Journal of Bioinformatics and Computational Biology
Vol. 8, No. 6 (2010) 1013–1026

DOI: 10.1142/S0219720010005087

Imperial College Press
www.icpress.co.uk

BIOINFORMATIC ANALYSIS OF THE NEUTRALITY OF RNA SECONDARY STRUCTURE ELEMENTS ACROSS GENOTYPES REVEALS EVIDENCE FOR DIRECT EVOLUTION OF GENETIC ROBUSTNESS IN HCV

ALEXANDER CHURKIN*, MORIAH COHEN*,
YONAT SHEMER-AVNI† and DANNY BARASH*,‡

*Department of Computer Science
Ben-Gurion University
Beer-Sheva 84105, Israel

†Department of Virology and Developmental Genetics
Faculty of Health Sciences
Ben-Gurion University
Beer-Sheva 84105, Israel
‡dbarash@cs.bgu.ac.il

Received 20 January 2010
Revised 18 August 2010
Accepted 18 August 2010

The properties and origin of genetic robustness have recently been investigated in several works that examined microRNA stem-loop structures, and a variety of conclusions have been reached without agreement. Considering that this is a universal phenomenon that is not restricted to miRNAs, we recall the original work on this topic that began from looking at viral RNAs of several types. We provide a link to this work by examining the neutrality of HCV structural elements, performing a detailed bioinformatic analysis using RNA secondary structure predictions across genotypes. This study provides supporting evidence for direct evolution of genetic robustness that is not limited to noncoding RNAs participating in gene regulation, but includes functionally important structural elements of the hepatitis C virus (HCV) that show excess of robustness beyond the intrinsic robustness of their stem-loop structure. These findings further support the adaptive behavior of genetic robustness in functional RNAs of various types that seem to have evolved with selection pressure towards increased robustness.

Keywords: Mutational robustness; RNA secondary structure prediction.

1. Introduction

The ability of an organism to exhibit increased neutrality and show a significant degree of robustness against mutations has been a topic of wide interest. The

‡Corresponding author.

evolution of mutational robustness in RNA viruses has been previously addressed by computer simulation.[1,2] Comparison between conserved and non-conserved elements using programs from the Vienna RNA package,[3] like alidot, has demonstrated an emergence of increased neutrality in RNA viruses.[1] The evolutionary origin of robustness, however, is still a subject of ongoing debate. Recently, in Ref. 4, evidence for the direct evolution of genetic robustness has been shown for microRNAs, by estimating their mutational robustness and comparing with random RNA sequences that fold into the same stem-loop structure. This was done on miRNAs using RNAinverse and other suitable programs from the Vienna RNA package.[3] Follow-up works[5,6] have remained in the scope of miRNAs, each suggesting a somewhat different divergence from the analysis in Ref. 4 that eventually led them to offer a contrasting viewpoint of congruent evolution of genetic and environmental robustness in miRNAs. Here, significant evidence that HCV secondary structural elements have evolved with selection pressure towards increased robustness is shown using the RNAinverse[3] program. This widens the scope of such observations that are certainly not limited to miRNAs, but can also be demonstrated with RNA secondary structure predictions in other functionally important systems.

Mutational robustness describes the extent to which the phenotype of an organism remains constant in spite of mutation. A general discussion can be found in Refs. 7 and 8. Mathematical models that consider robustness in the context of neutral networks and neutral evolution dynamics are available in Refs. 9–11. Computer simulations that relate to increased neutrality are also found in Ref. 12 among other works. Increased mutational robustness in RNA viruses was biologically demonstrated in Ref. 13 but might be attributed to effects other than to direct evolution of robustness (for example, in the experiment performed in Ref. 13, complementation weakens selection for robustness). In the original work of Wagner and Stadler,[1] it was shown that conserved RNA secondary structure elements are consistently more robust than non-conserved elements in different groups of RNA viruses. By this comparison it provided evidence for the first time in computer simulations that mutational robustness has evolved in RNA viruses. The materials and methods in Ref. 1 included as data several HCV sequences that were available at the time, along with several sequences from the other two RNA viral groups of Dengue virus and HIV1. The analysis presented here does not repeat the comparison between conserved and non-conserved elements available in Ref. 1. In our analysis, we provide further support for direct evolution of robustness in RNA viruses with up-to-date sequence data taken from the Los Alamos HCV database,[14] which was not available in the late 1990s. Furthermore, the methods used work well on RNA viruses and there is no need to be restricted to miRNAs as in recent works.[4,5,6]

We begin with a genome-wide study of the mutational robustness of representative stem-loops of the HCV. We show that the secondary structures of HCV stem-loops from IRES and core exhibit a significantly high level of mutational robustness when compared with random RNA sequences that have the same secondary structure. The secondary structures of the HCV stem-loops examined in

this work (IIIabc stem-loop from IRES and SL248, SL337 from core) are highly conserved between HCV genotypes, hypothesizing that they are essential in virus replication and have been under evolutionary pressure to conserve their secondary structure. This pressure favored robust configurations and may have led to the evolution of robustness.

