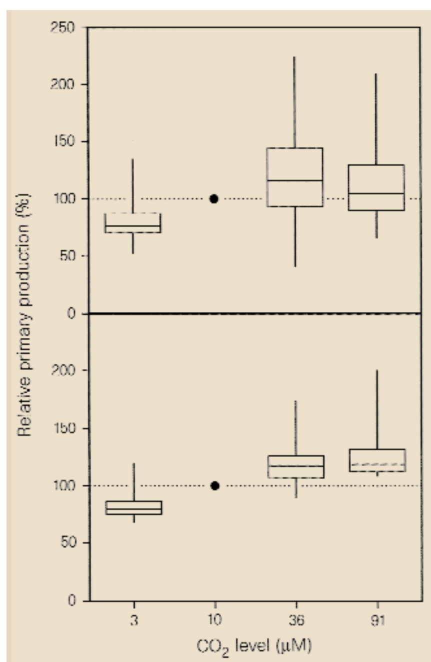

Figure 1 Oceanic primary production at reduced and increased CO_2 concentration (percentage of production at ambient CO_2). Measurements are triplicates at 18 stations along a transect in the Atlantic Ocean (35° S 49° W to 24° N 29° W) on board the Spanish BIO Hesperides. Upper panel, surface layer; lower panel, deeper layer of maximum chlorophyll content. CO_2 concentrations were: ambient 10 ± 0.3 μM, intermediate 36 ± 1.2 μM, high 91 ± 3.1 μM and low 3 ± 0.1 μM. Horizontal lines, median values; boxes, 25 and 75 percentiles; vertical lines include 10–90 percentiles. The coefficient of variation of triplicates for production rates averaged 11% in surface and 7% in deep layers. Primary production rates at elevated intermediate CO_2 concentration (36 μM) were significantly higher than those at ambient levels for surface and deeper layers ($P = 0.028$ and $P < 0.001$, respectively; one-tailed signed rank test).

in the enzyme carbonic anhydrase[5]. Culture experiments and diffusion models with marine phytoplankton have therefore suggested that photosynthesis and growth can be limited by the supply rate of CO_2 (ref. 4). But culture experiments with a few species cannot provide a fair description of the natural response of the mixed phytoplankton assemblage in the ocean where many species are involved and cell density, nutrient status and physiological acclimation may differ widely from those in culture.

We tested whether primary production was limited by ambient CO_2 concentration at 18 stations along a transect in the Atlantic Ocean. Primary production was measured as short-term ^{14}C-incorporation into organic matter in samples collected from the surface waters (~5 m) and from the deeper layer (30–150 m) with maximum chlorophyll content. We illuminated surface samples at 350 μmol photon m^{-2} s^{-1} and deeper samples at 50 μmol m^{-2} s^{-1} for 2 h in an incubator kept at ambient temperature.

We elevated or reduced the CO_2 concentration relative to the ambient level by slightly changing the pH, adding HCl or NaOH. This procedure is straightforward and can change CO_2 concentrations without altering the total inorganic carbon pool. The decline in pH along with the CO_2 rise may influence primary productivity, but this pH decline would also occur naturally with the future CO_2 rise. The alteration of pH may also influence trace metal bioavailability[5], but the duration of the experiments was so short (2 h) that they are unlikely to influence ^{14}C incorporation.

Overall, primary production at the 18 sites responded significantly to manipulation of the CO_2 concentration (Fig. 1). In surface waters, median primary production at low CO_2 (3 μM) was 75% of the level at ambient CO_2 (10 μM) and median primary production at elevated CO_2 (36 μM) was 115% of the ambient level. Likewise, in deeper chlorophyll-rich layers, median primary production was 78% at low CO_2 and 119% at elevated CO_2 as compared with ambient primary production. We saw no further increase in primary production at the highest CO_2 level tested (91 μM).

The response to CO_2 manipulation varied substantially among different sites, as expected considering that the composition of the phytoplankton communities was likely to vary along the transect. At some sites primary production did not respond to CO_2 elevation whereas at other sites primary production more than doubled. Rates of primary production at ambient CO_2 in the surface (0.02–0.44 mg C m^{-3} h^{-1}) and deeper layers (0.05–0.86 mg C m^{-3} h^{-1}) also varied considerably. There was no systematic relationship between the response of primary production to CO_2 manipulation and the geographical location or magnitude of primary productivity.

