In Silico Predicted Robustness of Viroids RNA Secondary Structures. I. The Effect of Single Mutations

Rafael Sanjuán, Javier Forment, and Santiago F. Elena

Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV), Avenida de los naranjos s/n, València, Spain

Viroids are plant subviral pathogens whose genomes are constituted by a single-stranded and covalently closed small RNA molecule that does not encode for any protein. Despite this genomic simplicity, they are able of inducing devastating symptoms in susceptible plants. Most of the 29 described viroid species fold into a rodlike or quasi-rodlike structure, whereas a few of them fold as branched structures. The shape of these RNA structures is perhaps one of the most characteristic properties of viroids and sometimes is considered their only phenotype. Here we use RNA thermodynamic secondary structure prediction algorithms to compare the mutational robustness of all viroid species. After characterizing the statistical properties of the distribution of mutational effects on structure stability and the wideness of neutral neighborhood for each viroid species, we show an evolutionary trend toward increased structural robustness during viroid radiation, giving support to the adaptive value of robustness. Differences in robustness among the 2 viroid families can be explained by the larger fragility of branched structures compared with the rodlike ones. We also show that genomic redundancy can contribute to the robustness of these simple RNA genomes.

Introduction

The plant pathogens known as viroids are small single-stranded RNAs (246–401 nt) about 10-fold smaller than the smallest RNA virus. They are circular, covalently closed molecule with a high degree of self-complementation resulting in a compact folding. Indeed, this RNA folding represents, together with the induced symptoms in susceptible plants, one of the few identifiable phenotypes of most viroids. Viroids, in contrast with viruses, do not encode for any proteins and entirely rely on the host's translational and transcriptional factors to complete their infectious cycle. Some viroids are catalytic RNAs that contain hammerhead-type ribozymes for self-cleavage of the multimeric RNAs generated during their rolling circle replication. The small size, circular structure, high G + C content, structural periodicities, and, specially, the catalytic activity exhibited by some viroids make them excellent candidates for being remnants of the precellular RNA world (Diener 1989, 1991, 2001). After the appearance of cellular organisms, some RNA molecules would have evolved to parasitize first certain cyanobacteria (the precursors of chloroplasts) and then eukaryotic cells, with mutation and recombination events contributing to their subsequent divergence (Elena et al. 1991, 2001). Viroids propagate in their plant hosts as populations of closely related variants (quasi-species), with usually one or few genotypes dominating the population.

Phylogenetic studies (Elena et al. 1991, 2001) along with structural and phenomenological properties support the current taxonomic classification of viroids (Flores, Randles et al. 2005) (table 1). Twenty-five out of the 29 known species belong to the family *Pospiviroidae* (named after PSTVd, the type member of the family) (table 1). Pospiviroids adopt in vitro a rodlike or quasi-rodlike secondary structure of minimal energy with 5 structural domains. Depending on the presence/absence of some of these domains, members of the family are assigned to 5 different genera. Pospiviroids replicate and accumulate in the nucleus. The other 4 viroids form the family *Avsunviroidae* (named after ASBVd, the type member of the family) (table 1). In general, avsunviroids' RNAs do not fold into rodlike structures but into less-organized branched shapes. Avsunviroids lack the characteristic structural motifs of the pospiviroids and contain a hammerhead ribozyme. Apart from the common hammerhead structures, sequence similarities among avsunviroids are remarkably low. Avsunviroids replicate and accumulate in the chloroplast. The comprehensive review of Flores, Hernández et al. (2005) is a good starting point to learn more about viroids' intriguing biology.

The folding of RNA sequences into secondary structures is a simple but biophysically well-grounded and powerful model for studying the mapping relationships between genotype and phenotype (Fontana 2002). These in silico studies have been of extreme help for illustrating how populations of RNA replicons can drift throughout regions of the sequence space where mutations are selectively neutral (neutral networks), hence becoming robust to deleterious mutational effects yet remaining evolvable (Schuster et al. 1994; Fontana and Schuster 1998; Schuster and Fontana 1999; Wagner and Stadler 1999; Ancel and Fontana 2000; Ancel Meyers et al. 2004); for evaluating whether mutational robustness can evolve as a correlated response to robustness against environmental changes (plastogenetic congruence) (Ancel and Fontana 2000; Ancel Meyers et al. 2004); for evaluating the contribution of compensatory mutations to the observed excess of antagonistic epistasis in RNA genomes (Wilke et al. 2003); for showing that most of the conformational order within biologically relevant RNA structures results not from the action of natural selection but from self-organization of RNA folding (Schultes et al. 1999); and for exploring the shape of the distribution of beneficial fitness effects (Cowperthwaite et al. 2005). In general, all these studies show that RNA structures can be relatively robust to changes in individual nucleotides and that the map between genotypic (sequence) and phenotypic (folding) spaces is not one-to-one because the sequence space contains extensive neutral networks connecting neighbors that fold into identical structures (Schuster et al. 1994; Schuster and Fontana 1999). However, the neutrality of RNA structures is lower than for proteins, where

Key words: mutational robustness, neutrality, plant pathogens, RNA folding, viroids.

E-mail: sfelena@ibmcp.upv.es.

Mol. Biol. Evol. 23(7):1427–1436. 2006
doi:10.1093/molbev/msl005
Advance Access publication May 5, 2006

Table 1
Sequences and Taxonomic Classification of the Viroid Species

Family	Genus	Species	Abbreviation	Accession	Size (nt)
Pospiviroidae	*Pospiviroid*	Potato spindle tuber viroid	PSTVd	V01465	359
		Tomato chlorotic dwarf viroid	TCDVd	AF162131	360
		Mexican papita viroid	MPVd	L78454	360
		Tomato planta macho viroid	TPMVd	K00817	360
		Citrus exocortis viroid	CEVd	M34917	371
		Chrysanthemum stunt viroid	CSVd	V01107	356
		Tomato apical stunt viroid	TASVd	K00818	360
		Iresine 1 viroid	IrVd-1	X95734	370
		Columnea latent viroid	CLVd	X15663	370
	Hostuviroid	Hop stunt viroid	HSVd	X00009	297
	Cocadviroid	Coconut cadang-cadang viroid	CCCVd	J02049	246
		Coconut tinangaja viroid	CTiVd	M20731	254
		Hop latent viroid	HLVd	X07397	256
		Citrus IV viroid	CVd-IV	X14638	284
	Apscaviroid	Apple scar skin viroid	ASSVd	M36646	329
		Citrus III viroid	CVd-III	AF184147	294
		Apple dimple fruit viroid	ADFVd	X99487	306
		Grapevine yellow speckle 1 viroid	GYSVd-1	X06904	367
		Grapevine yellow speckle 2 viroid	GYSVd-2	J04348	363
		Citrus bent leaf viroid	CBLVd	M74065	318
		Pear blister canker viroid	PBCVd	D12823	315
		Australian grapevine viroid	AGVd	X17101	369
	Coleviroid	*Coleus blumei* 1 viroid	CbVd-1	X52960	248
		C. blumei 2 viroid	CbVd-2	X95365	301
		C. blumei 3 viroid	CbVd-3	X95364	361
Avsunviroidae	*Avsunviroid*	Avocado sunblotch viroid	ASBVd	J02020	247
	Pelamoviroid	Peach latent mosaic viroid	PLMVd	M83545	337
		Chrysanthemum chlorotic mottle viroid	CChMVd	Y14700	399
	Elaviroid	Eggplant latent viroid	ELVd	AJ536613	333

