

***Oenothera*, a unique model to study the role  
of plastids in speciation**

Dissertation der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München

vorgelegt von  
**Stephan Greiner**

am 15. Mai 2008

Erstgutachter: Professor Reinhold G. Herrmann

Zweitgutachter: Professor Wolfgang Stephan

Tag der mündlichen Prüfung: 19. Juni 2008

# Table of Contents

<b>1. Introduction .....</b>	<b>1</b>
1.1. Eukaryotic genomes are integrated and compartmentalized.....	1
1.2. Dobzhansky-Muller incompatibilities and asymmetric hybridization barriers.....	2
1.2.1. The model of Dobzhansky-Muller incompatibility.....	4
1.2.2. “Speciation genes” have not yet been identified for PGI.....	5
1.3. Hybridization barriers formed by plastids.....	6
1.4. The occurrence of PGI in natural populations is underestimated .....	10
1.5. Physiology and cell biology of PGI .....	14
1.5.1. Albinotic phenotypes of PGI.....	15
1.5.2. PGI phenotypes with affected cell growth and function .....	15
1.6. <i>Oenothera</i> as a molecular model to investigate PGI.....	17
1.7. <i>Oenothera</i> genetics.....	19
1.7.1. Complete reciprocal translocations of whole chromosome arms in <i>Oenothera</i> .	19
1.7.2. Maintenance of complete permanent translocation heterozygosis.....	22
1.7.3. Exchanging plastomes between species .....	23
1.8. Aims of this work.....	25
<b>2. Material and Methods.....</b>	<b>26</b>
2.1. Material .....	26
2.1.1. Chemicals .....	26
2.1.2. Solutions, buffers and media.....	26
2.1.3. Antibodies .....	26
2.1.4. Oligonucleotides.....	27
2.1.5. Reference species for bioinformatic analysis.....	28
2.1.6. <i>Oenothera</i> strains .....	29
2.1.7. <i>Arabidopsis</i> strains .....	36

---

2.1.8. Bacterial strains .....	36
2.2. Methods .....	36
2.2.1. Growth of biological material .....	36
2.2.1.1. <i>Oenothera</i> growth conditions and breeding .....	36
2.2.1.1.1. Axenic culture of seedlings .....	36
2.2.1.1.2. Field experiments .....	36
2.2.1.1.3. Crossing experiments and seed storage .....	37
2.2.1.2. Bacterial growth conditions .....	37
2.2.2. Analysis of nucleic acids .....	37
2.2.2.1. Isolation of nucleic acids .....	37
2.2.2.1.1. Isolation of total DNA from <i>Oenothera</i> .....	37
2.2.2.1.2. Isolation of plasmid DNA from <i>Escherichia coli</i> .....	38
2.2.2.1.3. RNA isolation from <i>Oenothera</i> .....	38
2.2.2.2. Agarose gel electrophoresis .....	38
2.2.2.3. cDNA synthesis .....	38
2.2.2.4. PCR approaches .....	38
2.2.2.5. Sequencing approaches .....	39
2.2.2.5.1. Direct PCR product sequencing .....	39
2.2.2.5.2. Sequencing of cloned PCR products .....	39
2.2.2.5.3. Sequencing of inversion breakpoints in the <i>Oenothera</i> plastome .....	39
2.2.2.5.4. Plastome sequencing .....	39
2.2.2.6. SNP mapping by Nuclease S digestion .....	40
2.2.2.7. Design, digestion and analysis of CAPS markers .....	40
2.2.2.8. Gene expression analysis .....	41
2.2.2.8.1. Generation and application of macroarrays .....	41
2.2.2.8.2. Real-time PCR analysis .....	42
2.2.2.8.2.1. Analysis of nuclear gene expression <i>via</i> Real-time PCR .....	42
2.2.2.8.2.2. Analysis of plastid gene expression <i>via</i> Real-time PCR .....	43

---

2.2.3.	Analysis of proteins.....	43
2.2.3.1.	Preparation of thylakoid membranes.....	43
2.2.3.2.	Chlorophyll absorption measurements.....	44
2.2.3.3.	SDS polyacrylamide gel electrophoresis.....	44
2.2.3.4.	Western analysis.....	44
2.2.4.	Determination of chromosome configurations.....	44
2.2.5.	Chlorophyll <i>a</i> fluorescence analysis .....	45
2.2.6.	Bioinformatic analysis.....	46
2.2.6.1.	Calculation of genetic linkage.....	46
2.2.6.2.	Analysis of the <i>Oenothera</i> plastid genomes.....	46
2.2.6.2.1.	Repeat analysis.....	46
2.2.6.2.2.	Analysis of variable amino acid sites.....	46
2.2.6.2.3.	Computational prediction of sigma factor and T7 binding sites .....	47
2.2.6.2.4.	Prediction of Shine-Dalgarno sequences.....	48
2.2.6.2.5.	Calculation of phylogenetic trees .....	48
2.2.6.2.6.	Determination of Ka/Ks-values.....	48
<b>3.</b>	<b>Results .....</b>	<b>49</b>
3.1.	Molecular approaches in <i>Oenothera</i> genetics .....	49
3.1.1.	Marker systems for <i>Oenothera</i> genetics and breeding approaches.....	50
3.1.1.1.	Co-dominant markers discriminating A and B genomes .....	50
3.1.1.2.	Genotyping of Renner complexes.....	54
3.1.1.3.	Markers for basic plastomes and subplastomes .....	56
3.1.2.	New combination of genetic compartments.....	59
3.1.2.1.	Generation of interspecific AB-I and AB-III plastome-genome hybrids.....	59
3.1.2.2.	Generation of the interspecific incompatible BB-II hybrid .....	60
3.1.3.	Correlation of classical and molecular <i>Oenothera</i> maps .....	64
3.1.3.1.	The hybrid <sup>St</sup> albicans <sup>h</sup> tuscaloosa and its genetic behaviour.....	64
3.1.3.2.	Identification of marker alleles between <sup>St</sup> albicans and <sup>h</sup> tuscaloosa.....	65

---

3.1.3.3. Assignment of coupling group 7 to chromosome 9·8 .....	66
3.2. The complete sequences of the five basic <i>Oenothera</i> plastid genomes .....	67
3.2.1. Sequence analysis and annotation of the five plastid chromosomes.....	68
3.2.1.1. Size, gene content and design of the <i>Oenothera</i> plastid chromosomes .....	68
3.2.1.2. <i>CemA</i> is annotated with an alternative start codon ATA.....	73
3.2.2. Analysis of indels and sequence repetition in <i>Oenothera</i> plastomes .....	76
3.2.2.1.1. Indels within the five plastomes.....	76
3.2.2.1.2. Tandem and palindrome repeats .....	76
3.2.2.2. The large 56 kbp inversion.....	77
3.2.3. Differences in coding regions among <i>Oenothera</i> plastomes $I^{\text{Joh}} - V^{\text{doul}}$ .....	79
3.2.3.1. Summary of all gene coding differences.....	80
3.2.3.2. Genes with length polymorphisms.....	87
3.2.3.2.1. Insertions in <i>ndhF</i> and <i>clpP</i> without reading frame change.....	87
3.2.3.2.2. <i>NdhD</i> , <i>rpl22</i> and <i>rps18</i> contain frame changing insertions .....	89
3.2.3.2.3. Sequence divergence of <i>ycf1</i> and <i>ycf2</i> .....	92
3.2.3.2.4. The 5' end of <i>accD</i> is highly polymorphic .....	93
3.2.3.2.5. Alignments of <i>AtpA</i> and <i>PsbB</i> - proteins with mobility shifts .....	94
3.2.4. Differences in intergenic regions between plastomes $I^{\text{Joh}} - V^{\text{doul}}$ .....	95
3.2.4.1. Differences in promoter regions.....	96
3.2.4.2. Differences in ribosomal binding sites.....	96
3.3. Evolutionary analysis of the five basic plastomes .....	96
3.3.1. Calculation of phylogenetic trees.....	97
3.3.2. Estimation of evolutionary distances between the plastomes .....	98
3.3.3. Four basic types of PGI determine the strength of hybridization barriers .....	99
3.3.4. Selection pressure on the <i>Oenothera</i> plastomes.....	101
3.4. Investigation of plastome-genome incompatibility.....	103
3.4.1. Bioinformatic investigation.....	103
3.4.1.1. Search for candidate protein coding loci involved in PGI .....	104

---

3.4.1.2. Investigation of single amino acid exchanges.....	104
3.4.1.3. Investigation of genes showing length polymorphism.....	104
3.4.1.4. Search for candidate loci for PGI in intergenic regions .....	107
3.4.1.5. Differences in RNA editing sites .....	108
3.4.2. Molecular investigation of PGI.....	111
3.4.2.1. Analysis of nuclear gene expression in the three basic <i>Oenothera</i> lineages. 111	
3.4.2.2. Photosynthesis gene clusters are differentially regulated .....	123
3.4.2.3. Delineation of the AB-I incompatibility determinant .....	123
3.4.2.3.1. The incompatible AB-I hybrid shows a photosystem II phenotype.....	124
3.4.2.3.2. The effect on PSII is specific and does not notably affect PSI .....	124
3.4.2.3.3. Western analysis of AB-I thylakoid membrane .....	126
3.4.2.3.4. A deletion in the <i>clpP/psbB</i> spacer explains the AB-I phenotype .....	126
3.4.2.3.5. Bioinformatic and phylogenetic analysis of the <i>clpP/psbB</i> -spacer .....	127
3.4.2.3.6. Expression analysis of <i>psbB</i> and <i>clpP</i> .....	128
3.4.2.3.7. The <i>clpP/psbB</i> spacer of various subplastomes .....	129
<b>4. Discussion.....</b>	<b>130</b>
4.1. Benefit of co-dominant markers to <i>Oenothera</i> breeding .....	130
4.2. A complete alignment between the classical and molecular <i>Oenothera</i> maps .....	132
4.3. Organization and relationship of the <i>Oenothera</i> plastome sequences .....	133
4.3.1. <i>Oenothera</i> particularities of the plastome sequences.....	133
4.3.2. General divergence and repeat analysis .....	133
4.4. Evolutionary considerations based on the plastome sequence.....	134
4.4.1. Divergence time of the five basic plastomes.....	134
4.4.2. The phylogenetic tree of the five <i>Oenothera</i> plastomes .....	136
4.4.3. The large inversion of 56 kbp .....	136
4.4.4. Patterns of subplastome variation in <i>Oenothera</i> populations.....	137
4.5. Selection pressure and determinants of PGI .....	139
4.5.1. Selection pressures acting on the plastome .....	139

---

4.5.1.	Genetic architecture of PGI.....	141
4.5.2.	Molecular determinants of PGI suggest regulatory phenomena .....	142
4.5.3.	The incompatible hybrid AB-I of <i>Oenothera</i> .....	144
4.5.4.	PGI - a useful tool to identify mechanisms and driving forces of speciation ..	145
4.6.	The model <i>Oenothera</i> .....	146
<b>5.</b>	<b>Literature .....</b>	<b>148</b>
<b>6.</b>	<b>Summary .....</b>	<b>168</b>
<b>7.</b>	<b>Appendix .....</b>	<b>170</b>
7.1.	List of publications.....	170
7.2.	<i>Curriculum Vitae</i> .....	172
7.3.	Acknowledgement (Danksagung).....	173

## Table of Figures

<b>Figure 1.</b> Scheme of the thylakoid membrane system in higher plants.....	2
<b>Figure 2.</b> Hybrid variegation in interspecific crosses of <i>Passiflora</i> species .....	3
<b>Figure 3.</b> Summary of the Dobzhansky-Muller model.....	4
<b>Figure 4.</b> Plastome-genome compatibility/incompatibility in the subsection <i>Oenothera</i> .....	7
<b>Figure 5.</b> Distribution of the 11 North American species of subsection <i>Oenothera</i> .....	8
<b>Figure 6.</b> Chromosome configuration of 7 prs. in the strain johansen .....	20
<b>Figure 7.</b> Chromosome configuration $\odot$ 4, 5 prs. in the hybrid <sup>h</sup> johansen· <sup>G</sup> flavens.....	21
<b>Figure 8.</b> Chromosome configuration $\odot$ 14 in the strain suaveolens Grado .....	22
<b>Figure 9.</b> Maintenance of the complete permanent translocation heterozygous strain suaveolens Grado of <i>Oe. biennis</i> .....	23
<b>Figure 10.</b> Exchanging plastids and Renner complexes between <i>Oenothera</i> strains.....	24
<b>Figure 11.</b> Assignment of M40 alleles to the Renner complexes <sup>St</sup> laxans and <sup>St</sup> undans in the strain bauri Standard.....	54
<b>Figure 12.</b> <i>Bam</i> HI restriction pattern of the <i>Oenothera rrn16-trnI</i> <sub>GAU</sub> spacer region.....	56
<b>Figure 13.</b> Lutescent phenotype and somatic plastome segregation in F1 between <i>Oe. elata</i> subsp. <i>hookeri</i> strain johansen (AA-I) and <i>Oe. grandiflora</i> strain tuscaloosa (BB-III).....	60
<b>Figure 14.</b> Crossing scheme to exchange pastome III of <i>Oe. grandiflora</i> strain tuscaloosa with plastome II of <i>Oe. biennis</i> strain suaveolens Grado.....	62
<b>Figure 15.</b> Phenotypic and molecular discrimination of crossing intermediates and end products in the BB-II crossing program.....	63
<b>Figure 16.</b> Chromosome configuration $\odot$ 12, 1 pr. of the hybrid <sup>St</sup> albicans· <sup>h</sup> tuscaloosa.....	64
<b>Figure 17.</b> Assembly and segregation behaviour of the hybrid <sup>St</sup> albicans· <sup>h</sup> tuscaloosa .....	66

---

<b>Figure 18.</b> Gene map of the <i>Oenothera</i> plastid chromosomes .....	71
<b>Figure 19.</b> <i>CemA</i> alignment of the five basic plastomes and 50 reference species .....	74
<b>Figure 20.</b> Overall distribution of sequence divergence and repetition in the <i>Oenothera</i> plastid chromosomes.....	77
<b>Figure 21.</b> Scheme indicating the 56 kbp inversion of the <i>Oenothera</i> plastid chromosomes..	78
<b>Figure 22.</b> Sequence alignment of the 3' <i>ndhF</i> nucleotide sequence and translated gene product in <i>Oenothera</i> plastomes with no frame changing insertions.....	87
<b>Figure 23.</b> Sequence alignment of the 3' <i>clpP</i> nucleotide sequence and translated gene product uncovering a length polymorphism caused by larger deletions.....	88
<b>Figure 24.</b> Sequence alignment of the NdhD C-terminus and the frameshift causing sequence variety.....	89
<b>Figure 25.</b> Sequence alignment of the Rpl22 C-terminal end and the frameshift causing sequence variety .....	90
<b>Figure 26.</b> Sequence alignment of the Rps18 C-terminus and the frameshift causing sequence variety.....	91
<b>Figure 27.</b> Sequence alignment of AccD including sequences of eight reference species and all five <i>Oenothera</i> plastomes.....	93
<b>Figure 28.</b> Sequence alignment of the 3' <i>atpA</i> and polymorphism detected in <i>psbB</i> .....	94
<b>Figure 29.</b> Phylogenetic trees of the five <i>Oenothera</i> plastomes.....	97
<b>Figure 30.</b> Representative autoradiogram of an array filter representing a subset of nuclear genes for chloroplast function .....	112
<b>Figure 31.</b> Comparison of the transcript expression profiling of 187 nuclear genes for chloroplast function in the three naturally occurring plastome-genome combinations .....	113
<b>Figure 32.</b> Validation of macroarray expression analysis data <i>via</i> real-time PCR .....	114
<b>Figure 33.</b> Studies on photosystem II yield and redox kinetics of photosystem I.....	125

---

<b>Figure 34.</b> Immunoblot analysis of various thylakoid membrane proteins in green AB-III and incompatible AB-I tissue.....	126
<b>Figure 35.</b> Schematic overview of the <i>clpP/psbB</i> spacer region in <i>Oenothera</i> , <i>Spinacea</i> , <i>Nicotiana</i> , and <i>Arabidopsis</i> .....	128

## List of Tables

<b>Table 1.</b> Progenies and phenotypes of crosses between the three homozygous <i>Oenothera</i> lineages.....	9
<b>Table 2.</b> Plant taxa exhibiting plastome-genome incompatibility .....	11
<b>Table 3.</b> Antibodies used for Western analysis .....	26
<b>Table 4.</b> Oligonucleotide primers used for PCR amplification and mRNA quantification.....	27
<b>Table 5.</b> Names and accession numbers of referece plastomes .....	28
<b>Table 6.</b> <i>Oenothera</i> strains of the subsections <i>Oenothera</i> and <i>Munzia</i> used in this work.....	29
<b>Table 7.</b> PCR-based polymorphisms found between the Renner complexes of the A genome <sup>h</sup> <sub>johansen</sub> and the B genome <sup>h</sup> <sub>tuscaloosa</sub> .....	51
<b>Table 8.</b> SSLP and selected restriction endonuclease patterns of the M40 microsatellite region in various Renner complexes .....	55
<b>Table 9.</b> <i>Bam</i> HI restriction and SSLP pattern of the <i>rrn16-trnI</i> <sub>GAU</sub> spacer region in <i>Oenothera</i> plastomes and subplastomes used in this study.....	58
<b>Table 10.</b> Segmental arrangements and chromosome configurations of <i>Oe. biennis</i> strain suaveolens Grado, <i>Oe. grandiflora</i> strain tuscaloosa, and of their F1 twin hybrids.....	61
<b>Table 11.</b> Assignment of eight co-dominant markers to the seven coupling groups of <sup>h</sup> <sub>johansen</sub> · <sup>h</sup> <sub>tuscaloosa</sub> and their alleles in <sup>h</sup> <sub>tuscaloosa</sub> · <sup>h</sup> <sub>tuscaloosa</sub> and <sup>St</sup> <sub>albicans</sub> · <sup>h</sup> <sub>tuscaloosa</sub> . 67	67
<b>Table 12.</b> Sizes and base composition of the <i>Oenothera</i> plastome DNAs I <sup>joh</sup> – V <sup>doul</sup> .....	69
<b>Table 13.</b> Pairwise comparison of sequence similarity and identity in % of the five <i>Oenothera</i> plastomes.....	70
<b>Table 14.</b> List of genes found in the <i>Oenothera</i> plastomes .....	72
<b>Table 15.</b> Summary of all differences at the nucleotide level found in coding regions of the five <i>Oenothera</i> plastomes in pairwise comparison .....	81

---

<b>Table 16.</b> Summary of all differences at the protein level found in coding regions of the five <i>Oenothera</i> plastomes in pairwise comparison .....	84
<b>Table 17.</b> Sequence identity of <i>ycf1</i> in comparison with reference plastomes.....	92
<b>Table 18.</b> Sequence identity of <i>ycf2</i> in comparison with reference plastomes.....	92
<b>Table 19.</b> Evolutionary distances in years calculated for the five <i>Oenothera</i> plastomes.....	98
<b>Table 20.</b> Average Ka/Ks values calculated from the five <i>Oenothera</i> plastomes in pairwise comparison .....	102
<b>Table 21.</b> Estimation of possible determinants for plastome-genome incompatibility caused by single amino acid exchanges .....	105
<b>Table 22.</b> Estimation of possible determinants for plastome-genome incompatibility caused by indels .....	107
<b>Table 23.</b> Assessment of putative PEP promoters as candidate loci for plastome-genome incompatibility in <i>Oenothera</i> .....	109
<b>Table 24.</b> Assessment of putative NEP promoters as candidate loci for plastome-genome incompatibility in <i>Oenothera</i> .....	110
<b>Table 25.</b> Comparison of three transcriptomes of the genetic constitutions AA-I, BB-III and CC-V .....	115

## List of Abbreviations

AFLP	=	<b>a</b> mplified <b>l</b> ength <b>p</b> olymorphism
ATP	=	<b>a</b> denosine- <b>t</b> ri- <b>p</b> hosphate
atro	=	<b>a</b> trovirens
biM	=	<b>b</b> iennis <b>M</b> ünchen
BLAST	=	<b>b</b> asic <b>l</b> ocal <b>a</b> lignment <b>s</b> earch <b>t</b> ool
CAPS	=	cleavable <b>a</b> mplified <b>p</b> olymorphic <b>s</b> equences
CCD	=	charge coupled <b>d</b> evice
cDNA	=	complementary <b>D</b> N <b>A</b>
CFS	=	cytoplasmatic <b>f</b> emale <b>s</b> terility
CI	=	cytoplasmatic <b>i</b> ncompatibility
CMA <sub>3</sub>	=	chromomycin <b>A</b> <sub>3</sub>
CMS	=	cytoplasmatic <b>m</b> ale <b>s</b> terility
CoA	=	<b>c</b> o-enzyme <b>A</b>
Col	=	<b>C</b> olmar
DAPI	=	4',6- <b>d</b> iamidino-2- <b>p</b> henylindole
DAS	=	<b>d</b> ense <b>a</b> lignment <b>s</b> urface
DM	=	<b>D</b> obzhansky- <b>M</b> uller
DMI	=	<b>D</b> obzhansky- <b>M</b> uller <b>i</b> ncompatibility
DNA	=	<b>d</b> eoxyribonucleic <b>a</b> cid
doul	=	<b>d</b> outhat <b>1</b>
dV	=	<b>d</b> e <b>V</b> ries
EDTA	=	ethylenediamine- <b>t</b> etraacetic <b>a</b> cid
EMBOSS	=	<b>E</b> uropean <b>m</b> olecular <b>b</b> iology <b>o</b> pen <b>s</b> oftware <b>s</b> uite
EST	=	expressed sequence <b>t</b> ag
F1	=	<b>f</b> ilial generation <b>1</b>
F2	=	<b>f</b> ilial generation <b>2</b>

---

Fü	=	<b>F</b> ünfkirchen
G	=	<b>G</b> rado
GM	=	<b>G</b> rantham <b>m</b> atrix
h	=	<b>h</b> aplo
HAPPY	=	mapping of <b>h</b> aploid DNA samples using <b>p</b> olymerase chain reaction
HEPES	=	N-(2- <b>h</b> ydroxyethyl)- <b>p</b> iperazine-N'-(2-ethanesulfonic acid)
HMM	=	<b>h</b> idden <b>M</b> arkov <b>m</b> odel
HMMER	=	<b>h</b> idden <b>M</b> arkov <b>m</b> odel software
InterPro	=	<b>i</b> ntegrated <b>p</b> roteins
IR <sub>A/B</sub>	=	<b>i</b> nverted <b>r</b> epeat <b>A/B</b>
joh	=	<b>j</b> ohansen
lam	=	<b>l</b> amarckiana
LB	=	<b>L</b> uria- <b>B</b> ertani
LHC	=	chlorophyll-binding <b>l</b> ight- <b>h</b> arvesting complex
LOD	=	<b>l</b> ogarithm of <b>o</b> dds
LSC	=	<b>l</b> arge <b>s</b> ingle <b>c</b> opy
ML	=	<b>m</b> aximum <b>l</b> ikelihood
MOPS	=	3-(N- <b>m</b> orpholino) <b>p</b> ropansulfon <b>a</b> cid
MP	=	<b>m</b> aximum <b>p</b> arsimony
mRNA	=	<b>m</b> essenger <b>R</b> NA
MS	=	<b>M</b> urashige and <b>S</b> koog
mya	=	<b>m</b> illion <b>y</b> ears
N/A	=	<b>n</b> ot <b>a</b> pplicable
NADP	=	<b>n</b> icotinamide <b>a</b> denine <b>d</b> inucleotide <b>p</b> hosphate
NEP	=	<b>n</b> uclear <b>e</b> ncoded <b>p</b> olymerase
NJ	=	<b>n</b> eighbor <b>j</b> oining
NPQ	=	<b>n</b> on <b>p</b> hotochemical <b>q</b> uenching
P700	=	photosystem I primary electron donor chlorophyll <i>a</i>

---

PAM	=	<b>p</b> ulse <b>a</b> mplitude <b>m</b> odulated fluorometer
PAML	=	<b>p</b> hylogenetic <b>a</b> nalysis using <b>m</b> aximum <b>l</b> ikelihood
PCR	=	<b>p</b> olymerase <b>c</b> hain <b>r</b> eaction
PEP	=	<b>p</b> lastid encoded <b>p</b> olymerase
PFAM	=	<b>p</b> rotein <b>f</b> amilies
PGI	=	<b>p</b> lastome- <b>g</b> enome <b>i</b> ncompatibility
PHYLIP	=	<b>p</b> hylogeny <b>i</b> nference <b>p</b> ackage
pr.	=	<b>p</b> air
prs.	=	<b>p</b> airs
PSI	=	<b>p</b> hotosystem <b>I</b>
PSII	=	<b>p</b> hotosystem <b>II</b>
p-value	=	<b>p</b> robability- <b>v</b> alue
PVDF	=	<b>p</b> olyvinylidene <b>d</b> ifluoride
PVP 10,000	=	<b>p</b> olyvinylpyrrolidone (average molecular weight <b>10,000</b> )
R or r	=	<b>r</b> ed midripp
RFLP	=	<b>r</b> estriction <b>f</b> ragment <b>l</b> ength <b>p</b> olymorphism
RNA	=	<b>r</b> ibonucleic <b>a</b> cid
rRNA	=	<b>r</b> ibosomal <b>R</b> NA
RT-PCR	=	<b>r</b> everse <b>t</b> ranscription <b>P</b> CR
s	=	<b>S</b> weden
SDS	=	<b>s</b> odium <b>d</b> odecyl <b>s</b> ulfate
SNP	=	<b>s</b> ingle <b>n</b> ucleotide <b>p</b> olymorphism
SSC	=	<b>s</b> mall <b>s</b> ingle <b>c</b> opy
SSLP	=	<b>s</b> imple <b>s</b> equence <b>l</b> ength <b>p</b> olymorphism
St	=	<b>S</b> tandard
suavG	=	<b>s</b> uaveolens <b>G</b> rado
subsp.	=	<b>s</b> ub <b>s</b> pecies
TBE	=	<b>t</b> ris- <b>b</b> orate- <b>E</b> DTA

T-DNA	=	<b>transferred DNA</b>
TEM	=	<b>transmission electron microscopy</b>
Th	=	<b>Thorn</b>
TILLING	=	<b>targeting induced local lesions in genomes</b>
Tris	=	<b>tris-(hydroxymethyl)-aminomethane</b>
tRNA	=	<b>transfer RNA</b>
tusca	=	<b>tuscaloosa</b>
UTR	=	<b>untranslated region</b>
UV	=	<b>ultraviolet</b>

# 1. Introduction

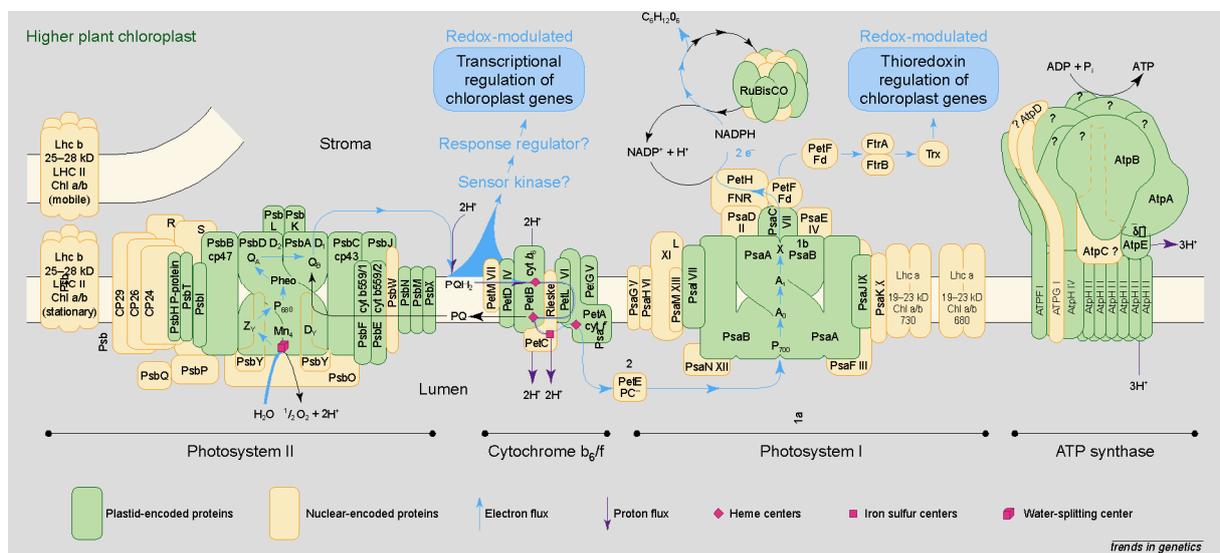
A fundamental characteristic of eukaryotes is their compartmentalized genome. Today, it is undisputed that eukaryotic cells and their partite genomes originated in endosymbiotic conglomerates from formerly free-living cells. Mitochondria and plastids are now known to be extant eubacterial derivatives of an oxygen-consuming  $\alpha$ -proteobacterium and an oxygen-producing cyanobacterium, respectively. Both ancestral symbionts were engulfed independently by an as yet unknown heterotrophic host cell. During evolution, the genetic potentials of the symbiotic partner cells converted into a single integrated compartmentalized genetic system with nucleus/mitochondria in animals and fungi and, in addition, plastids in plants. Both organelles still contain remnants of their ancient genomes. To date, the partite genome is regulated spatiotemporally and quantitatively in its entirety with nuclear regulatory dominance (Herrmann, 1997; Herrmann and Westhoff, 2001). This conversion occurred predominantly at the unicellular level (Martin *et al.*, 1998).

## 1.1. Eukaryotic genomes are integrated and compartmentalized

Co-evolution of the genetic compartments in eukaryotic cells represents one of three fundamental processes, which shaped eukaryotic genome evolution (Herrmann and Westhoff, 2001). Following endosymbiosis, most organelle genes were transferred to the nucleus or lost. Mitochondria and chloroplasts possess only rudimentary genomes, encoding parts of their expression machineries and the respiratory chain or thylakoid membrane (Herrmann, 1997; Martin *et al.*, 1998; Leister, 2003). However, the majority of organellar proteins are now encoded by the nuclear genome. Therefore, much of the nuclear coding capacity, in the order of 25 - 30%, is required for the management the cell organelles (Herrmann, 1997). The flux of genetic information into the nuclear genome led to the generation of novel genes and gene sets, predominantly of regulatory nature. The evolution of multicellularity, which included the development of specific tissues and a wide ontogenetic diversification of plastids, was accompanied by the evolution of novel genes and gene sets and again of complex regulatory networks. It required the invention of retrograde and anterograde signalling pathways and an embedding of organelle biogenesis into the respective ontogenetic programs that vary substantially between the different lineages of algae, mosses or higher plants (Lopez-Juez and Pyke, 2005). Interestingly, a substantial number of plastid proteins is not of cyanobacterial origin and nowadays fulfils functions, which are important for regulation of organelle

biogenesis and function. They are encoded by genes, which either evolved *de novo*, were already present in the host genome, or were transferred horizontally (Leister, 2003).

In consequence, the genetic composition of the eukaryotic cell is a mosaic of co-evolving organellar and nuclear compartments constituting a functional unit. Organelle biogenesis, maintenance, and adaptation are genetically and metabolically tightly embodied in the partite genetic machinery of the cell. The chimerical design and different levels of regulation are obvious, for example, in the thylakoid membrane system that harbours the primary photosynthetic machinery (Figure 1).

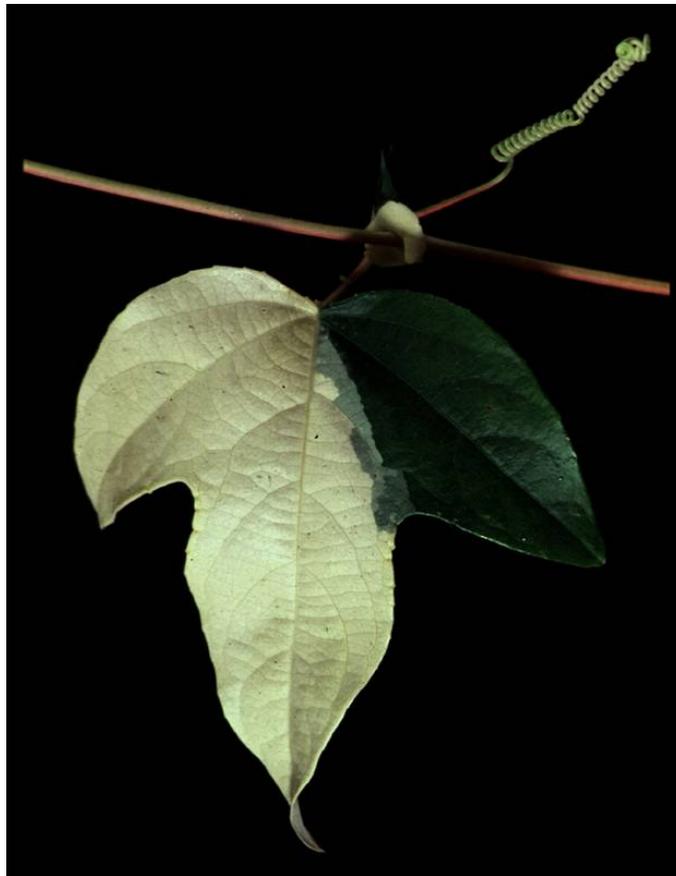


**Figure 1.** Scheme of the thylakoid membrane system in higher plants. Each complex is genetically of chimerical origin and consists of nuclear encoded (yellow) and plastid encoded (green) components (from Race *et al.*, 1999).

## 1.2. Dobzhansky-Muller incompatibilities and asymmetric hybridization barriers

Compartmental integration of co-evolving organelle and nuclear genomes becomes obvious after inter- and intraspecific organelle exchanges, for instance, of plastids and nuclei. Even between closely related species such an exchange can lead to serious developmental disturbances, so-called plastome-genome incompatibilities (PGI) (Stubbe, 1989; Herrmann *et al.*, 2003; Levin, 2003). Compartmental co-evolution influences a multitude of ontogenetic processes, like the photosynthetic machinery or the generative phase. It is often recognized in hybrid bleaching or hybrid variegation (Figure 2) and can cause hybrid sterility, hybrid inviability (hybrid weakness), or hybrid breakdown (Stebbins, 1950; Stubbe, 1989; Yao and Cohen, 2000; Levin, 2003). PGI reflects a disharmonic interaction between the cellular

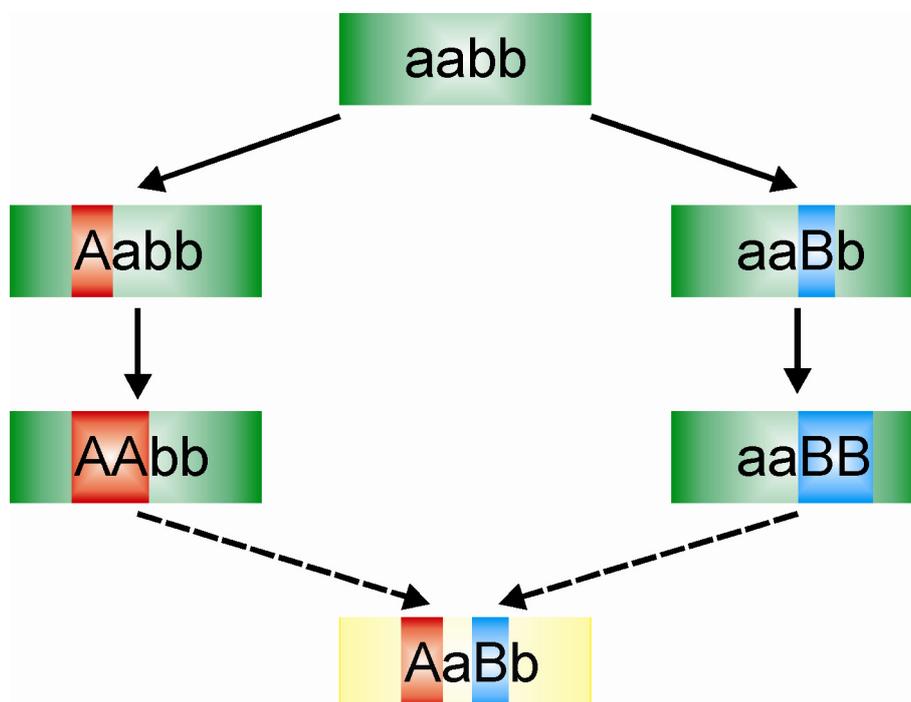
genetic compartments. Different from nuclear and plastid mutations affecting the organelle, PGI is reversible. An impaired foreign plastid will re-green if re-combined with its genuine genome. Apparently, organelles and nucleus share common, co-adapted genetic elements, shaped by species specific co-evolution (Dobzhansky, 1970; Rand *et al.*, 2004). Obviously, an exchange of the cellular compartments disturbs a co-adapted network. Therefore negative “cyto-nuclear” epistasis is frequently observed, typical for so-called Dobzhansky-Muller incompatibilities (DMIs) (Burke and Arnold, 2001; Tiffin *et al.*, 2001; Turelli and Moyle, 2007). Since DMI is the genetic base of speciation, compartmental co-evolution, characteristic to eukaryotes, appears as an important, often neglected, element in speciation processes.



**Figure 2.** Hybrid variegation in interspecific crosses of *Passiflora* species. Variegated tissue is generated by two plastomes transmitted from both sexes. Only one plastome is incompatible with the nuclear genome. Separation of the two tissue types is a consequence of the statistical process of sorting-out; the two plastid types separate during cell divisions (Kirk and Tilney-Bassett, 1978; Birky, 2001). Note that sorting-out of maternal and paternal plastids was not completed in the mating zone between green and white tissue.

### 1.2.1. *The model of Dobzhansky-Muller incompatibility*

The model of DMI was independently developed by Bateson (1909), Dobzhansky (1937) and Muller (1942) and is briefly summarized in Figure 3. The model assumes an ancestral population with the genotype  $aa/bb$ , in which gene products of  $a$  and  $b$  interact functionally and are therefore co-adapted. The population splits into two parts, which are temporally isolated from each other. If a new allele  $A$  arises in one of the subpopulations, individuals with the genetic constitution  $Aa/bb$  and  $aa/bb$  can freely breed with each other. Allele  $A$  may get fixed, and now the first subpopulation has the genotype  $AA/bb$ . In the second subpopulation a similar series of events produces genotype  $aa/BB$ . If the diverged subpopulations mate again, hybrids with the genotype  $Aa/Bb$  will be generated. However, since  $A$  and  $B$  have not co-evolved, their interaction can be maladaptive and allele  $A$  may be negative epistatic over allele  $B$ . Hybridization then results in reduced fitness of the  $Aa/Ba$  offspring implying that a post-zygotic hybridization barrier is established. Well documented examples for DMI are found hybrids of *Drosophila* or *Xiphophorus* and many other taxa [e.g. summarized in Bruke and Arnold (2001), Orr (2005) or Pennisi (2006)].



**Figure 3.** Summary of the Dobzhansky-Muller model, assuming a reproductive barrier based on gene interaction. Alleles  $A$  and  $B$  did not co-evolve and their interaction can lead to hybrid inviability.

Originally, the model of DMI was developed for autosomal gene interactions. Such interactions follow Mendelian inheritance and cause symmetric reproduction isolation in the sense that there is no fitness difference between reciprocal crosses. Hybrid necrosis observed in plants often illustrates such a symmetric DMI (Bomblies and Weigel, 2007). However, non-Mendelian inheritance, especially inheritance with reciprocal differences, is common in nature and can form post-zygotic hybridization barriers under the premises of DMI. In general they are called asymmetric DMIs (Burke and Arnold, 2001; Tiffin *et al.*, 2001; Turelli and Moyle, 2007).

Since organelles inherit preferentially maternal (Hagemann, 2004) and organelle-nuclear, especially mitochondrial-nuclear, incompatibilities were frequently described, the model of DMI was first extended to organelles (Lamprecht, 1944; Michaelis, 1954; Willett and Burton, 2001; Levin, 2003; Sackton *et al.*, 2003; Fishman and Willis, 2006). Probably the most prominent example for mitochondrial-nuclear incompatibility is the commercially important cytoplasmic male-sterility (CMS), observed in many taxa (Chase, 2007). But not only organelles are responsible for asymmetric DMIs. Asymmetric DMIs can also arise from endosymbiotic parasites like *Wolbachia*, sex specific suppression of transposable elements, epigenetic effects acting on one gender, transcripts present in the egg cell, sex chromosomes, and, commonly found in angiosperms, gametophyte-sporophyte interactions as well as triploid endosperm interactions (von Wangenheim, 1962; Preer, 1971; Grun, 1976; Turelli and Orr, 2000; Tiffin *et al.*, 2001; Turelli and Moyle, 2007).

For both cases, symmetric and asymmetric DMIs, theoretical models on population genetics and speciation forces exist, allowing to estimate speed and nature of speciation under both genetic prerequisites (Orr, 1995; Cruzan and Arnold, 1999; Turelli and Orr, 2000; Coyne and Orr, 2004; Turelli and Moyle, 2007).

### 1.2.2. “Speciation genes” have not yet been identified for PGI

In recent years, substantial progress has been made to evaluate the molecular basis of DMIs. Interacting nuclear DM gene pairs were identified in *Drosophila* (Brideau *et al.*, 2006; Presgraves and Stephan, 2007). In *Drosophila* and *Xiphophorus* at least one gene of further DM gene pairs could be cloned (reviewed in Orr, 2005). The molecular function of these genes indicates expectedly a complex picture of speciation. Transcription factors, chromatin binding proteins, a receptor tyrosine kinase and components of the nuclear pore could be

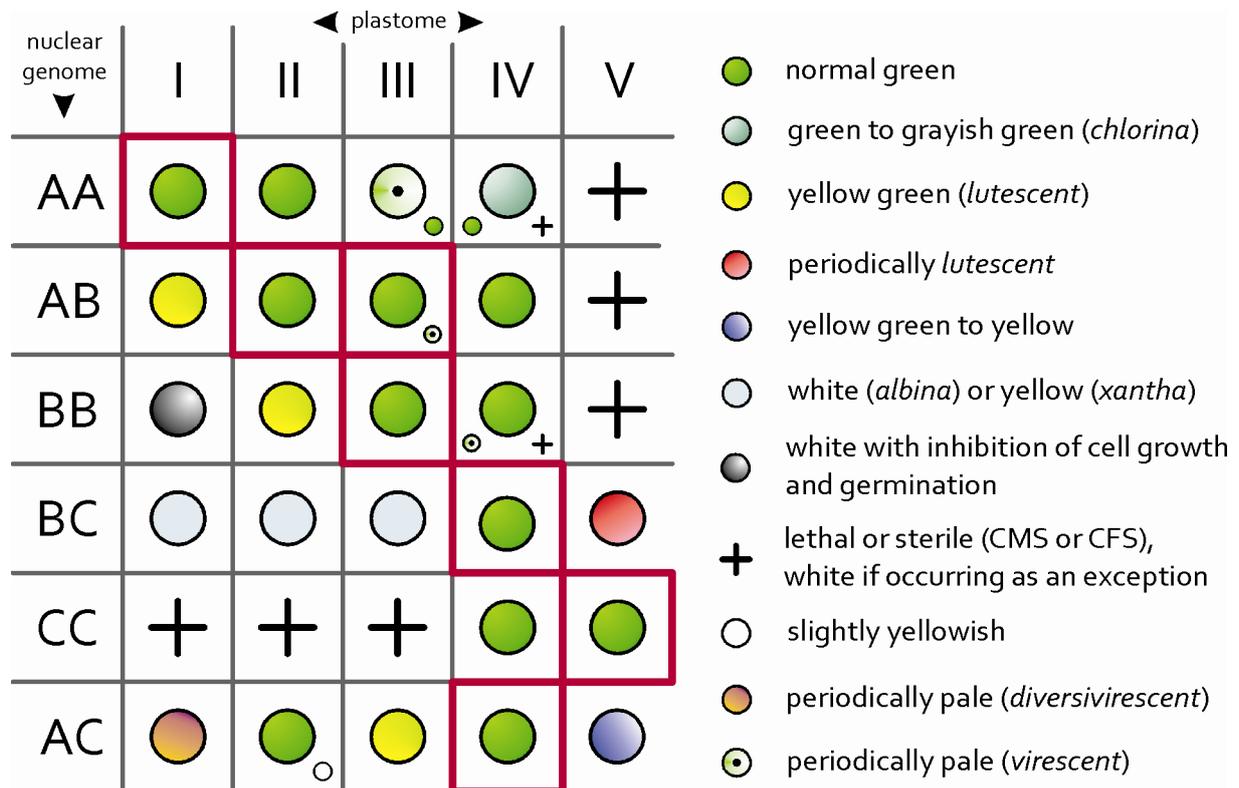
identified as DMI genes (Orr, 2005; Pennisi, 2006), indicating that DMI can be established at various levels. But not only DMIs caused by nuclear gene pairs, also “cyto-nuclear” incompatibilities (CI) are under study. In several plant species the mitochondrial determinants of CMS, short toxic polypeptides from reshuffled organelle genes or gene parts, and nuclear loci resorting fertility are known. These loci often code for pentatricopeptide-repeat proteins degrading the mRNA of a toxic mitochondrial component (Chase, 2007). Also in a few animal models CIs could be traced down to the molecular level. Disruptions of the cytochrome *c* oxidase gene, for instance, were identified to be responsible for hybrid inviability or breakdown in *Drosophila* and *Tigriopus* (Sackton *et al.*, 2003; Harrison and Burton, 2006). Surprisingly, this line of research does not adequately take the plastid as the third genetic compartment into consideration, although for the characteristic organelle of eukaryotic photoautotrophs quite a comprehensive literature is available (summarized in Kirk and Tilney-Bassett, 1978).

### 1.3. Hybridization barriers formed by plastids

That plastid can form important hybridization barriers and be substantially involved in speciation, becomes obvious from work on the genus *Oenothera*. Systematic studies on PGI including aspects of speciation were performed for more than a century. A comprehensive dataset is available notably from subsection *Oenothera* (= *Euoenothera*), the best studied of the five subsections in section *Oenothera* (Stubbe, 1989; Dietrich *et al.*, 1997). It is presented here in some detail as a showcase to illustrate the impact of plastids in speciation. The genus *Oenothera* is the only genus, from which such detailed information is available.

In subsection *Oenothera*, three basic nuclear genomes (A, B and C) occur in homozygous (AA, BB, CC) or stable heterozygous (AB, AC and BC) constitution and are associated with five basic genetically distinguishable plastome types (I - V). Only 12 of the 30 possible combinations of these genomes and plastomes are green, and of these, only seven exist naturally, in altogether 13 species. The remaining 18 combinations display plastome-genome incompatibility to various degrees and can be generated artificially or occur naturally as inviable hybrids (Figure 4). Detailed distribution maps of the subsection can be found in Dietrich *et al.* (1997). A summary for the 11 North American species, together with their subpopulations, is presented in Figure 5. The genetics of *Oenothera* has been reviewed by Cleland (1972) and Harte (1994) and will be explained in more detail in Chapter 1.7. A detailed taxonomic revision is given in Dietrich *et al.* (1997). The evolution of the subsection

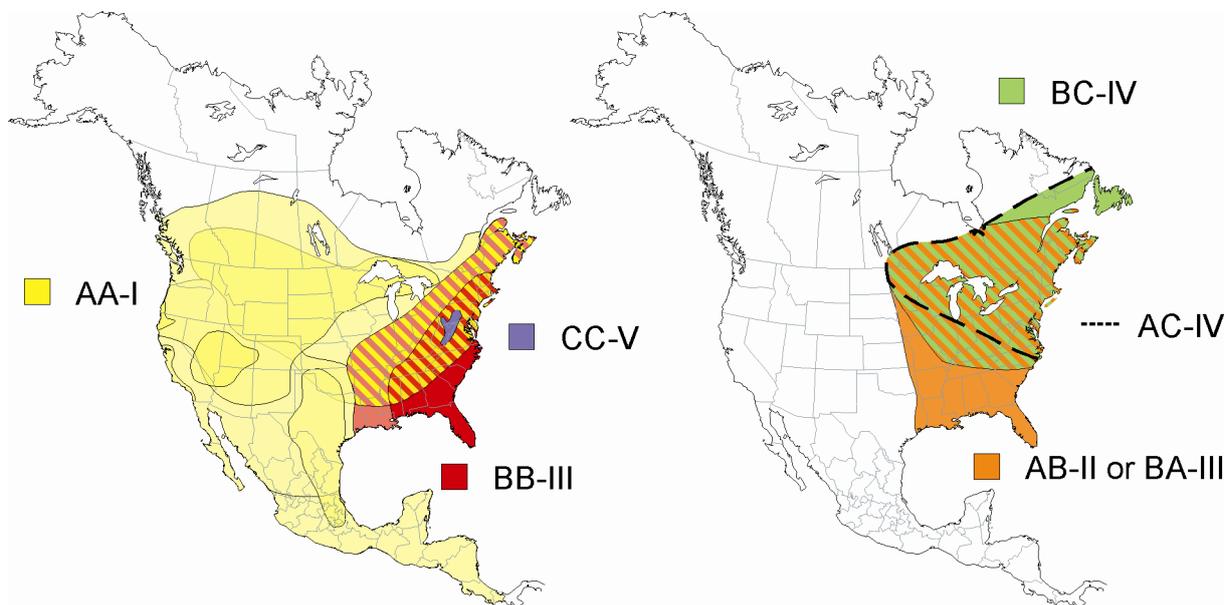
with respect to the plastome, population structure and hybridization behaviour has been outlined in Stubbe (1963a; 1964), Cleland (1972), Kirk and Tilney-Bassett (1978), Stubbe and Raven (1979), Wasmund and Stubbe (1986), Stubbe (1989), Wasmund (1990), Stubbe and Steiner (1999), and Dietrich *et al.* (1997). The following paragraphs briefly summarize relevant aspects.



**Figure 4.** Plastome-genome compatibility/incompatibility in the subsection *Oenothera*, redrawn from Stubbe (1959; 1989) with permission. A, B and C represent the basic nuclear genotypes, I - V the five genetically distinguishable basic plastomes. Genotypes boxed in red represent naturally occurring species. Minor symbols indicate variances noted for some nuclear subgenotypes.

Subsection *Oenothera* appears to be monophyletic and has been proposed to consist of three distinct evolutionary lineages. One lineage with the genetic constitution AA-I consists of five species: *Oenothera elata*, *Oe. jamesii*, *Oe. longissima*, *Oe. villosa* and *Oe. wolffii*, the first mentioned being presumably the most ancient species of the group. A second lineage with the genetic constitution BB-III consists of the ancient species *Oe. grandiflora* and its descendant *Oe. nutans*. The third lineage is represented by the single taxon *Oe. argillicola* that is endemic to the Appalachian mountains and considered as an early relict of the *Oenothera* evolution. Cleland (1972) assumed that the three lineages reached the North American continent,

originating from Middle or South America, in several waves, possibly 70,000 years ago starting at the beginning of the Wisconsin glaciations in Pleistocene.



**Figure 5.** Distribution of the 11 North American species of subsection *Oenothera* of the genus *Oenothera*. The map summarizes data presented in Dietrich *et al.* (1997) and includes information about the basic nuclear genomes (A, B and C) and their associated plastome types (I - V) of that subsection. Yellow and red gradations designate the distribution of distinct AA-I and BB-III genotypes. The left map shows the areas populated by the homozygous species, the right one that of their hybrids. Note that all genotypes overlap geographically.

The proposed ancient forms of the three lineages, *Oe. elata*, *Oe. grandiflora*, and *Oe. argillicola*, are now well separated geographically and post-zygotic hybridization barriers, in form of PGI, have evolved. As shown in Table 1, of all possible crosses between the basic genotypes AA-I, BB-III and CC-V only a single offspring, the combination AB-III, generates viable plants. Since no other pre- or post-zygotic isolation mechanism exists between the ancient species, exclusively PGIs isolate the three basic genotypes from each other - a text book example for Dobzhansky-Muller incompatibilities.

Another well documented example for sexual isolation of *Oenothera* species is the mating of *Oe. nutans*, *Oe. parviflora* and *Oe. argillicola*. The natural habitats of all three species overlap, and different flower morphology resulting in a pre-zygotic hybridization barrier, has already evolved. The different flower morphologies appear to be a consequence of different genetic constitutions and mating behaviours of the species. *Oe. argillicola* is a homozygous, *Oe. nutans* and *Oe. parviflora* are permanent translocation heterozygous species (Chapter

1.7). As a result, *Oe. argillicola* is open pollinating, *Oe. nutans* and *Oe. parviflora* are moderately out-crossing, but predominantly self-pollinating species. Nevertheless, their pre-zygotic hybridization barriers are not very pronounced. Therefore, natural hybrids between *Oe. argillicola* (CC-V) and *Oe. parviflora* (BC-IV) can be observed in the compatible combination CC-IV, and also between *Oe. nutans* (BB-III) and *Oe. parviflora* (BC-IV) in the compatible offspring BB-IV and BC-IV. Therefore, as long as plastome-genome combinations allow, pre-zygotic hybridization barriers like self-pollination are not strong enough to prevent gene flow between the species.

**Table 1.** Progenies and phenotypes of crosses between the three homozygous *Oenothera* lineages<sup>1)</sup>

cross	progeny	phenotype
<b>AA-I x BB-III</b>	AB-I	lutescent
	AB-III	green
<b>AA-I x CC-V</b>	AC-I	diversivirescent
	AC-V	yellow green to yellow
<b>BB-III x CC-V</b>	BC-III	white
	BC-V	periodically lutescent

<sup>1)</sup> Note that only one possible offspring (AB-III) is green and viable, all other ones show hybrid inviability due to their plastome-genome combination.

Gene flow is only prevented if plastome and nucleus are incompatible in interspecific F1 hybrids. Such incompatible F1 offspring is produced in the third possible cross between the three species. If *Oe. nutans* (BB-III) mates with *Oe. argillicola* (CC-V) no viable hybrids between these species are found, since the only possible F1 offsprings are the strongly incompatible combinations BC-III or BC-V (Table 1 and Figure 4). In the mating situation of these three species, the plastome provides the sole and strong hybridization barrier, since barriers like self-pollination are leaky. Only the plastome determines the reproductive isolation between *Oe. nutans* and *Oe. argillicola*. Comparable events occur in the hybridization zone of AA-I species and *Oe. biennis* (AB-II or BA-III). Although viable hybrids between these two groups were described (AA-I, AA-II, AB-II and AB-III), some possible hybridizations of this cross, such as AB-I or AA-III, result in incompatible combinations (Figure 4). Here the plastome built an asymmetric hybridization barrier, since depending on crossing mate and direction viable and inviable offspring is observed.

#### 1.4. The occurrence of PGI in natural populations is underestimated

As outlined before, it was clearly shown for *Oenothera* that plastids establish hybridization barriers and play a major role in speciation, but is this phenomenon of general importance in nature? An idea about the natural occurrence of PGI provides Table 2, presented for the first time in this thesis. To my knowledge the list includes all cases described, implying that PGI has been reported from 14 plant genera<sup>1)</sup>.

This relatively small number of taxa where PGI was detected may indicate that the aspect has not been representatively studied. Influence of plastids in speciation is probably underestimated, predominantly because of methodical and genetic reasons. The only reliable way to detect PGI is the occurrence of hybrid variegation (Figure 2). Only here a bleached phenotype can be correlated without any doubt to the plastome. The phenotype of hybrid variegation recognizes two different plastomes types, which are separated by the statistical phenomenon of sorting-out during cell division (Kirk and Tilney-Bassett, 1978; Birky, 2001). This observation allows to distinguish a bleached plastome based phenotype from bleached phenotypes caused by a hybrid nuclear genome or by another asymmetric effect like the mitochondria. If in F1 just a single bleached plastid type is observed, it is unclear whether the primary reason for the bleaching is related to the plastid genome (see below). Bleaching may reflect a secondary effect caused by another genetic compartment or may be developmentally or environmentally influenced (Kirk and Tilney-Bassett, 1978). However, for a hybrid variegated phenotype to occur, plastids have to be inherited biparentally and only one of the parental plastid types should be incompatible in a foreign nuclear background. As mentioned above, variegation results from sorting-out of maternal and paternal plastid types during cell division after fertilization (Kirk and Tilney-Bassett, 1978; Birky, 2001). The appearance and detection of hybrid variegation requires three basic prerequisites, besides biparental transmission of plastids and an incompatibility of one of the plastome types, also general viability of the interspecific hybrids. The number of taxa in which PGI can be readily detected is further significantly reduced, since only approximately a quarter of the angiosperm species studied transmit plastids biparentally (Corriveau and Corriveau, 1988; Harris and Ingram, 1991; Zhang *et al.*, 2003) and interspecific hybridization is possible with only a limited number of species. That PGI is more frequent, than the genetically detected cases,

---

<sup>1)</sup> Circumstantial evidence provided by formal genetic data from *Epilobium* suggests that incompatibility does not affect interactions with the nuclear genome alone and also exists between plastids and mitochondria (Michaelis, 1954).

**Table 2.** Plant taxa exhibiting plastome-genome incompatibility<sup>1)</sup>

genus	family	involved taxa	phenotype	reference
<i>Acacia</i>	<i>Fabaceae</i>	<i>decurrens</i> with <i>mearnsii</i>	yellow green to periodically pale	Moffett (1965)
<i>Campanula</i>	<i>Campanulaceae</i>	<i>americana</i> interspecific hybrids of different populations; interspecific hybrids of <i>carpatica</i> involving the variety <i>pelviformis</i>	no further described chlorophyll deficiency; white to periodically pale, mostly at the cotyledon stage	Pellew (1917), Galloway and Etersson (2005)
<i>Geranium</i>	<i>Geraniaceae</i>	<i>bohemicum</i> with <i>bohemicum</i> <i>deprehensum</i>	white to yellow green, altered flower morphology	Dahlgren (1923; 1925)
<i>Hypericum</i>	<i>Hypericaceae</i>	<i>acutum</i> , <i>montanum</i> , <i>pulchrum</i> and <i>quadrangulum</i> as well as further species involved in not clearly elaborated cases	depending on cross and direction, different chlorophyll deficiencies, occasionally altered flower colour	Farenholtz (1925), Noack (1931; 1934; 1937), Renner (1934), Herbst (1935)
<i>Medicago</i>	<i>Fabaceae</i>	<i>truncatula</i> Jawniel with Mount Tabor; <i>dzhawakhetica</i> with <i>sativa</i>	white or pale to periodically pale	Lesins (1961), Lilienfeld (1962)
<i>Menziesia</i>	<i>Ericaceae</i>	see <i>Rhododendron</i>		
<i>Oenothera</i>	<i>Onagraceae</i>	inter- and intraspecific hybrids involving most species in the five subsections of the section <i>Oenothera</i>	all types of phenotypes described in this thesis except of altered flower phenotypes and influence on pathogen resistance	Cleland (1972), Stubbe and Raven (1979), Stubbe (1989), Harte (1994), Dietrich <i>et al.</i> (1997) and others

Table 2. (continued)

genus	family	involved taxa	phenotype	reference
<i>Passiflora</i>	<i>Passifloraceae</i>	<i>menispermifolia</i> with <i>oestedii</i>	pale to white	Mráček (2005)
<i>Pelargonium</i>	<i>Geraniaceae</i>	<i>zonale</i> Roseum with <i>zonale</i> hort. Stadt Bern and <i>inquinans</i> ; <i>denticulatum</i> with <i>filicifolium</i> and <i>radula</i> , <i>citriodorum minor</i> and <i>cordatum</i> interspecific hybrids	white or pale to periodically pale; yellow white to yellow green; white	Smith (1915), Metzloff <i>et al.</i> (1982), Pohlheim (1986)
<i>Pisum</i>	<i>Fabaceae</i>	<i>sativum</i> subsp. <i>elatius</i> VIR320 with 41 accessions of the same subspecies	pale or yellow green to periodically pale; probably plastid transmission is altered and presumably CMS or CFS are plastid dependent	Bogdanova and Berdnikov (2001), Bogdanova and Kosterin (2006), Bogdanova (2007)
<i>Rhododendron</i>	<i>Ericaceae</i>	intergeneric hybrids between <i>Menziesia</i> and <i>Rhododendron</i> and various interspecific and intersectional hybrids in <i>Rhododendron</i>	depending on cross and direction, different types of chlorophyll deficiency	Noguchi (1932), Ureshino <i>et al.</i> (1999), Michishita <i>et al.</i> (2002), Ureshino and Miyajima (2002), Sakai <i>et al.</i> (2004), Kita <i>et al.</i> (2005)
<i>Silene</i>	<i>Caryophyllaceae</i>	<i>otites</i> with <i>pseudotites</i>	white to yellow or periodically yellow green	Newton (1931)

Table 2. (continued)

genus	family	involved taxa	phenotype	reference
<i>Trifolium</i>	<i>Fabaceae</i>	<i>repens</i> with <i>ambiguum</i> , <i>hybridum</i> , <i>nigrescens</i> and <i>uniflorum</i> ; <i>alpestre</i> with <i>heldreichianum</i>	white or periodically yellow green; sometimes reduced pollen viability	Pandey (1957), Kazimierski and Kazimierski (1970), Quesenberry and Taylor (1976), Pandey <i>et al.</i> (1987), Przywara <i>et al.</i> (1989), Meredith <i>et al.</i> (1995)
<i>Zantedeschia</i>	<i>Araceae</i>	one intra- and several interspecific hybrids between 4 species of the section <i>Aestivae</i> ; <i>ordorata</i> with <i>aethiopica</i> in the section <i>Zantedeschia</i> and hybrids between the two sections	white or pale to periodically yellow green; declined pathogen resistance; sometimes plastid transmission altered	New and Paris (1967), Yao <i>et al.</i> (1994; 1995), Yao and Cohen (2000), Snijder <i>et al.</i> (2004), Brown <i>et al.</i> (2005)

<sup>1)</sup> Only taxa, in which hybrid variegation occurs are presented, because only in this context hybrid bleaching can be correlated without doubt to the plastome (for details see text). For this reason *Impatiens*, in the newer literature commonly referred to as genus showing PGI, was excluded. Biparental transmission was not described for *Impatiens* (Pandey and Blaydes, 1957; Harris and Ingram, 1991) and in the commonly quoted reference (Arisumi, 1985) plastome chimerical seedlings were not described. Cases of PGI produced by cybrid technology or established by introgression breeding are also excluded, since the influence of other asymmetric effects like mitochondria is not clear and the use of this material in terms of identifying speciation barriers is doubtful. Especially PGIs gained by cybrid technology are artificial and do not reflect primary crossing barriers in nature, although the underlying mechanism may be comparable.

becomes obvious in cybrids, *i.e.* plants produced somatically from non-crossable species in tissue culture, which carry the nucleus of one species and the plastome of another one. Here PGI is observed frequently as summarized in Levin (2003) or Schmitz-Linneweber *et al.* (2005).

If hybrid variegation is not observed, PGI may exist but bleached hybrid phenotypes will not become visible or cannot be unambiguously recognized as PGI, basically due to the rules of plastid genetics. First, the process of sorting-out of two plastid types may already be completed in the embryo shortly after fertilization. Second, there is no straightforward means to assign unequivocally reciprocal phenotypes to the plastid; they may be of mitochondrial origin or be based on another asymmetric effect. Third, if an incompatible reaction of a plastome with a hybrid genome is observed in both crossing directions, formal genetics at first glance does not allow to discern this phenomenon from a bleaching effect caused by incompatible gene pairs in the hybrid nuclear genome alone. A representative example for these difficulties is provided by *Trifolium* genetics. Hybrid bleaching can frequently be detected in this genus, but in only a few instances the involvement of the plastome has unambiguously been demonstrated, although evidence suggests that its proportion is higher (see references in Table 2 and quotations therein). Finally, in various instances PGI may not display bleached phenotypes. Incompatibilities such as CMS (Stubbe and Steiner, 1999), embryo lethality (Stubbe, 1963b) or different photosynthetic performance of two plastid types in green tissue (Iwanaga *et al.*, 1978) caused by plastids are probably quite common in plants, but have been disregarded or overlooked in this context.

Taken together, PGI is probably a general phenomenon and widespread in plants, but generally underestimated because its detection is difficult, if genetic features are not appropriate. It is therefore not surprising that systematic investigations are scarce.

### **1.5. Physiology and cell biology of PGI**

For an in-depth analysis, which identifies relevant characteristics of “speciation genes”, the nature of PGI phenotypes must be known. A molecular determinant causing PGI was reported just from a single case (Schmitz-Linneweber *et al.*, 2005). In this case, RNA editing is the primary reason for incompatibility between the plastome of *Nicotiana* and the nucleus of *Atropa*. Therefore, one of the questions of the work is to evaluate whether this is a major molecular reason for PGI or whether the phenomenon is biologically more complex. The

large variety of PGI phenotypes (Table 2 and Figure 4) suggests that the latter possibility is more likely.

