J Mol Evol (2008) 67:560–569
DOI 10.1007/s00239-008-9174-5

In Silico Genetic Robustness Analysis of Secondary Structural Elements in the *miRNA* Gene

Wenjie Shu · Ming Ni · Xiaochen Bo · Zhiqiang Zheng · Shengqi Wang

Received: 9 July 2007 / Accepted: 6 October 2008 / Published online: 22 October 2008

Abstract Genetic robustness, insensitivity of the phenotype facing genetic mutations, is a fundamental and ubiquitously observed property of biological systems. In this study, we investigate the genetic robustness of the structural elements within native miRNA genes on a genome-wide scale. MicroRNAs (miRNAs) are a large family of endogenous noncoding RNAs that regulate gene expression at the posttranscriptional level. We examine the neutrality of the structural element in 1082 native pre-miRNAs from six species and demonstrate that the structural elements in native pre-miRNAs exhibit a significantly higher level of genetic robustness in comparison with structural elements within random pseudo pre-miRNAs. Hence, this excess robustness of structural elements in pre-miRNAs goes beyond the intrinsic robustness of the stem-loop structure. Furthermore, we show that it is not a by-product of a base composition bias. Interestingly, our data also demonstrate a difference in increased levels of average neutrality between structural elements. Remarkably, differential genetic robustness between structural elements is observed in both native and pseudo pre-miRNAs. Our results are much in agreement with previous experimental observations, and suggest that the genetic robustness of secondary structural elements in native pre-miRNAs, under different evolutionary selection pressures, may evolve due to its own selective advantage.

Keywords Genetic robustness · Secondary structural elements · MicroRNA · Evolution

The authors wish it to be known that, in their opinion, the authors Wenjie Shu and Ming Ni should be regarded as joint first authors.

Electronic supplementary material The online version of this article (doi:10.1007/s00239-008-9174-5) contains supplementary material, which is available to authorized users.

W. Shu · M. Ni · X. Bo (✉) · S. Wang (✉)
Beijing Institute of Radiation Medicine, Beijing 100850, China
e-mail: boxc@bmi.ac.cn

S. Wang
e-mail: sqwang@bmi.ac.cn

W. Shu · Z. Zheng
College of Electro-Mechanic and Automation,
National University of Defense Technology,
Changsha, Hunan 410073, China

Introduction

Robustness, a fundamental and ubiquitously observed phenomenon in biological systems, is defined as the ability to maintain stable functioning in the face of various perturbations (Kitano 2004). The robustness of phenotypes appears at various levels of biological systems, including gene expression, protein folding, metabolic flux, physiological homeostasis, development, and even organism fitness (de Visser et al. 2003). It is consequently not surprising that biologists have a long-standing interest in robustness, going back to Fisher's (1928a, b, 1931) work on dominance and Waddington's (1953, 1957) developmental canalization research. Kitano (2004) argues that the requirements for robustness and evolvability are similar, since robustness facilitates evolution and evolution favours robust traits. And thus, a proper understanding of the origin of robustness in biological systems will catalyze our understanding of evolution (Wagner 2005).

Depending on whether or not the perturbations are heritable, robustness is characterized as genetic or environmental robustness (Wagner et al. 1997). Here we primarily focus on the first kind of robustness—genetic

robustness—which describes insensitivity of the phenotype facing genetic mutations. Recently, genetic robustness has become the focus of numerous studies and has been found in RNA viruses (Elena et al. 2006; Montville et al. 2005; Wagner and Stadler 1999), viroids (Sanjuan et al. 2006a, b), and micro RNAs (miRNAs) (Bonnet et al. 2004; Borenstein and Ruppin 2006; Shu et al. 2007a). Despite the plethora of observations of genetic robustness, however, its evolutionary origins are still not clear. Whether it is a consequence of natural selection or a nonadaptive correlated side effect of other phenotypic traits is by and large unknown. Part of the problem lies in the difficulty in providing evidence for genetic robustness in natural biological systems (Gibson and Wagner 2000). The classical approach has inferred genetic robustness from increases in genetic variance observed after a major mutation or environmental challenge (de Visser et al. 2003), as exemplified by number of vibrissae in mice, number of ocelli in *D. subobscura*, and wing- and cross-vein interruptions and scutellar bristle numbers in *D. melanogaster* (Scharloo 1991). However, the evidence is often indirect and suffers from a lack of a natural reference genotype (Gibson and Wagner 2000). Experimental evolution is a more direct approach that has been applied to the study of robustness recently, which utilizes direct laboratory observation of short-term evolutionary processes, mostly in microbes (Fares et al. 2002; Elena and Lenski 2001). Although its evolutionary potential is limited by time constraints, this approach does not suffer from a lack of control and promises exciting new data and insights for a more comprehensive theory of the evolution of genetic robustness (de Visser et al. 2003).

