Katholieke Universiteit Leuven

Faculteit Bio-ingenieurswetenschappen

Departement Biosystemen

Afdeling Plantenbiotechniek

DISSERTATIONES DE AGRICULTURA

Doctoraatsproefschrift nr. 787 aan de faculteit Bio-ingenieurswetenschappen van de K.U.Leuven

Characterization and isolation of T-DNA tagged banana promoters active during *in vitro* regeneration and low temperature stress

Proefschrift voorgedragen tot het
behalen van de graad van
Doctor in de
Bio-ingenieurswetenschappen

door

Efrén Germán
SANTOS ORDOÑEZ

APRIL 2008

Doctoraatsproefschrift nr. 787 aan de faculteit Bio-ingenieurswetenschappen van de K.U.Leuven

ISBN 978-90-8826-046-9
Wettelijk depot D/2008/11.109/4

Katholieke Universiteit Leuven
Faculteit Bio-ingenieurswetenschappen
Departement Biosystemen
Afdeling Plantenbiotechniek

DISSERTATIONES DE AGRICULTURA

Doctoraatsproefschrift nr. 787 aan de faculteit Bio-ingenieurswetenschappen van de K.U.Leuven

Characterization and isolation of T-DNA tagged banana promoters active during *in vitro* regeneration and low temperature stress

Promotor:
Prof. R. Swennen, K.U.Leuven

Leden van de examencommissie:
Prof. R. Geers, voorzitter
Dr. S. Remy, K.U.Leuven, co-promotor
Prof. D. Van Der Straeten, UGent
Prof. J. Michiels, K.U.Leuven
Prof. B. Cammue, K.U.Leuven
Dr. M. Souza Júnior, EMBRAPA, Brazilië

Proefschrift voorgedragen tot het behalen van de graad van Doctor in de Bio-ingenieurswetenschappen

door

Efrén Germán SANTOS ORDOÑEZ

APRIL 2008

This work is dedicated to my wife Daniela

Acknowledgements

There were many mediators involved in the creation of this thesis and I would like to thank all of them for their contribution. First, my gratitude goes to the program established by the Vlaamse Interuniversitaire Raad (VLIR) and coordinated by Dr. Magda Vincx. This program was executed together with the Escuela Superior Politécnica del Litoral (ESPOL) from Ecuador, formerly coordinated by Dr. Jorge Calderón and now by Dr. Sergio Flores. I thank them all for the financial support to accomplish my Ph.D. Indeed, the research presented in this book was executed within the VLIR-ESPOL framework and more specifically, as part of CIBE (Centro de Investigaciones Biotecnológicas del Ecuador) which is located on the ESPOL campus in Guayaquil. This thesis research fits within the objective of CIBE to form researchers capable of using biotechnology as a tool for the social, economic and conservation development of the plant genetic resources in Ecuador.

I am deeply indebted to my promotor Prof. Rony Swennen for allowing me to conduct the research in the Laboratory of Tropical Crop Improvement at the Katholieke Universiteit Leuven. I am deeply grateful to my direct supervisor, Dr. Serge Remy, for introducing me to the practical techniques in the laboratory and for his guidance throughout the research and writing process. I have to stress my appreciation for all his efforts and my admiration for his patience. Also, I would like to thank Dr. László Sági for the interesting discussions, helpful suggestions and guidance during my research.

Words of thanks go to the members of my examination committee, Prof. Dominique Van Der Straeten, Prof. Jan Michiels, Prof. Bruno Cammue, Dr. Manoel Souza Júnior and Prof. Rony Geers, for critically reading my text and the constructive remarks.

The team of CIBE also contributed to this thesis and I sincerely thank the entire team formerly directed by Dr. Helga Rodríguez and presently lead by Dr. Esther Lilia Peralta. Especially, I am grateful to the first CIBE director, the late Dr. Rodolfo Maribona for his motivation and leadership. He was a driving force behind the development of the banana biotechnology research in Ecuador. He nurtured my enthusiasm for banana biotechnology, and believed in the capability of our team to deliver good results for developing countries. I am also grateful to Alberto Ortega, the coordinator of the Master Program in Agricultural Biotechnology (Programa de Maestría en Biotecnología Agrícola) of ESPOL for involving me in teaching activities at ESPOL during my Ph.D. research. I also appreciate the scientific interactions and company of my colleagues María Isabel Jiménez and Oscar Navarrate which are also following Ph.D. research in the framework of the VLIR-ESPOL project.

I also greatly appreciate the help of all the people of the Laboratory of Tropical Crop Improvement and especially the molecular biology group members Saskia Windelinckx, Els Thiry, Ann Krauss, Galein Kellens, Nancy Podevin, Wim Dillemans,

Steven Luypears, Liesbet Smeijers, Leandro Diniz, Bert Coemans and Geoffrey Arinaitwe. It is thanks to the proper administrative management that I never got lost in all the paperwork and thus, I sincerely thank Marleen Stockmans and Suzy Voets from the Laboratory of Tropical Crop Improvement, Annick Verheylezoon from VLIR and Ana Armijos from CIBE.

Friendship encompasses support at all times and that is why I want to thank all my friends. The company and continuous support of my new friends in Belgium made me feel at home.

Finally, one of the strongest forces throughout the work was my family's support. My wife Daniela has been essential for this thesis. Her unconditional and unrelenting support and energy encouraged me to finish. Asimismo, agradezco a mis padres por todo el apoyo que me dieron durante la realización de esta tesis.

Table of contents

List of Tables

List of Figures

Abbreviations

2,4-D	2,4-dichlorophenoxyacetic acid
ABA	abscisic acid
ABRE	abscisic acid-responsive element
ACC	1-Aminocyclopropane-1-carboxylate
A-PCR	anchored PCR
act	actin gene
AD	arbitrary degenerated
*Adh*1	alcohol dehydrogenase gene of maize
AMV	alfalfa mosaic virus
BA	benzyladenine
BAC	bacterial artificial chromosome
BBTV	banana bunchy top virus
BLA	baseline luciferase activity
bp	base pairs
BSA	bovine serum albumin
BSV	banana streak badnavirus
Caca	cooking banana 'Cacambou', ABB genomic group
CaMV	cauliflower mosaic virus
CAT	chloramphenicol acetyltransferase
CBF	C-repeat binding factor
CCD	charge-couple device
cDNA	complementary DNA
CIAP	calf intestinal alkaline phosphatase
COR	cold regulated
cos	constitutive expression of osmotically responsive genes
CRM	*cis*-regulatory module
CRT	C-repeat
Cv	BSV isolated from banana cultivar 'Williams'
DIG	digoxigenin
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate (adenine, cytosine, guanine and thymine)
DPE	downstream promoter element
DRE	dehydration-responsive element
DREB	dehydration-responsive element binding protein
DTT	dithiothreïtol
ECS	embryogenic cell suspension
EDTA	ethylenediaminetetraacetic acid
EFE	ethylene forming enzyme gene of banana
ELTLA	enhanced low temperature LUC activity
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EST	expressed sequence tag
EXO	*EXORDIUM* gene of *Arabidopsis thaliana*
*FEH*IIa	fructan 1-exohydrolase gene of *Cichorium intybus*
*Foc*4	*Fusarium oxysporum* f. sp. c*ubense* race 4
FW	fresh weight
gDNA	genomic DNA
gfp / GFP	green fluorescent protein of the Pacific jellyfish *Aequoria victoria*
GMGC	Global *Musa* Genomics Consortium
GN	dessert banana 'Grand naine', AAA genomic group
Gn^{50}	50 µg mL^{-1} geneticin
GSP	gene-specific primer
GUS	*β*-glucuronidase enzyme
HF	Expand High Fidelity PCR System, Roche
hos	high expression of osmotically responsive genes
hpt	hygromycin phosphotransferase gene
IAA	indolyl-3-acetic acid

ICE1	inducer of CBF expression 1
ICEr1	induction of CBF expression region 1
ICEr2	induction of CBF expression region 2
I-PCR	inverse PCR
ITC	International Transit Center
kb	kilo base pairs
LB	left T-DNA border sequence
LDC	lysine decarboxylase gene of *Arabidopsis thaliana*
los	low expression of osmotically responsive genes
LT	low temperature
LTLA	low temperature LUC activity
LTRE	low-temperature-responsive element
LUC	luciferase enzyme
luc	luciferase gene of the American firefly *Photinus pyralis*
luc^+	codon-optimized *luc*
luc^{+INT}	luc^+ containing an intron
MCS	multiple cloning site
MQ	milliQ
mRNA	messenger RNA
MS	Murashige and Skoog culture medium
MTA	material transfer agreement
MU	4-methylumbelliferone
MW	molecular weight
My	BSV isolated from banana cultivar 'Mysore'
NAA	α-naphthaleneacetic acid
NEB	New England Biolabs
neo	neomycin phosphotransferase II gene of transposon *Tn*5
nt	nuclotide
ocs	octopine synthase gene of *Agrobacterium tumefaciens*
ORF	open reading frame
OsP5CS2	Δ^1-pyrroline-5-carboxylate synthetase 2 gene of *Oryza sativa*
P4	half-strength MS medium supplemented with 1 µM IAA and 100 µM BA
PB	particle bombardment
PCR	polymerase chain reaction
PD	Petri dish
pe35S	enhanced promoter of the CaMV 35S RNA
phas	β-*phaseolin* gene of *Phaseolus vulgaris*
PIC	preinitiation complex
PROL	half-strength MS medium supplemented with 1 µM IAA and 10 µM BA
*PvChi*4	class IV chitinase gene of bean
PVP	polyvinyl pyrrolidone
RB	right T-DNA border sequence
RD1	half-strength MS medium supplemented with 100 mg L^{-1} myo-inositol
RD2	half-strength MS medium supplemented with 1 µM BA
RD29A	responsive to desiccation gene 29A of *Arabidopsis thaliana*
REG	half-strength MS medium supplemented with 1 µM IAA and 1 µM BA
RLU	relative light units
RNA	ribonucleic acid
ROI	region of interest
RT-PCR	reverse transcriptase PCR
ScBV	sugarcane bacilliform badnavirus
SDS	sodium dodecyl sulfate
SP	specific primer
SVC	settled cell volume
T35S	CaMV 35S RNA terminator
TaBV	taro bacilliform virus
TAIL-PCR	thermal asymmetric interlaced PCR
Taq	*Thermus aquaticus* bacterium
T-DNA	transferred DNA
TDZ	thidiazuron; half-strength MS medium supplemented with 10 µM thidiazuron

TF	transcription factor
TGA	transient GUS activity
THP	plantain 'Three Hand Planty', AAB genomic group
TL	time-lapse
Tm	melting temperature
Tm^{200}	200 µg mL^{-1} timentin
Tnos	nopaline synthase terminator
Tris	tris(hydroxymethyl) aminomethane
TSS	transcription start site
*Ubi*1	maize ubiquitin
*uid*A	β-glucuronidase gene of *Escherichia coli* (*gus*A)
uidA^{INT}	β-glucuronidase gene containing an intron
UTR	untranslated region
VR-ACS1	ACC synthase gene of *Vigna radiata* L.
Will	dessert banana 'Williams', AAA genomic group
X-Gluc	5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid cyclohexylammonium salt
ZmGLU1	*β*-glucosidase gene of maize
ZZ	half-strength MS medium supplemented with 5 µM 2,4-D and 1 µM zeatin

Abstract

A genome-wide T-DNA tagging strategy was pursued for the characterization and isolation of novel banana promoters. Embryogenic cell suspensions were transformed *via Agrobacterium tumefaciens* containing a promoterless, codon-optimized luciferase gene either without (*luc*$^{+}$) or with an intron (*luc*$^{+INT}$) next to the right T-DNA border. Approximately 89,000 transgenic cell colonies were first screened for baseline luciferase (LUC) activity at 26°C two to three months after transformation. A 1.6- to 4.3-fold higher tagging frequency was obtained with the *luc*$^{+INT}$ containing vector (pKCKUL1) than with the *luc*$^{+}$ vector (pETKUL2). Screening in real-time then continued under controlled temperature conditions in which LUC activity was monitored during a gradual decrease to different low temperature (LT) treatments including 18°C, 16°C, 12°C and 8°C. Luciferase activation frequency in cell colonies subjected to 26°C and different LT treatments ranged from 0.17% to 2.69%. Observed patterns included an enhanced, decreased or *status quo* LUC activity at LT relative to the 26°C LUC activity. Transgenic cell colonies responsive to 8°C were regenerated to *in vitro* plants with a real-time screening for LT regulated LUC activation at each developmental stage. The tagged lines ET2-17 and ET2-42 showed an enhanced LUC activity at 8°C at the cell colony stage. However, LUC activity of the latter line was not up-regulated by LT anymore at the shoot induction and *in vitro* plant stages. On the other hand, in the tagged line ET2-17 the up-regulation of LUC activity by LT at cell colony stage (10.7- fold) remained present at the shoot induction stage (2.5-fold up-regulation) but was absent at *in vitro* plant stage. In addition, other promoter-tagged lines with different levels of LUC activity at 26°C at cell colony stage showed a decrease of LUC activity at 8°C during the subsequent stages of the *in vitro* regeneration process (lines ET2-34, ET2-85 and ET2-156).

The number of T-DNA inserts in ten independent promoter-tagged lines tested averaged 3.6 with a range from 1 (line ET2-34) to 5 (lines ET2-17 and ET2-156). Isolation of T-DNA flanking sequences was accomplished *via* TAIL-PCR and I-PCR. Sequence analysis of these flanks revealed the presence of direct tandem repeats, vector backbone and/or T-DNA rearrangements in up to half of the lines analyzed. The lines ET2-49, ET2-89, ET2-111 and ET2-156 contained in one RB T-DNA flanking (5'-tagged) sequence the enhanced CaMV 35S promoter with part of the selectable marker gene *neo* from the tagging construct. Continuity of sequence between the corresponding right and left border sequences was revealed by linking PCR using flanking sequence specific primers. *In silico* analysis of the four 5'-tagged banana sequences in line ET2-17 suggested that two candidate promoters were tagged. The 5'-tagged sequence in line ET2-34 harbored a repetitive region, while in line ET2-85 three 5'-tagged sequences were retrieved with one most likely linked to a gene and containing a near-canonical TATA box.

An RT-PCR approach was followed to identify and confirm the sequence that activated LUC expression, which is of paramount importance in lines carrying multiple T-DNA copies. A forward primer annealing to the 5'-tagged sequence near (within 1 to 70 bp) the RB T-DNA sequence was employed in combination with a reverse primer complementary to the *luc*$^{+}$ gene. Transcriptional fusion between (the) 5'-tagged sequence(s) and the *luc*$^{+}$ gene was detected in tagged lines ET2-17, ET2-34, ET2-85 and ET2-156.

Candidate promoter sequences were cloned upstream of the *uid*A^{INT} reporter gene and back-transformed to banana. One (17-1) of two transcriptionally active tagged sequences in the promoter-tagged line ET2-17 was found active in back-transformed banana lines throughout *in vitro* development confirming its promoter characteristics. The LT up-regulation of promoter sequence 17-1 in back-transformed lines at early undifferentiated cell colony stage resembled that of the original promoter-tagged line ET2-17. In contrast, the 5'-tagged sequence 34-1 did not show promoter activity in back-transformed cultures irrespective of the temperature treatment and developmental stage. Finally, the promoter activity of the 5'-tagged sequence 85-1 in back-transformed lines was similar to the relatively strong activity in the original promoter-tagged line ET2-85.

In conclusion, T-DNA tagging has proven a reliable and reproducible method to characterize and isolate novel promoters in banana. Despite their relatively low activity in *in vitro* cultures compared to the maize ubiquitin promoter, the discovered promoters might be useful for banana improvement which warrants further research.

Chapter 1 Introduction

1.1 Banana

1.1.1 Importance

The global production of bananas and plantains (*Musa* spp.) was 70.7 and 34.0 millions tons, respectively, in 2006 (FAOSTAT 2006, faostat.fao.org). Bananas, henceforth including plantains, are the fourth most important crop in developing countries after rice, wheat and maize. It provides energy at low cost and concurrently it is an important source of income for millions of poor people in the humid tropics. Approximately 87% of the world production is produced by smallholders for home consumption or for sale in food markets, while the remaining 13% of the production is aimed at the export market. Therefore, banana plays an important role in poverty alleviation particularly in Africa while especially in Latin-America the banana export industry is part of the economic backbone. For instance, Ecuador is the leading banana export producer in the world.

1.1.2 Growing regions of banana delineated by temperature

Several abiotic stresses affect banana growth including wind, drought, salinity and cold. Low temperature (LT) is perhaps the most important environmental constraint for plant distribution on land (Ishitani *et al.* 1997) and an important factor that limits the geographical area suitable for growing a particular plant species (Ishitani *et al.* 1998; Lee *et al.* 2004a).

Temperature, humidity and rainfall need to be high (22-31°C, 2000 to 2500 mm/year, Robinson 1996) and evenly distributed for optimal growth and high yield of bananas. Due to the proximity of a market, bananas are also grown commercially in subtropical regions like the Canary Islands (Galán Saúco *et al.* 1992). However, temperatures are below optimum during several months of the year in these regions causing growth arrest and delay in fruit filling (Israeli and Lahav 2000). The subtropics are situated between 20° and 30° north and south of the equator. The climatic characteristics of the subtropics include broad fluctuations of temperatures between day and night and summer and winter (~ 25°C/10°C, Robinson 1996). In contrast, in tropical banana regions, geographically distributed between the equator and latitudes 20°N and 20°S, have favorable climatic conditions for growth as temperature fluctuations are limited between day and night, and between summer (dry) and winter (wet) seasons (Robinson 1996).

1.1.3 Effects of low temperature on banana development

The banana plant differs considerably from the majority of horticultural plants. Banana is a monocotyledonous, herbaceous, evergreen perennial plant (Robinson 1996). Under the assumption that water supply is sufficient, temperature determines the rate of banana growth and development. Cessation of new leaf emergence was detected when temperatures drops below 16°C (minimum temperature ~11°C) and growth is arrested below 14°C (minimum temperature ~9°C), while optimum temperature for leaf emergence, and growth and flower initiation are 31°C and 22°C, respectively (Turner and Lahav 1983; Robinson and Anderson 1991). Consequently, LT clearly affects banana plant development and decreases productivity (Robinson 1996).

The first symptom of low-temperature damage, a shortening of the distance between petioles, is observed when the day/night temperatures reached 17/10°C (Israeli and Lahav 2000). When temperatures drop below 9-10°C at the interior of the pseudostem (Ganry 1973) or in the growth chamber (Turner and Lahav 1983) banana growth is arrested and chilling injury occurs (Israeli and Lahav 2000). Chilling injury in fruit is already observed at an exposure of several hours to temperatures below 11-14°C (Israeli and Lahav 2000). Underpeel discoloration, caused by a redish-brown streaking in the vascular tissue below the epidermis of the fruit, is a symptom when the fruit is subjected to chilling temperatures (Robinson 1996). Plant physiological injuries due to low temperatures include choke throat, November dump and growth cessation (reviewed by Robinson 1996; Israeli and Lahav 2000). Choke throat is a phenomenon where the emergence of the inflorescence is hindered due to the short distance between leaf petioles and is caused when normal leaf emergence is restricted by LT (Robinson 1996). November dump (in the southern hemisphere) is synonymous of ‘May bunch’ in the northern hemisphere and happens when flower initiation coincides with low night temperatures (Robinson 1996). The resulting bunches are small and malformed. A gradual yellowing of the leaf is observed due to chilling injury at temperatures ranging from 5°C to 8°C (Israeli and Lahav 2000). Other symptoms of chilling injury include inhibition of metabolic processes and changes in membrane properties leading to an increased permeability and leakage resulting in reduced growth and yields (Kang *et al.* 2003). A decrease in the width and height of the plants is observed in LT environments, although for instance the number of leaves per cycle is identical between greenhouse and open-air conditions in the Canary Islands (Galán Saúco *et al.* 1992). The interval between successive harvests is shorter in greenhouse than in open-air bananas due to a reduction of winter leaf emission rate and flower emergence-harvest interval, which correlates with the increase in temperature under greenhouse conditions (Galán Saúco *et al.* 1992).

Frost damage is observed when banana plants are exposed even for few minutes below 0°C, which kills the leaves that turn black (Robinson 1996). Although temperatures

below 0°C are lethal for banana plants, successful cryopreservation of banana tissues including meristem cultures and apical meristems has been accomplished (Panis *et al.* 2005a). Lethal ice crystal formation is avoided by reducing the moisture content during an acclimation period in the presence of a high sucrose concentration (0.4 M) prior to cryopreservation (Panis *et al.* 1996, 2002). Enhanced expression of proteins involved in the energy-conserving glycolysis, the conservation of the cell wall and the maintenance of an osmoprotective intracellular sucrose concentration has been detected in meristem cultures during this acclimation period (Carpentier *et al.* 2007). Whether similar changes in protein expression occur in LT stressed *in vitro* cultures and field plants remains to be determined.

LT stress can render plants more susceptible to pests and diseases which has been unequivocally observed in subtropical banana growing regions. For instance, Fusarium wilt disease symptoms caused by the root pathogen *Fusarium oxysporum* f. sp. c*ubense* race 4 (*Foc*4), were detected in Australia only following LT stress of 'Williams' banana plants indicating an increased susceptibility after LT (Moore *et al.* 1993). Pegg *et al.* (1996) suggested that environmental factors such as LT may suppress the host responses to infection through a reduction of the energy efficiency of the roots, while others attribute the increased susceptibility to a disruption of the photoassimilation mechanisms compromising the resistance of 'Williams' to *Foc*4 populations (Moore *et al.* 1993; Moore 1994). Although Cavendish cultivars including 'Williams' are usually highly resistant to *Foc*4 in the tropics, Fusarium wilt symptoms are present on Cavendish cultivars in at least four different production areas in the subtropics (Australia, New South Wales and Queensland; Canary Islands; South Africa, Natal and Transvaal; and Taiwan; Ploetz *et al.* 1990).

1.2 Plant promoter activity during development and low temperature

1.2.1 Transcription initiation

Initial protein-encoding gene expression steps in eukaryotes involve a series of multiple events including decondesation of the locus, nucleosome remodeling, histone modifications, binding of transcriptional activators and coactivators to enhancers and promoters and recruitment of the basal machinery to the core promoter (see paragraph 1.2.2) for transcription initiation (reviewed by Smale and Kadonaga 2003). Chromatin remodeling controls the availability of DNA sequences known as promoters which have to be in contact with the transcription machinery in order to start transcription. DNA is packaged into chromatin and, therefore, the chromatin structure plays a critical role in gene expression and regulation (reviewed by Mellor 2006). Once the gene locus is accessible, the promoter is the most important factor controlling the expression of the corresponding gene. In this paragraph, the basic plant promoter structure, function, prediction and characterization *via* reporter gene assays are discussed.

1.2.2 Promoter structure and function

The basic structure needed for gene expression is denominated the promoter. A promoter can be catalogued as (1) a DNA region upstream of a transcription start site (TSS) of a gene; or as (2) a *cis*-acting element directing the rate of transcription initiation of a gene (Praz *et al.* 2002) which is included in (1). In the first category, gene regulatory sequences located in the DNA region upstream of the TSS are recognition sites for transcriptional activators and repressors (Wu *et al.* 2001). They are also denominated as *cis*-acting elements or transcription factor (TF) binding sites (Venter and Botha 2004). Hence, *cis*-acting elements play an important role in regulating gene specific patterns of expression (Wu *et al.* 2001), and can have a length from 5 to 20 nucleotides (nt) (Rombauts *et al.* 2003). Core promoter elements cover the second category (Wu *et al.* 2001) and they are in contact with the basal transcription machinery. Therefore, the core promoter is a minimal DNA region that is capable of initiating basal transcription (Lee and Young 2000). Transcriptional activators recruit the RNA polymerase II-containing transcription initiation apparatus to promoters of protein encoding genes (Lee and Young 2000). Prior to transcription initiation, the RNA polymerase II is involved in a stepwise assembly of general TFs such as the TFIID, which contains a TATA-box binding protein and therefore binds to the TATA-box of the core promoter, and in combination with other TFs the RNA polymerase II is then recruited to the core promoter to form a stable preinitiation complex (PIC) (reviewed by Venter and Botha 2004). The presence of TFs on a promoter allows the interaction of other TFs with other *cis*-acting elements in a tridimensional configuration to activate transcription (Berk 1999; Struhl 2001). To establish the tridimensional configuration needed for transcription initiation, *cis*-acting elements are grouped into modules which confer specific expression of the downstream gene (Rombauts *et al.* 2003). The *cis*-acting element modules or *cis*-regulatory modules (CRMs) can be arranged in different configurations including (1) the 'typical' CRM which contains a compact arrangement of clustered TF binding sites, (2) the 'diffuse' CRM in which several TF binding sites are distributed over a large DNA fragment, (3) the 'composite' CRM containing clusters of TF binding sites distributed over a large DNA fragment, and (4) the 'simple compact' CRM lacking a dense cluster of binding sites (Halfon 2006). A CRM is also known as an enhancer (Halfon 2006; Janssens *et al.* 2006). Specific activation of the common bean (*Phaseolus vulgaris*) β-*phaseolin* (*phas*) promoter during embryogenesis of transgenic tobacco lines is driven by the combinatorial effort of specific *cis*-acting elements needed for *phas* expression in certain tissues of the embryo such as the hypocotyl which are present in the 295 nt region proximal to the TSS (Li and Hall 1999). Although, upstream sequences modulate and enhance *phas* promoter activity, a 'typical' CRM needed for the high level spatially and developmentally regulated expression in embryos, cotyledon and hypocotyl tissues might be present in the -295 nt proximal region containing elements

(Bustos *et al.* 1991, Li and Hall 1999). Redundancy of function has been observed in several elements of the *phas* promoter and the combinatorial relation between *cis*-acting elements are important for the module (cotyledon, hypocotyl and radicle)-specific expression in the embryo (Chandrasekharan *et al.* 2003, see paragraph 1.2.4.2). In addition, expression patterns of the *phas* promoter in tobacco and Arabidopsis were identical for both species indicating a common interaction between transcription factors and DNA in distant related plants. Multiple *cis*-elements could act as activators, repressors, enhancers and/or chromatin modifiers which are crucial for the combinatorial transcriptional regulation in plants (Venter and Botha 2004).

A 'typical' eukaryotic core promoter consists of a TATA box located around 30 nt upstream from the TSS (Breathnach and Chambon 1981), an initiator region located at the TSS (Smale 1994) and a downstream promoter element (DPE) located approximately 30 nt downstream of the TSS (Kutach and Kadonaga 2000). Approximately 30-50% of known plant promoters contain a TATA box 45 to 25 nt upstream of the TSS (Shahmuradov *et al.* 2005). A high number of strongly expressed genes contain a TATA box, although some housekeeping and photosynthesis related genes lack a TATA box (Dynan 1986; Nakamura *et al.* 2002). Promoters without TATA boxes are denominated TATA-less promoters and their TSS might be controlled by the sequence surrounding the TSS (Smale 1997) or the DPE (Shahmuradov *et al.* 2005). Recently, a pyrimidine patch (composed of C and T) was identified around the TSS of plant promoters (Yamamoto *et al.* 2007a, 2007b, 2007c). The 'YR Rule', which consist of a pyrimidine (C/T) and a purine (A/G) at positions -1 and +1 at the TSS, respectively, is found in 77% (10,806) of the Arabidopsis promoters analyzed and is also applicable to rice promoters (Yamamoto *et al.* 2007a). The 'YR Rule' can be considered as a less stringent initiator region and therefore might be recognized by TFIID (Yamamoto *et al.* 2007a). In addition, the 'YR Rule' is in agreement with the consensus around the TSS found in 217 dicot promoters (Shahmuradov *et al.* 2003). Moreover, DNA elements further downstream of the TSS and located in gene regions, such as the 5' untranslated region (UTR), can (co-)regulate a specific level or pattern of expression. For instance, among the versions tested the maize ubiquitin promoter including 5' UTR and first intron has the highest activity (Christensen *et al.* 1992). The presence of the *Adh1* (maize) intron between the promoter of the phosphate transporter gene from barley and the reporter gene also increased the level of expression approximately 20-fold but did not appear to affect the specificity of expression (Schünmann *et al.* 2004).

1.2.3 Promoter prediction

Promoters are indirectly identified *via* gene expression data analyzes (cDNA/EST libraries, microarrays). For instance, comparison of expressed sequences to genomic sequences and further identification of 5' region of the expressed genes allowed the identification of

14,034 candidate promoters using full-length cDNAs in Arabidopsis (Seki *et al.* 2002). Identification of promoters prior to knowing the activity properties of a DNA sequence or expression pattern of the corresponding gene is not a simple task. Analysis of *cis*-acting elements in plant candidate promoter sequences can be performed using plant transcription factors databases such as PLACE (Higo *et al.* 1999) and PlantCARE (Lescot *et al.* 2002). However, any given sequence contains multiple putative *cis*-acting elements due to their short length which might not be all biologically functional. Therefore, analysis of sequences with bioinformatic tools in general is only indicative and not sufficient to prove their function (Radhamony *et al.* 2005). Indeed, for most plant promoters due to the limited information about transcription factors/*cis*-acting elements interactions and functional analyses of all the candidate promoter elements involved in a specific expression activity their identification is cumbersome. Knowledge of the balance between activators and repressors in specific tissues or developmental stages helps to elucidate the activity pattern of specific promoters. Despite the fact that the *phas* promoter inactivity in vegetative tissues is due to chromatin-mediated repression (reviewed by Li *et al.* 2001), dimethyl sulfate *in vivo* footprinting assays indicate that specific DNA-protein interactions may have an important role in *phas* promoter activity beyond the ubiquitous packing in nucleosomes (Li and Hall 1999). Most of the *cis*-acting elements identified in the *phas* promoter are protected in seed in contrast with the almost non-protected elements in leaf tissue as shown by *in vivo* footprinting assays (Li and Hall 1999). Furthermore, Li and Hall (1999) suggested that the overall expression of the *phas* promoter in the embryo is the combination of the different DNA-protein interactions that take place in different regions (modules) of the embryo (cotyledon, hypocotyl and radicle). Therefore, unique TFs may be active (or present) in certain tissues (regions) of the embryo and that these unique TFs together with other TFs that are shared in all the embryo tissues might interact with specific promoter elements and then start transcription from the *phas* promoter in a tissue specific manner in the embryo (Li and Hall 1999). Furthermore, Janssens *et al*. (2006) predicted the activity of the control region of the *Drosophila melanogaster* gene *even skipped* based on the sequence and the information about relevant transcription factors, including activators and short-ranged repressors (~150 nt), and their binding specificity.

The prediction of plant promoters has become an active area of research (Pandey and Krishnamachari 2006). Besides screening for *cis*-acting elements, other factors have to be considered in promoter prediction including determination of CpG islands (regions that are rich in CpGs, Rombauts *et al.* 2003) and DNA bending and curvature (Marilley and Pasero 1996; Schatz and Langowski 1997). CpG regions in promoters are often methylated and associated with transcriptional gene silencing. On the other hand, CpG islands in gene regions are often not methylated and are linked to active expression (reviewed by Rombauts *et al.* 2003). However, determination of promoters based on analyses of CpG islands was not straigthforward in Arabidopsis (Rombauts *et al.* 2003). CpG dimers (GC) were not

found at the TSS region in Arabidopsis and rice promoters indicating that plant promoters are not related to CpG islands (Yamamoto *et al.* 2007b). The three-dimensional conformation of a promoter region needed for gene expression might be related to DNA bending and curvature (Rombauts *et al.* 2003). Also the spacing and orientation between different or similar *cis*-acting elements, or the 'grammar' of regulatory sequences, is critical for the correct configuration of the transcription machinery to direct transcription and the balance between activators and quenchers or repressors contributes to a specific expression pattern (Halfon 2006). The distance between the TATA box and the initiator element is critical for transcription due to a proper interaction between the PIC which is anchored *via* the TATA box with TSSs. For instance, the rice phenylalanine-lyase promoter activity *in vitro* was eliminated when changing the distance between the TATA box and the initiator sequence (Zhu *et al.* 1995). Distance between the TATA box and the initiator sequence influence the accurate positioning and transcription initiation at the aforementioned *phas* promoter (Grace *et al.* 2004).

Different methods have been proposed for the identification of plant promoters, including thermodynamic calculations like the relative entropy values of candidate promoter sequences. Pandey and Krishnamachari (2006) analyzed several plant promoters in terms of sequence base composition, relative entropy, and periodicity and curvature properties. Distinguished DNA curvature was observed between promoter and non-promoter sequences and peaks of relative entropy values were detected in TATA and TATA-less promoters [0.35 and 0.15 bits (a bit is the unit of information content), respectively] at positions -30 and 0 relative to the TSS, respectively. In addition, several small entropy peaks (with relative values of 0.06 bits) were observed at different positions indicating that functional information is dispersed throughout the promoter sequences. Plant promoters are clearly different from promoters of yeast, *Escherichia coli* (Pandey and Krishnamachari 2006) and mammals (Yamamoto *et al.* 2007b). Promoter prediction programs like PromoterInspector (Scherf *et al.* 2000), Promoter 2.0 (Knudsen 1999), CorePromoter (Zhang 1998), Dragon Promoter Finder (Bajic *et al.* 2002), MacPromoter MM (Ohler *et al.* 1999, 2001) and NNPP (Reese 2001) that are developed for non-plant promoters were not able to produce satisfactory results for plant RNA Pol-II promoters (Pandey and Krishnamachari 2006). Hence, algorithms should be designed specifically for plant promoter prediction (Pandey and Krishnamachari 2006). Most of the promoter prediction programs can recognize around 50% of promoters with a false rate of 1 per 700-1000 bp when analyzing 33,120 base pairs (bp) containing 24 published mammalian promoters (Fickett and Hatzigeorgiou 1997). Pandey and Krishnamachari (2006) also obtained a relatively low frequency of false positives (<0.7%) by integrating in the prediction program other features like relative entropy. The newly discovered pyrimidine patches in plant promoters in combination with preferential sequences around the TATA-box (paragraph 1.2.2) and regulatory elements may elevate the accuracy of promoter

prediction (Yamamoto *et al.* 2007a). Finally, by restricting the search for promoters to known intergenic regions and avoiding analysis in coding sequences a further reduction of false positive predictions might be obtained (Shahmuradov *et al.* 2005; Pandey and Krishnamachari 2006).

Recently, a plant promoter prediction program was developed by constructing a library of known plant promoters. The TSSP plant promoter prediction program (www.softberry.com, Shahmuradov *et al.* 2003) is based on discriminate analysis of sequence features and plant regulatory motifs. The database comprises a total of 305 plant promoters of monocot and dicot species each with an experimentally reported TSS. The collected data correspond to sequences comprising the region -200 to +51 with the TSS in position +1. Candidate plant promoters from different species can be compared to this database to predict the presence of promoter sequences.

1.2.4 Reporter gene based analysis for promoter characterization

1.2.4.1 Advantages

Besides *in situ* studies *via* RT-PCR and Northern blot approaches, promoter activity during development or under specific stimuli can be assayed using reporter genes following transformation. The main advantage of the use of reporter genes is the analysis of gene expression in tissues from which isolation of transcripts is difficult (e.g. developing seeds), wherein transcript abundance is very low, or when promoter activity needs to be assayed throughout development avoiding time-consuming extraction of transcripts in different plant tissues. Furthermore, functional analysis of the *cis*-acting elements in the promoter requires fusion of deletion variants of the promoter to a reporter gene. Finally, to confirm that a specific activity pattern of the promoter is due to a specific *cis*-acting element, point mutations (base substitutions) need to be incorporated, again followed by fusion of the mutated promoter(s) to a reporter gene, back-transformation and reporter gene analysis.

In the next two paragraphs plant promoters responsive throughout development and to LT and the *cis*-acting elements involved in these responses are discussed since the aim of this work was to identify and characterize such promoters in banana.

1.2.4.2 Developmentally regulated promoters

Plant development involves tissue differentiation starting from embryogenesis and continuing throughout the life of a plant. Briefly, a zygote will develop into an embryo and consequently to a seedling after germination, which involves the creation of new plant structures. Development is regulated by genetically encoded programs involving differential and/or ubiquitous gene expression regulated by a multi-dimensional network system resulting in developmental change (Taiz and Zeiger 2002). Therefore, promoter

regions of differential expressed genes play an important role in dictating the rate and time of gene expression due to interactions with TFs (see paragraph 1.2.1 and 1.2.2). Many differentially expressed genes have been identified using Northern blot, RT-PCR and microarray based approaches. Several promoters of developmentally regulated genes have been tested in transgenic plants using reporter gene assays to confirm the gene expression pattern, but rarely promoter activity is followed in real-time and throughout *in vitro* plant development. Several developmentally regulated plant promoters have been discovered. Only few examples of promoters that were also characterized during embryo and/or plant development under *in vitro* or growth chamber conditions are described here. The activity of the *β*-glucosidase gene from maize (*ZmGLU1*) was increased during *in vitro* seed germination in tobacco, with a peak after 11 days as revealed by fluorometric GUS analysis (Gu *et al.* 2006). The highest level of GUS activity was observed in the roots. The mung bean (*Vigna radiata* L.) promoter which controls the *ACS1* (*VR-ACS1*), an auxin-inducible ACC synthase gene, showed constitutive though variable activity in several tissues and developmental stages of transgenic Arabidopsis and tobacco plants (Cazzonelli *et al.* 2005). The highest activity levels for the VR-ACS1 promoter as well for the CaMV35S promoter were measured in young and developing tissues of immature soil-grown tobacco plants but the levels were approximately 4 times higher with the former than with the latter promoter as shown by quantitative GUS and LUC assays. Fluorometric GUS assays revealed stronger VR-ACS1 promoter activity in leaves, petioles and stems than in roots in mature soil-grown tobacco plants similar to the trend observed in immature plants. In addition and surprisingly for a so-called constitutive promoter, a similar organ-specific trend was detected for the CaMV35S promoter but at a lower level of activity. The *phas* β-*phaseolin* promoter is highly active during embryogenesis and microsporogenesis but inactive during vegetative development in bean (Hall *et al.* 1999). The developmental regulation of this promoter was maintained in heterologous dicots plants like tobacco and Arabidopsis as assayed by measuring GUS activity (Chandrasekharan *et al.* 2003). The class IV chitinase bean (*PvChi*4) promoter, fused to the *uid*A reporter gene, was analyzed during greenhouse plant development in Arabidopsis and tobacco (Lima *et al.* 2002). The *PvChi*4 promoter is active in transgenic Arabidopsis and tobacco plants in the meristematic region and at the root tip. It is also highly active in the reproductive organs, in contrast with the activity measured in vegetative tissues in both species. During fruit development, promoter activity was detected in embryos of both species and in the septum in Arabidopsis, too. Transgenic Arabidopsis zygotic embryos were analyzed for GUS activity during development. At the globular stage no detectable GUS activity was observed. Promoter activity started at the heart-shaped stage in the radicle region. During the transition of heart-shaped to torpedo stage embryo GUS activity expanded until the hypocotyl and was maintained until embryo maturity when *PvChi*4 promoter activity was observed in the whole embryo with less staining intensity in the cotyledon region. During Arabidopsis seed germination, GUS

activity was detected in the root tips and the cotyledons, but was absent in the hypocotyl region unlike in mature embryos. To what extent promoter regulation in zygotic embryos is similar to that in somatic embryos, of much more importance in sterile plants such as banana, remains to be investigated.

Differential gene expression patterns under different developmental stages or stresses imply the interaction of TFs with *cis*-acting elements present in responsive promoters. Such elements then act as positive or negative regulators of gene expression. Identification of *cis*-acting elements involved in a specific developmental stage is difficult mainly due to the complex regulation of a developmental response. Activation of the common bean *phas* promoter during embryogenesis and its silent state during vegetative development in transgenic Arabidopsis lines is due to a combination of different *cis*-acting elements which altogether regulate the specific expression pattern during development (Chandrasekharan *et al.* 2003). Several elements are involved in the specific expression pattern during embryogenesis within the 295 nt of the proximal *phas* promoter (295-nt region upstream of the TSS) including RY motifs, CATGCA(C/T). One proximal RY motif (-70 to -64) allows expression in the hypocotyl of the embryo; while three other RY elements at specific positions (-277 to -271, -260 to -254, and -237 to -231) are necessary for radicle-specific expression. All the RY elements contribute to the expression in cotyledons but not in vascular tissue during embryogenesis. In addition, other elements are important for the level of activity and expression pattern including G-box (CACGTG), CCAAAT box, E-box (CACCTG), CACA box (AACACAT) and vicilin box (GCCACCTC). Multiple *cis*-acting elements driving specific promoter activity during embryogenesis reflects the existence of both redundancy and hierarchy in *cis*-acting element interactions (Chandrasekharan *et al.* 2003), and suggests that correct spacing of elements might contribute to the correct folding of the transcription machinery for developmental-specific expression. Therefore, different *cis*-acting elements might be combined in order to perform a specific developmental regulation according to TFs availability at specific developmental stage. The 416 nt upstream of the TSS in the aforementioned *PvChi*4 promoter contain the necessary *cis*-acting elements for the tissue specific expression in vegetative and reproductive organs (Lima *et al.* 2002). These include the seed specific enhancer [A(A/C)CCCA] at position -303 and five copies of the CANNTG element which is related to embryo specific expression (Lima *et al.* 2002). In addition, in the region between -2189 and -1445 relative to the TSS the *PvChi*4 promoter harbors a negative regulatory element because its deletion resulted in the activation of the promoter in the phloem and epidermal tissue of stems and leaves (Lima *et al.* 2002).

1.2.4.3 Low temperature responsive promoters

1.2.4.3.1 The well characterized low-temperature responsive promoter RD29A

The responsive to desiccation 29A (RD29A) promoter from Arabidopsis has been extensively studied under several abiotic stresses including LT (Yamaguchi-Shinozaki and Shinozaki 1994). Moreover, a *cis*-acting element which confers responsiveness to LT, drought and high salinity has been characterized in the RD29A promoter (paragraph 1.2.4.3.2) and therefore, this promoter is described in more detail. In addition, the promoters of other LT up-regulated genes are briefly described in paragraph 1.2.4.3.2. The RD29A promoter shows increased activity under LT stress, but also under dehydration, high salt and ABA treatments (Yamaguchi-Shinozaki and Shinozaki 1994). Induced LUC activity was observed in transgenic Arabidopsis plants containing the RD29A promoter fused to the *luc* gene after two days at 0°C, while LUC activity was absent in untreated samples (Ishitani *et al.* 1997). In addition, induction of LUC activity in these lines was also observed after treatments with NaCl (300 mM for 5 h) and ABA (100 µM for 3 h). The use of a RD29A::*luc* promoter::luciferase gene fusion in Arabidopsis ethyl methanesulfonate mutated plants led to the identification of upstream signaling components involved in cross-talk with related abiotic stresses such as dehydration in an ABA-dependent or -independent process (Ishitani *et al.* 1997). Mutants identified include *cos* (constitutive expression of osmotically responsive genes), *los* (low expression of osmotically responsive genes) and *hos* (high expression of osmotically responsive genes) (paragraph 1.2.4.3.2; Figure 1.1). A similar approach was employed by Chinnusamy *et al.* (2003) by screening for mutations that are impaired for cold-induction of a CBF3::*luc* (C-repeat binding factor 3 promoter::luciferase gene) fusion in transgenic Arabidopsis. This approach leads to the identification of the inducer of CBF expression 1 (ICE1) TF which anneals specifically to the MYC recognition sites in the CBF3 promoter. An overview of the regulation of genes during LT stress as unraveled in *Arabidopsis thaliana* is illustrated in Figure 1.1. Interaction between TFs and specific promoter elements allows the induction of specific genes upon LT stress leading to physiological changes to tolerate this stress.

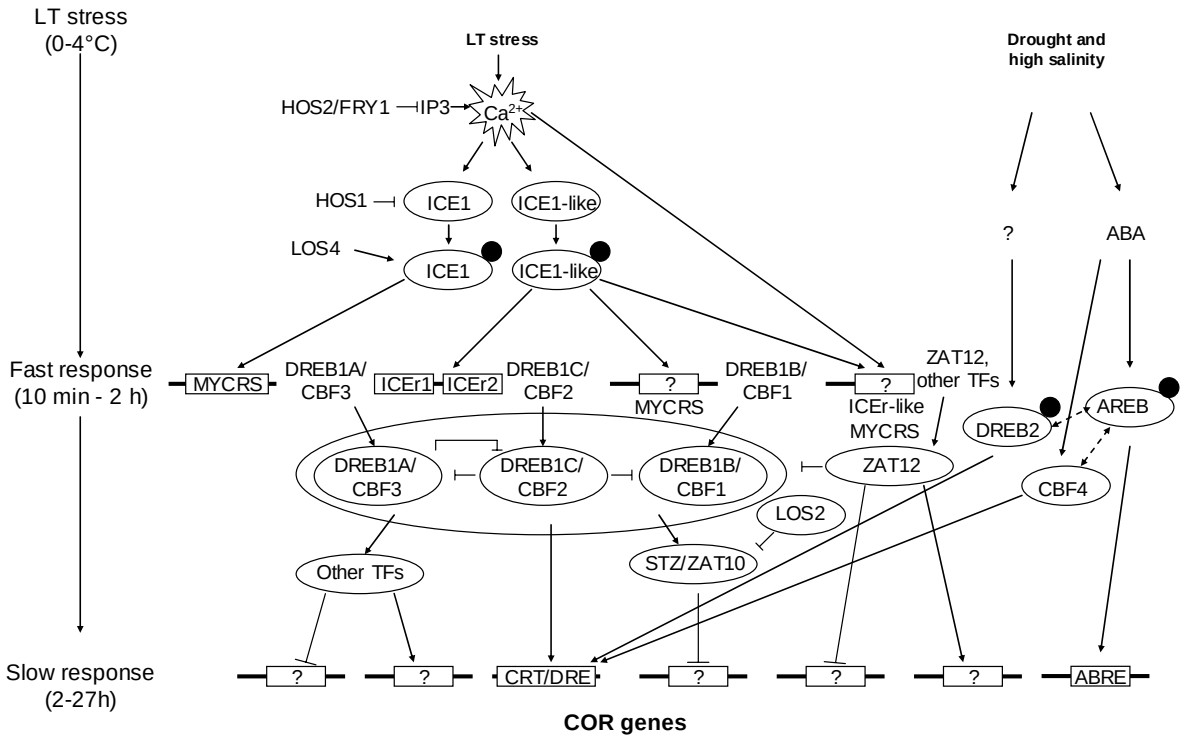

Figure 1.1. Regulatory network of gene expression under low temperature (LT, 0-4°C) stress in *Arabidopsis thaliana*. Boxes represent *cis*-acting elements involved in LT response. Transcription factors (TFs) that control LT gene expression are represented as ovals. Small closed circles indicate modification of the TFs for their activation. Interaction with drought and salinity stresses is depicted at the right. Lines with arrowhead indicate activation, whereas lines without arrowhead represent repression. Dotted lines indicate possible regulation. COR refers to cold regulated. Adapted from Heino and Palva (2003), Nakashima and Yamaguchi-Shinozaki (2006) and Chinnusamy *et al.* (2006).

1.2.4.3.2 Promoter elements involved in low temperature regulated gene expression in *Arabidopsis thaliana*

Promoters are characterized in more detail by creating deletion variants at their 5' and/or 3' end and by base substitutions to identify elements involved in a specific stress response. The dehydration responsive element (DRE) with the core region CCGAC confers responsiveness to LT (4°C for 10 h) but also to drought and high salinity but not to ABA (Yamaguchi-Shinozaki and Shinozaki 1994). An overview of *cis*-acting elements involved in LT responsive gene expression in *Arabidopsis thaliana* is shown in Table 1.1. The identification of the DRE element was performed by deletion analysis and base-substitution of promoter regions of the *RD29A/LTI178/COR78* gene from *Arabidopsis thaliana* (Yamaguchi-Shinozaki and Shinozaki 1994). Elements containing the core of the DRE sequence have been identified in cold-inducible promoters and are named C-repeat (CRT) and low-temperature-responsive element (LTRE), both containing an A/GCCGAC motif (reviewed by Yamaguchi-Shinozaki and Shinozaki 2005).

Table 1.1. *Cis*-acting regulatory elements involved in low temperature (LT) stress responsive gene expression in Arabidopsis

***Cis*-acting element**	**Sequence[a]**	**TF[b]**	**Promoter**	**Time course of response[c]**	**Reference**
MYCR	CANNTG	ICE1	CBF3/DREB1A	Early	Chinnusamy *et al.* 2003
ICEr1	GGACACATGTCAGA	Not known	CBF2/DREB1C	Early	Zarka *et al.* 2003
ICEr2	ACTCCG	Not known	CBF2/DREB1C	Early	Zarka *et al.* 2003
Novel[d]	CGCGT	Not known	Core sequence present in three novel elements enriched in promoters from genes induced in early cold response	Early	Vogel *et al.* 2005
DRE	TACCGACAT	CBF/DREB1	RD29A	Late	Yamaguchi-Shinozaki and Shinozaki 1994; Liu *et al.* 1998
CRT	GGCCGACAT	CBF/DREB1	Cor15A	Late	Baker *et al.* 1994; Stockinger *et al.* 1997
DRE-like	(A/G/T)(A/G)CCGACN(A/T)	CBF/DREB1	Overrepresented in promoters of genes induced in late cold response	Late	Chen *et al.* 2002
CRT/DRE	(A/G)CCGAC	CBF/DREB1	53% of promoters from genes induced in late cold response	Late	Vogel *et al.* 2005
Novel[d]	AGTNGGTCC	Not known	47% of promoters from genes induced in late cold response	Long term	Suzuki *et al.* 2005
Novel[d]	GATATNNT	Not known	74% of promoters from genes induced in late cold response	Long term	Suzuki *et al.* 2005
Novel[d]	(T/C)T(G/T)CCTCT	Not known	Overrepresented in promoters of genes induced in early and late response	Long term	Vogel *et al.* 2005
Novel[d]	(T/C)ANCTTCC	Not known	Overrepresented in promoters of genes induced in early and late cold response	Long term	Vogel *et al.* 2005

[a]Underlined indicates the core sequence.
[b]Transcription factor.
[c]Early, induction of promoter activity at 0-4°C starting after 10-30 min or 1.5 h; late, induction of promoter activity at 4°C starting after 2 to 27 h; long term, induction of promoter activity at 0.5, 1, 4, 8, 24 or 168 h at 4°C and which is maintained during 7 days of LT stress.
[d]Potential LT regulatory element.

By using yeast one-hybrid screening Stockinger *et al.* (1997) isolated an Arabidopsis gene encoding a DRE/CRT binding protein denominated CBF1 (Figure 1.1). The CBF homologous genes *CBF2* and *CBF3* were cloned from Arabidopsis and were called *DREB1C* and *DREB1A*, respectively (dehydration-responsive element binding protein 1C an 1A, respectively, Gilmour *et al.* 1998; Liu *et al.* 1998; Heino and Palva 2003), while the *CBF1* gene is identical to the *DREB1B* gene (Liu *et al.* 1998; Heino and Palva 2003). The *CBF/DREB1* genes are responsive to LT but not to drought stress (Liu *et al.* 1998; Shinwari *et al.* 1998; Gilmour *et al.* 1998). Overexpression of *CBF1/DREB1B*, *CBF3/DREB1A* or *CBF4*, a CBF homolog gene which is inducible by drought but not by LT (Haake *et al.* 2002), lead to constitutive expression of genes with promoters containing the DRE/CRT/LTRE element and to improved freezing and drought tolerance in non-acclimated plants (Jaglo-Ottosen *et al.* 1998; Kasuga *et al.* 1999; Haake et al. 2002). Hence,

identification of TFs that bind to DRE-like elements and enhanced expression of downstream genes upon LT stress allow the generation of plants with tolerance to several abiotic stresses (Kasuga *et al.* 2004; Pellegrineschi *et al.* 2004).

