

1 Zusammenfassung

1.1 Deutsche Zusammenfassung

Die Kolonisierung der terrestrischen Habitate durch streptophytische Algen gehört zu den wichtigsten Evolutionsschritten auf der Erde. Die Landpflanzen (Embryophyta) haben sich aus den Vorläufern der rezenten streptophytischen Grünalgen (früher Charophyceae) entwickelt. Viele typische Merkmale und Eigenschaften der Landpflanzen sind in dieser Gruppe von Süßwassergrünalgen entstanden. Zurzeit werden zwei Gruppen als mögliche Schwestergruppe der Embryophyten vorgeschlagen: die Charales oder die Coleochaetales.

Die Viridiplantae (Grünalgen und Landpflanzen) bestehen aus zwei monophyletischen Linien: den Chlorophyta und den Streptophyta. Die meisten Grünalgen gehören zu den Chlorophyten, während die Streptophyten alle Landpflanzen und eine diverse paraphyletische Ansammlung von Süßwassergrünalgen (Charophyceae) umfassen. Das Ziel der vorliegenden Arbeit war die Entwicklung der Landpflanzen aus den streptophytischen Algen zu untersuchen.

Zu Beginn dieser Arbeit wurden cDNA-Banken von drei streptophytischen Algen hergestellt: *Klebsormidium subtile* (4651 ESTs, 1554 Contigs), eine filamentöse Alge; *Coleochaete scutata* (5039 ESTs, 2303 Contigs), ebenfalls eine filamentöse Alge, die aber Verzweigungen und apikales Wachstum aufweist und *Chara vulgaris* (1431 ESTs, 563 Contigs), die eine sehr hoch organisierte streptophytische Alge repräsentiert. Für die Auswertung der streptophytischen ESTs sollte ein möglichst ursprünglicher Vertreter der Chlorophyten in die Analyse einbezogen werden. Daher wurde auch eine EST-Bank aus der einzelligen Grünalge *Pyramimonas parkeae* (5094 ESTs, 1580 Contigs), die an der frühen Evolutionslinie der Chlorophyten steht, etabliert. Die Vergleiche der Contigs mit den Genomen von *Arabidopsis thaliana* und *Chlamydomonas reinhardtii* zeigten wie erwartet, dass alle streptophytischen Algen eine größere Ähnlichkeit zum *Arabidopsis*- als zum *Chlamydomonas*-Genom haben. Erstaunlich ist nur, dass auch *Pyramimonas* eine größere Anzahl an Genen besitzt, die ähnlicher zum *Arabidopsis*- als zum *Chlamydomonas*-Genom sind. Diese Ergebnisse unterstreichen den ursprünglichen Charakter von *Pyramimonas*. In den weiteren Analysen wurden 515 Contigs der *Pyramimonas*-Bank mit signifikanter Ähnlichkeit zu bekannten oder hypothetischen Proteinen identifiziert. 920 Contigs der *Klebsormidium*-Bank konnten bekannten Proteinen zugeordnet werden. Es konnten 743 Contigs aus der *Coleochaete*-Bank und 138 Contigs aus der *Chara*-Bank Ähnlichkeiten zu bekannten Proteinen zugewiesen

werden. Die meisten Contigs der untersuchten Algen mit einer Ähnlichkeit zu bekannten Proteinen konnten dem Kohlehydrat-, Energie-, Lipid- und Aminosäure-Metabolismus, sowie Translation und Signalwegen zugeordnet werden (KEGG- und COG-Datenbank). Phylogenetische Analysen von 124 orthologen Proteinen (28765 Aminosäurepositionen) aus 40 verschiedenen Taxa der Viridiplantae zeigte, dass *Chara* die Schwestergruppe der Embryophyta zu sein scheint. Diese Position von *Chara* (Maximum-Likelihood-Analysen, Bayesische Analyse) sowie die Position von *Coleochaete* (Maximum-Likelihood-Analysen) werden in den Analysen jedoch nicht statistisch unterstützt.