We have previously developed a bioinformatics approach to quantitatively measure the genetic robustness of RNA secondary structure (Shu et al. 2007b). Here, we apply this method to investigate the genetic robustness of the structural elements in 1082 native pre-miRNAs from six species. miRNAs are abundant, endogenous, ~22-nucleotide (nt) noncoding RNAs that regulate gene expression at the post-transcriptional level for cleavage and/or translational repression through binding of a minimal-recognition 'seed' sequence (Bartel 2004; Lau et al. 2001; Lagos-Quintana et al. 2001; Lee and Ambros 2001). The short mature miRNAs (~22 nt) are cleaved from ~70-nt precursors (pre-miRNA) that fold into a stem-loop hairpin structure, through the action of the Dicer endonuclease (Hutvagner et al. 2001; Lee et al. 1993; Ketting et al. 2001). The stem-loop hairpin structure is critical for pre-miRNA recognition by Drosha (Lee et al. 2003) and Dicer (Hutvagner et al. 2001) RNAase III enzymes and for nuclear export of pre-miRNAs (Lund et al. 2004; Zeng and Cullen 2004). Recent comparative phylogenetic studies uncovered conserved miRNA-binding sequences in more than one-third of all genes, suggesting that miRNAs may regulate a large portion of cellular processes (Lewis et al. 2005; Brennecke et al. 2005; Farh 2005; Xie 2005; Stark et al. 2005; Grun et al. 2005).

The secondary structure of pre-miRNA provides an ideal test bed for studying genetic robustness. Two recent studies report that the stem-loop structures of pre-miRNAs exhibit a significantly higher level of genetic robustness, which goes beyond the intrinsic robustness of the stem-loop structure and is not a by-product of the base composition bias (Borenstein and Ruppin 2006; Shu et al. 2007a). Recently, researchers have focused on the study of differential genetic robustness. Using a phylogenetic shadowing approach (Boffelli et al. 2003), Berezikov et al. (2005) reported that strong conservation is observed in stem elements of miRNA stem-loops, known to be indispensable for miRNA biogenesis, as well as increased variation in loop elements, inferred to have reduced or no functional constraints on the sequence of these elements of pre-miRNA. While experimental studies have shown that pre-miRNAs are quite tolerant of mutations, influences on pre-miRNAs caused by point mutations vary greatly among stem-loop structural elements (Lee et al. 2003; Zeng et al. 2003; Zeng and Cullen 2003). Mutations in the loop are often inconsequential, however, alterations that disrupt base-pairing at the base of the stem have considerably deleterious effects. In this study, we hypothesize that this is a universal feature for all miRNA stem-loops. To our knowledge, no systematic effort has been made to test this hypothesis on a genome-wide scale, with the exception of a few experimental studies on limited pre-miRNAs (Lee et al. 2003; Zeng et al. 2003; Zeng and Cullen 2003).

Materials and Methods

Native Pre-miRNAs, Reference Sets, and RNA Folding

The 1082 native pre-miRNA sequences included in the analysis were selected from MicroRNA Registry release 7.1, and all have been experimentally verified to avoid a possible bias introduced by consideration of predicted pre-miRNAs (Table 1) (Griffiths-Jones 2004; Griffiths-Jones et al. 2006). The available sequences cover six species: *H. sapiens*, *D. melanogaster*, *D. rerio*, *C. elegans*, *M. musculus*, and *R. norvegicus*.

In addition to the native pre-miRNAs specified in Table 1, we generated a reference set for each species that consisted of 10 hairpin sequences with similar phenotypes to native pre-miRNAs (random pseudo pre-miRNAs) for each native pre-miRNA using a method in our previous study (Shu et al. 2007a). Furthermore, to rule out the effect of base composition bias in the analysis of structural elements robustness, we made four types of shuffling

Table 1 Native pre-miRNA data sets

Species	N_s	Overhang (%)	Length (nt)	%GC
H. sapiens	242	84 (34.7%)	85 ± 14	47.67 ± 7.88
C. elegans	112	83 (74.1%)	98 ± 6	44.61 ± 6.89
D. melanogaster	75	27 (36.0%)	88 ± 13	41.60 ± 5.39
D. rerio	350	112 (32.0%)	94 ± 18	45.11 ± 6.51
M. musculus	191	89 (46.6%)	80 ± 12	48.19 ± 8.26
R. norvegicus	112	86 (76.8%)	91 ± 9	50.34 ± 7.96

Note: N_S number of experimentally verified miRNAs in species. Overhang (%): number (percentage) of pre-miRNAs that contain overhang elements. Length: distribution of length in species. %GC: GC contents of sequence in species. The secondary structure of pre-miRNA was predicted by *RNAfold* in the Vienna RNA package (version 1.6)

reference sets for each species using four different sequence shuffling methods (Bonnet et al. 2004; Clote et al. 2005; Katz and Burge 2003; Workman and Krogh 1999; Shu et al. 2007b), namely, mononucleotide shuffling, dinucleotide shuffling, and shuffling based on a zero- and first-order Markov chain. Every shuffling reference set for each species also consisted of 10 shuffling pseudo pre-miRNA sequences that preserve not only the stem-loop structure, but also the exact or nearly exact mononucleotide and dinucleotide frequencies for each native pre-miRNA.

We apply the RNAfold in the Vienna RNA package (version 1.6) (Hofacker et al. 1994) utilizing default parameter values ($T = 37$°C) to predict secondary structures of pre-miRNA sequences, based on Zuker's minimum free energy algorithm (Zuker and Stiegler 1981), and only utilized the results of the optimal folding.

