Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 1930–1934, March 1996
Evolution

Evolution of chlorophyll and bacteriochlorophyll: The problem of invariant sites in sequence analysis

(evolutionary trees/sites free to vary/photosynthetic reaction centers/anciently diverged sequences)

PETER J. LOCKHART*, ANTHONY W. D. LARKUM†, MICHAEL A. STEEL‡, PETER J. WADDELL*, AND DAVID PENNY*

*Molecular Genetics Unit, School of Biological Sciences, Massey University, Palmerston North New Zealand; †School of Biological Sciences, University of Sydney, NSW 2006, Australia; and ‡Department of Mathematics, University of Canterbury, Christchurch, New Zealand

Communicated by Peter H. Raven, Missouri Botanical Garden, St. Louis, MO, October 30, 1995

ABSTRACT **Competing hypotheses seek to explain the evolution of oxygenic and anoxygenic processes of photosynthesis. Since chlorophyll is less reduced and precedes bacteriochlorophyll on the modern biosynthetic pathway, it has been proposed that chlorophyll preceded bacteriochlorophyll in its evolution. However, recent analyses of nucleotide sequences that encode chlorophyll and bacteriochlorophyll biosynthetic enzymes appear to provide support for an alternative hypothesis. This is that the evolution of bacteriochlorophyll occurred earlier than the evolution of chlorophyll. Here we demonstrate that the presence of invariant sites in sequence datasets leads to inconsistency in tree building (including maximum-likelihood methods). Homologous sequences with different biological functions often share invariant sites at the same nucleotide positions. However, different constraints can also result in additional invariant sites unique to the genes, which have specific and different biological functions. Consequently, the distribution of these sites can be uneven between the different types of homologous genes. The presence of invariant sites, shared by related biosynthetic genes as well as those unique to only some of these genes, has misled the recent evolutionary analysis of oxygenic and anoxygenic photosynthetic pigments. We evaluate an alternative scheme for the evolution of chlorophyll and bacteriochlorophyll.**

In the evolution of organisms on the early Earth, photosynthesis must have played a pivotal role in supplying a plentiful supply of fixed carbon upon which other organisms could subsist. The existence of photosynthesis on the early Earth perhaps as far back as 4 billion years ago (BYA) is supported by a number of lines of evidence (1). A more surprising suggestion is the proposal that oxygenic photosynthesis, an advanced stage of photosynthesis where oxygen is liberated as a by-product, evolved as early as 3.5 BYA—i.e., ≈1 BYA after the formation of the Earth. The evidence for this proposal comes from a number of sources but is based primarily on the existence of microfossils in stromatolitic rocks from the Archaean shield dating back to 3.5 BYA (2). However, because of the small and simple nature of the microfossils it may be impossible to phylogenetically identify these organisms precisely. Nevertheless, they do have features similar to modern cyanobacteria (blue-green algae), which are also found on modern stromatolites.

An outstanding problem for modern research is to understand how oxygenic photosynthesis evolved. Oxygenic photosynthesizers (chloroplasts, cyanobacteria, and prochlorophytes) use chlorophyll (Chl) as their major photosynthetic pigment. In contrast, all known anoxygenic eubacteria use bacteriochlorophyll (BChl). Until recently, it was widely accepted that the first photosynthetic organisms were eubacteria, which carried out a type of anoxygenic photosynthesis similar to that carried out by modern day anoxygenic eubacteria. However, other possibilities for the nature of early photosynthetic systems and their pigments have also been put forward. Some hypotheses suggest that Chl preceded the evolution of BChl in photosynthesis (3, 4). Under this view, some kind of anoxygenic photosynthesis may still have evolved before oxygenic photosynthesis. If so, it may well have been different from that found in anoxygenic eubacteria today.

In reconstructing the history of photosynthesis, molecular evidence is being used increasingly. At present, phylogenetic studies on the "tree of life" give conflicting histories for the evolution of oxygenic and nonoxygenic photosynthesizers (5–7). Improved methodology for sequence analysis may be of help here (8–10). At the level of evolution of biochemical pathway intermediates, there is the principle that, as organisms evolved, their metabolic pathways also evolved from simple to more complex in a series of simple steps. This principle has been called the Granick hypothesis after its proposer (11). One example, with which this contribution is concerned, is the evolution of the biochemical pathways of photosynthetic pigments. In this case, since Chl-a is generated at least one step earlier in the modern biosynthetic pathway than BChl-a (Fig. 1*A*), the expectation under the Granick hypothesis is that Chl-a evolved earlier than BChl-a.