2. Materials and Methods

2.1. *RNA folding prediction*

The RNA secondary structures in this study were predicted by an energy minimization approach[15] using RNAfold available in the Vienna RNA package,[3,16,17] noting that similar predictions can be done with mfold.[18] RNAinverse that is additionally available in the Vienna RNA package was used to generate a reference set of sequences that fold into the same secondary structure as the original sequence.

2.2. *HCV database*

The sequences of the IRES IIIabc, SL248 and SL337 stem-loops for each one of the six HCV genotypes were extracted from the Los Alamos HCV database[14] and these sequences were filtered such that for each genotype we collect only those sequences that differ from other sequences by at least one base. We particularly chose the stem-loops mentioned above because for these three stem-loops, we found enough (for statistical analysis) number of different sequences that fold using energy minimization prediction methods in a similar way to the secondary structures that were obtained by biological experiments. We also filtered several sequences that contained different letters than (A,T,U,C,G) and several sequences that failed to fold to the experimentally known biological structures of these stem-loops. Our aim was to choose a representative set for which there is complete information about their sequence composition and the folding prediction methods we use are tested in advance to work adequately on them.

2.3. *RNA secondary structure of HCV*

Having chosen segments from HCV as our model system, knowledge about the RNA secondary structure of HCV is taken into account throughout this work. Here, a basic background is provided about their overall structure.

Figure 1 depicts the general assembly of the HCV genome. The IRES, from which stem-loop IIIabc was taken in this study, is highlighted in the left. Its constituents can be viewed in more detail in Ref. 20. The core region is labeled as "C" in Fig. 1 and it contains stem-loops SL248 and SL337. More information about the core and the importance of its stem-loops for HCV replication can be found in Ref. 21. Another important region is the NS5B domain, for which information can be found in Ref. 22 about some of its functional elements. The whole assembly

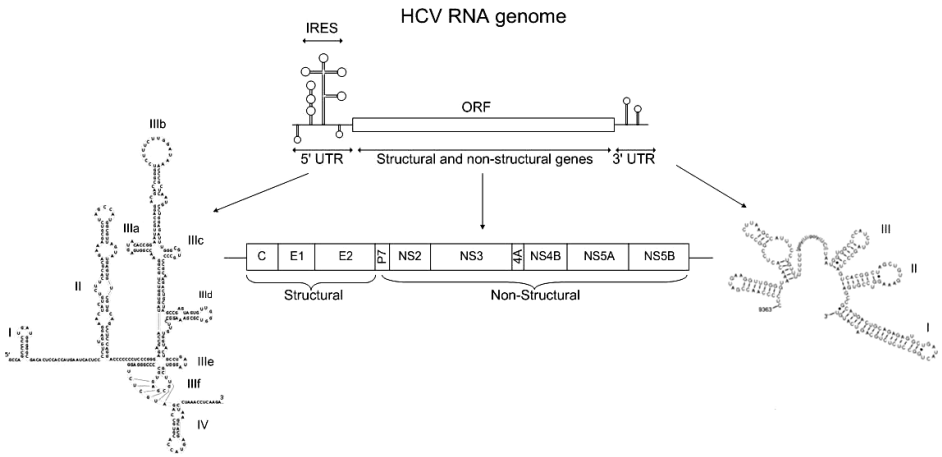

Fig. 1. The genome of HCV. The genome is composed of three major regions with high resemblance between different HCV genotypes. The 5′UTR and the extreme end of the 3′UTR are the most conserved regions of HCV RNA. There is more than 90% identity between different genotypes at the 5′UTR, which is involved in the initiation of both translation and RNA replication. The 3′UTR, however, differs in length in different genotypes but is highly conserved among viral strains of the same genotype. The IRES has a unique structure consisting of four domains, playing an important role in translation initiation. The core region, labeled "C", contains a number of stem-loops that are well conserved among all genotypes and are involved in replication.[19]

contains important stem-loop structures that have been a subject of considerable interest, some of which are better conserved than others and well predicted by energy minimization methods; hence we picked these particular ones for the analysis in our study.

It should be noted that RNA folding prediction by energy minimization mentioned above was already used in conjunction with phylogenetic and experimental methods to investigate the RNA secondary structure in the HCV genome. In a body of works, to name but a few, the RNA secondary structures in a variety of functional elements throughout the HCV genome were studied (e.g. Ref. 23) and also genome-scale ordered RNA structures (GORS) were detected and analyzed starting in Ref. 24.

