Parallel Duplication and Partial Subfunctionalization of β-Catenin/Armadillo during Insect Evolution

Riyue Bao, Tami Fischer, Renata Bolognesi, Susan J. Brown, and Markus Friedrich

1Department of Biological Sciences, Wayne State University
2Division of Biology, Kansas State University
3Department of Anatomy and Cell Biology, School of Medicine, Wayne State University

†Present address: Monsanto Company, St. Louis, Missouri
*Corresponding author: E-mail: friedrichm@wayne.edu.
Associate editor: Claudia Schmidt-Dannert

Abstract

β-Catenin is a multifunctional scaffolding protein with roles in Wnt signaling, cell adhesion, and centrosome separation. Here, we report on independent duplications of the insect β-Catenin ortholog armadillo (arm) in the red flour beetle Tribolium castaneum and the pea aphid Acyrthosiphon pisum. Detailed sequence analysis shows that in both species, one paralog lost critical residues of the α-Catenin binding domain, which is essential for cell adhesion, and accumulated a dramatically higher number of amino acid substitutions in the central Arm repeat domain. Residues associated with aspects of Wnt signaling, however, are conserved in both paralogs. Consistent with these molecular signatures, the effects of specific and combinatorial knockdown experiments in the Tribolium embryo indicate that the duplication resulted in redundant involvement in Wnt signaling of both β-Catenin paralogs but differential inheritance of the ancestral cell adhesion and centrosome separation functions. We conclude that the duplicated pea aphid and flour beetle β-catenin genes experienced partial subfunctionalization, which appears to be evolutionarily favored. Providing first evidence of genetic separability of the cell adhesion and centrosome separation functions, the duplicated Tribolium and Acyrthosiphon arm paralogs offer new inroads for context-specific analyses of β-Catenin. Our data also revealed the conservation of a C-terminally truncated Arm isoform in both singleton and duplicated homologs, suggesting an as yet unexplored role in Wnt signaling.

Key words: gene duplication, Drosophila, Tribolium, pea aphid, β-Catenin, armadillo, subfunctionalization, centrosome, cell adhesion, Wnt-signaling, redundancy.

Introduction

The Drosophila homolog of β-Catenin, Armadillo (Arm), was first discovered as transcriptional coactivator, which interacts with the Drosophila T-cell factor (dTCF) to stimulate the expression of Wingless (Wg) signaling target genes during development (Peifer et al. 1991; Peifer 1995). Subsequent work has shown that Arm is a critical transducer of the canonical Wnt signal transduction pathway (for review, see Peifer et al. 1991; Bienz 2005; Xu and Kimelman 2007). In the absence of Wnt signal, most Arm protein is membrane associated in the adherens junctions. The cytoplasmic Arm level is kept low because Arm is bound by the Adenomatous polyposis coli (APC) protein and targeted for degradation following phosphorylation at the N-terminus by Zeste-white 3 kinase (Zw3), the Drosophila homolog of Glycogen synthase kinase 3β (GSK-3β). Upon Wg signal activation, cytoplasmic Arm is stabilized by suppression of Zw3-mediated phosphorylation and targeted to the nucleus to associate with dTCF and elicit target gene transcription.

At the adherens junctions, Arm serves as a scaffold protein, directly binding to Cadherins and α-Catenin. In combination, these proteins form a multiprotein complex mediating cell adhesion and anchoring the actin filaments in epithelial cells (Oda et al. 1993, 1994; Peifer 1993). In the absence of Arm, the maintenance of cell–cell adhesion and the integrity of the actin cytoskeleton are disrupted (Cox et al. 1996; Orsulic and Peifer 1996; Pai et al. 1996). More recently, β-Catenin/Arm was also found essential for normal centrosome separation (McCartney et al. 2001). This function involves direct association with APC and possibly α-Catenin (McCartney et al. 2001; Kaplan et al. 2004; Liu et al. 2006; Bahmanyar et al. 2008).

The β-Catenin/Arm protein can be divided into three major domains: the acidic N-terminal tail domain, the highly conserved central or Arm repeat domain and the glycine-rich C-terminal tail domain (Riggelman et al. 1989; Peifer and Wieschaus 1990; Huber et al. 1997) (fig. 1). In the mouse β-Catenin homolog Catenin (cadherin-associated protein) β 1 (Ctnnb1), which is commonly used as reference sequence, the Arm repeat domain extends of 534 amino acids. This encompasses a chain of 12 imperfect 42 amino acids long Arm repeat motifs, interrupted by an insertion of 14 amino acids between the second and the third helix of Arm repeat 10.

Functional studies have shown that all three regions are important for both Wingless signaling and cell adhesion (Peifer et al. 1993; Castano et al. 2002). Modifications of residues in the terminal domains modulate the association of the central region with cell adhesive proteins (Castano et al. 2002). Through the central region, Arm forms mutually exclusive complexes with Cadherins, APC, or TCF (Hulsken et al. 1994; von Kries et al. 2000).
The sum of many structural and mutational studies developed a comprehensive model of β-Catenin/Arm protein function (for review, see Kimelman and Xu 2006; Gottardi and Peifer 2008). Low sequence conservation and high conformational flexibility, however, rendered functional studies of the Arm terminal domains difficult. In the central domain, by contrast, a high level of sequence conservation and overlapping requirements of different binding partners challenge the identification of context-specific residues and mechanisms. Here were present findings from studying the evolutionary variation of β-Catenin/Arm in insects, focusing on independent duplications in the red flour beetle Tribolium castaneum and the pea aphid Acyrthosiphon pisum. Comparative sequence analysis suggests that in both cases, the arm sister paralogs experienced similar subfunctionalization trajectories to the effect that the cell adhesion and centrosome separation functions were partitioned among the sister paralogs. The Wnt signaling–related functionalities, however, have remained preserved in both paralogs. Results from expression and RNAi knockdown analysis in Tribolium are consistent with a model of partial subfunctionalization of β-Catenin/Arm in Tribolium and Acyrthosiphon.