>90% of single amino acid substitutions is neutral (Bornberg-Bauer and Chan 1999). This difference has been interpreted as RNA occupying large, strongly ramified but sparsely connected networks in sequence space (Huynen et al. 1996) while proteins concentrate in dense clusters (Taverna and Goldstein 2002; Wroe et al. 2005).

As Wagner (2005) wrote : "RNA is an attractive molecule for studies of mutational robustness because of its importance in past and present life. It may have been life's earliest information carrier and biocatalyst. And although DNA and proteins have superseded RNA in these roles, RNA still is involved in most of life's key processes. Examples go well beyond the well-known messenger, transfer, and ribosomal RNAs. They include small nuclear RNAs, which are key parts of the splicing machinery; guide RNAs important in RNA editing; telomerase RNA necessary for maintaining chromosome ends; the signal recognition particle RNA involved in translocating proteins through membranes; and RNA molecules regulating gene expression through RNA interference." Viroids should be added to this list because they manipulate many of the cellular functions in their own benefit, and they do so by acquiring a well-defined secondary structure. Therefore, the robustness of viroids secondary structure can be taken as a proxy for the robustness of their fitness and virulence.

We are presenting the results of our analyses as a 2-paper series. In this first paper, we explore the average properties of point mutations and whether a trend for the evolution of increasing structural robustness can be detected in viroid species. Taking advantage of the existence of viroid genomes with partial duplications, we also explore the interplay between increasing genome redundancy and genome robustness. In the second paper, we will explore the interaction between deleterious mutations, particularly whether mutations interact synergistically, multiplicatively, or antagonistically, seeking for the existence of a correlation between the magnitude of mutational effects and the strength of epistatic interactions established between mutations.

Materials and Methods

Viroid Sequences

The sequences of each viroid species listed in table 1 were downloaded from the National Center for Biotechnology Information Web site (http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/12884.html).

RNA Folding and Folding Robustness

RNA minimum free-energy, E_i, secondary structures (MFESSs) were computed using the algorithm implemented in the RNAfold program of the ViennaRNA package v1.5 (Hofacker et al. 1994). Folding temperature was set to 25 °C, which corresponds to the physiological temperature at which viroids usually replicate. Comparison among structures was done using the RNAdistance program of the same package. This comparison gives a Hamming distance, $d_{i,j}$, between the bracket notations of both secondary structures.

For each viroid, the impact of all possible $3L$ (L, genome length) one-error mutants on the predicted MFESS was evaluated by calculating the Hamming distance to

the wild-type secondary structure. Mutational effects $s(d_{i,j})$ were scaled to genome length: $s(d_{i,j}) = d_{i,j}/L$. As a second measure for robustness, the size of the one-step mutational neutral neighborhood (Schuster et al. 1994) was estimated as the fraction of all $3L$ mutants that fold exactly into the same MFESS than the wild type.

The term plastogenetic congruence was coined by Ancel and Fontana (2000) to describe the correlation between environmental and mutational effects on RNA folding shape. The link between environmental and genetic robustness may provide an explanation for the evolution of the latter provided that environmental fluctuations are frequent. In the case of RNA folding, environmental robustness can be measured as the average structural Hamming distance between suboptimal folds and the MFESS (Ancel and Fontana 2000). The Boltzmann probability of the ith conformation, which determines the time a sequence spends folded into a given structure, is $b_i = \exp(-E_i/kT)/Z$, where k is the Boltzmann constant, T the folding absolute temperature, and $Z = \sum_j \exp(-E_j/kT)$ is the partition function (Ancel and Fontana 2000; Ancel Meyers et al. 2004). This probability can be estimated by obtaining all the suboptimal structures within a user-defined energy interval, δ (whose lower end corresponds with the MFESS), and using their energies to compute b_i. The average structural Hamming distances between suboptimal folds and the optimum, $d_{i,\mathrm{opt}}$, scaled to genome length, is then $s(d_{i,\mathrm{opt}}) = \sum_{i\in\delta} d_{i,\mathrm{opt}} b_i/LZ$. The number of suboptimal structures rapidly increases with sequence length. Given the length of viroid sequences and for computational reasons, $\delta = 0.5$ kcal/mol. These computations were done with the RNAsubopt program from the ViennaRNA package. The relationship between environmental and mutational robustness was assessed using the correlation coefficient between the respective estimates obtained for each viroid.

PERL scripts were written for automating all the necessary computations and comparisons involving programs from the ViennaRNA package as well for the purpose of the Monte-Carlo simulations described in the Discussion. These scripts are available on request. All experiments were performed on a 12-node Linux cluster.

Statistical analyses were done using the SPSS package v12.0 (http://www.spss.com). The independent contrasts method (Felsenstein 1985) implemented in the CONTRAST program (PHYLIP package v3.65, http://evolution.gs.washington.edu/phylip.html) was used for assessing correlations between traits while removing the problem of nonindependence of data due to the underlying phylogenetic relationships. The rationale of this contrast is as follows: imagine 2 traits X and Y under study; for each trait, the difference between adjacent tips in the phylogenetic tree is computed and scaled by the variance of the trait weighted by the length of the branches leading to each tip. These values are the so-called contrasts and have expectation zero and variance one. Contrasts are independent because the difference $X_i - X_j$ only depends on events in branches i and j of the tree (Felsenstein 1985). Obviously, evolutionary events on different branches are independent. For a set of n species, $n - 1$ of such contrasts exists. The correlation between the set of independent contrasts constructed for traits X and Y is then obtained.

Results

Characterizing the Statistical Properties of the Distributions of Mutational Effects

Table 2 shows the statistics of the distribution of mutational effects for each viroid. Average mutational effects (i.e., the average Hamming distance between the structure computed for the ith one-error mutant and the one obtained for the wild type scaled by genome length) ranged between 1.08% (for CEVd, IrVd-1, CLVd, and AGVd) and 10.16% (for PBCVd), whereas standard deviations ranged between 0.0126 (for CbVd-3) and 0.2945 (for PBCVd). In all cases, the mean of the distribution of mutational effects was larger than the median, suggesting a common pattern across viroids species of mutational effects skewed toward mildly effects. Accordingly, skewness coefficients were highly significant (t-test, $P < 0.0001$). TPMVd had the most skewed distribution and PLMVd the most symmetrical one. Table 2 also shows the size of the one-step neutral neighborhood (i.e., the fraction of one-error mutants that do not change the wild-type structure). Neutrality ranged between 17.46% (PBCVd) and 25.88% (CCCVd). Interestingly, PBCVd shows the largest fitness effects, the lowest neutrality, and the widest distribution of mutational effects.

