J Mol Evol (2008) 67:539–550
DOI 10.1007/s00239-008-9171-8

Does Hybridization Increase Evolutionary Rate? Data from the 28S-rDNA D8 Domain in Echinoderms

Anne Chenuil · Emilie Egea · Caroline Rocher · Hélène Touzet · Jean-Pierre Féral

Received: 10 June 2008 / Accepted: 22 September 2008 / Published online: 24 October 2008

Abstract The divergent domain D8 of the large ribosomal RNA is very variable and extended in vertebrates compared to other eukaryotes. We provide data from 31 species of echinoderms and present the first comparative analysis of the D8 in nonvertebrate deuterostomes. In addition, we obtained 16S mitochondrial DNA sequences for the sea urchin taxa and analyzed single-strand conformation polymorphism (SSCP) of D8 in several populations within the species complex *Echinocardium cordatum*. A common secondary structure supported by compensatory substitutions and indels is inferred for echinoderms. Variation mostly arises at the tip of the longest stem (D8a), and the most variable taxa also display the longest and most stable D8. The most stable variants are the only ones displaying bulges in the terminal part of the stem, suggesting that selection, rather than maximizing stability of the D8 secondary structure, maintains it in a given range. Striking variation in D8 evolutionary rates was evidenced among sea urchins, by comparison with both 16S mitochondrial DNA and paleontological data. In *Echinocardium cordatum* and *Strongylocentrotus pallidus* and *S. droebachiensis*, belonging to very distant genera, the increase in D8 evolutionary rate is extreme. Their highly stable D8 secondary structures rule out the possibility of pseudogenes. These taxa are the only ones in which interspecific hybridization was reported. We discuss how evolutionary rates may be affected in nuclear relative to mitochondrial genes after hybridization, by selective or mutational processes such as gene silencing and concerted evolution.

Keywords rRNA secondary structure · Evolutionary rate · Interspecific hybridization · Concerted evolution · Polyploidy · Nucleolar dominance/selective silencing · Positive selection · Effective size

A. Chenuil (✉) · E. Egea · C. Rocher · J.-P. Féral
Laboratoire DIMAR, diversité, évolution, écologie fonctionnelle marine, CNRS UMR6540, Université Aix-Marseille II, Université de la Méditerranée, Centre d'Océanologie de Marseille, Station marine d'Endoume, rue de la batterie des Lions, 13007 Marseille, France
e-mail: chenuil@univmed.fr

H. Touzet
LIFL, CNRS UMR 8022, Université Lille 1, Villeneuve d'Ascq, France

Introduction

Eukaryotic large subunit ribosomal RNA is composed of a mosaic of regions displaying considerable variation in rate of evolution (Gerbi 1985; Gorski et al. 1987; Hassouna et al. 1984). The conserved core displays a highly conserved primary and secondary structure and very limited length variation across eukaryotes. Several divergent domains, in contrast, vary extensively in length and sequence between species. These domains were used for molecular phylogeny reconstruction at various taxonomic levels (Ben Ali et al. 1999; Hillis and Dixon 1991; Lenaers et al. 1988; Qu 1986). These studies provided the opportunity to assess the amount of conservation of the secondary structure across taxonomic levels and allowed prediction of unknown secondary structures by comparative methods (e.g., Chenuil et al. 1997).

The D8 domain is one of the most rapidly evolving divergent domains. It displays particularly large, lineage-specific extensions in length, although a global topology of secondary structure common to many eukaryotes (fungi, plants, nematodes, and vertebrates) was evidenced by

Michot and Bachellerie (1987). Its functional importance was experimentally established by Sweeney et al. (1994). Vertebrates, the only deuterostomes in which a comparative study of the D8 domain was made, exhibit much more extension and variation than other eukaryotes (Michot and Bachellerie 1987).

In this paper, we analyze the evolution and the secondary structure of the D8 domain of another deuterostome phylum, Echinodermata, composed of five extant classes, sea urchins, starfishes, brittlestars, feather stars, and sea cucumbers. Using comparative methods, we establish the secondary structure of the echinoderm D8 domain and identify the zones where variation is concentrated. To study the pace of D8 evolution in sea urchins, (i) we compared D8 divergence with divergence of an independent mitochondrial marker, and (ii) we used divergence dates estimated from a joint analysis of molecular data and fossil records. Variations in evolutionary rates are discussed in the light of our knowledge of phylogeography, population genetics, and genome evolution.

Materials and Methods

The *Echinocardium cordatum* Species Complex

E. cordatum is a common spatangoid sea urchin displaying an antitropical distribution and absent from American coasts. The presence of distinct clades within this nominal species was established by mitochondrial 16S DNA and nuclear D8 and EF1 sequences (Chenuil and Féral 2003). The major split divides an Atlantic monophyletic group, clade A, and a Pacific-Mediterranean clade, divided into a Pacific clade and a clade B, itself divided into two sympatric clades, B1 and B2, found in the Mediterranean sea. Clade B1 is also found in one Atlantic location, Galicia (North Spain), where it hybridizes with clade A, as revealed by two nuclear markers, D8 and the first intron of Elongation Factor 1 α (unpublished data).

Samples and Molecular Techniques

The accession numbers of sequences downloaded from or submitted to GenBank are listed in Table 2. Most sea urchins were collected by scuba diving in the Mediterranean Sea, the Atlantic Ocean, the North Sea, and the Pacific Ocean. The few others were obtained from museum collections (Paris Museum of National History, University of Hamburg).

DNA was extracted using the Chelex method as by Chenuil and Féral (2003) and 1–3 µl of this solution was used in 20-µl PCRs. The D8 fragment was amplified by PCR with primers D8-F and D8-R1, D8-R2, or D8-R3 (Table 1). PCR was performed on an MJ Research gradient thermocycler PTC200 using the program: 5 min at 95°C, then 40 cycles of 40 s at 94°C, 30 s at 58°C, and 2 min at 72°C. Sequencing was performed by use of primer D8-F for nearly all samples and, also, for some samples, with the opposite primer. For all spatangoid and most regular sea urchins, we also sequenced a portion of the 16S rDNA mitochondrial gene. For samples from the species complex *E. cordatum* (Chenuil and Féral 2003) 16S-mtDNA sequences allowed us to assign individuals to a mitochondrial clade. We proceeded as did Chenuil and Féral (2003) with primers 16SF2 and 16SR. For a few specimens (old museum specimens), the same 16S fragment had to be amplified by parts, i.e., two smaller overlapping fragments, using primer 16SF2 with 16SR3 and primer 16SR with 16SF4. In such cases, if the 16S DNA sequence was identical to already existing ones, we did not use the sequence subsequently, even though the identity was probably due to low divergence among related species, since small fragments are more prone to contamination. We always used PCR-negative controls without genomic DNA template.

