



Isolation and characterization of genome differences in the indigenous grass *Monocymbium ceresiiforme*

By

MARIE CATHERINE ONANENA

Dissertation submitted in partial fulfillment of the requirement for the degree

MAGISTER SCIENTIA

Forestry and Agricultural Biotechnology Institute (FABI)

Department of Botany

In the

Faculty of Natural and Agricultural Sciences

University of Pretoria

SUPERVISOR: PROF KJ KUNERT

CO-SUPERVISOR: DR C VAN DER VYVER

March 2004

TABLE OF CONTENTS

	PAGES
ABSTRACT	I
RESEARCH OBJECTIVES	II
DISSERTATION COMPOSITION	III
AKNOWLEDGEMENT	IV
ABBREVIATIONS AND SYMBOLS	VI
INDEX OF FIGURES	X
INDEX OF TABLES	XX
Chapter 1: Plant genome and stress	1
1.1 Abstract	2
1.2 The plant genome	2
1.3 Plant genome variation	6
1.3.1 Induction of genome variation	6
1.3.2 Repetitive DNA and DNA sequence variation	7



1.3.3	Detection of genome variation	10
1.3.3.1	ITS	12
1.3.3.2	RFLP	13
1.3.3.3	RAPD	14
1.3.3.4	AFLP	14
1.3.3.5	SSR	15
1.3.3.6	RDA	17
1.4	Genome variation and grasses	21
1.4.1	The grass genome	21
1.4.2	Characterization of grass genome	22
1.4.2.1	<i>Monocymbium ceresiforme</i>	24
Chapter 2:	Materials and methods	27
2.1	Materials	28
2.1.1	Plant material	28
2.1.2	Microbial material	28
2.2	DNA isolation	29

2.2.1	Grass DNA	29
2.2.1.1	Sample preparation	29
2.2.1.2	Grass DNA isolation	30
2.2.1.3	Grass DNA quantification	30
2.2.2	Bacterial DNA	31
2.2.2.1	Bacterial DNA isolation	31
2.2.2.2	Bacterial DNA quantification	31
2.2.3	Bacterial endophyte	32
2.2.3.1	Isolation and identification of bacterial endophytes	32
2.3	DNA amplification and cloning	32
2.3.1	DNA amplification	32
2.3.2	DNA cloning	33
2.4	DNA analysis	33
2.4.1	DNA sequencing	33
2.4.2	Southern blot analysis	34
2.4.3	DNA slot blotting	34

2.4.4	Colony hybridization	35
2.5	DNA subtraction	36
2.5.1	Amplicon production	36
2.5.2	First round subtraction and amplification	37
2.5.3	Second and third round subtraction and amplification	39
2.5.4	Fourth and fifth round subtraction and amplification	39
2.5.5	Cloning of DNA subtraction products	40
Chapter 3:	Analysis of the ITS region	41
3.1	Abstract	42
3.2	Objective	42
3.3	Results	42
3.3.1	Grass material collection	42
3.3.2	Grass genomic DNA isolation	43
3.3.3	PCR amplification	45
3.3.4	Sequencing analysis	45

Chapter 4:	Execution of the RDA technology	47
4.1	Abstract	48
4.2	Objectives	48
4.3	Results	48
4.4	Production of subtraction products	50
4.4.1	First subtraction products	50
4.4.2	Second subtraction products	51
4.4.3	Third subtraction products	52
4.4.4	Fourth subtraction products	53
4.4.5	Fifth subtraction products	55
Chapter 5:	Characterization of Monocymbium subtraction products	57
5.1	Abstract	58
5.2	Objectives	58
5.3	Results	58
5.3.1	Analysis of first set of subtraction products	58

5.3.2	Genomic DNA amplification with primers designed from subtraction product S3cl2	61
5.3.3	Analysis of the fifth subtraction products	63
5.3.4	Genomic DNA amplification with primers designed from DP510	68
5.3.5	Analysis of fragments amplified with internal DP510 primers	71
Chapter 6:	Characterization of the grass endophytes	74
6.1	Abstract	75
6.2	Objective	75
6.3	Results	75
6.3.1	PCR amplification of DNA of the 16S rRNA region	75
6.3.2	PCR amplification with DP 510 primers	76
6.3.3	Colonies identification	77
Chapter 7:	Discussion and future perspectives	78
7.1	Discussion	79
7.2	Future perspectives	84