Phylogenetic Distribution of Robustness

Next, we explored whether the taxonomic relationships between viroids (table 1) would account for the differences in mutational effects. To do so, we computed a model II nested ANOVA in which species were nested within genera and genera within families. The error term of the model was constructed by looking at differences between all one-error mutants for each viroid species. Significant differences in mutational effects, that is, robustness, exist among species belonging to the same genus (table 3). Although this difference was significant, only 9.51% of the total observed variance for mutational effects was explained by differences between congenic species. No differences among sister genera existed, and, consequently, only 2.46% of the variance in robustness was explained by differences among cofamiliar genus. Finally, significant differences in the degree of mutational robustness existed among families, *Pospiviroidae* being twice more robust than *Avsunviroidae*. The average fitness effect of point mutations for *Pospiviroidae* was 0.0507 ± 0.0008, whereas for *Avsunviroidae*, it was twice as large (0.1050 ± 0.0030). The percentage of variation explained by differences among families was 19.94%. Finally, 68.09% of the total observed variation was explained by the differential effects that particular point mutations exert on each viroid's predicted RNA folding.

Next, we sought for an evolutionary trend toward increased robustness. To do so, we mapped mutational effects into the phylogenetic tree proposed by Elena et al. (2001). The root of this phylogenetic tree was placed by comparing viroid sequences with viroidlike (the so-called virusoids) and linear RNA satellites of viruses, with which they share certain structural properties (Diener 1989, 1991; Elena et al. 1991, 2001). Figure 1 shows the estimated average mutational effect for each genus along with the phylogenetic tree linking the different genera. According to this tree, the split between pelamoviroids and avsunviroids took place earlier

Table 2
Statistical Properties of the Distribution of Single Point Mutations on Viroids Structure (see text for explanation)

Family	Genus	Species	Mean	Median	Standard Deviation	Skewness	Neutral Neighborhood Size
Pospiviroidae	*Pospiviroid*	PSTVd	0.0168	0.0111	0.0278	7.0129	0.2358
		TCDVd	0.0179	0.0111	0.0235	5.1710	0.2250
		MPVd	0.0215	0.0111	0.0458	12.3952	0.2102
		TPMVd	0.0187	0.0111	0.0425	13.0593	0.2176
		CEVd	0.0573	0.0108	0.1052	2.0998	0.2022
		CSVd	0.0541	0.0112	0.1061	2.3615	0.2416
		TASVd	0.0639	0.0111	0.1149	2.0806	0.1972
		IrVd-1	0.0169	0.0108	0.0243	5.1383	0.2252
		CLVd	0.0319	0.0108	0.0808	3.9189	0.2333
	Hostuviroid	HSVd	0.0357	0.0135	0.0631	4.0391	0.2121
	Cocadviroid	CCCVd	0.0192	0.0163	0.0184	1.3414	0.2588
		CTiVd	0.0224	0.0157	0.0309	4.7726	0.2585
		HLVd	0.0680	0.0156	0.1384	2.6962	0.2526
		CVd-IV	0.0162	0.0141	0.0154	1.8219	0.2500
	Apscaviroid	ASSVd	0.0266	0.0122	0.0408	6.1418	0.2209
		CVd-III	0.0726	0.0136	0.1307	2.2358	0.1973
		ADFVd	0.0625	0.0131	0.1400	3.0708	0.1787
		GYSVd-1	0.1283	0.0109	0.2852	2.2167	0.2216
		GYSVd-2	0.0607	0.0110	0.1145	2.8578	0.2057
		CBLVd	0.0276	0.0126	0.0569	5.5226	0.2107
		PBCVd	0.2434	0.1016	0.2945	1.1441	0.1746
		AGVd	0.0299	0.0108	0.0678	3.8381	0.2448
	Coleviroid	CbVd-1	0.0992	0.0161	0.1706	1.8641	0.2016
		CbVd-2	0.0620	0.0133	0.1179	2.2679	0.1938
		CbVd-3	0.0128	0.0111	0.0126	4.4700	0.2318
Avsunviroidae	*Avsunviroid*	ASBVd	0.0507	0.0243	0.0739	2.2418	0.2524
	Pelamoviroid	PLMVd	0.1246	0.0119	0.1829	1.1067	0.2463
		CChMVd	0.1823	0.0201	0.2694	1.1523	0.2389
	Elaviroid	ELVd	0.0333	0.0119	0.0537	2.9960	0.2418

during the evolutionary history of viroids, whereas hostuviroids, pospiviroids, and cocadviroids were the most recently divergent genera. Notice that ELVd has not been included in this analysis because its exact location in the phylogenetic tree has not been determined yet. We assigned a rank order to each node, starting from the oldest one and using tide ranks for genera derived from unresolved nodes. Under the hypothesis of an evolutionary trend toward increasing mutational robustness, a negative correlation would be expected between the above ranks (i.e., phylogenetic deepness) and average mutational effects. A negative and significant partial correlation coefficient, controlling for family ($r = -0.4502$, 25 df, $P = 0.0185$), gave support to this hypothesis. Extra support to this evolutionary trend toward increased robustness comes from the average mutational effect for the viroidlike RNA satellites (0.1503 ± 0.0646) being as high as for the pelamoviroids.

Table 3
Model II Nested ANOVA Assessing for Differences in Mutational Robustness at Different Taxonomic Levels

Source of Variation[a]	Type III Sum of Squares	df	Mean Squares	Variance Component[b]	*F*	*P*
Family	66.9354	2	33.4677	4.754×10^{-3}	9.6077	0.0068
Genera	24.4577	6	4.0763	5.861×10^{-4}	1.9952	0.1119
Species	46.7227	21	2.2249	2.268×10^{-3}	137.0749	<0.0001
Error	461.7290	28447	0.0162	1.623×10^{-2}		

[a] The factor species was nested within genera and genera within family.
[b] Maximum likelihood estimates.

Correlation between Robustness and the Size of the Neutral Neighborhood

Would mutational robustness be associated with an increase in the size of the one-step neutral neighborhood? The straight forward way of looking for an association between average mutational effects and neutrality values (table 2) would be to compute a correlation coefficient between these 2 traits. Nonetheless, we want to stress out that a statistical correlation does not necessarily means a cause–effect relationship.