Despite the controversial issue of how far into the distant past currently implemented methods of primary sequence analysis can be used for reconstructing phylogenetic events (13–15), an attempt was recently made to test hypotheses of origin for photosynthetic pigments. This involved application of phylogenetic analysis methodology in an extreme situation. Here, the time of divergence between the compared sequences is so great that it must preclude a neutral model of sequence change from describing their evolution. The problem of analysis is even further complicated because the compared sequences have different biological roles and hence different functional constraints.

Burke *et al.* (12) carried out a maximum-likelihood analysis on Chl and BChl biosynthetic gene sequences in an attempt to answer the question as to whether the evolution of Chl preceded BChl. They concluded from their analyses that the first photosynthesizers contained BChl and not Chl. Their proposed scheme for the evolution of photosynthesis (Fig. 1*B*) envisaged that an ancestral nonspecific reductase (called *pchL* in Fig. 1*B*) carried out two sequential catalytic steps in the late stages of BChl synthesis. These steps are reduction of a porphyrin ring to a chlorin ring and reduction of the chlorin ring to a bacteriochlorin ring. In extant anoxygenic bacteria, these catalytic functions are encoded by two separate genes, *bchL* and *bchX*, respectively (Fig. 1). In the proposed scheme of Burke *et al.* (12), these two genes evolved from *pchL*.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: BYA, billion years ago; Chl, chlorophyll; BChl, bacteriochlorophyll; RC, reaction center; i.i.d., independent and evolving identically.

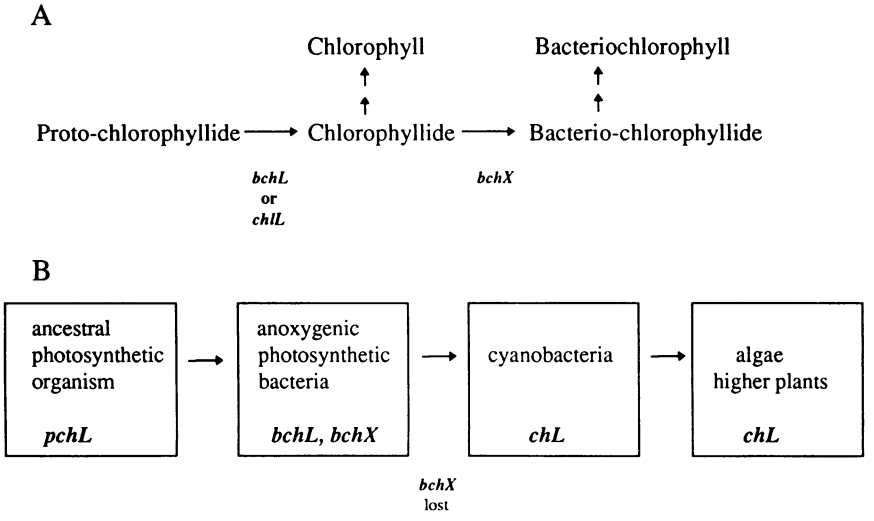

FIG. 1. (*A*) Scheme showing part of the modern biosynthetic pathway for synthesis of Chl and BChl. (*B*) Interpretation for evolution of Chl consistent with the results of Burke *et al.* (12) *bchX* has been lost in the lineage leading to cyanobacteria.

To test possible hypotheses, Burke *et al.* (12) started with a tree linking biosynthetic enzymes that are clearly homologous (Fig. 2*A*). They then examined where *nifH* outgroup sequences would join onto this tree. These sequences are assumed to be ancestral because of their wide distribution in archaebacteria and eubacteria. Generally, the placement of outgroups is expected to indicate the root of the ingroup sequences. However, outgroup sequences can be misplaced. For example, this can occur when different sequences change at unequal rates or when other systematic biases affect the data (16–18). Burke *et al.* (12) found that under a maximum-likelihood model the *nifH* genes always joined the *bchX* edge (Fig. 2*A*, edge 1). From this, they concluded that the *bchX* gene was the earliest diverged biosynthetic gene. Hence, based on the position of the root, they suggested that BChl must have preceded the evolution of Chl.

