The Light-Harvesting Complex of Photosystem I in *Chlamydomonas reinhardtii*: Protein Composition, Gene Structures and Phylogenetic Implications

Ryutaro Tokutsu 1, Haruhiko Teramoto 2, Yuichiro Takahashi 3, Taka-aki Ono 2 and Jun Minagawa 1, 4

1 Institute of Low Temperature Science, Hokkaido University, N19 W8, Sapporo, 060-0819 Japan
2 Laboratory for Photo-Biology, Photodynamics Research Center, The Institute of Physical and Chemical Research (RIKEN), Sendai, 980-0845 Japan
3 Department of Biology, Faculty of Science, Okayama University, 3-1-1 Tsushima-naka, Okayama, 700-8530 Japan

Light-harvesting chlorophyll *a/b*-binding proteins (LHCI) associated with photosystem I (PSI) and the genes encoding these proteins have been characterized in the unicellular green alga *Chlamydomonas reinhardtii*, extending previous studies of the PSII-LHCII [Teramoto et al. (2001) Plant Cell Physiol. 42: 849]. In order to assign LHCI proteins in the thylakoid membranes, the PSI-LHCI supercomplex that retains all of the major LHCI proteins was purified. Seven distinct LHCI proteins were resolved from the purified supercomplex by a high-resolution SDS polyacrylamide gel electrophoresis, and their N-terminal amino acid sequences were determined. One LHCI protein (band e) was newly found, although the other six LHCI proteins corresponded to those previously reported. Genomic clones encoding these seven LHCI proteins were newly isolated and the nucleotide sequences were determined. A comprehensive characterization of all members of *Lhc* gene family in this alga revealed that LHCI proteins are more highly diverged than LHCII, suggesting functional differentiation of the protein components in LHCI. Neighboring joining trees were constructed for LHC proteins from *C. reinhardtii* and those of *Arabidopsis thaliana* or *Galdieria sulphuraria* to assess evolutionary relationships. Phylogenetic analysis revealed that (1) green algal LHCI and LHCII proteins are more closely related to one another than to LHCI proteins in red algae, (2) green algae and higher plants possess seven common lineages of LHC proteins, and (3) Type I and III LHCI proteins are conserved between green algae and higher plants, while Type II and IV are not. These findings are discussed in the context of evolution of multiple diverse antenna complexes.

Keywords: *Chlamydomonas reinhardtii* — Light-harvesting chlorophyll- *a/b* protein complex — Photosystem I.

Abbreviations: Chl, chlorophyll; EST, expressed sequence tag; LHC, the light-harvesting complex; LHCI and LHCII, the light-harvesting complex of photosystem I and photosystem II; PSI and PSII, photosystem I and II.

The nucleotide sequences reported in this paper have been submitted to DDBJ/EMBL/GenBank under the accession numbers AB122114 (*LhcI-1* gene), AB122115 (*LhcI-2* gene), AB122116 (*LhcI-3 gene), AB122117 (*LhcI-4* gene), AB122118 (*LhcI-5* gene), AB122119 (*LhcI-6* gene), and AB122120 (*LhcI-7* gene).

Introduction

Plants produce metabolic energy by collecting light energy and transferring it to photosynthetic reaction centers. This process is facilitated by antennae or light harvesting complexes (LHC) which are composed of LHC proteins that bind light harvesting pigments. LHCs have been characterized from various organisms including chlorophyll *a/b*-containing Chlorophyta, chlorophyll *a*-containing Chromophyta, and chlorophyll *a*-containing Rhodophyta (Durnford et al. 1999). Structural studies of LHC proteins show that they are monophyletic and share a common architecture (Green and Durnford 1996).