Annexure	86
Buffers for DNA work	87
Buffers for Southern blotting	89
Buffers for DNA colony hybridization	92
Buffers for DNA electrophoresis	93
Buffers for RDA technique	95
Cloning reagents	96
Growth media	96
Sequence of primers used in experiments	98
References	100
Sommaire	116

ABSTRACT

In this study genome differences between three types of the grass *Monocymbium ceresiiforme* collected at different locations in South Africa were investigated. For identification and characterization of genome differences, PCR amplification of the ITS region with specific DNA primers designed to amplify the ITS region and Representational Difference Analysis (RDA), which is a rather new subtractive DNA technology for plants, were used. Although PCR products could be amplified with the ITS technique, these products were identified by bio-informatics tools to be of fungal origin possibly due to infestation of seed material with typical grass pathogens. By executing RDA on genomic DNA isolated from grass flowers and application of several rounds of DNA subtraction and kinetic enrichment by PCR reaction, several subtraction products derived from genomic DNA of individual types of grasses were identified and characterized. This included a subtraction product with homology to a highly repetitive maize retro-transposon and a second sequence, called DP510, with homology in part of the sequence to *Bacillus* DNA. Sequence analysis using bio-informatics tools further revealed that DP510 also had homology to genomic *Arabidopsis thaliana* DNA. However, by applying PCR amplification using DNA primers designed to amplify the individual subtraction products, none of the subtraction products was unique to one of the individual grass genomes but was able to identify several variants of DP510. Although experiments were carried out to demonstrate that DP510 has not derived from bacterial contamination of grass DNA, hybridization of labeled DP510 to isolated genomic DNA resulted only in a very weak signal. But, no experiments were carried out for hybridization of bacillus DNA with DP510 by Southern blotting technique. Consequently, there is still a lack of clear indication that DP510 is part of the grass genome.

RESEARCH OBJECTIVES

Molecular tools are increasingly applied for the characterization of plant genomes. Aim of this study was to characterize the genome of the inland grass species *Monocymbium cerasiiforme* occurring in South Africa at different locations with diverse environmental conditions. In particular, the RDA technique has been applied in this study with the objectives (i) to evaluate the potential of the RDA technique for genome characterization and (ii) to identify and characterize possible variations on the genomic level from the grass collected at different locations in South Africa without having any morphological differences, despite growing under different environmental conditions.

DISSERTATION COMPOSITION

Chapter 1 presents a short overview about our current knowledge of the composition of the plant genome including grasses, how stresses can cause variation in the plant genome and where such variation can occur in the genome. This chapter also focuses on the techniques that are widely applied to identify genome variation. Finally, a short introduction is given in this chapter regarding the characteristics of the grass species used in this study. In **Chapter 2** the experimental procedures that have been applied in this study are outlined. **Chapter 3** outlines the results of the different genomic DNA isolation techniques applied to obtain genomic DNA of sufficient quality to carry out a PCR amplification of the ITS region of the different types of grasses investigated. This technique, previously applied by scientists to characterize genome variation in grasses was used as a general molecular technique to identify any advantages of the RDA technique for characterization of genome variations. **Chapter 4** emphasizes on the application of the RDA technique for identification and cloning of putative DNA sequence differences after several rounds of subtractive hybridization and PCR amplification from the different types of grasses used in the study. In **Chapter 5** the results obtained by analyzing the different subtraction products using DNA sequencing and bio-informatics tools for sequence alignments and identification of DNA sequence homologies with known sequence data available in DNA sequence databases are outlined. **Chapter 6** outlines results obtained from PCR amplification of Bacillus DNA sequences and staining techniques for detection of bacterial endophytes in seed extracts to determine possible cross-contamination of grass genomic DNA with bacterial DNA. **Chapter 7** finally outlines the achievements made and also problems experienced in this study regarding the extension of the RDA technique to a further plant species, the identification of repetitive DNA in the grass and the detection of DNA sequences in the plant genome with homology to maize and Bacillus DNA. This chapter also outlines perspectives for future research activities. Finally relevant references listed in this study are listed in **References** and in the **Annexure** details about the composition of buffers, solutions, and other chemicals used in this study are provided.