2.4. *Neutrality calculation*

For a quantitative measure of mutational robustness, the neutrality η is calculated. The neutrality of an RNA sequence of length L that was used in Ref. 4 is defined as $\eta = (L - d)/L$, where d is the base-pair distance between the secondary structure of the original sequence and the secondary structure of the mutant, averaged over all $3L$ one-mutant neighbors. The base-pair distance that is available in the Vienna RNA package[3,17] is used as a distance measure between two RNA secondary structures. We argue against the more stringent neutrality measure that was suggested in Ref. 6, defined as the fraction of neutral single mutant neighbors, i.e. those that have an identical MFE structure to the original sequence. In many cases, experience

with RNAMute[25] shows that there are nearly identical MFE structures to the original sequence and those should not be discarded. For example, taking sequences that are about 100 nt long from the HCV IRES IIIabc stem-loops of various genotypes, we observe the following regarding their single-mutant neighborhood: about 30% of the mutants fold to an identical MFE structure as the original sequence; about 50% are grouped in base pair distance from 1 to 4, which are nearly identical, and only about 20% are with distances > 4. If we only count the number of identical MFE structures, the neutrality calculation will neglect important information. From our acquaintance with RNA folding prediction methods, we do not find deviations from the logical definition of neutrality given above necessary for the HCV segments we examined.

2.5. *Robustness and significance analysis*

For consistency and simplicity in the presentation, we use a similar labeling as in Ref. 4. For each RNA sequence, we measure the neutrality of the native stem-loop, η_m, and evaluate the neutrality of this stem-loop structure expected by chance, η_c, as the average neutrality of 100 random sequences that fold into the same stem-loop structure using RNAinverse that is available in the Vienna RNA package.[3,17] To determine whether the increased robustness observed for some genotype or some group of sequences is statistically significant, the magnitude and sign of the differences between the paired values η_m and η_c are obtained and analyzed by the nonparametric Wilcoxon signed rank test for paired data. The rank of the native stem-loop neutrality, r, among the neutrality of the 100 inversely folded sequences is calculated as well. The significance level of each RNA sequence robustness is then evaluated by $p = r/101$, providing an estimate for the probability of observing an equal or higher neutrality value by chance. Sequences for which $r \leq 5$ (corresponding to $p < 0.05$) are labeled as significantly robust.

2.6. *RNA hairpin structure intrinsic neutrality*

To quantify the level of the RNA neutrality that stems intrinsically from the RNA hairpin structure (as opposed to neutrality resulting from direct selection towards robustness measured above), an additional reference set is produced for each stem-loop. This set consists of 100 shuffled sequences produced by the Altschul–Erikson dinucleotide shuffle algorithm, which were made available by Clote and co-workers in Ref. 26, maintaining the RNA stem-loop mononucleotide and dinucleotide frequencies. The average neutrality of this set is used as a baseline neutrality value to evaluate the level of excess neutrality associated with the RNA hairpin structures.

3. Results

3.1. *Robustness results according to neutrality*

We calculated the neutrality of each one of the IRES IIIabc, SL248 and SL337 stem-loops in all six genotypes. For IRES IIIabc, 195 out of 200 sequences (97.5%)

Table 1. Percentage of robust sequences for IRES IIIabc, SL248 and SL337 stem-loops.

Genotype	No. of sequences	No. of robust sequences (%)	P-value	$\overline{\eta_m}$	$\overline{\eta_c}$
(a) Robustness analysis of IRES IIIabc stem-loop structures within each genotype.					
All genotypes	200	195 (97.5)	2.17×10^{-33}	0.93±0.03	0.83±0.02
1	1[a]	1 (100)	—	0.94	0.81
2	43	43 (100)	1.16×10^{-8}	0.95±0.02	0.82±0.01
3	48	47 (98)	6.91×10^{-9}	0.92±0.03	0.85±0.02
4	43	39 (90.7)	6.33×10^{-8}	0.90±0.04	0.82±0.02
5	8	8 (100)	0.031	0.91±0.01	0.81±0.01
6	57	57 (100)	5.31×10^{-11}	0.94±0.02	0.81±0.01
(b) Robustness analysis of SL248 stem-loop structures within each genotype.					
All genotypes	695	677 (97.4)	4.87×10^{-110}	0.96±0.01	0.93±0.03
1	59	59 (100)	2.47×10^{-11}	0.96±0.01	0.93±0.01
2	138	138 (100)	2.23×10^{-24}	0.95±0.02	0.91±0.03
3	164	159 (97)	4.95×10^{-27}	0.96±0.01	0.94±0.03
4	141	134 (95)	1.08×10^{-22}	0.97±0.01	0.95±0.02
5	12	12 (100)	4.88×10^{-4}	0.96±0.01	0.93±0.01
6	181	175 (96.7)	3.19×10^{-29}	0.96±0.01	0.94±0.03
(c) Robustness analysis of SL337 stem-loop structures within each genotype.					
All genotypes	401	394 (98.3)	4.95×10^{-67}	0.93±0.03	0.86±0.03
1	35	35 (100)	2.59×10^{-7}	0.96±0.01	0.90±0.02
2	109	107 (98.2)	1.43×10^{-19}	0.94±0.03	0.88±0.03
3	69	69 (100)	5.37×10^{-13}	0.94±0.02	0.83±0.02
4	49	49 (100)	1.15×10^{-9}	0.94±0.02	0.86±0.03
5	10	10 (100)	0.002	0.91±0.02	0.84±0.03
6	129	124 (96.1)	2.41×10^{-22}	0.92±0.04	0.84±0.03