In land plants, major LHCII proteins (associated primarily with PSII) consist of three closely related chlorophyll *a/b*-binding proteins including LHCII Type I, II, and III, which are encoded by *Lhcb1*, *Lhcb2* and *Lhcb3* genes, respectively (Jansson et al. 1992). Three minor monomeric LHC proteins or minor antennae, CP29, CP26, and CP24, encoded by *lhcb4*, *lhcb5* and *lhcb6* genes, respectively, are more closely associated with PSII than major LHCII proteins (Harrer et al. 1998). The major LHCII proteins are thus designated as peripheral LHCII proteins in the present study. In contrast to LHCII, the composition and structure of LHCI (associated with PSI) in land plants is not well established, but it is thought to be composed of four types of LHC proteins (Jansson 1999). LHCI Type I, II, III, and IV, encoded by *Lhca1*, *Lhca2*, *Lhca3* and *Lhca4* genes, respectively, have been suggested to dimerize into LHCI-730 (Type I and IV) and LHCI-680 (Type II and III) complexes (Lam et al. 1984). Two additional *Lhca* genes, *Lhca5* and *Lhca6*, were identified as EST clones in *Arabidopsis thaliana*, but their gene products have not yet been characterized (Jansson 1999).

Comprehensive understanding of the biology and genetics of photosynthesis in plants will require complete knowledge of the genes and gene products that mediate photosynthetic reac-

---

4 Corresponding author: E-mail, minagawa@lowtem.hokudai.ac.jp; Fax, +81-11-706-5493.
tions. Thus, additional studies are needed of the biochemistry, molecular biology and genetics of LHC proteins and their genes including studies of mechanisms of gene and protein regulation in response to environmental factors. The green alga *Chlamydomonas reinhardtii* has been widely used for studies of plant photosynthesis, metabolism, mating and flagellar movement (Harris 2001), and recently developed molecular techniques enhance the value of this unicellular alga as an excellent model organism (Rochaix 2002). Studies of photosynthesis and chloroplast biogenesis are particularly appropriate in *C. reinhardtii*, because of the availability of viable photosynthesis-deficient mutants that grow on acetate as sole carbon source. Genomics of *C. reinhardtii* are also rapidly improving and available resources include an EST library (Asamizu et al. 1999, Shrager et al. 2003), a high-density molecular map of the nuclear genome (Kathir et al. 2003), and an on-going genome sequencing project.

Earlier studies showed that four *LhcII* gene families, *LhcII-1*, *LhcII-2*, *LhcII-3*, and *LhcII-4*, encode peripheral LHCII proteins and *Lhcb4* and *Lhcb5* genes encode CP29 and CP26, respectively, in *C. reinhardtii* (Minagawa et al. 2001, Teramoto et al. 2001). However, homologs of *Lhcb6* gene encoding CP24 in higher plants have not been identified. The genetic phylogeny of LHCII proteins suggests that the ancestral LHC protein diverged into a gene family encoding peripheral LHCII proteins, CP29, and CP26, before green algae and higher plants diverged from each other. In contrast to *LhcII* genes, *C. reinhardtii* *LhcI* genes are relatively poorly characterized; two *LhcI* cDNAs have been isolated (Hwang and Herrin 1993, Durnford et al. 2003) and three putative cDNA sequences corresponding to *LhcI* genes have been deposited in GenBank to date (GenBank AF195794, AF104633, AF104632). However, four out of the five genes appear to be the same. Consequently, it remains to be confirmed that these reports correspond to unique genes and the gene products are biochemically associated with the LHCI complex.

This study presents the separation of seven distinct LHCI proteins from the purified PSI-LHCI supercomplex by high-resolution SDS-polyacrylamide gel electrophoresis, the isolation of the seven corresponding genomic clones, and the characterization of the complete *LhcI* gene family of *C. reinhardtii*. The phylogeny of entire gene family of *Lhc* including *LhcII* genes is discussed in the context of evolution of multiple diverse antenna complexes.

**Results and Discussion**

**Assignment of LHCI proteins in *C. reinhardtii***

In an earlier study, seven distinct *C. reinhardtii* ESTs were identified as possible members of the *C. reinhardtii* LHCI family (Teramoto et al. 2001). In this study, firstly the relationship between those ESTs and the LHCI proteins present in the thylakoid membranes was determined. Since the purified thylakoid membranes contain a large number of polypeptides that commonly migrate with LHCI proteins on SDS-polyacrylamide gel, we purified the PSI-LHCI supercomplex that retains all of the LHCI proteins as described in Materials and Methods. The LHCI proteins in the PSI-LHCI supercomplex were separated by SDS-PAGE with high resolution (Fig. 1). The purified fraction includes PSI polypeptides and seven bands (a through g) corresponding to LHCI. The three bands (b through d) migrated closely and the band e appeared to be faint and diffuse. However, it was confirmed by re-electrophoresis of each band that they are distinct protein bands (data not shown).