In their scheme for the evolution of photosynthesis, two inferences are needed to explain the evolution of cyanobacteria that possess Chl but not BChl. These are that (*i*) *bchL* evolved into the *chlL* gene enzyme now found in cyanobacteria and chloroplasts, and (*ii*) the ability to form BChl using *bchX* and incorporate it into reaction centres (RCs) was lost on the lineage leading to cyanobacteria and chloroplasts (Fig. 1*B*).

In this contribution, we examine the evidence of Burke *et al.* (12) that oxygenic photosynthesizers evolved from a lineage that first led to present day anoxygenic photosynthetic bacteria. In doing so, we examine the assertion that *bchX* is the earliest diverged biosynthetic pigment gene of those compared. Our conclusion is that, when sites not free to vary in the different genes are taken into account, a different hypothesis from that proposed by Burke *et al.* (12) is equally well supported by current data. This is that Chl is as old as or has even preceded BChl in the evolution of photosynthesis.

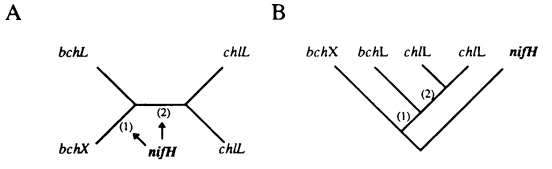

FIG. 2. (*A*) Unrooted tree describing the relationship between biosynthetic genes; two of five possible placements for attachment of the *nifH* outgroups are shown as (*1*) and (*2*). (*B*) Rooted tree obtained by Burke *et al.* (12); *nifH* genes join internode/internal branch 1 (edge 1), the *bchX* edge.

MATERIALS AND METHODS

Amino acid sequences for *nifH, bchL, bchX*, and *chlL* gene sequences were aligned using MULTALIN (19). The underlying nucleotide positions were then edited into these alignments, resulting in an alignment very similar to that used by Burke *et al.* (12). All the positions we analyzed were conserved within the alignment they report. However, since we aligned additional taxa not considered by them (Table 1), we removed all ambiguous positions around insertions and deletions. We then reconstructed a phylogeny for the sequences analyzed by Burke *et al.* (12) and made the assumption (as they did) that all sites in the sequences were equally free to vary. Although our alignment was very conservative, maximum-likelihood

Table 1. Nitrogenase reductase and RC pigment biosynthetic sequences

Locus	Taxon	GenBank accession nos.
nifH		
MTNIFHDK	*Methanococcus thermolithotrophicus*	X13830
MTNIFH	*Methanobacterium thermolithotropicus*	X07500
CPNIFH3	*Clostridium pasteurianum*	X07474, X06756
CPNIFH1	*C. pasteurianum*	X07472, X06756
AVINIFA	*Azotobacter vinelandii*	M11579
A7NIFH	*Anabaena* sp. PCC7120	V00001
RCNIFH	*Rhodobacter capsulatus*	X07866
MVNIFH	*Methanococcus voltae*	X03777
RSPNIFHD	*Rhodospirillum rubrum*	M33774
chlL		
CHMPXX	*Marchantia polymorpha*	X04465, Y00686
CHPCTRNX	*Pinus contorta*	X56200
CRCHPL	*Chlamydomonas reinhardtii*	X62905
SSU00733	*Synechococcus* sp. PCC7002	U00733
PEEFRXC	*Plectonema boryanum*	D00665
bchX and *bchL*		
RCPHSYNG	*R. capsulatus*	Z11165
RSBCHPUF	*Rhodospeudomonas sphaeroides*	X68795, S55638, S79406, X63320

The identity and source of sequences used in this study are shown.

analysis [DNAML3.5 (20)] on the first and second codon positions also gave us the same tree as reported by Burke *et al.* (12). The contribution of invariant sites in these data was then examined for any misleading effect they might have on this inference. We define invariant sites as positions in the alignment that are not free to vary between sequences of the same biological function. In addition, some of these positions are also invariant across all *nifH, bchL, bchX,* and *chlL* sequences. In our study, we have (*i*) implemented the maximum-likelihood model of DNAML3.5 (20) on data that excluded invariant positions and (*ii*) implemented a new option in DNAML (20), which allows a category of sites in the data to be invariant. This second strategy was possible because of a recently implemented option in DNAML3.5 (20). It was not available to Burke *et al.* (12) at the time of their work.