Materials and Methods

Sequence Retrieval and Ortholog Search

Drosophila melanogaster Arm protein sequence NCBI accession NP_476666.1 (Riggleman et al. 1989; Peifer and Wieschaus 1990) was used as the query in searches against the genome sequence databases of the mosquito species Anopheles gambiae, Aedes aegypti, and Culex quinquefasciatus (A. gambiae str. PEST genome database version 2.2, A. aegypti genome database version 1.0, C. quinquefasciatus genome database 1.0) (Holt et al. 2002; Zdobnov et al. 2002; Nene et al. 2007; Salzberg et al. 2009), the silkworm Bombyx mori (genome database version 1.0) (Mita et al. 2004; Xia et al. 2004), red flour beetle T. castaneum (T. castaneum Georgia GA2 genome database version 3.0) (Tribolium-Genome-Sequencing-Consortium 2008), the honeybee Apis mellifera (DH4 genome database version 4.0) (Honeybee-Genome-Sequencing-Consortium 2006), the jewel wasp Nasonia vitripennis (genome assembly 1.1) (Werren et al. 2010), the pea aphid A. pisum (genome assembly 1.0) (The-International-Aphid-Genomics-Consortium 2010), the water flea Daphnia pulex (genome database version 1.0) (Colbourne et al. 2011), the sea anemone Nematostella vectensis (genome assembly 1.0) (Putnam et al. 2007), and mouse (Mus musculus genome database version 37.1) (Waterston et al. 2002). Searches against the GenBank protein databases of the above listed species (both refseq and non-refseq) were performed with BLASTP (Altschul et al. 1997). The search for homologs in D. pulex was performed on the respective genome server of the Department of Energy Joint Genome Institute. Only candidate hits, which retrieved Drosophila arm as best hit in reciprocal Blast, were included for further analysis. The protein sequence searches were supplemented by searches against the genome sequence databases to reduce the possible omission of nonannotated arm homologs. The gene models retrieved from GenBank were examined regarding their correspondence to Drosophila arm. In the case of Tribolium and Acyrthosiphon, gene models were also examined for support by expressed sequence tag (EST) data available at NCBI. All resulting putative cDNA and protein sequences used in this study are provided in supplementary data file 1 (Supplementary Material online). In select cases, detailed investigation of the genome sequence resulted in gene structure corrections, which are documented in supplementary data file 1 (Supplementary Material online).

Multiple Alignment

Multiple protein sequence alignments were generated with Probabilistic Alignment Kit (PRANK, http://www.ebi.ac.uk/goldman-srv/prank/prank/) applying the Whelan and Goldman amino acid sequence residue numbers in text and figures correspond to the mouse β-Catenin homolog Ctnnb1. For gene tree reconstruction and sequence evolution analyses, variable regions were eliminated using the Gblocks algorithm (Castresana 2000) at “less stringent” setting.

Gene Tree Reconstruction

Bayesian analysis was carried out using MrBayes 3.12 with 4,000,000 generations (sample frequency = 100) and mixed protein distance models (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). The β-Catenin homolog of N. vectensis (Nvec_Ctnnb1) was set as outgroup.
Sequence Evolution Analyses
Relative rate tests were carried out with Poisson distances as implemented in PHYLTREE 2.0 (Kumar 1996). Nonsynonymous and synonymous substitution divergencies (dn and ds) were estimated with the yn00 algorithm of PAML version 4.4 (Yang 1997, 2007). The sliding window analysis of protein sequence conservation was performed by calculating protein distances for a window size of 10 amino acids using the protdist program of PHYLPACK package version 3.69 (Felsenstein 2005). To investigate rate heterogeneity among sites, we compared the Jones et al. (1992) model of protein sequence evolution without (Jones, Taylor, and Thorton [JTT]) and with Gamma distribution (JTT + Γ) (Yang 1993) by the likelihood ratio test (Felsenstein 1988). Amino acid frequencies and the shape parameter of gamma distribution were estimated from the data. Significance was assessed by calculating the likelihood ratio as \( \ln L_2 - \ln L_1 \), where \( \ln L_1 \) and \( \ln L_2 \) stand for the likelihoods of the simpler model (JTT) and the more complex model (JTT + Γ), respectively. A chi-square distribution was applied to obtain the corresponding \( P \) values, with degrees of freedom equal to the difference in the number of parameters between the two models. The Gblocks-processed PRANK alignment used in these analyses is available in supplementary data file 2 (Supplementary Material online). Amino acid substitutions were classified into conservative, semiconservative, and nonconservative as defined by Larkin et al. (2007).

Whole-Mount In Situ Hybridization
For the analysis of Tribolium arm1 and arm2 expression, DNA fragments covering exon 3 of Tcas_arm2 and exon 2 of Tcas_arm1 were polymerase chain reaction (PCR) amplified from Tribolium genomic DNA using primers GCAGTCCTTCGTTCGCTCTC and CTTCTGACATGCTCTCAATT, GTGTGTCGCCGACAAAGGC and CCTTAGTGTCGCAATTGGC, respectively. The PCR fragments were cloned into pGEM-T Vector (Promega) to serve as templates for in vitro transcription to generate digoxigenin-labeled RNA antisense probes. Whole-mount in situ hybridization on embryos was carried out as previously described (Wolff et al. 1995). Images were taken with a Zeiss Axioplan with differential interference optics. Digital images were recorded using a SPOT RT color camera (Diagnostic Instruments, Inc.).

Parental RNAi
Reverse transcriptase-PCR was carried out to clone template DNA corresponding to nucleotide 35-1693 in the cDNA model of Tcas_arm1 isoform A and nucleotide 14-379 in the cDNA model of Tcas_arm2 isoform A. Pairwise comparison of each probe against the complete mRNA model of the paralogous gene did not detect more than 91% sequence identity between the two sequences within a 21 bp sequence window ruling out paralog target promiscuity of the Tribolium arm1 and arm2 dsRNAs, which were transcribed and purified using the T7 Megascript and Megasclear kits (Ambion). The resulting dsRNA products were diluted to required concentrations and mixed with injection buffer (5 mM KCl, 0.1 mM KPO4, pH 6.8) in 1:1 ratio. Fertile Tribolium adult females were injected, and their progeny analyzed by collecting 0–72 h old eggs that were split into two samples; one was allowed to develop at 30 °C for 3 days before cuticle preparation while the other sample was immediately fixed and mounted in Slowfade (Molecular Probes/Invitrogen), which contains 4’,6-diamidino-2-phenylindole, to observe nuclei at various developmental stages.

Results
Ancient Lineage-Specific Duplications of Arm in Pea Aphid and Red Flour Beetle
To explore the evolutionary history of insect arm, we performed protein sequence BLAST searches in the genome databases of three mosquito species (A. aegypti, A. gambiae, C. quinquefasciatus), silkworm (B. mori), red flour beetle (T. castanemum), honeybee (A. mellifera), jewel wasp (N. vitripennis), pea aphid (A. pisum), water flea (D. pulex), and sea anemone (N. vectensis). Singleton arm orthologs were discovered in most of the insect species except the red flour beetle and the pea aphid. The Tribolium genome carried two homologs, which were linked by 2 kb on chromosome 9. Two homologs were also discovered in the pea aphid genome, which were likewise very closely linked (2-kb intergenic distance on supercontig NW_001933094.1). In both species, authenticity and functionality of the duplicated loci was supported by extensive EST evidence (supplementary data file 3, Supplementary Material online). The presence of two arm paralogs in select lineages could be explained by lineage-specific gene duplication events or early duplication in the ancestor of insects followed by independent losses. To distinguish between these scenarios, we constructed phylogenetic trees from protein sequence alignments with likelihood tree estimation methods. In this analysis, the species-specific paralogs were united to form species-consistent subclades, strongly supporting the independent duplication of arm in the lineages leading to pea aphid and red flour beetle (fig. 2).