CBF/DREB1 genes are fast but transiently induced by LT (Liu *et al.* 1998; Gilmour *et al.* 1998; Medina *et al.* 1999; Chinnusamy *et al.* 2003). Induction was detected after 10 to 30 min at 0-4°C. The mRNA level reached a maximun after 1 h and declined to the basal level within 6 h (Medina *et al.* 1999). Chinnusamy *et al.* (2003) isolated the TF ICE1 which bind specifically to the MYC recognition sequences (CANNTG) in the CBF3 promoter (Figure 1.1). ICE1 binds to all five MYC elements present within 1 kilo base pairs (kb) of the CBF3 promoter, but with a higher affinity to the MYC-2 site (CACATG) than the other four MYC elements (Chinnusamy *et al.* 2003). Activity of the ICE1 promoter was detected in roots, stem, leaves and floral parts when fused to the *uid*A reporter gene. Semi-quantitative RT-PCR analysis also revealed that the *ICE1* gene was expressed constitutively, but expression was stronger in leaves and stems than in other tissues (Chinnusamy *et al.* 2003). RNA blot analysis indicated a slight up-regulation of the *ICE1* transcript by LT, ABA and NaCl but not by dehydration. These data suggest that activation of ICE1 is required to enhance expression of downstream genes during LT treatment. In the *ice1* mutant the cold induction of *CBF3* is blocked and consequently the expression of CBF3-target genes is reduced (Chinnusamy *et al.* 2003). The *ice1* mutation has a minor effect on the expression of the other *CBF* genes (*CBF1* and *CBF2*) with a slight reduction in gene expression at the early stages of the LT treatment but at later time points up to 6 h the expression was not reduced anymore (Chinnusamy *et al.* 2003). However, because the *CBF2* expression was enhanced after 6 and 12 h of LT treatment (0°C) in the *ice1* mutant, it is suggested that CBF3 may repress *CBF2* expression (Chinnusamy *et al.* 2003). Moreover, because of the minor but different effect of the *ice1* mutation on *CBF1* and *CBF2* expression, a different mechanism of regulation is suggested for CBF1 and CBF2 probably with ICE1-like TFs (Zarka *et al.* 2003; Novillo *et al.* 2004; Chinnusamy *et al.* 2006; Van Buskirk and Thomashow 2006; Figure 1.1). Analysis of the CBF2 promoter revealed two *cis*-acting elements, ICEr1 and ICEr2, which in combination impart a robust cold response in *Arabidopsis thaliana* (Zarka *et al.* 2003, Figure 1.1). The ICEr1 element (GGACACATGTCAGA) contains the core sequence CACATG fitting the MYC-like element (CANNTG) although the expression of the *CBF2* gene is not strictly regulated by ICE1. Hence, the TF that binds to the ICEr1 element remains to be identified (Yamaguchi-Shinozaki and Shinozaki 2005). In addition, no known transcription factor-binding site was detected in the ICEr2 element (ACTCCG, Zarka *et al.* 2003).

Most but not all LT responsive genes are induced by exogenous ABA (Heino and Palva 2003) suggesting that both ABA-dependent and ABA-independent regulatory pathways are involved in LT responsive gene expression. Large scale transcriptome analysis revealed different LT responsive patterns in thousands of genes in Arabidopsis

(Chen *et al.* 2002). Furthermore, in promoter regions of genes with a specific LT response pattern certain *cis*-acting elements are overrepresented. ABRE-like [(C/G/T)ACGTG(G/T)(A/C)] and DRE-like [(A/G/T)(A/G)CCGACN(A/T)] elements are significantly overrepresented in 57 promoters of genes that were showing a late LT response (27 h at 4°C). The promoters of the cold regulated (COR) genes *Cor15b*, *Cor47*, *Cor78*, *RD29a*/*RD29b*, *Iti29* and *Kin1* were among them (Chen *et al.* 2002). Indeed, DRE-like elements are not enriched in promoters of genes induced at an early LT response (3 h at 4°C, Figure 1.1). The proteins that bind to the ABRE- and DRE-like elements belong to the bZIP and AP2/EREBP TF family and these types of TF genes were LT responsive within 3 h at 4°C (Chen *et al.* 2002). Apparently, this early LT-responsive TF genes participate in the regulation of late LT-responsive genes. Interaction between ABRE with DRE-like elements was observed in the RD29A promoter (Narusaka *et al.* 2003). One of the four DRE-like elements present in the RD29A promoter functions as a coupling element for the ABRE element (Narusaka *et al.* 2003; Yamaguchi-Shinozaki and Shinozaki 2005). Together they constitute an ABA responsive complex in ABA-inducible gene expression indicating the presence of cross-talk between DRE and ABRE in stress-inducible promoters (Yamaguchi-Shinozaki and Shinozaki 2005). A single copy of ABRE is not sufficient for an ABA-responsive promoter activity and the presence of at least one coupling element is therefore needed (reviewed by Yamaguchi-Shinozaki and Shinozaki 2005). Indeed, although the CBF signaling pathway leading to activation of CRT/DRE has been described as an ABA-independent route in LT-regulated gene expression (Liu *et al.* 1998), additional evidence has been presented that the CRT/DRE element can also be activated by ABA (Knight *et al.* 2004). The data show that ABRE- and DRE-like elements are two important elements involved in transcriptional regulations of late LT responsive genes (Chen *et al.* 2002). TATA boxes are not overrepresented in promoters of late LT responsive genes indicating the presence of TATA and TATA-less promoters responsive to LT stress (Chen *et al.* 2002).

Another *cis*-acting element involved in induced expression at LT was identified in barley. The LT responsive (6°C day/2°C night, 10-h day for 24 h) barley *blt*101.1 promoter contains a 10-bp motif (AAGAAGATGC) identified as a LT responsive element (Brown *et al.* 2001). The 10-bp motif is one base different, but in reverse complement to the enhancer-like TCA element (TCATCTTCTT) in a barley salicylic acid-responsive gene (Goldsbrough *et al.* 1993). In addition, the hexanucleotide CCGAAA which is proposed to be involved in the LT response (6°C day/2°C night, 10-h day for 5-7 days) of the barley *blt*4.9 promoter resembles the core sequence of the DRE element (Dunn *et al.* 1998).

Transcription factor genes identified as being up-regulated within 1 h at LT (Fowler and Thomashow 2002) were analyzed for the presence of ICEr-like elements in their promoter regions (Vogel *et al.* 2005). These genes were *ZAT12*, *ZAT10/STZ* (Figure 1.1), *MYB73*, *CZF1* and *CZF2*. The promoters of the genes analyzed contained four sequences

that overlap with ICEr1 and one sequence that overlaps with ICEr2. In addition, the bioinformatic MotifSampler tool (Thijs *et al.* 2002) was employed to search for overrepresented 8-bp sequences present in LT-responsive Arabidopsis genes (Vogel *et al.* 2005). These genes were identified through microarray analysis after 1, 24 and 168 h at 4°C and seven different LT responsive patterns were distinguished including two cold down-regulated patterns (Vogel *et al.* 2005). No potential *cis*-acting elements were identified in the promoters of LT down-regulated genes. In contrast, the promoters (1 kb upstream of genes) of 123 (53%) out of the 233 genes in cluster III, which are LT up-regulated with the highest expression level after 24 h of LT, contained the CRT/DRE element (A/GCCGAC, Vogel *et al.* 2005). Novel overrepresented *cis*-acting elements with core sequence CGCGT were discovered in promoters of cluster V genes, which are LT inducible with the highest expression level within 1 h after LT stress (Vogel *et al.* 2005). In addition, other potential *cis*-acting elements involved in induction of early and maintained in late LT response were identified with the highest overrepresented elements being (T/C)T(G/T)CCTCT and (T/C)ANCTTCC (Vogel *et al.* 2005). Additionally, Suzuki *et al.* (2005) detected overrepresented motifs in the promoters of 57 genes induced at LT in a long term response (induced expression either at 0.5, 1, 4, 8, 24 or 168 h at 4°C and maintained during 7 days of LT stress) including AGTNGGTCC (present in 47% out of 57 promoters) and GATATNNT (present in 74% out of 57 promoters and in at least two copies for 19 promoters). The aforementioned *cis*-acting elements involved in LT responsive gene expression in *Arabidopsis thaliana* are also shown in Table 1.1.

The presence of DRE- or ABRE-like sequences alone is not sufficient to obtain LT induction of certain promoters. The presence of negative regulators may also affect gene expression induction by LT. A fructan 1-exohydrolase promoter (*FEH*IIa) showed high activity in LT-stored roots and in etiolated leaves in *Cichorium intybus* (Michiels *et al.* 2004). The region -278 to -172 (relative to the ATG translation start site) in the *FEH*IIa promoter contains an ABRE element (TACGTGTC) and a sequence similar to the CRT/DRE element (TGGCCAACTT) that might be involved in the LT induced gene expression (4°C). LT also enhanced gene expression in green leaves, but reporter gene activity was much lower than in etiolated leaves in similar induction experiments. Further deletion analysis revealed the presence of negatively regulating *cis*-acting elements in the promoter regions -933 to -717 and -493 to -278 which decreased the level of LT induction as well as the strength of the promoter at 25°C (Michiels *et al.* 2004). Also the down-regulating *cis*-acting element S1F (G/ATTACCATG) involved in the regulation of nuclear *RPS*1 and *RPL*21 genes encoding the plastid ribosomal proteins S1 and L21, respectively (Zhou *et al.* 1992), is twice present in the positions -789 and -860 in the *FEH*IIa promoter. However, further analyses are necessary to confirm that the S1F box is responsible for the decreased LT induction or the relatively low level of expression at 25°C.

Several other genes are involved in the plant response to LT stress. The Arabidopsis mutant *hos1* (Ishitani *et al.* 1998) showed an enhanced induction of the *CBF2* and *CBF3* genes (Lee *et al.* 2001) and consequently, target genes of CBF were superinduced in the *hos1* mutant. In the wildtype plant the HOS1 protein accumulates in the nucleus upon LT stress suggesting that it negatively regulates the turnover of the ICE proteins which are the CBF inducers (Figure 1.1). The HOS1 protein is a novel RING finger protein containing E3 ubiquitin ligase activity by which the RING finger proteins transfer ubiquitin to proteins directing them to degradation (Gilmour *et al.* 1998; Thomashow 2001; Lee *et al.* 2001). Overexpression analysis of a C_2H_2 type zinc finger protein STZ/ZAT10 in Arabidopsis revealed a repression of the RD29A promoter (Lee *et al.* 2002). Furthermore, another Arabidopsis mutant, *los2*, showed a decreased LT induction of the RD29A promoter (Lee *et al.* 2002). The *LOS2* gene encodes a bi-functional protein with enolase activity and binds to the promoter of the *STZ/ZAT10* gene. The *STZ/ZAT10* expression was enhanced and prolonged in response to LT in the *los2* mutant suggesting that LOS2 is a negative regulator of the *STZ/ZAT10* gene (Lee *et al.* 2002, Figure 1.1). The LT inducible gene *ZAT12* encodes a TF which is also a negative regulator of the CBF cold responsive pathways and also participates in the induction and repression of cold-responsive genes (Vogel *et al.* 2005, Figure 1.1). In addition, the CBF2 TF acts as a negative regulator of the CBF1 and CBF3 TFs (Novillo *et al.* 2004, Figure 1.1).

Several other genes involved in the LT response pathways were isolated using mutants. Genes idenfied so far include *LOS1*, needed for the enhanced activity of the RD29A promoter (Guo *et al.* 2002), FRY2, a negative regulator for the CRT/DREB genes (Xiong *et al.* 2002), FRY1, a regulator of inositol triphosphate (IP3) which is involved in the regulation of Ca^{2+} channels in response to LT (Knight *et al.* 1996; Xiong *et al.* 2001; Sangwan *et al.* 2001, Figure 1.1) and LOS4, a positive regulator for the CBF genes (Gong *et al.* 2002) among others (reviewed by Heino and Palva 2003). Genes induced by LT are often also inducible by drought and high salinity suggesting the presence of a cross-talk between these stress-signaling pathways (Yamaguchi-Shinozaki and Shinozaki 2005, Figure 1.1).

1.3. T-DNA tagging to characterize developmental and low temperature regulated genes in plants

1.3.1 T-DNA promoter tagging: an insertional mutagenesis approach

Classical genetic approaches for gene identification rely on the disruption of a gene by insertional mutagenesis leading to a recognizable phenotype (Springer 2000). Gene inactivation is accomplished using a transposon (Sundaresan *et al.* 1995; Martienssen 1998) or T-DNA of *A. tumefaciens* (Feldmann 1991; Krysan *et al.* 1999). The resulting transgenic line should show a clear phenotype indicative of the disruption of an expressed gene. Once the desired phenotype is identified, the presence of the tag (inserted DNA fragment) greatly facilitates isolation of the affected gene in the unknown region starting from the known region or tag sequence (reviewed by Springer 2000). Insertional mutagenesis with transposable elements offers some advantages over T-DNA elements. Thanks to the remobilization of transposable elements germinal revertants can be obtained to verify that a phenotype is indeed caused by the inserted transposon (Springer 2000). This type of insertional mutagenesis is also useful for plants that are not highly competent for genetic transformation, since new transposon insertions are generated through crossing or propagation (Ramachandran and Sundaresan 2001). However, crossing in some plant species like banana is excluded and/or very time consuming because of sterility and a long life cycle.

Insertion of the transposon or T-DNA should be preferably as a single copy to ensure that the disruption of the identified gene is responsible for the phenotype obtained. However, insertional mutagenesis often leads to transgenic lines containing multiple copies of the inserted DNA, which complicates the interpretation of the phenotype or the expression pattern and the further identification of the authentic tagged sequence (Springer 2000). Nevertheless, insertional mutagenesis has become an important functional analysis tool in plants of which the whole genome has been sequenced (*Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*). T-DNA integrates randomly and therefore genome-wide saturation with a T-DNA insertion in most of the annotated genes has been achieved for model plants with highly transformation competence and a small genome size. Large T-DNA collections have been generated in Arabidopsis (Feldmann and Marks 1987; Bouchez *et al.* 1993; Campisi *et al.* 1999; Krysan *et al.* 1999; Weigel *et al.* 2000; Rios *et al.* 2002) and rice (Jeong *et al.* 2002, 2006; Chen *et al.* 2003). However, not all genes can be uncovered by classical insertional mutagenesis. For instance, redundant genes do not show a recognizable phenotype (Springer 2000).

Table 1.2. Plant T-DNA promoter tagging in different plant species

Selectable marker/ reporter gene	Plant species	Frequency of reporter gene activation (%)[a]	Stress/condition applied or tissue/development tested[b]	Reference
neo	*Nicotiana tabacum*	~0.1-5[c]	Kanamycin, calli, leaves, stems, roots	Teeri *et al.* 1986
neo	*Nicotiana plumbaginifolia*	~5-19[c]	Kanamycin, calli, leaves, stems, roots, flowers, leaves[d]	André *et al.* 1986
neo	*Arabidopsis thaliana*	28.5[e]	T_2 progeny calli, leaves, stems, roots	Koncz *et al.* 1989
neo	*Nicotiana* spp.	24.2[e]	T_2 progeny calli, leaves, stems, roots	Koncz *et al.* 1989
neo	*Nicotiana tabacum*	35[f]	Calli	Herman *et al.* 1990[g]
*uid*A	*Arabidopsis thaliana*	54[e]	Leaves, stems, roots, flowers	Kertbundit *et al.* 1991
*uid*A	*Nicotiana tabacum*	78	Leaves, roots	Topping *et al.* 1991
*uid*A	*Nicotiana tabacum*	5	Leaves, progeny seedlings	Fobert *et al.* 1991
*uid*A	*Arabidopsis thaliana*	16-30[h]	Leaves, roots, flowers	Lindsey *et al.* 1993
*uid*A	*Nicotiana tabacum*	22-92[h]	Leaves, roots, flowers, seed germination, seedling growth, mature plant, wound	Lindsey *et al.* 1993
*uid*A	*Solanum tuberosum*	9-25[h]	Leaves, roots, stems, tubers	Lindsey *et al.* 1993
*uid*A	*Arabidopsis thaliana*	ND	Developing nematode feeding structures, roots, *Heterodera schachtii* infection	Goddijn *et al.* 1993
*uid*A	*Arabidopsis thaliana*	17.2	T_1 siliques, T_2 siliques, T_2 seed development, T_3 seedlings	Topping *et al.* 1994
*uid*A	*Arabidopsis thaliana*	6	Light, dark, different organs and tissues	Mollier *et al.* 1995
*uid*A	*Arabidopsis thaliana*	15.4, 17.0	T_2 progeny leaves, low temperature (4°C, 0-24 h), ABA 0-24 h	Mandal *et al.* 1995
*uid*A, *uid*A^{INT}	*Arabidopsis thaliana*	7	Beet cyst nematode infection in roots, infection with other nematodes	Barthels *et al.* 1997
*uid*A	*Arabidopsis thaliana*	6	T_2 leaves, roots, flowers, siliques	Richardson *et al.* 1998
luc	*Nicotiana tabacum*	3.3	Regenerated shoots, greenhouse plants, F1 and F2 progeny of tagged line	Mudge and Birch 1998
*uid*A	*Arabidopsis thaliana*	7.9	*Meloidogyne incognita* infection in T_3 roots, other tissues from T_3 plants, infection with other nematodes	Favery *et al.* 1998
*uid*A	*Lotus japonicus*	9-28[h]	Roots, infection with *Mesorhizobium loti*	Martirani *et al.* 1999
*uid*A	*Arabidopsis thaliana*	1.1-10	Roots, aerial vegetative plant parts	Karimi *et al.* 1999
*uid*A	*Arabidopsis thaliana*	0.2	Guard-cells	Plesch *et al.* 2000
*uid*A	*Lotus japonicus*	3	Roots, nodules (Mesorhizobium loti infection), flowers	Webb *et al.* 2000

Table 1.2. (Continued)

Selectable marker/ reporter gene	Plant species	Frequency of reporter gene activation (%)[a]	Stress/condition applied or tissue/development tested[b]	Reference
*uid*A	*Oryza sativa*	1.6-2.1[h]	Primary transgenic plant leaves, roots, mature flowers, developing seeds	Jeon *et al.* 2000
*uid*A	*Arabidopsis thaliana*	ND	Roots, seedlings	Eastmond *et al.* 2000
*uid*A^{INT}	*Arabidopsis thaliana*	0.2	Leaves infected with *Botrytis cinerea*, seedlings, adult plants, wounding	Custers *et al.* 2002
*uid*A	*Medicago truncatula*	10.1	Roots, nodules induced by *Sinorhizobium meliloti*, different organs and tissues	Scholte *et al.* 2002
*uid*A::*neo*	*Brassica napus*	2.6-3.8[h, i]	Kanamycin, leaves, stems, roots, re-induced calli, node segments propagated on NAA containing medium	Bade *et al.* 2003
luc	*Arabidopsis thaliana*	3.7	T_1 seedlings, high sugar (400 mM Glc), salt (250 mM NaCl), cold (4°C, 6-8 h), mannitol, ABA, H_20_2, $CdCl_2$, 2,4-D, SA, paraquat, GA, T_2 seedlings	Alvarado *et al.* 2004
*uid*A, *uid*A^{INT}	*Arabidopsis thaliana*	15	Development of T_2 seeds, T_3 seedlings, mature plants	Resminath *et al.* 2005
luc, *luc*$^+$, *luc*$^{+INT}$	*Musa* spp.	0.06-3[h]	Embryogenic cell cultures, cell colonies, differentiated shoots, plants, bisected meristems, leaf discs	Remy *et al.* 2005
*uid*A^{INT}	*Arabidopsis thaliana*	ND	T_2 seedlings, T_3 plants	Prasad *et al.* 2005
*uid*A	*Lotus japonicus*	2.1-5.2[h]	Roots, nodules (*Mesorhizobium loti* infection)	Buzas *et al.* 2005
*uid*A	*Arabidopsis thaliana*	ND	Seeds, inflorescence, seedlings, leaves, siliques	Stangeland *et al.* 2005
AMV-*uid*A::*neo*	*Arachis hypogaea*	ND	T_0 plant leaves, roots, shoot parts, floral organs, T_1 plants	Anuradha *et al.* 2006

neo, neomycin phosphotransferase II; *uid*A, *β*-glucuronidase; *uid*A^{INT}, *uid*A containing an intron; AMV, Alfalfa mosaic virus translational enhancer; *luc*, firefly luciferase; *luc*$^+$, codon-optimized *luc*; *luc*$^{+INT}$, *luc*$^+$ containing an intron; ND, not determined; ABA, abscicic acid; NAA, naphthaleneacetic acid; Glc, glucose; 2,4-D, 2,4-dichlorophenoxyacetic acid; SA, salicylic acid; GA, gibberellic acid.

[a]The number of independent tagged lines showing reporter gene activation expressed in percentage of the total number of independent tagged lines screened.

[b]Main stress/condition applied or tissue/development tested used for the calculation of the frequency of selectable marker or reporter gene activation is underlined. Non-underlined stresses/conditions or tissues/developmental stages reflects additional analyses performed after the first screening.

[c]Estimated activation frequency calculated as the number of kanamycin-resistant calli of lines transformed with a construct containing a promoterless *neo* gene expressed in percentage of the total number of kanamycin-resistant calli of lines transformed with the *neo* driven by the nopaline synthase promoter.

[d]Leaf fragments were incubated in callus induction medium supplied with kanamycin.

[e]Transcriptional gene fusion vector employed.

[f]Minimun frequency obtained.

[g]Principal aim of the work was the study of T-DNA integration events.

[h]Range of activation frequency obtained.

[i]Relative activation frequency calculated as the number of kanamycin-resistant calli expressed in percentage over the number of hygromycin-resistant calli.

The isolation of promoters in plants using T-DNA tagging was first reported by Teeri *et al.* (1986) and André *et al.* (1986) as an alternative technique to discover new regulatory sequences instead of mRNA based methods (Table 1.2). The technique encompassed the creation of *in vivo* transcriptional fusions between plant promoters and a promoterless selectable marker which was placed near a T-DNA border. Because integration of T-DNA is random, only a portion of the transgenic lines screened will show activity of the selectable marker gene. Therefore, the frequency of marker gene activation is defined by the number of independent tagged lines showing marker gene activation expressed in percentage of the total number of independent tagged lines screened (Table 1.2). Important improvements of the method were the application of a second selectable marker gene for the selection of transgenic lines (Koncz *et al.* 1989) and the replacement of the promoterless selectable marker by a promoterless reporter gene (Kertbundit *et al.* 1991; Fobert *et al.* 1991; Table 1.2).

The second improvement enabled the isolation of promoters not only active during the selection period but also afterwards and in different tissues. Since activity of the reporter gene is monitored, tagged promoters are identified based on reporter gene expression. Hence, a mutant phenotype is not required (Springer 2000). An overview of T-DNA promoter tagging experiments performed in different plant species is presented in Table 1.2. Most of the promoter tagging experiments were performed in Arabidopsis. The T-DNA activation or tagging frequency varied greatly from 0.06% (Remy *et al.* 2005) and up to a 78% (Topping *et al.* 1991). The difference in activation frequency might be a result of different factors including explant/plant species used for transformation, stress/condition or tissue/development screened for selectable/reporter gene activity determination, selectable/reporter gene employed, and even the T-DNA plasmid/*A. tumefaciens* strain used.

Several modifications in T-DNA tagging constructs exist that target specific sequences like genes, promoters or enhancer elements (reviewed by Springer 2000). Promoter tagging constructs sometimes contain one or multiple stop codons in-frame upstream of the start codon of the reporter gene (Koncz *et al.* 1989; Mudge and Birch 1998), which allows transcriptional but not translational fusions. Gene tagging constructs may contain from one or more splice acceptor or donor sequences preceding the reporter gene allowing expression if the integration occurs in introns (Sundaresan *et al.* 1995; Springer 2000; Calderon-Villalobos *et al.* 2006). In addition, gene tagging constructs that lack the translational start codon of the reporter gene enrich for translational fusions with the tagged gene (Koncz *et al.* 1989). Only promoter tagging and not enhancer/gene tagging reports are shown in Table 1.2.

T-DNA tagging for the identification of promoters and genes overcomes some of the bottlenecks present in other promoter/gene discovery methods. Differential screening of cDNA libraries appears to preferentially isolate genes encoding relatively abundant

transcripts (Lindsey *et al.* 1993; Mollier *et al.* 2000). Furthermore, promoter recovery becomes complex if the cDNA is a member of a multigene family or if homologous pseudogenes are present (Mudge and Birch 1998). T-DNA mediated plant promoter tagging avoids these limitations, and has the potential to facilitate the isolation of a more diverse range of promoter types (Mudge and Birch 1998).

The main advantage of the T-DNA tagging approach is the characterization of promoter/gene activity prior to the knowledge of the exact DNA sequence. To characterize promoter activity, a reporter gene is preferred over a selectable marker gene in T-DNA tagging (Table 1.2) for ease and real-time detection of activity in tissues, different developmental stages or under specific stimuli. Different reporter genes have been used for promoter (Table 1.2) and gene tagging in plants such as the *β*-glucuronidase (*uid*A; *gus*) of *E. coli* (Jefferson *et al.* 1987), the green fluorescent protein (*gfp*; GFP) of the Pacific jellyfish *Aequoria victoria* (Heim *et al.* 1995; Chiu *et al.* 1996) and the firefly luciferase (*luc*) of the American firefly *Photinus pyralis* (Ow *et al.* 1986). The choice of the reporter gene for T-DNA promoter tagging depends on several factors but mainly on the specific experimental objectives or the type of promoters that is aimed at.

1.3.2 Reporter genes in T-DNA tagging

Most T-DNA promoter tagging reports rely on the use of the *uid*A reporter gene (Table 1.2) and spatial analysis of gene expression is commonly analyzed using *uid*A. Due to the long half-life of the GUS reporter gene system (~50 h in living mesophyll protoplasts, Jefferson *et al.* 1987), GUS activity assays should be carefully analyzed especially when executed during development or when different biotic or abiotic stress treatments have been applied. The importance of the half-life of the reporter gene system is exemplified in expression analysis of circadian-mediated gene expression. By using the luciferase (*luc*) gene, the cyclic pattern of gene expression was detected, whereas with the chloramphenicol acetyltransferase (*cat*) gene an almost constant signal was observed (Wood 1995) due to the half-life difference between the LUC and the CAT in mammalian cells (3 h and 50 h, respectively; Wood 1995; Thompson *et al.* 1991). Due to the short half-life of the LUC activity a circadian-regulated promoter (*cab*2) was also verified in Arabidopsis (Millar *et al.* 1992). Comparison of the GUS activity pattern with a mRNA based analysis of the *uid*A transcript is needed to confirm fast changes in gene expression. Interpretation of down-regulated gene expression using the *uid*A reporter gene is also problematic unless analysis of GUS activity can be performed at a time point where pre-treatment GUS enzyme is no longer available, for instance, after 5 days (Castle *et al.* 2005). Similarly, kinetic analyses of induction of GUS activity require sampling at several time points within a period of time. Such sampling is also necessary because GUS assays are detrimental (Mandal *et al.* 1995).

In summary, where change of gene expression is expected, the *uid*A reporter gene is not the best choice.

Alternatively, the *gfp* reporter gene system does not require toxic assays and avoids the use of exogenous substrates. Thus, detection of gene expression profiles in tissues and under specific stimuli can be accomplished using GFP (Chiu *et al.* 1996). However, few groups have reported the use of the GFP system for tagging genes and non for promoters (Table 1.2). Ryu *et al.* (2004) designed a dual T-DNA tagging system where the promoterless *uid*A and *gfp* reporter genes were located near the right and left T-DNA border, respectively. However, background fluorescence has been detected in plant tissues (Hraska *et al.* 2006) including in 2 days-old Arabidopsis seedlings (Castle *et al.* 2005) or in undifferentiated meristematic banana tissue (S. Remy, personal communication). To a certain extent this fluorescence background can be diminished by using appropriate filters (Maximova *et al.* 1998; Elliott et al. 1999; Hraska et al. 2006).

The most sensitive reporter gene that allows real-time *in planta* gene expression is the *luc* gene. The detection of LUC protein is, moreover, non-invasive and non-destructive (Ow *et al.* 1986). Together with the short half-life of LUC activity (~15.3 min after luciferin is applied, Van Leeuwen *et al.* 2000) these characteristics make the *luc* reporter gene ideal for tagging promoters and genes with inducible or developmentally regulated expression. On the other hand, the substrate luciferin has to be added to the samples to capture the emitted photons with sophisticated equipment such as a charge-couple device (CCD) digital camera (Remy *et al.* 2004). Fortunately, luciferin passes easily through the semi-permeable plant cell walls and membranes allowing direct interaction with the LUC enzyme (Van Leeuwen *et al.* 2000).

1.3.3 T-DNA tagging of developmentally and low temperature regulated genes and promoters

Multiple developmentally regulated promoters have been identified using T-DNA tagging since it requires simple screening for reporter gene activity during different developmental stages of the plant life-cycle (Table 1.2). Promoters tagged by Topping *et al.* (1994) showed different GUS activity patterns during seed development and in certain tissues (flowers and roots) of mature Arabidopsis plants indicating their developmental regulation. Similarly, a total of 133 rice T-DNA tagged lines showing GUS activity in the flower were analyzed for GUS activity during seed development 5 to 10 days after flowering which yielded lines with seed tissue specific GUS activity (Jeon *et al.* 2000). In a T-DNA tagged tobacco line monitored during seed germination, seedling growth and at maturity, GUS activity was tissue as well as developmentally regulated (Lindsey *et al.* 1993). For instance, in one of the screened lines no GUS activity was detected in the mature embryo, while during seed germination GUS activity was observed in the testa at the site of radicle

emergence (Lindsey *et al.* 1993). The gene *EXORDIUM* (*EXO*) was identified *via* T-DNA tagging using a promoterless *uid*A reporter gene in Arabidopsis (Farrar *et al.* 2003). The tagged promoter is active in embryos, shoots, root tips and young leaves. In addition, the *EXO* promoter is responsive to exogenous auxin and cytokinin, and is active in relatively young root vascular tissues and tissues of the silique other than the embryo again demonstrating a correlation between tissue specific and developmental promoter/gene regulation. A lysine decarboxylase (LDC) promoter, originally identified by T-DNA tagging, showed activity in the vascular tissue of flower petals and stamen filaments and in the vascular tissue and veins of rosette leaves in transgenic Arabidopsis plants carrying 1509 bp of the promoter region and 5' UTR of the LDC gene fused to the *uid*A reporter gene (Stangeland *et al.* 2005). Additionally, promoter activity was observed in different tissues during seed development. Lee *et al.* (2004b) tagged a gene which is expressed at a late anther developmental stage in Arabidopsis. GUS activity screening revealed expression in the anther locus but not in the vascular bundles or the connective tissues.

Developmentally regulated promoters were tagged using a *uid*A::*neo* promoterless fusion gene in *Brassica napus* (Bade *et al.* 2003). Selection of promoter-tagged lines in transgenic calli was performed on kanamycin selective medium ensuring that tagged promoters were active at the callus stage. Further GUS activity assays in the regenerated plants revealed different weak GUS activity patterns in different tissues, but often the activity was restricted to the vascular tissue of stem and leaves. GUS activity under greenhouse conditions was even weaker than in *in vitro* plants suggesting that the up-regulation of tagged promoters in undifferentiated tissue was due to the presence of one or more tissue culture components. Most of the promoters were found to be slightly active in reinitiated callus tissue from the *in vitro* plants and six out of 20 were found to be inducible by the auxin NAA. Transcriptome analysis revealed that even fresh medium alone induces significant expression changes in Arabidopsis (Kreps *et al.* 2002). Thus, selection or screening for tagged lines in undifferentiated tissue and/or under specific conditions can lead to the isolation of promoters that are developmentally regulated and/or inducible only under these specific conditions.

Relatively few research groups have exploited the LUC reporter system for the tagging and characterization of promoters during development or inducible by specific stimuli. Only recently were two gene-trap vectors containing the wild type *luc* gene constructed and successfully tested in the model plant *Arabidopsis thaliana* for identification of genes activated during 7 days of seedling development and by light (Yamamoto *et al.* 2003). Using a large-scale *in vivo* LUC screening system tagging of cold (6 to 8 h at 4°C)-responsive promoters has been reported in *Arabidopsis* seedlings (Alvarado *et al.* 2004), but no data were presented on the level of induction or repression and whether seedlings were monitored during or only after the cold treatment. Developmental regulation of these responses was not studied either. The rice Δ^1-pyrroline-

5-carboxylate synthetase 2 (*OsP5CS2*) gene was tagged in rice. GUS activity was undetectable but appeared under salt (250 mM NaCl) and LT (4°C) stress and in stamens of mature spikelets (Hur *et al.* 2004). Additionally, the *OsP5CS2* gene was inducible by ABA treatment (0.5 μM), water stress, salt and LT. In relation to tagging temperature-responsive promoter/genes, Mandal *et al.* (1995) first reported the identification of one (out of 1200 lines tested) tagged *Arabidopsis* line showing induced GUS activity after a 16 h treatment at 4°C. Screening for tagged cold-responsive genes was recently also performed in rice by subjecting plant samples for a period of time to low temperature before measuring GUS activity at room temperature (Lee *et al.* 2004a).

1.4 Promoters used in transgenic bananas

Genetic transformation of banana has been an important goal for different laboratories in the last 16 years. Transformation of monocot plant species with *A. tumefaciens* was for a long period of time thought to be impossible. However, research on the suitability of *Agrobacterium*-mediated genetic transformation for banana resulted in the successful generation of transgenic plants (May *et al.* 1995; Pérez Hernández *et al.* 1998, 2000, 2006a, 2006b; Ganapathi *et al.* 2001; Khanna *et al.* 2004; Arinaitwe 2008). Furthermore, genetic modified 'Gros Michel' banana plants, which were generated through *Agrobacterium*-mediated transformation, are being field-tested for their resistance to Black leaf streak in Uganda[1]. Other methods for genetic transformation of banana include particle bombardment (PB; Sági *et al.* 1995a, 1995b, 1995c; Becker *et al.* 2000) and electroporation of banana protoplasts (Sági *et al.* 1992, 1995b, 1995c).

Different banana varieties have been used for genetic transformation. The first reports of genetic transformation dealt with the cultivar 'Bluggoe' (genomic group ABB) (Sági *et al.* 1992, 1994, 1995a, 1995b, 1995c; Dugdale *et al.* 1998; Remy *et al.* 1998a) mainly because the first regenerable banana ECS starting from proliferating meristematic buds was developed for this cultivar (Dhed'a *et al.* 1991). In the following reports several other cultivars have been used for transformation including 'Three Hand Planty' (AAB, Sági *et al.* 1995a; Pérez Hernández *et al.* 1998, 2006a, 2006b; Remy *et al.* 1998a, 1998b, 2004, 2005; Schenk *et al.* 1999, 2001; Arinaitwe *et al.* 2004; Santos *et al.* 2007; Arinaitwe 2008), 'Grand naine' (AAA, May *et al.* 1995; Remy *et al.* 1998a; Schenk *et al.* 1999; Bosque-Pérez *et al.* 2000; Becker *et al.* 2000; Yang *et al.* 2003; Arinaitwe *et al.* 2004; Atkinson *et al.* 2004; Khanna *et al.* 2004; Acereto-Escoffie *et al.* 2005; Arinaitwe 2008) and Williams (AAA, Sági *et al.* 1995a; Remy *et al.* 1998a; Schenk *et al.* 1999; Bosque-Pérez *et al.* 2000) among others (Table 1.3).

Multiple promoters have been used in the transient and/or stable transformation of bananas (first reviewed by Swennen *et al.* 2003), most of them being heterologous. An

[1]www.promusa.org/index.php?option=com_content&task=view&id=116&Itemid=

updated overview of promoters used in transient and stable transformation experiments is shown in Table 1.3. Promoters from viral sources including the CaMV 35S RNA based promoters have been extensively used providing constitutive and high levels of transgene expression. In addition, the nopaline synthase promoter from *A. tumefaciens* has been employed in several banana genetic transformation events especially to drive the expression of the selectable marker gene (Table 1.3). Except for maize ubiquitin and rice actin promoters (Table 1.3) few plant promoters were used in banana transformation so far. Even less native promoters have been used. Only the promoters of the banana *actin* gene and genes involved in banana fruit development have been used in the generation of transgenic bananas.

Transient expression in banana ECS after PB revealed that the highest level of expression, as measured by histochemical and fluorometric GUS assays, was obtained with the maize *Ubi*1 promoter (Sági *et al.* 1995a). The following ranking of the tested promoters was obtained: *Ubi*1>e35S-AMV>Emu>e35S. Besides the CaMV 35S promoter and its derivates, promoters from several other plant-infecting viruses, including banana, have been tested in banana (Table 1.3). The intergenic regions of the six components of the banana bunchy top virus (BBTV; BT1.1 to BT6.1), revealed a very weak transient GUS activity after PB of banana ECS compared to the activity of the CaMV and *Ubi*1 promoters (Dugdale *et al.* 1998). However, GFP activity was detected for promoter BT6.1 in banana embryos and roots, especially in root meristems, and in leaf phloem and stomata (Dugdale *et al.* 1998). Further analyses of the BT1 to BT5 promoters at the plant level revealed that the highest expression was driven by the BT4 and BT5 promoters (Dugdale *et al.* 2000). Furthermore, the BBTV promoters apparently direct vascular-associated expression in transgenic banana plants and intron-mediated enhancement was detected when using the maize *Ubi*1 intron in a promoter dependent manner (Dugdale *et al.* 2000). A promoter derived from the BBTV-associated component S1 (S1.1) directs vascular tissue-specific expression in banana (Hermann *et al.* 2001b) similarly to the pattern observed for the other BBTV-derived promoters (Dugdale *et al.* 1998). Expression analysis revealed a near-constitutive expression of a taro bacilliform virus (TaBV) intergenic DNA fragment (T600 and T500) in banana with the strongest activity detected in vascular tissue (Yang *et al.* 2003). Analysis of promoter activity has also been tested in mature greenhouse banana plants. Activity of the *uid*A reporter gene driven by the Gelvin or Super promoter was detected in leaf segments of greenhouse plants and in ripe banana fruits (Ganapathi *et al.* 2001). Stable *Ace*-AMP1 expression driven by the *Ubi*1 promoter and measured by ELISA in leaf and fruit of greenhouse 'Williams' plants was detected at least throughout 5 generations (Remy 2000). Western blots and ELISA on roots of greenhouse plants revealed that the maize *Ubi*1 promoter drives higher gene expression than the Gelvin and Arabidosis tubulin promoters (Atkinson *et al.* 2004). Activity of two promoters isolated from the banana streak badnavirus (BSV) were analyzed in banana (Schenk *et al.* 2001; Remans

2002). Transgenic banana plants containing the BSV promoters isolated from the cultivars 'Mysore' (My) and 'Williams' (Cv) fused to the *uid*A reporter gene showed a near-constitutive GUS activity with strong activity in the corm and vascular tissues (Schenk *et al.* 2001). The My promoter was 29- to 3-fold more active in leaves of transgenic *in vitro* and greenhouse banana plants than the maize *Ubi*1 promoter, respectively, and similar differences in expression levels were detected in pseudostems and corm tissues of *in vitro* plants between the My and *Ubi*1 promoters. *In vitro* banana plants transformed with the Cv promoter revealed a comparable level of activity than that obtained with the *Ubi*1 promoter. Although the promoter activity of the bacterial nopaline synthase (*nos*) promoter was not compared with other promoters, its activity is confirmed due to the achievement of stable transgenic banana plants. The *nos* promoter drives the selectable marker gene in several reports and antibiotic-resistant transgenic banana plants have been obtained at high frequencies (Table 1.3).

Recombinant promoters have been used in few reports to achieve high level of transgene expression. The Emu promoter is a chimera containing a truncated maize alcohol dehydrogenase (*Adh*1) promoter with several copies of the anaerobic responsive element from the *Adh*1 promoter and octopine synthase (*ocs*) elements from the *ocs* gene (Last *et al.* 1991). Banana transient GUS expression assays revealed that the Emu promoter has a lower activity than the enhanced 35S promoter plus alfalfa mosaic virus leader sequence and maize *Ubi*1 promoter (Sági *et al.* 1995a). In addition, ELISA assays performed in stable transgenic banana plants for detection of transgene product confirmed the lower activity of the Emu promoter compared to the enhanced 35S and *Ubi*1 promoters (6- to 16-fold and 9- to 30-fold lower, respectively; Remy *et al.* 1998b).

Except for the BSV My promoter (Schenk *et al.* 2001) and the 5' deletion variant T600 of a TaBV promoter (Yang *et al.* 2003), the maize *Ubi*1 promoter showed at least similar or the highest activity when compared to other promoters (Sági *et al.* 1995a; Dugdale *et al.* 1998; Schenck *et al.* 1999; Hermann *et al.* 2001a, 2001b). Although a higher activity of the *Ubi*1 promoter was detected when fused to the *uid*A reporter gene than the six intergenic regions of the BBTV in transient assays (see above), higher activity was detected in transient assays for the BBTV intergenic regions BT5.1 and BT4.1 than the *Ubi*1 promoter when the promoters were fused to a codon-modified *gfp* reporter gene (Dugdale *et al.* 1998). Analysis of promoter activity in leaf tissue of banana plants stably transformed with 5' deletion variants of the TaBV promoters (T1200, T600, T500) fused to the *uid*A reporter gene revealed similar activities for the *Ubi*1 and the variants T1200 and T500 (Yang *et al.* 2003). However, the T600 variant revealed a four-fold higher GUS activity than the *Ubi*1 promoter (Yang *et al.* 2003).

Table 1.3. Promoters used for transient and/or stable gene expression in banana

Promoter					
Class[a]	**ID**	**Source organism**	**Description**	**Banana variety/cultivar/hybrid**[b]	**References**
Viral	35S	Cauliflower mosaic virus (CaMV)	CaMV 35S RNA promoter	'Bluggoe' (ABB), 'Grand Nain' (AAA), 'Three Hand Planty' (AAB), 'Calcutta 4' (AA), 'Williams (AAA), 'Mbwazirume' (AAA), *Musa balbisiana* (BB), 'TMPx 2796-5' (AAB X AA), 'TMPx 548-9' (AAB X AA), 'Dominico Hartón' (AAB), 'Lady finger' (AAB), 'Obino l'Ewai' (AAB), 'Orishele' (AAB), 'Taijiao' (AAA), 'Agbagba' (AAB), 'Rastali' (AAB), 'Maçã' (AAB), 'Mas' (AA), 'Gros Michel' (AAA)	Sági *et al.* 1992, 1994, 1995b, 1995c, 1998; Dugdale *et al.* 1998, 2000, 2001; Pérez Hernández *et al.* 1998, 2000, 2006a, 2006b; Schenk *et al.* 1999, 2001; Becker *et al.* 2000; Remy 2000; Bosque-Pérez *et al.* 2000; Remans 2002; Pineda *et al.* 2002; Yang *et al.* 2003; Atkinson *et al.* 2004; Khanna *et al.* 2004; Remy *et al.* 2004; Arinaitwe *et al.* 2004; Acereto-Escoffie *et al.* 2005; Pei *et al.* 2005; Tripathi *et al.* 2005; Panis *et al.* 2005b; Sreeramanan *et al.* 2006a, 2006b, 2006c; Matsumoto *et al.* 2007; Maziah *et al.* 2007a, 2007b; Huang *et al.* 2007; Arinaitwe 2008
	e35S	CaMV	Enhanced CaMV 35S promoter containing a duplication of the CaMV35S enhancer region	'Bluggoe' (ABB), 'Three Hand Planty' (AAB), 'Grand Nain' (AAA), 'Williams' (AAA), 'Lady finger' (AAB), 'Rastali' (AAB), 'Maçã' (AAB), 'Mas' (AA)	Sági *et al.* 1995a, 1995b, 1995c; Remy *et al.* 1998b; Remy 2000; Matsumoto *et al.* 2002; Wenck *et al.* 2003; Khanna *et al.* 2004; Acereto-Escoffie *et al.* 2005; Tripathi *et al.* 2005; Sreeramanan *et al.* 2006c; Huang *et al.* 2007; Santos *et al.* 2007
	35S-AMV	CaMV and alfalfa mosaic virus (AMV)	CaMV 35S promoter and AMV leader sequence	'Bluggoe' (ABB), 'Three Hand Planty' (AAB)	Sági *et al.* 1995c; Remy 2000
	e35S-AMV	CaMV and AMV	Enhanced CaMV 35S promoter and AMV leader sequence	'Bluggoe' (ABB), 'Three Hand Planty' (AAB), 'Grand Nain' (AAA), 'Maçã' (AAB)	Sági *et al.* 1994, 1995a, 1995b, 1995c, 1998; Schenk *et al.* 1999, 2001; Remy 2000; Pérez Hernández 2000; Matsumoto *et al.* 2002; Acereto-Escoffie *et al.* 2005
	EN435S	CaMV	CaMV 35S promoter containing four repeated enhancer regions (-290 to -90)	'Rastali' (AAB), 'Grand Nain' (AAA), 'Gros Michel' (AAA)	Sreeramanan *et al.* 2006a, 2006b; Maziah *et al.* 2007b; Arinaitwe 2008
	Sc	Sugarcane bacilliform badnavirus (ScBV)	First open reading frame promoter region of the ScBV	'Williams' (AAA), 'Grand Nain' (AAA), 'Three Hand Planty' (AAB),	Schenk *et al.* 1999
	My	Banana streak badnavirus (BSV) isolated from banana cultivar 'Mysore'	Sequence surrounding transcription start site of the BSV 'Mysore' isolate	'Grand Nain' (AAA), 'Williams' (AAA), 'Three Hand Planty' (AAB), 'Obino l'Ewai' (AAB), 'Orishele' (AAB)	Schenk *et al.* 2001; Remans 2002; Arinaitwe *et al.* 2004; Arinaitwe 2008

Table 1.3. Continued

Promoter					
Class[a]	**ID**	**Source organism**	**Description**	**Banana variety/cultivar/hybrid**	**References**
Viral	Cv	BSV isolated from banana cultivar 'Williams'	Sequence surrounding transcription start site of the BSV 'Williams' isolate	'Williams' (AAA), 'Three Hand Planty' (AAB)	Schenk *et al.* 2001; Remans 2002
	BT1.1	Banana Bunchy Top Virus (BBTV)	Intergenic region of BBTV DNA-1	'Bluggoe' (ABB)	Dugdale *et al.* 1998, 2000
	BT2.1	BBTV	Intergenic region of BBTV DNA-2 with and/or without maize *Ubi*1 intron	'Bluggoe' (ABB)	Dugdale *et al.* 1998, 2000
	BT3.1	BBTV	Intergenic region of BBTV DNA-3 with and/or without maize *Ubi*1 intron	'Bluggoe' (ABB)	Dugdale *et al.* 1998, 2000
	BT4.1	BBTV	Intergenic region of BBTV DNA-4 with and/or without maize *Ubi*1 intron	'Bluggoe' (ABB)	Dugdale *et al.* 1998, 2000
	BT5.1	BBTV	Intergenic region of BBTV DNA-5 with and/or without maize *Ubi*1 intron	'Bluggoe' (ABB)	Dugdale *et al.* 1998, 2000
	BT6.1	BBTV	Intergenic region of BBTV DNA-6 with and/or without maize *Ubi*1 intron	'Bluggoe' (ABB)	Dugdale *et al.* 1998
	BT6.3	BBTV	A 5' deleted version of the BT6.1 promoter plus first exon and intron of maize *Ubi*1	'Bluggoe' (ABB), 'Grand naine' (AAA), 'Lady finger' (AAB)	Dugdale *et al.* 1998, 2000, 2001; Becker *et al.* 2000; Yang *et al.* 2003
	BT6.1a	BBTV	BT6.1 promoter plus intron mediated enhancement of either maize U*bi*1, maize *adh*1, rice *act*1 or sugarcane *rbcs* genes	'Bluggoe' (ABB)	Dugdale *et al.* 2001
	S1	BBTV	Intergenic region of BBTV S1	'Bluggoe' (ABB)	Hermann *et al.* 2001b
	S2	BBTV	Intergenic region of BBTV S2 (with or without maize *Ubi*1 intron for the S2.2 variant)	'Bluggoe' (ABB)	Hermann *et al.* 2001b

Table 1.3. Continued

Promoter					
Class[a]	**ID**	**Source organism**	**Description**	**Banana variety/cultivar/hybrid**[b]	**References**
Viral	T1200, T600, T500	Taro bacilliform badnavirus (TaBV)	First open reading frame promoter region of TaBV	'Lady finger' (AAB)	Yang *et al.* 2003
Bact.	*nos*	*Agrobacterium tumefaciens*	Nopaline synthase gene promoter	'Grand Nain' (AAA), 'Three Hand Planty' (AAB), 'Williams' (AAA), 'Calcutta 4' (AA), 'Mbwazirume' (AAA), *Musa balbisiana* (BB), 'TMPx 2796-5' (AAB X AA), 'TMPx 548-9' (AAB X AA), 'Lady finger' (AAB), 'Obino l'Ewai' (AAB), 'Orishele' (AAB), 'Rastali' (AAB), 'Maçã' (AAB)	May *et al.* 1995; Sági *et al.* 1998; Pérez Hernández *et al.* 1998, 2000, 2006a, 2006b; Remy *et al.* 1998b, 2004; Bosque-Pérez *et al.* 2000; Remy 2000; Chakrabarti *et al.* 2003; Khanna *et al.* 2004; Arinaitwe *et al.* 2004; Sunil Kumar *et al.* 2005; Acereto-Escoffie *et al.* 2005; Panis *et al.* 2005b; Sreeramanan *et al.* 2006a, 2006b, 2006c; Matsumoto *et al.* 2007; Maziah *et al.* 2007a; Arinaitwe 2008
	mas	*Agrobacterium tumefaciens*	Mannopine synthase gene promoter	'Williams' (AAA), 'Three Hand Planty' (AAB), 'Orishele' (AAB)	Remy 2000; Arinaitwe 2008
Reco.	Emu	Maize, *Agrobacterium tumefaciens*	A chimera including six copies of the 41-bp ARE (anaerobic responsive element) from the Adh1 maize gene promoter plus four copies of the 40-bp ocs (octopine synthase) enhancer plus the 5' end of a truncated adh1 promoter linked to its first intron	'Bluggoe' (ABB), 'Three Hand Planty' (AAB), 'Rastali' (AAB)	Sági *et al.* 1995a, 1995b, 1995c, 1998; Remy 2000; Ganapathi *et al.* 2001
	(*ocs*)$_3$*mas* /Gelvin/ Super-promoter	*Agrobacterium tumefaciens*	A chimera including three repeats of the octopine synthase (ocs) promoter activator plus the mannopine synthase (mas) promoter activator and promoter	'Three Hand Planty' (AAB), 'Grand Nain' (AAA), 'Rastali' (AAB)	Remy *et al.* 1998b, 2000; Ganapathi *et al.* 2001; Atkinson *et al.* 2004
Plant	*Ubi*1	*Maize*	Maize ubiquitin gene promoter including first untranslated exon and intron	'Bluggoe' (ABB), 'Williams' (AAA), 'Grand Nain' (AAA), 'Three Hand Planty' (AAB), 'Lady finger' (AAB), 'Obino l'Ewai' (AAB), 'Orishele' (AAB), 'Rastali' (AAB)	Sági *et al.* 1995a, 1998; Dugdale *et al.* 1998, 2001; Remy *et al.* 1998b; Schenk *et al.* 1999, 2001; Becker *et al.* 2000; Remy 2000; Pérez Hernández 2000; Ganapathi *et al.* 2001; Hermann *et al.* 2001a, 2001b; Remans 2002; Yang *et al.* 2003; Atkinson *et al.* 2004; Remy *et al.* 2004; Arinaitwe *et al.* 2004; Sreeramanan *et al.* 2005, 2006a, 2006b; Arinaitwe 2008

Table 1.3. Continued

Promoter					
Class[a]	**ID**	**Source organism**	**Description**	**Banana variety/cultivar/hybrid**[b]	**References**
Plant	*Adh*1	Maize	Maize alcohol dehydrogenase 1	'Bluggoe' (ABB)	Dugdale *et al.* 2001
	*Act*1	Rice	Rice act1 gene promoter including first untranslated exon and intron	'Grand Nain' (AAA), 'Three Hand Planty' (AAB), 'Calcutta 4' (AA), 'Williams (AAA), 'Mbwazirume' (AAA), *Musa balbisiana* (BB), 'TMPx 2796-5' (AAB X AA), 'TMPx 548-9' (AAB X AA), 'Obino l'Ewai' (AAB), 'Orishele' (AAB), 'Rastali' (AAB)	May *et al.* 1995; Sági *et al.* 1998; Schenk *et al.* 1999, 2001; Remy 2000; Bosque-Pérez *et al.* 2000; Remans 2002; Arinaitwe *et al.* 2004; Sreeramanan *et al.* 2005, 2006a, 2006b
	TUB-1	Arabidopsis	Arabidopsis tubulin gene promoter	'Grand Nain' (AAA)	Atkinson *et al.* 2004
	*ubq*3	Arabidopsis	Arabidopsis ubiquitin gene promoter	'Maçã' (AAB), 'Rastali' (AAB)	Chakrabarti *et al.* 2003; Sunil Kumar *et al.* 2005; Matsumoto *et al.* 2007
	ahas	Arabidopsis	Arabidopsis acetohydroxyacid synthase gene promoter	'Maçã' (AAB)	Matsumoto *et al.* 2002; Matsumoto *et al.* 2007
	*act*2p	Arabidopsis	Arabidopsis actin 2 gene promoter including mRNA leader exon, an intron in the mRNA leader and first 19 codons	'Maçã' (AAB)	Matsumoto *et al.* 2002
	*ACT*1	Banana	Banana actin gene promoter plus untranslated leader	'Bluggoe' (ABB)	Hermann *et al.* 2001a
	EFE	Banana	Ethylene forming enzyme gene promoter[c]	'Rastali' (AAB)	Sunil Kumar *et al.* 2005
	ACC	Banana	1-Aminocyclopropane-1-carboxylate (ACC) synthase gene promoter		Wang and Peng 2001a
	*ACO*1	Banana	ACC oxidase gene promoter		Wang and Peng 2001b

ID refers to identification name. [a]Promoters were classified according to the source organism in viral, bacterial (Bact.), recombinant (Reco.) or plant. Recombinant refers to promoters which were modified by the fusion of two or more different promoter elements from similar or different species.