In helping to evaluate the effect of invariant sites, we used the capture–recapture method of Sidow *et al.* (21). This allowed us to estimate the proportions of invariant codons between the sequences of different biological function. Although not all the sequences shown in Table 1 were used in tree reconstruction, their inclusion in the overall alignment allowed us to make more reliable estimates of codons free to vary in these genes. Estimates from the capture–recapture method are known to have a possible bias if changes at first and second codon positions are not independent (e.g., upward for positive correlations and downward for negative correlations). However, this bias is expected to be small (21). We also estimated numbers of invariant sites by using the categories option of DNAML3.5 (20). We did this by holding increasing numbers of sites in our alignment invariant until we identified the point at which the log-likelihood for a given tree was optimal (P.J.W., unpublished data). The estimate of invariant codons is expected to be less than the estimated number of invariant sites. This has previously been discussed (21).

To separate codon positions in our alignments, we used the program PREPARE (22). Jackknifing was carried out using SEQBOOT3.5 (20). Evolutionary trees were tested under the Kishino–Hasegawa sites test (23) with DNAML3.5 (20). The results reported here used a transition/transversion ratio of 1, but altering this value had little affect on the analyses.

RESULTS AND DISCUSSION

Inferences from Anciently Diverged Data. The conclusions of Burke *et al.* (12) that *bchX* encodes the earliest diverged pigment reductase depend on the reliability of the *nifH* outgroup placement. They found, under a maximum-likelihood model and using a strategy that was careful to

