

SHORT COMMUNICATION

The 'rare allele phenomenon' in a ribosomal spacer

MENNO SCHILTHUIZEN,*† ROLF F. HOEKSTRA† and EDMUND GITTEBERGER*

*Institute for Evolutionary and Ecological Sciences, Leiden University, Leiden, The Netherlands, †Laboratory of Genetics, Wageningen University, Wageningen, The Netherlands

Abstract

We describe the increased frequency of a particular length variant of the internal transcribed spacer 1 (ITS-1) of the ribosomal DNA in a hybrid zone of the land snail *Albinaria hippolyti*. The phenomenon that normally rare alleles or other markers can increase in frequency in the centre of hybrid zones is not new. Under the term 'hybrizyme' or 'rare allele' phenomenon it has been recorded in many organisms and different genetic markers. However, this is the first time that it has been found in a multicopy locus. On the one hand, the pattern fits well with the view that purifying selection in hybrid populations works on many loci across the genome and should thus have its effect on many independent molecular markers. On the other hand, the results are puzzling, given that the multiple copies of rDNA are not expected to respond in unison. We suggest two possible explanations for these conflicting observations.

Keywords: *Albinaria hippolyti*, biased gene conversion, Gastropoda, hybrid zone, ITS-1, length polymorphisms, rDNA

Received 21 August 2000; revision received 8 December 2000; accepted 14 December 2000

Introduction

Hybrid zones are narrow regions where genetically distinct populations meet, mate, and produce hybrids (Barton & Hewitt 1985). They have been successfully exploited as 'natural laboratories' (Hewitt 1988) to study the genetics of species differences and reproductive isolation. In addition, hybrid zones have often been viewed as regions where evolutionary novelties arise. In plants, new species are known to have been produced from certain recombinant genotypes (Rieseberg *et al.* 1996), while in land snail hybrid zones, new colour morphs (Chiba 1997) and shell shapes (Goodfriend & Gould 1996) appear.

The examples just given apply to new and potentially advantageous combinations of already present alleles. However, many authors have asserted that entirely novel alleles might also evolve in the context of hybridization. A well known category of such allegedly newly formed alleles are the so-called 'hybrizymes' (Woodruff 1989). These are allozymes that are found in hybrid zones at high

frequencies, but which are rare or absent in the 'pure' parental populations. The phenomenon (also termed the 'rare allele phenomenon'; Sage & Selander 1979) is surprisingly common. Barton & Hewitt (1985) discussed 23 studies that included thorough allozyme data, 19 of which showed hybrizymes, and several additional examples have been published since (Keenan 1994; Schilthuizen 1995; Guiller *et al.* 1996).

The cause for the phenomenon was originally put down to intragenic recombination (Golding & Strobeck 1983; Woodruff 1989). However, sequencing studies (Bradley *et al.* 1993; Hoffman & Brown 1995) have shown that hybrizyme alleles are caused by simple point mutations, and can not be explained by either recombination or transposition. With the mutational origin of hybrizymes revealed, the cause for their maintenance was still unclear. Two hypotheses remained viable: increased rates of nucleotide substitution (Thompson & Woodruff 1978; Barton *et al.* 1983; Barton & Hewitt 1985; Woodruff 1989; Schilthuizen & Gittenberger 1994) or positive selection on otherwise slightly deleterious alleles (Barton & Hewitt 1985) acting in the hybrid zone.

Recently, we investigated the rare allele phenomenon by mutation screening in a neutral locus, an intron of the calmodulin (CaM) gene (Schilthuizen *et al.* 1999). The

Correspondence: Menno Schilthuizen.