The CO_2 concentration elevation to 36 μM corresponds roughly to the level expected in the surface ocean at atmospheric equilibrium (855 p.p.m., 25 °C) in year 2100, according to the IPCC 'business as usual' model[6]. Our results indicate that the median primary production could increase by about 15–19% in response to the CO_2 rise, with other factors remaining constant. The variable response to CO_2 elevation along the transect also suggests that the stimulation could be markedly higher (perhaps double) in some instances, probably dominated by phytoplankton species without CO_2-concentrating mechanisms. However, the long-term growth response to CO_2 elevation is usually much lower than the short-term photosynthetic response determined here. This has been shown frequently for terrestrial plants[3].

We propose that the overall response of oceanic primary production to the CO_2 rise would be relatively small, whereas the influence on the species composition of the phytoplankton assemblages could be profound, depending on the kinetics of carbon use.

Mette Hein, Kaj Sand-Jensen
Freshwater Biological Laboratory,
University of Copenhagen,
Helsingørsgade 51, DK-3400 Hillerød, Denmark
e-mail: mette_hein@dk-online.dk

1. Falkowski, P. G. *Photosynth. Res.* **39,** 235–258 (1994).
2. Raven, J. A. & Johnston, A. M. *Limnol. Oceanogr.* **36,** 1701–1714 (1991).
3. Bowes, G. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44,** 309–332 (1993).
4. Riebesell, U., Wolf-Gladrow, D. A. & Smetacek, V. *Nature* **361,** 249–251 (1993).
5. Morel, F. M. M. *et al. Nature* **369,** 740–742 (1994).
6. Houghton, J. T., Jenkins, G. J. & Ephramus, J. J. (eds) *The IPCC Scientific Assessment on Climate Change* (Cambridge Univ. Press, 1990).

Structural biology and phylogenetic estimation

When reconstructing evolutionary trees from DNA sequences, it is often assumed that increasing the amount of sequence will improve the phylogenetic estimate[1,2]. This is based on the notion that historical 'signal' will rise above misleading 'noise' as more sequence is gathered. Our analysis of mitochondrial genomes fails to support this assumption, but suggests a way to select objectively for data with maximum 'signal-to-noise' potential.

We carried out a phylogenetic parsimony analysis[3] of the entire protein-coding portion (12,234 base pairs) of the mitochondrial genome of 19 taxa whose interrelationships are widely accepted (Fig. 1a). Not only did we fail to obtain the expected phylogenetic tree from this large data set, but found compelling bootstrap support for the incorrect placements (Fig. 1b). This suggests that misleading signals will not always disappear as more data are collected.

When a subset of sites associated with residues important for protein folding was subjected to phylogenetic analysis, the expected tree resulted with strong bootstrap support. Phylogenetic analysis of DNA sequences might therefore be improved by incorporating structural and functional considerations into inference models[4–6].

To determine which of our 12,234 sites were responsible for the misleading signal, we superimposed the sequence data onto the accepted tree and measured, using the retention index[7,8], how well each site fitted that tree. We found that among codons, nucleotides at third positions produced the poorest fits; among genes, *NADH2* produced the poorest fit; and isoleucine, leucine and valine produced the poorest fits of the amino acids. Sites coding for residues with aliphatic side chains produced poorer fits than did those coding for any other type