Note: [a]All sequences in this genotype are the same.
P-values denote the probability of observing the increased neutrality found in each genotype (in SL248, SL337 and IRES IIIabc stem-loops) by chance and are calculated with the Wilcoxon signed rank test for paired data.

were found to be robust, according to the criterion that neutrality of the native sequence (η_m) is greater than the neutrality of the random sequences that have the same structure (η_c). The exception consists of several sequences from genotype 4 that folded slightly different from the common HCV IIIabc stem-loop (two out of three stems have a shift of several bases). The results for IRES IIIabc, SL248 and SL337 are summarized in Table 1.

3.2. *Neutrality distribution*

An increase in the neutrality of the HCV native stem-loops can also be observed in the different distributions of neutrality values in native versus inversely folded stem-loops. This can be viewed in Fig. 2(a) for IRES IIIabc, Fig. 2(b) for SL248 and Fig. 2(c) for SL337. Average neutralities of the native and inversely folded sequences for each stem-loop can be observed in Table 1 (first row in each one of the three table sections). From the table we can observe the relatively large differences between

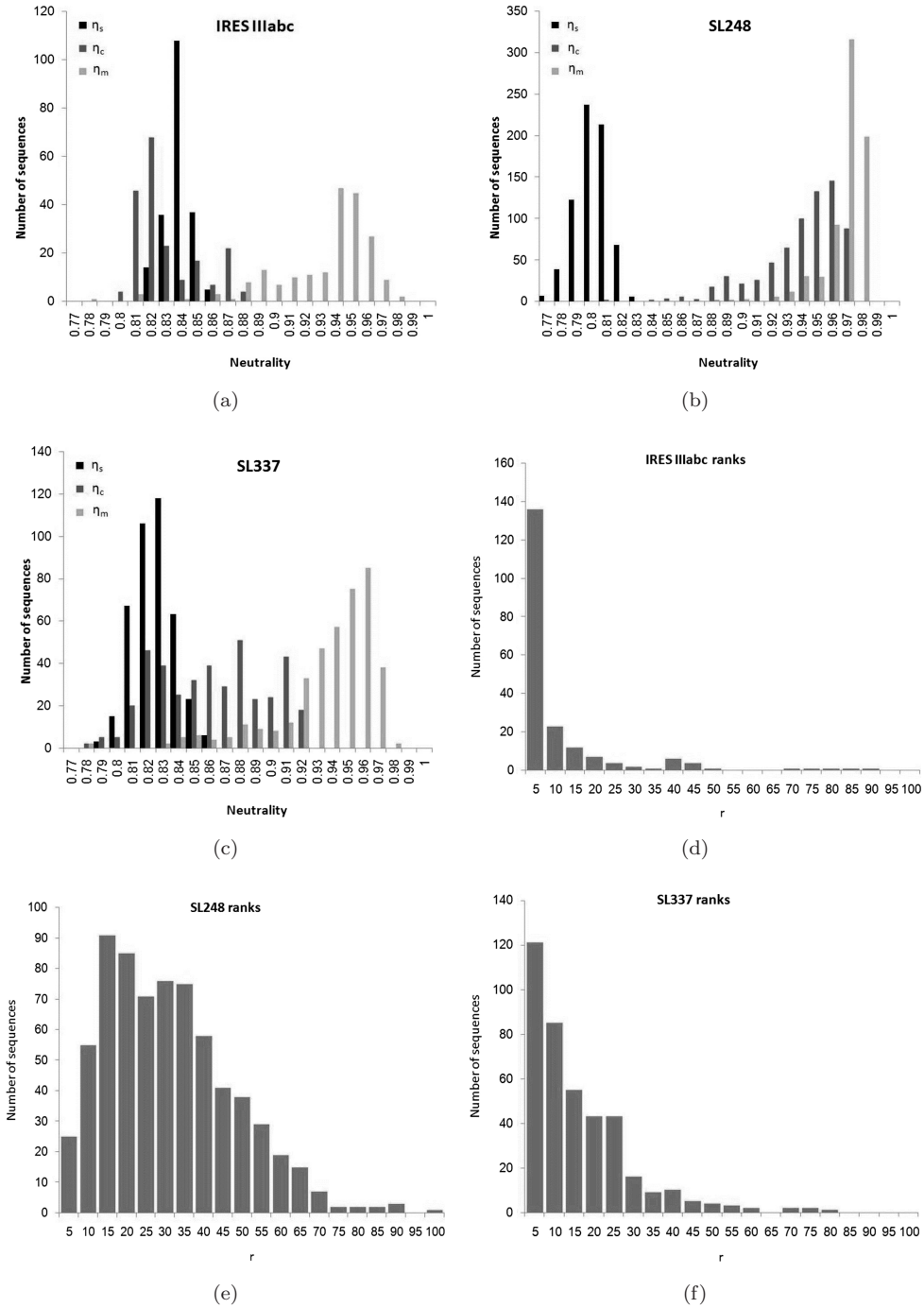

Fig. 2. Neutrality and rank distribution for sequences of IRES IIIabc, SL248 and SL337 stem-loops from HCV.

average neutralities in the IRES IIIabc stem-loop [0.93 versus 0.83 respectively, Table 1(a)] and in the SL337 stem-loop [0.93 versus 0.86 respectively, Table 1(c)], whereas there is a relatively small difference in the SL248 stem-loop [0.96 versus 0.93 respectively, Table 1(b)]. However, in all three examined stem-loops the two distributions were significantly different ($p < 2.17 \times 10^{-33}$ for IRES IIIabc, $p < 4.87 \times 10^{-110}$ for SL248 and $p < 4.95 \times 10^{-67}$ for SL337; Wilcoxon signed rank test for paired data), corroborating the hypothesis that evolved stem-loops of HCV are more robust than that expected by chance. Examining the robustness of HCV stem-loops for each one of the six genotypes provided a similar picture for each genotype separately (see also Table 1). The neutrality distribution graphs for each one of the genotypes separately can be found in the Supplementary Material.