ACKNOWLEDGEMENTS

First and foremost I would like to thank God, Almighty be all glory, without whom none of this would have been possible. You have guided me thus far; being a companion with your unfailing love, grace and guidance through all lives joys and sorrows. I look forward to traveling this path that you have set for me in your constant presence.

I would like to thank my supervisor, Prof. Karl Kunert for years of patience, enthusiasm and support in teaching. Also for the opportunities he gave to me to share his endless amount of knowledge.

A special thanks goes to M. Du Plessis, whose strong intervention allowed my acceptance at the University of Pretoria.

Thank you to my late mother who passed away the 20th of May 2003. Your death has disturbed deeply this work, but did not discourage me. I would like to dedicate this work to you.

Thank you to my father, my sisters and brother for years of fundamental support and guidance.

This work would not have been achieved without the help of lab colleagues: J Vorster, A. Prins, and A. Kiggundu. Thanks for all your editorial assistance and help with bioinformatics tools to analyze data.

My sincere gratitude goes to Dr C. Van der Vyver for her patience during my lab work, sharing her knowledge and spending her time teaching and guiding me.

I would like to thank the University of Pretoria that allowed me to conduct and complete my research.

My sincere gratitude and appreciation go to Prof. Ladislav (Laco) Mucina, Dept. of Botany, University of Stellenbosch for collecting grasses at Drakensberg.

A big thanks goes also to Prof Braam van Wyk for helping me with the grass distribution and Dr Teresa A. Coutinho for analyzing the bacteria.

Lastly but not the least, I want to thank my husband Boniface, for his constant love, encouragement and confidence in me.

Finally, to my children: Muriel, Raïssa, Glwadys, Donald, Audrey, and Bleurette. I would like through this achievement to be the model of endurance and determination for all of you.

ABBREVIATIONS AND SYMBOLS

°C	Degrees Celcius
%	Percentage
µg	Microgram
µl	Microlitre
A	Adenine
AFLP	Amplified fragment length polymorphism
Amp	Ampicillin
bp	Base pair
BHR	Broad host-range
C	Cytosine
CH ₃ COOHH ₄	Ammonium acetate
cm	Centimeters
CMP	Cytosine monophosphate
cpDNA	Chloroplast DNA
CP	Capside-like protein
CTAB	Cetyltrimethylammonium bromide
DG	Drakensberg grassland
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
dNTP	Deoxynucleoside triphosphate
DP	Difference product(s)
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherchia coli</i>
EDTA	Ethylenediamine tetra acetic acid
EPPS	N- (2-hydroxyethyl) piperazine-N- (3-propane sulfonic acid)
ETS	External transcribed spacer
g	Grams

G	Guanine
HGP	Horizontal gene pool
H	Hours
H ₂ O	Water
HG	Highveld grassland
HGT	Horizontal gene transfer
IGS	Intergenic spacer
IPTG	Isopropyl-[beta]-D-thiogalactopyranoside
IR _A and IR _B	Inverted repeats
INT	Integrase
ITS	Internal transcribed spacer
KDO	3-deoxy-D-manno-octulosonate
Kpa	Kilopascal
l	Liter
LB	Luria Bertani
LTR	Long terminal repeat
LSC	Large region of single copy gene(s)
M	Molar
mer	Oligomer
mg	Milligrams
MGEs	Mobile genetic element
ml	Millilitres
mm	Millimeters
mM	Millimolar
mtDNA	Mitochondrial DNA
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
ng	Nanogram
NaAc	Sodium acetate
NaOH	Sodium hydroxide