#### 1.5.1. *Albinotic phenotypes of PGI*

The predominant effects of PGIs are various kinds of albinism. Some albinotic phenotypes cause relatively fast necrosis already during germination, but the majority of alterations are embedded less severe into ontogenetic programs, and plants may even fully re-green temporarily or later in development (Table 1 and Figure 4). Schötz and co-workers have provided a detailed morphological and physiological analysis of bleached phenotypes from *Oenothera* (summarized in Harte, 1994). The data were confirmed by other groups (Glick and Sears, 1994; Dauborn and Brüggemann, 1996). Generally, reduced pigment content, lower rates of photosynthesis and an impaired thylakoid membrane structure in incompatible materials were found. Occasionally, plastids in leaf tissue degenerate completely. The observations are comparable to those made on other materials, such as *Passiflora* (Mráček, 2005) or *Zantedeschia* (Yao and Cohen, 2000). In principle, bleached phenotypes can be explained by altered RNA metabolism, protein synthesis, assembly, function and/or degradation of components of the photosynthetic machinery. Mutants involved in these processes display similar phenotypes. If, for example, photosynthetic complexes do not operate correctly, thylakoid membrane structure is often changed (*e.g.* Amann *et al.*, 2004). The phenomenon of delayed bleaching and/or re-greening, observed in nearly all taxa, from which PGI is known (Table 2), resembles a class of mutations common to plants that cause a delay in greening due to deficiencies in photosynthesis, chloroplast biogenesis or regulation of plastid gene expression (*e.g.* Walbot and Coe, 1979; Bondarava *et al.*, 2003; Sjogren *et al.*, 2006).

#### 1.5.2. *PGI phenotypes with affected cell growth and function*

Another spectrum of PGI phenotypes suffers from general cellular dysfunction. Inhibition of cell growth can cause embryo lethality, lack of germination and changes of leaf morphogenesis or of other organs (Stubbe, 1963b). The haploid ontogenetic phase, the gametophyte, can be affected as well causing CFS (cytoplasmic female sterility), CMS or reduced pollen vigor (Stubbe *et al.*, 1978). Both, gametophytic and sporophytic effects have a strong impact on biparental or uniparental transmission of plastids. In *Oenothera*, in extreme cases maternal transmission of a strongly incompatible plastome can be suppressed

completely (Chiu and Sears, 1993). Furthermore, male sterile anthers are frequently associated with round starch grains in the pollen (Stubbe and Steiner, 1999). Such pollen grains generally show changes in respiration, lipid and starch metabolism (Göpel, 1976). In some instances, plastome dependent pollen abortion is correlated statistically significantly with chromosome breaks, asymmetric anaphase chromosome distributions and trinucleated tetrads (Chapman and Mulcahy, 1997). In *Hypericum*, flower colour may depend on plastome-genome interaction (Farenholtz, 1925). An involvement of PGI in pathogen defence has been reported from *Zantedeschia* (Snijder *et al.*, 2004) as well as from bleached mutants (Jain *et al.*, 2004). For a summary and references of the comprehensive literature in *Oenothera* see Harte (1994), the literature quoted above and references therein. For literature addressing other taxa, see citations in Table 2.

Molecular determinants for PGI causing altered cell growth and function have not been identified so far. Genes responsible for gametophyte or sporophyte development are generally encoded by the nuclear genome and no plastid interaction partners are known. However, various plastome knock-out lines have been described, which presumably exert secondary or pleiotropic effects on plant cell growth, leaf or flower morphology. These include disruptions of *accD* (Kode *et al.*, 2005), *clpP* (Shikanai *et al.*, 2001), *rps18* (Rogalski *et al.*, 2006), *ycf1* and *ycf2*, two ORFs of unknown function (Drescher *et al.*, 2000), or a general inhibition of plastid translation (Ahlert *et al.*, 2003).

However, another strategy to identify molecular components involved in PGI is based on an analysis of retrograde and anterograde signalling. For example, the *Oenothera* plastome-genome combination AA-III (Figure 4) bleaches reversibly and circumstantial evidence suggests that this is due to a temporally dysfunctional differentiation process of the chloroplast. It can be partially suppressed by growth hormones (Glick and Sears, 1994) or cured genetically, if plants carry the incompatible plastome III and a compatible plastome II. In so-called mixed cells, in which sorting-out of the two plastome types has not been completed, and in cell layers containing plastome III oriented ad- or abaxially to at least one cell layer with the compatible plastome II, the incompatible plastome III is able to re-green in an AA background (Stubbe, 1958); for pictures see p. 167, plate 2e, in Harte (1994). The missing component of the AA-III incompatibility seems to be a metabolite or a gene product, which can cross cell borders and move from plastid to plastid, perhaps *via* the stromules (Kwok and Hanson, 2004a; b).

The AA-III phenotype described by Glick and Sears (1994), or phenotypes involving sterility resemble those observed in programmed cell death (van Doorn, 2005; Chase, 2007). Developmental disturbances in incompatible combinations are likely due to plastid malfunctions, since the plastid is a crucial compartment of the metabolic and signalling network of the plant cell, such as tetrapyrrole, redox and reactive oxygen signalling, organelle gene expression or sugar signalling (Laloi *et al.*, 2006; Pesaresi *et al.*, 2007). Anyhow, components of these signalling cascades coded on the plastome are unknown, and quite likely not present. Therefore, PGI phenotypes involved in these pathways may reflect a secondary effect of a global plastid dysfunction in plastid gene expression. This may be similar to sterility phenotypes of CMS, affecting flower morphology. They are possibly the result of reduced ATP levels due to mitochondrial dysfunction, causing misexpression of floral regulators (Chase, 2007).

#### **1.6. *Oenothera* as a molecular model to investigate PGI**

The model of choice to study PGI in molecular terms is limited to the species listed in Table 2, since only here the phenomenon was observed. Unfortunately, none of them, except *Medicago*, has been established so far as a model in modern plant research. For most potential models, the genetic basis to investigate PGI is meager and often restricted to single crosses (Table 2). For a potential utility, classical genetics, phylogeography as well as comparative molecular analysis between species must be developed first. Strictly speaking at present different types of PGI have been reported only from *Oenothera*, *Rhododendron*, *Hypericum*, *Trifolium* and *Zantedeschia* (Table 2), and only these genera would allow more detailed access to speciation forces acting on PGI. For obvious reasons, *Rhododendron* is not a suitable model, and many hybrids in *Hypericum*, *Trifolium* and *Zantedeschia* suffer from additional hybridization barriers, such as uneven chromosome numbers, hybrid sterility, or embryo abortion, combined with limited success of interspecific crosses (see references in Table 2). As sole material only *Oenothera* lacks these serious genetic limitations. Only in *Oenothera*, genetically different plastome types, their distribution in various species and impact on evolution were investigated, including an intense phenotypic, genetic and physiological characterization of PGI (Cleland, 1972; Harte, 1994; Dietrich *et al.*, 1997). In comparison to other potential models, therefore, only *Oenothera* remains as a suitable material for molecular investigations.

In evening primroses morphologically different and interbreedable species together with biparental transmission of organelles and a general fertility of interspecific plastome-genome incompatible offspring are the rule. If sterile offspring occurs as an exception, sterility is plastome dependent (CMS or CFS) and can be cured genetically while equipping an incompatible, sterile combination with a compatible plastome type (Stubbe *et al.*, 1978). Other hybridization barriers, except of the plastome, do not play a notable role in speciation, at least in subsection *Oenothera*, on which most of work on PGI was performed (Dietrich *et al.*, 1997). However, work on PGI is not limited to that subsection. It was extended to all five subsections within the section *Oenothera*, including interspecific hybrids as well as hybrids between subsections (Stubbe and Raven, 1979). Furthermore, the knowledge of genetics and taxonomy in the genus, as for the whole family of *Onagraceae*, is unique (Raven, 1988; Levin *et al.*, 2003; Levin *et al.*, 2004).

The genus *Oenothera* has an outstanding genetic tradition, summarized in Lehmann (1922), Cleland (1972) and Harte (1994). It is under study for more than a century and a classical example for the study of hybrid variegation (Kirk and Tilney-Bassett, 1978), biparental transmission of plastids (Hagemann, 2004), and permanent translocation heterozygosity in plants (Holsinger and Ellstrand, 1984; Levin, 2002; Golczyk *et al.*, 2005; also see Chapter 1.7). Fundamental discoveries, namely the rediscovery of the Mendelian rules (de Vries, 1900a; b), the genetic independence of plastids (Renner, 1934), the first description of polyploidy (Lutz, 1907), and the mutation theory (de Vries, 1901 - 1903) were made on *Oenothera*. There is also commercial interest, since *Oenothera* species produce high amounts of  $\gamma$ -linolenic acids in seeds. The fatty acid is used in food supplies and in medical applications (Morse and Clough, 2006; Fieldsend, 2007). Additionally, *Oenothera* tissue culture (Stubbe and Herrmann, 1982; Kuchuk *et al.*, 1998; Mehra-Palta *et al.*, 1998) has been applied for the industrial production of pharmaceutically active secondary metabolites (Taniguchi *et al.*, 2002). Consequently commercial cultivars and attempts for genetic manipulation do exist (de Gyves *et al.*, 2004; Fieldsend, 2007).

The unique position of *Oenothera* in addressing the role of compartmental co-evolution, as well as of the nuclear genome in speciation processes is due to two facts: a favourable combination of genetic features and the existence of a well developed taxonomy, cytogenetic and formal genetics in the genus (Cleland, 1972; Harte, 1994; Dietrich *et al.*, 1997). In the subsection *Oenothera*, the best studied of all subsections, more than 300 genetically analyzed

strains are known (Stubbe and Diers, 1958; Cleland, 1972; Steiner and Stubbe, 1984; 1986; Wasmund and Stubbe, 1986; Wasmund, 1990; Schumacher *et al.*, 1992; Schumacher and Steiner, 1993). As mentioned above, the genetic features of the genus include the possibility of wide interspecific crossing, biparental transmission of organelles, fertility of plastome-genome hybrids and a system of balanced lethal factors, self-incompatibility or selective fertilization in combination with reciprocal translocations of whole chromosome arms, resulting in partial or complete permanent translocation heterozygosity (Cleland, 1972; Stubbe, 1989; Harte, 1994; Levin, 2002 and Chapter 1.7). These features are relatively common throughout all eukaryotic organisms, but their combination is unique in *Oenothera*. Taken together, they allow the exchange of plastids, individual or more chromosome pairs, and even entire haploid chromosome sets between species and the production of plastome-genome incompatible plants (Chapter 1.7). As mentioned above, this possibility led to the identification of five basic, genetically distinguishable plastomes (I, II, III, IV and V) and three haploid nuclear genomes (A, B and C) (Stubbe, 1959; 1960; Stubbe, 1989 and Figure 4 and 5). The rich source of genetic material and detailed knowledge about population structures and natural distribution of genotypes also allow the study of pre-speciation processes, the diversification of populations and the successive evolution from one plastome type into another (Cleland, 1972; Stubbe, 1989; Harte, 1994; Dietrich *et al.*, 1997, and Figure 5).

## 1.7. *Oenothera* genetics

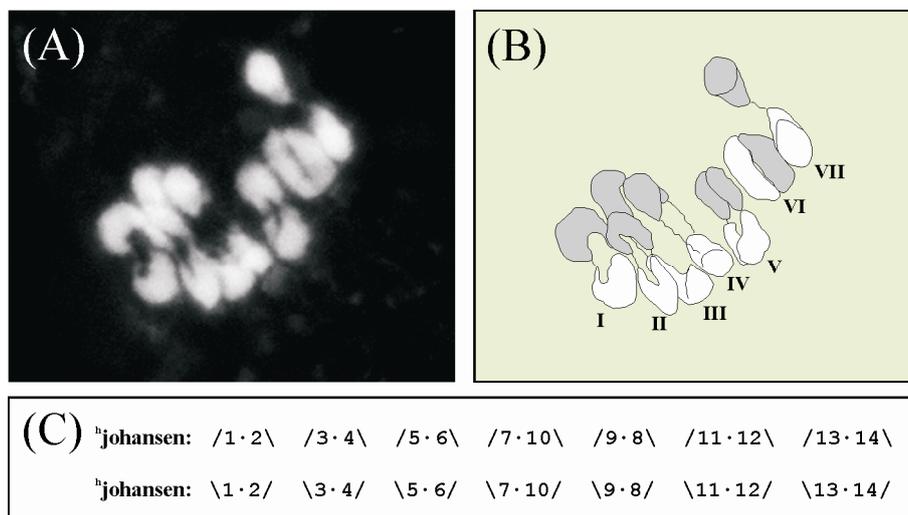
Since the genetics of *Oenothera* is a major reason that renders the genus for suitable investigation on PGI, the following chapter will address the unique combinations of features in the genetics of *Oenothera*. It allows to exchange plastids, but also single or multiple chromosome pairs (Chapter 3.1.3), as well as entire haploid genomes (Renner complexes) between strains and species. Its principles and details were reviewed in various publications (Lehmann, 1922; Stubbe, 1960; Cleland, 1972; Stubbe, 1989; Harte, 1994; Levin, 2002).

### 1.7.1. *Complete reciprocal translocations of whole chromosome arms in Oenothera*

A principal aspect of *Oenothera* genetics are entire haploid genomes, so-called Renner complexes. They are entitled with names such as <sup>h</sup>johansen, <sup>G</sup>albicans or <sup>G</sup>flavens, and inherit as single units. All loci within a Renner complex are in linkage disequilibrium. How these superlinkage groups assemble will be developed as follows.

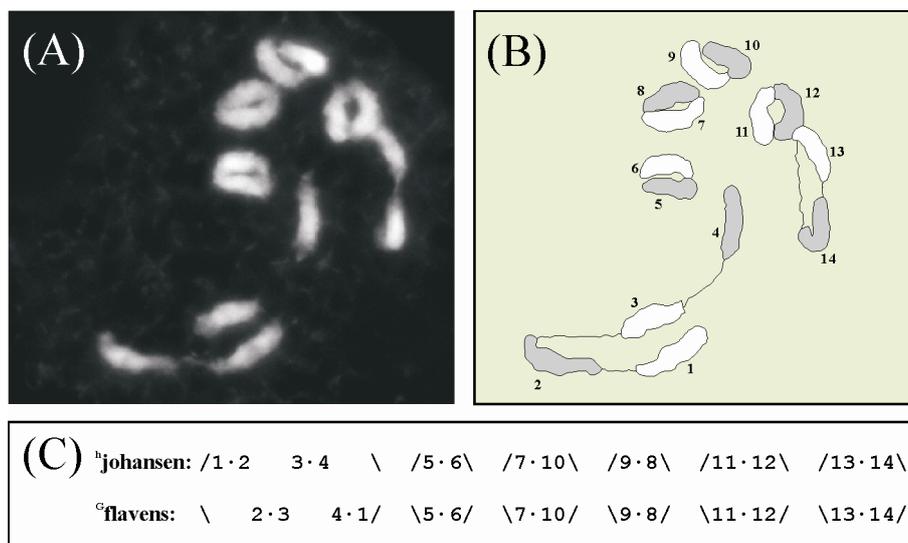
*Oenothera* is a diploid organism with seven chromosomes ( $7n$ ). However, it is useful in *Oenothera* genetics not to number chromosomes but chromosome arms. For example, chromosome I consists of arms 1·2, chromosome II of arms 3·4 and so on, up to chromosome VII with arms 13·14. *Oenothera* chromosomes have the structural particularity that *reciprocal* translocations of *whole* chromosome arms at the *centromer* of a chromosome can occur. Basically reciprocal translocations are rare within the same Renner complex and their occurrence is phenotypically neutral. The reciprocal exchange of chromosome arms leads to an altered, so called segmental arrangement or chromosome formula of a Renner complex. For example, the chromosomes 1·2 and 3·4 display three different segmental arrangements, including 1·4 3·2 and 1·3 2·4. In nature all possible 91 combination of chromosome arms were found. However, chromosome arm combinations differ in their occurrence frequency (Cleland, 1972).

A strain, harbouring two Renner complexes with identical chromosomal arrangements, is a homozygous strain, forming 7 bivalents or pairs in meiosis. An example is *Oe. elata* subsp. *hookeri* strain johansen with the Renner complex combination  ${}^h\text{johansen}\cdot{}^h\text{johansen}$ . In principle, diakinesis and meiotic segregation in the strain johansen looks identical to that of any other diploid organism (Figure 6).



**Figure 6.** Chromosome configuration of 7 prs. in the strain johansen ( ${}^h\text{johansen}\cdot{}^h\text{johansen}$ ): determination in diakinesis *via* DAPI staining (panel A) and graphical interpretation (panel B). The configuration can be predicted by the chromosome formulas of the Renner complexes involved (panel C). DAPI staining and graphical interpretation was done in co-operation with Hieronim Golczyk.

However, if reciprocal translocations of chromosome arms occurred in one of the Renner complexes, meiotic pairing behaviour is altered. An example is the hybrid  $^h\text{johansen} \cdot ^G\text{flavens}$ . The exchange of arms 2 with 4 of chromosomes 1·2 and 3·4 in  $^G\text{flavens}$  led to the altered segmental arrangement  $1 \cdot 4^{\text{flavG}}$  and  $3 \cdot 2^{\text{flavG}}$  relative to  $^h\text{johansen}$  still carrying chromosomes  $1 \cdot 2^{\text{joh}}$  and  $3 \cdot 4^{\text{joh}}$ . Although the remaining chromosomes 5·6 7·10 9·8 11·12 and 13·14 of  $^h\text{johansen}$  and  $^G\text{flavens}$  pair as bivalents, chromosomes  $1 \cdot 2^{\text{joh}} \cdot 2 \cdot 3^{\text{flavG}} \cdot 3 \cdot 4^{\text{joh}} \cdot 4 \cdot 1^{\text{flavG}}$  still conduct and arrange in a ring of 4 chromosomes ( $\odot 4$ ) in diakinesis. The hybrid then has the chromosome configuration  $\odot 4, 5$  prs. (Figure 7).

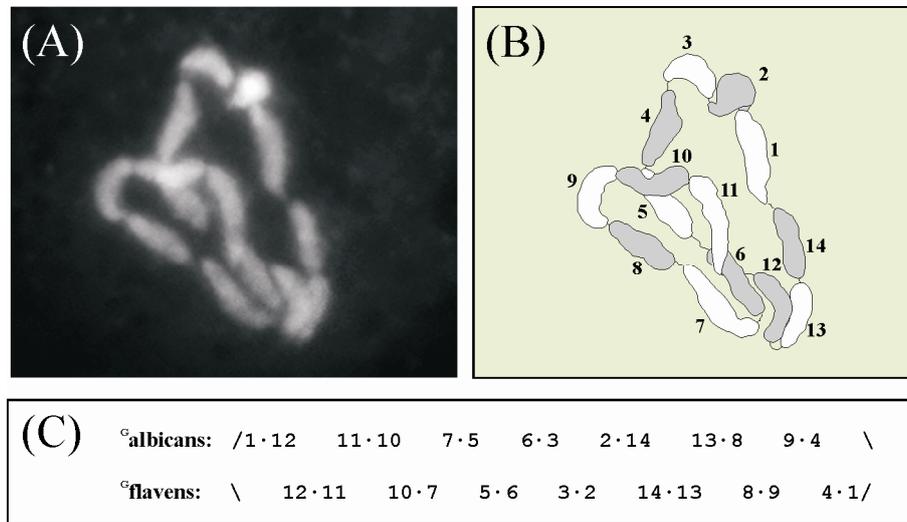


**Figure 7.** Chromosome configuration  $\odot 4, 5$  prs. in the hybrid  $^h\text{johansen} \cdot ^G\text{flavens}$ : determination in diakinesis *via* DAPI staining (panel A) and graphical interpretation (panel B). The configuration can be predicted by the chromosome formulas of the Renner complexes involved (panel C). DAPI staining and graphical interpretation was done in co-operation with Hieronim Golczyk. Note that chromosomes but not chromosome arms are numbered in panel B.

In strains, where reciprocal translocations affect more or even all chromosomes, larger rings arrange in meiosis (*e.g.*  $\odot 6, \odot 4, 2$  prs.). If at least one free bivalent is formed, the situation is designated partial permanent translocation heterozygosis. In case the two Renner complexes assemble in a ring of fourteen ( $\odot 14$ ), incorporating all chromosomes, this is called terminal or complete permanent translocation heterozygosis. An example is *Oe. biennis* strain *suaveolens* Grado with the Renner complexes  $^G\text{albicans}$  and  $^G\text{flavens}$  (Figure 8).

The chromosome configuration, number and size of rings and pairs in diakinesis of a strain or hybrid can be predicted, if the arrangement of chromosome arms, the chromosomal formula, is known for the complexes involved (see Figures 6 - 8). Theoretically, 15 chromosome

configurations are possible. Since chromosomal formulas were determined for about 300 experimental strains, each meiotic configuration can be generated (Cleland, 1972).



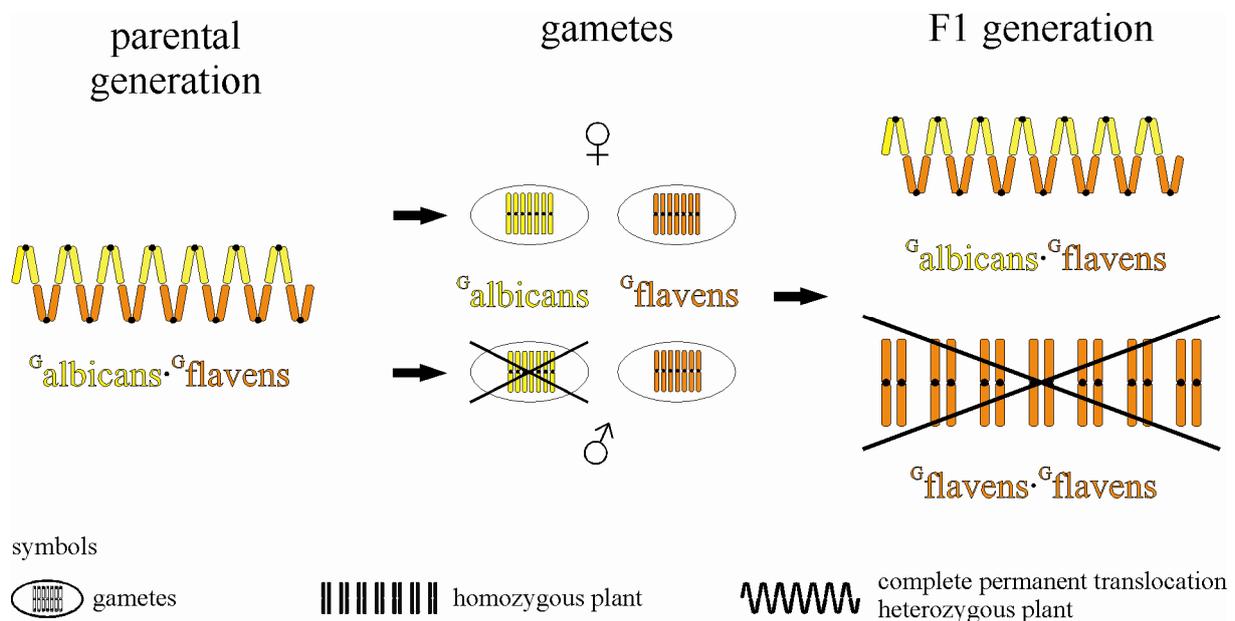
**Figure 8.** Chromosome configuration  $\odot 14$  in the strain *suaveolens* Grado ( $G^{\text{albicans}} \cdot G^{\text{flavens}}$ ): determination in diakinesis *via* DAPI staining (panel A) and graphical interpretation (panel B). The configuration can be predicted by the chromosome formulas of the Renner complexes involved (panel C). DAPI staining and graphical interpretation was done in co-operation with Hieronim Golczyk. Note that chromosomes but not chromosome arms are numbered in panel B.

### 1.7.2. Maintenance of complete permanent translocation heterozygosis

Rings of chromosomes in diakinesis have consequences in inheritance, since they change linkage equilibrium. Each ring or free pair constitutes a single linkage group. For example, a hybrid with the chromosome configuration  $\odot 6$ ,  $\odot 4$ , 2 prs. has 4 linkage groups. Two linkage groups are represented by the free chromosomes and two by the ring of 6 and ring of 4. In contrast, a hybrid with the chromosome configuration 7 prs. has seven linkage groups, one for each chromosome. If chromosomes assemble in a ring of 14 chromosomes, only one superlinkage group is found, involving both haploid genomes entirely.

The occurrence of just one superlinkage group leads to the genetics of complete permanent translocation heterozygosis. Figure 9 illustrates the maintenance of the complete permanent translocation heterozygous *Oe. biennis* strain *suaveolens* Grado. As mentioned above, it contains the Renner complexes  $G^{\text{albicans}}$  and  $G^{\text{flavens}}$  and has the chromosome configuration  $\odot 14$ . The ring of fourteen chromosomes in meiosis *should widely prevent free segregation of chromosomes and practically homologous recombination*. It ensures that  $G^{\text{albicans}}$  and

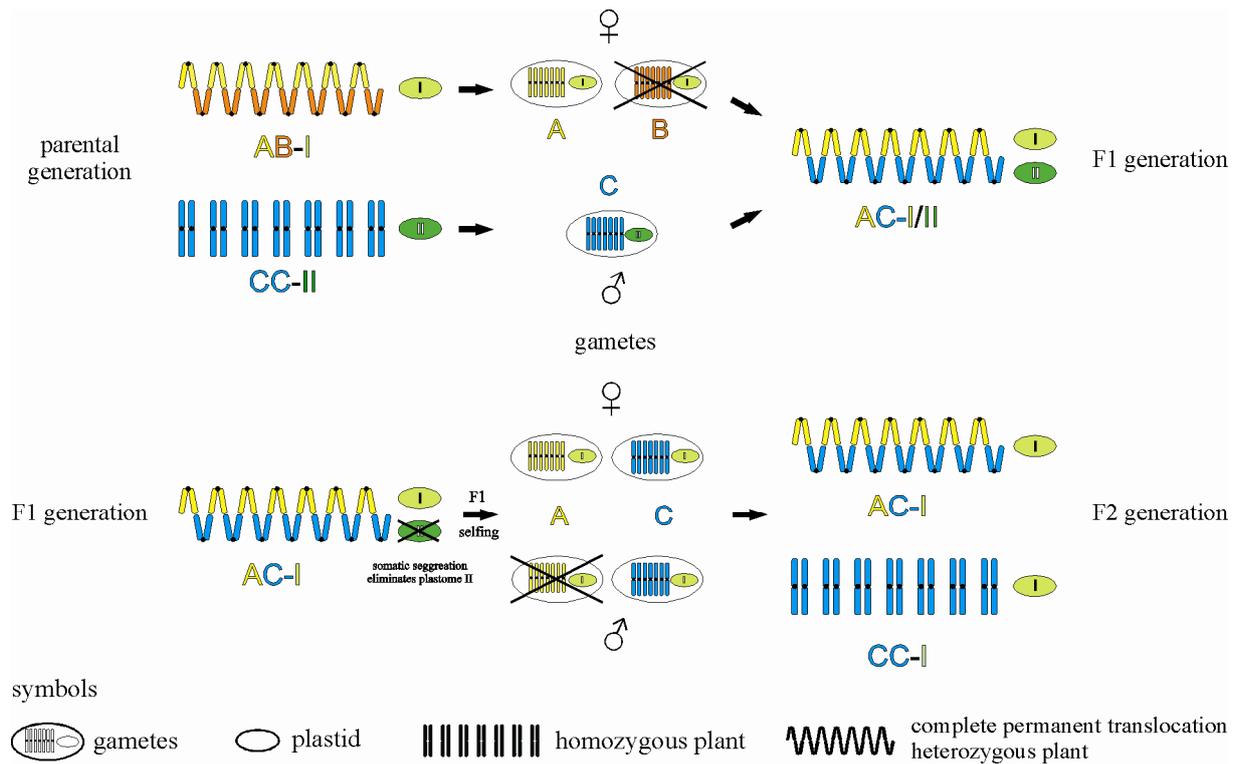
$G^{\text{flavens}}$  are not mixed and inherit as units of a single linkage group. Male and female gametes exclusively contain chromosomes of the haploid set  $G^{\text{albicans}}$  or  $G^{\text{flavens}}$ . Because of gametophytic lethal factors the complex  $G^{\text{albicans}}$  is inherited strictly maternally by the egg cell (♀), the  $G^{\text{albicans}}$  pollen is abortive and eliminated.  $G^{\text{albicans}}$  is therefore called the  $\alpha$ - or egg cell complex. The  $\beta$ -complex  $G^{\text{flavens}}$  can be inherited biparentally *via* egg and pollen (♂♀) and is designated the pollen complex. Because of sporophytic lethal factors  $G^{\text{flavens}} \cdot G^{\text{flavens}}$  homozygotes are not found in the offspring of *Oe. biennis* strain suaveolens Grado. Consequently, the F1 generation is identical to the paternal generation without segregation of any traits. Similar to apomixis a clone is produced. Species following this pattern are designated true breeding species.



**Figure 9.** Maintenance of the permanent translocation heterozygote *Oe. biennis* strain suaveolens Grado. Free segregation of chromosomes and homologous recombination is suppressed in the complete permanent translocation heterozygote between the haploid Renner complexes  $G^{\text{albicans}}$  (yellow) and  $G^{\text{flavens}}$  (orange). Gametophytic lethal factors repress germination of  $G^{\text{albicans}}$  pollen, sporophytic lethal factors the raise of homozygous F1  $G^{\text{flavens}} \cdot G^{\text{flavens}}$  offspring. Referred to the parental generation, identical offspring is produced in F1. For detailed explanation see text.

### 1.7.3. Exchanging plastomes between species

Knowledge about chromosomal formulas, together with biparental transmission of plastids offers the possibility to exchange plastids and/or single chromosome pairs or sets between species in *Oenothera*. Figure 10 outlines the exchange of plastids between species *via* sexual



**Figure 10.** Exchange of plastids and genome rearrangement between *Oenothera* strains. Repression of homologous recombination in a complete permanent translocation heterozygote and somatic segregation leading to sorting out of the two plastid types in F1 (I and II), allow the exchange of plastomes as well as haploid genomes (A, B, and C) in F2. For detailed explanation see text.

crosses in just two generations. It also illustrates how Renner complexes can be recombined. In the chosen example, a complete permanent translocation heterozygous hybrid AB associated with plastome I as seed parent is crossed with the homozygous strain CC-II. The homozygous Renner complex C lacks gametophytic or sporophytic lethal factors and self incompatibility alleles. It forms 7 bivalents in meiosis. In a cross between AB and CC only the hybrid AC occurs in F1, since complex B cannot be transmitted by the pollen. The chromosomal formulas of the complexes A and C are chosen that way that the hybrid AC assembles a meiotic ring of 14 chromosomes. This meiotic configuration allows now an exchange of plastids. Since plastid inheritance is biparental, the F1 generation carries two plastome types, plastome I and II. Somatic segregation of the two plastome types leads to sorting-out of plastids resulting in flowers on a plant, carrying exclusively plastome I (or II). Selfing such flowers ensures that only plastome I is passed on to the next generation. However, the nuclear genomes segregate as splitting F2 progenies (CC-I and AC-I), due to the rules of *Oenothera* genetics for a ring of 14 chromosomes: Complex A is exclusively inherited by the egg cell, but complex C is free of gametophytic or sporophytic lethals. The

ring of 14 chromosomes again should inhibit free segregation of single chromosomes and recombination in the hybrid AC between haploid genomes. Therefore, the coding potentials of A and C are not mixed and an unchanged, homozygous CC genotype, now associated with plastome I, can occur in the F2. With this breeding material, plastids as well as whole haploid genomes have been exchanged. Starting in the parental generation with the hybrid AB-I and the homozygous combination CC-II, crossing end products in the F2 generation were AC-I and CC-I.

### 1.8. Aims of this work

The outlined genetics renders the genus *Oenothera* to an unrivalled model to study processes of speciation, especially the impact of PGI. Only in *Oenothera* systematic investigations of PGI, as illustrated by the compatibility chart in Figure 4, are possible and available for a large and representative number of strains in almost all *Oenothera* subpopulations (Stubbe, 1959; 1960; Cleland, 1962; Stubbe, 1963a). This comprehensive source of classical and population genetics is an ideal base for development of molecular approaches. One aim of this thesis was to identify first determinants responsible for PGI in *Oenothera* and to investigate their role in speciation processes. This required a relatively broad molecular characterization of both, plastome and nuclear genome, including the evaluation of the complete sequences of the five basic plastomes types of the subsection. Furthermore, the sequence of the five plastid genomes and genotyping of *Oenothera* subpopulations allowed achieving a second aim, a close evolutionary look at phylogeny coherences of PGI in *Oenothera*. Finally, a third aim of this thesis was to contribute to a more global understanding of PGI. Therefore, the thesis summarizes for the first time comprehensively the phenomena of PGI and attempts to identify responsible selection forces and general molecular reasons. It appears that PGI deserves wider attention. It could contribute substantially to an understanding of speciation processes, but its impact on speciation is underestimated and often neglected.

## 2. Material and Methods

### 2.1. Material

#### 2.1.1. Chemicals

Chemicals used in this work were all of *p.a.* quality and, if not otherwise mentioned, purchased from the companies Biozym Scientific GmbH (Oldendorf, Germany), Fluka (Basel, Switzerland), Merck (Darmstadt, Germany), Roth (Karlsruhe, Germany), Serva (Heidelberg, Germany) or Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany).

#### 2.1.2. Solutions, buffers and media

Solutions, buffers or media, for which the composition is not given in the Method section were prepared as described in Sambrook *et al.* (1989).

#### 2.1.3. Antibodies

**Table 3.** Antibodies used for Western analysis

protein/protein complex	subunit	source
photosystem I	PsaA/B	Nechushtai (Hebrew University, Jerusalem, Israel)
	Lhca1	Agrisera (Vännäs, Sweden)
photosystem II	PsbB	Berzborn (Ruhr-University Bochum, Bochum, Germany)
	D1	Agrisera (Vännäs, Sweden)
ATP synthase	AtpA	Berzborn (Ruhr-University Bochum, Bochum, Germany)
cytochrome <i>b<sub>6</sub>f</i> complex	PetA	Herrmann (Ludwig-Maximilians-University, Munich, Germany)
CemA	CemA	Soll (Ludwig-Maximilians-University, Munich, Germany)

2.1.4. *Oligonucleotides***Table 4.** Oligonucleotide primers used for PCR amplification and mRNA quantification<sup>1)</sup>

<b>oligonucleotide</b>	<b>sequence (5-to-3')</b>
16S SEQ (+)	CTTGACACACCGCCCGTCACACT
cemA5'rev	GAGTGCGTAGTATTCCACCA
cemAfor2	CTGATTTATGTATCGGATTCC
clpP_IIP31rev	AATGATACATCAGCTCGAGTCC
Ijeto	GATTCACATCATCTCTTACAACC
M101for	GTGCTTCTAAGTGTGGAGCAACA
M101rev	CATCAGACCTTTCTTCTCCATACAGA
M13forward	CCCAGTCACGACGTTGTAAAACG
M13reverse	AGCGGATAACAATTCACACAGG
M34for	GAGACTCCTGTCTGACGCCAG
M34for	CCATGGCGTGTTACGGACAC
M60for	GCAACCAACAATGGCGGTCTG
M60rev	CTCTTACCGCAGCCGGAATCC
ndhDint/2	GCTTGTATCCGTAGAACATT
P11for	TCAAATGGTTCTCCCAAAGACC
P38for	TACCAAGTCTGAAACCGAGTGG
P7rev	GCATAAGGGTTTCTTCGTAGGA
psaC5'	TGATAGATCCAATGTCGCATT
psaI_IVP11rev	GGAGAAATCCATTCTTGTCGTC
psbB_IVP15rev	TAGTCCATAAGGATCGGACACC
psbBfor	ATTTTCTGATGAACGCACAG
psbL7in1	TTGATCCATTGAGGTATCTG
rbcLfor	TGTGGCATATGCCTGCTCTG
RPL20R5M	ATTTGGCTTCGGTTGCTGTC
rps16_IIP3for	GAACAGAAGAAAGGGTGTGCGAG
rps18for	ATTGCTATAAAACAAGCTCG
rps18revM	CTGTTGGTCTTAGAACCAGA
SP6rev	ATTTAGGTGACACTATAGAATAC
T7for	TAATACGACTCACTATAGGGCGA
trnI PCR (+)	CCAGGCACAACGACGCAATTATCA
trnQ_IVP37for	CACTGGAATTGACGAATAACC
VP9for	CATCTTCTTCGTCTTCGTCTCC
VP10rev	AATACACCCAATGCCAGATAGC

<sup>1)</sup> Oligonucleotides used for co-dominant nuclear markers are listed in Table 7, p. 51.

## 2.1.5. Reference species for bioinformatic analysis

**Table 5.** Names and accession numbers of reference plastomes

reference species	accession no.	reference species	accession no.
<i>Acorus calamus</i>	NC_007407	<i>Lobularia maritima</i>	NC_009274
<i>Aethionema grandiflorum</i>	NC_009266	<i>Lotus japonicus</i>	NC_002694
<i>Agrostis stolonifera</i>	NC_008591	<i>Marchantia polymorpha</i>	NC_001660
<i>Amborella trichopoda</i>	NC_005086	<i>Morus indica</i>	NC_008359
<i>Arabidopsis thaliana</i>	NC_000932	<i>Nandina domestica</i>	NC_008336
<i>Arabis hirsuta</i>	NC_009268	<i>Nasturtium officinale</i>	NC_009275
<i>Atropa belladonna</i>	NC_004561	<i>Nicotiana tabacum</i>	NC_001879
<i>Barbarea verna</i>	NC_009269	<i>Nuphar advena</i>	NC_008788
<i>Calycanthus floridus</i>	NC_004993	<i>Nymphaea alba</i>	NC_006050
<i>Capsella bursa-pastoris</i>	NC_009270	<i>Olimarabidopsis pumila</i>	NC_009267
<i>Citrus sinensis</i>	NC_008334	<i>Oryza indica</i>	NC_008155
<i>Coffea arabica</i>	NC_008535	<i>Oryza japonica</i>	NC_001320
<i>Crucihimalaya wallichii</i>	NC_009271	<i>Panax ginseng</i>	NC_006290
<i>Cucumis sativus</i>	NC_007144	<i>Pelargonium x hortorum</i>	NC_008454
<i>Daucus carota</i>	NC_008325	<i>Phalaenopsis aphrodite</i>	NC_007499
<i>Draba nemorosa</i>	NC_009272	<i>Phaseolus vulgaris</i>	NC_009259
<i>Drimys granadensis</i>	NC_008456	<i>Piper cenocladum</i>	NC_008457
<i>Epifagus virginiana</i>	NC_001568	<i>Platanus occidentalis</i>	NC_008335
<i>Eucalyptus globulus</i>	NC_008115	<i>Populus alba</i>	NC_008235
<i>Glycine max</i>	NC_007942	<i>Ranunculus macranthus</i>	NC_008796
<i>Grossypium hirsutum</i>	NC_007944	<i>Saccharum officinarum</i>	NC_006084
<i>Helianthus annuus</i>	NC_007977	<i>Solanum tuberosum</i>	NC_008096
<i>Hordeum vulgare</i>	NC_008590	<i>Sorghum bicolor</i>	NC_008602
<i>Jasminum nudiflorum</i>	NC_008407	<i>Spinacia oleracea</i>	NC_002202
<i>Lactuca sativa</i>	NC_007578	<i>Triticum aestivum</i>	NC_002762
<i>Lepidium virginicum</i>	NC_009273	<i>Vitis vinifera</i>	NC_007957
<i>Liriodendron tulipifera</i>	NC_008326	<i>Zea mays</i>	NC_001666

2.1.6. *Oenothera* strains**Table 6.** *Oenothera* strains of the subsections *Oenothera* and *Munzia* used in this work. For a detailed description of the taxonomy see Dietrich (1977) and Dietrich *et al.* (1997)

strain	genetic constitution	reference for plastome	Renner complex ( $\alpha$ - $\beta$ )	diakinesis	reference for complex ( $\alpha$ and $\beta$ )	strain first described
<i>Oenothera elata</i> subsp. <i>elata</i>						
chapultepec	AA-I	Stubbe (1963a)	<sup>h</sup> chapultepec	7 prs.	Steiner (1951)	Steiner (1951)
cholula	AA-I	Stubbe (1963a)	<sup>h</sup> cholula	7 prs.	Steiner (1955)	Steiner (1955)
puebla	AA-I	Stubbe (1963a)	<sup>h</sup> puebla	7 prs.	Steiner (1955)	Steiner (1955)
toluca	AA-I	Stubbe (1963a)	<sup>h</sup> toluca	7 prs.	Steiner (1951)	Steiner (1951)
<i>Oenothera elata</i> subsp. <i>hookeri</i>						
franciscana de Vries <sup>1)</sup>	AA-I	Stubbe (1959), Stinson (1960)	<sup>h</sup> franciscana de Vries	7 prs.	Cleland (1935)	Davis (1916), Renner (1941)
franciscana E. & S. <sup>1)</sup>	AA-I	Stubbe (1959)	<sup>h</sup> franciscana E. & S.	7 prs.	Cleland (1935)	Davis (1916), Renner (1941)
hookeri de Vries	AA-I	Stubbe (1959)	<sup>h</sup> hookeri de Vries	7 prs.	Cleland and Blakeslee (1931)	de Vries (1913)
johansen	AA-I	Stinson (1960)	<sup>h</sup> johansen	7 prs.	Cleland (1935)	Cleland (1935)
<i>Oenothera villosa</i> subsp. <i>villosa</i>						
bauri Standard <sup>2)</sup>	AA-I	Stubbe, 1959	<sup>St</sup> laxans- <sup>St</sup> undans	⊙14	Baerecke (1944)	Renner (1937)

Table 6. (continued)

strain	genetic constitution	reference for plastome	Renner complex ( $\alpha\beta$ )	diakinesis	reference for complex ( $\alpha$ and $\beta$ )	strain first described
<i>Oenothera biennis</i>						
biennis de Vries <sup>3,4)</sup>	AB-II	de Vries (1913), Renner (1924)	<sup>dV</sup> albicans· <sup>dV</sup> rubens	⊙8, ⊙6	Catcheside (1940)	de Vries (1901 - 1903) de Vries (1913)
biennis München <sup>3,4)</sup>	AB-II	Stubbe (1959)	<sup>biM</sup> albicans· <sup>biM</sup> rubens	⊙8, ⊙6	Catcheside (1940)	Renner (1917)
chicaginensis Colmar	BA-III	Stubbe (1963a)	<sup>Col</sup> excellens· <sup>Col</sup> punctulans	⊙12, 1 pr.	Renner (1956), Stubbe (1963a)	Renner (1950)
lawrenceville 3 <sup>5)</sup>	AB-II	Cleland (1962)	$\alpha$ -lawrenceville 3· $\beta$ -lawrenceville 3	⊙14	Cleland (1958)	Cleland (1958)
nuda Standard	AB-II	Stubbe (1963a)	<sup>St</sup> calvans· <sup>St</sup> glabrans	⊙14	Jean <i>et al.</i> , (1966)	Renner (1956)
purpurata	AA-II	Stubbe (1959)	<sup>h</sup> purpurata	7 prs.	Catcheside (1940)	Klebahn (1914)
suaveolens Grado	AB-II	Stubbe (1914; 1959)	<sup>G</sup> albicans· <sup>G</sup> flavens	⊙14	Stubbe (1953) Stubbe and Diers (1958)	Stubbe (1953)
suaveolens Fünfkirchen	AB-II	Stubbe (1959)	<sup>Fü</sup> albicans· <sup>Fü</sup> flavens	⊙10, 2 prs.	Stubbe (1953)	Stubbe (1953)
suaveolens Standard	AB-II	Stubbe (1959)	<sup>St</sup> albicans· <sup>St</sup> flavens	⊙12, 1 pr.	Cleland and Blakeslee (1931), Catcheside (1940)	Blaringhem (1914)
<i>Oenothera biennis</i> x <i>Oenothera glazioviana</i>						
conferta Standard	AB-II	Stubbe, 1963	<sup>St</sup> convelans· <sup>St</sup> aemulans	⊙12, 1 pr.	Renner and Hirmer (1956)	Renner (1950)

Table 6. (continued)

strain	genetic constitution	reference for plastome	Renner complex ( $\alpha\cdot\beta$ )	diakinesis	reference for complex ( $\alpha$ and $\beta$ )	strain first described
<i>Oenothera glazioviana</i> <sup>6)</sup>						
coronifera Standard <sup>7)</sup>	BA-II	Stubbe (1963a)	<sup>St</sup> quaerans- <sup>St</sup> paravelans	⊙12, 1 pr.	Rossmann (1963)	Renner, (1937)
blandina <sup>8, 9)</sup>	A/B-III	de Vries (1913), Renner (1924)	<sup>h</sup> blandina	7 prs.	Catcheside (1940)	de Vries (1917)
decipiens <sup>8, 9)</sup>	A/B-III	de Vries (1913), Renner (1924)	<sup>h</sup> decipiens	7 prs	Catcheside (1939)	de Vries (1919a)
deserens <sup>8, 9)</sup>	A/B-III	de Vries (1913), Renner, (1924)	<sup>h</sup> deserens	7 prs.	Renner (1943a)	de Vries (1919b)
rr-lamarckiana Sweden <sup>10)</sup>	AB-III	Stubbe (1959)	r- <sup>S</sup> velans-r- <sup>S</sup> gaudens	⊙12, 1 pr.	Cleland and Blakeslee (1931), Emerson and Sturtevant (1931), Catcheside (1940)	Heribert-Nilsson (1912)
<i>Oenothera grandiflora</i>						
bellamy A	BB-III	Schumacher <i>et al.</i> (1992)	<sup>h</sup> bellamy A	7 prs.	Steiner and Stubbe (1984)	Steiner and Stubbe (1984)
B <sup>A</sup> B castleberry A-4	BB <sup>A</sup> -III	Schumacher <i>et al.</i> (1992)	<sup>h</sup> B <sup>A</sup> -castleberry A-4·B-castelberry A-4	7 prs.	Schumacher <i>et al.</i> (1992)	Steiner and Stubbe (1986)
B <sup>A</sup> B chastang 7	BB <sup>A</sup> -III	Schumacher <i>et al.</i> (1992)	<sup>h</sup> B <sup>A</sup> -chastang 7·B-chastang 7	7 prs.	Schumacher <i>et al.</i> (1992)	Steiner and Stubbe (1986)
castleberry B-8	BB-III	Schumacher <i>et al.</i> (1992)	<sup>h</sup> castleberry B-8	7 prs.	Schumacher and Steiner (1993)	Steiner and Stubbe (1986)
stockton 1	BB-III	Schumacher <i>et al.</i> (1992)	<sup>h</sup> stockton 1	7 prs.	Cleland (1972)	Cleland (1972)
tuscaloosa	BB-III	Schumacher <i>et al.</i> (1992)	<sup>h</sup> tuscaloosa	7 prs.	Steiner and Stubbe, (1984)	Steiner and Stubbe (1984)

Table 6. (continued)

strain	genetic constitution	reference for plastome	Renner complex ( $\alpha\beta$ )	diakinesis	reference for complex ( $\alpha$ and $\beta$ )	strain first described
<i>Oenothera nutans</i>						
elkins 2	BB-III	Wasmund (1990)	$\alpha$ -elkins 2· $\beta$ -elkins 2	⊙14	Wasmund (1990)	Wasmund (1990)
horsesheads 2	BB-III	Cleland (1972), Wasmund (1990)	$\alpha$ -horseshead 2· $\beta$ -horseshead 2	⊙14	Cleland (1972), Wasmund (1990)	Cleland (1972)
marienville 3	BB-III	Cleland (1962), Wasmund (1990)	$\alpha$ -marienville 3· $\beta$ -marienville 3	⊙14	Cleland (1958), Wasmund (1990)	Cleland (1958)
mitchell	BB-III	Cleland (1962), Wasmund (1990)	$\alpha$ -mitchell· $\beta$ -mitchell	⊙14	Peer (1950)	Peer (1950)
<i>Oenothera oakesiana</i>						
ammophila Standard	AC-IV	Stubbe (1959)	<sup>St</sup> rigens· <sup>St</sup> percurvans	⊙12, 1 pr.	Baerecke (1944)	Hoeppener and Renner (1929)
<i>Oenothera parviflora</i>						
atrovirens Standard <sup>11)</sup>	BC-IV	Stubbe (1959)	<sup>St</sup> pingens· <sup>St</sup> flectens	⊙14	Catcheside (1940), Baerecke (1944)	de Vries (1901 - 1903) de Vries (1913)
silesiaca Standard	BC-IV	Stubbe (1959)	<sup>St</sup> subpingens· <sup>St</sup> subcurvans	⊙14	Baerecke (1944)	Renner (1942a; 1943b)
st. stephen	BC-IV <sup>4)</sup>	Cleland (1962), Cleland (1972)	$\alpha$ -st. stephen· $\beta$ -st. stephen	N/A <sup>12)</sup>	Cleland (1972)	Cleland (1972)

Table 6. (continued)

strain	genetic constitution	reference for plastome	Renner complex ( $\alpha\cdot\beta$ )	diakinesis	reference for complex ( $\alpha$ and $\beta$ )	strain first described
<i>Oenothera argillicola</i>						
douthat 1	CC-V	Stinson, 1960	<sup>h</sup> douthat 1	7 prs.	Stinson, 1953	Stinson, 1953
williamsville	CC-V	Stubbe, unpublished	<sup>h</sup> williamsville	7 prs.	Stubbe, unpublished	Stubbe, unpublished
wilson creek 1	CC-V	Greiner, this work	N/A <sup>12)</sup>	N/A <sup>12)</sup>	Greiner, this work	Greiner, this work
<i>Oenothera villaricae</i>						
berteriana Schwemmler <sup>13)</sup>	N/A <sup>12)</sup>	N/A <sup>12)</sup>	B·1	⊙14	Haustein (1952)	Schwemmler <i>et al.</i> (1938)

<sup>1)</sup> The strains franciscana de Vries and franciscana E. & S. are derivatives of Davis's franciscana B (Davis, 1916). The connection of these strains is summarized in Renner (1941).

<sup>2)</sup> A very similar strain is known under the name "*Oe. hungarica*".

<sup>3)</sup> Based on crossing data published by de Vries (1901 - 1903; 1913) and own data, Renner (1924) determined the plastid type of biennis de Vries to be identical to that of suaveolens. All suaveolens strains investigated carry plastome type II (Stubbe, 1959).

<sup>4)</sup> The two strains of *Oe. biennis* differ in the inheritance of their rubens complex, <sup>dV</sup>rubens ( $\sigma$ ) and <sup>biM</sup>rubens ( $\sigma$ ) (de Vries, 1913; Renner, 1917), but not in their meiotic configurations (Cleland and Oehlkers, 1930). With respect to chromosomal arrangement of their  $\alpha$ - and  $\beta$ -complexes (albicans and rubens) no

distinction can be made between both strains (Renner, 1938; Catcheside, 1940; Renner, 1950). Both strains, as well as all strains of *Oe. biennis* in the strict genetical sense of *albicans-rubens*, are phenotypically highly constant throughout Europe (Renner, 1937; 1950).

<sup>5)</sup> For the plastome type of this strain no primary data have been published in the literature. For its identification see Cleland (1962), Cleland (1972), Raven and Stubbe (1979), Dietrich *et al.*, (1997) and this work.

<sup>6)</sup> *Synonyms: Oe. erythrosepala and ex genetica: Oe. lamarckiana*

<sup>7)</sup> This strain of *Oe. glazioviana* is not identical with *Oe. lamarckiana* in a strict genetical sense. In general, it is a BA-II species (Rossmann, 1963; Stubbe, 1963a), which basically does not exist in the taxonomic systems of Cleland (1962) or Dietrich *et al.* (1997). Cleland (1972) assumes that *coronifera* is a hybrid between *Oe. biennis* and *Oe. glazioviana* and not a strain of *Oe. glazioviana* as suggested in Dietrich *et al.* (1997).

<sup>8)</sup> *Decipiens* and *deserens* are a so called “full”-mutants (de Vries, 1917; 1919a; b). For review and possible mechanism of “full”-mutant appearance see Cleland (1942) and (Renner, 1943a). Both originated from *Oe. glazioviana* strain *lamarckiana* de Vries (1919a; b), and are a mixture of the complexes *velans* (A) and *gaudens* (B). With *blandina* the situation is more complex. It did not arise directly from *lamarckiana* de Vries but from a hybrid of two mutants of *lamarckiana* de Vries, *mut. laxa* x *mut. semilata* (de Vries, 1917). Because the original material was lost, it is nearly impossible to reconstruct the mechanism of chromosome rearrangement, how <sup>h</sup>*blandina* could have been developed. Dietrich *et al.* (1997) state <sup>h</sup>*blandina* as an A-complex, which is true in various respects, especially as de Vries calls the mutant species “*Oe. blandina*” also “*Oe. lamarckiana mut. veluntina*”. He saw *blandina* as some kind of stable “*velans* species”. The reason why <sup>h</sup>*blandina* is described in this work as an A/B-complex is its association with plastome III (see footnote 9). *Blandina* does not show a typical *virescent* phenotype, expected from an AA-III genotype Stubbe (1959). Therefore, it is assumed that there are still elements of *gaudens* (B) present in the genome of <sup>h</sup>*blandina*.

<sup>9)</sup> As stated in footnote 8, *blandina*, *deserens* and *decipiens* are mutants that originated in *Oe. glazioviana* strain *lamarckiana* de Vries (de Vries, 1917; 1919a; b). As mutants they carry the same plastome as the strain they are originated from. The crossing data presented by de Vries (1901 - 1903; 1913; summarized in

Renner, 1924) do not indicate a different plastome type in *lamarckiana* de Vries than plastome III. It is therefore likely, that the genetic behaviour of *blandina*, *deserens* and *decipiens* plastomes is identical to that of *lamarckiana* Sweden, although the former strains were not investigated directly with respect to their plastome types.

<sup>10)</sup> In the *Oenothera* literature basically just two strains of true *Oe. glazioviana* were described, the strain, which de Vries found at Hilversum, and the Swedish strain of Heribert-Nilsson. The Swedish strain differs from that of de Vries, among other characters, in the absence of the phenotypical marker R (red midrip) (Heribert-Nilsson, 1912; de Vries, 1916; Renner, 1917). In the literature it is not always clearly stated, which strain was used by a particular author. The chromosomal arrangement of *velans* and *gaudens* was described in at least four publications (Cleland and Blakeslee, 1931; Emerson and Sturtevant, 1931; Catcheside, 1940; Cleland and Hammond, 1950). Emerson and Sturtevant (1931) and Cleland and Hammond (1950) used the Swedish strain to determine its chromosomal arrangement, for the other authors this is unclear. Additionally, in the work of Emerson und Sturtevant (1931) and in earlier publications of Cleland, diakinesis configurations of *lamarckiana* are cited from not clearly stated strains. However, it is questionable whether a distinction between *lamarckiana* de Vries and *lamarckiana* Sweden is necessary in this context. De Vries argues that the Swedish strain represents an offspring of the strain he used (de Vries, 1916). Stubbe has tested a large variety of strains of *Oe. glazioviana* (in the strict genetic sense of *Oe. lamarckiana*) collected from all over the world. He found continuously the Renner complexes *velans*·*gaudens* associated with plastome III, as in the Swedish strain (Stubbe, unpublished). Cleland (1972, p. 226) presents similar data, and also Kappus (1957) found in the strain *lamarckiana* Altenheim *velans* and *gaudens*, but interestingly associated with plastome II. Therefore, with respect to chromosomal formulas or complex constitution no difference appears to exist between *lamarckiana* de Vries and Sweden, and between strains of “*Oe. lamarckiana*” in general.

<sup>11)</sup> alternative designation of the strain: “*Oe. cruciata* de Vries”

<sup>12)</sup> N/A = not applicable

<sup>13)</sup> alternative designation of the strain: *berteriana* Erlangen

### 2.1.7. *Arabidopsis* strains

Leaf material of *Arabidopsis* ecotype Columbia-0 was obtained from PD Dr. Jörg Meurer (Ludwig-Maximilians-University, Munich, Germany).

### 2.1.8. *Bacterial* strains

*E. coli* DH5 $\alpha$  (Bethesda Research Laboratories, 1986)

## 2.2. Methods

### 2.2.1. *Growth of biological material*

#### 2.2.1.1. *Oenothera* growth conditions and breeding

##### 2.2.1.1.1. *Axenic culture of seedlings*

All *Oenothera* species and hybrids, including the CC genotype, were cultivated in growth chambers at 24°C with 8/16 h dark light cycles at 100  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , supplied by Osram L85W/25 Universal White fluorescent lamps (Osram GmbH, Munich, Germany). Plants were grown from seeds sterilized in 70% EtOH and 30% H<sub>2</sub>O<sub>2</sub> essentially on 1/2 MS-media (Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) by a protocol adapted from de Gyves *et al.* (2001).

##### 2.2.1.1.2. *Field experiments*

For field growth *Oenothera* seeds were sown in January, choked at the 2-leaf to 4-leaf state in trays of 54 seedlings and kept in a greenhouse to the end of April. After one or two weeks of hardening, rosettes were planted to the field. Flowering of evening primroses started during summer. A few strains such as pupurata, douthat 1 or biennis München, behave biennial depending on seasonal variation.

Strains of *Oe. grandiflora* were short-day treated to induce early flowering in the season. This treatment, performed in a greenhouse after plants had broken rosette stage at a height of 40 cm, ensured that the material flowered and set seeds simultaneously with the other strains under study. Without this treatment *Oe. grandiflora* starts blooming later in autumn, and does not complete its generation cycle before winter in the climate of Munich. Alternatively, seeds

of *Oe. grandiflora* were sown in autumn in a greenhouse to reach flower shooting before winter. During winter the material grew under a natural short-day period. Some of these plants were not transferred to the field later in spring and kept in the greenhouse until flowering, since seasonal cold interference may inhibit flower induction.

#### 2.2.1.1.3. *Crossing experiments and seed storage*

For crossing experiments with *Oenothera* plants immature anthers were removed the day before maturation and flowering. The emasculated flower was guarded and pollinated with the desired pollen of the male crossing mate during the following day and guarded again. Pollination was repeated the next day. Mature seeds were harvested approximately six weeks after pollination and dried at room temperature. No dormancy is needed for sowing the successive generation. Dry seeds can be stored at -20°C for decades. Without freezing, seeds lose the ability for germination after about four years.

For checking the plastome type of a flower bud, leaf material of three successive bracts on a stem was pooled for DNA isolation. *Oenothera* bracts are arranged in an angle of 120° around the stem. Material pooling ensures to recognize stems carrying two different plastome types.

#### 2.2.1.2. *Bacterial growth conditions*

*E. coli* cells were cultivated on LB-media with appropriate antibiotics according to Miller (1987).

### 2.2.2. *Analysis of nucleic acids*

#### 2.2.2.1. *Isolation of nucleic acids*

##### 2.2.2.1.1. *Isolation of total DNA from Oenothera*

Total DNA was isolated from green leaves using the DNeasy™ Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol with minor modifications: Approximately 50 - 100 mg of plant material (fresh weight) was homogenized. To 400 µl AP1-buffer supplied by the manufacturer 4.0 µl of 10% PVP 10,000 and 0.4 µl 1 M sodium ascorbate solution was added. To increase DNA yield, DNA was eluted twice from the DNeasy Mini Spin Column with 50 µl included AE buffer. DNA yields were in the range of 100 ng/µl and decreased with the age of material.

#### 2.2.2.1.2. *Isolation of plasmid DNA from Escherichia coli*

“Miniprep isolation” from *E. coli* DH5 $\alpha$  cells (Invitrogen, Carlsbad, CA) was done with “QIAprep<sup>TM</sup> Spin Miniprep Kit” (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

#### 2.2.2.1.3. *RNA isolation from Oenothera*

Total RNA was isolated from *Oenothera* using the “RNeasy<sup>TM</sup> Plant Mini Kit” (Qiagen, Hilden Germany) with minor adjustments required for *Oenothera* tissue. To cope with interfering mucilage and polyphenols, only up to 20 mg of mature leaf material, for younger leaves up to 100 mg, were extracted. Two additional washing steps with RW1 buffer and five with buffer RPE, both supplied by the manufacturer, were included. To remove contaminating DNA, RNA was digested with DNAase I (Roche, Mannheim, Germany).

#### 2.2.2.2. *Agarose gel electrophoresis*

DNA was analyzed on 2 to 3% SeaKem LE agarose (Biozym Scientific GmbH, Oldendorf, Germany) in 1 x TBE buffer, RNA on 0,8% agarose gels in 1 x MOPS buffer according to standard methods. As size standard a 100 bp ladder (New England Biolabs, Ipswich, MA) was used.

#### 2.2.2.3. *cDNA synthesis*

To remove contaminating DNA, the RNA to be used was digested with DNase I (Roche, Mannheim, Germany). Synthesis of cDNA was performed with Super-Script III RNase H<sup>-</sup> (Invitrogen, Karlsruhe, Germany) according to the manufacturer's instructions using random primers (Roche, Mannheim, Germany).

#### 2.2.2.4. *PCR approaches*

PCR was performed using standard protocols with a prolonged initial denaturation time of five minutes, if total *Oenothera* DNA was used as template. Depending on the final application of the PCR products Taq DNA Polymerase (Qiagen, Hilden, Germany) was used for “standard” PCRs, for a higher accuracy or extreme long templates the PCR reactions were performed with Phusion High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland) or

BIO-X-ACT™ Long DNA Polymerase (Bioline GmbH, Luckenwalde, Germany), respectively, according to the manufacturer's protocol.

#### 2.2.2.5. Sequencing approaches

##### 2.2.2.5.1. Direct PCR product sequencing

Individual PCR products were directly sequenced after cleaning with QIAquick™ PCR Purification Kit (Qiagen, Hilden, Germany). Nucleotide sequences were determined with an ABI 377 robot (Applied Biosystems, Foster City, CA). Cycling reactions were performed using DYE™ ET terminator Cycle Sequencing Kit (GE Healthcare UK Ltd, Little Chalfont Buckinghamshire, UK) and cycling conditions recommended by the supplier.

##### 2.2.2.5.2. Sequencing of cloned PCR products

A genomic locus, amplified by PCR in a permanent translocation heterozygous strain, customarily yields two bands originating in the  $\alpha$ - and  $\beta$ -complexes, respectively. Both bands were cloned individually with the pGEM®-T Easy Vector System (Promega, Madison, WI) according to the supplier's manual. For sequencing analysis the primer pair T7for and SP6rev was used. Several, independent clones of the same PCR product were sequenced.

##### 2.2.2.5.3. Sequencing of inversion breakpoints in the *Oenothera* plastome

Total DNA of *Oe. villaricae* strain berteriana Schwemmler (*syn.*: berteriana Erlangen) was isolated as described above. The primer pair *rbcL*for and *psaI\_IVP11*rev deduced from a highly conserved region in the *Oenothera* plastomes was used to sequence the interval equivalent of the inversion breakpoint between *rbcL* and *accD* in the berteriana Schwemmler plastome. PCR was performed with BIO-X-ACT™ Long DNA Polymerase (Bioline GmbH, Luckenwalde, Germany) and the product was sequenced by primer walking. The same strategy was used with the conserved primer pair *rps16\_IIP3*for and *trnQ\_IVP37*for for the corresponding region of the inversion between *rps16* and *trnQ*.

##### 2.2.2.5.4. Plastome sequencing

To complete the nucleotide sequences of the five basic *Oenothera* plastomes, already generated and assembled chromatograms by Hupfer (2002) were inspected manually for gaps or sequences of poor quality in all five basic plastomes, namely *Oenothera elata* subsp.

*hookeri* strain johansen (plastome I<sup>joh</sup>), *Oe. biennis* strain suaveolens Grado (plastome II<sup>suavG</sup>), *Oe. glazioviana* strain rr-lamarckiana Sweden (*ex genetica*: *Oe. lamarckiana*, plastome III<sup>lam</sup>), *Oe. parviflora* strain atrovirens Standard (plastome IV<sup>atro</sup>) and *Oe. argillicola* strain douthat 1 (plastome V<sup>doul</sup>). PCR derived fragments or fragments subcloned from plastid chromosomes (Hupfer, 2002) served as sequencing templates for both DNA strands. The fragments overlapped to a minimum of 200 bp. Almost the same, approximately 450, oligonucleotide primers (MWG Biotech, Ebersberg, Germany) employed for the already published plastome I<sup>joh</sup> were used (Hupfer *et al.*, 2000). After determination of nucleotide sequences, the data were subjected to the BLAST algorithm (Altschul *et al.*, 1990) provided by the National Centre for Biotechnology Information (Bethesda, MD). Assembly and evaluation of sequences were performed with the SeqMan 7.1 program (DNASTAR Inc, Madison, WI) using plastome I<sup>joh</sup> (AJ271079.2) as template for the other plastomes. All plastome sequences were aligned using the MegAlin 7.1 program (DNASTAR Inc, Madison, WI) and the program BioEdit 5.0.9 (North Carolina State University) as alignment editor (Hall, 1999).