[b]The name of the banana cultivar, variety or hybrid is mentioned and the genomic group is indicated between brackets. A refers to *Musa acuminata* and B to *Musa balbisiana*.

[c]The *EFE* promoter has 88.8% similarity to the *ACO*1 banana promoter (Wang and Peng 2001b).

The banana ethylene forming enzyme gene promoter (*EFE*) driving the 's' gene of the hepatitis B surface antigen (HBsAg), showed a 3-fold higher activity as determined by ELISA in transgenic *in vitro* banana lines than the Arabidopsis *ubq*3 gene promoter (Sunil Kumar *et al.* 2005). However, expression analysis at the greenhouse plants revealed a higher expression of the *ubq*3 promoter than in the *EFE* promoter. The difference in *EFE* promoter activity between *in vitro* and greenhouse banana plants might be due to the production of ethylene during tissue culture (Biddington 1992) leading to a higher expression in *in vitro* than in greenhouse plants (Sunil Kumar *et al.* 2005).

A 1.2 kb banana *actin* 1 promoter deleted variant drove a 2-fold higher transient *uid*A expression in bombarded banana ECS than the CaMV 35S promoter but showed 4-fold less activity than the maize *Ubi*1 promoter (Hermann *et al.* 2001a). Similarly in leaf extracts of transgenic plants, GUS activity was 2-fold higher for the 1.2 kb actin 1 promoter than the CaMV35S promoter but 2-fold lower than the maize *Ubi*1 promoter (Hermann *et al.* 2001a).

In most cases, constitutive expression of genes including stress tolerance genes leads to undesirable pleiotropic effects during plant development. Therefore, identification and employment of promoters active only at certain stages, in certain organs or under specific stresses are needed to direct expression of stress tolerance genes (Sreenivasulu *et al.* 2007). Promoters of banana could be used not only to transform other plant species but also for the transformation of other banana cultivars. The introduction of genes with their native promoters from a crossable plant or from the crop plant itself through genetic transformation has been called cisgenesis (Schouten *et al.* 2006; Jacobsen and Schouten 2007). Cisgenic plants should provide more public acceptance than transgenic plants and might be exempted from legislations concerning genetically modified organism (Jacobsen and Schouten 2007). However, the combination of banana promoters fused to unrelated banana genes should also provide more public acceptance. Almost all cultivated bananas are sterile and thus, genetic transformation will undoubtfully become a useful tool for their genetic improvement.

1.5 The Global *Musa* Genomics Consortium

Due to the importance of genetic information for the improvement in *Musa*, the Global *Musa* Genomics Consortium (GMGC) was created in 2002 by scientists around the world (www.musagenomics.org; The Global *Musa* Genomics Consortium 2002). With the advent of new, fast and high-throughput genomics technologies that cover the sequence and the analysis of whole genomes the GMGC believes that these technologies should be applied for the sustainable improvement of banana. The GMGC is an international network of researchers dedicated to the establishment of banana as a model crop for comparative genomics and gene discovery leading eventually to the development of improved banana

varieties. Currently, a total of 33 institutions in 23 countries are committed to collaborate and agree to share materials and resources which includes sequence information and technologies (Roux *et al.* 2008). The information developed within the framework of the GMGC includes (i) genetic maps, (ii) BAC libraries that are available under a material transfer agreement (MTA), (iii) the status of cytogenetics and genome organization, (iv) the sequencing status of the banana genome and (v) the status on the development of functional genomics tools including cDNA (EST) libraries, gene/promoter trapping and mutation induction, among others. Due to the GMGC duplication of efforts is minimized and valuable banana genomics information is accessible to the *Musa* research community.

1.6 Objective and outline of this study

A high throughput T-DNA tagging strategy was followed for the characterization and isolation of novel banana promoters. The main objective of this work was to characterize and isolate banana promoters that are responsive throughout *in vitro* development and LT stress. First, promoters were characterized in T-DNA tagged banana lines that showed LUC activity throughout *in vitro* development in combination with LT stress. Then, the promoter(s) of a selected number of such T-DNA tagged lines were isolated and analyzed. Finally, the activity of the candidate promoters was validated by cloning them upstream of the *uid*A^{INT} reporter gene for back-transformation to confirm promoter activity at early *in vitro* regeneration stages and under LT stress.

First a general overview of (i) the effect of LT in banana development, (ii) the role of plant promoters in gene expression during LT stress and development, (iii) the T-DNA tagging approach to isolate plant promoters, (iv) the promoters used for transient and/or stable gene expression in banana, and (v) the genome information and genomics tools developed for *Musa* is given (Chapter 1). The materials and methods followed in this study are described in Chapter 2. Embryogenic cell suspensions were transformed by *Agrobacterium tumefaciens* carrying a promoterless luciferase (*luc*$^+$) gene next to the right T-DNA border (Chapter 3). Screening for LUC activity was performed in promoter-tagged lines throughout *in vitro* development and under LT stress (Chapter 3). Once regenerated promoter-tagged lines were selected, molecular characterization including Southern blot analysis and isolation of banana promoter sequences using PCR-based methods was performed (Chapter 4). Most of the promoter-tagged lines contained more than one T-DNA, which makes the identification and isolation of the banana promoter(s) complex. Selection of candidate promoters in multiple T-DNA copy lines was based on *in silico* analysis to identify putative *cis*-acting promoter elements in the isolated banana sequences (Chapter 4) and on the analysis of transcriptional fusion between the isolated banana sequences and the *luc*$^+$ gene *via* a RT-PCR approach (Chapter 5). Finally, candidate banana promoter sequences were fused to the *uid*A^{INT} reporter gene and back-transformed into

banana to confirm promoter activity in early regeneration stages and under LT stress (Chapter 6). A general discussion and perspectives of this study is presented in Chapter 7. An overview of the different stages of this study is shown in Figure 1.2.

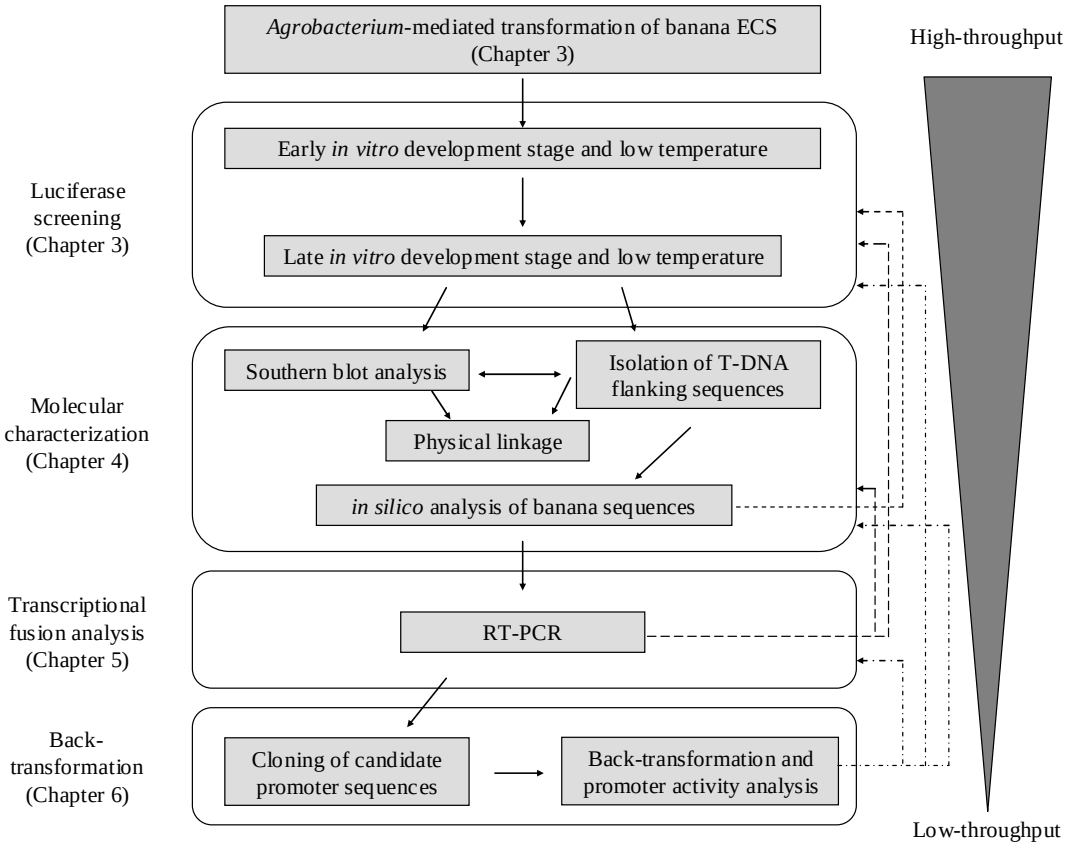

Figure 1.2. Flow chart of the conducted research. ECS refers to embryogenic cell suspension. Dotted (and dashed) arrows represent reflection at each step during the process to find correlation(s) with previous steps.

Chapter 2 Materials and methods

2.1 *Agrobacterium*-mediated transformation

2.1.1 Plant material and culture conditions

Embryogenic cell suspensions (ECS) of different banana cultivars initiated from proliferating *in vitro* meristems (Dhed'a *et al.* 1991; Strosse *et al.* 2003; Strosse 2005) were subcultured every two weeks in ZZ medium [half-strength MS medium (Murashige and Skoog 1962) supplemented with 5 µM 2,4-D and 1 µM zeatin; Table 2.1]. The main banana cultivar employed for promoter tagging was 'Three Hand Planty' (THP, AAB genomic group, International Transit Center accession number, ITC.0185), which is an important plantain-type banana in Africa. Since high quality ECS of THP (Strosse 2005) with a very high transformation competence (S. Remy, personal communication) was available at the Laboratory of Tropical Crop Improvement, it was the cultivar of choice in the promoter tagging experiments described here. On a smaller scale, the dessert banana 'Williams' (Will, AAA genomic group, ITC.0365) and the cooking banana 'Cacambou' (Caca, ABB genomic group, ITC.0058) were also transformed. Will is one of the most important cultivars in the export banana industry, while Caca is an important food source in Africa.

Table 2.1. Plant culture media

	Medium						
Component	**ZZ**	**RD1**	**RD2**	**PROL**	**REG**	**TDZ**	**P4**
½ MS (g L^{-1})	2.17	2.17	2.17	-	-	-	
MS with vitamins (g L^{-1})	-	-	-	4.4	4.4	4.4	4.4
MS Vitamins (mL L^{-1})	1	1	1	-	-	-	-
Ascorbic acid (mg L^{-1})	10	10	10	10	10	10	10
IAA (µM)	-	-	-	1	1	-	1
BA (µM)	-	-	1	10	1	-	100
Myo-inositol (mg L^{-1})	-	100	-	-	-	-	-
Zeatin (µM)	1	-	-	-	-	-	-
2,4-D (µM)	5	-	-	-	-	-	-
TDZ (µM)	-	-	-	-	-	10	-
Sugar (g L^{-1})	30	30	30	30	30	30	30
Gelrite (g L^{-1})	3	2.5	2.5	3	3	3	3
pH (after autoclave)	5.8	5.8	5.8	5.8	5.8	5.8	5.8

MS, Murashige and Skoog culture medium (Murashige and Skoog 1962); IAA, indolyl-3-acetic acid; BA, benzyladenine; 2,4-D, 2,4-dichlorophenoxyacetic acid; TDZ, thidiazuron.

2.1.2 Transformation vectors

Maps for promoter tagging and corresponding positive control vectors are shown in Figure 2.1. The T-DNA of the first promoter tagging vector pETKUL2 (Remy *et al.* 2005) contains the promoterless codon-optimized firefly luciferase gene (*luc*$^{+}$; Sherf and Wood 1994) fused to the nopaline synthase terminator (Tnos) and the neomycin phosphotransferase II (*neo*) selectable marker gene driven by the enhanced 35S RNA promoter (pe35S) and followed by the 35S RNA terminator (T35S) from the Cauliflower Mosaic Virus (CaMV). As control the vector pETKUL3 was used in which the *luc*$^{+}$ is driven by the pe35S and the Tobacco Mosaic Virus untranslated leader sequence (T) followed by the T35S (Remy *et al.* 2005). The selectable marker gene cassette is identical to that in pETKUL2. The second promoter tagging vector pKCKUL1 (Remy *et al.* unpublished) is similar to pETKUL2 but contains the promoterless *luc*$^{+}$ gene with intron 4 of the maize *RpoT* gene inserted at nucleotide position +165 of the *luc*$^{+}$ open reading frame (ORF) (*luc*$^{+INT}$; Bourdon *et al.* 2001). The control vector pKCKUL2 (Remy *et al.* unpublished) is identical to pKCKUL1 except that the maize ubiquitin (*Ubi*1) promoter plus leader intron (L) is driving the *luc*$^{+INT}$ gene. Vector pFAJ3000 (not shown) contains the *uid*A gene with the potato PIV2 intron driven by the 35S RNA promoter from CaMV (Debondt *et al.* 1994). All vectors were transferred to *E. coli* TOP10® cells (Invitrogen, Merelbeke, Belgium) by heat shock and to *A. tumefaciens* strain EHA105 (Hood *et al.* 1993) by electroporation.

2.1.3 Transformation protocol

Agrobacterium-mediated transformation of banana ECS was performed as described by Remy *et al.* (2005) and Pérez Hernández *et al.* (2006b). Agrobacteria were co-cultivated with embryogenic suspension cells of the plantain THP, the dessert banana Will and the cooking banana Caca. The different media used during the transformation, selection and regeneration of ECS are shown in Table 2.1.

2.1.3.1 *Agrobacterium* infection and co-cultivation

In the standard protocol one infected sample contains 200 µL of a 33% settled cell volume (SCV) of ECS which equals approximately 50 mg fresh weight (FW) cells mixed with 1000 µL induced bacteria (SCV:agrobacteria ratio of 1:5). In the initial experiment 10 different ratios of SCV:agrobacteria were tested ranging from 1:10 (100 µL:1000 µL) to 1:40 (25 µL:1000 µL), or the ratio was kept constant at 1:5 but volumes were lowered (100 µL:500 µL, 50 uL:250 µL, 25 µL:125 µL). In experiment 11 a ratio of 1:10 was used (100 µL:1000 µL), while the standard volumes and ratio were used in the remaining experiments 15 and 16. The promoter tagging experiments 1-9 and 12-14 are not part of the work described here.

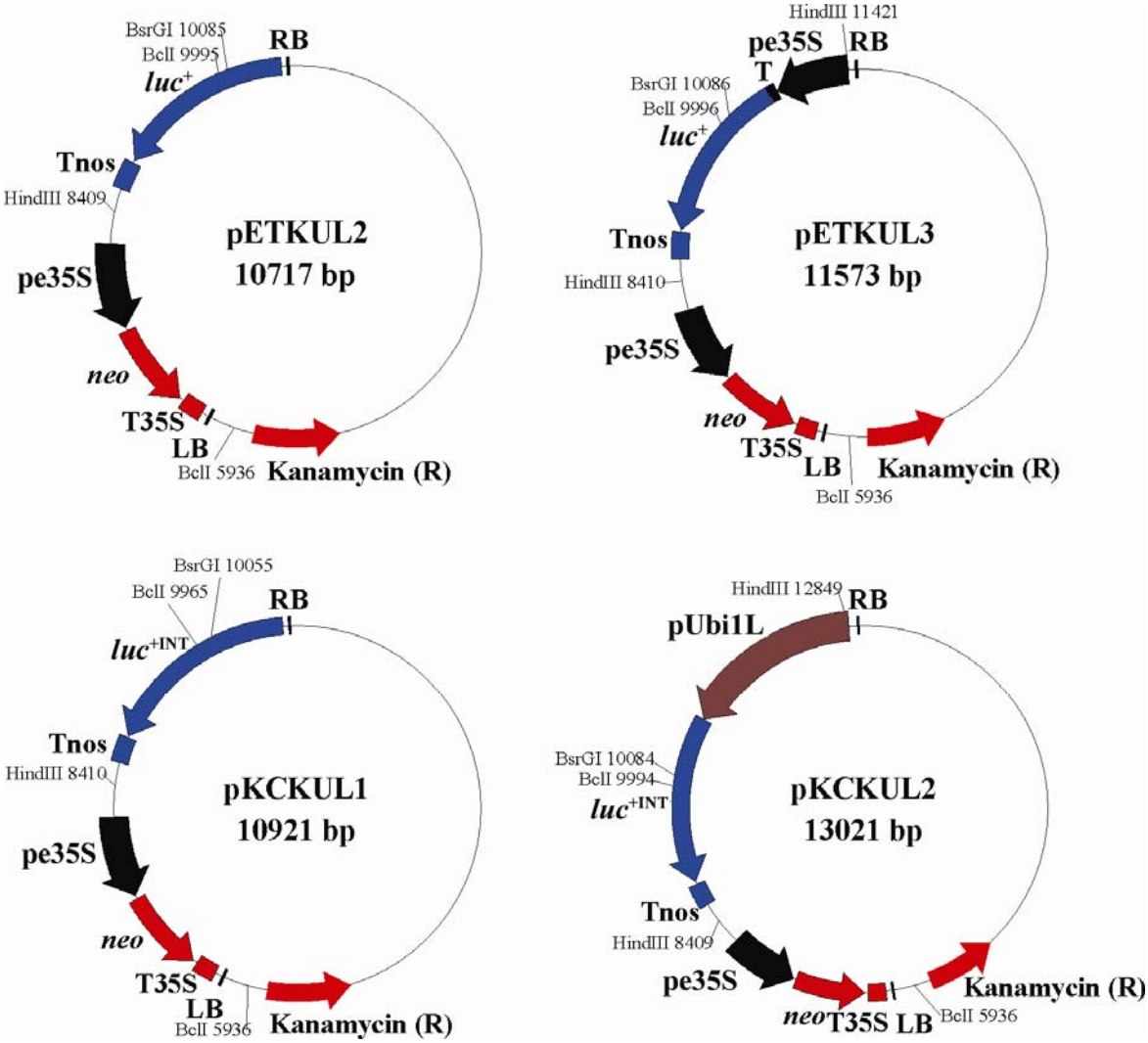

Figure 2.1. **Promoter tagging vectors pETKUL2 and pKCKUL1, and their corresponding positive control vectors pETKUL3 and pKCKUL2, respectively.** Kanamycin (R), kanamycin resistance gene; LB, T-DNA left border; T35S, Cauliflower Mosaic Virus (CaMV) 35S RNA terminator; *neo*, neomycin phosphotransferase II gene; pe35S, enhanced CaMV 35S RNA promoter; Tnos, nopaline synthase terminator; luc^+, codon-optimized firefly luciferase gene; RB, T-DNA right border; T, Tobacco Mosaic Virus untranslated leader sequence; luc^{+INT}, luc^+ with intron 4 of maize *RpoT* gene inserted at postion +165 of the luc^+ ORF; pUbi1L, maize ubiquitin promoter plus leader intron. Relevant restriction sites and their positions are indicated.

2.1.3.2 Selection of transgenic cell colonies

The standard selection procedure and the selection procedure applied in experiment 10 are shown in Figure 2.2. The composition of the media is given in Table 2.1. After six days of co-cultivation, the cell mixtures were transferred on solid ZZ medium containing 50 µg mL^{-1} geneticin (for the selection for transgenic cultures) and 200 µg mL^{-1} timentin (for the elimination of agrobacteria). The cells were incubated in darkness at 26 ± 2°C with a biweekly subculture for two to three months. Selection of transgenic lines in experiments 11, 15 and 16 was also done (not shown) but luciferase screening during *in vitro* development was not performed.

To calculate the transformation frequency (the number of putative transgenic cell cultures per sample of 50 mg FW cells) and the promoter tagging frequency (the number of independent candidate promoter tagged lines expressed in percentage of the total number of independent transgenic lines) the number of independent putative transgenic lines was counted two months after transformation in at least three samples. Counting was performed using a stereoscope with a magnification of 12 X to 25 X.

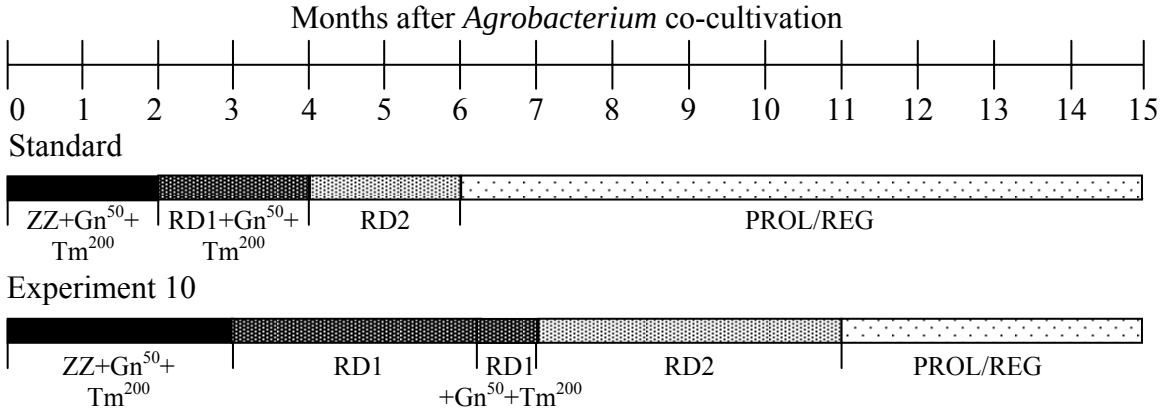

Figure 2.2. Schematic representation of selection and regeneration procedures. Selection for transgenic cultures was performed on ZZ and RD1 media containing geneticin (Gn) and timentin (Tm). Shoot formation was induced on RD2 medium and further regeneration was accomplished upon transfer onto PROL or REG medium. Transgenic lines were maintained on the latter medium under a three monthly subculture regime. ZZ, cell proliferation medium [half-strength MS (Murashige and Skoog 1962), containing 5 µM 2,4-D and 1 µM zeatin]; Gn^{50}, geneticin (50 µg mL^{-1}); Tm^{200}, timentin (200 µg mL^{-1}); RD1, embryo induction medium (half-strength MS containing 100 mg L^{-1} *myo*-inositol); RD2, shoot induction medium (half-strength MS medium containing 10 µM benzyladenine); PROL, shoot multiplication medium (MS medium containing 10 µM benzyladenine and 1 µM indole acetic acid); REG, plant regeneration medium (MS medium containing 1 µM benzyladenine and 1 µM indole acetic acid).

2.1.3.3 Regeneration of transgenic lines

The standard regeneration procedure and the regeneration procedures applied in experiment 10 are shown in Figure 2.2, while the composition of the media is indicated in Table 2.1. First, transgenic cell colony lines (a line is an independent enumerated transgenic event or cell that regenerates into a plant) maintained on ZZ medium were transferred onto RD1 medium for embryo induction, and then placed on RD2 medium for shoot induction. Finally, regenerating cultures were transferred onto PROL or REG medium for shoot multiplication or plant regeneration including root formation, respectively. Cultures were kept for a limited period of time under selective pressure on RD1 medium, but from RD2 medium onwards all cultures were kept free of antibiotics (Figure 2.2). All cultures were kept at 26 ± 2°C in the dark until shoot formation and under a 16 h photoperiod with a photosynthetic photon flux density of 50 µE m^{-2} s^{-1} provided by Cool White fluorescent lamps (TLD 58W/33, Philips, France) when transferred to PROL or REG medium.

2.1.3.4 Re-initiation of proliferating meristem cultures and induction of embryogenic calli from *in vitro* plants

After *in vitro* plant regeneration, proliferating meristem cultures were re-initiated from the apical meristem according to a modified procedure of Strosse *et al*. (2003). *In vitro* apical meristems were excised from regenerated plants and transferred to TDZ medium (Table 2.1) for two months with an intermittent subculture after one month. Meristem cultures were then transferred and maintained on P4 medium (Table 2.1) for at least six months until good quality, cauliflower-like proliferating cultures consisting of multiple meristems were obtained. Small-sized (1-2 mm^3) explants of these meristem cultures called scalps were placed on ZZ medium for at least five months for induction of embryogenic calli.

2.1.3.5 Transient GUS activity

After six days of co-cultivation with *A. tumefaciens* transient β-glucuronidase (GUS) gene activity was assessed. Polyester meshes with cell mixtures transformed with pFAJ3000 (paragraph 2.1.2) were transferred to 5-cm Petri dishes (PDs) each containing an autoclaved filter paper and histochemically stained for transient GUS activity (TGA) as described (Jefferson *et al.* 1987). One millilitre of a modified X-gluc solution (Mendel *et al.* 1989) containing 1 mg mL^{-1} of 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid cyclohexylammonium salt (X-gluc; Duchefa, Haarlem, The Netherlands) was applied per sample. Following a 5 h incubation at 37°C the samples were kept overnight at room temperature. TGA frequency was expressed as the number of blue foci per sample of 50 mg FW cells averaged over five replicates. Counting of blue foci was done under a stereoscope and images were acquired with a digital charge-couple device (CCD) camera (SpotRT-slider, Diagnostic Instruments Inc., Sterling Heights, MI, USA).

2.2 Screening for luciferase activity

2.2.1 Digital image acquisition and quantification of luciferase activity

Detection of LUC activity was performed with a liquid nitrogen-cooled slow-scan CCD camera (Versarray™ 512 B LN, Roper Scientific, Vianen, The Netherlands) equipped with a light sensitive camera lens (Nikkor F 50mm f/1.2, Nikon, Tokyo, Japan) as described by Remy *et al.* (2004). Briefly, samples were placed in a light-tight box, a live image was first captured as reference under light conditions, and then LUC images were recorded in complete darkness with an integration time of 20 min. Time-lapse (TL) images were taken at an interval of 30 or 20 min and image processing including quantification of light emission within the region(s) of interest (ROI) was performed with the Metamorph® 5.0r3 software (Universal Imaging, Downington, PA, USA). The integrated light intensity value of transgenic lines showing LUC activity was corrected for background by subtracting the

integrated light intensity of a non-transformed control line. Care was taken to keep the size of the ROI constant during this background correction.

2.2.2 Substrate

Luciferin (beetle luciferin potassium salt, Promega, Leiden, The Netherlands) was dissolved in milliQ (MQ) water or in half-strength MS medium at a final concentration of 0.1 mM, filtered sterilized and stored at -20°C. Application of luciferin took place by pipetting or spraying with a type 250-2 airbrush (Badger Airbrush Company, Franklin Parck, IL, USA). The volume applied per sample on ZZ, RD1, RD2 and REG medium was 400-600 µL, 50-100 µL, 400-600 µL and 600-800 µL, respectively. Repeated pipetting of the luciferin solution was performed when lines were screened on ZZ, RD1 and RD2 medium. Luciferin was included in the ZZ medium at a final concentration of 0.1 mM for the cell colonies of the 16°C → 4°C treatment (paragraph 2.2.4.2).

2.2.3 Baseline luciferase activity

Baseline luciferase activity (BLA) screening was performed either at room temperature or under temperature controlled conditions at 26°C (paragraph 2.2.4). At all developmental stages luciferin was applied once (1- 0.5 h) or twice (26-14 h and 1-0.5 h) prior to the screening. Cell cultures maintained on RD1 and RD2 medium were screened for BLA in either 5-cm PD or 24-well plates. *In vitro* plantlets maintained in test tubes containing PROL or REG media were transferred to 5-cm PD by placing them horizontally on filter paper saturated with liquid half-strength MS medium before spraying with luciferin. BLA screening of reinitiated proliferating meristem cultures and embryogenic calli (paragraph 2.1.3.4) was performed in 5-cm PD. Transgenic lines were screened in real-time for BLA at the different developmental stages for at least 2 h and up to 8 h. BLA acquired images were set in pseudocolor and analyzed by scaling the number of different grey values, and scored as very strong (VS), strong (S), moderate (M), weak (W) and not detectable (N) at an upper greyscale limit setting to more than 10,000, between 5000 to 10,000, 3000 and 5000, 500 and 3000 and to less than 500, respectively, over a total of 65,536 gray levels. Transgenic lines which were transformed with the promoter tagging vector pETKUL2 or the positive control vector pETKUL3 were coded as ET2 and ET3, respectively, and followed by an identification number.

2.2.4 Low temperature (LT) luciferase activity

2.2.4.1 Temperature control system

The temperature control system consisted of three parts (Figure 2.3). Firstly, a thin (2 mm thickness) metal plate (31 x 31 cm) was constructed on a top of a twisted copper tube (0.6 cm outer diameter) resulting in direct contact between both elements over the whole surface of the plate. This construction was then placed (metal plate facing upwards) inside a metal box of 31 x 31 x 25 cm with the in- and outlet of the copper tube extending from the box. To avoid condensation inside the culture vessels and guarantee visibility at the same time, the metal box was covered with a transparent plexiglass (polymethyl methacrylate) lid. Secondly, temperature was controlled by a circulating thermostatic bath (1140S, VWR International Inc., San Diego, CA, USA) containing a 1:1 mixture of distilled water and ethylene glycol (Acros Organics, Geel, Belgium). Thirdly, silicon tubes covered with foam for insulation connected the thermostatic bath outside the light-tight box with the temperature-controlled plate placed inside the light-tight box. The tubes entered through a light-tight opening in the back of the light-tight box. Temperature of the circulating thermostatic bath was calibrated to reach the desired temperatures within the culture vessels by monitoring the temperature with a wired thermometer (indoor/outdoor model No. EM899, Oregon Scientific, Portland, OR, USA).

Figure 2.3. **Temperature control system for real-time monitoring of luciferase activation during temperature changes.** (A) Circulating thermostatic water bath. (B) Silicon tubes covered with foam for insulation, connecting the thermostatic bath outside the light-tight box and the temperature-controlled plate inside the light-tight box. (C) Metal box containing the temperature-controlled metal plate. (D) In- and outlet of copper tube underneath the metal plate. (E) Transparent plexiglass lid covering the metal box. (F) Samples in 5-cm PD on the temperature-controlled metal plate inside the metal box.

2.2.4.2 Cell colony stage (ZZ medium)

Two to three months after transformation (Figure 2.2), transgenic cell colonies maintained on ZZ medium were screened in real-time for BLA at 26°C (paragraph 2.2.3), and subsequently at a LT of 18°C, 16°C, 12°C, 8°C or 16°C→4°C (an acclimation period at 16°C followed by a further decrease in temperature to 4°C). LUC screening performed after 3 months of transformation at cell colony stage is denominated as I.

Several TLs each of six to seven hours were recorded. After the application of luciferin, samples were placed on the temperature-controlled metal plate and the first TL covered the first two hours at 26°C when the temperature was set to reach the lower temperature as indicated above. This LT was reached approximately one hour later, and the first TL screening continued for another 3 h at the LT. Alternatively, the first TL recording was entirely done with the cultures under 26°C incubation (i.e. screening for BLA) and only during the second TL the temperature regime as described above was followed. After refilling the liquid nitrogen container of the camera the following TL was immediately initiated with the samples remaining at the LT. Eventually, one or several TLs were taken during the LT regime. Occasionally, a TL was conducted where the temperature was kept low for the first two hours and then the temperature was increased back to 26°C for the remaining hours. The latter increase was realized in three consecutive steps by raising the temperature in the thermostatic water bath by 1/3 of the difference between 26°C and the corresponding LT every 20 min. A final TL was performed after raising the temperature to 26°C. For the LT treatment 16°C→4°C the cultures were maintained at 16°C for seven days and then the temperature was set to reach 4°C. Following screening for BLA, one TL was recorded during the temperature drop from 26°C to 16°C, throughout the 7 day period at 16°C (one TL was done per day), during the second drop in temperature to 4°C and finally during the temperature increase back to 26°C.

Cell colonies showing LUC activity were scored with ↑, ↓ and = signifying an increase, decrease or status quo relative to the LUC activity at 26°C. Several cell colonies within the region of LUC activity were picked with a forcep and transferred to RD1 medium (paragraph 2.2.4.3) in a 5-cm PD and spread out in rows. For final identification of the temperature responding colony one to two months later a second screening following again the above outlined procedure was performed. In experiment 10, BLA positive cell colonies were identified at room temperature two months after transformation, removed and one month later the remaining cell colonies were screened under temperature-controlled conditions following the above outlined procedure. For experiments 11, 15 and 16 the LUC screening was not continued and only for lines of experiment 10, LUC screening was followed during the next stages of the *in vitro* regeneration process (paragraphs 2.2.4.3, 2.2.4.4, 2.2.4.5 and 2.2.4.6).

2.2.4.3 Embryo induction stage (RD1 medium)

Cell cultures on RD1 medium were screened six weeks after the last ZZ screening in experiment 10. TLs, temperature regimes and scoring of LUC activity were executed as described in paragraph 2.2.4.2. Responsive cultures were transferred onto RD1 or RD2 medium depending on their development (Figure 2.2). The LUC screenings performed at embryo induction stage five and seven months after transformation are denominated as II and III, respectively.

2.2.4.4 Shoot induction stage (RD2 medium)

Regenerating cultures placed on RD2 medium in 5-cm PD underwent a screening for LT LUC activation as described in paragraph 2.2.4.2. However, irrespective of their LUC activity pattern at this stage all independent lines were transferred to either PROL or REG medium. The LUC screenings performed at shoot induction stage eight and 10 months after transformation are denominated as IV, V (both at eight months) and VI, respectively.

2.2.4.5 Plantlet stage (PROL and REG medium)

Regenerated *in vitro* plantlets maintained on REG medium were screened for LT LUC activation in a horizontal position after removal from the solid medium and placing the samples on filter paper saturated with liquid half-strength MS medium in a 5-cm PD (paragraph 2.2.3) as described in paragraph 2.2.4.2. At least three independent screenings using more than nine different clones in total for each line were performed under LT stress. The LUC screenings performed at *in vitro* plantlet stage 14 to 20 months after transformation are denominated as VII.

2.2.4.6 Re-initiated proliferating meristem cultures and induction of embryogenic calli (P4 and ZZ media, respectively)

Before screening for LT LUC activation (paragraph 2.2.4.2) the re-initiated P4 and ZZ cultures were established for at least nine and five months, respectively. Similar to the BLA screening (paragraph 2.2.3) cultures were incubated in 5-cm PD during the LT screening procedure.

2.3 Molecular characterization of tagged lines

2.3.1 Total DNA isolation

Total DNA was isolated from 150-250 mg or 1-1.2 g of *in vitro* or greenhouse leaf tissue, respectively, with a protocol based on the methods of Dellaporta *et al.* (1983) and Aljanabi and Martinez (1997). Leaf material was crushed with the RETSCH MM301 Mixer Mill (Retsch Inc., Haan, Germany) twice at 30 Hz for 45 s using a stainless steel ball for homogenization. Four millilitre of freshly prepared extraction buffer containing 100 mM Tris-HCl pH 8.0, 50 mM ethylenediaminetetraacetic acid (EDTA), 500 mM NaCl, 10 mM β-mercaptoethanol and 2% (w/v) polyvinyl pyrrolidone (PVP, MW 40,000 or 10,000) was added to the powdered tissue. Sodium dodecyl sulfate (SDS) was added to a final concentration of 2% (w/v) and incubated for 10 min at 55°C. Then, 1300 µL of 5 M potassium acetate was added to obtain a final concentration of 1.17 M and after vortexing for a few seconds, the samples were centrifuged for 10 min at 12,000 rpm and 4°C. The supernatant was collected avoiding any debris, 3 mL of 6 M NaCl was added and vortexed at maximum speed for 30 s. RNaseA (Qiagen, Hilden, Germany) was added to a final concentration of 200 µg mL^{-1} and the mixture was incubated for 15 min at 37°C and centrifuged for 10 min at 12,000 rpm and 4°C. The supernatant was transferred to a chloroform resistant tube, an equal volume of chloroform:isoamylalcohol (24:1) was added with a glass pipet and the tube was inverted several times. Samples were centrifuged at 12,000 rpm for 5 min and the upper aqueous phase was transferred to a new tube. Chloroform extraction was occasionally repeated. An equal volume of isopropanol was finally added to the aqueous phase, mixed well and incubated at -20°C for at least 30 min. Following centrifugation at 12,000 rpm and 4°C for 20 min, the pellet was washed with 70% ethanol and briefly dried in a laminar flow before dissolving in 1 mL of autoclaved (MQ) water and stored at -20°C.

After adding loading buffer to reach a final concentration of 1 x [6 x: 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 30% (v/v) glycerol] the DNA was loaded together with a 1 kb or 100 bp reference ladder (Eurogentec, Seraing, Belgium) on a 0.8% (w/v) agarose gel. The gel was prepared in either 1 x TAE buffer (40 mM Tris-acetate, 1 mM EDTA) or 1 x SB buffer [10 mM NaOH, pH adjusted to 8.5 using H_3BO_3, (Brody and Kern 2004)], and run at 80 V for 80-100 min or at 160 V for 25-40 min, respectively. Following staining with 0.5 µg mL^{-1} ethidium bromide for 30 min and destaining for 30 min in destilled water (dH_2O) a picture was taken with a digital camera (Camedia C-3030, Olympus, Tokyo, Japan) when placing the gel on a UV transilluminator. DNA concentration and quality ($A_{260/230}$ and $A_{260/280}$) were measured using the NanoDrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA).

2.3.2 Polymerase Chain Reaction (PCR)

2.3.2.1 Standard PCR

All PCR reactions in this work were performed in 0.2 mL eppendorf tubes with the Mastercycler Gradient PCR machine (Eppendorf, Hamburg, Germany). PCR mixtures contained 1 x thermopolymerase buffer [2 mM Tris-HCl pH 8.8, 1 mM KCl, 1 mM $(NH_4)_2SO_4$, 0.2 mM $MgSO_4$, 0.01% (v/v) Triton X-100, New England Biolabs, Hertfordshire, United Kingdom], 100 µM of each dNTP, 0.5 µM of both the forward and reverse primer, 0.025 U μL^{-1} *Taq* DNA polymerase (New England Biolabs, Hertfordshire, United Kingdom) and 2.5-10 ng μL^{-1} total DNA in a total volume of 20 µL. The PCR program consisted of an initial denaturation at 95.0°C for 2 min, then 35 cycles of denaturation at 95.0°C for 30 s, annealing of primers at 60.0°C for 30 s, and elongation at 68.0°C for 1 min and ending with a final elongation at 68.0°C for 2 min. The primers used and their characteristics are shown in Table 2.2.

Table 2.2. Standard PCR primers, characteristics and expected amplicon length

Primer	Transgene	Primer sequence (5'-3')	T_m (°C)[a]	GC (%)	Amplicon length (bp)
Luc+L	*luc*$^+$	GGAGAGCAACTGCATAAGGC	59.98	55.00	1070
LucR		CCGGTATCCAGATCCACAAC	60.20	55.00	
NeoSHR	*neo*	GGCCATTTTCCACCATGATA	60.53	45.00	554
Neo5		GAGGCTATTCGGCTATGACTG	58.97	52.38	
35SL	pe35S-*luc*$^+$	AATATCGGGAAACCTCCTCG	60.28	50.00	739-412[b]
Luc-R3		TCTTCCAGCGGATAGAATGG	60.17	50.00	110[c]
35SLRT1	pe35S-*luc*$^+$	GACCATGGGGATTGAACAAG	60.17	50.00	

[a]The melting temperature (T_m) of a primer was calculated with the primer3 software available at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).

[b]Amplicon lengths if the enhanced 35S promoter (pe35S) was relocated during T-DNA integration upstream of the *luc*$^+$ in the correct orientation for driving the transcription of the *luc*$^+$. The 35SL primer anneals twice in pe35S (positions -649 and -322 relative to the transcription start site). The amplicon sizes of 739 and 412 bp are expected if the pe35S up to the position +1 (transcription start site) is fused to the theoretical cleavage site of the T-DNA RB.

[c]Amplicon length if the pe35S was relocated during T-DNA integration upstream of the *luc*$^+$ in the correct orientation for driving the transcription of the *luc*$^+$. The 35SLRT1 primer anneals downstream of the pe35S and covers the first 16 bp of the *neo* ORF in promoter tagging vector pETKUL2. The amplicon size of 110 bp is expected if the pe35S and the first 16 bp of the *neo* is fused to the theoretical cleavage site of the T-DNA RB.

Two forward primers 35SL and 35SLRT1 (Table 2.2) annealing in the enhanced 35S promoter and at the 5' end of the *neo* selectable marker gene in pETKUL2, respectively, were used independently with a reverse primer (Luc-R3) annealing to the *luc*$^+$ transgene for the detection of a specific T-DNA rearrangement whereby the enhanced 35S promoter was relocated during T-DNA integration upstream of the *luc*$^+$ in the correct orientation for driving the transcription of the *luc*$^+$. The PCR program was as described above but with an annealing temperature of 60°C or 55°C and with an elongation step of 1.5 min.

Detection of amplicons by agarose gel electrophoresis was performed as described in paragraph 2.3.1

2.3.2.2 Linking PCR

To verify sequence continuity of right and left T-DNA border flanking sequences (5' and 3' regions, respectively) specific primers annealing at the 5' region (F-link) and at the 3' region (R-link) were designed for these sequences (Figure 2.4, Table 2.3). The PCR mix and program were as described in paragraph 2.3.2.1.

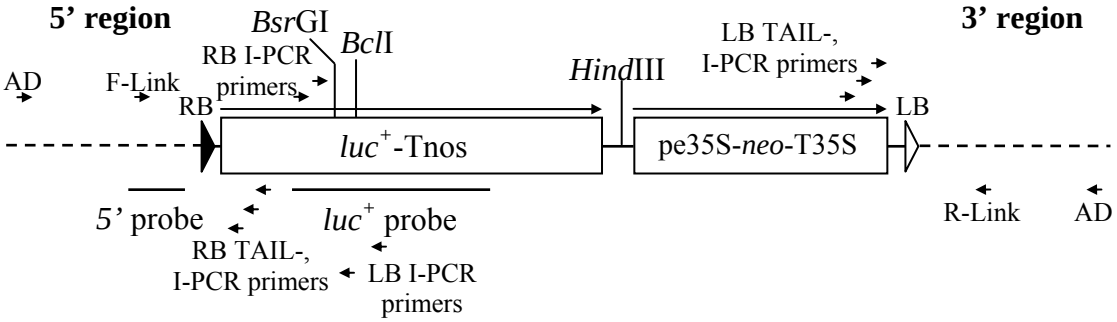

Figure 2.4. Schematic representation of T-DNA insertion of promoter tagging vector pETKUL2 with probe (thick lines), restriction enzyme and primer (short arrows) positions for Southern hybridization and PCR analyses (see also Figure 2.1). The position of the codon-optimized luciferase (*luc*$^+$) and neomycin phosphotransferase II (*neo*) gene cassettes are shown with respect to the right (black triangle, RB) and left (open triangle, LB) T-DNA border. Long arrows mark the direction of transcription. Dotted lines represent plant genomic DNA flanking the right and left T-DNA border, denominated 5' and 3' region, respectively. AD, arbitrary degenerated primer for TAIL-PCR. F-link, forward primer used for linking PCR. R-link, reverse primer used for linking PCR. The restriction sites of the enzymes used for I-PCR is depicted (*Bsr*GI and *Bcl*I) as well for the primers for I- and/or TAIL-PCR. The drawing is not precisely according to scale.

Table 2.3. Linking PCR primers, characteristics and expected amplicon length

ET2 Line-seq[a]	Primer	Primer sequence (5'-3')	T_m (°C)[b]	GC (%)	Amplicon length (bp)
17-1	17-1ProF	GATCGGAACTCTCATTGATCG	59.64	47.62	958
	17-1LBR	AATGCTGGCTTCGCATAGAT	59.84	45.00	
17-2	17-14RT2	AGAATTCTCATGCGCGTTG	59.96	47.37	645
	17-1LB1	ACCACTGTATGGAGCCAAGC	60.14	55.00	
17-3	17-3ProF	CATGGGAGCAAAGTAAGAGG	57.43	50.00	1007
	17-3LBR	AACATTAGCGCTTGCTGTCA	59.64	45.00	
17-4	17-4RT1	TCCAAGTGCAACATTTAAAACC	59.03	36.36	580
	17-2LB1	TTTGTGAATCCCGTTCTTGTC	59.96	42.86	
34-1	34-RTF	TGAATGATGGTTTTCCTGCTC	60.07	42.86	742
	34-R-LB	GGCGGATCAAAGTTCGATTA	60.04	45.00	
85-1	85-1RT3	AACAGCTTGGGTGCATGTG	60.74	52.63	493
	85-1LB	CCTTCTTCCCTGTGGATTGA	60.04	50.00	
85-2	85-2RT3	GGGCAAGAGCAATTCTAGC	57.65	52.63	563
	85-2LB	GGCGATCCCTTTCTCTCTCT	59.92	55.00	

[a]pETKUL2 tagged line number - candidate promoter sequence number.
[b]The melting temperature (T_m) of a primer was calculated with the primer3 software available at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).

2.3.3 Southern hybridization

Southern hybridization analysis was essentially performed as described by Remy (2000). Ten microgram of total DNA per sample was digested overnight at 37°C using 60 units of *Hind*III (New England Biolabs, Hertfordshire, United Kingdom) in a volume of 100 µL containing 10 µL of 10 x NEB2 buffer [10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreïtol (DTT), pH 7.9, New England Biolabs, Hertfordshire, United Kingdom]. Digestion was stopped by placing the samples in an oven at 70°C for 15 min and the reaction mixture was finally vacuum concentrated for 30-40 min to approximately 30 µL. Following the addition of 6 x loading buffer (paragraph 2.3.1) to each sample the restricted fragments were separated by gel electrophoresis in a 0.8% (w/v) ultrapure agarose molecular grade (GIBCO-BRL Life Technologies, Merelbeke, Belgium) gel in 1 x TAE buffer (paragraph 2.3.1) for 5 h at 40 V. Sixty nanograms digoxigenin (DIG)-labelled DNA molecular marker III (Roche, Vilvoorde, Belgium) was also loaded on the gel together with one copy (77.5 pg) and five copy (387.5 pg) reconstructions of *Hind*III digested pETKUL2 (paragraph 2.1.2). A picture of the gel was taken as described in paragraph 2.3.1. The rest of the protocol including stripping of membranes for rehybridization was done according to Remy (2000). DIG-labelled probes were prepared by PCR using the PCR program described in paragraph 2.3.2.1 and the following components: 1 x thermopolymerase buffer [2 mM Tris-HCl pH 8.8, 1 mM KCl, 1 mM $(NH_4)_2SO_4$, 0.2 mM $MgSO_4$, 0.01% (v/v) Triton X-100, New England Biolabs, Hertfordshire, United Kingdom], 1 x DIG labeling mix (200 µM dATP, 200 µM dCTP, 200 µM dGTP, 190 µM dTTP, 10 µM DIG-11-dUTP, pH 7.0, Roche, Vilvoorde, Belgium), 1 µM of each primer, 0.04 U μL^{-1} *Taq* DNA polymerase (New England Biolabs, Hertfordshire, United Kingdom) and 50-200 ng μL^{-1} total DNA in a total volume of 25 µL. First, a probe specific for the *luc*$^+$ transgene was used and following stripping of the membrane, hybridization was done with one or more probe(s) specific for the candidate promoter sequence(s). The characteristics of all primers to prepare the DIG labeled probes are shown in Table 2.4. Labeling and concentration of the probes were estimated by agarose gel electrophoresis employing a 100 bp reference ladder (Eurogentec, Seraing, Belgium) as described (paragraph 2.3.1).

Table 2.4. Primers and their characteristics for the production of DIG labeled probes

Primer	Seq[a]	Primer sequence (5'-3')	T_m (°C)[b]	GC (%)	Probe length (bp)	Position from RB[c]
LucR	*luc*$^+$	CCGGTATCCAGATCCACAAC	60.20	55.00	863	+1169
LucL2		GTGTTGGGCGCGTTATTTAT	59.86	45.00		+307
17-1ProF	17-1	GATCGGAACTCTCATTGATCG	59.64	47.62	422	-427
17-1R1		AACACGCGCATAGTCCAAGT	60.72	50.00		-6
17-2ProF	17-2	GGCTGATTTCGACTCGTAGC	59.98	55.00	425	-437
17-2R1		TTCCTGCTGCCGTATGAAG	59.96	52.63		-13
17-3ProF	17-3	CATGGGAGCAAAGTAAGAGG	57.43	50.00	435	-522
17-3ProR		AACCAGAACTTGGCTCTAAAGTG	59.01	43.48		-88
17-4ProF	17-4	AAGCCATGCACTTTAAGGAA	57.52	40.00	165	-237
17-4ProR		TGGTTTTAAATGTTGCACTTGG	59.95	36.36		-73
17-5F	17-5	CCAATTCACTGTTCCTTGCAT	59.99	42.86	302	-357
TPL575		ATACAGGCAGCCCATCAGTC	60.10	55.00		-56
34-1F	34-1	TTGCTGCATAGAGGTCATGG	59.82	50.00	1216 and 828[d]	-1244 and -856
34-1R		ATTCACTGAGGCATGAGCAA	59.40	45.00		-29
TAILRBluc1	*luc*$^+$	ATAGCTTCTGCCAACCGAAC	59.34	50.00	268[e]	+235
34-RTF	34-1	TGAATGATGGTTTTCCTGCTC	60.07	42.86		-33

[a]Transgene or tagged sequence where the primers anneal with pETKUL2 tagged line number-candidate promoter sequence number.
[b]The melting temperature (T_m) of a primer was calculated with the Primer3 software at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).
[c]The position of the 5' nucleotide of the primer from the RB junction of the T-DNA insertion with positive (+) positions downstream of the insertion site in the T-DNA sequence and negative (-) positions upstream of the insertion site in the banana sequence except for sequence 17-5 where the tagged sequence was not from banana origin but from the vector backbone sequence of pETKUL2.
[d]Primer 34-1F anneals twice because of the presence of a 306 bp direct repeat sequence and is resulting in two amplicons of different size.
[e]Corresponds to 235 bp of the T-DNA near the RB and 33 bp of the T-DNA RB flanking sequence in line ET2-34.

