The EMBO Journal vol.10 no.3 pp.499-503, 1991

An analysis of the origin of metazoans, using comparisons of partial sequences of the 28S RNA, reveals an early emergence of triploblasts

R.Christen, A.Ratto, A.Baroin[1], R.Perasso[1], K.G.Grell[2] and A.Adoutte

URA 671 CNRS, Station Zoologique, 06230 Villefranche sur mer, [1]Laboratoire de Biologie Cellulaire 4, URA 1134 CNRS, Bâtiment 444, Université Paris XI, 91405 Orsay Cedex, France and [2]Institut für Biologie, Universität Tübingen, D7400 Tübingen, Germany

Communicated by P.Slonimski

In order to study the origin of metazoans, we have compared sequences from the 5′ end of the large subunit ribosomal RNA of a number of protists, fungi, plants and metazoans, including all diploblastic phyla (sequences of 10 new species have been determined, including that of the placozoan, *Trichoplax adhaerens*). These sequences were analyzed using distance matrix, maximum parsimony and maximum likelihood methods, and the validity of the results was ascertained with bootstrapping and species removal or addition. Triploblasts and diploblasts formed two clearly separated monophyletic units; this divergence, which apparently preceded the diversification of diploblastic animals (i.e. the successive sponge, ctenophore, cnidarian radiations), showed a much more ancient origin of triploblasts with respect to diploblasts than classically assumed. These results do not exclude the possibility that triploblasts and diploblasts arose independently from different protists.

Key words: evolution/metazoans/protists/ribosomal RNA

Introduction

On the basis of long and thorough studies of comparative morphology, metazoans (animals) have been divided into three major taxonomic units, namely: animals with three embryonic layers (triploblasts), animals with two embryonic layers (diploblasts) and animals with extremely loose tissue differentiation (sponges), the latter being sometimes included within diploblasts.

These taxonomic units have received general agreement; in contrast, the phylogeny of metazoans, that is the reconstruction of the evolutionary pathways followed from unicellular organisms to these diverse body plans, remains highly controversial. In particular, analyses of the origin of metazoans using methods of 'molecular phylogeny' and ribosomal RNA (rRNA) as a phylogenetic index have recently revived the controversy as to whether all metazoans can be viewed as successive offshoots within a single monophyletic unit, or if different processes of cell aggregation led to parallel radiations and different body plan organizations (i.e. sponges, diploblasts and triploblasts). From a set of molecular data (Field *et al.*, 1988) that comprised partial sequences of 18S rRNA from a number of protists, several triploblastic phyla and one diploblastic phylum (cnidarians), several mathematical analyses have been derived, albeit with conflicting interpretations (Field *et al.*, 1988; Gouy and Li, 1989; Patterson, 1989; Lake, 1990).

In this study, we have reinvestigated this problem by bringing two new and important experimental data sets. First, we have sequenced another molecule, 28S rRNA, in a domain that had been described as particularly well suited for deep phylogenetic analyses (Qu *et al.*, 1988). Second, we included several representatives of each diploblastic phylum (sponges, cnidarians, ctenophores and placozoa). Because it comprised all diploblastic phyla, our study revealed relationships between diploblasts and triploblasts that were strikingly different from the modern view as described in most zoological textbooks, but that rejoined older theories which had recognized triploblastic structure as extremely ancient. Also, our data suggest that the investigation of the problem of metazoan emergence might be particularly difficult because it occurred at a period of intense phyletic diversification.

Results and Discussion

We have sequenced the 5′ end of the 28S rRNA (~450 nucleotides) of ten organisms (Figure 1) belonging to all diploblastic phyla (sponges, cnidarians, ctenophores and placozoans) and analyzed these sequences by comparison with a large database comprising protists, fungi, triploblasts and metaphytes (Baroin *et al.*, 1988; Perasso *et al.*, 1989). In a broad phylogeny of eukaryotes (Figure 2), and in agreement with previous studies (Sogin *et al.*, 1986; Baroin *et al.*, 1988; Perasso *et al.*, 1989), the earliest emerging species were flagellates and a slime mold; later radiations comprised an amoeba, ciliates, a dinoflagellate, a heliozoan, fungi and chromophytes. Metazoa emerged from the tree slightly later in the form of two distinct branches, respectively triploblasts and diploblasts, with green algae, metaphytes and a few additional groups intermingled at their basis. This trichotomy, leading respectively to diploblasts, triploblasts and metaphytes, is in general agreement with traditional views which assign independent origins to metaphytes and metazoa, but is surprising with respect to the depth of the split between diploblasts and triploblasts and their lack of a common stem, since triploblasts are generally assumed to be derived from diploblasts.