Because viroid species are part of a hierarchically structured phylogeny, the pairs of data cannot be considered as independently drawn from the same distribution. However, this nonindependence problem can be circumvented by considering the phylogenetic relationships between species (fig. 2 in Elena et al. 2001). Twenty-seven independent contrasts were constructed, and a correlation coefficient between them was computed. The value of the correlation coefficient was significantly negative ($r = -0.4885$, 25 df, $P = 0.0097$), as hypothesized. Therefore, we conclude that neutrality goes hand-by-hand with mutational robustness for viroid species.

Differences in Robustness among Rodlike and Branched Structures

Next, we explored whether the magnitude of the mutational effects was homogeneous for all regions in the molecule, or, alternatively, mutations at some sites were

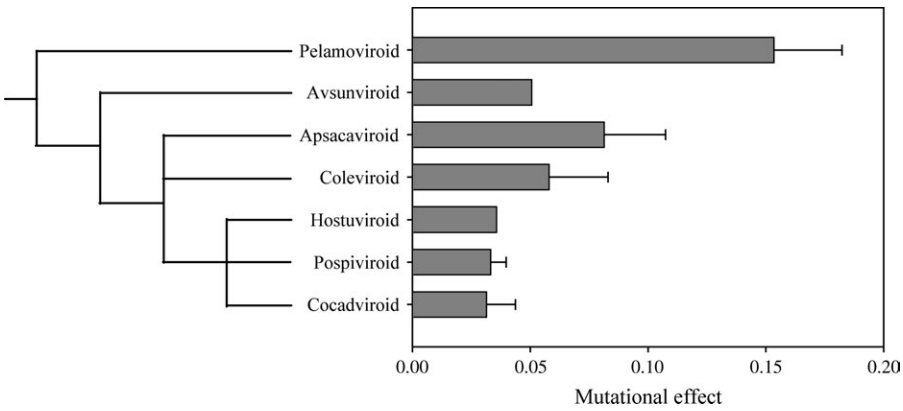

FIG. 1.—Phylogenetic distribution of mutational effects. The phylogenetic tree shown here was adapted from figure 2 in Elena et al. (2001). The location of the root was estimated by comparing viroid sequences with viroidlike RNA satellites. Error bars represent standard error of the mean.

particularly deleterious, whereas mutations at other sites were of milder effect. For illustrative purposes, figure 2 shows the distribution of average mutational effects along the genome for 2 viroids with radically different folding structures; the pospiviroid PSTVd, with its characteristic rodlike MFESS, and the avsunviroid PLMVd, with a highly branched predicted MFESS. In PSTVd, most sites have really minor effects, whereas a few sites have moderately large effects (always smaller than 20%). The latter appear to be randomly scattered along the molecule. By contrast, for PLMVd, the situation is quite different. A large fraction of sites have strong effects (up to 60%) on the predicted folding. Furthermore, these very sensitive sites are located around the right part of the molecule (fig. 2*B*), where most of the hairpin loops concentrate. These distinctions between rodlike and branched structures can be generalized to all viroid molecules (data not shown).

As it has been shown above, mutational effects were larger for *Avsunviroidae* than for *Pospiviroidae*. Figure 2 suggests that this difference strongly depends on whether the molecule is rodlike or branched. This being the case, a positive correlation is expected between the magnitude of mutational effects and the number of hairpin loops predicted for the MFESS. Figure 3 shows the relationship

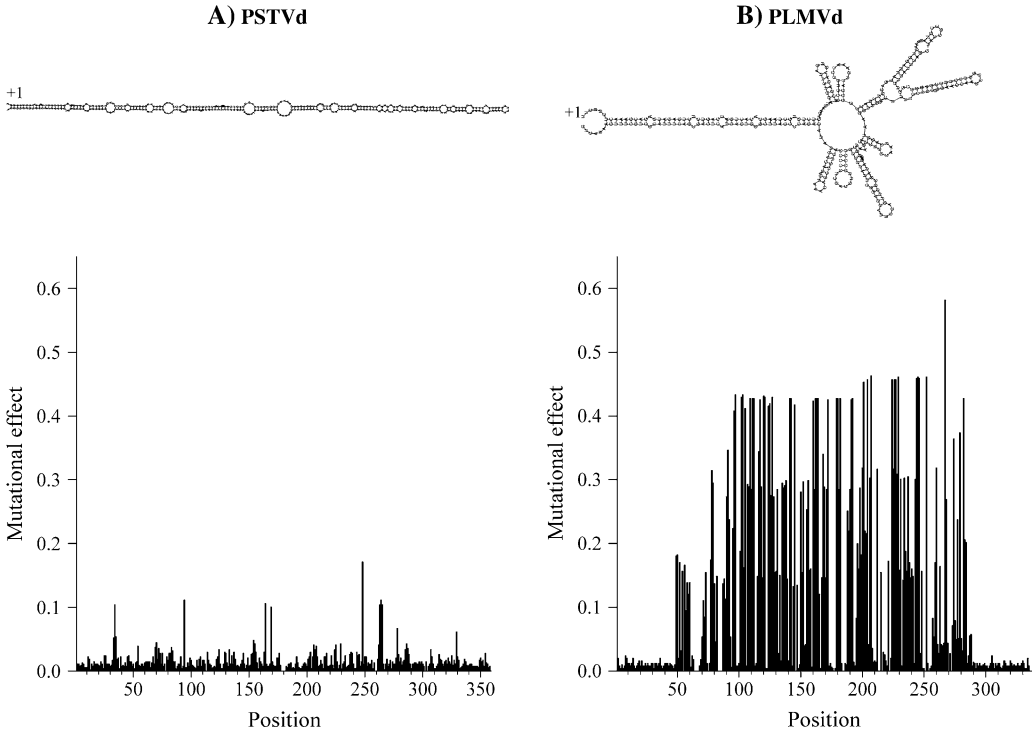

FIG. 2.—The magnitude of mutational effects do not distribute evenly along viroid molecules. Predicted minimum free-energy structure and distribution of mutational effects along the molecule for (*A*) PSTVd, the prototypic viroid with rodlike structure and (*B*) PLMVd, a pelamoviroid with a highly branched secondary structure.