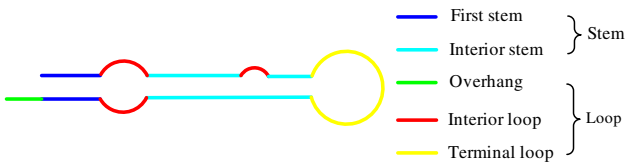

Fig. 1 Structural elements of the miRNA stem-loop. The first stem element and the internal stem make up the stem element, while the loop element is composed by the overhang element, internal loop, and terminal loop elements

Neutrality of Structural Elements in the Stem-Loop Structure

The secondary structure of pre-miRNA can be viewed as a combination of basic structural elements, namely, the first stem element, internal stem element, overhang element, internal loop element, and terminal loop element (Fig. 1). The first stem element and the internal stem element make up the stem element, while the loop element is composed of the overhang, internal loop, and terminal loop elements. To investigate the genetic robustness of the structural elements in pre-miRNAs, we define the structural elements set of the miRNA stem-loops, in which $E_1, \ldots, E_6$ represent the stem element, first stem element, loop element, overhang element, internal loop element, and terminal loop element, respectively. The distributions of length of structural elements in native pre-miRNAs of different species are listed in Table 2.

For a secondary structure of RNA, a much more strict definition of neutrality, which does not assume any structure distance metric, is defined as the fraction of the $3 \times l$ one-mutant neighbors that perfectly preserves the original structure (Borenstein and Ruppin 2006; Shu et al. 2007a, b). In this study, this definition is applied to the structural elements of the miRNA stem-loop. For each element of a hairpin, the neutrality ξ_k of the element E_k is defined as the fraction of the $3 \times l_k$ structures in the element E_k that perfectly preserves the original structure after a mutation occurs. Formally, the neutrality ξ_k is defined as

$$\xi_k = \frac{N_k}{3 \times l_k}, \quad k = 1, 2, \cdots, 6 \tag{1}$$

where N_k is the number of mutants in element E_k that perfectly preserves the original structure, and l_k is the length of the element E_k. Greater genetic robustness of a given stem-loop element E_k can be inferred from a higher value of neutrality ξ_k of that element.

Table 2 Distribution of length of secondary structural element in native pre-miRNAs for different species

Species	Stem (bp)	First stem (bp)	Loop (nt)	Overhang (nt)	Internal loop (nt)	Terminal loop (nt)
H. sapiens	31 ± 5	5 ± 3	23 ± 7	3 ± 3	13 ± 6	7 ± 3
C. elegans	33 ± 4	5 ± 2	32 ± 8	8 ± 5	17 ± 7	8 ± 4
D. melanogaster	30 ± 4	5 ± 3	27 ± 8	5 ± 5	15 ± 7	8 ± 4
D. rerio	33 ± 7	5 ± 2	28 ± 8	4 ± 4	16 ± 8	8 ± 4
M. musculus	29 ± 5	5 ± 3	22 ± 6	3 ± 3	12 ± 6	7 ± 3
R. norvegicus	32 ± 4	5 ± 2	27 ± 7	5 ± 5	15 ± 6	7 ± 3

Note: Mean ± SD length of stem elements (bp) and loop elements (nt) in native pre-miRNAs

Table 3 Robustness comparisons

	Stem	First stem	Loop	Overhang	Internal loop	Terminal loop
p^r	0	4.10×10^{-106}	1.18×10^{-95}	5.55×10^{-187}	1.56×10^{-156}	1.80×10^{-62}
p^{m0}	0	3.54×10^{-102}	6.73×10^{-89}	1.04×10^{-177}	2.95×10^{-154}	1.70×10^{-53}
p^{m1}	0	8.47×10^{-115}	14.99×10^{-91}	7.55×10^{-186}	4.84×10^{-143}	7.28×10^{-55}
p^{mono}	0	2.31×10^{-109}	1.63×10^{-78}	2.40×10^{-151}	1.16×10^{-126}	1.01×10^{-38}
p^{di}	0	6.42×10^{-148}	6.05×10^{-62}	2.15×10^{-177}	3.01×10^{-96}	1.11×10^{-54}

Note: The neutrality of the structural elements in native pre-miRNAs is compared with that of corresponding elements in random pseudo pre-miRNA by *t*-tests. p^r, p^{m0}, p^{m1}, p^{mono}, and p^{di} represent *p*-values comparing random pseudo pre-mi10RNAs and shuffling pseudo pre-miRNAs based on the zero-Markov, first-Markov, mononucleotide, and dinucleotide model, respectively

Robustness and Significant Analysis

For each structural element, we evaluate the neutrality, ξ_k^m ($k = 1, 2, \cdots, 6$), in all native pre-miRNAs included in the analysis (Table 3), and measure the neutrality of pseudo pre-miRNAs, ξ_k^r ($k = 1, 2, \cdots, 6$), to verify whether genetic robustness stemmed intrinsically from the miRNA stem-loops. To analyze the robustness of the structural element in the native miRNA stem-loops, we compare the distribution of the neutrality of the structural elements in native pre-miRNAs with that of corresponding elements in random pseudo pre-miRNAs (i.e., comparing the distribution of ξ_k^m in native pre-miRNAs with the corresponding distribution of ξ_k^r in random pseudo pre-miRNAs, $k = 1, 2, \cdots, 6$, respectively) and apply a *t*-test assuming that the two samples come from normal distributions with unknown and possibly unequal variances (Behrens-Fisher problem). The *t*-test uses Satterthwaite's approximation for the effective degrees of freedom.