To examine this point in more detail, we restricted our analysis to the cluster that contained the multicellular organisms. This approach was useful because with sequences belonging to more closely related organisms, multiple substitutions were avoided and the information/noise ratio was improved (in particular, parallel substitutions were reduced). Independently of the method used to construct the trees, three main sytems were detected, namely plants, triploblasts and diploblasts, separated from each other by a few radiations of protists (see Figure 3A and 3B), confirming the results of the global tree (Figure 2). No protist could be consistently related with the metazoan radiations, in

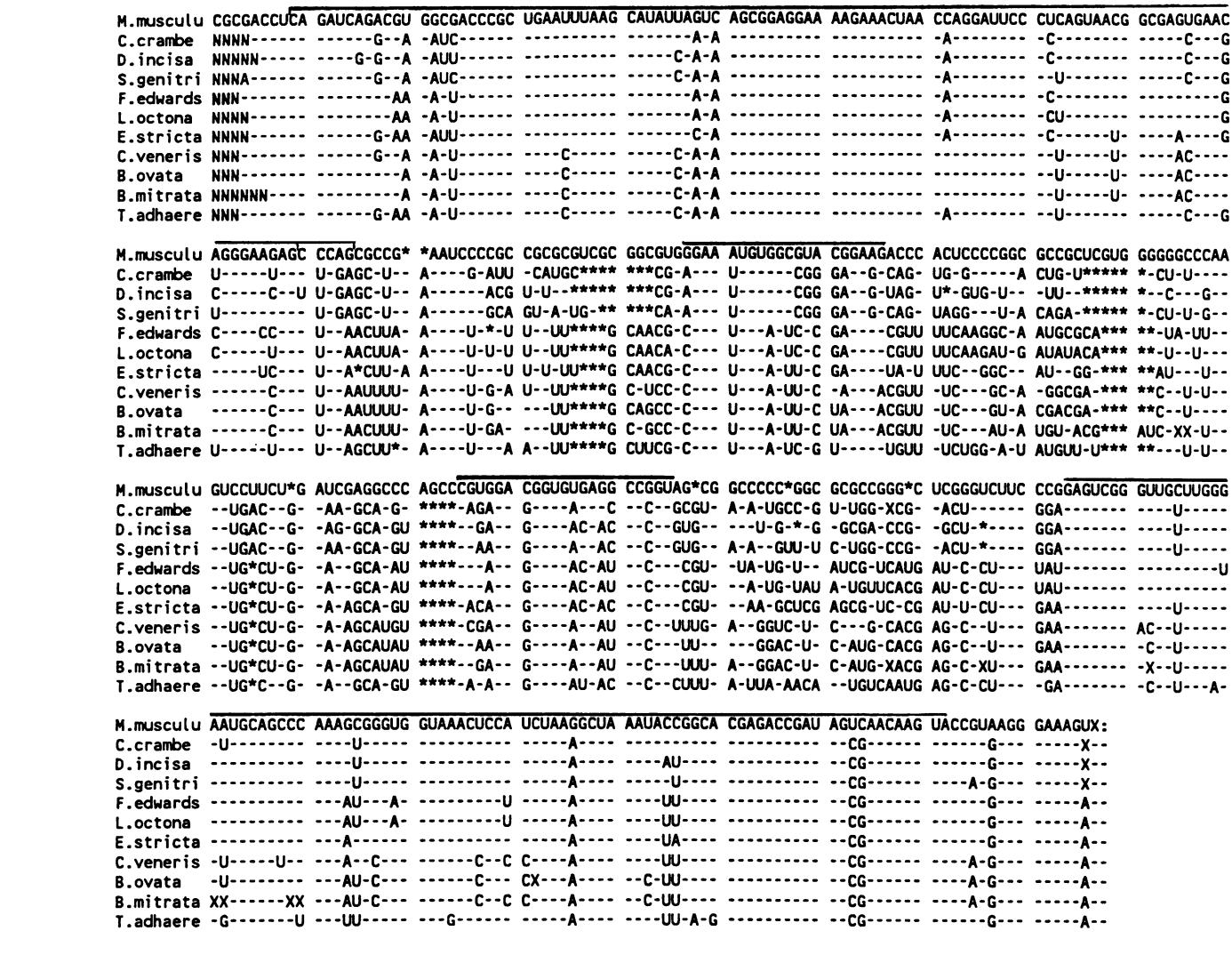

Fig. 1. Sequences of diploblastic metazoans used in the paper. Only nucleotides that differ from those of the mouse are indicated (identities are denoted by hyphens, deletions by stars and nucleotide positions that could not be identified by 'X'). The portions of sequences used for the tree shown in Figure 2 are overlined. Other sequences of protists, plants and triploblastic metazoans were previously published.

contrast to plants which, as expected, were closely associated with chlorophyte algae. It is noteworthy that independently of the method used, sponges, cnidarians and ctenophores respectively clustered as distinct units, testifying of the reliability and robustness of the analysis. *Trichoplax adhaerens* clearly belonged to the diploblastic radiation. Topologies such as those shown in Figure 4B and 4C, but not 4A, were thus derived from analyses (see Figures 2 and 3) using distance matrix (FITCH), parsimony (PAUP), and also more sophisticated but time consuming approaches such as maximum likelihood (DNAML), branch and bound options for parsimony (PAUP) and bootstrapping methods (Hendy and Penny, 1982; Felsenstein, 