

# MicroRNA evolution by arm switching

Sam Griffiths-Jones<sup>\*+,</sup> Jerome H.L. Hui<sup>\*</sup>, Antonio Marco & Matthew Ronshaugen<sup>++</sup>

Faculty of Life Sciences, University of Manchester, Manchester, UK

 This is an open-access article distributed under the terms of the Creative Commons Attribution Noncommercial No Derivative Works 3.0 Unported License, which permits distribution and reproduction in any medium, provided the original author and source are credited. This license does not permit commercial exploitation or the creation of derivative works without specific permission.

**MicroRNAs (miRNAs) modulate transcript stability and translation. Functional mature miRNAs are processed from one or both arms of the hairpin precursor. The miR-100/10 family has undergone three independent evolutionary events that have switched the arm from which the functional miRNA is processed. The dominant miR-10 sequences in the insects *Drosophila melanogaster* and *Tribolium castaneum* are processed from opposite arms. However, the duplex produced by Dicer cleavage has an identical sequence in fly and beetle. Expression of the *Tribolium* miR-10 sequence in *Drosophila* S2 cells recapitulates the native beetle pattern. Thus, arm usage is encoded in the primary miRNA sequence, but outside the mature miRNA duplex. We show that the predicted messenger RNA targets and inferred function of sequences from opposite arms differ significantly. Arm switching is likely to be general, and provides a fundamental mechanism to evolve the function of a miRNA locus and target gene network.**

Keywords: microRNA; evolution; Hox; *Drosophila*; *Tribolium*  
EMBO reports (2011) 12, 172–177. doi:10.1038/embor.2010.191

## INTRODUCTION

A key step in microRNA (miRNA) processing is the cleavage of the precursor stem-loop by the Dicer enzyme to produce an approximately 22-nucleotide RNA duplex (Lee *et al*, 2003). Data from cloning and expression studies have shown that for many miRNAs, one arm of the RNA duplex preferentially accumulates. This mature miRNA is often assumed to be the dominant functional product that is incorporated into the RNA-induced silencing complex to direct translational repression or transcriptional degradation of messenger RNA (mRNA) targets (Hutvagner, 2005). High-throughput sequencing can be used to quantify the production of miRNAs generated from each arm of the duplex

in different species, tissues and times in development. These data show that some miRNA precursors are processed to produce a single dominant mature miRNA, with a small proportion of reads originating from the opposite arm—often called the miR\* sequence. Previous models have suggested that the choice of dominant miRNA arm is determined by thermodynamic and structural properties of the processed duplex (Khvorova *et al*, 2003; Schwarz *et al*, 2003). Other miRNA precursors are processed to produce significant quantities of mature miRNAs from both arms. It has also been shown that miR\* sequences might function to downregulate targeted mRNAs (Okamura *et al*, 2008). Recent studies have demonstrated that sequences from opposite arms of a precursor miRNA can be sorted into different Argonaute (Ago) complexes, in which the dominant arm directs translational repression (by means of Ago1) and the miR\* sequence directs transcriptional degradation (by means of Ago2; Czech *et al*, 2009; Ghildiyal *et al*, 2009; Okamura *et al*, 2009).

Data from several large-scale miRNA sequencing studies have highlighted the possibility that the arm from which the dominant mature miRNA is processed can switch in different tissues and at different developmental times (Ro *et al*, 2007; Ruby *et al*, 2007; de Wit *et al*, 2009; Chiang *et al*, 2010). Although these studies have suggested that a change in arm usage provides a mechanism to evolve the function of a miRNA gene locus, the nature of the molecular events that lead to such changes has been largely unexamined. The potential functional consequences for evolutionary changes in the processing of the miRNA duplex are profound, as mature miRNAs from opposite arms might target different mRNAs. Here, we combine comparative genomics and deep sequencing to reconstruct the evolutionary history of changes in arm usage in the ancient miR-100/10 miRNA family. We use a cell culture-based assay to quantify miRNA abundance and determine that the sequences that direct this choice must be outside the miRNA duplex. Furthermore, we use computational prediction of miRNA targets in this model system to gain functional insights into the process of arm choice and evolution.