Furthermore, to rule out the effect of base composition bias on robustness analysis of the structural elements in the native pre-miRNAs, we measure the neutrality of the structural elements of shuffling pseudo pre-miRNAs in every type of reference set for each species, namely, ξ_k^{m0}, ξ_k^{m1}, ξ_k^{mono}, and ξ_k^{di} ($k = 1, 2, \cdots, 6$), in which the superscripts represent the shuffling methods based on zero-Markov chain, first-Markov chain, mononucleotide, and dinucleotide, respectively. We also compare the neutrality of structural elements in native pre-miRNAs with that of corresponding elements in shuffling pseudo pre-miRNAs to further verify the genetic robustness of the miRNA stem-loops (i.e., comparing the distribution of ξ_k^m in native pre-miRNAs with the corresponding distribution of ξ_k^{m0}, ξ_k^{m1}, ξ_k^{mono}, and ξ_k^{di} in shuffling pseudo pre-miRNAs, $k = 1, 2, \cdots, 6$, respectively).

For each element, the relative increased level of average neutrality is defined as the ratio of the difference between the average neutrality of native pre-miRNAs and that of pseudo pre-miRNAs in the corresponding reference sets to the average neutrality of random miRNAs, i.e.,

$$l_k = \frac{\overline{\xi_k^m} - \overline{\xi_k^r}}{\overline{\xi_k^r}} \times 100\%, \quad k = 1, 2, \cdots, 6 \tag{2}$$

which reflects the degree of the increased level of average neutrality relative to that of pseudo pre-miRNAs.

To further test the difference of genetic robustness between various structural elements, we apply a *t*-test for each pair distribution of the neutrality of the structural elements in all native pre-miRNAs. A *t*-test is also performed for each pair distribution of the neutrality of the structural elements in random and shuffling pseudo pre-miRNAs, as done for native stem-loops, to verify whether the differential genetic robustness stemmed intrinsically from the miRNA stem-loops and to rule out the effect of base composition bias on differential genetic robustness analysis.

Results

Excess Robustness of Secondary Structural Elements in Native Pre-miRNAs

Data comparing the neutrality of structural elements in native pre-miRNAs with that of corresponding elements in random pseudo pre-miRNAs demonstrate that the structural elements of native pre-miRNAs are robust ($\xi_k^m > \xi_k^r$, for $k = 1, 2, \cdots, 6$; i.e., more robust than those of random pseudo pre-miRNAs; Table 3). The increased neutrality of the structural elements in native pre-miRNAs is also evident from the distribution of neutrality values of different structural elements in native versus random pseudo pre-miRNAs (Fig. 2). Although the difference in the mean neutrality of stem elements in native versus random pseudo pre-miRNAs is relatively small (0.14 vs. 0.09, respectively), the two distributions are significantly different from one another ($p = 0$; Table 3).

The mononucleotide and dinucleotide frequencies of an RNA sequence (which are not preserved in random pseudo pre-miRNAs) are critical for secondary structure stability

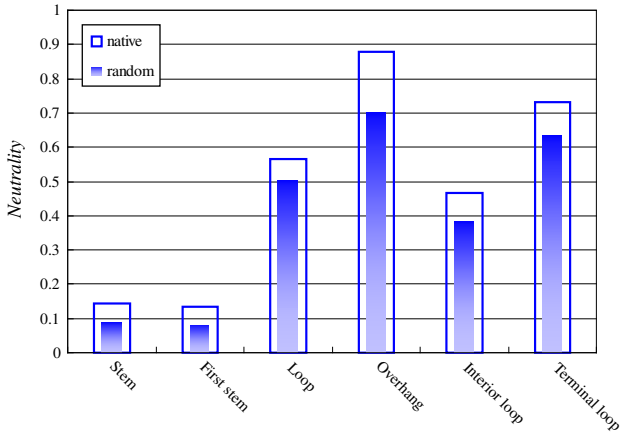

Fig. 2 Distribution of neutrality ξ of different structural elements in 1082 native pre-miRNAs, and in corresponding random pseudo pre-miRNAs. Two-dimensional histogram plots of the distribution of neutrality ξ in different secondary structural elements of native and random pseudo pre-miRNAs. Each bar is constituted by an outer, open subbar and an inner, filled subbar, which represents the distribution of neutrality ξ for native pre-miRNAs and random pseudo pre-miRNAs, respectively

(Bonnet et al. 2004; Clote et al. 2005; Katz and Burge 2003; Workman and Krogh 1999). It is, consequently, important to verify that the observed increased neutrality of structural elements in native pre-miRNAs is not a by-product of a bias in the base composition relative to random pseudo pre-miRNAs. To this end, we generate four different shuffling pseudo pre-miRNAs that preserve not only the similar stem-loop structure, but also the exact or nearly exact mononucleotide and dinucleotide base composition of the native pre-miRNAs. Additionally, we compare the neutrality ξ of structural elements in native pre-miRNAs with that of corresponding elements in different types of shuffling pseudo pre-miRNAs. Our data suggest that the structural elements of native pre-miRNAs

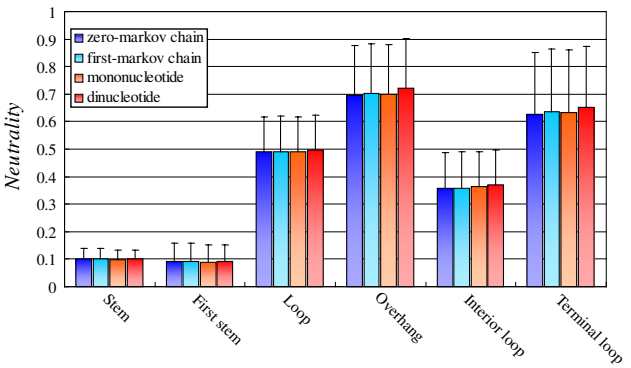

Fig. 3 Distribution of neutrality ξ of different structural elements in shuffling pseudo pre-miRNAs. Two-dimensional histogram plots of the distribution of neutrality ξ in different secondary structural elements of shuffling pseudo pre-miRNAs

are more robust than those of the shuffling pseudo pre-miRNAs (Table 3). The different types of sequence shuffling methods are indistinguishable (Table 3, Fig. 3).

