Molecular Phylogenetics and Evolution 54 (2010) 542–552

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Exceptional cryptic diversity and multiple origins of parthenogenesis in a freshwater ostracod

S.N.S. Bode [a,b,*], S. Adolfsson [c,d,e], D.K. Lamatsch [a,b,f], M.J.F. Martins [c,d,g,h], O. Schmit [g,i], J. Vandekerkhove [g,h,i], F. Mezquita [i], T. Namiotko [h], G. Rossetti [g], I. Schön [a,1], R.K. Butlin [b,1], K. Martens [a,1]

[a] RBINS, Dept. of Freshwater Biology, Rue Vautier 29, 1000 Brussels, Belgium
[b] University of Sheffield, Dept. of Animal and Plant Sciences, Western Bank Sheffield, S10 2TN, UK
[c] EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Überlandstrasse 133, 8600 Dübendorf, Switzerland
[d] ETH-Zürich, Institute of Integrative Biology (IBZ), Universitätstrasse 16, 8092 Zürich, Switzerland
[e] GEMI, UMR CNRS IRD 2724, IRD, 911 Avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France
[f] Institute of Limnology, Austrian Academy of Sciences, 5310 Mondsee, Austria
[g] University of Parma, Dept. of Environmental Sciences, Viale G.P. Usberti 33A, 43100 Parma, Italy
[h] University of Gdansk, Dept. of Genetics, Lab. of Limnozoology, Kładki 24, 80-822 Gdańsk, Poland
[i] University of Valencia, Dept. of Microbiology and Ecology, Avenue Dr. Moliner 50, 46100 Burjassot, Spain

ARTICLE INFO

Article history:
Received 15 May 2009
Revised 5 August 2009
Accepted 18 August 2009
Available online 22 August 2009

Keywords:
Cryptic species
Species complex
Geographical parthenogenesis
Asexual reproduction
Evolution of sex
Coalescence
mtDNA
Ostracoda
Eucypris virens

ABSTRACT

The persistence of asexual reproduction in many taxa depends on a balance between the origin of new asexual lineages and the extinction of old ones. This turnover determines the diversity of extant asexual populations and so influences the interaction between sexual and asexual modes of reproduction. Species with mixed reproduction, like the freshwater ostracod (Crustacea) morphospecies *Eucypris virens*, are a good model to examine these dynamics. This species is also a geographic parthenogen, in which sexual females and males co-exist with asexual females in the circum-Mediterranean area only, whereas asexual females occur all over Europe. A molecular phylogeny of *E. virens* based on the mitochondrial COI and 16S fragments is presented. It is characterised by many distinct clusters of haplotypes which are either exclusively sexual or asexual, with only one exception, and are often separated by deep branches. Analysis of the phylogeny reveals an astonishing cryptic diversity, which indicates the existence of a species complex with more than 40 cryptic taxa. We therefore suggest a revision of the single species status of *E. virens*. The phylogeny indicates multiple transitions from diverse sexual ancestor populations to asexuality. Although many transitions appear to be ancient, we argue that this may be an artefact of the existence of unsampled or extinct sexual lineages.

1. Introduction

Sex is the most common mode of reproduction in the animal kingdom (see for example, Bell, 1982). It is also the ancestral mode of all metazoa and probably all eukaryotes. Yet it is easily lost by mutation, hybridisation, or polyploidisation and it also comes with costs, such as the twofold cost of males (Maynard Smith, 1978), vulnerability to sexually transmitted infections or predators and many more (Kondrashov, 1993). The reason behind the success of sex is one of the most intriguing paradoxes of evolution still waiting to be solved. Many hypotheses have been proposed but, so far, none appears capable of providing a general and inclusive explanation for widespread sexual reproduction (West et al., 1999; Butlin, 2002). The best-supported hypotheses implicate mutation accumulation due to lack of recombination in asexuals (Kondrashov, 1988; Lynch et al., 1993) or rapidly changing selection pressures due to the 'Red Queen' process (Van Valen, 1973), particularly as a result of host–parasite interactions (Hamilton, 1980). It is becoming increasingly evident that several factors must be involved simultaneously (West et al., 1999) and that their contributions may differ among organisms (Burt, 2000; Schön et al., 2009).

Asexual mutants can have significant selective advantages due to superior population growth and colonising abilities. However, they have not generally replaced their sexual relatives but instead they are typically short-lived in evolutionary terms (Maynard Smith, 1978; Bell, 1982). It is difficult to estimate the ages of asexual lineages (Butlin, 2002). Therefore apparent exceptions to the rule are controversial, ironically called 'ancient asexual scandals' (Judson and Normark, 1996). The most well-known putative ancient asexuals are bdelloid rotifers (Welch and Meselson, 2000), darwinulid ostracods (Martens et al., 2003; Martens and Schön,

* Corresponding author. Address: University of Sheffield, Dept. of Animal and Plant Sciences, Western Bank Sheffield, S10 2TN, UK. Fax: +44 (0) 1142220002.
E-mail address: s.bode@sheffield.ac.uk (S.N.S. Bode).
[1] These authors contributed equally to this work.

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.08.022

2008), and oribatid mites (Heethoff et al., 2007; Laumann et al., 2007). These are entire families that have been fully asexual for millions of years, a conclusion supported by several lines of evidence, such as genomic characteristics in the case of bdelloids (Welch and Meselson, 2001; Welch et al., 2004) and extensive fossil records without males in the case of the darwinulids (Martens et al., 2003; Martens and Schön, 2008; but see Smith et al., 2006) and oribatid mites (Maraun and Scheu, 2000).

Ancient asexual lineages are genetically diverse because of an accumulation of mutations since their origin. In other taxa, genetically diverse asexual populations may be generated by a high rate of origin resulting in a standing diversity that is dependent, at least in part, on population size (Janko et al., 2008). High rates of origin of asexuals occur in cyclical parthenogens through repeated loss of the sexual phase, as in aphids for example (Delmotte et al., 2001; Loxdale and Lushai, 2003), and where there is repeated hybridisation, as in *Poeciliopsis* fish (Mateos and Vrijenhoek, 2002) or *Daphnia* (Paland et al., 2005). In some cases, frequent origin of asexual lineages from an ancestral sexual population clearly occurs, but the mechanism remains uncertain, as in *Potamopyrgus* snails (Neiman et al., 2005). It is important to understand both the mechanism and rate of origin of asexual lineages because they influence models for the maintenance of co-existing sexual and asexual populations. For example, Janko et al. (2008) show that asexual lineages may be short-lived purely because of turn-over in a finite population, without the need to invoke declining fitness due to mutation accumulation or environmental change. In contrast, in *Potamopyrgus*, frequent origin of novel asexual lineages is necessary to maintain the asexual population in the face of co-evolving parasites (Lively et al., 2004).

The freshwater ostracod (Crustacea) *Eucypris virens* is a model taxon containing distinct and non-cyclical sexual and asexual lineages. The reproductive mode is the only known difference between these groups of lineages. This species belongs to the Cyprididae, a family that harbours many species with either fully asexual, fully sexual or mixed modes of reproduction (Horne et al., 1998; Martens et al., 2008). It belongs to the superfamily Cypridoidea, which is related to the superfamily Darwinuloidea whose extant representatives are the putatively ancient asexual ostracod family of the Darwinulidae, mentioned above.

Eucypris virens is characterised by high levels of variability, both at the morphological (Martens, 1998; Baltanas et al., 2002) and genetic level (Rossi et al., 1998; Schön et al., 2000; Schön, 2007). Schön et al. (2000) suggested that both multiple origins of asexual lineages and hybridisation might contribute to this high diversity. They also proposed that some asexual lineages might be old, with sequence divergence of the mitochondrial COI locus up to 21% indicating that the oldest asexual lineages might have originated 10 million years ago. However, wider sampling of lineages was needed to test these ideas and fossil evidence is scarce, as this particular species occupies only temporary pools.

Eucypris virens is also known for its distinct geographical distribution in which males are restricted to the circum-Mediterranean area and have never been recorded north of the Alps (Horne et al., 1998) whereas asexual lineages occur in Northern Europe as well as in sympatry with sexual lineages. This pattern is termed "geographical parthenogenesis" (Vandel, 1928) and has been recorded in a number of organisms (Law and Crespi, 2002; van Dijk, 2003). This phenomenon may be explained by the greater colonising ability of asexuals, their potential to adapt to novel environments because they are free of the retarding effects of gene flow (see Butlin et al., 2003, for further discussion), or the difference in environmental stability between regions (see Horne and Martens, 1999, for a review).

Here, we investigate the number of sexual and asexual lineages of *E. virens* in Europe and how they are related to each other phylogenetically, using partial mitochondrial cytochrome oxidase I (COI) and 16S ribosomal DNA gene sequences. This phylogenetic information provides a framework for understanding the origin of the genetic diversity within asexuals, the rate of appearance of asexual lineages and their ages, as well as their geographic distribution. Asexual organisms are thought to have originated from sexual ancestors without reversals. This assumption is supported by multiple lines of evidence (Bell, 1982) and by the argument that it requires more evolutionary steps to re-acquire sexual reproduction than it does to lose it (but see Domes et al., 2007). Although most species of the Cyprididae are asexual or include asexual lineages, it is very likely that the ancestral state of the family was sexual reproduction. Therefore, we assume that the ancestor of *E. virens* was sexual and test whether transitions from sexual to asexual reproduction are frequently encountered across the phylogeny or whether a single ancient event has led to diverse asexual lineages. Furthermore, we determined the number of independent sexual lineages (potential cryptic species) within the morphospecies because origin of asexuals from distinct sexual ancestors may contribute to clonal diversity.