## RESULTS AND DISCUSSION

### Ancient duplications of miR-100 led to four subfamilies

Although the animal genomes studied so far contain hundreds of miRNA genes, deep sequencing and phylogenetic reconstruction

Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK

\*These authors contributed equally to this work

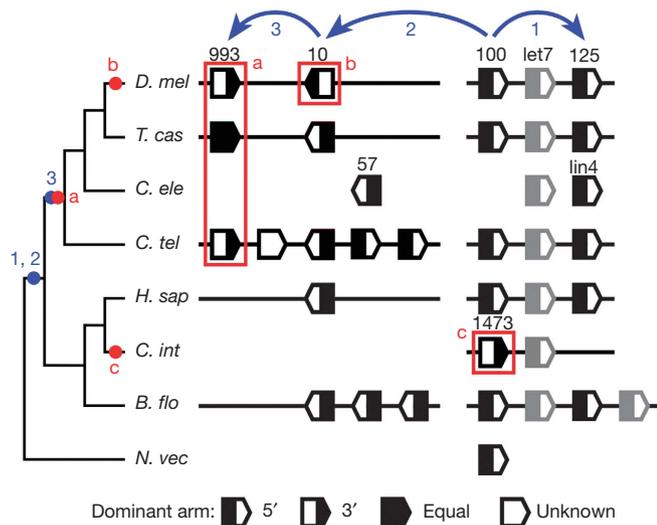
+Corresponding author. Tel: +44 (0)161 275 5673; Fax: +44 (0)161 275 1505;

E-mail: sam.griffiths-jones@manchester.ac.uk

++Corresponding author. Tel: +44 (0)161 275 3685; Fax: +44 (0)161 275 1505;

E-mail: matthew.ronshaugen@manchester.ac.uk

Received 6 August 2010; revised 21 October 2010; accepted 4 November 2010; published online 7 January 2011



**Fig 1** | The evolution of the miR-100/10 family. miR-10 family members are located in two clusters. Boxes represent the hairpin precursor sequences, with the direction of transcription shown by arrowheads and the dominant mature sequence filled. The left-hand cluster is located in the Hox complex, with Hox genes transcribed from 3' to 5' (not shown). Inferred duplication events (blue) are shown by arrows (top) and their evolutionary time is marked on the species tree (left). Arm switches in the ancestral miR-993 (a), *Drosophila* miR-10 (b), and *Ciona* miR-1473 (c) are boxed and labelled on the tree (red). Example genomes are shown. Genomic distances are not to scale. *B. flo*, *B. floridae*; *C. ele*, *C. elegans*; *C. int*, *C. intestinalis*; *C. tel*, *C. teleta*; *D. mel*, *D. melanogaster*; *H. sap*, *H. sapiens*; *N. vec*, *N. vectensis*; *T. cas*, *T. castaneum*.