Interestingly, there are differences in the relative increased levels of structural elements in native pre-miRNAs, as noted by the examination of the relatively increased level of average neutrality. While the average neutrality of native pre-miRNA stem elements increased by $\sim$60% relative to corresponding stem elements within random miRNAs, the average neutrality increase for different loop elements is not greater than 30% (Table 4). Examining the relatively increased level of average neutrality of pre-miRNAs within species provides a similar picture for each species separately, and the base composition does not have any influence on the difference in relatively increased level (Tables 4, 5 and supplementary materials). The differential relative increased level of average robustness of various structural elements indicates that different structural elements may be under a different selection pressure in the evolution process.

Differential Genetic Robustness Between Structural Elements in Pre-miRNAs

Examination of the neutrality ξ of each element in the 1082 native stem-loops suggests that there is greater neutrality in the loop element in comparison to the stem element of miRNA stem-loops ($\sim$0.14 vs. $\sim$0.60, respectively; Fig. 2). There is a significant difference between the neutrality ξ of the stem element and the loop element of native miRNA stem-loops ($\xi_1 < \xi_3$, $p = 0$; Table 7), as well as between the paired distributions of the neutrality ξ of other stem and loop elements (at the level of significance of 0.05, $\xi_i < \xi_j$ for $i = 1, 2$; $j = 3, 4, 5, 6$; Table 6). Although the neutrality ξ is similar and minor for different stem elements (Fig. 2), the differences between the first stem element and the stem element are significant ($p = 9.65 \times 10^{-8}$; Table 6). For different loop elements, the neutrality ξ value varied largely, ranging from 0.00 to 1.00. The overhang and terminal loop elements are the highest robustness elements ($\sim$0.88 and $\sim$0.73, respectively; Fig. 2), with significant differences noted between the distributions of neutrality of the loop elements (Table 6). Recently, experimental studies have demonstrated that miRNA precursors are tolerant of mutations in the body of the stem and, to a lesser extent, the loop (Lee et al. 2003; Zeng and Cullen 2004; Zeng et al. 2003). Mutations at the base of the loop have little effect, indicating that loops are not essential for processing. On the contrary, disrupting the base-pairing at the base of the stem has a marked deleterious effect on miRNA processing, indicating that the precursor stem is critical for mature miRNA production. Our results are in good agreement with experimental observations.

Table 4 The relative increased level of average neutrality of different secondary structural elements for each species

Note: Average values of different secondary structural elements from 1082 miRNAs of six species

Species	Stem	First stem	Loop	Overhang	Internal loop	Terminal loop
H. sapiens	58.34	61.69	10.73	27.99	21.53	17.58
C. elegans	59.07	62.09	22.86	16.15	26.49	18.70
D. melanogaster	59.57	61.13	11.55	30.17	18.81	20.99
D. rerio	65.62	64.90	9.42	29.27	20.30	8.20
M. musculus	63.88	68.36	8.63	21.83	17.50	17.53
R. norvegicus	59.51	60.73	15.31	12.68	22.66	22.13
Average	61.95	63.78	11.68	24.67	20.85	15.35

Table 5 The relative increased level of average robustness for each structural element with different randomization methods

Note: Average values of different secondary structural elements from 1082 miRNAs of six species

Species	Stem	First stem	Loop	Overhang	Internal loop	Terminal loop
Random	38.25	38.94	38.20	10.46	19.79	17.25
Zero-Markov	30.23	32.21	29.45	13.43	20.56	23.10
First-Markov	29.63	31.39	28.89	12.91	19.87	22.88
Mononucleotide	31.60	33.37	31.40	13.04	20.47	21.75
Dinucleotide	30.96	32.48	30.76	11.97	18.00	20.19

Table 6 *p*-values of *t*-tests comparing structural elements in native pre-miRNAs

	Stem	First stem	Loop	Overhang	Internal loop	Terminal loop
Stem	1.00					
First stem	9.65×10^{-8}	1.00				
Loop	0	0	1.00			
Overhang	0	0	0	1.00		
Internal loop	0	0	7.62×10^{-104}	0	1.00	
Terminal loop	0	0	2.15×10^{-139}	2.65×10^{-84}	1.75×10^{-281}	1.00

Differential genetic robustness between various structural elements is also observed in random pseudo pre-miRNAs, which coincides with the observations made in native pre-miRNAs (Table 7). Furthermore, the base composition bias does not have any influence on the differential genetic robustness between structural elements in pre-miRNAs (Fig. 3 and supplementary materials). These results indicate that differential genetic robustness may stem intrinsically from the special stem-loop structures of pre-miRNAs. While it is possible in theory that simple differences between structural elements, such as the length of structural elements, may underlie differences in genetic robustness, we do not observe any correlation between the length and the neutrality of the structural elements in this study (Pearson's correlation coefficients between length and neutrality were 0.0027 and −0.0648 for stem and loop elements, respectively).