2.3.4 Isolation of T-DNA flanking sequences

2.3.4.1 Thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR)

TAIL-PCR was performed according to Liu *et al*. (1995) with some modifications. The standard mixtures are shown in Table 2.5 for the three TAIL-PCR reactions. For sequence verification, TAIL-PCR was occasionally repeated using high fidelity polymerases that require the preparation of two master mixtures to avoid degradation of the primers by the 3'→5' exonuclease activity of the proofreading polymerase. High fidelity TAIL-PCR reactions were performed using the Extensor Hi-Fidelity PCR Enzyme Mix (ABgene, Epsom, UK) or the Expand High Fidelity PCR System (Roche, Vilvoorde, Belgium) with final polymerase concentrations of 0.05 U μL^{-1} or 0.026 U μL^{-1}, 0.032 U μL^{-1} or 0.017 U μL^{-1} and 0.015 U μL^{-1} for the primary, secondary and tertiary reactions, respectively. These high fidelity master mixtures also contain the other PCR components at the final concentrations as shown in Table 2.5. For each polymerase the corresponding buffer was used. Upon preparing the standard TAIL-PCR mixtures or combining the high fidelity

TAIL-PCR master mixtures, samples were kept on ice until the PCR program was initiated and the initial denaturation temperature was reached.

Table 2.5. Standard TAIL-PCR mixtures

	Reaction					
	Primary		Secondary		Tertiary	
Component (conc.)	**Vol. (µL)**	**Final conc.**	**Vol. (µL)**	**Final conc.**	**Vol. (µL)**	**Final conc.**
PCR buffer (10 x)[a]	2.0	1 x	2.5	1 x	5.0	1 x
dNTPs (2 mM)	1.0	100 µM[b]	1.25	100 µM[b]	2.5	100 µM[b]
AD primer (20 µM)[c]	2.0	2 µM	2.5	2 µM	5.0	2 µM
SP primer (20 µM)[d]	0.15	0.15 µM	0.25	0.20 µM	0.5	0.20 µM
Taq polymerase (5 U μL^{-1})[e]	0.20	0.05 U μL^{-1}	0.16	0.032 U μL^{-1}	0.3	0.030 U μL^{-1}
MQ H_2O	12.65		17.34		35.7	
DNA (25-100 ng μL^{-1})	2.0	2.5-10 ng μL^{-1}	1.0[f]	1000 x	1.0[f]	1000 x
Total volume	20		25		50	

[a]Thermopolymerase buffer [20 mM Tris-HCl pH 8.8, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100; New England Biolabs, Hertfordshire, United Kingdom].
[b]Occasionally the concentration of dNTPs was raised to 200 µM.
[c]Arbitrary degenerated primer.
[d]Specific primer.
[e]Occassionally the concentration of the *Taq* polymerase (New England Biolabs, Hertfordshire, United Kingdom) was halved.
[f]One µL of a 40 and 20 times dilution of the previous reaction was included in the secondary and tertiary reaction, respectively, resulting in a one thousand times dilution of the previous reaction.

All TAIL-PCR primers are shown in Table 2.6. Six different degenerated primers were used independently: AD1, AD2, AD4, AD2-1, AD2-4 and AD2-5. T-DNA specific primers for the isolation of T-DNA RB flanking sequences were TAILRBLUC1, TAILRBLUC2, and Luc-R3 for the primary, secondary and tertiary TAIL-PCR reaction, respectively. For the recovery of T-DNA LB flanking sequences, the corresponding primers were TAILLBpET2n1, TAILLBpET2n2, and TAILLBpET2n3, respectively. A schematic overview of the annealing sites of the primers is shown in Figure 2.4. The PCR programs for the three TAIL-PCR reactions are shown in Table 2.7.

Table 2.6. TAIL- and I-PCR primers and their characteristics

Primer[a]	Primer sequence (5'-3')[b]	Deg[c]	T_m (°C)[d]	GC (%)	Border[e]	Reaction step	Application	Reference
AD1	SCACNTCSTNGTNTCT	256	46-53	44-63	RB, LB	I, II, III	TAIL-PCR	
AD2	NGTCGASWGANAWGAA	128	43-48	38-50	RB, LB	I, II, III	TAIL-PCR	Liu *et al.* 1995
AD4	NGTCGASWGANAAGAA	64	43-48	38-50	RB, LB	I, II, III	TAIL-PCR	
AD2-1	NTCGTSWGANAWGTT	128	38-44	33-47	RB, LB	I, II, III	TAIL-PCR	Qin *et al.* 2003
AD2-4	NGTCGASWCANTWCTA	128	43-48	38-50	RB, LB	I, II, III	TAIL-PCR	Qin *et al.* 2003
AD2-5	NGTCGASWCTNAWCAA	128	43-48	38-50	RB, LB	I, II, III	TAIL-PCR	Qin *et al.* 2003
TAILRBluc1	ATAGCTTCTGCCAACCGAAC	NA[f]	59.34	50	RB	I	TAIL-, I-PCR	Remy *et al.* 2005
TAILRBluc2	TCCACCTCGATATGTGCATC	NA	59.48	50	RB	II	TAIL-, I-PCR	Remy *et al.* 2005
Luc-R3	TCTTCCAGCGGATAGAATGG	NA	60.17	50	RB	III	TAIL-, I-PCR	Remy *et al.* 2005
TAILLBpET2n1	TTCTTCTGAGCGGGACTCTG	NA	60.67	55	LB	I	TAIL-, I-PCR	
TAILLBpET2n2	GGTTTCGCTCATGTGTTGAG	NA	59.29	50	LB	II	TAIL-, I-PCR	Remy *et al.* 2005
TAILLBpET2n3	TTAAAAACGTCCGCAATGTG	NA	59.6	40	LB	III	TAIL-, I-PCR	Remy *et al.* 2005
TAILLBpET2n4	CAGATCCCCCGAATTAATTC	NA	58.32	45	LB	II	I-PCR	Remy *et al.* 2005
LucL2	GTGTTGGGCGCGTTATTTAT	NA	59.86	45	RB	I	I-PCR	
LucL3	CTACCGTGGTGTTCGTTTCC	NA	60.41	55	RB	II, III	I-PCR	
Luc+R	AGAATCTCACGCAGGCAGTT	NA	60.02	50	LB	I, II, III[g]	I-PCR	
LucR5	GGACTCTGGCACAAAATCGT	NA	60.12	50	LB	II, III[h]	I-PCR	

[a]Arbitrary degenerated (AD) and specific primers.
[b]S refers to G or C; N refers to A, T, C, or G and W refers to A or T.
[c]Level of degeneracy.
[d]The melting temperature (T_m) of the AD primers was calculated by the Oligo Calculator version 3.09 available at www.basic.northwestern.edu/biotools/oligocalc.html using the formula T_m= 100.5 + (41 * (yG+zC)/(wA+xT+yG+zC)) - (820/(wA+xT+yG+zC)) + 16.6*log10([Na^+]). The T_m of the specific primers was calculated by Primer3 software at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).
[e]Primer used for the isolation of the mentioned T-DNA border flanking sequence. RB refers to right T-DNA border and LB to left T-DNA border.
[f]Not applicable.
[g]The primer Luc+R was only used in the primary I-PCR reaction following *Bsr*GI restriction, and in the primary, secondary and tertiary reaction following *Bcl*I restriction.
[h]The primer LucR5 was used in the secondary and tertiary reactions following *Bsr*GI restriction.

Table 2.7. TAIL-PCR program

Reaction	Step	Temperature (°C)	Time (min:s)	Cycles
Primary	Initial denaturation	95	2:00	
	DNA denaturation	95	1:00	
	Annealing of specific primers	64	1:00	5
	Elongation	68	3:30	
	DNA denaturation	95	1:00	
	Annealing of primers	27.5 → 72 [a]	3:00	
	Elongation	68	3:30	
	DNA denaturation	95	0:30	
	Annealing of specific primers	64	1:00	2
	Elongation	68	3:30	
	DNA denaturation	95	0:30	15
	Annealing of AD primers	42	1:00	
	Elongation	68	3:30	
	Final elongation	68	5:00	
	Reaction stop	4	hold	
Secondary and tertiary	DNA denaturation	95	0:30	
	Annealing of specific primers	64	1:00	2
	Elongation	68	3:30	
	DNA denaturation	95	0:30	12/10[b]
	Annealing of AD primers	42	1:00	
	Elongation	68	3:30	
	Final elongation	68	5:00-10:00[c]	
	Reaction stop	4	hold	

[a]Temperature increase from 27.5°C till 72°C with an increment of 0.3°C/s.
[b]For the secondary and tertiary TAIL-PCR reactions the number of supercycles (combination of two high-stringency and one low-stringency cycle) was 12 and 10, respectively.
[c]Final elongation step was 5 to 10 min.

Secondary and tertiary TAIL-PCR products (10 µL) were loaded usually next to each other on agarose gel to check for specificity by known size shifts, together with a 1 kb reference ladder (Eurogentec, Seraing, Belgium). Agarose gel electrophoresis was performed as described in paragraph 2.3.1.

2.3.4.2 Inverse PCR (I-PCR)

In a volume of 100 µL, 250 to 315 ng of total DNA was digested for 2 h at 37°C with *Bsr*GI or *Bcl*I (New England Biolabs, Hertfordshire, United Kingdom). A schematic representation of the restriction sites of the enzymes used for I-PCR is shown in Figure 2.4. One hundred µL of MQ water and 200 µL of phenol:chloroform:isoamylalcohol (25:24:1) were then added, the mixture was vortexed for five seconds and centrifuged (5 min at 13,000 rpm and 4°C). The upper aqueous phase was transferred to a new eppendorf tube and 100 µL of 10 M ammonium acetate, 3 µL of 20 mg mL^{-1} glycogen and 700 µL of 100% ethanol were added. The mixture was vortexed, incubated at -80°C for one hour and centrifuged (40 min at 13,000 rpm and 4°C). The supernatant was removed and the pellet was washed with 70% ethanol, dried in a laminar flow or by vacuum drying for 10 min and finally dissolved in 10 µL MQ water.

The digested DNA was then prepared for self-ligation by adding 87 µL MQ water, 1 µL T4 DNA ligase (400,000 U mL^{-1}; New England Biolabs, Hertfordshire, United Kingdom) and 10 µL 10 x ligation buffer (500 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 100 mM DTT, 10mM ATP, 250 µg mL^{-1} BSA) to 2 µL of digested DNA. Following overnight incubation at 16°C, the reaction was stopped and purified by phenol:chloroform extraction as described above.

Two to three nested or semi-nested PCR reactions were performed. In the primary I-PCR reaction the components were combined in the following order and final concentrations in a volume of 25 or 50 µL: 1 x thermopolymerase buffer [2 mM Tris-HCl pH 8.8, 1 mM KCl, 1 mM $(NH_4)_2SO_4$, 0.2 mM $MgSO_4$, 0.01% (v/v) Triton X-100; New England Biolabs, Hertfordshire, United Kingdom], 200 µM of each dNTP, 0.2 µM of both the forward and reverse primer, 0.026 U μL^{-1} *Taq* DNA polymerase (New England Biolabs, Hertfordshire, United Kingdom) and finally 2 µL self-ligated DNA fragments. Reaction mixtures in the secondary and tertiary PCR steps were identical except that 1 µL of a 40 x or 20 x dilution of the previous step reaction was used as template in a final volume of 25 µL or 50 µL, respectively. Similar to TAIL-PCR (paragraph 2.3.4.1), I-PCR was sometimes repeated using the Extensor Hi-Fidelity PCR Enzyme Mix (ABgene, Epsom, UK) or the Expand High Fidelity PCR System (Roche, Vilvoorde, Belgium). A final high fidelity polymerase concentration of 0.025 U μL^{-1} was employed in the three I-PCR reactions with the same final concentrations of the other components of the PCR mixture.

For the primary, secondary and tertiary I-PCR the primer combinations for the isolation of T-DNA RB flanking sequences were LucL2/TAILRBluc1, LucL3/TAILRBluc2 and LucL3/Luc-R3, respectively (Table 2.6). The corresponding primer sets for the isolation of the T-DNA LB flanking sequences the primers used were Luc+R/TAILLBpET2n1, LucR5/TAILLBpET2n4 and LucR5/TAILLBpET2n3 respectively (Table 2.6). In some instances, primer TAILpET2n4 was replaced by TAILpET2n2. When using the *Bcl*I enzyme the primer LucR5 could not be used because it anneals beyond the restriction site, thus the primer Luc+R was used instead. A schematic representation of the annealing sites of the primers used for I-PCR is shown in Figure 2.4.

To improve specificity, a touchdown PCR program was performed for all three I-PCR reactions with an initial denaturation step of 95°C for 2 min and four steps of 3 cycles each consisting of 95°C for 15 s; 68°C, 66°C, 64°C and 62°C for 20 s for each step, respectively; and 68°C for 3.5 min followed by 30 cycles of 95°C for 15 s, 60°C for 20 s, and 68°C for 3.5 min, with a final elongation step of 68°C for 5 min.

PCR products were run on agarose gel and visualized as described (paragraph 2.3.1).

2.3.5 Reverse transcriptase (RT)-PCR

2.3.5.1 RNA isolation

Sixty to 100 mg of leaf, pseudostem, corm and root tissues from *in vitro* plantlets or proliferating meristem cultures initiated from *in vitro* apical meristems (paragraph 2.1.3.3) were crushed using the RETSCH MM301 Mixer Mill (Retsch Inc., Haan, Germany) as described above (paragraph 2.3.1). Total RNA was extracted using the RNeasy® Total RNA Isolation Kit (Qiagen, Hilden, Germany) according to the instructions of the manufacturer. The RLT lysis buffer was supplemented with β–mercaptoethanol (10 µL mL^{-1}) and the RNA was finally eluted from the spin column with 50 µL RNase free MQ water. RNA concentration and quality ($A_{260/230}$ and $A_{260/280}$) was assessed using the NanoDrop® ND-1000 spectrophotometer (Nanodrop technologies, Wilmington, DE, USA).

2.3.5.2 First Strand cDNA synthesis

First strand cDNA synthesis was performed using the RevertAid^{TM} H Minus First Strand cDNA Synthesis Kit (Fermentas, St-Leon Rot, Germany). One microgram of total RNA was prepared for cDNA synthesis by adding 1 µL of oligo$(dT)_{18}$ primer (0.5 µg μL^{-1}) or 0.5-1 µL of the specific primer Luc+R (20 µM; Table 2.6) and RNase free MQ water until a total volume of 12 µL. After adding the remaining components according to the manufacturer's instructions, the samples were incubated at 42°C or 45°C for one hour and the reaction was stopped by heating at 70°C for 10 min. The reaction was placed on ice for immediate use in RT-PCR or stored at -20°C.

2.3.5.3 RT-PCR reaction

All RT-PCR reactions were performed in a total volume of 20 µL. Two or 10 µL of 1 x or 5 x diluted first strand cDNA reaction, respectively, was mixed with 18 or 10 µL master mix consisting of 1 x thermopolymerase buffer [2 mM Tris-HCl pH 8.8, 1 mM KCl, 1 mM $(NH_4)_2SO_4$, 0.2 mM $MgSO_4$, 0.01% (v/v) Triton X-100; New England Biolabs, Hertfordshire, United Kingdom], 200 µM of each dNTP, 0.5 µM of both the forward and reverse primer and 0.025 U μL^{-1} *Taq* DNA polymerase (New England Biolabs, Hertfordshire, United Kingdom). To exclude false positive results due to the presence of DNA, PCR was performed on the corresponding RNA sample. As positive control, PCR was done on DNA in parallel together with at least two different negative controls i.e. water and RNA and/or cDNA from a non-transformed control plant.

The primers for the housekeeping *actin*1 gene (*act*, GeneBank No. AF285176) and *luc*$^{+}$ transgene are indicated in Table 2.8. The *act* primers ActinF3 and ActinR2 span an intron and hence, allow checking for gDNA contamination. Different *luc*$^{+}$ primer pairs were used to confirm LUC expression as detected in the LUC screening assays. To identify

the banana promoter sequence responsible for the LUC activation, RT-PCR was performed with a candidate promoter specific forward primer (Table 2.9) in combination with the *luc*$^+$ specific reverse primer TAILRBluc1 (Table 2.6).

Table 2.8. RT-PCR primers and their characteristics for the *actin*1 and *luc*$^+$ genes

Gene	Primer	Primer sequence (5'-3')	T_m (°C)[a]	GC (%)	Amplicon size (bp)
*actin*1	ActinF3	CCCAAGGCAAACCGAGAGAAG	65.39	57.14	225[b]-150[c]
	ActinR2	GTGGCTCACACCATCACCAG	62.65	60.00	
luc$^+$	LucR5	GGACTCTGGCACAAAATCGT	60.12	50.00	282
	LucL2	GTGTTGGGCGCGTTATTTAT	59.86	45.00	

[a]The melting temperature (T_m) of the primers was calculated by Primer3 software at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).
[b]Expected size of the amplicon when genomic DNA was used as template.
[c]Expected size of the amplicon when cDNA was used as template.

The following PCR program was initiated: a first denaturation at 95.0°C for 2 min, then 35 cycles of denaturation at 95.0°C for 30 s, annealing of primers at 60.0°C for 30 s, elongation at 68.0°C for 30 s or 1 min, and ending with a final elongation at 68.0°C for 2 min. Loading buffer at a final concentration of 1 x (paragraph 2.3.1) was added to 10 or 20 µL RT-PCR product and the mixture was run and visualized on agarose gel as described (paragraph 2.3.1).

For semi-quantitative RT-PCR analysis, 10 µL of 5 x diluted first strand cDNA reaction mixture was used as template in a total volume of 20 µL as described above and replicate samples were removed at 18, 22, 26, 30, 35; 20, 25, 30 and 35 cycles or at 25, 30, 35 and 40 cycles. The above described PCR program was applied except that elongation and final elongation were executed at 72°C for 30 s and 72°C for 2 min, respectively. Consequently, semi-quantitative RT-PCR analysis as presented in this work refers to non-competitive RT-PCR.

Table 2.9. RT-PCR primers and their characteristics to detect transcriptional fusion between candidate promoter sequences and the *luc*$^+$ transgene

Promoter specific primer	ET2 Line-seq[a]	Primer sequence (5'-3')	T_m (°C)[b]	GC (%)	Size (bp)[c] TAIL RBluc1	Size (bp)[c] TAIL RBluc2
17-RT-1	17-1	GCCAGGAAACATGACACTTG	59.14	50.00	273	222
17-RT-2	17-2	TCATACGGCAGCAGGAAAG	59.96	52.63	258	207
17-RT-3	17-3	ACATGATGTTTCCTAATACATATCTGA	57.31	29.63	302	251
17-RT-4	17-4	TCCAAGTGCAACATTTAAAACC	59.03	36.36	328	277
34-RT-1	34-1	TGAATGATGGTTTTCCTGCTC	60.07	42.86	268	217
85-RT-1	85-1	AACAGCTTGGGTGCATGTG	60.74	52.63	256	205
85-RT-2	85-2	GGGCAAGAGCAATTCTAGC	57.65	52.63	270	219
85-RT-3	85-3	TTTATACATCCAAAACATTGATTGA	57.61	24.00	270	219
156-RT-2	156-2	TTAGACATGTCGACCGAATGA	59.14	42.86	256	205
156-RT-3[d]	156-3	ATTGCACGCAGGTTCTCC	59.79	55.56	295	244
156-RT-4	156-4	AATCAACGACACCGAACTCTG	60.16	47.62	289	238

[a]pETKUL2 tagged line number-candidate promoter sequence number.
[b]The melting temperature (T_m) of a primer was calculated with the Primer3 software at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi (Rozen and Skaletsky 2000).
[c]Amplicon size when using the *luc*$^+$ primers TAILRBluc1 or TAILRBluc2.
[d]ET2-line 156 contains a T-DNA flanking sequence with part of the pe35S and the *neo* selectable marker gene to which primer 156-RT-3 anneals (positions +21 to +38 relative to the transcription start site of the *neo*).

2.3.6 Cloning of T-DNA flanking sequences

2.3.6.1 DNA Gel extraction

Thirty-six to 120 µL mixtures (30-100 µL PCR sample mixed with 6-20 µL 6 x loading buffer) were loaded together with a 1 kb reference ladder (Eurogentec, Seraing, Belgium) on a 0.8-1.2% (w/v) agarose gel in the presence of 1 x TAE or 1 x SB buffer (paragraph 2.3.1). Amplicons were separated at 90 V for 100-140 min or at 110-220 V for 110-40 min, respectively. Occasionally, the electrophoresis buffer contained 1 mM guanosine to protect the amplicons from UV breakdown (Grundemann and Schomig 1996). The gels were stained for 30 min with ethidium bromide at a final concentration of 0.5 µg mL^{-1} or 1 µg mL^{-1} when using guanosine, and destained in dH_2O for 30 min. The amplicons were then excised using a scalpel while illuminating the gel with a UV transilluminator (312 nm) or a portable UV lamp (254 nm) for a few seconds. The weight of the excised gel piece was determined and the DNA was extracted with the NucleoSpin® Extract II kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. The DNA was eluted in two steps with 50 and 30 µL of prewarmed (70°C) MQ water and each time centrifuged (11,000 rpm for 1 min at room temperature) after a one min incubation at room temperature. Then, samples were vacuum concentrated at 40°C to a final volume of 2-10 µL.

2.3.6.2 TA cloning and transformation

The gel-extracted PCR products were cloned into the pCR®4-TOPO® vector using the TOPO TA Cloning Kit (Invitrogen, Merelbeke, Belgium). The ligation reaction and transformation and culture of *E. coli* TOP10® cells was performed according to the manufacturer's instructions except for the volumes recommended for the ligation reaction that were halved. The transformation mixture was plated on LB medium (per L: 10 g bacto trypton, 5 g bacto yeast extract, 10 g NaCl, 15 g bacteriological agar, pH 7.0) containing 50 μg mL^{-1} of kanamycin for selection of transformed colonies. The plates were incubated overnight upside down at 37°C and stored at 4°C until further analysis.

2.3.6.3 Analysis of the colonies and sequencing

To check for successful cloning, colony PCR was performed on transformed *E. coli* TOP10® cells. The following PCR master mix with a total volume of 19 μL was prepared: 1 x thermopolymerase buffer [2 mM Tris-HCl pH 8.8, 1 mM KCl, 1 mM $(NH_4)_2SO_4$, 0.2 mM $MgSO_4$, 0.01% (v/v) Triton X-100, New England Biolabs, Hertfordshire, United Kingdom], 200 μM of each dNTP, 0.2 μM of both the forward and reverse primer, 0.0125 U μL^{-1} *Taq* DNA polymerase (New England Biolabs, Hertfordshire, United Kingdom). After picking the colony with a sterile 20 μL pipet tip, a small amount of bacterial cells was suspended in the PCR master mix and with the same tip a 200 μL LB culture containing 50 mg L^{-1} kanamycin was inoculated and incubated at 37°C. At least five independent colonies per cloning reaction were tested.

The forward T7 primer (5'-TAATACGACTCACTATAGGG-3') and the reverse pUCR primer (5'-CAGGAAACAGCTATGAC-3') which anneal up- and downstream of the cloning site in pCR®4-TOPO® were used for colony PCR. The PCR program consisted of an initial denaturation at 95.0°C for 10 min, then 35 cycles of denaturation at 95.0°C for 30 s, annealing of primers at 60.0°C for 30 s, and elongation at 68.0°C for 1 or 3 min and ending with a final elongation at 68.0°C for 2 min.

A larger selective LB culture of 6 mL was initiated for the positive colonies using the above mentioned 200 μL culture as inoculum. After overnight incubation at 37°C and 200 rpm, plasmid DNA was isolated from 4.5 mL culture with the Qiaprep® plasmid miniprep kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and the DNA was diluted with prewarmed (70°C) MQ water. The concentration and quality of the DNA was measured with the NanoDrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). The remaining 1.5 mL culture was used to establish glycerol stocks for storage at -80°C.

The isolated plasmids were sent for commercial sequencing at the Genetic Service Facility of the VIB (University of Antwerp, Wilrijk, Belgium). For samples containing

amplicons smaller than 900 bp, 10 µL of 50 ng μL^{-1} plasmid DNA was prepared, whereas 20 µL of 50 ng μL^{-1} plasmid DNA was made for larger amplicons.

2.3.7 *In silico* analysis of T-DNA flanking sequences

The sequences were analyzed with the blast 2 sequences software (www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi; Tatusova and Madden 1999) searching for the specific primer annealing regions in the T-DNA. Sequences which showed no homology to the T-DNA of pETKUL2 were catalogued as non-specific.

The banana flanking sequences were analyzed with blastn and blastx (www.ncbi.nlm.nih.gov/BLAST; Altschul *et al.* 1997) programs in the GeneBank database using "nr" (all GenBank + RefSeq Nucleotides + EMBL + DDBJ + PDB sequences, excluding HTGS0,1,2, EST, GSS, STS, PAT, WGS) and "est" (database of GenBank + EMBL + DDBJ sequences from EST Divisions) settings, and with the fasta3 program (www.ebi.ac.uk/fasta33/nucleotide.html; Pearson and Lipman 1988). The sequences were also analyzed using the blast software in a banana EST database donated by Syngenta to the Global *Musa* Genomics Consortium (www.musagenomics.org).

The 5' tagged putative promoter sequences were queried for *cis*-acting elements in the PlantCARE (bioinformatics.psb.ugent.be/webtools/plantcare/html; Lescot *et al*. 2002) and PLACE (www.dna.affrc.go.jp/PLACE; Higo *et al*. 1999) databases. More specific promoter analysis was performed with the TSSP software (www.softberry.com). The *in silico* analyses of the T-DNA flanking sequences was performed in the period between July 2005 and February 2007.

2.4 Cloning and back-transformation of candidate promoter sequences

2.4.1 High fidelity PCR amplification

PCR was performed on ET2 lines using primers specific for the amplification of putative promoter sequences as shown in Table 2.10. Following initial PCR amplification with standard *Taq* Polymerase, candidate promoter sequences were PCR amplified with a high fidelity polymerase mixture (Extensor Hi-Fidelity PCR Enzyme Mix, ABgene, Epsom, UK). A total PCR master mix of 50 µL was obtained by combining two master mixes of 25 µL each that were prepared separately. Master mix 1 contained 400 µM of each dNTP, 1.0 µM of both the forward and reverse primer and 100-200 ng genomic DNA, while master mix 2 consisted of 2 x Extensor Buffer 1 (4.5 mM $MgCl_2$) and 0.06 U μL^{-1} of Extensor Hi-Fidelity PCR Enzyme Mix.

After combining both master mixtures, samples were processed as described (paragraph 2.3.4.1) and the following PCR program was initiated: a first denaturation at 94.0°C for 2 min, then 35 cycles of denaturation at 94.0°C for 30 s, annealing of primers at 60.0°C for 30 s, elongation at 68.0°C for 2 min, and ending with a final elongation at 68.0°C for 10 min.

Detection of amplicons was performed by agarose gel electrophoresis as described in paragraph 2.3.1.

Table 2.10. Primers and their characteristics for the amplification of putative promoter sequences

Primer	ET2 line-seq[a]	Primer sequence (5'-3')[b]	T_m (°C)[c]	GC (%)[d]	Position from RB[e]	Restriction enzyme	Size (bp) of putative promoter sequence	Plasmid
17-1F	17-1	CCGTAGAGATGATCCATGAGC	59,67	52,38	-1742	*Eco*RI	1738	pESKUL1
17-1R1		AACACGCGCATAGTCCAAGT	60.72	50.00	-5			
17-1FSacI	17-1	TAGTT**GAGCTC**CCGTAGAGATGATCCATGAGC	59.67	52.38	-1742	*Eco*RI	1350	pESKUL2
17-1R4AscI		AAGTT**GGCGCGCC**AACCCGAGAGAAACGATCAA	59.67	45.00	-393			
17-2FHindIII	17-2	CGTCGG**AAGCTT**TGAATCGTGAGTGGCGTATG	60.69	50.00	-1210	*Hind*III/ *Bam*HI	1198	pESKUL3
17-2R1BamHI		TACTT**GGATCC**TTCCTGCTGCCGTATGAAG	59.96	52.63	-13			
17-3FHindIII	17-3	GCCGCC**AAGCTT**CCTCAGGACATGAATCATTAGC	59.67	45.45	-686	*Hind*III/ *Bam*HI	681	pESKUL4
17-3R1BamHI		GGCCC**GGATCC**CACTTGTAAAACATTCTACAGTTCAAA	58.1	29.63	-6			
34-1F	34-1	TTGCTGCATAGAGGTCATGG	59.82	50.00	-856, -1244[f]	*Eco*RI	828, 1216[f]	pESKUL5/6[g]
34-1R		ATTCACTGAGGCATGAGCAA	59.4	45.00	-29			
85-1F2PstI	85-1	TAATCGT**CTGCAG**GAACAACCAGGTGAGAGAGAGC	60.44	54.55	-604	*Pst*I/ *Eco*RI	599	pESKUL7
85-1R2		ATGCACCCAAGCTGTTGAAG	61.24	50.00	-6			

[a]pETKUL2 tagged line number - candidate promoter sequence number.

[b]In bold are the restriction sites indicated for cloning of the amplified putative promoter sequences in pCAMBIA 1391Z. The restriction sites are **GAGCTC** for *Sac*I, **GGCGCGCC** for *Asc*I, **AAGCTT** for *Hind*III, **GGATCC** for *Bam*HI and **CTGCAG** for *Pst*I.

[c]The melting temperature (T_m) of a primer was calculated with the Primer3 software at frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi; (Rozen and Skaletsky 2000). For primers containing a restriction site, the T_m was calculated only for the sequence downstream (3' side) of the restriction site.

[d]The GC content was calculated based only on the sequence downstream of the restriction site if applicable.

[e]The position of the 5' nucleotide of the primer from the RB junction of the T-DNA insertion with negative (-) positions upstream of the insertion site in the original ET2 line.

[f]The primer 34-1F anneals twice because of the presence of a direct repeat resulting in two amplicons of different size.

[g]The pESKUL5 and pESKUL6 should harbor the 828 bp and 1216 bp putative promoter sequence, respectively.

2.4.2 Cloning

Amplified candidate promoter sequences were cloned first in the pCR®4-TOPO® vector by adding 2 µL of the PCR product to the ligation reaction using the TOPO TA Cloning Kit (Invitrogen, Merelbeke, Belgium) and then transferred into chemically competent *E. coli* cells as described (paragraph 2.3.6.2). Colony PCR and plasmid isolation was performed as described (paragraph 2.3.6.3). Cloned candidate promoter sequences were then inserted in the multiple cloning site of pCAMBIA 1391Z (www.cambia.org) which is located upstream of the *uid*A^{INT} reporter gene (Figure 2.5). The promoter sequences that contained a different restriction enzyme site at each end of the sequence (Table 2.10) were digested sequentially (first and second digestion in 50 µL and 100 µL, respectively) in parallel with the pCAMBIA 1391Z vector. For the sequences 17-1 and 34-1 a single enzyme digestion was performed with the *Eco*RI enzyme (Table 2.10), which cuts in the flanking regions of the cloning site in the pCR®4-TOPO® vector.

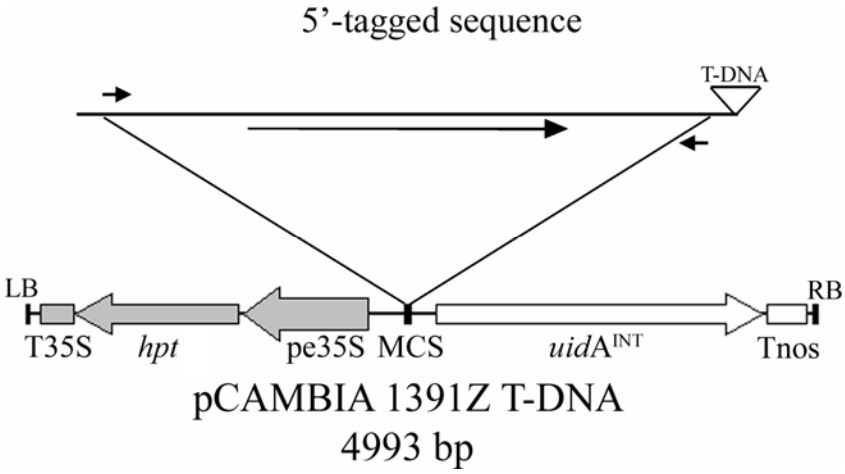

Figure 2.5. T-DNA of vector pCAMBIA 1391Z (www.cambia.org) used for back-transformation. 5'-tagged sequence represents a candidate promoter sequence. Short arrows represent the specific primers used for the amplification of the candidate promoter sequence, while the large arrow indicates the 5' to 3' orientation of this sequence. LB, T-DNA left border; T35S, Cauliflower Mosaic Virus (CaMV) 35S RNA terminator; *hpt*, hygromycin phosphotransferase gene; pe35S, enhanced CaMV 35S RNA promoter; MCS, multiple cloning site; *uid*A^{INT}, *uid*A reporter gene containing the castorbean catalase intron; Tnos, nopaline synthase terminator; RB, T-DNA right border.

Digestions of 5 µg purified pCR®4-TOPO® vector DNA containing the putative promoter sequences and 5-10 µg of pCAMBIA 1391Z with the appropriate enzyme and buffer (NEBEcoRI for *Eco*RI, NEB2 for *Hind*III, NEBBamHI for *Bam*HI and NEB3 for *Pst*I) were performed overnight at 37°C in a total volume of 50 µL and 100 µL, respectively. Reactions were stopped by incubation at 70°C for 20 min. For fragments that needed a second digestion and for the linearized pCAMBIA 1391Z vector, the restriction mix was purified by adding 100 µL and 200 µL of NT buffer to the 50 µL and 100 µL reaction, respectively. The mixture was then applied onto a NucleoSpin® Extract II column

(Macherey-Nagel, Düren, Germany) for centrifugation according to the manufacturer's instructions. Elution was done with 50 µL and 80 µL of prewarmed (70°C) Elution Buffer or MQ, respectively. The second digestion of fragments that were digested sequentially was performed in a total volume of 100 µL and then the reactions were purified by adding 200 µL of NT buffer to the reaction prior adding the samples to the NucleoSpin® Extract II column. Finally, elution was performed by adding 50 µL of prewarmed (70°C) Elution Buffer. Single enzyme restricted pCAMBIA 1391Z vector was dephosphorylated by CIAP (Promega, Leiden, The Netherlands) incubation according to the manufacturer's instructions and purified using the NucleoSpin® Extract II kit (Macherey-Nagel, Düren, Germany) as described above.

The purified restricted candidate promoter sequence was run on a 0.8% (w/v) agarose gel in 1 x TAE buffer containing 1 mM guanosine after adding 10 µL of 6 x loading buffer (paragraph 2.3.1) to 50 µL purified solution. Extraction of the desired promoter fragments from gel was done as described in paragraph 2.3.6.1, but performing only one elution with 50 µL prewarmed (70°C) Elution Buffer.

After measuring the concentrations with the NanoDrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA), the pCAMBIA 1391Z linearized vector and promoter fragments were each time mixed in a 1:1 and 1:3 molar ratio together with 2 µL T4 DNA Ligase Reaction Buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg mL^{-1} BSA; New England Biolabs, Hertfordshire, United Kingdom) and 1 µL T4 DNA Ligase (400 U μL^{-1}; New England Biolabs, Hertfordshire, United Kingdom) in a final volume of 20 µL. Samples were incubated overnight at 16°C. Five µL of the ligation mix was added to 100 µL chemically competent *E. coli* cells (Sambrook *et a*l. 1989). The transformation procedure, colony PCR and plasmid DNA isolation were done as described (paragraphs 2.3.6.2 and 2.3.6.3). To check for the correct orientation of the promoter fragment, colony PCR was performed as described in paragraph 2.3.6.3 using a specific forward primer for each promoter fragment (Table 2.10) with the reverse primer GUS1R (5'-CTTGTAACGCGCTTTCCCACC-3') or GUSIntR1 (5'-TAGAAATTTACCCTCAGATC-3'). The obtained plasmids were commercially sequenced and are indicated in Table 2.10.

2.4.3 *Agrobacterium* transformation

Competent *A. tumefaciens* strain EHA105 cells were electroporated with the different pESKUL plasmids (Table 2.10) as follows. Fifty µL electrocompetent EHA105 cells was added to one µL of pESKUL plasmid (250-600 ng μL^{-1}), incubated for 2 min on ice and transferred to a 2 mm ice-cold cuvette. Electroporation was done at 2.26 kV for approximately 5.93 msec and immediately 800 µL of YM medium (per L: 0.1 g NaCl, 0.2 g $MgSO_4 \cdot 7H_2O$, 0.5 g $K_2HPO_4 \cdot 3H_2O$, 0.4 g yeast extract, 10.0 g mannitol, pH 7.0) was

added. Following gentle transfer to a 15-mL tube the sample was incubated for 3 h at 28°C and 180 rpm on a rotary shaker. Electroporated bacteria were then plated on YM medium [1.3% (w/v) bactoagar] containing 50 µg mL^{-1} kanamycin and incubated at 28°C for 36-48 h upside-down. Single colonies were grown in 20 mL YEP medium (per L: 10 g bactopeptone, 10 g bacto yeast extract, 5 g NaCl, pH 7.5) containing 50 µg mL^{-1} kanamycin for 48 h to establish glycerol stocks for storage at -80°C.

Starting with 17 mL *Agrobacterium* culture and following the standard alkaline lysis plasmid purification protocol (Sambrook *et al.* 1989) and further purification with the Qiaprep Miniprep kit (Qiagen, Hilden, Germany) the pESKUL plasmids were verified by restriction analysis. Five µg plasmid DNA was digested with the enzyme(s) used to clone the different putative promoter fragments into pCAMBIA 1391Z (paragraph 2.4.2).

2.4.4 *Agrobacterium*-mediated back-transformation

Agrobacterium-mediated transformation of ECS of the banana cultivars THP and 'Grand naine' (GN, AAA, ITC.1256) with the candidate promoter-*uid*A^{INT} transcriptional fusions (pESKUL plasmids) called back-transformation and TGA were done as described (paragraph 2.1.3 and 2.1.3.5, respectively). After two to three months of selection on ZZ medium containing 50 mg L^{-1} hygromycin and 200 mg L^{-1} timentin, proliferating cell colonies maintained at 26±2°C were histochemically stained or fluorometrically analysed for GUS activity according to modified procedures of Jefferson *et al.* (1987). Staining was done in X-Gluc solution containing 100 mM Tris-HCl pH 8.0, 1% (w/v) 5-bromo-4-chloro-3-indolyl-β-glucuronic acid dissolved in DMSO, 0.5 mM ferrocyanide, 0.5 mM ferricyanide, 1% (w/v) ascorbic acid, 10 mM EDTA and 2% (w/v) CHAPS detergent (Roche, Vilvoorde, Belgium), while 20% (v/v) methanol was included in the fluorometric assay buffer. Five months after transformation, back-transformed cell colonies and cell cultures on selective ZZ and RD1 medium, respectively, were histochemically stained and fluorometrically analysed for GUS activity at 26°C and immediately (approximately 5 min) after an 8°C temperature period of 12-16 h. To obtain a sufficient amount of material (from 50 to 300 mg fresh weight per sample) for the fluorometric GUS analyses several independent transgenic lines were pooled per construct.

Chapter 3 Promoter tagging during *in vitro* development and low temperature stress

3.1 Introduction

New generations of transgenic plants require a more accurate regulated expression of transferred genes, which calls for the identification and characterization of novel promoters in higher plants. Several methods for promoter isolation are based on mRNA differential expression analysis and usually require prior knowledge of the nucleotide sequence of the candidate gene(s), which is however mostly lacking in banana.

An alternative method for the isolation of promoters in plants is T-DNA tagging, where a promoterless reporter or selectable marker gene integrates into the nuclear genome by *Agrobacterium*-mediated transformation (Teeri *et al.* 1986, André *et al.* 1986). The T-DNA integrates randomly in the genome, although hot spots for integration have been observed in Arabidopsis and rice, mainly in transcribed regions of the genome (Alonso *et al.* 2003; Schneeberger *et al.* 2005; Jeong *et al.* 2006) increasing the probability for the isolation of a large number of promoters by T-DNA tagging. A prerequisite for a successful T-DNA tagging method is an efficient genetic transformation system, which has been established in banana (Pérez Hernández *et al.* 2006b). The main advantages of promoter T-DNA tagging are that promoter sequences are retrieved and their functional analysis to a certain extent are done without prior sequence information, and in their natural environment (original promoter position in the banana genome). These are crucial advantages when dealing with banana because only around 1% of its genome is currently sequenced (bioinfo.inibap.org/statusdb/stats.php?page=clone; 21 March 2008). In addition, few promoters from banana have so far been identified and characterized (Hermann *et al.* 2001a; Wang and Peng 2001a, 2001b; Sunil Kumar *et al.* 2005, Table 1.3). Recently, an efficient promoter T-DNA system has been developed for banana yielding the first generation of constitutive banana promoters (Remy *et al.* 2005). Thus, promoter T-DNA tagging provides valuable information on functional banana regulatory DNA sequences.

Application of the firefly luciferase (*luc*) gene as reporter gene allows real-time detection of the luciferase (LUC) enzyme in a non-invasive and non-destructive manner combined with high sensitivity (Ow *et al.* 1986). Furthermore, the short half-life of LUC activity (Van Leeuwen *et al.* 2000) allows monitoring of dynamic gene expression changes, which makes the *luc* reporter gene ideal for tagging promoters and genes with inducible or developmentally regulated expression (paragraphs 1.3.2 and 1.3.3).

Here, by taking advantage of the *luc* reporter gene features, real-time monitoring of LUC activity was performed in transgenic banana lines transformed with a promoter tagging vector under low temperature (LT) stress during *in vitro* development. Following large scale screening for LUC activity at cell colony stage under baseline (paragraph 3.2.1)

and LT (paragraph 3.2.2) conditions including a detailed quantitative time course analysis (paragraph 3.2.3), LUC activity was monitored throughout the *in vitro* regeneration of a limited number of selected lines (paragraph 3.3). The main objective of the work presented in this chapter was to identify promoter-tagged lines with different responses in LUC activity to LT stress during development from undifferentiated tissue into *in vitro* plants.

3.2 Screening for luciferase activity in large T-DNA tagged cell colony populations

3.2.1 Screening for baseline luciferase activity

3.2.1.1 Baseline luciferase activity frequency and level

Two months after *Agrobacterium*-mediated transformation with the promoter tagging vector, independent antibiotic resistant transgenic cell colonies were screened for baseline LUC activity (BLA) under non-stress *in vitro* conditions. BLA frequencies in THP ranged from 0.17% to 0.98% and from 0.73% to 1.38% using vector pETKUL2 and pKCKUL1, respectively (Table 3.1). Tagging frequencies in THP transformed with the pETKUL2 vector in three different experiments (10, 11, and 15) remained constant (0.98%, 0.70%, 0.89%, respectively). However, in experiment 16 the tagging frequency dropped significantly to 0.17% using the same cultivar and tagging vector and might be attributed to the low number of cell colonies screened in that experiment. The BLA frequencies reached in Will and Caca were somewhat lower (0.31% and 0.70%, respectively) with the pKCKUL1 tagging vector.

Table 3.1. Number and frequency of cell colonies showing baseline luciferase activity (BLA)

Experiment	Cultivar	Tagging vector	Total No. of cell colonies screened[a]	Number and frequency (%)[b] of cell colonies showing BLA
10	THP[c]	pETKUL2	15,887	155 (0.98)
11	THP	pETKUL2	11,973	84 (0.70)
15	THP	pETKUL2	11,375	101 (0.89)
		pKCKUL1	13,930	192 (1.38)
16	THP	pETKUL2	3570	6 (0.17)
		pKCKUL1	8925	65 (0.73)
	Will[d]	pKCKUL1	19,250	60 (0.31)
	Caca[e]	pKCKUL1	4833	34 (0.70)
		TOTAL	89,743	697 (0.78)

[a]Based on the average number of independent cell colonies calculated from at least three samples per construct and per cultivar.
[b]The number of independent candidate promoter-tagged cell colonies showing BLA expressed in percentage of the total number of independent transgenic cell colonies screened.
[c]Plantain cultivar 'Three Hand Planty' (genotype AAB).
[d]Dessert banana cultivar 'Williams' (genotype AAA).
[e]Cooking banana cultivar 'Cacambou' (genotype ABB).

A large number of independent transgenic cell colonies (up to ca. 89,000) were screened (Table 3.1). This large-scale screening was possible by screening thousands of independent cell colonies in a single image (Figure 3.1). In total, 697 candidate promoter-tagged lines showing BLA were detected (Table 3.1). Processed LUC images revealed several cell colonies with LUC activity per sample transformed with one of the promoter tagging vectors or the positive control vector pETKUL3, whereas non-transformed cell colonies showed no luminescence (Figure 3.1). These results indicate that banana promoters are tagged in the pETKUL2 and pKCKUL1 transformed lines showing BLA. However, the number of candidate promoter-tagged lines was higher with pKCKUL1 than with pETKUL2 (Figure 3.1). A first qualitative assessment of the level of LUC activity in each cell colony is feasible using the processed LUC images depicted in pseudocolors and the corresponding color bar as shown in Figure 3.1. It is obvious that irrespective of the tagging vector a wide range of LUC activity was obtained in the different candidate promoter-tagged cell colonies.

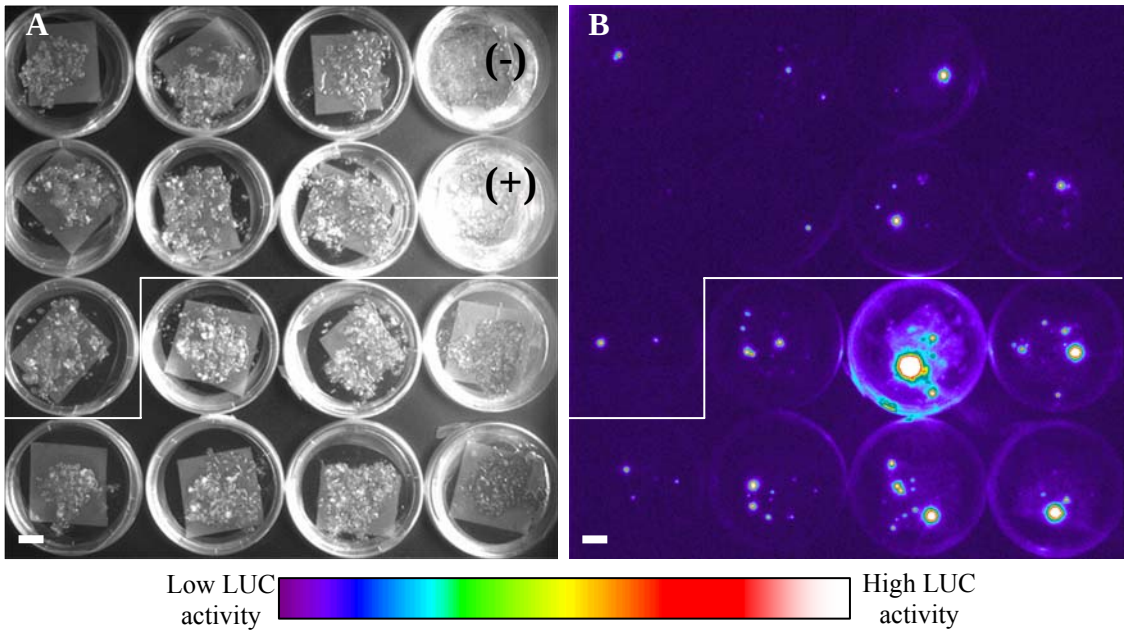

Figure 3.1. Large scale screening for baseline LUC activity (BLA) in thousands of cell colonies transformed with promoter tagging vectors pETKUL2 and pKCKUL1, and the positive control vector pETKUL3. (A) Live image under normal light conditions. (B) LUC image in complete darkness with an integration time of 20 min and depicted in pseudocolors with a maximum greyscale value of 500. The level of BLA can be assessed by the color bar. The symbols (-) and (+) refer to control, non-transformed and positive control cell colonies transformed with the *luc*$^{+}$ gene under control of the enhanced CaMV 35S RNA promoter (pETKUL3), respectively. The remaining upper seven Petri dishes, above the white line, correspond to samples transformed with pETKUL2 while the seven Petri dishes below the white line were transformed with pKCKUL1. Bar represents 1 cm.

3.2.1.2 Effect of an intron in the luciferase gene on the frequency and level of BLA

As can be observed in Table 3.1 for the cultivar THP, a higher tagging frequency was obtained with pKCKUL1 (*luc*$^{+INT}$ gene) than with pETKUL2 (intronless *luc*$^{+}$ gene) in experiments 15 (1.38% *vs*. 0.89%, respectively) and 16 (0.73% *vs.* 0.17%, respectively). The processed LUC images acquired during the large scale activation screenings confirmed these findings and also suggested that higher LUC activities were reached with pKCKUL1 (Figure 3.2). For further investigation only the candidate promoter-tagged lines of experiment 15 showing strong BLA, i.e. luminescence visually detectable at a maximum greyscale setting of 10,000, were considered. A total of 12 and 49 promoter-tagged lines transformed with pETKUL2 and pKCKUL1, respectively, meet this *criterium* which amounts to 12% and 26% of the total number of lines showing BLA (Table 3.1), respectively. After normalization for the total number of transgenic lines screened, the number of cell colonies with such strong BLA was 3.3-fold higher in pKCKUL1 than in pETKUL2 tagged cell colonies (49 *vs.* 15, respectively). These results suggest that more tagged lines with strong BLA are obtained with an intron containing *luc*$^{+}$ gene than with an intronless *luc*$^{+}$ gene. Among the positive control pETKUL3 (*luc*$^{+}$ driven by the enhanced 35S promoter) transformed lines, BLA was visually detectable at the threshold level indicated above in only one line. Therefore, an additional five independent pETKUL3 lines

with BLA that became visible at a maximum greyscale setting of at least 1000 were included for further analysis. Using the image analysis software the luminescence of the independent tagged lines of this group with strong BLA was quantified and expressed in relative light units (RLU) after correction for the background measured in control, non-transformed lines. The mean RLU values of the pETKUL2 and pKCKUL1 transformed lines revealed a 3.5-fold higher luminescence in the latter lines (Figure 3.2) confirming that higher LUC activities are obtained in candidate promoter-tagged lines with luc^{+INT} than with luc^{+}.

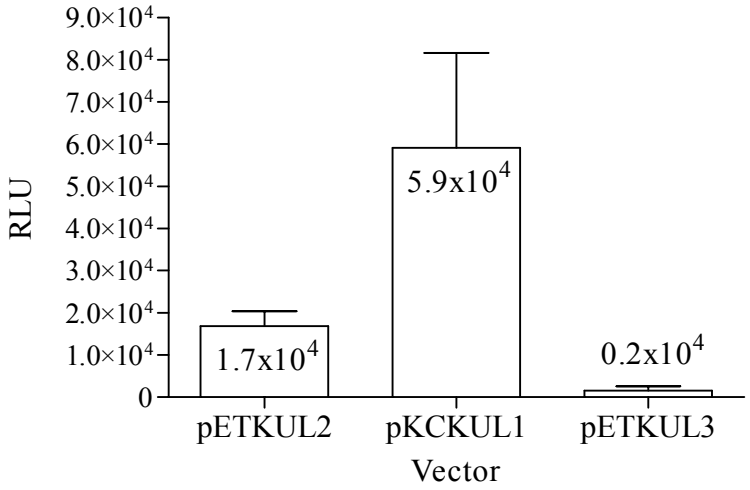

Figure 3.2. Baseline LUC activity (BLA) in cell colonies transformed with promoter tagging vectors pETKUL2 and pKCKUL1, and the positive control vector pETKUL3. Average (± standard error) BLA per vector expressed in relative light units (RLU) of the cell colonies of experiment 15 with strong LUC activity which is visually detectable at a maximum greyscale setting of 10,000 (pETKUL2 and pKCKUL1) or at least 1,000 (pETKUL3). The RLU of 12, 49 and 6 independent transgenic cell colonies were measured for vectors pETKUL2, pKCKUL1 and pETKUL3, respectively. Correction was done for the background measured in control non-transformed lines. BLA was measured in an area of 0.14 cm^2 for all cell colonies.

3.2.2 Screening for luciferase activity during low temperature treatment

3.2.2.1 Temperature-controlled real-time screening

Two to three months after *Agrobacterium* transformation, transgenic cell colonies were screened for LUC activity under temperature-controlled conditions using an in-house designed temperature control system (paragraph 2.2.4.1). In combination with the luc^{+} reporter gene it allowed the real-time monitoring of LUC activity during different temperature regimes in a non-invasive manner. The cell colonies were recovered after the LUC screening for *in vitro* regeneration. By lowering the temperature of the thermostatic water bath of the temperature control system, the desired temperature was reached for the samples placed on the plate.