†Present address: Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Locked Bag 2073, 88999 Kota Kinabalu, Malaysia. Fax: + 60 88 435311; E-mail: schilthuizen@excite.com

frequency of a normally rare haplotype was seen to increase to high levels in the centre of a hybrid zone in the land snail *Albinaria hippolyti*, without any novel mutants associated with the hybrid zone. This pattern strongly suggested the action of natural selection, either on the gene itself or on loci linked with it. This led to a third explanation (besides increased mutation rates and positive selection on the hybridzyme locus itself): presumably, purifying selection against especially unfit multilocus hybrid genotypes continually purges hybrid populations. This results in altered frequencies of markers which are in linkage disequilibrium with the selected loci. In the case of rare marker alleles, such selection could lead to a drastic increase in frequency in the area where selection takes place (i.e. the hybrid zone). The hybridzyme phenomenon is thought to be one of the reflections of this population genetic process. However, because recombination will eventually break up linkages, patterns of hybridzymes are likely to be transient in nature. The concept is further elaborated in Schilthuizen *et al.* (1999).

In this paper, we describe another example of the rare allele phenomenon in the same land snail hybrid zone. In this case, however, the locus involved is the tandemly repeated ribosomal DNA (rDNA). We studied the distribution of sequence variation in the first internal transcribed spacer (ITS-1) of the rDNA across the hybrid zone between *A. h. aphrodite* and *A. h. holtzi*, which revealed a pattern very similar to that in hybridzymes and the *CaM* intron.

On the one hand, this pattern fits well with the view expounded above, that purifying selection in hybrid populations works on many loci across the genome and should thus have its effect on many independent molecular markers. On the other hand, the results are puzzling, given that the many copies of rDNA are not expected to respond in unison. We discuss the possible explanations for our new observations.

The study organism and its hybrid zones

A. hippolyti is a cauliid land snail that lives in large parts of the Greek island of Crete (Schilthuizen *et al.* 1993). It is subdivided into six subspecies, four of which are separated by hybrid zones. The zones, most of which are not associated with obvious ecotones, are characterized by steep coincident clines (20–260 m wide) in numerous conchological, anatomical and biochemical traits. Probably, the clines are chiefly maintained by a balance between dispersal and selection against hybrids (Schilthuizen 1995). In the centre of most of the hybrid zones, hybridzyme alleles have been shown. The most spectacular case of these is the increase of a very rare *sAat* allele to frequencies above 50% in the hybrid zone between *A. h. holtzi* and *A. h. aphrodite* near the town of Kroussónas (Schilthuizen & Gittenberger 1994). It is this particular zone that the present study focuses on as well.

Materials and methods

Sampling, DNA extraction, amplification, and sequence analysis

Live adult snails were collected in November 1997 near the town of Kroussónas in central Crete. Four sample sites were located 1–2 km away on either side of the morphologically determined hybrid zone, and belonged to the typical *Albinaria hippolyti holtzi* (samples HE and HW) and *A. h. aphrodite* (samples AE and AW). The remaining two samples (CE and CW) were collected in the centre of the hybrid zone (Fig. 1). Each of the six samples was taken within an area of 10 × 50 m. In addition, a sample of *A. h. hippolyti* from Damánia (Crete) was used for comparison. Voucher specimens have been deposited in the collection of the National Museum of Natural History 'Naturalis', Leiden, the Netherlands. See Schilthuizen *et al.* (1999) for further details on the sampling procedure.

DNA extractions were carried out on single snails as described in Schilthuizen *et al.* (1999), using the protocol by van Moorsel *et al.* (2000). We did separate extractions for each of 18–35 individuals per sample, with the exception of the *A. h. hippolyti* sample, from which a single individual was studied as reference. Individuals were referred to by the sample locality and a unique number, e.g. AE-1, AE-2, and so on.

On the basis of an unpublished set of sequences for *A. hippolyti* (see below), the oligonucleotide 5'CGGGTCGTCAGGTTCAAAGAGCG3' was designed to match a conserved region in ITS-1, and used in combination with primer 5.8c from Hillis & Dixon (1991), 5'GTGCGTTCGAAATGTCGATGTTCAA3', to amplify an ≈ 400 bp fragment. Five pmol of each of these primers was then used in amplification reactions of 50 µL, with a final Mg²⁺