Table 1 Primers used in the study

Primer	5′-3′ sequence
D8-F	AACTTCGGGATAAGGATTGGCTC
D8-R1	ATTAAACAGTCGGATTCCCCTTGTC
D8-R2	GGGATAAGGATTGGCTCTGAGGG
D8-R3	GTCCGTGCCAGTTCTGAGTCAGC
16SF2	GTTTCGCCTGTTTACCAAAAACAT
16SR	CGAACAGACCAACCCTTAAAAGC
16SR3	CGGAGGATTTTCTTTCTCCG
16SF4	YAMRCTTTGATTGGGGCAATC

To understand the pattern of variability of D8 in *Echinocardium cordatum*, and establish whether the variants found within an individual were alleles of one diploid segregating locus, we used the SSCP technique to genotype a high number of individuals. We proceeded as follows: 9% PAGE-glycerol gels (200 ml of gel was prepared adding 100 ml of 1 × TBE, 45 ml of acrymamide-bisacrymalide 40% [37:5], 10 ml of glycerol, and 45 ml of distilled water, with 1 ml of 10% ammonium persulfate and 200 µl of Temed added just before pouring) were run on a Hoeffer 660 electrophoresis apparatus using the double-gel sandwich system (four gels, 24 cm long, per run). We used 4 combs of 28 wells, loading the 24 central wells only, and the thinnest spacers available (0.75 mm). The system was linked to a cooling circuit (set at 2°C). Six liters of 0.5 × TBE were used as electrophoresis buffer for the inferior tank, in which polymerized gels were immerged at least 1 h before loading to homogenize temperature. We mixed 5 µl of each PCR product with 5 µl of formamide

blue, denatured at 96°C for 4–6 min, and immediately put the plate on ice. Wells were rinsed using a syringe before installing the upper tank. After loading samples, electrophoresis was started at a power of 15 W. During the first 2 h we noted the position of the dark and light blue relative to time to build the linear curve, allowing prediction of the time necessary for migration, and set up the timer to stop overnight migration (9 to 11 h). The light blue migrates exactly as D8 double-strand DNA fragments and must nearly reach the bottom of the gel; all other bands, corresponding to SSCP or heteroduplex conformation, migrate more slowly. After electrophoresis, gels were separated from the glass plates and placed on a tray using a Saran film (a narrow band of Whatman paper placed on top of the gel under the Saran wrap helped in recovering the gel from the glass plate). We prepared 100 ml of 0.5 × TBE with 6 µl of Gelstar dye and dispatched 25 ml on top of each gel in a tray, using a plastic bacteriology spreader. After 20 min, we rinsed the excess of Gelstar dye with tap water, then rinsed the colored gels in distilled water or 0.5 × TBE for 0–45 min to eliminate background. Photographs were taken using a Geldoc camera (BioRad) with identical settings for all gels. To facilitate the interpretation we run artificially heterozygous samples, mixing PCR products from individuals known to be homozygous for each type of sequence.

Data Analyses

Secondary Structure

For each sequence, the minimal free energy secondary structure of the D8 domain was determined with the RNAfold software (Hofacker 2003). We used the "partition function and pair probabilities" option for the fold algorithm and selected temperature parameters corresponding to mean ecological conditions for each species. We then built a multiple alignment taking account of both the primary and the secondary structures. We also evaluated the thermodynamical stability of the D8a helical regions identified by comparing their free energy level with that of randomized sequences following the protocol of Clote (2006). For each sequence, we generated 1000 shuffled sequences with the same dinucleotide frequency. We then folded each shuffled sequence with RNAfold. This allowed us to construct expected distributions of optimal free energy, one for each distinct sequence, and obtain their *Z*-score value. The *Z*-score is the standardized distance of the observed value to the mean of the inferred distribution (X0 − E)/s.

Evolutionary Rate

We compared DNA sequence variation among lineages and markers with regard to estimated divergence times between sea urchins. Only the D8 portion sequenced for all taxa was taken into account, removing 36 bp of the 5′ end, invariant at the family level. Three approaches were used. The first one was based on pairwise distances. The other two were based on a reliable topology of sea urchins and consisted in optimizing the branch lengths of 16S and D8 trees relative to the respective sets of sequence data (see below); the second approach used maximum likelihood and a mixed model dedicated to sequences forming secondary structure for the D8 data set; the third approach used the parsimony criterion for the D8 dataset.

Pairwise Distances

To measure divergence among D8 sequences, due to the limited sequence variation for most comparisons and the importance of secondary structure, we used the following simple rule: a substitution or a 1-base insertion or deletion counts as 1, whereas one pair of compensatory substitutions or indel counts for 1.5. Alternatively, we measured D8 genetic distance more conventionally (D8cor), as the proportion of nucleotidic differences after removal of indel positions in a variable region of the D8a helix (Fig. 3). We allocated an equal weight for each species (clade A and clade B of *E. cordatum* are considered two species here), thus species sharing identical D8 types were counted individually.

Divergence times were inferred from Lee (2003) for *Strongylocentrotus* species, from Metz et al. (1998) for *Arbacia*, and from Smith et al. (2006) for the other nodes. The latter study employed maximum likelihood and bayesian phylogenetic methods, using paleontological estimates of minimum or absolute divergence times, to infer divergence times of each node in a phylogeny of sea urchins, taking into account substitution rate variation among branches and using 28S rDNA and mitochondrial sequence data. To compare D8 and mitochondrial divergence for identical pairs of taxa, the Jukes and Cantor distance was inferred for 16S mt-DNA using the BioEdit software (Hall 1999). The 16S divergence was also estimated including numerous additional species (not sequenced for D8), but the results were similar and are not shown.

Tree-Based Methods

The phylogeny of sea urchins was reconstructed from the litterature (Stockley et al. 2005; Smith et al. 2006; Chenuil and Féral 2003; this study) (see Fig. 5) assuming starlike topology among variants within species. Branch lengths were optimized using the PHASE software (Hudelot et al. 2003). For D8 data, we accounted for secondary structures by using a mixed model including a substitution matrix for

pairs of sites involved in stems (we used the RNA7D model and a gamma distribution with three classes of evolutionary rates) and a substitution matrix for unpaired sites for loops (we used the HKY85 model with a gamma distribution and three classes). The model used for the 16S dataset was the HKY85 model (mixed models could not be used because secondary structure cannot be inferred for the partial 16S region sequenced). PHASE considers indels as ambiguities. Insertions and deletions, however, are prominent in our D8 data set. To analyze indel events together with nucleotide substitutions, we used a parsimony algorithm on the D8 dataset (DNApars [Felsenstein 1981]) to optimize branch lengths using the same tree topology. The parsimony analyses were also performed on two split data sets, one corresponding to the 5′ half of each stem and the loops, and the other corresponding to the 3′-half data of the stems, to check whether the nonindependence of sites involved in secondary structure biased the global analysis. Then D8 and 16S optimized branch lengths and their ratio were compared among taxa. We also used the DNApars program to infer the topology from D8 sequences to check whether the hypothesis used for the imposed topology, that *E. cordatum* D8 haplotypes formed a starlike topology, was supported. Then we also optimized branch lengths from the D8 data on the topology corrected according to the robust nodes obtained for the D8 parsimony tree.

Results

GenBank accession numbers are EF999826 to EF999854 for 16S and EF999855 to EF999883 for D8 sequences.