#### 2.2.2.6. SNP mapping by Nuclease S digestion

To confirm additional single base pair insertions in *ndhD* and *rpl22* the primers P11for and ndhDint/2 for *ndhD* and P38for and P7rev for *rpl22* were used. PCR products were obtained by a proof reading polymerase. The products of different lengths (*e.g.* the 393 bp *ndhD* fragment of plastome I<sup>joh</sup> and the 392 bp *ndhD* fragment of plastome II<sup>suavG</sup>) were mixed. The templates were denatured, reannealed and cut at mismatches with Surveyor Nuclease S (Transgenomic, Elancourt, France) according to the manufacturer's protocol. The cleaved PCR products were analyzed on 2% agarose gels.

#### 2.2.2.7. Design, digestion and analysis of CAPS markers

CAPS markers were designed using the software SNP2CAPS (Thiel *et al.*, 2004). The PCR products obtained were digested with appropriate restriction endonucleases supplied by Fermentas International Inc, (Ontaria, Canada) or New England Biolabs (Ipswich, MA) according to the manufacturer's protocol and analyzed on agarose gels.

#### 2.2.2.8. Gene expression analysis

##### 2.2.2.8.1. Generation and application of macroarrays

Three *Oenothera* species *Oe. elata* subsp. *hookeri* strain johansen (AA-I), *Oe. grandiflora* strain tuscaloosa (BB-III) and *Oe. argillicola* strain douthat 1 (CC-V), from which fertile plastome-genome hybrids and cybrids can be produced were chosen to compare individual expression profiles. The universal vector primers M13forward and M13reverse were used to amplify PCR products of a subset of 187 selected cDNAs known or predicted to encode chloroplast proteins. All PCR products were subjected to agarose gel electrophoresis to confirm sizes and amplification quality. Once checked, each individual amplicon was adjusted to three different concentrations of 3.5, 14 and 56 ng/ $\mu$ l, respectively. Each dilution was spotted in duplicate onto a 7.8 x 11.9 cm positively charged Hybond N<sup>+</sup> nylon membrane (Amersham Biosciences, New York, NJ) by 20-fold repetition to the same points using robotics equipped with a 0.4 mm 96-pin gridder (BioRobotics, Cambridge, UK). As a negative control pBluescript vector (Stratagene, La Jolla, CA) was also spotted onto the filters. After spotting, filters were denatured in 1.5 M NaCl, 0.5 M NaOH and neutralized in 0.5 M Tris, pH 7.2, 1 M NaCl. After drying, filters were cross-linked with 120 mJ of 302 nm UV light by a UV-Stratalinker 1800 (Stratagene, La Jolla, CA). <sup>32</sup>P-dCTP labeled cDNA probes were synthesized from 10  $\mu$ g total RNA as described above. The labeled cDNAs were incubated for 20 min at 37°C with RNase H<sup>-</sup> (Invitrogen, Karlsruhe, Germany) to remove RNA. The labeled cDNAs were purified using MicroSpin<sup>TM</sup> G-50 columns (Amersham Biosciences, New York, NJ). The arrays were pre-hybridized for 2 h at 60°C in 0.25 M phosphate buffer and 7% SDS. The labeled cDNAs were hybridized to filters overnight at 60°C. Filters were washed twice at 60°C in 2 x SSC, 0.1% SDS and twice in 1 x SSC and 0.1% SDS. Filters were then exposed on Fuji Film imaging plates (Fuji, Tokyo, Japan). The radioactive images were obtained with the FLA-3000 phosphoimager (Fuji, Tokyo, Japan). Array images were imported into the program AIDA Image Analyzer 4.0 (Raytest GmbH, Straubenhardt, Germany) and signals were deduced. For normalization, the mean value of the selected background within each sub-grid was averaged and subtracted to calculate the intensity of all spots. The duplicate signals from 3 different concentrations were averaged and the expression profiles obtained were compared to calculate the ratios with program AIDA Array Compare 4.0 (Raytest GmbH, Straubenhardt, Germany). Histograms were generated using Microsoft Excel 2003 (Microsoft, Redmond, WA).

#### 2.2.2.8.2. Real-time PCR analysis

Real-time PCR was performed using a commercially available master mix containing Taq DNA polymerase, SYBR-Green I dye and deoxyribonucleoside triphosphates (LightCycler - FastStart DNA master SYBR-Green I, Roche Molecular Biochemicals, Mannheim, Germany). PCR products were followed by measuring SYBR Green I fluorescence. SYBR Green I dye emits a fluorescence signal at 530 nm only when bound to double-stranded DNA. Therefore, during PCR the increase in SYBR Green I fluorescence is directly proportional to the amount of double-stranded DNA generated. After addition of primers (0.5 mM), MgCl<sub>2</sub> (4 mM), and template cDNA to the master mix, an initial denaturation step at 95°C for 10 min, followed by 45 cycles of denaturation (95°C for 15 sec), annealing (58°C for 5 sec) and extension (72°C for 10 sec) were performed. All ramp rates were set to 20°C per sec. Detection of the fluorescent product was performed at the end of the extension period.

To prove that only the desired PCR product had been amplified, a melting curve analysis was performed after completion of PCR. For this, PCR products were denatured at 95°C, annealed at 55°C, and gradually heated to 95°C, whereas SYBR-Green I fluorescence was detected stepwise every 0.1°C. During such slow heating of the reaction mixture, melting of double-stranded DNA and a corresponding decrease of SYBR Green I fluorescence occurred. When the temperature of the reaction mixture reached the characteristic mean melting temperature of a particular DNA product (where the DNA is 50% double-stranded and 50% single-stranded), the first derivative presents a peak of a melting curve. If PCR generates only one amplicon, melting curve analysis shows only one melting peak. If primer dimers or other non-specific products are present, they cause additional melting peaks. To estimate primer-dimer formation, a control without template DNA was included in each experiment. The template quantification was determined by the crossing point using the LightCycler analysis software, as described in Wittwer *et al.* (1997).

##### 2.2.2.8.2.1. Analysis of nuclear gene expression via Real-time PCR

To normalize cDNAs of *Oe. elata* subsp. *hookeri* strain johansen (AA-I), *Oe. grandiflora* strain tuscaloosa (BB-III) and *Oe. argillicola* strain douthat 1 (CC-V), the expression of actin was measured with the primer pairs M101for/ M101rev, derived from EST cluster: S\_2275-22-F04. Subsequently the expression of the EST clusters C\_936-9-B11 (transketolase), C\_2590-26-F11 (phosphoribulokinase) and C\_4066-89-H09 (chlorophyll *a/b* binding family,

Elip2) was measured with the primer pairs M34for/M34rev, M60for/M60rev and M75for/M75rev, respectively. Real-time PCR was performed as described above. Due to sequence dissimilarity (1.59%) among different *Oenothera* species primer pairs designed for the A genome did not always amplify products in BB or CC species. This was also evident by temperature shifts of melting curves indicating unspecific PCR products.

#### 2.2.2.8.2.2. Analysis of plastid gene expression via Real-time PCR

AB-I (<sup>h</sup>johansen·<sup>h</sup>tuscaloosa I<sup>loh</sup>) and AB-III (<sup>h</sup>johansen·<sup>h</sup>tuscaloosa III<sup>lam/tusca</sup>) cDNAs were normalized to *psaC* using the primer pairs Ijeto and *psaC5'*. *PsaC* was chosen for normalization as no expression difference was detected in a marcoarry between the two genotypes (Geimer and Meurer, unpublished). The expression levels of *clpP* and *psbB* were determined with the primer pairs RPL20R5M and *clpP\_IIP31rev* for *clpP*, and *psbBfor* and *psbB\_IVP15rev* for *psbB*. Real-time PCR was performed as described above.

### 2.2.3. Analysis of proteins

#### 2.2.3.1. Preparation of thylakoid membranes

Leaves from adult plants were used for the isolation of thylakoid membranes according to Ossenbühl *et al.* (2004) with adjustments for *Oenothera* tissue. The modifications are essential to remove interfering mucilage. The entire procedure including centrifugation was performed at 4°C and in dim light. Approximately 4 g of leaf material were homogenized in 80 ml isolation buffer (330 mM sorbitol, 2 mM EDTA, 1 mM MgCl<sub>2</sub>, 5 mM ascorbate, 10 mM NaF, 50 mM HEPES/KOH, pH 7,5). 20 ml of the homogenized material was filtered through two layers of Miracloth. Then Miracloth was renewed and a further 20 ml were filtered. When filtering was completed, the lysate was centrifuged at 1,400 g for five minutes. Supernatants containing soluble proteins and mucilage were discarded. The pellet was resuspended in 3 ml washing buffer (5 mM sorbitol, 10 mM NaF, 50 mM HEPES/KOH, pH 7,5) and centrifuged for four minutes at 1,000 g. The washing step was repeated at least three times. Depending on mucilage content, the amount of washing buffer can be increased. The final thylakoid pellet was resuspended in 1 ml TMK buffer (100 mM sorbitol, 5 mM MgCl<sub>2</sub>, 10 mM NaF, 50 mM HEPES/KOH, pH 7,5), incubated for 10 min on ice for lysis, centrifuged again for 3 min and finally resuspended in 500 µl of the same buffer.

### 2.2.3.2. *Chlorophyll absorption measurements*

To estimate thylakoid membrane quantities chlorophyll concentrations were measured according to Arnon (1949).

### 2.2.3.3. *SDS polyacrylamide gel electrophoresis*

Proteins were separated by SDS polyacrylamide gel electrophoresis with a stacking gel of 8% acrylamide and a separating gel of 15% acrylamide. The protein samples were mixed with the same volume of loading buffer and loaded onto gels after denaturation for 1 min at 80°C.

### 2.2.3.4. *Western analysis*

For immunodetection, proteins were transferred to polyvinylidene difluoride (PVDF) membranes (Amersham Buchler, Braunschweig, Germany) according to the manufacturer's instructions, incubated with specific antisera, and visualized using the enhanced chemiluminescence technique (Amersham Buchler, Braunschweig, Germany).

### 2.2.4. *Determination of chromosome configurations*

Chromosome configurations of various *Oenothera* strains or hybrids were investigated with anther tissue. Although some variation was observed during the flowering season, anther length roughly correlates with the meiotic stage. There is also variation between anther length and developmental stage between strains or hybrids. In the strain *suaveolens* Grado the following rough correlation between anther length and meiotic stage was determined: approximately 2.0 mm - premeiotic interphase, 2.5 - 3.0 mm - leptotene-zygotene-pachytene, and 3.2 - 5.0 mm - from pachytene to tetrads (also diakinesis).

Inflorescences were collected from several plants of each strain, fixed in three parts 96% EtOH and one part 100% acetic acid and stored at -20°C. Anthers were washed several times in distilled water to remove the fixative and macerated for several minutes in 45% acetic acid. The anther loculi were separated under stereoscope microscope, cut transversely into halves and then squashed gently in 45% acetic acid by tapping the cover glass with a needle. After freezing using the dry-ice method, the cover glasses were removed, the preparations air-dried and stored at -20°C until use. The preparations were mounted in a drop of the DAPI staining solution (2 mg/1 ml; dissolved in a mixture of glycerol:McIlvaine buffer, 1:2), sealed with the

rubber cement “Fixogum” (Marabu GmbH & Co. KG, Tamm, Germany), and stored at 4 °C in the dark. Alternatively, CMA<sub>3</sub>-staining solution was applied according to Schweizer and Ambros (1994). Briefly, a drop of this solution was placed on a preparation and covered with a plastic cover slip. After three days incubation in the dark of a humid chamber at room temperature, preparations were drained, mounted in the CMA<sub>3</sub>-medium, and analyzed. Five to seven picture frames of the same object were taken, each at a different focus plane, by turning the micrometer screw of a fluorescence microscope equipped with a cooled CCD monochrome camera. They were quickly stacked and processed using standard macro commands ("do stack" and "do weighted average") of the CombineZM software (<http://www.hadleyweb.pwp.blueyonder.co.uk>) created by Alan Hadley.

#### 2.2.5. Chlorophyll *a* fluorescence analysis

Efficiency and functional state of photosystem II are reflected by chlorophyll *a* fluorescence parameters at room temperature (Schreiber *et al.*, 1986). They were calculated from wild-type *Oe. elata* subsp. *hookeri* strain johansen (AA-I), <sup>h</sup>johansen·<sup>h</sup>tuscaloosa III<sup>lam/tusca</sup> (AB-III), and incompatible <sup>h</sup>johansen·<sup>h</sup>tuscaloosa I<sup>oh</sup> (AB-I) rosette leaves using a pulse-modulated fluorimeter (PAM 101, Waltz, Effeltrich, Germany). The light intensity of the modulated measuring beam (1.6 kHz) was 0.5 μmol m<sup>-2</sup> sec<sup>-1</sup>. Leaves, dark adapted for 10 min, were used to detect the intrinsic (F<sub>o</sub>) and maximal (F<sub>m</sub>) fluorescence yields, the latter being determined by application of a saturating light pulse (0.8 sec, 7,000 μmol m<sup>-2</sup> sec<sup>-1</sup>). The potential maximum quantum yield was calculated as (F<sub>m</sub>-F<sub>o</sub>)/F<sub>m</sub> = F<sub>v</sub>/F<sub>m</sub>. Red actinic light (650 nm, 50 μmol m<sup>-2</sup> sec<sup>-1</sup>) was used for measurements of fluorescence quenching. Non-photochemical quenching (NPQ) was determined by applying repetitive saturation pulses with 20 sec intervals and calculated as (F<sub>m</sub>'-F)/(F<sub>m</sub>'-F<sub>o</sub>) (Kooten and Snel, 1990).

The light-dependent redox state of photosystem I was measured on leaves as absorption changes at 830 nm in the absence or presence of actinic (650 nm, 50 μmol m<sup>-2</sup> sec<sup>-1</sup>) and far red light (12 Wm<sup>-2</sup>) using the photosystem I attachment of PAM 101 (Klughammer and Schreiber, 1994). Saturating light pulses (0.8 s, 7,000 μmol m<sup>-2</sup> sec<sup>-1</sup>) were applied to follow photosystem II-dependent reduction of photosystem I in far-red background light.

## 2.2.6. Bioinformatic analysis

### 2.2.6.1. Calculation of genetic linkage

LOD scores and genetic linkage were calculated with the Kosambi function using the Joinmap program (van Ooijen and Voorrips, 2001).

### 2.2.6.2. Analysis of the *Oenothera* plastid genomes

#### 2.2.6.2.1. Repeat analysis

Applying the programs palindrome and etandem of the EMBOSS suite (Rice *et al.*, 2000) two different types of repeats, palindromes and direct tandems, were analyzed. The minimal cut-off identity between two copies was set to 90% for both repeat types. In case of multiple copies for one tandem, each copy was required to have at least one other member matching this constraint. 16 to 100 bp for palindromic and 10 to 100 bp for tandem repeats, respectively, were investigated for minimal and maximal copy sizes. Gap size between palindromes was restricted to a maximal length of 3 kb. Overlapping repeats with sequence similarity were grouped into one repeat motif. Both, the direct and inverted part of two repeats, had to overlap for palindromes. The longest element is provided as its representative for each repeat motif. Inversion breakpoints were analyzed separately in *berteriana* Schwemmler the plastomes of the subsection *Oenothera*.

#### 2.2.6.2.2. Analysis of variable amino acid sites

The degree of conservation of a single amino acid site between 30 reference species covering di- as well as several monocotyledonous plant species was investigated to estimate the impact of single amino acid exchanges to PGI within *Oenothera*. To exclude variable positions that were computationally deduced from misaligned indel regions, multiple alignments made were also inspected manually. In general, mutations exchanging amino acids with highly different biochemical properties have an increased likelihood to alter or destroy protein function. In contrast, highly conserved sites indicated by alignments of the reference species are less likely to undergo a drastic change. The distribution of biochemical properties of each *Oenothera* site was compared with the property distribution of the corresponding site in all reference species, to enrich candidate sites/proteins responsible for PGI. Applying the Grantham distance matrix (Grantham, 1974) biochemical properties were measured. From an all-against-all comparison of the *Oenothera* and reference species sites, respectively, distance distributions for a

particular site were derived. Statistically significant differences between the two distributions were tested by a non-parametric Wilcoxon rank sum test (<http://www.r-project.org/index.html>). P-values  $\leq 0.05$  were considered as significant. Sites representing an *Oenothera* specific adaptation, *i.e.* sites that are similar within each data set but dissimilar between both sets, are excluded by the test. To gather additional evidence for selected sites, the *Oenothera* mutations were checked whether they are located within known functional regions. Protein domains were detected applying HMMER 2.3.1 and the PFAM database, release of July 22, 2005 (Bateman *et al.*, 2004). Only alignments with e-values  $\leq 1e^{-10}$  were considered. Whether the corresponding domain position is highly conserved was checked by manual inspection of HMM-logos of the PFAM domain (<http://www.pfam.org>). Transmembrane domains were deduced by the InterPRO database (<http://www.ebi.ac.uk/interpro>) and analyzed with the online DAS server (Cserzo *et al.*, 1997). Since an alignment of single amino acids is not possible in regions of size variation, variant *Oenothera* proteins were inspected manually for functional domains.

#### 2.2.6.2.3. Computational prediction of sigma factor and T7 binding sites

To search for polymerase binding sites, multiple alignments of intergenic regions were delimited either by the 5' neighbouring gene or a maximum size of 600 bp. No attempt was made to differentiate between individual  $\sigma$ -factors, due to overlapping binding specificity of different bacterial polymerase-like (PEP)  $\sigma$ -factors in plastids (Liere and Börner, 2006). Therefore candidate sites were predicted by a consensus sequence, TYRMNN(N)<sub>16-20</sub>WANNWT, a search pattern, which covers a wide range of experimentally reported sites. The regular expression found was similar to but less specific than the consensus suggested by Homnn and Link (2003) and Kanamaru and Tanaka (2004). Matches to the consensus ATA<sub>0-1</sub>N<sub>0-1</sub>GAA(N)<sub>15-23</sub>YRT (Silhavy and Maliga, 1998; Kapoor and Sugiura, 1999) were defined as binding sites of the phage type polymerase (NEP-promoters), representing the NEP type Ib-promoters. The sequences were not investigated for the type Ia and type II NEP-promoters. The consensus of type Ia (YRTa), which can be considered as derivatives of the type Ib consensus, is too low for computational predictions. The type II NEP-promoter element is known just from a single case (Liere and Börner, 2006). Candidate binding sites were positioned within the multiple alignments and edited by manual supervision to correct for misaligned regions, *e.g.* due to small repeats.

#### 2.2.6.2.4. Prediction of Shine-Dalgarno sequences

To search for candidate Shine-Dalgarno regions sequences 50 bp upstream of the start codon were investigated, using the program free2bind (Starmer *et al.*, 2006) and the 3' 16S RNA sequence of *Oenothera*. A minimum free energy of 4.4 kcal and a maximum distance to the start codon of 23 bp was required for the reported matches.

#### 2.2.6.2.5. Calculation of phylogenetic trees

To generate phylogenetic trees, multiple codon-based alignments of the 47 genes variable in the five *Oenothera* plastomes were used. Including the corresponding sequences of the *Lotus japonicus* plastome (Kato *et al.*, 2000) as outgroup for tree rooting, the dataset comprises 44,472 aligned characters present in six species. Neighbor-Joining (NJ), Maximum-Likelihood (ML) and Maximum Parsimony (MP) from the PHYLIP package (Felsenstein, 1993) were applied to infer trees. With 1000 random samples each, bootstrap analysis for NJ and ML was performed. In addition, gene specific phylogenetic trees for all variable genes were determined by NJ and ML. Using the consense program of PHYLIP, a species tree was then built from individual gene trees. Trees for non-coding sequences were derived from 76 intergenic regions, which showed nucleotide substitutions between the five *Oenothera* plastomes.

#### 2.2.6.2.6. Determination of Ka/Ks-values

Applying the yn00 program of the PAML package (Yang, 1997), synonymous and non-synonymous substitution rates were estimated. Ka and Ks were determined by the Nei-Gojobori method as implemented in yn00 and F3x4 were selected as substitution matrix. From pairwise codon-based alignments, rates for protein-coding genes variable among at least two of the five *Oenothera* plastomes were estimated. For five different plastomes, there are 10 pairwise combinations for each gene, resulting in a total of 780 rates for all and 470 rates for variable genes. The computation of  $\omega = Ka/Ks$  was not always applicable (*e.g.* for  $Ks = 0$ ). Therefore,  $\omega$  could be determined for only 215 pairwise combinations. A concatenated alignment of individual protein coding regions was analyzed, to compare average Ka and Ks rates between species.

### 3. Results

A major aim of this study was to introduce *Oenothera* as a molecular model for novel, fundamental biological questions. At the beginning of this work *Oenothera* cell and tissue culture as well as nuclear transformation had been established (Stubbe and Herrmann, 1982; Kuchuk *et al.*, 1998; Mehra-Palta *et al.*, 1998; de Gyves *et al.*, 2001), there was access to an EST library (Mráček *et al.*, 2006), to an AFLP map consisting of 202 dominant markers assigned to seven coupling groups (Mráček, 2005), and to the complete sequence of a plastid chromosome (Hupfer *et al.*, 2000). Furthermore, a rich stock of experimental strains and species, including various interspecific plastome-genome incompatible lines, or crossing intermediates to produce such lines, was available. The collection, predominantly based on the work of Wilfried Stubbe and Werner Dietrich (Stubbe, 1959; Dietrich *et al.*, 1997), included various strains with a long lasting genetic tradition that originated from the collections of Ralph E. Cleland, Otto Renner and Hugo de Vries.

Two major molecular approaches were performed and applied to the material. First, a marker system for both, nucleus and plastid was developed that could be used for a large number of strains and species. The utility of the markers was checked with breeding approaches and employed, *e.g.* to assemble plastome-genome incompatible plants, to improve the quality of the existing AFLP map (Mráček, 2005), or to follow up distinct steps of interspecific chromosome exchanges. Second, the nucleotide sequences of representatives of the five basic plastome types were generated or completed, respectively, and compared. Based on the evaluated sequence, a novel strategy was developed that rests on a comparison of plastid chromosomal sequence differences with formal genetic data, to identify molecular determinants that are involved in PGI in the subsection *Oenothera*.

#### 3.1. Molecular approaches in *Oenothera* genetics

This chapter describes attempts to generate an efficient SSLP and CAPS marker system for *Oenothera* genetics. It allows the discrimination not only of all basic plastomes and nuclear genomes, but also of individual Renner complexes and subplastomes, in sexual crosses. The application of the marker system will be documented with the generation of the artificial and incompatible plastome-genome combinations AB-I and BB-II and, in a connection of both, classical and molecular genetics - assigning linkage groups to distinct chromosomes in genetic crosses. In a pilot experiment chromosome 9·8 of the classical map could be merged

in this thesis with linkage group 7 of the molecular one (Cleland, 1972; Mráček, 2005). Taken together these findings provide fundamental progress in *Oenothera* genetics, since they significantly improve the use of the genus as a molecular model and open wide possibilities in *Oenothera* research and breeding.

### 3.1.1. Marker systems for *Oenothera* genetics and breeding approaches

Up to now, only a few phenotypic markers for nucleus and plastome (in terms of plastome mutants) were available and used in *Oenothera* genetics (Cleland, 1972; Stubbe and Herrmann, 1982; Stubbe, 1989; Dietrich *et al.*, 1997). Also only a small number of molecular markers were described for a very limited number of strains (Mráček, 2005; Larson *et al.*, 2008).

#### 3.1.1.1. Co-dominant markers discriminating A and B genomes

The EST library (Mráček *et al.*, 2006) was utilized as a source to generate co-dominant markers that could be used for mapping and genotyping. The 5' terminal sequences of selected cDNAs from the available hookeri de Vries EST library were used to generate primers and to amplify distinct regions from genomic DNA by PCR from two homozygous species namely, *Oe. elata* subsp. *hookeri* strain johansen (genome AA; <sup>h</sup>johansen·<sup>h</sup>johansen) and *Oe. grandiflora* strain tuscaloosa (BB; <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa). 98 selected primer pairs generated from the hookeri de Vries (AA) library were able to amplify 75 products from strain johansen (AA) and strain tuscaloosa (BB), reflecting the close relationship between these species. From 46 PCR products each, sequences were determined and compared between the A and B genotypes. Marker sequences are deposited in GenBank under the accession numbers listed in Table 7. The average length of the products was approximately 300 bp, twelve products contained an intron; one gene duplication was detected. Within the 46 PCR products investigated eleven pairs showed size polymorphisms, 29 gave at least one CAPS marker appropriate to distinguish both genotypes by analyzing cleaved products on agarose gels. Only a single size polymorphism could directly be analyzed on agarose gels. In 35 products a minimum of one polymorphism was present (Table 7). Also the duplicated locus showed polymorphism between <sup>h</sup>johansen and <sup>h</sup>tuscaloosa but was excluded from Table 7. The less conserved, occasionally intron-containing 5' UTR and 5' gene coding regions of the AA and BB nuclear genomes used in this study displayed 98% sequence identity. These findings demonstrate that the *Oenothera* EST library is a rich source for PCR based markers.

**Table 7.** PCR-based polymorphisms found between the Renner complexes of the A genome <sup>h</sup>johansen (above) and the B genome <sup>h</sup>tuscaloosa (below)

marker	accession number	intron	SSLP	no. SNPs	marker type (enzyme)	predicted PCR products [bp] <sup>1)</sup>	predicted restriction fragments [bp] <sup>1)</sup>	primer	primer sequence (5' to 3')	EST/cluster accession	closest <i>Arabidopsis</i> homologue (blastX)	protein function and localization in <i>Arabidopsis</i>																																																																																																																																																																																																																															
<b>M02<sup>2)</sup></b>	EU483117	no	no	2	CAPS ( <i>ApeKI</i> )	288	250, 38	M02for	tggccatggcgacacaagcctc	C_4044-89-F11	At1g03130	photosystem I reaction centre subunit (PsaD2), putative, chloroplast																																																																																																																																																																																																																															
	EU483119					288	186, 64, 38	M02rev	cctcaacctgagccttacggag				<b>M03<sup>2)</sup></b>	EU483121	yes	no	1	SNP	466	N/A	M03for	atatcacctggctactgtagct	C_4496-95-C09	At2g39470	photosystem II oxygen-evolving complex subunit (PsbP-like), chloroplast	EU483123	466	M03rev	aactccctccaatctgaagggt	<b>M07<sup>2)</sup></b>	EU483125	no	no	2	CAPS ( <i>PstI</i> )	356	356	M07for	accataccatataccagctgc	S_4170-90-H07	At5g64380	fructose-bisphosphatase-like protein, mitochondrium	EU483127	356	299, 57	M07rev	tcaagcgctctcggtgcatctc	<b>M08<sup>2)</sup></b>	EU483129	no	no	2	CAPS ( <i>BsuRI</i> )	282	247, 35	M08for	ctcagccaggaggacctcaagc	S_3501-84-D07	At2g01290	ribose 5-phosphate isomerase, localization unknown	EU483131	282	173, 74, 35	M08rev	gaggtgggtatgacctctgctg	<b>M13</b>	EU447201	no	no	3	CAPS ( <i>NciI</i> )	328	241, 86, 1	M13for	atcgatcatcatggcca	C_1649-15-A03	At1g29930	chlorophyll <i>a/b</i> -binding protein like, photosystem II 5 kD protein, chloroplast	EU447202	328	123, 118, 86, 1	M13rev	cgagaatggatcacctcca	<b>M15</b>	EU447203	yes	8 bp	17	CAPS ( <i>StyI</i> )	399	162, 149, 88	M15for	ttgaggagggtgcagttacaga	C_179-81-E03	At1g61520	Lhca2 protein, PSI type III chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447204	391	242, 149	M15rev	cgtccaagtaggtcaggctct	<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447206	503	403, 85, 15	M17rev	catctgcagtggtagatctctga	<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for
<b>M03<sup>2)</sup></b>	EU483121	yes	no	1	SNP	466	N/A	M03for	atatcacctggctactgtagct	C_4496-95-C09	At2g39470	photosystem II oxygen-evolving complex subunit (PsbP-like), chloroplast																																																																																																																																																																																																																															
	EU483123					466		M03rev	aactccctccaatctgaagggt				<b>M07<sup>2)</sup></b>	EU483125	no	no	2	CAPS ( <i>PstI</i> )	356	356	M07for	accataccatataccagctgc	S_4170-90-H07	At5g64380	fructose-bisphosphatase-like protein, mitochondrium	EU483127	356	299, 57	M07rev	tcaagcgctctcggtgcatctc	<b>M08<sup>2)</sup></b>	EU483129	no	no	2	CAPS ( <i>BsuRI</i> )	282	247, 35	M08for	ctcagccaggaggacctcaagc	S_3501-84-D07	At2g01290	ribose 5-phosphate isomerase, localization unknown	EU483131	282	173, 74, 35	M08rev	gaggtgggtatgacctctgctg	<b>M13</b>	EU447201	no	no	3	CAPS ( <i>NciI</i> )	328	241, 86, 1	M13for	atcgatcatcatggcca	C_1649-15-A03	At1g29930	chlorophyll <i>a/b</i> -binding protein like, photosystem II 5 kD protein, chloroplast	EU447202	328	123, 118, 86, 1	M13rev	cgagaatggatcacctcca	<b>M15</b>	EU447203	yes	8 bp	17	CAPS ( <i>StyI</i> )	399	162, 149, 88	M15for	ttgaggagggtgcagttacaga	C_179-81-E03	At1g61520	Lhca2 protein, PSI type III chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447204	391	242, 149	M15rev	cgtccaagtaggtcaggctct	<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447206	503	403, 85, 15	M17rev	catctgcagtggtagatctctga	<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg									
<b>M07<sup>2)</sup></b>	EU483125	no	no	2	CAPS ( <i>PstI</i> )	356	356	M07for	accataccatataccagctgc	S_4170-90-H07	At5g64380	fructose-bisphosphatase-like protein, mitochondrium																																																																																																																																																																																																																															
	EU483127					356	299, 57	M07rev	tcaagcgctctcggtgcatctc				<b>M08<sup>2)</sup></b>	EU483129	no	no	2	CAPS ( <i>BsuRI</i> )	282	247, 35	M08for	ctcagccaggaggacctcaagc	S_3501-84-D07	At2g01290	ribose 5-phosphate isomerase, localization unknown	EU483131	282	173, 74, 35	M08rev	gaggtgggtatgacctctgctg	<b>M13</b>	EU447201	no	no	3	CAPS ( <i>NciI</i> )	328	241, 86, 1	M13for	atcgatcatcatggcca	C_1649-15-A03	At1g29930	chlorophyll <i>a/b</i> -binding protein like, photosystem II 5 kD protein, chloroplast	EU447202	328	123, 118, 86, 1	M13rev	cgagaatggatcacctcca	<b>M15</b>	EU447203	yes	8 bp	17	CAPS ( <i>StyI</i> )	399	162, 149, 88	M15for	ttgaggagggtgcagttacaga	C_179-81-E03	At1g61520	Lhca2 protein, PSI type III chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447204	391	242, 149	M15rev	cgtccaagtaggtcaggctct	<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447206	503	403, 85, 15	M17rev	catctgcagtggtagatctctga	<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																											
<b>M08<sup>2)</sup></b>	EU483129	no	no	2	CAPS ( <i>BsuRI</i> )	282	247, 35	M08for	ctcagccaggaggacctcaagc	S_3501-84-D07	At2g01290	ribose 5-phosphate isomerase, localization unknown																																																																																																																																																																																																																															
	EU483131					282	173, 74, 35	M08rev	gaggtgggtatgacctctgctg				<b>M13</b>	EU447201	no	no	3	CAPS ( <i>NciI</i> )	328	241, 86, 1	M13for	atcgatcatcatggcca	C_1649-15-A03	At1g29930	chlorophyll <i>a/b</i> -binding protein like, photosystem II 5 kD protein, chloroplast	EU447202	328	123, 118, 86, 1	M13rev	cgagaatggatcacctcca	<b>M15</b>	EU447203	yes	8 bp	17	CAPS ( <i>StyI</i> )	399	162, 149, 88	M15for	ttgaggagggtgcagttacaga	C_179-81-E03	At1g61520	Lhca2 protein, PSI type III chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447204	391	242, 149	M15rev	cgtccaagtaggtcaggctct	<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447206	503	403, 85, 15	M17rev	catctgcagtggtagatctctga	<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																													
<b>M13</b>	EU447201	no	no	3	CAPS ( <i>NciI</i> )	328	241, 86, 1	M13for	atcgatcatcatggcca	C_1649-15-A03	At1g29930	chlorophyll <i>a/b</i> -binding protein like, photosystem II 5 kD protein, chloroplast																																																																																																																																																																																																																															
	EU447202					328	123, 118, 86, 1	M13rev	cgagaatggatcacctcca				<b>M15</b>	EU447203	yes	8 bp	17	CAPS ( <i>StyI</i> )	399	162, 149, 88	M15for	ttgaggagggtgcagttacaga	C_179-81-E03	At1g61520	Lhca2 protein, PSI type III chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447204	391	242, 149	M15rev	cgtccaagtaggtcaggctct	<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447206	503	403, 85, 15	M17rev	catctgcagtggtagatctctga	<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																															
<b>M15</b>	EU447203	yes	8 bp	17	CAPS ( <i>StyI</i> )	399	162, 149, 88	M15for	ttgaggagggtgcagttacaga	C_179-81-E03	At1g61520	Lhca2 protein, PSI type III chlorophyll <i>a/b</i> -binding protein, chloroplast																																																																																																																																																																																																																															
	EU447204					391	242, 149	M15rev	cgtccaagtaggtcaggctct				<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast	EU447206	503	403, 85, 15	M17rev	catctgcagtggtagatctctga	<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																	
<b>M17</b>	EU447205	yes	9 bp	11	CAPS ( <i>TaqI</i> )	512	427, 85	M17for	gagatcacagtagtaatggctcca	C_1955-18-D05	At3g54890	chlorophyll <i>a/b</i> -binding protein, chloroplast																																																																																																																																																																																																																															
	EU447206					503	403, 85, 15	M17rev	catctgcagtggtagatctctga				<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast	EU447208	396	396	M19rev	cacctgcctcaccggaact	<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																			
<b>M19</b>	EU447207	no	no	1	CAPS ( <i>PflMI</i> )	396	208, 188	M19for	aatcctaattggctcctctaca	C_2501-25-D11	At1g29920	chlorophyll <i>a/b</i> -binding protein – like, chloroplast																																																																																																																																																																																																																															
	EU447208					396	396	M19rev	cacctgcctcaccggaact				<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast	EU447210	362	231, 131	M23rev	ggagtgatgacctcgagct	<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																																					
<b>M23</b>	EU447209	no	2 bp	10	CAPS ( <i>StyI</i> )	364	364	M23for	ccacgcgaactctttaaact	C_46-4-A11	At2g34420	photosystem II type I chlorophyll <i>a/b</i> binding protein, chloroplast																																																																																																																																																																																																																															
	EU447210					362	231, 131	M23rev	ggagtgatgacctcgagct				<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	EU447212	278	212, 39, 27	M28rev	gcgactaaggggacgctatcg	<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																																																							
<b>M28</b>	EU447211	no	no	1	CAPS ( <i>BsuRI</i> )	278	182, 39, 30, 27	M28for	ggctccgacatccttgggag	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast																																																																																																																																																																																																																															
	EU447212					278	212, 39, 27	M28rev	gcgactaaggggacgctatcg				<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	EU447214	310	M33rev	acctcaagtggggagctctg	<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																																																																									
<b>M33</b>	EU447213	yes	1 bp	1	SSLP	309	not specifiable on agarose gels	M33for	tcaggcctcaagagctcagcc	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast																																																																																																																																																																																																																															
	EU447214					310		M33rev	acctcaagtggggagctctg				<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	EU447216	213	200, 13	M38rev	gtccgaccaagcagcgacgtt	<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																																																																																										
<b>M38</b>	EU447215	no	no	2	CAPS ( <i>BsuRI</i> )	213	156, 44, 13	M38for	ggcaaaagctatggccactctc	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast																																																																																																																																																																																																																															
	EU447216					213	200, 13	M38rev	gtccgaccaagcagcgacgtt				<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	EU447218	680	374, 306	M39rev	ggaaccagtagctagtagcttgc	<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																																																																																																												
<b>M39</b>	EU447217	yes	no	2	CAPS ( <i>BclI</i> )	680	374, 262, 44	M39for	ccaaagtggatcgcggtgtc	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast																																																																																																																																																																																																																															
	EU447218					680	374, 306	M39rev	ggaaccagtagctagtagcttgc				<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	EU432401	500	M40rev	tcagccctttgtccgaagtgcg																																																																																																																																																																																																														
<b>M40</b>	EU432390	yes	117 bp	N/A	SSLP	583	N/A	M40for	accgtctctccaagcactgc	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast																																																																																																																																																																																																																															
	EU432401					500		M40rev	tcagccctttgtccgaagtgcg																																																																																																																																																																																																																																		

Table 7. (continued)

marker	accession number	intron	SSLP	no. SNPs	marker type (enzyme)	predicted PCR products [bp] <sup>1)</sup>	predicted restriction fragments [bp] <sup>1)</sup>	primer	primer sequence (5' to 3')	EST/cluster accession	closest <i>Arabidopsis</i> homologue (blastX)	protein function and localization in <i>Arabidopsis</i>																																																																																																																																																																																																																																		
<b>M41</b>	EU447219	no	no	1	CAPS ( <i>EarI</i> )	267	161, 106	M41for	acaccctcttatccaatggc	C_1234-11-B07	At2g45290	transketolase, putative, chloroplast																																																																																																																																																																																																																																		
	EU447220					267	267	M41rev	ttccacgagagtgtccgtgg				<b>M43</b>	EU447221	no	no	11	CAPS ( <i>BsuRI</i> )	275	206, 47, 22	M43for	accacattctcaaaagctccg	S_1221-11-A06	At3g63140	mRNA binding protein precursor – like, chloroplast	EU447222	275	118, 88, 47, 22	M43rev	cggagcaagaagctctttgg	<b>M44</b>	EU447223	yes	6 bp	6	SSLP	305	not specifi- able	M44for	tcaacaatggctgccgcagtg	C_1431-12-H01	At1g20020	ferredoxin-NADP+ reductase-like protein, chloroplast	EU447224	311	on agarose gels	M44rev	agtgtctcaccctgccggag	<b>M46</b>	EU447225	no	no	6	CAPS ( <i>XhoI</i> )	194	108, 86	M46for	aaatggcgtccatggcgctta	S_1491-13-D02	At5g12860	2-oxoglutarate/malate translocator-like protein, chloroplast	EU447226	194	194	M46rev	ctgggactcaagctcggcag	<b>M47</b>	EU447227	no	no	1	CAPS ( <i>TaqI</i> )	257	257	M47for	tgggtgggattgccctacgtg	S_1494-13-D05	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	EU447228	257	175, 82	M47rev	gcgacaaccttaacctgtcg	<b>M48</b>	EU447229	no	3 bp	5	CAPS ( <i>BseRI</i> )	247	97, 95, 52, 3	M48for	tctctgactccactccactgc	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), endomembrane system	EU447230	244	189, 52, 3	M48rev	agcttctgtgtgagcttggct	<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	EU447232	225	148, 77	M50rev	accaacgaaccgtctagccag	<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac
<b>M43</b>	EU447221	no	no	11	CAPS ( <i>BsuRI</i> )	275	206, 47, 22	M43for	accacattctcaaaagctccg	S_1221-11-A06	At3g63140	mRNA binding protein precursor – like, chloroplast																																																																																																																																																																																																																																		
	EU447222					275	118, 88, 47, 22	M43rev	cggagcaagaagctctttgg				<b>M44</b>	EU447223	yes	6 bp	6	SSLP	305	not specifi- able	M44for	tcaacaatggctgccgcagtg	C_1431-12-H01	At1g20020	ferredoxin-NADP+ reductase-like protein, chloroplast	EU447224	311	on agarose gels	M44rev	agtgtctcaccctgccggag	<b>M46</b>	EU447225	no	no	6	CAPS ( <i>XhoI</i> )	194	108, 86	M46for	aaatggcgtccatggcgctta	S_1491-13-D02	At5g12860	2-oxoglutarate/malate translocator-like protein, chloroplast	EU447226	194	194	M46rev	ctgggactcaagctcggcag	<b>M47</b>	EU447227	no	no	1	CAPS ( <i>TaqI</i> )	257	257	M47for	tgggtgggattgccctacgtg	S_1494-13-D05	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	EU447228	257	175, 82	M47rev	gcgacaaccttaacctgtcg	<b>M48</b>	EU447229	no	3 bp	5	CAPS ( <i>BseRI</i> )	247	97, 95, 52, 3	M48for	tctctgactccactccactgc	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), endomembrane system	EU447230	244	189, 52, 3	M48rev	agcttctgtgtgagcttggct	<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	EU447232	225	148, 77	M50rev	accaacgaaccgtctagccag	<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag										
<b>M44</b>	EU447223	yes	6 bp	6	SSLP	305	not specifi- able	M44for	tcaacaatggctgccgcagtg	C_1431-12-H01	At1g20020	ferredoxin-NADP+ reductase-like protein, chloroplast																																																																																																																																																																																																																																		
	EU447224					311	on agarose gels	M44rev	agtgtctcaccctgccggag				<b>M46</b>	EU447225	no	no	6	CAPS ( <i>XhoI</i> )	194	108, 86	M46for	aaatggcgtccatggcgctta	S_1491-13-D02	At5g12860	2-oxoglutarate/malate translocator-like protein, chloroplast	EU447226	194	194	M46rev	ctgggactcaagctcggcag	<b>M47</b>	EU447227	no	no	1	CAPS ( <i>TaqI</i> )	257	257	M47for	tgggtgggattgccctacgtg	S_1494-13-D05	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	EU447228	257	175, 82	M47rev	gcgacaaccttaacctgtcg	<b>M48</b>	EU447229	no	3 bp	5	CAPS ( <i>BseRI</i> )	247	97, 95, 52, 3	M48for	tctctgactccactccactgc	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), endomembrane system	EU447230	244	189, 52, 3	M48rev	agcttctgtgtgagcttggct	<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	EU447232	225	148, 77	M50rev	accaacgaaccgtctagccag	<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																												
<b>M46</b>	EU447225	no	no	6	CAPS ( <i>XhoI</i> )	194	108, 86	M46for	aaatggcgtccatggcgctta	S_1491-13-D02	At5g12860	2-oxoglutarate/malate translocator-like protein, chloroplast																																																																																																																																																																																																																																		
	EU447226					194	194	M46rev	ctgggactcaagctcggcag				<b>M47</b>	EU447227	no	no	1	CAPS ( <i>TaqI</i> )	257	257	M47for	tgggtgggattgccctacgtg	S_1494-13-D05	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	EU447228	257	175, 82	M47rev	gcgacaaccttaacctgtcg	<b>M48</b>	EU447229	no	3 bp	5	CAPS ( <i>BseRI</i> )	247	97, 95, 52, 3	M48for	tctctgactccactccactgc	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), endomembrane system	EU447230	244	189, 52, 3	M48rev	agcttctgtgtgagcttggct	<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	EU447232	225	148, 77	M50rev	accaacgaaccgtctagccag	<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																														
<b>M47</b>	EU447227	no	no	1	CAPS ( <i>TaqI</i> )	257	257	M47for	tgggtgggattgccctacgtg	S_1494-13-D05	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast																																																																																																																																																																																																																																		
	EU447228					257	175, 82	M47rev	gcgacaaccttaacctgtcg				<b>M48</b>	EU447229	no	3 bp	5	CAPS ( <i>BseRI</i> )	247	97, 95, 52, 3	M48for	tctctgactccactccactgc	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), endomembrane system	EU447230	244	189, 52, 3	M48rev	agcttctgtgtgagcttggct	<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	EU447232	225	148, 77	M50rev	accaacgaaccgtctagccag	<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																
<b>M48</b>	EU447229	no	3 bp	5	CAPS ( <i>BseRI</i> )	247	97, 95, 52, 3	M48for	tctctgactccactccactgc	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), endomembrane system																																																																																																																																																																																																																																		
	EU447230					244	189, 52, 3	M48rev	agcttctgtgtgagcttggct				<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	EU447232	225	148, 77	M50rev	accaacgaaccgtctagccag	<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																		
<b>M50</b>	EU447231	no	no	3	CAPS ( <i>HhaI</i> )	225	225	M50for	ctgtccaccacaatggctgc	C_1598-14-D12	At2g40100	photosystem II chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast																																																																																																																																																																																																																																		
	EU447232					225	148, 77	M50rev	accaacgaaccgtctagccag				<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	EU447234	235	235	M52rev	tccatgtaggcatcgagctcc	<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																				
<b>M52</b>	EU447233	no	no	1	CAPS ( <i>HpyCH4III</i> )	235	120, 115	M52for	aagcagccatggcgacatctc	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast																																																																																																																																																																																																																																		
	EU447234					235	235	M52rev	tccatgtaggcatcgagctcc				<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol	EU447236	193	193	M57rev	agaatgaccacggaatgtcc	<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																						
<b>M57</b>	EU447235	no	no	3	CAPS ( <i>BsrI</i> )	193	126, 67	M57for	ctgatgttctctccaagatg	C_2295-22-H03	At3g55440	triosephosphate isomerase, cytosol																																																																																																																																																																																																																																		
	EU447236					193	193	M57rev	agaatgaccacggaatgtcc				<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast	EU447238	544	280, 155, 109	M58rev	ctgaactgccacggctgttgg	<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																																								
<b>M58</b>	EU447237	yes	11 bp	15	CAPS ( <i>BsrI</i> )	533	424, 109	M58for	gatccggaggatggaagtctc	S_2302-22-H10	At4g02510	protein import component Toc159-like, chloroplast																																																																																																																																																																																																																																		
	EU447238					544	280, 155, 109	M58rev	ctgaactgccacggctgttgg				<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast	EU447240	699	517, 178, 4	M59rev	caaacctctgtgtggccacac	<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																																																										
<b>M59</b>	EU447239	yes	4 bp	51	CAPS ( <i>RsaI</i> )	703	322, 199, 178, 4	M59for	tgtctccgccacaatgtccg	C_2346-23-D11	At1g67090	ribulose bisphosphate carboxylase, small subunit (RuBisCO), chloroplast																																																																																																																																																																																																																																		
	EU447240					699	517, 178, 4	M59rev	caaacctctgtgtggccacac				<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol	EU447242	231	on agarose gels	M60rev	ctcttaccgcagccggaatcc	<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																																																																												
<b>M60</b>	EU447241	no	no	7	CAPS ( <i>AluI</i> )	231	not specifi- able	M60for	gcaaccaacaatggcggtctg	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgl1), mitochondrion, cytosol																																																																																																																																																																																																																																		
	EU447242					231	on agarose gels	M60rev	ctcttaccgcagccggaatcc				<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	EU447244	250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg	<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																																																																																														
<b>M74</b>	EU447243	no	no	3	CAPS ( <i>HhaI</i> )	250	208, 40, 2	M74for	aatggcgctctccagcagac	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast																																																																																																																																																																																																																																		
	EU447244					250	104, 54, 50, 40, 2	M74rev	tgtttctgagagtaccgttgg				<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	EU447246	149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																																																																																																																
<b>M75</b>	EU447245	no	no	1	CAPS ( <i>AluI</i> )	149	120, 29	M75for	gtctgttatcagagtctgggac	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast																																																																																																																																																																																																																																		
	EU447246					149	64, 56, 29	M75rev	cctgatcagccatgcactctgag																																																																																																																																																																																																																																					

Table 7. (continued)

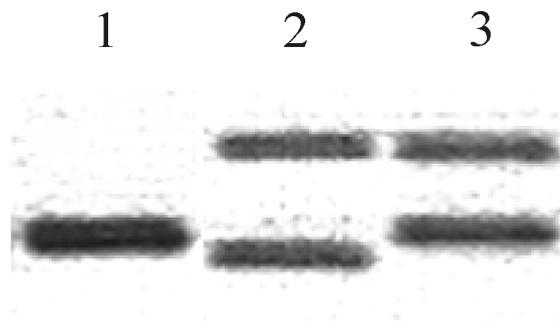
marker	accession number	intron	SSLP	no. SNPs	marker type (enzyme)	predicted PCR products [bp] <sup>1)</sup>	predicted restriction fragments [bp] <sup>1)</sup>	primer	primer sequence (5' to 3')	EST/cluster accession	closest <i>Arabidopsis</i> homologue (blastX)	protein function and localization in <i>Arabidopsis</i>																																																																																		
<b>M82</b>	EU447247	no	no	2	CAPS ( <i>NaeI</i> )	223	223	M82for	agcaccatgggtgagcacctcc	C_5307-94-E09	At1g60950	ferredoxin (FedA), chloroplast																																																																																		
	EU447248					223	137, 86	M82rev	agtagggcaaatgattccctc				<b>M86</b>	EU447249	no	no	4	CAPS ( <i>DdeI</i> )	208	84, 66, 45, 13	M86for	tcctcattctctacctccagag	C_4643-96-G12	At2g21170	triosephosphatisomerase, chloroplast	EU447250	208	150, 58	M86rev	accagccatagcaacgacgcc	<b>M88</b>	EU447251	no	no	1	CAPS ( <i>DdeI</i> )	176	176	M88for	accacagtctccgagtaact	C_4753-98-B06	At5g50250	RNA binding protein (rbp31), chloroplast	EU447252	176	152, 24	N88rev	tggtgagcccaatccgaggtc	<b>M95</b>	EU447253	yes	no	2	CAPS ( <i>HhaI</i> )	315	299, 16	M95for	tcggactcagcaatggcgctc	C_5102-112-A10	At5g54190	NADPH:protochlorophyllide oxidoreductase A, chloroplast	EU447254	315	252, 47, 16	M95rev	tggtggctgtctgtgctcgaa	<b>M97<sup>2)</sup></b>	EU483132	no	18 bp	23	CAPS ( <i>ApeKI</i> )	357	246, 71, 29, 11	M97for	atgaaagcacaaggagtccctc	S_1348-12-C09	At5g47560	malate/fumarate transporter, tonoplast	EU483134	375	346, 29	M97rev	cgagaatgaagctgcctaaga	<b>M98<sup>2)</sup></b>	EU483136	no	11 bp	24	CAPS ( <i>PfI</i> )	459	231, 184, 44	M98for	aagccgagatcctcctgcaatgg
<b>M86</b>	EU447249	no	no	4	CAPS ( <i>DdeI</i> )	208	84, 66, 45, 13	M86for	tcctcattctctacctccagag	C_4643-96-G12	At2g21170	triosephosphatisomerase, chloroplast																																																																																		
	EU447250					208	150, 58	M86rev	accagccatagcaacgacgcc				<b>M88</b>	EU447251	no	no	1	CAPS ( <i>DdeI</i> )	176	176	M88for	accacagtctccgagtaact	C_4753-98-B06	At5g50250	RNA binding protein (rbp31), chloroplast	EU447252	176	152, 24	N88rev	tggtgagcccaatccgaggtc	<b>M95</b>	EU447253	yes	no	2	CAPS ( <i>HhaI</i> )	315	299, 16	M95for	tcggactcagcaatggcgctc	C_5102-112-A10	At5g54190	NADPH:protochlorophyllide oxidoreductase A, chloroplast	EU447254	315	252, 47, 16	M95rev	tggtggctgtctgtgctcgaa	<b>M97<sup>2)</sup></b>	EU483132	no	18 bp	23	CAPS ( <i>ApeKI</i> )	357	246, 71, 29, 11	M97for	atgaaagcacaaggagtccctc	S_1348-12-C09	At5g47560	malate/fumarate transporter, tonoplast	EU483134	375	346, 29	M97rev	cgagaatgaagctgcctaaga	<b>M98<sup>2)</sup></b>	EU483136	no	11 bp	24	CAPS ( <i>PfI</i> )	459	231, 184, 44	M98for	aagccgagatcctcctgcaatgg	C_1202-10-G11	At2g06520	PsbX (photosystem II subunit X), chloroplast	EU483138	470	426, 44	M98rev	aggcaaaataaacggggatacagc										
<b>M88</b>	EU447251	no	no	1	CAPS ( <i>DdeI</i> )	176	176	M88for	accacagtctccgagtaact	C_4753-98-B06	At5g50250	RNA binding protein (rbp31), chloroplast																																																																																		
	EU447252					176	152, 24	N88rev	tggtgagcccaatccgaggtc				<b>M95</b>	EU447253	yes	no	2	CAPS ( <i>HhaI</i> )	315	299, 16	M95for	tcggactcagcaatggcgctc	C_5102-112-A10	At5g54190	NADPH:protochlorophyllide oxidoreductase A, chloroplast	EU447254	315	252, 47, 16	M95rev	tggtggctgtctgtgctcgaa	<b>M97<sup>2)</sup></b>	EU483132	no	18 bp	23	CAPS ( <i>ApeKI</i> )	357	246, 71, 29, 11	M97for	atgaaagcacaaggagtccctc	S_1348-12-C09	At5g47560	malate/fumarate transporter, tonoplast	EU483134	375	346, 29	M97rev	cgagaatgaagctgcctaaga	<b>M98<sup>2)</sup></b>	EU483136	no	11 bp	24	CAPS ( <i>PfI</i> )	459	231, 184, 44	M98for	aagccgagatcctcctgcaatgg	C_1202-10-G11	At2g06520	PsbX (photosystem II subunit X), chloroplast	EU483138	470	426, 44	M98rev	aggcaaaataaacggggatacagc																												
<b>M95</b>	EU447253	yes	no	2	CAPS ( <i>HhaI</i> )	315	299, 16	M95for	tcggactcagcaatggcgctc	C_5102-112-A10	At5g54190	NADPH:protochlorophyllide oxidoreductase A, chloroplast																																																																																		
	EU447254					315	252, 47, 16	M95rev	tggtggctgtctgtgctcgaa				<b>M97<sup>2)</sup></b>	EU483132	no	18 bp	23	CAPS ( <i>ApeKI</i> )	357	246, 71, 29, 11	M97for	atgaaagcacaaggagtccctc	S_1348-12-C09	At5g47560	malate/fumarate transporter, tonoplast	EU483134	375	346, 29	M97rev	cgagaatgaagctgcctaaga	<b>M98<sup>2)</sup></b>	EU483136	no	11 bp	24	CAPS ( <i>PfI</i> )	459	231, 184, 44	M98for	aagccgagatcctcctgcaatgg	C_1202-10-G11	At2g06520	PsbX (photosystem II subunit X), chloroplast	EU483138	470	426, 44	M98rev	aggcaaaataaacggggatacagc																																														
<b>M97<sup>2)</sup></b>	EU483132	no	18 bp	23	CAPS ( <i>ApeKI</i> )	357	246, 71, 29, 11	M97for	atgaaagcacaaggagtccctc	S_1348-12-C09	At5g47560	malate/fumarate transporter, tonoplast																																																																																		
	EU483134					375	346, 29	M97rev	cgagaatgaagctgcctaaga				<b>M98<sup>2)</sup></b>	EU483136	no	11 bp	24	CAPS ( <i>PfI</i> )	459	231, 184, 44	M98for	aagccgagatcctcctgcaatgg	C_1202-10-G11	At2g06520	PsbX (photosystem II subunit X), chloroplast	EU483138	470	426, 44	M98rev	aggcaaaataaacggggatacagc																																																																
<b>M98<sup>2)</sup></b>	EU483136	no	11 bp	24	CAPS ( <i>PfI</i> )	459	231, 184, 44	M98for	aagccgagatcctcctgcaatgg	C_1202-10-G11	At2g06520	PsbX (photosystem II subunit X), chloroplast																																																																																		
	EU483138					470	426, 44	M98rev	aggcaaaataaacggggatacagc																																																																																					

<sup>1)</sup> Not each PCR product was fully sequenced; length in bp is derived from the position of the primer in the EST sequence used as template.

<sup>2)</sup> Polymorphisms first published in Mráček (2005) were confirmed, converted to CAPS, annotated and submitted to GenBank in this study.

### 3.1.1.2. Genotyping of Renner complexes.

The SSLP marker M40 was considered to be a suitable marker for genotyping a large variety of Renner complexes. It was derived from the EST cluster C\_1231-11-B04 of *Oe. elata* subsp. *hookeri* strain *hookeri* de Vries. The probe codes for a chloroplast-located sedoheptulose-bisphosphatase, which contains two introns. The two marker alleles known so far are intron spanning and highly polymorphic between the A genome <sup>h</sup>johansen and the B genome <sup>h</sup>tuscaloosa (Table 7) indicating that the region was suitable to detect further polymorphisms in different Renner complexes.



**Figure 11.** Assignment of M40 alleles to the Renner complexes <sup>St</sup>laxans and <sup>St</sup>undans in the strain bauri Standard: in tuscaloosa a single band of 500 bp was amplified (lane 1); in bauri Standard two bands of 474 bp and 579 bp were detected (lane 2); confirmation of the two bands, 500 bp and 579 bp in the F1 hybrid <sup>h</sup>tuscaloosa·<sup>St</sup>undans assigned the band of 579 bp to <sup>St</sup>undans (lane 3).

PCR products of M40 were amplified in 20 homozygous and five permanent translocation heterozygous strains. Homozygous strains harbour two identical haploid Renner complexes, so-called haplo-complexes or alethal complexes. For example, strain tuscaloosa of *Oenothera grandiflora* carries twice the haplo-complex <sup>h</sup>tuscaloosa. Consequently, the primer pair M40for and M40rev amplifies only a single band of 500 bp (Figure 11, lane 1, and Table 8). The strain bauri Standard, a permanent translocation heterozygous strain, has two different Renner complexes (<sup>St</sup>laxans and <sup>St</sup>undans). Since M40 alleles differ, the primer pair amplified two bands of 474 and 579 bp (Figure 11, lane 2, and Table 8). The different bauri Standard bands could be assigned to either <sup>St</sup>laxans or <sup>St</sup>undans by crosses with <sup>h</sup>tuscaloosa (Figure 11). Tuscaloosa taken as the seed parent generates the F1 hybrid <sup>h</sup>tuscaloosa·<sup>St</sup>undans. The combination <sup>h</sup>tuscaloosa·<sup>St</sup>laxans is not produced in this crossing direction, due to gametophytic lethal factors. The <sup>St</sup>laxans complex is inherited strictly maternally. Taken bauri Standard as pollen donor the only realizable hybrid is <sup>h</sup>tuscaloosa·<sup>St</sup>undans; the <sup>St</sup>undans complex inherits strictly paternally. Investigating <sup>h</sup>tuscaloosa·<sup>St</sup>undans for the marker M40

**Table 8.** SSLP and selected restriction endonuclease patterns of the M40 microsatellite region in various Renner complexes<sup>1)</sup>

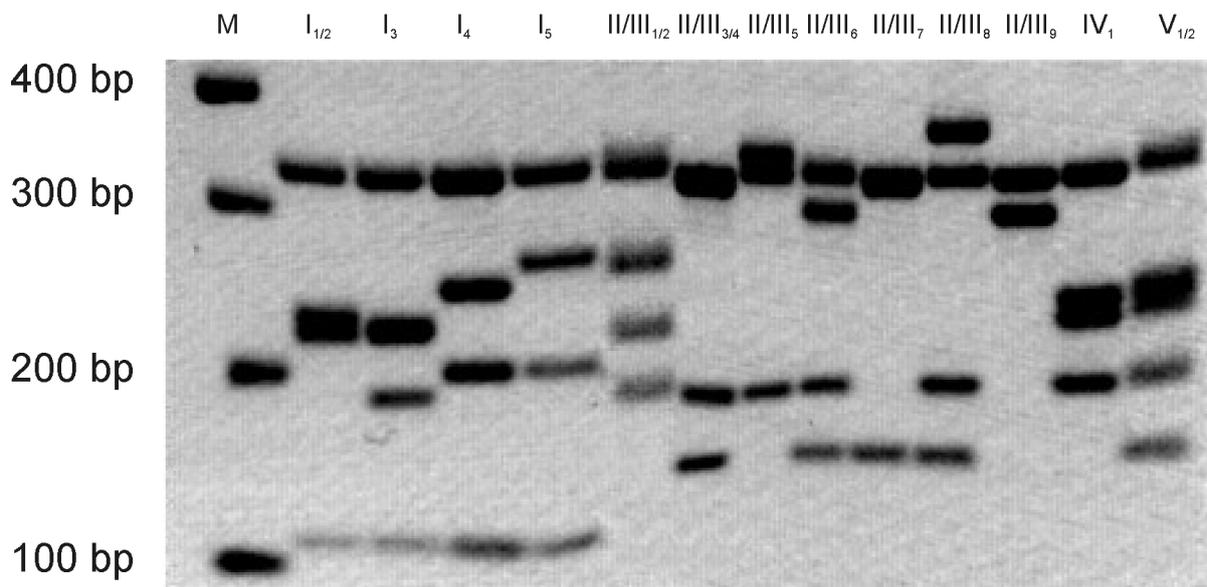
M40 allele	Renner complex	SSLP [bp]	CAPS [bp] ( <i>MboII</i> )	CAPS [bp] ( <i>MspI</i> )	CAPS [bp] ( <i>SpeI</i> )	accession number
A <sub>1</sub>	<sup>G</sup> albicans, <sup>St</sup> albicans, <sup>h</sup> blandina, <sup>h</sup> franciscana de Vries, <sup>h</sup> franciscana E. & S., <sup>h</sup> hookeri de Vries, <sup>St</sup> laxans, <sup>h</sup> purpurata, <sup>St</sup> rigens	474	210, 170, 67, 27	319, 155	249, 126, 99	EU432376, EU432377, EU432378, EU432379, EU432380, EU432381, EU432382, EU432383, EU432384
A <sub>2</sub>	<sup>h</sup> chapultepec, <sup>h</sup> cholula, <sup>h</sup> puebla, <sup>h</sup> toluca, <sup>St</sup> undans	579	285, 200, 67, 27	431, 148	352, 135, 92	EU432385, EU432386, EU432387, EU432388, EU432389
A <sub>3</sub>	<sup>h</sup> johansen, r- <sup>S</sup> velans	583	349, 207, 27	428, 155	352, 132, 99	EU432390, EU432391
B <sub>1</sub>	<sup>h</sup> decipiens, <sup>h</sup> deserens, r- <sup>S</sup> gaudens	470	207, 169, 67, 27	315, 155	371, 99	EU432392, EU432393, EU432394
B <sub>2</sub>	<sup>h</sup> bellamy A, <sup>h</sup> B <sup>A</sup> -castleberry A-4, <sup>h</sup> B <sup>A</sup> -chastang 7, <sup>G</sup> flavens, <sup>h</sup> stockton 1	499	226, 179, 67, 27	325, 174	364, 135	EU432395, EU432396, EU432397, EU432398, EU432399
B <sub>3</sub>	<sup>St</sup> flavens	500	227, 179, 67, 27	325, 175	365, 135	EU432400
B <sub>4</sub>	<sup>h</sup> tuscaloosa	500	227, 179, 67, 27	325, 175	365, 135	EU432401
C <sub>1</sub>	<sup>h</sup> douthat 1, <sup>h</sup> williamsville, <sup>h</sup> wilson creek 1	496	243, 159, 67, 27	496	364, 132	EU432402, EU432403, EU432404

<sup>1)</sup> corresponding *Oenothera* strains and species are listed in Table 6, p. 29.

displays two PCR products of 500 bp and 579 bp (Figure 11). Since the band of 500 bp is amplified from the <sup>h</sup>tuscaloosa complex, the band of 579 bp must obviously be linked to <sup>St</sup>undans. Therefore, the 474 bp PCR product detected in the strain bauri Standard can serve as a marker for <sup>St</sup>laxans (Figure 11). Comparable approaches were performed for the permanent translocation heterozygous strains ammophila Standard (<sup>St</sup>rigens·<sup>St</sup>percurvans), rr-lamarckiana Sweden (r-<sup>S</sup>velans·r-<sup>S</sup>gaudens), suaveolens Grado (<sup>G</sup>albicans·<sup>G</sup>flavens), and suaveolens Standard (<sup>St</sup>albicans·<sup>St</sup>flavens). Hybrids were confirmed by phenotypic markers (Dietrich *et al.*, 1997).