3.2.2.2 Luciferase activity under different low temperature regimes

Cell colonies transformed with pETKUL2 or pKCKUL1 were screened for LUC activity under temperature controlled conditions. Real-time monitoring of LUC activity was performed while lowering the temperature and 3-10 h later LUC activity pattern was visually scored relative to the LUC activity at 26°C. The number and frequency of the different LUC activity patterns was calculated (Table 3.2). The frequencies calculated for the low temperature LUC activity (LTLA) are identical or higher than the BLA frequencies shown in Table 3.1, except for experiment 10. The higher LTLA frequencies are due to cell colonies showing LUC activity at LT but not displaying BLA at 26°C that is visually detectable above a maximum greyscale level of 500.

As expected, the differences in LTLA tagging frequency between promoter tagging vectors were comparable to those in BLA tagging frequencies with higher tagging frequencies obtained with pKCKUL1 than with pETKUL2, except in experiment 15 for the 26°C → 18°C treatment. In experiment 10, BLA screening was performed at room temperature two months after transformation and the 155 cell colonies showing LUC activity (Table 3.1) were removed. One month later the remaining cell colonies were screened for LTLA at 8°C and 96 cell colonies that were not BLA positive one month earlier showed different LTLA patterns. The proportion of cell colonies showing an enhanced low temperature LUC activity [ELTLA, Table 3.2 under the increase (↑) heading], relative to the BLA at the different LT treatments ranged from 0.0% to 48.7% (Table 3.2). No obvious relation with the severity of the LT treatment was observed. However, despite the one week acclimatization to 16°C a subsequent drop in temperature to 4°C caused a strong decrease in the number of cell colonies showing LTLA (from 32 to 18 and from 75 to 51 for pETKUL2 and pKCKUL1 transformed THP cell colonies in experiment 15, respectively). Moreover, in the large majority (more than 90%) of these lines the LTLA decreased. Within the cultivar THP and within a temperature treatment comparable LTLA tagging frequencies were obtained with the same tagging vector in different experiments (e.g. pETKUL2, 26°C → 8°C: 0.61%, 0.75% and 0.53% in experiment 10, 11 and 15, respectively).

Table 3.2. Number of candidate promoter-tagged cell colonies showing an increase, decrease or a *status quo* in LUC activity under different low temperature treatments relative to the baseline LUC activity (BLA) at 26°C and the corresponding low temperature LUC activity (LTLA) frequencies

				Number of cell colonies showing different LUC activity patterns at low temperature (%)[a]				
Exp.	**Cv.[b]**	**Tagging vector**	**Temperature treatment**	**↑**	**↓**	**=**	**Total LTLA[c]**	**LTLA frequency (%)[d]**
10	THP	pETKUL2	26°C → 8°C	10 (10.4)	84 (87.5)	2 (2.1)	96	0.61
11	THP	pETKUL2	26°C → 8°C	14 (15.6)	69 (76.7)	7 (7.8)	90	0.75
15	THP	pETKUL2	26°C → 18°C	4 (12.9)	10 (32.3)	17 (54.8)	31	1.36
		pKCKUL1	26°C → 18°C	8 (32.0)	2 (8.0)	15 (60.0)	25	0.90
		pETKUL2	26°C → 16°C	7 (41.2)	4 (23.5)	6 (35.3)	17	0.75
		pKCKUL1	26°C → 16°C	19 (48.7)	5 (12.8)	15 (38.5)	39	1.40
		pETKUL2	26°C → 12°C	6 (25.0)	10 (41.7)	8 (33.3)	24	1.05
		pKCKUL1	26°C → 12°C	12 (25.5)	12 (25.5)	23 (48.9)	47	1.69
		pETKUL2	26°C → 8°C	4 (33.3)	5 (41.7)	3 (25.0)	12	0.53
		pKCKUL1	26°C → 8°C	14 (45.2)	10 (32.3)	7 (22.6)	31	1.11
		pETKUL2	26°C → 16°C → 4°C[e]	11 (34.4)[f] – 0 (0.0)[g]	9 (28.1)[f] – 17 (94.4)[g]	12 (37.5)[f] – 1 (5.6)[g]	32[f] – 18[g]	1.40[f] – 0.78[g]
		pKCKUL1	26°C → 16°C → 4°C[e]	31 (41.3)[f] – 3 (5.9)[g]	11 (14.7)[f] – 48 (94.1)[g]	33 (44.0)[f] – 0 (0.0)[g]	75[f] – 51[g]	2.69[f] – 1.83[g]
16	THP	pKCKUL1	26°C → 16°C	20 (43.5)	14 (30.4)	12 (26.1)	46	0.86
		pETKUL2	26°C → 12°C	1 (16.7)	5 (83.3)	0 (0.0)	6	0.17
		pKCKUL1	26°C → 12°C	8 (25.0)	13 (40.6)	11 (34.4)	32	0.90
	Will	pKCKUL1	26°C → 16°C	16 (34.0)	9 (19.1)	22 (46.8)	47	0.38
		pKCKUL1	26°C → 12°C	6 (23.1)	17 (65.4)	3 (11.5)	26	0.37
	Caca	pKCKUL1	26°C → 16°C	0 (0.0)	4 (23.5)	13 (76.5)	17	0.53
		pKCKUL1	26°C → 12°C	3 (17.6)	10 (58.8)	4 (23.5)	17	1.05

[a]LUC activity pattern was scored relative to the BLA at 26°C with ↑, ↓ or = indicating an increase, decrease or *status quo*, respectively. Between brackets the frequency of each pattern is expressed in percentage of the total number of cell colonies showing LTLA.
[b]Cultivars 'Three Hand Planty' (THP), 'Williams' (Will) and 'Cacambou' (Caca).
[c]Total number of cell colonies showing LTLA.
[d]Total number of cell colonies showing LTLA expressed in percentage of the total number of transgenic cell colonies screened.
[e]Cell colonies were cultured at 26°C, then subjected to 16°C for one week and finally to 4°C.
[f]Results for the transition from 26°C to 16°C.
[g]Results for the transition from 16°C to 4°C.

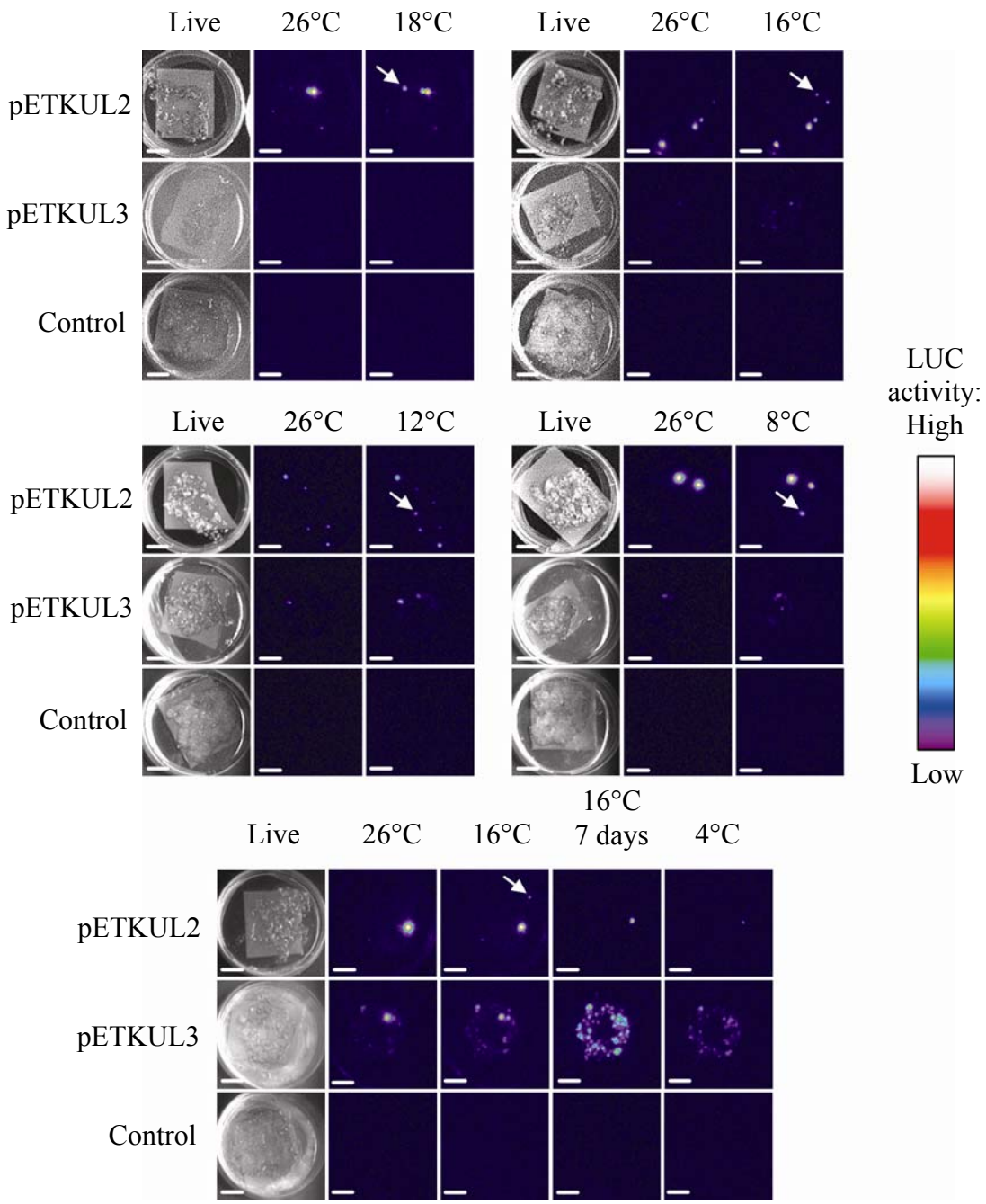

Figure 3.3. Luciferase activity at 26°C and at different low temperature (LT) treatments in candidate promoter-tagged cell colonies of experiment 15 screened two months after *Agrobacterium* transformation. Representative images were taken under normal light (Live) and dark conditions of cell colonies transformed with pETKUL2 (promoterless *luc*$^+$ gene) and pETKUL3 (positive control vector carrying the *luc*$^+$ gene under control of the enhanced CaMV 35S RNA promoter), and non-transformed cell colonies (Control). LUC images were recorded in the dark after 4 h and 3 h at 26°C and the corresponding LT, respectively, and are depicted in pseudocolors (see color bar) with an upper greyscale limit setting of 500. Arrows indicate promoter-tagged cell colonies showing an enhanced low temperature LUC activity compared to the 26°C image. For cell colonies subjected to the 16°C→4°C treatment (lower images), images were also recorded at the 7th day of acclimatization at 16°C and 3 h after the colonies were subjected at 4°C. Scale bars represent 1 cm.

For the different LT treatments of experiment 15, representative processed images used for the analyses are shown in Figure 3.3. Per temperature treatment a non-transformed control, a candidate promoter-tagged pETKUL2 transformed and a positive control pETKUL3 transformed sample are shown. Cell colonies showing an ELTLA are indicated with an arrow and were identified under all LT treatments. Cell colonies transformed with pETKUL3 were showing a low, almost undetectable BLA sometimes comparable to the negative control non-transformed sample (e.g. 18°C treatment, Figure 3.3). The LUC activity of the pETKUL3 cell colonies at LT (16°C, 12°C and 8°C) remained similar or was higher than at 26°C. A strong enhancement of LUC activity was observed for the positive control pETKUL3 lines after 7 days at 16°C in the 26°C→16°C→4°C treatment. Moreover, already after 3 h at 16°C LTLA reached above BLA. However, the LUC activity drastically dropped when the pETKUL3 transformed samples were subjected to 4°C. At 7 days at 16°C no enhancement in LTLA was observed in the pETKUL2 transformed sample in Figure 3.3. In addition, the line that showed ELTLA after 3 h at 16°C was not showing detectable LUC activity after 7 days indicating the transient response to 16°C of the candidate promoter tagged in that line.

3.2.3 Time course analysis of LUC activity in independent candidate promoter-tagged cell colonies

3.2.3.1 Qualitative analysis

Two months after *Agrobacterium* transformation with promoter trap vector pETKUL2, screening of 15,887 independent cell colonies at 26°C revealed 155 (0.98%) cell colonies showing BLA in experiment 10 (Table 3.1). After removing LUC activated colonies, the remaining cell colonies were re-screened for BLA (26°C) one month later during 2 h followed by a LT treatment of 8°C while monitoring LUC activity in real-time for up to 10 h (paragraph 3.2.2.2). This type of screening at cell colony stage (I) was repeated during the *in vitro* regeneration process (paragraph 3.3) and the LUC activity patterns for 23 responsive lines are listed in Table 3.3. Despite the early removal of BLA positive colonies one month before real-time screening at 26°C and 8°C, BLA (26°C) was detected at developmental stage I for all lines except line 42 though at variable levels from weak to very strong reflecting the strength of the tagged promoters. Representative cell colony lines showing an up-regulation of LUC activity at 8°C are illustrated (Table 3.3) and include pETKUL2 transformed tagged lines 17 (ET2-17) and 42 (ET2-42). In addition, cell colony lines showing a down-regulation of LUC activity with different levels of BLA were also selected for further LUC screening during *in vitro* development (paragraph 3.3).

Table 3.3. Scoring of luciferase (LUC) activation at 26°C and 8°C in promoter-tagged lines of banana during three different *in vitro* developmental stages

	Developmental stage					
	I (cell colony)[b]		IV (shoot induction)[c]		VII (*in vitro* regenerated plant)[d]	
ET2 line[a]	26°C	8°C	26°C	8°C	26°C	8°C
17	W	↑	M	↑	M	↓
42	N	↑	M	↓	W	↓
41	M	↓	S	↓	W	↓
156	VS	↓	VS	↓	VS	↓
64	VS	↓	S	↓	VS	↓
34	VS	↓	S	↓	M	↓
49	S	↓	S	↓	VS	↓
28	S	↓	S	↓	N	N
176	M	↓	M	↓	VS	↓
133	M	↓	M	↓	W	↓
85	W	↓	VS	↓	S	↓
179	W	↓	VS	↓	S	↓
130	W	↓	VS	↓	M	↓
123	W	↓	S	↓	VS	↓
132	W	↓	S	↓	W	↓
10	W	↓	M	↓	M	↓
82	W	↓	M	↓	M	↓
102	W	↓	M	↓	M	↓
114	W	↓	M	↓	M	↓
89	W	↓	M	↓	W	↓
111	W	↓	W	↓	VS	↓
37	W	↓	W	↓	N	N
62	W	↓	W	↓	N	N
+[e]	M	↓	M	↓	W	↓
-[f]	N	N	N	N	N	N

Based on the direct correlation between the number of greyscale levels as detected by the CCD camera and the level of LUC activity lines were ranked at 26°C in five different classes. When LUC activity was detectable with an upper greyscale limit setting of more than 10,000, between 5000 and 10,000, 3000 and 5000, 500 and 3000 or less than 500 it was considered very strong (VS), strong (S), moderate (M), weak (W) or not detectable (N), respectively. At each developmental stage real-time monitoring of LUC activity took place while lowering the temperature to 8°C and 3-10 h later LUC activity was scored again with ↑ and ↓ indicating an increase and decrease, respectively, relative to the LUC activity at 26°C.

[a]Embryogenic cell suspensions of 'Three Hand Planty' were transformed with T-DNA promoter tagging vector pETKUL2 (promoterless *luc*$^+$ gene), and geneticin selection for transgenic cultures started 1 week after cocultivation.

[b]Cell colony stage on ZZ medium three months after transformation.

[c]Shoot induction stage on RD2 medium eight months after transformation.

[d]*In vitro* regenerated plant on REG medium 15 months after transformation.

[e]Line transformed with positive control vector pETKUL3 (*luc*$^+$ gene fused to the enhanced CaMV 35S RNA promoter).

[f]Non-transformed line

Comparison of LUC images at cell colony stage showed a clear increase of LUC activity at 8°C in promoter-tagged line ET2-17, while the LUC activity of lines ET2-34, ET2-85 and ET2-156 was lower at 8°C than at 26°C (Figure 3.4 I). A similar LUC activity pattern was observed in positive control cell colonies transformed with vector pETKUL3.

LUC activity was absent in the negative, non-transformed control line (Figure 3.4) indicating that promoters were tagged in these pETKUL2 transformed lines. The promoter-tagged line ET2-17 was chosen as a representative up-regulated line at LT stress due to the repetitive up-regulated LUC pattern detected in early developmental stages in further screenings during *in vitro* development (Figure 3.4, paragraph 3.3).

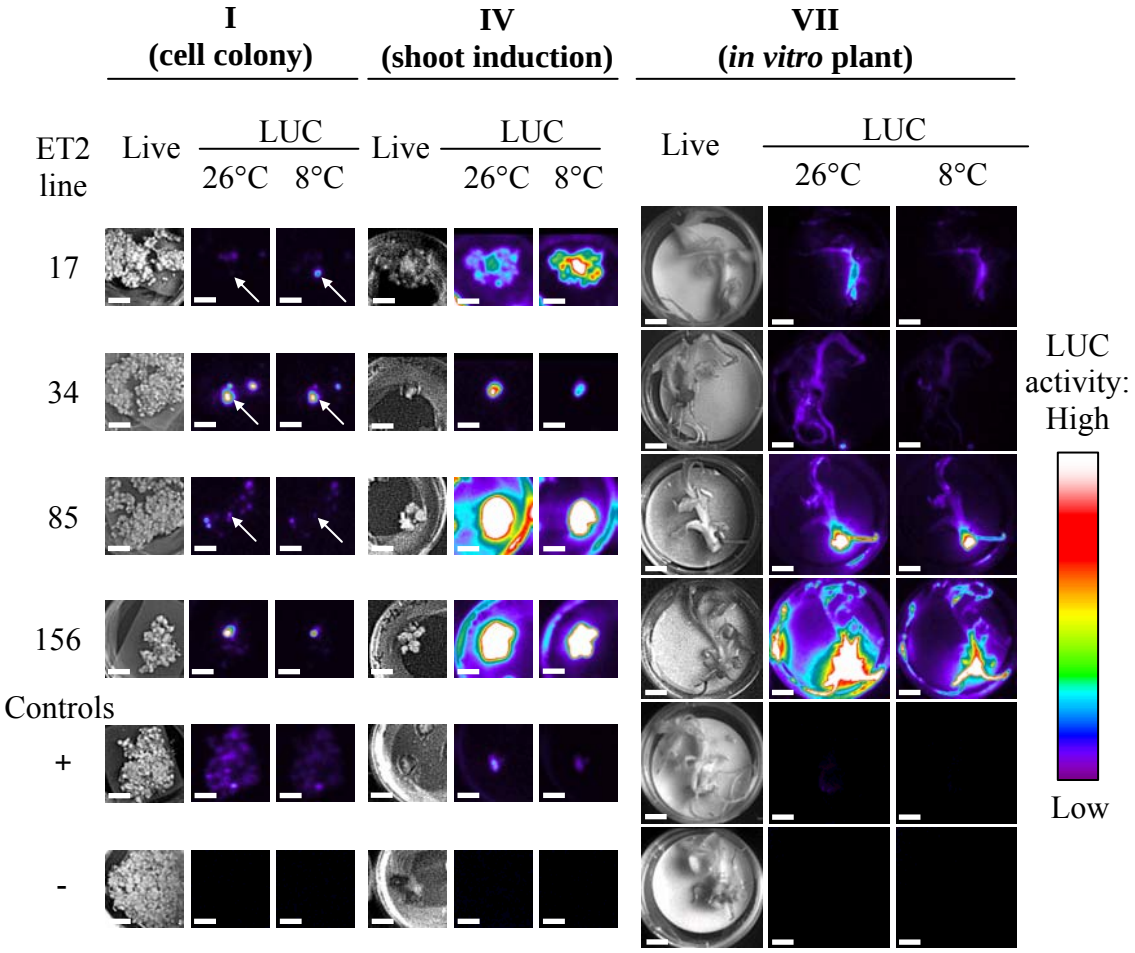

Figure 3.4. Luciferase activation at 26°C and 8°C in candidate promoter tagged lines throughout *in vitro* regeneration. Representative images were taken under normal light (Live) and dark (LUC) conditions of four candidate tagged lines (17, 34, 85 and 156) transformed with pETKUL2 (promoterless *luc*$^+$ gene), a positive control line carrying the *luc*$^+$ gene under control of the enhanced CaMV 35S RNA promoter (+), and a negative non-transformed control line (-). Screenings I, IV and VII correspond to cell colony, shoot induction and plant stage, respectively. LUC images were recorded after 2 h and 3 h at 26°C and 8°C, respectively, and are depicted in pseudocolors (see color bar) with an upper greyscale limit setting of 1000. Arrows indicate the corresponding line. Scale bars represent 1 cm.

3.2.3.2 Quantitative analysis

Besides the qualitative analysis of LTLA during the temperature decrease as explained in the previous paragraph, a quantitative analysis was performed to verify the LTLA pattern of promoter-tagged lines of experiment 10 by software-based measurement of the light emission within the region of interest (ROI). The ROI comprises an area somewhat larger than that covered by the largest promoter-tagged line at each stage, except for screening at stage IV where the smallest culture line was taken as a reference (paragraph 3.3). Measurements were performed for two hours at 26°C, during the temperature change to 8°C which took approximately one hour and for a period of three hours at 8°C (Figure 3.5). The lines ET2-17 and ET2-42 showed an increase in LUC activity at 8°C with a much higher up-regulation in the former than in the latter line (excluding the measurements between the temperatures 26°C and 8°C), from 1128 ± 192 to 12,071 ± 773 RLU and from 320 ± 65 to 766 ± 54 RLU, respectively (Figure 3.5 I). On the other hand, the promoter-tagged lines ET2-28, ET2-34, ET2-85 and ET2-156 showed a decrease in LUC activity upon a change to LT similar to the LUC activity pattern of a positive control line (+) ET3 containing the T-DNA of the pETKUL3 vector (Figure 3.5 I). The LUC activity of line ET2-34 at cell colony stage (screening I) increased transiently during the temperature drop from 26°C to 8°C, but was finally lower at 8°C than at 26°C. The LTLA patterns including those of the lines presented in Figure 3.5 were confirmed by calculation of the $RLU_{8°C}/RLU_{26°C}$ ratio using the average LUC activity at 8°C and 26°C, excluding the measurements between the temperatures 26°C and 8°C from Figure 3.5 (Table 3.4). Consequently, a $RLU_{8°C}/RLU_{26°C}$ above 1.0 indicates an up-regulation of LUC activity by LT, whereas a repression of LUC activity by LT results in a $RLU_{8°C}/RLU_{26°C}$ smaller than 1.0. Of the six independent pETKUL2 transformed lines shown in Figure 3.5 I, two lines (ET2-17 and ET2-42) showed an increase in LUC activity under LT stress (10.7- and 2.4-fold, respectively). The drop in temperature caused in the other candidate promoter-tagged ET2-lines showed a decrease in LUC activity ranging from 0.2- to 0.4-fold (Table 3.4). Quantification of lines carrying the *luc*$^+$ driven by the enhanced 35S promoter (pETKUL3) was performed at a single cell colony level (Figure 3.5 I and Table 3.4) or at an overall cell colony level (all cell colonies in a 5-cm PD, data not shown).

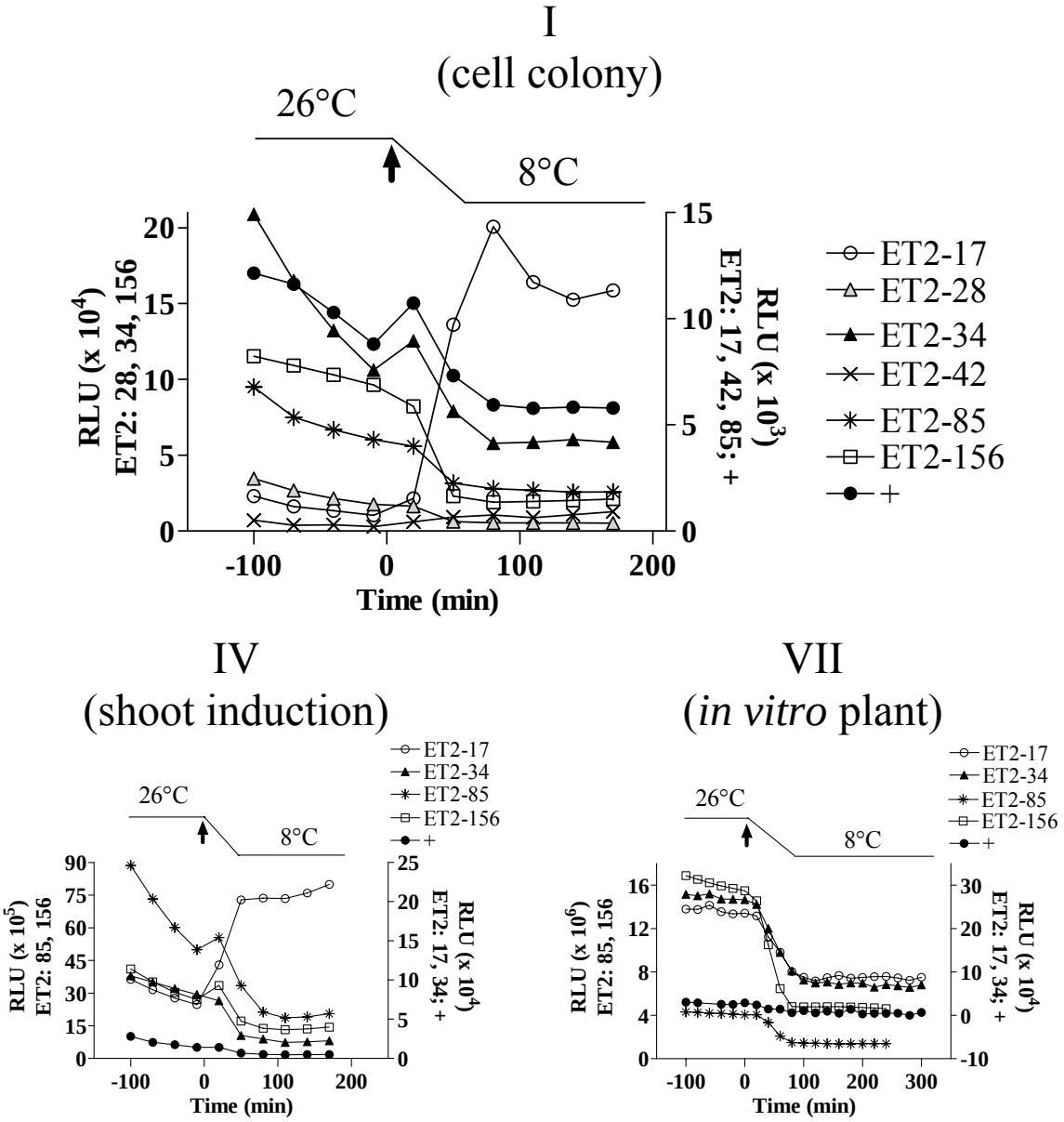

Figure 3.5. Time course of luciferase (LUC) activity during temperature regime in promoter-tagged lines of experiment 10 transformed with the promoter tagging vector pETKUL2 or the positive control vector pETKUL3 (+) throughout the *in vitro* regeneration process. Screenings I, IV and VII correspond to cell colony, shoot induction and plant stage, respectively. LUC activity was monitored for 2 h at 26°C, at time point zero (indicated by an arrow) temperature was set to 8°C, which was reached 1-1.5 h later, and then maintained for 2-4 h (solid line above the graphs). Two images of 20 min exposure time were recorded per hour at screenings I and IV, while three images were acquired per hour at screening VII. In all cases, the LUC activity expressed in relative light units (RLU) was corrected for the background measured in a negative, non-transformed control line. The region of interest for quantification of LUC activity was 0.34, 0.58, and 23.19 cm^2 at screening I, IV, and VII, respectively. The Y axis scale for each line is indicated.

Table 3.4. Fold changes in LUC activity in candidate promoter-tagged lines of experiment 10 in response to low temperature stress at several stages during *in vitro* development

		$RLU_{8°C}/RLU_{26°C}$				
Stage - screening[a]	**Feature**	**ET2-17**[b]	**ET2-34**[b]	**ET2-85**[b]	**ET2-156**[b]	**+**[c]
Cell colony - I	Change	10.7	0.4	0.4	0.2	0.5
	cChange[d]	19.6	0.7	0.7	0.3	1.0
Shoot induction - IV	Change	2.5	0.2	0.3	0.4	0.2
	cChange	10.3	1.0	1.2	1.7	1.0
In vitro plant - VII	Change	0.4	0.3	0.3	0.3	0.2
	cChange	1.5	1.2	1.4	1.3	1.0

Real-time screening for LUC activity in four candidate promoter-tagged lines and one positive control line during the temperature decrease from 26°C to 8°C at several stages of the *in vitro* regeneration process. The average of the last four to six measurements of the LUC activity expressed in relative light units (RLU) at each temperature (26°C and 8°C) was used to calculate the fold change in LUC activity under 8°C temperature stress. A $RLU_{8°C}/RLU_{26°C}$ value above one indicates an up-regulation of LUC activity at 8°C, whereas a value less than one represents a down-regulation of LUC activity at 8°C

[a]Developmental stage – screening number. I: First screening of antibiotic-resistant cell colonies on ZZ medium three months after transformation. IV: Fourth screening of shoot forming cultures on RD2 medium eight months after transformation. VII: Seventh screening of regenerated *in vitro* plants maintained in REG medium 14-20 months after transformation.

[b]Independent lines transformed with promoter tagging vector pETKUL2 which contains a promoterless *luc*$^+$ gene

[c]Line transformed with the positive control vector pETKUL3 which contains the enhanced CaMV 35S RNA promoter driving the *luc*$^+$.

[d]The LUC activity at 8°C was corrected by multiplying with the reverse of the fold change ($RLU_{26°C}/RLU_{8°C}$) of the positive control line (ET3) and used to recalculate the corrected $RLU_{8°C}/RLU_{26°C}$.

Despite the fact that the LUC activity level differed among the positive control (+) ET3-lines at all temperatures tested in experiment 15 (Figure 3.3), the decrease in LUC activity at 8°C was similar at a single cell colony level (Figure 3.5 I) and the overall cell colony level (data not shown) and amounted up to approximately 0.5-fold (Table 3.4). This negative effect on LUC activity is of the same order of magnitude than that recorded for the ET2-lines displaying the same LUC activity pattern. In addition, when the $RLU_{8°C}/RLU_{26°C}$ ratios were corrected or normalized for the apparent decrease in LUC activity of the positive control line at 8°C (Table 3.4), a 19.6- up-regulation of LUC activity at 8°C occurred in this line at the cell colony stage. Few exceptions were observed in some cell colonies where the LUC activity pattern at LT was different than in the majority of the other cell colonies of the positive control lines (data not shown) and might be attributed to the integration of the T-DNA nearby banana *cis*-acting elements responsible for the different LUC activity pattern at LT.

The effect of different LT regimes on the LUC activity of cell colonies transformed with the positive control vector pETKUL3 (pe35S-T::*luc*$^+$, paragraph 2.1.2) or pKCKUL2 (p*Ubi*1L::*luc*$^{+INT}$, paragraph 2.1.2) was studied in several experiments two to three months after transformation. The LUC activity at the different temperatures was averaged over several images recorded for a period of two hours (Figure 3.6A). Light emission

measurements involved all the cell colonies contained in one PD per cultivar-temperature-vector combination and one to four replicate PDs were measured within an experiment. The RLU ratio between the LUC activity at 26°C and the different LTs (RLU at 26°C/RLU at LT) was then calculated (Figure 3.6B). A value below 1.0 indicates that LUC activity was higher at the LT than at 26°C.

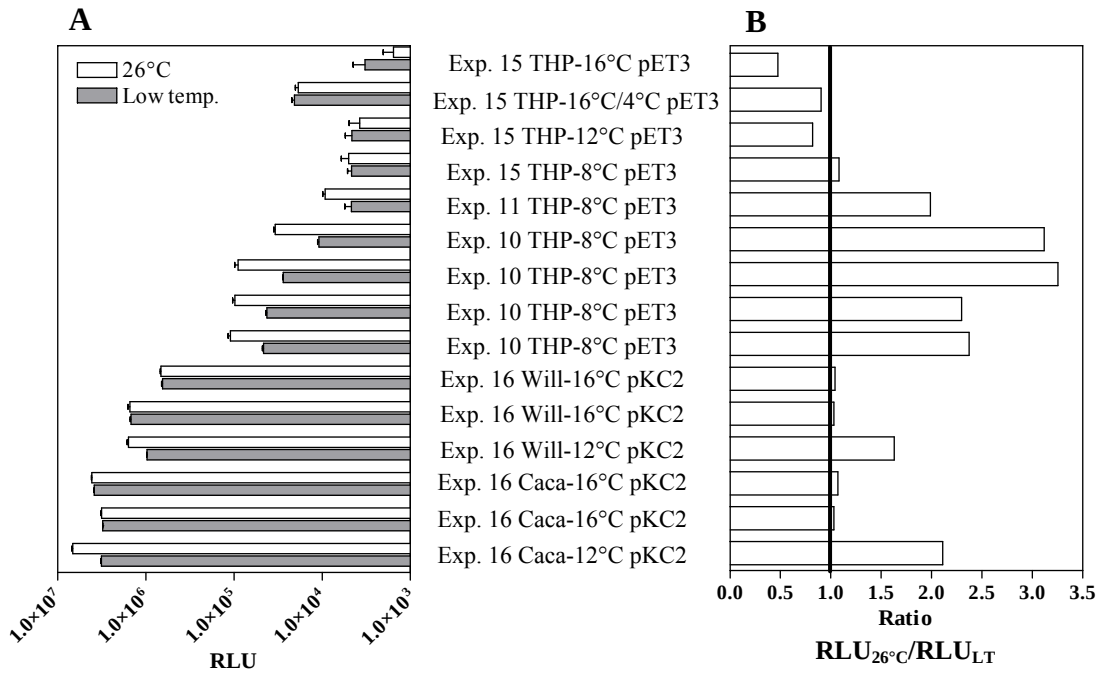

Figure 3.6. Quantitative analysis of the LUC activity pattern under different low temperature (LT) regimes in cell colonies transformed with the positive control vector pETKUL3 (pET3) or pKCKUL2 (pKC2) two to three months after transformation. The *luc*$^{+}$ or *luc*$^{+INT}$ gene in these vectors is driven by the enhanced CaMV 35S RNA promoter with TMV leader sequence or the maize *Ubi*1 promoter plus leading intron, respectively. The LUC activity was monitored in real-time throughout the temperature regime, corrected for the background measured in the negative, non-transformed control and expressed in relative light units (RLU). Experiment, cultivar, LT treatment and positive control vector are shown in the legend between the graphs. (A) LUC activity at 26°C and the corresponding LT as indicated in the legend between the graphs with one Petri dish (PD) analyzed per entry. From one to four replicate PDs were measured per cultivar-temperature-vector combination within an experiment. The approximate number of cell colonies per PD varies according to the experiment (Exp. 10: 981, Exp. 11: 491, Exp. 15: 403, Exp. 16: Will, 1750; Caca, 537. Cell colonies of the different experiments were screened for LUC activity after 9-10 weeks (Exp. 15), 8-9 weeks (Exp. 11), 12-13 weeks (Exp. 10) and 9 weeks (Exp. 16) after genetic transformation. The ROI was 7.15 cm^2 measuring all cell colonies in a 5-cm PD. The RLU value at 26°C is the average (± standard error) of four images recorded for a period of 2 h before the LT treatment, while for the LT it is the average (± standard error) activity of four to six images recorded between 1 and 3 h after reaching the indicated LT. For the LT regime of 16°C for one week acclimatization and then a drop to 4°C in experiment 15 the quantification for LT was performed when the material was 1 to 3 h at 16°C. X axis is at logarithmic scale. (B) The RLU ratio was calculated by dividing the LUC activity value at 26°C by the activity at the LT ($RLU_{26°C}/RLU_{LT}$). Ratios below one (thick vertical line) indicate higher LUC activity at LT than at 26°C.

The LUC activity at 26°C (BLA) in the pETKUL3 lines of experiment 10 was comparable in three of the four replicates (34,111 ± 1336; 90,024 ± 8765; 97,761 ± 6106 and 110,152 ± 6937 RLU). The LUC activity pattern at 8°C in the pETKUL3 lines of this experiment showed a decrease in LUC activity when compared to the BLA in the four

replicates, ranging from 2.3- to 3.1-fold ($RLU_{26°C}/RLU_{LT}$). The three replicate pKCKUL2 lines of the cultivar 'Williams' (Will) showed high BLA levels (1,525,360 ± 68,521; 675,740 ± 13,296 and 1,580,021 ± 56,822 RLU; Figure 3.6A). A lack of LT LUC activity response was observed for the two replicate pKCKUL2 Will samples that were subjected to 16°C (1.0-fold, Figure 3.6B). In the cultivar 'Cacambou' (Caca) the three replicates of the pKCKUL2 lines show a strong BLA (4,131,387 ± 22,086; 3,177,623 ± 50,313 and 6,763,359 ± 162,381 RLU, Figure 3.6A). In addition, the LUC activity of the two replicates of the pKCKUL2 Caca lines that were subjected to 16°C also showed a *status quo* in the activity at LT when compared to the BLA (1.1- and 1.0-fold, Figure 3.6B).

The BLA in experiment 15 was comparable for the treatments 16°C, 12°C and 8°C (1560 ± 475, 3774 ± 1178 and 5023 ± 1044, respectively, Figure 3.6A), although processed LUC images revealed a general weak LUC activity among the pETKUL3 lines (Figure 3.3). The BLA for the treatment 16°C/4°C (18,683 ± 1466) revealed a higher LUC activity than the other treatments of experiment 15. The BLA in experiment 11 is lower than the BLA of the treatment 16°C/4°C of experiment 15 (9291 ± 591 and 18,683 ± 1466, respectively) but still higher than the BLA of the other LT treatments of experiment 15 (Figure 3.6A). Furthermore, the BLAs in experiment 10 revealed higher activity than the highest BLA of the treatment 16°C/4°C of experiment 15 (Figure 3.6A). In general, there was nearly a 70-fold difference in average LUC activity at 26°C (110,152 *vs*. 1560 RLU) in the different experiments for lines transformed with pe35S-T::*luc*$^{+}$ (Figure 3.6A)

Low temperature LUC activity in experiment 15 was comparable among the LT treatments 16°C, 12°C and 8°C (3278 ± 1179; 4610 ± 874 and 4646 ± 517 RLU, respectively). However, the highest LT LUC activity in experiment 15 was detected for the treatment 16°C/4°C (20,678 ± 1425). This treatment showed also the highest LUC activity at 26°C (Figure 3.6A). A comparable LUC activity at LT in experiment 11 (4671 ± 517 RLU) and a higher LUC activity in experiment 10 (10,937 ± 348; 27,656 ± 422; 42,524 ± 952 and 46,394 ± 1209 RLU) was observed compared to the LT LUC activity in the different treatments of experiment 15 (Figure 3.6A).

Different LUC activity patterns were observed when decreasing the temperature from 26°C (Figure 3.6B). The LUC activity pattern for the LTs 16°C (including 16°C and 16°C/4°C treatments) and 12°C in experiment 15 revealed an increase in LUC activity at LT ($RLU_{26°C}/RLU_{LT}$: 0.5, 0.9 and 0.8, respectively, Figure 3.6B). A slight decrease in LUC activity was detected in experiment 15 between 26°C and 8°C ($RLU_{26°C}/RLU_{LT}$ of 1.1). A 2.0-fold decrease in LUC activity by 8°C was observed in experiment 11 which is much more pronounced than in experiment 15. In experiment 10, a clear down-regulation of LUC activity was detected at 8°C ranging from 2.3- to 3.1-fold difference (Figure 3.6B). Furthermore, the lines of the cultivars Will and Caca transformed with p*Ubi*1::*luc*$^{+INT}$ showed a similar temperature dependent LUC activity pattern as observed in pe35S::*luc*$^{+}$

lines, but a *status quo* occurred between 26°C and 16°C, while a decrease of LUC activity was observed at 12°C.

Direct comparison between the enhanced 35S promoter and the *Ubi*1 promoter could not be performed in the present study. The *Ubi*1 promoter drives the intron containing *luc*$^+$ (pKCKUL2) while the enhanced 35S promoter drives the *luc*$^+$ (pETKUL3). The data indirectly suggest that higher BLA were obtained with the *Ubi*1 than with the enhanced 35S promoter (Figure 3.6A).

3.2.4 Discussion

The combination of the improved T-DNA promoter tagging vectors pETKUL2 (Remy *et al.* 2005) and pKCKUL1 (Remy *et al.* unpublished results) and careful real-time *in vivo* monitoring of activated LUC expression allowed sensitive screening for candidate tagged banana promoters. Two to three months after *Agrobacterium* infection of embryogenic suspension cells high-throughput screening for promoter-activated lines was performed at the cell colony stage. Large populations of T-DNA tagged lines have been produced in model plant species for reverse genetic studies or high-throughput T-DNA flanking sequence analyses like Arabidopsis (150,000, Alonso *et al.* 2003; ca. 100,000, Sessions *et al.* 2002; 60,480, Krysan *et al.* 1999; 59,979, Rosso *et al.* 2003) and rice (30,000, Jeon and An 2001; 29,482, Sallaud *et al.* 2004). Usually less T-DNA tagged lines have been screened for promoter activity using a promoterless reporter gene in these plants (Arabidopsis: 20,261, Alvarado *et al.* 2004; 2000, Mollier *et al.* 1995; 700, Calderon-Villalobos *et al.* 2006; and rice: 18,358, Jeon *et al.* 2000; 15,586, Lee *et al.* 2004; 3140, Ryu *et al.* 2004). Another model plant assessed for promoter tagging but involving much less lines is tobacco (200, Koncz *et al.* 1989; 234, Topping *et al.* 1991; 1000, Fobert *et al.* 1991; 300, Mudge and Birch 1998). In addition, other plant species in which promoters have been successfully isolated *via* T-DNA tagging include potato (Lindsey *et al.* 1993), *Brassica napus* (Bade *et al.* 2003), *Lotus japonicus* (Webb *et al.* 2000; Buzas *et al.* 2005) and *Arachis hypogaea* (Anuradha *et al.* 2006). Again, relatively few independent transgenic lines (87 to 821) were screened for reporter gene expression in these plants. The limited number of lines screened for promoter activation is mainly due to the laborious and time consuming procedures involving reporter gene activity assays besides the handling and maintenance of regenerated tagged lines. In contrast, the promoter tagging system described for the non-model organism banana allows the screening of tens of thousands of independent transgenic lines at the early cell colony stage (Remy *et al.* 2005). In this work a population of approximately 89,000 independent transgenic lines was screened for LUC activity.

Here, banana T-DNA-tagging frequency ranged from 0.17% to 2.69% (Table 3.2). Promoter tagging in several plant species is variable and ranges from 0.06% to 78%.

Different tagging frequencies obtained in different reports (Table 1.2) might be the result of a number of factors, including stage of development or tissue screened, (a)biotic stress treatments prior or during screenings, and reporter gene used. The highest promoter tagging frequency obtained in banana in the present study (2.69%, Table 3.2) is comparable to other luciferase based tagging experiments including tobacco (3.3%, Mudge and Birch 1998) and Arabidopsis (3.7%, Alvarado *et al.* 2004). Similar tagging frequencies were also observed in reports where the promoterless *uid*A reporter gene was used including rice (1.1% in leaves and 0.9% in roots, Jeon *et al.* 2000), *Brassica napus* (2.6% to 3.8% using a *uid*A::*neo* promoterless fusion, Bade *et al.* 2003), *Lotus japonicus* (3%, Webb *et al.* 2000; 2.4%, Buzas *et al.* 2005) and *Arachis hypogaea* (3.5% using an AMV-*uid*A::*neo* promoterless fusion, Anuradha *et al.* 2006). However, a direct comparison with the promoter tagging frequencies obtained in banana is not possible because screening in the other plant species was performed in different tissues of regenerated plants, while screening in banana was performed at an early undifferentiated stage (paragraph 3.2.1.1).

Comparable T-DNA tagging frequencies obtained in two plant species (Arabidopsis and tobacco) containing significant different genome sizes suggest that T-DNA is preferentially integrated in transcribed regions of the genome (Koncz *et al.* 1989). Schneeberger *et al.* (2005) obtained a correlation between gene expression and T-DNA integration in expressed regions of the genome in Arabidopsis. This observation is strengthened by reports indicating that the T-DNA is integrated in the chromosomal set of the female gametophyte when performing the *in planta* transformation method in Arabidopsis (Desfeux *et al.* 2000; Bechtold *et al.* 2000). Similarly, T-DNA integration is preferred in regions with highly expressed genes in rice (Jeong *et al.* 2006). The T-DNA in ECS of banana may be preferentially integrated into regions of the genome active during embryogenesis (developmentally regulated promoters) or induced by one or a combination of the components in the medium used for the maintenance of cells (inducible promoters), like the hormones zeatin (cytokinin) and 2,4-D (auxin) present in the ZZ medium (Table 2.1) when *Agrobacterium* transformation was performed. Bade *et al.* (2003) succeeded to trap and isolate new *Brassica napus* promoters restricted in activity to callus tissue with a promoterless *uid*A::*neo* fusion gene by first selecting for kanamycin resistant calli and then screening for callus-specific GUS expression. Further studies in the developmentally regulated *Brassica napus* promoters performed in the original tagged lines revealed that 30% of the promoters were inducible by auxin. Integration of the T-DNA tagging vector occurred during culture on auxin-containing medium (Bade *et al.* 2003).

In two independent experiments a higher BLA frequency was obtained with pKCKUL1 than with pETKUL2. Most likely this is due to the presence of the intron in the *luc*$^{+}$ gene in pKCKUL1, which is for the rest identical to pETKUL2 (Figure 2.1). However, we expected the frequency of promoter-tagged lines under baseline conditions to be comparable between these vectors. An increase of 3.3-fold in the number of promoter-

tagged lines was observed with pKCKUL1 compared to pETKUL2 for the lines showing strong BLA. The total tagging frequency using pKCKUL1 was 1.6- and 4.3-fold higher than with pETKUL2 in the two independent tagging experiments performed in the THP banana cultivar (Table 3.1). T-DNA tagging using a modified tagging construct including the presence of an intron in the *uid*A reporter gene, 3' intron splice sites upstream of the start codon of the *uid*A, and the kozak consensus around the start codon, showed an increase of 2.7-fold in reporter gene activation frequency in *Medicago truncatula* (from 15% to 41%, Scholte 2002). Similarly, the activation frequency of GUS activity was increased by 2-fold in rice when using an intron with splicing sites upstream of the *uid*A ORF (from 2% to 4.1% in leaves and roots, Jeon *et al.* 2000).

An increase of 3.5-fold in luminescence was observed in the strongest LUC activated pKCKUL1 tagged lines compared to the strongest LUC activated pETKUL2 lines (Figure 3.2). A comparable 2.4-fold increase in LUC activity was obtained by Bourdon *et al.* (2001) when comparing the *luc*$^{+INT}$ with the *luc*$^{+}$ in maize, both driven by the maize *Ubi*1 promoter plus leading intron. The strong LUC activity and the high LUC activation frequency in the promoter-tagged lines transformed with pKCKUL1 indicate that tagged promoters with weak activity might be overlooked in pETKUL2 transgenic cell colonies. The high LUC activity in lines transformed with the pKCKUL1 is probably due to an enhanced translation efficiency of the *luc*$^{+INT}$. Different steps including transcription, polyadenylation, mRNA export, translational efficiency, and the rate of mRNA decay are affected by the presence of introns (Nott *et al.* 2004). Bourdon *et al.* (2001) observed that the steady-state mRNA levels were comparable between the *luc*$^{+}$ and *luc*$^{+INT}$ genes indicating that the higher LUC activity with *luc*$^{+INT}$ was due to an increase in translation efficiency.

Two to three months after *Agrobacterium* transformation, antibiotic-resistant cell colonies were screened for LUC activation when undergoing a temperature drop from 26°C to a LT (from 18°C to 8°C, and even 4°C after a week acclimation at 16°C). The proportion of THP cell colonies showing an enhanced low temperature LUC activity (ELTLA) (of the total number of cell colonies showing LUC activity at LT) at the different LT treatments was ranging from 10.4% to 41.2% and from 25.0% to 48.7% when using the pETKUL2 (*luc*$^{+}$, Exps. 10, 11, and 15) and pKCKUL1 (*luc*$^{+INT}$, Exp. 15) tagging constructs, respectively (Table 3.2). When comparing the tagging constructs used in experiment 15, where the same number of samples (5-cm PD) were used for the different LT treatments (18°C, 16°C, 12°C, 8°C and 16°C→4°C), comparable proportions of ELTLA were obtained in most of the treatments (except for the 18°C treatment). However, a slightly higher frequency was obtained in the pKCKUL1 than in the pETKUL2 tagged lines. Therefore, the results suggest that the presence of an intron is not interfering drastically with the proportion of ELTLA tagged lines.

Different proportions of ELTLA were observed in experiment 15 for the different LT treatments (16°C and 16°C→4°C>18°C>12°C and 8°C). Similarly, for the *status quo* lines different proportions were obtained at different LTs (18°C>12°C, 16°C and 16°C→4°C>8°C) for both tagging constructs (Table 3.2). However, it is difficult to estimate until what level the effects of either the promoter activity or of the luciferase-luciferin enzymatic light reaction, which is temperature-dependent (see below), contribute to the different LUC activity patterns observed. Independent of the luciferase-luciferin enzymatic reaction, the different patterns of LUC activity detected in the different experiments and at different temperatures tested reflect that banana promoters with different response to LT have been tagged.

Promoter-tagged lines of experiment 10 were characterized for LUC activity at an early undifferentiated stage for BLA and their pattern at 8°C was analyzed. Promoter-tagged lines showed different levels of BLA, indicating the tagging of banana promoters with different strength. After removal of the baseline LUC activity (BLA) colonies in this experiment, the remaining cell colonies were screened one month later for LUC activity at 26°C for two hours before subjecting them to a LT of 8°C while monitoring LUC activity in real-time. Despite the early removal of BLA positive colonies, BLA (26°C) was detected at developmental stage I for all lines except line ET2-42 though at variable levels from weak to very strong reflecting the strength of the tagged promoters. Most likely, one month earlier the BLA of these colonies was masked by surrounding colonies or they did not show any BLA. In addition, up-regulation of LUC activity was detected in the promoter-tagged lines ET2-42 and ET2-17 when subjected to 8°C. The cell colonies transformed with the enhanced 35S promoter driving the *luc*$^{+}$ showed a decrease of LUC activity at 8°C. Similarly, most of the promoter-tagged lines showed the same activity pattern at 8°C. Promoter-tagged lines showed similar or higher BLA than cell colonies transformed with the enhanced 35S promoter. The pattern of LUC activity at 8°C was confirmed by quantitative analysis *i.e.* software-based measurements of the light emission by the cell colonies. Real-time measurements of LUC activity confirmed the different LUC activity patterns observed in the promoter-tagged lines. A constant decrease of LUC activity was observed at 26°C in the ET2-34 tagged line at early undifferentiated stage (Figure 3.5 I) prior to the onset of the LT treatment. This might be attributed to the lack of equilibrium between the LUC activity and the luciferin inflow (Van Leeuwen *et al.* 2000) although the LUC activity stabilized once the LT was reached. Kim *et al.* (2001) showed that the activity of the CaMV 35S promoter fused to the *uid*A reporter gene in transgenic tobacco as measured by northern blot analysis remained constant after a 24 h cold (4°C) treatment. A comparable result was observed by Kim *et al.* (2002) in Arabidopsis plants transformed with the *uid*A reporter gene driven by the CaMV 35S promoter. Histochemical GUS expression was comparable between the 23°C and 3°C treatments. These data suggest that the different LUC activity patterns obtained in the banana promoter-tagged lines at the LT

treatments are not completely due to the promoter activity, because a down-regulation of LUC activity was observed when decreasing the temperature to 8°C in lines containing the enhanced 35S promoter driving the *luc*$^{+}$. The light intensity of the firefly luciferase reaction as measured in solution is optimal in the temperature range of 22.5°C-25°C, while the luminescence decreases when the temperature deviates from this optimal range (McElroy and Seliger 1961; Dickinson *et al.* 1993; Ueda *et al.* 1994). Similarly, different reports indicate that the luminescence production is temperature dependant using the bacterial *luxCDABE* reporter system *in vivo* (Maoz *et al.* 2002; Bresolin *et al.* 2006). Therefore, an important factor that may explain the decrease in LUC activity in the cell colonies transformed with the enhanced 35S promoter driving the *luc*$^{+}$ might be the temperature dependency of the luciferase-luciferin light reaction. A correction on the LUC activity during the temperature regime was performed using the data obtained from the positive control line for experiment 10. Therefore, a higher up-regulation of LUC activity was obtained in line ET2-17 than originally calculated (from 10.7- to 19.6-fold). Both qualitative scoring and quantitative measurements using image analysis software revealed a decrease of LUC activity in lines transformed with the enhanced 35S promoter driving the *luc*$^{+}$. Because of the abovementioned temperature dependency of the LUC reaction, the promoters tagged in lines showing a comparable decrease in LUC activity upon LT treatment to the lines transformed with the enhanced 35S promoter might not be responsible for the decrease in LUC activity. In other words, the LUC enzymatic activity rather than the promoter activity is affected by low temperature.