Primary Structure of the D8 Domain

The D8 haplotypes obtained for each species are reported in Table 2. For individuals sequenced by the forward primer only, we did not obtain the first 5′ 36 bases of the D8 sequence, which is extremely conserved. In every case, we obtained a portion of the conserved 28S ribosomal domain C9. Several species or even genera share the same D8 haplotype (*Arbacia lixula* and *A. punculata*, several genera within the Schizasteridae family). Conversely, some species display polymorphism or paralogues: we found two D8 haplotypes in the whole-genome sequence of *Strongylocentrotus purpuratus*. We do not know the exact structure of the rDNA cluster(s) since repetitive regions are difficult to resolve by whole-genome sequencing. Each clade of *Echinocardium cordatum* displayed two D8 haplotypes (at least). Sequence variation and size variation are mostly localized to a central region (Fig. 1), which corresponds to the tip of helix D8a (see below). Except for *Asterias forbesi*, whose D8 domain is 376 bp long, size variation is limited to a few base pairs, the length of the D8 domain ranges between 164 bp (for the crinoid *Florometra*, which has no nucleotide in this variable zone) and 183 bp in the ophiuroid *O. cinereum* (Fig. 2). In *E. cordatum*, haplotype IV appears to be ancestral to the other three haplotypes (see alignment in Fig. 3 and parsimony reconstruction in Fig. 5c); haplotype II′ may be interpreted as a product of recombination between haplotype II and haplotype III since the 5′ (respectively, 3′) strand of the stem is identical to haplotype II (respectively, III). In *Strongylocentrotus droebachiensis*, we observe a tandem duplication of the (GUCA) sequence and of the complementary motif (UGGC). All species display perfect (except for one pair of compensatory substitutions in *Arbacia*) disjoint repeats of the eight-nucleotide sequence "GGCGAG CC," which together form a part of stem D8a (Figs. 1 and 2). Perfect repeats of the four-nucleotide sequence "UCGG" are also observed in all species, with two adjacent repeats in the basal helix located at the branching of the three stems, and, in sea urchins, two additional disjoint but close repeats, in the 3′ side of the D8b and in the basal stem.

Secondary Structure of D8 in Echinoderms

All sequences show a common evolutionary conserved secondary structure obtained by RNAfold. Figure 1 shows the optimal structure obtained for *E. cordatum*. This is a Y-shaped structure with two arms, D8a and D8b. Stem D8a is significantly longer than stem D8b. It exhibits a series of bulges consisting of one or two unpaired nucleotides. The basal helix (in yellow) contains an asymmetrical internal loop. These results are consistent with previously published structures in other eukaryote species (Michot and Bachellerie 1987; Sweeney et al. 1994).

Further evidence suggests that these putative structures are valid. For helix D8a, 8 out of 21 haplotypes exhibit a UUCG tetraloop motif, an abundant structural element in ribosomal RNA (Uhlenbeck 1990; Woese et al. 1990). Although this region contains a lot of bulges, all sequences show a significant stability with respect to randomized sequences, with *Z*-scores ranging from 3.5 to 5.8. *E. cordatum* haplotypes II, II′, and III, which all appear to be derived relative to haplotype IV, display an increased stability relative to haplotype IV (Fig. 5c). Each of them also displays a distinct asymmetric bulge of one nucleotide, lowering the stability of stem D8a (Fig. 3). The stability of the putative reciprocal recombination product of II′, not found in our sample, would be even higher than that of the observed haplotypes (Fig. 3). Among the 20 haplotypes identified, these are the only cases of asymmetric bulges in the tip of D8a.

Regarding stem D8b, the stem is supported by several compensatory mutations that retain the potential for Watson-Crick pairings. In addition, the hairpin loop in sea urchins is

Table 2 Origin of sequence data and D8 haplotype names (number of specimens). Accessions provided by this study are in boldface

Species name (family)	16S accession no.	D8 accession no.	D8 haplotype
Abatus cavernosus (Schizasteridae)	**EF999841**	**EF999867**	Schi-II (3)
A. cordatus (Schizasteridae)	**EF999842**	**EF999866**	Schi-II
A. ingens (Schizasteridae)	**EF999838**	**EF999863**	Schi-I
A. nimrodi (Schizasteridae)	**EF999839**	**EF999862**	Schi-I (2)
A. shackletoni (Schizasteridae)	**EF999840**	**EF999864**	Schi-I
Amphipneustes rostratus (Schizasteridae)		**EF999865**	Schi-I[a]
A. similis (Schizasteridae)	**EF999836**	**EF999868**	Schi-II
Arbacia lixula (Arbacoidae)	**EF999826**	**EF999855**	A-lix
A. punctulata (Arbacoidae)	DQ073733	AY026367	A-lix
Asterias forbesi (Asteridae) AST		AF212169	
Brachisternaster chesheri (Schizasteridae)	**EF999837**	**EF999861**	Schi-I
Brissopsis lyrifera (Brissidae)	**EF9998232**	**EF999872**	B-lyr (5)
Echinocardium capense (Lovenidae)		**EF999876**	I (5)
E. cordatum (Lovenidae)	**EF999848–EF999854**	**EF999879–EF99983**	II(3), II(3)′, III(10), IV(33)[b]
E. flavescens (Lovenidae)	**EF999846**		
E. mediterraneum (Lovenidae)	**EF999847**	**EF999875**	I (5)
E. mortenseni (Lovenidae)		**EF999877**	I
E. pennatifidum (Lovenidae)	**EF999845**	**EF999878**	I (5)
Echinus esculentus (Echinidae)	**EF999829**	**EF999858**	E-esc
Florometra serratissima (Florometridae) CRI		AF212168	
Hypselaster limicolus (Schizasteridae)	**EF999835**	**EF999870**	Schi-II
Lovenia elongata (Lovenidae)	**EF999844**	**EF999874**	L-elo
Moira atropos (Schizasteridae)		**EF999871**	Schi-II[c]
Ophioderma cinereum (Ophiodermatidae) OP		AY859643	
Paracentrotus lividus (Echinidae)	**EF999828**	**EF999857**	P-liv
Schizaster canaliferus (Schizasteridae)	**EF999833**		
Spatangus purpureus (Spatangidae)	**EF999843**	**EF999873**	I′ (4)[d]
Sphaerechinus granularis (Toxopneustidae)	**EF999830– EF999831**	**EF999856**	S-gra
Strongylocentrotus pallidus (Strongylocentrotidae)	**EF999827**–AM900392	**EF999860**	S-pal
S. droebachiensis (Strongylocentrotidae)	AM900391 & AY652514	**EF999859**	S-dro
S. purpuratus (Strongylocentrotidae)[e]	X12631	AF212171	S-pur, S-pur’
Tripylus excavatus (Schizasteridae)	**EF999834**	**EF999869**	Schi-II

[a] Missing 3′ end (this region is invariant within the family Schizasteridae)

[b] Many individuals were sequenced for *E. cordatum,* most of which were heterozygous. Here only homozygous unambiguous sequences are considered. D8 haplotypes found in mitochondrial clade A are II (3 individuals) and III (10 individuals); in clade B, II′ (3 individuals) and IV (23 individuals); and in the Pacific clade, IV (10 individuals). Less frequent haplotypes, II and II′, coexisted with the common types III and IV, respectively, in numerous heterozygotes

[c] Ambiguous nucleotides at 5′ diagnostic indel site and 3′ missing end

[d] Only one nucleotide, located in the 5′-end portion, which is not available for all taxa, differs from D8 haplotype I

[e] See also the whole-genome sequence for the two variants: http://www.hgsc.bcm.tmc.edu/projects/seaurchin

GCAA, which belongs to the family of GNRA tetraloops (where N is any nucleotide and R is a purine). In other species, the hairpin loop is a variation of the UUCG motif.