3.3. *Rank distribution*

The level of neutrality for each sequence in the analysis was also evaluated by calculating the rank (r) of the neutrality of native sequence among the neutralities of all 100 random inversely folded sequences with the same secondary structure (see Sec. 2). High ranks indicate robust sequences. Significantly robust sequences have $r < 6$ (e.g. p-value < 0.05) and significantly non-robust sequences have $r > 95$. By inspecting rank distribution, a clear tendency towards high ranks can be viewed in Fig. 2(d) for IRES IIIabc, Fig. 2(e) for SL248 and Fig. 2(f) for SL337.

3.4. *Robustness results according to rank*

Some observations from the distributions of ranks can be made. First, in IRES IIIabc, 136 of 200 (68%) sequences were found to be significantly robust and zero sequences were found to be significantly non-robust. For 195 out of 200 (97.5%) sequences, $r < 25$. Second, a similar procedure was performed for SL248 and SL337, and the results are summarized in Table 2. Taken together, we find a relatively large number of significantly robust sequences while there are no significantly non-robust sequences in IRES IIIabc and SL337 stem-loops. The SL248 is less robust than others but still has 25 significantly robust sequences versus only one significantly non-robust sequence. From these results we may conclude that stem-loop IIIabc from IRES is much more robust than SL248 and SL337 stem-loops from core, while within the core, the SL337 stem-loop is more robust than the SL248 stem-loop. The graphs of the rank distributions for each one of the six genotypes separately can be found in the Supplementary Material.

3.5. *Ratio between adaptive and intrinsic robustness*

In comparing our finding to that in Ref. 4, we quantified the ratio between the adaptive robustness of HCV stem-loops, $\eta_m - \eta_c$, and its intrinsic robustness, $\eta_c - \eta_s$. The values that were obtained for all genotypes are listed in Table 3. As can be observed from the results, for core stem-loops we obtain positive values for intrinsic robustness: in SL337 most of the robustness comes from the adaptive robustness, while

Table 2. Robustness analysis of IRES IIIabc (a), SL248 (b) and SL337 (c) stem-loops, for each one of the six HCV genotypes.