To gage the age of the underlying duplication events, we determined synonymous substitution divergencies (ds) between the arm sister paralogs in both the pea aphid and red flour beetle as an approximate measure of gene duplicate age (Lynch and Conery 2000). The ds values of 2.0708 and 3.8667 for the red flour beetle and pea aphid, respectively, indicated that the duplications in both species dated back deep in time, to an early point in the evolution of these lineages. Consistent with the predicted effect of purifying selection on the long-term preserved gene duplicates, the replacement substitution divergencies (dn) between sister paralogs were low, ranging between 0.2310 in Tribolium and 0.1886 in the pea aphid.

Asymmetric Protein Sequence Diversification in the Duplicated Arm Paralogs
The gene tree also indicated very strong differences in protein sequence change between the arm sister paralogs in both Tribolium and Acyrthosiphon (fig. 2). In both cases, relative rate
testing confirmed the significance of the between-paralog substitution differences (Triboylum: $Z = 12.2492$, $P < 0.05$; Acyrsphilon: $Z = 7.65892$, $P < 0.05$). The comparison of terminal branch lengths to other insect homologs in the arm gene tree indicated that the less substitution-affected sister paralogs evolved at a rate comparable to that of singleton arm homologs, whereas the second paralog was characterized by a number of substitutions strongly exceeding the average of singleton arm homologs. We named the more conservatively evolved paralogs of the red flour beetle and pea aphid Tcas_arm1 and Aphis_arm1, respectively, in contrast to the strongly modified paralogs Tcas_arm2 and Aphis_arm2. The five-letter acronym Aphis was chosen for the pea aphid (A. pism) genes to avoid possible confusion with genes from the honeybee (A. mellifera).

### Conservation of Amino Acid Residues Involved in the Regulation of Cytosolic Arm Levels

Given the evidence of dramatic sequence change in the strongly modified arm paralogs, we probed the protein sequence alignment for clues in regards to paralog-specific modifications of β-Catenin functions (supplementary data file 4, Supplementary Material online). We started out focusing on sites that are involved in the regulation of cytosolic Arm levels. We identified 11 sites characterized in mammalian systems that were conserved in all insect arm homologs. This included the GSK-3 including all the pea aphid and red flour beetle homologs.

- **Sequence Alignment for Clues in Regards to Paralog-Specific Modifications:**
  - **Acyrsphilon**
  - **Tribolium**
  - **Anopheles gambiae**
  - **Apis mellifera**
  - **Acyrsphilon pism**
  - **Mus musculus**
  - **Nasonia vitripennis**
  - **Daphnia pulex**
  - **Bombyx mori**
  - **Culex quinquefasciatus**
  - **Drosophila melanogaster**

### Conservation of Amino Acid Residues Related to Transcriptional Control Functions

We next examined the conservation of residues that are critical in the context of transcriptional target gene control (supplementary data file 4, Supplementary Material online). The phenylalanine TCF-interaction residues F253 and F293 were conserved in almost all Arm homologs (Graham et al. 2000). The only exception was the replacement of the F253 residue by tyrosine and arginine in Aphis_arm1 and Aphis_arm2, respectively. The Ser/Akt phosphorylation target site S29 appeared conserved with the caveat that this assessment was compromised in the pea aphid due to alignment ambiguity (Bek and Kemler 2002; van Noort et al. 2002) (fig. 3). Overall, however, the pervasive conservation of residues involved in cytosolic Arm level control suggested long-term conservation of Wnt-signaling function in both the moderately and strongly evolved arm paralogs of the red flour beetle and pea aphid.

- **Ser/Akt Phosphorylation Target Site S29:**
- **SCF Ubiquitylation Target Site T393:**
- **CBP Acetylation Target Site K49:**
- **APC Interaction Sites K312 and K435**
- **K19, the CBP Acetylation Target Site K49, and the APC Interaction Sites K312 and K435** (Coluccia et al. 2007; Hart et al. 1999; Kitagawa et al. 1999; Latres et al. 1999; Liu et al. 1999; Winston et al. 1999; Graham et al. 2000; Ryo et al. 2001;

### Conservation of Amino Acid Residues Related to Transcriptional Control Functions

We next examined the conservation of residues that are critical in the context of transcriptional target gene control (supplementary data file 4, Supplementary Material online). The phenylalanine TCF-interaction residues F253 and F293 were conserved in almost all Arm homologs (Graham et al. 2000). The only exception was the replacement of the F253 residue by tyrosine and arginine in Aphis_arm1 and Aphis_arm2, respectively. The Ser/Akt phosphorylation target site S29, which has been hypothesized to enhance transcription by interaction with histone acetylases (Fang et al. 2007), was unambiguously preserved in the conservatively organized arm homologs. Also the strongly modified Tcas_arm2 and Aphis_arm2 genes encoded alignable serine residues. Residue homology, however, was tentative due to the poor sequence conservation in the adjacent insertion region (fig. 3). A similar situation applied to Ser/Akt phosphorylation target S675 (Fang et al. 2007). In this case, homology assessment was cautioned in the Tcas_arm2 paralog due to an N-terminally adjacent single amino acid deletion. The possible Met phosphorylation target Y670, which has been associated with nuclear translocation in mammalian
systems (Zeng et al. 2006), was conserved in all conserva-
vatively organized insect Arm proteins but replaced by leucine
in Aphis_arm2. In Tcas_arm2, the same residue was replaced
with a serine, which may serve as an alternate phosphory-
lation target. Most importantly, the Legless (Lgs) interaction
residues D162 and D164 and the P300 acetylation target site
K345, which enhances TCF interaction in mammalian sys-
tems (Levy et al. 2004), were universally conserved across

**Fig. 3.** Protein sequence conservation of insect Arm homologs. Triple asterisks indicate omission of poorly alignable sequence stretch. Numbers on top of alignment indicate select alignment sites discussed in the text with reference to the mouse β-Catenin amino acid sequence. Select ambiguous alignment regions in functional domains discussed in the text are indicated by light print font. Hatched boxes indicate α-helices in the Arm repeat domains as annotated at the bottom of the alignment. Residues that have been experimentally associated with specific protein–protein interactions are highlighted by colored background. Colors indicate functional context: Red = Regulation of cytosolic Arm levels, Green = β-Catenin binding, Brown = Interaction with transcriptional coregulators, Blue = Interaction between the Helix C domain and α helix 3 of Arm repeat 12, Tangerine = E-Cadherin binding. In select cases, overlapping functions are indicated color gradient. For references, see text and supplementary data file 4 (Supplementary Material online). Sliding window relief of sequence conservation averaged from Dmel_arm, Agam_arm, Tcas_arm1, Amel_arm, and Aphis_arm1 is shown on top of the Arm repeat region (supplementary data file 6, Supplementary Material online). The Y axis indicates averaged amino acid sequence divergence. Background lines represent 0.2% divergence increments.
insect species (Hoffmans and Basler 2007). The sum of these findings indicated the conservation of transcription regulating functionality in singleton Arm homologs and also the duplicated Arm homologs of the red flour beetle and pea aphid, consistent with the evidence of Wnt-signaling context preservation.