Evolutionary Rate of D8 in Sea Urchins

Figure 4a illustrates for each marker the relationship between the amount of molecular divergence accumulated and the divergence time (Table 3). Globally 16S distances and our crude D8 distance estimates are well correlated, with the notable exceptions of the distances among *Echinocardium cordatum* clades and among *Strongylocentrotus* species. In both cases a striking increase in D8 divergence relative to 16S-mt divergence is observed (Fig. 4b). The low molecular divergence for the split between Schizasteridae and other Spatangoida (95 mya), affecting both

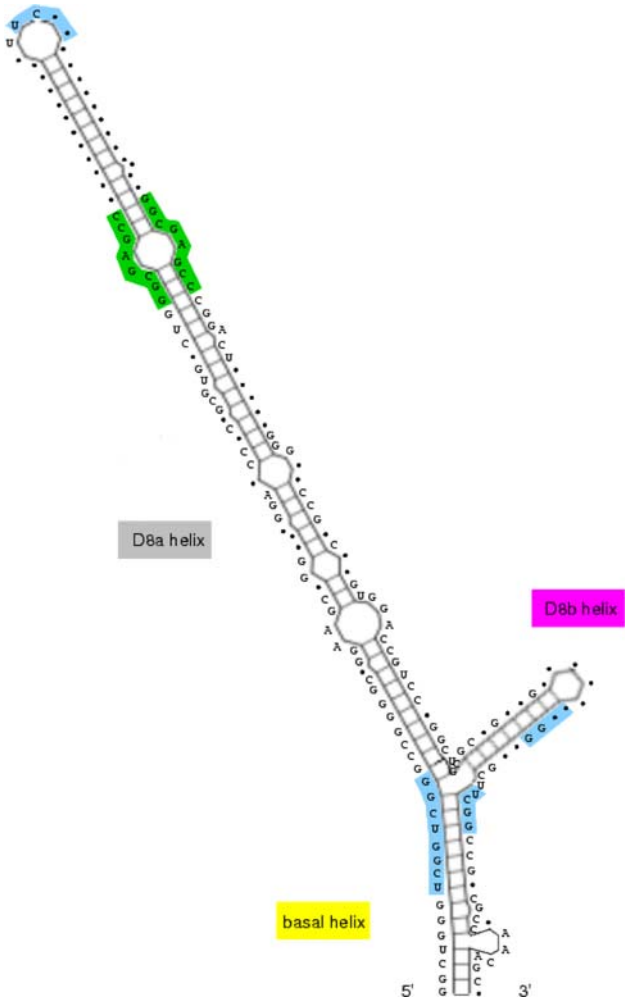

Fig. 1 Folding of the D8 domain in *Echinocardium cordatum* obtained with RNAfold. We report the consensus primary structure on the folding. We consider that a nucleotidic site is conserved as long as it is present in more than 80% of the complete echinoderm sequences (five of six sequences). Varying sites are represented by a dot. Repeated elements "GGCGAGCC" and "UCGG" reported under Results are highlighted in green and blue, respectively

nuclear and mitochondrial DNA, is not due to an error in paleontological estimates but is explained by the substitution rate, which, in the Schizasteridae family, is half the average value in the Spatangoida order, as shown by relative rate tests (Chenuil et al. 2008). The 16S divergence appears to be increased in *Strongylocentrotus* and, though less markedly, in *Echinocardium cordatum,* compared to *Arbacia*, for similar divergence times (Table 3), and in the general curve (Fig. 4a). This was already known for *Strongylocentrotus* (Biermann et al. 2003) and appears to be due to *S. purpuratus*, not *S. pallidus and S. droebachiensis*, in branch length optimized trees (Fig. 5b). The increase in D8 relative to 16S divergence is, nevertheless, very salient in *S. pallidus, S. droebachiensis*, and *E.cordatum* (Fig. 5a, b). The ratio of D8 to 16S-mtDNA divergence appears to be positively correlated with D8 length, even when data from *E. cordatum* (clade A vs. clade B) and *Strongylocentrotus* (*S. purpuratus* vs. *S. droebachiensis*) are not taken into account (the linear correlation coefficient r^2 is 0.693), as expected due to (i) our crude way of measuring D8 divergence and (ii) the fact that longer stem regions have more possibility to vary while maintaining a given threshold of stability (Fig. 4b). However, the D8/16S-mt divergence ratios are disproportionately high in these two groups, establishing that the acceleration of evolution in the D8 is not an artifact and deserves attention (Figs. 4b and 5).

The approach based on maximum likelihood and secondary structure models (not parsimony), theoretically the best modeled, dismissed the major part of the information since variation among sea urchins consists almost exclusively in insertions or deletions (Figs. 1 and 2), which are not considered by these programs. Nevertheless, the ratio of optimized branch lengths D8/16S provides a pattern similar to that of the pairwise distances within spatangoids: *Echinocardium cordatum* ancestral and terminal branches display the maximum D8/16S ratio among Spatangoida sea urchins, and the branches leading to *Strongylocentrotus droebachiensis* and *S. pallidus* display the highest ratios among regular sea urchins (data not shown; NB: the *S. pallidus* 16S branch length being null, the ratio can be considered infinite). We emphasize that, the amount of information being rather limited in D8 without indels, the longest branch among *S. pallidus* and *S. droebachiensis* varied according to the chosen parameters of the model of evolution.

When D8 branch lengths were optimized using parsimony, both split data sets (18 parsimony informative sites in the 5′ strands of stems and the loop of D8a, 16 parsimony informative sites for the 3′ strands) yielded very similar trees, also very similar to the global data set, so we present the global data set (34 parsimony informative sites) (Fig. 5a). The D8 branch length ratios are very high in *Strongylocentrotus droebachiensis*, *S. pallidus* (infinite), and *E. cordatum* (data not shown; they can be deduced by comparing Fig. 5a vs. b). An infinite *S. pallidus D8*/16S branch length ratio is due to the null 16S branch length obtained by the optimization process on the imposed topology. These results appear more likely to be caused by an increase in the D8 rate (Fig. 5a) in these taxa than by a decrease in the 16S rate (Fig. 5b). Branch lengths appear to be highly variable among some closely related species for D8 but not for 16S, except in *Strongylocentrotus*, where *S. purpuratus* has a much longer 16S branch than *S. pallidus* and *S. droebachiensis*. But in this case, an independent study of 16S sequences in 10 strongylocentrotid species identified the outlier as being *S. purpuratus* for highly increased branch lengths (Biermann et al 2003). The two