1985; Penny and Hendy, 1986; Felsenstein, 1988) to analyse up to 25 species within a single tree (see details in Materials and methods). These results have a bearing on two very broad biological questions: which protist gave rise to metazoans and which morphological pathway was followed throughout the early evolution of metazoans?

The emergence of metazoans has been explained by two major theories: the syncitial theory (from a multinucleated ciliate, see Hadzi, 1963; Hanson, 1977) or the colonial theory (from a colonial flagellate, see Haeckel, 1874; Hyman, 1940). Present molecular data (see Field *et al.*, 1988 and this work) do not allow clarification of this point. Diploblasts and triploblasts indeed always emerged close to the points of branching of ciliates and of some photosynthetic flagellates, but the precise branching orders of the various protists in this zone of the dendrograms could not be established with certainty, because the topologies of these branching points changed according to which metazoan species were selected to construct the tree. Maximum likelihood analyses also showed that no protist could be significantly associated with either the diploblasts or the triploblasts or both, in marked contrast to the constant monophyletic unit obtained between vascular plants and the two unicellular green algae *Chlorogonium elongatum* and *Pyramimonas parkeae*.

In a first analysis of partial 18S rRNA sequences (Field *et al.*, 1988), a biphyletic origin of metazoans was described, with the two cnidarians that had been sequenced forming a monophyletic radiation with a ciliate, a yeast and a metaphyte; in contrast, new analyses of the same data have led to the conclusions that diploblasts and triploblasts could (Patterson, 1989) or did (Lake, 1990) form a monophyletic unit. However, it is possible that the appropriate protists have