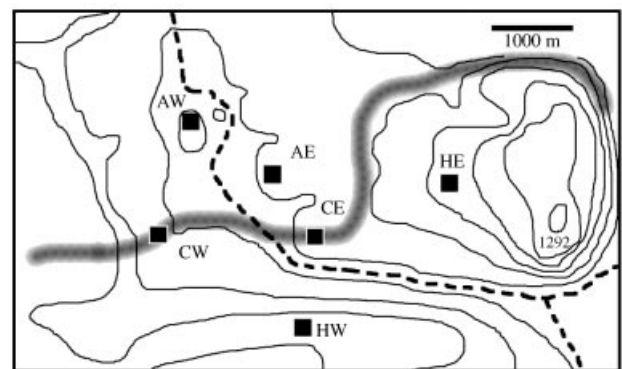


Fig. 1 Map of the study area, 5 km W of the town of Kroussónas, Central Crete. Contours are given at 50 m altitude intervals. The thick shaded line indicates the approximate width and position of the morphological hybrid zone between *Albinaria hippolyti aphrodite* (to the north) and *A. h. holtzi* (to the south). The interrupted line is a path.

concentration of 2.8 mM, 5 µL of undiluted genomic DNA and otherwise standard conditions. We used the following temperature profile on the BIO-RAD (Veenendaal, The Netherlands) Gene Cyclor: denaturation of 2 min at 94 °C, then 30 cycles of 1 min at 94 °C, 1 min at 65 °C, 1 min 30 s at 72 °C, ending with a final extension of 5 min at 72 °C.

Polymerase chain reaction (PCR) products were run overnight at 90 V on non-denaturing 7.5% polyacrylamide (PAGE) gels with 7% glycerol, and stained with ethidium bromide. Individuals were scored by assigning them with the length variant(s) they produced. Lengths were measured by comparing them with a comigrating ProMega 100 bp ladder. For 17 individuals, length variants were excised from the gel and, after purification following Sambrook *et al.* (1989), cloned in a T-vector. The correct clones were identified by performing the PCR reaction on transformed *Escherichia coli* colonies. Plasmids were isolated using QIAPREP columns and sequenced on an ABI automated sequencer in both directions using the T7 and SP6 priming sites on the vector. In a few cases, PCR products were sequenced directly, using the PCR primers as sequencing primers. The nucleotide sequences have been submitted to the GenBank database under accession numbers AF136039 and AF321498–AF321515.

The sequences were aligned by eye. Phylogenetic relationships among them were analysed using the branch-and-bound algorithm in PAUP*4.0b2 (Swofford 1999). In the sequences, gaps were treated as missing characters, but 12 informative indels were coded as characters in a separate presence/absence matrix. One hundred bootstrap replicates were performed using a heuristic search option. The trees were rooted with the *A. h. hippolyti* sequence as the outgroup.

Results

We identified four different length variants, estimated at 360, 390, 410, and 430 bp. Sequencing subsequently showed that their actual lengths were 357, 388, 418, and 394 bp, respectively. The considerable discrepancy between estimated and actual length of the latter is presumably due to a secondary structure, slowing its migration in the non-denaturing PAGE. This is commonly observed in GC-rich sequences like those from rDNA (Sambrook *et al.* 1989). We will refer to the variants by their actual lengths throughout the remainder of this paper.

Many individuals, particularly those from samples CE and CW, carried two length variants. However, the fact that no more than two variants were found per individual suggests that the separate copies in each rDNA array are tightly linked and evolve in a concerted fashion, as expected for ribosomal loci (Hillis & Dixon 1991). In Table 1, the frequencies of length variants in each sample are given. The *Albinaria hippolyti holtzi* samples carried predominantly the 388 variant, while the other variants occurred at low to

Table 1 Frequencies of internal transcribed spacer 1 length variants at each of the localities. The numbers of individuals typed are given in parentheses. If an individual was distinctly dimorphic for band size, each band was counted as 1/2