Nature © Macmillan Publishers Ltd 1997

Nature © Macmillan Publishers Ltd 1997

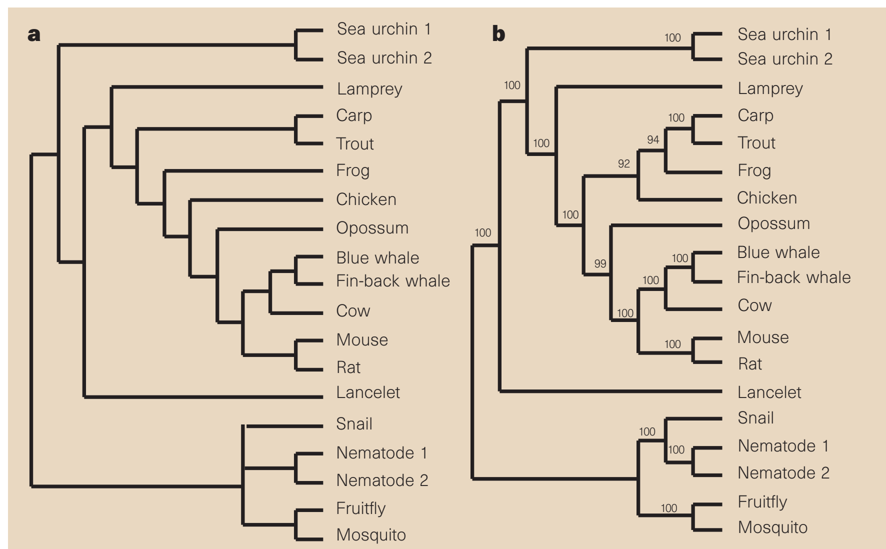

Figure 1 a, The widely accepted pattern of relationships for the 19 taxa analysed. Species names: mouse, *Mus musculus*; rat, *Rattus norvegicus*; cow, *Bos taurus*; fin-back whale, *Balaenopterus physalus*; blue whale, *Balaenopterus musculus*; opossum, *Didelphis virginiana*; chicken, *Gallus gallus*; frog, *Xenopus laevis*; carp, *Cyprinus carpio*; trout, *Oncorhynchus mykiss*; lamprey, *Petromyzon marinus*; lancelet, *Branchiostoma floridae*; sea urchin 1, *Paracentrotus lividus*; sea urchin 2, *Strongylocentrotus purpuratus*; fruitfly, *Drosophila yakuba*; snail, *Cepaea nemoralis*; mosquito, *Anopheles gambiae*; nematode 1, *Ascaris suum*; nematode 2, *Caenorhabditis elegans*. **b,** Bootstrap consensus tree based on a parsimony analysis of 12,234 nucleotide sites coding for the 13 mitochondrial genes. Numbers show percentage bootstrap support at each node.

of side chain; sites coding for charged residues produced better fits than did those for uncharged residues; and sites coding for hydrophilic residues produced better fits than did those coding for hydrophobic ones (Fig. 2).

On the basis of these analyses, we were able to identify functional classes of sites that were phylogenetically reliable. We found that parsimony analysis of a subset of sites comprising first and second codon positions, modally coding for proline, cysteine, methionine, glutamine and asparagine, yielded the expected tree with strong bootstrap support for most nodes. Although this cannot serve as an independent test (the expected tree was used to discover sites that were resilient), it is noteworthy that most of the resilient sites are associated with amino-acid residues important for tertiary protein structure.

Tree-based explorations of sequence variation such as this show considerable promise for both evolutionary and structural biology. The distributions and incidences of change implied by an accepted set of evolutionary relationships can be physically mapped onto a structural model of a protein or RNA product[9,10] to shed light on the constraints and opportunities that have characterized the evolution of the underlying sequence. This information can be used both to refine phylogenetic models and to provide structural biologists with clues about the relative importance of particular co-varying combinations of residues for protein structure, function and folding.

Comparative morphologists have long acknowledged that character co-variation among organisms is affected not only by a shared history, but also by functional requirements. To work as an integrated unit, the sub-components of a complex morphological character must necessarily co-vary. There is every reason to believe that molecules are similar. As the database of protein structures continues to grow, and as sequences from an ever broader taxonomic range of organisms become available for a number of proteins, we can look forward to a better understanding of the historical and functional constraints that act on macromolecules and, as a consequence, to more realistic biochemically based models of change from which to infer evolutionary trees.