The seven LHCI proteins were eluted from the gel and subjected to N-terminal amino acid sequencing. It was revealed that the seven proteins contain unique N-terminal sequences (Table 1), indicating that *C. reinhardtii* LHCI is composed of at
least seven distinct polypeptides. Bassi et al. (1992) identified
seven putative \( C. \) reinhardtii LHCI polypeptides using two-
dimensional gel electrophoresis and reported their N-terminal
sequences. The N-termini of those proteins, designated p14,
p14.1, p15.1, p18, p18.1, p22, and p22.1, were unique with
the exception of polypeptides p22 and p22.1, which had different
immunological and electrophoretic properties but identical N-
termini (Bassi et al. 1992). The N-terminal sequence of the
bands \( a \), \( b \), \( c \), \( d \), \( e \), \( f \), and \( g \) match the reported sequences of p14,
p18.1, p14.1, p15.1, p18, and p22/p22.1, respectively (Table 1).

The results presented here and those of Bassi et al. (1992)
would be well consistent, if p22 and p22.1 are the products of
a single gene. This possibility is supported by the fact that the
\( C. \) reinhardtii EST database includes only a single gene contain-
ing the sequence of KAGNWLPGD (the N-terminus of p22
and p22.1). These observations are not without precedent in \( C. \)
reinhardtii, since a recent study by two-dimensional gel elec-
trophoresis coupled with mass spectrometric analysis showed
that a \( C. \) reinhardtii LHCCI protein has multiple forms (Hippler
et al. 2001). Furthermore, Hippler et al. (2001) identified eight-
een distinct \( C. \) reinhardtii protein spots (based on immuno-
reactivity) to LHCI on a two-dimensional immunoblot. They
suggested that differential processing and/or post-translational
modification could generate this heterogeneity. It is thus likely
that p22 and p22.1 might also be generated by differential
processing and/or post-translational modification.

In addition to these six distinct LHCI proteins, a new
LHCI protein was identified as band \( e \). The N-terminal
sequence (AAVRPVWFPG) has not been previously reported
in PSI and LHCI preparations presumably because this band is
diffuse and thus stained faintly, and migrates very closely with
the band \( f \) on the gel. We thus conclude that employment of the
purified PSI-LHCI supercomplex and the high-resolution SDS-
polyacrylamide gel electrophoresis enabled us to detect all
major LHCI proteins.

### Table 1: Properties of LHCI polypeptides in \( C. \) reinhardtii

<table>
<thead>
<tr>
<th>Band</th>
<th>N-terminal sequence</th>
<th>Gene</th>
<th>Protein (^a)</th>
<th>Transit peptide (^b)</th>
<th>Mature protein (^b)</th>
<th>Molecular weight (^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a )</td>
<td>SAVPE</td>
<td>( LhcI-1 )</td>
<td>p14 (^d)</td>
<td>28</td>
<td>236</td>
<td>25,770</td>
</tr>
<tr>
<td>( b )</td>
<td>ASSRPLWLP</td>
<td>( LhcI-5 )</td>
<td>p18.1</td>
<td>30</td>
<td>237</td>
<td>25,699</td>
</tr>
<tr>
<td>( c )</td>
<td>EEKSIAKVDR</td>
<td>( LhcI-2 )</td>
<td>p14.1 (Type III)</td>
<td>28</td>
<td>229</td>
<td>25,242</td>
</tr>
<tr>
<td>( d )</td>
<td>AADRKLWAPG</td>
<td>( LhcI-3 )</td>
<td>p15.1</td>
<td>26</td>
<td>217</td>
<td>23,271</td>
</tr>
<tr>
<td>( e )</td>
<td>AAVRPVWFPG</td>
<td>( LhcI-7 )</td>
<td>–</td>
<td>25</td>
<td>232</td>
<td>25,158</td>
</tr>
<tr>
<td>( f )</td>
<td>RQSWLPGSQI</td>
<td>( LhcI-4 )</td>
<td>p18</td>
<td>34</td>
<td>194</td>
<td>20,321</td>
</tr>
<tr>
<td>( g )</td>
<td>KAGNWLPGD</td>
<td>( LhcI-6 )</td>
<td>p22/p22.1 (Type I)</td>
<td>26</td>
<td>215</td>
<td>23,426</td>
</tr>
</tbody>
</table>