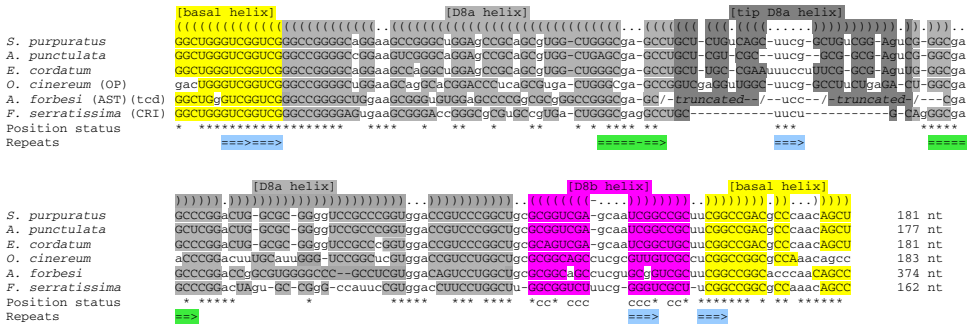

Fig. 2 Alignment and secondary structure of the total D8 domain in echinoderms. Capital letters represent pairing bases and are overlined. The basal helix is overlined in yellow; helix D8a, in gray; and helix D8b, in pink. Repeats are represented by arrows of the same colors as in Fig. 1. In the position status row, conserved positions are indicated by an asterisk, compensatory substitutions in stem D8b are indicated by a "c," and in zones of uncertain alignment (homology) no indications are given. The alignment of the variable tip of helix D7a in 17 sea urchin D8 haplotypes is presented in Fig. 3. Outside the tip of D8a (in dark gray), all other sea urchins (not fully sequenced, thus not shown here) have nearly identical sequences. All regular sea urchins have a 'G' in position 32 (helix D8a), whereas spatangoids have an 'A.' *Arbacia punctulata* is identical to *A. lixula* and relatively divergent in the 5′ end of D8. All other sea urchins are very similar to *Strongylocentrotus purpuratus:* the only differences are in position 24 (helix D8a), where Arbacoidae, Schizasteridae, Brissidae, and Spatangidae have a 'C,' whereas other regular sea urchins and Lovenidae have an 'A.' *Paracentrotus lividus* has the same compensatory substitution in D8b as *E. cordatum,* and *Lovenia elongata* has a 'U' instead of a 'C' maintaining pairing with a 'G,' also in D8b

Fig. 3 Alignment and secondary structure of the tip of D8a in sea urchins. Symbols for category of sites: (*) invariant site pairing in the stem; (c) site of compensatory substitution or indel; (l) site of the loop; (−) aligment gap inserted by us in all species to preserve symmetry with respect to a nonpaired nucleotide insertion (occurring in the opposite strand of the stem); (i) noncompensated insertion forming a bulge. In the consensus row, conserved nucleotides are in boldface when they are invariant in normal cases when they appear in the majority of genera. Putative, not observed, haplotype II″ (reciprocal recombination II × III product of II′) is also represented in gray characters. Free energy of the corresponding helix tip computed using RNAfold is indicated (*Z*-scores range between 4 and 6)

species of Echinidae, *P. lividus* and *E. esculentus*, that also display long terminal branches for D8 (Fig. 5a) present the same pattern for 16S (Fig. 5b), thus there is no marker-specific increase in evolutionary rate.

When a parsimony phylogenetic reconstruction was performed on the D8 data set without imposing a defined topology (Fig. 5c), haplotype IV of *E. cordatum* clearly appeared to be ancestral to the other three haplotypes (the topology of this subtree is supported by very high bootstrap values, contrary to the rest of the tree). Therefore it is likely that Ecord-IV was the ancestral haplotype. When this topology was used instead of the one used in Fig. 5a, b, and the 16S branch lengths adequately corrected, the optimized branch length ratios D8/16S were very similar (not shown), except that the Ecord-IV terminal branch length was shorter. The ancestor to haplotypes Ecord-II, -II′, and -III also had an elongated branch, but less than the terminal branch leading to Ecord-III. Therefore the global pattern remains an increase in the branch length D8/16S ratios in *E.cordatum*, more pronounced in some haplotypes than others, and less pronounced in the common stem.

Table 4 displays the genotypes inferred from SSCP patterns in *E. cordatum*. These frequencies are not compatible with the hypothesis of a single mendelian biallelic

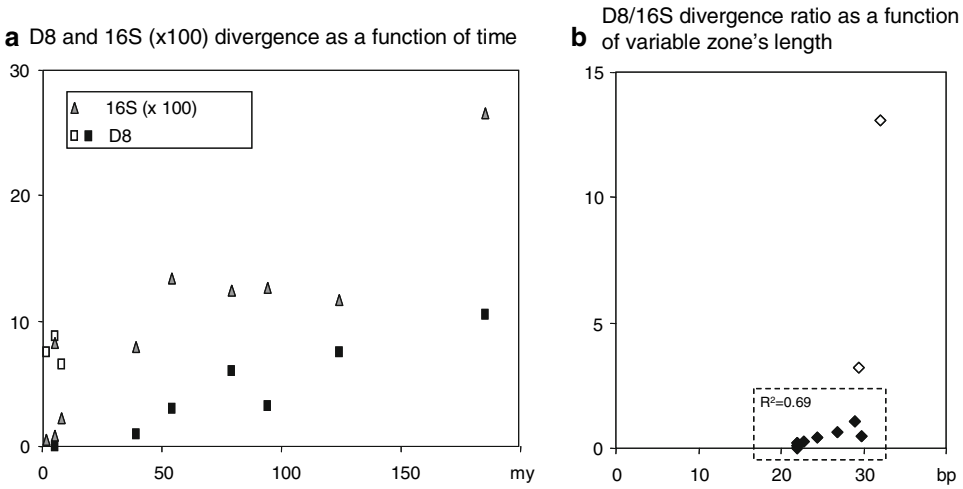

Fig. 4 Evolutionary rate of D8 and 16S mt-DNA in spatangoid and regular sea urchins. Open symbols are for pairs of species for which D8 divergence is higher than 16S divergence. They correspond to *E. cordatum* and *Strongylocentrotus*. (**a**) Divergence of D8 (rectangles) and 16S-mtDNA (triangles) sequences as a function of divergence time (mya). (**b**) Ratio of D8/16S divergence (divided by 100) as a function of mean length of the tip of D8a (within group). This figure allows us to check that the higher D8 divergence (relative to 16S) among the outlier taxa is not due to an increased length of the D8a helix

Table 3 Divergence times and molecular divergence in 16S-mt and D8-rDNA[a]

Divergence node	MY	16S	D8	D8cor[b]
E. cordatum within clade A (II/III)	0	NR	5	2/22
E. cordatum within clade B (II′/IV)	0	NR	9.5	1/22
E. cordatum				
A/B (II/IV)	9	NR	6.5	1/22
A/B (II/II′)	9	NR	4	0/22
A/B (III/IV)	9	NR	9.5	2/22
A/B (II′/III)	9	NR	6	2/22
A/B (average)	9	0.018	**6.5**	2/22
Spatangidae[c]/Lovenidae (average)	40	0.09	1	0.36/22 (0–1/22)
Brissidae/(Lovenidae + Spatangidae)	55	0.136	3	1.2/22 (1–2/22)
Schizasteridae/other Spatangoids	95	0.121	3.25	1.33/22 (2–3/22)
S. pallidus/S. droebachiensis	2.5 (2–3)	ND	**7.5**	2/24
S. purpuratus/ (*S. droebachiensis* + *S. pallidus*)	6 (5–7)	0.066	**8.75**	2/24
A. lixula/A. punctulata	6 (5–7)	0.007	0	0
Strongyloncentrotidae/Echinidae	80 (61–100)	0.146	6	2/22
S. granularis/(Echinidae + Strongyloncentrotidae)	125 (111–139)	0.129	7.5	4/22
Arbacia sp./other regular sea urchins	186 (171–202)	0.248	10.5	6/22