The effect of LT on positive control lines, which contains the enhanced 35S promoter driving the *luc*$^{+}$ (pETKUL3) or the maize *Ubi*1 promoter driving the *luc*$^{+INT}$ (pKCKUL2), was tested (Figure 3.6). LUC activity analysis of pETKUL3 lines was performed in three independent experiments. An unexpected low BLA (LUC activity at 26°C) was detected in experiment 15. Up to a 70-fold difference in BLA was detected in the different experiments and might be the result of other factors besides the size effect of the cell colonies assayed (larger cell colonies in experiment 10 than in experiments 15 and 11 which should lead to higher LUC activity). Almost half of cell colonies were present in the experiments 15 and 11 than in experiment 10 per sample analyzed and might be also a reason for obtaining less BLA in the former experiments. Thus, further quantitative analysis in extracts from cell colonies should be performed in order to normalize the values for total protein concentration and avoid the cell colony size and number differences. Within each experiment, the pattern of the LUC activity at LT could vary according to the temperature tested. In general terms, when the LTs applied were still relatively high (16°C or 12°C), an up-regulation or a *status quo* was observed in the pETKUL3 lines. Low temperatures reaching 8°C showed in most of the samples a down-regulation of LUC activity. The temperature dependency of the luciferase-luciferin enzymatic light reaction might also be an important factor for the patterns observed here (see above). The positive control lines of

the banana cultivars Will and Caca containing the *luc*$^{+INT}$ under control of the maize *Ubi*1 promoter showed a similar temperature-LUC enzymatic activity dependency pattern as observed in the pETKUL3 lines. LUC activity of pKCKUL2 lines is comparable between 16°C and 26°C, but is lower at 12°C. Similarly, promoter activity studies using the *uid*A reporter gene under control of the maize *Ubi*1 promoter revealed lower GUS activity at 8°C and 4°C than at 26°C and 24°C in banana transformed cell colonies (Figure 6.7) and in wheat (4°C, Oullet *et al.* 1998), respectively. Further analysis of the activity of the maize *Ubi*1 and CaMV35S promoters at LT should be verified to compare the different activity patterns obtained in promoter-tagged lines.

LUC screening analysis revealed a fast up-regulation of LUC activity upon LT. The promoter-tagged line ET2-17 showed an immediate increase in LUC activity when changing the temperature from 26°C to 8°C (Figure 3.5 I). Therefore, the promoter tagged in the ET2-17 line should be responsive within 20 min, which was the integration time for the acquisition of one LUC image. Low temperature induced genes (and their promoters) could be classified as early or late responsive (Table 1.1 and Figure 1.1). Induction of CBF/DREB1 genes could be detected after 30 min of LT reaching a maximum level after 1 h (4°C, Medina *et al.* 1999). On the other hand, late responsive genes such as the *RD29A* gene could be induced by LT (4°C) within two to five hours (Shinwari *et al.* 1998; Yamaguchi-shinozaki and Shinozaki 1994). Therefore, the banana promoters showing a fast up-regulation in LUC activity at different LTs are classified as early LT responsive promoters.

3.3 Screening for luciferase activity during development and low temperature treatment

3.3.1 A selection of candidate promoter-tagged lines screened during embryo induction

Cell colonies of experiment 10 were regenerated into *in vitro* plants by subculturing through different media (paragraph 2.1.3.3) and screened at each stage for LUC activity in real-time during a drop in temperature of 26°C to 8°C. As explained in paragraph 3.2.2.2, after removal of the BLA positive lines, the remaining cell colonies were screened one month later during the transition from 26°C to 8°C. A total of 96 (0.61%) independent cell colonies showed either an increase (10), decrease (84) or *status quo* (2) relative to the LUC activity at 26°C when they were subjected to 8°C (Table 3.2). These 96 lines showing low temperature LUC activity (LTLA) were screened during the *in vitro* regeneration process (Table 3.5). Seventeen lines with LUC activity in the first screening at cell colony stage (I) were not active in the second screening (II) at embryo induction stage. And subsequently, 15 lines showing luminescence in the second screening did not maintain the activity in the third screening (III) which was also performed at the embryo induction phase (Table 3.5). As a result, 64 independent candidate promoter-tagged lines or 0.41% of the total number of independent lines screened were displaying LTLA at the time of embryo induction.

Table 3.5. Transgenic candidate promoter-tagged banana lines showing low temperature LUC activity (LTLA) when subjected to a temperature decrease from 26°C to 8°C at several stages during *in vitro* regeneration

Number of lines showing LTLA during the *in vitro* regeneration process (%)[1]						
I	**II**	**III**	**IV**	**V**	**VI**	**VII**
96 (0.61)	79 (0.50)	64 (0.41)	33 (0.21)	33 (0.21)	22 (0.14)	19 (0.12)

[1]The lines showing either an increase, decrease or *status quo* during the temperature change were counted and between brackets expressed in percentage of the total number of independent lines that were screened at stage I in experiment 10 which amounted up to 15,732. I: First screening of antibiotic-resistant cell colonies on ZZ medium three months after transformation. II and III: Second and third screening of cultures on RD1 medium five and seven months after transformation, respectively. IV, V and VI: Fourth, fifth and sixth screening of cultures on RD2 medium eight (IV and V) and ten months (VI) after transformation. VII: Seventh screening of regenerated *in vitro* plants maintained in REG medium 14-20 months after transformation.

Screening for LTLA in cell colonies was repeated during the *in vitro* regeneration process and the LUC activity patterns for 23 responsive lines are listed in Table 3.3. The results of the screenings II and III (embryo induction stage, RD1 medium) are not shown in this table because LUC activity at 26°C was not stable and the 8°C temperature was not reached, respectively.

3.3.2 Low temperature luciferase activity in shoot forming cultures

Differentiated cell cultures placed on RD2 medium for shoot formation were again screened in real-time for LUC activity under the same temperature regime as before (from 26°C to 8°C). Among the 64 lines showing LUC activity at stage III, only 33 lines were selected for further screenings (Table 3.5). These were tagged lines with high BLA and lines showing an up-regulation of LUC activity at LT. Most of the tagged lines showing very weak or no BLA were discarded. Eight lines did not survive after screening V and 22 out of the remaining 25 lines displayed LTLA at the sixth screening (Table 3.5, VI). Analysis of the fourth screening (IV) is shown in detail in Table 3.3 and Figure 3.4 as a representative screening for the shoot formation stage. The screening V was performed only 7 days later resulting in the loss of 8 lines (data not shown) due to the short period between these screenings. The qualitative LTLA analysis of the 23 independent lines for screening IV is shown in Table 3.3. Although tagged line ET2-42 showed an increase in LUC activity upon lowering the temperature to 8°C at the cell colony stage (Table 3.3 I and Figure 3.5 I), at the stage of shoot induction (screening IV) the LUC activity was lower at 8°C than at 26°C (Table 3.3). On the contrary, line ET2-17 was still showing an increase in LUC activity at 8°C as it was the case with the undifferentiated cell colony stage (Figures 3.4 IV and 3.5 IV). In the remaining candidate promoter-tagged lines the decrease in LUC activity observed at the early screening I also occurred at screening IV (Table 3.3 and Figure 3.4). It should be mentioned that the apparent lack of LUC activity at screening I and 26°C for line ET2-17 in Figure 3.4 I is due to the upper greyscale setting of 1,000 in the LUC images, while at screening IV a larger cell culture was screened in which the LUC activity was detectable at a 1,000 greyscale setting. Negative non-transformed control lines showed no detectable LUC activity. The positive control line (enhanced 35S promoter driving the *luc*$^{+}$) showed a moderate LUC activity, while a decrease of the activity was detected when lowering the temperature to 8°C (Table 3.3, Figures 3.4 IV and 3.5 IV).

The level of BLA varied greatly for almost all tagged lines when comparing the activity between the screenings I (cell colony) and IV (shoot formation; Table 3.3). A high proportion of tagged lines showed an increase in the BLA (13/23 lines), while in only two lines the BLA was decreased at shoot formation. Finally, eight tagged lines showed a similar BLA activity at screening I and IV. For instance, in tagged line ET2-34, LUC activation at 26°C reached a very strong and strong level at cell colony stage and shoot formation stage, respectively. However, in tagged line ET2-85 the reverse pattern was observed in which the BLA level increased with development since a weak and very strong BLA was detected at screening I and IV, respectively. In contrast, BLA remained very strong in line ET2-156 throughout all stages so far.

3.3.3 Low temperature luciferase activity in regenerated *in vitro* plants

Fully regenerated *in vitro* plants were screened for LTLA (Screening VII, Table 3.3 and Figure 3.5). Three lines, which showed luminescence during the whole regeneration process (ET2-28, ET2-37, and ET2-62), were not preserving LUC activity at the *in vitro* plant stage (Table 3.5, VII). A high proportion of promoter-tagged lines (12/23) showed a decrease in BLA at developmental stage VII (*in vitro* plant) when compared with the BLA at stage IV (shoot induction). A comparable number of tagged lines showed either a higher (5/23) or a similar (6/23) BLA at screening VII compared with screening IV (Table 3.3). The analysis of the 23 selected tagged ET2 lines is shown in Table 3.3 and the corresponding images of five candidate promoter-tagged lines are shown in Figure 3.4. Upon cold treatment (8°C) at screenings I and IV an increase in LUC activity was observed in line ET2-17 which was lost when screened at *in vitro* plant stage (Table 3.3 and Figure 3.4, screening VII). In contrast, all other lines including the positive control line showed a consistent decrease in LUC activity when lowering the temperature throughout the *in vitro* regeneration stages shown in Table 3.3 in all tissues (Figure 3.4). However, the level of LUC activity was not homogenous within the tagged lines ET2-17, ET2-85 and ET2-156, and in the positive control line at *in vitro* plant stage (Figure 3.4 VII). LUC activity was limited to the pseudostem of the *in vitro* plant for line ET2-17 (Figure 3.4). The highest LUC activity was confined to the basal parts of the *in vitro* plant including the corm and roots tissues (Figure 3.4 VII) in tagged lines ET2-85 and ET2-156, which showed a strong and very strong LUC activity at BLA, respectively (Table 3.3 VII, Figure 3.4 VII). In addition, LUC activity within the leaves of tagged lines was also not homogenous as illustrated for line ET2-156, where regions in the leaves with strong LUC activity could be observed (Figure 3.4 VII). Screenings of at least 10 different clones from the same tagged line were performed to confirm the LUC activity pattern.

Analysis in several screenings at *in vitro* plant stage revealed a common expression pattern for the tagged lines. The line ET2-34 showed a consistent decrease of LUC activity at 8°C measured in six screenings and in at least 24 clones of the same line. However, two clones showed an increase of LUC activity (data not shown). One of the clones depicted a transient increase of LUC activity once 8°C was reached and localized in the pseudostem reaching a peak after five hours at 8°C. The other clone, which was a not fully regenerated *in vitro* plant (a shoot-like structure), showed an increase of LUC activity immediately after changing the temperature to 8°C and remained high when 8°C was reached. However, further LUC screenings using the same ET2-34 clones that showed an increase in LUC activity at 8°C revealed no increase of the activity at LT and their LUC activity pattern remained similar to that observed in Figure 3.4 VII for the tagged line ET2-34. Furthermore, screening of other shoot-like clones of line ET2-34 showed also a decrease in LUC activity at 8°C.

3.3.4 Quantitative time course analysis of LUC activity throughout development

In order to verify the LTLA pattern of promoter-tagged lines during development (Table 3.3 and Figure 3.4), a more detailed quantitative time course analysis of LUC activity was performed. This analysis is shown for four tagged lines and a positive control line in Figure 3.5. These lines are shown since they represent different LUC activity patterns (Table 3.3) and their candidate promoter sequences were isolated and characterized (Chapter 4). The data at the undifferentiated cell colony (screening I), shoot induction (screening IV) and plant (screening VII) stages are presented (Figure 3.5). Measurements of light emission from the region of interest (ROI) were performed during the last 2 h at 26°C, during the temperature change to 8°C which took approximately one hour and for a period of 3 h at 8°C (paragraph 3.3.3.2). Confirmation of LTLA patterns were performed by calculation of the $RLU_{8°C}/RLU_{26°C}$ ratio using the average LUC activity at 8°C and 26°C excluding the measurements during the temperature decrease (Table 3.4).

The LUC activity in line ET2-17 increased immediately after changing the temperature to 8°C at cell colony (from 1128 ± 192 to 12,071 ± 773 RLU, 10.7- fold) and shoot induction stage (from 83,476 ± 6917 to 210,277 ± 4208 RLU, 2.5-fold), while at *in vitro* plant stage the LUC activity decreased (from 241,812 ± 3011 to 88,295 ± 904, 0.4-fold). In addition, when these ratios were corrected or normalized for the apparent decrease in LUC activity of the positive control line at 8°C (Table 3.4), a 19.6- , 10.3-, and 1.5- fold up-regulation of LUC activity at 8°C occurred in this line at cell colony, shoot induction, and *in vitro* plant stage, respectively. Although having different LUC activity levels at 26°C, the promoter tagged lines ET2-34, ET2-85, and ET2-156 showed a lower LUC activity at 8°C during all the developmental stages comparable to the positive control line which contains the luc^+ driven by the enhanced 35S promoter. A transient increase of LUC activity was observed in lines ET2-85 and ET2-156 during the temperature decrease from 26°C to 8°C at shoot induction stage (screening IV). However, in none of the lines this transient increase in LUC activity occurred at plant stage (screening VII). A strong decrease in BLA at 26°C was observed for line ET2-85 at screening IV, while the LUC activity at 8°C stabilized at a lower level (Figure 3.5). The BLA level of the positive control line at cell colony stage (screening I) was above that of lines ET2-17 and ET2-85, while it was lower than the BLA of all four tagged lines at the shoot induction phase (screening IV) and almost undetectable at *in vitro* plant stage (screening VII).

3.3.5 Restoration of baseline luciferase activity after low temperature treatment

Promoter-tagged lines were screened in real-time for LTLA throughout the *in vitro* regeneration process (sections 3.2.3 and 3.3.1 to 3.3.4). To check whether the LUC activity level returned to the original BLA, the temperature was raised again to 26°C after the 8°C temperature treatment. The complete time course of LUC activity during the temperature

regime for four promoter-tagged lines at two different developmental stages (shoot induction and plant stage) is shown in Figure 3.7. The up-regulated LUC activity in line ET2-17 at shoot induction stage (screening IV) remained high for 10 h at 8°C and then slowly decreased but remained more than twofold higher than the BLA before the temperature decrease after 23 h at 8°C. Upon setting the temperature to 26°C again, the LUC activity strongly decreased within one hour to a level below that of the BLA before the 8°C treatment (Figure 3.7A) demonstrating the up-regulation of the tagged promoter in this line by LT. The decrease in LUC activity in line ET2-34 induced by 8°C at this developmental stage reversed to the BLA level when switching the temperature back to 26°C (Figure 3.7A). The LUC activity pattern during the 24 h period of 8°C was similar to that of the control. However, LUC activity in lines ET2-85 and ET2-156 did not change upon increasing the temperature following the 8°C treatment. Only a small transient increase in LUC activity immediately after the temperature raise was observed and was comparable to the positive control line. On the contrary, at plant stage (screening VII) the LUC activity, which decreased in all the tagged lines as well as in the positive control line by the 8°C treatment, was restored in all lines when setting the temperature back to 26°C albeit to a level below that of the BLA (Figure 3.7B). The reaction to the increasing temperature was immediate in the tagged lines as well as in the positive control line. For tagged lines ET2-34, ET2-17 and the positive control line the restored LUC activity did not reach a stable level yet, whereas in lines ET2-85 and ET2-156 it reached a stable level 5-6 h after the temperature raise. LUC activity for the positive control line showed an unstable pattern throughout the temperature regime in developmental stage VII. As shown in Figure 3.4 VII, the positive control line showed a very weak LUC activity at BLA and was undetectable when reaching 8°C. Due to the very weak LUC activity in the positive control line (close to the values of the non-transformed plant, data not shown) that was limited to the basal area of the *in vitro* plant (Figure 3.4 IV) and to the large ROI, the RLU values obtained after subtracting the RLU of a negative non-transformed control line were variable, although a pattern could be observed. The LUC activity patterns during the two changes in the temperature are distinctive in the *in vitro* plant stage. When restoring the temperature to 26°C after the LT period, a rapid enhancement of LUC activity was detected reaching a peak immediately after 26°C. The peak of LUC activity is the highest during the temperature regime in ascending order for the lines ET2-17, ET2-34, ET2-156 and the positive control line, while for the ET2-85 line the activity was increased but to a lower level than in the first 26°C period (Figure 3.7B).

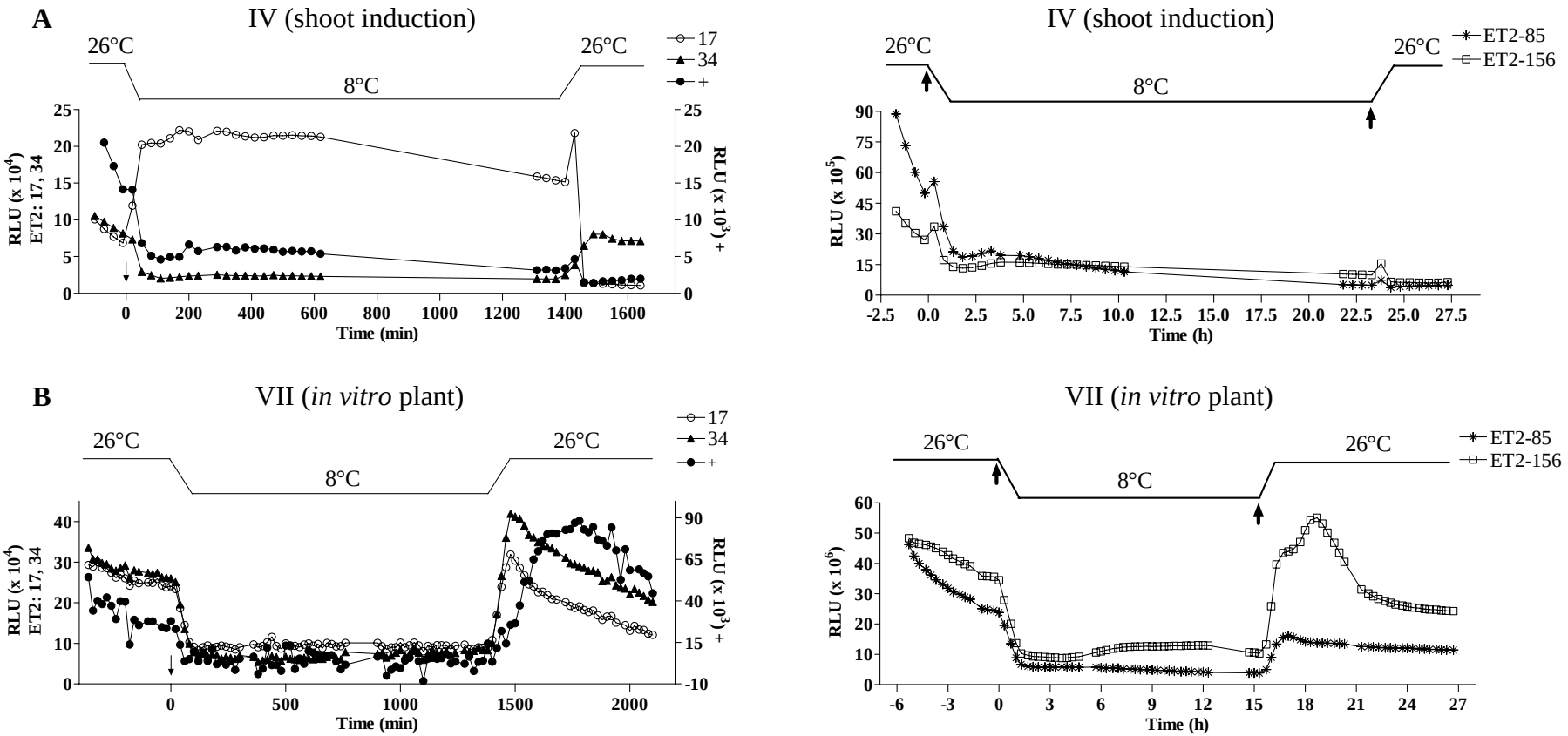

Figure 3.7. Time course of luciferase (LUC) activity during temperature regime in four candidate promoter-tagged lines (ET2-17, ET2-34, ET2-85 and ET2-156) and a positive control line (+, enhanced 35S promoter) at two *in vitro* developmental stages (IV and VII). Three images of 20 min exposure time were recorded per hour. The LUC activity expressed in relative light units (RLU) was corrected for the background measured in a negative, non-transformed control line. The region of interest for quantification of LUC activity was 0.58 and 23.19 cm^2 for shoot induction (A, screening IV) and plant (B, screening VII) stage, respectively. The Y axis scale for each line is indicated between brackets when necessary. LUC activity was monitored for 2 h at 26°C, at time point zero (indicated by an arrow), temperature was set to 8°C, which was reached 1 h later. The 8°C temperature was maintained for 23 h and then raised again to 26°C (indicated by an arrow). This temperature regime is schematically shown by the solid line above the graphs. At plant stage (B) LUC activity was monitored for 6 h at 26°C, the temperature was set to 8°C and maintained for 23 h (ET2-17, ET2-34 and +) or 15 h (ET2-85 and ET2-156) before changing again to 26°C.

3.3.6 Recapitulation of low temperature LUC activity pattern in re-initiated cell cultures from candidate promoter-tagged lines

To corroborate the expression pattern under LT stress obtained during the early *in vitro* regeneration stages, undifferentiated cell cultures resembling cell colonies of screening I on ZZ medium (paragraph 3.2) were induced from the regenerated promoter-tagged lines. First proliferating meristem cultures were re-initiated from the apical meristems of the *in vitro* multiplied plants essentially as described by Strosse *et al.* (2003), but during the two initial subculture cycles the meristems were kept on TDZ medium before transfer onto P4 medium (Table 2.1, paragraph 2.1.3.4) for further maintenance. Small-sized (1-2 mm^3) explants consisting of multiple meristems were then excised and subcultured on ZZ medium (paragraph 2.1.3.4) for embryogenic callus induction. Eleven months after starting the re-initiation process the differentiated P4 cultures (IP4) together with the re-established undifferentiated ZZ cultures (IZZ) of four independent candidate promoter-tagged lines were screened for LUC activity under LT stress (Figure 3.8). Due to high temperatures in the room where the LUC screening was performed, the circulating water responsible for the temperature control of the temperature-controlled plate (paragraph 2.2.4.1) was not able to maintain the samples at 8°C. Instead, a temperature ranging from 11°C to 9°C was obtained. At the IP4 stage of proliferating meristems the LT of 11-9°C did not affect the LUC activity in any of the lines tested similar to the positive control line (Figure 3.8A). It should be noted that in these differentiated meristem cultures LUC activity driven by the enhanced CaMV 35S promoter in the positive control remained weak. The quantitative time course analysis supported these findings except for line ET2-156 that showed a somewhat lower LUC activity after 6 h at 11-9°C than at 26°C (20,538,600 ± 493,256 RLU *vs*. 28,174,683 ± 310,402, respectively; Figure 3.8B). A minor transient increase in LUC activity was detected in the other three candidate promoter-tagged lines during the drop in temperature. Precise calculation of the fold change in LUC activity under LT stress expressed as $RLU_{11\text{-}9°C}/RLU_{26°C}$ agrees with this qualitative assessment (Table 3.6).

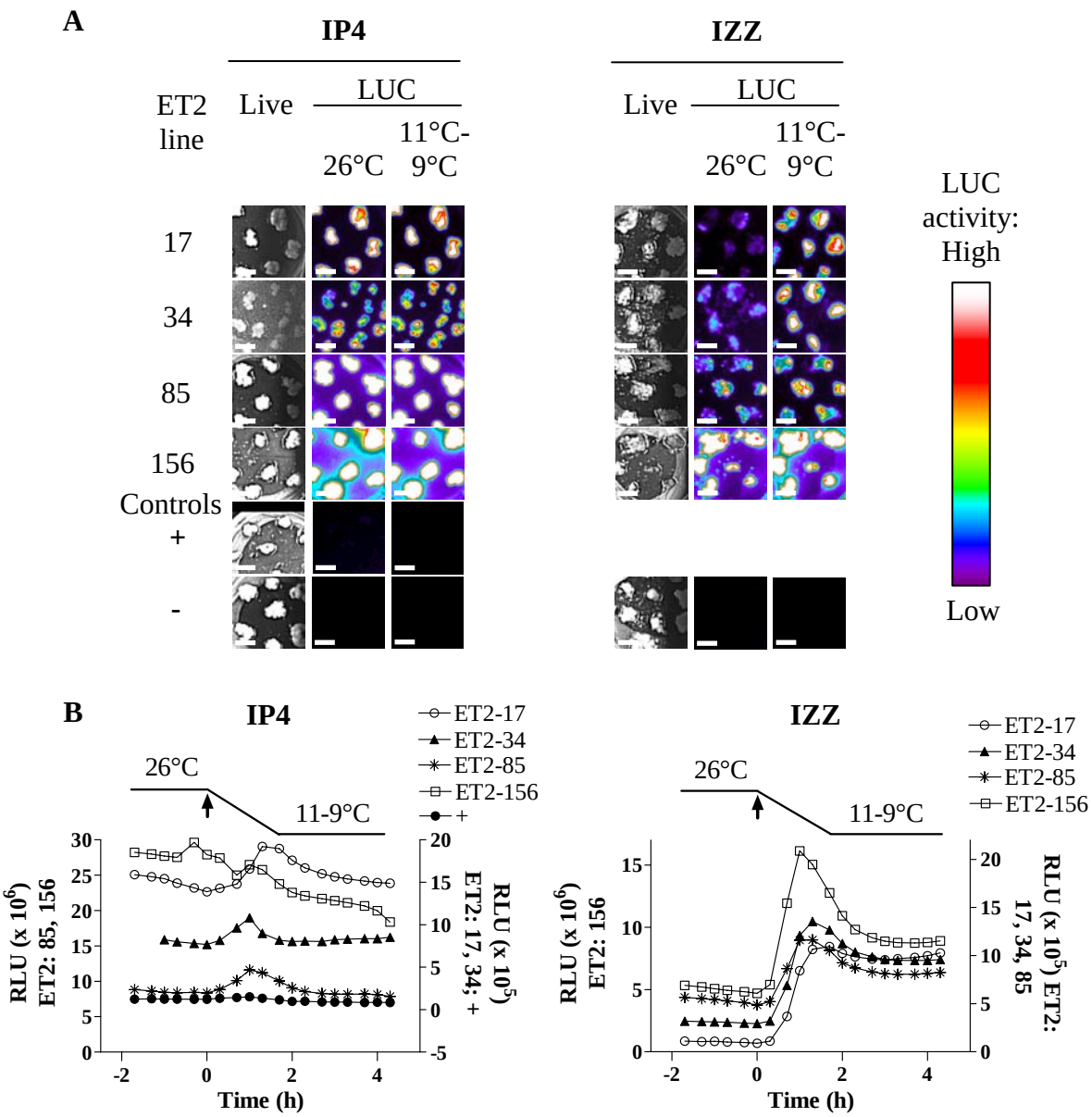

Figure 3.8. Luciferase activation at 26°C and 11°C-9°C in proliferating meristem cell cultures of candidate promoter-tagged lines maintained in P4 or ZZ medium and re-initiated from the apical meristems of the *in vitro* multiplied promoter-tagged plants. Four promoter-tagged lines (ET2-17, ET2-34, ET2-85 and ET2-156) transformed with pETKUL2 (promoterless *luc*$^+$ gene), a positive control line carrying the *luc*$^+$ gene under control of the enhanced CaMV 35S RNA promoter (+), and a negative non-transformed control line (-) were analyzed. Re-initiated cell culture stage IP4 and IZZ correspond to cell cultures maintained at least 9 and 5 months on P4 and ZZ medium, respectively. (A) Representative images were taken under normal light (Live) and dark (LUC) conditions. Luciferase (LUC) images were recorded after 1-2 h and 3 h at 26°C and 11°C-9°C, respectively, and are depicted in pseudocolors (see color bar) with an upper greyscale limit setting of 1000. Scale bars represent 1 cm. (B) Time course of LUC activity was monitored for 1-2 h at 26°C, at time point zero (indicated by an arrow) temperature was set to 11°C-9°C, which was reached 1.5-2 h later and then maintained for 3 h (solid lines above the graphs). Three images were recorded per hour at 20 min interval. LUC activity expressed in relative light units (RLU) was corrected for the background measured in a negative, non-transformed control line. The region of interest for quantification of LUC activity was 14.82 cm^2. The Y axis scale for each line is indicated between brackets.

Following correction for the apparent decrease in LUC activity in the positive control line the $RLU_{11\text{-}9°C}/RLU_{26°C}$ values for ET2-17, ET2-34 and ET2-85 IP4 cultures (1.5, 1.5 and 1.4, respectively; Table 3.6) indicated a small up-regulation in LUC activity by LT. On the other hand, real-time monitoring clearly revealed a stimulatory effect of the 11-9°C temperature stress on the LUC activity in all the candidate promoter-tagged lines at the stage of undifferentiated IZZ cultures (Figure 3.8A). However, the degree of up-regulation was higher in line ET2-17 than in the other three candidate promoter-tagged lines. Six hours after the change in temperature the up-regulation amounted up to approximately 10-fold and 2- to 4-fold, respectively, with an even stronger up-regulation during the temperature transition (Table 3.6). These findings were again supported by the $RLU_{11\text{-}9°C}/RLU_{26°C}$ values although the level of up-regulation was below 2.0 in ET2-85 and ET2-156. Unfortunately, no IZZ cultures of positive control lines were finally established (Figure 3.8) preventing the calculation of a corrected $RLU_{11\text{-}9°C}/RLU_{26°C}$ value (Table 3.6). Comparison of the non-corrected fold change in LUC activity ($RLU_{11\text{-}8°C}/RLU_{26°C}$) between the originally tagged cell colony (ZZ, screening I, Table 3.4) and the re-initiated cell colony-like culture (IZZ, Table 3.6) revealed a strong correlation for line ET2-17 (10.7 *vs.* 9.6, respectively). In contrast, for the remaining three candidate promoter-tagged lines ET2-34, ET2-85 and ET2-156 the repression of LUC activity in the original cell colony was reversed to a relatively small up-regulation in the re-established cell colony-like culture (0.4, 0.4 and 0.2 *vs.* 3.1, 1.5 and 1.8, respectively).

Table 3.6. Fold changes in LUC activity in response to low temperature (LT) stress in proliferating meristem cell cultures re-initiated from the apical meristems of the *in vitro* multiplied plants of candidate promoter-tagged lines of experiment 10

		$RLU_{11\text{-}9°C}/RLU_{26°C}$				
Stage[a]	**Feature**	**ET2-17[b]**	**ET2-34[b]**	**ET2-85[b]**	**ET2-156[b]**	**+[c]**
IP4	Change	1.0	1.1	1.0	0.7	0.7
	cChange[d]	1.5	1.5	1.4	1.0	1.0
IZZ	Change	9.6	3.1	1.5	1.8	NT[e]

Real-time screening for LUC activity in four candidate promoter-tagged lines and one positive control line during the temperature decrease from 26°C to 11°C-9°C at developmental stage I-like proliferating cell cultures re-initiated from the apical meristems of *in vitro* multiplied plants. The average of the last six measurements of the LUC activity expressed in relative light units (RLU) at each temperature (26°C and 11°C-9°C) was used to calculate the fold change in LUC activity under LT. A $RLU_{11°C\text{-}9°C}/RLU_{26°C}$ value above one indicates an up-regulation of LUC activity, whereas a value less than one represents a down-regulation of LUC activity at LT.

[a]Re-initiated cell culture stages IP4 and IZZ correspond to cell cultures maintained at least 9 and 5 months on P4 and ZZ medium, respectively.

[b]Independent lines transformed with promoter tagging vector pETKUL2 which contains a promoterless *luc*$^+$ gene

[c]Line transformed with the positive control vector pETKUL3 which contains the enhanced CaMV 35S RNA promoter driving the *luc*$^+$.

[d]The LUC activity at LT was corrected by multiplying with the reverse of the fold change ($RLU_{26°C}/RLU_{11°C\text{-}9°C}$) of the positive control line (ET3) and used to calculate the corrected $RLU_{11°C\text{-}9°C}/RLU_{26°C}$.

[e]NT refers to not tested.

3.3.7 Discussion

In the present section (3.3) the LUC activity pattern under LT stress (8°C) was assessed during *in vitro* development in 23 candidate promoter-tagged ET2 lines of experiment 10. Some tagged lines that showed LUC activity in screenings at early developmental stages were not showing activity in late developmental stages (paragraphs 3.3.1, 3.3.2 and 3.3.3). The loss of LUC activity during the regeneration process indicates that the tagged promoters were active in certain developmental stages only. In addition, the first screening was performed on culture medium containing the hormones 2,4-D and zeatin, which were absent during the second and third screening (RD1 medium), suggesting that growth regulators may also affect the activity. On the other hand, the lack of LUC activity in lines active during the second but not the third screening may be due to developmental regulation of the promoters because both screenings were performed on the same RD1 culture medium that does not contain plant hormones. Nevertheless, the observed BLA profiles of some candidate tagged lines might be the result of regulation by one or more tissue culture components rather than by the developmental program of *in vitr*o regeneration. Bade *et al.* (2003) addressed this issue in promoter tagged *Brassica napus* lines where 6 out of 20 tagged promoters with callus-specific activity were also found to be auxin inducible when node segments of the regenerated plants were placed on NAA-containing medium. Since LUC screening is initiated at an early stage in cell colonies grown on medium containing a high 2,4-D concentration (5 µM), this kind of verification might be worth to perform on the tagged banana lines. Interpretation of the results will be however difficult because other components of the medium may play a role as well.

Analysis of the LUC activity at 8°C of selected tagged lines of experiment 10 during *in vitro* development revealed different LUC activity responses. The tagged line ET2-42 which showed an increase of LUC activity at the cell colony stage was not up-regulated at the shoot induction and *in vitro* plant stages (Table 3.3), suggesting that *in vitro* development and/or medium components in combination with LT stress regulate the activity of the promoter tagged in this line. On the other hand, in the tagged line ET2-17 which showed a strong up-regulation of LUC activity during LT at cell colony stage (10.7-fold, I), the LUC activity was also up-regulated at the shoot induction stage (2.5-fold, IV). However, the up-regulation of LUC activity at LT was lost at *in vitro* plant stage for this line. Comparison of the LUC activity throughout *in vitro* regeneration indicates that the strongest up-regulation of LUC activity occurred at the undifferentiated cell colony stage in line ET2-17. These real-time findings strongly indicate a developmental regulation of the tagged promoters in this line. In conclusion, as line ET2-17 became more differentiated, less up-regulation of LUC activity occurred upon LT stress confirming the developmental regulation of LUC activity. The other tagged lines showed a consistent down-regulation of LUC activity at the different stages screened indicating the stable activity pattern at LT of

the promoters tagged in these lines during *in vitro* development. A similar LUC activity pattern was observed in the lines transformed with the enhanced 35S promoter driving the *luc*$^{+}$.

A homogenous LUC activity was not detected in all tissues of *in vitro* plants for most of the tagged lines screened. LUC activity remained stronger in the pseudostem region for the tagged line ET2-17 and may reflect a tissue specificity of the promoter (Figure 3.4 VII). On the other hand, the tagged line ET2-34 depicted a homogenous but weak LUC activity in all the tissues of the *in vitro* plants. Luciferase activity is dependent on the presence of the substrate luciferin, Mg^{2+}, ATP and oxygen (DeLuca *et al.* 1974; Aflalo 1991) and their availability could be an important factor influencing the LUC activity in plants. The possible contribution in differences of the availability of each of the components needed for the luciferase-dependent light reaction was tested by Van Leeuwen *et al.* (2000) in order to determine their effect in the variegated *in planta* LUC activity in *Petunia* leaves. Application of three-times pre-sprayed 1 mM luciferin solution is sufficiently for the luciferin availability in plant tissues (Van Leeuwen *et al.* 2000). Two applications of 0.1 mM luciferin 26-14 h and 1-0.5 h was applied to the banana lines prior screening for LUC activity (paragraphs 2.2.2 and 2.2.3). The fact that LUC activity was homogenous in all tissues of the tagged line ET2-34 *in vitro* plant suggests that luciferin penetration was homogenous in all the *in vitro* plant tissues. Variegated LUC activity was almost not detected in undifferentiated tissues of banana tagged lines (Figures 3.4 and 3.8A). However, *in vitro* plants reflect different LUC activity according to the tissue. The availability of the other components (Mg^{2+}, ATP and oxygen) are not critical for luciferase light reaction in plant tissues and the effect of variegated LUC activity is not due to local difference of the substrates (Van Leeuwen *et al.* 2000). Variegated LUC activity was related to variegated *luc* transgene transcription (Van Leeuwen *et al.* 2001). This variegated *luc* transgene transcription might be a combination of different factors including promoter activity, influence of T-DNA flanking sequences in independent transformants and local modulations of the temporal regulation of transgene activity due to a different integration site in the genome. All these factors are commanded by chromatin remodeling or availability of transcription factors (Van Leeuwen *et al.* 2001). Therefore, specific strong activity of the promoters of the tagged lines ET2-85 and ET2-156 might be delimited to regions of the basal part of the *in vitro* plants including corm and roots (Figure 3.4 VII). In addition, variegated LUC activity was also detected when comparing the activity within the leaves of *in vitro* plants (Figure 3.4 VII, ET2-156). Some regions of the leaves showed higher LUC activity than others and might be attributed to a variegated *luc* transcription (see above). However, the difference in LUC activity in different regions of the leaves might be also attributed to an effect of injury allowing better penetration of luciferin and/or higher LUC activity if the tagged promoter is wound inducible. Leaves of promoter tagged lines might be accidentally wounded during the transfer of the *in vitro* plants from the test

tubes containing REG medium to the 5-cm PD (paragraph 2.2.4.5). Therefore, analysis of protein extracts might clarify if the promoter of the tagged lines showing variegated LUC activity has tissue specific activity, or is wound inducible or that the increased LUC activity was due to a better penetration of luciferin at the wounded site.

The level of BLA in the positive control lines was relatively weak throughout *in vitro* regeneration (Figures 3.4 and 3.5) with often an undetectable light emission in the regenerated plant pointing in the direction of possible silencing of the enhanced 35S promoter. The following findings support this hypothesis: (i) the enhanced 35S promoter has an activity level in banana between that of the weak CaMV 35S promoter and the strong maize *Ubi*1 promoter (Ságі *et al.* 1995a, 1995b; Remy *et al.* 1998b); (ii) transgenic banana lines carrying the luc^{+INT} gene under control of the constitutive maize *Ubi*1 promoter did not show remarkable changes in BLA since it remained very strong up to the regenerated plant (data not shown); (iii) besides the luc^{+} gene the CaMV 35S promoter drives the *neo* selectable marker gene in pETKUL3 (Remy *et al.* 2005), which makes it more prone to transcriptional silencing by methylation and this is worsened in case of multiple T-DNA copies mainly when these are arranged in inverted repeats *via trans* silencing (reviewed by Muskens *et al.* 2000).

A markedly constant decrease of LUC activity was observed at 26°C in some tagged lines including ET2-85 and the positive control line at shoot induction stage (Figure 3.7A) prior to the onset of the LT treatment. This may be attributed to the lack of equilibrium between the LUC activity and the luciferin inflow (Van Leeuwen *et al.* 2000). However, the LUC activity stabilized once the LT was reached at all developmental stages studied irrespective of the reaction (up- or down-regulated) and the level of activity (Figure 3.6 and 3.7). In the promoter tagged lines ET2-17, ET2-34 and ET2-156, and in the positive control line at *in vitro* plant stage (Figure 3.7B) in which the LUC activity was restored after the LT treatment, maximum LUC activity occurred during the temperature rise and then the reaction stabilized in the following hours during the second 26°C treatment. This LUC activity pattern might be due to a post-transcriptional regulation. The post-transcriptional regulation of gene expression includes mRNA splicing and editing, transcript stability, translation efficiency, and modification of protein turnover rate (Gallie 1993). In addition, mRNA stability is believed to play a major role in regulation of gene expression in higher plants (Phillips *et al.* 1997). In prokaryotes, low temperatures stabilize the secondary structure of mRNA particularly the 5'- untranslated region (UTR) preventing the ribosomes to recognize the Shine-Dalgarno sequence. Consequently, translation is blocked (Ermolenko and Makhatadze 2002) and a pool of non-translating ribosomes is obtained (Farewell and Neidhardt 1998). Since the rate of translation decreases at lower temperatures, a subsequent increase to 26°C will result in a strong enhancement of light emission in the first hours before reaching again a stable level of LUC activity.

The up-regulation of LUC activity at the undifferentiated cell colony stage for line ET2-17 was further corroborated by the recapitulation of the LT up-regulated LUC activity profile in cell colony-like cultures maintained on ZZ medium (IZZ). Cell colony-like cultures were re-initiated from the apical meristems of *in vitro* multiplied plants *via* proliferating meristem cultures as described (Strosse *et al.* 2003). However, the re-initiated proliferating meristem cultures of line ET2-17 induced on TDZ medium and maintained on P4 medium (IP4) did not show any LT up-regulated LUC activity, which indicates again that the LUC activity pattern at LT is developmentally regulated. Furthermore, the other three tested promoter tagged lines (ET2-34, ET2-85 and ET2-156) displayed a different LUC activity pattern at LT in the IP4 and ZZ screenings. The tagged lines showed a minor increase in LUC activity under LT stress at IZZ for all the promoter tagged lines tested, corroborating the LUC activity pattern dependency of the medium and cell type of the cultures. LUC images revealed an apparent higher BLA in the IP4 cultures than in the IZZ cultures of the tagged lines. This higher LUC activity may indicate that the promoters tagged are developmentally regulated and/or induced by one of the components of the P4 medium, which includes a high ratio of cytokining/auxin concentration (BA 100 µM / IAA 1 µM, respectively) *vs*. a low ratio auxin/cytokinin concentration in the ZZ medium (2,4-D 5 µM / zeatin 1 µM, respectively). Further quantification of LUC activity in extracts may allow a more reliable comparison between different developmental stages of a tagged line and between different lines and should discard differences due to size differences between the cultures.

In summary, approximately 89,000 independent transgenic banana cell colonies were screened for LUC activity in four independent experiments thereby obtaining an activation frequency from 0.17% to 2.69% with the tagging vectors pETKUL2 (promoterless *luc*$^{+}$) and pKCKUL1 (promoterless *luc*+^{INT}). A higher tagging frequency was obtained with pKCKUL1 than with pETKUL2. Cell colonies were screened under controlled temperature conditions in which LUC activity was monitored in real-time at 26°C followed by a gradual decrease to different temperatures including 18°C, 16°C, 12°C and 8°C. The LUC activity pattern was followed during the temperature regime and the different cell colonies were showing an enhanced, decreased or stable LUC activity at LT relative to that at 26°C. Of experiment 10, transgenic cell colonies responsive to 8°C were regenerated to plants and LUC screening was repeated during the *in vitro* regeneration process. At each regeneration step LT was applied and LUC activity was followed in real-time. The tagged lines ET2-17 and ET2-42 showed an enhanced LUC activity at 8°C at the cell colony stage. However, the LUC activity of the tagged line ET2-42 was not up-regulated at the shoot induction and *in vitro* plant stages. On the other hand, the tagged line ET2-17 showed a strong up-regulation of LUC activity by LT at the cell colony stage (10.7-fold) and shoot induction stage (2.5-fold), but not anymore at the *in vitro* plant stage. In addition, other tagged lines with different levels of LUC activity at 26°C showed a decrease

of the activity at 8°C during the different developmental stages of the *in vitro* regeneration process (ET2-34, ET2-85 and ET2-156). In conclusion, several T-DNA tagged lines were identified showing different levels and patterns of LUC activity at LT during different *in vitro* developmental stages.

In the following chapter estimation of T-DNA copy number and isolation of the flanking T-DNA sequences is first described for a selected number of promoter-tagged lines of experiment 10 (Table 3.4) which show different LUC activity levels and patterns at LT. Second, the T-DNA flanking sequences retrieved from the promoter-tagged lines ET2-17, ET2-34, ET2-85, ET2-156 and ET2-179 were analyzed with bioinformatic tools.

Chapter 4 Isolation and analysis of T-DNA flanking sequences in candidate promoter-tagged lines

4.1 Introduction

The advent of several banana genetic transformation systems in the last decade opened the possibility for a diversity of new applications including the introduction of genes conferring resistance against biotic or abiotic stresses. In addition, the employment of *Agrobacterium* for the isolation of genes or promoters is a new component for the functional analysis toolkit in banana (Remy *et al*. 2005). Relatively low T-DNA copy numbers are obtained *via Agrobacterium*-mediated transformation facilitating the recovery of T-DNA flanking regions for the identification of functional sequences. PCR-based methods are widely used for the isolation of the unknown T-DNA flanking sequences including inverse (I)-PCR (Triglia *et al.* 1988, Ochman *et al.* 1988), thermal asymmetric interlaced (TAIL)-PCR (Liu *et al.* 1995) and adaptor/cassette/linker-ligation PCR (Mueller and Wold 1989; Pfeifer *et al.* 1989; Espelund and Jakobsen 1992; Warshawsky and Miller 1994) and its derivatives. Long amplicons are required for successful identification of genes and promoters, especially in not fully sequenced genomes like that of banana.

Single T-DNA copy number lines are preferred in T-DNA tagging experiments because it allows a relatively easy identification of the tagged sequence. This is also demonstrated by the fact that in all reports dealing with T-DNA tagging, researchers analyze only single T-DNA copy lines in greater detail (Wei *et al.* 1997; Mudge and Birch 1998; Ökrész *et al.* 1998; Mollier *et al.* 2000; Webb *et al.* 2000; Farrar *et al.* 2003). However, more than one T-DNA copy is often integrated in the plant genome. Therefore, molecular characterization of such T-DNA tagged lines with an interesting reporter gene activity pattern is performed but at a lower scale than single T-DNA copy lines. For instance, Lee *et al*. (2004a) analyzed rice T-DNA tagged lines responsive to LT with two or even three T-DNA copies. In addition, further gene expression analysis of a multi T-DNA copy line was performed in order to determine which gene was responsible for the reporter gene activity observed in the tagged rice line (Lee *et al*. 2004a). Fifty percent or more of *Agrobacterium*-mediated transgenic banana lines contain more than one T-DNA copy (Ganapathi *et al.* 2001; Khanna *et al.* 2004; Pérez Hernández *et al.* 2006a; Huang *et al.* 2007; Remy *et al.* unpublished results). Thus, multiple T-DNA flanking sequences were expected in at least half of the candidate promoter-tagged banana lines (Chapter 3) making the isolation of the candidate promoter sequences more complex. Furthermore, complex T-DNA integrations have been reported in a wide variety of plant species including dicots [Arabidopsis (De Buck *et al.* 1999; Meza *et al.* 2002); tobacco (Kononov *et al.* 1997)] and monocots [bentgrass (Fu *et al.* 2006); rice (Kim *et al.* 2003; Zhu *et al.* 2006); barley (Stahl *et al.* 2002)]. Complex T-DNA integrations refer to tandem T-DNA repeats, vector

backbone and/or T-DNA rearranged integration events. Naturally, these types of T-DNA integrations also greatly complicate the recovery and identification of the correct flanking sequences.

Promoter-tagged lines showing LUC activity were identified and characterized during *in vitro* development and LT (Chapter 3). In this chapter, the T-DNA copy number determination and the isolation and analysis of T-DNA flanking sequences of 'Three Hand Planty' (THP) promoter-tagged lines of experiment 10 transformed with the pETKUL2 tagging vector are described. Southern blot hybridization analysis of DNA isolated from *in vitro* or greenhouse leaf tissue of putative promoter-tagged lines was employed for the estimation of the T-DNA copy number (paragraph 4.2), while TAIL- and I-PCR were used for the recovery of T-DNA flanking sequences (paragraph 4.3). In promoter-tagged lines with multiple T-DNA copies, isolation of multiple T-DNA flanking sequences was necessary to allow identification of the promoter sequence(s) responsible for the reporter gene activation. The physical linkage between the 5'-tagged flanking sequences and T-DNA integration was investigated by Southern blot analysis in the promoter-tagged line ET2-17 which showed an up-regulation of LUC activity by LT at early developmental stages and contained multiple T-DNAs. In addition, the physical linkage was also performed in the single T-DNA copy line ET2-34 showing a decrease of LUC activity by LT during the different developmental stages (paragraph 4.4.1.1). In addition, several types of complex T-DNA integration patterns were detected in the promoter-tagged lines (paragraph 4.4.1.2) and T-DNA border junction sites were analyzed in detail (paragraph 4.4.1.3). Finally, bioinformatic analysis of the T-DNA border flanking sequences revealed the presence of putative cryptic and functional promoters and their elements (paragraph 4.4.2).

4.2 Southern hybridization analysis

A selected number of promoter-tagged lines of experiment 10 were analyzed for T-DNA copy number by Southern hybridization (Table 3.3). The tagged lines analyzed showed either an up-regulation of LUC activity at early developmental stages (ET2-17) or a decrease of the activity by LT during the *in vitro* regeneration process. Most of these selected lines were showing a strong or very strong LUC activity (Table 3.3) in at least one developmental stage (ET2-34, ET2-49, ET2-85, ET2-111, ET2-123, ET2-156 and ET2-179) and only two promoter-tagged lines with a moderate LUC activity (ET2-82 and ET2-102; Table 3.3) were also analyzed. Leaf material from *in vitro* or greenhouse transgenic THP plants was excised for DNA isolation, and total DNA was digested with *Hind*III before separation by gel electrophoresis. A DIG-labeled probe specific for the luc^+ transgene was used for hybridization. Southern hybridization revealed a T-DNA copy number of at least one (ET2-34) to five (ET2-17 and ET2-156) averaging 3.6 in 10 independent lines tested (Figure 4.1). Despite the presence of multiple T-DNA copies in promoter-tagged lines, LUC activity either remained stable or even increased throughout *in vitro* regeneration (Table 3.3). On the other hand, the multiple T-DNA copy positive control line (+) that contained the luc^+ under control of the enhanced CaMV 35S RNA promoter displayed low or undetectable LUC activity (data not shown). No bands were present in the negative non-transformed line confirming the integration of the T-DNA in the promoter-tagged lines. More than one T-DNA copy is probably integrated at a single integration site in lines showing one or more band(s) with an intensity higher than that of the one copy plasmid reconstruction, which occurs in the majority of the tested lines (ET2-179, ET2-85, ET2-123, ET2-49, ET2-82, ET2-102 and ET2-156; Figure 4.1). A detailed explanation for the presence of more than one T-DNA copy per band is given in paragraph 4.4.1.2. Reproducible T-DNA integration patterns were obtained in promoter-tagged lines ET2-17 and ET2-156 using total DNA from pooled leaves of different *in vitro* plant clones (Figure 4.1A) and total DNA from a leaf of a greenhouse plant (Figure 4.1B) demonstrating the stability of the integrated T-DNA copies during development.