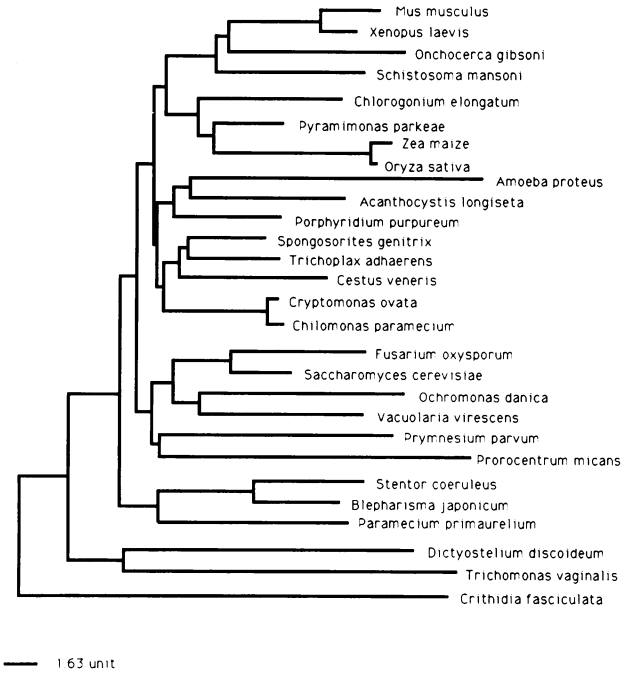

Fig. 2. General phylogenetic tree of eukaryotes. Our molecular index was the 5′ end of the RNA from the large ribosomal subunit which has a rate of mutation particularly well suited for phylogenetic analyses within the eurkaryotic kingdom (Qu *et al.*, 1983). Using a least square distance method ('Fitch' program from the Phylip package with the G global-rearrangement option), a tree was derived from aligned sequences of 25 eukaryotes. Distances were calculated using the K_{nuc} correction of Kimura (1980). Distances between two species are equal to the sums of the projections on the x axis (K_{nuc}), whereas lengths of the branches along the y axis are arbitrary (only the most conserved parts of the sequences were retained: positions 8–117, 149–168, 245–275 and 317–410, after alignment on the sequence of *M.musculus* as numbered in Perasso *et al.*, 1989). Scale bar: unit K_{nuc}. Identical topologies were obtained when other species of protists, plants and metazoans were selected, when subdomains of the sequences were used, and also when using different algorithms such as a dynamic clustering method to analyze distance matrices (Hénaut and Delorme, 1988) or following an analysis by parsimony (see also Figure 3).

not yet been found or that unicellular organisms which gave rise to metazoans did not survive to the present day; for these reasons and because of the early separation between triploblasts and diploblasts, any interpretation of known molecular data as proof for monophyly of metazoans could be challenged by the discovery of a group of protists branching between diploblasts and triploblasts. The many cases of cellular aggregations existing within the protists indeed suggest that multiple derivations of multicellularity are possible (Bütschli, 1910; Jägersten, 1955; Barnes *et al.*, 1988; Bonner, 1988). Finally, it is likely that we are facing a situation which is difficult to resolve for methodological reasons: both our data and those of Field *et al.* (1988) suggest that the lines leading to diploblasts and triploblasts originated at a period of intense diversification of protists. Indeed, the branching points of all the phyla corresponding to this radiation (belonging to both the metazoa and protists) are very close to each other on the trees and their topology is variable. If this diversification occurred during a relatively short period of time, reconstruction of the branching orders is expected to be difficult, whatever approach (morphological or molecular) is used.

Our analysis confirmed the classical distinction between diploblastic and triploblastic metazoans, but always showed triploblastic metazoans as a monophyletic unit with a well individualized stem, which separated early from the diploblasts (see Figure 3). In all analyses, the diversification of diploblasts into placozoa, sponges, cnidarians and ctenophores occurred after the split between diploblasts and triploblasts, a striking result not previously observed because only a single diploblastic phylum had been analyzed.