NTR	Non-transcribed repeat(s)
NTS	Non-transcribed spacer
nDNA	Nuclear DNA
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphism DNA
RDA	Representational difference analysis
rDNA	Ribosomal DNA
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNAase-H	Ribonuclease H
rRNA	Ribosomal ribonucleic acid
rpm	Revolutions per minute
RT	Reverse transcriptase
SDS	Sodium dodecyl sulphate
sdH ₂ O	Sterile distilled water
sec	Second(s)
SG	Savannah grassland
SP	Subtraction product(s)
ssDNA	Single stranded DNA
SSC	Small region of single copy gene(s)
spp	Species (plural)
SSR	Single sequence repeats
STMS	Sequence tagged marked site
T	Thymine
TAE	Tris-acetate EDTA
T _m	Melting temperature
Tris	2-amino-2- (hydromethyl) propane-1, 3 diol
tRNA	Transfer RNA
UV	Ultraviolet
VNTR	Variable number tandem repeat(s)



X-gal

5-bromo-4-chloro-indol- [beta]-D galactoside

INDEX OF FIGURES

	Page
Figure 1.1	3
The plant cell genomes and their characteristics (V = Vacuole).	
Figure 1.2	5
Genome organization and gene distribution in a plant genome (Sandhu and Gill, 2002).	
Figure 1.3	7
Stress-induced genome variations.	
Figure 1.4	8
General structure of a retro-element (TY1-copia retro-transposon) with long terminal repeats (LTR) in direct orientation at each end. Within the LTRs are U3, R and U5 regions that contain signals for initiation and termination of transcription. Also showing the unique gene order within two open reading frames encoding for a capsid-like proteins (CP); Protease (PR); Integrase (INT); Reverse transcriptase (RT); and RNAase H. Other sequences featured are PBS (primer binding sites), PPT (polypurine tracts), NA (nucleic acid binding moiety) (after Kumar and Bennetzen, 1999).	
Figure 1.5	11
A summary of the qualitative characteristics of DNA based techniques to detect genome variation in plants. Cost = financial requirements to prepare a laboratory work and obtain results. Development time = time required to develop genetic assays that depend on the availability of primers. Precision = diversity present within a sample. (Adapted from O'Hanlon <i>et al.</i> , 2000).	
Figure 1.6	12
Ribosomal RNA genes in the cytosol. Arrangement of the 18S-5.8S-25S RNA gene complexes. IGS = intergenic spacer; ITS = internal transcribed spacer (Henry, 1997).	

Figure 1.7 17
Comparative patterns of some DNA based techniques used to detect genome variation in plants. (A) represents an example of a RFLP obtained from five different and unrelated, inbred maize lines digested with *EcoRI* (five tracks on the left) or *HindIII* (five tracks on the right) and probed with a ³²P- labeled single copy maize sequence (Karp *et al.*, 1996). (B) shows some RAPD patterns obtained from two near isogenic lines of maize amplified with one of eight operon primers (Karp *et al.*, 1996). (C) represents a typical AFLP profile obtained from different breeding lines of sugar beet amplified with a ³³P-labeled (*PstI*) primer and a non-labeled (*MseI*) primer (Karp *et al.*, 1996). (D) shows an example of SSRs profile obtained from different inbred lines of sunflower using a single ³³P-labeled micro-satellite primer pair (Karp *et al.*, 1996).

Figure 1.8 20
Different steps involved in representational difference analysis. (—) represents adaptor sequence; (—) represents tester DNA and (=) represents driver DNA.

Figure 1.9 23
A summary of some DNA based techniques used so far to characterize grass genome variation.

Figure 1.10 24
Vegetative parts of *Monocymbium ceresiiforme*.

Figure 1.11 25
Monocymbium ceresiiforme distribution map in South Africa (National Botanical Institute).

Figure 3.1 43
Different parts of the inland grass *Monocymbium ceresiiforme*. The vegetative parts shown are: the flower part (A), the Emerging leaf of the grass (B) and the grass seed (C).