Amino Acid Replacements in the Helix C Domain
It has been proposed that some of the interactions of Arm with transcriptional coregulators are mediated through the capping of Arm repeat 12 by the folding back of the N-terminally adjacent Helix C domain (Xing et al. 2008). We probed for the evolutionary conservation of this mechanism by examining the conservation of amino acid residues that have been associated with the interaction of Arm repeat 12 and the basic Helix C domain (Xing et al. 2008) (supplementary data files 4 and 5, Supplementary Material online). The respective sites were widely conserved in our alignment (fig. 3). Nonconservative amino acid replacement differences were only found in the mosquito species and the strongly sequence modified arm paralogs of the red flour beetle and the pea aphid.

In the mosquitoes, the nonconservative replacements were restricted to residues in the Helix C domain (S681L, L682P, R684F), suggesting a possible reorganization of the backfolding mechanism during the evolution of this dipteran subclade. A similar Helix C-restricted amino acid replacement distribution was found in the Tcas_arm2 paralog. In this case, at least one residue was lost by deletion (L674) and one site experienced nonconservative replacement (L678A), in addition to one conservative amino acid change (R664Q) and one semi-conservative residue change (K671S). Aphis_arm2 was characterized by three nonconservative amino acid replacements that were equally distributed between Arm repeat 12 and Helix C (S646C, F660S, R684S). In addition, the amino acid residue K671 either experienced a deletion or conservative replacement to glutamine. Aphis_arm2 was further distinguished by the deletion of three amino acids N-terminally adjacent to the Helix C domain (fig. 3).

Taken together, these findings suggested conservation of the Helix C backfolding mechanism in the conservatively evolving insect arm homologs. The residue differences in the strongly modified paralogs of the pea aphid and the red flour beetle, however, indicated loss or reorganization of the interaction between the Helix C domain and Arm repeat 12.

Paralog-Specific Loss of the α-Catenin Interaction Residues
The α-Catenin binding domain extends from the N-terminal domain into the first Arm repeat (figs. 1 and 3) (Pai et al. 1996) (supplementary data file 4, Supplementary Material online). As expected, this region was highly conserved in all singleton arm orthologs. The α-Catenin interaction domain was also conserved in the conservatively organized arm paralogs of the red flour beetle and pea aphid. The strongly modified arm paralogs of pea aphid and red flour beetle, however, were characterized by extensive deletions, suggesting the loss of amino acid residues that have been found essential for α-Catenin binding (Aberle et al. 1996) (Cox et al. 1996; Orsulic and Peifer 1996; Pai et al. 1996; Bek and Klemm 2002).

In the case of the Tcas_arm2 five (R90, T120, R124, L125, E127) of 16 investigated α-Catenin interaction residues were located in indel-diverged regions rendering homology assignments problematic. Notwithstanding, sequence conservation could be ruled out for at least two of these residues (T120, R124). Of note, the suggestive power of the loss of R124 is reduced in light of the only very mild reduction of α-Catenin binding in mutational analysis of the corresponding site in mouse Plakoglobin (Aberle et al. 1996). However, we also noted the deletion-induced loss of the conserved L125 residue, which is essential for α-Catenin binding in mouse Plakoglobin (Aberle et al. 1996). Furthermore, three α-Catenin binding associated residues that were embedded in indel-conserved regions of Tcas_arm2 had accumulated semi-conservative (T112P) or nonconservative amino acid replacements (R93T, R95P).

In the case of Aphis_arm2, the essential α-Catenin binding residues R90, R93, and R95 mapped to deletion-diverged regions, but no alignment variant was compatible with their conservation. Moreover, three sites corresponding to essential α-Catenin binding residues had accumulated nonconservative amino acid replacements (T120A, R124N, E127G). Finally, the critical L125 residue was replaced by methionine in Aphis_Arm2 (fig. 3). These findings were strongly suggestive of paralog-specific loss of α-Catenin binding capacity in the highly strongly modified arm paralogs.

Evolutionary Conservation of Residues Associated with Cadherin Binding
In the cell adhesion context, β-Catenin/Arm associates with Cadherins in addition to α-Catenin. The preliminary evidence of α-Catenin binding incompetence of the strongly modified arm paralogs in the red flour beetle and the pea aphid prompted us to investigate the possibility of correlated changes in 15 residues known to be associated with Cadherin binding (supplementary data file 4, Supplementary Material online). Most of these residues were conserved in the strongly modified arm paralogs in the red flour beetle and the pea aphid. The lysine residue 335, however, which has been shown to participate in E-Cadherin as well as TCF binding (Roura et al. 1999; Huber and Weis 2001; Brembeck et al. 2004; Taurin et al. 2006; Hoffmans and Basler 2007), was replaced by arginine in Tcas_arm2. Even more intriguingly, arginine residue 474, which has likewise been shown to participate in E-Cadherin and TCF binding (Roura et al. 1999; Huber and Weis 2001; Brembeck et al. 2004; Taurin et al. 2006; Hoffmans and Basler 2007), was replaced by lysine in Tcas_arm2 and glycine in Aphis_arm2. A second notable change in Aphis_arm2 concerned the unique nonconservative replacement of tyrosine with cysteine on site 654, which has been implicated to bind E-Cadherin residue 665 (Huber and Weis 2001).
Evolutionary Conservation of the Drosophila C-terminal Domain-Truncated Arm Isoform

In Drosophila, *arm* is expressed in two isoforms: the first *arm* mRNA species discovered, isoform A (*Dmel_armRA*), produces a transcript that encodes the ubiquitous 843 amino acid long Arm protein *Dmel_ArmPA*. Differential splicing of exon 6 generates a less abundant *arm* mRNA species: isoform C (*Dmel_armRC*), which translates into the 721 amino acid long *Dmel_ArmPC* protein (*DCTD-Arm*) that lacks the C-terminal domain. This and the axonal localization of this isoform have been taken as evidence that *DCTD-Arm* is a cell adhesion–specific isoform (Loureiro and Peifer 1998). Considering the evidence of cell adhesion function loss in *Tcas_arm2* and *Aphis_arm2*, we explored genomic evidence of whether the differential splicing of the *DCTD-Arm* isoform was conserved in other insect species and, most importantly, in the *arm* paralogs of red flour beetle and pea aphid.

Comparative analysis of gene structures revealed a number of *Dmel_armRA* splicing sites that were conserved across a subset of *arm* homologs (fig. 4). Most importantly, exon 6 was conserved across all insect homologs including the strongly modified paralogs of pea aphid and red flour beetle. Furthermore, consistent conservation of the seven C-terminal amino acids long ΔCTD-Arm terminus suggested that the underlying differential splicing mechanism was preserved as well (fig. 4). This was confirmed by extensive EST coverage of the predicted isoforms A and C in both *arm1* and *arm2* of pea aphid (supplementary data file 3, Supplementary Material online). We therefore concluded that ΔCTD-Arm processing was conserved in both *Aphis_arm1* and *Aphis_arm2* despite the evidence of a reduced role in cell adhesion of *Aphis_arm2*.