Concerning the pattern of early morphological evolution in metazoans, the prevailing view is that during the transition from unicellular eukaryotes to the highly complex living metazoans, life evolved from simple to more complex body plans, namely: from colony-like unicellular organisms through diploblastic body plan to triploblastic body plan, progressively yielding the ~35 extant animal phyla. Most theories follow the idea of a monophyletic origin of metazoans (with perhaps the exception of sponges) as for example the trochaea theory presented by Nielsen (1985) who proposed a phylogeny based mainly on the structure of larval ciliary bands and the nervous system of living animals, ctenophores being described as a rather late radiation. In such schemes, sponges could be an example of the type of organisms that arose primitively from choanoflagellate-like protists (Salvini-Plawen, 1978). Mesozoans, which have a simple morphology and a low number of differentiated tissues, have also been taken as an example of the kind of simple tissue organization that must have characterized the first metazoans. However, this view is now usually rejected since the morphological simplicity of mesozoans is thought to result from parasitism (Whittaker and Margulis, 1978). Finally, *Trichoplax adhaerens*, a small dorso–ventrally flattened animal ~1 mm in diameter, made of only four different cell types grouped into a continuous single epithelial layer and a loose mesenchymatous internal sheet, classified within its own phylum (the Placozoa, see Grell and Benwitz, 1971) and one of the most simple metazoans (Grell, 1971), could have retained a morphology characteristic of earlier forms. In contrast to these theories, molecular data seem to imply that an embryonic organization in three layers arose early in the history of metazoans. The complex features shared by diploblasts and triploblasts, such as constants in patterns of early embryology or for example the neuromuscular system, must either have been present in a common ancestor or independently derived. Considering the relatively short distance that separates the protozoan world from the triploblast–diploblast divergence, the first hypothesis implies a rapid evolution from unicellular organisms to complex structures, whereas an independent derivation would reveal a preexisting potential for differentiation into several specific cell types.

The difficulty of discovering relationships of metazoans to protists by molecular methods or of solving the puzzle of morphological evolution by classical comparative studies seems to result from rapid diversification processes, the study of which probably requires a combined analysis of morphological and molecular data. Despite these uncertainties, the present work indicates that ancestors of extant triploblasts emerged much earlier than usually assumed. This idea is not incompatible with data from the earliest known palaeontological faunas (i.e. Ediacara and Burgess Shale, see Cloud and Glaessner, 1982; Conway Morris, 1989) which contain fossils identified as triploblasts. As these earliest fossils are dated at ~700 million years, an estimate