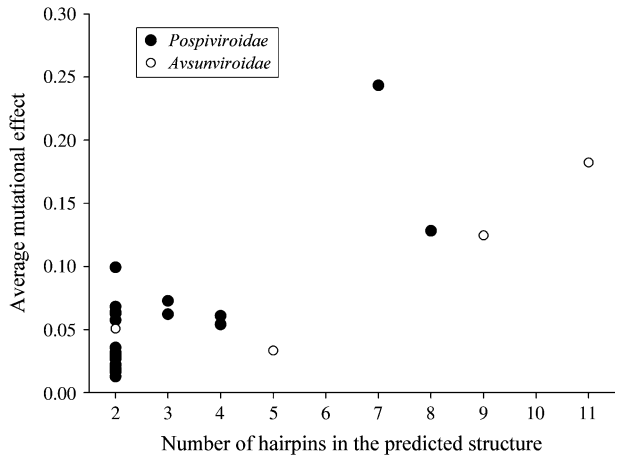

FIG. 3.—Relationship between the number of hairpins present in the predicted minimum free-energy structure and the average mutational effect.

between the average mutational effect estimated for each viroid species (from table 1) and the number of hairpins in their predicted structure. A positive and highly significant correlation was found between the independent contrasts ($r = 0.6957$, 25 df, $P < 0.0001$), giving support to the idea that structural robustness in viroids is inversely proportional to the number of hairpin loops in their structures.

The Effect of Genome Duplications on Robustness

Redundancy has been suggested to increase mutational robustness (Krakauer and Plotkin 2002). Viroids allow testing this prediction because isolates of several species contain partial genomic duplications encompassing functional domains or enlarged genomes. Therefore, direct comparisons of the mutational effects for nonredundant genomes (i.e., unitary length genomes) and partially redundant ones originated from the former are possible. The cases of CCCVd, CEVd, and CbVd will be used to illustrate this point.

In addition to a point mutation at position 70, the slow and fast variants of CCCVd differ in that the former has a duplication of 41 nt at the right terminal domain (Mohamed et al. 1982). The median effects of point mutations for CCCVd-fast and -slow are, respectively, 1.22% and 1.05%, the difference being statistically significant (fig. 4*A*; Mann–Whitney's U test, $P = 0.0002$). However, the reduction in the severity of mutational effects associated with partial genome duplication was not in hand with an increase in the size of the one-step neutral neighborhood. The frequency of neutral mutations for CCCVd-fast and -slow variants was, respectively, 38.57% and 38.81%, the difference between both being not significant (Fisher's exact test, $P = 0.7533$).

Compared with the reference variant listed in table 1 (also known as variant C), variant D-104 of CEVd, in addition to 7 point mutations, contains a duplication of 104 nt (Semancik and Durán-Vila 1999). This duplication also affects the right terminal domain of the folded molecule. The median mutational effect for the unitary length CEVd

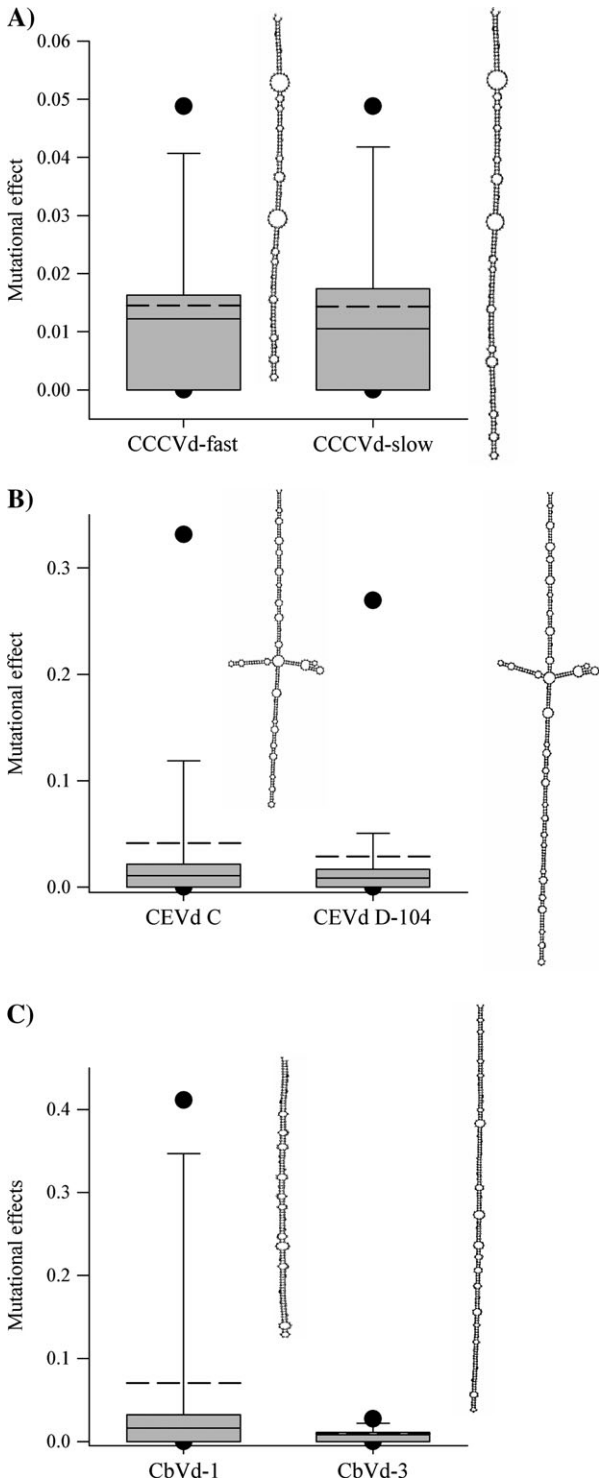

FIG. 4.—Box plot comparing the distribution of the fitness effects caused by point mutations in viroids with unitary and redundant genomes. (*A*) CCCVd-slow and -fast variants. (*B*) CEVd C and D-104 variants. (*C*) CbVd-1 and CbVd-3. The box represents the 25% and 75% percentiles; error bars correspond with the 5% and 95% percentiles; the continuous line within the box represents the median effect and the dashed line the average effect. Outliers are represented by dots. The predicted minimum free-energy structure for each sequence is presented next to the corresponding box.

C was 1.08%, whereas the median effect for D-104 was 0.84%, a significantly smaller figure (fig. 4*B*; *U* test, $P < 0.0001$). The observed increase in mutational robustness associated with the large duplication carried by CEVd D-104 was not concomitant with an increase in the size of the neutral neighborhood because it only changed from 36.19% to 35.81% (Fisher's exact test, $P = 0.7797$).

CbVd-1 is the shortest and CbVd-3 the largest among the coleviroids (116 nt longer than CbVd-1). In addition to extensive nucleotide differences, including insertions of short stretches of nucleotides in both the upper and lower strands of the predicted rodlike folded structure, the 2 terminal domains, including both terminal loops, of CbVd-3 resemble additions to the shorter CbVd-1 molecule, probably result of a recombination event in planta (Spieker 1996). Although this case does not strictly represent a duplication event of the shorter molecule, at least it represents another instance of increased genome complexity by adding extra genetic material. The median mutational effect for CbVd-1 is 1.61% and 0.82% for CbVd-3 (fig. 4*C*; *U* test, $P < 0.0001$). In this case, the size of the one-step neutral neighborhood associated with CbVd-3 (39.13%) was significantly increased compared with that of CbVd-1 (34.69%) (Fisher's exact test, $P = 0.0020$).