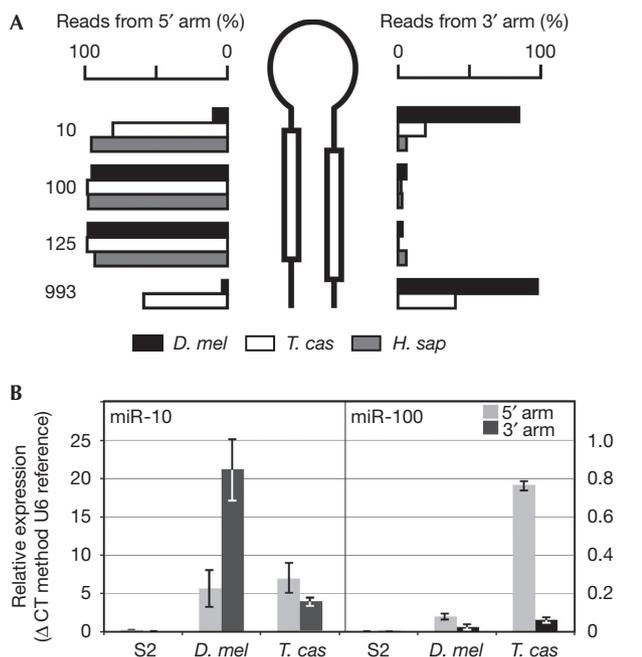
suggest that miR-100 is the only miRNA shared by all bilateria; conservation in the cnidarian *Nematostella vectensis* suggests that it appeared near the origin of multicellularity, around 650 million years ago (Grimson et al, 2008; Wheeler et al, 2009). Early duplication of miR-100 gave rise to at least three other ancient and conserved miRNA subfamilies in animals: miR-125, miR-10 and miR-993, collectively called the miR-100/10 family in this study. We used covariance model approaches, alignment inspection and synteny to exhaustively identify miR-100/10 family homologues in representative sequenced genomes (supplementary Fig S1 online). miR-100 and miR-125 are clustered in most bilaterian genomes with another ancient miRNA, let-7 (Sempere et al, 2003). miR-10 is located in a conserved position within the Hox complex of animals. The Hox complex is an ancient and highly conserved cluster of related transcription factors that specify segment identity in almost all animals, and originated before the divergence of protostomes and deuterostomes. The conserved genomic organization of the four miR-100/10 subfamilies allowed us to date the three duplication events that gave rise to these miRNAs to before the expansion of the Eumetazoa. The most parsimonious interpretation of patterns of conservation and genomic organization suggests that miR-10 and miR-125 derive from independent duplications of the ancestral miR-100 before diversification of protostomes and deuterostomes, whereas miR-993 arose from a later duplication of miR-10, soon after the divergence of protostomes (Fig 1).

Additional independent duplication events have expanded the miR-100/10 family. For example, multiple duplications of miR-10 have occurred in the lancelet *Branchiostoma floridae* and additional duplications of miR-10 and miR-993 can be found in the sequenced lophotrochozoan genomes (for example, *Capitella teleta*). Interestingly, these duplicates are located within Hox complexes. Whole-genome duplication events on the vertebrate lineage have also resulted in additional miR-100, miR-125 and miR-10 family members. Some miR-10 paralogues have been subsequently lost from duplicated Hox complexes, presumably through pseudogenization, facilitated by functional redundancy. In a few cases, the Hox complex structure has disintegrated, for example, in tunicates (*Ciona intestinalis*), nematodes (*Caenorhabditis elegans*) and planarians (*Schmidtea mediterranea*). In these genomes, the association of miR-100/10 family members with Hox genes is also lost.

### Arm choice has evolved in the miR-100/10 family

We sequenced a library of small RNAs from *Tribolium* using the ABI SOLiD platform (Marco et al, 2010). We further validated the relative expression of mature miRNAs of the miR-100/10 family in *Tribolium* by using quantitative PCR (supplementary Fig S3 online), and combined these data with those from next-generation sequencing experiments (Ruby et al, 2007; Stark et al, 2007; Wheeler et al, 2009; Marco et al, 2010) and high-throughput cloning (Landgraf et al, 2007) in other species. The data show that most mature miRNAs from the miR-100/10 family originate from the 5' arm of the precursor hairpin (Fig 2A). However, the dominant sequence from the *Drosophila* miR-10 locus originates from the 3' arm (Schwarz et al, 2003; Ruby et al, 2007; Stark et al, 2007), as do all previously validated miR-993 sequences in *D. melanogaster* (Ruby et al, 2007), *Locusta migratoria* (Wei et al, 2009), *Procambarus clarkii* (Wheeler et al, 2009) and the *Ciona* miR-100 orthologue miR-1473 (Shi et al, 2009; Fig 1). This indicates that the miR-100/10 family has undergone a minimum of three independent arm switching events (Fig 1). Evidence from the evolution of the miR-100/10 family shows that arm usage is more labile than thought previously and that the function of some miRNAs is likely to have evolved through changes that control the production of the dominant arm.