The marker regions were analyzed by sequencing. The sequences are deposited in GenBank (Table 8). In all 29 Renner complexes investigated the marker allele M40 displays two introns. Basically, the introns contain microsatellites of different lengths resulting in eight SSLPs among the Renner complexes under study. For all three basic nuclear genome types, A, B, and C, at least one specific SSLP was found. PCR products converted into CAPS markers allow discrimination of single genotypes on 2 to 3% agarose gels (Table 8). Work on the marker M40 was done in co-operation with Uwe Rauwolf.

### 3.1.1.3. Markers for basic plastomes and subplastomes



**Figure 12.** *Bam*HI restriction pattern of the *Oenothera rrn16-trnI<sub>GAU</sub>* spacer region. So far, 17 alleles could be identified, of which 13 can be distinguished by a *Bam*HI digest on 2%-agarose gels (lanes 2-14; headings represent allele names presented in Table 9). The first lane (M) shows a 100 bp-ladder (New England BioLabs, Ipswich, MA).

An interspecific plastome exchange also requires plastome specific markers. For this purpose, the plastid *rrn16-trnI<sub>GAU</sub>* spacer region was investigated from 41 *Oenothera* strains. A high degree of polymorphism in this region was already reported for a limited number of strains (Hornung *et al.*, 1996; Sears *et al.*, 1996). Amplification with the primer pair 16S SEQ (+) and *trnI* PCR (+) and sequence analysis of PCR products derived, uncovered a *Bam*HI restriction polymorphism in the *rrn16-trnI<sub>GAU</sub>* spacer. In the subsection *Oenothera*, 17 different *Bam*HI restriction patterns could be detected, of which 13 can be easily distinguished on 2% agarose gels (Figure 12 and Table 9). The marker alleles *rrn16-trnI<sub>GAU</sub>* I<sub>1</sub> and *rrn16-trnI<sub>GAU</sub>* I<sub>2</sub> present in the strains chapultepec, cholula, puebla or toluca, cannot be discerned *via* a *Bam*HI, but by a *Bsm*BI digestion (I<sub>1</sub>: 619 bp, 261 bp and I<sub>2</sub>: 870 bp).

The patterns reflect the phylogenetic relationship between the chosen strains and the basic plastome types. In both species with plastome IV, *Oe. oakesiana* and *Oe. parviflora*, only a single allele (*rrn16-trnI<sub>GAU</sub>* IV<sub>1</sub>) could be detected in altogether four strains. Also alleles *rrn16-trnI<sub>GAU</sub>* V<sub>1</sub> and V<sub>2</sub> of three strains of *Oe. argillicola* are identical, except for a single base pair polymorphism. For plastome I the closely related strains of *Oe. elata* subsp. *elata* (chapultepec, cholula, puebla and toluca) have two very similar alleles, *rrn16-trnI<sub>GAU</sub>* I<sub>1</sub> and I<sub>2</sub>, which differ only in a single deletion. In *Oe. elata* subsp. *hookeri* two alleles were confirmed in four different strains. The investigated strain bauri Standard of *Oe. villosa* subsp. *villosa* also displays a distinct allele. Taken together, the alleles described in the basic plastomes I, IV and V reflect a marker for the genetic behaviour of plastomes they originated from, since they are only found in one basic plastome type each. Additionally, they confirm evolutionary tendencies of subplastome evolution. Within the AA-I clade, the *Oe. elata* subsp. *elata*, *Oe. elata* subsp. *hookeri* and *Oe. villosa* subsp. *villosa* represent different evolutionary lineages (Dietrich *et al.*, 1997), a fact also supported by the *rrn16-trnI<sub>GAU</sub>* marker allele.

The situation is very different for alleles of subplastomes of plastome II or plastome III, naturally found in *Oe. biennis*, *Oe. glazioviana*, *Oe. grandiflora* and *Oe. nutans*. Apparently, no specific allele related to the genetic behaviour of those plastome types exists (Table 9). Allele *rrn16-trnI<sub>GAU</sub>* II/III<sub>3</sub> is found in plastomes II or III of *Oe. biennis* (AB-II or BA-III), *Oe. grandiflora* (BB-III) and *Oe. nutans* (BB-III). Allele *rrn16-trnI<sub>GAU</sub>* II/III<sub>1</sub> is not characteristic for plastome type II since it is also found in the strain castleberry B-8 of *Oe. grandiflora* (BB-III). These data indicated gene flow between the four species (Chapter

**Table 9.** *Bam*HI restriction and SSLP pattern of the *rrn16-trnI<sub>GAU</sub>* spacer region in *Oenothera* plastomes and subplastomes used in this study<sup>1)</sup>

<i>rrn16-trnI<sub>GAU</sub></i> allele	strain	species	plastome type	SSLP [bp]	CAPS [bp] ( <i>Bam</i> HI)	accession number
I <sub>1</sub>	chapultepec	<i>Oe. elata</i> subsp. <i>elata</i>	I	880	322, 229, 220, 109	EU262892
I <sub>2</sub>	cholula, puebla, toluca	<i>Oe. elata</i> subsp. <i>elata</i>	I	870	322, 220, 218, 110	EU262893, EU282392, EU282393
I <sub>3</sub>	franciscana de Vries, franciscana E.& S., johansen	<i>Oe. elata</i> subsp. <i>hookeri</i>	I	1058	322, 224, 220, 182, 110	EU282394, EU282395, EU262894,
I <sub>4</sub>	hookeri de Vries	<i>Oe. elata</i> subsp. <i>hookeri</i>	I	876	322, 247, 197, 110	EU262895
I <sub>5</sub>	bauri Standard	<i>Oe. villosa</i> subsp. <i>villosa</i>	I	891	322, 262, 197, 110	EU262896
II/III <sub>1</sub>	biennis München, castleberry B-8, conferta Standard, purpurata, suaveolens Fünfkirchen, suaveolens Grado, suaveolens Standard	<i>Oe. biennis</i> , <i>Oe. grandiflora</i> , <i>Oe. biennis</i> x <i>Oe. glazioviana</i>	II or III	977	322, 257, 216, 182	EU282396, EU282397, EU282398, EU282399, EU282400, EU262897, EU282401
II/III <sub>2</sub>	coronifera Standard, nuda Standard	<i>Oe. biennis</i> , <i>Oe. glazioviana</i>	II	981	322, 257, 220, 182	EU282402, EU262898
II/III <sub>3</sub>	bellamy A, biennis de Vries, chastang 7, chicaginensis Colmar, horsesheads 2, marienville 3, stockton 1	<i>Oe. biennis</i> , <i>Oe. grandiflora</i> , <i>Oe. nutans</i>	II or III	963	322, 318, 182, 141	EU282404, EU262899, EU282405, EU282403, EU282406, EU282407, EU282408
II/III <sub>4</sub>	mitchell	<i>Oe. nutans</i>	III	980	335, 322, 182, 141	EU262900
II/III <sub>5</sub>	castleberry A-4	<i>Oe. grandiflora</i>	III	845	341, 322, 182	EU262901
II/III <sub>6</sub>	elkins 2	<i>Oe. nutans</i>	III	940	322, 295, 182, 141	EU262902
II/III <sub>7</sub>	lawrenceville 3	<i>Oe. biennis</i>	II	781	322, 318, 141	EU262903
II/III <sub>8</sub>	tuscaloosa	<i>Oe. grandiflora</i>	III	1009	364, 322, 182, 141	EU262904
II/III <sub>9</sub>	blandina, decipiens, deserens, rr-lamarckiana Sweden	<i>Oe. glazioviana</i>	III	617	322, 295	EU282409, EU282410, EU282411, EU262905
IV <sub>1</sub>	ammophila Standard, atrovirens Standard, silesiaca Standard, st. stephen	<i>Oe. oakesiana</i> , <i>Oe. parviflora</i>	IV	961	322, 236, 221, 182	EU282415, EU262906, EU282412, EU282413
V <sub>1</sub>	douthat 1	<i>Oe. argillicola</i>	V	1101	322, 236, 221, 182, 140	EU262907
V <sub>2</sub>	williamsville, wilson creek 1	<i>Oe. argillicola</i>	V	1102	322, 236, 221, 182, 141	EU282414, EU262908

<sup>1)</sup> Table 6, p. 29, lists references describing species, strain, and plastome type for each accession.

4.4.4). Variation within plastomes II and III appears to be more diverged as in other plastome types. Alleles *rrn16-trnI*<sub>GAU II/III<sub>4</sub> - II/III<sub>8</sub></sub> are specific to single strains, but alleles *rrn16-trnI*<sub>GAU II/III<sub>1</sub></sub> and II/III<sub>3</sub> may reflect the two major patterns for plastomes II and III (Table 9).

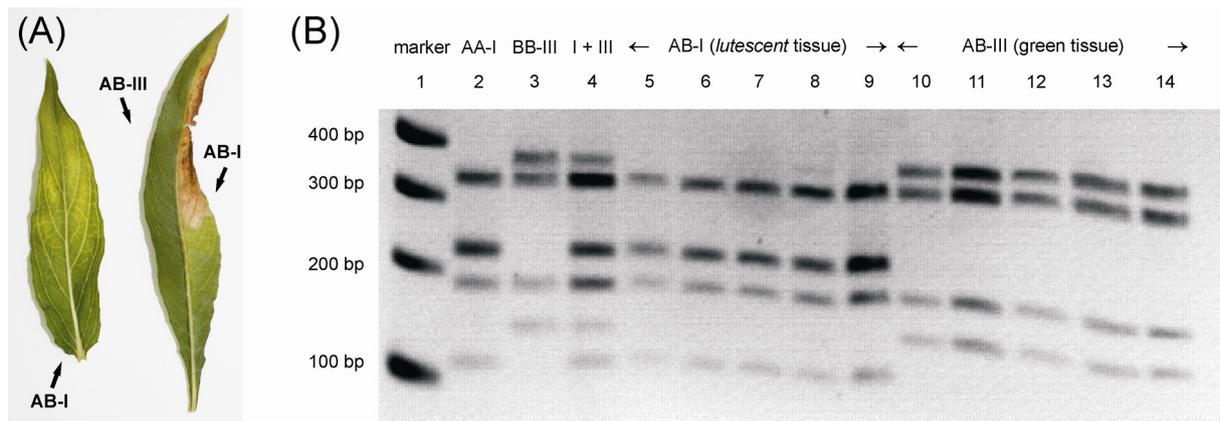
### 3.1.2. *New combination of genetic compartments*

To demonstrate the power of the above described marker system, two different incompatible combinations (AB-I and BB-II) as well as genetically compatible control plants (AB-III) were established in crossing experiments.

#### 3.1.2.1. *Generation of interspecific AB-I and AB-III plastome-genome hybrids*

With appropriate parental lines, the production of dominant interspecific plastome-genome incompatible hybrids, such as the combination AB-I, can be rather easy. They already arise in the F1 generation (Chapter 3.3.3). The homozygous strains *Oe. elata* subsp. *hookeri* strain johansen (AA-I; <sup>h</sup>johansen·<sup>h</sup>johansen I<sup>joh</sup>) and *Oe. grandiflora* strain tuscaloosa (BB-III; <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa III<sup>tusca</sup>) were used to generate the dominant incompatible AB-I and a compatible AB-III hybrid. Since plastids are transmitted biparentally in *Oenothera* and segregate somatically in the F1 generation, AB-I and AB-III hybrids, <sup>h</sup>johansen·<sup>h</sup>tuscaloosa I<sup>joh</sup> and <sup>h</sup>johansen·<sup>h</sup>tuscaloosa III<sup>tusca</sup>, respectively, can be generated directly from a single cross of both parental species. Compatible and incompatible tissue usually segregates on the same individual. The segregation of plastomes was checked by the PCR polymorphism described above (Figure 13).

An AB hybrid with a second subplastome (III<sup>lam</sup>) was selected as a control to confirm plastome III specificity of the genetic pattern. The hybrid containing plastome III of *Oe. glazioviana* strain rr-lamackiana Sweden was generated from a cross between *Oe. elata* subsp. *hookeri* strain johansen equipped with plastome III<sup>lam</sup> (AA-III<sup>lam</sup>; <sup>h</sup>johansen·<sup>h</sup>johansen III<sup>lam</sup>; Stubbe, unpublished) and *Oe. grandiflora* strain tuscaloosa. No notable difference between the two green AB-III F1 hybrid lineages (AB-III<sup>lam</sup> and AB-III<sup>tusca</sup>) could be detected phenotypically or in later molecular investigation (Chapter 3.4.2.3).



**Figure 13.** *Lutescent* phenotype and somatic plastome segregation in F1 of a cross between *Oe. elata* subsp. *hookeri* strain johansen (AA-I) and *Oe. grandiflora* strain tuscaloosa (BB-III). Note that plants can contain only a single plastome type, but also plastome chimerical tissue (panel A). Incompatible and green tissue was correlated with the plastome type via the *Bam*HI CAPS marker (Table 9). Marker alleles *rrn16-trnI*<sub>G<sub>AU</sub></sub> I<sub>3</sub> and II/III<sub>8</sub> were amplified from plastome I<sup>joh</sup> and III<sup>tusca</sup> and digested with *Bam*HI (lanes 2 and 3). Lane 4 shows a mixture of cleaved I<sup>joh</sup> and III<sup>tusca</sup> PCR products. Applying the marker to tissues of individuals of a F1 offspring correlates plastome I<sup>joh</sup> exclusively with the *lutescent* phenotype (panel B, lanes 5 – 9) and plastome III<sup>tusca</sup> with green tissue (panel B, lanes 10 – 14).

### 3.1.2.2. Generation of the interspecific incompatible BB-II hybrid

A different strategy was chosen to produce the incompatible combination BB-II. In this case not a simple F1 hybrid is produced, but a foreign plastid has to be incorporated into a homozygous nuclear background. During the crossing program the gene content of the B genome must not change. The possibility of such a plastid exchange is unique to *Oenothera* genetics and requires a complex crossing program as it will be explained subsequently.

The first crossing mate, the strain suaveolens Grado of *Oenothera biennis* (AB-II), is a complete permanent translocation heterozygous strain. It forms, due to its segmental arrangement, a ring of 14 chromosomes in diakinesis (Table 10, first row). Consequently, this strain produces two different egg cells. The  $\alpha$ -complex (<sup>G</sup>albicans) is exclusively inherited maternally. The  $\beta$ -complex (<sup>G</sup>flavens) is predominantly inherited paternally and found facultatively in egg cells. The second crossing partner, *Oe. grandiflora* strain tuscaloosa (BB-III), displays a different breeding behaviour. Tuscaloosa is a bivalent former with 7 bivalents in meiosis, and has twice the complex <sup>h</sup>tuscaloosa (Table 10, second row). The complex <sup>h</sup>tuscaloosa is inherited biparentally and is free of gametophytic or sporophytic lethal factors.

**Table 10.** Segmental arrangements and chromosome configurations of *Oe. biennis* strain suaveolens Grado (<sup>G</sup>albicans·<sup>G</sup>flavens), *Oe. grandiflora* strain tuscaloosa (<sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa) and of their F1 twin hybrids <sup>G</sup>albicans·<sup>h</sup>tuscaloosa and <sup>G</sup>flavens·<sup>h</sup>tuscaloosa

strain or hybrid	Renner complex	basic genome	segmental arrangement of chromosome arms										chromosome configuration
suaveolens Grado	<sup>G</sup> albicans	A	/1·12	11·10	7·5	6·3	2·14	13·8	9·4	\		⊙14	
	<sup>G</sup> flavens	B	\	12·11	10·7	5·6	3·2	14·13	8·9	4·1/			
tuscaloosa	<sup>h</sup> tuscaloosa	B	/1·2\	/3·4\	/5·6\	/7·10\	/9·8\	/11·12\	/13·14\		7 prs.		
	<sup>h</sup> tuscaloosa	B	\1·2/	\3·4/	\5·6/	\7·10/	\9·8/	\11·12/	\13·14/				
<sup>G</sup> flavens· <sup>h</sup> tuscaloosa	<sup>G</sup> flavens	B	/1·4	3·2	\	/5·6\	/7·10\	/9·8\	/11·12\	/13·14\	⊙4, 5 prs.		
	<sup>h</sup> tuscaloosa	B	\	4·3	2·1/	\5·6/	\7·10/	\9·8/	\11·12/	\13·14/			
<sup>G</sup> albicans· <sup>h</sup> tuscaloosa	<sup>G</sup> albicans	A	/1·12	11·10	7·5	6·3	4·9	8·13	14·2	\	⊙14		
	<sup>h</sup> tuscaloosa	B	\	12·11	10·7	5·6	3·4	9·8	13·14	2·1/			

Figure 14 presents the crossing scheme how to exchange plastomes between these two species. In a cross of *Oe. biennis* strain suaveolens Grado as mother plant and *Oe. grandiflora* strain tuscaloosa as pollen donor, the F1 generation obtained is not uniform. Two different egg cells, with the genetic constitution <sup>G</sup>albicans or <sup>G</sup>flavens, are produced by suaveolens Grado. Each of them can be combined with <sup>h</sup>tuscaloosa, giving rise to a non-Mendelian splitting F1 generation. The two segregation populations are called twin hybrids, namely with the genotypes <sup>G</sup>albicans·<sup>h</sup>tuscaloosa (AB) and <sup>G</sup>flavens·<sup>h</sup>tuscaloosa (BB). Due to biparental plastid transmission in *Oenothera*, both plastome types are inherited (plastome II<sup>suavG</sup> from suaveolens Grado and plastome III<sup>tusca</sup> from tuscaloosa). Therefore, all F1 offspring is chimerical for its plastid genotype (II<sup>suavG</sup>/III<sup>tusca</sup>). For further breeding the twin hybrid <sup>G</sup>flavens·<sup>h</sup>tuscaloosa is discarded. The diakinesis configuration of this hybrid is a small ring of 4 and 5 bivalents (⊙4, 5 prs.) (Table 10, third row). This chromosome configuration allows six linkage groups; the hybrid is not constant in successive generations. The second twin hybrid, <sup>G</sup>albicans·<sup>h</sup>tuscaloosa, is different. The chromosome formulas of <sup>G</sup>albicans and <sup>h</sup>tuscaloosa allow a ring of 14 chromosomes (⊙14) (Table 10, fourth row). Therefore, the complexes <sup>G</sup>albicans and <sup>h</sup>tuscaloosa are not mixed, due to repression of recombination and free segregation by the meiotic ring (Chapter 1.7.2). If <sup>G</sup>albicans·<sup>h</sup>tuscaloosa is now selfed, the progeny splits into two populations in F2 generations with the genetic constitution <sup>G</sup>albicans·<sup>h</sup>tuscaloosa and <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa, a consequence on the maternal inheritance of <sup>G</sup>albicans and the biparental inheritance of the freely segregating alethal complex <sup>h</sup>tuscaloosa. The interspecific exchange of plastids again takes place by somatic segregation

and sorting-out of the two plastomes II<sup>suavG</sup> and III<sup>tusca</sup> in the F1 hybrid <sup>G</sup>albicans·<sup>h</sup>tuscaloosa II<sup>suavG</sup>/III<sup>tusca</sup>. Flower buds carrying exclusively plastome II<sup>suavG</sup>, when selfed, lead to the desired offspring <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa II<sup>suavG</sup> (BB-II) (Figure 15, panel C).

#### parents

*Oe. biennis* strain suaveolens Grado (<sup>G</sup>albicans·<sup>G</sup>flavens II<sup>suavG</sup>) [AB-II; α-complex ♀, β-complex ♀♂; ⊙14]  
*Oe. grandiflora* strain tuscaloosa (<sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa III<sup>tusca</sup>) [BB-III; haplo-complex ♀♂; 7 prs.]

#### crossing scheme

P: <sup>G</sup>albicans·<sup>G</sup>flavens II<sup>suavG</sup> [⊙14] x <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa III<sup>tusca</sup> [7 prs.]  
 F1: <sup>G</sup>albicans·<sup>h</sup>tuscaloosa II<sup>suavG</sup>/III<sup>tusca</sup> [⊙14]      <sup>G</sup>flavens·<sup>h</sup>tuscaloosa II<sup>suavG</sup>/III<sup>tusca</sup> [⊙4, 5 prs.]  
     ↓ selection of flower buds just carrying platome II      ↓ unstable hybrid discarded  
     <sup>G</sup>albicans·<sup>h</sup>tuscaloosa II<sup>suavG</sup> [⊙14] x self      <sup>G</sup>flavens·<sup>h</sup>tuscaloosa II<sup>suavG</sup>/III<sup>tusca</sup> [⊙4, 5 prs.]  
 F2: <sup>G</sup>albicans·<sup>h</sup>tuscaloosa II<sup>suavG</sup>      <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa II<sup>suavG</sup> [BB-II]

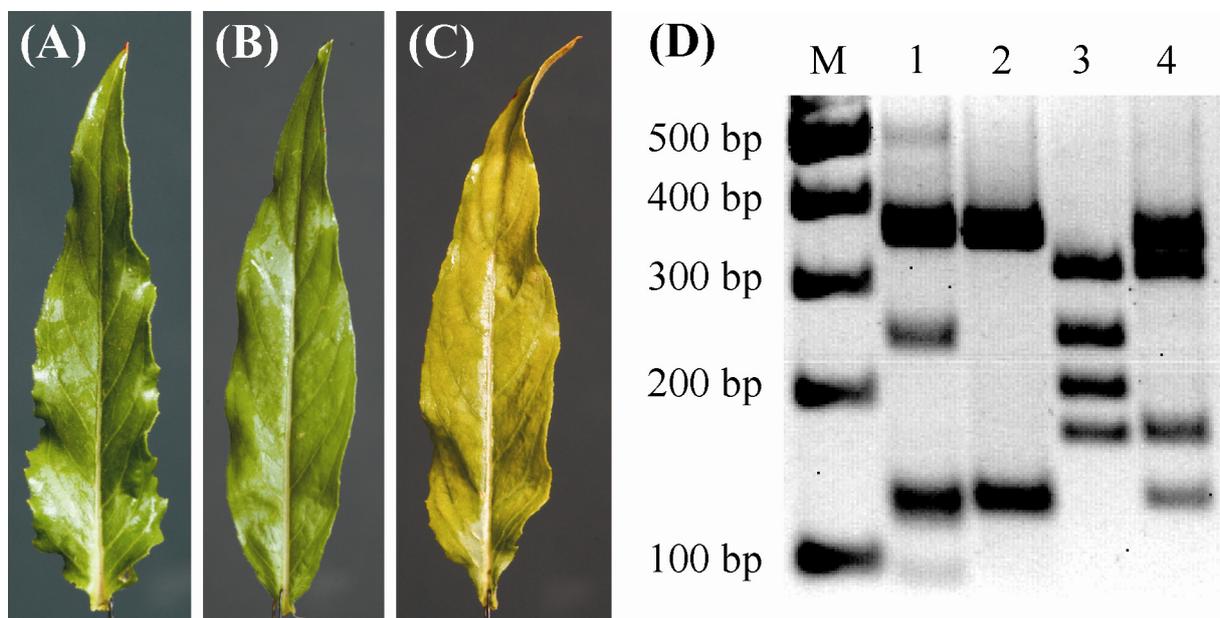
**Figure 14.** Crossing scheme to exchange pastome III of *Oe. grandiflora* strain tuscaloosa with plastome II of *Oe. biennis* strain suaveolens Grado. A detailed description is given in the text.

During the crossing program, two basic genome combinations (AB and BB) and two basic plastome types (II and III) have to be distinguished. At least the nuclear genotypes can be separated phenotypically. Panel A of Figure 15 shows the leaf shape of the hybrid <sup>G</sup>albicans·<sup>h</sup>tuscaloosa (AB). It is clearly distinguishable from the shape of a BB leaf of strain tuscaloosa (<sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa) or hybrid <sup>h</sup>tuscaloosa·<sup>G</sup>flavens (Figure 15, panel B). The phenotypic marker could be confirmed at the molecular level. Line 1 of panel D in Figure 15 shows <sup>G</sup>albicans·<sup>h</sup>tuscaloosa being heterozygous for the M40 alleles A<sub>1</sub> (<sup>G</sup>albicans) and B<sub>4</sub> (<sup>h</sup>tuscaloosa) as confirmed by a *SpeI* digestion. Line 2 shows just a single *SpeI* restriction pattern derived from homozygous M40-B<sub>4</sub> alleles (<sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa; BB). The BB combination <sup>G</sup>flavens·<sup>h</sup>tuscaloosa gives the same pattern. M40 alleles are listed in Table 8. Taking phenotypic and especially molecular markers into consideration, the progeny in the splitting generations <sup>G</sup>albicans·<sup>h</sup>tuscaloosa and <sup>G</sup>flavens·<sup>h</sup>tuscaloosa in F1, as well as <sup>G</sup>albicans·<sup>h</sup>tuscaloosa and <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa in F2 can be reliably monitored.

The crucial point during the crossing program is the selfing of the F1 hybrid <sup>G</sup>albicans·<sup>h</sup>tuscaloosa (Figure 14). Selection has to happen on flower buds carrying exclusively plastome II<sup>suavG</sup>. A phenotypic discrimination is not possible, since both plastids

carrying plastomes II or III are green in the AB nuclear background (Figure 4). With the plastidic CAPS marker plastome identity of a flower bud can be reliably and easily checked: Figure 15, panel D, line 3 (*Bam*HI digest of the *rrn16-trnI*<sub>GAU</sub> allele II/III<sub>1</sub> for plastome II<sup>suavG</sup>) and panel D, line 4 (*Bam*HI digest of the *rrn16-trnI*<sub>GAU</sub> allele II/III<sub>8</sub> for plastome III<sup>tusca</sup>).

The marker system is also suitable to confirm the genetic identity of the incompatible combination BB-II, which differs phenotypically from its compatible counterpart BB-III by a yellow-green leaf phenotype (*lutescent*) (Figure 15, panels B and C).



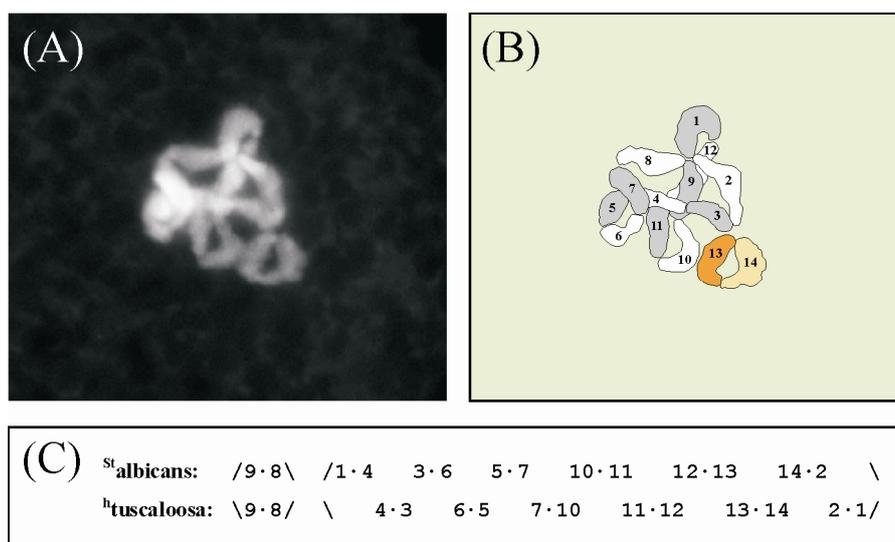
**Figure 15.** Phenotypic and molecular discrimination of crossing intermediates and end products in the BB-II crossing program: <sup>G</sup>albicans<sup>h</sup>tuscaloosa with plastome II<sup>suavG</sup> or III<sup>tusca</sup> (AB-II/III) (panel A); compatible, native combination <sup>h</sup>tuscaloosa<sup>h</sup>tuscaloosa or the hybrid <sup>G</sup>flavens<sup>h</sup>tuscaloosa with plastome III<sup>tusca</sup> (BB-III) (panel B); incompatible combination <sup>h</sup>tuscaloosa<sup>h</sup>tuscaloosa II<sup>suavG</sup> (BB-II) (panel C); molecular discrimination of twin hybrids and plastomes (panel D); <sup>G</sup>albicans<sup>h</sup>tuscaloosa (AB), heterozygous for M40 alleles A<sub>1</sub> and B<sub>4</sub>, digested with *Spe*I (lane 1); <sup>h</sup>tuscaloosa<sup>h</sup>tuscaloosa (or <sup>G</sup>flavens<sup>h</sup>tuscaloosa) (BB) homozygous for M40 allele B<sub>4</sub>, digested with *Spe*I (lane 2); plastome II<sup>suavG</sup> *rrn16-trnI*<sub>GAU</sub> allele II/III<sub>1</sub>, digested with *Bam*HI (lane 3) and plastome III<sup>tusca</sup>, *rrn16-trnI*<sub>GAU</sub> allele II/III<sub>8</sub> digested with *Bam*HI (lane 4); first lane (M) shows a 100 bp-ladder (New England BioLabs, Ipswich, Massachusetts). Detailed explanations are given in the text.

### 3.1.3. Correlation of classical and molecular *Oenothera* maps

Many of the co-dominant markers introduced under Chapter 3.1.1.1 were assigned to seven coupling groups of the AFLP map representing the seven *Oenothera* chromosomes (Greiner and Rauwolf, unpublished). However, molecular and classical *Oenothera* maps are not integrated. The classical maps are based on the segmental arrangement of chromosome arms and their relative location. These chromosomal arrangements or chromosomal formulas have been determined for more than 300 *Oenothera* strains (Cleland, 1972 and others; see also Introduction) and phenotypic characters have been assigned to individual chromosome arms. The extension of the classical map with molecular markers is clearly one of the most important future steps in *Oenothera* research and breeding. It would allow an immediate genetic access to a huge variety of strains.

#### 3.1.3.1. The hybrid *St*<sup>albicans</sup>·*h*<sup>tuscaloosa</sup> and its genetic behaviour

To address this question, attempts were made to combine methods of molecular and classical *Oenothera* genetics, using chromosome 9·8 as example. Chromosomes of the classical map can only be identified by their genetic and cytological behaviour. Therefore, an appropriate cross was chosen to distinguish chromosome 9·8 from the rest of the genome by segregation analysis.



**Figure 16.** Chromosome configuration ©12, 1 pr. of the hybrid *St*<sup>albicans</sup>·*h*<sup>tuscaloosa</sup>: determination in the diakinesis *via* DAPI staining (panel A) and graphical interpretation, the free bivalent is marked in red (panel B). The configuration can be predicted by the chromosome formulas of the Renner complexes involved (panel C). DAPI staining and graphical interpretation were done in co-operation with Hieronim Golczyk. Note that chromosomes but not chromosome arms are numbered in panel B.

The strains suaveolens Standard ( $^{St}albicans \cdot ^{St}flavens$ ) and tuscaloosa ( $^{h}tuscaloosa \cdot ^{h}tuscaloosa$ ) were crossed, resulting in the F1 hybrid  $^{St}albicans \cdot ^{h}tuscaloosa$  ( $\odot 12$ , 1 pr.). This hybrid resembles closely the hybrid  $^{G}albicans \cdot ^{h}tuscaloosa$  ( $\odot 14$ ) that was already introduced for plastid exchange (Figure 14). However, as a significant difference it has the chromosome configuration of a ring of 12 chromosomes and one free bivalent ( $\odot 12$ , 1 pr.). The free bivalent is formed by chromosome 9·8, illustrated in Figure 16.

The diakinesis configuration of  $\odot 12$ , 1 pr. expects a specific segregation pattern of the chromosomes involved. It allows only two linkage groups. The twelve chromosomes assemble in a single linkage group, which does not segregate in a Mendelian way. Like  $^{G}albicans$  in Figure 14,  $^{St}albicans$  is exclusively transmitted by the egg cell. A selfed F1 hybrid  $^{St}albicans \cdot ^{h}tuscaloosa$ , therefore, splits into  $^{St}albicans \cdot ^{h}tuscaloosa$  and  $^{h}tuscaloosa \cdot ^{h}tuscaloosa$  in a theoretical ratio of 1:1. The lethal combination of homozygous  $^{St}albicans$  complexes ( $^{St}albicans \cdot ^{St}albicans$ ) is not found. However, segregation of the Renner complexes  $^{St}albicans$  and  $^{h}tuscaloosa$  involves only the *six* chromosomes each, which are part of the ring. The seventh chromosome, in this instance chromosome 9·8, is free and does not co-segregate with the torso complexes  $^{St}albicans$  and  $^{h}tuscaloosa$ . Since chromosome 9·8 lacks lethal factors (Renner, 1942b; Stubbe, 1953)<sup>2)</sup>, it can segregate independently in a Mendelian manner ( $9 \cdot 8^{albSt} 9 \cdot 8^{albSt} / 9 \cdot 8^{albSt} 9 \cdot 8^{tusca} / 9 \cdot 8^{tusca} 9 \cdot 8^{tusca}$ ; ratio 1:2:1) (Figure 17).

The presence of two linkage groups can be monitored with molecular markers. The large linkage group, containing lethal factors and involving 12 chromosomes, can be easily distinguished from the small lethal factor free one, containing just a free bivalent. Since identification of the seven *Oenothera* linkage groups in the AFLP map, representing all seven *Oenothera* chromosomes, was carried out with the  $^{h}tuscaloosa$  complex (Mráček, 2005), the freely segregating bivalent of  $^{St}albicans \cdot ^{h}tuscaloosa$ , namely 9·8, must be identical with one of the seven coupling groups in the map of  $^{h}johansen \cdot ^{h}tuscaloosa$ .

### 3.1.3.2. Identification of marker alleles between $^{St}albicans$ and $^{h}tuscaloosa$

Eight co-dominant markers, representing all seven *Oenothera* chromosomes as individual linkage groups, were checked for heterozygosity in the hybrid  $^{St}albicans \cdot ^{h}tuscaloosa$  (AB).

<sup>2)</sup> The statement of Cleland (1972, p. 110) that the phenotypic marker *pil*, which is located on chromosome arm 8 of  $^{St}albicans$ , is homozygous lethal obviously arose from a misinterpretation of Renner's sometimes complex German (Renner, 1942b).

The strategy was to digest PCR products with appropriate restriction endonucleases shown to distinguish Renner complexes <sup>h</sup>johansen (A) and <sup>h</sup>tuscaloosa (B) (Table 7). All markers described resembled the restriction pattern of the A genome <sup>h</sup>johansen in the A genome <sup>St</sup>albicans (Table 11).

**parents:**

*Oe. biennis* strain suaveolens Standard (<sup>St</sup>albicans·<sup>St</sup>flavens) [α-complex ♀, β-complex ♂]

*Oe. grandiflora* strain tuscaloosa (<sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa) [haplo-complex ♀]

**crossing scheme:**

P: <sup>St</sup>albicans·<sup>St</sup>flavens x <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa

F1: <sup>St</sup>albicans·<sup>h</sup>tuscaloosa [⊙12, 1 pr.] x self <sup>St</sup>flavens·<sup>h</sup>tuscaloosa [⊙4, ⊙4, 3 pr.]

F2:  $\left\{ \begin{array}{l} \text{<sup>St</sup>albicans·<sup>h</sup>tuscaloosa (50\%)} \\ \text{<sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa (50\%)} \end{array} \right\} \left\{ \begin{array}{lll} \text{9·8<sup>albSt</sup>/9·8<sup>albSt</sup> (25\%)} & \text{9·8<sup>albSt</sup>/9·8<sup>tusca</sup> (50\%)} & \text{9·8<sup>tusca</sup>/9·8<sup>tusca</sup> (25\%)} \\ \text{9·8<sup>albSt</sup>/9·8<sup>albSt</sup> (25\%)} & \text{9·8<sup>albSt</sup>/9·8<sup>tusca</sup> (50\%)} & \text{9·8<sup>tusca</sup>/9·8<sup>tusca</sup> (25\%)} \end{array} \right\}$

linkage group 1

linkage group 2

**Figure 17.** Assembly and segregation behaviour of the hybrid <sup>St</sup>albicans·<sup>h</sup>tuscaloosa. The F2 population was used to identify chromosome 9·8 (linkage group 2), which segregates independently from the large strongly linked coupling groups of the torso complexes <sup>St</sup>albicans or <sup>h</sup>tuscaloosa (linkage group 1). Detailed explanations are given in the text.

### 3.1.3.3. Assignment of coupling group 7 to chromosome 9·8

Analysis of 38 <sup>St</sup>albicans·<sup>h</sup>tuscaloosa F2 plants uncovered an astonishing linkage of markers M19, M40, M41, M46, M47, M50 and M74 with the <sup>St</sup>albicans·<sup>h</sup>tuscaloosa and <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa phenotype, respectively (LOD = 11,13). In the large coupling group only <sup>St</sup>albicans·<sup>h</sup>tuscaloosa or <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa genotypes could be detected in the expected proportion of about 50% (Figure 17). A second, freely segregating coupling group detected with marker M58 was confirmed with a LOD > 10. All allelic combinations (M58-A<sub>1</sub>/M58-A<sub>1</sub>, M58-A<sub>1</sub>/M58-B<sub>1</sub> and M58-B<sub>1</sub>/M58-B<sub>1</sub>) were found in this linkage group. These data assign coupling group 7, of which M58 is part of, to chromosome 9·8 (Table 11). Furthermore, the strong linkage of M19, M40, M41, M46, M47, M50 and M74, representing six chromosomes in a ring of 12, provides a first direct molecular evidence for the suppression of homologous recombination and free segregation through meiotic rings in *Oenothera* (Cleland, 1972; Harte, 1994).

**Table 11.** Assignment of eight co-dominant markers to the seven coupling groups of <sup>h</sup>johansen·<sup>h</sup>tuscaloosa (Greiner and Rauwolf, unpublished) and their alleles in <sup>h</sup>tuscaloosa·<sup>h</sup>tuscaloosa and <sup>St</sup>albicans·<sup>h</sup>tuscaloosa<sup>1)</sup>

marker	coupling group	alleles found in	
		<sup>h</sup> tuscaloosa· <sup>h</sup> tuscaloosa	<sup>St</sup> albicans· <sup>h</sup> tuscaloosa
M19	CG1	M19-B <sub>1</sub> /M19-B <sub>1</sub>	M19-A <sub>1</sub> /M19-B <sub>1</sub>
M40	CG2	M40-B <sub>4</sub> /M40-B <sub>4</sub>	M40-A <sub>1</sub> /M40-B <sub>4</sub>
M74	CG3	M74-B <sub>1</sub> /M74-B <sub>1</sub>	M74-A <sub>1</sub> /M74-B <sub>1</sub>
M47	CG4	M47-B <sub>1</sub> /M47-B <sub>1</sub>	M47-A <sub>1</sub> /M47-B <sub>1</sub>
M46	CG5	M46-B <sub>1</sub> /M46-B <sub>1</sub>	M46-A <sub>1</sub> /M46-B <sub>1</sub>
M41	CG6	M41-B <sub>1</sub> /M41-B <sub>1</sub>	M41-A <sub>1</sub> /M41-B <sub>1</sub>
M50	CG6	M50-B <sub>1</sub> /M50-B <sub>1</sub>	M50-A <sub>1</sub> /M50-B <sub>1</sub>
M58	CG7	M58-B <sub>1</sub> /M58-B <sub>1</sub>	M58-A <sub>1</sub> /M58-B <sub>1</sub>

<sup>1)</sup> The combination <sup>St</sup>albicans·<sup>St</sup>albicans could not be investigated directly, since it is not realizable genetically. All alleles described in <sup>St</sup>albicans (A genome) resemble the restriction patterns of <sup>h</sup>johansen (A genome) in Table 7.

### 3.2. The complete sequences of the five basic *Oenothera* plastid genomes

Studies of PGI in the genus *Oenothera* require not only markers for the assembling plastome-genome incompatible hybrids. A major prerequisite is the availability of the complete sequences of the five basic plastome types to pinpoint potential plastome determinants responsible for PGI. Although at the beginning of this thesis the complete sequence of plastome I<sup>ioh</sup> as well as substantial parts of the remaining *Oenothera* plastomes already existed (Hupfer, 2002), an incredible poor sequence quality and serious errors in *ycf2* (Rice and Palmer, 2006), made it necessary, however, to re-sequence plastome I<sup>ioh</sup> entirely together with the remaining plastomes. This work was continued by Xi Wang and me. Originally, it was initiated by Rainer Meier, Martina Silber, Helena Funk, and Peter Poltnigg. Sequence annotation and bioinformatics were done in co-operation with Xi Wang and Georg Haberer.

Using a primer-based strategy (Chapter 2.2.2.5.4) the sequences of the plastid chromosomes representing the five genetically distinguishable basic plastomes of subsection *Oenothera*

were finally established. As “standard” basic plastomes served *Oe. elata* subsp. *hookeri* strain johansen (plastome I<sup>loh</sup>), *Oe. biennis* strain suaveolens Grado (plastome II<sup>suavG</sup>), *Oe. glazioviana* strain rr-lamarckiana Sweden (plastome III<sup>lam</sup>), *Oe. parviflora* strain atrovirens Standard (plastome IV<sup>atro</sup>) and *Oe. argillicola* strain douthat 1 (plastome V<sup>doul</sup>). The sequences are available at GeneBank under the accession numbers AJ271079.3, EU262887, EU262889, EU26890 and EU262891.

### 3.2.1. Sequence analysis and annotation of the five plastid chromosomes

The following paragraphs give an overview about structure, base pair composition, coding capacity, and sequence annotation of the five plastid chromosomes. This approach was followed by a comparative analysis to pinpoint functional differences for PGI.

#### 3.2.1.1. Size, gene content and design of the *Oenothera* plastid chromosomes

The *Oenothera* plastid chromosomes are circular molecules of 165,728 bp (plastome I<sup>loh</sup>), 164,807 bp (plastome II<sup>suavG</sup>), 165,225 bp (plastome III<sup>lam</sup>), 163,365 bp (plastome IV<sup>atro</sup>), and 165,055 bp (plastome V<sup>doul</sup>) in size. Over-all, 56.6% are coding, 43.4% are non coding regions (spacers and introns) with a G+C content of 39.1%; 41.7% in coding and 36.7% in non-coding sequence intervals (Table 12). The G+C content is slightly higher than that of *Nicotiana* (37.8%), *Spinacea* (36.8%;) or *Arabidopsis* (36.3%) plastomes, but similar to that of *Oryza* (39.0%) and *Zea* (38.5%) ([http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/plastids\\_tax.html](http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/plastids_tax.html)). Expectedly, the characteristic anatomy found in most of the plastid chromosomes is also displayed in *Oenothera* (Gordon *et al.*, 1981; 1982): A pair of large inverted repeats separates two single copy regions, the large single copy region (LSC) and the small single copy region (SSC) (Figure 18). The overall divergence of the chromosomes is expectedly low, between 96.3% and 98.6% sequence similarity and 96.1% to 98.5% sequence identity (Table 13). These numbers are comparable to those found among various *Nicotiana* species and *Atropa* (96.0% to 98.5% identity; Schmitz-Linneweber *et al.*, 2002; Yukawa *et al.*, 2006).

Within the *Oenothera* plastid chromosomes order and clusters of genes are identical. Each plastome codes for 113 unique genes and two pseudo genes. 18 genes are duplicated in the IR (Figure 18). Basically, genes, gene order and gene clusters are colinear with those of chromosomes from *Nicotiana* (Yukawa *et al.*, 2006), *Lotus* (Kato *et al.*, 2000), *Atropa*

**Table 12.** Sizes and base composition of the *Oenothera* plastome DNAs I<sup>joh</sup> – V<sup>dou1</sup>

plastome	entire chromosome		LSC		IR		SSC		coding regions		non-coding regions	
	size [bp]	GC content [%]	size [bp]	GC content [%]	size [bp]	GC content [%]	size [bp]	GC content [%]	size [bp]	GC content [%]	size [bp]	GC content [%]
I <sup>joh</sup>	165,728	39.1	89,342	37.3	28,683	43.4	19,020	34.4	95,030	41.6	70,698	37.4
II <sup>suavG</sup>	164,807	39.1	88,964	37.4	28,471	43.4	18,901	34.5	92,526	41.7	72,281	36.7
III <sup>lam</sup>	165,225	39.0	89,591	37.3	28,376	43.1	18,882	34.6	92,846	41.8	72,379	36.5
IV <sup>atro</sup>	163,365	39.1	87,732	37.2	28,369	43.4	18,895	34.5	92,152	41.8	71,213	36.6
V <sup>dou1</sup>	165,055	39.1	88,511	37.3	28,772	43.4	19,000	34.5	94,143	41.7	70,912	36.1

LSC = large single copy region

SSC = small single copy region

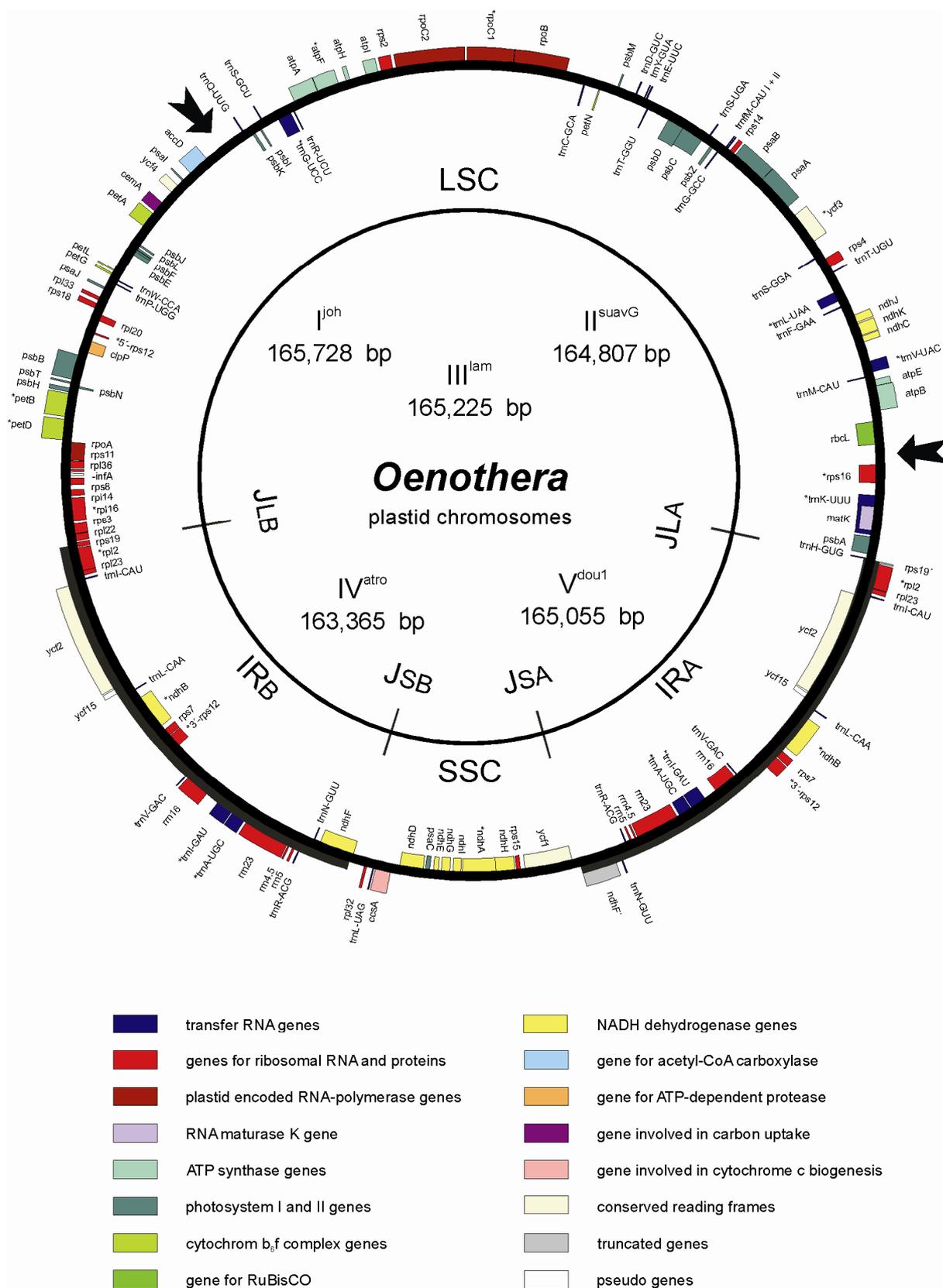
IR = inverted repeat, duplicated in the plastome (IR<sub>A</sub> and IR<sub>B</sub>)

(Schmitz-Linneweber *et al.*, 2002), *Spinacea* (Schmitz-Linneweber *et al.*, 2001a), *Arabidopsis* (Sato *et al.*, 1999) and *Eucalyptus* (Steane, 2005). An exception is a large inversion of approximately 56 kbp in the LSC region (Gordon *et al.*, 1982; Hachtel *et al.*, 1991; Systma *et al.*, 1993; Hupfer *et al.*, 2000). It reverses gene order between *rbcL* and *trnQ*<sub>UUG</sub> (Figure 21 and Chapter 3.2.2.2) since it occurred in the intergenic regions between the *accD/rbcL* and *rps16/trnQ*<sub>UUG</sub>, respectively. The inversion breakpoints are polymorphic and contain repeats (Chapter 3.2.2.2). Specific also for the *Oenothera* plastid chromosomes are two copies of the initiator tRNA *trnI*<sub>CAU</sub>, which differ by a single nucleotide polymorphism in plastomes I<sup>loh</sup>, II<sup>suavG</sup>, III<sup>lam</sup> and IV<sup>atro</sup> and are part of a tandem repeat structure (Chapter 3.2.2).

**Table 13.** Pairwise comparison of sequence similarity and identity in % of the five *Oenothera* plastomes

similarities [%]	II <sup>suavG</sup>	III <sup>lam</sup>	IV <sup>atro</sup>	V <sup>doul</sup>	identity [%]	II <sup>suavG</sup>	III <sup>lam</sup>	IV <sup>atro</sup>	V <sup>doul</sup>
I <sup>loh</sup>	98.6	97.5	96.3	96.3	I <sup>loh</sup>	98.5	97.3	95.9	95.9
II <sup>suavG</sup>	---	97.9	97.2	96.7	II <sup>suavG</sup>	---	97.8	96.9	96.4
III <sup>lam</sup>	---	---	96.5	96.3	III <sup>lam</sup>	---	---	96.2	96.1
IV <sup>atro</sup>	---	---	---	97.7	IV <sup>atro</sup>	---	---	---	97.6

Taken together the *Oenothera* plastid chromosomes encode 4 rRNA genes (16S, 23S, 5S, 4.5S), a total of 31 distinct tRNA genes and 78 protein-coding loci including *yef1*, *yef2*, *yef3* and *yef4* (Table 14). A single intron is found in sixteen genes, one gene (*yef3*) contains two. *Rps12* mRNA is generated by transsplicing as also observed in other species. All five species share the same set of introns, one class I intron (*trnL*<sub>UUA</sub>) and 17 of class II. *ClpP* lacks both introns in comparison to *Nicotiana*, the surrounding coding sequences are conserved. For both introns, the deleted sequences overlap precisely with intron borders established in reference species. This indicates that the mechanism of intron loss may have involved a processed RNA intermediate. As in various other plastid chromosomes, two pairs of genes overlap, *atpB-atpE* (4 bp) and *psbC-psbD* (52 bp); *matK* is located within the intron of *trnK*<sub>UUU</sub>. The 31 tRNA species are sufficient to satisfy all the requirements for protein synthesis in the organelle and represent 20 amino acid species. The standard plastid/bacterial code with a predicted methionine ATG start codon is used for all protein-coding genes. Nevertheless, two exceptions have been noted: *NdhD* starts with ACG and is edited as shown for plastome I<sup>loh</sup> and IV<sup>atro</sup> (Hupfer, 2002); *cemA* was tentatively annotated with an ATA start codon (for details see Chapter 3.2.1.2). As in other plastid chromosomes, the most common stop codon is



**Figure 18.** Gene map of the *Oenothera* plastid chromosomes. Arrows mark the inversion breakpoints. Genes drawn on the outside are transcribed clockwise, on the inside counter clockwise.

**Table 14.** List of genes found in the *Oenothera* plastomes<sup>1)</sup>

gene categories	no. genes	genes
<b>transcription</b>	4	<i>rpoA</i> , <i>rpoB</i> , <i>*rpoC1</i> , <i>rpoC2</i>
<b>RNA processing</b>	1	<i>matK</i>
<b>ribosomal RNA</b>	4	<i>rrn5</i> , <i>rrn4.5</i> , <i>rrn16</i> , <i>rrn23</i>
<b>transfer RNA</b>	31	<i>*trnA<sub>UGC</sub></i> , <i>trnC<sub>GCA</sub></i> , <i>trnD<sub>GUC</sub></i> , <i>trnE<sub>UUC</sub></i> , <i>trnF<sub>GAA</sub></i> , <i>trnG<sub>GCC</sub></i> , <i>*trnG<sub>UCC</sub></i> , <i>trnH<sub>GUG</sub></i> , <i>trnI<sub>CAU</sub></i> , <i>*trnI<sub>GAU</sub></i> , <i>*trnK<sub>UUU</sub></i> , <i>trnL<sub>CAA</sub></i> , <i>*trnL<sub>UAA</sub></i> , <i>trnL<sub>UAG</sub></i> , <i>trnM<sub>CAU</sub>I</i> , <i>trnM<sub>CAU</sub>II</i> , <i>trnM<sub>CAU</sub></i> , <i>trnN<sub>GUU</sub></i> , <i>trnP<sub>UGG</sub></i> , <i>trnQ<sub>UUG</sub></i> , <i>trnR<sub>ACG</sub></i> , <i>trnR<sub>UCU</sub></i> , <i>trnS<sub>GCU</sub></i> , <i>trnS<sub>GGA</sub></i> , <i>trnS<sub>UGA</sub></i> , <i>trnT<sub>GGU</sub></i> , <i>trnT<sub>UGU</sub></i> , <i>trnV<sub>GAC</sub></i> , <i>*trnV<sub>UAC</sub></i> , <i>trnW<sub>CCA</sub></i> , <i>trnY<sub>GUA</sub></i>
<b>ribosomal proteins</b>	21	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> , <i>rps8</i> , <i>rps11</i> , <i>*rps12</i> , <i>rps14</i> , <i>rps15</i> , <i>*rps16</i> , <i>rps18</i> , <i>rps19</i> ( <i>rps19'</i> ), <i>*rpl2</i> , <i>rpl14</i> , <i>*rpl16</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> , <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
<b>photosystem I</b>	5	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
<b>photosystem II</b>	15	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
<b>ATP synthase</b>	6	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>*atpF</i> , <i>atpH</i> , <i>atpI</i>
<b>cytochrome <i>b6f</i> complex</b>	6	<i>petA</i> , <i>*petB</i> , <i>*petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
<b>Calvin cycle</b>	1	<i>rbcL</i>
<b>cytochrome <i>c</i> synthesis</b>	1	<i>ccsA</i>
<b>NADPH dehydrogenase</b>	11	<i>*ndhA</i> , <i>*ndhB</i> , <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> ( <i>ndhF'</i> ), <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
<b>carbon metabolism</b>	1	<i>cemA</i>
<b>fatty acid synthesis</b>	1	<i>accD</i>
<b>proteolysis</b>	1	<i>clpP</i>
<b>PSI assembly/stability</b>	2	<i>*ycf3</i> , <i>ycf4</i>
<b>conserved reading frames</b>	2	<i>ycf1</i> , <i>ycf2</i>
<b>pseudo genes</b>	2	$\psi$ - <i>infA</i> , $\psi$ - <i>ycf15</i>

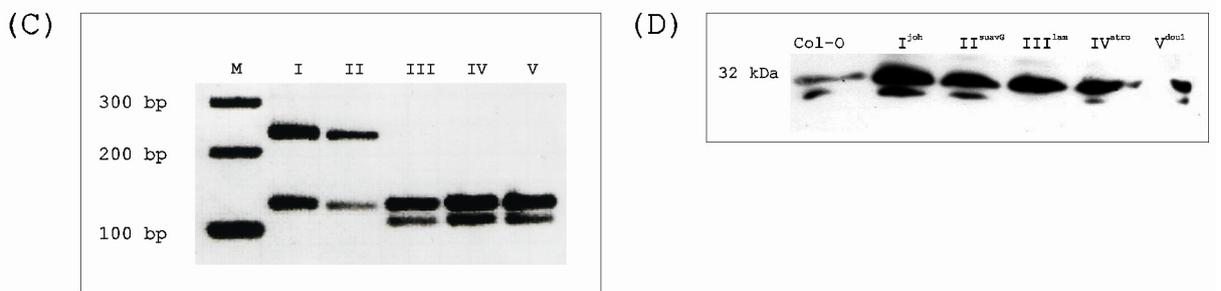
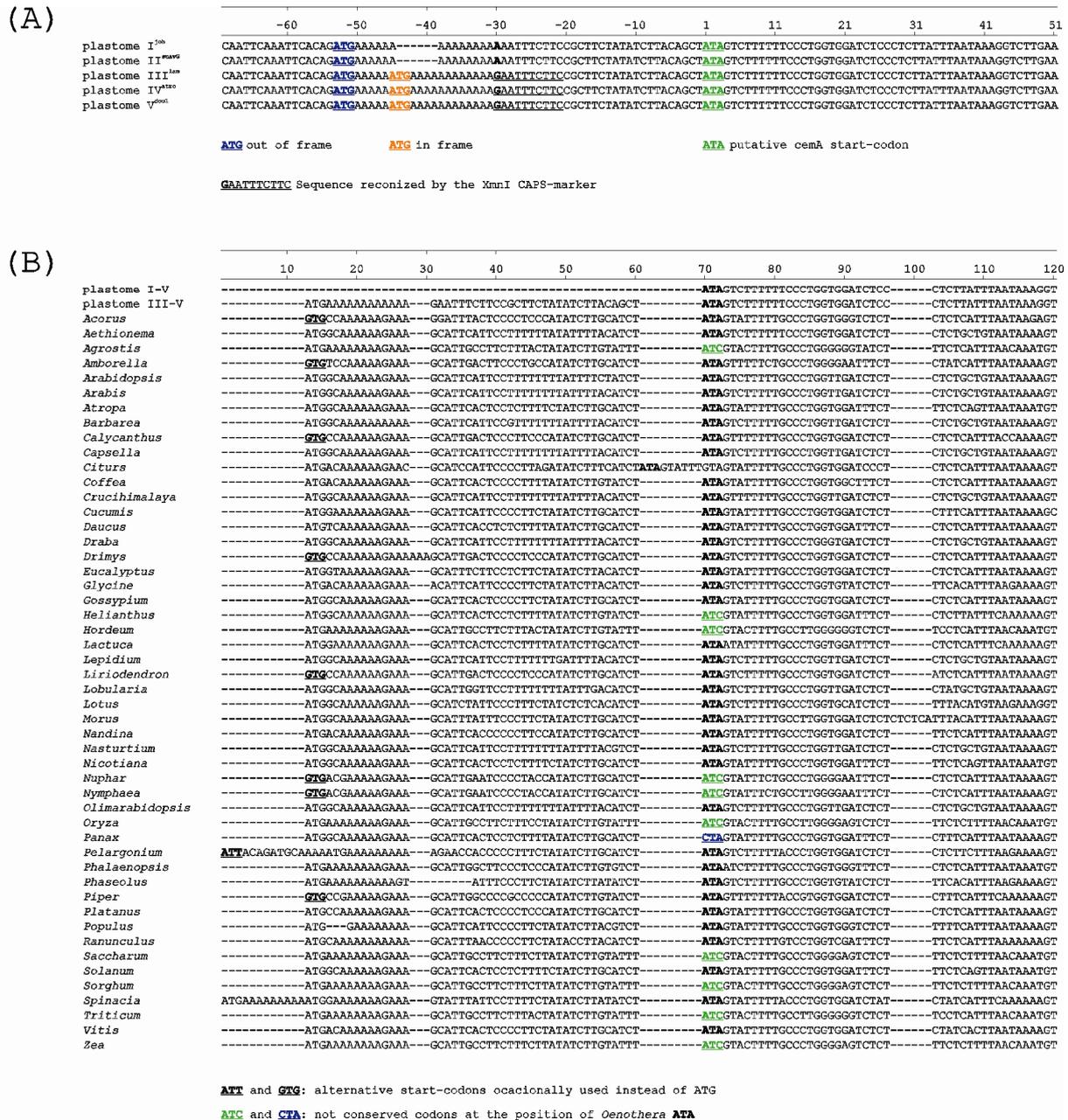
<sup>1)</sup> Intron-containing genes are marked by *asterisks* (\*). Note that two copies of the gene *trnM<sub>CAU</sub>* are located in tandem in the LSC region and truncated versions of *ndhF* (*ndhF'*) and *rps19* (*rps19'*) are found at the borders of the IR<sub>A</sub>.

TAA (51.3%) (Meurer *et al.*, 2002). Within the 78 protein coding genes five, namely *accD*, *rpl22*, *rpl23*, *ycf1* and *ycf2*, are not generally found in angiosperms, but appear to be functional in *Oenothera*. In contrast, *sprA* is missing. Its presence was reported from various dicots, *e.g.* *Solanaceae*, *Arabidopsis* or *Spinacea* (Vera and Sugiura, 1994; Schmitz-Linneweber *et al.*, 2002). Like in *Nicotiana*, *Arabidopsis* and *Eucalyptus*, *infA* present in various species appears as pseudo gene. An additional pseudo gene ( $\psi$ -*ycf15*) is also present in all plastomes. It has premature in-frame stop codons generated by frequent insertions of variable sizes. The significance of such less conserved ORFs with usually regional homology is unknown (Schmitz-Linneweber *et al.*, 2001a). The junctions of IR<sub>A</sub>/LSC, LSC/IR<sub>B</sub> and IR<sub>B</sub>/SSC are identical within the five plastomes. A minor exception is the SSC/IR<sub>A</sub> junction, which is identical in plastomes I<sup>loh</sup>, II<sup>suavG</sup> and III<sup>lam</sup>, but differs by two additional bp in plastomes IV<sup>atro</sup> and V<sup>doul</sup>. The region downstream of *ycf1* until *ndhF'* is highly polymorphic among all five plastomes, and *ycf1* entirely located within SSC. At the border of IR<sub>B</sub> additional, truncated versions of *ndhF* and *rps19* are found at the SSC/IR<sub>A</sub> and IR<sub>A</sub>/LSC junctions, respectively. *NdhF'* lacks the 5' end, *rps19'* the 3' end. It is not known whether the truncated versions are functional.

### 3.2.1.2. *CemA* is annotated with an alternative start codon ATA

*CemA* provided a special annotation problem within its 5' terminus, not only for the five *Oenothera* plastomes but also for the reference species (Figure 19). *CemA* (*ycf10*) encodes an inner envelope polypeptide involved in CO<sub>2</sub> uptake (Rolland *et al.*, 1997). Figure 19 presents the sequence alignment of the 5' *cemA* region from *Oenothera*. The sequence context precludes an unequivocal assignment of the N-terminus of the reading frame *per se* (Figure 19, panel A) and by comparison with 50 reference plastomes (Figure 19, panel B).

The *cemA* 5' regions of plastomes III<sup>lam</sup>, IV<sup>atro</sup> and V<sup>doul</sup> contain two ATG codons embedded in a polyA stretch, of which only the second one is in frame. The first one would require a two bp frame shift. Possibly none of them is being used. The presence of two ATG sequences and polyA stretches are shared with other taxa, but the corresponding sequence interval is not well conserved. In plastomes I<sup>loh</sup> and II<sup>suavG</sup> a deletion of 6 bp removes the second, in-frame ATG (Figure 19, panel A). The ATG common to all five plastomes, 9 bases upstream of the deletion, would necessitate a frame shift for correct translation (Figure 19, panel A). The differences between the plastomes could be confirmed not only by sequence analysis but also by a CAPS marker. A PCR product amplified with the primer pair *cemA5'rev* and *psbL7in1*



**Figure 19.** *CemA* alignment of the five basic plastomes (panel A) and 50 reference species (panel B). Verification of the sequence polymorphism in panel A via a CAPS marker (panel C), and immunological confirmation of the *CemA* protein in different species harbouring the five basic plastome types (panel D). Detailed explanations are given in the text.

and digested with *XmnI* provided the indicated restriction polymorphism (Figure 19, panel C). Western analysis with isolated thylakoid membranes using a polyclonal monospecific CemA (Ycf10) antiserum elicited against the *Arabidopsis* protein showed that a CemA protein of 32 kDa is present in all five plastid types (Figure 19, panel D). Their molecular masses are virtually identical and of the same size as the corresponding *Arabidopsis* protein (strain Col-0) suggesting comparable translational initiation. The use of the only ATG alternative for *cemA* in plastomes I and II at nucleotide positions 66,631 (plastome I<sup>ioh</sup>) or 66,106 (plastome II<sup>suavG</sup>) between the 3' end of *ycf4* and the *cemA* stop codon would generate a predicted protein of 21 kDa rather than of 27 kDa as in *Arabidopsis* or plastomes III<sup>lam</sup>, IV<sup>atro</sup> and V<sup>doul</sup>. It is therefore unlikely that this ATG serves as an initiation codon. Furthermore, cDNA sequences, amplified with the primer pair *cemA*5'rev and *psbL7int1* 200 bp upstream and 150 bp downstream of the ATG motif, contain no edited start codon in all five plastomes. Finally, RT-PCR analysis to search for a nuclear copy of *cemA* was performed from all five species used for plastome sequencing with the 5' primer *cemA*for2 and three polyT<sub>18</sub> primer with an additional 3' base, A, C or G, but no signal was obtained. These findings indicate that also no functional gene translocation to the nucleus of *Oenothera* species carrying plastome I<sup>ioh</sup> or II<sup>suavG</sup> has occurred.