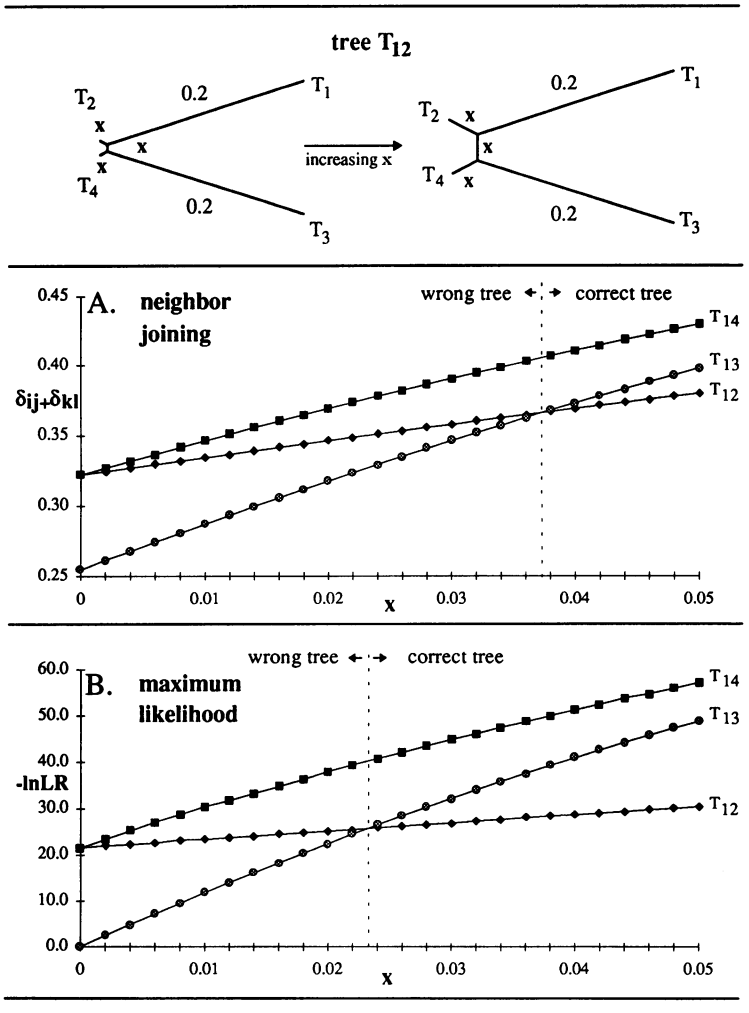

FIG. 3. A four taxon case where tree building methods will select the wrong tree in the presence of invariant sites. The tree building methods are an i.i.d. correction with neighbor joining (*A*) and a maximum-likelihood model (*B*). Given certain edge lengths, both methods will become inconsistent when invariant sites are present. In this model, two edges are fixed with a 20% probability of change at any site [0.2 substitution/all sites (including invariant sites)]. The three other edges have the length x and this is varied. To the left of the dotted lines are the smaller values of x for which the tree building methods will select the wrong tree; 50% of the sites are actually invariant but the analyses assume that all sites are equally free to vary. The mechanism of sequence substitution at the variable sites is a two-state equifrequency Poisson process. The y axes are $\delta_{ij} + \delta_{kl}$, which is the sum of the corrected distances between taxa i, j, k, and l (*A*), 2lnLR, which is the likelihood ratio goodness of fit statistic (*B*).

minimize any effect of base composition bias (10, 24), evidence for placing *bchX* with *nifH* (Fig. 2). Their study, by using a maximum-likelihood procedure, was also rigorous in that it allowed evaluation of all possible trees for a prescribed mechanism of sequence evolution. However, the validity of their analysis rested upon a sufficiently good fit between the assumptions of the maximum-likelihood model and evolution of the sequences.

An important and untested assumption in the model used by Burke *et al.* (12) was that the evolution of the sequences was sufficiently well described by the model that all the sites in the sequences were independent and evolving identically at the same rate (that is, this mechanism is said to be i.i.d.). With coding sequences, these assumptions are unlikely to be true since many positions are functionally constrained, in some cases to the extent of being invariant (13, 15, 25). Fig. 3 illustrates the problem for phylogenetic inference for a case in which there is a large proportion of invariant sites in the data. It shows that neighbor-joining and maximum-likelihood methods can select the wrong tree when invariant sites are present. This problem is part of a general inconsistency problem due to model violations that can occur when positions in the sequences have differences in their substitution probabilities (P.J.W., unpublished data).

That inconsistency, due to invariant sites, is part of the problem in the analysis of Burke *et al.* (12) is indicated from maximum-likelihood analysis of the data when a proportion of the sites are treated as invariant. When no sites are assumed to be invariant then DNAML3.5 (20) gives significant support for the tree of Burke *et al.* (12) over a tree in which the outgroup edge joins between *chlL* and *bchL/bchX* (Fig. 2*A*, edge 2). However, when >30% of the sites are considered invariant, then the Burke *et al.* (12) tree is no longer significant at the 0.05 level. Our estimates using the capture–recapture method (21) for the number of invariant codons across all sequences in the data are ≈28%. Using a maximum-likelihood procedure, the number of invariant sites is estimated between 40% and 50%. Thus, at the 0.05 level, the tree of Burke *et al.* (12) is not significantly better than the alternative tree we test.

The presence of invariant sites across all compared sequences is only part of the problem. An additional complexity arises when invariant sites occur in all genes of similar function but not in homologues of different function. In the data used by Burke *et al.* (12), this phenomena also occurs. That is, while some sites are invariant across all enzymes, others are only invariant in enzymes of a particular function. Estimates of codons free to vary in the reductase genes of different biological function are shown in Table 2. The values reported here indicate (*i*) that there are codons that are invariant between all sequences and (*ii*) that there is also a large proportion of invariant codons occurring between *chlL* and *bchL* genes, which are not invariant at the same positions in *bchX* and *nifH* genes. This unequal distribution of invariant sites is such that, even if nucleotide positions at sites free to vary are saturated by multiple substitutions, phylogenetic analyses will still be strongly biased toward placing the *nifH* outgroups outside the partition that separates *bchl* and *chlL* from *bchX*. Thus, *nifH* will tend to join with *bchX* irrespective of the true phylogeny. This problem will not be detected by the analyses carried out by Burke *et al.* (12) and the observation cautions against using simple models to study such ancient divergences.