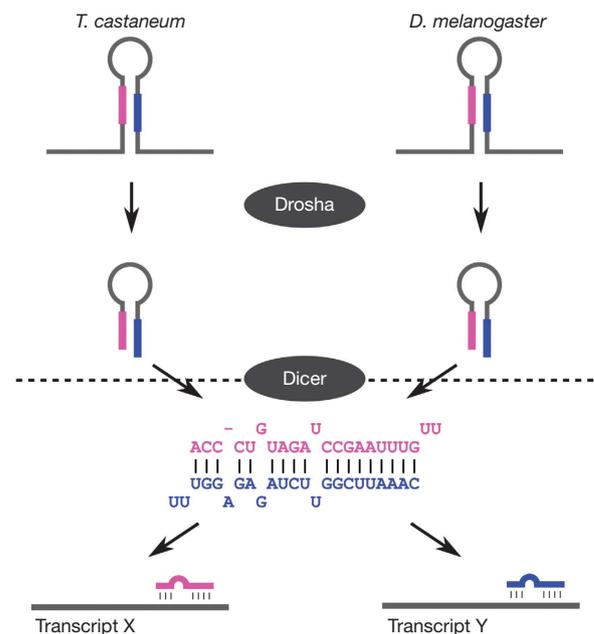
Previous studies have suggested that arm choice is governed by the asymmetrical stability of the miR/miR\* duplex resulting from Dicer cleavage (Schwarz et al, 2003; Hutvagner, 2005). This model implies a passive mechanism driven primarily by thermodynamic properties of the processed duplex. We have found that although the *D. melanogaster* and *T. castaneum* miR-10 sequences show opposite dominant arm usage, the miR/miR\* duplexes are identical in sequence and, as such, have identical thermodynamic properties. This observation suggests that the thermodynamic model alone is unlikely to control the choice of arm. To examine the influence of the cellular environment on RNA stability and arm choice, we quantified mature miRNA levels in *Drosophila* S2 cells. We constructed and expressed plasmids to drive the expression of 300-nucleotide transcripts derived from the miRNA primary transcript centred around the miR-10 precursor from both *Tribolium* and *Drosophila*. Relative steady-state production of mature miRNA sequences from 5' and 3' arms of both sequences in *Drosophila* S2 cells was assayed by using TaqMan miRNA quantitative PCR assays (Fig 2B). These experiments show



**Fig 2** | Arm-specific miRNA processing in miR-100/10 family members. (A) Relative expression of mature miRNAs from 5' and 3' arms. Schematic of a precursor hairpin (centre). The relative abundance of the 5' and 3' arms of the four miR-100/10 subfamilies is shown, estimated by high-throughput sequencing reads (*D. mel*; Ruby et al, 2007 and *T. cas*; supplementary Fig S2 online) and cloning frequencies (*H. sapiens*; Landgraf et al, 2007). (B) Arm-specific miRNA processing in *Drosophila* S2 cells. *Tribolium* and *Drosophila* miR-10/miR-100 constructs were transfected into S2 cells and expressed under the control of the actin promoter. TaqMan miRNA real-time PCR assays were used to estimate 5' and 3' miRNA arm levels. Error bars show the s.e.m. of three technical replicates of three biological experiments. *D. mel*, *D. melanogaster*; *H. sap*, *H. sapiens*; miRNA, microRNA; *T. cas*, *T. castaneum*.

that the dominant *D. melanogaster* miR-10 sequence is processed from the 3' arm, in accordance with the endogenous profile observed in high-throughput sequencing experiments. The profile of arm use from the *T. castaneum* miR-10 sequence in *Drosophila* S2 cells shows a near 2:1 dominance of the 5' arm, similar to the pattern observed in *Tribolium* RNA deep-sequencing experiments. We further expressed constructs containing *Tribolium* and *Drosophila* miR-100 precursors in *Drosophila* S2 cells, and the endogenous profile of arm use was recapitulated in both cases. The observation that identical miRNA duplexes can have opposite arm usage profiles does not support a purely thermodynamic model of arm choice (Khvorova et al, 2003; Schwarz et al, 2003). These data support a model in which the sequence signals determining this choice are located outside this duplex (in the loop and/or flanking regions), and that control of arm use is regulated before Dicer processing (Fig 3).