Figure 3.2

44

Isolated genomic DNA from grasses collected at three different locations, visualized on a 1% agarose gel stained with ethidium bromide. In (A) and (B) lane M represents molecular weight marker III (Roche, Switzerland), while lanes 1 and 2 represent λ DNA of 125 ng and 250 ng respectively. In (A) lane 3 represents 500 ng of λ DNA. (A) Degraded genomic DNA isolated from the leaves and stems indicated by a smear (lanes 4-6). Lane 4 represents DNA isolated from SG grass, lane 5 represents DNA isolated from DG grass, and lane 6 represents DNA isolated from HG grass. (B) Non-degraded genomic DNA isolated from the flowers shown as a high molecular weight band (lanes 3-5). Lane 3 represents DNA isolated from SG grass, lane 4 represents DNA isolated from DG grass, and lane 5 represents DNA isolated from HG grass.

Figure 3.3

45

The ITS region amplified from genomic DNA of the grasses collected at three different areas visualized on an agarose gel stained with ethidium bromide. Lane M represents molecular weight marker III (Roche, Switzerland). Lane 1 represents the ITS amplified product from the SG grass genomic DNA. Lane 2 represents the ITS amplified product from the DG grass genomic DNA. Lane 3 represents the ITS amplified product from the HG grass genomic DNA. The size of the amplification product (532 bp) is indicated with an arrow.

Figure 3.4

46

Sequence data of ITS region amplified from genomic DNA of SG grass. ITS1 primer and ITS4 used for amplification (Table A.7) are underlined.

Figure 4.1

49

Digestion of genomic DNA isolated from Monocymbium collected from three different areas with restriction enzyme *Hind*III (A) and amplicons obtained from genomic DNA after adaptor ligation and amplification with adaptor pair RHind 12 and 24 (B), visualized on an agarose gel after staining with ethidium bromide.

Lane M (A and B) represents a 100 bp ladder (Roche, Switzerland) lane 1 represents (A) SG genomic DNA and (B) amplicon, lane 2 represents (A) HG genomic DNA and (B) amplicon, lane 3 represents (A) DG genomic DNA and (B) amplicon, and lane 4 (A and B) represents 2 µg herring sperm DNA for DNA quantification.

Figure 4.2

50

First subtraction products (SP1) of *Monocymbium cerasiiforme* genomic DNA representations visualized on an agarose gel stained with ethidium bromide. Lane M represents a 100 bp DNA ladder (Roche, Switzerland). Lanes 1 and 2 represent the subtraction product resulting from the subtraction of SG and DG grass representations, where SG grass representation was used as tester and DG grass representation as driver (lane 1), and DG grass representation was used as tester and SG grass representation as driver (lane 2). Lanes 3 and 4 represent the resulting products from the subtraction of DG and HG grass representations, where the HG grass representation was used as tester and DG grass representation as driver (lane 3), and DG grass representation was used as tester and HG grass representation as driver (lane 4).

Figure 4.3

52

Second subtraction products (SP2) of *Monocymbium cerasiiforme* genomic DNA representations at a 150:1 driver to tester ratio, visualized on an agarose gel stained with ethidium bromide. Lanes 1 and 2 represent the resulting subtraction products from the subtraction of the DG grass representation and the SG grass representation, where the DG grass representation was used as driver and the SG grass representation as tester (lane 1), and DG grass representation was used as tester and the SG grass representation used as driver (lane 2). Lanes 3 and 4 represent the reciprocal subtraction of DG grass representation and HG grass representation, where the HG grass representation was used as tester and the DG grass representation was used as driver (lane 3), and the DG grass representation was used as tester and the HG grass representation was used as driver (lane 4). Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

Figure 4.4

53

Third subtraction products (SP3) of *Monocymbium cerasiiforme* genomic DNA representations at a 1500:1 driver to tester ratio, visualized on an agarose gel stained with ethidium bromide. Lanes 1 and 2 represent the resulting products from the hybridization of the DG grass representation and the SG grass representation where the DG grass representation was used as driver and the SG grass representation as tester (lane 1), and DG grass representation was used as tester and the SG grass representation used as driver (lane 2). Lanes 3 and 4 represent the hybridization between the DG grass representation and HG grass representation, where the HG grass representation was used as tester and the DG grass representation was used as driver (lane 3), and the DG grass representation was used as tester and the HG grass representation was used as driver (lane 4). Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

