Jc, *Junonia coenia*; Ms, *Manduca sexta*; Fr, *Fugu rubripes*; Bf, *Branchiostoma floridae*; Mm, *Mus musculus*; Ci, *Ciona intestinalis.*

(see the Supplementary information in [5]), we can avoid the assumption of two independent duplications and suggest that *Ubx* and *abdA* orthologs were also present in the protostome common ancestor. This would mean that the number of Hox genes in the Precambrian common ancestor of ecdysozoans and lophotrochozoans must have been certainly nine and probably all ten of the Hox genes found in extant arthropods.

Supplementary material

Supplementary material including a figure showingalternative scenarios for relationships of *ftz* and *Lox5* and additional methodological details is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

Thanks to Michael Akam and Richard Thomas for helpful comments, Richard Thomas for sharing results, Jackie MacKenzie-Dodds and Deborah Goode for practical assistance, Jeremy Austin for advice on PAUP analyses and Ruth Dewel for providing tardigrades. This work has been supported by the BBSRC and the Natural History Museum, London.

References

1. McGinnis W, Levine M, Hafen E, Kuroiwa A, Gehring WJ: **A conserved DNA sequence in homeotic genes of the *Drosophila* Antennapedia and bithorax complexes.** *Nature* 1984, **308:**428-433.
2. Brown SJ, Hilgenfeld RB, Denell RE: **The beetle *Tribolium castaneum* has a *fushi tarazu* homolog expressed in stripes during segmentation.** *Proc Natl Acad Sci USA* 1994, **91:**12922-12926.
3. Dawes R, Dawson I, Falciani F, Tear G, Akam M: **Dax, a locust Hox gene related to *fushi-tarazu* but showing no pair-rule expression.** *Development* 1994. **120:**1561-1572.
4. Grenier JK, Garber TL, Warren R, Whitington PM, Carroll S: **Evolution of the entire arthropod Hox gene set predated the origin and radiation of the onychophoran/arthropod clade.** *Curr Biol* 1997, **7:**547-553.
5. de Rosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, *et al.*: **Hox genes in brachiopods and priapulids: implications for protostome evolution.** *Nature* 1999, **399:**772-776.
6. Giribet G, Carranza S, Baguñá J, Riutort M, Ribera C: **First molecular evidence for the existence of a Tardigrada plus Arthropoda clade.** *Mol Biol Evol* 1996, **13:**76-84.
7. Aguinaldo AA, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, *et al.*: **Evidence for a clade of nematodes, arthropods and other moulting animals.** *Nature* 1997, **387:**489-493.
8. Maier D, Sperlich D, Powell JR: **Conservation and change of the developmentally crucial *fushi tarazu* gene in *Drosophila.*** *J Mol Evol* 1993, **36:**315-326.
9. Lawrence PA, Johnston P: **Pattern formation in the *Drosophila* embryo: allocation of cells to parasegments by *even-skipped* and *fushi tarazu.*** *Development* 1989, **105:**761-767.
10. Telford MJ, Thomas RH: **Expression of homeobox genes show chelicerate arthropods retain their deutocerebral segment.** *Proc Natl Acad Sci USA* 1998, **95:**10671-10675.
11. Damen WGM, Hausdorf M, Seyfarth EA, Tautz D: **A conserved mode of head segmentation in arthropods revealed by the expression pattern of Hox genes in a spider.** *Proc Natl Acad Sci USA* 1998, **95:**10665-10670.
12. Hirth F, Hartmann B, Reichert H: **Homeotic gene action in embryonic brain development of *Drosophila.*** *Development* 1998, **125:**1579-1589.
13. Abzhanov A, Kaufman TC: **Novel regulation of the homeotic gene *Scr* associated with a crustacean leg-to-maxilliped appendage transformation.** *Development* 1999, **126:**1121-1128.
14. Telford MJ, Thomas RH: **Of mites and *zen*: expression studies in a chelicerate confirm *zen* is a divergent Hox gene.** *Dev Genes Evol* 1998, **208:**591-594.
15. Damen WGM, Tautz D: **A Hox class 3 orthologue from the spider *Cuppiennius salei* is expressed in a Hox-like fashion.** *Dev Genes Evol* 1998, **208:**586-590.
16. Copeland JWR, Nasiadka A, Dietrich BH, Krause HM: **Patterning of the *Drosophila* embryo by a homeodomain deleted Ftz polypeptide.** *Nature* 1996, **379:**162-165.
17. Furukubo-Tokunaga K, Flister S, Gehring WJ: **Functional specificity of the *Antennapedia* homeodomain.** *Proc Natl Acad Sci USA* 1993, **90:**6360-6364.