Genotype	No. of sequences	$\overline{m - \eta_c \eta}$	$\overline{c - \eta_s \eta}$	Pearson correlation coefficient	P-value
(a) IRES IIIabc					
All genotypes	200	0.10±0.04	-0.01 ± 0.02	−0.70318	5.31×10^{-31}
1	1	0.13	−0.01	—	—
2	43	0.13±0.02	-0.02 ± 0.02	−0.71796	5.98×10^{-08}
3	48	0.07±0.04	0.01 ± 0.02	−0.7687	1.77×10^{-10}
4	43	0.09±0.05	-0.02 ± 0.02	−0.73411	2.12×10^{-08}
5	8	0.10±0.01	-0.02 ± 0.01	−0.52735	0.0896
6	57	0.12±0.02	-0.01 ± 0.02	−0.5233	2.96×10^{-05}
(b) SL248					
All genotypes	695	0.03±0.02	0.14±0.03	−0.79632	1.88×10^{-153}
1	59	0.03±0.02	0.14±0.02	−0.62862	9.77×10^{-08}
2	138	0.04±0.02	0.11±0.04	−0.81647	3.02×10^{-34}
3	164	0.03±0.03	0.14±0.03	−0.85447	6.14×10^{-48}
4	141	0.02±0.02	0.16±0.02	−0.80399	3.50×10^{-33}
5	12	0.03±0.02	0.13±0.02	−0.84403	0.0006
6	181	0.03±0.02	0.14±0.03	−0.81946	3.88×10^{-45}
(c) SL337					
All genotypes	401	0.08±0.03	0.04±0.04	−0.58857	9.6×10^{-39}
1	35	0.06±0.02	0.09±0.02	−0.70531	2.23×10^{-06}
2	109	0.06±0.02	0.05±0.03	−0.56761	1.23×10^{-10}
3	69	0.11±0.02	0.01±0.02	−0.44659	0.0001
4	49	0.07±0.02	0.04±0.03	−0.79646	7.76×10^{-12}
5	10	0.07±0.02	0.02±0.04	−0.71781	0.0194
6	129	0.08±0.04	0.02±0.03	−0.43117	3.38×10^{-07}

Table 3. Comparison of adaptive versus intrinsic robustness for IRES IIIabc, SL248 and SL337 stem-loops.

	IRES IIIabc	SL248	SL337
Percent of robust sequences	97.5%	97.4%	98.3%
Percent of non-robust sequences	2.5%	2.6%	1.7%
Percent of significantly robust sequences	68%	3.6%	30.2%
Percent of significantly non-robust sequences	0%	0.1%	0%

in the SL248 stem-loop most of the robustness comes from the intrinsic robustness. In the IRES IIIabc stem-loop we obtain intrinsic robustness close to zero and even negative, which means that all of the robustness in this stem-loop comes from the adaptive robustness. In IRES IIIabc the dinucleotide composition of the sequence appears to be as robust as the secondary structure of this stem-loop (or even more,

$\eta_s \geq \eta_c$), and indeed, in HCV, IRES is highly conserved in both sequence and structure. For example, an indication for the important role of primary sequence in IRES can be found in several works, e.g. Ref. 27. In order to verify that these results do not come from the long sequence of the IRES IIIabc stem-loop in our analysis, we partitioned the IIIabc (112 nt), and checked the smaller stem-loop IIIb (46 nt). The results obtained are the same as that for the full IIIabc stem-loop.

4. Discussion

We have reported on a detailed statistical analysis for the IIIabc, SL248 and SL337 functional stem-loops, which encompass all six HCV genotypes. The core protein is a multifunctional protein that besides composing the viral nucleocapsid, which packages the viral genomic RNA, is found to be modulating a number of cellular regulatory functions like apoptosis (promoting host's cell death) and cell proliferation by physical interaction and inhibition of tumor suppression genes (p53).[28] In addition, a more specific role has been attributed to SL248 in replication. It has been shown that a combination of mutations induced in a number of stem-loops in the core encoding region, including a mutation in SL248, delayed viral replication kinetics and reduced viral titers.[21] Considering the known location and properties of all three stem-loops analyzed, no evidence was found to speculate in advance that one stem-loop should be more mutationally robust than the others. The computational results show that stem-loop III from IRES is more mutationally robust than core stem-loops SL248 and SL337, and that for IRES most of the robustness calculated comes from the adaptive robustness. This might be attributed to an evolutionary origin, but alternative hypotheses cannot be ruled out when comparing in between these stem-loops.