Length variant	AE(29)	AW(18)	CE(35)	CW(26)	HE(24)	HW(35)
357	—	—	0.07	0.06	0.19	0.21
388	0.47	0.17	0.33	0.11	0.68	0.79
394	0.02	—	0.43	0.31	0.10	—
418	0.52	0.83	0.17	0.52	0.03	—

very low frequencies. In *A. h. aphrodite*, 388 and 418 were the major variants. As expected, samples CE and CW, from the centre of the hybrid zone, showed mixtures of all parental length variants. Surprisingly, however, the 394 variant, which is absent from AW and HW, and rare in AE and HE (0.02 and 0.10, respectively), reached higher frequencies in the central populations, 0.43 in CE and 0.31 in CW. A comparison of the pooled data from the four parental sample sites with those from the two central sample sites, show a statistically significant association of this variant with the latter ($\chi^2 = 69.02$; d.f. = 1; $P < 0.01$).

The phylogenetic analysis (Fig. 2) of the 17 sequences resulted in four most parsimonious trees of length = 41 steps and a retention index (Farris 1989) of 0.98. The trees differ only in the positions of the 418 sequence of CW-4 and the 394 sequence of CW-7. They also show that, although each length variant is phylogenetically well-defined, individual sequences within each length variant are not identical. This means that a size variant actually represents a cluster of related sequences. There is a basal split between the 418 variant on one branch, and the three other variants on the other. Within the latter branch, the three length variants 388, 394, and 357 successively branch off.

Discussion

The pattern of a single, normally rare, sequence variant increasing in frequency in the centre of the hybrid zone, is very similar to what has been observed in *sAat* allozymes and in CaM intron haplotypes in this hybrid zone. Again, only a single ITS-1 variant appears in the zone, rather than the plethora of new length variants that would be expected if there were an elevated mutation rate. Selection acting directly upon the ITS-1 region itself is not plausible either, given that it is a region that, in spite of a few conserved cores (Liu & Schardl 1994; Armbruster *et al.* 2000) evolves largely neutrally (Schlötterer *et al.* 1994).

This leaves us with two possible explanations for the increase in the 394 length variant. As it is a largely neutral locus, it is tempting to invoke the reasoning that was applied to explain the increase of the CaM haplotype, in other words, linkage of the rare length variant with a locus that

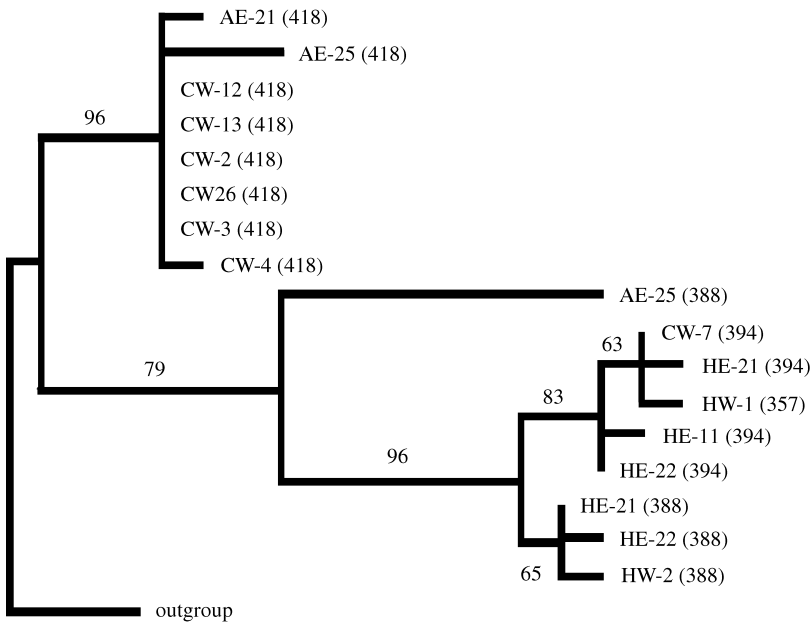


Fig. 2 A representative most parsimonious tree of the internal transcribed spacer 1 sequences. The codes for the four length variants are given with the sample codes. Numbers above the nodes give percentages of 100 bootstrap replicates.

suffers less from selection against hybrids. By this scenario, the intense selection against unfit hybrids purges the hybrid zone from other alleles and their associated ITS-1 length variants (Schilthuisen *et al.* 1999).

