ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Gene 394 (2007) 96–104
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Intron distribution in Plantae: 500 million years of stasis during land plant evolution

René Teich [a,b], Carina Grauvogel [a,c], Jörn Petersen [a,*]

[a] *Institut für Genetik, Technische Universität Braunschweig, D-38106 Braunschweig, Germany*
[b] *Department of Clinical Chemistry and Molecular Diagnostics, Biomedical Research Center, Philipps-University of Marburg, D-35043 Marburg, Germany*
[c] *Georg-Speyer-Haus, Johann Wolfgang Goethe Universität Frankfurt, D-60596 Frankfurt am Main, Germany*

Received 18 January 2007; received in revised form 12 February 2007; accepted 13 February 2007
Available online 24 February 2007
Received by M. Di Giulio

Abstract

Little is known about the evolution of the intron–exon organization in the more primitive groups of land plants, and the intron distribution among Plantae (glauco-, rhodo-, chloro- and streptophytes) has not been investigated so far. The present study is focused on some key species such as the liverwort *Marchantia polymorpha*, representing the most ancient lineage of land plants, and the streptophycean green alga *Mesostigma viride*, branching prior to charophycean green algae and terrestrial plants. The intron distribution of six genes for sugar phosphate metabolism was analyzed including four different glyceraldehyde-3-phosphate dehydrogenases (*GAPDH*), the sedoheptulose-1,7-bisphosphatase (*SBP*) and the glucose-6-phosphate isomerase (*GPI*). We established 15 new sequences including three cDNA and twelve genomic clones with up to 24 introns per gene, which were identified in the *GPI* of *Marchantia*. The intron patterns of all six genes are completely conserved among seed plants, lycopods, mosses and even liverworts. This intron stasis without any gain of novel introns seem to last for nearly 500 million years and may be characteristic for land plants in general. Some unique intron positions in *Mesostigma* document that a uniform distribution is no common trait of all streptophytes, but it may correlate with the transition to terrestrial habitats. However, the respective genes of chlorophycean green algae display largely different patterns, thus indicating at least one phase of massive intron rearrangement in the green lineage. We moreover included rhodophyte and glaucophyte reference sequences in our analyses and, even if the well documented monophyly of Plantae is not reflected by a uniform intron distribution, at least one *GPI* intron is strictly conserved for 1.5 billion years.

Keywords: Intron gain and loss; Gene duplication; Glyceraldehyde-3-phosphate dehydrogenase; Glucose-6-phosphate isomerase; Sedoheptulose-1,7-bisphosphatase; *Marchantia polymorpha*; *Mesostigma viride*

1. Introduction

Introns are regions of genes that are removed from the corresponding RNA. Spliceosomal introns are universally and exclusively distributed among eukaryotes. Their ancestry is unclear, but a common origin with selfsplicing group II introns that have been identified in bacterial and organellar genomes has been discussed (reviewed by Roy and Gilbert, 2006). Introns were initially regarded as junk DNA, but especially in metazoa and plants, they mediate a more complex gene regulation via alternative splicing, leading to a broad variety of products from a given gene set (Xing and Lee, 2006). Moreover, it had been proposed that primordial genes originated from the rearrangement of functional modules by the recombination of exons (exon shuffling; e.g. Pathy, 1999). Irrespective of the long-standing “intron early” versus “intron late” debate (e.g. Gilbert, 1987; Stoltzfus et al., 1994), both massive intron gain and loss have been documented in eukaryotic evolution (Rogozin et al., 2003; Nielsen et al., 2004). Several mechanisms for both events were proposed, e.g. recombination with genomic DNA or reverse transcribed RNA (see Roy and

---

Abbreviations: cDNA, DNA complementary to RNA; EST, expressed sequence tag; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GPI, glucose-6-phosphate isomerase; kb, kilobases; Myr, million years; PCR, polymerase chain reaction; RT, reverse transcription; SBP, sedoheptulose-1,7-bisphosphatase.

* Corresponding author. Tel.: +49 531 391 5785; fax: +49 531 391 5765.
E-mail address: j.petersen@tu-bs.de (J. Petersen).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.02.011

Gilbert, 2006). Gain and loss occur with specific frequencies and each may predominate in different evolutionary periods (this study). Large scale comparisons of intron positions between available genomes of metazoa, fungi and plants propose remarkable double digit conservation rates (10%–15%) between these lineages (Fedorov et al., 2002; Rogozin et al., 2003). However, such studies essentially base on the parsimonious assumption that shared intron positions reflect ancestral characters, but they do not account for multiple events of intron gain (Nielsen et al., 2004). The extent of independent intron insertion into the same position possibly facilitated by so-called proto-splice sites (Dibb and Newman, 1989) is difficult to determine. A broad taxon sampling is helpful to investigate the origin and evolution of specific introns, since a widespread presence would support a common ancestry, whereas a sporadic occurrence rather favors independent gains.