2. Materials and methods

2.1. Specimen collection and DNA extraction

We conducted a large-scale sampling campaign covering most of Europe and the Mediterranean including North Africa (Fig. 1). Between 50 and 100 specimens of *E. virens* were collected with a hand net from each of 135 temporary pools during winter and early spring of 2005/2006 and 2006/2007 and were kept alive in mineral water overnight to eliminate stomach contents. The animals were slowly killed in diluted ethanol, in order that their carapace valves stayed open, and specimens were subsequently transferred into 100% ethanol and kept at 4 °C. Details of the sampling sites utilised in this study, selected from a database of field sampling data compiled by all members of the SexAsex project (http://evirens.group.shef.ac.uk/), are reported in Table 1S and have been visualised in Fig. 1 using *DIVA-GIS v.5.4* software (Hijmans et al., 2001).

Specimens were chosen for analysis at random except that we preferentially used individuals whose valves were open and intact, as those were good indications that specimens were alive and healthy when they were killed. Genomic DNA was extracted from the ostracod softparts using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's protocol, after washing each specimen with ultrapure water and PBS (1× concentration) buffer. Valves were retained for morphometric analysis (to be reported elsewhere). A subset of ostracods was prepared for both allozyme analysis and DNA extraction with subsequent sequencing of COI (Adolfsson et al., submitted for publication).

2.2. DNA sequencing

Our choice of the first marker, the mitochondrial cytochrome oxidase I (COI), was based on the availability of universal primers, its high rate of evolution (Lunt et al., 1996) and its widespread use for DNA barcoding which make it especially suited to examine intra-specific relationships in a taxon where there is limited genetic information. A 657 bp fragment of the COI gene was amplified using the following universal invertebrate primers: HCO2198 (forward) 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′ and LCO1490 (reverse) 5′-GGTCAACAAATCATAAAGATATTGG-3′ (Folmer et al., 1994). A few populations from Corfu and Italy could not be amplified with this COI primer pair and therefore another pair of specific primers (FMCO 5′-TAGGACAGCCRGGATCWCT-3′ and RMCO 5′-

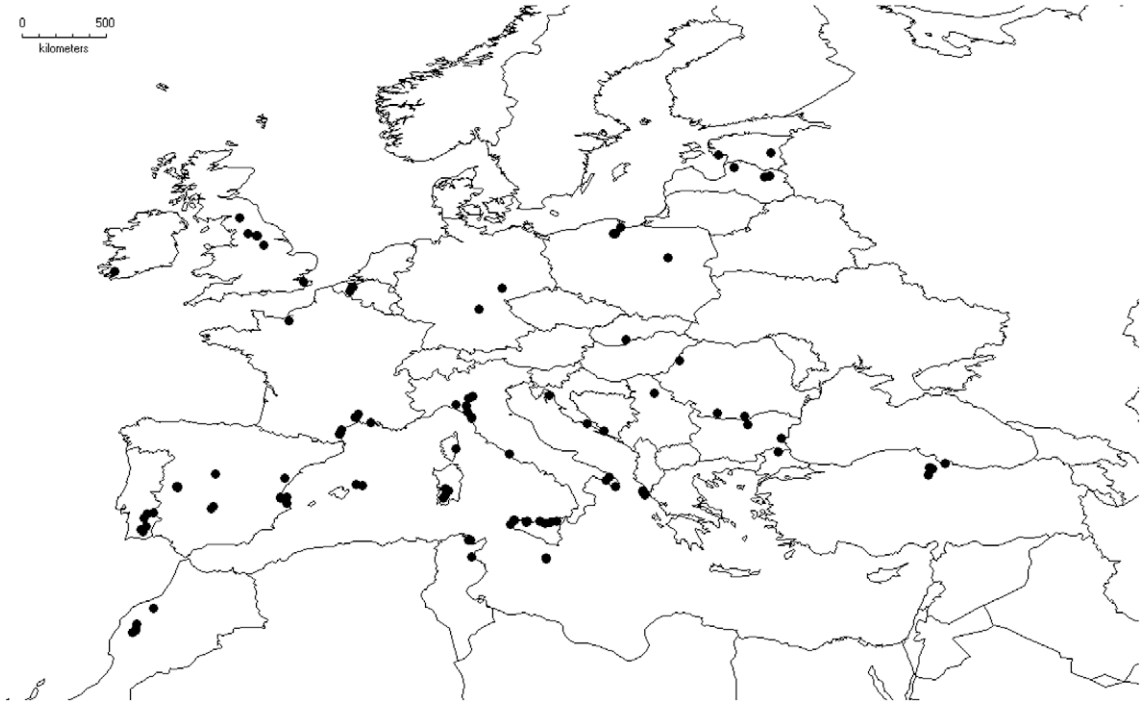

Fig. 1. Sampling sites of *Eucypris virens*, corresponding to *Table* 1S. The map was produced using *DIVA-GIS v.5.4.*

CGGTCTGTTAAWAGCATWGTGA-3′), resulting in a 476 bp fragment, was designed using *Primer3 v.0.4.0* (Rozen and Skaletsky, 2000).

PCR amplification was performed using a Westburg Biometra T Personal thermal cycler starting with 3 min of denaturation at 95 °C followed by 35 cycles of 30 s at 95 °C, 54.5 °C for 30 s of annealing time, 1.5 min at 72 °C of extension and a final extension stage of 10 min at 72 °C. The 20 μl volume reactions contained 0.5U recombinant *Taq* DNA polymerase (Invitrogen), 1× PCR Buffer (200 mM Tris–HCl [pH 8.4], 500 mM KCl), 2 mM dNTP, 10 μM of each primer, 1.5 mM $MgCl_2$, and ~8 ng template DNA, although DNA yield was variable. Difficult template DNA was amplified using the HotStar Taq Master Mix Plus Kit (Qiagen) applying the manufacturer's protocol with the same conditions as above, but a total of 37 cycles were used with an annealing temperature of 48 °C, or 45 °C for the primer pair FMCO and RMCO.

The second marker, 16S, was chosen to complement the COI information with a more conserved marker (DeSalle et al., 1987). A 480 bp fragment was amplified from the mitochondrial 16S rRNA gene using primer 16SL (forward) 5′-CGCCTGTTTAACAAAAACAT-3′ and as reverse 16SH 5′-CCGGTCTGAACTCAGATCACGT-3′ or 16SBr 5′-CCGGTCTGAACTCAGATCACGT-3′ (Palumbi, 1996). Amplification was performed with the HotStar Taq Master Mix Plus Kit and 35 cycles of 50 s denaturation at 94 °C, 50 s annealing at 50 °C, and 1 min 20 s extension at 72 °C. Conditions were otherwise as described above, except that less DNA was used. PCR products were run on 1.4% agarose gels, stained with SYBR® Safe (Invitrogen) and photographed. The remaining products were cleaned using the GFX™ PCR DNA and Gel Band purification kit (GE Healthcare) following the manufacturer's protocol and eluting according to concentration estimates from the gel images. The cleaned PCR products were directly sequenced in both directions with the PCR primers and the ABI BigDye Terminator v.1.1. kit and then further purified with an ethanol/EDTA precipitation method (ABI cycle sequencing kit manual) to be run on an ABI 3130xl Genetic Analyser.

Chromatograms were edited in *CodonCode Aligner v.1.6.3.* and unambiguously aligned through the ClustalW algorithm in *MEGA v.4.0* (Tamura et al., 2007). Sequences were checked for identity by BLAST search for highly similar sequences (megablast) (Altschul et al., 1990). Putative COI fragments scored highly with other COI sequences of *E. virens* in Genbank (Schön et al., 2000). So far, there have been no submissions of 16S from *E. virens* to Genbank. However, the 16S sequences generated had high similarity with other 16S ostracod sequences. COI was also translated in *MEGA* using the invertebrate mitochondrial genetic code to test for the presence of amplified pseudogenes (Bensasson et al., 2001) identifiable through stop codons within the reading frame (Song et al., 2008). A total of 449 *E. virens* sequences was obtained for COI (Table 1S). In order to reduce the computational effort, we used 374 sequences in phylogenetic analyses, excluding 75 that were identical to sequences already present in the data set. For 16S, 219 sequences were obtained. Three outgroups from the same subfamily (Eucypridinae) (Meisch, 2000), *Eucypris pigra*, *Tonnacypris lutaria* (COI and 16S) and *Eucypris crassa* (16S only) were also sequenced using the same conditions as for *E. virens*. All edited sequences are available from Genbank, Accession Nos. GQ914057 – GQ914280 (16S) and GQ914281 – GQ914731 (COI).