Here, we are interpreting these results as a consequence of redundant genomes being more robust. However, an alternative explanation exists: mutational effects may be milder in longer genomes simply because more Watson–Crick bonds can be stablished. To test this alternative hypothesis, we computed a correlation coefficient between the independent contrasts obtained for mutational effects and genome length. If the alternative hypothesis was true, then a negative correlation is expected among these variables. Obviously, to avoid confounding effects, CbVd-3 was excluded from this computation. A negative although not significant correlation was found ($r = -0.1633$, 24 df, one-tail $P = 0.2127$), which suggests that, if any, the effect of genome length was not large enough to explain the increased robustness observed for CCCVd-slow, CEVd D-104, and CbVd-3.

All in all, in good agreement with the redundancy hypothesis, mutational effects are milder for viroid genomes showing certain degree of redundancy than for their nonredundant counterparts. At least for the 2 viroids containing partial duplications, increased robustness was not achieved by increasing neutrality.

Assessing Plastogenetic Congruence

The term plastogenetic congruence refers to the correlation between environmental and mutational effects on the RNA folding shape. For each viroid, the suboptimal structures contained in an energy interval $\delta = 0.5$ kcal/mol were computed and compared with the MFESS used in previous analyses. Then a weighted average effect was computed, as described in the Material and Methods, to quantify the extent of thermal stability. Figure 5 shows that plastogenetic congruence is weak at best (correlation between independent contrasts: $r = 0.1864$, 25 df, $P = 0.3519$), implying that mutational robustness could hardly evolve as a consequence of selection promoting environmental robustness.

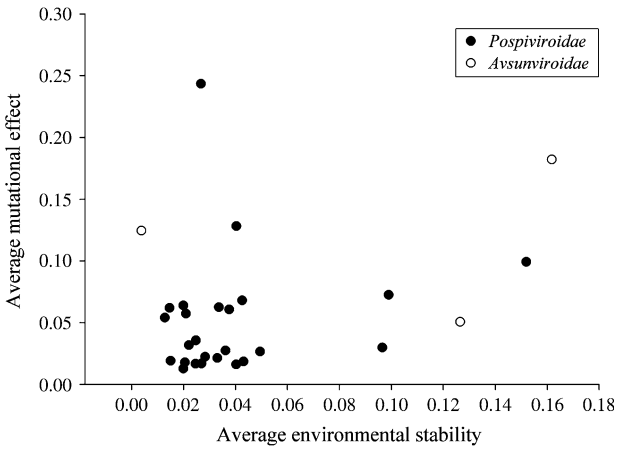

FIG. 5.—Correlation between mutational effects and environmental (or thermal) stability.

Discussion

The interest of evolutionary, molecular, and developmental biologists on robustness has increased in the recent years (Barkai and Leibler 1997; Bornholdt and Sneppen 2000; von Dassow et al. 2000; de Visser et al. 2003; Wilke and Adami 2003; Hermisson and Wagner 2004; Kitano 2004; Stelling et al. 2004; Wagner 2004; Elena et al. 2006). Robustness, broadly understood as the ability of a system to continue functioning in the face of environmental or genetic perturbations, has different underlying mechanisms ranging from thermodynamic stability of RNA and protein folding to redundancy in gene networks or animal behavior. Robustness is critical for evolution because it determines the phenotypic expression of underlying genetic variation. Because adaptive evolution requires phenotypic variation, robustness mechanisms may impose a limit to the rate of adaptation. Understanding how robustness mechanisms have evolved and how robustness, short-term, and long-term evolution are interconnected has become one of the most attractive areas of research for evolutionary biologists. Indeed, it is still not clear whether genetic robustness may evolve as a direct consequence of natural selection (adaptive robustness), as a by-product of stabilizing selection acting on fitness-related traits (intrinsic robustness), or as a correlated response to environmental robustness (congruent robustness or plastogenetic congruence) (de Visser et al. 2003). RNA folding represents an invaluable tool for studying the evolution of robustness (Fontana 2002): folding can be seen as a phenotype encoded by an RNA sequence and the thermodynamics rules governing folding as the genotype-to-phenotype map. These sort of in silico studies have shown that RNA structures possess some degree of robustness to changes in nucleotide sequence and environmental perturbations. They have also proved the existence of neutral regions in sequence space, wherein many sequences fold exactly into the same structure. Importantly, naturally occurring molecules have been shown to be more thermodynamically and mutationally robust than artificial ones and that RNA folding robustness can be congruent (Wagner and Stadler 1999; Ancel Meyers et al. 2004). Most of these studies have focused on artificial

RNA molecules, whereas only few of them have analyzed natural RNA sequences.

Viroids provide a unique opportunity for assessing robustness in biologically relevant RNA molecules and more importantly for understanding the evolution of mutational robustness. However, and despite their clear interest, viroids have never been used in any of the aforementioned studies, with the exception of a passing-through mention to PSTVd by Ancel Meyers et al. (2004). Here, we have tried to close this gap by exploring the robustness of all known viroid species. However, how biologically relevant are our in silico analyses? First, it is thought that RNA structures are essential for viroid pathogenicity (Keese and Symons 1985) and in the case of the *Avsunviroidae*, for adopting the hammerhead ribozyme conformation implicated in processing multimeric intermediates of replication into unitary genomes (Flores, Hernández et al. 2005). Therefore, viroid fitness critically depends on the right folding of their genomes. In fact, natural variation usually involves compensatory changes that preserve the MFESS. Examples are abundant (reviewed in Diener 2001), and here we are only mentioning a representative example from each viroid family. Site-directed mutagenesis experiments with PSTVd have shown that only double mutants in the hairpin II (important for replication) that did not alter the native rodlike structure were viable and stably recovered after several passages in tomato plants (Qu et al. 1993). By contrast, single mutants were either nonviable or reverted to the wild-type sequence (Qu et al. 1993). Recent studies with CChMVd have shown that mutations affecting stems relevant for maintaining a functionally important kissing loop either reverted to the wild-type sequence after inoculation of susceptible host plants or led to rearrangements that still allowed correct establishment of the kissing loop (de la Peña and Flores 2002; Gago et al. 2005). Thus, the RNA folding phenotype is preserved despite accumulating changes at the nucleotide level.