Parallel Amino Acid Substitutions in the Repeat Domain of the Duplicated Insect Arm Paralogs

In light of the evidence that the strongly modified *arm* sister paralogs of the red flour beetle and pea aphid experienced parallel loss of α-Catenin interaction, we investigated the possible impact outside the α-Catenin binding region in the form of parallel amino acid replacements. We focused on a total of 468 sites in the central Arm repeat domain that were highly

**Fig. 4.** Conservation of insect *arm* gene splice isoforms. Comparison of intron–exon structure in the coding region of select insect *arm* genes. *Drosophila arm* exon numbers based on Loureiro and Peifer (1998). Boxes show multiple amino acid sequence alignment of the *Drosophila* Arm PA and PC C-termini to the corresponding protein sequences in other insect species. Arrowheads indicate splice junctions.
conserved as indicated by amino acid differences in not more than two taxa (supplementary data file 6, Supplementary Material online). In this subsample of sites, Tcas_arm2 and Aphis_arm2 paralogs had accumulated 119 (25.4%) and 106 (22.6%) amino acid substitutions, respectively, revealing a similar degree of overall sequence change. Thirty of these sites had accumulated amino acid replacements in both Tcas_arm2 and Aphis_arm2. Compared with the ancestral Arm sequence reconstructed from conserved insect Arm homologs, 33.3% out of these replacements constituted conservative substitutions and 16.7% constituted semiconservative substitutions. Most significantly, five of these replacements had resulted in identical amino acids changes (K180N, V414I, A576T, V600I, V658I) (fig. 3 and supplementary data file 7, Supplementary Material online).

In contrast to the equivalent degree of sequence change in the strongly modified Tcas_arm2 and Aphis_arm2 paralogs, we noted differences in the rate of substitution accumulation in the conservatively organized Tcas_arm1 and Aphis_arm1 paralogs. The total of 30 (6.4%) of Aphis_arm1-specific substitutions at conserved sites compared with six (1.3%) substitutions that were specific for Tcas_arm1. This finding revealed a higher relative amount of amino acid substitution in the moderately evolving paralog of the pea aphid following gene duplication. Interestingly, seven of the amino acid replacements in the conservatively organized arm1 paralog of the pea aphid took place at sites that had also accumulated replacement substitutions in the strongly modified arm2 paralog of the red flour beetle. Moreover, five substitutions led to the same amino acid residues (D267E, E479D, V485I, V537I, I610V) (supplementary data file 6, Supplementary Material online). No such patterns were found in the reverse comparison between the strongly modified pea aphid arm2 paralog and the conservatively organized red flour beetle arm1 paralog.

Differential Accumulation of Arm Paralog Transcripts during Tribolium Embryogenesis

To gain insights into the functional consequences of β-Catenin/arm duplication, we took advantage of the well-developed accessibility of Tribolium for developmental experimentation. First, we investigated the embryonic expression patterns of Tcas_arm1 and Tcas_arm2 by whole-mount in situ hybridization (fig. 5). These experiments were carried out with antisense RNA probes against regions that were only 64% identical, ruling out cross-reactivity of one probe against both paralog transcripts. The comparison of embryos labeled for Tcas_arm1 and Tcas_arm2 revealed that both paralogs were ubiquitously expressed but differentially enriched in select tissues. Early germ band extension embryos were characterized by the preferential accumulation of Tcas_arm2 in a segmental expression pattern that was compatible with a function in Wnt signaling (fig. 5B). This pattern was only marginally detectable with the Tcas_arm1 probe (fig. 5A).

A second conspicuous difference was the preferential accumulation of Tcas_arm1 signal in the developing nerve cord of early dorsal wall embryos (fig. 5E). In most other regions of the embryo including the developing legs, Tcas_arm2 appeared to be expressed more strongly (compare fig. 5C with D). Taken together, these observations suggested a stronger involvement of Tcas_arm1 in neuronal development and of Tcas_arm2 in Wnt-signaling processes.

Paralog-Specific Embryonic Knockdown Phenotypes in Tribolium

To study the functional consequences of β-Catenin/arm duplication, we examined the effects of reducing transcript levels in the embryo by parental RNAi experiments (table 1). Depletion of Tcas_arm1 by injecting a high (0.5–1.5 µg/µl) concentration of dsRNA resulted in immediate loss of egg production. In addition, adult lethality was observed within 4 weeks. When a lower dose (10 ng/µl) of Tcas_arm1 was injected, the adults exhibited normal viability and egg laying. Embryos harvested over an extended period of time displayed a graded series of knockdown phenotypes as the RNAi effect dissipated (fig. 6). Nuclear staining of a 0–72 h sample of 33 Tcas_arm1 RNAi embryos collected 1 week after injection revealed five embryos with no nuclei, four with fairly normal blastodermers (fig. 6B), four embryos in which the embryonic tissue appeared to be separating from the serosa (not shown), 19 post-blastoderm embryos containing amorphous clumps of cells instead of germ bands (compare fig. 6H and G) and one deteriorating germ band embryo (not shown). As the effects of RNAi began to weaken further over time, it was possible to score cuticular phenotypes. Tcas_arm1 RNAi embryos consistently showed defects in dorsal closure but contained the normal complement of segments and complete appendages (fig. 6K). Occasionally,
the distal tips of a single antennae or leg developed abnormally (arrowhead in fig. 6L and arrow in fig. 6M).

Parental RNAi of Tcas_arm2 had very different consequences. Injection at high dsRNA concentration (0.5–1.5 µg/µl) affected neither viability nor egg production but the resulting eggs failed to produce cuticle structures, indicating that Tcas_arm2 fulfills at least one vital function in early development prior to cuticle deposition. A 0–72 h sample of 53 Tcas_arm2 RNAi embryos collected 2 weeks after injection at a lower dsRNA concentration (500 ng/µl) contained 43 embryos with no or few nuclei (not shown), five blastoderm stage embryos (fig. 6C), four embryos with nuclei clumped together at the posterior pole (fig. 6I), and one fairly wild-type (wt) germ rudiment (not shown).

Closer examination of the Tcas-Arm2 RNAi embryos revealed improperly oriented nuclear divisions (arrowhead, fig. 6F), and many nuclei that had sunk into the yolk (arrow, fig. 6E). This contrasted with the nuclei in wt and Tcas-Arm1 RNAi blastoderm embryos, which were found in a single layer near the surface (fig. 6D and E). Furthermore, nuclear divisions in Tcas-Arm1 RNAi embryos were oriented parallel to the surface of the eggs (arrow in fig. 6E), as in wt. We also recovered a few Tcas_arm2 RNAi cuticles, which were truncated, containing only a few abdominal segments, and their appendages were often crumpled. In one case, the distal tips of all appendages were missing (arrow in fig. 6N), reminiscent of the appendage phenotype reported for wg-knockdown embryos in Tribolium (Ober and Jockusch 2006). The segments that did form, appeared wider than normal (arrowhead in fig. 6N), similar to the segments formed in Tcas-pangolin RNAi embryos (Bolognesi et al. 2008).