We used evolutionary and expression analysis to identify examples of arm-switching events between both paralogous and orthologous members of the miR-100/10 family. Data also demonstrate that arm switches can occur in different tissues and across a developmental time course. Indeed, although the 5' arms



**Fig 3** | A model of the differential processing of the miR-10 transcript. The precursor miRNA hairpin is excised from the primary miRNA transcript by Drosha. Dicer cleavage yields the mature miRNA duplex, and the dominant arm targets the RNA-induced silencing complex to mRNA transcripts. The sequences of the mature miRNA duplex produced by Dicer cleavage are identical in *D. melanogaster* and *T. castaneum*; yet, the 5' arm (magenta) dominates in *Tribolium* and the 3' arm (blue) dominates in *Drosophila*. The sequence signals controlling arm dominance are therefore outside the mature miRNA duplex, and the arm choice must be specified before Dicer cleavage (that is, above the dotted line). miRNA, microRNA.

of miR-100 and miR-125 are dominant in mammals, some tissues have a profile with the 3' mature sequence favoured (Landgraf et al, 2007). Many chicken miRNA families show switches in dominant arm use in a developmental time course (Glazov et al, 2008). Different mouse tissues have also been shown to have switches in dominant miRNA arm usage (Ro et al, 2007; Chiang et al, 2010). The miR-10 sequence does not seem to have significant tissue- or stage-specific arm switching in *Drosophila* or *Tribolium* (supplementary Fig S4 and Table S1 online). Nonetheless, these observations support our model that arm choice is specified outside the mature duplex.

### Arm switching impacts mRNA target regulation

The sequences of mature miRNA products from 5' and 3' arms are different. Opposite arms might therefore regulate the expression of either distinct sets of mRNA transcripts or common targets through different sites in the target mRNA. We used two methods (seed-based and non-seed-based) to predict targets of the 5' and 3' arms of miR-10, miR-993, miR-100 and miR-125 in *Drosophila* untranslated regions (UTRs; Table 1). It is interesting to note that, in each case, the number of target sites predicted for the dominant miRNA arm was fewer than those for the other arm. This suggests a selective pressure against the occurrence of miRNA target sites. We found that target transcripts for each *D. melanogaster* 5'/3'

**Table 1** | Predicted targets of mature miRNAs derived from the 5' and 3' arms of miR-100/10 family members

Target type	miRNA	5' Targets	3' Targets	Common targets	Ln HP*	P-value** (random)	P-value** (shuffling)
Seed based	miR-10	2,130	252	67	-25.56	0.515	0.389
	miR-100	156	3,262	29	-1.06	0.979	0.984
	miR-125	209	1,680	29	-4.61	0.887	0.891
	miR-993	247	430	6	-0.72	0.986	0.991
miRanda	miR-10	742	602	95	-72.5	0.319	0.723
	miR-100	519	1,732	168	-117.15	0.272	0.516
	miR-125	864	843	166	-140.42	0.622	0.453
	miR-993	3,691	1,153	592	-321.15	0.957	0.518

We observed no evidence for 5' and 3' miRNA pairs having more targets in common than is expected by chance ( $P \gg 0.05$ ). Additionally, the probability values for the four miR/miR\* pairs studied here were not significantly different to the set of all other possible 5' and 3' pairwise comparisons in our data set ( $P = 0.22$ ; one-tailed Mann-Whitney test). \*Ln HP: logarithm of the cumulative hypergeometric probability. \*\*P-value: unbiased probability that both arms of the miRNA binds to more genes in common than expected by chance for two sets of null distributions, namely shuffling and random. For miR-100, miR-125 and miR-993 3' arms, we took the sequence in the complementary arm slipped by two nucleotides, as no information for miR\* exists in miRBase. miRNA, microRNA.

pair do not overlap more than expected by chance. The functional consequences of arm switching are therefore profound: homologous miRNA loci can generate distinct functions by regulating arm usage.