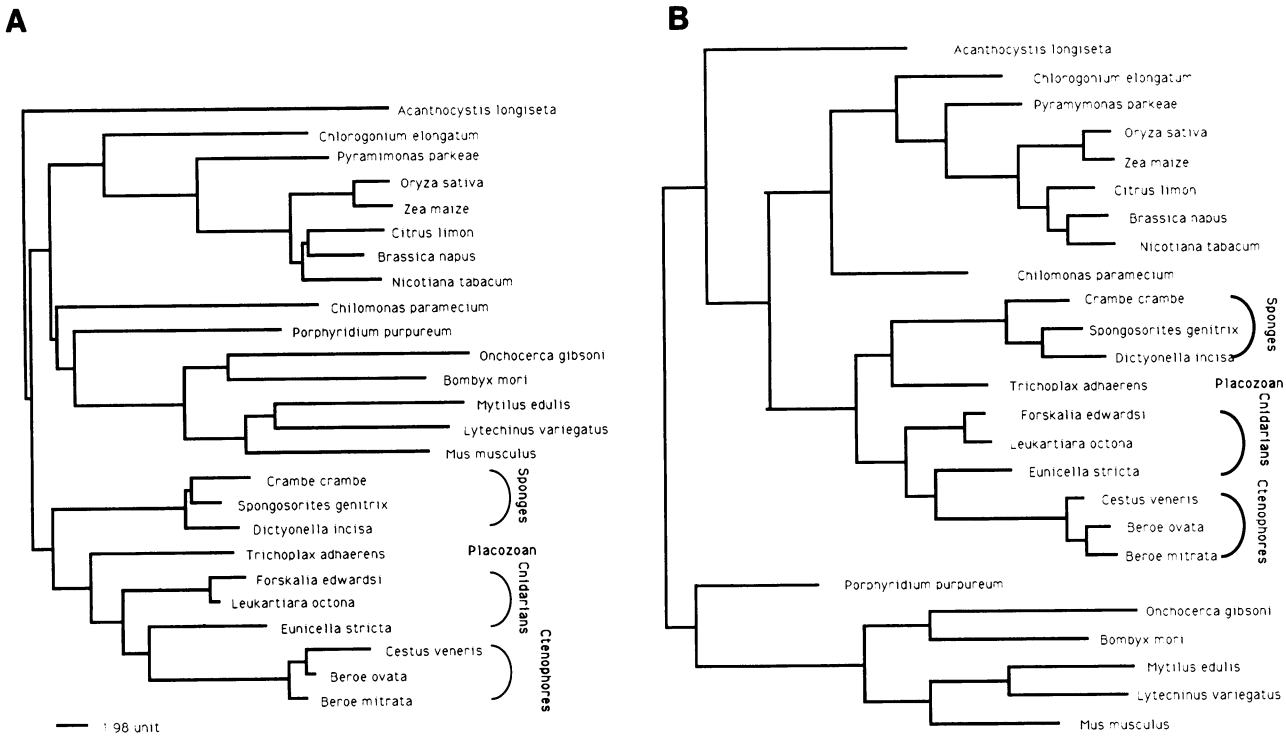

Fig. 3. Phylogenetic relationships between a restricted set of species. Sequences of eukaryotes belonging to the same monophyletic group (that comprised the metazoans) were analyzed using a distance matrix method ('Fitch', **A**) or the maximum parsimony program provided by Swofford ('PAUP', **B**), two methods widely used to reconstruct molecular phylogenies (Nei, 1987). Due to the restriction in the variety of species analyzed, additional domains could be aligned and used (positions 7–134, 144–171, 200–220, 224–250 and 278–372 of Figure 1). In the two analyses, plants, triploblastic metazoans and diploblastic metazoans were seen as three different monophyletic units, separated by the radiations of a few protists: *Chilomonas paramecium, Porphyridium purpureum* and *Acanthocystis longiseta.* Exact positions of the deepest nodes (i.e. relative branchings of plants, diploblasts, triploblasts and associated protists) varied slightly according to the method of analysis or the species selected and therefore cannot be considered as significant. A is represented as in Figure 2; in B, the lengths of the horizontal branches are proportional to the number of nucleotide substitutions required to explain the sequence differences (vertical branches are of arbitrary length). PAUP was used with the global, mulpars, maxtree = 25 and hold = 25 options. A single most parsimonious tree was obtained.

of 1–1.2 million years for an emergence of triploblastic ancestors is within reasonable estimates.

Materials and methods

Extraction of ribosomal RNA

Animals were collected in the Mediterranean sea, near the marine stations of Villefranche sur mer and Endoume, and were formally identified by C.Carré (cnidarians), C.Mills and M.-L.Nicaise (ctenophores) and J.Vacelet (sponges). The placozoan *Trichoplax* was grown in sea water using flagellates as food organisms. Starved animals were individually isolated and pelleted before extraction.

Total RNA was extracted from fresh tissues by homogenization in guanidinium thiocyanate (4 M guanidinium thiocyanate, 50 mM Tris–HCl pH 7.6, 4 mM EDTA, 2% N-laurylsarcosyl, 1% 2-mercaptoethanol). Proteins were removed by phenol–chloroform extractions (3×), followed by chloroform extractions (2×). Total RNA was then precipitated by ethanol addition and centrifugation and the pellet was resuspended in distilled water at 2 μg/μl.

Sequencing

Reverse transcriptions in the presence of dideoxynucleotides were primed with ^{32}P-labelled synthetic oligonucleotides to conduct a 'Sanger type' sequencing procedure but with reverse transcriptase in place of DNA polymerase (Qu *et al.*, 1983). The three primers used were complementary to evolutionarily conserved regions of the eukaryotic large subunit rRNA (Baroin *et al.*, 1988; Perasso *et al.*, 1989). Hybridization occurred efficiently with all metazoa analysed including *Trichoplax*. Since each primer hybridizes only to a single region and since they are within ~100–150 nucleotides of each other at the 5′ end of the molecule, they allow for overlapping sequences to be obtained rapidly.