Neben der globalen Analyse der gesamten Genome wurden im zweiten Teil der Arbeit zwei spezielle Fragestellungen bearbeitet:

- Bei den meisten eukaryotischen Genen wird an das 3'-Ende der kernkodierte prä-mRNA ein Poly(A)-Schwanz angehängt. In Tieren, Pilzen und Embryophyten enthält das Polyadenylierungssignal eine A-reiche Sequenz (meistens AAUAAA oder verwandte Sequenz), die 13 bis 30 bp oberhalb der „cleavage site“ (CS) zu finden ist, welche allgemein als das „near upstream element“ (NUE) bezeichnet wird. Bei einigen Chlorophyceen wird das Pentanukleotid UGUAA als Polyadenylierungssignal genutzt. ESTs von streptophytischen und chlorophytischen Algen wurden auf Unterschiede in Polyadenylierungssignalen untersucht. Die Polyadenylierungssignale variieren weitgehend in den Grünalgen. In abgeleiteten Chlorophyten wurde ein neues NUE (UGUAA) erfunden und ersetzte nicht nur das A-reiche NUE, sondern das komplette Poly(A)-Signal in allen untersuchten Chlorophyten außer *Scherffelia* (nur NUE ersetzt), *Pyramimonas* und *Ostreococcus* (UGUAA fehlt vollständig). Das UGUAA-Element fehlt komplett in Streptophyten. Die Struktur des Poly(A)-Signals wurde in streptophytischen Algen jedoch oft modifiziert. Bei den meisten untersuchten Arten fehlt ein A-reiches NUE; stattdessen scheinen diese Arten hauptsächlich U-reiche Elemente zu nutzen.
- Saccharose findet man in vielen Pflanzen, in denen sie unterschiedliche Funktionen erfüllt: Einerseits ist Saccharose ein Hauptprodukt der Photosynthese. Pflanzen lagern zusätzlich Saccharose ein, damit sie umweltbedingtem Stress wie z.B. niedrigen Temperaturen oder Trockenheit entgegenwirken können. Im Gegensatz dazu wird in Cyanobakterien und anderen Algen vermutet, dass die Saccharose eine Rolle bei der Bewältigung von Umweltstress spielt. Gene für die Saccharose-Synthese wurden in *Klebsormidium* (Saccharose-Phosphat-Synthase (SPS)) und in *Coleochaete* und *Chara* (Saccharose-Synthase (SuS), SPS) identifiziert. In öffentlichen Datenbanken konnten Proteine mit Ähnlichkeit zu SPS in den Grünalgen *Chlamydomonas*, *Volvox* und *Chlorella*

und in den Rotalgen *Cyanidischyzon* und *Galdieria* gefunden werden. Die phylogenetischen Analysen haben gezeigt, dass die Proteine sowohl der Grünalgen und Landpflanzen (SuS) als auch der Rotalgen, Grünalgen und Landpflanzen (SPS) monophyletisch sind. Überraschend war aber auch, dass die Algen und Landpflanzen beide Proteine (SuS und SPS) von einem Proteobakterium und nicht von einem Cyanobakterium geerbt haben. Ebenfalls wurde der Saccharosegehalt in verschiedenen Algen biochemisch bestimmt. Saccharose ist sowohl in Cyanobakterien, in chlorophytischen und streptophytischen Algen als auch in Landpflanzen vorhanden. Um die Funktion von Saccharose in streptophytischen Algen zu untersuchen, wurden Tagesprofile des Saccharosegehalts von verschiedenen Algen erstellt. Nur in den untersuchten streptophytischen Algen konnte eine Veränderung des Saccharosegehalts im Laufe des Tages beobachtet werden. Im Gegensatz dazu blieb der Saccharosegehalt in allen anderen untersuchten Algen während der 24h konstant. Diese Ergebnisse deuten auf eine andere Funktion in den streptophytischen Algen als in den übrigen Algen hin.

1.2 Englische Zusammenfassung/Abstract

The colonization of the terrestrial habitat by streptophyte algae was one of the most important steps during evolution on earth. Land plants (embryophytes) evolved from the ancestors of the extant streptophyte algae. Many typical characteristics of the Embryophyta were developed within this group of fresh water algae. Currently two groups are considered to be the sister group of the embryophytes: the Charales or the Coleochaetales.

The Viridiplantae (green algae and land plants) consist of two monophyletic lineages: the Chlorophyta and the Streptophyta. Most green algae belong to the Chlorophytes, while the Streptophytes include all land plants and a small group of freshwater algae known as Charophyceae. The aim of this work was to analyze the evolution of the land plants from the streptophyte algae.