Second, if secondary structure can be taken as a proxy to viroid fitness, our in silico analyses should be able of predicting to some extent the in vivo mutational effects. Specifically, it would be expected that sequence variants drawn from natural populations should have smaller effects in the predicted secondary structure than randomly generated changes because the former are filtered by natural selection. To test this prediction, we browsed sequence databases for natural variants differing in a single point mutation from the standard wild-type sequence. Unfortunately, these data are scarce, and we were only able of gathering 6 natural one-substitution variants from CVd-III (Owens et al. 2000). The average mutational effect for these variants was 0.0159 ± 0.0082, which was 4.57 times lower than the average random effect for CVd-III (table 2). To assess whether these values were statistically different from the average effect of random mutations, the following Monte-Carlo approach was used: 1000 random samples of 6 single-nucleotide mutants were taken with replacement, and the average mutational effect for each sample was computed; then, the number of cases for which the estimated average mutational effect was ≤ 0.0159 was recorded. The probability of getting an average effect of this magnitude was lower than 6.80%. Despite its inconclusiveness, this result suggests that in silico predicted mutational effects on CVd-III RNA folding are a good approximation to in vivo mutational effects.

Our main results can be summarized as follows. First, the distribution of mutational effects on secondary RNA structures was highly skewed, with many mutations having small effects and a flat tail of strong mutational effects. This parallels the results obtained with a wide variety of species including the vesicular stomatitis RNA virus (Sanjuán et al. 2004), *Escherichia coli* (Elena et al. 1998), *Caenorhabditis elegans* (Keightley and Caballero 1997), and *Drosophila melanogaster* (Fernández and López-Fanjul 1996; Lyman et al. 1996). Neutrality, that is, the fraction of all point mutations that did not change structure, ranged between $\sim$17% and $\sim$26%.

Second, mutational effects significantly varied across viroid families and among viroid species from the same family, although most of the observed interspecies variability was explained by differences at the highest taxonomic level. The average mutational effects were larger for the members of the family *Avsunviroidae* than for the members of the *Pospiviroidae*. At a structural level, this was easily explained by the larger number of hairpin loops in the former because branched shapes were particularly prone to strong mutational effects (figs. 2 and 3). When phylogenetic information was incorporated into the analysis, a clear pattern of increasing robustness along evolutionary time was observed (fig. 1). It seems unlikely that this robustness has come up as a consequence of selection favoring stability to environmental perturbations because we failed to detect significant plastogenetic congruence (fig. 5). This suggests that robustness in viroid RNA structures might be adaptive per se, possibly driven by the thigh link existing between the structure and the biochemical functions of the molecule. In this sense, the hammerhead ribozyme structure characteristic of avsunviroids has been proved to be very sensitive to changes in the nucleotides involved in keeping the right structure (Ambrós et al. 1998; Flores et al. 1999).

Third, robustness can evolve by increasing redundancy in genomes or by modifying the network of genetic interactions (Krakauer and Plotkin 2002). The existence of several viroids containing either partial genome duplications (CCCVd and CEVd) or sequences able of forming redundant structures (CbVd), as result of recombination events, provide a good opportunity for seeking whether redundancy increases robustness. The comparison of these 3 viroid genomes with their corresponding nonredundant counterparts revealed that, in fact, the average effect of point mutations was significantly reduced in redundant genomes. This larger robustness was not a mere consequence of having longer genomes. In all 3 species, duplications are incomplete, and thus, redundancy is only partial. This might explain why, in these cases, the reduction in mutational effects was not necessarily accompanied by an increase in the size of the neutral neighborhood. In contrast, for the ensemble of viroid species, a significant negative correlation was detected between average mutational effects and the size of the one-step neutral neighborhood. Some viroid species may have more such neutral neighbors, and hence be more robust, whereas other may have fewer neighbors and be less robust. This correlation suggests that neutrality is, at least in part, responsible for the observed robustness. During

evolution and diversification of viroids, selection may have been driving viroid populations toward regions of sequence space, wherein neutrality is the highest possible. As it was mathematically proved (Huynen et al. 1996; Bornberg-Bauer and Chan 1999; van Nimwegen et al. 1999; Reidys et al. 2001; Wilke 2001), neutral networks are not homogenous but consist of regions (or islands) where many sequences fold into the same structure connected by bridges containing few sequences that fold into the same structure. Molecules folding into the functionally correct structure are favored by natural selection, whereas mutant sequences that fall out the neutral neighborhood are quickly eliminated from the population due to their low fitness. The concept of neutral networks may have interesting consequences for viroid evolution. A neutral network ensures that mutation can transform a sequence variant into another without fitness consequences, allowing for exploration of sequence space and hence fostering evolvability. Once a viroid species has expanded in sequence space and filled out one of such neutral islands, it may start moving throughout bridges to reach a different island and spread out there. It is conceivable that each such island in a neutral landscape may correspond to a different viroid species that may still show little variation at the phenotypic level (e.g., the rodlike structure). If so, viroid evolution could occur without traversing low-fitness valleys as assumed by the shifting-balance view of speciation (Wright 1982).

Acknowledgments

We are grateful to José A. Daròs, Ricardo Flores, and all the virology crew at the Instituto de Biología Molecular y Celular de Plantas for comments and suggestions. The manuscript was improved by the comments of 2 anonymous reviewers. This research was supported by grants BMC2003-00066 and BFU2005-23720-E/BMC from the Spanish Ministerio de Educación y Ciencia–FEDER and by the European Molecular Biology Organization Young Investigator Program to S.F.E.

Literature Cited

Ambrós S, Hernández C, Desvignes JC, Flores R. 1998. Genomic structure of 3 phenotypically different isolates of peach latent mosaic viroid: implications of the existence of constraints limiting the heterogeneity of viroid quasispecies. J Virol 72: 7397–406.

Ancel LW, Fontana W. 2000. Plasticity, evolvability, and modularity in RNA. J Exp Zool 288:242–83.

Ancel Meyers LW, Lee JF, Cowperthwaite M, Ellington AD. 2004. The robustness of naturally and artificially selected nucleic acid secondary structures. J Mol Evol 58:681–91.

Barkai N, Leibler S. 1997. Robustness in simple biochemical networks. Nature 387:913–7.

Bornberg-Bauer E, Chan HS. 1999. Modeling evolutionary landscapes: mutational stability, topology, and superfunnels in sequence space. Proc Natl Acad Sci USA 96:10689–94.

Bornholdt S, Sneppen K. 2000. Robustness as an evolutionary principle. Proc R Soc Lond B Biol Sci 267:2281–6.

Cowperthwaite MC, Bull JJ, Ancel Meyers L. 2005. Distributions of beneficial fitness effects in RNA. Genetics 170:1449–57.

de la Peña M, Flores R. 2002. Chrysanthemum chlorotic mottle viroid RNA: dissection of the pathogenicity determinant and comparative fitness of symptomatic and non-symptomatic variants. J Mol Biol 321:411–21.

de Visser JAGM, Hermisson J, Wagner GP et al. (19 co-authors). 2003. Evolution and detection of genetic robustness. Evolution 57:1959–72.

Diener TO. 1989. Circular RNAs: relics of precellular evolution? Proc Natl Acad Sci USA 86:9370–4.