Focusing on the mutational robustness of the three representative HCV stem-loop structures, this study provides further evidence for direct evolution of increased robustness in HCV. Theoretically, it is possible that selection might have favored a more stable folding and these more stable structures may happen to be more robust to mutations. In such a case the increased robustness is a secondary effect of selection for another feature (for example, thermodynamic stability, which is positively associated with mutational robustness[10,12]). However, viral RNA secondary structure motifs in general and HCV motifs in particular, such as the stem-loops taken in this study, need to be sufficiently flexible to unfold for processes like replication or translation in HCV,[21] making thermodynamic stability less likely as the primary effect. Moreover, deleterious mutations have been shown in HCV functional elements that lead to increased thermodynamic stability (e.g. Ref. 20). In addition, no correlation was found between GC content and mutational robustness in the HCV sequences examined in our study. Even though these arguments support direct selection for mutational robustness in HCV, a larger set of experimentally studied RNA secondary structure elements is required in order to reach a conclusive statement. Nevertheless, by using a larger number of sequences than

previous studies,[1] our findings strongly suggest the possibility of direct selection for robustness in an important class of viral RNAs. Furthermore, it suggests that HCV RNA sequences are more suitable than miRNAs for robustness studies that involve reliable folding predictions by energy minimization. This is because as shown in this study, enough HCV sequences that fold in a similar way to what is obtained in biological experiments can be collected for performing a proper statistical analysis, whereas for miRNAs it was shown in Ref. 29 that out of a representative set of 10 microRNA precursors, the majority of experimentally determined structures differ from those predicted. In principle, more types of biologically relevant RNA molecules other than viral RNAs can be taken for assessing their mutational robustness, but caution is advised and the reliability of the predictions should also be considered.

Supplementary Material

The Supplementary Material consists of four files: a general figure file, a table for IRES file, a table for SL337 file, and a table for SL248 file. The files can be downloaded from http://www.worldscinet.com/jbcb/.

Acknowledgments

The work was supported by the Lynn and William Frankel Center for Computer Sciences and the Kreitman Foundation Fellowship for A.C. at Ben-Gurion University.

References

1. Wagner A, Stadler PF, Viral RNA and evolved mutational robustness, *J Exp Zool* **285**:119–127, 1999.
2. Sanjuán R, Forment J, Elena SF, In silico predicted robustness of viroids RNA secondary structures. I. The effect of single mutations, *Mol Biol Evol* **23**(7):1427–1436, 2006.
3. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P, Fast folding and comparison of RNA secondary structures, *Monatsh Chem* **125**:167–188, 2006.
4. Borenstein E, Ruppin E, Direct evolution of genetic robustness in microRNA, *Proc Natl Acad Sci USA* **103**(17):6593–6598, 2006.
5. Shu W, Ni M, Bo X, Zheng Z, Wang S, In silico genetic robustness analysis of secondary structural elements in the miRNA gene, *J Mol Evol* **67**:560–569, 2008.
6. Szollosi GJ, Derenyi I, Congruent evolution of genetic and environmental robustness in microRNAs, *Mol Biol Evol* **26**:867–874, 2009.
7. Kitano H, Biological robustness, *Nat Rev Genet* **5**:826–837, 2004.
8. Wagner A, Robustness and evolvability: A paradox resolved, *Proc R Soc B* **275**: 91–100, 2008.
9. Schuster P, Fontana W, Stadler PF, Hofacker IL, From sequences to shapes and back: A case study in RNA secondary structures, *Proc R Soc London Ser B* **255**:279–284, 1994.