Figure 4.5

54

Fourth round subtraction products (SP4) of *Monocymbium cerasiiforme* genomic DNA representations at a 200000:1 driver to tester ratio, visualized on an agarose gel stained with ethidium bromide, where the third subtraction products were either digested with *MseI* (A), or were undigested with *MseI* (B) before amplification. In both (A) and (B), lanes 1 and 2 represent the subtraction of DG grass representation with the SG grass representation, where the DG grass representation was used as driver and the SG grass representation as tester (lane 1), and the SG grass representation was used as driver and the DG grass representation as tester (lane 2). In both (A) and (B) lanes 3 and 4 represent the subtraction of DG grass representation and HG grass representation where the HG grass representation was used as tester and the DG grass was used as driver (lane 3), and reciprocally in lane 4 the DG grass representation was used as tester and the HG grass representation was used as driver. Lane M represents a 100 bp DNA ladder (Roche, Switzerland). Lane 5 (B) represents 2 µg of herring sperm DNA for DNA quantification.

Figure 4.6

55

Fifth round subtraction products (SP5) of *Monocymbium cerasiiforme* genomic DNA

representations at a 2000000:1 driver to tester ratio, visualized on an agarose gel stained with ethidium bromide. Lanes 1 and 2 represent the subtraction of the DG grass representation and the SG grass representation where the DG grass representation was used as driver and the SG grass representation as tester (lane 1), and the DG grass representation was used as tester and the SG grass representation as driver (lane 2). Lanes 3 and 4 represent the subtraction of the DG grass representation and the HG grass representation, where the HG grass representation was used as tester and the DG grass representation as driver (lane 3), and the DG grass representation was used as tester and the HG grass representation as driver (lane 4). Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

Figure 5.1

59

Screening of plasmid vector *pMosBlue* for insertion of subtraction product obtained after two rounds of subtractive hybridization with *Monocymbium cerasiiforme* representations. Lanes 1-5 represent respectively fragments S3cl2, S3cl12, S3cl18, S4cl27, and S4cl39. Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

Figure 5.2

60

Nucleotide sequence of the subtraction product S3cl2 indicating the sequences used as primers for amplification of SG, DG, HG grass genomic DNAs (underlined).

Figure 5.3

60

Subtraction product S3cl2 sequence (A) and *Zea mays* line LH82 transposon Ins2 sequence (B).

Figure 5.4

61

PCR amplification of grass genomic DNA at 55°C with primers S3cl2L and S3cl2R visualized on an agarose gel stained with ethidium bromide. Lanes 1, 2, and 3 represent the amplification products from SG, DG, and HG grass genomic DNA respectively.

Lane 4 represents herring sperm DNA. Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

Figure 5.5

62

Sequence alignment of the isolated subtraction product S3cl2 and the PCR amplification products clones D22, D25, S40, S45, and H36 showing consensus sequence, with red indicates 100% similarity, 75% similarity in blue and 50% similarity in black. All black dots (*) showed the regions with no similarities. Empty spaces indicate base pairs deletions.

Figure 5.6

64

Sequence alignment of the five sequences obtained as subtraction products after five rounds of subtractive hybridization by using representation DNA from SG grass as tester and representation DNA from DG grass as driver. Nucleotide changes are indicated in bold, and nucleotide deletions with a dash (-).

Figure 5.7

65

Sequence analysis of DP510 (A, B, C, and D). Identified hyper-variable DNA region indicated in green. Perfect homology to *Arabidopsis* sequence indicated in pink. Partial homology to *Arabidopsis* genomic DNA indicated in red. Significant homology to *Bacillus* DNA indicated in blue. E-value-14 was deemed as a lower significant homology.

Figure 5.8

67

Hybridization of first subtraction products (A); second subtraction products (B); third subtraction products (C); fourth subtraction products (D) and fifth subtraction products (E) with labeled DP510 fragment. Lanes 1 and 2 represent the subtraction of SG and DG grass representations, where the SG grass representation was either used as tester (lane 1) or the DG grass representation was used as tester (lane 2). Lanes 3 and 4 represent the subtraction of HG and the DG grass representations, where the HG grass representation

was either used as tester (lane 3) or the DG grass representation was used as tester (lane 4).