However, this scenario is somewhat problematic in a tandemly repeated gene like the rDNA unit. Ribosomal DNA clusters are often scattered over the genome. Even if they are restricted to a single cluster on a single chromosome, there should still be crossing-over within the rDNA unit, which would rapidly disturb any linkage disequilibrium between the rDNA array and a selected locus.

At present, we have no knowledge of the distribution of rDNA tandem repeats in the *Albinaria* genome. However, studies in 10 other species of molluscs, including a pulmonate, show that silver-stained nucleolar organizer regions (which are sites where ribosomal loci are concentrated) are always present on a single, nonrecombining satellite DNA (X. X. Li, personal communication; Vitturi *et al.* 1991). If the situation in *Albinaria hippolyti* is similar, then this could explain how linkage between the rDNA and a selected locus is retained.

However, in the concertedly evolving rDNA, there may be a second explanation for the 'rare allele phenomenon' as well. Concerted evolution is thought to result either from unequal crossing over among tandem repeats, or from biased gene conversion (Elder & Turner 1995). The latter process might also result in the pattern observed here if, in hybrids, conversion is biased towards the 394 variant. Data from hybrid lizards have shown that rDNA in these animals does indeed undergo biased conversion in the direction of only one of the parental types (Hillis *et al.* 1991). In *Albinaria* hybrids, the same might be true. However, it is not clear why gene conversion should be biased towards a variant that is normally quite rare.

In conclusion, it is striking that the 'rare allele phenomenon' should also occur in a multicopy locus like rDNA. At present, we cannot distinguish between the two possible explanations: (i) selection as in the CaM intron; or (ii) biased gene conversion. More information is needed about the cytogenetics and the rates and mechanisms of concerted evolution in *Albinaria* rDNA.

Acknowledgements

Evert Gutteling provided the outgroup sequence. Coline van Moorsel, Peter van Baarlen and Hans de Jong helped with literature, gave useful suggestions and participated in discussions on rDNA concerted evolution. Tony van Kampen operated the automated sequencer. The participants in the E-mail newsgroup Mollusc-Molecular-News helped explain the discrepancy between fragment lengths on gels and in sequences. Li Xiao Xu of the South Australian Research and Development Institute and Katja Hora of the University of Amsterdam provided information on mollusc rDNA. The manuscript was improved by the comments of an anonymous referee. This work was supported by the Netherlands Organization for Scientific Research (NWO grants 805-41.152 and SIR-14-2236).

References

- Armbruster GFJ, van Moorsel CHM, Gittenberger E (2000) Conserved sequence patterns in the non-coding ribosomal ITS-1 of distantly related snail taxa. *Journal of Molluscan Studies*, **66**, 570-573.
- Barton NH, Halliday RB, Hewitt GM (1983) Rare electrophoretic variants in a hybrid zone. *Heredity*, **50**, 139-146.
- Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113-148.
- Bradley RD, Bull JJ, Johnson AD, Hillis DM (1993) Origin of a novel allele in a mammalian hybrid zone. *Proceedings of the National Academy of Science of the USA*, **90**, 8939-8941.