Diener TO. 1991. Subviral pathogens of plants: viroids and viroid-like satellite RNAs. FASEB J 5:2808–13.

Diener TO. 2001. The viroid: biological oddity or evolutionary fossil? Adv Virus Res 57:137–84.

Elena SF, Carrasco P, Daròs JA, Sanjuán R. 2006. Mechanisms of genetic robustness in RNA viruses. EMBO Rep 7: 168–73.

Elena SF, Dopazo J, Flores R, Diener TO, Moya A. 1991. Phylogeny of viroids, viroidlike satellite RNAs, and the viroidlike domain of hepatitis δ virus RNA. Proc Natl Acad Sci USA 88:5631–4.

Elena SF, Dopazo J, de la Peña M, Flores R, Diener TO, Moya A. 2001. Phylogenetic analysis of viroid and viroid-like satellite RNAs from plants: a reassessment. J Mol Evol 53:155–9.

Elena SF, Ekunwe L, Hajela N, Oden SA, Lenski RE. 1998. Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. Genetica 102/103:349–58.

Felsenstein J. 1985. Phylogenies and the comparative method. Am Nat 125:1–15.

Fernández J, López-Fanjul C. 1996. Spontaneous mutational variances and covariances for fitness-related traits in *Drosophila melanogaster*. Genetics 143:829–37.

Flores R, Hernández C, Martínez de Alba E, Daròs JA, di Serio F. 2005. Viroids and viroid-host interactions. Annu Rev Phytopathol 43:117–39.

Flores R, Navarro JA, de la Peña M, Navarro B, Ambrós S, Vera A. 1999. Viroids with hammerhead ribozymes: some unique structural and functional aspects with respect to other members of the group. Biol Chem 380:849–54.

Flores R, Randles JW, Owens RA, Bar-Joseph M, Diener TO. 2005. Viroids. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball AL, editors. Virus taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses. London: Elsevier Academic Press. p 1145–59.

Fontana W. 2002. Modelling 'evo-devo' with RNA. Bioessays 24:1164–77.

Fontana W, Schuster P. 1998. Continuity in evolution: on the nature of transitions. Science 280:1451–5.

Gago S, de la Peña M, Flores R. 2005. A kissing-loop interaction in a hammerhead viroid RNA critical for its *in vitro* folding and *in vivo* viability. RNA 11:1073–83.

Hermisson J, Wagner GP. 2004. The population genetic theory of hidden variation and genetic robustness. Genetics 168: 2271–84.

Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P. 1994. Fast folding and comparison of RNA secondary structures. Monatsh Chem 125:167–88.

Huynen MA, Stadler PF, Fontana W. 1996. Smoothness within ruggedness: the role of neutrality in adaptation. Proc Natl Acad Sci USA 93:397–401.

Keese P, Symons RH. 1985. Domains in viroids: evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. Proc Natl Acad Sci USA 82:4582–6.

Keightley PD, Caballero A. 1997. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. Proc Natl Acad Sci USA 94:3823–7.

Kitano H. 2004. Biological robustness. Nat Rev Genet 5:826–37.

Krakauer DC, Plotkin JB. 2002. Redundancy, antiredundancy, and the robustness of genomes. Proc Natl Acad Sci USA 99:1405–9.

Lyman RF, Lawrence F, Nuzhdin SV, Mackay TF. 1996. Effects of single P-element insertions on bristle number and viability in *Drosophila melanogaster*. Genetics 143:277–92.

Mohamed NA, Haseloff J, Imperial JS, Symons RH. 1982. Characterization of the different electrophoretic forms of the cadang-cadang viroid. J Gen Virol 63:181–92.

Owens RA, Yang G, Gundersen-Rindal D, Hammond RW, Candresse T, Bar-Joseph M. 2000. Both point mutation and RNA recombination contribute to the sequence diversity of citrus viroid III. Virus Genes 20:243–52.

Qu F, Heinrich C, Loss P, Steger G, Tien P, Riesner D. 1993. Multiple pathways of reversion in viroids for conservation of structural elements. EMBO J 12:2129–39.

Reidys C, Forst CV, Schuster P. 2001. Replication and mutation on neutral networks. Bull Math Biol 63:57–94.

Sanjuán R, Moya A, Elena SF. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc Natl Acad Sci USA 101:8396–401.

Schultes EA, Hraber PT, LaBean TH. 1999. Estimating the contribution of selection and self-organization in RNA secondary structure. J Mol Evol 49:76–83.

Schuster P, Fontana W. 1999. Chance and necessity in evolution: lessons from RNA. Physica D 133:427–52.

Schuster P, Fontana W, Stadler PF, Hofacker IL. 1994. From sequences to shapes and back: a case study in RNA secondary structures. Proc R Soc Lond B Biol Sci 255: 279–84.

Semancik JS, Durán-Vila N. 1999. Viroids in plants: shadows and footprints of a primitive RNA. In: Domingo E, Webster R, Holland JJ, editors. Origin and evolution of viruses. San Diego, CA: Academic Press. p 37–64.

Spieker RL. 1996. *In vitro*-generated 'inverse' chimeric *Coleus blumei* viroids evolve *in vivo* into infectious RNA replicons. J Gen Virol 77:2839–46.

Stelling J, Sauer U, Szallasi Z, Doyle III FJ, Doyle J. 2004. Robustness of cellular functions. Cell 118:675–85.

Taverna DM, Goldstein RA. 2002. Why are proteins so robust to site mutations? J Mol Biol 315:479–84.

van Nimwegen E, Crutchfield JP, Huynen M. 1999. Neutral evolution of mutational robustness. Proc Natl Acad Sci USA 96:9716–20.

von Dassow G, Meir E, Munro EM, Odell GM. 2000. The segment polarity network is a robust developmental module. Nature 406:168–92.

Wagner A. 2004. Distributed robustness versus redundancy as causes of mutational robustness. Bioessays 27:176–88.

Wagner A. 2005. Robustness and evolvability in living systems. Princeton, NJ: Princeton University Press.

Wagner A, Stadler PF. 1999. Viral RNA and evolved mutational robustness. J Exp Zool 285:119–27.

Wilke CO. 2001. Adaptive evolution in neutral networks. Bull Math Biol 63:715–30.

Wilke CO, Adami C. 2003. Evolution of mutational robustness. Mutat Res 522:3–11.

Wilke CO, Lenski RE, Adami C. 2003. Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding. BMC Evol Biol 3:1–14.

Wright S. 1982. The shifting balance theory and macroevolution. Annu Rev Genet 16:1–19.

Wroe R, Bornberg-Bauer E, Chan HS. 2005. Comparing folding codes in simple heteropolymer models of protein evolutionary landscapes: robustness of the superfunnel paradigm. Biophys J 88:118–31.

Edward Holmes, Associate Editor

Accepted May 1, 2006