Moreover, miRNA arm switching is probably frequent. We observed two types of arm switching in the miR-100/10 family alone. First, the early duplication of miR-10 giving rise to miR-993 is associated with a switch of dominant arm in the resulting miR-993 family. In the general case, duplication of an ancestral precursor that generates mature sequences from both arms—with each arm having unique targets—might lead to partitioning of targets between paralogues by regulated arm choice (sub-functionalization). Alternatively, duplication of a miRNA with a single dominant product, followed by arm switching, generates a new function (neo-functionalization; Ruby *et al*, 2007). Indeed, de Wit *et al* (2009) recently demonstrated an arm-switching event in paralogous copies of the *Caenorhabditis*-specific miR-246 family. The second type of change in arm usage is seen within the miR-10 family, in which the dominant arm has switched in *Drosophila*. This event provides a mechanism to alter the targets and function of orthologous sequences in related organisms. In the general case, the change in miRNA arm use might be a near-complete switch—as in miR-10—or a more subtle tuning of steady-state levels of miRNAs. Comparison between high-throughput sequencing data suggests that over 10% (5/46) of orthologous miRNAs in *D. melanogaster* and *T. castaneum* have shown a pronounced switch in the arm from which the dominant mature sequence is produced (Marco *et al*, 2010). Further, more than 40% of one-to-one orthologous miRNAs show a tuning in arm use in which there is a difference of more than tenfold (Marco *et al*, 2010). Understanding the mechanisms that control the switching or tuning events is vital for correct experimental interpretation in model systems; for example, *Drosophila* miR-10 will be an excellent model for understanding the functional consequences of arm switching, but is not likely to be a good model for the conserved functions of miR-10 in other animals.

The concept of a single functional miRNA processed from a precursor is clearly an oversimplification. Functional sequences from both arms might be produced, with differential regulation of

orthologous sequences in different organisms, and even of the same sequence in different tissues and stages. The sequences from opposite arms have distinct targets, possibly mediated by alternative Ago complexes (Czech *et al*, 2009; Ghildiyal *et al*, 2009; Okamura *et al*, 2009). The ability to regulate the expression of sequences from different arms will have a profound effect on the targets and functions of a subset of miRNAs.

## METHODS

**miR-100/10 family homologue detection and alignment.** Members of the miR-100, miR-10, miR-993 and miR-125 subfamilies and the let-7 family were retrieved from the miRBase database (Griffiths-Jones *et al*, 2008) and aligned with ClustalW (Larkin *et al*, 2007). A consensus secondary structure was predicted using RNAalifold from the ViennaRNA package (Bernhart *et al*, 2008). A covariance model was built from the alignment and searched against fully sequenced bilaterian genomes using INFERNAL 1.0 (Nawrocki *et al*, 2009). The genomes and assemblies searched were *N. vectensis* (JGI 1.0), *D. melanogaster* (BDGP 5), *T. castaneum* (Baylor 2.0), *C. teleta* (JGI 1.0), *B. floridae* (JGI 2.0), *Homo sapiens* (NCBI36), *C. intestinalis* (JGI 2.0) and *C. elegans* (Wormbase WS200). Newly identified homologues were realigned to the model using calign from INFERNAL 1.0. The alignment, structure and model were refined by an iterative procedure of covariance model searches, realignment, secondary structure prediction and manual editing (RALEE; Griffiths-Jones, 2005). Subfamilies were classified by analysis of syntenic regions: genomic regions containing the Hox complex were manually annotated using BLASTX and BLASTP to identify orthologous Hox genes by position, sequence conservation of homeodomains and diagnostic amino-acid motifs (for example, the hexapeptide and UBDA motif). miR-10 and miR-993 homologues were classified by position in the Hox complex. miR-100 and miR-125 homologues were classified by proximity to members of the let-7 miRNA family.