Therefore, for the annotation of *cemA* in the five *Oenothera* plastomes, one must either postulate an unknown frame-correcting mechanism, which overcomes the missing base, or CemA translation begins at an alternative start codon. An appropriate candidate would be a conserved ATA motif at positions 66,526 (plastome I<sup>ioh</sup>), 66,001 (plastome II<sup>suavG</sup>), 66,535 (plastome III<sup>lam</sup>), 65,401 (IV<sup>atro</sup>) and 65,948 (V<sup>doul</sup>), which would result in a predicted polypeptide of 25 kDa (Figure 19, panel A and B). This motif is conserved, for instance, in *Arabidopsis*, *Atropa*, *Calyanthus*, *Citrus*, *Cucumbis*, *Eucalyptus*, *Gossypium*, *Lotus*, *Spinacea*, *Nicotiana* and *Vitis* but not in *Zea mays* (Figure 19, panel B). Non-ATG start codons occasionally utilized, e.g. GUG, UUG and AUU, operate less efficiently than AUG (Kozak, 1983) and a near optimal initiation codon context may be required to compensate. It seems possible that plastids utilize ATA as an initiation codon. Therefore, *cemA* was tentatively annotated in all five plastomes with the alternative start codon ATA, but this point needs to be settled.

### 3.2.2. Analysis of indels and sequence repetition in *Oenothera* plastomes

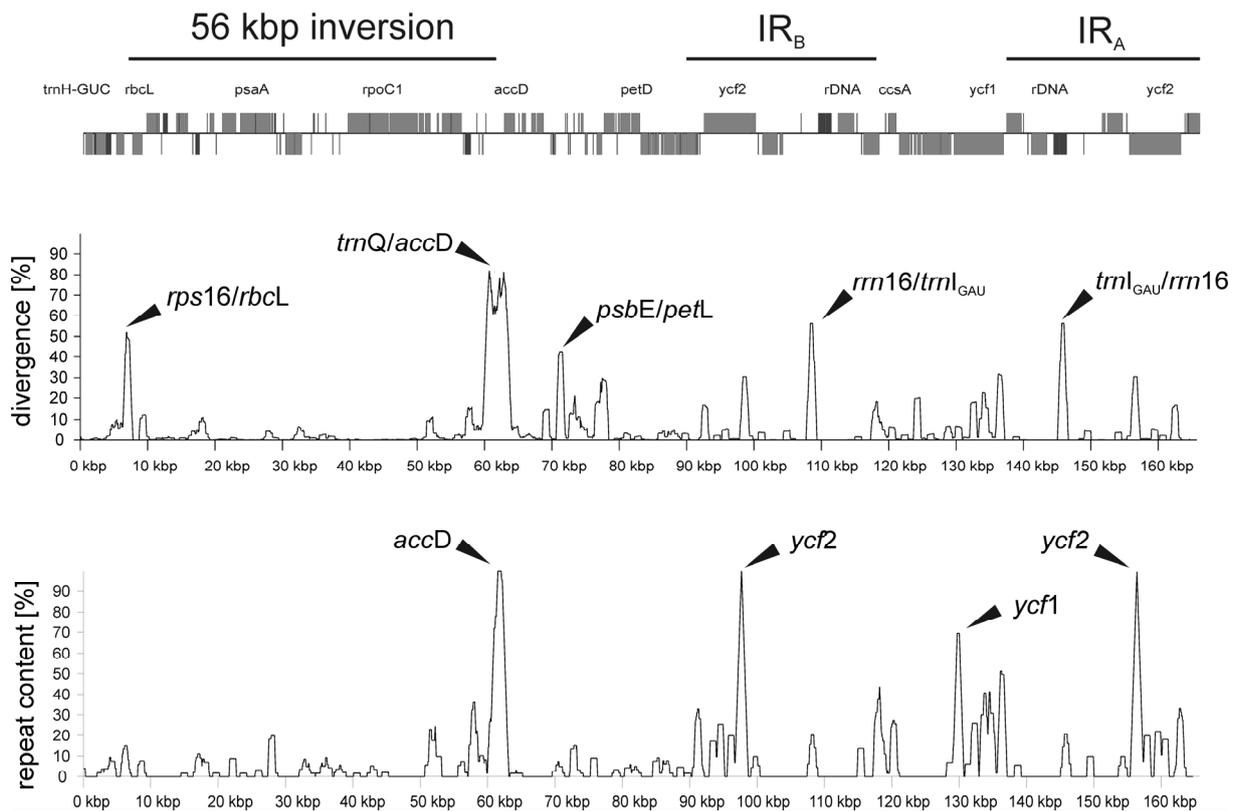
Differences within the five plastomes mostly depend on repeat structures, at least in non-coding regions. PGI may therefore also reflect a result of structural evolution of the plastid chromosome. Furthermore, since the five plastomes are perfectly syntenic, a unique opportunity for a comparative analysis of indels and sequence repetition is given. Such an analysis is usually difficult between distant species.

#### 3.2.2.1.1. Indels within the five plastomes

Insertions, deletions and repetitions are relatively frequent within the five plastomes. Relative to plastome IV<sup>atro</sup>, 1456 nucleotide insertions and 3819 deletions are found in plastome I<sup>oh</sup>, for plastome II<sup>suavG</sup>, III<sup>lam</sup> and V<sup>doul</sup> numbers are 1156/2598, 1701/3561, 864/2557, respectively. As expected, indels occur less often in genes and are present in only relatively few polypeptide genes. Only *accD*, *clpP*, *ndhD*, *ndhF*, *rps18*, *rpl22*, *ycf1* and *ycf2* are affected. These changes in genes will be discussed separately (Chapters 3.2.3.2.1 - 3.2.3.2.4).

#### 3.2.2.1.2. Tandem and palindrome repeats

Both repeat types, tandem and palindrome repeats, are distributed highly similar between all five plastomes. In terms of tandem repeats, on average 61 were detected, within a range of 55 (plastome IV<sup>atro</sup>) up to 70 (plastome I<sup>oh</sup>). The mean copy number was 4.5 copies per tandem and an average size of 41 bp per copy. In all plastomes the largest tandem repeat regions spanning more than 1 kbp were found in the two *ycf2* genes in the IR. They consist of variants of an AAG/TTC trinucleotide sequence. Expanded tandem repeats are also frequent and overlapping in *accD* and *ycf1* contributing to the substantial sequence divergence of these genes (Chapters 3.2.3.2.3 and 3.2.3.2.4). Associated with small identical tandem repeats in all five plastomes are *rpl32*, *ndhF* and the tandemly repeated *trnM<sub>CAU</sub>I* and *trnM<sub>CAU</sub>II*. More interesting are tandem repeats, which change coding potential specific to distinct plastomes. These repeats, generally of moderate size and low copy number, are found in *clpP*, *ndhF* and *rps18* (Chapters 3.2.3.2.1 and 3.2.3.2.2). The plastome-specific repeat differences associated with *ccsA*, *rpl22*, *rpl32*, *rps19* and *trnS<sub>GCU</sub>* do not change coding context, since the differing repetitive elements are located outside coding sequences.



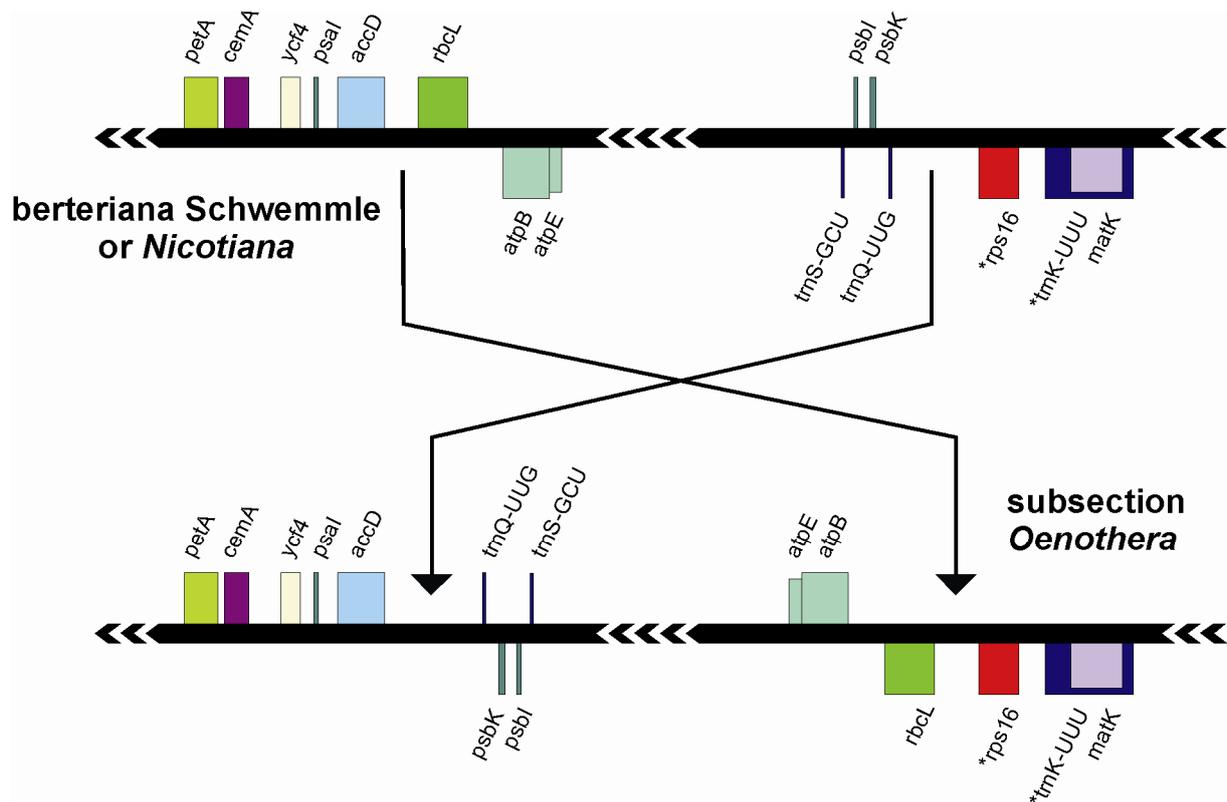
**Figure 20.** Overall distribution of sequence divergence and repetition in the *Oenothera* plastid chromosomes. Repeat regions and areas of high divergence are often, but not always, correlated.

Analyses of palindrome repeats give similar results. Approximately 70 palindrome repeats on average, with a maximal gap size of 3 kb or less, were detected within the five plastome sequences. However, these repeats are generally smaller and far less variable. Sizes ranged from 32 bp (detection limit given by the threshold applied) up to 56 bp. They were detected in *accD*, *ccsA*, *matK*, *ndhD*, *ndhF*, *ndhJ*, *petD*, *psaA*, *psaB*, *psbH*, *rpl32*, *rpoA*, *rpoB*, *rps18*, *ycf1*, *ycf2* and *ycf4*, but no notable change of coding sequence associated with palindrome repeats was observed. The highly polymorphic genes *accD* and *ycf2* are exceptions, as it is *ndhD*. Here, frame changing insertion of a single nucleotide in a polyA-tail (Chapter 3.2.3.2.2) is located in a palindrome repeat. Presence and absence of repetitive elements and divergent regions correlate well, but repeat content and high divergence are not strongly linked (Figure 20).

### 3.2.2.2. The large 56 kbp inversion

As shown in Figure 20, the breakpoints of the large 56 kbp inversion (Figure 21) are associated with divergence and repeat content. The junctions of the inversion in the intergenic

regions between *trn*<sub>Q<sub>UUG</sub></sub>/*accD* and *rps16/rbcL* are highly divergent and contain palindromes as well as tandem repeats. Therefore, a closer investigation of the breaking points in the *Oenothera* plastomes was performed. This analysis also included the study of the ancestral situation in the subsection *Munzia*, lacking the inversion (Gordon *et al.*, 1982; Hachtel *et al.*, 1991; Systma *et al.*, 1993; Hupfer *et al.*, 2000).



**Figure 21.** Scheme indicating the 56 kbp inversion that occurred in the sequence intervals *rps16/rbcL* and *trn*<sub>Q<sub>UUG</sub></sub>/*accD*, respectively, in the large single copy segment of the *Oenothera* plastid chromosomes. The inversion flanked by *trn*<sub>Q<sub>UUG</sub></sub> and *rbcL* changes the orientation of 35 genes. Transcription of genes above lines is counter clockwise, below lines clockwise.

In summary, within the *Oenothera* plastomes more tandem repeats are detected in the *trn*<sub>Q<sub>UUG</sub></sub>/*accD* spacer, than in the *rps16/rbcL* spacer. In detail, 5 to 8 tandem repeats were detected between the *trn*<sub>Q<sub>UUG</sub></sub> and *accD* spacer depending on the plastome. Only a maximum of two tandem repeats were found in plastome IV<sup>atro</sup> and none in plastome V<sup>doul</sup> in the *rps16/rbcL* spacer. In terms of palindrome repeats, three were detected at the *rps16/rbcL* junction of plastomes I<sup>oh</sup> and II<sup>suav<sup>G</sup></sup>, and none in the remaining plastomes. Between two and no palindrome(s) were identified in the *trn*<sub>Q<sub>UUG</sub></sub>/*accD* spacer region. Elements of clustered palindromes, split among both spacer regions, are of particular note. Such elements were

detected in all five plastomes. Their number varies enormously, between one (plastome V<sup>doul</sup>) and 17 (plastome III<sup>lam</sup>). Furthermore, the same sequence motifs can be present in tandem and palindrome arrangement. The repetition patterns preclude an accurate demarcation of the insertion breakpoints and also affect the highly variable N-terminal region of AccD (Chapter 3.2.3.2.4).

In an attempt to better understand the inversion breakpoints and underlying processes, the corresponding regions in *Oe. villaricae* strain berteriana Schwemmler, a member of the closely related sister subsection *Munzia*, were sequenced (*rps16/trnQ*<sub>UUG</sub>, accession no. EU255777 and *rbcL/accD*, accession no. EU255778). Subsection *Munzia* lacks this inversion (Hachtel *et al.*, 1991). The berteriana Schwemmler regions corresponding to the *Oenothera* breakpoints between *rps16/trnQ*<sub>UUG</sub> and *rbcL/accD* do not display the pronounced divergence. The spacer region between *rps16* and *trnQ* in that plastid chromosome lacks tandem and palindrome repeats. Nevertheless, the entire region is conserved and present in two parts in all five *Oenothera* plastomes, separated by an interspersed *Oenothera*-specific sequence interval. The berteriana Schwemmler region between *rbcL* and *accD*, in turn, lacks palindromes, but contains two tandem repeats. Approximately 1.5 kb of the berteriana Schwemmler *rbcL/accD* spacer are unique to the berteriana Schwemmler plastome. Conversely, *trnQ/accD* spacer sequences between 1.5 and 2.5 kb depending on the *Oenothera* plastome have no equivalent in the berteriana Schwemmler plastome. The same holds true for the *Oenothera rps16/rbcL* spacer, in which the number of unique nucleotides differs between approximately 50 and 500 bp. The repeat structure in the *rps16/rbcL* spacer of the *Oenothera* plastomes (Figure 20) appears to be linked to the inversed arrangement of *rbcL* and the result of duplication and relocation during the inversion process, as the berteriana Schwemmler equivalent is missing. As mentioned, palindrome repeat copies split among both spacers are not rare in the *Oenothera* plastomes, but no such cases were detected in berteriana Schwemmler with the selected threshold.

### 3.2.3. Differences in coding regions among *Oenothera* plastomes $V^{oh} - V^{doul}$

As shown in the last paragraphs coding capacity and genome structure of the five *Oenothera* plastomes resemble closely that of known plastid chromosome. Although the *Oenothera* plastomes have their particularities like the large inversion (Chapter 3.2.2.2), mechanisms which cause PGI do not seem to be linked with the presence/absence of singular sets of genes or with an uncommon genome structure. Therefore, a search of PGI determinants must

concentrate on smaller differences in coding or regulatory regions. Differences in coding regions are of special interest. In the *Oenothera* plastomes they are found in two classes, single codon exchanges and genes showing size polymorphism.

### 3.2.3.1. Summary of all gene coding differences

All three gene coding sequence classes (protein, tRNA and rRNA coding loci) were compared to detect changes between the five plastomes. Of the 78 protein-coding genes, the nucleotide sequence of 31 is identical, while 47 vary between at least one plastome pair (Table 15). Single base pair substitutions were detected in all 47 polymorphic genes and in 8 of them additionally indels are present. In sixteen genes substitutions do not change coding context, namely in *atpE*, *cemA*, *ndhI*, *ndhJ*, *petA*, *petB*, *psaB*, *psaC*, *psbC*, *rpl16*, *rpl33*, *rpoC1*, *rps4*, *rps11*, *rps14* and *rps19*. Only synonymous substitutions were detected in 11 genes: *ndhB*, *ndhC*, *ndhH*, *psbB*, *rpl32*, *rpoA*, *rpoB*, *rpoC2*, *rps2*, *rps8*, and *rps15* (Tables 15 and 16). A synonymous substitution and a single bp insertion/deletion were found in *rpl22* (Chapter 3.2.3.2.2). Furthermore, synonymous substitutions and a multiple base pair indel were present in *ndhF* (Chapter 3.2.3.2.1), as well as multiple indels and a non-synonymous substitution in *rps18* (Chapter 3.2.3.2.2). Both types of substitutions were confirmed for *ndhD*, in plastome I<sup>oh</sup> together with a single base pair insertion, and a multiple base pair insertion in plastome V<sup>dou1</sup> (Chapter 3.2.3.2.2). In the remaining 16 genes (*accD*, *atpA*, *atpB*, *atpF*, *ccsA*, *clpP*, *matK*, *ndhA*, *ndhE*, *petD*, *psaA*, *psbA*, *rps3*, *ycf1*, *ycf3* and *ycf4*) both, synonymous and non-synonymous substitutions are present. Five of them, *clpP*, *accD*, *ndhD*, *ycf1* and *ycf2*, differ by multiple indels among plastomes (Chapters 3.2.3.2.1 - 3.2.3.2.4). Of the 31 tRNA genes, 30 are identical and one, *trnfM<sub>CAU</sub>II*, is variable in one nucleotide. The mutation should have no or only a negligible effect on the folding of the tRNA, and hence not influence function, since it is not part of the anticodon. In addition, a second, not mutated copy, *trnfM<sub>CAU</sub>I*, is present in the *Oenothera* plastomes. All four *rrn* genes are identical among the *Oenothera* plastomes.

**Table 15.** Summary of all differences at the nucleotide level found in coding regions of the five *Oenothera* plastomes in pairwise comparison.

gene	length (alignment)	DNA																			
		# pairwise substitution (bp)										# pairwise insertion/deletion (bp)									
		I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V	I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V
<i>accD</i>	1711	16	39	48	133	23	36	109	29	115	57	144	177	333	170	57	189	314	228	317	461
<i>atpA</i>	1518	1	2	1	1	3	2	2	3	3	-	-	-	-	-	-	-	-	-	-	-
<i>atpB</i>	1497	-	1	2	2	1	2	2	1	1	2	-	-	-	-	-	-	-	-	-	-
<i>atpE</i>	402	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-
<i>atpF</i>	555	1	1	1	2	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-
<i>ccsA</i>	960	-	1	4	3	1	4	3	5	4	3	-	-	-	-	-	-	-	-	-	-
<i>cemA</i>	645	-	1	-	-	1	-	-	1	1	-	-	-	-	-	-	-	-	-	-	-
<i>clpP</i>	759	-	7	27	28	7	27	28	26	27	1	-	3	45	6	3	45	6	48	9	51
<i>matK</i>	1539	1	3	6	6	2	5	5	3	3	4	-	-	-	-	-	-	-	-	-	-
<i>ndhA</i>	1092	-	3	5	6	3	5	6	4	5	1	-	-	-	-	-	-	-	-	-	-
<i>ndhB</i>	1533	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-
<i>ndhC</i>	363	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
<i>ndhD</i>	1551	-	1	3	4	1	3	4	2	3	1	21	21	21	27	-	-	48	-	48	48
<i>ndhE</i>	306	-	2	1	-	2	1	-	3	2	1	-	-	-	-	-	-	-	-	-	-
<i>ndhF</i>	2334	1	2	1	2	1	-	1	1	2	1	15	15	15	15	-	-	-	-	-	-
<i>ndhH</i>	1182	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
<i>ndhI</i>	498	-	-	1	-	-	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-
<i>ndhJ</i>	477	-	1	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 15. (continued)

gene	length (alignment)	DNA																			
		# pairwise substitution (bp)										# pairwise insertion/deletion (bp)									
		I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V	I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V
<i>petA</i>	957	-	1	1	1	1	1	1	2	2	2	-	-	-	-	-	-	-	-	-	
<i>petB</i>	648	-	-	1	1	-	1	1	1	1	2	-	-	-	-	-	-	-	-	-	
<i>petD</i>	483	3	-	1	1	3	4	4	1	1	2	-	-	-	-	-	-	-	-	-	
<i>psaA</i>	2253	-	-	1	3	-	1	3	1	3	2	-	-	-	-	-	-	-	-	-	
<i>psaB</i>	2205	-	1	1	1	1	1	1	2	2	-	-	-	-	-	-	-	-	-	-	
<i>psaC</i>	246	-	1	-	-	1	-	-	1	1	-	-	-	-	-	-	-	-	-	-	
<i>psbA</i>	1062	-	-	1	1	-	1	1	1	1	2	-	-	-	-	-	-	-	-	-	
<i>psbB</i>	1527	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	
<i>psbC</i>	1422	-	1	-	-	1	-	-	1	1	-	-	-	-	-	-	-	-	-	-	
<i>rpl16</i>	408	-	1	2	2	1	2	2	1	1	2	-	-	-	-	-	-	-	-	-	
<i>rpl22</i>	429	-	-	-	1	-	-	1	-	1	1	-	-	15	-	15	-	15	-	15	
<i>rpl32</i>	156	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
<i>rpl33</i>	201	1	-	1	-	1	2	1	1	-	1	-	-	-	-	-	-	-	-	-	
<i>rpoA</i>	1101	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	
<i>rpoB</i>	3219	-	1	5	5	1	5	5	4	4	-	-	-	-	-	-	-	-	-	-	
<i>rpoC1</i>	2040	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	
<i>rpoC2</i>	4152	1	1	1	4	2	2	5	2	5	3	-	-	-	-	-	-	-	-	-	
<i>rps2</i>	711	-	-	2	2	-	2	2	2	2	-	-	-	-	-	-	-	-	-	-	

Table 15. (continued)

gene	length (alignment)	DNA																			
		# pairwise substitution (bp)										# pairwise insertion/deletion (bp)									
		I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V	I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V
<i>rps3</i>	660	-	1	2	2	1	2	2	3	3	-	-	-	-	-	-	-	-	-	-	
<i>rps4</i>	612	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	
<i>rps8</i>	417	-	2	1	1	2	1	1	3	3	-	-	-	-	-	-	-	-	-	-	
<i>rps11</i>	435	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	
<i>rps14</i>	303	-	-	-	2	-	-	2	-	2	2	-	-	-	-	-	-	-	-	-	
<i>rps15</i>	264	-	1	1	2	1	1	2	-	1	1	-	-	-	-	-	-	-	-	-	
<i>rps18</i>	324	-	5	1	1	5	1	1	6	6	-	-	48	18	18	48	18	18	30	30	
<i>rps19</i>	300	-	1	2	1	1	2	1	1	2	1	-	-	-	-	-	-	-	-	-	
<i>ycf1</i>	7446	7	38	55	70	32	54	65	56	61	18	234	231	348	219	279	250	231	153	138	
<i>ycf2</i>	7185	5	18	27	31	12	23	24	26	30	11	45	108	144	164	153	101	209	252	108	
<i>ycf4</i>	555	-	1	1	1	1	1	1	2	2	-	-	-	-	-	-	-	-	-	-	

plastome I = I<sup>loh</sup>, plastome II = II<sup>suavG</sup>, plastome III = III<sup>lam</sup>, plastome IV = IV<sup>atro</sup>, plastome V = V<sup>dou1</sup>

**Table 16.** Summary of all differences at the protein level found in coding regions of the five *Oenothera* plastomes in pairwise comparison.

gene	length (alignment)	protein																			
		# pairwise substitution (amino acids)										# pairwise insertion/deletion (amino acids)									
		I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V	I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V
<i>accD</i>	612	5	22	23	55	19	21	48	21	53	7	48	59	139	136	19	91	104	98	105	153
<i>atpA</i>	508	1	1	1	1	2	2	2	2	2	-	-	2	-	-	2	-	-	2	2	-
<i>atpB</i>	499	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-
<i>atpE</i>	134	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>atpF</i>	185	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-
<i>ccsA</i>	320	-	1	3	2	1	3	2	4	3	1	-	-	-	-	-	-	-	-	-	-
<i>cemA</i>	215	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>clpP</i>	253	-	6	13	14	6	13	14	12	13	1	-	1	15	2	1	15	2	16	3	17
<i>matK</i>	513	-	2	4	3	2	4	3	2	1	1	-	-	-	-	-	-	-	-	-	-
<i>ndhA</i>	364	-	3	4	5	3	4	5	3	4	1	-	-	-	-	-	-	-	-	-	-
<i>ndhB</i>	511	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-	-
<i>ndhC</i>	121	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
<i>ndhD</i>	517	2	2	3	9	-	1	4	1	4	3	7	7	7	9	-	-	16	-	16	16
<i>ndhE</i>	102	-	2	-	-	2	-	-	2	2	-	-	-	-	-	-	-	-	-	-	-
<i>ndhF</i>	778	1	1	1	1	-	-	-	-	-	-	5	5	5	5	-	-	-	-	-	-
<i>ndhH</i>	394	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
<i>ndhI</i>	166	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>ndhJ</i>	159	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 16. (continued)

gene	length (alignment)	protein																		
		# pairwise substitution (amino acids)									# pairwise insertion/deletion (amino acids)									
		I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V	I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V
<i>petA</i>	319	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>petB</i>	216	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>petD</i>	161	2	-	1	-	2	3	2	1	-	1	-	-	-	-	-	-	-	-	-
<i>psaA</i>	751	-	-	-	1	-	-	1	-	1	1	-	-	-	-	-	-	-	-	-
<i>psaB</i>	735	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>psaC</i>	82	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>psbA</i>	354	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>psbB</i>	509	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-
<i>psbC</i>	474	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>rpl16</i>	136	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>rpl22</i>	143	-	-	1	-	-	1	-	1	-	1	-	-	5	-	-	5	-	5	-
<i>rpl32</i>	52	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>rpl33</i>	67	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>rpoA</i>	367	-	-	1	1	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-
<i>rpoB</i>	1073	-	1	5	5	1	5	5	4	4	-	-	-	-	-	-	-	-	-	-
<i>rpoC1</i>	680	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>rpoC2</i>	1384	1	1	1	4	2	2	5	2	5	3	-	-	-	-	-	-	-	-	-
<i>rps2</i>	237	-	-	2	2	-	2	2	2	2	-	-	-	-	-	-	-	-	-	-

Table 16. (continued)

gene	length (alignment)	protein																			
		# pairwise substitution (amino acids)										# pairwise insertion/deletion (amino acids)									
		I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V	I/II	I/III	I/IV	I/V	II/III	II/IV	II/V	III/IV	III/V	IV/V
<i>rps3</i>	220	-	1	1	1	1	1	1	2	2	-	-	-	-	-	-	-	-	-	-	
<i>rps4</i>	204	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
<i>rps8</i>	139	-	2	1	1	2	1	1	3	3	-	-	-	-	-	-	-	-	-	-	
<i>rps11</i>	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
<i>rps14</i>	101	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
<i>rps15</i>	88	-	1	1	2	1	1	2	-	1	1	-	-	-	-	-	-	-	-	-	
<i>rps18</i>	111	-	7	2	15	7	2	15	5	3	13	-	10	6	6	10	6	6	10	10	6
<i>rps19</i>	100	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
<i>ycf1</i>	2489	10	36	40	46	28	37	41	41	44	14	78	77	116	87	81	66	73	65	46	67
<i>ycf2</i>	2410	3	17	22	30	13	21	25	20	32	10	17	38	50	56	53	35	69	80	64	82
<i>ycf4</i>	185	-	1	-	-	1	-	-	1	1	-	-	-	-	-	-	-	-	-	-	

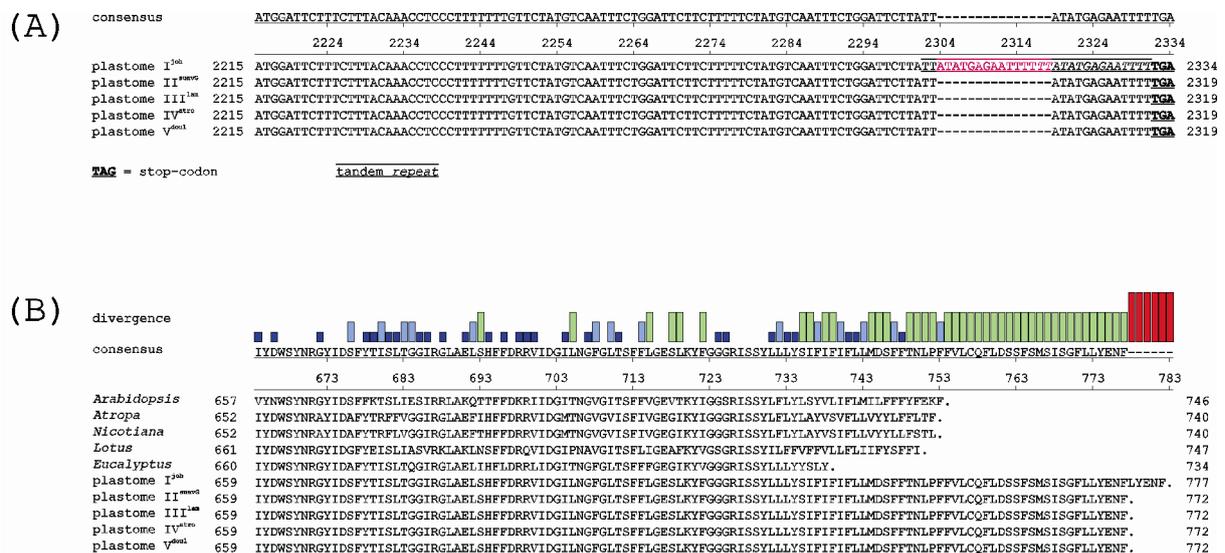
plastome I = I<sup>loh</sup>, plastome II = II<sup>suavG</sup>, plastome III = III<sup>lam</sup>, plastome IV = IV<sup>atro</sup>, plastome V = V<sup>doul</sup>

### 3.2.3.2. Genes with length polymorphisms

Protein coding genes with length polymorphism are of special interest as potential candidates for PGI. However, most of the differences described reside in polypeptide regions known to be highly variable in plastid chromosomes in general, often at the very C-terminus of a polypeptide. Therefore, it is unlikely that they are of functional and/or evolutionary relevance. Nevertheless, they will be presented in more detail in the following chapters.

Length polymorphisms found in genes between the *Oenothera* plastomes are exclusively in protein-coding genes. They can be grouped into three categories: Loci with reading frame shifts (*ndhD*, *rpl22* and *rps18*), loci without reading frame shifts (*ycf1*, *ycf2*, *accD*, *clpP*, and *ndhF*) and, as third class, *atpA* and *psbB* from which gene products in plastome III and plastomes IV and V, respectively, are known to differ in electrophoretic mobility. These changes are independent of the genotype, with which the plastid types are associated with (Herrmann *et al.*, 1980).

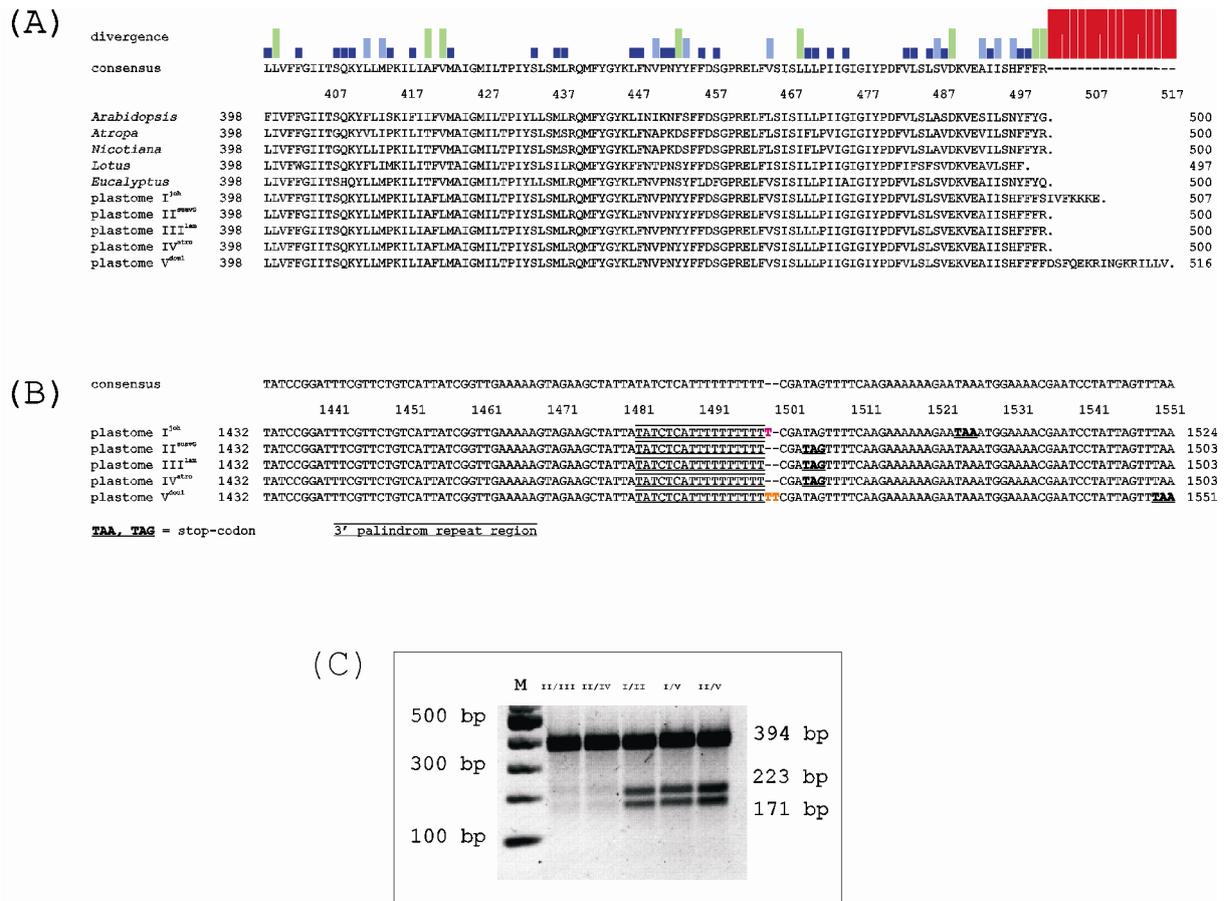
#### 3.2.3.2.1. Insertions in *ndhF* and *clpP* without reading frame change



**Figure 22.** Sequence alignment of the 3' *ndhF* nucleotide sequence (panel A) and translated gene product (panel B) in *Oenothera* plastomes with no frame changing insertions.

Sequence alignment of the *ndhF* 3' region in *Oenothera* and reference plastomes is shown in Figure 22. In plastome I<sup>oh</sup> a 15 bp tandem repeat (marked in red), resulting from a duplication of the nucleotide positions 2304 - 2318 and not present in the other plastomes, does not

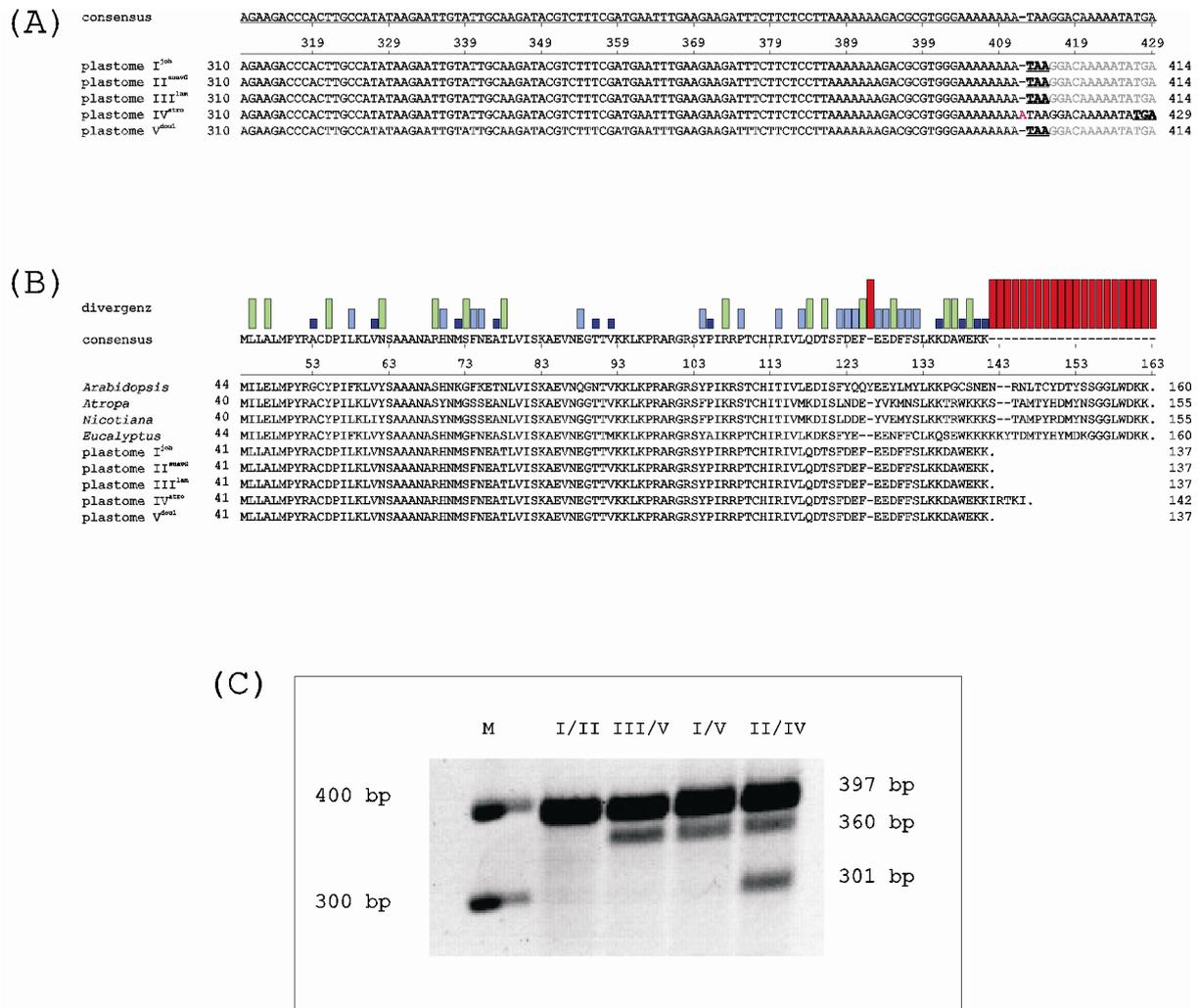


3.2.3.2.2. *NdhD*, *rpl22* and *rps18* contain frame changing insertions

**Figure 24.** Sequence alignment of the *NdhD* C-terminus (panel A) and the frameshift causing sequence variety (panel B). Insertions shown in panel B were confirmed by a Nuclease S1 digest. PCR products with the polymorphism at position 223 and a size of 392 bp ( $II^{suavG}$ ,  $III^{lam}$ ,  $IV^{atro}$ ), 393 bp ( $I^{oh}$ ) or 394 bp ( $V^{doul}$ ) were mixed. Since PCR products of  $II^{suavG}$ ,  $III^{lam}$ ,  $IV^{atro}$  are identical, S1 Nuclease cleavage was not detected (lanes II/III and II/IV). Lanes I/II, I/V and II/V confirm the polymorphisms. Nuclease S1 digests in fragment mixtures detect the expected cleavage products of 223 bp and 171 bp for PCR products of  $I^{oh} + II^{suavG}$  (lane I/II),  $I^{oh} + V^{doul}$  (lane I/V) and  $II^{suavG} + V^{doul}$  (lane II/V), respectively (panel C).

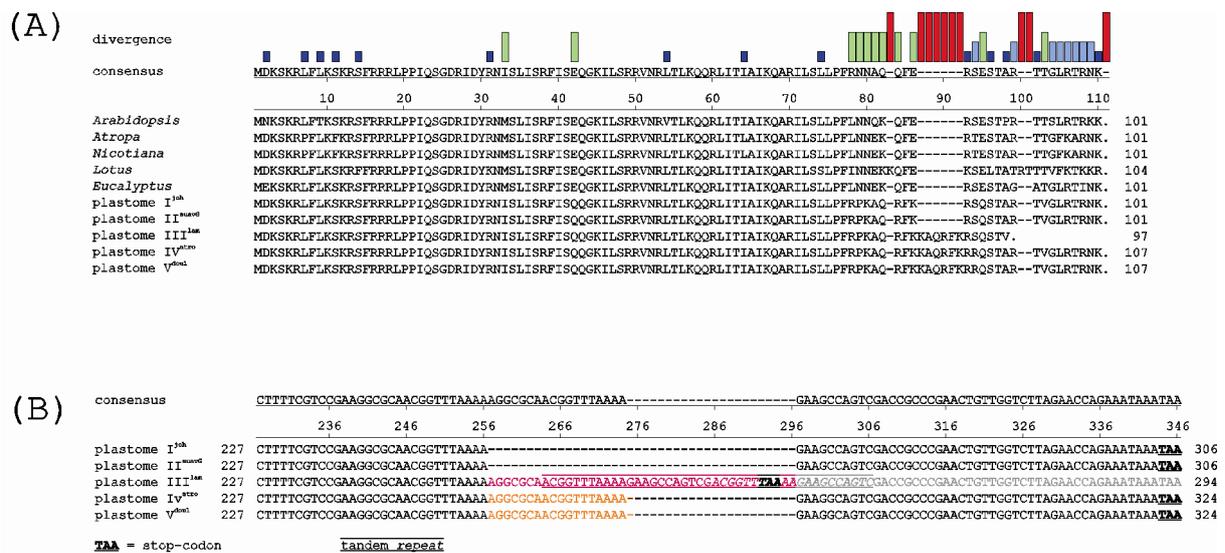
The *NdhD* protein is a 500 amino acid subunit of the thylakoid-located NAD(P)H dehydrogenase, and of equivalent size in plastomes  $II^{suavG}$ ,  $III^{lam}$  and  $IV^{atro}$  as well as in *Arabidopsis*, *Atropa*, *Nicotiana* and *Eucalyptus*. The corresponding *Lotus* polypeptide is predicted to be 3 amino acid residues shorter (Figure 24, panel A). Additional T residues, one in plastome  $I^{oh}$  at position 1498 and two in plastome  $V^{doul}$  at positions 1498/1499 in a T-rich stretch, generate an extension of 7 and 16 amino acid residues, respectively. The region is part of a palindromic repeat (Figure 24, panel B). To confirm sequencing results in the polyT

stretch, PCR products of the plastome regions including the insertions were generated with primers P11for and ndhDint/2. Appropriate templates with sequence polymorphism were mixed, denatured, reannealed, and cut at mismatches with Surveyor Nuclease S1 (Transgenomic, Elancourt, France). Nuclease S1 specifically cuts inaccurate base pairing of any kind. The pattern observed confirmed the polymorphisms detected by sequence data (Figure 24, panel C).



**Figure 25.** Sequence alignment of the Rpl22 C-terminal end (panel A) and the frameshift causing sequence variety (panel B). The insertion could be confirmed by a Nuclease S1 digest. PCR products of plastome I<sup>loh</sup>, II<sup>suavG</sup>, III<sup>lam</sup>, V<sup>doul</sup> (397 bp) or IV<sup>atro</sup> (396 bp), the latter carrying the frame changing insertion at position 301, were mixed and digested with Nucleases S1. Lane I/II confirms identical sequences of I<sup>loh</sup> and II<sup>suavG</sup> PCR products. Lanes III/V and I/V serve as internal controls to validate the method. They confirm an SNP at position 360 in PCR products of IV<sup>atro</sup> and V<sup>doul</sup>, relative to I<sup>loh</sup>, II<sup>suavG</sup> and III<sup>lam</sup>. Lane II/IV confirms the frame changing inversion since the expected band of 301 bp is present in a digested mixture of II<sup>suavG</sup> and IV<sup>atro</sup> PCR products (panel C).

In *rpl22*, an additional A residue at position 411 in plastome IV<sup>atro</sup> upstream of the virtual stop codon causes an elongation by four amino acid residues (Figure 25, panels A and B). The C-terminal part of *rpl22* is not conserved between the *Oenothera* plastomes and the chosen reference plastomes (Figure 25, panel B). In the plastid chromosome of *Lotus* *rpl22* is missing. Again, the sequence divergence has been confirmed by Nuclease S1 mapping using primers P38for and P7rev. PCR assays were performed as described above. The resulting polymorphism confirmed the sequence divergence between the plastomes (Figure 25, panel C).



**Figure 26.** Sequence alignment of the Rps18 C-terminus (panel A) and the frameshift causing sequence variety (panel B). Insertions could be confirmed by a SSLP (panel C). Details are given in the text.

*Rps18* is a ribosomal protein of 101 amino acid residues, as in *Arabidopsis*, *Atropa*, *Nicotiana* and *Eucalyptus*. The highest degree of sequence divergence between plants is found in the C-terminal part (Figure 26, panel A). A 40 bp insertion caused by a tandem repeat in the 3' region of *rps18* in plastome III introduces a C-terminal TAA stop codon that leads to 9

different C-terminal residues and a predicted polypeptide that is four residues shorter than in the reference plastomes (Figure 26, panel A and B). A shorter insertion of 18 bp at the same site in plastomes IV<sup>atro</sup> and V<sup>doul</sup> generates six additional amino acid residues as compared to plastomes I<sup>joh</sup> and II<sup>suavG</sup>. The larger insertions in *rps18* of plastome III<sup>lam</sup> (40 bp) and plastomes IV<sup>atro</sup> and V<sup>doul</sup> (18 bp) compared to plastomes I<sup>joh</sup> and II<sup>suavG</sup> were directly analyzed as length polymorphic PCR products in 3% agarose gels using the primer pair *rps18for* and *rps18revM* (Figure 26, panel C).

### 3.2.3.2.3. Sequence divergence of *ycf1* and *ycf2*

Analysis of *ycf1* and *ycf2* is sophisticated, since these genes are only moderately conserved in plastid genomes in general and function or functional domains are unknown (Drescher *et al.*, 2000). The overall conservation of *ycf1* and *ycf2* is low between the reference plastomes. However, conservation of the two genes among the five *Oenothera* plastomes is comparable to that between some reference plastomes, but in general conservation is somewhat higher (Tables 17 and 18).

**Table 17.** Sequence identity of *ycf1* in comparison with reference plastomes

identity [%]	<i>Atropa</i>	<i>Nicotiana</i>	<i>Lotus</i>	<i>Eucalyptus</i>	I <sup>joh</sup>	II <sup>suavG</sup>	III <sup>lam</sup>	IV <sup>atro</sup>	V <sup>doul</sup>
<i>Arabidopsis</i>	58.5	59.1	50.9	61.4	35.6	35.6	35.6	35.7	35.6
<i>Atropa</i>	---	92.5	52.7	64.2	36.3	36.3	36.5	36.6	36.5
<i>Nicotiana</i>	---	---	53.1	64.4	36.3	36.4	36.5	36.7	36.5
<i>Lotus</i>	---	---	---	50.7	33.0	33.1	33.0	33.3	33.0
<i>Eucalyptus</i>	---	---	---	---	38.7	38.7	38.8	39.0	38.8
I <sup>joh</sup>	---	---	---	---	---	99.6	98.8	98.2	98.8
II <sup>suavG</sup>	---	---	---	---	---	---	98.8	98.3	98.8
III <sup>lam</sup>	---	---	---	---	---	---	---	98.2	100.0
IV <sup>atro</sup>	---	---	---	---	---	---	---	---	98.2

**Table 18.** Sequence identity of *ycf2* in comparison with reference plastomes

identity [%]	<i>Atropa</i>	<i>Nicotiana</i>	<i>Lotus</i>	<i>Eucalyptus</i>	I <sup>joh</sup>	II <sup>suavG</sup>	III <sup>lam</sup>	IV <sup>atro</sup>	V <sup>doul</sup>
<i>Arabidopsis</i>	89.9	90.0	86.4	91.1	71.3	71.3	71.1	71.5	71.4
<i>Atropa</i>	---	99.2	87.8	92.4	73.0	73.0	72.9	73.1	73.0
<i>Nicotiana</i>	---	---	88.1	92.6	73.1	73.1	72.9	73.2	73.1
<i>Lotus</i>	---	---	---	89.3	69.8	69.8	69.6	70.0	69.9
<i>Eucalyptus</i>	---	---	---	---	75.0	75.0	74.8	75.2	75.1
I <sup>joh</sup>	---	---	---	---	---	99.9	99.3	99.0	98.9
II <sup>suavG</sup>	---	---	---	---	---	---	99.4	99.0	99.0
III <sup>lam</sup>	---	---	---	---	---	---	---	99.0	98.8
IV <sup>atro</sup>	---	---	---	---	---	---	---	---	99.7

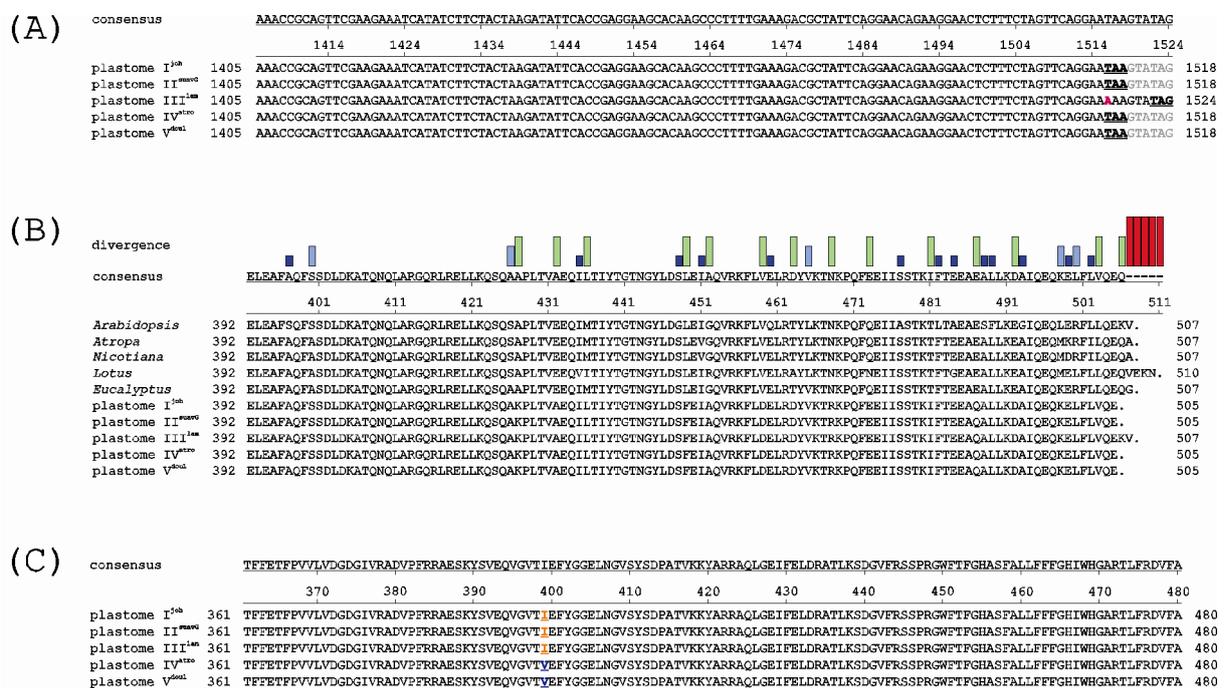
3.2.3.2.4. The 5' end of *accD* is highly polymorphic



Figure 27. Sequence alignment of *AccD* including sequences of eight reference species and all five *Oenothera* plastomes.

*AccD* encodes one of the four subunits of the plastid enzyme acetyl-CoA carboxylase (ACCase), which catalyses the first step of *de novo* fatty acid synthesis in the plastid carboxylating acetyl-CoA to malonyl-CoA. A knockout of *accD* in *Nicotiana* is lethal and affects leaf morphology (Kode *et al.*, 2005). While the first half and the very C-terminus of the deduced protein sequence is quite diverged among the taxa investigated in this study, the second half of the protein, which houses the carboxyl transferase domain, is well conserved, not only among all five *Oenothera* plastomes and the *Munzia* plastome of *Oe. villaricae* strain berteriana Schwemmler (accession number EU255778), but also among the reference plastomes of *Atropa*, *Nicotiana*, *Arabidopsis*, *Lotus*, *Glycine*, *Pisum* and *Eucalyptus* (Figure 27). The diverged region presumably is not directly involved in catalytic function. For that reason it seems unlikely that the substantial differences among AccD polypeptides contribute notably to interspecific plastome-genome incompatibility. However, it is conceivable that the single amino acid exchanges between the five plastomes found in the functional domain of the conserved AccD region affect the interaction with the nuclear partner subunits.

### 3.2.3.2.5. Alignments of *AtpA* and *PsbB* - proteins with mobility shifts



**Figure 28.** Sequence alignment of the 3' *atpA* region (panel A and B) and polymorphism detected in *psbB* (panel C) presumably giving rise to known variance of polypeptides in *Oenothera* (Herrmann *et al.*, 1980).

The TAA stop codon of *atpA* in plastome III<sup>lam</sup> has been changed by a T-to-A transversion (Figure 28, panel A) resulting in an extension of two amino acid residues including lysine as compared to all other plastomes (Figure 28, panel B). It is likely that these additional residues account for the somewhat slower electrophoretic mobility that has been noted for the thylakoid located ATP synthase subunit  $\alpha$  of that plastome independent of the nuclear genome type (Herrmann *et al.*, 1980). It is relevant to note that the 3' region of *atpA* is also variant among the reference plastomes of *Lotus*, *Arabidopsis*, *Nicotiana*, *Atropa* and *Eucalyptus*.

CP47, an inner antenna chlorophyll protein of photosystem II encoded by *psbB*, is variant between plastomes IV<sup>atro</sup> and V<sup>doul</sup> compared to plastomes I<sup>boh</sup> – III<sup>lam</sup> (Figure 28, panel C). The genes contain a single G (IV<sup>atro</sup>, V<sup>doul</sup>) to A (I<sup>boh</sup>, II<sup>suavG</sup>, III<sup>lam</sup>) conversion at position 1195, which leads to a conservative change from valine to isoleucine, consistent with previous work (CP47 of *Oe. argillicola* strain douthat 1, accession no. X55899, CP47 of *Oe. elata* subsp. *hookeri* strain hookeri de Vries, accession no. X55900). It remains to be verified whether the mutation accounts for the electrophoretic mobility shift (Herrmann *et al.*, 1980).

#### 3.2.4. Differences in intergenic regions between plastomes I<sup>boh</sup> – V<sup>doul</sup>

For investigation of PGI not only coding regions but also regulatory elements have to be taken into consideration. Promoters, terminators, processing signals or ribosome binding sites are located within the intergenic sequence intervals and are less conserved among plastomes. From altogether 114 regions, 76 are variable and only 38 identical. 24.1 indels and 5.4 substitutions have been detected on average for each intergenic region. Taken together, 2,743 indels and 616 substitutions were detected for all 114 intergenic regions. This contrasts observations made for coding regions, in which only few indels have been found.

A functional analysis of the non-coding regions is of course more difficult than that of coding regions, since conservation of regulatory elements is generally lower. Bioinformatic identification of chloroplast mRNA processing signals or terminators is nearly impossible. However, consensus strength of plastid promoters as well as ribosomal binding sites is high enough and allow a comparative analysis.

#### 3.2.4.1. *Differences in promoter regions*

To search for potential promoters of the bacterial type (-10 and -35 boxes) and phage type polymerases, the region approximately 600 bp upstream of each gene was searched for consensus sequences (Chapter 2.2.6.2.3). The analysis confirmed various sites previously reported from other species (Hess and Börner, 1999; Kanamaru and Tanaka, 2004; Shiina *et al.*, 2005; Liere and Börner, 2006; Swiatecka-Hagenbruch *et al.*, 2007) also in the *Oenothera* plastomes. Putative sigma factor binding sites were found in 75 genes, and 69 contained type Ib-like NEP promoters. In total, at least one polymerase binding site could be deduced for 88 genes. Predictions for both promoter types were found for 56 genes. Among the predicted PEP and NEP sites, 39 and 27 promoters had at least one difference in the binding site between one of the five plastomes, respectively. A relatively wide spectrum of changes, including additional or lacking binding sites, single point mutations or spacing differences were found between the predicted promoters. The functional relevance of these changes, such as binding strength or transcriptional start sites remain to be verified and will be discussed later (Chapter 3.4.1.4).

#### 3.2.4.2. *Differences in ribosomal binding sites*

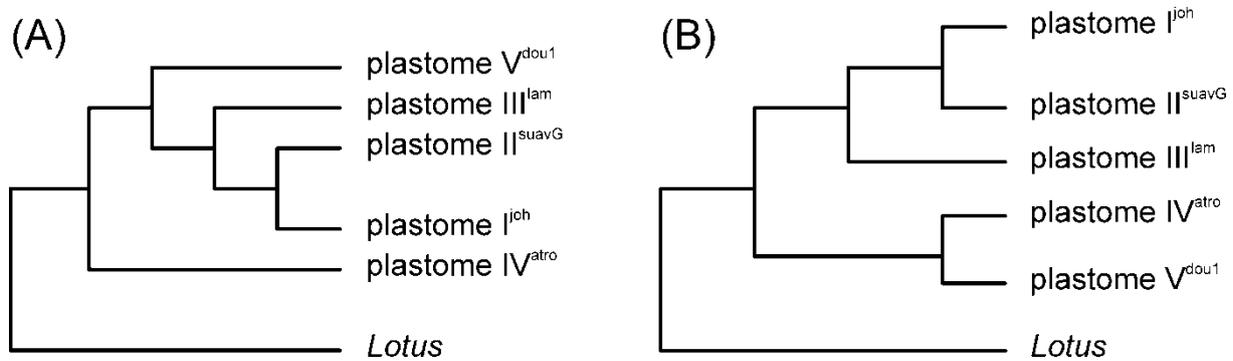
To evaluate a potential contribution of translation initiation to PGI, the binding capacity of ribosomal binding sites was studied. Such binding sites were detected by their match to the 3' end of the 16S rRNA in 30 genes within the first 23 bp upstream of the initial start codon. The Shine-Dalgarno sequences show no significant differences among the plastomes, excluding the possibility that translational efficiency contributes to PGI.

### 3.3. **Evolutionary analysis of the five basic plastomes**

The complete sequence of the five basic *Oenothera* plastomes offers the unique opportunity to correlate sequence diversification with functional and evolutionary aspects. The evolutionary succession of the five plastomes was analyzed by appropriate algorithms generating phylogenetic trees, together with a rough estimation of the possible divergence times between the five plastomes. Furthermore, closer inspections of the compatibility chart (Figure 4) allowed to deduce for the first time four genetically different types of PGI. Different types of PGI build hybridization barriers of different strength on which natural selection acts. Estimation of selection pressure, again for the first time, suggested that

plastomes underlie notable selection forces and indicate a distinct set of plastome encoded genes contributing to speciation.

### 3.3.1. Calculation of phylogenetic trees



**Figure 29.** Phylogenetic trees of the five *Oenothera* plastomes. Different tree topologies appear depending on the method. NJ and MP place plastomes I<sup>joh</sup>, II<sup>suavG</sup>, III<sup>lam</sup> and V<sup>dou1</sup> in one clade and plastome IV<sup>atro</sup> in a separate branch (panel A), whereas ML puts I<sup>joh</sup>, II<sup>suavG</sup> and III<sup>lam</sup> vs. IV<sup>atro</sup> and V<sup>dou1</sup> as two separate clades (panel B).

To evaluate the evolutionary succession of the five plastome types, tree calculation was performed using whole plastome sequences. Classical genetic data describing the evolutionary succession of the five plastome types are already available (Stubbe, 1963a; 1964; Stubbe and Steiner, 1999), but molecular investigations were so far limited to a small set of plastome sequences (Hornung *et al.*, 1996; Sears *et al.*, 1996; Levin *et al.*, 2003; Levin *et al.*, 2004). To generate phylogenetic trees all three methods, Neighbor-Joining (NJ), Maximum Likelihood (ML) and Maximum Parsimony (MP) were used. They were calculated for both, the 47 variable genes as well as the 76 variable intergenic regions. Trees were rooted by orthologous genes or sequence intervals, respectively, with the *Lotus japonicus* plastome sequence (Kato *et al.*, 2000) as an outgroup. Plastomes I<sup>joh</sup>, II<sup>suavG</sup> and III<sup>lam</sup> were always grouped in one clade with plastomes I<sup>joh</sup> and II<sup>suavG</sup> being close relatives (bootstrap values 100% for NJ and ML) (Figure 29). A common ancestor for clade I<sup>joh</sup> - III<sup>lam</sup> and a clade consisting of IV<sup>atro</sup> and V<sup>dou1</sup> (bootstrap 99%) was supported by ML. Using MP one maximal parsimonious tree was generated, placing plastome V<sup>dou1</sup> in one clade with plastomes I<sup>joh</sup> - III<sup>lam</sup> while plastome IV<sup>atro</sup> is located in a separate branch. This tree is supported by both, NJ analysis of the concatenated alignments (bootstrap 96.7% for separation of plastomes V<sup>dou1</sup> and IV<sup>atro</sup>) as well as by a species/consensus tree of the 47 individual NJ gene trees. Only 28.9% result in a branching pattern identical to the ML topology of the consensus tree, while

40% of all gene trees support a separate branch for plastome IV<sup>atro</sup>. For classical genetic reasons (Stubbe and Steiner, 1999) it is proposed that the latter tree topology reflects the phylogenetic relationship between the five *Oenothera* plastomes (Chapter 4.4.2).

### 3.3.2. Estimation of evolutionary distances between the plastomes

A long standing and unsolved question in *Oenothera* taxonomy is the divergence time of species within subsection *Oenothera*. Since molecular data are largely missing and fossil records not available, the only statements found in the *Oenothera* literature is a guess of Cleland (1972, pp. 299 - 302) and calculations based on nuclear allozyme variation made by Levy and Levin (1975). Cleland assumes a possible invasion of the North American continent by the common ancestor of subsection *Oenothera* from Middle America in several waves, possibly 70,000 years ago, at the end of the Pleistocene. However, Cleland does not rule out the possibility that the subsection is older. Levy and Levin (1975) estimate the age of subsection presumably to 1 mya. These statements were never examined.

**Table 19.** Evolutionary distances in years calculated for the five *Oenothera* plastomes

plastome	I <sup>joh</sup>	II <sup>suavG</sup>	III <sup>lam</sup>	IV <sup>atro</sup>
II <sup>suavG</sup>	83,195	---	---	---
III <sup>lam</sup>	471,436	415,973	---	---
IV <sup>atro</sup>	831,947	804,215	831,947	---
V <sup>doul</sup>	998,336	998,336	942,871	693,289

For a molecular estimation of the divergence time between the five basic plastomes in the section *Oenothera*, average substitution rates of protein coding genes between the five plastomes were investigated. Average Ks values between the five plastomes varied over more than one order of magnitude, from  $3 \times 10^{-4}$  for plastomes I<sup>joh</sup> and II<sup>suavG</sup> up to  $3.6 \times 10^{-3}$  for plastomes I<sup>joh</sup>/II<sup>suavG</sup> and V<sup>doul</sup>. Divergence for plastomes ranged generally between 416,000 (II<sup>suavG</sup> and III<sup>lam</sup>) and 830,000 years, if a calibration derived for dicotyledoneous chloroplast genes (Chaw *et al.*, 2004) was applied. The divergence of plastomes I<sup>joh</sup> and II<sup>suavG</sup> is estimated to be as recent as 83,000 years ago, while the most distant pair (I<sup>joh</sup> and V<sup>doul</sup>) has diverged approximately 1 mya (Table 19). These data differ from Cleland's assumption that the subsection *Oenothera* has arisen at the end of the Pleistocene and do not reconstitute Levy and Levin's time frame of divergence (see Chapter 4.4.1).