Some PCR amplifications were only successful after several attempts, probably because of low quantity or quality of DNA or, in some cases, because of sequence divergence in the primer regions (especially for COI). This increased the risk of amplifying contaminating DNA. Therefore, we checked for compatibility of COI and 16S sequences where both were available and removed 18 individuals where there was evidence of contamination (identical COI sequences for individuals with divergent 16S sequences, or vice versa). Eight of the excluded individuals were from one Turkish site (LAD) which had divergent 16S sequences and inconsistent COI sequences. Some individuals from this site were also divergent morphologically and probably do not belong to *E. virens*.

2.3. Phylogenetic analyses

Bayesian phylogenetic analyses were performed using *Beast v.1.4.8.* (Drummond and Rambaut, 2007) and maximum likelihood (ML) analyses using *PhyML v.2.4.4.* (Guindon and Gascuel, 2003), on

the COI and 16S datasets separately, and on the concatenation of the two as a partitioned supermatrix. *Modeltest v.3.7* (Posada and Crandall, 1998) was run on both separate datasets to test for the most adequate model of evolution. The programs *Dnaml* and *Dnamlk* (Felsenstein and Churchill, 1996) from the *Phylip package v.3.68* (http://Evolution.genetics.washington.edu/phylip.html) were run in order to test whether the data conform to a molecular clock. In addition, the following three models implemented in *Beast*, were applied to a preliminary COI dataset and their likelihoods were compared: General time reversible with gamma distribution and invariable sites (GTR + Γ + I); SRD06 (Hasegawa–Kishino–Yano [HKY] with the first and second position partitioned from the third position for each codon); and HKY with three partitions for the three bases per codon (HKY + Γ + 3). Xml files for the *Beast* runs were prepared using *Beauti v.1.4.8.* The searches were run assuming a strict molecular clock but no clock rate was specified. For the remaining parameters, the default settings of *Beauti* were used, except that Jeffrey's prior distribution parameter was set to 5.0. We applied a Yule algorithm describing a pure-birth speciation process rather than a population-based coalescent process (this decision was based on preliminary runs which showed that deep branches were present in the dataset).

For each dataset, three independent searches were run for 10,000,000 generations each. *Tracer v.1.4.* (Rambaut and Drummond, 2007) was used for visualising the output of the Bayesian tree searches and the trees from all runs were combined using *LogCombiner v.1.4.8,* re-sampling every 10,000th tree and thereby obtaining 3,000 trees. *TreeAnnotator v.1.4.8.* was used to discard a burn-in of the first 1,000 trees and to choose one tree out of 2,000 remaining trees based on the likelihoods of their topologies. Trees were manipulated and rotated to facilitate comparisons of the topologies between the various datasets (and methods) using *FigTree v.1.2* (http://tree.bio.ed.ac.uk/software/figtree/).

A partition homogeneity test (Farris et al., 1994), also called incongruence length difference test (ILD), was performed with 1,000 replicates on the concatenated COI + 16S dataset using *PAUP v.4.0 b10.* (Swofford, 2003). In addition, the two separate datasets, with congruent individuals only, were run in separate Beast searches and their combined likelihoods were compared with the likelihood from the concatenated dataset to confirm that the two sets of sequences are compatible with a common tree, as expected for two sections of the mitochondrial genome. The concatenated dataset consisted of 156 individuals of *E. virens* and two outgroups, for which both regions were sequenced (with ends trimmed: COI bp67-bp605; 16S bp82-bp411). The Bayesian search parameters and models were as above with the single gene datasets, but manually combined in Xml format to allow partitioned evolutionary models but a common tree search.

Two ML searches in *PhyML* with neighbour-joining starting tree, substitution model parameters estimated during the search and 500 non-parametric bootstrap replicates each were performed for the COI and the concatenated datasets. Their topologies were compared and bootstrap values were combined, to provide 1,000 bootstrap replicates.

2.4. Quantitative cluster analysis

In order to test whether our set of samples can be divided into distinct clusters of related haplotypes, we applied a maximum likelihood method developed by Pons et al. (2006, provided as an R script by T. Barraclough and M.T. Monaghan). This method tests the proposition that the gene tree can be described by two processes, a pure-birth process generating a set of populations and a coalescent process within populations, and determines the most likely transition time between these phases of evolution. The transition point defines a set of clusters or singleton sequences each of which can be considered as a sample from an independently evolving population (see Section 4). The method was applied to the strict clock tree of the COI dataset obtained from *Beast,* using the package *R v.2.8.0.* (R Development Core Team, 2008) with the extension *ape v.2.2–2* (Paradis et al., 2004) for phylogenetic applications. The numbers of lineages were plotted against time and a threshold between speciation and coalescence was fitted using the general mixed Yule coalescent (GMYC) model. A χ^2 test was performed (M.T. Monaghan, pers. comm.) to test for significance of application of this clustering model against the null model of a uniform branching pattern. Support limits for the number of clusters were estimated from the likelihood surface for the time of transition between phases. Clusters of sequences, or singletons, were labelled numerically (with prefix 'p') and classified as either sexual, in the presence of at least one male, asexual in the absence of males, or mixed in the case of presence of males in the same cluster as females from asexual populations (sites where males have not been observed). Assignments of reproductive modes to clusters were aided by allozyme and flow cytometry analyses of a subset of the sequenced individuals (Adolfsson et al., submitted for publication). The clusters defined by this analysis of the COI data set were fitted onto the topologies of the 16S and ML trees. From here on, we will refer to the sequence clusters inferred by this analysis simply as 'clusters'.

A limitation of the Pons et al. (2006) method is that it fits a single transition time between the pure-birth and coalescent phases. This may make the outcome sensitive to over-sampling of individual genotypes. Therefore, we applied the method not only to the COI dataset used in the phylogenetic analysis (374 sequences) but also to the full set of 449 sequences and to a set with two large groups of identical haplotypes reduced to a single sequence per site (406 sequences, excluding 26 sequences from cluster p34 collected at sites COA and COB and 17 sequences from cluster p25 collected at sites DR1, BER, BRC, CC3, CC4, CC5 and MF2).

2.5. Other statistical analyses

Mean sequence diversities (π) within and between clusters, and Tajima's *D* statistics within clusters, were calculated in *MEGA v.4.1.* (Tamura et al., 2007) using the complete deletion option. Diversity values were used to test whether sister-pairs of clusters conform to the 4X rule (Birky et al., 2005), which aims to delineate independent populations in a simpler, but conceptually similar way to the method of Pons et al. (2006). AT content in COI was also calculated with *MEGA.*

Although we sampled widely across Europe, it is almost certain that our dataset does not include all extant lineages of *E. virens.* This is especially true for the extant sexual lineages, since sexual populations have restricted distributions (see below) and no sexual populations to the east of the Mediterranean basin were found. Therefore, we conducted rarefaction analyses, based on the COI data set. We simulated rarefaction for all asexuals (249 individuals; 96 localities), all sexuals (127 individuals; 28 localities), and the two reproductive modes together (376 individuals; 117 localities of which 6 contain both sexual and asexual individuals) using the program *EstimateS v.8.0.0.* (Colwell, 2005) in order to estimate how many more clusters we would expect to find if the sampling effort were increased, given the existing data and sampling information. Samples were randomized 1,000 times without replacement. The following estimators for the total number of clusters to be expected, and their respective confidence intervals (where applicable), were extracted: Chao 1 and Chao 2 with bias-correction (Chao, 1987), Jackknife 2 (Smith and van Belle, 1984) and Michaelis–Menten (Colwell and Coddington, 1994), using the recommended 'MMMeans' option whereby the estimates of each sample pooling level are computed just once.

In order to summarise the spatial distribution of each cluster, we calculated the centroid of all sites where the cluster was present and the mean distance of individual sample sites from that central position, using the Haversine formula (see http://www.movable-type.co.uk/scripts/gis-faq-5.1.html).

3. Results

3.1. Phylogenetic analysis of COI data

Increasing evidence for frequent nuclear insertions of mitochondrial DNA, known as numts, has lead to the conclusion that the high diversity of COI lineages detected in DNA barcoding projects may be an overestimate of the true diversity (Song et al., 2008). This phenomenon is particularly frequent in crustaceans (Bensasson et al., 2000, 2001). However, the sequences obtained here contain neither stop codons nor indels and can be aligned unambiguously, despite a large number of base pair differences. It has also been suggested (Song et al., 2008) that a lower AT content can indicate the presence of numts due to a different compositional bias between mitochondrial and nuclear DNA. In our data, the base composition of the COI sequences was biased towards A + T (A = 27.4; T = 33.0; G = 17.3; C = 22.3), which is consistent with expectations for the mitochondrial genome.