In contrast to various studies of the gene structure among metazoa and fungi (e.g. Müller et al., 2002; Rogozin et al., 2003; Nielsen et al., 2004), the number of large-scale intron analyses in plants is limited and restricted to seed plants (Knowles and McLysaght, 2006; Roy and Penny, 2007). However, single gene comparisons of glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), phosphoribulokinase (*PRK*) and sedoheptulose-1,7-bisphosphatase (*SBP*) revealed conspicuous differences of the intron pattern between the chlorophycean green alga *Chlamydomonas reinhardtii* and land plants (Kersanach et al., 1994; Petersen et al., 2006a; Teich et al., in press). Since chlorophytes and land plants have a common origin (Viridiplantae; Fig. 1), these findings accordingly propose an active phase of intron rearrangement. Land plants emerged from charophycean green algae and both are designated as Streptophyta, whereas Chlorophyta comprise the remaining lineages of green algae (Chlorophyceae, Ulvophyceae, Trebouxiophyceae, Prasinophyceae). With respect to streptophytes, the liverwort *Marchantia polymorpha* and the green alga *Mesostigma viride* represent two evolutionary key species (Fig. 1). Liverworts are proposed to be the most ancient lineage of land plants that arose more than 475 Myr ago (Qiu et al., 2006; Wellman et al., 2003) and *Marchantia* is probably the best investigated representative of liverworts (Ohyama et al., 1986; Oda et al., 1992). The phylogenetic position of *Mesostigma* was controversially discussed for a long time (McCourt et al., 2004), but in recent studies it has been classified as a basally branching streptophyte (Simon et al., 2006; Petersen et al., 2006b; Rodriguez-Ezpeleta et al., 2007).

In the present study we compared the intron distribution of altogether six genes for sugar phosphate metabolism in order to investigate the extent of intron gain and loss among Plantae. Controversial assessments of gain versus loss that are based on large-scale genome comparisons (e.g. Fedorov et al., 2002; Qiu et al., 2004) may either be caused by different statistic approaches or by the limited amount of completely sequenced taxa. We established six more comprehensive data sets with crucial reference species in order to investigate the intron distribution in (i) terrestrial plants (including the liverwort *Marchantia*) and (ii) streptophytes (incl. *Mesostigma*). Since green plants are an integral part of the monophyletic superensemble Plantae (Glaucophyta, Rhodophyta, Viridiplantae; Fig. 1) also designated as Archaeplastida (Adl et al., 2005), which originated in a single primary endosymbiosis with a cyanobacterium (Rodriguez-Ezpeleta et al., 2005), we also investigated the intron distribution in (iii) chlorophytes, rhodophytes and in the glaucophyte *Cyanophora paradoxa*.

2. Materials and methods

2.1. Plant material

The liverwort *M. polymorpha* and the glaucophyte *C. paradoxa* (strain SAG 29.80 [Pringsheim isolate]) were obtained from the Deutsche Sammlung für Mikroorganismen and Zellkulturen (DSMZ, Braunschweig) and the Sammlung für Algenkulturen (SAG, Göttingen), respectively. Culturing was

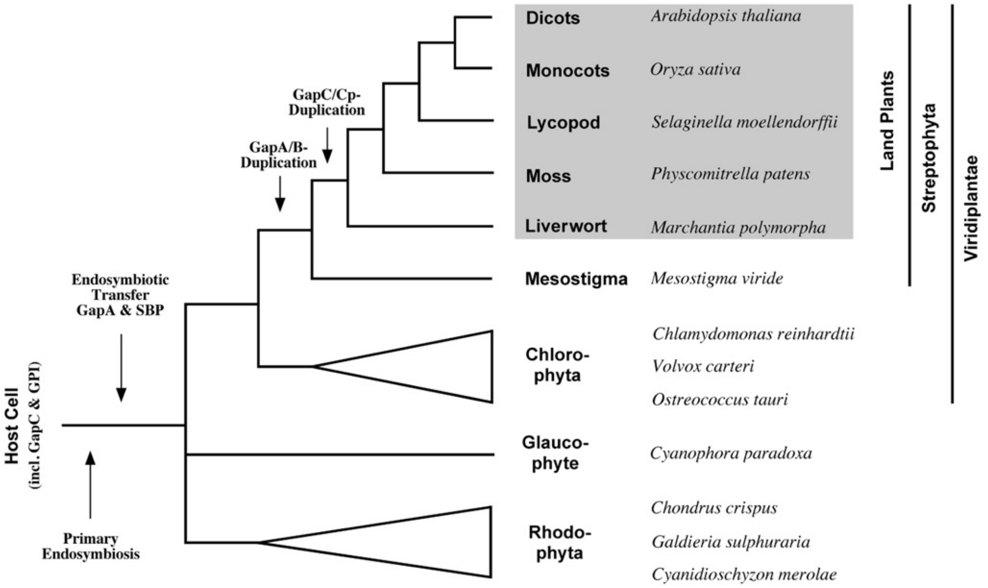

Fig. 1. Evolution of Plantae after the primary plastid endosymbiosis. Lineages of land plants are highlighted with a gray box.