**Short RNA read mapping.** Wild-type *Tribolium* were reared at 28 °C in wholewheat flour supplemented with 1% yeast extract. Small RNA was extracted from mixed-sex adults using the

miRVana miRNA isolation kit (Ambion). RNA was size selected (for less than 40 nucleotide) with a flashPAGE fractionator (Ambion). Purification was then performed with a flashPAGE reaction clean-up kit (Ambion). A library was constructed using the SOLiD Small RNA Expression Kit, according to the manufacturer's instructions (Ambion). To meet the sample quantity for SOLiD sequencing, the complementary DNA libraries were then further amplified using 15 cycles of PCR. The final products ranging from approximately 105 to 150 bp were purified and size selected. SOLiD sequencing was carried out at the Centre for Genomic Research at the University of Liverpool. Adaptor sequences were removed from the 3' ends of reads, and reads were mapped to the Baylor release 3.0 of the *Tribolium* genome using Bowtie 0.11, with no mismatches allowed (Langmead et al, 2009). Reads mapping to more than one genomic position were discarded.

**Prediction of target genes of miR/miR\* pairs.** We scanned the collection of 3' UTRs in *D. melanogaster* (available at [http://flybase.org/genomes/Drosophila\\_melanogaster/current/fasta](http://flybase.org/genomes/Drosophila_melanogaster/current/fasta)) for canonical miRNA target sites (Bartel, 2009) and potential targets on the basis of RNA–RNA duplex stability, using mature miRNA sequences derived from the 5' and 3' arms of all four miR-100/10 family members (Enright et al, 2003). Where no miR\* sequence is reported in the miRBase database, the miR\* sequence was deduced from the hairpin structure. We then calculated the probability that 5' and 3' mature sequences (the miR/miR\* duplex) share more common gene targets than expected by chance. This probability is given by the cumulative hypergeometric distribution (Sokal & Rohlf, 1995), and is equivalent to Fisher's exact test. As these probability values are highly dependent on the length and nucleotide composition of the 3' UTRs, we estimated unbiased *P*-values by simulating, for each miR/miR\* pair, sets of 1,000 synthetic miR/miR\* pairs to generate a null distribution of probability values. Two independent sets of null distributions were generated. In the first, a synthetic mature miRNA was produced by shuffling the sequence of the known mature miRNA. Thereafter, a synthetic miR\* was generated by taking the reverse complement of the synthetic miR sequence, adding a random two-nucleotide 3' overhang sequence and introducing bulges in the duplex by random mutation of the synthetic miR\* sequence, with a probability of 0.2 per nucleotide. These parameters were based on empirical observations of animal miRNAs from the miRBase database. In the second set, synthetic miR/miR\* pairs were randomly generated.

**Construction of miR-10- and miR-100-expression constructs and transfection of S2 cells.** Expression constructs containing 250 bp flanking the miR-10 and miR-100 precursor sequences from *D. melanogaster* and *T. castaneum* were cloned into the pAc5.1 vector (Invitrogen) using the following primers: dme-miR-10-FP-KpnI 5'-TATATTTAGGGGTACCGCGATTGCCTAGCGGACTT CATT-3'; dme-miR-10-RP-EcoRI 5'-TTAAGGGAATCCACTTT CCGCTTGCCATCAGCAACACTT-3'; tca-miR-10-FP-KpnI 5'-ATA AGGGGTACCTTTGTTGCAGGTTTTCTCAAGA-3'; tca-miR-10-RP-EcoRI 5'-GGGGAATCCGATTTATGTCTAAGGACAATCT-3'; dme-100-pac-FP-KpnI 5'-TTTGGGTACCCTTCGATTCATCGAGACACATTGAAGTTCACGA-3'; dme-100-pac-RP-EcoRI 5'-CAGG GAATTCCTTGACTCGACTTTTTATCCTTACTCCGCCATT-3'; tca-100-pac-FP-EcoRI 5'-CAGGGAATCCCATGCTAAAAATATATAT CTCCTGGACTGA-3'; tca-100-pac-RP-NotI 5'-AATTGCGGCCGC

CTCAAACAGACCCTGCCACAAATCTAGTACTT-3'; sequence-verified constructs were transfected into S2 cells using 1.5 µg pAc5.1 plasmids and Effectene (Qiagen) according to the manufacturer's instructions. Small RNAs were extracted from both transfected and control S2 cells 48 h post-transfection, using the miRVana miRNA isolation kit (Applied Biosystems). Quantification of respective miRNA arm use (5' and 3') was carried out with TaqMan miRNA real-time PCR assays (Applied Biosystems). Samples from three independent transfections were assayed in triplicate. The U6 small RNA TaqMan assay (Ambion) was used for normalization.