Note: MY, million years; ND, not determined

[a] The values used in the analyses in Fig. 4 are given in columns 2–4. When values correspond to averages, they are followed by the range in parentheses (minimum–maximum) for column 5. NR (not relevant): since mitochondrial DNA and D8 segregate independently it is not relevant to compute 16S-mt DNA distances among D8 types within a biological species. The values displaying disproportionately high D8 divergence relative to time and to 16S divergence (see Figs. 4 and 5) are in boldface

[b] Divergence in the D8a variable zone after removing indel positions in spatangoids and regular sea urchins separately (Fig. 2), so that we retained 22 nucleotides for spatangoids and 24 for regular sea urchins: number of differences/number of compared nucleotides

[c] The D8 sequence for the Spatangidae, *Spatangus purpureus*, is identical to that for all *Echinocardium* species except *E. cordatum* in the portion of the alignment analyzed for all taxa, thus the molecular distances equal 0 (cf. minimum range); other values refer to comparisons with *Lovenia*

locus in each clade: this hypothesis would imply a large excess in heterozygote frequency, a phenomenon for which we do not see biologically plausible explanations. Variability within individuals may also be a consequence of the organization of ribosomal DNA in clusters of more or less homogeneous repeats (see Discussion).

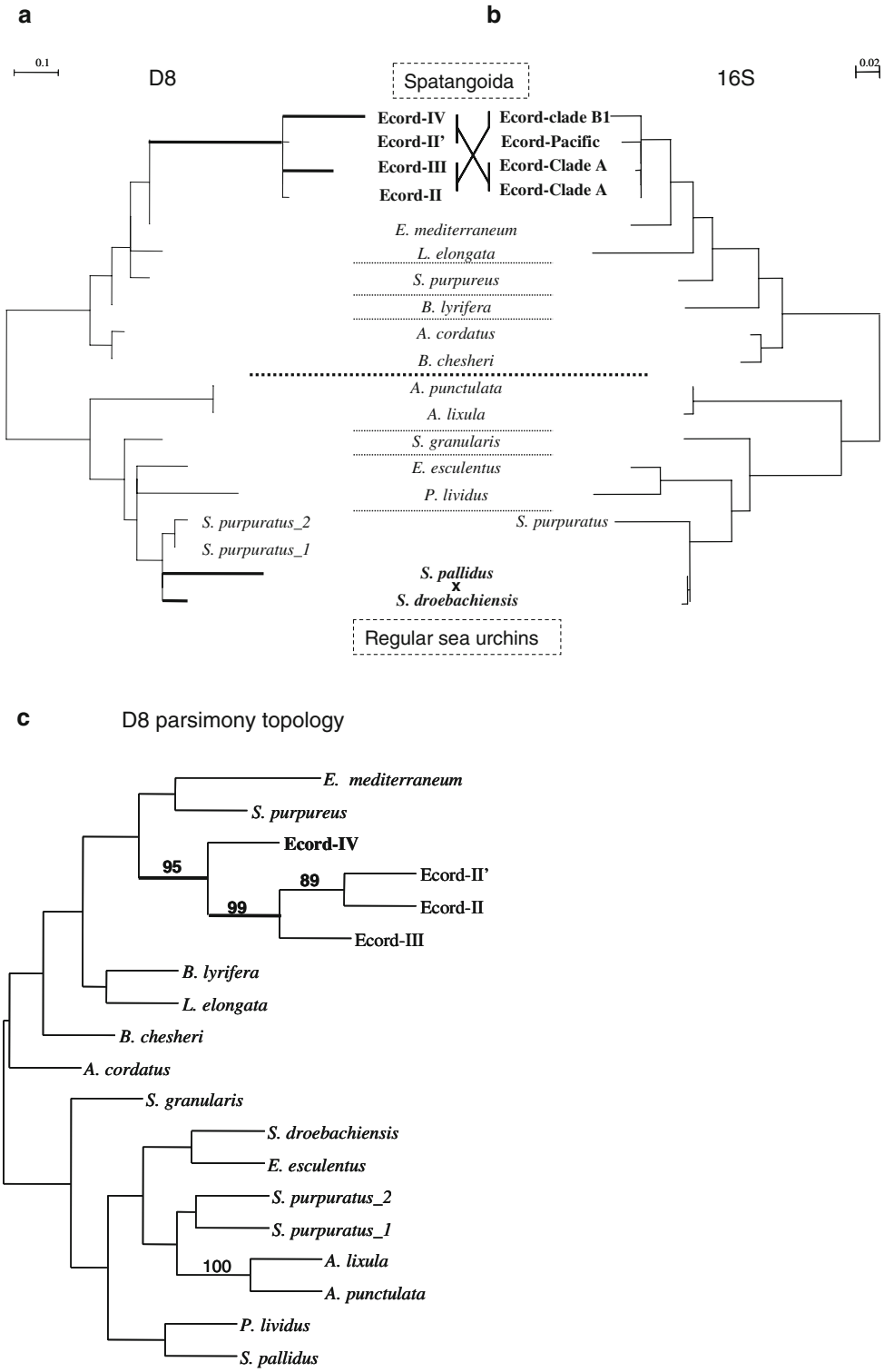

Fig. 5 Branch lengths and phylogenetic trees for D8 and 16S-mt sequence data among sea urchins. Branch lengths are optimized on an imposed topology where intraspecific variants are considered *a priori* as polyfurcations (**a**, **b**), using the parsimony criterion for D8 (to keep information from indels) and using maximum likelihood for 16S-mt. Hybridizing taxa are in boldface and marked with an X. "Ecord-XX" refers to haplotypes from *Echinocardium cordatum*. (**a**) D8 branch lengths. (**b**) 16S branch lengths. The thick dotted line separates the regular sea urchins from the irregular sea urchins, order Spatangoida; the thin dotted lines separate the distinct families. (**c**) Phylogenetic tree and bootstrap values (%) obtained from the D8 data set using parsimony and without imposing the topology. Among regular sea urchins, relationships are not robust and valid clades are not found. Notice that D8 haplotypes in *E. cordatum* appear to branch successively (no polyfurcation), with haplotype IV appearing to be ancestral with high bootstrap support (99%). The possibility that haplotype II′ is a recombinant from hapotypes II and III suggests that relationships among these three haplotypes are not tree-like

Discussion

The General Mode of Evolution of the D8 Domain

The folding of the D8 domain in echinoderms is strongly supported by energy values, compensated substitutions, and compensated indels. It follows the common eukaryotic model. Large extensions as observed in vertebrates may occur in echinoderms, as illustrated by the D8 domain of starfish *Asterias forbesi* (376 bp). However, in the other classes, the stems are about the same size as in the non-vertebrate metazoan studied by Michot and Bachellerie

Table 4 Number of individuals per D8-SSCP genotype in each clade in *Echinocardium cordatum* (summed over several populations)

	SSCP (II)	SSCP (III)	SSCP (II–III)
Clade A	5	27	53
	SSCP (II′)	SSCP (IV)	SSCP (II′–IV)
Clade B	6	89	211

(1987), *Caenorhabditis elegans*. The second stem (D8b) is extremely conserved in size (6–8 bp) in all echinoderms, including starfish. In *Florometra serratissima* the variable zone corresponding to the tip of D8a is absent, suggesting that the D8 domain may be extremely conserved in crinoids. The brittle star has a 185-bp-long D8, slightly longer than the longest sea urchin D8 (183 bp), thus we expect variation to occur within the class.