Modeltest v.3.7 found TIM + Γ + I to be the preferred model of substitution for the total COI dataset according to the AIC and TVM + Γ + I according to the hierarchical likelihood ratio tests (−ln Lk = 9226.437 and −ln Lk = 9229.16 for models chosen by AIC and hLRTs, respectively). The TIM + Γ + I model requires less parameters to be estimated ($K = 8$) than TVM + Γ + I ($K = 9$). Both models are nested within the HKY model, which is implemented in the packages used and so was our chosen model. In ML tree searches, we used HKY + Γ + I with 6 rate categories for the gamma distribution. In the independent model comparison implemented in *Beast*, the HKY model with three codon positions partitioned was preferred over the SRD06 and GTR + Γ + I models (Bayes factors calculated in *Tracer* (Suchard et al., 2001) 300.6nats and 52.5nats, respectively, where 'nats' are Bayes factor units on a natural logarithm scale) and was therefore chosen for subsequent analyses except where otherwise indicated. We used 6 rate categories. Because this model tends to fit a low substitution rate for 2nd positions of codons, there is no requirement for a separate category of invariant sites. Three combined independent *Beast* runs were sufficient to achieve effective sample size (ESS) parameter values above the recommended threshold of 200 (Drummond and Rambaut, 2007) for confident parameter estimation and a smooth unimodal trace curve. The likelihoods of each run were similar and on this basis it was concluded that the three independent runs converged to a similar optimum. The test of the molecular clock with *Dnaml*/*Dnamlk* showed a higher likelihood for the non-clocklike mode of evolution, but the improvement was only marginally significant ($-2\Delta LL = 420.13$, df = 372, $p = 0.043$). Therefore we have assumed clock-like evolution in subsequent analyses.

The COI data are characterised by groups of closely related haplotypes which are separated by surprisingly large distances (*p*-distance up to 14.5% within *E. virens*). This is reflected in the phylogenetic analysis (Fig. 2) which shows multiple clades of closely related sequences, well supported by both Bayesian and ML analyses, which are connected by long branches. Support for deeper nodes was relatively weak because of saturation of COI sequence divergence, leading to some inconsistencies between the

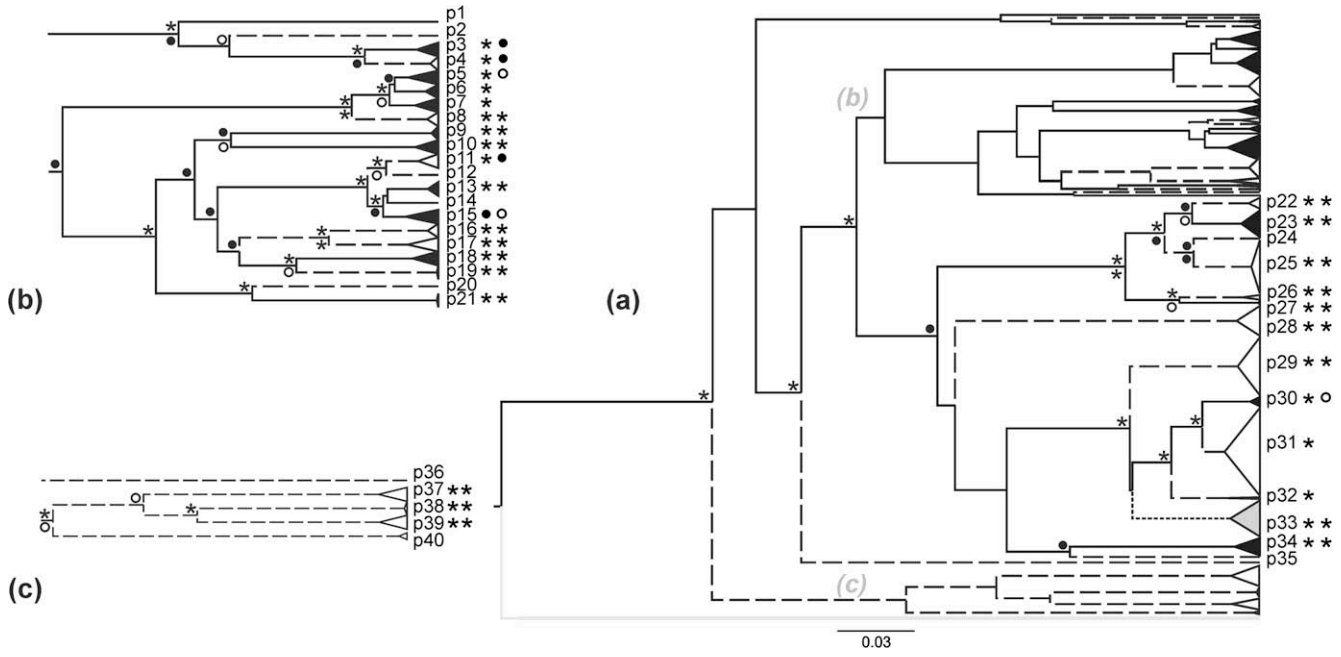

Fig. 2. Bayesian phylogenetic strict clock tree of COI dataset. Triangles at the tips represent clusters determined with the Pons et al. (2006) method. Lines without triangles are singletons. Clusters are labelled numerically, p1 to p40, p standing for 'provisional species', and labels are next to the triangles on the right. Widths of triangles reflect the numbers of individuals in the sample belonging to the clusters (part a only), whereas depth indicates maximum divergence within clusters (parts a–c). Full lines and black fills represent purely sexual groups. White triangles refer to asexual clusters with dashed lines, which indicate that the transition from sexual to asexual reproductive mode must have happened somewhere along this branch. A dotted line with a grey triangle indicates a mixed cluster. Outgroups are indicated as grey lines. Values of posterior probabilities (above nodes) and bootstrap percentages (below nodes) are represented by symbols (posterior probability; bootstrap): star (>0.95; 95%), black dot (>0.7; 70%) white dot (>0.5; 50%). Values for the basal nodes of clusters are placed beside the cluster names (posterior probability; bootstrap). The bootstrap values obtained by *PhyML v.2.4.4.* with 1000 bootstrap replicates, were superimposed from congruent ML topologies. (a) Full Bayesian COI tree. The bar indicates units of substitutions per site. (b) Expansion of the top part of the phylogeny, indicated as (b) in (a). The same applies to (c) which is the bottom part of the phylogeny, indicated as (c) in (a). This tree and both ML trees with bootstrap supports are available in Newick format as Supplementary data.

topology of the tree shown and the ML tree (available in Supplementary data).

3.2. Cluster analysis

Application of the Pons et al. (2006) method to the COI tree generated by the *Beast* analysis (374 sequence data set) revealed that the model of two different processes, Yule branching and coalescent events within populations, was a significantly better fit than the null model of a single process ($-2\Delta LL = 49.3$, df = 3, $p = 1.114e\text{-}10$; branching rates $\lambda_1 = 16.12$ and $\lambda_2 = 210.49$, and model parameters p_1=0.93 and p_2=0.60). Branching rates are measured relative to COI distance. Parameters p_1 and p_2 indicate whether branching rates are speeding up or slowing down. The p_1 value close to 1 indicates approximately constant branching rate in the diversification phase whereas $p_2 < 1$ suggests growing populations in which coalescent events are biased towards earlier times (Pons et al., 2006). The transition between these two phases is clearly visible in a plot of COI lineages through time (Fig. 3). The ML threshold between the two phases reveals the existence of 40 separate *E. virens* clusters including 8 singletons (support limits 34–43). Clusters are indicated and labelled in Fig. 2. The average p-distance between clusters was 9.5% (standard error 2.2%).

Out of 11 sister pairs of clusters, excluding singletons, 8 were found to be consistent with the 4X rule (p3–p4, p9–p10, p13–p15, p18–p19, p22–p23, p26–p27, p30–p31, p38–p39), while in 3 cases (p5–p6, [p5 + p6]–p7, p16–p17) the mean distance between clusters was less than four times the larger of the mean distances within clusters. This rule has been found to be more conservative than the ML method (Birky and Barraclough, 2009), which is consistent with our results.

Analysis of the complete data set (449 sequences), including all repeated haplotypes, resulted in 38 clusters (support limits 35–40). Clusters p5 + p6, and p14 + p15 were combined but there were no other differences from the reduced data set. When two large groups of identical haplotypes, from sites in Spain, Corfu, the UK and Belgium, were reduced to one haplotype per region (406 sequence data set), the number of clusters was further reduced to 33 (support limits 27–38). Cluster combinations were: p13 + p14 + p15, p5 + p6, p11 + p12, p24 + p25, p30 + p31. Other model parameters changed little, except p_1 where the lower value

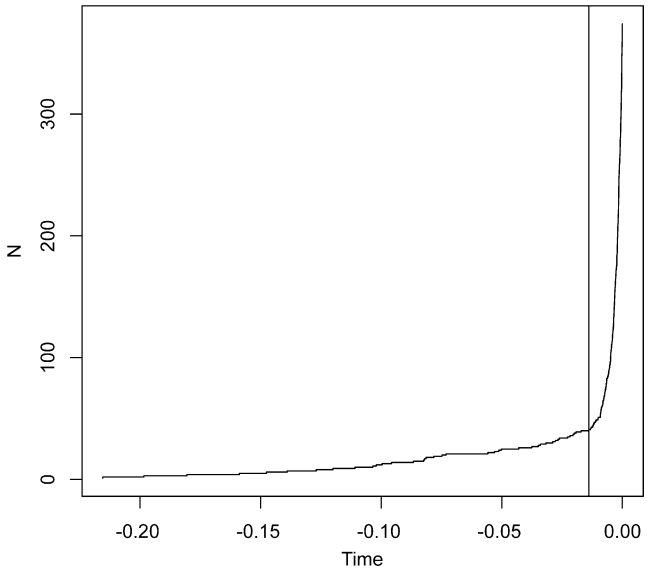

Fig. 3. Lineage-through-time plot of the COI Bayesian tree. The threshold between Yule speciation and coalescence within populations, determined with the GMYC model, is indicated here by a vertical line. The time axis is in substitutions per site.