\(^a\) Bassi et al. (1992).
\(^b\) Deduced from nucleotide sequence and N-terminal sequence of mature protein.
\(^c\) Estimated molecular weight of mature protein.
\(^d\) According to Bassi et al. (1992), p14 has the N-terminal sequence of AA VPE. However, the mobility on a gel and the DNA sequence data sug-
gests that band \( a \) most likely corresponds to p14.

### Genes encoding LHCI proteins

The N-terminal sequences from the seven distinct LHCI
proteins in Table 1 have a one-to-one correspondence with
seven \( LhcI \) EST clones reported previously (Teramoto et al.
2001). In particular, bands \( a \) through \( g \) are the products of
\( LhcI-1 \), \(-5\), \(-2\), \(-3\), \(-7\), \(-4\), and \(-6\), respectively (Table 1). Based
on sequence similarity and stable association of the gene prod-
ucts to PSI-LHCI supercomplex, it can be concluded that \( C. \)
reinhardtii contains seven \( LhcI \) genes and the gene products
(bands \( a \) through \( g \)) are chlorophyll \( a/b \)-binding proteins
(Fig. 2). However, we cannot exclude the possibility that \( C. \)
reinhardtii has additional \( LhcI \) gene(s) that are expressed at a
very low level and therefore not represented in the EST library,
or the gene product(s) do not appear in the isolated PSI-LHCI
supercomplex. These yet-to-be-identified \( LhcI \) gene(s) may
encode polypeptides that were co-purified with some LHCI
preparation, such as p14.2, p15, p17.2, and p22.2 (Bassi et al.
1992). It will be possible to resolve these questions more com-
pletely when the complete sequence of the \( C. \) reinhardtii
nuclear genome becomes available and all \( Lhc \)-homologous
genes can be identified by direct homology searching. This
information will help establish the precise structure of the pho-
system antenna complexes in \( C. \) reinhardtii.

EST sequences are usually short so that they do not cover
entire coding regions, and are not sufficiently reliable to estab-
lish subtle sequence variations. In particular when we deal with
a very closely related multigene family, such as \( Lhc \) genes,
ESTs would not be suitable resources for further studies.
Therefore, we isolated genomic clones for each of the seven \( C. \)
reinhardtii \( LhcI \) genes. Genomic sequences including introns
and exons were amplified by PCR, cloned and sequenced. The
experiments were triplicated in order to exclude any errors dur-
ing the manipulations. The resulting sequences were deposited
in GenBank as \( LhcI-1 \) through \( LhcI-7 \) (GenBank AB122114–
AB122120). The cDNA sequences of several \( C. \) reinhardtii
\( LhcI \) genes were already present in GenBank including the fol-
lowing: cabI-1 (Hwang and Herrin 1993, GenBank X65119),

---
C. reinhardtii LHCI family 141

Transit peptides
LhcI-1
LhcI-2
LhcI-3
LhcI-4
LhcI-5
LhcI-6
LhcI-7

CAB6A (GenBank AF19579), Lhca1 (GenBank AF104633), Lhca2 (GenBank AF104632), and mRNA for Lhca (Durnford et al. 2003, GenBank AY171231). The first four sequences are identical to LhcI-6 and the last sequence is identical to LhcI-7. The sequences of LhcI-1, -2, -3, -4, and -5 were novel at the time of submission to the GenBank database. The untranslated regions of these genes are not homologous to each other; however, the coding regions of these genes are well conserved (Fig. 2). Nucleotide sequences of the introns were also determined.

LhcI-1 to LhcI-7 genes contain 8, 7, 5, 6, 5, 6, and 7 introns, respectively, whose locations vary (see Fig. 4). All the introns fulfill the 'GT-AG' rule.

The mature LHCI proteins encoded by LhcI-1 to LhcI-7 are composed of 236, 237, 229, 217, 194, 215 amino acid residues (20.3–25.8 kDa, Table 1), respectively. Genomic nucleotide sequences and N-terminal amino acid sequences (Table 1) were used to deduce that the putative transit peptides of the seven LHCI proteins are 25–34 residues long (Fig. 2, Table 1).