Table 2. Estimates of the number of codons free to vary in an alignment of biosynthetic and nitrogenase reductase genes

Sequences used to estimate codons free to vary	Observed no. of variable patterns in 242 (1st + 2nd codon) positions			Proportion of codons free to vary
	1st	2nd	1st + 2nd	
chlL + *bchL*	38	20	15	0.42 ± 0.04
chlL + *bchL* + *bchX*	64	43	32	0.71 ± 0.05
chlL + *bchL* + *bchX* + *nifH*	75	58	50	0.72 ± 0.02

A test of the difference in the number of codons free to vary between *chlL* + *bchL* and *chlL* + *bchL* + *bchX* shows a 4.74× SE difference.

Since the maximum-likelihood method used by Burke *et al.* (12) should examine only sites free to vary between all sequences, we have applied the same tree selection criteria (maximum-likelihood) to sites estimated as being free to vary. In an attempt to identify such sites, we first chose an initial subset of data columns which showed variable character states in the *chlL/bchL* genes (there were 86 such positions). Fig. 4 shows the support with these data for the Burke *et al.* (12) tree (T_1; Fig. 2*A*, edge 1) over an alternative tree (T_2; Fig. 2*A*, edge 2) in which BChl did not precede evolution of Chl. Next, we examined the relative support for these hypotheses as additional columns of data were included from the original alignment. These were randomly sampled without replacement (jackknife sampled). The tree length difference and standard error were calculated for the two hypotheses when 10, 20, 30, 40, 50, and 60 randomly sampled columns of data were added. For each data set size, 10 replicates were made. Fig. 4 shows that as the number of sites included in the analysis increased so did the support for the hypothesis proposed by Burke *et al.* (12). Greatest support for this hypothesis occurred with all 242 sites from our alignment.

If the inference of Burke *et al.* (12) is valid, then an expectation is that significant support will be found for their tree (T_1), over the alternative one (T_2), when the appropriate number of sites free to vary is included in the analysis. Examination of the log-likelihood scores found with DNAML3.5 (20) for *chlL/bchL* sequences suggests that between 60% and 70% sites are invariant (this corresponds to the 48% of codons that were estimated as invariant). Hence, <15 sites should be added to the 86 variable columns of data in order to obtain a data set that contains the appropriate number of sites free to vary. Significant support for the tree of Burke *et al.* (12) is found only when a much larger number of violating sites are included (Fig. 4).

An important consideration here is our estimate of codons free to vary. These are extremely low in the oxygenic photosynthetic *ChlL* sequences (≈18%) and still small between *chlL* and *bchL* genes (≈42%). If these estimates are artificially low, we will be discounting valid information used by Burke *et al.*

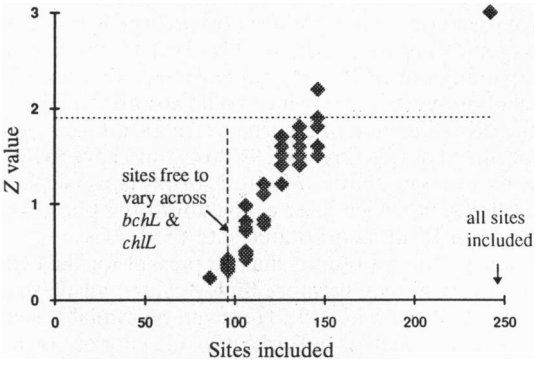

FIG. 4. Effect of adding in invariant sites. Significant support is found only for the Burke *et al.* (12) tree when such sites are present. Shown is the relative support under the Kishino–Hasegawa sites test (23) for two trees, T_1 over T_2. These trees differ only in the placement of the outgroups (see Fig. 2*A*, edges 1 and 2). T_1 is the tree reported by Burke *et al.* (12). T_2 places the outgroups on an edge that separates *chlL* from the other genes. The data sets analyzed are of sizes 86, 96, 106, 116, 126, 136, 146, and 242 positions. For each data set, the *Z* value—tree length difference (lnLR tree T_1 − lnLR tree T_2) divided by the standard error—has been plotted. Only when sites are included in excess of those estimated as being free to vary is there support for the tree of Burke *et al.* (12) at the 0.05 level.