10. Huynen M, Stadler PF, Fontana W, Smoothness with ruggedness: The role of neutrality in adaptation, *Proc Natl Acad Sci* **93**:397–401, 1996.
11. Van Nimwegen E, Crutchfield JP, Huynen M, Neutral evolution of mutational robustness, *Proc Natl Acad Sci* **96**:9716–9720, 1999.
12. Ancel LW, Fontana W, Plasticity, evolvability, and modularity in RNA, *J Exp Zool* **288**(3):242–283, 2000.
13. Montville R, Froissart R, Remold SK, Tenaillon O, Turner PE, Evolution of mutational robustness in an RNA virus, *PLoS Biology* **3**:1939–1945, 2005.
14. Kuiken C, Yusim K, Boykin L, Richardson R, The los alamos HCV sequence database, *Bioinformatics* **21**(3):379–384, 2005.
15. Zuker M, Stiegler P, Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information, *Nucleic Acids Res* **9**:133–148, 1981.
16. Mathews DH, Sabina J, Zuker M, Turner DH, Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J Mol Biol* **288**:911–940, 1999.
17. Hofacker IL, Vienna RNA secondary structure server, *Nucleic Acids Res* **31**: 3429–3431, 2003.
18. Zuker M, Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res* **31**:3406–3415, 2003.
19. Shi ST, Lai MC, HCV 5′ and 3′ UTR: When translation meets replication, in Tan S-L (ed.), *Hepatitis C Viruses: Genomes and Molecular Biology*, Horizon Scientific Press, Norfolk, UK, pp. 49–88, 2006.
20. Tang S, Collier AJ, Elliot RM, Alterations to both the primary and predicted secondary structure of stem-loop IIIc of the Hepatitis C virus 1b 5′ Untranslated Region (5′ UTR) lead to mutants severely defective in translation which cannot be complemented in trans by the wild-type 5′ UTR sequence, *J Virol* **73**(3):2359–2364, 1999.
21. Vassilaki N, Friebe P, Meuleman P, Kallis S, Kaul A, Paranhos-Baccala G, Leroux-Roels G, Mavromara P, Bartenschlager R, Role of the hepatitis C virus core+1 open reading frame and core cis-acting RNA elements in viral RNA translation and replication, *J Virol* **82**(23):11503–11515, 2008.
22. You S, Stump DD, Branch AD, Rice CM, A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for Hepatitis C Virus RNA replication, *J Virol* **78**:1352–1366, 2004.
23. Tuplin A, Wood J, Evans DJ, Patel AH, Simmonds P, Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus, *RNA* **8**:824–841, 2002.
24. Simmonds P, Tuplin A, Evans DJ, Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: Implication for virus evolution and host persistence, *RNA* **10**:1337–1351, 2004.
25. Churkin A, Barash D, RNAMute: RNA secondary structure mutation analysis tool, *BMC Bioinformatics* **7**:221, 2006.
26. Clote P, Ferre F, Kranakis E, Krizanc D, Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency, *RNA* **11**(5):578–591, 2005.
27. Barria MI, Gonzalez A, Vera-Otarola J, Leon U, Vollrath V, Marsac D, Monasterio O, Perez-Acle T, Soza A, Lopez-Lastra M, Analysis of natural variants of hepatitis C virus internal ribosome entry site reveals that primary sequence plays a key role in cap-independent translation, *Nucleic Acids Res* **37**(3):957–971, 2003.
28. Alisi A, Giambartolomei S, Cupelli F, Merlo P, Fontemaggi G, Spaziani A, Bolsano C, Physical and functional interaction between HCV core protein and the different p73 isoforms, *Oncogene* **22**(17):2573–2580, 2003.

29. Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczysnska D, Krzyzosiak WJ, Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design, *J Biol Chem* **279**(40):42230–42239, 2004.

Alexander Churkin received his B.Sc. and M.Sc. degrees from the Department of Computer Science at Ben Gurion University of the Negev, Beer-Sheva, Israel, in 2004 and 2006, respectively. Since 2006, he has been pursuing his Ph.D. degree in the Department of Computer Science at Ben Gurion University of the Negev as a Kreitman Doctoral Fellow. His research interests include bioinformatics, RNA structure prediction, and scientific computing.

Moriah Cohen received her B.Med.Sc. degree from the Sackler School of Medicine at Tel Aviv University, Israel, in 2008. Since then, she has been a bioinformatics student in the Department of Computer Science at Ben Gurion University of the Negev, Beer-Sheva, Israel. Her research interests include bioinformatics and biological sciences.

Yonat Shemer-Avni received her Ph.D. degree in Life Sciences in 1989 from Ben Gurion University of the Negev, Beer Sheva, Israel. Between 1990 and 1992 she was a postdoctoral fellow at the Weizmann Institute of Science, Rehovoth, identifying the murine receptor for tumor necrosis factor type I. From 1992 to 1996 she worked as co-investigator in M.E.R.I.D. project "Natural history of chronic viral hepatitis." From 1995 to 2001, she was the head of the Laboratory for Molecular Diagnostic and Research of Viral Hepatitis. From 2001 to present she is the head of Clinical Virology Unit and Frozen Tissues Service, of Soroka Academic Medical Center, Beer Sheva, Israel. Currently, her main research focus is viruses and their microRNAs genes.

Danny Barash received his Ph.D. degree in applied science in 1999 from the University of California, Davis. From 1999 to 2001, he was employed at Hewlett-Packard Laboratories in the Technion, Israel, pursuing research on image processing and computer vision. From 2001 to 2003, he was a Howard Hughes Medical Institute Postdoctoral Fellow at New York University and a Research Fellow at the Institute of Evolution in the University of Haifa, Israel, where he made a transition to computational biology. Since 2004, he has been with the Department of Computer Science at Ben Gurion University of the Negev, Beer-Sheva, Israel, where he is currently an Associate Professor. His research interests include computational biology, RNA structure predictions, computational imaging, and numerical analysis.