### 3.3.3. Four basic types of PGI determine the strength of hybridization barriers

The five basic plastomes in *Oenothera* differ in their evolutionary diversification, as shown in tree or divergence time calculations. The functional consequence can be easily recognized by the compatibility chart (Figure 4). In *Oenothera*, the C genome is probably the most ancient genotype of the subsection and exerts a dominant negative effect on A and B genomes. Plastome I, considered to be the most advanced plastome, is compatible exclusively with its natural AA background. Studies on *Rhododendron* uncover a similar pattern (Noguchi, 1932). This suggests that the strength of post-zygotic barriers and the ability to produce PGI tend to be a function of “cytoplasmic” divergence, a finding also noted by Levin (2003). This observation builds the base for four genetically different classes of PGI.

The compatibility chart of *Oenothera* (Figure 4) presents various phenotypically and genetically different PGIs, which suggest strongly that plastome dependent hybridization barriers can have different strengths and reasons, depending on the genetic design of the incompatibility. Closer inspection of the plastome-genome combinations underlying these incompatibilities uncover that PGIs can be grouped into four classes. The impact of each of them on hybridization barriers is different. In principle, these classes cover all CIs and are not restricted to plants or plastids, but only the comprehensive data set of *Oenothera* allows general considerations.

A first PGI type, designated *dominant PGI*, is found in F1 nuclear hybrids with the plastome of one parent (F1-PGI). In this instance, a single copy of a compatible genome in heterozygous constitution cannot prevent PGI. In *Oenothera* AB-I, AC-I, AC-V, BC-III and BC-V fall into this class (Figure 4). Most of the examples listed in Table 2 are dominant PGIs as well, since they appear in the F1. Dominant PGIs build strong hybridization barriers, immediately affecting the F1 generation.

A second class is represented by so-called *recessive PGIs*. In *Oenothera*, the combinations AA-III, AA-IV, BB-II and CC-II belong to this type. To illustrate their genetic design, the incompatible combination AA-III will be considered. Recessive incompatibilities are homozygous in their nuclear background (here: AA), carry a foreign plastome (as plastome III with that genotype), and, different from the previous class, can be cured as heterozygotes with the genome of the plastome donor. The plastid harbouring plastome III, which is native to genotype BB, is also green in combination with AB. The evolutionary consequence of

recessive PGIs is hybrid breakdown. The compatible AB-III F1 hybrid segregates 25% AA-III (incompatible), 50% AB-III (green) and 25% BB-III (green) individuals in the F2 generation. This hybridization barrier is generally weaker than that of dominant incompatibilities. Recessive PGIs have been also observed in *Pelargonium* and *Trifolium* as breakdown of viable F1 generations (Smith, 1915) or F1 backcrosses (Smith, 1915; Meredith *et al.*, 1995). Their occurrence is probably underestimated because of the lack of hybrid variegation in higher generations. Usually, sorting-out is completed in all flower organs during life cycle in F1 and therefore hybrid variegation is not transmitted to successive generations (Kirk and Tilney-Bassett, 1978; Birky, 2001). Since hybrid variegation is the only reliable way to detect PGI (Chapter 1.4), recessive PGI may be overlooked frequently.

The third category includes *co-dominant PGIs*. Examples in the compatibility scheme are AA-V, BB-I, BB-V, CC-I and CC-III. They can be caused by a single, co-dominant nuclear factor, or suffer from at least two genome conflicts, a dominant and a recessive one. Combination BB-I may illustrate a co-dominant incompatibility (Figure 4). In general, BB-I suffers from a dominant maladaptive factor between the B genome and plastome I that already appears in an AB-I background. Replacement of the A genome by a further B genome, enhances the relatively weak AB-I phenotype, resulting in the strong incompatibility BB-I. Genetically, there are two possible explanations: First, BB-I can be caused by two factors, a dominant one, already responsible for the AB-I phenotype, plus a recessive one, which becomes notable in the homozygous BB background. Both factors, as sole responsible for dominant or recessive PGIs, respectively, together cause the strong BB-I phenotype. The second explanation is a co-dominant inheritance of a single determinant that is possibly monogenic. In this model, the incompatible BB-I phenotype is caused by a dose effect. One B genome in the AB-I background displays a weaker phenotype than two in the combination BB-I. Both models assume that AB-I and BB-I share, at least fractional, identical molecular reasons. Co-dominant incompatibilities can play a limited role in evolution, as they enforce an already existing hybridization barrier in F2. In the chosen example the incompatible hybrid AB-I, when selfed, splits into AA-I (green), AB-I (incompatible) and BB-I (strongly incompatible) in F2.

The fourth case, *chimeric PGI*, is characterized by a heterozygous nucleus and a plastome that has an evolutionary history different from both haploid genomes. An example is the combination BC-I. Plastome I is naturally associated only with A genomes and its

combination with BB or CC is disharmonic. Consequently, if BB-I and CC-I are crossed, an incompatible BC-I offspring is not surprising. It is very likely that chimeric PGIs are of polygenic origin. Other incompatibilities of that kind in the compatibility chart are AB-V and, in some respect, AC-III and BC-II<sup>3)</sup>. The significance of *chimeric PGIs* to generate hybridization barriers is limited. Their occurrence in nature is improbable, since both parental lines have to be already incompatible.

### 3.3.4. Selection pressure on the *Oenothera* plastomes

PGI leads to hybridization barriers of different strengths, but which selection forces produce them? Is PGI a consequence of selection at all, or just a phenomenon of reinforcement? Can the plastome with its limited set of genes contribute to speciation? These questions can easily be tested by measuring selection forces on *Oenothera* plastome sequences. Genes causative for speciation have been suggested to be under positive selection for a limited period (Gillespie, 1991). To investigate whether there is selection at all on plastome sequences and also to derive candidate genes for incompatibility factors, ratios of non-synonymous ( $K_a$ ) versus synonymous ( $K_s$ ) substitutions were determined for variable genes using alignments of their entire coding sequences.

Out of 233 pairwise comparisons, for which the method was applicable, 33 (14.1%) exhibited elevated  $K_a/K_s$  rates above 1.0. However, an excess of non-synonymous substitutions was not equally distributed between the pairs under study but clustered predominantly to five genes, *ycf1*, *ycf2*, *accD*, *clpP* and *ndhA*. For almost all plastome pairs these genes displayed fast evolution and  $\omega$  values higher than 1, indicative of positive selection. The first three genes contain extended repetitive regions that are only weakly conserved in other species. It is therefore unclear whether divergence of these regions is functionally relevant or whether the increase of non-synonymous substitution rates is simply the result of observed high variability. The highest rates ( $\omega = 4.1$ ) calculated for *clpP* seem to differentiate plastomes I<sup>loh</sup> and II<sup>suavG</sup> versus III<sup>lam</sup>, while maximal rates for *accD* ( $2.2 < \omega < 4$ ) were observed between clade I<sup>loh</sup>/II<sup>suavG</sup> and IV<sup>atro</sup> and V<sup>doul</sup>, respectively. A similar, though less pronounced,

<sup>3)</sup> The combinations AC-III and BC-II are chimerical incompatibles only if they originate from crosses between their incompatible parents AA-III x CC-III and BB-II x CC-II, respectively. As an exception, the genetics of permanent translocation heterozygosity in *Oenothera* (Chapter 1.7) allows the assembly of AC-III or BC-II from compatible parents (e.g. AB-III x CC-V or BA-II x CC-V). In this case, AC-III or BC-II represent dominant incompatibles.

**Table 20.** Average Ka/Ks values calculated from the five *Oenothera* plastomes in pairwise comparison

<b>Ka/Ks (Ka; Ks)</b>	<b>plastome I<sup>joH</sup></b>	<b>plastome II<sup>suavG</sup></b>	<b>plastome III<sup>lam</sup></b>	<b>plastome IV<sup>atro</sup></b>
<b>plastome II<sup>suavG</sup></b>	0.6266 (0.0002; 0.0003)	---	---	---
<b>plastome III<sup>lam</sup></b>	0.5107 (0.0009; 0.0017)	0.5945 (0.0009; 0.0015)	---	---
<b>plastome IV<sup>atro</sup></b>	0.4176 (0.0013; 0.0030)	0.4567 (0.0013; 0.0029)	0.4446 (0.0013; 0.0030)	---
<b>plastome V<sup>dou1</sup></b>	0.4608 (0.0017; 0.0036)	0.4961 (0.0018; 0.0036)	0.4309 (0.0015; 0.0034)	0.2371 (0.0006; 0.0025)

observation has been made for *ndhA*. Together, elevated rates in the five genes comprise 26 of 33 pairs with positive selection. For *ccsA*, *petD*, and *matK* each, only one pairwise  $\omega$  value exceeded the chosen criterion for positive selection. However, it was noted that several pairs of *matK* and *ccsA* indicate elevated evolutionary rates well above the median  $\omega = 0.405$ . Clustering of elevated rates has also been observed for *ndhD*, *rps18* and *rps3*. Compared to a mean of 0.157 deduced for pairwise comparisons of 30 angiosperm chloroplast genomes that was widely constant in most branches of the mono- and eudicotyledonous plants investigated, mean Ka/Ks ratios between the five plastomes are relatively high. Most ratios varied between 0.4 and 0.5 with a minimum of 0.24 between plastomes IV<sup>atro</sup> and V<sup>doul</sup> and a maximum of 0.63 for plastomes I<sup>joh</sup> and II<sup>suavG</sup> (Table 20). The unusually high mean Ka/Ks ratios between most of the five plastomes and the relatively large number of genes for the first time indicate positive selection or fast evolutionary rates and are consistent with a significant contribution of plastomes to speciation.

### 3.4. Investigation of plastome-genome incompatibility

The complete sequences of the five basic plastomes, combined with the comprehensive molecular and ecological knowledge about the photosynthetic process, provide a solid framework to investigate PGI in *Oenothera*. In the following chapters bioinformatic and molecular attempts are presented to evaluate differences found among the five basic plastomes, in respect to functional and evolutionary aspects. Some factors identified by bioinformatic analysis are complementary with molecular and phenotypic data. This approach provided strong evidence that PGI reflects predominantly a regulatory phenomenon. In a distinct case, the interspecific hybrid AB-I, a single locus contributes substantially to PGI. A deletion in the divergently operating promoter region between *psbB* and *clpP* can explain the AB-I phenotype. Down-regulation of *psbB* mRNA leading to reduced levels of CP47 chlorophyll *a* apoprotein results in a disturbance of photosystem II.

#### 3.4.1. Bioinformatic investigation

Bioinformatic investigations were used to estimate the significance of variations detected among the five plastomes in functional terms. Three possible causes for PGI were tested: coding regions, regulatory regions and RNA editing. Editing is of special interest, since it was identified as the primary reason for PGI in the *Atropa/Nicotiana* cybrid (Schmitz-Linneweber *et al.*, 2005).

#### 3.4.1.1. Search for candidate protein coding loci involved in PGI

Two attempts were made to estimate the impact of plastome-specific differences in genes. First, in the case of single amino acid exchanges, biochemical properties of a substitution and its location in a functional domain were considered (Table 21). Second, in proteins with length polymorphism, direct alignment for single amino acids in the insertion/deletion is not possible since alignment partners are missing. Therefore, only the presence of the length variation in a functional domain was tested. All variations were compared to the distinct plastome-genome combinations in the compatibility chart (Figure 4). The findings are described in the two following chapters. Collectively, the data indicate that PGI reflects predominantly a regulatory, rather than a mere structural phenomenon.

#### 3.4.1.2. Investigation of single amino acid exchanges

A putative functional impact of non-synonymous sites in polypeptides, *i.e.* degree of conservation and differences of biochemical properties, between *Oenothera* and those of 30 reference species was estimated as described in Material and Methods (Chapter 2.2.6.2.2). Starting from 388 non-synonymous replacements (excluding *ycf1*, *ycf2* and the highly variable *accD* N-terminus; Chapters 3.2.3.2.3 and 3.2.3.2.4), 35 sites in 19 polypeptides were identified that showed a significant difference between the distribution of biochemical properties within *Oenothera* and the reference species ( $p < 0.05$ ). 25 of the significant sites were located within a known PFAM domain (Table 21), and therefore represent putative factors causing PGI. It is noteworthy that almost all genes with elevated  $K_a/K_s$  rates (Chapter 3.3.4) are present in the set of 19 proteins with significantly differing amino acid exchanges.

#### 3.4.1.3. Investigation of genes showing length polymorphism

Protein coding genes with length polymorphisms are of intrinsic interest for PGI. Ten loci predictably should generate variant polypeptides between the evening primrose plastomes. These are represented by *ndhD*, *rpl22* and *rps18* with reading frame shifts, *ycf1*, *ycf2*, *accD*, *clpP*, and *ndhF* without reading frame shifts, and two genes, *atpA* and *psbB*, from which the electrophoretic mobility is known to differ in plastomes III<sup>lam</sup> and plastomes IV<sup>atro</sup> and V<sup>doul</sup>, respectively, independent of the nuclear genome the plastomes are associated with (Herrmann *et al.*, 1980). Details of the variations were presented earlier (Chapter 3.2.3.2). All loci were compared mutually, and aligned with the plastome/genome compatibility scheme of the

**Table 21.** Estimation of possible determinants for plastome-genome incompatibility caused by single amino acid exchanges

gene	aa <sup>1)</sup> <i>Oenothera</i> (I; II; III; IV; V)	consensus in reference species	GD <sup>2)</sup>	p-value	aa <sup>1)</sup> exchange in a functional domain	transmembrane domain	possible incompatibility in
<i>accD</i>	H;H;P;P;P	P	76	1,33E-02	+	-	BB-I; BB-II
	R;R;R;G;G	G	125	3,10E-03	+	-	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
	W;W;W;R;R	S or N	174; 85	3,69E-02	N/A	-	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
	P;P;P;P;S	P	73	2,64E-06	+	-	AA-V; AB-V; BB-V
	N;N;N;N;L	N	153	6,16E-05	+	-	AA-V; AB-V; BB-V
	S;S;S;S;L	Q or K	68; 107	1,75E-04	+	-	AA-V; AB-V; BB-V
	F;F;F;F;L	L	22	5,31E-04	+	-	AA-V; AB-V; BB-V
<i>atpA</i>	S;Y;S;S;S	S or V	55	8,94E-04	+	-	BB-II
<i>atpB</i>	R;R;R;R;G	R	125	4,15E-08	+	-	AA-V; AB-V; BB-V
<i>atpF</i>	G;G;G;G;V	G	109	1,82E-13	+	not affected	AA-V; AB-V; BB-V
<i>ccsA</i>	K;K;N;K;K	K or Q	46	1,57E-02	N/A	-	AA-III; AC-III
<i>clpP</i>	L;L;M;L;L	W	61	3,29E-11	N/A	-	AA-III; AC-III
	G;G;A;A;A	A	60	6,52E-07	-	-	BB-I; BB-II
<i>matK</i>	W;W;W;L;W	L or F	40	6,85E-04	+	-	AA-IV
	R;R;R;W;W	L	102; 61	1,64E-02	+	-	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>ndhA</i>	R;R;G;R;R	R or K	125	1,64E-04	+	not affected	AA-III; AC-III
	S;S;Y;Y;Y	Y	143	1,48E-04	+	not affected	BB-I; BB-II
	T;T;T;I;I	I	89	2,49E-02	N/A	weakly affected	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>ndhB</i>	R;R;R;R;I	R	98	3,65E-04	+	not affected	AA-V; AB-V; BB-V
<i>ndhC</i>	I;I;K;K	K or E	107	5,52E-03	+	not affected	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>ndhD</i>	A;A;A;A;S	A	99	3,29E-11	-	not affected	AA-V; AB-V; BB-V
<i>ndhE</i>	A;A;T;A;A	A	58	3,29E-11	+	weakly affected	AA-III; AC-III
<i>ndhH</i>	R;R;R;I;I	R or K	98	2,85E-02	+	-	AA-V; AB-V; BB-V
<i>petD</i>	A;A;A;S;A	A	99	1,90E-06	+	not affected	AA-IV
	V;W;V;V;V	V	88	2,44E-14	+	not affected	BB-II
	N;K;N;N;N	N	94	8,43E-09	N/A	not affected	BB-II
<i>rpoB</i>	C;C;S;S;S	S	112	2,50E-04	N/A	-	BB-I; BB-II
	R;R;R;G;G	R	125	1,94E-04	+	-	AA-V; AB-V; BB-V

Table 21. (continued)

gene	aa <sup>1)</sup> <i>Oenothera</i> (I; II; III; IV; V)	consensus in reference species	GD <sup>2)</sup>	p-value	aa <sup>1)</sup> exchange in a functional domain	transmembrane domain	possible incompatibility in
<i>rpoC2</i>	R; I; R; R; R	K or T	26; 89	8,70E-03	+	not present	BB-II
	H;H;H;P;P	H	76	2,97E-02	+	not present	AA-V; AB-V; BB-V
	H;H;H;H;P	H	76	5,03E-05	N/A	not present	AA-V; AB-V; BB-V
	D;D;D;D;G	D or E	94	1,07E-03	+	not present	AA-V; AB-V; BB-V
<i>rps3</i>	R;R;T;R;R	R or E	66	2,72E-05	N/A	not present	AA-III; AC-III
<i>rps8</i>	H;H;P;H;H	H	76	9,37E-04	+	not present	AA-III; AC-III
<i>rps15</i>	S;S;S;S;I	S, R or N	98	1,98E-03	+	not present	AA-V; AB-V; BB-V

<sup>1)</sup> aa = amino acid      <sup>2)</sup> GD = Grantham Distance

plastome I = I<sup>oh</sup>, plastome II = II<sup>suavG</sup>, plastome III = III<sup>lam</sup>, plastome IV = IV<sup>atro</sup>, plastome V = V<sup>dou1</sup>

subsection (Figure 4). Length polymorphisms in *accD*, *atpA*, *clpP*, *ndhD*, *ndhF*, *rpl22* and *rps18* could evidently not be compared with distinct amino acid residues. However, none of their variant regions, except the altered *ndhD* 3' terminus, is located within a functional domain (Table 22). Computational analysis of transmembrane domains was used also to check each candidate for aberrations in its transmembrane domain architecture. Marginal effects appeared only for components of the NAD(P)H dehydrogenase complex, for *ndhA*, *ndhD* and *ndhE*. Apart from the fact that it is unclear whether the predicted relatively small perturbations alter notably the stability of the transmembrane regions, disruption of *ndh* genes in *Nicotiana* does not result in a pronounced phenotype (Burrows *et al.*, 1998; Kofer *et al.*, 1998). Since almost all of the outlined polypeptide differences reside in the very C-termini that are also known to be quite variable in general, it is unlikely that the outlined changes are of major functional and/or evolutionary relevance.

**Table 22.** Estimation of possible determinants for plastome-genome incompatibility caused by indels<sup>1)</sup>

gene	indel	indel present in plastome					possible incompatibility	functional domain	transmembrane domain
		I <sup>joh</sup>	II <sup>suavG</sup>	III <sup>lam</sup>	IV <sup>atro</sup>	V <sup>dou1</sup>			
<i>accD</i>	N/A	various indels					N/A	-	-
<i>atpA</i>	I <sup>506-507</sup>	-	-	+	-	-	AA-III; AC-III	-	-
<i>clpP</i>	I <sup>17-18</sup>	-	-	-	-	+	AA-V; AB-V; BB-V	-	-
	I <sup>26</sup>	-	-	+	-	-	AA-III; AC-III	-	-
	Δ <sup>222-336</sup>	-	-	-	+	-	AA-IV	-	-
<i>ndhD</i>	I <sup>501-507</sup>	+	-	-	-	-	AB-I	+	weakly affected
	I <sup>501-516</sup>	-	-	-	-	+	AA-V; AB-V; BB-V	+	weakly affected
<i>ndhF</i>	I <sup>778-782</sup>	+	-	-	-	-	AB-I	-	-
<i>rpl22</i>	I <sup>142-146</sup>	-	-	-	+	-	AA-IV	-	-
<i>rps18</i>	I <sup>87-92</sup>	-	-	+	+	+	BB-I; BB-II	-	-
	Δ <sup>99-110</sup>	-	-	+	-	-	AA-III; AC-III	-	-

<sup>1)</sup> *Ycf1* and *ycf2* were excluded, since both loci are highly variable in general and their function and functional domains are unknown (Drescher *et al.*, 2000).

#### 3.4.1.4. Search for candidate loci for PGI in intergenic regions

PGI may be a regulatory phenomenon. To pinpoint potential determinants, the sequences of the five *Oenothera* plastomes were searched for differences in promoter regions as well as for potentially different mRNA editing sites. The differences located in promoter sequences of

the *Oenothera* plastomes were compared with the genetically determined compatibility relationships (Figure 4).

Two types of RNA polymerases operate in plastids of higher plants, the ancient eubacterial type polymerase PEP (-10 and -35 boxes) and a phylogenetically acquired phage-type polymerase (NEP), that use different promoters (Shiina *et al.*, 2005; Liere and Börner, 2006). To evaluate possible roles of differences within plastome promoter regions to PGI, 38 putative PEP and 25 NEP promoters, that were altered in at least one of the plastomes (Chapter 3.2.4.1) were investigated according to three criteria: (i) their similarity to an ideal consensus (Silhavy and Maliga, 1998; Kapoor and Sugiura, 1999; Homann and Link, 2003; Kanamaru and Tanaka, 2004; Shiina *et al.*, 2005; Liere and Börner, 2006), (ii) number and (iii) position of predicted polymerase binding sites relative to a translational start site (Tables 23 and 24). With this selection scheme 9 putative PEP promoters, notably of *clpP*, *psbB*, *rpl16*, *rpl33*, *rps12*, *rps15*, *trnG<sub>GCC</sub>*, *trnL<sub>CAA</sub>*, *trnS<sub>UGA</sub>*, and 7 predicted NEP promoters, namely of *atpH*, *clpP*, *ndhG*, *psbB*, *psbK*, *rps4* and *trnG<sub>GCC</sub>*, were deduced as candidates causing PGI. Three promoter sequences indicated drastic changes for both polymerases: the promoter of *trnG<sub>GCC</sub>* harbouring mutations specific for plastome V<sup>doul</sup>, and the divergently operating bidirectional promoters for *clpP* (encoding a catalytic subunit of the protease Clp) and *psbB* (encoding the chlorophyll *a* binding protein CP47 of the photosystem II core complex) that contain a large deletion specific for plastome I<sup>loh</sup>. The latter region will subsequently be presented as a case study, because an independent approach uncovered it as a likely and major locus of incompatibility in the interspecific combination of the AB genome with plastome I (Chapter 3.4.2.3).

#### 3.4.1.5. Differences in RNA editing sites

The findings that editotypes in plastids may differ between related species and even between ecotypes (Tillich *et al.*, 2005) and that plant-specific editing sites can often not be modified heterologously (Schmitz-Linneweber *et al.*, 2005; Shikanai, 2006), has suggested that this kind of RNA maturation plays a crucial role not only in translation, but also in speciation processes (Schmitz-Linneweber *et al.*, 2001b; Schmitz-Linneweber *et al.*, 2005). Editing of the *ndhD* ACG start codon in plastomes I<sup>loh</sup> and IV<sup>atro</sup> (Hupfer, 2002) as in *Spinacea*, *Nicotiana* and *Antirrhinum* (Neckermann *et al.*, 1994) established the presence of an editing system in evening primroses. Comparison of the protein coding sequences with those of the liverwort *Marchantia polymorpha*, which lacks RNA editing, uncovered 320 potential sites in

**Table 23.** Assessment of putative PEP promoters as candidate loci for plastome-genome incompatibility in *Oenothera*<sup>1)</sup>

gene	estimated effect	PGI considered
<i>accD</i>	N/A	N/A
<i>atpB</i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>atpF</i>	unlikely	AA-V; AB-V; BB-V
<b><i>clpP</i></b>	likely	<b>AB-I</b> ; AC-I
<i>ndhD</i>	unlikely	BB-I; BB-II
<i>ndhF</i>	N/A	N/A
<b><i>ndhG</i></b>	possible	<b>AB-I</b> ; AC-I
<i>petL</i>	N/A	N/A
<i>psaI</i>	possible	N/A
<b><i>psbB</i></b>	likely	<b>AB-I</b> ; AC-I
<i>psbD</i>	unlikely	BB-I; BB-II
<i>psbE</i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<b><i>psbI</i></b>	possible	<b>AB-I</b> ; AC-I
<i>rbcL</i>	possible	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>rpl16</i>	likely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>rpl20</i>	possible	AA-IV
<i>rpl22</i>	possible	AA-IV
<i>rpl32</i>	possible	N/A
<i>rpl33</i>	likely	BB-I; BB-II
<i>rps12</i>	likely	AA-V; AB-V; BB-V
<i>rps15</i>	likely	AA-V; AB-V; BB-V; BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>rps16</i>	unlikely	BB-I; BB-II
<i>rps18</i>	possible	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>rps4</i>	unlikely	AA-IV
<i>trnF<sub>GAA</sub></i>	unlikely	AB-I; AC-I
<i>trnG<sub>GCC</sub></i>	likely	AA-V; AB-V; BB-V
<i>trnG<sub>UCC</sub></i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>trnH<sub>GUG</sub></i>	unlikely	AA-V; AB-V; BB-V
<i>trnL<sub>CAA</sub></i>	likely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>trnL<sub>UAG</sub></i>	unlikely	AA-V; AB-V; BB-V; BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>trnP<sub>UGG</sub></i>	N/A	N/A
<i>trnQ<sub>UUG</sub></i>	possible	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>trnS<sub>GCU</sub></i>	N/A	N/A
<i>trnS<sub>UGA</sub></i>	likely	BB-II
<i>trnT<sub>UGU</sub></i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>ycf2</i>	possible	N/A
<i>petN</i>	unlikely	AA-IV

<sup>1)</sup> a putative effect on AB-I is marked in bold

**Table 24.** Assessment of putative NEP promoters as candidate loci for plastome-genome incompatibility in *Oenothera*<sup>1)</sup>

gene	estimated effect	PGI considered
<i>ycf2</i>	unlikely	AA-III; AC-III
<i>atpB</i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>atpF</i>	unlikely	AA-III; AC-III
<b><i>atpH</i></b>	likely	<b>AB-I</b> ; AC-I
<i>atpI</i>	possible	AA-V; AB-V; BB-V
<i>clpP</i>	likely	AA-III; AC-III; AA-V; AB-V; BB-V
<i>ndhF</i>	possible	AA-V; AB-V; BB-V; BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<b><i>ndhG</i></b>	likely	<b>AB-I</b> ; AC-I
<i>psaI</i>	unlikely	BB-II
<i>psaJ</i>	unlikely	AA-IV
<i>psbB</i>	likely	AA-III; AC-III
<i>psbK</i>	likely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>rpl20</i>	unlikely	AA-IV
<i>rpl33</i>	possible	AA-V; AB-V; BB-V
<i>rpoB</i>	unlikely	AA-V; AB-V; BB-V
<i>rps12</i>	possible	AA-V; AB-V; BB-V
<i>rps15</i>	unlikely	AA-IV
<i>rps16</i>	unlikely	AA-III; AC-III
<i>rps18</i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>rps4</i>	likely	AA-IV
<i>trnG<sub>GCC</sub></i>	likely	AA-V; AB-V; BB-V
<i>trnI<sub>CAU</sub></i>	possible	AA-III; AC-III
<i>trnP<sub>UGG</sub></i>	unlikely	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>trnQ<sub>UUG</sub></i>	possible	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III
<i>trnT<sub>UGU</sub></i>	possible	BC-I; BC-II; BC-III; CC-I; CC-II; CC-III

<sup>1)</sup> a putative effect on AB-I is marked in bold

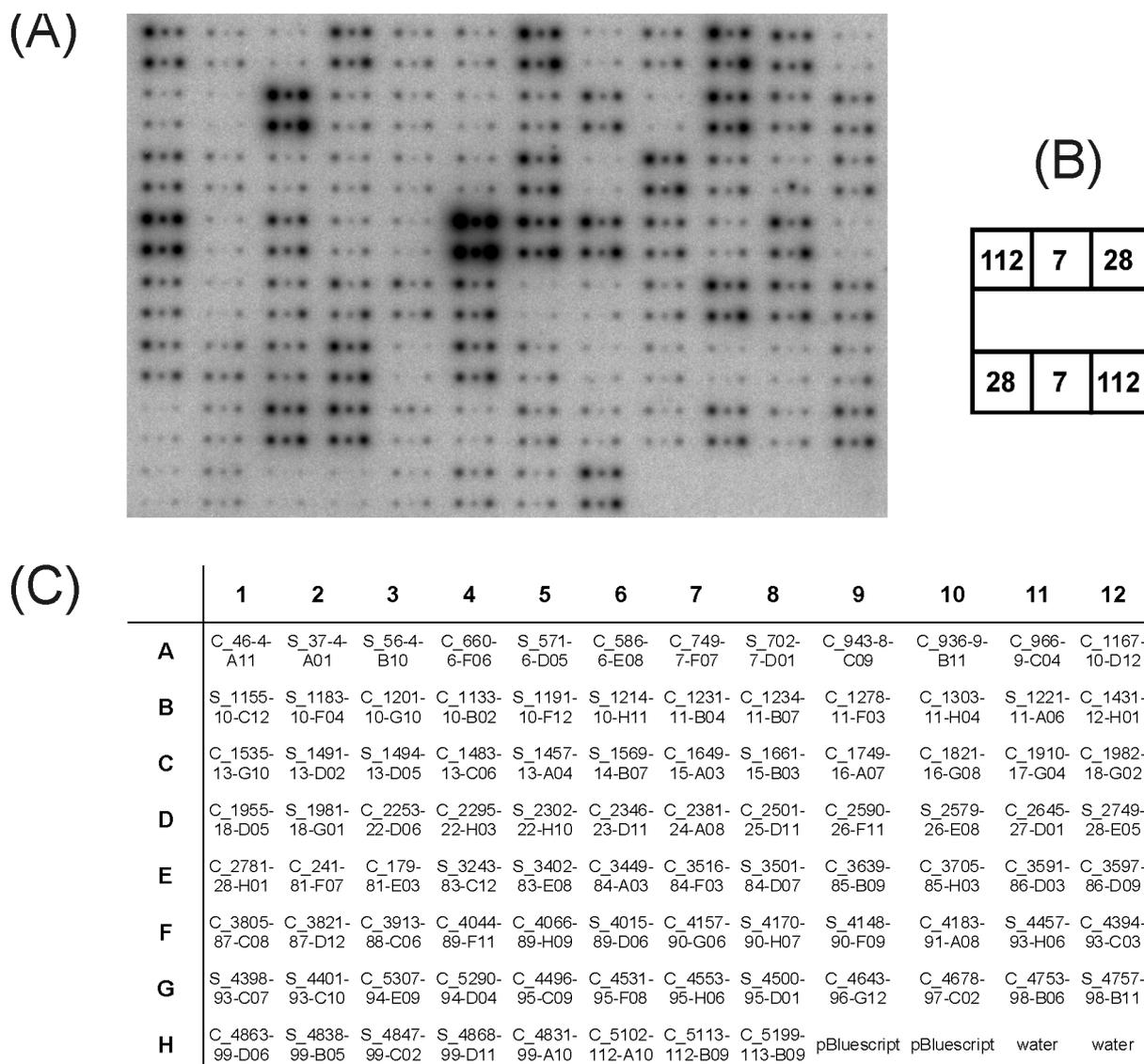
coding regions in *Oenothera* (Wang, 2006), which is an order of magnitude higher than editotypes found usually in higher plants and includes therefore false positives. However, only a single nucleotide substitution differs among the potentially edited sites in the Onagracean plastomes. *NdhA* of plastomes I<sup>oh</sup>, II<sup>suavG</sup> and III<sup>lam</sup> contains a C-to-T conversion at amino acid position 309 compared to plastomes IV<sup>atro</sup> and V<sup>doul</sup>, which would result in a T-to-I change. Since knock-out lines of genes for NADPH dehydrogenase subunits in *Nicotiana* display no or only weak phenotypes (Burrows *et al.*, 1998; Kofler *et al.*, 1998), editing - different from the Solanacean model (Schmitz-Linneweber *et al.*, 2002; Schmitz-Linneweber *et al.*, 2005) - does not play a crucial role in speciation of the *Oenothera* clade nor in the generation of interspecific plastid-nuclear incompatibility.

### 3.4.2. Molecular investigation of PGI

The data set presented above can be used to search for molecular determinants causing PGI in the subsection. In principle, PGI can be caused by structural or regulatory processes. Co-evolution of polypeptides with their interaction partners (polypeptides or nucleic acids), is a well-known phenomenon (Goh *et al.*, 2000). Diverging protein coding loci are therefore of intrinsic interest, but the overwhelming majority of such loci found in *Oenothera*, do not seem to be of obvious functional relevance nor causative for interspecific compartmental incompatibility. In protein coding genes, out of 388 diverging loci, only 35 are located in functional domains (Chapter 3.4.1.2), and even for those 35 exchanges it is unlikely, that they can explain the enormous variety of PGI phenotypes in *Oenothera*, as discussed in more detail in Chapter 4.5.2.

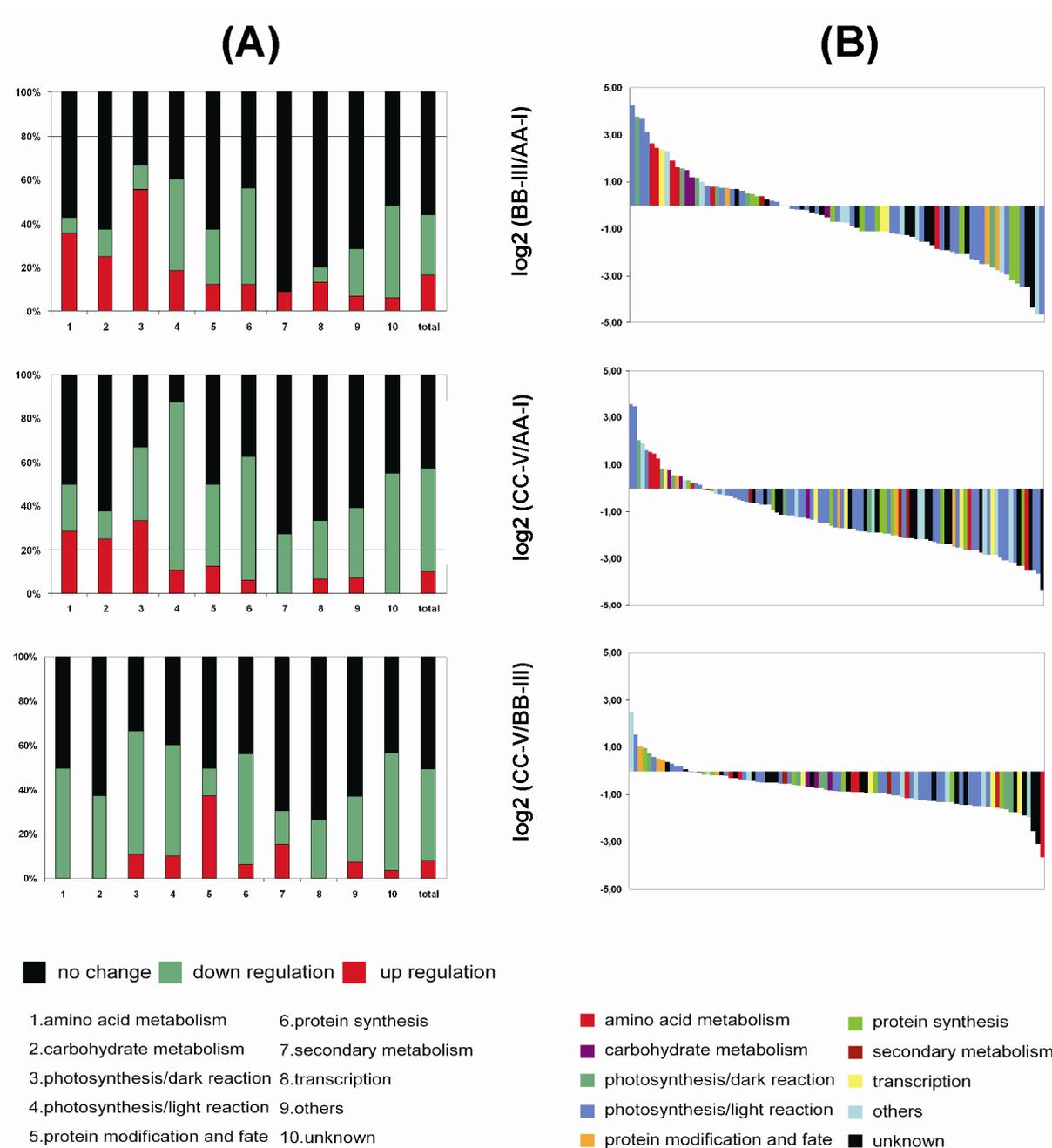
#### 3.4.2.1. Analysis of nuclear gene expression in the three basic *Oenothera* lineages

If differences in coding regions cannot or not fully explain PGI in the genus, gene regulation must be of particular importance. This should be reflected first in different plastid promoter sequences as deduced in Chapter 3.4.1.4 but also by variance of nuclear gene expression in species whose hybrids show plastome-genome incompatible phenotypes. Variance in nuclear gene expression may be particularly relevant, since regulation of plastid and nuclear mRNA levels have to be intimately coordinated to allow stoichiometric delivery of chloroplast-destined components (Figure 1). Therefore, these differences could potentially reflect species-specific expression patterns and plastome-genome interactions.



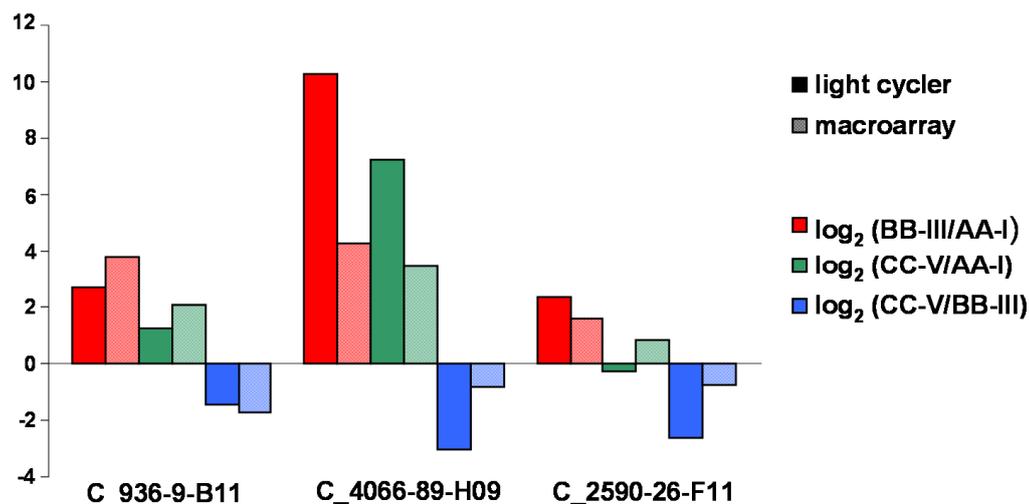
**Figure 30.** Representative autoradiogram of an array filter representing a subset of nuclear genes for chloroplast function. Signals were generated by hybridizing labeled cDNA from the AA genome (panel A). Order and quantities (ng/spot) of immobilized probes (panel B) as well as EST clusters of spotted probes (panel C) are given. For description of the EST clusters under study see Table 25.

To address the questions of co-regulated gene clusters, and to which extent transcriptomes differ between naturally occurring species residing in different habitats, macroarrays were equipped with EST-derived probes of 187 nuclear genes that contribute to known and unknown chloroplast functions (Table 25). The expression profiles of leaves from three different *Oenothera* species *Oe. elata* subsp. *hookeri* strain johansen (AA-I), *Oe. grandiflora* strain tuscaloosa (BB-III), and *Oe. argillicola* strain douthat 1 (CC-V), kept under the same physiological conditions, were investigated. Hybrids between these three species give usually plastome-genome incompatible offspring (Table 1). For precise quantifications, the spots on



**Figure 31.** Comparison of the transcript expression profiling of 187 nuclear genes for chloroplast function in the three naturally occurring plastome-genome combinations AA-I, BB-III and CC-V. All expression ratios are converted to  $\log_2$  for simplicity. The histogram of 10 major functional categories shows the proportion of identically expressed genes (black), higher- (red) and lower-expressed (green) genes (panel A). Histogram of average ratios of transcripts from plants of the three species (panel B). The ratio for each spot represents the average of six independent quantities per gene and two replicates. The colours of the bars represent the categories as indicated.

each filter contained 112, 28 and 7 ng of each PCR product in duplicate. A representative array is illustrated in Figure 30. The hybridization signals were statistically normalized using standard methods (Data Range, AIDA Array Compare program, version 4.0, Raytest Isotopenmeßgeräte GmbH). Data Range is a global normalization method, normalizing all spots using the same reference value. After normalization the hybridization signals were compared to each other - BB-III *versus* AA-I, CC-V *versus* AA-I and CC-V *versus* BB-III. The selected nuclear genes were grouped into 10 different major functional categories of the chloroplast including amino acid metabolism, carbohydrate metabolism, photosynthesis (light and dark reactions), protein modification and fate, protein biosynthesis, secondary metabolism, transcription, unknown proteins and others (Table 25). The proportions of genes differentially or identically expressed were determined and histograms of the corresponding categories were generated (Figure 31). Macroarray analysis was done in co-operation with Won Kyong Cho.



**Figure 32.** Validation of expression data obtained by macroarray analysis *via* real-time PCR. Explanation is given in the text. Numbers represent expression of different EST clusters (C\_936-9-B11 stands for a putative transketolase, C\_4066-89-H09 for a chlorophyll *a/b* binding family protein and C\_2590-26-F11 for a phosphoribulokinase).

To validate the expression data obtained by the macroarray-based approach, mRNA levels of three gene clusters were quantified by real-time RT-PCR. A comparison of the expression levels of three gene clusters between the three genomes confirmed the differential expression of the genes and highlight gene clusters, which were higher and lower expressed in the respective

**Table 25.** Comparison of three transcriptomes of the genetic constitutions AA-I, BB-III and CC-V

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
1	S_56-4-B10	At3g48560	acetolactate synthase, chloroplast	amino acid metabolism	0,00	0,00	0,00
2	C_1133-10-B02	At5g63570	glutamate-1-semialdehyde aminotransferase, chloroplast	amino acid metabolism	0,00	0,00	0,00
3	C_86-4-D12	At5g35630	glutamate-ammonia ligase, chloroplast	amino acid metabolism	-1,84	-3,47	-1,56
4	S_4232-91-E09	At4g33680	unknown protein (transaminase activity), chloroplast	amino acid metabolism	1,91	0,00	-3,64
5	S_1491-13-D02	At5g12860	2-oxoglutarate/malate translocator protein, chloroplast	carbohydrate metabolism	0,00	0,00	0,00
6	S_4757-98-B11	At3g62410	CP12 domain-containing protein, chloroplast	carbohydrate metabolism	0,00	-0,59	0,00
7	S_4457-93-H06	At1g43670	fructose 1,6-bisphosphatase, localisation unknown	carbohydrate metabolism	N/A	N/A	N/A
8	C_1167-10-D12	At3g52930	fructose bisphosphate aldolase, cytosol	carbohydrate metabolism	0,00	0,00	0,00
9	S_4170-90-H07	At5g64380	fructose-bisphosphatase protein, mitochondrium	carbohydrate metabolism	N/A	N/A	N/A
10	S_2748-28-E04	At2g01140	fructose-bisphosphate aldolase, chloroplast	carbohydrate metabolism	-0,49	-1,29	-0,81
11	C_2253-22-D06	At4g38970	fructose-bisphosphate aldolase, chloroplast	carbohydrate metabolism	1,19	0,51	-0,67
12	C_4157-90-G06	At3g26650	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	carbohydrate metabolism	2,44	1,55	-0,89
13	C_943-8-C09	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	carbohydrate metabolism	1,64	1,27	-0,36
14	C_1278-11-F03	At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast	carbohydrate metabolism	N/A	N/A	N/A
15	C_1535-13-G10	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	carbohydrate metabolism	0,79	-0,11	-0,89
16	S_1494-13-D05	At1g42970	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast	carbohydrate metabolism	0,00	0,00	0,00
17	C_3639-85-B09	At1g13440	glyceraldehyde-3-phosphate dehydrogenase, cytosolic	carbohydrate metabolism	2,64	1,49	-1,15
18	C_4863-99-D06	At1g79750	malate oxidoreductase, (AtNADP-Me4), chloroplast	carbohydrate metabolism	N/A	N/A	N/A
19	S_1183-10-F04	At2g13560	malate oxidoreductase, mitochondrion	carbohydrate metabolism	N/A	N/A	N/A
20	C_4553-95-H06	At2g22780	NAD-dependent malate dehydrogenase, chloroplast	carbohydrate metabolism	0,99	-0,24	-1,22
21	S_1981-18-G01	At3g47520	NAD-dependent malate dehydrogenase, chloroplast	carbohydrate metabolism	0,00	0,00	0,00
22	S_4401-93-C10	At5g09660	NAD-dependent malate dehydrogenase, microbody	carbohydrate metabolism	2,31	1,88	-0,43
23	C_3805-87-C08	At3g12780	phosphoglycerate kinase protein, cytosol, mitochondrion	carbohydrate metabolism	1,16	0,55	-0,62
24	C_2590-26-F11	At1g32060	phosphoribulokinase (Pgk1), mitochondrion, cytosol	carbohydrate metabolism	1,58	0,84	-0,74

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
25	S_4838-99-B05	At3g55800	sedoheptulose-bisphosphatase, chloroplast	carbohydrate metabolism	0,00	0,00	0,00
26	C_4643-96-G12	At2g21170	triosephosphatisomerase, chloroplast	carbohydrate metabolism	N/A	N/A	N/A
27	C_2295-22-H03	At3g55440	triosephosphatisomerase, cytosol	carbohydrate metabolism	0,00	0,00	0,00
28	S_1221-11-A06	At3g63140	mRNA binding protein, chloroplast	gene expression	0,00	-1,36	0,00
29	S_4148-90-F09	At5g50250	RNA binding protein (rbp31), chloroplast	gene expression	N/A	N/A	N/A
30	C_4753-98-B06	At5g50250	RNA binding protein (rbp31), chloroplast	gene expression	0,00	0,00	0,00
31	C_5113-112-B09	At5g50250	RNA binding protein (rbp31), chloroplast	gene expression	0,00	0,00	0,00
32	C_5186-113-A08	At1g60000	RNA-binding protein (Cp29), chloroplast	gene expression	-3,18	-3,32	-0,18
33	S_1191-10-F12	At2g37220	RNA-binding protein (Cp29), chloroplast	gene expression	0,00	-0,12	0,00
34	C_2645-27-D01	At2g37220	RNA-binding protein (Cp29), chloroplast	gene expression	2,36	0,78	-1,56
35	S_4500-95-D01	At2g37220	RNA-binding protein (Cp29), chloroplast	gene expression	0,00	0,00	0,00
36	S_4269-91-H10	At2g39140	unknown protein (RNA binding), chloroplast	gene expression	N/A	N/A	N/A
37	S_3801-87-C04	At4g34730	unknown protein (RNA binding), chloroplast	gene expression	0,00	-2,56	-0,92
38	C_660-6-F06	At2g21330	fructose-bisphosphate aldolase, chloroplast	photosynthesis/dark reaction	1,49	0,76	-0,74
39	S_3501-84-D07	At2g01290	ribose 5-phosphate isomerase, localisation unknown	photosynthesis/dark reaction	N/A	N/A	N/A
40	C_2346-23-D11	At1g67090	RuBisCO small subunit, chloroplast	photosynthesis/dark reaction	0,53	-1,13	-1,64
41	C_3597-86-D09	At1g67090	RuBisCO small subunit, chloroplast	photosynthesis/dark reaction	0,00	-0,03	0,00
42	C_1857-17-B09	At5g61410	ribulose-5-phosphate-3-epimerase, chloroplast	photosynthesis/dark reaction	-0,76	-1,89	-1,12
43	C_766-7-G12	At2g39730	RuBisCO activase, chloroplast	photosynthesis/dark reaction	-2,64	-1,89	0,72
44	C_1231-11-B04	At3g55800	sedoheptulose-bisphosphatase, chloroplast	photosynthesis/dark reaction	0,00	0,21	0,00
45	S_571-6-D05	At2g45290	transketolase, chloroplast	photosynthesis/dark reaction	0,00	0,00	0,00
46	C_936-9-B11	At2g45290	transketolase, chloroplast	photosynthesis/dark reaction	3,78	2,05	-1,74
47	C_1234-11-B07	At2g45290	transketolase, chloroplast,	photosynthesis/dark reaction	0,79	0,00	-0,79
48	C_3913-88-C06	At4g09650	ATP synthase delta subunit (AtpD), chloroplast	photosynthesis/light reaction	0,00	-0,27	0,00

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
49	C_4678-97-C02	At4g04640	ATP synthase gamma subunit 1 (AtpC1), chloroplast	photosynthesis/light reaction	0,77	-0,26	-1,03
50	C_966-9-C04	At4g32260	ATP synthase subunit (AtpG), chloroplast	photosynthesis/light reaction	0,00	-1,24	0,00
51	C_4066-89-H09	At4g14690	chlorophyll <i>a/b</i> binding family protein (Elip2), chloroplast	photosynthesis/light reaction	4,25	3,47	-0,86
52	C_3916-88-C09	At4g03280	cytochrome B <sub>6</sub> -F complex subunit (PetC), chloroplast	photosynthesis/light reaction	-1,12	-1,25	-0,12
53	C_5307-94-E09	At1g60950	ferredoxin protein (FedA), chloroplast	photosynthesis/light reaction	-0,03	-0,51	-0,47
54	C_241-81-F07	At4g14890	ferredoxin protein, chloroplast	photosynthesis/light reaction	N/A	N/A	N/A
55	C_1431-12-H01	At1g20020	ferredoxin-NADP <sup>+</sup> reductase protein, chloroplast	photosynthesis/light reaction	0,00	-0,53	0,00
56	C_1982-18-G02	At5g66190	ferredoxin-NADP <sup>+</sup> reductase protein, chloroplast	photosynthesis/light reaction	0,00	-0,36	0,00
57	C_4394-93-C03	At1g19150	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	0,00	0,00	0,00
58	S_37-4-A01	At1g45474	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	0,00	0,00	0,00
59	C_179-81-E03	At1g61520	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	0,00	-0,28	0,00
60	C_2781-28-H01	At3g47470	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	0,00	-1,45	0,00
61	C_5199-113-B09	At3g47470	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	-0,34	-1,82	-1,47
62	C_1955-18-D05	At3g54890	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	-2,25	-3,10	-0,84
63	C_1749-16-A07	At3g61470	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	0,00	-2,66	0,00
64	C_3705-85-H03	At5g54270	PSI chlorophyll <i>a/b</i> -binding protein (Lhca), chloroplast	photosynthesis/light reaction	0,00	-2,30	0,00
65	C_4044-89-F11	At1g03130	PSI reaction centre subunit (PsaD2), chloroplast	photosynthesis/light reaction	0,63	-0,69	-1,32
66	C_4896-99-G04	At4g28750	PSI reaction centre subunit (PsaE1), chloroplast	photosynthesis/light reaction	-2,06	-3,06	-1,03
67	C_5189-113-A11	At1g31330	PSI reaction centre subunit (PsaF), chloroplast	photosynthesis/light reaction	-4,64	-2,94	1,56
68	C_5220-113-D06	At1g55670	PSI reaction centre subunit (PsaG), chloroplast	photosynthesis/light reaction	-2,32	-1,74	0,59
69	C_3443-83-H09	At3g16140	PSI reaction centre subunit (PsaH1), chloroplast	photosynthesis/light reaction	-3,47	-3,18	0,32
70	C_469-111-E06	At1g30380	PSI reaction centre subunit (PsaK), chloroplast	photosynthesis/light reaction	-2,94	-3,47	-0,56
71	C_5290-94-D04	At4g12800	PSI reaction centre subunit (PsaL), chloroplast	photosynthesis/light reaction	-1,18	-1,67	-0,49
72	C_2287-22-G06	At5g64040	PSI reaction centre subunit (PsaN), chloroplast	photosynthesis/light reaction	-2,47	-3,64	-1,15

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
73	C_1201-10-G10	At4g10340	PSII chlorophyll <i>a/b</i> -binding protein (Cp26), chloroplast	photosynthesis/light reaction	-0,69	-2,11	-1,43
74	C_749-7-F07	At5g01530	PSII chlorophyll <i>a/b</i> -binding protein (Cp29), chloroplast	photosynthesis/light reaction	3,66	3,59	-0,06
75	S_4015-89-D06	At5g01530	PSII chlorophyll <i>a/b</i> -binding protein (Cp29), chloroplast	photosynthesis/light reaction	0,00	-1,48	0,00
76	C_1821-16-G08	At1g15820	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	0,00	-1,83	0,00
77	C_46-4-A11	At2g34420	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	0,68	0,15	-0,54
78	C_1303-11-H04	At2g34430	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	0,12	-1,17	-1,29
79	C_1598-14-D12	At2g40100	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	N/A	N/A	N/A
80	C_2501-25-D11	At1g29920	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	-0,89	-0,70	0,20
81	C_1649-15-A03	At1g29930	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	0,00	-1,35	0,00
82	C_2381-24-A08	At2g05070	PSII chlorophyll <i>a/b</i> -binding protein (Lhcb), chloroplast	photosynthesis/light reaction	-0,15	-1,49	-1,32
83	C_1098-9-G03	At1g44575	PSII chlorophyll <i>a/b</i> -binding protein (PsbS), chloroplast	photosynthesis/light reaction	-0,22	-1,15	-0,94
84	C_3591-86-D03	At3g50820	PSII oxygen-evolving complex subunit (PsbO), chloroplast	photosynthesis/light reaction	0,00	-0,65	0,00
85	C_391-111-F03	At1g06680	PSII oxygen-evolving complex subunit (PsbP), chloroplast	photosynthesis/light reaction	-1,94	-2,32	-0,40
86	C_4496-95-C09	At2g39470	PSII oxygen-evolving complex subunit (PsbP), chloroplast	photosynthesis/light reaction	0,00	0,00	0,00
87	C_4183-91-A08	At3g01440	PSII oxygen-evolving complex subunit 3 (PsbQ), chloroplast	photosynthesis/light reaction	N/A	N/A	N/A
88	C_3449-84-A03	At4g05180	PSII oxygen-evolving complex subunit 3 (PsbQ), chloroplast	photosynthesis/light reaction	0,21	0,21	0,00
89	C_1462-13-A09	At1g79040	PSII reaction centre subunit (PsbR), chloroplast	photosynthesis/light reaction	-1,56	-2,84	-1,25
90	C_311-111-B07	At2g30570	PSII reaction centre subunit (PsbW), chloroplast	photosynthesis/light reaction	-1,22	-2,64	-1,43
91	S_1155-10-C12	At4g28660	PSII reaction centre subunit (PsbW), chloroplast	photosynthesis/light reaction	3,13	1,61	-1,51
92	C_5178-112-H12	At1g67740	PSII reaction centre subunit (PsbY), chloroplast	photosynthesis/light reaction	-1,89	-1,69	0,19
93	C_2191-21-D05	At1g20340	plastocyanin (PetE2), chloroplast	photosynthesis/light reaction	-1,12	-2,56	-1,47
94	C_2282-22-F12	At2g05620	protein involved in electron flow in PSI (Pgr5), chloroplast	photosynthesis/light reaction	0,83	-0,42	-1,25
95	C_1508-13-E07	At5g45390	ATP-dependent Clp protease protein, chloroplast	protein modification and fate	0,00	-2,47	0,51
96	S_1457-13-A04	At1g49970	ClpP protease complex subunit (ClpR1), chloroplast	protein modification and fate	0,00	0,00	0,00

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
97	S_1607-14-E09	At1g09130	ClpP protease complex subunit (ClpR3), chloroplast	protein modification and fate	0,00	0,00	0,00
98	S_3243-83-C12	At5g20720	Cpn21 protein, chloroplast	protein modification and fate	0,75	0,54	-0,20
99	C_2667-27-E12	At5g20720	Cpn21 protein, chloroplast	protein modification and fate	-2,47	-2,00	0,45
100	S_4398-93-C07	At5g17710	GrpE protein, chloroplast	protein modification and fate	N/A	N/A	N/A
101	C_273-81-C12	At5g23120	PSII stability/assembly factor HCF136, chloroplast	protein modification and fate	-1,12	-1,74	-0,62
102	C_4106-90-C03	At5g56500	RuBisCO subunit binding-protein beta subunit, chloroplast	protein modification and fate	-2,74	-1,69	1,03
103	C_2715-28-B07	At1g68590	plastid ribosomal protein 3 (Psrp-3), chloroplast	protein synthesis	-0,14	0,00	-1,60
104	C_2722-28-C02	At5g17870	Plastid ribosomal protein 6 (Psrp-6), chloroplast	protein synthesis	0,40	-0,94	-1,32
105	C_1910-17-G04	At5g48760	ribosomal protein L13 (CL13a), chloroplast	protein synthesis	N/A	N/A	N/A
106	S_3402-83-E08	At3g54210	ribosomal protein L17 (CL17), chloroplast	protein synthesis	0,49	0,34	-0,15
107	C_927-8-C04	At5g47190	ribosomal protein L19 (CL19), chloroplast	protein synthesis	-2,06	-2,64	-0,58
108	C_939-7-H02	At3g63490	ribosomal protein L1 (CL1), chloroplast	protein synthesis	N/A	N/A	N/A
109	S_413-5-A07	At1g35680	ribosomal protein L21 (CL21), chloroplast	protein synthesis	-1,12	-1,94	-0,86
110	C_4381-93-B02	At2g24090	ribosomal protein L35 (CL35), chloroplast	protein synthesis	-0,67	-1,60	-0,92
111	C_583-6-E05	At1g07320	ribosomal protein L4 (CL4), chloroplast	protein synthesis	N/A	N/A	N/A
112	C_1498-13-D09	At1g05190	ribosomal protein L6 (CL6), chloroplast	protein synthesis	-1,09	-2,00	-0,94
113	C_586-6-E08	At3g44890	ribosomal protein L9 (CL9), chloroplast	protein synthesis	0,00	0,00	0,00
114	S_1569-14-B07	At5g30510	ribosomal protein S1 (CS1), chloroplast	protein synthesis	N/A	N/A	N/A
115	C_3821-87-D12	At5g30510	ribosomal protein S1 (CS1), chloroplast	protein synthesis	0,00	-1,90	0,00
116	C_873-8-F06	At4g20360	translation elongation factor (EF-Tu), chloroplast	protein synthesis	-3,32	-2,40	0,97
117	C_804-7-E12	At3g62910	translation releasing factor RF-1 protein, chloroplast	protein synthesis	N/A	N/A	N/A
118	S_2749-28-E05	At5g54500	1,4-benzoquinone reductase, localisation unknown	secondary metabolism	N/A	N/A	N/A
119	S_4868-99-D11	At4g36250	betaine aldehyde dehydrogenase, chloroplast	secondary metabolism	N/A	N/A	N/A
120	S_2579-26-E08	At3g51820	chlorophyll synthetase protein, chloroplast	secondary metabolism	0,00	0,00	0,00

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
121	S_5270-94-B06	At5g18660	divinyl protochlorophyllide 8-vinyl reductase, chloroplast	secondary metabolism	0,00	-2,12	-0,56
122	S_1214-10-H11	At3g63410	MPBQ/MSBQ methyltransferase, chloroplast	secondary metabolism	0,00	0,00	0,00
123	C_5102-112-A10	At5g54190	NADPH:protochlorophyllide oxidoreductase A, chloroplast	secondary metabolism	0,00	-2,08	0,00
124	C_4531-95-F08	At2g43710	stearoyl-acyl carrier protein desaturase, chloroplast	secondary metabolism	N/A	N/A	N/A
125	S_4845-99-B12	At3g14110	TPR-containing protein (Flu), chloroplast	secondary metabolism	N/A	N/A	N/A
126	S_67-4-C05	At4g34350	isopentenyl diphosphate biosynthesis protein, chloroplast	secondary metabolism	N/A	N/A	N/A
127	S_3870-87-H04	At4g35250	vestitone reductase-related protein, chloroplast	secondary metabolism	0,39	-0,62	-1,00
128	C_4231-91-E08	At3g04730	auxin-induced protein (Iaa16), nucleus	transcription	N/A	N/A	N/A
129	C_574-6-D08	At5g65670	auxin-induced protein (Iaa9), nucleus	transcription	-2,84	-2,84	-0,07
130	S_578-6-D12	At3g19290	bZIP transcription factor (Abf4), nucleus	transcription	N/A	N/A	N/A
131	S_2696-27-H12	At1g66230	myb-related transcription factor, nucleus	transcription	N/A	N/A	N/A
132	S_4280-92-A09	At5g17260	NAM (no apical meristem) protein, localisation unknown	transcription	N/A	N/A	N/A
133	C_2267-22-E08	At2g46820	unknown protein (DNA binding), chloroplast	transcription	-1,12	-2,84	-1,79
134	C_80-4-D06	At4g28440	unknown protein (DNA-binding protein-related), chloroplast	transcription	N/A	N/A	N/A
135	S_464-5-D09	At5g24930	unknown protein (transcription factor activity), chloroplast	transcription	N/A	N/A	N/A
136	S_4847-99-C02	At4g21090	adrenodoxin ferredoxin , mitochondrion	other	N/A	N/A	N/A
137	C_4831-99-A10	At1g19920	ATP-sulfurylase (Aps2), chloroplast	other	N/A	N/A	N/A
138	S_2650-27-D06	At3g03850	auxin-induced protein, mitochondrion	other	N/A	N/A	N/A
139	C_3439-83-H05	At3g01500	carbonic anhydrase, chloroplast	other	-4,64	-2,18	2,49
140	C_1023-9-C06	At5g55280	cell division protein FtsZ homolog, chloroplast	other	N/A	N/A	N/A
141	S_702-7-D01	At4g39090	drought-inducible cysteine proteinase (Rd19A), ER	other	0,00	0,00	0,00
142	C_1483-13-C06	At4g39090	drought-inducible cysteine proteinase (Rd19A), ER	other	0,00	0,00	0,00
143	S_710-7-D09	At4g02440	EID1, localisation unknown	other	N/A	N/A	N/A
144	C_299-111-A07	At1g75750	gibberellin-regulated protein 1 (Gasa1), cell wall	other	0,00	0,34	-0,15

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
145	S_5361-115-C05	At2g36640	AtEcp63, localisation unknown	other	N/A	N/A	N/A
146	S_231-81-E08	At4g30950	omega-6 fatty acid desaturase (Fad6), chloroplast	other	-0,71	-2,18	-1,47
147	S_4725-97-G04	At3g44880	pheide A oxygenase (Ac1), chloroplast	other	N/A	N/A	N/A
148	S_1661-15-B03	At3g10290	phosphate/phosphoenolpyruvate translocator, chloroplast	other	N/A	N/A	N/A
149	S_3972-88-G11	At2g47590	photolyase/blue-light receptor (PHR2), localisation unknown	other	0,00	-1,94	-0,22
150	S_1248-11-C09	At2g21280	protein evolved in plastid division (AtSulA), chloroplast	other	N/A	N/A	N/A
151	S_327-111-D03	At4g01800	protein import component SecA-Type, chloroplast	other	N/A	N/A	N/A
152	S_3300-82-F01	At3g23710	protein import component Tic22, chloroplast	other	N/A	N/A	N/A
153	S_2302-22-H10	At4g02510	protein import component Toc159, chloroplast	other	N/A	N/A	N/A
154	S_2673-27-F10	At2g28800	protein translocase (Albino3), chloroplast	other	0,00	-1,22	0,37
155	S_685-7-B08	At3g52180	protein tyrosine phosphatase/kinase (PTPKIS1), chloroplast	other	N/A	N/A	N/A
156	C_674-7-A09	At5g36120	YGGT family protein (TARZAN), chloroplast	unknown	-0,94	-2,25	-1,29
157	C_1202-10-G11	At2g06520	unknown protein (2 transmembrane domains), chloroplast	unknown	-1,56	-2,40	-0,86
158	C_5285-94-C11	At1g77090	unknown protein (contains a PsbP domain), chloroplast	unknown	-0,40	-2,12	-1,74
159	C_4132-90-E05	At5g57040	unknown protein (glyoxalase I family), chloroplast	unknown	N/A	N/A	N/A
160	S_2315-23-B03	At1g52510	unknown protein (hydrolase activity), chloroplast	unknown	-0,29	0,00	-3,06
161	S_1235-11-B08	At3g04760	unknown protein (PPR repeat-containing protein), chloroplast	unknown	N/A	N/A	N/A
162	S_4837-99-B04	At5g14660	unknown protein (peptide deformylase), chloroplast	unknown	N/A	N/A	N/A
163	C_3516-84-F03	At4g22310	unknown protein (Pfam-domain), mitochondrion	unknown	N/A	N/A	N/A
164	C_861-9-A01	At5g19390	unknown protein (pleckstrin homology domain), chloroplast	unknown	N/A	N/A	N/A
165	C_3695-85-G05	At1g24020	unknown protein (pollen allergen), localisation unknown	unknown	0,25	-1,12	-1,40
166	S_3742-86-H03	At3g17800	unknown protein (response to UV-B), chloroplast	unknown	-1,89	-3,32	-1,43
167	C_1374-12-E11	At4g01050	unknown protein (rhodanese domain-containing protein)	unknown	N/A	N/A	N/A
168	C_494-5-F05	At4g27700	unknown protein (rhodanese domain-containing protein)	unknown	-4,32	-4,32	0,00

Table 25. (continued)

no.	EST/cluster accession	closest <i>Arabidopsis</i> homolog (blastX)	protein function and localisation in <i>Arabidopsis</i>	functional group	relative expression log <sub>2</sub> (BB-III/AA-I)	relative expression log <sub>2</sub> (CC-V/AA-I)	relative expression log <sub>2</sub> (CC-V/BB-III)
169	S_564-6-C09	At2g37240	unknown protein (weak similarity to FmHP), chloroplast	unknown	N/A	N/A	N/A
170	S_614-6-G12	At1g08070	unknown protein, chloroplast	unknown	N/A	N/A	N/A
171	C_3535-84-G10	At1g08380	unknown protein, chloroplast	unknown	-1,69	-2,18	-0,49
172	C_482-5-E05	At1g15370	unknown protein, chloroplast	unknown	-2,06	-2,74	-0,67
173	C_862-8-E07	At1g15980	unknown protein, chloroplast	unknown	0,00	-2,64	-0,29
174	C_719-7-E06	At1g45688	unknown protein, chloroplast	unknown	N/A	N/A	N/A
175	S_3927-88-D02	At1g54520	unknown protein, chloroplast	unknown	0,00	-2,40	-0,43
176	C_239-81-F05	At1g74730	unknown protein, chloroplast	unknown	-1,32	-1,84	-0,49
177	S_3881-87-E11	At2g36145	unknown protein, chloroplast	unknown	0,00	-1,03	-0,89
178	S_4357-92-H02	At3g20680	unknown protein, chloroplast	unknown	-3,47	-3,47	0,08
179	S_45-4-A10	At3g56140	unknown protein, chloroplast	unknown	-0,20	-0,69	-0,49
180	C_2290-22-G09	At4g01150	unknown protein, chloroplast	unknown	-1,25	-2,18	-0,92
181	S_2708-28-A12	At4g13220	unknown protein, chloroplast	unknown	0,67	0,00	-2,56
182	S_3688-85-F09	At4g30620	unknown protein, chloroplast	unknown	-0,01	-1,89	-1,89
183	C_4596-96-D01	At5g08050	unknown protein, chloroplast	unknown	0,00	-0,64	-0,20
184	S_2535-25-G10	At5g14970	unknown protein, chloroplast	unknown	-1,47	-2,84	-1,32
185	C_3554-86-A04	At5g42765	unknown protein, chloroplast	unknown	N/A	N/A	N/A
186	S_2327-23-C03	At5g52110	unknown protein, chloroplast	unknown	N/A	N/A	N/A
187	S_4836-99-B03	At5g57345	unknown protein, chloroplast	unknown	0,00	-1,74	-0,32

species (Figure 32). However, the extent of the ratios, irrespective of whether higher or lower than 1, differed somehow depending on the cluster and the method used.