- Chiba S (1997) Novel colour polymorphisms in a hybrid zone of *Mandarina*. *Biological Journal of the Linnean Society*, **61**, 369–384.
- Elder JF, Turner BJ (1995) Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology*, **70**, 297–320.
- Farris JS (1989) The retention index and the rescaled consistency index. *Cladistics*, **5**, 417–419.
- Golding GB, Strobeck C (1983) Increased number of alleles found in hybrid populations due to intragenic recombination. *Evolution*, **37**, 17–29.
- Goodfriend GA, Gould SJ (1996) Paleontology and chronology of two evolutionary transitions by hybridization in the Bahamian land snail *Cerion*. *Science*, **274**, 1894–1897.
- Guiller A, Coutellec-Vreto M-A, Madec L (1996) Genetic relationships among suspected contact zone populations of *Helix aspersa* (Gastropoda: Pulmonata) in Algeria. *Heredity*, **77**, 113–129.
- Hewitt GM (1988) Hybrid zones – natural laboratories for evolutionary studies. *Trends in Ecology and Evolution*, **3**, 158–167.
- Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology*, **66**, 411–453.
- Hillis DM, Moritz C, Porter CA, Baker RJ (1991) Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science*, **251**, 308–310.
- Hoffmann SMG, Brown WM (1995) The molecular mechanism underlying the ‘rare allele phenomenon’ in a subspecific hybrid zone of the California Field Mouse, *Peromyscus californicus*. *Journal of Molecular Evolution*, **41**, 1165–1169.
- Keenan CP (1994) Recent evolution of population structure in Australian Barramundi, *Lates calcarifer* (Bloch): an example of isolation by distance in one dimension. *Australian Journal of Marine and Freshwater Research*, **45**, 1123–1148.
- Liu J-S, Scharl CL (1994) A conserved region in Internal Transcribed Spacer 1 of plant nuclear rRNA genes. *Plant Molecular Biology*, **26**, 775–778.
- van Moorsel CHM, van Nes WJ, Megens H-J (2000) A quick, simple, and inexpensive mollusc DNA extraction protocol for PCR-based techniques. *Malacologia*, **42**, 203–206.
- Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM (1996) Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science*, **272**, 741–745.
- Sage RD, Selander RK (1979) Hybridization between species of the *Rana pipiens* complex in central Texas. *Evolution*, **33**, 1069–1088.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual* 2nd edn. Cold Spring Harbor Laboratory Press, New York.
- Schilthuisen M (1995) A comparative study of hybrid zones in the polytypic land snail *Albinaria hippolyti* (Pulmonata: Clausiliidae). *Netherlands Journal of Zoology*, **45**, 261–290.
- Schilthuisen M, Gittenberger E (1994) Parallel evolution of an *sAat* ‘hybrizyme’ in hybrid zones in *Albinaria hippolyti* (Boettger). *Heredity*, **73**, 244–248.
- Schilthuisen M, Hoekstra RF, Gittenberger E (1999) Selective increase of a rare haplotype in a land snail hybrid zone. *Proceedings of the Royal Society of London, Series B*, **266**, 2181–2185.
- Schilthuisen M, Welter-Schultes FW, Wiese V (1993) A revision of the polytypic *Albinaria hippolyti* (Boettger, 1878) from Crete (Gastropoda Pulmonata: Clausiliidae). *Zoologische Mededelingen, Leiden*, **67**, 137–157.
- Schlötterer C, Hauser M-T, von Haeseler A, Tautz D (1994) Comparative evolutionary analysis of rDNA ITS regions in *Drosophila*. *Molecular Biology and Evolution*, **11**, 513–522.
- Swofford DL (1999) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Thompson JN, Woodruff RC (1978) Mutator genes – pacemakers of evolution. *Nature*, **274**, 317–321.
- Vitturi R, Colombera D, Catalano E, Amico FP (1991) Spermatocyte chromosome analysis of *Helicella virgata* (Pulmonata: Helicidae): silver-stained and C-banded chromosomes. *Journal of Heredity*, **82**, 339–343.
- Woodruff DS (1989) Genetic anomalies associated with *Cerion* hybrid zones: the origin and maintenance of new electrophoretic variants called hybrizymes. *Biological Journal of the Linnean Society*, **36**, 281–294.

This work is part of an ongoing research project on the evolution, biogeography and taxonomy of the eastern Mediterranean clausiliid land snails, conceived and supervised by Edmund Gittenberger. This paper resulted from a postdoctoral fellowship awarded to Menno Schilthuisen by the Netherlands Organization for Scientific Research. The strong population genetic component of this project, which aimed at elucidating the enigmatic ‘rare allele phenomenon’, was the reason it was largely carried out in close cooperation with Rolf F. Hoekstra.