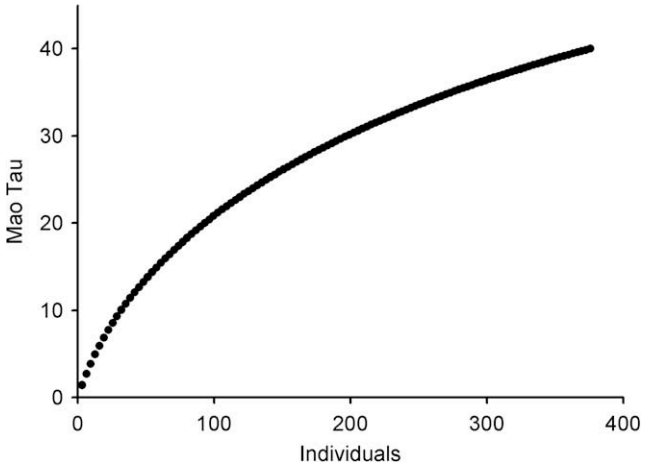

Fig. 4. Individual-based rarefaction curve of COI data for all clusters. Mao Tau is the species richness estimator.

indicates slowdown in speciation rate or incomplete sampling of extant species ($\lambda_1 = 61.830$ and $\lambda_2 = 123.866$, $p_1 = 0.411$ and $p_2 = 0.693$). Hereafter, we use the 40 cluster solution, unless otherwise specified.

Rarefaction simulation curves of the *E. virens* COI dataset show that expected plateaus of cluster numbers have not yet been reached (Fig. 4). Extrapolation from the data, with re-sampling, provided estimates of the total number of clusters between 44 (Chao 1) and 64 (Jackknife 2) (Table 1). Because sexual and asexual clusters are differently distributed spatially (see below), we also conducted rarefaction analyses separately for the two sets of samples (Table 1). We conclude that many clusters, of both reproductive modes, remain to be sampled.

3.3. Analysis of 16S sequences

For 16S, the optimal model of substitution, according to the AIC criterion, was the model K81uf + Γ + I ($-\ln Lk = 2111.542$, $K = 7$) and, according to hLRTs, was HKY + Γ + I ($-\ln Lk = 2112.925$, $K = 6$). Again, both models are nested within the HKY + Γ + I model, which we implemented in *Beast*. For 16S and also for the concatenated data, three independent runs were sufficient, as for COI, and all chains converged to optimal unimodal curves for all parameters.

The ILD test was not significant (p = 0.83) indicating that the two congruent datasets were not conflicting. Comparison between independent analyses of the two data sets and joint analysis, in *Beast*, showed that combined analysis produced a higher marginal likelihood than the separate analyses, indicating that data for the two loci support the same tree. Inspection of the trees shows that the differences are in deep nodes that are poorly supported by the COI data,

Table 1
Mean asymptotic values of four extrapolation estimators and their confidence intervals (CI), where applicable, as applied to the COI data.

		Chao 1	Chao 2	MM	Jack 2
All clusters	Mean	44	53.22	51.14	63.79
	95% CI Lower bound	40.81	44.04	na	na
	95% CI Upper bound	59.8	83.24	na	na
Sexual clusters	Mean	18.6	32.46	36.46	35.14
	95% CI Lower bound	18.05	21.37	na	na
	95% CI Upper bound	25.11	80.06	na	na
Asexual clusters	Mean	29.0	26.97	28.06	31.97
	95% CI Lower bound	24.86	24.54	na	na
	95% CI Upper bound	52.91	40.32	na	na

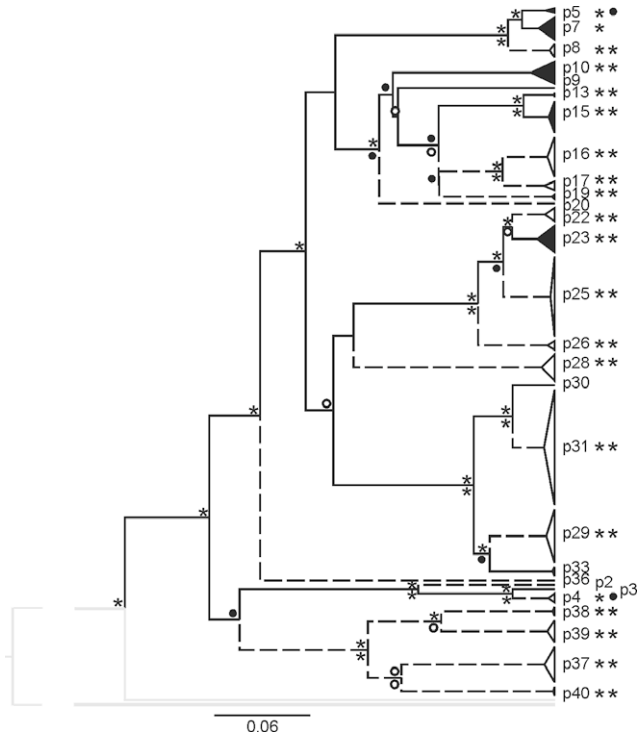

Fig. 5. Bayesian phylogenetic strict clock tree of the concatenated COI and 16S dataset. Annotations as in Fig. 1(a). The branches reaching to the last node between the outgroups are longer than shown here but were shortened in order to save space. This tree is available in Newick format as Supplementary data.

such as the placement of p2–p4 (compare Figs. 2 and 5). There was no disagreement in terms of placing individuals into clusters.

Separate Bayesian trees were obtained (available in Supplementary data) for 16S sequences and for the concatenated dataset (Fig. 5). Individuals included in the same cluster in the COI analysis always grouped together in the 16S and combined trees. The majority of individuals sequenced only for 16S clearly fell into the clusters defined by the Pons et al. (2006) method applied to the COI tree. However, 2 and 7 individuals fell into two additional groups of similar sequences which we identify as clusters p41 and p42. The larger cluster (p42) contains males and so appears to be sexual. The smaller cluster (p41) is all female and is tentatively interpreted as asexual. Both were sampled on the island of Corfu, Greece (Table 2S).

3.4. Comparison of sexual and asexual Eucypris virens clusters

All clusters, except one, either contained a mix of males and females from one or a few nearby sites or were exclusively female. Some females from sites with males fell into exclusively female clusters suggesting these were asexual females coexisting with sexuals in the same pond. All individuals from exclusively female sites fell into exclusively female clusters. Therefore, we interpret clusters containing males as sexually reproducing lineages (Fig. 2, black triangles) and clusters containing only females as asexual lineages (Fig. 2, white triangles). Allozyme genotypes and estimation of ploidy using flow cytometry fully support these inferences (Adolfsson et al., submitted for publication).

One cluster, p33 (Fig. 2, grey triangle), contains males, diploid females and females that are likely to belong to an asexually reproducing lineage or lineages because they are triploid (Adolfsson et al., submitted for publication). For this cluster, we tested whether a single origin of asexuality was compatible with the COI sequence data by comparing the marginal likelihood of an unconstrained tree with the marginal likelihood of a tree in which all putative asexual individuals were constrained to belong to a single group. The analysis was conducted in *Beast* with the same setting as above except that we used the simpler HKY + Γ + I model, rather than partitioning codon positions, because of the smaller dataset. The unconstrained tree was a better fit, but only marginally so (Bayes factor, 2.42nats). Therefore, we concluded that a single origin of asexual reproduction was compatible with the data and this cluster was treated as one sexual and one asexual lineage in subsequent comparisons.

Table 2
Comparison of sexual and asexual clusters of *Eucypris virens* COI data from the 40 cluster solution (excluding singletons, doubletons and clusters with individuals from one site only for the spatial summary). p42 was included in the spatial summary. See Table 3S for equivalent results based on the 33 cluster solution.

	Sexual clusters	Asexual clusters	Comparison
Genetic diversity (π)			
Mean (SD)	0.0046 (0.0034)	0.0050 (0.0042)	$t = 0.22$, $p = 0.80$
(n)	(11)	(16)	
Range	0, 0.011	0.0008, 0.016	
Tajima's D			
Mean (SD)	−1.12[a] (0.40)	−0.4 (1.14)	$t = 2.28$, $p = 0.04$
(n)	(9)	(15)	
Range	−1.88, 0.61	−1.86, 1.89	
Central latitude			
Mean (SD)	38.0 N (3.15)	45.3 N (5.44)	$t = 4.50$, $p << 0.01$
(n)	(9)	(16)	
Dispersion (km)			
Mean (SD)	8.30 (9.61)	583 (303)	$t = 7.58$, $p << 0.01$
Range	0.04, 26.62	8.8, 1027	

[a] This mean is significantly different from zero.

Genetic diversity within clusters was similar, on average, for sexual and asexual clusters (Table 2). Average Tajima's D was significantly negative for sexual clusters, and remarkably consistent, whereas D was highly variable for asexual clusters. These results were similar for the 33 cluster solution based on the larger, 406 sequence data set (mean $D_{asexual} = -0.43$; mean $D_{sexual} = -1.05$, see Table 3S). Sexual clusters were centred around the Mediterranean whereas the geographical centres of asexuals clusters tended to be further north and were more variable, as expected from the previously described pattern of geographical parthenogenesis (Horne and Martens, 1999). Sexual clusters were narrowly distributed. Three out of 19 sexual clusters were, in fact, collected from only a single site and the most widely distributed sexual cluster occurred in seven sites, separated by <20 km. By contrast, several asexual clusters were very widely distributed. These genetic and spatial patterns will be analysed in more detail elsewhere.