Discussion

Genetic Robustness of miRNA Genes

To explore genetic robustness of the structural elements in pre-miRNAs, we examine the neutrality of the structural element in 1082 native pre-miRNAs from six species. We demonstrate that the secondary structural elements in native pre-miRNAs exhibit a significantly higher level of

Table 7 *p*-values of *t*-tests comparing structural elements in random pseudo pre-miRNAs

	Stem	First stem	Loop	Overhang	Internal loop	Terminal loop
Stem	1.00					
First stem	7.50×10^{-12}	1.00				
Loop	0	0	1.00			
Overhang	0	0	0	1.00		
Internal loop	0	0	0	0	1.00	
Terminal loop	0	0	6.18×10^{-284}	1.43×10^{-72}	0	1.00

genetic robustness than the corresponding structural elements in random pseudo pre-miRNAs, indicating that this excess robustness is beyond the intrinsic robustness of the stem-loop structure. Furthermore, the excess robustness is not a by-product of a bias in the base composition of the native pre-miRNAs. Remarkably, we find differences in the increased level of average neutrality between various structural elements, indicating that different structural elements may be under a different selection pressure in the evolution process. The data suggest that the secondary structural elements in native pre-miRNAs have indeed been under evolutionary pressures that favored robust configuration elements. Our results also demonstrate that differential genetic robustness between various structural elements is observed in native pre-miRNAs, in good accordance with experimental observations (Lee et al. 2003; Zeng and Cullen 2004; Zeng et al. 2003). The greatest deleterious effect on miRNA processing results from disruption of the base-pairing at the base of the predicted precursor stem, indicating that the precursor stem is critical for mature miRNA production. However, the changes in the loop sequence do not yield any perceptible effects, indicating that loops are not essential for processing. Interestingly, striking differences in neutrality between various structural elements are also observed in random pseudo pre-miRNAs, with no effects observed from the base composition bias, indicating that the differential genetic robustness may stem intrinsically from special stem-loop structures.

The excess robustness of structural elements may have led to the genetic robustness of the integrated stem-loop structures of pre-miRNAs. A recent study reports that the stem-loop structures of miRNA stem-loops show excess robustness with respect to mutational perturbation, compared with random RNA sequences with similar stem-loop structures (Borenstein and Ruppin 2006). Also, the in silico genetic robustness analysis of miRNA stem-loops in our previous study demonstrates that the stem-loop hairpin structures of pre-miRNAs exhibit a significantly higher level of genetic robustness at different FDR-controlled *p*-values (Shu et al. 2007a). Furthermore, both of these two studies demonstrate that this excess robustness goes beyond the intrinsic robustness of the stem-loop structure and is not a by-product of the base composition bias. These results in both studies are in good agreement with each other, although the quantitative measures of genetic robustness defined in these two studies are quite different.

Evolutionary Origin of Genetic Robustness

Our results indicate that genetic robustness is ubiquitous in pre-miRNAs. Yet the principles and mechanisms that lead to the emergence of the observed robustness are far less clear. Whether it is a consequence of natural selection or a nonadaptive correlated side effect of other phenotypic traits is by and large unknown. A recent review article categorized the theories addressing the evolution of genetic robustness into three main classes: adaptive, intrinsic, and congruent (de Visser et al. 2003). To explore the evolutionary origin of the excess genetic robustness observed in pre-miRNAs, the authors in both of these two studies additionally examine the environmental robustness of pre-miRNAs (Borenstein and Ruppin 2006; Shu et al. 2007a). However, we come to a different conclusion about the origin of this excess genetic robustness. The data in the study by Borenstein and Ruppin suggest that the excess robustness of miRNA stem-loops is the result of direct evolutionary pressure toward increased robustness. Yet our results demonstrate that the stem-loop structures of pre-miRNAs buffer against genetic perturbations and, at the same time, are insensitive to environmental perturbations, suggesting that increased genetic robustness may evolve as a correlated side effect of the evolution for environmental robustness. The different conclusions (adaptive robustness versus congruent robustness) may have resulted from differences in the reference backgrounds employed. The reference backgrounds in our study are made up of random and shuffled pseudo pre-miRNA sequences with preserved phenotypes that are similar to real pre-miRNAs (Shu et al. 2007a). Their reference backgrounds, on the other hand, are produced by inverse folding (Borenstein and Ruppin 2006).

According to the classification of Hermisson and Wagner (2004), who classify robustness as adaptive and intrinsic, however, both of these two studies come to the same conclusion about adaptive robustness. They consider robustness to be adaptive if the buffering of that trait with respect to some source of variation has the target of natural selection, i.e., robust character states are selected because of their reduced variability (Hermisson and Wagner 2004). Adaptive robustness evolves due to its own selective advantage. The natural force assumed to be responsible for its evolution is stabilizing selection acting directly on the character. The adaptive robustness in this definition encompasses both adaptive and congruent scenarios in the classification system established by de Visser et al. (2003). The difference is that the natural forces, assumed to be responsible for its evolution, function as a kind of stabilizing selection acting directly on a character or on some highly correlated pleiotropic trait (Hermisson and Wagner 2004). Taken together, the results of the two studies suggest that the genetic robustness of miRNA stem-loops, under different evolutionary selection pressures, may evolve due to its own selective advantage. While the findings support this hypothesis, additional theoretical and experimental work is required to fully elucidate the mechanisms of the evolution of robustness. A greater

understanding of the evolution of robustness will require quantitative knowledge of the forces producing robustness, such as the distribution of fitness effects of mutations (Meiklejohn and Hartl 2002).