Combined Knockdown of the Tribolium Arm Paralogs Uncovers Redundant Involvement in Wnt Signaling

The conservation of Wnt signaling–related residues in Tcas_arm1 and Tcas_arm2 combined with the comparatively mild Wnt-signaling phenotypes obtained in the single gene knockdown experiments prompted us to hypothesize that the sister paralogs functioned redundantly in Wnt signaling, providing mutual compensatory rescue in the single knockdown embryos. To test this idea, we investigated whether the simultaneous knockdown of both paralogs would yield stronger Wnt-signaling phenotypes as the result of the breakdown of compensatory rescue activity. Injecting Tcas_arm1 at 10 ng/µl together with Tcas_arm2 at 500 ng/µl yielded only very few eggs and no cuticles. In a second attempt, we injected 54 adult females with Tcas_arm1 and Tcas_arm2 both at 10 ng/µl. Also in this case, adult lethality was high. Only 22 individuals were still alive after 2 weeks from which eggs were collected every 3 days and aged for 5 days at 30 °C before cleared cuticles were prepared. All of the small number of eggs collected during the first 2 weeks after injection (10–30 per collection) were empty. The following collection of 24 eggs, however, included 10 eggs, which contained cuticle phenotypes that ranged from mild as obtained by the Tcas_arm2 single knockdown to more severe (fig. 7). In the latter class, the cuticles were small, spherical, dorsally open and ventrally smooth (fig. 7A). Evidence of segmentation around the lateral edges indicated the presence of rudimentary head and thoracic segments. These abnormalities corresponded to the phenotypes produced by knockdown of other components of the Wnt-signaling pathway such as Tcas_porcupine (Tcas_por6) (fig. 7B) (Bolognesi et al. 2008, 2009). Thus, the synergistic effect of simultaneously knocking down Tcas_arm2 and Tcas_arm2 confirmed the idea that the both paralogs contributed to Wnt signaling, resulting in developmental redundancy.

Discussion

The evolution of duplicated genes is often associated with some form of paralog-specific inheritance of select ancestral functionality, leading to what is commonly referred to as subfunctionalization (Force et al. 1999). Alternatively, gene duplicates may continue to execute ancestral functionality in an overlapping manner. If the sister paralogs continue to function at reduced dosage such that the combined dosage reproduces the ancestral level of singleton gene expression, this outcome can be considered quantitative subfunctionalization. In this specific form, quantitative subfunctionalization would not lead to or increase genetic redundancy. Moreover, it does not reconcile with long-term preservation of gene duplicates. Neutrally drift-inspired changes of the relative expression levels are expected to eventually lead to the complete silencing and loss of one paralog. However, there are many known cases of long-term preserved duplicates where the shared functional inheritance is associated with an increase in developmental redundancy and, as a consequence, robustness. This can be due to an overall increase of functional transcript or additional mechanisms that regulate compensatory rescue. In regards to the case presented here, we conclude that the combined data suggest that the β-Catenin/Arm homologs of pea aphid and red flour beetle experienced in part sub-functionalization and in part redundant conservation of an ancestral functionality (fig. 8). Below, we will refer to this outcome as partial subfunctionalization.

Redundancy Buffering of β-Catenin/Arm’s Wnt-Signaling Function by Gene Duplication

The shared preserved ancestral function of the duplicated paralogs concerns β-Catenin/Arm’s role in Wnt-signal transduction. This is indicated by the conservation of all known Wnt signaling–related amino acid residues in both paralogs and further substantiated by the data in Tribolium. In the latter, both sister paralogs are ubiquitously coexpressed, differing only quantitatively in select regions of the embryo. Functional proof for preserved Wnt-signaling functionality of both paralogs in Tribolium comes in the form of the synergistic effect of knocking down Tcas_arm1 and Tcas_arm2 in combination. This experiment generated a stronger Wnt-signaling deficiency phenotype than single knockdown of Tcas_arm1 and Tcas_arm2, implying
compensatory rescue in the single knockdown embryos. The discovery of functional redundancy in the context of Wnt signaling is of little surprise given the critical significance of this signaling pathway during development and the importance of redundant control in developmental regulation (Cooke et al. 1997; Wagner 2008). The similarly conserved Wnt signaling–related residues in the pea aphid arm paralogs suggest that the latter provide the same form of developmental buffering in the pea aphid, which will be interesting to test experimentally.

Our knockdown data also provide evidence of differential paralog-specific sensitivities of later Wnt-signaling processes, indicating a more complex picture, at least in Tribolium. This evidence comes in the form of the mild segmentation and appendage phenotypes in Tcas_arm2 knockdown animals, which were not observed in the Tcas_arm1 knockdown animals. The paralog specificity of these phenotypes is consistent with the stronger enrichment of Tcas_arm2 transcripts in the segments of early germband embryos as well as in the thoracic appendages of older embryos compared with Tcas_arm1. This corollary identifies differential expression level regulation as candidate mechanism of the varied contribution to Wnt signaling by the sister paralogs. Alternatively or in addition, it is possible that the two paralogs differ by facilitating different Wnt-signaling dynamics in a tissue-specific manner. A precedent of this has been found in the zebrafish, where the very closely related β-Catenin-1 and β-Catenin-2 paralogs act redundantly in neuroectoderm suppression but the earlier patterning of the organizer is specifically dependent on normal β-Catenin-2 levels (Bellipanni et al. 2006). Recent work suggests that the difference in cytosolic turnover dynamics is the underlying cause for the differential function (Mo et al. 2009).

In this context, it is further interesting to note the evolutionary conservation of the differential splicing of exon 6 in all the arm homologs of red flour beetle and pea aphid. This finding suggests a critical Wnt signaling–specific function of

### Table 1. Summary of Tcas_arm1 and Tcas_arm2 Knockdown Results.

<table>
<thead>
<tr>
<th>Tcas_arm1 (0.5–1.5 μg/μl) high</th>
<th>Tcas_arm2 (0.5–1.5 μg/μl) high</th>
<th>Tcas_arm2 (500 ng/μl) low</th>
<th>Tcas_arm1 (10 ng/μl) + Tcas_arm2 (500 ng/μl)</th>
<th>Tcas_arm1 (10 ng/μl) + Tcas_arm2 (10 ng/μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult Lethality</td>
<td>Egg Production</td>
<td>Cuticle Formation</td>
<td>Cell Clumping</td>
<td>Aborted Cell Divisions</td>
</tr>
<tr>
<td>wt</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>arm1 (10 ng/μl) low</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>arm2 (0.5–1.5 μg/μl) high</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>arm1 (10 ng/μl) + arm2 (500 ng/μl)</td>
<td>Yes</td>
<td>Yes (few)</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
| NOTE.—na, not applicable if earlier phenotype precludes observation; nd, not determined.