Assuming sexuality to be the ancestral mode of reproduction in *E. virens*, the COI tree implies 18 transition events from sexual to asexual reproduction followed by 6 branching events to generate the 24 asexual clusters included in our sample. The different topologies and samples of the 16S and combined trees suggest 13 or 11 transitions to asexual reproduction and 5 or 4 branching events, resulting in 18 or 17 sampled asexual lineages. Nodes linking asexual lineages to their most closely related sexual relatives range in depth from the very recent event within cluster p33 to a genetic distance of 21% at the node connecting asexual clusters p37–p40 to the basal node of the ingroup.

4. Discussion

4.1. Deep phylogenetic divergence between sexual lineages and implications for intra-specific systematics

Our results reveal exceptionally high mitochondrial genetic variation within the morphospecies *Eucypris virens*. This is in line with

the extensive allozyme diversity reported previously (Rossi et al., 1998, 2008) and with an earlier analysis of mitochondrial and nuclear (ITS) sequence data (Schön et al., 2000) both of which were based on relatively restricted sampling of the species' distribution. Not only is diversity high but it is strongly structured into distinct groups of related sequences, separated by long internal branches. We found no evidence for nuclear insertions of mitochondrial sequences and, therefore, we have no doubt about the authenticity of this surprisingly high divergence.

Considering first the 18 sexual clusters (including p33 and one cluster identified only in the 16S data set, p42), most are restricted to one or a few sites and are allopatric with respect to other sexual clusters. The sexual clusters tend to be concentrated around the Mediterranean. They have restricted distributions, suggesting low dispersal and/or colonisation ability, and yet there is substantial genetic variation within clusters (Table 2), indicating that effective population sizes are large, probably in part due to banks of resting eggs (Brendonck and De Meester, 2003). Consistently negative Tajima's *D* values, and the low p_2 parameter estimate, might be due to population expansion but this could be partly an artefact of the Pons et al. (2006) clustering method: if deeply branched mitochondrial lineages within a population tend to be allocated to different clusters, this will shift D towards negative values. However, the overall patterns remained the same when two abundant haplotypes were reduced to one representative per region resulting in 33 clusters (see Table 3S).

Closely related clusters p5, p6 and p7 have overlapping distributions in Morocco, even occurring in the same pond, and their divergence did not satisfy the 4× rule of Birky et al. (2005). They were combined in some versions of the cluster analysis. Clusters p13, p14 and p15 occur in the Iberian Peninsula but have not been found together. These two groups are separated by COI genetic distances of less than 3% and the clustering may be interpreted as strong spatial population structure within a species, as it is in the alternative analyses where these clades form a single cluster. All other sexual clusters are separated by much greater genetic distances for COI (5–18%, Fig. 2). They are likely to have been isolated for long periods and may be considered separate species. However, there is no generally agreed threshold genetic distance for delimitation of allopatric species. A p-distance of 3% COI has been advocated for many taxa (Hebert et al., 2003), and this is surpassed by many pairs of *E. virens* sexual clusters, but the discontinuity between intra- and inter-specific divergence on which this threshold is based is unlikely to be universal (see, for example, Elias et al., 2007) and deep mitochondrial structure does not preclude nuclear gene flow (e.g. King et al. 2008). Divergence levels for 16S have confirmed the deep separation among *E. virens* sexual lineages.

These findings suggest that revision of the species status of *E. virens* is required. We argue that the large genetic distances between sexual lineages, and the statistical support of the cluster analysis (Pons et al., 2006), justify giving the clusters provisional species status. However, further analysis of nuclear markers will be required to confirm lack of gene exchange between sexual clusters and uncertainty around the cut-off generated by the cluster analysis will have to be reduced before their status can be confirmed. The status of the sexual clusters under the Biological Species Concept is likely to remain uncertain if they are genuinely allopatric, as our current samples suggest. We propose to refer to them by their 'p' numbers until further description is completed. This means that the morphospecies *E. virens* comprises an extensive species complex with at least 14 extant sexual members (if p5–7 and p13–15 are considered intra-specific groups). The rarefaction analysis strongly suggests that further sampling will reveal additional lineages, with a likely total of >30 sexual clusters present. In addition to further sampling, morphological studies, mating experiments and investigation of additional genetic markers are clearly important to establish the status of these putative cryptic species (these studies are in hand and will be reported elsewhere). For the present, it is essential to take account of this diversity of sexual lineages in considering the origin and nature of asexual diversity in *E. virens*.

Hidden biological diversity, in the form of cryptic species, appears to be widespread, especially in less extensively studied and morphologically conservative invertebrate taxa (Hebert et al., 2004; also see Saez and Lozano, 2005, for a review). However, situations similar to the extreme level of cryptic diversity observed here remain rare, one example being the diversity discovered within one species of the sexually reproducing amphipod genus, *Hyalella,* in North America (Witt et al., 2006). As our results and others (Witt et al., 2006; Pfenninger and Schwenk, 2007) indicate, much more cryptic diversity is yet to be found. We are in the middle of an unprecedented global freshwater crisis due to climatic changes and detrimental anthropogenic activities. One of the many consequences is an enormous loss in biodiversity (Dudgeon et al., 2006). The North American freshwater fauna, for example, is predicted to go extinct at a rate of 4% per decade (Ricciardi and Rasmussen, 1999). This model was based on 'classic' species, so if several of these harbour cryptic species, the loss of biodiversity will be even more dramatic.

4.2. Multiple transition events from sexual to asexual reproductive mode

Our cluster analysis identified 25 asexual lineages (23 in the most conservative cluster analysis), including one within cluster p33 and one identified only in the 16S data set (p41). All phylogenetic reconstructions imply numerous transition events from sexual to asexual reproductive mode (Figs. 2 and 5) although the precise number depends on the topologies of the deeper parts of the trees, which are not confidently resolved. We have assumed that the ancestral state was sexual reproduction and that transitions to asexual reproduction are overwhelmingly more likely than reversals. This assumption is based on the idea that mutations in loci essential for sexual reproduction tend to accumulate in asexual lineages making the re-acquisition of sex increasingly difficult. The assumption has rarely been challenged although a possible example of a reversal from asexual to sexual reproduction was recently reported in oribatid mites based on indirect evidence from phylogenetic reconstructions (Domes et al., 2007). Asexual reproduction in ostracods is believed to be apomictic and asexual lineages may be polyploid (Butlin et al., 1999). There is evidence for a hybrid origin of asexual lineages in other non-marine ostracods (Turgeon and Hebert, 1994; Chaplin and Hebert, 1997), and this may also occur in *E. virens* (Schön et al., 2000), but, in general, the mechanisms underlying the transitions between reproductive modes are not known. This makes it difficult to rule out reversals to sexual reproduction in young asexual lineages. However, allowance for reversals would not alter our basic conclusion that extant asexual lineages of *E. virens* are the result of many different origins.

Rarefaction analysis (Table 1) suggests that the true number of asexual lineages in Europe is higher than the number observed in our samples, despite the large sample size and wide distribution of sample sites (Fig. 1). This is true despite the fact that many of the observed asexual lineages have wide geographic distributions and multiple lineages co-occur within some sites. The wide distribution of asexual lineages suggests that they have much greater colonising ability than sexual lineages, presumably because a single individual is able to establish a new population, whereas in sexual lineages, males and females have to find each other in spatially and temporally highly dispersed environments. It is surprising, therefore, that asexual lineages generally show less evidence of population expansion than sexual lineages, although some do have strongly negative Tajima's *D* indices (Table 2). Diversity within

asexual lineages was similar to diversity within sexual lineages, despite the fact that their larger ranges lead to an expectation of larger census population sizes. This could be partly an artefact of the clustering method which currently fits a common transition point to all lineages whereas, in reality, one expects deeper coalescence in lineages with larger populations (Pons et al., 2006). However, it may be caused by background selection and/or selective sweeps which are expected to reduce the effective size of an asexual lineage relative to its census size. Alternatively, limited diversity may suggest that the asexual lineages are young and have not yet accumulated the genetic diversity that would be expected from their population size.

We have tentatively interpreted most sexual clusters identified by the method of Pons et al. (2006) as sexual species. This is not without problems because the criterion of intrinsic reproductive isolation cannot be tested satisfactorily for allopatric populations. In such cases, it is common to use other criteria, especially evidence of genetic independence such as reciprocal monophyly (Avise and Ball, 1990; Hey, 2009). These criteria may also be applied to asexual organisms which show cohesive groups as a result of shared ancestry and shared selective constraints (Templeton, 1989). On this basis, Barraclough et al. (2003) and Birky et al. (2005) have argued that asexual populations that coalesce independently are, in some sense, equivalent to sexual species (see Barraclough et al., 2009 for further discussion). Clusters identified by the 4X rule or by the Pons et al. (2006) approach have been interpreted as 'asexual species' in bdelloid rotifers (Birky et al., 2005; Fontaneto et al., 2007). In *Eucypris virens*, all sister pairs of asexual clusters identified by the Pons et al. (2006) method meet the 4× rule, except p16 and p17. On this basis, the morphospecies *E. virens* is an even more impressive species complex with at least 40 cryptic species and potentially as many as 60 or more (Table 1). However, it remains to be seen how this picture will be modified by analysis of nuclear loci, which are expected to have deeper coalescence.