Gavin J. P. Naylor
Department of Zoology and Genetics,
Iowa State University, Ames, Iowa 50011, USA
e-mail: gnaylor@iastate.edu
Wesley M. Brown
Department of Biology, University of Michigan,
Ann Arbor, Michigan 48109-1048, USA

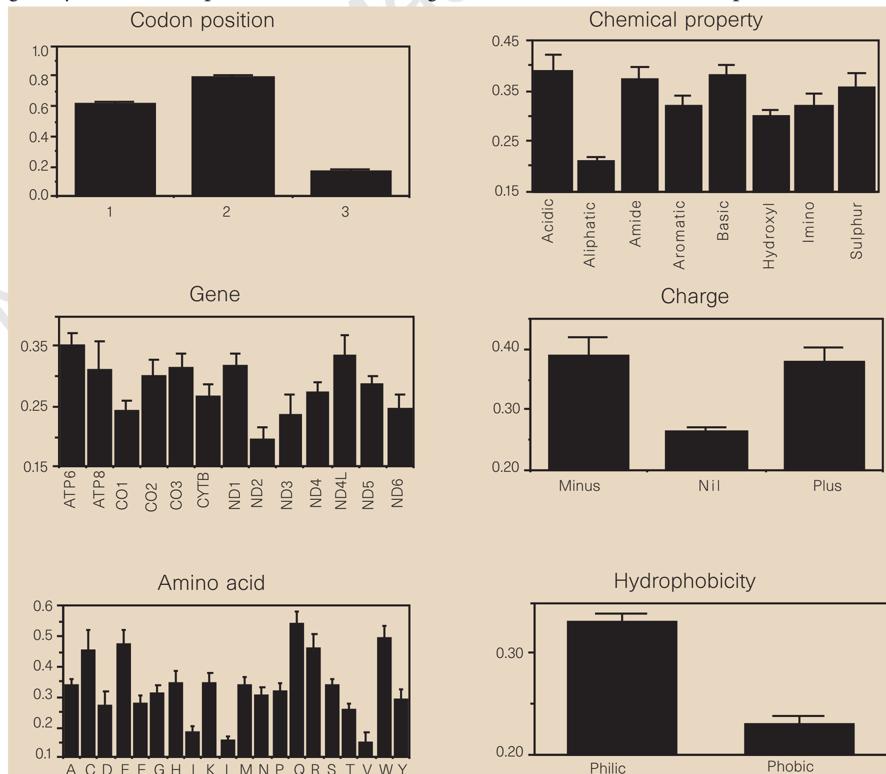

Figure 2 Relationship between functional characteristics and phylogenetic fit of a site for the accepted tree. Fit was assessed using the retention index[7,8]. The relative effects of the different levels of each influence are plotted against log retention index (ordinate) as response sample means. Bars show s.e.m. Gene abbreviations: ATP6, ATP8, ATPase subunits 6 and 8; CO1–3, cytochrome oxidase subunits 1–3; CYTB, cytochrome B; ND1–6 and ND4L, NADH dehydrogenase subunits 1–6 and 4L. Amino acids are listed as single-letter code.

1. Churchill, G. A., von Haessler, A. & Navidi, W. C. *Mol. Biol. Evol.* **9,** 753–769 (1992).
2. Huelsenbeck, J. P. & Hillis, D. M. *Syst. Biol.* **42,** 247–264 (1993).
3. Swofford, D. L. *PAUP 3.1.1* (Smithsonian Inst., Washington DC, 1993).
4. Irwin, D. M., Kocher, T. D. & Wilson, A. C. *J. Mol. Evol.* **32,** 128–144 (1991).
5. Adkins, R. M., Honeycutt, R. L. & Disotell, T. R. *Mol. Biol. Evol.* **13,** 1393–1404 (1996).
6. Perutz, M. F. *Mol. Biol. Evol.* **1,** 1–28 (1983).
7. Archie, J. W. *Syst. Zool.* **38,** 253–269 (1989).
8. Farris, J. S. *Cladistics* **5,** 417–419 (1989).
9. Golding, B. in *Non-neutral Evolution* (ed. Golding, B.) 126–139 (Chapman and Hall, New York, 1994).
10. Schnare, M. N., Damberger, S. H., Gray, M. W. & Gutell, R. R. *J. Mol. Biol.* **256,** 701–719 (1996).

Nature © Macmillan Publishers Ltd 1997