#### 3.4.2.2. *Photosynthesis gene clusters are differentially regulated*

In summary, the nuclear gene expression data obtained revealed distinct expression signatures in the different *Oenothera* wild-type species demonstrating the applicability of the array-based technology in judging stationary RNA levels and in analyzing the interspecific variance of nuclear gene expression. Clusters of nuclear genes could be highlighted that are potential targets for species-specific co-regulation. While the stationary RNA levels of a subset of genes from all categories was unchanged, mutual comparison of all hybridization signals revealed that about 50% of all genes studied were differentially expressed. Remarkably, most differences could be found in the category photosynthesis, its light as well as dark reaction, ranging from 61% to 84% differentially expressed genes. On the other hand, less than 22% of genes for the secondary metabolism were differentially expressed (Figure 31). These data indicate that the regulation of genes for photosynthesis is an adaptorial phenomenon in species from different habitats. In consequence, regulation of photosynthetic genes may contribute to PGI, and implicitly also to speciation.

#### 3.4.2.3. *Delineation of the AB-I incompatibility determinant*

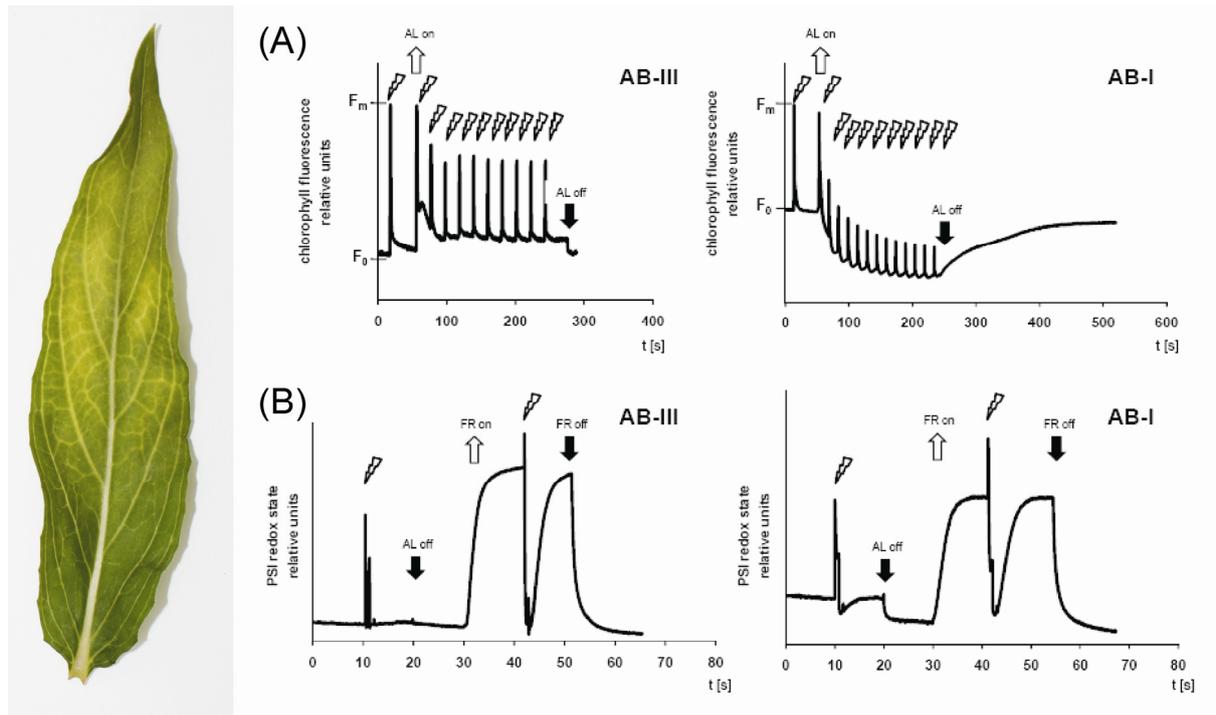
To get an idea about the nature of regulatory factors for photosynthesis in plastid gene expression that could be responsible for PGI, the incompatible interspecific hybrid AB-I was taken as a showcase. The hybrid was taken for two reasons: first, the combination AB-I displays a clear phenotype, affecting photosynthesis (following chapters) and second its genetics is easy accessible. It can be deduced from the incompatibility chart that plastome I must be different in some respect to plastomes II, III, and IV, since only plastome I is incompatible in AB background (Figure 4). Therefore, all changes common to other plastomes could be disregarded. Plastome V was excluded from the analysis because the combination AB-V is extremely disharmonic and differs substantially from AB-I. It is fully bleached, largely pollen sterile and displays severe inhibition of cell division (Stubbe, 1963b). Representing a chimerical incompatibility the genetic determinants responsible for its phenotypes are presumably complex and different from those causing bleaching of AB-I individuals (Chapter 3.3.3).

#### 3.4.2.3.1. *The incompatible AB-I hybrid shows a photosystem II phenotype*

To test if the incompatibly AB-I displays a phenotype, which can be correlated with the photosynthetic machinery, spectroscopic analyses were performed to quantify photosystem II and photosystem I activity (Figure 33). Chlorophyll fluorescence induction was measured on dark-adapted leaves of incompatible  ${}^h\text{johansen}\cdot{}^h\text{tuscaloosa}$   $F^{\text{oh}}$  (AB-I) and compatible  ${}^h\text{johansen}\cdot{}^h\text{tuscaloosa}$   $\text{III}^{\text{lam/tusca}}$  (AB-III) and  ${}^h\text{johansen}\cdot{}^h\text{johansen}$   $F^{\text{oh}}$  (AA-I) plants. Maximum photosystem II quantum efficiency was reduced to  $0.52 \pm 0.04$  in bleached AB-I leaf tissue compared to  $0.79 \pm 0.03$  in AB-III or AA-I consistent with a deficient photosystem II activity. The low Fv/Fm ratio observed was caused by an elevated Fo fluorescence level in AB-I (Figure 33). The Fm level elicited by a saturating pulse was quenched by moderate actinic light ( $50 \mu\text{E m}^{-2} \text{sec}^{-1}$ ) in all plants. However, the steady state fluorescence of AB-I dropped far below the initial Fo level depending on light intensity. Actinic light-induced quenching was slower in AB-I than in AB-III or AA-I, and a longer period was required to reach steady state. NPQ increased dramatically from  $0.39 \pm 0.14$  in AA-I and AB-III to  $2.47 \pm 0.42$  in AB-I. After switching off actinic light, the fluorescence decayed to the former Fo level in AA-I and AB-III but increased in AB-I within several minutes to finally reach the increased Fo already observed during the dark adaptation prior to the measurement, again indicating malfunction of photosystem II. Application of far red light, which preferentially excites photosystem I, had no notable effect on the half-life of the fluorescence rise upon light/dark switches nor on the elevated Fo levels suggesting that the AB-I incompatibility reflects a direct effect on photosystem II driven electron transport, such as stable accumulation of  $Q_A^-$  species in the dark, which induces an increased Fo level. The AB-I patterns clearly appeared as a consequence of altered photosystem II.

#### 3.4.2.3.2. *The effect on PSII is specific and does not notably affect PSI*

The extent of photosystem I oxidation in terms of balanced electron flow from photosystem II to photosystem I was monitored using absorption changes at 830 nm on actinic background light in order to further substantiate the primary lesion in the incompatible hybrid (Figure 33). Although photosystem I signal intensity in AB-I was generally slightly reduced, the results clearly showed a much higher actinic light-induced oxidation state of photosystem I in AB-I compared to AB-III. At  $50 \mu\text{E m}^{-2} \text{sec}^{-1}$  about 15% of P700 was oxidized in AB-I whereas the reaction centre of photosystem I remained almost completely reduced in AB-III, a response likely due to limited electron flow towards photosystem I of a deficient photosystem II rather



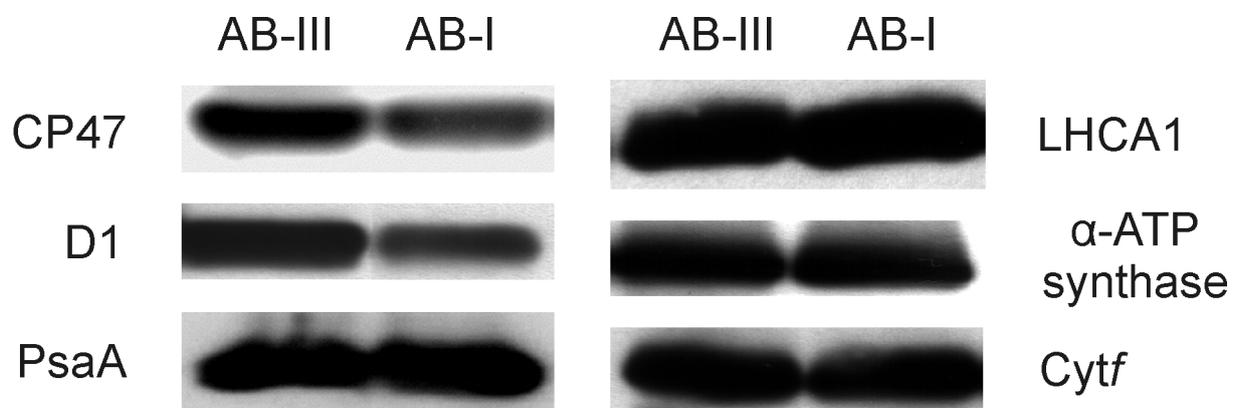
**Figure 33.** Studies on photosystem II yield and redox kinetics of photosystem I. Fluorescence induction kinetics of compatible AB-III and incompatible AB-I leaves. Fluorescence induction traces induced by saturating white-light pulses showed the maximal fluorescence raise during the light pulse ( $F_m$ ). The  $F_m$  levels were normalized to equal heights. Dark-adapted leaves were exposed to consecutive saturating light pulses during application of continuous actinic light (panel A). The P700 oxidized state of compatible and incompatible leaves exposed to  $50 \mu\text{E m}^{-2} \text{sec}^{-1}$  actinic red light was recorded. The signal level of AB-III was not affected after switching off the actinic light (downward black arrow) indicating that photosystem I is largely reduced. However, significant absorbance changes were recorded after switching off actinic light in AB-I indicating that a substantial part of photosystem I was oxidized. Application of FR light (upward open arrow) oxidized photosystem I in both compatible and incompatible leaves. Subsequent saturating light pulses (squiggled arrows) on the FR background light transiently reduced photosystem I completely in compatible but only partially in incompatible leaves. AA-I plants resembled in their photosystem II fluorescence and photosystem I redox characteristics those of AB-III (not shown) (panel B). Left: Phenotype of the incompatible hybrid AB-I with the genetic constitution  $h^{\text{johansen}} \cdot h^{\text{tuscaloosa}} P^{\text{oh}}$  (AB-I).

than of downstream effects in the AB-I incompatible hybrid. Furthermore, short light pulses at far-red background light were sufficient to completely reduce photosystem I in AB-III and AA-I, but not in AB-I, indicating a slower photosystem II mediated electron transport. Taken together, expression and chlorophyll fluorescence analyses are fully consistent with a decreased photosystem II activity relative to photosystem I and suggest that the primary

lesion in the incompatible AB-I hybrids affects photosystem II function. Photosystem I measurements were done in co-operation with Uwe Rauwolf.

#### 3.4.2.3.3. Western analysis of AB-I thylakoid membrane

Western analysis with appropriate antisera was performed to check whether the specific reaction of photosystem II can be confirmed at the protein level (Figure 34). Indeed, levels of the photosystem II proteins CP47 (gene: *psbB*) and D1 (gene: *psbA*) were significantly reduced whereas those of the photosystem I subunit PsaA, LHCA1, cytochrome *f* and of the ATP synthase subunit  $\alpha$  remained unchanged in AB-I leaves compared to those of AB-III.



**Figure 34.** Immunoblot analysis of various thylakoid membrane proteins in green AB-III and incompatible AB-I tissue. Levels of CP47 apoprotein and D1 polypeptide (photosystem II) are altered in AB-I tissue whereas those of components of photosystem I (PsaA, LHCA), ATP synthase subunit  $\alpha$  and cytochrome *f* of the cytochrome *b<sub>6</sub>f* complex are not changed.

#### 3.4.2.3.4. A deletion in the *clpP/psbB* spacer explains the AB-I phenotype

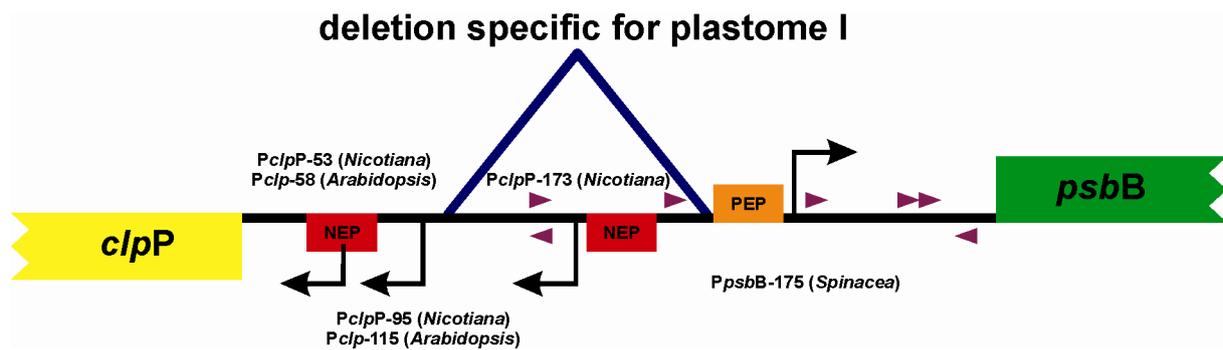
Since AB-I can be confined to a major effect on photosystem II, plastome  $I^{joh}$  was examined for specific variation, affecting components of photosystem II. Only two regions are present in plastome  $I^{joh}$  fulfilling this criterion: First, a small deletion in the promoter of *psbI* in the intergenic spacer of *psbI* and *psbK*, and second, a large deletion in the *clpP/psbB* spacer, affecting both, NEP and PEP promoters (Tables 23 and 24 and Chapter 3.4.2.3.5). The involvement of the *psbI* promoter in the AB-I phenotype is rather unlikely since, (i) mRNA levels of *psbI* are not changed (Geimer and Meurer, unpublished) and (ii) a knock-out of *psbI* shows no apparent phenotype (Schwenkert *et al.*, 2006). Consequently, only the 148 bp deletion at position 77,080 of plastome  $I^{joh}$  (accession number AJ271079.3) in the intergenic

*clpP/psbB* region, remained as the sole potential plastid component responsible of the AB-I incompatibility.

To confirm this assumption, all further 20 specific deletions of plastome I<sup>oh</sup> were investigated. They included also minor changes such as single nucleic acid exchanges in non-coding regions. Six of these regions are polymorphic within all plastomes. A contribution of these regions to the AB-I PGI phenotype is unlikely, since plastomes II, III and IV remain fully compatible in an AB background (Figure 4). This excludes the genes *accD*, *ycf1* and *ycf2*, the intergenic regions *rps16/rbcL* and *trn<sup>Q</sup><sub>UUG</sub>/accD*, as well as the SSC/IR<sub>A</sub> junction. 14 regions specifically altered in plastome I remain. All regions involving the NADPH complex can be disregarded, because knock-outs of individual NDH subunits in tobacco lack a conspicuous phenotype (Burrows *et al.*, 1998; Kofer *et al.*, 1998). Therefore, the length variance of NdhD and NdhF (Chapters 3.2.3.2.1 and 3.2.3.2.2) as well as of the intergenic regions *ndhG/ndhI*, *ndhI/ndhH*, *ndhF/rpl23* or *ndhF/trn<sup>N</sup><sub>GUU</sub>* can be excluded. Additional non-coding regions, such as the *atpI/atpH* spacer and a small deletion downstream of the *petN* stop codon are not involved in the AB-I phenotype, since neither the ATP synthase nor the cytochrome complex are affected in AB-I tissue (Figure 34). Specific variations left are small deletions in the intron of *rpl16*, the 5' region of *rpl32*, in the *trn<sup>G</sup><sub>UCC</sub>/trn<sup>S</sup><sub>GCU</sub>*, *trn<sup>L</sup><sub>UAA</sub>/trn<sup>T</sup><sub>UGU</sub>*, and *rps8/rpl14* spacers, and a single amino acid exchange in Rpl32. They are not of functional relevance either, since the translational apparatus is not notably affected in AB-I (Figure 34).

#### 3.4.2.3.5. Bioinformatic and phylogenetic analysis of the *clpP/psbB*-spacer

To test the assumption of an involvement of the *psbB/clpP* spacer region in the AB-I phenotype, a phylogenetic footprinting analysis of the intergenic region between the two genes delimited distinct boxes in the promoter region described for *clpP* (Sriraman *et al.*, 1998) and *psbB* (Westhoff, 1985; Westhoff and Herrmann, 1988) and uncovered substantial changes to *Oenothera* and among its plastomes. Notably, a large deletion in plastome I<sup>oh</sup> directly upstream of the highly conserved *psbB* PEP promoter and the *clpP* 5' region eliminated a putative and a confirmed *clpP* promoter and two putative *psbB* promoters (Figure 35). Conservation as well as similarity to known sites in other species suggested that an element missing in plastome I<sup>oh</sup> caused the specific adaptation of plastome I<sup>oh</sup> to the AA genome, and that the change of this region leads to an altered *psbB* and/or *clpP* transcription in the AB-I combination.



**Figure 35.** Schematic overview of the *clpP/psbB* spacer region in *Oenothera*, *Spinacea*, *Nicotiana*, and *Arabidopsis*. Positions of the indicated transcription start sites (black arrows) of NEP and PEP promoters (*PclpP* and *PpsbB*) relative to the start codons were determined experimentally in *Arabidopsis*, *Nicotiana* and *Spinacea* (Westhoff, 1985; Hajdukiewicz *et al.*, 1997; Sriraman *et al.*, 1998; Swiatecka-Hagenbruch *et al.*, 2007). Putative, not experimentally verified promoters in *Oenothera* are marked with filled triangles. The experimentally verified *PclpP*-173 and *PpsbB*-175 are highly conserved and confirmed bioinformatically in *Oenothera* and all reference species. The deletion (open triangle) is not present in *Oenothera* plastomes II<sup>suavG</sup> – V<sup>doul</sup> or plastomes of reference species and is therefore specific for plastome I<sup>oh</sup> in *Oenothera*.

#### 3.4.2.3.6. Expression analysis of *psbB* and *clpP*

To estimate the functional relevance of the identified deletion, expression of *psbB* and *clpP* relative to *psaC* was compared by quantitative real time RT-PCR between AB-I and the compatible (wild-type) AA-I and AB-III constitutions. *PsaC* was chosen for normalization, since no expression difference could be detected between the two genotypes in macroarray analysis (Geimer and Meurer, unpublished).

Indeed, transcript levels of *clpP* were increased approximately two-fold (1.98 +/- 0.018), but those of *psbB* were down-regulated three- to six-fold (0.32 +/- 0.008 to 0.14 +/- 0.034) in AB-I<sup>oh</sup> compared to AB-III<sup>tusca</sup> or AB-III<sup>lam</sup>. Down-regulation to about 0.23 +/- 0.056 compared to AA-I<sup>oh</sup> proves that low *psbB* expression in AB-I<sup>oh</sup> is not a matter of plastome I<sup>oh</sup> alone, but result of its disharmonic interaction with the B genome. The data confirm the functional relevance of the deletion. In consequence, the AB-I phenotype can be explained as a likely result of changes in *psbB* expression, since *Arabidopsis* mutants predominantly affected in *psbB* expression display comparable fluorescence characteristics as observed for AB-I in Chapter 3.4.2.3.1 (Meurer *et al.*, 1996; Meurer *et al.*, 2002).

#### 3.4.2.3.7. *The clpP/psbB spacer of various subplastomes*

If a specific deletion of the *psbB/clpP* spacer region causes the AB-I phenotype, the deletion must be present not only in subplastome I of the strain johansen, but also in other plastome I variants, which all behave in a comparable way genetically. For references describing the genetic behaviour of the investigated subplastomes see Table 6. The region was amplified using the primer pair VP9for and VP10rev. Sequencing the derived PCR products in the strains chapultepec, cholula, puebla, toluca, franciscana de Vries, franciscana E. & S., hookeri de Vries, and bauri Standard, all harbouring plastome I, uncovered no polymorphism to plastome I of strain johansen (accession numbers EU449954 - EU449961, and AJ271079.3). The same result was obtained in variants of plastome II, III and IV. Strains with plastome II (nuda Standard and suaveolens Grado, accession numbers EU449962 and EU262889), plastome III (rr-lamarckiana Sweden and tuscaloosa, accession numbers EU262890 and EU449963) and plastome IV (ammophila Standard and atrovirens Standard, accession number EU449964 and EU262891) were all found identical for the investigated region to the corresponding plastomes.

## 4. Discussion

Traditional work on *Oenothera* was largely restricted to classical genetic and cytogenetic approaches (Cleland, 1972; Harte, 1994). The present thesis describes attempts to establish *Oenothera* as model for speciation, especially for PGI in molecular terms, aimed to investigate the phylogenetic dynamics of plastome-genome interaction and genome changes on a microevolutionary scale and their role in speciation processes. To achieve this aim (i) a co-dominant marker system for both genetic compartments, the plastome and the nucleus, was established. The markers substantially simplify the analysis of hybrids and plastome-genome incompatible combinations in the genus. In parallel, (ii) the sequence of the five basic plastid genomes was completed and evaluated to search for potential plastidic determinants causing PGI. A novel combination of classical genetic, bioinformatic and molecular approaches allowed in a case study (iii) the identification of a plastid determinant responsible for the plastome-genome incompatible combination AB-I. The work presented also resolves various technical problems, since application of molecular approaches is not trivial for *Oenothera*. The genus contains unusual high amounts of mucilage and polyphenols that adversely affect isolation of cellular components and enzymatic reactions. Optimized protocols described in this work allow reproducibly an efficient isolation of high quality DNA, RNA and thylakoid membrane proteins from a wide range of tissue, including adult plants.

### 4.1. Benefit of co-dominant markers to *Oenothera* breeding

Molecular co-dominant markers are of intrinsic interest in *Oenothera* breeding but were largely missing (Mráček, 2005; Larson *et al.*, 2008). They are of importance for assembling plastome-genome incompatible plants, for genome restructurations, but also for commercial *Oenothera* breeding (Fieldsend, 2007). Until now, neither basic nor subplastome types could be distinguished without time consuming RFLP analyses (Herrmann *et al.*, 1980; Gordon *et al.*, 1981; 1982; Chapman *et al.*, 1999). In crossing experiments to exchange plastomes between species, different plastome types were marked using bleached plastome mutants. Time consuming crossing steps were necessary to establish/remove these mutants during such crossing programs (Stubbe, 1960; 1989). The designated *rrn16-trnI<sub>GAU</sub>* marker allele (Table 9) allows an easy monitoring of all basic plastome types and a large variety of subplastomes. The use of plastome mutants in crossing experiments to exchange plastomes is now dispensable. The marker allele reduces the number of crosses significantly.

Substantial progress was also made for the nuclear genome with the marker allele M40 (Table 8). Crosses, involving ring forming hybrids, can now be monitored for all basic genotypes and a large number of individual Renner complexes. The major benefit is the possibility to screen splitting generations of different Renner complex combinations already at the seedling stage by a single PCR and marker. Up to date, only mature plants could be screened, using phenotypic markers. The advantages of the molecular approaches are obvious, especially for annual herbs, such as *Oenothera*.

Further co-dominant markers presented in Table 7 were applied to *Oenothera* genomes. Surprisingly, already minor divergence at the nucleotide level was appropriate to generate CAPS markers from almost all genes studied (~ 60% at average of only 300 bp). From altogether 34 markers, 22 could be mapped to the seven coupling groups of the hybrid <sup>h</sup>johansen-<sup>h</sup>tuscaloosa, representing chromosomes 1·2 3·4 5·6 7·10 9·8 11·12 13·14 of the classical *Oenothera* map (Greiner and Rauwolf, unpublished). With the nuclear marker M58 it was possible to correlate chromosome 9·8 with linkage group 7 of the molecular linkage map (Chapter 3.1.3.3). This is a breakthrough in *Oenothera* genetics, since it combines 100 years of classic genetic research on *Oenothera* chromosomes with modern, molecular approaches. A complete merge of the classical and molecular map is essential and will be possible in future *Oenothera* breeding (Chapter 4.2).

With the classical map prediction of the segregation behaviour in superlinkage groups, involving single or multiple chromosomes is possible (Cleland, 1972 and Chapters 1.7 and 3.1.3). It allows a directed exchange of single chromosomes at the diploid level between genomes. Consequently, single chromosomes can be specifically placed in a breed without disturbing the combination of characters located in the remaining genome (Figure 17). This can be applied for example to study the impact of single linkage groups on genome evolution in *Oenothera* but also to introduce single chromosomes with particular traits into commercial cultivars. So far, monitoring of such crosses was realizable only with phenotypic markers, which is generally difficult. Some chromosomes in many strains even lack phenotypic markers (Cleland, 1972; Fieldsend, 2007). Therefore, molecular markers have to be assigned on the classical map. Mapping of molecular markers to the classical map provides immediate access to (i) various phenotypic markers located on the classical *Oenothera* map, and (ii) to 300 strains, for which the chromosomal formulas are known (Cleland, 1972).

In summary, the markers described in this work represent a significant progress in *Oenothera* genetics. They allow a precise and easy molecular identification of plastomes, Renner complexes and single chromosomes in crossing programs. They will also have an impact on commercial *Oenothera* breeding.

#### 4.2. A complete alignment between the classical and molecular *Oenothera* maps

None of the molecular nuclear markers described so far were assigned to the classical *Oenothera* map. Assigning linkage group 7 to chromosome 9·8 (Chapter 3.1.3.3) is therefore a breakthrough for future breeding approaches. The same approach leading to the identification of chromosome 9·8 as coupling group 7, can be applied for all other chromosome pairs. The prerequisite for such an approach, different hybrids of <sup>h</sup>johansen or <sup>h</sup>tuscaloosa with various Renner complexes, having each of the seven chromosomes as single, free and identifiable pairs, is available.

To give an example: The hybrid <sup>St</sup>laxans·<sup>h</sup>tuscaloosa possesses the chromosome configuration ⊙4, ⊙4, ⊙4, 1 pr.; chromosome 1·2 is free and the only bivalent (1 pr.). Segregation analysis with molecular markers in the F<sub>2</sub> generation should display four coupling groups, three consisting of four chromosomes (⊙4), and one with a single chromosome, identical to chromosome 1·2. All further chromosomes, following the same principle, can be assigned using, e.g. the hybrids <sup>St</sup>pingens·<sup>h</sup>johansen (chromosome 3·4), <sup>h</sup>johansen·<sup>r-S</sup>gaudens (chromosome 5·6), <sup>h</sup>johansen·<sup>St</sup>percurvans (chromosome 7·10,) <sup>h</sup>tuscaloosa·<sup>St</sup>undans (chromosome 11·12), and <sup>Th</sup>tingens·<sup>h</sup>tuscaloosa (chromosome 13·14).

If all seven coupling groups of <sup>h</sup>johansen·<sup>h</sup>tuscaloosa were associated with chromosomes, in a next step, *identification of chromosome arms* can be addressed. A possible hybrid to work with is <sup>St</sup>albicans·<sup>h</sup>cholula. It identifies the freely segregating pair 1·4. Since chromosome 1·2 was already identified from <sup>St</sup>laxans·<sup>h</sup>tuscaloosa, as outlined above, all markers, assigned to chromosome 1·2 in <sup>St</sup>laxans·<sup>h</sup>tuscaloosa, which also map to chromosome 1·4 of <sup>St</sup>albicans·<sup>h</sup>cholula must be part of arm 1. Consequently, the remaining markers on 1·2 of <sup>St</sup>laxans·<sup>h</sup>tuscaloosa, which are not part of coupling group 1·4 in <sup>St</sup>albicans·<sup>h</sup>cholula, map on arm 2. Conversely, if arm 1 will be identified in chromosome 1·4, also arm 4 will be characterized. Since arm 4 is characterized, it becomes possible, in turn, to assign arm 3 on chromosome 3·4.

The principle can be applied to all further chromosome arms. To identify arms 5, 6, 8 and 9, e.g. the hybrid <sup>h</sup>cholula·<sup>St</sup>percurvans could be used, which has chromosome 6·8 as free pair. In accessing the rich source of analyzed *Oenothera* strains, determination of all chromosome arms is readily possible.

### 4.3. Organization and relationship of the *Oenothera* plastome sequences

Besides a basic molecular characterization of the nuclear genome the complete sequences of the five basic plastome types in the subsection *Oenothera* was a major requirement to establish the material as a molecular model. Plastid chromosomes, because of their endosymbiotic ancestry, limited coding potential, relatively conserved organization and well-defined structure, provide a unique source of information to address a wide range of fundamental questions.

#### 4.3.1. *Oenothera* particularities of the plastome sequences

The five basic *Oenothera* plastomes have a coding potential of 113 unique genes, which is nearly identical and comparable to those of plastid chromosomes of vascular plants in size, organization, gene clustering and conservation. They deviate from the “ancestral form” of plastid chromosomes by a single kilobase-magnitude inversion in the large single-copy segment. The *Oenothera* plastomes belong to the largest plastid chromosomes known from vascular plants and are moderately larger than those found generally for higher plants. Usually plastid chromosomes range in size between 130 and 160 kbp (reviewed in Palmer, 1992). Genes are well definable, except of the N-terminus of *cemA* (*ycf10*) encoding an inner envelope polypeptide involved in CO<sub>2</sub> uptake. The annotations of its N-termini were not consistent, even after comparison with *cemA* loci from 50 reference species (Chapter 3.2.1.2). The locus does not appear to be a pseudo gene, as judged from PCR and Western analysis. Also no evidence was found for a nuclear copy of this gene. This point remains to be settled in general, not only in *Oenothera* but for all plastid chromosomes.

#### 4.3.2. General divergence and repeat analysis

Despite their gross conservation and close relationship *Oenothera* plastid genomes are remarkably diverged, compared, for instance, to those of the Solanacean taxa *Atropa belladonna* and *Nicotiana* (reviewed in Palmer, 1992; Schmitz-Linneweber *et al.*, 2002). Point mutations including changes in restriction sites, repetitions, insertions/deletions and

inversions occurred within all five plastomes. They are scattered all over the circular chromosome, but found with different frequency in different parts of the plastome sequences (Figure 20). Probably due to copy-correction of the repeated segments (*e.g.* Wolfe *et al.*, 1987), somewhat less changes within the five plastomes are expectedly found in the IR. Insertions/deletions and repetitions, in direct or inverse orientation, are relatively frequent (Wolfson *et al.*, 1991; Hupfer *et al.*, 2000). Illegitimate recombination, repetitive and divergent regions are often correlated, suggesting a role of repeats in generating divergent regions. Slipped mispairing may cause or contribute to changes in a number of instances as well (Winter and Herrmann, 1987; Wolfson *et al.*, 1991; Sears *et al.*, 1996). Generally, coding sequences are well conserved among the plastomes. Eight genes, almost exclusively non-photosynthetic, differ by indels. Nearly a dozen genes differ by repeats including plastome-specific changes. Although non-protein and non-RNA coding regions evolve faster than genes, they still are quite well conserved. This indicates that functional elements within these sequence intervals are conserved and compactly bundled. Almost invariant are the junctions of the IR regions that often cause size differences between spermatophyte plastid chromosomes (*e.g.* Goulding *et al.*, 1996).

#### 4.4. Evolutionary considerations based on the plastome sequence

The complete sequences of the five *Oenothera* plastomes as well as the study of the occurrence of *Oenothera* subplastomes in different populations and species (Chapter 3.1.1.3) also contributed to solve several open and long standing questions in *Oenothera* genetics.

##### 4.4.1. Divergence time of the five basic plastomes

The age and divergence time of the subsection *Oenothera* is of particular interest and not satisfactorily solved. The common ancestor of the subsection *Oenothera* originated presumably from Mexico and Central America and invaded the North American continent in several waves (Cleland, 1972). Its closest recent relative, *Oe. maysillesii*, a species of the subsection *Emersonia* (Stubbe and Raven, 1979), is still found within this area (Dietrich *et al.*, 1985). Although models of speciation and colonization history proposed by Cleland (1972) were modified accounting to more recent data (Dietrich *et al.*, 1997), time frame of radiation in the subsection *Oenothera* is still unclear.

In the comprehensive literature, speculations about time scales are found only in Cleland (1972, pp. 299 - 302) and Levy and Levin (1975). In fact, to answer this question with customary approaches is difficult, since fossil records are missing. However, geological and general considerations let Cleland to assume that the origin to the four ancient *Oenothera* populations, which constitute his colonization model, may be correlated with the four major periods of glaciations in the Pleistocene (Nebraskan, Kansan, Illinoian and Wisconsin). This assumption determined the age of the subsection to about one million years or probably little more. Nevertheless, Cleland (1972) does not rule out the possibility that an origin of the subsection can be placed earlier, at the beginning of the Pleistocene, since all species for instance are still crossable, except that they display PGI. Therefore, Cleland suggests it more likely that the divergence of the subsection *Oenothera* started at the beginning of the Wisconsin, some 70,000 years ago.

Levy and Levin (1975) do not agree with Cleland's assumption of a very young subsection *Oenothera*. Their rough estimation of divergence time based on nuclear allozyme variation substantiates Cleland's first surmise of a possible age of the subsection of about 1 mya. In accord with Cleland they assume that *Oe. argillicola* (plastome V) diverged as basal taxon from the rest of subsection. However they disagree with the idea that subsequent evolution of the following *Oenothera* species happened more or less equidistant in time due to the major glaciation periods.

The data about the divergence time of the five plastomes presented in this thesis (Table 19) seem to place the appearance of the subsection *Oenothera* to the middle of the Pleistocene, about one million years ago. However, they are contrast to Levy and Levin (1975) calculated time frames of subsequent evolution. The data of this thesis support Cleland's equidistant evolution model. The model is substantiated by the fact that time scales of divergence between the five plastome types correlate roughly with the major periods of glaciation.

Although the absolute calibration of the molecular clock derived from the plastid sequences may require correction, especially since there seemed to be a high, probably temporary adaptation pressure onto *Oenothera* plastomes (Chapters 3.3.4 and 4.5.1), which distorts molecular clock calculations, the more than tenfold difference of divergence times of the five plastomes will remain (Table 19). Nevertheless, the calculated divergent time fits realistically

into geological reality, natural history and population structure of the subsection (Cleland, 1972; Dietrich *et al.*, 1997).

Time frame of plastome evolution does not necessarily correlated with that of the nucleus in the subsection *Oenothera*. Many of the taxa under study are permanent translocation heterozygotes and arose from hybridization events (Dietrich *et al.*, 1997). Plastomes and nuclear genomes in *Oenothera* species may have therefore diverging evolutionary histories. For example plastome IV and the basic nuclear genome C in *Oe. oakesiana* (AC-IV) and *Oe. parviflora* (BC-IV) are monophyletic, whereas their nuclear genomes A and B are not (Dietrich *et al.*, 1997). It would be therefore interesting to substantiate the findings of Levy and Levin (1975) with more recent methods of molecular clock calculation, to estimate the divergence time of the nuclear genomes in the subsection *Oenothera*.

#### 4.4.2. *The phylogenetic tree of the five Oenothera plastomes*

Divergence time and natural history of the subsection *Oenothera* was also reflected by phylogenetic trees. Calculations using the entire sequences of the five plastomes as data set, lead to two different trees (Figure 29). Plastomes I<sup>loh</sup>, II<sup>suavG</sup> and III<sup>lam</sup> were always found in one clade, with plastomes I<sup>loh</sup> and II<sup>suavG</sup> as closest relatives, but trees were different for plastomes IV<sup>atro</sup> and V<sup>dou1</sup>. NJ and MP supported a tree with plastome IV<sup>atro</sup> in an own clade and added V<sup>dou1</sup> to the clade of I<sup>loh</sup>, II<sup>suavG</sup> and III<sup>lam</sup>. ML generated a new clade with its members IV<sup>atro</sup> and V<sup>dou1</sup> (Chapter 3.3.1). The first tree is better supported, but interpretation of classic genetic data suggests that the second tree is probably more likely (Stubbe and Steiner, 1999). However, both trees fit into the natural history and the models of successive colonization of the North American continent, proposed for the genus (Stubbe, 1963a; 1964; Cleland, 1972; Dietrich *et al.*, 1997; Stubbe and Steiner, 1999). A possible solution to the two tree variants could be the sequence of a much closer related outgroup to the five *Oenothera* plastomes than *Lotus*. Rooting the plastome tree of the subsection *Oenothera* with its next relative, *Oe. maysillesii*, subsection *Emersonia* (Stubbe and Raven, 1979), may satisfactorily resolve the relation between the basic plastome types IV and V.

#### 4.4.3. *The large inversion of 56 kbp*

Also consistent with the natural history of the genus *Oenothera* is the occurrence of a 56 kbp inversion in the large single-copy region of the plastid chromosome. It provides a robust

phylogenetic marker, since it is shared among all *Oenothera* plastomes. It is absent in the closely related South American subsections *Munzia* and *Raimannia* (Hachtel *et al.*, 1991) as well as in the sister group *Epilobium* (Schmitz and Kowallik, 1986). This observation and its almost identical endpoints in all five plastomes suggest that the inversion is not caused by one of the proposed rare parallel inversions (Downie and Palmer, 1994; Johansson, 1999; Tsumura *et al.*, 2000). It has arisen monophyletically within the *Oenothera* clade and late in the history of the Onagracean complex, in the common ancestor of the subsection. Hence, the inversion marks a recent split in the history of the genus and predates the divergence of the subsection.

#### 4.4.4. Patterns of subplastome variation in *Oenothera* populations

Other regions in the *Oenothera* plastomes reveal a higher phylogenetic resolution than the large inversion. The *rrn16-trnI<sub>GAU</sub>* spacer region, used as marker allele for the identification of basic and subplastome types in crossing experiments (Chapters 3.1.1.3, 3.1.2.1 and 3.1.2.2), is also an indicator for gene flow and recent hybridization events within the subsection *Oenothera*. The 17 alleles of the derived CAPS marker (Table 9) give an interesting spotlight on the phylogenetic relationship between plastomes and species. It could be shown that the region is suitable for such a phylogenetic analysis (Hornung *et al.*, 1996; Sears *et al.*, 1996). The marker is an indicator how much variation or subplastomes exist within the subgenus, without performing laborious RFLP analysis of the whole plastid chromosome (Herrmann *et al.*, 1980). Nevertheless, the degree of variation is important to know, if differences between five sequenced reference plastomes, described in this thesis, are going to be generalized in terms of speciation events.

The *rrn16-trnI<sub>GAU</sub>* allele provides a good confirmation of the monophyletic origin of plastome IV in *Oe. oakesiana* (AC-IV) and *Oe. parviflora* (BC-IV). The current model claims that *Oe. parviflora* (BC-IV) arose as a hybrid between a hypothetical ancestor, with the genomic constitution CC-IV, and *Oe. nutans* (BB-III). *Oe. parviflora* (BC-IV) itself hybridized with *Oe. biennis* (AB-II or BA-III) resulting in the AC-IV species *Oe. oakesiana* (Dietrich *et al.*, 1997). Therefore, plastome IV of both species should be identical and monophyletic, which could be confirmed at the molecular level with the described marker allele IV<sub>1</sub>. All strains tested for plastome IV had an identical sequence (Table 9), suggesting that plastome IV was already fixed in the population of the hypothetical ancestor CC-IV.

For plastomes of type I, the clearly distinguishable and specific alleles of *Oe. elata* subsp. *elata* (AA-I<sub>1/2</sub>), *Oe. elata* subsp. *hookeri* (AA-I<sub>3</sub>, AA-I<sub>4</sub>) and *Oe. villosa* subsp. *villosa* (AA-I<sub>5</sub>) (Table 9) indicate that gene flow between these three populations no longer occurs. This is likely due to their geographic distribution and propagation strategies (Dietrich *et al.*, 1997). Although these species still carry genetically plastome I, the subplastomes seem to evolve in different directions. This is a clear example that *Oenothera* provides a suitable material not only to study speciation, but also pre-speciation processes.

Cleland (1962) mentioned that the border between plastome type II and III in *Oenothera* strains, naturally occurring at the North American continent, is not sharp. The genetic behaviour of these plastomes in AB, BA and BB-species was somehow in between the plastomes of types II and III of Stubbe (1959). The distribution of *rrn16-trnI<sub>GAU</sub>* alleles in different *Oenothera* strains of *Oe. biennis* (AB-II and BA-III), *Oe. glazioviana* (AB-III), *Oe. grandiflora* (BB-III) and *Oe. nutans* (BB-III) is consistent with the observation of Cleland (1962). Although for nearly all strains tested the genetic plastome type is known but no specific sequence for the *rrn16-trnI<sub>GAU</sub>* spacer region could be found for plastomes II or III (Table 9). If the evolutionary history of the above mentioned species is considered, nearly all of them arose from hybridization events and there is more than one model present in the literature, how these species may have evolved (Cleland, 1972; Wasmund, 1990; Schumacher and Steiner, 1993; Dietrich *et al.*, 1997). The data presented in this work indicate that there is still substantial gene flow within the genotypes under study, since alleles *rrn16-trnI<sub>GAU</sub>* II/III<sub>1</sub> and II/III<sub>3</sub> are present in more than one species. This is of particular interest, because the plastome appears to play an inferior role in establishing hybridization barriers between *Oe. grandiflora* (BB-III), *Oe. nutans* (BB-III) and *Oe. biennis* (AB-II or BA-III), since the majority of all possible crosses between the three species result in compatible offspring (Figure 4). The data support the hypothesis that the plastome is the only strong hybridization barrier in the genus (Chapter 1.3), and if it is not very pronounced or missing, gene flow can occur in hybridization zones.

Detecting so many different *rrn16-trnI<sub>GAU</sub>* alleles of plastome II/III in the same species, but also identical ones in different species (Table 9), suggests that plastomes II or III are not fixed so far in the different *Oenothera* populations. Consequently, their genetic behaviour is not fixed as well. This suggestion is conducted by the findings of Wasmund (1980) and Drillisch (1975) that the *biennis*-1 strains shuswap lake, mc call, birch tree 1, birch tree 2, citronelle

and delaware are associated with plastome II, but the *biennis-2* strain micaville carries plastome III. Usually, strains of the microspecies *biennis-1* and *biennis-2* in Cleland's taxonomy, are associated with plastomes III and II, respectively (Cleland, 1972). The data indicate that plastome types II and III in AB, BA or BB species still adapt to their nuclear background, offering the possibility to study very recent, ongoing evolutionary processes.

In summary, plastomes in *Oenothera* are an excellent tool to monitor hybridization events and their impact for speciation. Functional genetic analysis of fast evolving parts of the *Oenothera* plastome, related to physiological characterization of plastome-genome incompatibility in the genus should allow to draw a complete and definite picture of the evolution and mechanisms of speciation in the North American subsection *Oenothera*.

#### 4.5. Selection pressure and determinants of PGI

What causes the differences between the *Oenothera* plastomes and what are the functional molecular determinants responsible for PGI? Incompatibility between nuclear and plastid genomes can lead to hybridization barriers of different strengths, with remarkable impact on speciation (Chapter 3.3.3), but which selection forces produce them?

##### 4.5.1. Selection pressures acting on the plastome

That strong selection forces can act on the plastome is evident and well studied in *Oenothera*. In this thesis, bioinformatic comparisons of amino acid substitutions rates in the five basic *Oenothera* plastomes uncovered remarkably higher mean  $K_a/K_s$  values, compared to averaged values found in angiosperms (Chapter 3.3.4). This indicates a substantial selection pressure on *Oenothera* plastome sequences. Similar results were recently obtained for *clpP* in *Silene* and other genera (Erixon and Oxelman, 2008). Remarkably, the genus *Silene* also displays PGI phenotypes (Table 2).

The consequence of selection on a certain plastid type is chloroplast capture, the introgression of a plastid of one species into another one. This can happen relatively quickly, already if a small fitness advantage exists for the female (Tsitroni *et al.*, 2003). The *Triticeae* tribe, with its members *Triticum* and *Hordeum*, represents a well studied example. In its interspecific hybrids strong preferences for distinct plastid genomes can be found (Redinbaugh *et al.*, 2000). The outlined examples unequivocally demonstrate selection on plastomes, which can

be monitored in nature, as in *Oenothera*, *Silene* or *Triticaceae*, and has been described by the theoretical models of chloroplast capture.

A major selection force acting on plastomes is obviously related to the photosynthetic process. It is one of the most important functions of the organelle and photosynthesis itself is the principal energy supplying reaction of a plant cell that influences important processes such as water balance and drought tolerance (see below). Indeed, differences in photosynthetic performance have been detected with different “cytoplasms” in introgression lines of *Triticum* and *Aegilops* (Iwanaga *et al.*, 1978). A similar finding was reported from *Oenothera*, in which different photosystem II yields in green plants could be genetically linked to two different compatible plastomes in the same nuclear background (Glick and Sears, 1994). Cold stress acting on photosynthesis could be excluded as a selection force (Dauborn and Brüggemann, 1996). Thus, although the coding potential of closely related plastomes is almost identical (Chapter 3.2), photosynthetic efficiency conveyed by closely related plastome types may differ, since the adaptation to their nuclear components is different. It is obvious, that these differences are result and subject of selection.

Regardless of whether the origin of genetic differences of photosynthesis traits is located in nuclear or in plastid genomes, these differences are found in natural populations and are influenced by selection. Higher photosynthetic rates are usually associated with growth and fitness advantages (Arntz and Delph, 2001). Selection often acts indirectly on photosynthetic traits *via* drought or oxidative stress. It can be monitored, *e.g.* by water-use efficiency or photosynthetic parameters (Arntz and Delph, 2001; Hura *et al.*, 2007; Yang *et al.*, 2007). The link of photosynthesis to drought stress consequently implies that periods of climate changes may introduce PGI and therefore built hybridization barriers resulting in speciation. *It will be a fascinating future task to examine this hypothesis.* For the genus *Oenothera*, drought stress is a very likely selection pressure for speciation, since appearance of the subsection *Oenothera* was accompanied by a fluctuating climate for both, precipitation and temperature, during the Pleistocene (Cleland, 1972).

Another driving force could reside in the intrinsic structure of the plastid genome. Genes like *accD* and *ycf2* are highly divergent among all plant lineages and also between closely related taxa as they are located at so called “hot spots of divergence” in plastid chromosomes (Herrmann *et al.*, 1980; Salts *et al.*, 1984 and Figure 20). The driving force towards increased

evolution rates in these regions could act on the DNA structure itself and perhaps provide reasons for PGI.

#### 4.5.1. Genetic architecture of PGI

At present it is difficult to assess how many nuclear genes or plastid targets are involved in distinct incompatible combinations. Molecular data are rare and incompatible F1 hybrids are sometimes sterile or exhibit reduced fertility (Table 2) complicating genetic analysis of the nuclear genome. At least for *Acacia*, *Pelargonium*, *Pisum* and the BB-II combination in *Oenothera* monogenic nuclear determinants were reported (Smith, 1915; Stubbe, 1953; Moffett, 1965; Bogdanova and Berdnikov, 2001). The two instances, from which the plastid determinants for PGI are known, are monogenic. Mutation of just a single nucleotide causes bleaching in the cybrid with the nuclear genome of *Atropa* and the plastome of *Nicotiana* (Schmitz-Linneweber *et al.*, 2005). Data from the incompatible *Oenothera* AB-I hybrid, in turn, suggest that a single regulatory region of plastome I causes compartmental incompatibility (Chapter 3.4.2.3). On the other hand, other instances indicate a more complex genetics. Segregation analysis in *Zantedeschia*, for example, uncovered two or three nuclear loci (New and Paris, 1967; Yao and Cohen, 2000), as it is also the case for different PGIs of *Oenothera* (van der Meer, 1974; Jean, 1984; Rauwolf and Greiner, unpublished). In evening primroses, the phenotypic marker *lor* combined with an incompatible plastome is involved in embryo lethality (Renner, 1943c), but only if linked to marker *Fl*. Only a *lor lor Fl Fl* genotype leads to embryo lethality. Individuals homozygous for *lor* alone combined with a foreign plastome suffer only from chlorophyll deficiency. This example illustrates that nuclear genes involved in compartmental incompatibility can be dissected and separated genetically and are amenable to molecular mapping approaches.

Polygenic nuclear determinants do not necessarily claim for multiple plastid targets. Except for RbcL, all plastome encoded polypeptides are part of multisubunit assemblies that consist of more than one nucleo-plastidic gene pair and plastid encoded polypeptides may interact with more than one partner polypeptide of nuclear origin (Race *et al.*, 1999). Thus, altered expression or function of one component in the plastid may generally affect accumulation and/or function of more than one nuclear encoded protein, since regulation of nuclear-plastidic networks occurs at various levels.

#### 4.5.2. *Molecular determinants of PGI suggest regulatory phenomena*

Obviously, PGI could basically reflect a regulatory and/or a structural phenomenon. Diverging loci are therefore of intrinsic interest, but the overwhelming majority of such loci found in the plastome sequences of *Oenothera* do not seem to be of functional relevance nor causative for interspecific compartmental incompatibility. Diverging loci with non-synonymous amino acid replacements or length polymorphisms with or without frame shifts (Table 22) presumably play only a minor role in PGI, if any, since all changes occurred in parts of polypeptide chains, which are generally highly variable and do not affect conserved domains (Chapter 3.2.3.2). In terms of single amino acid exchanges, within all 388 detected non-synonymous substitutions in polypeptides, only 35 sites indicate possible biochemical differences in altogether 19 genes, namely *accD*, *atpA*, *atpB*, *atpF*, *ccsA*, *clpP*, *matK*, *ndhA*, *ndhB*, *ndhC*, *ndhD*, *ndhE*, *ndhH*, *petB*, *rpoB*, *rpoC2*, *rps3*, *rps8*, and *rps15*. However only 25 of these altered amino acid sites, are located in functional domains (Table 21), and could contribute to PGI, if at all. Many of them are unlikely to contribute. For instance, variance of RpoB between *Nicotiana* and *Atropa* did not turn out to influence PGI notably (Herrmann *et al.*, 2003; Schmitz-Linneweber *et al.*, 2005) and gene deletions in *Nicotiana* suggest that disruption of genes for subunits of the NADH dehydrogenase complex (Burrows *et al.*, 1998; Kofer *et al.*, 1998) are not of relevance either. Furthermore, a biochemically possibly significant amino acid change in *petD* fits with the compatibility chart in principle (Figure 4), but appears to be neutral, because BB-II does not show a cytochrome phenotype (Rauwolf and Meurer, unpublished). All in all, plastome encoded genes for structural components of the photosynthetic machinery are usually highly conserved and it appears that compartmental co-evolution in *Oenothera* causing hybrid bleaching influences predominantly regulatory processes.

In other DMIs investigated gene regulation plays an important role (Ellison and Burton, 2006; Haerty and Singh, 2006; Ortiz-Barrientos *et al.*, 2007). Also a major involvement of RNA metabolism in PGI is likely, since compared to cyanobacteria RNA metabolism of photosynthetic genes in higher plants was substantially changed, but the structural components are highly conserved (Herrmann and Westhoff, 2001; Liere and Börner, 2006 and citations therein). RNA metabolism is a rapidly evolving process. Apart from editotype differences in *Arabidopsis* (Tillich *et al.*, 2005), ecotypes of *Arabidopsis* may show different 5' and 3' ends of mitochondrial mRNA (Forner *et al.*, 2007), and in the case of *Zea mexicana*

(*Teosinte*) and *Zea mays*, an altered gene regulation rather than gene function is a driving force in speciation (Doebley, 2004).

Available data on PGI are commensurate with these observations. Bleached sectors of *Passiflora* hybrids (Figure 2), containing only maternal plastids, appear to possess an altered plastid gene regulation, and fail to accumulate ribosomal RNAs (Mráček, 2005). Studies on nuclear gene expression in the three basic homozygous *Oenothera* genotypes (AA-I, BB-III and CC-V) uncovered different transcription profiles especially for photosynthetic genes (Chapter 3.4.2.2), indicating that species-specific gene regulation could be involved in PGI. Altered expression patterns in incompatible *Zantedeschia* hybrids also fit to this picture (Yao and Cohen, 2000). In one instance, nuclear genome dosage was probably responsible for overcoming PGI in *Rhododendron* (Sakai *et al.*, 2004), again supporting the idea that altered gene regulation, rather than structural aspects of the protein machinery, is frequently responsible for PGI. This idea is consistent with the two cases, from which the plastidic PGI determinants are known. In both cases, RNA metabolism causes or is involved in PGI, respectively. In the *Atropa/Nicotiana* cybrids it is based on RNA editotype differences, which also explains the substantial phenotypic difference of the reciprocal cybrids (Schmitz-Linneweber *et al.*, 2001b; Schmitz-Linneweber *et al.*, 2005), and in the AB-I cybrid of *Oenothera* the intergenic region between the divergently transcribed *clpP* and *psbB* operons appears to play a major role in the incompatible phenotype (Chapter 3.4.2.3).

A possible reason for this delimitation is found in the photosynthetic and translation machineries, which are both of dual genetic origin (Herrmann *et al.*, 2003). Structural components were streamlined and optimized during billions of years and hence are not expected to be any more modified extensively in microevolution. However, adaptation of at least photosynthetic traits is under selection. Therefore, selection more likely acts on the fine tuning of the photosynthetic process for distinct ontogenetic situations and/or habitats. Molecular data from *Oenothera* strongly support the view that regulation of photosynthesis is a driving force of PGI. Sequence comparison between the five basic plastome types uncovered that coding regions are highly conserved, while differences were noted in genes responsible for the regulation of photosynthesis or in intergenic regions (Tables 21, 22, 23 and 24).

#### 4.5.3. *The incompatible hybrid AB-I of Oenothera*

In this thesis the study of the hybrid AB-I proved that the strategy of systematic filtering on genetically well defined material, as available from *Oenothera*, is of convincing value (Chapter 3.4.2.3). It correlated the bleached AB-I phenotype with a distinct major locus, a plastome I-specific deletion in the *clpP-psbB* intergenic region, with a down-regulation of *psbB* transcripts, reduced CP47 polypeptide and photosystem II activity in AB-I. As expected, the locus is present in several subplastomes with identical genetic behaviour. Biochemical and biophysical data are consistent with a primary lesion in photosystem II and reminiscent to photosystem II down-regulation in a bleached *Arabidopsis* mutant with substantially reduced *psbB* transcript levels (Meurer *et al.*, 2002) and photosystem II mutants of *Nicotiana* (Swiatek *et al.*, 2003), suggesting that the locus contributes to hybrid bleaching with high probability.

Species-specific editotype differences in plastid DNA and co-evolving nuclear encoded *trans* factors have been shown to be important in compartmental co-evolution between *Atropa* and *Nicotiana*. They play a crucial role in harmonious nucleo-plastid interaction between both species and also explain the pronounced phenotypic difference of their reciprocal cybrids (Schmitz-Linneweber *et al.*, 2001a; 2002; 2005). RNA editing obviously does not influence compartmental co-evolution in evening primroses (Chapter 3.4.1.5). Thus, other aspects must cause or be involved in compartmental divergence, such as transcription and/or transcript stability of photosynthetic genes *via* interaction with *trans* factors of nuclear origin and corresponding *cis* elements in the *psbB* promoter or of stabilizing elements in 5' UTRs of its mRNA species in the AB-I hybrid. Collectively, these findings suggest that despite of photosynthetic defects in both, the Solanaceen and *Oenothera* materials, the determinants causing plastid-nuclear incompatibility are different. Obviously, the ways, in which individual species or genera have evolved, their histories, and adaptation to their present-day habitats are diverse, and include changes at quite different molecular levels.

The deletion in plastome I causing the AB-I phenotype, is a species specific adaptation of plastome I to its nuclear genotype A. Obviously, the B genome exerts a dominant negative effect on the expression of *psbB*, since AA-I is compatible and expression of the relevant components encoded by the A genome in the combination AB-I is not sufficient to fully complement the compartmental incompatibility. Loci such as the *clpP-psbB* intergenic region deduced by this approach are therefore potential candidates that deserve further study. Expression analysis of the indicated region, varying abiotic parameters such as dry stress in

related compatible and incompatible hybrids like AA-I, AA-II, AB-I, and AB-II, could identify and characterize selection forces responsible for PGI and speciation (Chapter 4.5.1).

#### 4.5.4. PGI - a useful tool to identify mechanisms and driving forces of speciation

Plastidic determinants of PGI are known from just two cases, mRNA editing in the *Atropa/Nicotiana* cybrids (Schmitz-Linneweber *et al.*, 2005) and *psbB/clpP* transcription in AB-I of *Oenothera* (Chapter 3.4.2.3). However, the respective nuclear partner genes have not yet been identified. As outlined above the situation for nuclear/nuclear or nuclear/mitochondrial DMIs is quite different (Chapter 1.2.2). Molecular data on DM gene pairs are only spotlights of probably more general evolutionary scenarios and more data from different models are needed to achieve a generalization of those phenomena leading to speciation. PGI promises both, an easier and broader access to such determinants, as models currently under study. The basic problem with working on functional genetics of DMI was stated by Orr (2005): “If species are taxa that are reproductively isolated, a genetics of speciation must, almost by definition, be a genetics where such a thing is not possible, between organisms that do not exchange genes.” As outlined by Coyne (1992) there is no way to distinguish, whether the so-called “speciation genes” are the initial course of a reproductive barrier or just a by-product of speciation.

Studies on PGI could, at least in part, bypass these problems. As shown above, PGI can display much weaker effects on hybrid fitness than the so far predominantly studied hybrid sterility. In general, an important category of PGI phenotypes are numerous types of different chlorophyll deficiencies (Table 2). This class of phenotypes with little direct effect on fertility provides excellent material for mapping approaches. Since it is quite diverse, a relatively broad picture on the nature of distinct “speciation genes” and underlying mechanisms may be achieved. On the other hand, if mild PGIs are monitored in closely related taxa, there is a fair chance to look at a primary effect of speciation. These taxa are usually at the branch point of speciation and, as described for *Oenothera* and conceivable for various other taxa, PGI is the only strong hybridization barrier (Dietrich *et al.*, 1997).

The second benefit, working with PGI is a limited number of genes and gene functions involved potentially. Chloroplast genomes encode in the order of 120 genes, predominantly related to and almost equally distributed for photosynthesis and gene expression in the organelle (Table 14). As plastid gene expression is closely linked to photosynthesis, many

PGI phenotypes reflect disturbances of the photosynthetic machinery. This provides a solid framework to investigate the molecular basis of responsible determinants and processes. The substantial knowledge about photosynthesis and chloroplast genomes may help to cope the next and most challenging step in understanding speciation, *the identification of driving forces acting on the selection of speciation genes*. To relate molecular data to an ecological and evolutionary context, detailed knowledge about molecular function and regulation of a speciation gene is needed, a requirement given if genes involved in photosynthesis are responsible for PGI.

#### 4.6. The model *Oenothera*

Identification of genes causing PGI is of particular interest for understanding molecular aspects of the evolution of the compartmentalized eukaryotic genome, specifically of plastome-genome interaction and the impact of the chloroplast in pre-speciation and speciation processes. As outlined in this thesis, with appropriate experimental material, there is a reasonable chance to deduce the selection pressures and speciation forces that act on photosynthesis and cause PGI. Ironically, the plastome, a genetic compartment which so far was largely neglected in studying speciation processes, allows an easier access to the selection pressures acting on speciation genes than the nuclear genome alone (see above).

The phenomenon of PGI itself may also be of particular interest in cell biology. Since it reflects disturbance of specific co-evolved networks between plastome and genome (Herrmann *et al.*, 2003), PGIs can be considered as “network mutants”, which allow studying genetic compartments in molecular terms. This is a fundamentally different strategy than mutant approaches or high-throughput analysis, employed in molecular biology so far. Studies of network disturbances between species not only promise to deduce principles of how biological networks are designed, but also how they interact and change during speciation. Network regulation is a fundamental, unsolved question in current molecular biology and PGI may contribute, in addition to its value for exploring speciation, also to this line of research.

To achieve this goal, during the past decade basic cell and molecular biological approaches were developed for the evening primroses, notably nuclear transformation via *Agrobacterium tumefaciens*, protoplast regeneration and tissue culture approaches (Stubbe and Herrmann, 1982; Kuchuk *et al.*, 1998; Mehra-Palta *et al.*, 1998), the complete sequences of the five basic

genetically distinct plastome types (Hupfer *et al.*, 2000 and this thesis), the construction and application of an EST library (Mráček *et al.*, 2006 and this thesis), the identification of molecular markers to distinguish strains and species (this thesis), the first molecular map of the genus (Mráček, 2005; Rauwolf, unpublished), and substantially improved cytogenetic approaches (Golczyk, unpublished).