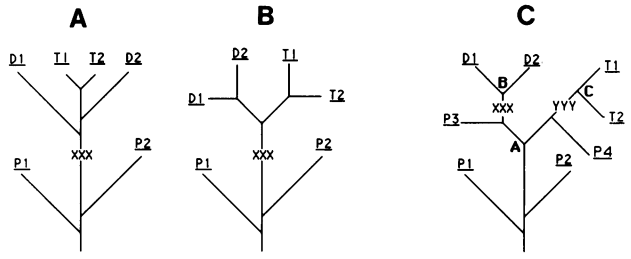

Fig. 4. Three different hypotheses for the early diversification of metazoans. **A**: 'Classical theory': within a eukaryotic world consisting of unicellular protistan phyla (P1 and P2), a single multicellularization event occurred (XXX), followed by the radiations of different diploblastic phyla (D1 and D2), one of which led to extant triploblastic phyla (T1 and T2). **B**: A single multicellularization event was followed by the early divergence of two main stems, one leading to the radiations of diploblasts, the other to the radiations of triploblasts. **C**: Two multicellularization events occurred (XXX and YYY), one protist led to diploblasts and a different one gave rise to the triploblasts. (Extant organisms are underlined).

Sequence comparisons

Alignment of sequences, computation of the observed as well as the corrected numbers of nucleotide differences (using Kimura's Knuc correction, see Kimura, 1980), and derivation of the resulting distance matrices were carried out as previously described (Baroin *et al.*, 1988; Perasso *et al.*, 1989). Dendrograms were obtained using several methods: ATD (M.-O.Delorme and A.Henaut); FITCH and DNAML (Phylip package, J.Felstenstein); PAUP (Swofford). 20–25 species were first selected among the whole data base in order to determine which protists could be associated most closely

with metazoans. Subanalyses were then run using these protists and species selected among the various phyla of metazoans. This process was repeated by changing the species selected, as well as the domains selected for the analysis, and all methods were used to derive topologies. A topology was considered significant when obtained by all procedures, significant ($P < 0.01$) with DNAML and shown > 95% by bootstrapping. For example, in order to examine the validity of our results which seemed to rule out the hypothesis of Figure 4A, we proceeded as follows: (i) to obtain the most parsimonious tree, we used the branch and bound option of PAUP. Due to the duration of the computation, this analysis was restricted to 12 species (three diploblasts, three triploblasts, two plants and four protists, with options hold = 25, maxtree = 50). A single most parsimonious tree was obtained, the topology of which was exactly that of Figure 3B. (ii) We examined by bootstrapping (using the Global option of the DNABOOT program from the Phylip package; 60 simulations with 20–50 replicates) the validity of the two monophyletic units, i.e. triploblasts and diploblasts. Analyses were repeated with various subsets of species (protists, diploblasts and triploblasts). In most analyses, diploblasts and triploblasts always formed separate monophyletic units; in a few cases (and in each case not exceeding at the most 5% of the replicates and detected only by invoking the R option in the 3.22 version of the program) a triploblast was occasionally seen nested within diploblasts. (iii) Finally, studies were effected using maximum likelihood analysis. Many simulations were first obtained with a low number of species, that preceded three simulations using the 25 species of Figure 3. The DNAML computer program (PHYLIP package) was recompiled for use with a 80387 arithmetic coprocessor and run on a Compaq 80386 at 20 MHz. Each analysis took >100 h of uninterrupted calculation and led to trees identical in topology to Figure 3A, with distances AB and AC (see Figure 4C) significantly positive ($P < 0.01$). However, the branching points of the protists which radiated between diploblasts and triploblasts (P3 and P4 of Figure 4C) could not be significantly positioned. Further analysis conducted with two invariant methods (after Lake or Cavender, see Felsenstein's Phylip package), led to non-significant topologies.

Acknowledgements

We thank our colleagues C.Carré and M.-L.Nicaise for their help in identifying cnidarians and ctenophores, J.Vacelet for collection and identification of the sponges as well as discussion of the results, J.Felsenstein for kindly providing us with his latest version of his Phylip package and discussions. H.Philippe computer programs helped greatly in dealing with our sequences. This work has been supported by the CNRS, the universities of Paris XI and Paris VI and a short term EMBO fellowship to A.Baroin. Sequences have been deposited with EMBL data bank and are available from the authors.