Putative structural elements in LHCI proteins were assigned based on the reported high resolution three-dimensional structure of pea LHClII (Kühlbrandt et al. 1994). Fig. 2 indicates putative structural elements of the seven LHCI proteins, including a transit peptide, three transmembrane helices, an amphipathic helix near the C-terminus, and chlorophyll ligands in LHCI (Kühlbrandt et al. 1994). Transit peptides were deduced from N-terminal sequences of mature proteins.

Fig. 2  Structures of seven C. reinhardtii LhcI proteins. Sequences were aligned using ClustalW (Thompson et al. 1994). Positions of introns are indicated by vertical bars. Gaps (—) were introduced to maximize the extent of homology. Hatched regions correspond to putative membrane-spanning (A–C) and amphiphatic (D) helices. Shaded letters indicate residues identified as chlorophyll ligands in LHCI (Kühlbrandt et al. 1994). Transit peptides were deduced from N-terminal sequences of mature proteins.
between helix C and A, which suggests that this protein may have a unique structural/functional role in LHCI.

**Characterization of the Lhc gene family in *C. reinhardtii***

Revealing the complete set of the *LhcI* genes in *C. reinhardtii* enabled us to carry out a comprehensive study on all members of a *Lhc* gene family from a chlorophyte. Fig. 3 shows a neighbor-joining phylogenetic tree (Saitou and Nei 1987) for *C. reinhardtii* LHCI and II protein sequences. LI818 protein, distantly related to the LHC protein family (Savard et al. 1996) was used as an outgroup for this phylogeny. While LHCII proteins are relatively homogeneous, which have a high overall level of sequence identity of 70–80% (Teramoto et al. 2001), LHCI proteins are rather heterogeneous and more diverged from LHCII proteins (Fig. 3). For instance, the overall identity between LhcI-1 and LhcI-2 through -7 (including transit peptide sequences) was, 41.1%, 50.3%, 46.3%, 48.9%, 38.9%, and 46.3%, respectively. The large diversity of LHCI proteins could indicate that the LHCI family diversified before the divergence of LHCII (Hwang and Herrin 1993). However, the extent of divergence is not always proportional to rate of evolution, and higher diversity in the LHCI protein family, which is also observed in higher plants (Jansson 1999), might reflect functional specialization of LHCI proteins.

**Phylogenetic implications**

LHC protein sequences of the green alga *C. reinhardtii* were also compared to those from vascular plants and red algae, to consider the divergence of these genes in the context of the phylogenetic separation between Chlorophyceae (green algae) and Streptophyta (terrestrial plants and their closest green algal relatives, the charophytes; Bremer 1985), or Chlorophyta (green plants) and Rhodophyta (red algae). For this comparison, it is important to have complete sets of sequences to exclude a possibility that paralogous genes are compared. The identification of the true LHC members in the PSI-LHCI supercomplex and the thorough nucleotide sequencing of the relevant genes in this study enabled us to accomplish this comparison for the first time.

Fig. 4 shows a neighbor-joining tree summarizing the relationships of representative LHC proteins from *A. thaliana* (vascular plant) whose nuclear genome has been completely sequenced, and *C. reinhardtii* (green alga). Seven major clades were observed including four for the components of LHCII complex (labeled CP29, CP26, CP24 and peripheral LHCII in Fig. 4) and three for the LHCI complex (labeled LHCI Type I, LHCI Type III, and peripheral LHCI in Fig. 4). These clades were thus suggested to have emerged before Chlorophyceae and Streptophyta diverged from each other. The three minor antennae are encoded by the *Lhcb4*, *Lhcb5*, and *Lhcb6* genes.

While the *Lhcb5* genes encoding CP26 were not duplicated or diverged in *C. reinhardtii* and *A. thaliana*, the *Lhcb4* gene encoding CP29 was triplicated only in *A. thaliana*. The *Lhcb6* gene encoding CP24 is not found in *C. reinhardtii*, which may indicate that this gene was originally present but then lost, or that it never diverged in green algae. The longer branch for At.Lhcb6 in Fig. 4 tends to support the former hypothesis. As has been previously described, the peripheral LHCII in *C. rein-
are encoded by the LhcII genes, which diverged into several “types” after the phylogenetic separation of Chlorophyceae and Streptophyta (Teramoto et al. 2001).