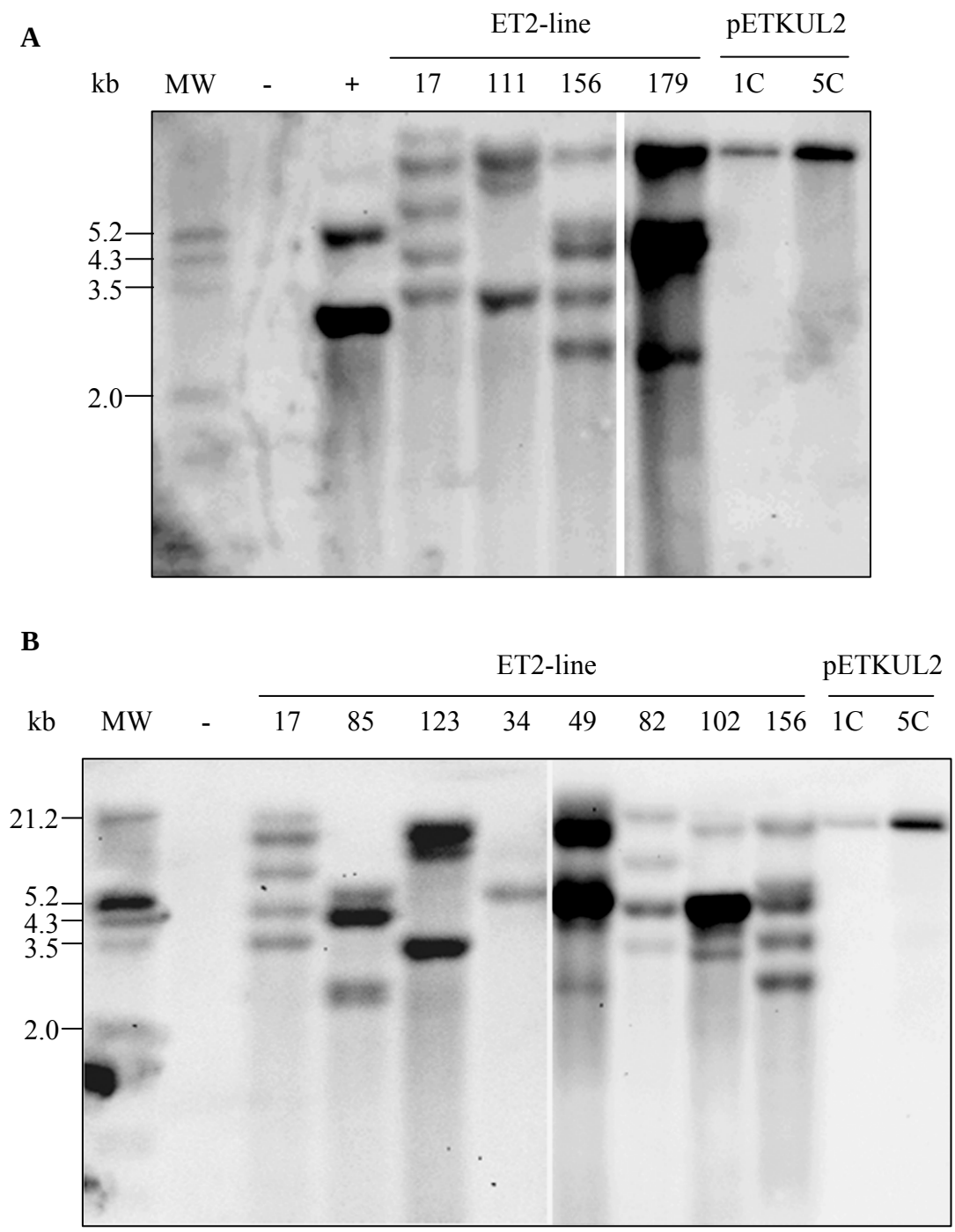

Figure 4.1. Southern blot analysis for the integration of the *luc*$^+$ transgene. Ten micrograms of total DNA were digested with *Hind*III and separated fragments were hybridized with a DIG-labeled *luc*$^+$ probe of 862 bp. (A) Total DNA isolated from leaf material of several *in vitro* plant clones per line. (B) Total DNA isolated from one leaf of a greenhouse-grown plant per line. Symbols – and + refer to a non-transformed control line and a positive control line carrying the *luc*$^+$ gene under control of the enhanced CaMV 35S RNA promoter, respectively. ET2-line refers to promoter-tagged lines transformed with promoter tagging vector pETKUL2. 1C and 5C refer to one and five copy number pETKUL2 plasmid (*Hind*III linearized) reconstructions, respectively. MW: DIG-labeled DNA molecular marker III (Roche). kb: kilo basepair.

4.3 Isolation of T-DNA flanking sequences by TAIL-PCR and I-PCR

Irrespective of the number of integrated T-DNA copies, both TAIL- and I-PCR were performed for the isolation of the T-DNA flanking sequences. Amplification of RB T-DNA flanking sequences (5'-tagged sequences) allows the isolation of putative banana promoters located upstream of the *luc*$^+$ present in the T-DNA of pETKUL2, while the left border (LB) T-DNA flanking sequences (3'-tagged sequences) should contain a coding sequence unless a cryptic promoter has been tagged. The positions of the primers and restriction enzymes relative to the T-DNA border sequences are shown in Figure 4.2. Amplicons for the isolation of 5'-tagged sequence showing the expected size shift of 94 bp between the secondary and tertiary TAIL-PCR reactions or the 145 bp size difference between the primary and secondary I-PCR reactions (Figure 4.2) indicate the specific amplification of sequences flanking the RB T-DNA. Similarly, the expected size differences between the consecutive TAIL- and I-PCR reactions for the isolation of 3'-tagged sequences of 150 bp and 234 bp, respectively (Figure 4.2), strongly suggest specific amplification.

As shown in Figure 4.3, the multiple T-DNA copy line ET-17 had a higher number of amplicons than the single T-DNA copy line ET2-34 (4 *vs*. 2, respectively) irrespective of the PCR walking method. In addition, amplicon sizes ranged from less than 200 bp until 3 kb with both methods. Because they consist of more unknown T-DNA flanking sequence that might contain most or all of the *cis*-acting elements of the candidate promoter sequence (paragraph 4.4.2), the longer amplicons are preferred for sequencing. In turn, this avoids a time consuming second PCR walking step. Occasionally, amplicons were not observed in the primary I-PCR reaction as illustrated in Figure 4.3B for the isolation of the 3'-tagged region in line ET2-34, but they were subsequently detected in the secondary I-PCR reaction.

RB T-DNA region

```
     1 cacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatatc

    51 cgattattct aataaacgct cttttctctt agGTTTACCC GCCAATATAT
                                          ..........RB.....

   101 CCTGTCAAAC ACTGATAGTT TAAACTTCGA ACGCCACCAT GGAAGACGCC
       ..RB....

   151 AAAAACATAA AGAAAGGCCC GGCGCCATTC TATCCGCTGG AAGATGGAAC
                             <---------Luc-R3---------

   201 CGCTGGAGAG CAACTGCATA AGGCTATGAA GAGATACGCC CTGGTTCCTG

   251 GAACAATTGC TTTTACAGAT GCACATATCG AGGTGGACAT CACTTACGCT
                  <---------TAILRBluc2---------

   301 GAGTACTTCG AAATGTCCGT TCGGTTGGCA GAAGCTATGA AACGATATGG
               <---------TAILRBluc1---------

   351 GCTGAATACA AATCACAGAA TCGTCGTATG CAGTGAAAAC TCTCTTCAAT

   401 TCTTTATGCC GGTGTTGGGC GCGTTATTTA TCGGAGTTGC AGTTGCGCCC
                  ---------LucL2--------->

   451 GCGAACGACA TTTATAATGA ACGTGAATTG CTCAACAGTA TGGGCATTTC

   501 GCAGCCTACC GTGGTGTTCG TTTCCAAAAA GGGGTTGCAA AAAATTTTGA
       ---------LucL3--------->

   551 ACGTGCAAAA AAAGCTCCCA ATCATCCAAA AAATTATTAT CATGGATTCT

   601 AAAACGGATT ACCAGGGATT TCAGTCGATG TACACGTTCG TCACATCTCA
                                 BsrGI

   651 TCTACCTCCC GGTTTTAATG AATACGATTT TGTGCCAGAG TCCTTCGATA
                             <---------LucR5---------

   701 GGGACAAGAC AATTGCACTG ATCATGAACT CCTCTGGATC TACTGGTCTG
                    BclI

   751 CCTAAAGGTG TCGCTCTGCC TCATAGAACT GCCTGCGTGA GATTCTCGCA...
                             <---------Luc+R---------
```

LB T-DNA region

```
…4201 GAGTTCTTCT GAGCGGGACT CTGGGGTTCG GATCGATCCT CTAGCTAGAG
      ---------TAILLBpET2n1--------->

  4251 TCGATCGACA AGCTCGAGTT TCTCCATAAT AATGTGTGAG TAGTTCCCAG

  4301 ATAAGGGAAT TAGGGTTCCT ATAGGGTTTC GCTCATGTGT TGAGCATATA
                             ---------TAILLBpET2n2--------->

  4351 AGAAACCCTT AGTATGTATT TGTATTTGTA AAATACTTCT ATCAATAAAA

  4401 TTTCTAATTC CTAAAACCAA AATCCAGTAC TAAAATCCAG ATCCCCCGAA
                                        <---------TAILLBpET2n4

  4451 TTAATTCGGC GTTAATTCAG TACATTAAAA ACGTCCGCAA TGTGTTATTA
       ---------> ---------TAILLBpET2n3--------->

  4501 AGTTGTCTAA GCGTCAATTT GTTTACACCA CAATATATCC TGCCAccagc
                             ..........LB.....

  4551 cagccaacag ctccccgacc ggcagctcgg cacaaaatca ccactcgata
```

Figure 4.2. Right (RB) and left border (LB) T-DNA regions in pETKUL2. RB and LB are shown above dotted lines. Lowercase and uppercase indicate pETKUL2 backbone and T-DNA sequences, respectively. Bold uppercase in the RB and LB T-DNA regions indicate the 5' partial *luc*$^+$ coding region and the 3' partial neomycin phosphotransferase II selectable marker gene with the CaMV 3' UTR, respectively. The recognition sites for the restriction enzymes *Bsr*GI and *Bcl*I used in I-PCR are boxed. Arrows indicate the annealing site and direction of primers used in TAIL- and I-PCR (Table 2.6). TAILRBluc1, TAILRBluc2 and Luc-R3 were employed in the primary, secondary and tertiary RB TAIL-PCR reaction, respectively, while in the primary, secondary and tertiary LB TAIL-PCR reaction TAILLBpET2n1, TAILLBpET2n4 or TAILLBpET2n2 and TAILLBpET2n3 were used, respectively. Following digestion with *Bsr*GI primer pairs LucL2/TAILRBluc1, LucL3/TAILRBluc2 and LucL3/Luc-R3 amplified the 5'-tagged region(s) and Luc+R/TAILLBpET2n1, LucR5/TAILLBpETn2 or TAILLBpET2n4 and LucR5/TAILLBpET2n3 the 3'-tagged region(s). When using *Bcl*I, the corresponding primer pairs were LucL2/TAILRBluc1, LucL3/TAILRBluc2 and LucL3/TAILRBluc3 for the 5'-tagged region(s) and Luc+R/TAILLBpET2n1, Luc+R/TAILLBpETn2 or TAILLBpET2n4 and Luc+R/TAILLBpET2n3 for the 3'-tagged region(s). Uppercase underlined italics refer to a stop codon 18 bp in-frame with the start codon of the *luc*$^+$.

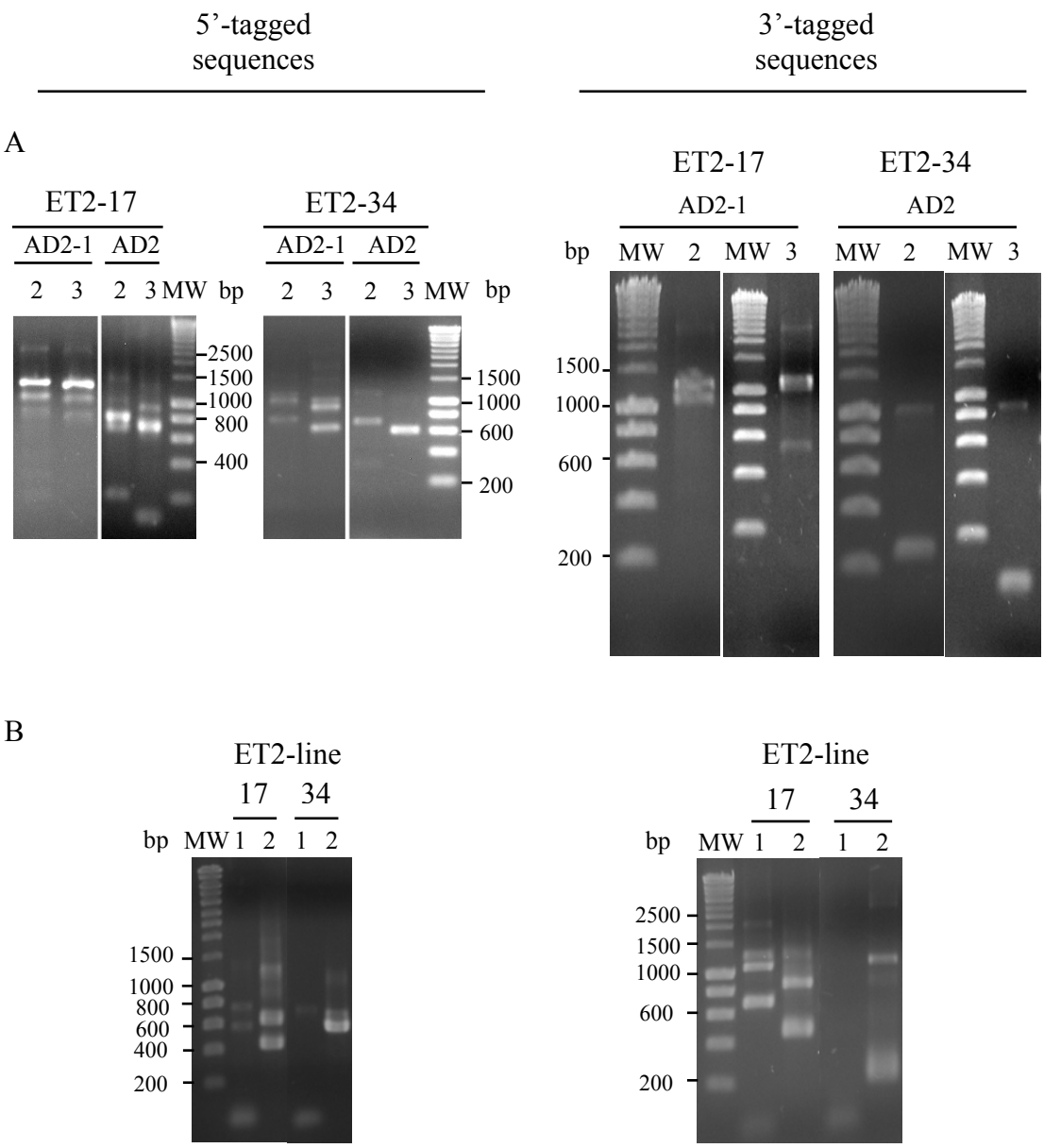

Figure 4.3. TAIL- and I-PCR analysis of the promoter-tagged lines ET2-17 and ET2-34 for the isolation of 5'- and 3'-tagged sequences flanking the right and left T-DNA borders, respectively. (A) TAIL-PCR was performed using two different AD primers: AD2-1 and AD2. Numbers 2 and 3 refer to the secondary and tertiary TAIL-PCR reaction, respectively. (B) I-PCR was performed using the *Bcl*I restriction enzyme. Numbers 1 and 2 refer to primary and secondary I-PCR reaction. MW: Smart Ladder (1 kb, Eurogentec, Seraing, Belgium). bp: base pair.

In TAIL-PCR similar amplicon patterns were obtained with a proofreading polymerase mixture (HF, Expand High Fidelity PCR System, Roche) and the standard non-proofreading *Taq* polymerase (New England Biolabs, NEB; data not shown). However, using I-PCR these polymerases generated different amplicon patterns in promoter-tagged line ET2-17. Larger size amplicons were produced with HF polymerase than with the standard NEB polymerase (above 2 kb and below 1 kb, respectively; Figure 4.4). On the other hand, in the three other promoter-tagged lines tested amplicons were of comparable

sizes with either polymerase although band intensities were more clear with the standard NEB polymerase (Figure 4.4). No amplicons were produced by I-PCR using DNA of a control non-transformed plant indicating that the amplicons in the promoter-tagged lines are specific and linked to T-DNA.

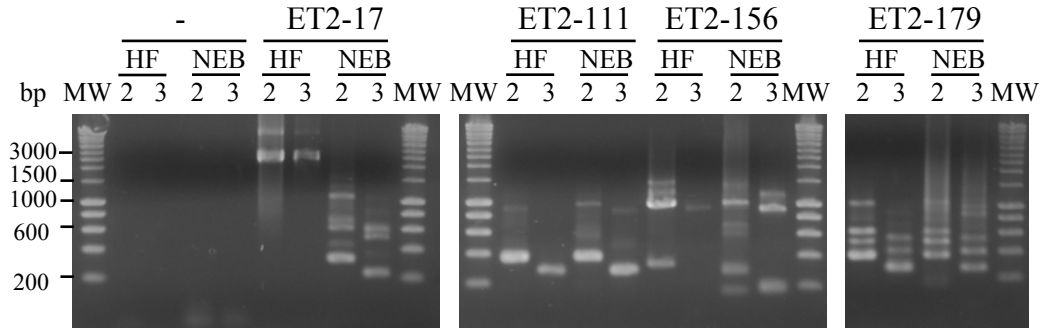

Figure 4.4. Comparison of the proofreading Expand High Fidelity PCR System (HF, Roche, Vilvoorde, Belgium) and the non-proofreading standard *Taq* DNA polymerase (NEB, New England Biolabs, Hertfordshire, UK) in I-PCR analysis using the *Bsr*GI restriction enzyme for the isolation of the 5'-tagged sequences flanking the T-DNA right border. Samples included promoter-tagged lines transformed with vector pETKUL2 and a non-transformed control line (-). Numbers 2 and 3 refer to secondary and tertiary I-PCR reaction, respectively. MW: Smart Ladder (1 kb, Eurogentec, Seraing, Belgium). bp: base pair.

For certain tagged lines a discrepancy existed between the T-DNA copy number and the number of amplicons produced by PCR walking. For instance, three amplicons specific for the 5'-tagged sequences were produced in tagged line ET2-156 (Figure 4.4), whereas five T-DNA copies were integrated in this line (Figure 4.1). In an attempt to match the amplicon number with the T-DNA copy number, the high stringency annealing temperature of the TAIL-PCR program (Table 2.7) was decreased from 64°C to 60°C when using the AD2 primer. This drop in annealing temperature, however, did not cause any change in the amplicon patterns of the different promoter-tagged lines (Figure 4.5). Surprisingly, several amplicons were present in the secondary TAIL-PCR reaction of the control, non-transformed line at both annealing temperatures but the signal intensities were higher at the lower annealing temperature of 60°C. In the tertiary TAIL-PCR reaction a smear occurred at the annealing temperature of 60°C while annealing at 64°C did not yield any amplicon. Obviously, an annealing temperature of 60°C allows non-specific binding of the primers in the present experiment. It is worth noting that the lowest band in the tertiary reaction of the promoter-tagged line ET2-179 is of the same size than an amplicon obtained with the vector pETKUL2 indicating the presence of vector backbone sequences in this line. Amplicons were absent in the water control demonstrating that the TAIL-PCR mix was free from contaminating sequences.

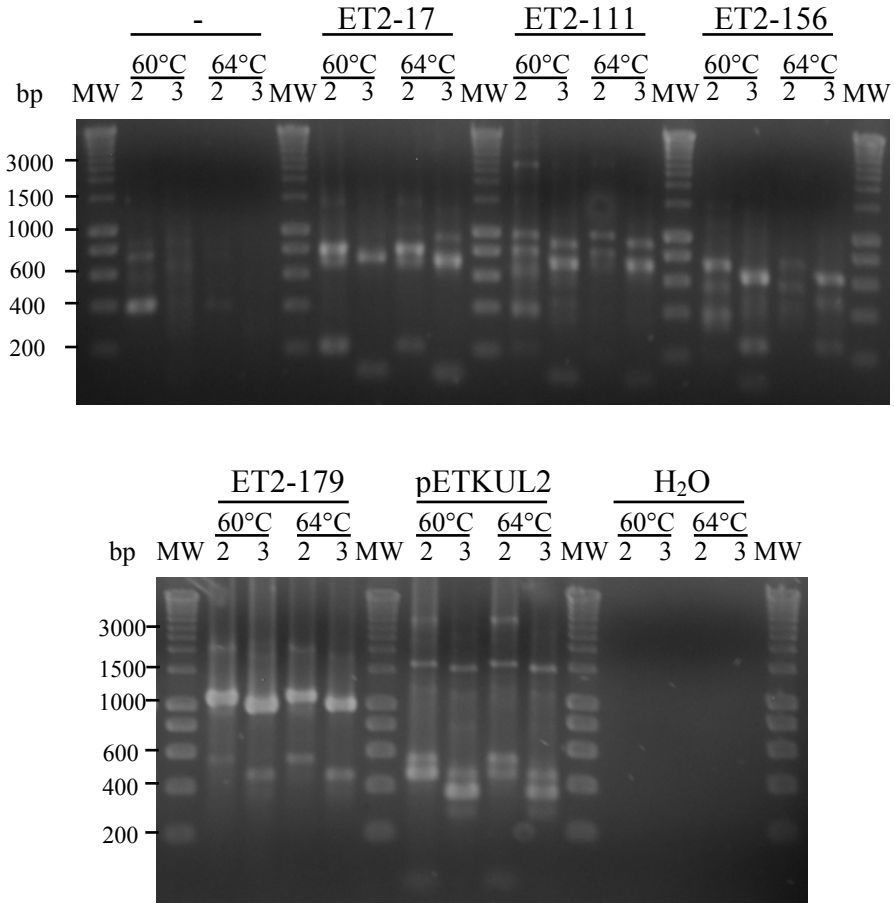

Figure 4.5. Effect of the high stringency annealing temperature (60°C *vs.* 64°C) on TAIL-PCR analysis using the degenerate AD2 primer for the isolation of the 5'-tagged sequence flanking the T-DNA right border. Samples included promoter-tagged lines transformed with vector pETKUL2, a control non-transformed line (-), the positive control tagging vector pETKUL2 and the negative water (MQ) control. Numbers 2 and 3 refer to secondary and tertiary TAIL-PCR reaction, respectively. MW: Smart Ladder (1 kb, Eurogentec, Seraing, Belgium). bp: base pair.

To further increase the success rate of isolating all flanking sequences in T-DNA multicopy lines, TAIL-PCR and I-PCR were performed with six different degenerate primers and two different restriction enzymes, respectively. TAIL-PCR using the degenerated primers AD1, AD4 and AD2-4 showed lower number of amplicons and/or amplicons below 500 bp in length than the other arbitrary degenerated primers (data not shown). In TAIL-PCR degenerated primer AD2-5 yielded less RB T-DNA flanking sequences than the other two degenerated primers AD2 and AD2-1 (1.6 *vs*. 2.3 and 2.9, respectively, Table 4.1), while in I-PCR more RB T-DNA flanking sequences were obtained with *Bcl*I than with *Bsr*GI (3.0 *vs*. 2.0, respectively, Table 4.1). Although usually different tagged sequences were obtained with the two PCR walking methods, identical sequences were also retrieved in several lines (paragraph 4.4.2, Annex). The number of isolated RB T-DNA flanking sequences corresponded well with the number of T-DNA copies except for the five copy T-DNA tagged line ET2-156, for which only three cloned

amplicons were sequenced (Table 4.1) and line ET2-49 in which more different sequences were found than the estimated T-DNA copy number (6 and 4, respectively). The number of cloned amplicons was higher than the different sequences obtained per line (on average 6.3 and 3.4, respectively; Table 4.1). These data strongly suggested that identical sequences were amplified by the two different PCR walking methods and/or with the same method using different degenerated TAIL-PCR primers or different restriction enzymes in I-PCR (paragraph 4.4.2, Annex). Although their T-DNA copy number was determined (Figure 4.1), only limited PCR walking was accomplished in the three tagged lines ET2-82, ET2-102 and ET2-123 explaining why they were omitted from these results.

Table 4.1. Number of T-DNA copies and 5'-tagged sequences flanking the T-DNA right border in 'Three Hand Planty' lines transformed with promoter tagging vector pETKUL2. The minimum number of T-DNA copies was determined by Southern blot analysis using a DIG-labeled *luc*$^+$ probe (paragraph 4.2). Isolation of 5'-tagged sequences was accomplished using three different arbitrary degenerated (AD) primers in TAIL-PCR and two different restriction enzymes in I-PCR

		No. of 5'-tagged sequences						
		TAIL-PCR			**I-PCR**			
ET2-line	**No. of T-DNA copies**	**AD2**	**AD2-1**	**AD2-5**	***Bsr*GI**	***Bcl*I**	**Different seq.[a]**	**No. of amplicons cloned for sequencing[b]**
17	5	3	5	2	2	2	5	10
156	5	2	1	2	2	2	3	3
49	4	4	5	1	NT	4	6	10
179	4	2	3	3	3	5	4	9
85	3	2	1	1	2	NT	3	3
111	3	2	3	1	1	3	2	6
34	1	1	2	1	2	2	1	3
Avg	3.3	2.3	2.9	1.6	2.0	3.0	3.4	6.3

[a]Number of different 5'-tagged sequences obtained.
[b]Number of amplicons that were cloned and sequenced.
Avg and NT refer to average and not tested.

4.4 Sequence analysis of T-DNA flanking sequences

4.4.1 T-DNA integration site analysis

4.4.1.1 Physical linkage of T-DNA flanking sequences

The amplicons which showed the expected size shifts in semi-nested amplification for the RB and LB T-DNA flanking sequences produced *via* TAIL- and I-PCR (paragraph 4.3) were gel excised, cloned into the pCR®4-TOPO® vector (Invitrogen, Merelbeke, Belgium) and commercially sequenced. Isolation of T-DNA flanking sequences revealed multiple sequences in multicopy T-DNA lines (paragraphs 4.2 and 4.3.1). To check for the physical linkage between the integrated T-DNA copies and the 5'-tagged sequences in a multicopy and single copy T-DNA lines (ET2-17, five T-DNA copies and ET2-34, one T-DNA copy, respectively; Figure 4.1) Southern blot hybridization was performed using first a probe specific to the *luc*$^+$ transgene (paragraph 4.2) and then specific probes for the different 5'-tagged sequences. In line ET2-34, re-hybridization was performed with a probe overlapping part of the 5'-tagged sequence and part of the T-DNA. Previous Southern blot results revealed five *luc*$^+$ inserts in line ET2-17 (Figure 4.1). On the same blots a comparison between the hybridization patterns obtained with the *luc*$^+$ probe and the 5'-tagged sequence-specific probes (indicated as 5' probe in Figure 4.6A and with numbers 1 to 5 above the blots in Figure 4.6B) revealed common fragments. For instance, the probes for *luc*$^+$ and the 5'-tagged sequence 17-1 showed a common band of approximately 4.5 kb, demonstrating that the *luc*$^+$ transgene is physically linked to integration 17-1. For four out of five *luc*$^+$ inserts in line ET2-17 the physical linkage with a cloned 5' region was established as shown in Figure 4.6B.

The 5'-tagged sequences 4 and 5 were linked to the same *luc*$^+$ insert and the remaining *luc*$^+$ copy could not be linked. The 5'-tagged sequence 17-5 corresponds to a fragment of the pETKUL2 vector backbone (paragraph 4.4.1.2) and explains the absence of a signal in the negative non-transformed control plant when the specific 5' probe was used. Furthermore, a total of two bands were observed with this probe in line ET2-17 indicating the presence of another vector backbone copy. On the other hand, the 5' probes of the remaining sequences of line ET2-17 annealed to fragments of the negative non-transformed control plant indicating the presence of the 5'-tagged sequences in the banana genome. Two signals were detected when using the probes 17-1 and 17-3 in line ET2-17 of which one corresponds well with a signal in the negative non-transformed control plant. The presence of the two signals might be attributed to the presence of duplicates of the 5'-tagged sequences 17-1 and 17-3 in the banana genome of which one was affected by the integration of the T-DNA. Furthermore, two signals were observed in the negative non-transformed control plant using the 17-3 probe and might be also explained by duplication

of the 5'-tagged sequence 17-3 in the banana genome. Hybridization with the probe spanning 33 and 235 bases in the 5'-tagged sequence and in the T-DNA, respectively, demonstrated the physical link between the 5'-tagged sequence and the *luc*$^+$ in line ET2-34.

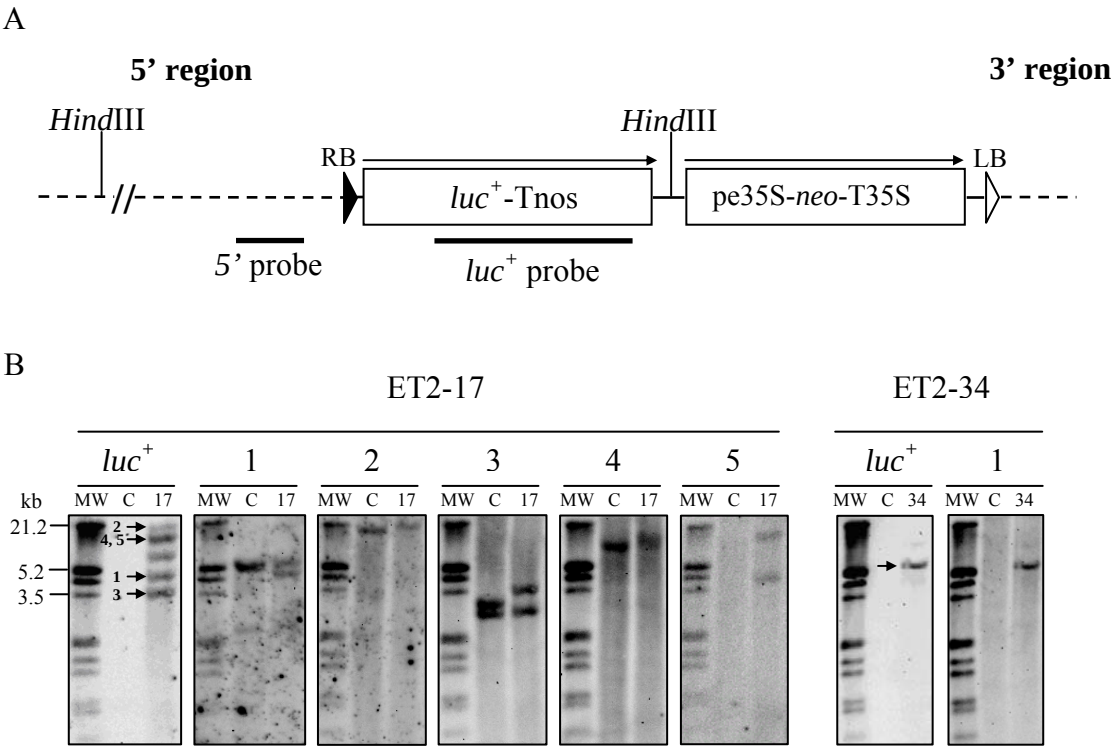

Figure 4.6. Physical linkage between 5'-tagged sequences and T-DNA in promoter-tagged lines ET2-17 and ET2-34. (A) Schematic representation of probe (thick lines) positions. The position of the codon-optimized luciferase (*luc*$^+$) and *neo* gene cassettes are shown with respect to the right (closed triangle) and left (open triangle) T-DNA border. Long arrows mark the direction of transcription. Dotted lines represent plant genomic DNA flanking the right (RB) and left (LB) T-DNA border, denominated 5' and 3' region, respectively. The restriction sites for the *Hind*III enzyme are indicated. The drawing is not precisely according to scale. (B) Southern hybridization analysis for the integration of the *luc*$^+$ gene and the cloned 5' regions. Ten micrograms of total DNA were digested with *Hind*III, separated fragments were hybridized with a DIG-labeled *luc*$^+$ probe (862 bp) and re-hybridized with a 5'-tagged sequence-specific probe (seq. 17-1: 422 bp, seq. 17-2: 425 bp, seq. 17-3: 435 bp, seq. 17-4: 165 bp and seq. 17-5: 302 bp for line ET2-17, and seq. 34-1: 268 bp for line ET2-34 which overlaps the 5'-tagged sequence and the T-DNA). C: negative non-transformed control plant. 17: tagged line ET2-17. MW: DIG-labeled DNA molecular marker III (Roche). The tagged *luc*$^+$ inserts are marked by arrows.

Whereas isolation of 5'-tagged sequences was performed for the identification of putative promoter sequences, isolation of 3'-tagged sequences allowed the identification of putative coding regions in promoter-tagged lines, but the latter was performed on a smaller scale (Table 4.2). For instance, the number of amplicons cloned for sequencing reached up to 4.3 and 6.3 for the 3'- and 5'-tagged region, respectively. Since most of the promoter-tagged lines contained more than one T-DNA copy (paragraph 4.2) multiple 5'- and 3'-tagged sequences were obtained in T-DNA multicopy lines (Table 4.2). The number of different 5'- and 3'-tagged sequences were lower than the number of amplicons cloned indicating the isolation of redundant sequences. For the isolation of 3'-tagged sequences for

line ET2-17 it is clear that only three non-redundant sequences were identified despite the high number of amplicons cloned for sequencing (3 *vs.* 12, respectively). The lower number of non-redundant sequences identified in the 3'-tagged region than in the 5'-tagged region in line ET2-17 might be attributed to truncations at the LB region of the T-DNA avoiding the annealing of the specific primers used for the isolation of 3'-tagged sequences.

Table 4.2. Number of T-DNA copies and T-DNA flanking sequences of 'Three Hand Planty' lines transformed with promoter tagging vector pETKUL2

		5'-tagged sequences		3'-tagged sequences	
ET2-lines	**No. T-DNA copies**	**Different sequences[a]**	**No. of amplicons cloned for sequencing[b]**	**Different sequences[a]**	**No. of amplicons cloned for sequencing[b]**
17	5	5	10	3	12
156	5	3	3	3	3
49	4	6	10	1	1
179	4	4	9	2	2
85	3	3	3	4	5
111	3	2	6	NT	NT
34	1	1	3	1	3
Average		3.4	6.3	2.3	4.3

The minimum number of T-DNA copies was determined by Southern hybridization using a DIG-labeled *luc*$^+$ probe. Isolation of 5'- and 3'-tagged sequences was accomplished by TAIL- and I-PCR.
[a]Number of different T-DNA flanking sequences obtained.
[b]Number of amplicons that were cloned and sequenced.
NT refers to not tested.

To determine which 5'-tagged sequence formed a continuous sequence with which 3'-tagged sequence, a series of PCR reactions were performed per 5'-tagged sequence using each time the same forward primer specific for the 5'-tagged sequence in combination with a reverse primer specific for one of the 3'-tagged sequences (Figure 4.7A) and the positive combinations are shown in Figure 4.7B. This type of PCR is called here linking-PCR. For line ET2-17 only three different 3'-tagged sequences were isolated by TAIL-PCR using specific primers annealing in the T-DNA near the LB (Table 4.2). One 3'-tagged sequence corresponded to a rearranged T-DNA insertion (paragraph 4.4.1.2) and the remaining two sequences were linked to the 5'-tagged sequences 17-2 and 17-4 (Figure 4.7B). To identify the 3'-tagged regions linked to the 5'-tagged sequences 17-1 and 17-3, TAIL-PCR was performed using specific primers annealing to the cloned 5'-tagged sequences to walk downstream. The amplicon pattern in TAIL-PCR of the non-transformed control line was similar to that of line ET2-17, although small differences were observed as expected. For instance, longer but weak amplicons were detected with the AD2 and AD2-1 primers for sequence 17-1 in the tagged line ET2-17 than in the control (Figure 4.8A). An additional amplicon of ~1300 bp was detected in the tagged line ET2-17 when compared to the non-transformed line when using the AD2-5 primer (Figure 4.8A). Amplicons of the non-

transformed control line were cloned and sequenced, which confirmed that they were linked to the 5'-tagged sequences 17-1 and 17-3. Finally, by using an F-link and R-link primer specific for the 5'-tagged sequences 17-1 and 17-3 and their respective downstream sequences (Figure 4.7A), proof of their linking was obtained in line ET2-17 (Figure 4.7B)

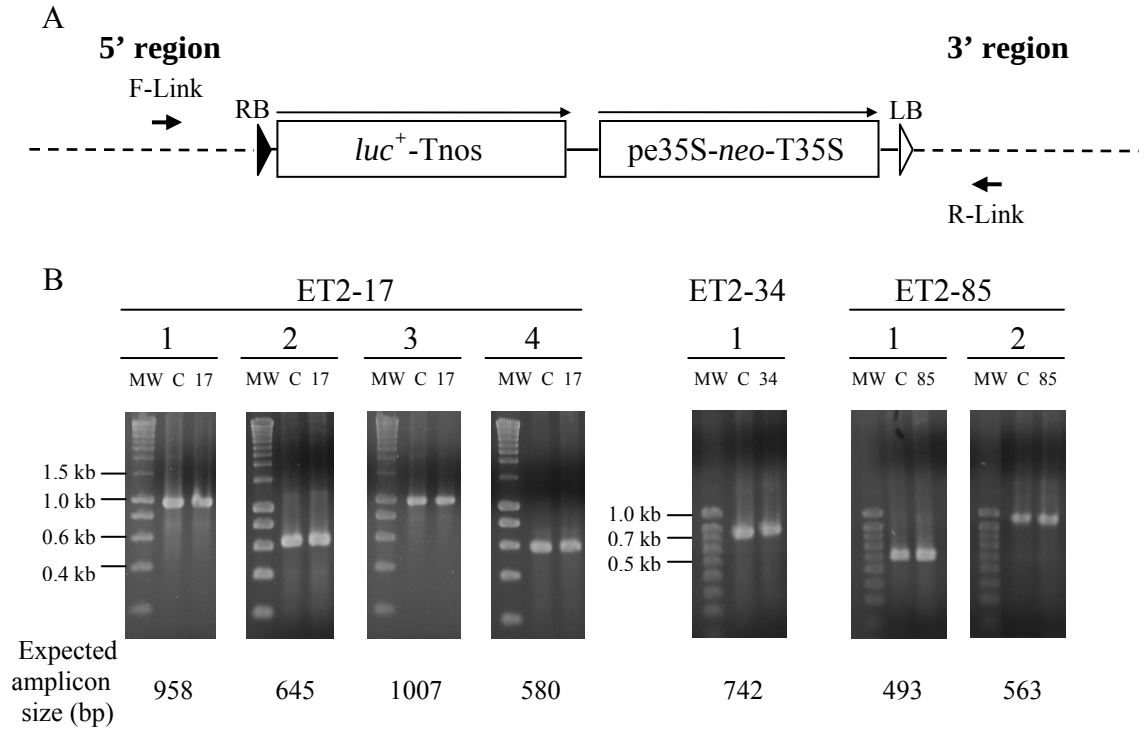

Figure 4.7. Linking-PCR performed on promoter-tagged lines ET2-17, ET2-34 and ET2-85 between the 5'- and 3'-tagged sequences. (A) Schematic representation of positions of primers F-link and R-Link (short arrows) used in linking PCR and annealing in the 5'- and 3'-tagged sequences, respectively. (B) PCR confirmation in lines ET2-17 (17), ET2-34 (34) and ET2-85 (85) of tagged sequences forming a continuous sequence with primers specific for 5'-tagged sequences in combination with reverse primers specific for 3'-tagged sequences. The expected sizes of amplicons are shown below the corresponding gel picture. Numbers above the gels indicate the corresponding tagged sequences (1, 2, 3, 4). MW: Smart Ladder (1 kp, Eurogentec, Seraing, Belgium). kb: kilo base pair. Primers used for the different tagged sequences were: 17-1, 17-1ProF/17-1LBR; 17-2, 17-14RT2/17-1LB1; 17-3, 17-3ProF/17-3LBR; 17-4, 17-4RT1/17-2LB1; 34-1, 34-RTF/34-R-LB; 85-1, 85-1RT3/85-1LB; 85-2, 85-2RT3/85-2LB.

In conclusion, linking-PCR revealed specific amplicons with the calculated length in non-transformed control plants and in tagged plants of line ET2-17 for all four 5'-tagged sequences demonstrating the sequence continuity between the corresponding cloned 5' and 3' regions (Figure 4.7B). Similarly, the RB and LB T-DNA flanking sequences in the single T-DNA copy line ET2-34 were linked in the non-transformed control plant and the promoter-tagged line ET2-34 as shown by the presence of the amplicon with the calculated length of 742 bp (Figure 4.7). For the promoter-tagged line ET2-85, only two out of three 5'-tagged sequences were linked with two out of three 3'-tagged sequences (Table 4.2 and Figure 4.7B). This result implies the presence of at least four 5'- and 3'-tagged sequences in line ET2-85 yet only three 5'-tagged sequences were identified and Southern hybridization indicates the integration of at least three T-DNA copies (Figure 4.1).

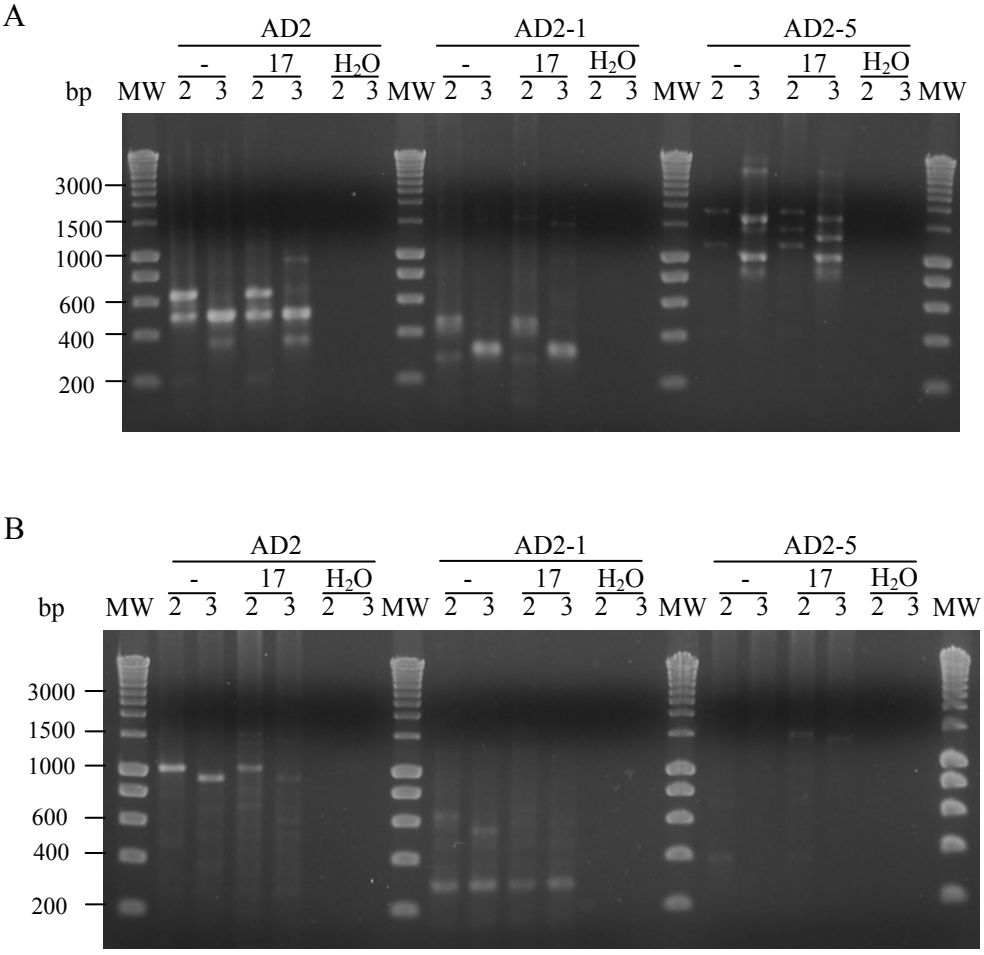

Figure 4.8. Secondary (2) and tertiary (3) TAIL-PCR using three arbitrary degenerated (AD) primers for the isolation of 3'-tagged sequences by walking from two cloned 5' sequences of promoter-tagged line ET2-17. Samples included a control, non-transformed (-) line, line ET2-17 (17) and a negative control of milliQ water (H_2O). (A) Specific primers annealed to the 5'-tagged sequence 17-1. (B) Specific primers annealed to the 5'-tagged sequences 17-3. MW: Smart Ladder (1 kb, Eurogentec, Seraing, Belgium). bp: base pair.

Amplicons from the linking-PCR of non-transformed lines were sequenced and analyzed for the integrity of the banana DNA during T-DNA insertion by comparison with 5' and 3'-tagged sequences isolated by TAIL- or I-PCR from the tagged lines. For the 5'-tagged sequence 85-2 and the tested LB flanking sequence shown in Figure 4.7B an amplicon of 563 bp was expected, but the amplicon obtained was approximately 900 bp. Sequence analysis revealed the presence of 347 bp between these 5'- and 3'-tagged sequences (Figure 4.9), which was apparently removed during T-DNA integration. Microhomology between the RB or LB T-DNA junction region and the banana DNA comprised 1-5 bp and 2-7 bp, respectively (Figure 4.9). Discarding the microhomology between the T-DNA border junction regions and the genomic DNA and taking only the base pairs into account that are not homologous to the T-DNA borders, the removal of

banana DNA during integration occurred in four out of five T-DNA integration events (80.0%) by the excision of 6 bp, 20 bp, 26 bp and 344 bp of banana DNA (Figure 4.9). Analysis for the removal of banana DNA during T-DNA integration was not possible for the flanking sequences 17-1 and 17-3 since the 3'-tagged sequences were not isolated by walking from the T-DNA LB (Figure 4.9B).

A	3'-tagged sequence	5'-tagged sequence
17-2	…CCGAATTAGGCAGACAAAAAACATTGCCACCAgtgATGACCGACCCTTTCCTGCTGCCGT…	
17-4	…GTGCTCATGGACATTTTGTTTTGAAgctctcactaAGGCAGGTAGTCTGCTATAACTAAG…	
34-1	…AATGTTGACTAGGAggctacatagtcatggtaaagatggtCCATAACAAAAAAGAGCAGG…	
85-1	…GATATTGGTATTTAGCCTAtagacgatcacaaaaatggcaACCACATGCACCCAAGCTGT…	
85-2	…AAGCAGTTCTGCGCACcaaaaggt(331 bp)tttttgcaGCATTCTTTTGAAACTAGCT…	

B	3'-tagged sequence	5'-tagged sequence
17-1	…CAAGTGAAAAATCTCCACCTCGTGGATAG	::AAGCAACACGCGCATAGTCCAAGTGTCAT…
17-3	…TCTAGATGATAAACAAAAGCAAAATTCCA	::TGTAAAACATTCTACAGTTCAAATTCTAG…

Figure 4.9. Sequence analysis of the linking-PCR amplicons for the 5'- and 3'-tagged sequences in three promoter-tagged lines. The 5'- and 3'-tagged sequences are shown in uppercase and shaded. (A) The T-DNA flanking sequences were isolated by PCR walking from the T-DNA into the genomic DNA. Homologous sequences between the RB or LB T-DNA junction region and the banana genomic DNA is underlined or double underlined, respectively. The basepairs in lowercase letters might have been removed during T-DNA integration. (B) The 5'-tagged sequences were isolated by PCR walking from the T-DNA while their corresponding 3'-tagged sequences were isolated by TAIL-PCR using primers specific to the 5'-tagged sequence (Figure 4.8). Double colons (::) indicate the insertion site of the T-DNA.

4.4.1.2 Occurrence of T-DNA rearrangement, vector backbone and tandem repeats in promoter-tagged lines

Isolation of promoters *via* T-DNA tagging has been accomplished in different plant species. However, multiple T-DNA copies (paragraph 4.2) and hence, multiple T-DNA flanking sequences complicate the further identification of the promoter candidate(s) (paragraph 4.4.2 and Chapter 5) in the promoter-tagged lines (paragraph 4.3). Besides the presence of multiple T-DNA copies in the majority of the promoter-tagged lines, integration of rearranged T-DNA copies, vector backbone and tandem repeats were observed by sequence analysis of 5'- and 3'-tagged regions (overview in Table 4.3). To make this analysis more reliable another six lines were included besides the core group of three lines (ET2-17, ET2-34 and ET2-85). The T-DNA of pETKUL2 spans 4438 bp and thus, T-DNA tandem repeats are expected in the promoter-tagged lines ET2-49, ET2-82, ET2-102, ET2-156 and ET2-179 as illustrated by a band of high intensity located at this position (paragraph 4.2, Figure 4.1). In addition, integration of the complete pETKUL2 vector in a tandem repeat configuration is expected in lines ET2-17, ET2-49, ET2-102, ET2-111, ET2-123, ET2-156 and ET2-179 due to the presence of a band with similar size of the *Hind*III linearized

pETKUL2 vector (Figure 4.1). Sequence analysis confirmed the presence of tandem repeats and vector backbone in lines ET2-49 and ET2-156, and in lines ET2-17, ET2-49, ET2-123, ET2-156 and ET2-179, respectively (Table 4.3) corroborating the observed data in the Southern blots. In addition, the vector backbone sequence isolated in flanking sequence 17-5 of line ET2-17 was physically linked to a band of the size of the *Hind*III linearized pETKUL2 (compare Figures 4.1 and 4.6B), which suggested the integration of a complete pETKUL2 vector.

Table 4.3. Minimum number of T-DNA copies and different pETKUL2 DNA sequences discovered in T-DNA flanking regions

ET2-line	No. T-DNA copies[1]	Vector backbone[2]	Tandem repeats[3]	Promoter/gene rearrangements[4]
17	5	1 (RB)[5], 1 (LB)[6]	-	1 (LB)
156	5	1 (RB, LB)[6, 7], 1 (RB)[6]	1 (RB, LB)[7], 1 (LB)[8]	1 (RB)
49	4	1 (RB)[9], 1 (LB)[10]	2 (RB)	1 (RB)
179	4	1 (LB)[10]	-	-
85	3	-	-	-
111	3	-	-	1 (RB)
123	3	1 (RB)[5]	-	-
34	1	-	-	-
89	NT	-	-	1 (RB)
Average	3.5	0.4	0.4	0.6

In brackets the isolated flanking region(s) of the T-DNA is (are) shown: right border (RB) or left border (LB).
[1]Minimun number of T-DNA copies determined by Southern blot hybridization using a DIG labeled *luc*$^+$ probe.
[2]DNA of the vector backbone, which carries the T-DNA.
[3]Direct T-DNA tandem repeats (RB-LB/RB-LB T-DNA configuration).
[4]T-DNA rearrangement: partial sequences of the enhanced CaMV 35S promoter and the neomycin phosphotransferase II selectable marker gene upstream of the T-DNA border.
[5]The vector backbone region near the LB was located upstream of the RB T-DNA region (see Figure 4.10B).
[6]Vector backbone found as filler DNA.
[7]The sequence was isolated by walking from the RB and LB T-DNA regions.
[8]A DNA fragment of 998 bp comprising 715 bp homologous to the insertion element IS113 from *A. tumefaciens* was found between the T-DNA copies of the T-DNA tandem repeat found in the LB T-DNA flanking sequence (3'-tagged region).
[9]Presence of a rearranged vector backbone and T-DNA.
[10]Vector backbone sequence most likely due to the readthrough of the VirD2 at the LB.
NT refers to not tested.