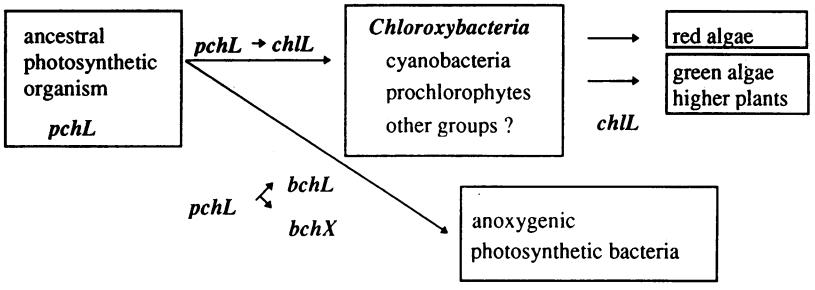

FIG. 5. Alternative scheme examined for evolution of Chl and BChl and which is compatible with tree T_2 (Fig. 4). It incorporates Prochlorophyta as well as Cyanobacteria under the one name of Chloroxybacteria for describing all oxygenic prokaryotic organisms. This scheme incorporates one proposal of Burke *et al.* (12) that the earliest reductase enzyme (encoded by *pchl*) for reducing the tetrapyrrole ring was nonselective. Nevertheless, it does not exclude the possibility that a subsequent specific enzyme later evolved for the reduction of the protochlorophyllide (the precursor to *chlL* gene). In this scheme, *bchL* and *bchX* also evolved from the nonspecific ancestral protochlorophyllide reductase (*pchL*) gene in anoxygenic photosynthetic bacteria at the time diversification of photosystems I and II was occurring.

(12). However, our findings are reinforced by observations elsewhere, suggesting low numbers of sites free to vary in other oxygenic photosynthetic genes (26). Our observations are also supported by calculations using the "Bealey" theorem (27). This theorem provides a bound (violated for these sequences) on the expected number of constant sites that should be present under a simple i.i.d. mechanism from the histogram of parsimony scores across sites. Its implementation also suggests that with these data there are far too many constant sites for evolution of the sequences to have occurred as a result of an i.i.d. mechanism. Based on these estimates, it appears that the apparent support for the tree of Burke *et al.* (12) is spurious and the result of including patterns in the data that violate the assumptions of the method used.

An Alternative Hypothesis for the Evolution of Photosynthetic Pigments. The scheme adopted by Burke *et al.* (12) is one favored by many biologists in the field; however it is not the only possibility. As mentioned in the Introduction, others have suggested that Chl may have preceded BChl (3, 4, 28, 29) and at least two of these (3, 4) have suggested an evolutionary scheme in which oxygenic photosynthetic bacteria evolved simultaneously with the anoxygenic photosynthetic bacteria. We have outlined such a scheme in Fig. 5. According to this scheme, the possibility that the early nonspecific reductase caused the production of some BChl (as well as Chl) could mean that some BChl might also be incorporated into RCs of the ancestral photosynthetic organism and even of the early Chloroxybacteria (those photosynthetic organisms that carry out oxygenic photosynthesis; see Fig. 5). The infrared absorption capabilities of BChls may not have been very useful to the earliest photosynthesizers, which could absorb visible light, but with the diversification of light-harvesting strategies that must have occurred at this time (4, 29) this would have favored the development of shade-adapted organisms with light-harvesting BChl, which later evolved into the photosynthetic bacteria with BChl incorporated into their RCs.

Summary. The hypothesis that we present for the evolution of Chl and BChl may be more biologically realistic than that proposed by Burke *et al.* (12). However, on available sequence data neither the analysis of Burke *et al.* (12) nor our own favors a particular hypothesis. If correct, ours suggests that the *bchX* edge in reconstructed evolutionary trees may be very long not because it represents an early diverged lineage but because after gene duplication and the development of its present biological function it may have accumulated relatively more changes than the *bchL* and *chlL* lineages. Earlier theoretical work (16, 17) together with our reported findings suggest that *bchX* and the outgroups could have wrongly joined in the analyses of Burke *et al.* (12). As a consequence, this interpretation may be misleading our understanding of early events in the evolution of photosynthesis.