The clades for LHCI complex include one for Cr.LhcI-6 and At.Lhca1 (labeled LHCI Type I), and one for Cr.LhcI-2 and At.Lhca3 (labeled LHCI Type III, Fig. 4). These two branches are sufficiently divergent to be considered different types. Since At.Lhca1 and At.Lhca3 encode LHCI Type I and Type III protein, respectively, which are well conserved among higher plants (Jansson 1999), Cr.LhcI-6 and Cr.LhcI-2 are most likely those counterparts in C. reinhardtii. This could be further supported by the facts that Cr.LhcI-6 and Cr.LhcI-2 are more abundant than “peripheral LHCII” in C. reinhardtii (Hippler et al. 2001) and that both Cr.LhcI-6 and Cr.LhcI-2 proteins conserve a specific feature of the respective ‘Type’ observed in higher plants such as a very short sequence between Helix C and A (Type I) and a six-residue insertion at the beginning of Helix B (Type III) (Fig. 2, Pichersky and Jansson 1996). Another large LHCI lineage includes five C. reinhardtii LHCI proteins, LhcI-1, -3, -4, -5, and -7, and four of the A. thaliana LHCI proteins, Lhca2, Lhca4, Lhca5, and Lhca6. A comparable organization is observed for the proteins in LHCII complex, where minor antennae form separate branches from other peripheral LHCII proteins. As an analogy to the minor antennae, which lie in-between peripheral LHCII and PSII core groups (Harrer et al. 1998), LHCI Type I and Type III might function inner antennae for PSI, and the other five LHCI polypeptides in C. reinhardtii might work as peripheral antennae for LHCI. Recently, a three-dimensional model of the PSI-LHCI supercomplex from C. reinhardtii was reported based on a single particle analysis of images of the purified complex (Kargul et al. 2003). The total number of chlorophylls in the complex and the volume of the electron density suggested that 11 LHCI polypeptides lie on one side of the PSI core. It will be of interest to correlate the information in this study with this structural model.
One of the remaining questions about chloroplast diversity concerns the origin of LHCII. Rhodophyta, like cyanobacteria, only have chlorophyll a and use phycobilisomess as the light-harvesting antennae for PSII, while PSI utilizes chlorophyll-based LHC proteins as antennae (Wolfe et al. 1994). It has been proposed that Rhodophyta and Chlorophyta are directly descended from a single common ancestor (Burger et al. 1999, Moreira et al. 2000), although this idea is controversial (Stillier and Hall 1997). If Rhodophyta arose after the split of LHCII, the LHCI proteins in Rhodophyta would be clustered with those of Chlorophyta in a phylogenetic tree. Fig. 5 shows phylogenetic relationships between LHC proteins of \textit{C. reinhardtii} and \textit{Galdieria sulphuraria}, the thermoacidophilic red alga whose \textit{Lhc} genes are best characterized. Fig. 5 clearly shows that \textit{C. reinhardtii} genes encoding LHCI and LHCII proteins are more closely related to one another than to any of the chlorophyll \textit{a}-binding LHC proteins from \textit{G. sulphuraria} including \textit{Lhc}r-1, -2, -3, -4, and -5 (Marquardt et al. 2001). This implies that the chlorophyll \textit{a}-binding LHC proteins diverged from the ancestral LHC lineage prior to the functional separation of the LHCI and LHCII components. This in turn suggests that the ancestral LHCI polypeptides have replaced the phycobilisomes during evolution, and thus evolved LHCII in Chlorophyta. This hypothesis agrees with previous reports suggesting that the lineage of fucoxanthin chlorophyll \textit{a/c}-binding proteins in Chromophyta emerged prior to the separation of LHCII (Durnford et al. 1996, Caron et al. 1996); thus, the chromophytic plastid may have evolved secondarily from a red algulike eukaryote (Durnford et al. 1999).