It is difficult to infer the timing of transitions from sexual to asexual reproduction (Butlin, 2002). For any one lineage, the transition may have occurred at any point between the most recent common ancestor (MRCA) of individuals within an asexual cluster and the last node connecting the asexual lineage to an extant sexual lineage. For *E. virens*, the MRCAs of asexual lineages range in depth between 0.16% and 1.35% while the nodes connecting them to sexual lineages range in depth from 1% (within cluster p33) to 20% COI divergence (4% 16S divergence; the most basal node within *E. virens*). We have no calibration for rates of substitution in *Eucypris* but application of the general invertebrate rate of about 2% divergence per million years (Brower, 1994) which is similar to a crustacean rate (Knowlton, 1993) emphasises how wide a range of time-scales is implied by these figures:$<10^4$ years to 10^7 years.

The asexual lineage that branches at the base of the *E. virens* phylogeny (clusters p37–p40) was sampled exclusively in Eastern Europe and Turkey. Because we have not yet found sexual lineages east of the island of Corfu (Greece), there is a real possibility that undetected sexual lineages exist which are more closely related to these asexual clusters. In general, the inference from rarefaction analysis is that many lineages remain to be discovered. Our sampling may have missed more narrowly distributed sexual lineages than widespread asexual lineages, although the rarefaction analysis suggests that lineages of both reproductive modes remain undiscovered. The probable existence of unsampled sexual lineages suggests that the nodes connecting asexual lineages to sampled sexual lineages overestimate the age of the transition in reproductive mode. It is also very likely that both sexual and asexual lineages have gone extinct over the long time-scale of the *E. virens* phylogeny. Asexual lineages derived from extinct sexual lineages will appear to have origins at the time of speciation of the extinct sexual lineage, older than the true transition in reproductive mode. Given these considerations, and the observation that some asexual lineages are clearly young (as in cluster p33), our phylogenetic analyses do not provide evidence for long-term persistence of asexually reproducing lineages.

The genetic diversity of asexual lineages in *E. virens* appears to be maintained by frequent origin of new asexual lineages, combined with high dispersal and colonisation abilities which allow clones to spread and result in large effective population sizes. Large population size can reduce the turn-over rate of clones, even without changes in fitness (Janko et al., 2008). High dispersal and colonisation ability in the patchy and ephemeral environment of temporary pools may also allow clones to escape from co-evolving parasites (Ladle et al., 1993). But probably the greatest paradox in such a system is the persistence of narrowly distributed sexual populations within the range of the widespread and genetically diverse asexual assemblage.

Acknowledgments

This work was funded by the EU Marie Curie Research Training Network 'SexAsex' (from Sex to Asex: a case study on interactions between sexual and asexual reproduction, contract MRTN-CT-2004-512492). K.M. and I.S. also acknowledge the financial contribution of projects 1.5.172.09 (krediet aan navorsers) and G.0118.03N (projectonderzoek) of the FWO Vlaanderen (Fund for Scientific Research, Flanders). We are very grateful to all members of the network for their practical contributions to the work presented here and for many stimulating discussions. We wish to thank Tim Barraclough and Michael T. Monaghan for providing the R scripts used for performing the cluster analysis, and also Albert Phillimore for providing a useful R script for tree editing. We are also much indebted to Metin Sezer for his help with sampling in Turkey as well as to Tamara Karan-Žnidaršič for providing samples from Serbia. We thank Emilie R.D. Bautz for primer design and sequencing as well as Bruno Nevado and Zohra Elouaazizi for help with sampling and in the laboratory. We also thank Vincent Zintzen for his help with the extrapolation of the rarefaction analysis and David J. Horne for managing our network database. This paper has benefitted from the helpful comments of Joachim Mergeay and C. William Birky.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.08.022.

References

Adolfsson, S., Michalakis, Y., Paczesniak, D., Bode, S.N.S., Butlin, R.K., Lamatsch, D.K., Martins, M.J.F., Schmit, O., Vandekerkhove, J., Jokela, J., submitted for publication. Evaluation of elevated ploidy and asexual reproduction as alternative explanations for geographic parthenogenesis in Eucypris virens ostracods. Evol. Int. J. Org. Evol.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Avise, J.C., Ball Jr., R.M., 1990. Principles of genealogical concordance in species concepts and biological taxonomy. In: Futuyama, D., Antonovics, J. (Eds.), Oxford Surveys in Evolutionary Biology. Oxford University Press, Oxford, pp. 45–67.

Baltanas, A., Alcorlo, P., Danielopol, D.L., 2002. Morphological disparity in populations with and without sexual reproduction: a case study in *Eucypris virens* (Crustacea: Ostracoda). Biol. J. Linn. Soc. 75, 9–19.

Barraclough, T.G., Birky Jr., C.W., Burt, A., 2003. Diversification in sexual and asexual organisms. Evolution 57, 2166–2172.

Barraclough, T.G., Fontaneto, D., Herniou, E.A., Ricci, C., 2009. The evolutionary nature of diversification in sexuals and asexuals. In: Butlin, R., Bridle, J., Schluter, D. (Eds.), Speciation and Patterns of Diversity. Cambridge University Press, United Kingdom, Cambridge, pp. 29–45.