Biological Implications of Robustness

Recognition of robustness in miRNA stem-loops has direct consequences for miRNA research. First, the excess robustness in pre-miRNAs may facilitate the in silico identification of novel miRNAs on a single genome. Since most current computational methods for prediction of miRNA genes rely heavily on phylogenetic conservation and the structural characteristics of pre-miRNAs (Bartel and Chen 2004; Berezikov et al. 2006; Kim and Nam 2006), most of the identified miRNAs are highly conserved among species and most research has focused on these highly conserved miRNAs (Berezikov et al. 2005; Lim et al. 2003a, b; Pang et al. 2006; Xie 2005). However, nonconserved miRNAs represent a potentially important source of functional novelties during evolution. Recently, various nonconserved miRNAs have been discovered and experimentally verified (Bentwich et al. 2005; Pfeffer et al. 2005). The property of excess robustness is probably not sufficient by itself to identify pre-miRNAs, however, it can serve as a complementary method to filter out random pseudo pre-miRNA sequences and to facilitate improvement of miRNA prediction on a single genome.

Second, properties of robustness may also be utilized for the optimal design of nucleic acid sequences and, furthermore, for the improvement of in vitro selection or SELEX (Systematic Evolution of Ligands by Exponential enrichment). SELEX is an experimental method for selecting functional RNAs from a large pool (10^{15}) of random sequences (Tuerk and Gold 1990; Ellington and Szostak 1990). The use of designed sequences with properties of robustness, in lieu of random sequences, may increase the probability of identifying novel functional RNAs.

Finally, a greater understanding of robustness in miRNA genes may also facilitate the future research on robustness. As the secondary structure of miRNA stem-loops embodies many of the properties controlling molecular evolution (Borenstein and Ruppin 2006; Shu et al. 2007a), it forms a promising framework for studying the evolutionary origin of genetic robustness. The excess robustness of miRNA genes examined in the current study can be regarded as primary robustness based on genotype-phenotype mapping. The simplicity of this form of robustness, full tractability of RNA secondary structure, and complete control of reference background facilitate the exploration of its evolutionary origins. Protein structures, which are much more complicated than RNA secondary structures, may possess a similar tendency for sequence-based robustness as well as additional principles and mechanisms contributing to their robustness. With the prediction algorithms for protein folding (Baker 2000), our methodology can be applied to the robustness analysis of protein structures without any difficulties. Furthermore, our methodology may be heuristic in the study of higher-level robustness: developmental robustness. Hornstein and Shomron (2006) suggest that miRNA interactions with the network of protein-coding genes evolved to buffer stochastic perturbations and thereby confer robustness to developmental genetic programs. The relationship between genetic and developmental robustness is derived quantitatively through the variance of phenotypic fluctuations, which are directly measurable experimentally (Kaneko 2007).

Acknowledgments The authors would like to thank the reviewers and the editors of the paper for their constructive comments, which contributed to an improved presentation. The authors would also like to thank the Super Biomed Computation Center at Beijing Institute of Health Administration and Medicine Information for providing computing resources. This work was supported by grants from the National High Technology Research and Development Program of China (No. 2007AA02Z311 and No. 2006AA02Z304) and Grants from the National Nature Science Foundation of China (No. 30700139 and No. 30600120).

References

Baker D (2000) A surprising simplicity to protein folding. Nature 405:39–42

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297

Bartel DP, Chen CZ (2004) Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat Rev Genet 5:396–400

Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, Barzilai A, Einat P, Einav U, Meiri E, Sharon E, Spector Y, Bentwich Z (2005) Identification of hundreds of conserved and nonconserved human microRNAs. Nat Genet 37:766–770

Berezikov E, Guryev V, van de BJ, Wienholds E, Plasterk RH, Cuppen E (2005) Phylogenetic shadowing and computational identification of human microRNA genes. Cell 120:21–24

Berezikov E, Cuppen E, Plasterk RH (2006) Approaches to microRNA discovery. Nat Genet 38(Suppl):S2–S7

Boffelli D, McAuliffe J, Ovcharenko D, Lewis KD, Ovcharenko I, Pachter L, Rubin EM (2003) Phylogenetic shadowing of primate sequences to find functional regions of the human genome. Science 299:1391–1394

Bonnet E, Wuyts J, Rouze P, Van de PY (2004) Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics 20:2911–2917

Borenstein E, Ruppin E (2006) Direct evolution of genetic robustness in microRNA. Proc Natl Acad Sci USA 103:6593–6598

Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3:e85

Clote P, Ferre F, Kranakis E, Krizanc D (2005) Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency. RNA 11:578–591

de Visser JA, Hermisson J, Wagner GP, Ancel ML, Bagheri-Chaichian H, Blanchard JL, Chao L, Cheverud JM, Elena SF,

Fontana W, Gibson G, Hansen TF, Krakauer D, Lewontin RC, Ofria C, Rice SH, von Dassow G, Wagner A, Whitlock MC (2003) Perspective: evolution and detection of genetic robustness. Evol Int J Org Evol 57:1959–1972

Elena SF, Lenski RE (2001) Epistasis between new mutations and genetic background and a test of genetic canalization. Evol Int J Org Evol 55:1746–1752