**FIG. 6.** Knockdown analysis of Tribolium arm sister paralog function. All embryos are oriented with anterior to the left. (A–I) 4′,6-diamidino-2-phenylindole-stained nuclei. (J–N) cuticles. (A, D, G, J) wild type. (B, E, H, K, L, M) Tcas_arm1 parental RNAi embryos. (C, F, I, N) Tcas_arm2 parental RNAi embryos. Boxes in A, B, and C indicate regions magnified in D, E, and F, respectively. Arrow in E points to nuclei dividing parallel to surface. Similar arrow in F points to interphase nuclei just below surface of egg, and arrowhead in F points to a nucleus dividing perpendicular to egg surface. Arrow in J points to terminal claw in wild-type cuticle. Similar arrow in M points to misshapen distal leg and in N point to leg missing distal segments. Arrowhead in K points to normal antenna; similar arrowhead in L points to deformed antenna. Arrowhead in N points to segments wider than normal.
the DCTD-Arm isoform in contrast to the originally hypothesized cell adhesion related subfunction (Loureiro and Peifer 1998). There may thus be diverse mechanisms in place, which fine-tune Wnt-signal transduction dynamics at the level of cellular Arm turnover. Indeed, it is tempting to speculate that yet another level of differential Wnt-signaling dynamics might exist in pea aphid and the red flour beetle through the reorganization of the capping interaction between the Helix C domain and the Arm repeat 12 in the strongly modified arm paralogs (Xing et al. 2008).

Select Preservation of Cell Adhesion Function in Tcas_arm1

Parallel functionalization fates of the duplicated pea aphid and flour beetle arm homologs are indicated by the loss of β-Catenin binding residues and enhanced sequence change in one paralog of each species. On their own, the sequence evolution data can be explained by one of two outcomes: subfunctionalization or neofunctionalization. In the subfunctionalization scenario, the duplication event released one paralog from the cell adhesion function thus permitting increased substitution accumulation. The latter may be due to the relaxation of purifying selection, resulting from the loss of ancestral cell adhesion function, or positive selection-driven optimization of ancestral functionalities that were constrained by adaptive conflicts in the ancestral singleton protein. In the neofunctionalization scenario, the faster paralog acquired a novel function after relinquishing the cell adhesion function. Also in this case, substitution accumulation could be driven by positive selection, relieve from purifying selection, or a combination thereof.

For discussing how the results from the knockdown analysis in Tribolium support the subfunctionalization scenario, it is helpful to keep in mind that the injection experiments with high doses of dsRNA uncover early developmental processes that are tolerant to substantial decrease of normal Tcas_arm1 or Tcas_arm2 levels. The phenotypes resulting from the injection at a low dose of dsRNA or appearing several weeks after injection, when the arm transcript reduction gradually wears off, reveal processes that are sensitive to milder reduction of arm transcript levels. A second important component of the discussion is the relation to known phenotypes in Drosophila, which provides further clues for the mechanistic interpretation of the phenotypes seen in Tribolium. In addition, the comparison with Drosophila allows for determining whether a given function is conserved between the two species and thus ancestral consistent with subfunctionalization or unique for Tribolium and thus derived and indicative of neofunctionalization.

Females injected with high dosage of Tcas_arm1 dsRNA were characterized by severely reduced viability and immediate loss of egg laying capacity. By contrast, no effect on these traits was noted in the high dsRNA dosage Tcas_arm2 knockdown animals. One can interpret these findings as the consequences of paralog-specific systemic breakdown of cell

![Fig. 7](https://example.com/f7.png) Double knockdown analysis of Tribolium arm sister paralog function. (A) Cuticle of Tcas_arm1 + Tcas_arm2 parental RNAi embryo. (B) Cuticle of Tcas_parce depleted embryo. In these ventral views anterior is to the left.

![Fig. 8](https://example.com/f8.png) Parallel duplication and hypothesized partial subfunctionalization of insect arm. Protein domains are coded by color fills. Protein functions are given in right column where S, A and C represent Wnt signaling, cell adhesion, and centrosome separation, respectively. Bar represents time scale in millions of years. See text for details.
adhesion in Tcas_arm1 animals. In support of this conclusion, Drosophila arm-null mutant oocytes fail to complete oogenesis due to the loss of contact with the follicle cells (Orsulic and Peifer 1996). This defect is not rescued by the mutant Drosophila arm allele S14, which lacks the α-Catenin binding domain, demonstrating the specific requirement of Arm-mediated cell adhesion in the context of oogenesis. While Tcas wg (wingless) RNAi also interferes with egg production, similar to Tcas_arm1 RNAi, no reduction in egg production is seen in RNAi experiments with most other Wnt pathway components including Wntless, porc, and arrow (Bolognesi et al. 2008, 2009). It is thus seems reasonable to conclude that the egg production phenotype reflects abolishment of cell adhesion instead of the consequences of reduced Wnt signaling.

The defects in embryos obtained from animals injected with a low concentration of Tcas_arm1 dsRNA included severe tissue organization problems, the separation of embryonic from serosal tissue and the failure to complete dorsal closure. One common denominator of these defects is the deficiency in maintaining or establishing cell contacts that are exposed to heightened mechanical stress. These data thus lend further support to the assumed paralog-specific cell adhesion function of Tcas_arm1. This interpretation is backed by similar phenotypes in zygotic null arm Drosophila embryos, which are largely rescued by the α-Catenin binding domain-deleted arm allele S14 except for the failure to complete head inversion and, most intriguingly, dorsal closure (Orsulic and Peifer 1996). None of these defects were encountered in the low concentration Tcas_arm2 dsRNA knockdown progeny.

Select Preservation of Centrosome Regulation Function in Tcas_arm2

Tcas_arm2 dsRNA knockdown embryos were characterized by a cell division and chromosome maintenance phenotype not found in Tcas_arm1 embryos. This effect of Tcas_arm2 knockdown is highly reminiscent of the “nuclear fallout phenotype” (Buttrick et al. 2008) in arm-deficient syncytial Drosophila embryos, which was initially proposed to be caused by the failure to build an Arm-dependent molecular bridge between the actin cortex and the centrosomes (McCartney et al. 2001). It was further postulated that Drosophila Arm facilitated this link by binding to α-Catenin and APC2 (McCartney et al. 2001). Early observations linking α-Catenin to Arm in the context of spindle tethering in Drosophila included the GSK-3β-dependent colocalization of Arm, α-Catenin and also E-Cadherin to pseudofurrows (McCartney et al. 2001). However, the dependence of spindle organization on a direct interaction of Arm with α-Catenin was not specifically tested. Subsequent studies in Drosophila and mammalian systems revealed an interaction of β-Catenin/Arm with APC2 in the context of centrosome separation (Kaplan et al. 2004; Bahmanyar et al. 2008). The requirement of APC2 for normal centrosome separation was confirmed in studies testing APC’s presumed roles in cell adhesion and spindle formation with complete null-mutant forms of APC1 and APC2 (McCartney et al. 2006) or investigating the role of the GSK-3β regulator AKT (Buttrick et al. 2008). Most recently, the protein complex containing APC2 and Arm has been implicated in transferring to the centrosome the effect of cortical actin rearrangement, which is essential for centrosome separation (Cao et al. 2010). Of note, no further evidence of the postulated interaction between Arm and α-Catenin in the context of centrosome regulation has been reported. This is significant given that the Tcas_arm2-specific nuclear fallout phenotype is most parsimoniously explained by the preservation of an ancestral role in centrosome regulation in Tcas_arm2 that is not dependent on α-Catenin. Moreover, these data suggest for the first time that the roles of β-Catenin/Arm in centrosome separation and cell adhesion can be genetically separated.