Bell, G., 1982. The Masterpiece of Nature: The Evolution and Genetics of Sexuality. Croom Helm, London.
Bensasson, D., Zhang, D.X., Hartl, D.L., Hewitt, G.M., 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol. Evol. 16, 314–321.
Bensasson, D., Zhang, D.X., Hewitt, G.M., 2000. Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. Mol. Biol. Evol. 17, 406–415.
Birky, C. W., Barraclough, T. G., 2009. Asexual Speciation. In: I. Schön, K. Martens, P. J. Van Dijk (Eds.), Lost Sex. The Evolutionary Biology of Parthenogenesis. Springer, Dordrecht, in press.
Birky, C.W., Wolf, C., Maughan, H., Herbertson, L., Henry, E., 2005. Speciation and selection without sex. Hydrobiologia 546, 29–45.
Brendonck, L., De Meester, L., 2003. Egg banks in freshwater zooplankton: evolutionary and ecological archives in the sediment. Hydrobiologia 491, 65–84.
Brower, A.V.Z., 1994. Rapid morphological radiation and convergence among races of the butterfly Heliconius–Erato inferred from patterns of mitochondrial-DNA evolution. Proc. Natl. Acad. Sci. USA 91, 6491–6495.
Burt, A., 2000. Perspective: sex, recombination, and the efficacy of selection—Was Weismann right? Evolution 54, 337–351.
Butlin, R., 2002. The costs and benefits of sex: new insights from old asexual lineages. Nat. Rev. Genet. 3, 311–317.
Butlin, R.K., Bridle, J.R., Kawata, M., 2003. Genetics and the boundaries of species' distributions. In: Blackburn, T., Gaston, K.J. (Eds.), Macroecology: Concepts and Consequences. Blackwell Science, Oxford, pp. 274–295.
Butlin, R.K., Schön, I., Martens, K., 1999. Origin, age and diversity of clones – commentary. J. Evol. Biol. 12, 1020–1022.
Chao, A., 1987. Estimating the population size for capture–recapture data with unequal catchability. Biometrics 43, 783–791.
Chaplin, J.A., Hebert, P.D.N., 1997. Cyprinotus incongruens (Ostracoda): an ancient asexual? Mol. Ecol. 6, 155–168.
Colwell, R. K., 2005. Estimates: statistical estimation of species richness and shared species from samples. http://purl.oclc.org/estimates.
Colwell, R.K., Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. Philos. Trans. R. Soc. Lond., B 345, 101–118.
Delmotte, F., Leterme, N., Bonhomme, J., Rispe, C., Simon, J.C., 2001. Multiple routes to asexuality in an aphid species. Proc. R. Soc. Lond., B 268, 2291–2299.
DeSalle, R., Freedman, T., Prager, E.M., Wilson, A.C., 1987. Tempo and mode of sequence evolution in mitochondrial-DNA of Hawaiian *Drosophila*. J. Mol. Evol. 26, 157–164.
Domes, K., Norton, R.A., Maraun, M., Scheu, S., 2007. Reevolution of sexuality breaks Dollo's law. Proc. Natl. Acad. Sci. USA 104, 7139–7144.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z., Knowler, D.J., Leveque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L., Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biol. Rev. Camb. Philos. Soc. 81, 163–182.
Elias, M., Hill, R.I., Willmott, K.R., Dasmahapatra, K.K., Brower, A.V.Z., Malllet, J., Jiggins, C.D., 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proc. R. Soc. Lond., B 274, 2881–2889.
Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.
Felsenstein, J., Churchill, G.A., 1996. A hidden Markov Model approach to variation among sites in rate of evolution. Mol. Biol. Evol. 13, 93–104.
Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3, 294–299.
Fontaneto, D., Herniou, E.A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., Barraclough, T.G., 2007. Independently evolving species in asexual bdelloid rotifers. PLoS Biol. 5, 914–921.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
Hamilton, W.D., 1980. Sex versus non-sex versus parasites. Oikos 35, 282–290.
Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond., B 270, 313–321.
Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proc. Natl. Acad. Sci. USA 101, 14812–14817.
Heethoff, M., Domes, K., Laumann, M., Maraun, M., Norton, R.A., Scheu, S., 2007. High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite Platynothrus peltifer (Acari, Oribatida). J. Evol. Biol. 20, 392–402.
Hey, J., 2009. On the arbitrary identification of real species. In: Butlin, R., Bridle, J., Schluter, D. (Eds.), Speciation and Patterns of Diversity. Cambridge University Press, United Kingdom, Cambridge, pp. 15–28.
Hijmans, R.J., Guarino, L., Cruz, M., Rojas, E., 2001. Computer tools for spatial analysis of plant genetic genetic resources data: 1. DIVA-GIS. Plant Gen. Resour. Newsl. 127, 15–19.
Horne, D.J., Baltanás, A., Paris, G., 1998. Geographical distribution of reproductive modes in living non-marine ostracods. In: Martens, K. (Ed.), Sex and Parthenogenesis: Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods. Blackhuys Publishers, Leiden, Netherlands, pp. 77–99.
Horne, D.J., Martens, K., 1999. Geographical parthenogenesis in European non-marine ostracods: post-glacial invasion or Holocene stability? Hydrobiologia 391, 1–7.
Janko, K., Drozd, P., Flegr, J., Pannell, J.R., 2008. Clonal turnover versus clonal decay: a null model for observed patterns of asexual longevity, diversity and distribution. Evolution 62, 1264–1270.
Judson, O.P., Normark, B.B., 1996. Ancient asexual scandals. Trends Ecol. Evol. 11, 41–46.
King, R. A., Tibble, A. L., Symondson, W. O. C., 2008. Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. 17, 4684-4698.
Knowlton, N., 1993. Sibling species in the Sea. Annu. Rev. Ecol. Syst. 24, 189–216.
Kondrashov, A.S., 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336, 435–440.
Kondrashov, A.S., 1993. Classification of hypotheses on the advantage of amphimixis. J. Hered. 84, 372–387.
Ladle, R.J., Johnstone, R.A., Judson, O.P., 1993. Coevolutionary dynamics of sex in a metapopulation - escaping the Red Queen. Proc. R. Soc. Lond., B 253, 155–160.
Laumann, M., Norton, R.A., Weigmann, G., Scheu, S., Maraun, M., Heethoff, M., 2007. Speciation in the parthenogenetic oribatid mite genus Tectocepheus (Acari, Oribatida) as indicated by molecular phylogeny. Pedobiologia 51, 111–122.
Law, J.H., Crespi, B.J., 2002. The evolution of geographic parthenogenesis in Timema walking-sticks. Mol. Ecol. 11, 1471–1489.
Lively, C.M., Dybdahl, M.F., Jokela, J., Osnas, E.E., Delph, L.F., 2004. Host sex and local adaptation by parasites in a snail-trematode interaction. Am. Nat. 164, S6–S18.
Loxdale, H.D., Lushai, G., 2003. Rapid changes in clonal lines: the death of a 'sacred cow'. Biol. J. Linn. Soc. 79, 3–16.
Lunt, D.H., Zhang, D.X., Szymura, J.M., Hewitt, G.M., 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. Insect Mol. Biol. 5, 153–165.
Lynch, M., Burger, R., Butcher, D., Gabriel, W., 1993. The mutational meltdown in asexual populations. J. Hered. 84, 339–344.
Maraun, M., Scheu, S., 2000. The structure of oribatid mite communities (Acari, Oribatida): patterns, mechanisms and implications for future research. Ecography 23, 374–383.
Martens, K., 1998. General morphology of non-marine ostracods. In: Martens, K. (Ed.), Sex and Parthenogenesis: Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods. Blackhuys Publishers, Leiden, Netherlands, pp. 57–75.
Martens, K., Rossetti, G., Horne, D.J., 2003. How ancient are ancient asexuals? Proc. Biol. Sci. 270, 723–729.
Martens, K., Schön, I., 2008. Ancient asexuals: darwinulids not exposed. Nature 453, 587.
Martens, K., Schön, I., Meisch, C., Horne, D.J., 2008. Global diversity of ostracods (Ostracoda, Crustacea) in freshwater. Hydrobiologia 595, 185–193.
Mateos, M., Vrijenhoek, R.C., 2002. Ancient versus reticulate origin of a hemiclonal lineage. Evolution 56, 985–992.
Maynard Smith, J., 1978. The evolution of sex. Cambridge Univ. Press.
Meisch, C., 2000. Freshwater Ostracoda of Western and Central Europe. Süsswasserfauna von Mitteleuropa 8/3. Spektrum Akademischer Verlag, Heidelberg, Berlin.
Neiman, A., Jokela, J., Lively, C.M., 2005. Variation in asexual lineage age in Potamopyrgus antipodarum, a New Zealand snail. Evolution 59, 1945–1952.
Paland, S., Colbourne, J.K., Lynch, M., 2005. Evolutionary history of contagious asexuality in Daphnia pulex. Evolution 59, 800–813.
Palumbi, S.R., 1996. Nucleic acids, II: the polymerase chain reaction. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, Mass, pp. 205–247.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.
Pfenninger, M., Schwenk, K., 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol. Biol. 7.
Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
R Development Core Team, 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Rambaut, A., Drummond, A. J., 2007. Tracer v1.4. http://beast.bio.ed.ac.uk/Tracer.
Ricciardi, A., Rasmussen, J.B., 1999. Extinction rates of North American freshwater Fauna. Conserv. Biol. 13, 1220–1222.
Rossi, V., Piotti, A., Baltanas, A., Benassi, G., Menozzi, P., 2008. Genetic diversity and mixed reproduction in *Eucypris virens* (Crustacea: Ostracoda). Fund. Appl. Limn. 172, 147–159.
Rossi, V., Schön, I., Butlin, R., Menozzi, P., 1998. Clonal genetic diversity. In: Martens, K. (Ed.), Sex and Parthenogenesis Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods. Backhuys Publishers, Leiden, Netherlands, pp. 257–274.
Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Meth. Mol. Biol. 132, 365–386.
Saez, A.G., Lozano, E., 2005. Body doubles. Nature 433, 111.
Schön, I., 2007. Did Pleistocene glaciations shape genetic patterns of European ostracods? A phylogeographic analysis of two species with asexual reproduction. Hydrobiologia 575, 33–50.
Schön, I., Gandolfi, A., Di Masso, E., Rossi, V., Griffiths, H.I., Martens, K., Butlin, R.K., 2000. Persistence of asexuality through mixed reproduction in *Eucypris virens* (Crustacea, Ostracoda). Heredity 84, 161–169.

Schön, I., Martens, K., Van Dijk, P. (Eds.), 2009. Lost sex. The Evolutionary Biology of Parthenogenesis. Springer, Dordrecht, in press.
Smith, R. J., Kamiya, T., Horne, D. J., 2006. Living males of the 'ancient asexual' Darwinulidae (Ostracoda: Crustacea). 273, 1569-1578.
Smith, E.P., van Belle, G., 1984. Nonparametric estimation of species richness. Biometrics 40, 119–129.
Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. USA.
Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol. Biol. Evol. 18, 1001–1013.
Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Sinauer Associates, Sunderland, Massachusetts.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
Templeton, A.R., 1989. The meaning of species and speciation: a genetic perspective. In: Otte, D., Endler, J.A. (Eds.), Speciation and Its Consequences. Sinauer Associates Inc., Massachusetts, Sunderland, pp. 3–27.
Turgeon, J., Hebert, P.D.N., 1994. Evolutionary interactions between sexual and all-female Taxa of Cyprinotus (Ostracoda, Cyprididae). Evolution 48, 1855–1865.
van Dijk, P.J., 2003. Ecological and evolutionary opportunities of apomixis: insights from Taraxacum and Chondrilla. Philos. Trans. R. Soc. Lond., B 358, 1113–1120.
Van Valen, L.M., 1973. A new evolutionary law. Evol. Theory 1, 1–30.
Vandel, A., 1928. La parthénogenèse géographique. Contribution à l'étude biologique et cytologique de la parthénogenèse naturelle. Bull. Biol. France Belg. 62, 164–281.
Welch, D.B.M., Meselson, M.S., 2001. Rates of nucleotide substitution in sexual and anciently asexual rotifers. Proc. Natl. Acad. Sci. USA 98, 6720–6724.
Welch, D.M., Meselson, M., 2000. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Science 288, 1211–1215.
Welch, J.L.M., Welch, D.B.M., Meselson, M., 2004. Cytogenetic evidence for asexual evolution of bdelloid rotifers. Proc. Natl. Acad. Sci. USA 101, 1618–1621.
West, S.A., Lively, C.M., Read, A.F., 1999. A pluralist approach to sex and recombination. J. Evol. Biol. 12, 1003–1012.
Witt, J.D.S., Threloff, D.L., Hebert, P.D.N., 2006. DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. Mol. Ecol. 15, 3073–3082.