Figure 5.9

68

Copy number determination in Southern-blotting of *Eco*RI-digested genomic DNA from DG, HG, and SG grass with labeled DP510. DG, HG, and SG represent genomic DNA from the three types of grass. Reference bands indicate the intensity of 100 and 1000 copies of DP 510 (top two rows).

Figure 5.10

69

Sequence of DP510, indicating the sequences used as primers for amplification of SG grass genomic DNA (underlined).

Figure 5.11

69

Colony blot of *E. coli* colonies carrying cloned DP510 amplification products from genomic DNA of SG grass and using labeled DP510 as a probe for hybridization. Colonies Bh31, Bh48, and Bh79 gave the lowest hybridization signal, which were used for further characterization.

Figure 5.12

70

Sequence alignment of difference products obtained after fifth round of subtractive hybridisation (DP56, DP57, DP58, DP59, and DP510), with amplified and selected DNA fragments from SG grass genomic DNA using primers Bhal5A and Bhal3A designed from DP510 (Bh31, Bh48, and Bh79). Detected nucleotide differences are indicated in bold.

Figure 5.13

71

Sequence of DP510 with primers Bhal5R and Bhal3L (underlined) designed to amplify an internal segment of sequence in grass genomic DNA.

Figure 5.14

72

Amplification products of grass genomic DNA amplified with primers designed from internal part of DP510 visualized on an agarose gel stained with ethidium bromide. Lane 1 represents the amplified fragment from DG grass genomic DNA; lane 2 the amplified fragment from SG grass genomic DNA and lane 3 represents the amplified fragment from HG grass genomic DNA. Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

Figure 5.15

73

Sequence data of the amplified fragments obtained after amplification of genomic DNA with primers designed for amplification of the internal part of DP510. Sequence of amplified fragments obtained from the SG grass, DG grass and HG grass genomic DNA are indicated by SG3, DG1, and HG4 respectively.

Figure 6.1

76

Amplification of *Bacillus subtilis* 16S rRNA region with specific primers designed for a region of the 16S rRNA region and products visualized on a 1.5% agarose gel stained with ethidium bromide. Lane 1 represents amplification product from *Bacillus subtilis* DNA. Lanes 2, 3, 4 represent isolated DG, HG, SG grass genomic DNA derived from grass flowers. Lane 5 represents H₂O control. Lane 6 represents genomic DNA isolated from a mixed bacterial culture derived from crashed SG seeds. Lane M represents 100 bp DNA ladder (Roche, Switzerland).

Figure 6.2

77

Amplification of DP510 fragment covering the region with homology to the *Bacillus* DNA from grass genomic DNAs and the *Bacillus subtilis* DNA visualized on a 1.5% agarose gel stained with ethidium bromide. Lane 1 represents genomic DNA from DG grass, lane 2 HG grass and lane 3 SG grass. Lane 4 represents H₂O control without DNA addition and lane 5 represents amplification of the predicted 157 bp DP510 fragment from isolated *Bacillus subtilis* DNA.



Lane M represents a 100 bp DNA ladder (Roche, Switzerland).

INDEX OF TABLES

	Page
Table 5.1	66
Sequence analysis of DP 510.	
Table A.1	87
DNA extraction buffer (pH 8)	
Table A.2	91
Buffers/Southern blotting	
Table A.3	93
TAE buffer (50xstock)	
Table A.4	94
DNA loading buffer	
Table A.5	94
Agarose gel composition	
Table A.6	97
Growth medium composition	
Table A.7	98
Sequences of all the primers used for amplification of the DNA regions. “ITS” represents the sequence of the primer set used to amplify the internally transcribed spacer sequence (ITS region); “NTS” primer set used to amplify the non-transcribed spacer (NTS region);	



“Retrotransposon” primer used to amplify a retro-transposon–like region from the grass species *Monocymbium cerasiiforme*. “DP510” primer set used to amplify a fragment with homology to the *Bacillus halodurans* region and “*Bacillus subtilis*” the primer set used to amplify the *Bacillus subtilis* 16s rRNA region.

Table A.8

99

Sequence of the three adaptor sets used for execution of the RDA