Parallel Partial Subfunctionalization of β-Catenin/Arm in Pea Aphid and Red Flour Beetle

In combination, the shared contribution to Wnt signaling and the paralog-specific execution of the cell adhesion and centrosome regulation functions suggest a scenario of partial subfunctionalization in Tribolium (fig. 8). Neofunctionalization can be ruled out based on the ancestrality of the affected functions. Our data further suggest that the same trajectory unfolded in the pea aphid in parallel, considering the similarities in sequence evolution and the phylogenetic evidence of independent duplication events. Although it will be interesting to follow whether future genome projects will unravel additional examples, it seems already reasonable to conclude that this trajectory is evolutionarily favored. This is at least further backed by the outcomes of β-Catenin/Arm duplication in other metazoan groups discussed below.

Our model assigns differential losses of ancestral functions to Tcas_arm1 and Tcas_arm2. Only Tcas_arm2, however, is characterized by strong sequence divergence as predicted from release from or reduction of adaptive conflicts. A likely explanation is the continued interaction of APC with both paralogs, because APC interacts with β-Catenin/Arm in the context of Wnt signaling, which appears to be conserved in both paralogs. Thus, although Tcas_arm1 putatively lost the centrosome separation function, it continues to interact with APC like Tcas_arm2. The loss of cell adherens functionality in Tcas_arm2, by contrast, implies the release from Cadherin interaction allowing for the observed enhanced substitution accumulation. Nonetheless, considering the specificity of the cause and the evidence for many continued interactions (TCF, APC), the dramatic sequence change in Tcas_arm2 remains remarkable.

Our model makes a number of testable predictions. Most obviously, the differential deployment in cell adhesion and centrosome regulation should be reflected in differential binding affinities to cell adhesion and centrosome regulation-specific interacting proteins. We predict that the conservatively evolving Arm paralogs bind the cell adhesion complex components α-Catenin and E-Cadherin, whereas the highly modified paralogs lack this ability. If confirmed, the pea aphid and red flour beetle arm paralogs
would offer new avenues to dissect interaction-specific residues in the Arm repeat domain.

For instance, a particularly attractive explanation for excess substitutions in the Arm-repeat domains of the more divergent paralogs is that most of these substitutions resulted from the release of purifying selection on sites that are specifically responsible for the binding to E-Cadherin (fig. 8). Amino acid residues, which specifically experienced substitution in Tcas_arm2 and Aphis_arm2, therefore mark candidate E-Cadherin interaction sites in the conservatively evolved arm paralogs. Since E-Cadherin competes with other proteins for Arm binding such as APC and TCF (Hulsken et al. 1994; Roura et al. 1999; Huber and Weis 2001; Brembeck et al. 2004; Taurin et al. 2006; Hoffmans and Basler 2007), there are likely two classes of residues: One that is shared by competing proteins and remained conserved in the putative α-Catenin binding deficient paralogs. The second is specific for E-Cadherin and more variable in the arm2 paralogs. Intriguingly, this mix of conserved and nonconserved E-Cadherin interaction sites is seen in our data set. The five parallel substitutions in Tcas_arm2 and Aphis_arm2 therefore represent prime candidates for Cadherin-specific interaction.

Comparison to β-Catenin/Arm Duplications in Other Animal Groups

Given the extensive body of experimental data on β-Catenin, the independent duplications of this important gene in the pea aphid and the red flour beetle have also exceptional potential to further our understanding of causes and consequences gene duplicate subfunctionalization. The value of this system is further enhanced by comparison with other examples of β-catenin duplication, which range from the ancient subfunctionalized duplicates in vertebrates and nematodes (Korswagen et al. 2000) to a recent duplication in fish (Bellipanni et al. 2006). In the vertebrates, an early gene duplication produced the desmosome-mediated and thus cell adhesion–specific paralog plakoglobin. Since vertebrate β-Catenin remained responsible for all ancestral bilaterian β-Catenin functions, the outcome of this duplication resulted in an expanded cell adhesion repertoire. Another duplication occurred very recently in fish, generating the above mentioned β-Catenin-1 and β-Catenin-2 paralogs, which are characterized by an interesting mix of redundant and paralog-specific functions (Bellipanni et al. 2006).

Three ancient β-Catenin homologs have been described in the nematode Caenorhabditis elegans: wrm-1, bar-1, and hmp-2 (Korswagen et al. 2000). A fourth even more diverged homolog, sys-1, was recently discovered based on structural conservation and seems specific for Wnt signaling based on binding to the C. elegans TCF homolog POP-1 (Kidd et al. 2005; Liu et al. 2008). wrm-1 is involved in a diverged Wnt pathway in a manner that is not homologous to that of Drosophila arm (Korswagen et al. 2000). The ancestral cell adhesion and Wnt-signaling functions of β-Catenin have been partitioned between hmp-2 and bar-1, respectively. Remarkably, sys-1, wrm-1, and even hmp-2 can still rescue the transcriptional function of bar-1 despite their differential deployment during normal development (Natarajan et al. 2001; Liu et al. 2008). An intriguing question is whether the functional equivalence of sys-1, wrm-1, and even hmp-2 in this context reflects the coincidental effect of overlapping functional constraints in these paralogs or as yet unknown aspects of developmental buffering. Notwithstanding this open question, the available data suggest similar subfunctionalization scenarios of the duplicated β-Catenin genes in insects and nematodes that resulted in the separation of the cell adhesion and Wnt-signaling modalities. No role of the nematode β-Catenin homologs in centrosome regulation has been reported thus far, however.

Since Ohno’s (1970) pioneering investigation of the critical role of gene duplication in genetic and phenotypic diversification, substantial progress has been made in elucidating the mechanisms underlying the fixation, preservation, and functionalization of genes in duplicate (Dopman and Hartl 2007; Hahn et al. 2007; Lynch 2007; Emerson et al. 2008). At this time, many aspects of gene duplication have remained insufficiently understood (Hahn 2009). Further progress in this area will benefit from genome-wide analyses, but also detailed gene-specific case studies such as the one presented here.

Supplementary Material

Supplementary data files 1–7 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The Daphnia sequence data were produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) in collaboration with the Daphnia Genomics Consortium (http://daphnia.cgb.indiana.edu). R.B. was supported by a Wayne State Graduate Enhancement Research Fellowship. Work in the laboratory of M.F. was supported by National Science Foundation grant 0951886. The work of R.B., T.F., and S.J.B. was supported by National Institute of Health R01 HD029594. We thank Michelle Gordon for expert technical assistance and the two anonymous reviewers for excellent comments.

References