We identified how novel haplotypes may arise by mutation, with a putative recombination event in *E. cordatum*, and a possible duplication in *S. droebachiensis* on each side of the stem. Insertion-deletion mutational events may involve more than one base on each side of the stem: this is suggested by the alignment (Fig. 3), where most closely related haplotypes even within a species or a genus may differ by several contiguous indel positions (*E. cordatum, Strongylocentrotus*). There may be a minimum level of stability required for a mutant haplotype to subsist until a compensatory mutation restoring pairing in the stem arises, thus we expect that longest haplotypes evolve more rapidly. Though our first measure of D8 divergence overestimates divergence for taxa with the longest D8 regions, an observation not affected by this artifact supports this hypothesis. Haplotype IV of *E. cordatum,* ancestral to the other three haplotypes, which subsequently arose very rapidly, has an increased stability relative to haplotype I, from which it derived, and which is found in all other *Echinocardium* species (Figs. 4a and 5c).

In the absence of selection against helixes that are too stable or too long, one would expect that positive retroaction leads to an accelerated increase in stem length. Short stems tend to evolve very slowly because mutations generally are eliminated rapidly due to loss of stability, before the compensatory mutation occurs and stabilizes the structure. However, our data, though not conclusive, suggest the presence of selection against overly stable helixes in echinoderms: in *E. cordatum* bulges are reported in the stems of the most stable haplotypes, and the putative reciprocal recombination product of type II′, more stable than all other *E. cordatum* haplotypes, was never observed, as if it had been eliminated by natural selection.

Putative Causes of Lineage-Specific Accelerated D8 Evolution

We found that when the two outlying taxa are removed, the molecular divergence for D8 and for mitochondrial 16S are globally well correlated (Fig. 4b), and variation in evolutionary rates among lineages (Figs. 4a and 5) affects both molecules in a similar way. These new data from the D8 region therefore confirm the trend observed for mitochondrial DNA (Chenuil et al. 2008), that is, a slower molecular evolution in the Schizasteridae family. However, in three taxa, *Echinocardium cordatum, Strongylocentrotus pallidus*, and *S. droebachiensis,* there is a marked increase in D8 evolutionary rate compared to mitochondrial 16S DNA. A common characteristic of these taxa is the presence of interspecific hybridization: *S. droebachiensis* and *S. pallidus* are known to hybridize (Addison and Hart 2005; Biermann et al. 2004; Vasseur 1952), as well as clades A and B in the *E. cordatum* species complex (Chenuil and Féral 2003; Egea et al., unpublished). It is not rigorously proven that all the other species do not hybridize, but phylogeographic studies or preliminary data are available for most of them and do not reveal such a phenomenon. In the order Spatangoida, we sequenced mitochondrial DNA for several populations of *B. lyrifera, S. purpureus, E. flavescens*, and *E. capense*, from both the Atlantic and the Mediterranean, and *E. mediterraneum* from several Mediterranean populations (unpublished); contrary to *E. cordatum*, no split occurs in any of these haplotype networks or trees, which suggests the presence of a species complex or the possibility of hybridization. In *Arbacia spp.* (Metz et al 1998) mitochondrial sequences and analyzes of the bindin sequence evolution rule out the possibility of hybridization events in the recent history of the genus. Phylogeographic analyses are also available for the commercially harvested sea urchin *Paracentrotus lividus* (Duran et al 2004, Iuri et al 2007).

Since the acceleration is visible in ribosomal DNA, but not, or much less, in mitochondrial DNA, we envisaged two types of processes: a selective and two mutational ones. (1) In the context of a hybrid (backcrossed) genome where genetic cohesion is suboptimal and has to be restored (Kondrashov et al. 2002), mutations are expected to appear more frequently than usual as selectively favorable and less frequently as deleterious or slightly deleterious. Since substitution rates of favorable mutations are proportional to the effective size, whereas substitution rates of slightly deleterious mutations are inversely related to the effective size (Ohta 1992), hybridization may lead to an increase in evolutionary rates more pronounced in nuclear than in mitochondrial DNA. This is the pattern observed in this study. Under this hypothesis one expects that cases of asymmetrical hybridization lead to rate

increase limited to one of the two hybridizing clades. The introgression among *S. pallidus* and *S. droebachiensis* is supposed to be assymmetrical (though nuclear genetic data are lacking to confirm this), involving sperm from *S. droebachiensis* only. This should lead to a relative D8/16S rate increase in *S. droebachiensis* but not in *S. pallidus*; however, both lineages, but mostly *S. pallidus*, appear to be affected. Clearly, quantitative predictions are not straightforward; theoretical models on the evolution of substitution rates after a hybridization event, contrary to the case of genome duplication, are lacking.

(2) Hybridization may lead to selective silencing (or nucleolar dominance) of the ribosomal gene cluster (Pikaard 2000): in numerous studies involving experimental or natural hybrids, the ribosomal cluster from a single parental species is transcribed in the hybrid cell. This leads, in subsequent generations, to relaxation of selective constraints on the silenced rDNA clusters. New variants can then arise and rapidly lose their function, producing a tremendous variability of pseudogenes (e.g., Muir et al. 2001; Marquez et al. 2003). We suggest that these variants may, by recombination with functional copies, increase the mutation rate, and, as a consequence, the substitution rate, of the selectively constrained rDNA copies. Kovarik et al. (2008) suggested that silenced rDNA genes may be less efficiently homogenized by concerted evolution (see below) due to their association with condensed chromatin and heterochromatin. Despite the fact that our analysis of evolutionary rate considered only functional copies, with very stable secondary structures, we occasionally amplified D8 fragments of smaller size, in addition to the larger fragment, in two individuals of *E. flavescens*, with easily identifiable deletions leading to an absence of secondary structure (data not shown), supporting the existence of pseudogenes.

(3) In multicellular organisms, ribosomal DNAs are generally organized in one or several clusters of tandem repeats whose number vary between taxa. The action of local mutations generates divergence within and among clusters, but concerted evolution, mainly by the way of unequal crossing-over and gene conversion, efficiently counteracts mutation, homogenizing the whole repetitive family (Dover 1989). In a recombined individual issued from a hybrid, if the two main parental repeat types are different, the efficiency of concerted evolution is reduced (e.g., unequal crossing-overs are less frequent between divergent copies), increasing heterogeneity within and among individuals. This has the same effect on evolutionary rates as an increase in effective size: substitution rates of slightly deleterious mutations decrease, while substitution rates of selectively advantageous mutations increase. The observed pattern (rate increase after hybridization) would correspond to a case where favorable mutations are more frequent than slightly deleterious ones. This seems unlikely (Ohta 1992; but see Bazin et al. 2006), in particular, for a highly constrained molecule such as rDNA.