We thank Joe Felsenstein and the anonymous reviewers for their constructive criticism. This work was supported by a grant from the New Zealand Public Good Science Fund.

1. Schidlowski, M. J. (1991) *Southeast Asian Earth Sci.* **5,** 333–337.
2. Schopf, J. W. (1993) *Science* **260,** 640–646.
3. Olson, J. M. (1970) *Science* **168,** 438–446.
4. Larkum, A. W. D. (1991) in *Chlorophylls*, ed. Scheer, H. (CRC, Boca Raton, FL), pp. 367–383.
5. Olsen, G. J. & Woese, C. R. (1993) *FASEB J.* **7,** 113–123.
6. Klenk, H.-P., Palm, P. & Zillig, W. (1994) *Syst. Appl. Microbiol.* **16,** 638–647.
7. Gogarten, J. P. (1995) *Trends Ecol. Evol.* **10,** 147–151.
8. Adachi, J. & Hasegawa, M. (1992) *Comput. Sci. Monogr. Inst. Statist. Math.* **27,** 1–77.
9. Lake, J. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 1455–1459.
10. Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. (1994) *Mol. Biol. Evol.* **11,** 605–612.
11. Granick, S. (1965) in *Evolving Genes and Proteins*, eds. Bryson, V. & Vogel, H. J. (Academic, New York), pp. 67–68.
12. Burke, D. H., Hearst, J. E. & Sidow, A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7134–7138.
13. Meyer, T. E., Cusanovich, M. A. & Kamen, M. D. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 217–220.
14. Lockhart, P. J., Howe, C. J., Bryant, D. A., Beanland, T. J. & Larkum, A. W. D. (1992) *J. Mol. Evol.* **34,** 153–162.
15. Albert, V. A., Backlund, A., Bremer, K., Chase, M. W., Manhart, J. R., Mishler, B. D. & Nixon, K. C. (1994) *Ann. Missouri. Bot. Gard.* **81,** 534–555.
16. Felsenstein, J. (1978) *Syst. Zool.* **27,** 401–410.
17. Hendy, M. D. & Penny, D. (1989) *Syst. Zool.* **38,** 297–309.
18. Penny, D., Lockhart, P. J., Steel, M. A. & Hendy, M. D. (1994) in *Models in Phylogeny Reconstruction*, eds. Scotland, R. W., Siebert, D. J. & Williams, D. M., Systematics Association Special Volume No. 52 (Clarendon, Oxford), pp. 211–230.
19. Corpet, F. (1988) *Nucleic Acids Res.* **16,** 10881–10891.
20. Felsenstein, J. (1993) PHYLIP 3.5 Manual (Univ. of California Herbarium, Berkeley).
21. Sidow, A., Nguyen, T. & Speed, T. P. (1992) *J. Mol. Evol.* **35,** 253–260.
22. Penny, D., Watson, E. E., Hickson, R. E. & Lockhart, P. J. (1993) *N.Z. J. Bot.* **31,** 275–288.
23. Kishino, H. & Hasegawa, M. (1989) *J. Mol. Evol.* **29,** 170–179.
24. Sidow, A. & Wilson, A. C. (1990) *J. Mol. Evol.* **31,** 51–68.
25. Shoemaker, J. S. & Fitch, W. M. (1989) *Mol. Biol. Evol.* **6,** 270–289.
26. Buttner, M., Xie, D.-L., Nelson, H., Pinther, W., Hauska, G. & Nelson, N. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8135–8139.
27. Steel, M. A., Hendy, M. D. & Penny, D. (1995) *Universtat Bielefeld Forschungsschwerpunk Mathematisierung-Strukturbildingsprozesse*, preprint XCIV.
28. Mauzerall, D. (1973) *Ann. N.Y. Acad. Sci.* **206,** 483–492.
29. Larkum, A. W. D. (1992) in *Research in Photosynthesis*, ed. Murata, N. (Kluwer, Dordrecht, The Netherlands), Vol. 3, pp. 475–480.