**Supplementary information** is available at EMBO reports online (<http://www.emboreports.org>).

#### ACKNOWLEDGEMENTS

We thank Kevin Peterson for sharing data before publication, the University of Liverpool Centre for Genomic Research for SOLiD sequencing, Ana Kozomara for helping with miRNA expression data sets and Casey Bergman for helpful comments on the manuscript. This work was funded by the University of Manchester (fellowships to S.G.-J. and M.R.), the Wellcome Trust VIP scheme and the Biotechnology and Biological Sciences Research Council (BB/G011346/1).

#### CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

#### REFERENCES

- Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* **136**: 215–233
- Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF (2008) RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics* **9**: 474
- Chiang HR et al (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* **24**: 992–1009
- Czech B, Zhou R, Erlich Y, Brennecke J, Binari R, Villalta C, Gordon A, Perrimon N, Hannon GJ (2009) Hierarchical rules for Argonaute loading in *Drosophila*. *Mol Cell* **36**: 445–456
- de Wit E, Linsen S, Cuppen E, Berezikov E (2009) Repertoire and evolution of miRNA genes in four divergent nematode species. *Genome Res* **19**: 2064–2074
- Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in *Drosophila*. *Genome Biol* **5**: R1
- Ghildiyal M, Xu J, Seitz H, Weng Z, Zamore PD (2009) Sorting of *Drosophila* small silencing RNAs partitions microRNA\* strands into the RNA interference pathway. *RNA* **16**: 43–56
- Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP, Tizard ML (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* **18**: 957–964
- Griffiths-Jones S (2005) RALEE–RNA ALignment editor in Emacs. *Bioinformatics* **21**: 257–259
- Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**: D154–D158
- Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, Degnan BM, Rokhsar DS, Bartel DP (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**: 1193–1197
- Hutvagner G (2005) Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. *FEBS Lett* **579**: 5850–5857
- Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**: 209–216
- Landgraf P et al (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**: 1401–1414
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25
- Larkin MA et al (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948
- Lee Y et al (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**: 415–419

- Marco A, Hui JH, Ronshaugen M, Griffiths-Jones S (2010) Functional shifts in insect microRNA evolution. *Genome Biol Evol* **2**: 686–696
- Nawrocki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25**: 1335–1337
- Okamura K, Phillips MD, Tyler DM, Duan H, Chou YT, Lai EC (2008) The regulatory activity of microRNA\* species has substantial influence on microRNA and 3' UTR evolution. *Nat Struct Mol Biol* **15**: 354–363
- Okamura K, Liu N, Lai EC (2009) Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. *Mol Cell* **36**: 431–444
- Ro S, Park C, Young D, Sanders KM, Yan W (2007) Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res* **35**: 5944–5953
- Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res* **17**: 1850–1864
- Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**: 199–208
- Sempere LF, Sokol NS, Dubrovsky EB, Berger EM, Ambros V (2003) Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-complex gene activity. *Dev Biol* **259**: 9–18
- Shi W, Hendrix D, Levine M, Haley B (2009) A distinct class of small RNAs arises from pre-miRNA-proximal regions in a simple chordate. *Nat Struct Mol Biol* **16**: 183–189
- Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn. New York: W. H. Freeman & Co
- Stark A et al (2007) Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* **450**: 219–232
- Wei Y, Chen S, Yang P, Ma Z, Kang L (2009) Characterization and comparative profiling of the small RNA transcriptomes in two phases of locust. *Genome Biol* **10**: R6
- Wheeler BM, Heimberg AM, Moy VN, Sperling EA, Holstein TW, Heber S, Peterson KJ (2009) The deep evolution of metazoan microRNAs. *Evol Dev* **11**: 50–68



EMBO reports is published by Nature Publishing Group on behalf of European Molecular Biology Organization. This article is licensed under a Creative Commons Attribution Noncommercial No Derivative Works 3.0 Unported License [<http://creativecommons.org/licenses/by-nc-nd/3.0>]