More evidence is needed to check the association between hybridization and increased nuclear DNA evolutionary rate, as well as the trend toward a (slighter) acceleration in mitochondrial evolutionary rates. Data from nuclear DNA regions not involved in multigenic families (ongoing work on EPIC [exon primed intron crossing] loci) may help to differentiate selective and mutational explanations.

Acknowledgments The European Network of Excellence, Marine Genomics Europe, provided funds and access to sequencing facility. The DIMAR laboratory also belonged to the European Network of Excellence, MARBEF (marine biodiversity and ecosystem functioning), which provided funds for meetings and travels. Subject editor Nicolas Galtier and two anonymous reviewers greatly helped to improve the manuscript.

References

Addison JA, Hart MW (2005) Colonization, dispersal, and hybridization influence phylogeography of North Atlantic sea urchins (*Strongylocentrotus droebachiensis*). Evolution 59:532–543

Bazin E, Glémin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. Science 312:570–572

Ben Ali A, Wuyts J, De Wachter R, Meyer A, Van de Peer Y (1999) Construction of a variability map for eukaryotic large subunit ribosomal RNA. Nucleic Acids Res 27:2825–2831

Biermann CH, Kessing BD, Palumbi SR (2003) Phylogeny and development of marine model species: strongylocentrotid sea urchins. Evol Dev 5:360–371

Biermann CH, Marks JA, Vilela-Silva AC, Castro MO, Mourao PA (2004) Carbohydrate-based species recognition in sea urchin fertilization: another avenue for speciation? Evol Dev 6:353–361

Chenuil A, Féral J-P (2003) Sequences of mitochondrial DNA suggest that *Echinocardium cordatum* is a complex of several sympatric or hybridizing species. A pilot study. In: Féral J-P, David B (eds) Echinoderm research 2001. Proc 6th Eur Conf Echinoderm. Banyuls-sur-mer, France. Swets & Zeitlinger, Lisse, Netherlands, pp 15–21

Chenuil A, Solignac M, Michot B (1997) Evolution of the large-subunit ribosomal RNA binding site for protein L23/25. Mol Biol Evol 14:578–588

Chenuil A, Egea E, Rocher C, Féral J-P (2008) Comparing substitution rates in spatangoid sea urchins with putatively different effective sizes, and other echinoderm data sets. In: Harris LG (ed) 12th international Echinoderm conference. Balkema, Durham, NH

Clote P (2006) Combinatorics of saturated secondary structures of RNA. J Comput Biol 13:1640–1657

Dover GA (1989) Linkage disequilibrium and molecular drive in the rDNA gene family. Genetics 122:249

Duran S, Palacin C, Becerro MA, Turon X, Giribet G (2004) Genetic diversity and population structure of the commercially harvested sea urchin *Paracentrotus lividus* (Echinodermata, Echinoidea). Mol Ecol 13:3317–3328

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376

Gerbi SA (1985) Evolution of ribosomal DNA. In: McIntyre RJ (ed) Molecular evolutionary genetics. Plenum, New York, pp 419–517

Gorski JL, Gonzalez IL, Schmickel RD (1987) The secondary structure of human 28S rRNA: the structure and evolution of a mosaic rRNA gene. J Mol Evol 24:236–251

Kovarik A, Dadejova M, Lim YK, Chase MW, Clarkson JJ, Knapp S, Leitch AR (2008) evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann Bot 101:815–823

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hassouna N, Michot B, Bachellerie J-P (1984) The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in eukaryotes. Nucleic Acids Res 12:3563–3583

Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. Q Rev Biol 66:411–427

Hofacker IV (2003) Vienna RNA secondary structure server. Nucleic Acids Res 31:3429–3431

Hudelot C, Gowri-Shankar V, Jow H, Rattray M, Higgs P (2003) RNA-based phylogenetic methods: application to mammalian mitochondrial RNA sequences. Mol Phylogenet Evol 28:241–252

Iuri V, Patti FP, Procaccini G (2007) Phylogeography of the sea urchin *Paracentrotus lividus* (Lamarck) (Echinodermata:Echinoidea): first insights from the South Tyrrhenian Sea. Hydrobiologia 580:77–84

Kondrashov AS, Sunyaev S, Kondrashov FA (2002) Dobzhansky-Muller incompatibilities in protein evolution. Proc Natl Acad Sci USA 99:14878–14883

Lee Y-H (2003) Molecular phylogenies and divergence times of sea urchin species of Strongylocentrotidae, Echinoida. Mol Biol Evol 20:1211–1221

Lenaers G, Nielsen H, Engberg J, Herzog M (1988) The secondary structure of large-subunit rRNA divergent domains, a marker for protist evolution. Biosystems 21:215–22

Marquez LM, Miller DJ, MacKenzie JB, van Oppen MJH (2003) Pseudogenes contribute to the extreme diversity of nuclear ribosomal DNA in the hard coral *Acropora*. Mol Biol Evol 20:1077–1086

Metz EC, Gomez-Gutierrez G, Vacquier VD (1998) Mitochondrial DNA and bindin gene sequence evolution among allopatric species of the sea urchin genus *Arbacia*. Mol Biol Evol 15: 185–195

Michot B, Bachellerie J-P (1987) Comparisons of large subunit rRNAs reveal some eukaryote-specific elements of secondary structure. Biochimie 69:11–23

Muir G, Fleming CC, Schlotterer C (2001) Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. Mol Biol Evol 18:112–119

Ohta T (1992) The nearly neutral theory of molecular evolution. Annu Rev Ecol Syst 23:263–286

Pikaard CS (2000) Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids. Plant Mol Biol 43:163–177

Qu LH (1986) Structuration et évolution de l'ARN ribosomique 28S chez les eucaryotes. Etude systématique de la région 5′ terminale. Ph.D. thesis, Université Paul Sabatier de Toulouse

Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood DTL (2006) Testing the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea (Echinodermata). Mol Biol Evol 23:1832–1851

Stockley B, Smith AB, Littlewood T, Lessios HA, Mackenzie-Dodds JA (2005) Phylogenetic relationships of spatangoid sea urchins (Echinoidea): taxon sampling density and congruence between morphological and molecular estimates. Zool Scr 34:447–468

Sweeney R, Chen LH, Yao MC (1994) An ribosomal-RNA variable region has an evolutionarily conserved essential role despite sequence divergence. Mol Cell Biol 14:4203–4215

Uhlenbeck AC (1990) Tetraloops and RNA folding. Nature 346: 613–614

Vasseur E (1952) Geographic variation in the Norwegian sea-urchins, *Strongylocentrotus droebachiensis* and *S. pallidus*. Evolution 6:87–100

Woese CR, Winker S, Gutell RR (1990) Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". Proc Natl Acad Sci USA 87:8467–8471