It has been suggested that the LHC of the primitive green alga \textit{Mantoniella} (Prasinophyceae) delivers excitation energy to both PSI and PSII (Schmitt et al. 1993), and that \textit{Mantoniella} LHC sequences are not particularly related to LHCI or LHCII (Durnford et al. 1999). This implies that the same LHC proteins might have functioned as antenna systems for PSI and PSII in the ancestral green algae. Our current hypothesis for the evolution of LHC proteins is as follows: (1) Rhodophyta diverged from an ancestral photosynthetic eukaryote with ancestral LHC proteins, which served as an antenna system for PSI, while phycobilisomes served as an antenna system for PSII; (2) when Chlorophyta diverged, the ancestral LHC proteins assumed the functional role of phycobilisomes, and served as antenna systems for both photosystems; (3) subsequently, primitive LHC proteins acquired functional differentiation, which included inner antennae for PSI (LHCI Type I and type III), peripheral LHCI, inner antennae for PSII (CP29, CP26, and CP24), and peripheral LHCII; and (4) after the Chlorophyceae/Streptophyta split, \textit{Lhc} genes differentiated further by divergence and/or gene duplications in order to adapt to various light regimes.

### Materials and Methods

**Strains and culture conditions**

Wild type \textit{C. reinhardtii} (strain 2137) was obtained from Chlamydomonas Genetics Center (Durham, NC, U.S.A.). Cells were grown in Tris-acetate-phosphate medium (Gorman and Levine 1965) under continuous illumination (5 \(\mu\text{E m}^{-2}\text{s}^{-1}\)) at 26°C.

**Purification of PSI-LHCI supercomplex and SDS-polyacrylamide gel electrophoresis**

Thylakoid membranes (0.8 (mg Chl) ml\(^{-1}\)) were purified and solubilized with \textit{n}-dodecyl \(\beta\)-\textit{o}-maltoside (0.8% w/v) as described in Takahashi et al. (1991). The thylakoid extracts were subjected to sucrose density gradient centrifugation and the resulting A-3 fraction enriched in PSI-LHCI supercomplex was collected (Takahashi et al. 1991). The PSI-LHCI supercomplex was furthermore separated by DEAE column chromatography with the buffer containing 50 mM Tris-\(\text{HCl, pH}\) 8.0, 0.05% DM and a linear gradient of NaCl (25–175 mM). SDS-polyacrylamide gel electrophoresis that separates LHCI proteins with high resolution was carried out as described previously (Fling and Gregerson 1986). The resolving gel contained 15–22.5% gradient acrylamide.

**Microsequencing of polypeptides**

LHCI proteins were detected by staining, excised from the gel, eluted electrophoretically and absorbed on to Prosorb membranes (Applied Biosystems, Foster City, CA, U.S.A.). Peptides were analyzed by automated Edman degradation on a Protein Sequencer model PPSQ-21A (Shimadzu Corp., Kyoto, Japan).

**Genomic DNA preparation**

\textit{C. reinhardtii} cells were grown in liquid culture, harvested by centrifugation, and resuspended in buffer containing 50 mM Tris-\(\text{HCl, pH}\) 8.0, 100 mM EDTA, and 1% SDS. The cell suspension was incubated at 50°C for 2 h, extracted twice with phenol-chloroform-isooamyl alcohol (25 : 24 : 1), and precipitated with isopropanol.

**Cloning of genomic clones for \textit{C. reinhardtii LhcI} genes**

Genomic clones for \textit{C. reinhardtii LhcI} genes were isolated by PCR as described previously (Teramoto et al. 2001) and ligated into pBluescriptII SK+ (Stratagene, La Jolla, CA, U.S.A.). Three independent clones were determined for each \textit{LhcI} gene. Sequence alignment and phylogenetic analysis was performed with CLUSTAL W (Thompson et al. 1994) using the Neighbor-Joining method (Saitou and Nei 1987). Phylogenetic trees were displayed using the program NJ plot (Petrie and Gouy 1996).

**Acknowledgments**

We would like to thank Dr. B.R. Green for valuable discussion. I.M. acknowledges financial support from The Akiyama Foundation and CREST of JST (Japan Science and Technology Corporation).
References


Chlamydomonas reinhardtii LHCI family


(Received October 15, 2003; Accepted November 30, 2003)