Elena SF, Carrasco P, Daros JA, Sanjuan R (2006) Mechanisms of genetic robustness in RNA viruses. EMBO Rep 7:168–173

Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. Nature 346:818–822

Fares MA, Ruiz-Gonzalez MX, Moya A, Elena SF, Barrio E (2002) Endosymbiotic bacteria: groEL buffers against deleterious mutations. Nature 417:398

Farh KK (2005) The widespread impact of mammalian microRNAs on mRNA repression and evolution. Science 310:1817–1821

Fisher RA (1928a) The possible modifications of the response of the wild type to recurrent mutations. Am Nat 62:115–116

Fisher RA (1928b) Two further notes on the origin of dominance. Am Nat 62:571–574

Fisher RA (1931) The evolution of dominance. Biol Rev 6:345–368

Gibson G, Wagner G (2000) Canalization in evolutionary genetics: a stabilizing theory? Bioessays 22:372–380

Griffiths-Jones S (2004) The microRNA Registry. Nucleic Acids Res 32:D109–D111

Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34:D140–D144

Grun D, Wang YL, Langenberger D, Gunsalus KC, Rajewsky N (2005) microRNA target predictions across seven Drosophila species and comparison to mammalian targets. PLoS Comput Biol 1:e13

Hermisson J, Wagner GP (2004) Evolution of phenotypic robustness. In: Jen E (ed) Robust design: a repertoire from biology, ecology, and engineering. Oxford University Press, New York, pp 47–70

Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P (1994) Fast folding and comparison of RNA secondary structures. Monatsh Chem [Chem Month] 125:167–188

Hornstein E, Shomron N (2006) Canalization of development by microRNAs. Nat Genet 38(Suppl):S20–S24

Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293:834–838

Kaneko K (2007) Evolution of robustness to noise and mutation in gene expression dynamics. PLoS ONE 2:e434

Katz L, Burge CB (2003) Widespread selection for local RNA secondary structure in coding regions of bacterial genes. Genome Res 13:2042–2051

Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. Genes Dev 15:2654–2659

Kim VN, Nam JW (2006) Genomics of microRNA. Trends Genet 22:165–173

Kitano H (2004) Biological robustness. Nat Rev Genet 5:826–837

Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294:853–858

Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294:858–862

Lee RC, Ambros V (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. Science 294:862–864

Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75:843–854

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN (2003) The nuclear RNase III Drosha initiates microRNA processing. Nature 425:415–419

Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15–20

Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP (2003a) Vertebrate microRNA genes. Science 299:1540

Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP (2003b) The microRNAs of *Caenorhabditis elegans*. Genes Dev 17:991–1008

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. Science 303:95–98

Meiklejohn CD, Hartl DL (2002) A single mode of canalization. Trends Ecol Evol 17:468–473

Montville R, Froissart R, Remold SK, Tenaillon O, Turner PE (2005) Evolution of mutational robustness in an RNA virus. PLoS Biol 3:e381

Pang KC, Frith MC, Mattick JS (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends Genet 22:1–5

Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, van Dyk LF, Ho CK, Shuman S, Chien M, Russo JJ, Ju J, Randall G, Lindenbach BD, Rice CM, Simon V, Ho DD, Zavolan M, Tuschl T (2005) Identification of microRNAs of the herpesvirus family. Nat Methods 2:269–276

Sanjuan R, Forment J, Elena SF (2006a) In silico predicted robustness of viroids RNA secondary structures. I. The effect of single mutations. Mol Biol Evol 23:1427–1436

Sanjuan R, Forment J, Elena SF (2006b) In silico predicted robustness of viroids RNA secondary structures. II. Interaction between mutation pairs. Mol Biol Evol 23:2123–2130

Scharloo W (1991) Canalization: genetic and developmental aspects. Annu Rev Ecol Syst 22:65–93

Shu W, Bo X, Ni M, Zheng Z, Wang S (2007a) In silico genetic robustness analysis of microRNA secondary structures: potential evidence of congruent evolution in microRNA. BMC Evol Biol 7:223

Shu W, Bo X, Zheng Z, Wang S (2007b) RSRE: RNA structural robustness evaluator. Nucleic Acids Res 35:W314–W319

Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3′UTR evolution. Cell 123:1133–1146

Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505–510

Waddington CH (1953) The genetic assimilation of an acquired charcter. Evolution 7:118–126

Waddington CH (1957) The strategy of the genes. Macmillan, New York

Wagner A (2005) Robustness and evolvability in living systems. Princeton studies in complexity. Princeton University Press, Princeton, NJ

Wagner A, Stadler PF (1999) Viral RNA and evolved mutational robustness. J Exp Zool 285:119–127

Wagner GP, Booth G, Bagheri-Chaichian H (1997) A population genetic theory of canalization. Evolution 51:329–347

Workman C, Krogh A (1999) No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution. Nucleic Acids Res 27:4816–4822

Xie X (2005) Systematic discovery of regulatory motifs in human promoters and 3′UTRs by comparison of several mammals. Nature 434:338–345

Zeng Y, Cullen BR (2003) Sequence requirements for micro RNA processing and function in human cells. RNA 9:112–123

Zeng Y, Cullen BR (2004) Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. Nucleic Acids Res 32:4776–4785

Zeng Y, Yi R, Cullen BR (2003) MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci USA 100:9779–9784

Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res 9:133–148