At the beginning of this work cDNA libraries were prepared from three streptophyte algae: *Klebsormidium subtile* (4651 ESTs, 1554 contigs) a filamentous alga; *Coleochaete scutata* (5039 ESTs, 2303 contigs) a highly developed streptophyte alga; and *Chara vulgaris* (1431 ESTs, 563 contigs) morphologically, the most highly developed streptophyte alga. For comparison with chlorophytes, ESTs from *Pyramimonas parkeae* (5094 ESTs, 1580 contigs), an early branching chlorophyte, were also sequenced. Comparison of the contigs with the genomes of *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* showed that the contigs of

all analysed streptophyte algae were more similar to sequences in *Arabidopsis* than to those in *Chlamydomonas*. Surprisingly, *Pyramimonas* has a higher number of genes, that are more similar to *Arabidopsis* than to *Chlamydomonas*. These results revealed the early branching of *Pyramimonas*. The analyses identified 515 contigs in the *Pyramimonas* library, 920 contigs in the *Klebsormidium* library, 743 contigs in the *Coleochaete* library and 138 contigs in the *Chara* library with significant similarity to known sequences. Most of the contigs of the analysed algae with similarity to known proteins are associated with the carbohydrate-, energy-, lipid- and amino acid metabolisms. Others are related to Translation and Signal Transduction (KEGG and COG Database). Phylogenetic analysis of 124 orthologous proteins (28765 amino acid positions) from 40 different taxa of Viridiplantae showed that *Chara* (Charales) seemed to be the sister group of the embryophytes. However, neither this position of *Chara* (in the maximum-likelihood and Bayesian analysis) nor the position of *Coleochaete* (in the maximum-likelihood analysis) is statistically supported (in the present study).

The EST information has been used also in two different projects:

- Eukaryotes attach poly-A tails to the 3' ends of most nuclear-encoded mRNAs. In animals, fungi and embryophytes, the signal for polyadenylation contains an A-rich (often AAUAAA or related sequence) sequence 13 to 30 nucleotides upstream from the cleavage site, which is commonly referred to as the near upstream element (NUE). However, it has been reported that the pentanucleotide UGUAA is used as polyadenylation signal for some genes in volvoclean algae. Expressed genes (ESTs) from streptophyte and chlorophyte algae were analyzed for differences in polyadenylation signals. Polyadenylation signals vary widely in green algae. A new NUE (UGUAA) was invented in advanced chlorophytes and replaced not only the A-rich NUE but the complete poly(A) signal in all chlorophytes investigated except *Scherffelia* (only NUE replaced), *Pyramimonas* and *Ostreococcus* (UGUAA completely missing). The NUE is completely absent from streptophytes. However, the structure of the poly(A) signal was often modified in streptophyte algae. In most species investigated, an A-rich NUE is missing; instead, these species seem to rely mainly on U-rich elements.
- Sucrose is universal in plants and fulfils many roles: It is a main product of photosynthesis and used as transport metabolite to sustain heterotrophic tissues. In addition, plants accumulate large amounts of sucrose in response to environmental stress such as low temperature and draught. In contrast, in cyanobacteria and algae, sucrose has been mainly proposed to function as an osmo- and cryoprotectant. ESTs encoding proteins for the sucrose biosynthesis were obtained for *Klebsormidium* (sucrose phosphate synthase

(SPS)) and *Coleochaete* and *Chara* (sucrose synthase (SuS), SPS). Database searches showed that proteins with similarity to SPS are also present in some green algae *Chlamydomonas*, *Volvox* und *Chlorella* and red algae *Cyanidischyzon* and *Galdieria*. A phylogenetic analysis of all available SPS and SuS was performed. In all phylogenetic analyses the obtained trees show that the proteins from green algae and land plants (SuS), and red algae, green algae and land plants (SPS) are monophyletic. Surprisingly the trees indicate that both plant and algal proteins (SuS and SPS) were inherited from proteobacteria and not from cyanobacteria. The sucrose content of various algae was also biochemically determined. Sucrose is present in cyanobacteria, chlorophyte and streptophyte algae and land plants. To test the function of sucrose in streptophyte algae, a daily profile of the cellular sucrose content in various algae were analyzed. The only algae where changes in the sucrose content during the day were observed were streptophyte algae. In contrast, in all other tested algae and cyanobacteria the sucrose content remained constant during a 24h daily cycle. This is in agreement with a different function of sucrose in streptophyte algae than in chlorophyte algae and cyanobacteria.