References

Barnes,R.S.K., Calow,P. and Olive,P.J.W. (1988) *The Invertebrates: a New Synthesis*. Blackwell Scientific Publications, Oxford.

Baroin,A., Perasso,R., Qu,L.H., Brugerolle,G., Bachellerie,J.-P. and Adoutte,A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3474–3478.

Bonner,J.T. (1988) *The Evolution of Complexity by Means of Natural Selection*. Princeton University Press, Princeton, NJ.

Bütschli,O. (1910) *Vorlesungen über vergleichende Anatomie*. 1. Lieferung. Verlag von Wilhelm. Engelman, Leipzig.

Cloud,P. and Glaessner,M.F. (1982) *Science*, **217**, 783–792.

Conway Morris,S. (1989) *Science*, **246**, 339–346.

Felsenstein,J. (1985) *Evolution*, **39**, 783–791.

Felsenstein,J. (1988) *Annu. Rev. Genet.*, **22**, 521–565.

Field,K.G., Olsen,G.J., Lane,D.J., Giovannoni,S.J., Ghiselin,M.T., Raff,E.C., Pace, N.R. and Raff,R.A. (1988) *Science*, **239**, 748–753.

Gouy,M. and Li,W.-H. (1989) *Mol. Biol. Evol.*, **6**, 109–122.

Grell,K.G. (1971) *Naturev. Rundschau*, **42**, 160–161.

Grell,K.G. and Benwitz,G. (1971) *Cytobiologie*, **4**, 216–240.

Hadzi,J. (1963) *The Evolution of Metazoa*. Pergamon Press, New York.

Haeckel,E. (1874) *Q. J. Microbiol. Sci.*, **14**, 142–165.

Hanson,E.D. (1977) *The Origin and Early Evolution of Animals*. Wesleyan University Press, Middletown, CT.

Hénaut,A. and Delorme,O. (1988) *Patterns Recog. Lett.*, **7**, 207–213.

Hendy,M.D. and Penny,D. (1982) *Math. Biosci.*, **59**, 277–290.

Hyman,L. (1940) *The Invertebrates: Protozoa through Ctenophora*, Vol. I. McGraw Hill, New York.

Jägersten,G. (1955) *Zoologiska Bidrag* (Uppsala), **30**, 321–354.

Kimura,M. (1980) *J. Mol. Evol.*, **16**, 111–120.

Lake,J.A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 763–766.

Nei,M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nielsen,C. (1985) *Biol. J. Linn. Soc.*, **25**, 243–299.

Patterson,C. (1989) In Fernholm,B., Bremer,K. and Jörnvall,H. (eds). *The Hierarchy of Life*. Elsevier Science Publishers B.V., pp. 471–488.

Penny,D. and Hendy,M.D. (1986) *Mol. Biol. Evol.*, **3**, 403–417.

Perasso,R., Baroin,A., Qu,L.H., Bachellerie,J.-P. and Adoutte,A. (1989) *Nature*, **339**, 142–144.

Qu,L.H., Michot,B. and Bachellerie,J.-P. (1983) *Nucleic Acids Res.*, **11**, 5903–5920.

Qu,L.H., Nicoloso,M. and Bachellerie,J.-P. (1988) *J. Mol. Evol.*, **28**, 113–124.

Salvini-Plawen,L.V. (1978) *Z. Zool. Syst. Evol.*, **16**, 40–88.

Sogin,M.L., Elwood,H.J. and Gunderson,J.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1383–1387.

Whittaker,R.H. and Margulis,L. (1978) *BioSystems*, **10**, 3–18.

Received on November 5, 1990; revised on December 6, 1990

Note added in proof

The sequence data reported here have been submitted to the EMBL/GenBank/DDBJ sequence databases and are available under the following accession numbers: *Trichoplax adhaerens*, X57253; *Cestus veneris*, X57254; *Bunicella stricta*, X57255; *Porskalia edwardsi*, X57256; *Berce mitrata*, X57257; *Berce ovata*, X57258; *Leukartiara octona*, X57259; *Spongosorites genitrix*, X57260; *Dictyonella incisa*, X57261 and *Crambe crambe*, X57262.