Two aspects remain to be solved for or are unique to the evening primroses model. Availability of plastid transformation would be highly desirable and remains to be established. It is advisable to focus molecular analysis on the plastid first, because the plastome, due to its limited and well defined coding potential, allows easier access to determinants of PGI. Furthermore, in *Oenothera* homologous recombination is limited not only to permanent translocation heterozygous (Cleland, 1972; Levin, 2002), but also to homozygous species (Rauwolf, unpublished). Although this phenomenon is valuable for an exchange of plastids and/or nuclei between species, and a principal and unique advantage of *Oenothera* in investigating PGI, it prevents application of some of the customary methods in plant genomics such as mapping approaches. Therefore, the model requires alternative strategies, already applied for organisms like mouse or men, where mapping strategies are not applicable. Establishing whole genome sequencing, transposon or T-DNA tagging, TILLING, and mapping strategies involving artificial chromosomes or DNA breaks, such like HAPPY or radiation hybrid mapping, will bypass this virtual drawback.

## 5. Literature

- Ahlert D, Ruf S, and Bock R (2003) Plastid protein synthesis is required for plant development in tobacco. *Proceedings of the National Academy of Sciences of the United States of America*. 100:15730-15735.
- Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*. 215:403-410.
- Amann K, Lezhneva L, Wanner G, Herrmann RG, and Meurer J (2004) ACCUMULATION OF PHOTOSYSTEM ONE1, a member of a novel gene family, is required for accumulation of [4Fe-4S] cluster-containing chloroplast complexes and antenna proteins *Plant Cell*. 16:3084-3097.
- Arisumi T (1985) Rescuing abortive *Impatiens* hybrids through aseptic culture of ovules. *Journal of the American Society of Horticultural Science*. 110:273-276.
- Arnon DI (1949) Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. *Plant Physiology*. 24:1-15.
- Arntz MA and Delph LF (2001) Pattern and process: evidence for the evolution of photosynthetic traits in natural populations. *Oecologia*. 127:455-467.
- Baerecke M (1944) Zur Genetik und Cytologie von *Oenothera ammophila* Focke, *Bauri* Boedijn, *Beckeri* Renner, *parviflora* L., *rubricaulis* Klebahn, *silesiaca* Renner. *Flora*. 138:57-92.
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, and Eddy SR (2004) The Pfam protein families database. *Nucleic Acids Research*. 32(Database Issue):D138-D141.
- Bateson W (1909) Heredity and variation in modern lights. pp. 85-101. in Seward AC, ed. *Darwin and modern science*. Cambridge. Cambridge University Press.
- Birky CW, Jr. (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics*. 35:125-148.
- Blaringhem L (1914) L'*Oenothera Lamackiana* Seringe et les *Oenothères* de la forêt de Fontainebleau. *Revue générale de Botanique*. 25:35.
- Bogdanova VS (2007) Inheritance of organelle DNA markers in a pea cross associated with nuclear-cytoplasmic incompatibility. *Theoretical and Applied Genetics*. 114:333-339.
- Bogdanova VS and Berdnikov VA (2001) Observation of a phenomenon resembling hybrid dysgenesis, in a wild pea subspecies *Pisum sativum* ssp. *elatius*. *Pisum Genetics*. 33:5-8.
- Bogdanova VS and Kosterin OE (2006) A case of anomalous chloroplast inheritance in crosses of garden pea involving an accession of wild subspecies. *Doklady Biological Sciences*. 406:44-46.

- Bomblies K and Weigel D (2007) Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nature Reviews Genetics*. 8:382-393.
- Bondarava N, De Pascalis L, Al-Babili S, Goussias C, Golecki JR, Beyer P, Bock R, and Krieger-Liszkay A (2003) Evidence that cytochrome b559 mediates the oxidation of reduced plastoquinone in the dark. *Journal of Biological Chemistry*. 278:13554-13560.
- Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, and Barbash DA (2006) Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *Science*. 314:1292-1295.
- Brown FS, Snijder RC, and van Tuyl JM (2005) Biparental plastid inheritance in *Zantedeschia albomaculata* (Araceae). *Acta Horticulturae*. 673:463-468.
- Burke JM and Arnold ML (2001) Genetics and the fitness of hybrids. *Annual Review of Genetics*. 35:31-52.
- Burrows PA, Sazanov LA, Svab Z, Maliga P, and Nixon PJ (1998) Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid *ndh* genes. *EMBO Journal*. 17:868-876.
- Catcheside DG (1939) A position effect in *Oenothera*. *Journal of Genetics*. 38:345-352.
- Catcheside DG (1940) Structural analysis of *Oenothera* complexes. *Proceedings of the Royal Society Biological Sciences Series B*. 128:509-535.
- Chapman MJ, Mulcahy DL, and Stein DB (1999) Identification of plastid genotypes in *Oenothera* subsect. *Munzia* by restriction fragment length polymorphisms (RFLPs). *Theoretical and Applied Genetics*. 98:47-53.
- Chapman MJ and Mulcahy DL (1997) Effect on genome-plastome interaction on meiosis and pollen development in *Oenothera* species and hybrids. *Sexual Plant Reproduction*. 10:288-292.
- Chase CD (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. *Trends in Genetics*. 23:81-90.
- Chaw SM, Chang CC, Chen HL, and Li WH (2004) Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of Molecular Evolution*. 58:424-441.
- Chiu WL and Sears BB (1993) Plastome-genome interactions affect plastid transmission in *Oenothera*. *Genetics*. 133:989-997.
- Cleland RE (1935) Cyto-taxonomic studies on certain *Oenotheras* from California. *Proceedings of the American Philosophical Society*. 75:339-429.
- Cleland RE (1942) The origin of hDecipiens from the complexes of *Oenothera Lamarckiana* and its bearing upon the phylogenetic significance of similarities in segmental arrangement. *Genetics*. 27:55-83.

- Cleland RE (1958) The evolution of the North American *Oenotheras* of the “*biennis*” group. *Planta*. 51:378-398.
- Cleland RE (1962) Plastid behaviour of the North American *Euoenotheras*. *Planta*. 57:699-712.
- Cleland RE (1972) *Oenothera* cytogenetics and evolution in Sutcliffe JF and Mahlberg P, eds. *Experimental botany*. London, New York. Academic Press.
- Cleland RE and Blakeslee AF (1931) Segmental interchange, the basis of chromosomal attachments in *Oenothera*. *Cytologia*. 2:175-233.
- Cleland RE and Hammond BL (1950) Analysis of segmental arrangements in certain race of *Oenothera*. Indiana University Publications. Science Series. 16:10-72.
- Cleland RE and Oehlkers F (1930) Erbllichkeit und Zytologie verschiedener *Oenotheren* und ihrer Kreuzungen. *Jahrbuch wissenschaftliche Botanik*. 73:1-124.
- Corriveau JL and Corriveau AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *American Journal of Botany*. 75:1443-1458.
- Coyne JA (1992) Genetics and speciation. *Nature*. 355:511-515.
- Coyne JA and Orr HA (2004) *Speciation*. Sunderland. Sinauer.
- Cruzan MB and Arnold ML (1999) Consequences of cytonuclear epistasis and assortative mating for the genetic structure of hybrid populations. *Heredity*. 82:36-45.
- Cserzo M, Wallin E, Simon I, von Heijne G, and Elofsson A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Engineering*. 10:673-676.
- Dahlgren KVO (1923) *Geranium bohemicum* L. x *G. bohemicum depressum* Erk Alm., ein grün-weiss-marmorierter Bastard. *Hereditas*. 4:239-250.
- Dahlgren KVO (1925) Die reziproken Bastarde zwischen *Geranium bohemicum* L. und seiner Unterart *deprehensum* Erik Alm. *Hereditas*. 6:237-256.
- Dauborn B and Brüggemann W (1996) Genome and plastome effects on photosynthesis parameters in *Oenothera* species of differential low-temperature tolerance. *Physiologia Plantarum*. 97:79-84.
- Davis BM (1916) Hybrids of *Oenothera biennis* and *Oenothera franciscana* in the first and second generations. *Genetics*. 1:197-251.
- de Gyves EM, Sparks CA, Fieldsend AF, Lazzeri P, and Jones HD (2001) High frequency of adventitious shoot regeneration from commercial cultivars of evening primrose (*Oenothera* spp.). *Annals of Applied Biology*. 138:329-332.

- de Gyves EM, Sparks CA, Sayanova O, Lazzeri P, Napier JA, and Jones HD (2004) Genetic manipulation of gamma-linolenic acid (GLA) synthesis in a commercial variety of evening primrose (*Oenothera* sp.). *Plant Biotechnology Journal*. 2:351-357.
- de Vries H (1900a) Das Spaltungsgesetz der Bastarde. *Berichte der Deutschen Botanischen Gesellschaft*. 18:83-90.
- de Vries H (1900b) Sur la loi de disjonction des hybrides. *Comptes Rendus de l'Académie des Sciences Serie II A Sciences de la Terre et des Planetes*. 130:845-847.
- de Vries H (1901 - 1903) *Die Mutationstheorie*, v. I (1901), II (1903). Leipzig. von Veit.
- de Vries H (1913) *Gruppenweise Artbildung - Unter spezieller Berücksichtigung der Gattung Oenothera*. Berlin. Gebrüder Borntraeger.
- de Vries H (1916) Gute, harte, und leere Samen von *Oenothera*. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 16:239-292.
- de Vries H (1917) *Oenothera Lamarckiana* mut. *velutina*. *Botanical Gazette*. 63:1-24.
- de Vries H (1919a) *Oenothera Lamarckiana erythrina*, eine neue Halbmutante. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 21:91-118.
- de Vries H (1919b) *Oenothera rubrinervis*, a half mutant. *Botanical Gazette*. 67:1-26.
- Dietrich W (1977) The South American species of *Oenothera* sect. *Oenothera* (*Raimannia*, *Renneria*; *Onagraceae*). *Annals of the Missouri Botanical Garden*. 64:425-626.
- Dietrich W, Raven PH, and Wagner WL (1985) Revision of *Oenothera* sect. *Oenothera* subsect. *Emersonia* (*Onagraceae*). *Systematic Botany*. 10:29-48.
- Dietrich W, Wagner WL, and Raven PH (1997) Systematics of *Oenothera* section *Oenothera* subsection *Oenothera* (*Onagraceae*) in Anderson C, ed. *Systematic botany monographs*. The American Society of Plant Taxonomists.
- Dobzhansky T (1937) *Genetics and the origin of species*. New York. Columbia University Press.
- Dobzhansky T (1970) *Genetics of the evolutionary process*. New York. Columbia University Press.
- Doebley J (2004) The genetics of maize evolution. *Annual Review of Genetics*. 38:37-59.
- Downie SR and Palmer JD (1994) A chloroplast DNA phylogeny of the Caryophyllales based on structural and inverted repeat restriction site variations. *Systematic Botany*. 19:236-252.
- Drescher A, Ruf S, Calsa T, Jr., Carrer H, and Bock R (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant Journal*. 22:97-104.

- Drillisch M (1975) Vergleichende Untersuchungen an "A-Genotypen" von *Oenothera*. PhD thesis. Heinrich-Heine-University. Düsseldorf.
- Ellison CK and Burton RS (2006) Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. *Evolution*. 60:1382-1391.
- Emerson SH and Sturtevant AH (1931) Genetical and cytological studies in *Oenothera*. III The translocation interpretation. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 59:395-419.
- Erixon P and Oxelman B (2008) Whole-gene positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast *clpP1* gene. *PLoS ONE*. 3:e1386.
- Farenholtz H (1925) Über Rassen- und Artkreuzungen in der Gattung *Hypericum*. pp. 23-32. Festschrift Schauinsland. Bremen.
- Felsenstein J (1993). Phylogeny Inference Package (PHYLIP) 3.5. Department of Genetics, University of Washington. Seattle.
- Fieldsend AF (2007) The impact of plant breeding on seed oil content and quality in evening primrose crops. pp. 29-36. Proceedings of the joint international conference on long-term experiments. Agricultural Research and Natural Resources. Debrecen-Nyírlugos.
- Fishman L and Willis JH (2006) A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*. 60:1372-1381.
- Forner J, Weber B, Thuss S, Wildum S, and Binder S (2007) Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5' and 3' end formation. *Nucleic Acids Research*. 35:3676-3692.
- Galloway LF and Etterson JR (2005) Population differentiation and hybrid success in *Campanula americana*: geography and genome size. *Journal of Evolutionary Biology*. 18:81-89.
- Gillespie JH (1991) The causes of molecular evolution. Oxford University Press, Inc.
- Glick RE and Sears BB (1994) Genetically programmed chloroplast dedifferentiation as a consequence of plastome-genome incompatibility in *Oenothera*. *Plant Physiology*. 106:367-373.
- Goh CS, Bogan AA, Joachimiak M, Walther D, and Cohen FE (2000) Co-evolution of proteins with their interaction partners. *Journal of Molecular Biology*. 299:283-293.
- Golczyk H, Hasterok R, and Joachimiak AJ (2005) FISH-aimed karyotyping and characterization of Renner complexes in permanent heterozygote *Rhoeo spathacea*. *Genome*. 48:145-153.
- Göpel G (1976) Stoffwechseländerungen in plastomabhängig inaktivem Pollen von *Oenothera*. *Zeitschrift für Pflanzenphysiologie*. 78:411-415.

- Gordon KHJ, Crouse EJ, Bohnert HJ, and Herrmann RG (1981) Restriction endonuclease cleavage site map of chloroplast DNA from *Oenothera parviflora* (*Euoenothera* plastome IV). *Theoretical and Applied Genetics*. 59:281-296.
- Gordon KHJ, Crouse EJ, Bohnert HJ, and Herrmann RG (1982) Physical mapping of differences in chloroplast DNA of the five wild-type plastomes in *Oenothera* subsection *Euoenothera*. *Theoretical and Applied Genetics*. 61:373-384.
- Goulding SE, Olmstead RG, Morden CW, and Wolfe KH (1996) Ebb and flow of the chloroplast inverted repeat. *Molecular and General Genetics*. 252:195-206.
- Grantham R (1974) Amino acid difference formula to help explain protein evolution. *Science*. 185:862-864.
- Grun P (1976) *Cytoplasmic genetics and evolution*. New York. Columbia University Press.
- Hachtel W, Neuss A, and vom Stein J (1991) A chloroplast DNA inversion marks an evolutionary split in the genus *Oenothera*. *Evolution*. 45:1050-1052.
- Haerty W and Singh RS (2006) Gene regulation divergence is a major contributor to the evolution of Dobzhansky-Muller incompatibilities between species of *Drosophila*. *Molecular Biology and Evolution*. 23:1707-1714.
- Hagemann R (2004) The sexual inheritance of plant organelles. pp. 93-114 in Daniell H and Chase CD, eds. *Molecular biology and biotechnology of plant organelles - chloroplasts and mitochondria*. Berlin, Heidelberg, New York. Springer.
- Hajdukiewicz PT, Allison LA, and Maliga P (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO Journal*. 16:4041-4048.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 41:95-98.
- Harris AS and Ingram R (1991) Chloroplast DNA and biosystematics: The effects of intraspecific diversity and plastid transmission. *Taxon*. 40:393-412.
- Harrison JS and Burton RS (2006) Tracing hybrid incompatibilities to single amino acid substitutions. *Molecular Biology and Evolution*. 23:559-564.
- Harte C (1994) *Oenothera* Contributions of a plant to biology in Frankel R, Grossman M, Linskens HF, Maliga P and Riley R, eds. *Monographs on Theoretical and Applied Genetics*. Berlin, Heidelberg, New York. Verlag.
- Haustein E (1952) Die Endenbezeichnung der Chromosomen einiger *Oenotheren* aus dem Subgenus *Raimannia*. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 84:417-453.
- Herbst W (1935) Über Kreuzungen in der Gattung *Hypericum*, mit besonderer Berücksichtigung der Buntblättrigkeit. *Flora*. 129:235-259.

- Heribert-Nilsson N (1912) Die Variabilität der *Oenothera Lamarckiana* und das Problem der Mutation. Zeitschrift für induktive Abstammungs- und Vererbungslehre. 9:89-231.
- Herrmann RG (1997) Eukaryotism, towards a new interpretation. pp. 73-118. in Schenk HEA, Herrmann RG, Jeon KW, Müller NE and Schwemmler W, eds. Eukaryotism and symbiosis. Berlin, Heidelberg, New York. Springer.
- Herrmann RG, Maier RM, and Schmitz-Linneweber C (2003) Eukaryotic genome evolution: rearrangement and coevolution of compartmentalized genetic information. Philosophical Transactions of the Royal Society of London Biological Sciences Series B. 358:87-97.
- Herrmann RG, Seyer P, Schedel R, Gordon K, Bisanz C, Winter P, Hildebrandt JW, Wlaschek M, Alt J, Driesel AJ, and Sears BB (1980) The plastid chromosomes of several dicotyledons. pp. 97-112. in Bücher T, Sebald W and Weiß H, eds. Biological chemistry of organelle formation. Berlin, Heidelberg, New York. Springer.
- Herrmann RG and Westhoff P (2001) Thylakoid biogenesis: the result of a complex phylogenetic puzzle. pp. 1-28. in Aro E-M and Andersson B, eds. Regulation of photosynthesis. Dordrecht. Kluwer Academic Publishing.
- Hess WR and Börner T (1999) Organellar RNA polymerases of higher plants. International Review of Cytology. 190:1-59.
- Hoepfener E and Renner O (1929) Genetische und zytologische Oenotherenstudien I. Zur Kenntnis der *Oenothera ammophila* Focke. Zeitschrift für induktive Abstammungs- und Vererbungslehre. 49:1-25.
- Holsinger KE and Ellstrand NC (1984) The evolution and ecology of permanent translocation heterozygotes. American Naturalist. 124:48-71.
- Homann A and Link G (2003) DNA-binding and transcription characteristics of three cloned sigma factors from mustard (*Sinapis alba* L.) suggest overlapping and distinct roles in plastid gene expression. European Journal of Biochemistry. 270:1288-1300.
- Hornung S, Fulgosi H, Dörfel P, and Herrmann RG (1996) Sequence variation in the putative replication origins of the five genetically distinct basic *Euoenothera* plastid chromosomes (plastomes). Molecular and General Genetics. 251:609-612.
- Hupfer H (2002) Vergleichende Sequenzanalyse der fünf Grundplastome der Sektion *Oenothera* (Gattung *Oenothera*) – Analyse des Cytochrom-Komplexes. PhD thesis. Ludwig-Maximilians-University. Munich.
- Hupfer H, Swiatek M, Hornung S, Herrmann RG, Maier RM, Chiu WL, and Sears B (2000) Complete nucleotide sequence of the *Oenothera elata* plastid chromosome, representing plastome I of the five distinguishable *Euoenothera* plastomes. Molecular and General Genetics. 263:581-585.
- Hura T, Grzesiak S, Hura K, Thiemt E, Tokarz K, and Wedzony M (2007) Physiological and biochemical tools useful in drought-tolerance detection in genotypes of winter

- triticales: accumulation of ferulic acid correlates with drought tolerance. *Annals of Botany*. 100:767-775.
- Iwanaga M, Mukai Y, Panayotov I, and Tsunewaki K (1978) Genetic diversity of the cytoplasm in *Triticum* and *Aegilops*. VII. Cytoplasmatic effects on respiratory and photosynthetic rates. *Japanese Journal of Genetics*. 53:387-396.
- Jain SK, Langen G, Hess W, Börner T, Huckelhoven R, and Kogel KH (2004) The white barley mutant *albastrians* shows enhanced resistance to the biotroph *Blumeria graminis* f. sp. *hordei*. *Molecular Plant-Microbe Interactions*. 17:374-382.
- Jean R (1984) The genetics of pollen lethality in the complex heterozygote *Oenothera nuda*. *Biologisches Zentralblatt*. 103:515-527.
- Jean R, Lambert AM, and Linder R (1966) Analyse cytogénétique de l'*Oenothera nuda* Renner. *Bulletin de la Société Botanique du Nord de la France*. 19:6-26.
- Johansson JT (1999) Three large inversions in the chloroplast genomes and one loss of the chloroplast gene *rps16* suggest an early evolutionary split in the genus *Adonis* (*Ranunculaceae*). *Plant Systematics and Evolution*. 218:133-143.
- Kanamaru K and Tanaka K (2004) Roles of chloroplast RNA polymerase sigma factors in chloroplast development and stress response in higher plants. *Bioscience, Biotechnology, and Biochemistry*. 68:2215-2223.
- Kapoor S and Sugiura M (1999) Identification of two essential sequence elements in the nonconsensus type II *PatpB-290* plastid promoter by using plastid transcription extracts from cultured tobacco BY-2 cells. *Plant Cell*. 11:1799-1810.
- Kappus A (1957) Wilde *Oenotheren* in Südwestdeutschland. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 88:38-55.
- Kato T, Kaneko T, Sato S, Nakamura Y, and Tabata S (2000) Complete structure of the chloroplast genome of a legume, *Lotus japonicus*. *DNA Research*. 7:323-330.
- Kazimierski T and Kazimierski EM (1970) Investigations of the hybrids of the genus *Trifolium* L. III. Morphological traits and cytogenetics of the hybrid *Trifolium repens* L. x *T. nigrescens* Viv. *Acta Societatis Botanicorum Poloniae*. 41:127-147.
- Kirk JTO and Tilney-Bassett RAE (1978) The plastids. Their chemistry, structure, growth and inheritance. Amsterdam. Elsevier.
- Kita K, Kurshige Y, Yukawa T, Nishimura S, and Handa T (2005) Plastid inheritance and plastome-genome incompatibility of intergeneric hybrids between *Menziesia* and *Rhododendron*. *Journal of the Japanese Society of Horticultural Science*. 74:318-323.
- Klebahn H (1914) Formen, Mutationen, und Kreuzungen bei einigen *Oenotheren* aus der Lüneburger Heide. *Jahrbuch der Hamburger Wissenschaftlichen Anstalten*. 31, 3. Beiheft:1-64.

- Klughammer C and Schreiber U (1994) An improved method, using saturating light pulses, for the determination of photosystem I quantum yield via P700<sup>+</sup>-absorbance changes at 830 nm. *Planta*. 192:261-268.
- Kode V, Mudd EA, Iamtham S, and Day A (2005) The tobacco plastid accD gene is essential and is required for leaf development. *Plant Journal*. 44:237-244.
- Kofer W, Koop HU, Wanner G, and Steinmuller K (1998) Mutagenesis of the genes encoding subunits A, C, H, I, J and K of the plastid NAD(P)H-plastoquinone-oxidoreductase in tobacco by polyethylene glycol-mediated plastome transformation. *Molecular and General Genetics*. 258:166-173.
- Kooten O and Snel JFH (1990) The use of chlorophyll fluorescence nomenclature in plant stress physiology. *Photosynthesis Research*. 25:147-150.
- Kozak M (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiological Reviews*. 47:1-45.
- Kuchuk N, Herrmann RG, and Koop H-U (1998) Plant regeneration from leaf protoplasts of evening primrose (*Oenothera hookeri*). *Plant Cell Reports*. 17:601-604.
- Kwok EY and Hanson MR (2004a) GFP-labelled Rubisco and aspartate aminotransferase are present in plastid stromules and traffic between plastids. *Journal of Experimental Botany*. 55:595-604.
- Kwok EY and Hanson MR (2004b) Stromules and the dynamic nature of plastid morphology. *Journal of Microscopy*. 214:124-137.
- Laloi C, Przybyla D, and Apel K (2006) A genetic approach towards elucidating the biological activity of different reactive oxygen species in *Arabidopsis thaliana*. *Journal of Experimental Botany*. 57:1719-1724.
- Lamprecht H (1944) Die genisch-plasmatische Grundlage der Artbarriere. *Agri Hortique Genetica*. 2:75-142.
- Larson EL, Bogdanowicz SM, Agrawal AA, Johnson MTJ, and Harrison RG (2008) Isolation and characterization of polymorphic microsatellite loci in common evening primrose (*Oenothera biennis*). *Molecular Ecology Resources*. 8:434-436.
- Lehmann E (1922) Die Theorien der Oenotheraforschung - Grundlagen zur experimentellen Vererbungs- und Entwicklungslehre. Jena. Gustav Fischer.
- Leister D (2003) Chloroplast research in the genomic age. *Trends in Genetics*. 19:47-56.
- Lesins K (1961) Interspecific crosses involving Alfalfa. I. *Medicago dzhawakhetica* (Bordz.) Vass. x *M. sativa* L. and its peculiarities. *Canadian Journal of Genetics and Cytology*. 3:135-152.
- Levin DA (2002) The role of chromosomal change in plant evolution: Oxford Series in Ecology and Evolution. New York. Oxford University Press, Inc.

- Levin DA (2003) The cytoplasmic factor in plant speciation. *Systematic Botany*. 28:5-11.
- Levin RA, Wagner WL, Hoch PC, Hahn WJ, Rodriguez A, Baum DA, Katinas L, Zimmer EA, and Sytsma KJ (2004) Paraphyly in tribe *Onagreae*: Insights into phylogenetic relationships of *Onagraceae* based on nuclear and chloroplast sequence data. *Systematic Botany*. 29:147-164.
- Levin RA, Wagner WL, Hoch PC, Nepokroff M, Pires JC, Zimmer EA, and Sytsma KJ (2003) Family-level relationships of *Onagraceae* bases on chloroplast *rbcL* and *ndhF* data. *American Journal of Botany*. 90:107-115.
- Levy M and Levin DA (1975) Genic heterozygosity and variation in permanent translocation heterozygotes of the *Oenothera biennis* complex. *Genetics*. 79:493-512.
- Liere K and Börner T (2006) Transcription of plastid genes. pp. 184-233. *in* Grasser KD, ed. Regulation of transcription in plants. Oxford. Backwell Publishing.
- Lilienfeld FA (1962) Plastid behaviour in reciprocally different crosses between two races of *Medicago truncatula* Gaertn. *Seikon Zihô*. 13:3-38.
- Lopez-Juez E and Pyke KA (2005) Plastids unleashed: their development and their integration in plant development. *International Journal of Developmental Biology*. 49:557-577.
- Lutz AM (1907) A preliminary note on the chromosomes of *Oenothera lamarckiana* and one of its mutants, *Oenothera gigas*. *Science*. 27:151-152.
- Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, and Kowallik KV (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*. 393:162-165.
- Mehra-Palta A, Koop H-U, Goes S, Troidl E-M, Nagy G, Tyagi S, Kofer W, and Herrmann RG (1998) Tissue culture of wild-type, interspecific genome/plastome hybrids and plastome mutants of evening primrose (*Oenothera*): controlled morphogenesis and transformation. *Plant Cell Reports*. 17:605-611.
- Meredith MR, Michaelson-Yeates TPT, Ougham HJ, and Thomas H (1995) *Trifolium ambiguum* as a source of variation in the breeding of white clover. *Euphytica*. 82:185-191.
- Metzlaff M, Pohlheim F, Börner T, and Hagemann R (1982) Hybrid variegation in the genus *Pelargonium*. *Current Genetics*. 5:245-249.
- Meurer J, Lezhneva L, Amann K, Gödel M, Bezhani S, Sherameti I, and Oelmüller R (2002) A peptide chain release factor 2 affects the stability of UGA-containing transcripts in *Arabidopsis* chloroplasts. *Plant Cell*. 14:3255-3269.
- Meurer J, Meierhoff K, and Westhoff P (1996) Isolation of high-chlorophyll-fluorescence mutants of *Arabidopsis thaliana* and their characterisation by spectroscopy, immunoblotting and northern hybridisation. *Planta*. 198:385-396.
- Michaelis P (1954) Cytoplasmic inheritance in *Epilobium* and its theoretical significance. *Advances in Genetics*. 6:287-401.

- Michishita A, Ureshino K, and Miyajima I (2002) Plastome-genome incompatibility of *Rhododendron serpyllifolium* (A. Gray) Miq. to evergreen *Azalea* species belonging to series *Kaempferia*. *Journal of the Japanese Society of Horticultural Science*. 71:375-381.
- Miller H (1987) Practical aspects of preparing phage and plasmid DNA: growth, maintenance, and storage of bacteria and bacteriophage. *Methods in Enzymology*. 152:145-170.
- Moffett AA (1965) Genetical studies in Acacias. III. Chlorosis in interspecific hybrids. *Heredity*. 20:609-620.
- Morse NL and Clough PM (2006) A meta-analysis of randomized, placebo-controlled clinical trials of Efamol® evening primrose oil in atopic eczema. Where do we go from here in light of more recent discoveries? *Current Pharmaceutical Biotechnology*. 7:503-524.
- Mráček J (2005) Investigation of genome-plastome incompatibility in *Oenothera* and *Passiflora*. PhD thesis. Ludwig-Maximilians-University. Munich.
- Mráček J, Greiner S, Cho WK, Rauwolf U, Braun M, Umate P, Altstätter J, Stoppel R, Mlčochová L, Silber MV, Volz SM, White S, Selmeier R, Rudd S, Herrmann RG, and Meurer J (2006) Construction, database integration, and application of an *Oenothera* EST library. *Genomics*. 88:372-380.
- Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposia*. 6.
- Neckermann K, Zeltz P, Igloi GL, Kossel H, and Maier RM (1994) The role of RNA editing in conservation of start codons in chloroplast genomes. *Gene*. 146:177-182.
- New EH and Paris CD (1967) Virescence in *Zantedeschia*. *Proceedings of the American Society of Horticultural Science*. 92:685-695.
- Newton WCF (1931) Genetical experiments with *Silene otites* and related species. *Journal of Genetics*. 24:109-120.
- Noack KL (1931) Über *Hypericum*-Kreuzungen. I. Die Panaschüre der Bastarde zwischen *Hypericum acutum* Moench und *Hypericum montanum* L. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 59:77-101.
- Noack KL (1934) Über *Hypericum*-Kreuzungen. IV. Die Bastarde zwischen *Hypericum acutum* Mönch, *montanum* L. *quandragulum* L., *hirsutum* L. and *pulchrum* L. *Zeitschrift für Botanik*. 28:1-71.
- Noack KL (1937) Über *Hypericum*-Kreuzungen. V. Weitere Untersuchungen über die Buntblättrigkeit bei Artbastarden. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 73:108-130.
- Noguchi Y (1932) Studies on the species crosses of Japanese *Rhododendron*. *Japanese Journal of Botany*. 6:103-124.

- Orr HA (1995) The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*. 139:1805-1813.
- Orr HA (2005) The genetic basis of reproductive isolation: insights from *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. 102 Suppl 1:6522-6526.
- Ortiz-Barrientos D, Counterman BA, and Noor MA (2007) Gene expression divergence and the origin of hybrid dysfunctions. *Genetica*. 129:71-81.
- Ossenbühl F, Gohre V, Meurer J, Krieger-Liszkay A, Rochaix JD, and Eichacker LA (2004) Efficient assembly of photosystem II in *Chlamydomonas reinhardtii* requires Alb3.1p, a homolog of *Arabidopsis* ALBINO3. *Plant Cell*. 16:1790-1800.
- Palmer JD (1992) Comparison of chloroplast and mitochondrial genome evolution in plants. pp. 99-133. in Herrmann RG, ed. *Plant gene research*. Berlin, Heidelberg, New York. Springer.
- Pandey KK (1957) A self-compatible hybrid from a cross between two self-incompatible species in *Trifolium*. *Journal of Heredity*. 48:278-281.
- Pandey KK and Blaydes GW (1957) Cytoplasmic inheritance of plastids in *Impatiens sultanii* Hook, F., *Petunia violacea* Lindl. and *Chlorophytum elatum* R. BR. *Ohio Journal of Science*. 57:135-147.
- Pandey KK, Grant JE, and Williams EG (1987) Interspecific hybridisation between *Trifolium repens* and *T. uniflorum*. *Australian Journal of Botany*. 35:171-182.
- Peer LB (1950) A cytogenetic study of certain broad-leaved races of *Oenothera*. Indiana University Publications. Science Series. 16:82-159.
- Pellew C (1917) Types of segregation. *Journal of Genetics*. 6:317-339.
- Pennisi E (2006) Two rapidly evolving genes spell trouble for hybrids. *Science*. 314:122-123.
- Pesaresi P, Schneider A, Kleine T, and Leister D (2007) Interorganellar communication. *Current Opinion in Plant Biology*. 10:600-606.
- Pohlheim F (1986) Hybrid variegation in crosses between *Pelargonium zonale* (L.) l'Herit. ex. Ait. and *Pelargonium inquinans* (L.) l'Herit. ex. Ait. *Plant Breeding*. 97:93-96.
- Preer JR, Jr. (1971) Extrachromosomal inheritance: hereditary symbionts, mitochondria, chloroplasts. *Annual Review of Genetics*. 5:361-406.
- Presgraves DC and Stephan W (2007) Pervasive adaptive evolution among interactors of the *Drosophila* hybrid inviability gene, Nup96. *Molecular Biology and Evolution*. 24:306-314.
- Przywara L, White DWR, Sanders PM, and Maher D (1989) Interspecific hybridization of *Trifolium repens* with *T. hybridum* using *in ovule* embryo and embryo culture. *Annals of Botany*. 64:613-624.

- Quesenberry KH and Taylor NL (1976) Interspecific hybridization in *Trifolium* L., sect. *Trifolium* Zoh. I. Diploid hybrids among *T. alpestre* L., *T. rubens* L., *T. heldreichianum* Hausskn., and *T. noricum* Wulf. *Crop Science*. 16:382-386.
- Race HL, Herrmann RG, and Martin W (1999) Why have organelles retained genomes? *Trends in Genetics*. 15:364-370.
- Rand DM, Haney RA, and Fry AJ (2004) Cytonuclear coevolution: the genomics of cooperation. *Trends in Ecology & Evolution*. 19:645-653.
- Raven PH (1988) *Onagraceae* as a model of plant evolution. pp. 85-107. in Gottlieb LD and Jain SK, eds. *Plant evolutionary biology: a symposium honouring G. Ledyard Stebbins*. London. Chapman and Hall.
- Redinbaugh MG, Jones TA, and Zhang Y (2000) Ubiquity of the St chloroplast genome in St-containing *Triticeae* polyploids. *Genome*. 43:846-852.
- Renner O (1917) Versuche über die gametische Konstitution der Oenotheren. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 18:121-294.
- Renner O (1924) Die Scheckung der Oenotherenbastarde. *Biologisches Zentralblatt*. 44:309-336.
- Renner O (1934) Die pflanzlichen Plastiden als selbständige Elemente der genetischen Konstitution. *Berichte der mathematisch-physikalischen Klasse der sächsischen Akademie der Wissenschaften zu Leibzig*. 86:241-266.
- Renner O (1937) Wilde Oenotheren in Norddeutschland. *Flora*. 31:182-226.
- Renner O (1938) Kurze Mitteilung über Oenothera. II. Zu den Chromosomenformel der Komplex albicans, curvans, flectens, gaudens (rubens), rigens. *Flora*. 32:319-324.
- Renner O (1941) Über die Entstehung homozygotischer Formen aus komplexheterozygotischen Oenotheren. *Flora*. 135:201-238.
- Renner O (1942a) Europäische Wildarten von *Oenothera*. *Berichte der Deutschen Botanischen Gesellschaft*. 60:448-466.
- Renner O (1942b) Über das Crossing-over bei *Oenothera*. *Flora*. 136:117-214.
- Renner O (1943a) Über die Entstehung homozygotischer Formen aus komplexheterozygotischen Oenotheren II. Die Translokationshomozygoten. *Zeitschrift für Botanik*. 39:49-106.
- Renner O (1943b) Zur Kenntnis von *O. silesiaca* n. sp., *parviflora* L., *ammophila* Focke, *rubricaulis* Kleb. *Flora*. 36:325-335.
- Renner O (1943c) Zur Kenntniss des Pollenkomplexes flectens der *Oenothera atrovirens* Sh. et Bartl. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 81:391-483.

- Renner O (1950) Europäische Wildarten von *Oenothera* II. Berichte der Deutschen Botanischen Gesellschaft. 63:129-138.
- Renner O (1956) Europäische Wildarten von *Oenothera* III. Planta. 47:219-254.
- Renner O and Hirmer U (1956) Zur Kenntnis von *Oenothera*. I. Über *Oe. conferta* n. sp. II. Über künstliche Polyploidie. Biologisches Zentralblatt. 75:513-531.
- Rice DW and Palmer JD (2006) An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. BMC Biology. 4:31.
- Rice P, Longden I, and Bleasby A (2000) *EMBOSS*: The european molecular biology open software suite. Trends in Genetics. 16:276-277.
- Rogalski M, Ruf S, and Bock R (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucleic Acids Research. 34:4537-4545.
- Rolland N, Dorne AJ, Amoroso G, Sultemeyer DF, Joyard J, and Rochaix JD (1997) Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. EMBO Journal. 16:6713-6726.
- Rossmann G (1963) Analyse der *Oenothera coronifera* Renner. Flora. 153:451-468.
- Sackton TB, Haney RA, and Rand DM (2003) Cytonuclear coadaptation in *Drosophila*: disruption of cytochrome c oxidase activity in backcross genotypes. Evolution. 57:2315-2325.
- Sakai K, Ozaki Y, Ureshino K, Miyajima I, and Okubo H (2004) Effectiveness of inter-ploid crosses for overcoming plastome-genome incompatibility in intersectional crosses of Azaleas. Acta Horticulturae. 651:47-53.
- Salts Y, Herrmann RG, Peleg N, Lavi U, Izhar S, Frankel R, and Beckmann JS (1984) Physical mapping of plastid DNA variation among eleven *Nicotiana* species. Theoretical and Applied Genetics. 69:1-14.
- Sambrook J, Fritsch EF, and Maniatis T (1989) Molecular cloning – a laboratory manual. Cold Spring Harbor. Cold Spring Harbor Laboratory Press.
- Sato S, Nakamura Y, Kaneko T, Asamizu E, and Tabata S (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Research. 6:283-290.
- Schmitz-Linneweber C, Kushnir S, Babiychuk E, Poltnigg P, Herrmann RG, and Maier RM (2005) Pigment deficiency in nightshade/tobacco cybrids is caused by the failure to edit the plastid ATPase alpha-subunit mRNA. Plant Cell. 17:1815-1828.
- Schmitz-Linneweber C, Maier RM, Alcaraz JP, Cottet A, Herrmann RG, and Mache R (2001a) The plastid chromosome of spinach (*Spinacia oleracea*): complete nucleotide sequence and gene organization. Plant Molecular Biology. 45:307-315.

- Schmitz-Linneweber C, Regel R, Du TG, Hupfer H, Herrmann RG, and Maier RM (2002) The plastid chromosome of *Atropa belladonna* and its comparison with that of *Nicotiana tabacum*: the role of RNA editing in generating divergence in the process of plant speciation. *Molecular Biology and Evolution*. 19:1602-1612.
- Schmitz-Linneweber C, Tillich M, Herrmann RG, and Maier RM (2001b) Heterologous, splicing-dependent RNA editing in chloroplasts: allotetraploidy provides trans-factors. *EMBO Journal*. 20:4874-4883.
- Schmitz UK and Kowallik KV (1986) Polymorphism and gene arrangement among plastomes of ten *Epilobium* species. *Plant Molecular Biology*. 7:115-127.
- Schreiber U, Schliwa U, and Bilger W (1986) Continuous recording of photochemical and non-photochemical chlorophyll fluorescence quenching with a new type of modulation fluorometer. *Photosynthesis Research*. 10:51-62.
- Schumacher E and Steiner EE (1993) Cytological analysis of complex-heterozygotes in populations of *Oenothera grandiflora* (*Onagraceae*) in Alabama. *Plant Systematics and Evolution*. 184:77-87.
- Schumacher E, Steiner EE, and Stubbe W (1992) The complex-heterozygotes of *Oenothera grandiflora* L'Her. *Botanica Acta*. 105:375-381.
- Schweizer D and Ambros PF (1994) Chromosome banding. *Methods in Molecular Biology*. 29:97-111.
- Schwemmler J, Haustein E, Sturm A, and Binder M (1938) Genetische und zytologische Untersuchungen an Eu-Oenotheren. Teil I bis VI. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 75:358-800.
- Schwenkert S, Umate P, Dal Bosco C, Volz S, Mlcochova L, Zoryan M, Eichacker LA, Ohad I, Herrmann RG, and Meurer J (2006) PsbI affects the stability, function, and phosphorylation patterns of photosystem II assemblies in tobacco. *Journal of Biological Chemistry*. 281:34227-34238.
- Sears BB, Stoike LL, and Chiu WL (1996) Proliferation of direct repeats near the *Oenothera* chloroplast DNA origin of replication. *Molecular Biology and Evolution*. 13:850-863.
- Shiina T, Tsunoyama Y, Nakahira Y, and Khan MS (2005) Plastid RNA polymerases, promoters, and transcription regulators in higher plants. *International Review of Cytology*. 244:1-68.
- Shikanai T (2006) RNA editing in plant organelles: machinery, physiological function and evolution. *Cellular and Molecular Life Sciences*. 63:698-708.
- Shikanai T, Shimizu K, Ueda K, Nishimura Y, Kuroiwa T, and Hashimoto T (2001) The chloroplast clpP gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. *Plant and Cell Physiology*. 42:264-273.

- Silhavy D and Maliga P (1998) Mapping of promoters for the nucleus-encoded plastid RNA polymerase (NEP) in the iojap maize mutant. *Current Genetics*. 33:340-344.
- Sjogren LL, Stanne TM, Zheng B, Sutinen S, and Clarke AK (2006) Structural and functional insights into the chloroplast ATP-dependent Clp protease in *Arabidopsis*. *Plant Cell*. 18:2635-2649.
- Smith L (1915) Variegation in *Pelargonium*. *Proceedings of the Royal Horticultural Society*. 41:36.
- Snijder RC, Lindhout P, and van Tuyl JM (2004) Genetic control of resistance to soft rot caused by *Erwinia carotovora* subsp. *carotovora* in *Zantedeschia* spp. (*Araceae*), section *Aestivae*. *Euphytica*. 136:319-325.
- Sriraman P, Silhavy D, and Maliga P (1998) The phage-type PclpP-53 plastid promoter comprises sequences downstream of the transcription initiation site. *Nucleic Acids Res*. 26:4874-4879.
- Starmer J, Stomp A, Vouk M, and Bitzer D (2006) Predicting Shine-Dalgarno sequence locations exposes genome annotation errors. *PLoS Computational Biology*. 2:e57.
- Steane DA (2005) Complete nucleotide sequence of the chloroplast genome from the Tasmanian blue gum, *Eucalyptus globulus* (*Myrtaceae*). *DNA Research*. 12:215-220.
- Stebbins G (1950) *Variation and Evolution in plants*. New York. Columbia University Press.
- Steiner EE (1951) Phylogenetic relationships of certain races of *Oenothera* from Mexico and Guatemala. *Evolution*. 5:265-272.
- Steiner EE (1955) A cytogenetic study of certain races of *Oenothera elata*. *Bulletin of the Torrey Botanical Club*. 82:292-297.
- Steiner EE and Stubbe W (1984) A contribution to the population biology of *Oenothera grandiflora* L'Her. *American Journal of Botany*. 71:1293-1301.
- Steiner EE and Stubbe W (1986) *Oenothera grandiflora* revisited: a new view of its population structure. *Bulletin of the Torrey Botanical Club*. 113:406-412.
- Stinson HT (1960) Extranuclear barriers to interspecific hybridization between *Oenothera Hookeri* and *Oe. argillicola*. *Genetics*. 45:815-838.
- Stubbe W (1953) Genetische und zytologische Untersuchungen an verschiedenen Sippen von *Oenothera suaveolens*. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*. 85:180-209.
- Stubbe W (1958) Dreifarbenpanschierung bei *Oenothera*. II. Wechselwirkungen zwischen Geweben mit zwei erblich verschiedenen Plastidensorten. *Zeitschrift für Vererbungslehre*. 89:189-203.
- Stubbe W (1959) Genetische Analyse des Zusammenwirkens von Genom und Plastom bei *Oenothera*. *Zeitschrift für Vererbungslehre*. 90:288-298.

- Stubbe W (1960) Untersuchungen zur genetischen Analyse des Plastoms von *Oenothera*. Zeitschrift für Botanik. 48:191-218.
- Stubbe W (1963a) Die Rolle des Plastoms in der Evolution der Oenotheren. Berichte der Deutschen Botanischen Gesellschaft. 76:154-167.
- Stubbe W (1963b) Extrem disharmonische Genom-Plastom-Kombinationen und väterliche Plastidenvererbung bei *Oenothera*. Zeitschrift für Vererbungslehre. 94:392-411.
- Stubbe W (1964) The role of the plastome in evolution of the genus *Oenothera*. Genetica. 35:28-33.
- Stubbe W (1989) *Oenothera* – An ideal system for studying the interaction of genome and plastome. Plant Molecular Biology Reporter. 7:245-257.
- Stubbe W and Diers L (1958) Die Chromosomenformel des albicans-Komplexes der Sippe *Grado* von *Oenothera suaveolens*. Zeitschrift für Vererbungslehre. 89:320-322.
- Stubbe W and Herrmann RG (1982) Selection and maintenance of plastome mutants and interspecific genome/plastome hybrids from *Oenothera*. pp. 149-165. in Edelman M, Hallick RB and Chua N-H, eds. Methods in chloroplast molecular biology. Amsterdam. Elsevier.
- Stubbe W, Pietsch B, and Kowallik KV (1978) Cytologische Untersuchungen über väterliche Plastidenvererbung und plastomabhängige Degradation der Samenanlagen bei einem *Oenothera*-Bastard. Biologisches Zentralblatt. 97:39-52.
- Stubbe W and Raven PH (1979) A genetic contribution to the taxonomy of *Oenothera* sect. *Oenothera* (including subsection *Euoenothera*, *Emersonia*, *Raimannia* and *Munzia*). Plant Systematics and Evolution. 133:39-59.
- Stubbe W and Steiner E (1999) Inactivation of pollen and other effects of genome-plastome incompatibility in *Oenothera*. Plant Systematics and Evolution. 217:259-277.
- Swiatecka-Hagenbruch M, Liere K, and Börner T (2007) High diversity of plastidial promoters in *Arabidopsis thaliana*. Molecular Genetics and Genomics. 277:725-734.
- Swiatek M, Regel RE, Meurer J, Wanner G, Pakrasi HB, Ohad I, and Herrmann RG (2003) Effects of selective inactivation of individual genes for low-molecular-mass subunits on the assembly of photosystem II, as revealed by chloroplast transformation: the psbEFLJ operon in *Nicotiana tabacum*. Molecular Genetics and Genomics. 268:699-710.
- Systema KJ, Hahn WJ, Smith JF, and Wagner WL (1993) Characterisation and phylogenetic utility of a large inversion in the chloroplast genome of some species in *Oenothera* (*Onagraceae*). American Journal of Botany Supplement. 80:79.
- Taniguchi S, Imayoshi Y, Hatano T, Yazaki K, and Yoshida T (2002) Hydrolysable tannin production in *Oenothera tetraptera* shoot tissue culture. Plant Biotechnology. 19:357-363.

- Thiel T, Kota R, Grosse I, Stein N, and Graner A (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Research*. 32:e5.
- Tiffin P, Olson MS, and Moyle LC (2001) Asymmetrical crossing barriers in angiosperms. *Proceedings of the Royal Society Biological Sciences Series B*. 268:861-867.
- Tillich M, Funk HT, Schmitz-Linneweber C, Poltnigg P, Sabater B, Martin M, and Maier RM (2005) Editing of plastid RNA in *Arabidopsis thaliana* ecotypes. *Plant Journal*. 43:708-715.
- Tsitrone A, Kirkpatrick M, and Levin DA (2003) A model for chloroplast capture. *Evolution*. 57:1776-1782.
- Tsumura Y, Suyama Y, and Yoshimura K (2000) Chloroplast DNA inversion polymorphism in populations of *Abies* and *Tsuga*. *Molecular Biology and Evolution*. 17:1302-1312.
- Turelli M and Moyle LC (2007) Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics*. 176:1059-1088.
- Turelli M and Orr HA (2000) Dominance, epistasis and the genetics of postzygotic isolation. *Genetics*. 154:1663-1679.
- Ureshino K and Miyajima I (2002) The study on the relationship between leaf colour and ptDNA inheritance in intersectional crosses of *Rhododendron kiusianum* x *R. japonicum* f. *flavum*, resulting in an unexpected triploid progeny. *Journal of the Japanese Society of Horticultural Science*. 71:214-219.
- Ureshino K, Miyajima I, Ozaki Y, Kobayashi N, Michishita A, and Akabane M (1999) Appearance of albino seedlings and ptDNA inheritance in interspecific hybrids of *Azalea*. *Euphytica*. 110:61-66.
- van der Meer JP (1974) Hybrid chlorosis in interspecific crosses of *Oenothera*: Polygenic inheritance of the nuclear component. *Canadian Journal of Genetics and Cytology*. 16:193-201.
- van Doorn WG (2005) Plant programmed cell death and the point of no return. *Trends in Plant Science*. 10:478-483.
- van Ooijen JW and Voorrips RE (2001). JoinMap® 3.0. Software for the calculation of genetic linkage maps. Plant Research International. Wageningen, the Netherlands.
- Vera A and Sugiura M (1994) A novel RNA gene in the tobacco plastid genome: its possible role in the maturation of 16S rRNA. *EMBO Journal*. 13:2211-2217.
- von Wangenheim KH (1962) Zur Ursache der Abortion von Samenanlagen in Diploid-Polyploid-Kreuzungen. *Molecular and General Genetics*. 93:319-334.
- Walbot V and Coe EH (1979) Nuclear gene *iojap* conditions a programmed change to ribosome-less plastids in *Zea mays*. *Proceedings of the National Academy of Sciences of the United States of America*. 76:2760-2764.

- Wang X (2006) Comparative bioinformatic analysis of *Oenothera* plastome types I-V - a bioinformatic approach towards elucidating genome/plastome incompatibilities and speciation processes. Diplomina thesis. Ludwig-Maximilians-University. Munich.
- Wasmund O (1980) Cytogenetische Untersuchung zur Systematik einiger Sippen der Subsektion *Euoenothera* der Gattung *Oenothera* (*Onagraceae*). State Examination thesis. Heinrich-Heine-University. Düsseldorf.
- Wasmund O (1990) Cytogenetic investigation on *Oenothera nutans* (*Onagraceae*). *Plant Systematics and Evolution*. 169:69-80.
- Wasmund O and Stubbe W (1986) Cytogenetic investigations on *Oenothera wolffii* (*Onagraceae*). *Plant Systematics and Evolution*. 154:79-88.
- Westhoff P (1985) Transcription of the gene encoding the 51 kd chlorophyll a-apoprotein of the photosystem II reaction centre from spinach. *Molecular and General Genetics*. 201:115-123.
- Westhoff P and Herrmann RG (1988) Complex RNA maturation in chloroplasts. The psbB operon from spinach. *European Journal of Biochemistry*. 171:551-564.
- Willett CS and Burton RS (2001) Viability of cytochrome c genotypes depends on cytoplasmic backgrounds in *Tigriopus californicus*. *Evolution*. 55:1592-1599.
- Winter P and Herrmann RG (1987) A five-base-pair-deletion in the gene for the large subunit causes the lesion in the ribulose biphosphate carboxylase/oxygenase-deficient plastome mutant sigma of *Oenothera hookeri*. *Botanica Acta*. 101:42-48.
- Wittwer CT, Herrmann MG, Moss AA, and Rasmussen RP (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. *BioTechniques*. 22:130-131, 134-138.
- Wolfe KH, Li WH, and Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America*. 84:9054-9058.
- Wolfson R, Higgins KG, and Sears BB (1991) Evidence for replication slippage in the evolution of *Oenothera* chloroplast DNA. *Molecular Biology and Evolution*. 8:709-720.
- Yang X, Chen X, Ge Q, Li B, Tong Y, Li Z, Kuang T, and Lu C (2007) Characterization of photosynthesis of flag leaves in a wheat hybrid and its parents grown under field conditions. *Journal of Plant Physiology*. 164:318-326.
- Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences*. 13:555-556.
- Yao J-L and Cohen D (2000) Multiple gene control of plastome-genome incompatibility and plastid DNA inheritance in interspecific hybrids of *Zantedeschia*. *Theoretical and Applied Genetics*. 101:400-406.

- Yao J-L, Cohen D, and Rowland RE (1994) Plasmid DNA inheritance and plastome-genome incompatibility in interspecific hybrids of *Zantedeschia* (*Araceae*). *Theoretical and Applied Genetics*. 88:255-260.
- Yao J-L, Cohen D, and Rowland RE (1995) Interspecific albino and variegated hybrids in the genus *Zantedeschia*. *Plant Science*. 109:199-206.
- Yukawa M, Tsudzuki T, and Sugiura M (2006) The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. *Molecular Genetics and Genomics*. 275:367-373.
- Zhang Q, Liu Y, and Sodmergen (2003) Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. *Plant and Cell Physiology*. 44:941-951.

## 6. Summary

The subject of this thesis was to develop molecular approaches appropriate to investigate speciation processes. The genus *Oenothera* was chosen for study, since it offers the unique possibility to exchange plastids, individual or more chromosomes and/or even entire haploid genomes (so-called Renner complexes) between species. In addition, a rich stock of information in taxonomy, cytogenetics and formal genetics is available, collected for more than a century of research. Interspecific exchange of plastids, nuclear genomes or chromosomes often leads to mis-development of the resulting hybrids. These inviable hybrids form hybridization barriers responsible for speciation. In the case of plastid and nuclear genome exchange, hybrid bleaching is frequently observed, which results from plastome-genome incompatibility (PGI) due to compartmental co-evolution. Traditional work on *Oenothera* was almost exclusively restricted to classical genetic and cytogenetic approaches. Subsection *Oenothera*, the best studied of the five subsections in the section *Oenothera*, was used in this work. It is comprised of three basic nuclear genomes, A, B and C, which occur in homozygous (AA, BB, CC) or stable heterozygous (AB, AC, BC) combination. In nature, the nuclear genomes are associated with five basic, genetically discernible plastid types (I - V) in distinct combinations. The following results were obtained:

(i) Biochemistry with *Oenothera* is not trivial due to exceedingly high amounts of mucilage and tannins which adversely interfere with the isolation of macromolecules and enzymatic reactions. A basic biochemistry for the material was therefore developed initially, notably to obtain appropriate subcellular fractions, restrictable, amplifiable and clonable DNA, RNA, supramolecular protein assemblies and proteins of appropriate purity. (ii) Evaluation of the PGI literature clearly indicates that PGI can form hybridization barriers according to the Dobzhansky-Muller gene pair model of speciation, even if the genes reside in different cellular compartments. (iii) *Oenothera* PGIs could be classified into four genetically distinct categories, which influence hybridization barriers in different ways. (iv) Co-dominant marker systems (SSLP and CAPS) were generated for both, nuclear genome and plastome. Their potential was successfully evaluated with crossing programs designed to exchange plastomes, genomes, or individual chromosomes between species. (v) The plastome markers allowed to genotype 41 subplastomes to judge inter- and intraplastome diversity and displayed molecular loci linked to the genetic behaviour of basic plastome types I - V. (vi) A single, highly polymorphic marker (M40) was sufficient to genotype 29 different Renner complexes of the

basic genome types A, B and C. (vii) Markers specific for all seven *Oenothera* chromosomes were selected. Combined with the genetics of a partial permanent translocation heterozygote (ring of 12 chromosomes *plus* 1 bivalent, which behave as two distinct linkage groups) they allowed the assignment of molecular linkage group 7 to chromosome 9·8 of the classical *Oenothera* map. Material for the assignment of the remaining chromosomes and their arms was produced or selected so that both map types can now be fully integrated. (viii) In parallel to work on the nuclear genome, the sequences of the five basic *Oenothera* plastomes were completed (in cooperation). Elaborated in this thesis, due to its limited coding potential, conserved nature, and substantial knowledge about photosynthesis, plastid chromosomes provide relatively easy access to “speciation genes” and selection pressures causing speciation. (ix) Phylogenetic analysis of the sequences provided a plastome pedigree, and also an idea about the age of the subsection, i.e. back to the middle of Pleistocene, approximately 1 mya ago. This contributed to solve a long lasting question in the *Oenothera* literature. (x) Application of appropriate algorithms uncovered for the first time that plastomes are subject to natural selection and hence contribute to speciation. This was questioned repeatedly. (xi) A novel weighting strategy, combining classical genetic data on plastome-genome compatibility/incompatibility with molecular data and bioinformatic approaches, was applied to deduce potential plastid determinants for PGI. (xii) In a case study it could be shown that a single plastid locus contributes substantially to PGI in the interspecific hybrid AB-I, which was found to be defective in photosystem II. A plastome I-specific deletion in the bidirectional promoter region between *psbB* and *clpP* was found to be responsible for the phenotype observed. The finding is consistent with reduced levels of *psbB* mRNA and its product CP47 chlorophyll *a* apoprotein of photosystem II, with spectroscopic data and phenotype. (xiii) Available data indicate that interspecific plastome-genome hybrids represent some sort of “network mutants”. This would imply that speciation is predominantly a regulatory phenomenon. In the studied cases PGIs are involved in the fine-tuning of regulation of photosynthesis, rather than in an adaptation of its structural components. This is considered as a major finding of this thesis.

## 7. Appendix

### 7.1. List of publications

*Part of this work has already been published or is in preparation to be published*

Mráček J, **Greiner S**, Cho WK, Rauwolf U, Braun M, Umate P, Altstätter J, Stoppel R, Mlčochová L, Silber MV, Volz SM, White S, Selmeier R, Rudd S, Herrmann RG, and Meurer J (2006) Construction, database integration, and application of an *Oenothera* EST library. *Genomics*. 88:372-380.

**Greiner S**, Wang X, Rauwolf U, Silber M, Maier K, Haberer G, Meurer J, and Herrmann RG (2008) The complete nucleotide sequences of the five genetically distinct plastid genomes of *Oenothera*, subsection *Oenothera*: I. Sequence evaluation and plastome evolution. *Nucleic Acids Research*. 36:2366-2378.

**Greiner S**, Herrmann RG, Wang X, Rauwolf U, Mayer K, Haberer G, and Meurer J (2008) The complete nucleotide sequences of the five genetically distinct plastid genomes of *Oenothera*, subsection *Oenothera*: II. A microevolutionary view using bioinformatics and formal genetic data. Submitted.

Rauwolf U, Golczyk H, Meurer J, Herrmann RG, and **Greiner S** (2008) Molecular marker systems for *Oenothera* genetics. Submitted.

Golczyk H, Musiał K, Rauwolf U, Meurer J, Joachimiak AJ, Herrmann RG, and **Greiner S** (2008) Meiotic events in *Oenothera* – a non-standard pattern of chromosome behaviour. Submitted.

**Greiner S**, Rauwolf U, Meurer J, and Herrmann RG (2008) The impact of plastids on speciation. In preparation.

Rauwolf U, **Greiner S**, Braun M, Golczyk H, Mracek J, Mohler V, Herrmann RG, and Meurer J (2008) Sexual reproduction with homologous recombination limited to telomeric regions of chromosomes. In preparation.

*Further publications*

Swiatek M, **Greiner S**, Kemp S, Drescher A, Koop H-U, Herrmann RG, and Maier RM (2003) PCR analysis of pulsed-field gel electrophoresis-purified plastid DNA, a sensitive tool to judge the hetero-/homoplastomic status of plastid transformants. *Current Genetics*. 43:45-53.

Rauwolf U, Golczyk H, **Greiner S**, and Herrmann RG (2008) Variable amounts of DNA related to the size of chloroplasts: II. Biochemical determination of DNA amounts per organelle and cell. In preparation.

## **Ehrenwörtliche Versicherung**

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt habe.

Zudem erkläre ich, dass ich keine früheren Promotionsversuche unternommen habe.

München, 15.05.2008

(Stephan Greiner)

## 7.2. *Curriculum Vitae*

### *Persönliche Daten*

---

Geburtsdatum                      13. Februar 1979  
 Geburtsort                              Ulm

### *Ausbildung*

---

09/1989 – 06/1998                      ***Maria-Theresia-Gymnasium***, Augsburg, Abitur

seit 10/1999                              ***Ludwig-Maximilians-Universität München (LMU München)***  
*Studium der Biologie*

10/2001                                      • Erwerb des Vordiploms

10/2001 – 10/2002                      • Erstellung des experimentellen Teils der Diplomarbeit, mit  
 Unterbrechung, am John Innes Centre, Norwich, UK

04/2002 - 07/2004                      • Hauptstudium der Fächer  
 Physiologische Botanik im Hauptfach, sowie  
 Systematische Botanik, Biochemie und Zoologie  
 (Zellbiologie) im Nebenfach

07/2004                                      • Erwerb des Diploms (mit Auszeichnung)  
 • Diplomarbeit: “Die Bedeutung des Genes *SERRATE* in der  
 Ontogenese von *Arabidopsis thaliana*.”

10/2004-05/2008                      • Dissertation am Lehrstuhl Botanik I, Department Biologie I  
 bei Professor Herrmann

### 7.3. Acknowledgement (Danksagung)

Diese Arbeit entstand am Lehrstuhl I des Botanischen Instituts der Ludwig-Maximilians-Universität München in der Arbeitsgruppe von Herrn Professor Reinhold G. Herrmann. Ihm danke ich für die Bereitstellung des Arbeitsplatzes, die hervorragenden Arbeitsbedingungen und die interessante Themenstellung. Darüber hinaus möchte ich ihm dafür danken, dass er meinen wissenschaftlichen Werdegang seit dem Grundstudium begleitet hat und mir immer mit Rat und Tat zur Seite gestanden hat. Auch hat er maßgeblich mein wissenschaftliches Verständnis und meine Entwicklung geprägt.

In diesem Zusammenhang sei auch Herrn Professor Dario Leister gedankt, der mich als neuer Inhaber des Lehrstuhls I immer in meiner Arbeit unterstützt hat.

Auch sei Frau Professor Suanne Renner recht herzlich gedankt, dass Sie es ermöglicht hat, unser *Oenothera*-Feld im Botanischen Garten anzulegen.

Außerdem geht mein Dank an die Hanns-Seidel-Stiftung, durch deren Promotionsstipendium das Erstellen dieser Arbeit erst ermöglicht wurde.

Mein Dank geht auch an Herrn PD Dr. Jörg Meurer für die vielen Gespräche über „Sinn und Unsinn der *Oenothera*“, aber auch für die schöne Zeit, die ich in seinem Labor verbringen durfte.

Des Weiteren danke ich Herrn Professor Wolfgang Stephan für die Erstellung des Zweitgutachtens und dafür, dass er mich im Studium an die molekulare Evolution herangeführt und immer ein offenes Ohr für mich hatte.

Ein besonderer Dank geht an Uwe Rauwolf dafür, dass er vor gut drei Jahren mit mir das Wagnis aufgenommen hat, überhaupt an *Oenothera* zu arbeiten. Ohne ihn wäre es sicher nicht gelungen „die Bestie“ niederzuwerfen, und das Projekt wäre nicht in dem guten Zustand, wie es sich heute darstellt.

Hieronim Golczyk möchte ich aufrichtig dafür danken, dass er sich der *Oenothera*-Zytologie angenommen hat. Eine funktionierende Zytologie war von immenser Wichtigkeit für das Projekt und wäre ohne ihn sicher nicht etabliert worden.

Mein Dank geht auch an Georg Haberer und Xi Wang, die im MIPS-Institut die Bioinformatik für das Projekt betreut und durchgeführt haben und deren Ergebnisse nicht unwesentlich in diese Arbeit eingeflossen sind.

Außerdem möchte ich Serena Schwenkert und Won Kyong Cho danken, Serena für ihre Hilfe bei allem was mit Proteinen zu tun hatte und Won, ohne den es die *Oenothera*-Macro-„arrays“ nicht gegeben hätte.

Dank geht auch an Ulrich Wißnet, Elisabeth Gerick und Martina Reymers. Uli, herzlichen Dank dafür, dass Du jahrelang vorbildlich die *Oenotheren* im Gewächshaus und im Freiland betreut hast. Elli danke ich für die präzise Arbeit bei *Oenothera*-Kreuzungen und -Ernten, aber auch für die Sequenzierung eines großen Teils der fünf Plastome. Martina, danke für Deinen unermüdlichen Einsatz in der *Oenothera*-Gewebekultur.

Ganz herzlich möchte ich mich bei Lena Wasmuth bedanken für all die Jahre der Unterstützung und Anteilnahme, aber auch für Verständnis und Toleranz, wenn ich am Wochenende im Institut war, was eigentlich immer der Fall war.

Mein letzter und größter Dank geht an meine Familie, meine Eltern und meinen Bruder, die mich die ganzen Studienjahre unterstützt haben, und mir immer die Freiheit verschafft haben, das zu tun, was ich für richtig gehalten habe.