Besides the vector backbone sequences (Figures 4.10A and 4.10B) and T-DNA direct tandem repeats (Figure 4.10C) found in the T-DNA flanking regions, a specific T-DNA rearrangement was identified in which the enhanced 35S promoter (pe35S) with a truncated neomycin phosphotransferase II (*neo*) selectable marker gene was located in the T-DNA flanking regions. In four T-DNA rearrangements, the pe35S and truncated *neo* was located upstream of the *luc*$^+$ in the RB T-DNA region (RB sequence was not present, Figure 4.10D), while a similar T-DNA rearrangement was located upstream of the LB

region of the T-DNA in line ET2-17 (LB sequence was not present, Figure 4.10E; Table 4.4).

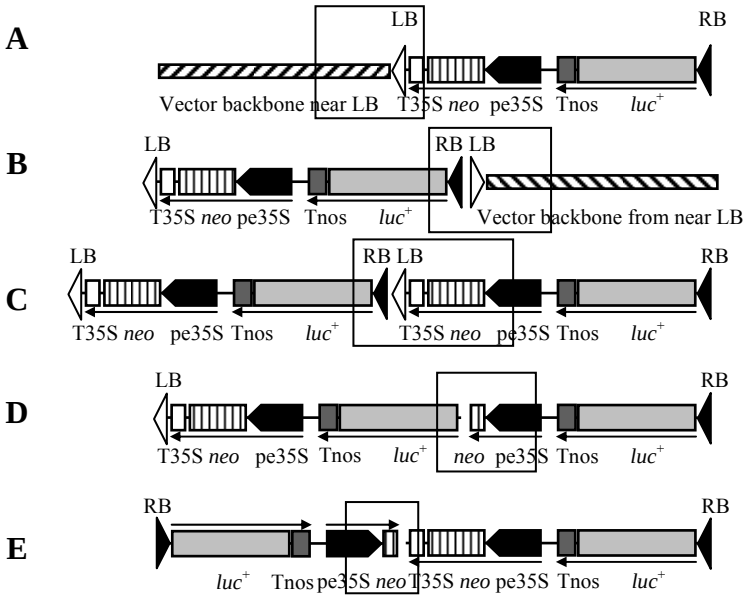

Figure 4.10. Schematic representation of different T-DNA integration patterns proposed to be found in certain promoter-tagged lines based on sequence analysis of the T-DNA flanking sequences. (A) T-DNA single copy with pETKUL2 vector backbone from near the LB. (B) Vector backbone from near the LB region located in the RB T-DNA flanking region. (C) T-DNA direct tandem repeat. (D) T-DNA rearrangement in which the enhanced 35S promoter (pe35S) with the truncated neomycin phosphotransferase II gene (*neo*) was located upstream of the RB region of the T-DNA or a truncated T-DNA direct tandem repeat. (E) The rearranged pe35S and truncated *neo* were located upstream of a truncated inverted T-DNA repeat. RB and LB are represented by closed and open triangles, respectively. The codon-optimized luciferase gene (*luc*$^+$) and the nopaline synthase terminator (Tnos) are illustrated by light and dark grey boxes, respectively. The pe35S and the *neo* are depicted in black box and box with vertical lines, respectively, while the 35S RNA terminator (T35S) is represented by an open box. Arrows indicate direction of transcription. Rectangles showed the fragments that were retrieved and sequenced. The diagrams are not precisely according to scale.

The rearranged pe35S::*neo* located upstream of the RB T-DNA in four promoter-tagged sequences (49-6, 89-1, 111-1 and 156-3) opens the possibility that the *luc*$^+$ is activated by the pe35S, although the presence of a stop codon (TAA) in-frame 18 bp upstream of the start codon of the *luc*$^+$ may prevent a translational fusion with the *neo* gene if an in-frame fusion occurs (Figure 4.2). The fusion transcript *neo-luc*$^+$ may be translated in the lines ET2-49 and ET2-89, in which the nicking site was 26 bp (tagged sequence 49-6) and 33 bp (tagged sequence 89-1) downstream from the RB theoretical cleavage site, respectively (paragraph 4.4.1.3), resulting in the absence of the stop codon which is located 16 bp from the RB cleavage site. However, only the rearranged T-DNA in line ET2-49 revealed that the *neo* is in frame with the *luc*$^+$ ORF, suggesting the presence of a NEO-LUC fusion protein in line ET2-49 and not in the other lines (ET2-89, ET2-111 and ET2-156). The length of the pe35S promoter fragment ranged from 314 bp to 458 bp (Table 4.4), but the presence of the complete pe35S (and remaining T-DNA region) was expected as shown in Figure 4.10D. The limited amplicon size of the isolated flanking sequences did

not allow the analysis of more upstream sequences (data not shown) and a second walking step was not done. The size of the truncated *neo* varied from 54 bp to 482 bp (Table 4.4) suggesting that the integration of the rearranged T-DNA might be due to a truncated T-DNA direct repeat in the lines ET2-49, ET2-89, ET2-111 and ET2-156. Comparison of the diagrams of the T-DNA direct tandem repeat (Figure 4.10C) with the rearranged T-DNA integration (Figure 4.10D) supports this hypothesis. On the other hand, a truncated inverted T-DNA tandem repeat might have occurred in the line ET2-17 (Figure 4.10E).

Table 4.4. Isolated T-DNA flanking sequences in promoter-tagged lines that contain fragments of the enhanced 35S promoter and neomycin phosphotransferase II selectable marker gene

		Size (bp)[a]			
ET2-line-Seq[b]	**T-DNA border flanking sequence**	**pe35S**	**pe35S from TSS[c]**	***neo***	**Predicted pe35S::*neo* PCR amplicon size[d]**
49-6	RB	458	418	482	564
89-1	RB	458	418	61	122
111-1	RB	374	334	148	257
156-3	RB	314	274	54	175
17-1	LB	458	418	67	NA

[a]Size of the enhanced 35S promoter (pe35S) and neomycin phosphotransferase II (*neo*) selectable marker gene found in the isolated T-DNA flanking sequences.
[b]Promoter-tagged line and the identification number of the sequence.
[c]Size of the pe35S from its transcription start site (TSS).
[d]*In silico* PCR with the primer 35SLRT1, which anneals from the position -4 to +16 relative to the start codon of the *neo*, and the primer Luc-R3, which anneals from the position +56 to +37 relative to the start codon of the *luc*$^+$.
NA refers to not applicable.

In order to confirm the presence of the pe35S and truncated *neo* upstream of the *luc*$^+$ in the promoter-tagged lines ET2-49, ET2-89, ET2-111 and ET2-156, PCR was performed using primers annealing at the pe35S::*neo* fusion cassette and the *luc*$^+$ (Figure 4.11). In addition, other promoter-tagged lines were also included to screen for the presence of this type of rearrangement. Firstly, primer 35SLRT1 which anneals from the position -4 to +16 relative to the start codon of the *neo*, and the primer Luc-R3, which anneals from the position +56 to +37 relative to the start codon of the *luc*$^+$ were used for PCR on total DNA of the promoter-tagged lines (Figure 4.11A). Secondly, primer 35SL which anneals twice in the pe35S was used in combination with the Luc-R3 primer (Figure 4.11B).

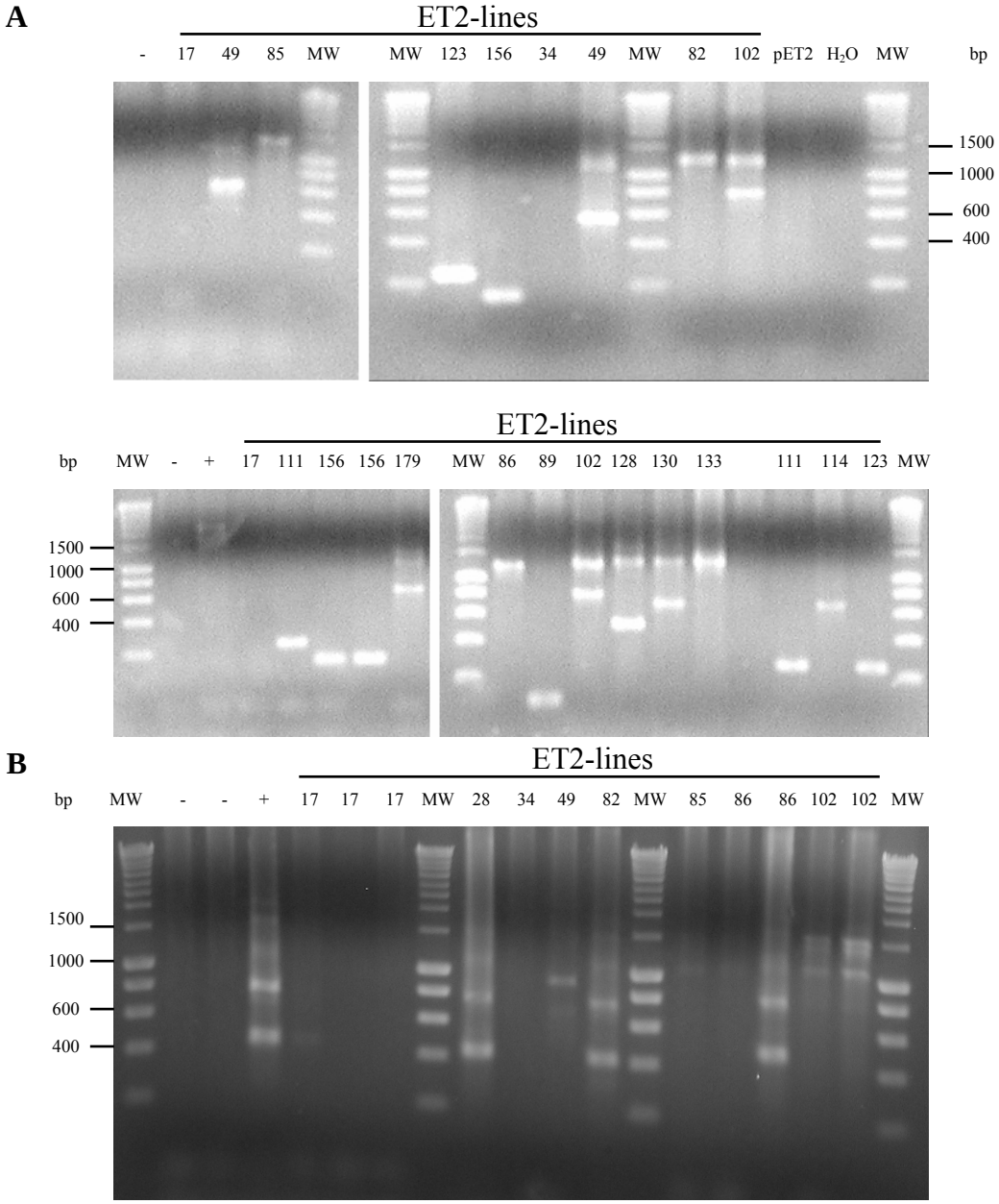

Figure 4.11. PCR performed on total DNA of several promoter-tagged lines for the presence of the enhanced 35S promoter and neomycin phosphotransferase II selectable marker gene in the 5' region of the *luc*$^+$. (A) The primer 35SLRT1, which anneals from the position -4 to +16 relative to the start codon of the neomycin phosphotransferase II, and the primer Luc-R3, which anneals from the position +56 to +37 relative to the start codon of the *luc*$^+$ were employed in the PCR. (B) The primers 35SL, which anneals twice in the enhanced 35S promoter, was employed in combination with the Luc-R3 primer. -: negative control non-transformed plants. +: positive control lines containing the enhanced 35S promoter driving the *luc*$^+$. pET2: pETKUL2 vector. H_2O: milliQ water. MW: Smart Ladder (1 kb, Eurogentec, Seraing, Belgium). bp: base pair.

The expected amplicon sizes using the first pair of primers (Table 4.4) correlate with the obtained amplicons (Figure 4.11A) confirming the presence of the pe35S::*neo* upstream of the *luc*$^+$. At least an amplicon of less than 200 bp and up to 1200 to 1400 bp was detected in 14 of 16 lines assayed (Figure 4.11A). The promoter-tagged lines ET2-49,

ET2-82, ET2-85, ET2-86, ET2-102, ET2-128, ET2-130, ET2-133 and ET2-179 showed a common amplicon of approximately 1200 bp (Figure 4.11A). The presence of this amplicon most likely indicates the presence of a direct tandem repeat because the expected amplicon size for a tandem repeat without intermittent genomic DNA (Figure 4.10C) is 1222 bp using the primer pair 35SLRT1/Luc-R3 if the RB and LB are nicked at the theoretical nicking sites and linked. Among the lines showing the 1200 bp amplicon, tandem repeats were also detected by Southern hybridization in the lines ET2-49, ET2-82, ET2-85, ET2-102 and ET2-179 as illustrated by the signal of 4438 bp (paragraph 4.2, Figure 4.1). The remaining lines were not investigated by Southern hybridization. No amplicons were observed in the promoter-tagged line ET2-17 and ET2-34 indicating the absence of the pe35S::*neo* upstream of the *luc*$^+$ and T-DNA direct tandem repeats (Figure 4.11A). Hence, the observed LUC activity profile in these two lines was not influenced by the CaMV 35S promoter which explains why amongst others these lines were analyzed in greater detail. In the positive control line (pe35S::*luc*$^+$) a clear amplicon was not detected when using the primer 35SLRT1 which anneals from the position -4 to +16 relative to the start codon of the *neo*, although a weak signal of approximately 3 kb was detected (Figure 4.11A). However, when using the primer that anneals completely within the enhanced 35S promoter amplicons were detected (Figure 4.11B). No detectable amplicons were observed in the control non-transformed line for any of the primers employed. Amplicons below 1222 bp reflect the presence of the pe35S and truncated *neo* upstream of the *luc*$^+$ in the tagged lines ET2-102, ET2-123, ET2-179, ET2-128 and ET2-130 (Figure 4.11A). Duplicate amplicons were observed in the positive control line (pe35S::*luc*$^+$) and the promoter-tagged lines ET2-28, ET2-49, ET2-82, ET2-86 and ET2-102 when using the primer 35SL confirming the presence of the pe35S upstream of the *luc*$^+$ (Figure 4.11B). In conclusion, PCR detection for the presence of the pe35S upstream of the *luc*$^+$ and T-DNA direct tandem repeats is a powerful screening tool *a priori* to the isolation of candidate tagged promoters.

4.4.1.3 T-DNA border junction sites

Sequence analysis of the RB T-DNA/genomic DNA junction regions of non-rearranged T-DNA integrations revealed that in 62.5% (10/16) the nick and integration of the T-DNA occurred between the third and fourth base of the RB sequence (TGA↓C…; Figure 4.12), which coincides with the predicted theoretical nicking site of the RB (Yanofsky *et al.* 1986). In addition, the nicking site in the RB of the remaining cloned fragments was near the predicted theoretical nicking site (nucleotide positions -1 to +2 relative to the theoretical nicking site) except for the sequences 17-2 and 156-4R (Figure 4.12). On the other hand, only 11.1% (1/9, *i.e.* T-DNA insertion with flanking sequence 85-2) of the LB T-DNA/genomic DNA junction regions suggest that the nicking and integration of the T-

DNA occurred between base three and four of the LB repeat (TGG↓C…) which corresponds to the predicted theoretical nicking site of the LB T-DNA (Yanofsky *et al.* 1986). Comparison of the RB and LB junctions regions belonging to the same T-DNA insertion as determined by 'linking'-PCR (paragraph 4.4.1.1) for the T-DNA insertions with the expected nicking site at the RB (*i.e.* T-DNA insertions with RB T-DNA flanking sequences 17-4, 34-1, 85-1 and 85-2) revealed that the corresponding LB nicking site was not located between the third and fourth base except for the T-DNA with flanking sequence 85-2 (Figure 4.12). Moreover, the cleaving site in the LB differed among the T-DNA insertions. In conclusion, the RB was more correctly nicked than the LB of the T-DNA.

Isolation of a 3'-tagged sequence in promoter-tagged line 156 (156-1L, Figure 4.12) revealed 998 bp downstream the presence of a RB T-DNA region (156-4R, Figure 4.12), which indicated the integration of a T-DNA tandem repeat separated by 998 bp (Figure 4.13B, 156-1LB/156-4RB). This region comprised 715 bp with high homology to the insertion element IS113 from *A. tumefaciens* (M82888.1, 99% identity at E = 0.0). Both T-DNA border repeats were not nicked at the predicted cleavage site (Figure 4.12). The tandem T-DNA integration also explains the presence of at least one additional 5'-tagged sequence that was not identified in the isolation of 5'-tagged sequences by PCR walking (Table 4.1).

For the flanking sequences 17-1, 17-2, 17-3, 17-4, 34-1, 85-1 and 85-2 analysis for the presence of filler DNA and sequence identity between the RB junction region in the tagged lines and the corresponding 5'-tagged genomic DNA region in a non-transformed line became possible due to the linking-PCR sequence analysis for the corresponding 5'- and 3'-tagged sequences in non-transformed plants (Figure 4.12 and paragraph 4.4.1.1). However, such analysis between the LB junction region and 3'-tagged sequence in insertions 17-1 and 17-3 could not be included because the isolation of their 3'-tagged sequences was accomplished by TAIL-PCR starting from the 5'-tagged sequences and not from the T-DNA (paragraph 4.4.1.1). As linking-PCR was not executed, tagged lines ET2-156 and ET2-179 are not incorporated in these analyses.

Filler DNA was observed in one (17-3) out of seven (14.3%) 5'-tagged/T-DNA junctions and comprised nine nucleotides; while the two T-DNA/3'-tagged junctions 17-2 and 85-1 (2/5 or 40%) contained filler DNA of 3 and 9 nucleotides in length, respectively (Figure 4.12). Microhomology between the tagged sequences and the T-DNA end was observed in four out of seven (57.1%) 5'-tagged/T-DNA junction regions and consisted of 1 or 3 nucleotides. At the LB/banana DNA junction region three (85-2, 34-1 and 17-4) out of five (60.0%) 3'-tagged sequences contained over 2, 4 and 7 nucleotides homology to the T-DNA, respectively (Figure 4.12). To summarize, only two out of seven (28.6%) banana/RB T-DNA junction regions did not contain microhomology or filler DNA, while none of the LB T-DNA/banana DNA junctions was free from filler DNA and microhomology.

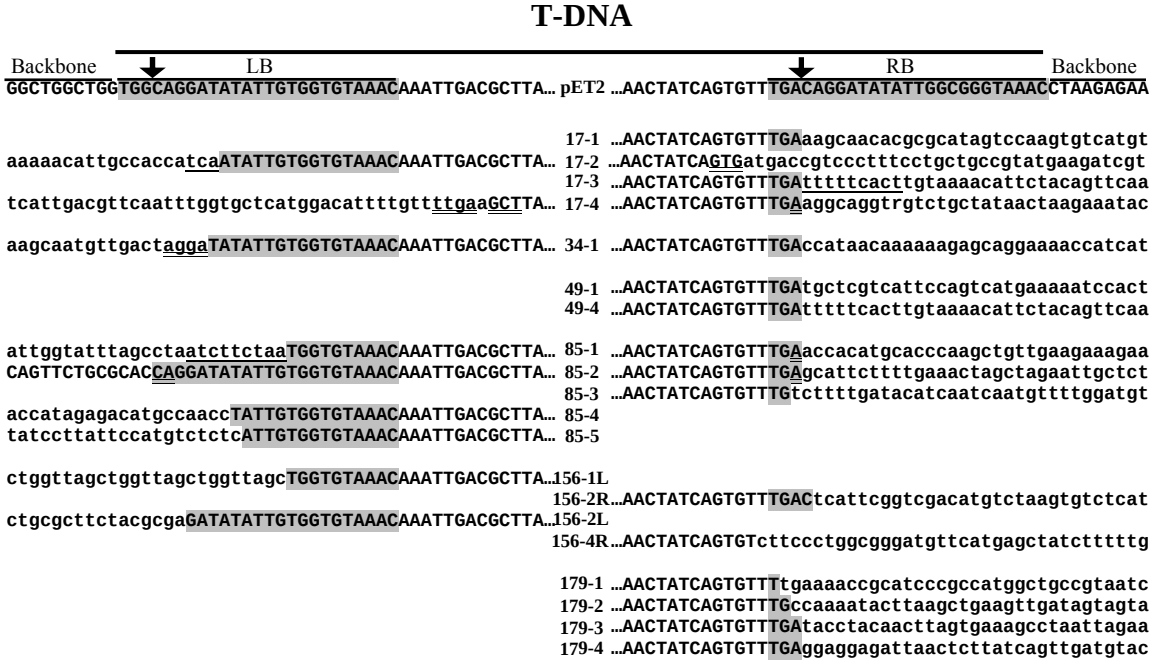

Figure 4.12. Sequence alignment of right (RB) and left border (LB) T-DNA/genomic DNA junction regions in promoter-tagged lines to the T-DNA/vector backbone junction regions in the tagging vector pETKUL2 (pET2). Only junction regions of non-rearranged T-DNA insertions are shown. RB and LB sequences in pET2 sequence are shaded, and shaded nucleotides in the junction regions reflect identities with RB and LB sequences. Nucleotides corresponding to pETKUL2 sequence are in uppercase, while lowercase letters correspond to banana DNA from the different cloned T-DNA border flanking sequences in the promoter-tagged lines except for the sequences 156-1L, 156-2L and 156-4R. Cloned fragments are numbered first by the promoter-tagged line number followed by the number of the flanking sequence isolated. Underlined nucleotides indicate filler DNA, while double underlined nucleotides are identical in the T-DNA and the non-transformed banana genomic DNA. The corresponding RB and LB flanking sequences for the different T-DNA copies were not determined in lines ET2-49, ET2-156 and ET2-179. Arrow indicates the location of the theoretical cleavage site in the RB and LB of the T-DNA. The letter r in the sequence 17-4 indicates a purine (A or G).

Sequence analysis at the junction sites of rearranged T-DNA integration patterns illustrated in Table 4.3 and Figure 4.10 was performed for identification of filler DNAs and microhomologies, too (Figure 4.13). Integration of the LB vector backbone sequence was probably due to the readthrough of the LB repeat in the T-DNA 49-1LB of line ET2-49 and in the T-DNA 179-1LB of line ET2-179 (Figure 4.13A; schematic representation Figure 4.10A). Only T-DNA direct repeats were observed in four T-DNA flanking sequences of the promoter-tagged lines ET2-49 and ET2-156 (Figure 4.13B; schematic representation Figure 4.10C). Two independent T-DNA tandem junctions of line ET2-49 showed 3 nucleotides microhomology at the T-DNA ends. One T-DNA direct repeat junction region in line ET2-156 (156-3LB/156-1RB) revealed the presence of filler DNA originated from the vector backbone (25 bp) which joined the two T-DNAs (Figure 4.13B). In addition, microhomology of 1 bp at 156-3LB and 2 bp at 156-1RB was observed between this filler DNA and the T-DNA ends in the corresponding isolated flanking sequences of the LB (156-3LB) and RB (156-1RB; Figure 4.13B). A DNA fragment of 998 bp was observed

between two other T-DNA copies in line ET2-156 with 715 bp homologous to the insertion element IS113 from *A. tumefaciens* (M82888.1, 99% identity at E = 0.0; 16 February 2007) demonstrating the integration of two T-DNA copies in direct orientation with probably part of the *A. tumefaciens* genome in the same locus (156-1LB/156-4RB, Figure 4.13B) corresponding to the junction regions 156-1L and 156-4R in Figure 4.12. The cleaving at the border repeats was not at the predicted site in the different T-DNA direct repeats analyzed and only at the flanking sequence 49-5 it coincided with the theoretical cleaving site at the LB (Figure 4.13B). The T-DNA rearrangement observed in RB T-DNA flanking sequences in which the enhanced 35S promoter and the truncated *neo* was located upstream in the correct orientation for driving the *luc*$^+$ (flanking sequences 49-6, 156-3RB, 111-1 and 89-1; schematic representation shown in Figure 4.10D), is characterized by the presence of filler DNA originating from the T-DNA, vector backbone or unknown sources in three out of four rearrangement events (Figures 4.13C). Similarly, the rearranged T-DNA in the LB flanking sequence 17-5LB was fused to the T-DNA (schematic representation shown in Figure 4.10E) by filler DNA (Figure 4.13D). The vector backbone near the LB region fused to the RB T-DNA region (schematic representation shown in Figure 4.10B) was found in two 3'-tagged sequences and the LB was nicked at the theoretical cleavage site in both isolated sequences analyzed, while the RB repeats were nicked near the theoretical cleavage site (Figure 4.13E). The results demonstrate that for most of the complex rearranged T-DNA integration patterns the cleavage sites in T-DNA borders were not at the theoretical sites. In conclusion, T-DNA border junction sites of rearranged T-DNA copies are highly variable and complex.

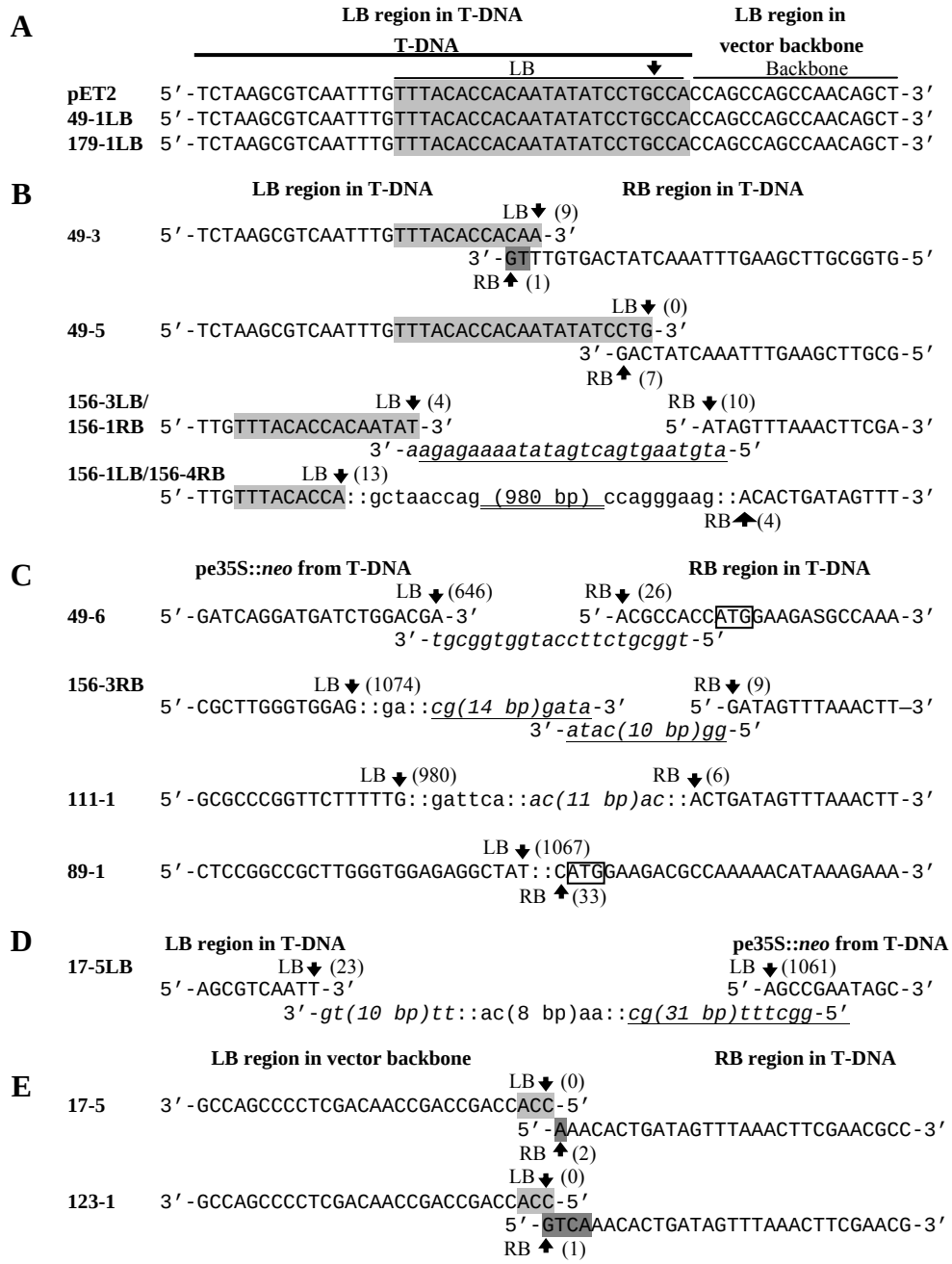

Figure 4.13. Sequence analysis of the T-DNA border junction regions in different rearranged T-DNA copy insertions of promoter-tagged lines. (A) Junction between T-DNA and vector backbone in the 3'-tagged region of insertions 49-1LB and 179-1LB of lines ET2-49 and ET2-179, respectively. (B) Junction regions of T-DNA direct repeats. (C) T-DNA border junctions at T-DNA rearrangement in which the enhanced 35S promoter (pe35S) with truncated neomycin phosphotransferase II (*neo*) selectable marker gene were located in the RB T-DNA flanking sequences. (D) Identical but the pe35S and truncated *neo* gene were found in the LB T-DNA flanking sequence. (E) Vector backbone near the LB region located in the RB T-DNA flanking region. Light and dark grey backgrounds refer to LB and RB repeats, respectively. The arrow indicates the theoretical cleavage site in the complementary strand of the LB in (A) and between brackets the distance in base pairs of the nicking site to the predicted theoretical cleaving site of the corresponding border in (B), (C), (D) and (E). Lowercase represents filler DNA while lowercase italics and underlined lowercase italics represent filler DNA from T-DNA and vector backbone origin, respectively. The 715 bp region homologous to the insertion element IS113 from *A. tumefaciens* is comprised in the double underlined 980 bp in the sequence 156-1LB. Double colons (::) illustrate junction positions between different DNA sequences. Start codon of the *luc*$^{+}$ is boxed.

4.4.2 *In silico* analysis of T-DNA flanking sequences

The banana T-DNA flanking sequences in promoter-tagged lines ET2-17, ET2-34, ET2-85 and ET2-179 were analyzed for the presence of *cis*-acting elements and promoter characteristics in the 5'-tagged sequences, and for putative coding regions in the 5'- and 3'-tagged sequences (Figure 4.14 and Annex). Databases were queried in the period comprising July 2005 until February 2007.

As a first indication for the presence of promoter sequences within the 5'-tagged regions, an *in silico* search for *cis*-acting elements was performed and only elements involved in drought and/or low temperature (dehydration-responsive element, DRE; induction of CBF expression region 1 or ICEr1) and abscisic acid response (abscisic acid-responsive element, ABRE) together with candidate TATA boxes are indicated (Figure 4.14). The 1.74 kb 5'-tagged sequence 17-1 of line ET2-17 contains four DRE-like, one ICEr1-like and four ABRE elements, which is significantly more than in any of the other 5'-tagged regions. The two candidate TATA boxes are located at positions -390 and -200 relative to the T-DNA right border junction with all other elements upstream of them. However, a TATA-less promoter was identified by the TSSP software providing a candidate transcription start site at position -512 (Figures 4.14 and 4.15). The lack of homology to any available sequence in the databases for this 5'-tagged sequence plus the corresponding linked 3' region suggests that a cryptic promoter might be tagged in sequence 17-1. Analysis of the sequence 17-1 revealed a *Hind*III restricted fragment size of 4597 bases comprising part of the 5'-tagged sequence and the T-DNA (data not shown) corresponding well to the signal detected in the Southern blot analysis of the sequence 17-1 (Figure 4.6B). Only one DRE-like element and two putative TATA boxes were located in the 1.28 kb 5'-tagged sequence 17-2. Analysis of the 5'- and 3'-tagged sequence revealed that the T-DNA was integrated in a transcribed region (Figures 4.14 and 4.16) showing high homology over a length of 281 bp to a 596 bp banana EST (6000092615T1; 96% identity at $E < 1 \times 10^{-135}$). In addition, two other regions from the same EST accession (6000092615T1, 93% and 96% identity at $E < 3 \times 10^{-53}$ and $E < 3 \times 10^{-04}$, respectively) were found in the 5'- and 3'- flanking regions of the high homology region. Two of the sequences homologous to the banana EST accession contain microhomology at the ends (AGGTT) and 219 bp between both sequences has no homology to any known sequence and may indicate the presence of an intron (Figure 4.16).

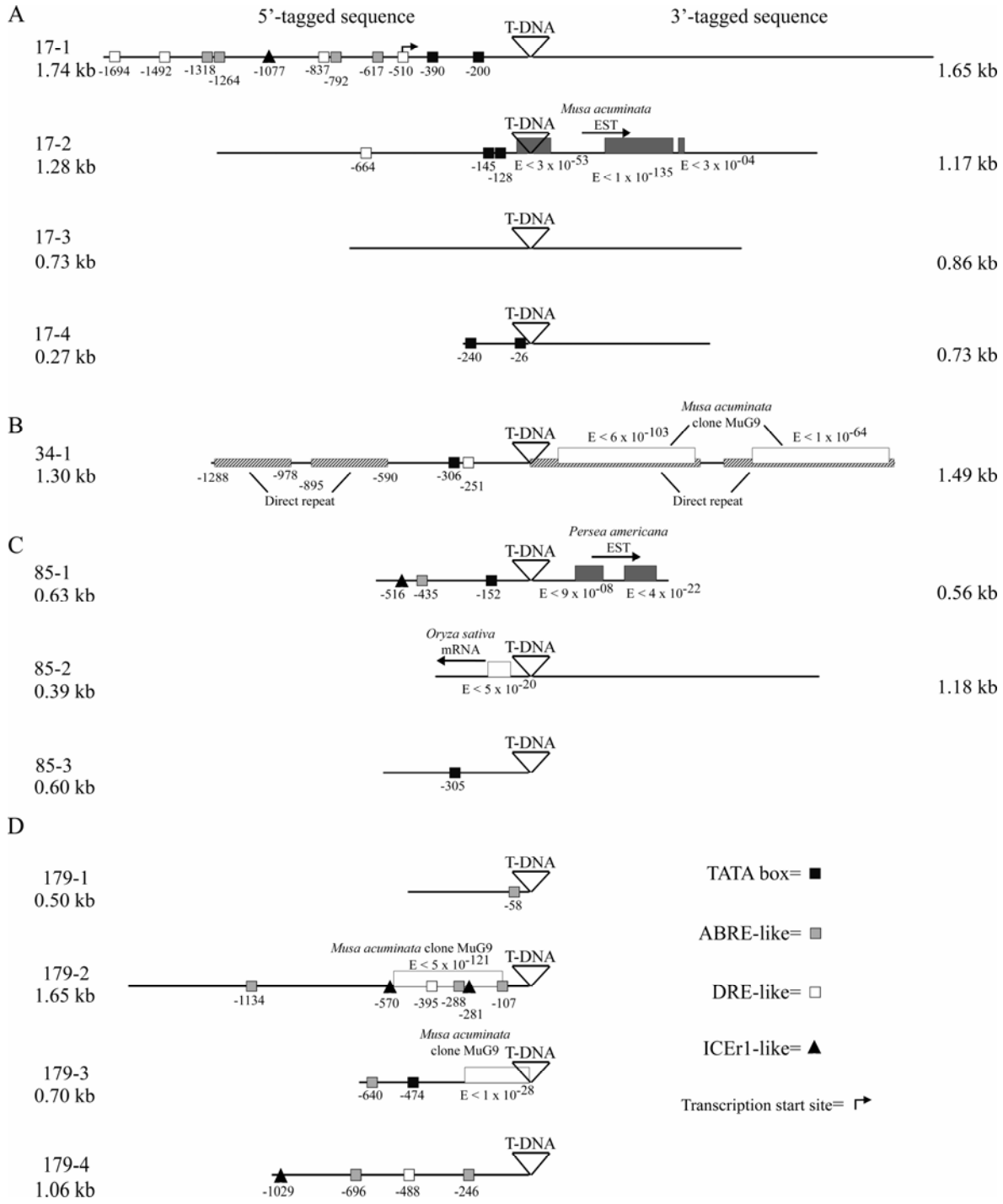

Figure 4.14. Presence of promoter *cis*-acting elements in the 5'-tagged regions and sequence homology in the 5'- and 3'-tagged regions in the promoter-tagged lines ET2-17 (A), ET2-34 (B), ET2-85 (C) and ET2-179 (D). The presence of putative TATA boxes, ABRE-, and DRE-like *cis*-acting elements in the 5'-tagged sequences are illustrated and were determined by querying the PlantCARE and/or PLACE databases. The core sequence of the ICEr1 *cis*-acting element (CACATG) was located manually. The distance of the *cis*-acting elements to the T-DNA insertion site is indicated. Open triangle represents T-DNA insertion site. The small arrow indicates a transcription start site defined by the TSSP software in sequence 17-1 (position -512). Dark grey and white boxes indicate homology to EST and to genomic sequence accessions as queried in the GenBank, respectively, and the corresponding E value is indicated. Arrows indicate the direction of transcription if applicable. The presence of direct repeats is illustrated in boxes with diagonal lines. The size of the 5'- and 3'-tagged sequences are indicated in the left and right side of the figure, respectively. The diagram is not precisely according to scale. kb: kilo base pairs.

Putative 5' and 3' intron splice sites were found in the microhomology region between the two exons of the 3'-tagged sequence 17-2 (AG/GTTAGC, 5' splice site; GTGATCAG/GT, 3' splice site; Figure 4.16) showing homology to the intron splice sites consensus of dicots and monocots (AG/GTAAGT, 5' splice site; TTTTGCAG/GT, 3' splice site; Simpson and Filipowicz 1996). The highly homologous region of 281 bp shows homology to the last 90 amino acids of an unknown rice protein (GenBank accession no. BAD87356, 74% identity and 87% positives at $E < 2 \times 10^{-32}$) and 85 amino acids of another unknown rice protein (NP_916242, 71% identity and 87% positives at $E < 2 \times 10^{-29}$). Besides two candidate TATA boxes in sequence 17-4, no other relevant promoter elements were identified in the two remaining 5'-tagged sequences 17-3 and 17-4 (Figure 4.14) and database searches did not reveal homology to any known sequence for these 5' regions and their linked 3' regions.

```
   1 GTAGAGATGA TCCATGAGCA ATTTTGAAGA GATCGCCAAA TATGAACCGA CAGAAGATAT ATAGTGACAG AACCAACCTT CCGCTCCCAG CTTGGATCGG
 101 GCCCTGTGGT CGTTCAAGAT CGACAGGCCG TGCTTGAGGA CCTTGTTGGA TCGCCGGGCC TTGAGCTTGC GCACGAGGCA GAGGTACTCC CACGCCCCCC
 201 CGCCGCCGTT CTCGACCTGG CTCTCGAGCC GCTCCATCCG CGCTGCCTCC GACGCCCCCA CCATCGGCCT GCGCTTGCCG GCGAAAACCT CCTTGCGGGC
 301 GAGATGTAGA AGGAAGAGAG AGGAAACAGA GCAACGCCGG CCGCCGTGAT TCCAATAGCG ATCGTCCTCG AAAGCAATGG GAGATCCAAG GAAAACGTAC
 401 CCCAGAACGG ATCCTCTCCA GCACGTGTCG GATCCAATAG CTCCTCTCCA CAGCCCAGCC GTCCACAGTG GGATCCACGT GCATCAATGA TACCAATGGA
 501 GGGTTACGAT AGAAGAATGA GCGGATGGAG GGTGGAGATC GAGAGAGGTT AATCATCGGA CAGGTTCGAT TTGGAGAGGA TCACATAAAC CTTCGTCCAA
 601 GGCGGATCCG ATGACATGAC TGGAGTTGGG TTATTGGACT TGGGGCTCGA TCTAGACTCG CAACACATGA GATATGAACT CGGGTGCAAA TGAGTCCAAT
 701 CCATCTGAAA ACAGTTCAAA TTTATTCGTT AATATTAATG GTGTTTAATT AATTATTTAT TTATCTTAAC CATGATTGCT TGTATTGGAG ACCGAGGCAG
 801 AAAAATTAAT AGATTAAAAA TAAAATGAAA TAACAAAAAA TAGAAATCTT AAATCAATTA AATGGAATCC GGACCCGATT AGAGCGGGTC GCGTATGATA
 901 AATCCGACCC GCTGCCATCC GCTATCGCAT CGGGTCAGCG TGAGACTGAC GTGGCACGCG CGCCACTGTT ACTTGGTACA CCAAGCGTCA CAGCATCACA
1001 CAAAGGGAGT GCTTCGCGCA CTCGCAGAAA CGCCCGTCAT CTCCTCGATT AAAATTTATA TGCGGACAAG GTCGGTCCTT CGATTCCAAC CGTTCGATGA
1101 TCTCGCATAG GGATCGCTTC CGGACGTACG CCGTCGGATG TTATCGTTGG TTAAATGACG TTAATACCCT TGCCCTCGGT TTCTTGTGCC TCCCTCACGC
1201 CTCTCGCATG TGATCAGTTG CTAATCCAAC CCGACCCGCC CACACCTCTC TCTCTCCCTC CCTCCCTTCC TCACCCCACC CACCCACCCC CCGGGGAGAC
1301 TACGGTAGAT CTCAGATCGG AACTCTCATT GATCGTTTCT CTCGGGTTAA TATACGCGGT CGCGAGGACG GAGGAAGCTC GAATCGGTTC CTTTCTTGCT
1401 CAATTCAGCC AGGTACAGCT CTCCTGACTT CTTTTCCGGG GATTTTCCCC TCTTTTCGAT GTTATTAATG TGGTGCTGAA TTTTATGAGT TCCTGTTGGT
1501 GGATCTGCAC GCTGAGATGA GTAGGATTCA TTTCCGGTTC TATATTTCCG TAGTTTACTT TGTAGTTCGT CTCATCGGAT TAGGGCTTAG TTTTTTCTTA
1601 ATGTAAGGTT TAGGGATTTG TTATGCTCTG TAAGTAATTG GTGATGTTAC ACGCTTCGGA ACAATAAATT AGGTCTGTTG CGAGATCGGA AACGATTGGA
1701 TGCCAGGAAA CATGACACTT GGACTATGCG CGTGTTGCTT▼CTATCCACGA GGTGGAGATT TTTCACTTGG ACTACTCAAA TCTGGGACAC TTAATTATTG
1801 CTAAAATGGA GGGACAGCCT TTTTCATGAA TCCTCAGTAA TGCTTCCAGC ATTTTGATGA ATAATGTAGC CTGAGAAGTT TATGAAAGGC GTGATGATTT
1901 GCATATAGTA ACGGTTAGGT TAAGACTTGA GAATAGCAAG AAGCGATAAC RAAGACTTGG ATATCTTAAA TTATCTTCAA AGATGTTCTA TTTCTTCAAA
2001 AATTATTATT AACATCTGTA TCGTGATCCC ATTTGCTGCA TTTATTACTT CTCTTTGGTC GGCAATGCCT GAATTATCCA GGACAAATCT GAAGCACCAA
2101 GTCTACTGCG TTGGCTTGTT CCAAAAACAC TTARCTGACT GATTTGTGGC AACCTTCCTA TATACAAAAT AGTTCAATTA ATTATTGCAT CAGTCATATA
2201 ATTTCATGTC ACAGCATTGG TAGTGTARAT TATATAACTT CTGCCGATGG TTATCTATGC GAAGCCAGCA TTACTAACAG GTTAGGGACT ACTCTCATCT
2301 AACACAATAC TGTTTGAGGA TCACAATGAA TAAAGGTACA AGGTTTAGAT TGCCCATTGA GATAAATTTC AATTATCATT GGTCTCACTG CACATTTGTT
2401 TTATGAAATA TTACATACTG TATCTTCTTG TTTCGTTCTA ATGATCTTTT ATTCCTTCTA TTGGATTTCT TGCCTTCTGT TTCTTATTTG GCATTTGGTG
2501 AAGCGCCAGA CTTAGATGCC AAATGCATGA TCACCATGTT GGTTCTAGTC CATGAAGGCT GCAAATAGAG TGCATGTTAG TGCTCAATTA TCAAGGTTAA
```

Figure 4.15. 5'- and 3'- tagged region of insertion 17-1 of the promoter-tagged line 17. Double underlined uppercase indicate DRE-like elements, while underlined uppercase indicate ABRE-like elements. Bold uppercase italics represent the core sequence of the ICEr1 *cis*-acting element. Bold uppercase represents putative TATA boxes. Closed triangle represents T-DNA insertion site. Boxed nucleotide indicates transcription start site according to the TSSP software. R refers to A or G.

Figure 4.16. **5'- and 3'- tagged region of insertion 17-2 of the promoter-tagged line ET2-17.** Double underlined uppercase indicate a DRE-like element. Bold uppercase represents putative TATA boxes. Closed triangle represents T-DNA insertion site. Bold lowercase italics indicate sequences homologous to a *Musa acuminata* EST (600092615T1) from the Syngenta EST database. Boxed nucleotide indicates homologous sequences at the end of two putative exons. Y refers to T or C.

Analysis of the 1.3 kb 5'-tagged sequence of tagged line ET2-34 revealed the presence of one TATA and one DRE-like element together with a 306 bp direct repeat starting 590 bp upstream of the T-DNA insertion site (Figures 4.14 and 4.17). The direct repeat region lacked homology to any known sequence, but its presence suggests that integration of the T-DNA occurred in a repetitive region of the banana genome, which is strengthened by the repetitive presence of 694 bp within the 1494 bp 3'-tagged sequence (Figures 4.14 and 4.17) and indicates the tagging of a cryptic promoter. The latter two repetitive regions show over a length of 557 bp and 421 bp, and 437 bp and 412 bp high homology to the BAC clones MuG9 (AY4845883, 84% and 82% identity at $E < 6 \times 10^{-103}$ and 1×10^{-64}, respectively) and MA4_54B05 (AC186753, 80% and 81% identity at $E < 1 \times 10^{-43}$ and 6×10^{-43}, respectively) of the wild banana *Musa acuminata* 'Calcutta 4' (AA genomic group, ITC.0249), respectively. Thus, the repetitive region in the 3'-tagged region was also present in the genome of 'Calcutta 4', and additional homology of the two regions over a length of 430 bp and 432 bp to the BAC clone MuH9 (AY484589, 81% and 80% identity at $E < 2 \times 10^{-42}$ and 9×10^{-42}, respectively) of this cultivar was found. Four homologous regions were also detected in the BAC clone MBP_31O07 (AP009334) of the *Musa balbisiana* 'Pisang Klutuk Wulung' (BB genomic group) over a length of 261 bp (87% and 86% identity at $E < 2 \times 10^{-62}$ and $E < 1 \times 10^{-57}$, respectively), and 221 bp and 251 bp (86% and 84% at $E < 2 \times 10^{-47}$ and $E < 3 \times 10^{-46}$, respectively). In addition, the two regions of the repetitive sequence of the 3'-tagged sequence of line ET2-34 show homology to different regions of a partial cDNA clone isolated from leaves of *Musa acuminata* subsp. burmannicoides subjected to heat stress (DN238884.1, 84% and 83% at $E < 3 \times 10^{-22}$ and $E < 2 \times 10^{-17}$, respectively). Finally, the 3'-tagged sequence in ET2-34 shows homology to a

banana EST clone over 195 bp (600170679T1, 82% identity at E < 2 x 10^{-22}) suggesting that transcription may occur despite the presence of a repetitive region or homologous sequences might be expressed in other loci of the banana genome.

Figure 4.17. 5'- and 3'- tagged region of insertion 34-1 of the promoter-tagged line ET2-34. Double underlined uppercase indicates a DRE-like element. Bold uppercase represents a putative TATA box. Closed triangle represents T-DNA insertion site. Lowercase italics and bold lowercase italics indicate direct repeats at the 5'- and 3'-tagged sequences, respectively. Boxed are the nucleotides removed during T-DNA integration. K refers to G or T.

Southern hybridization with a probe specific to the 5'-tagged region of the promoter-tagged line ET2-34 was performed to discover how common the tagged repetitive region is in the banana genome and to verify the link between the inserted T-DNA and the 5'-tagged sequence after hybridization with a *luc*$^{+}$-specific probe (Figure 4.18). The 5'-tagged sequence-specific probe was DIG-labeled using a forward primer which anneals in the repetitive region and a reverse primer annealing near the T-DNA insertion site yielding two amplicons (827 bp and 1215 bp; Figure 4.18A). The use of this probe mixture revealed the abundance of the repetitive region in the banana genome as shown by the smear observed in the control non-transformed line and tagged line ET2-34 although a specific band around 3 kb was detected (Figure 4.18B).

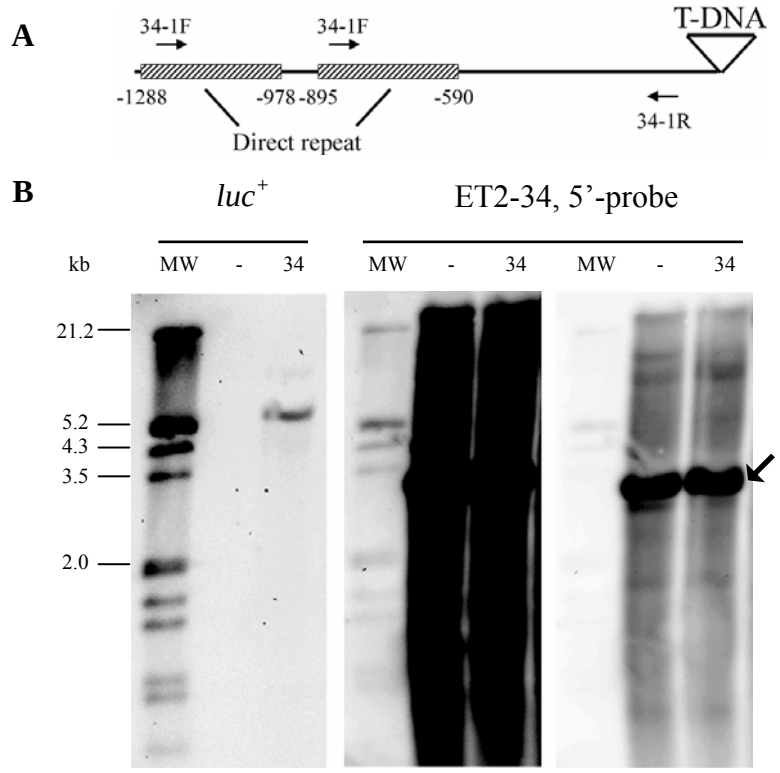

Figure 4.18. Southern hybridization analysis for the integration of the *luc*$^+$ gene and the cloned 5'-tagged region of the promoter-tagged line ET2-34. (A) Schematic representation of the 5'-tagged sequence. The presence of repetitive region is illustrated as boxes with diagonal lines. Numbers indicate the distance to the T-DNA insertion site. Arrows indicate the annealing site of the primers used for the labeling of the 5'-tagged sequence-specific probe for Southern hybridization. (B) Ten micrograms of total DNA were digested with *Hind*III, separated fragments were hybridized with a DIG-labeled *luc*$^+$ probe (862 bp) and re-hybridized with a 5'-tagged sequence-specific probe of line ET2-34 (mixture of 827 bp and 1215 bp). Image at the right is less scaled to allow the ~3 kb signal to be distinguished. -: negative non-transformed control plant. 34: tagged line ET2-34. MW: DIG-labeled DNA molecular marker III (Roche). Arrow indicate signal of ~3 kb. kb: kilo base pair.

Analysis of the 5'-tagged sequence 85-1 revealed the presence of an ICEr1- and ABRE-like sequence, and a near-canonical candidate TATA box (Figures 4.14 and 4.19). The respective 3'-tagged sequence contained two regions with homology to a *Persea americana* EST (CK758727.1, 82% and 86% identity at $E < 9 \times 10^{-08}$ and $E < 4 \times 10^{-22}$, respectively), while the 5'-tagged sequence 85-2 showed homology to an *Oryza sativa* mRNA in the reverse orientation (AB071299, 89% identity at $E < 5 \times 10^{-19}$). No relevant *cis*-acting elements were found in the 5'-tagged sequence of insertion 85-2 and the 3'-tagged sequence showed no homology to any known sequence (Figure 4.14). For the insertion 85-3, no homology was found in the 5'-tagged sequence, while the 3'-tagged region could not be isolated. Only a candidate TATA box was located 305 bp upstream from the T-DNA insertion site in the 5'-tagged sequence 85-3 (Figure 4.14).

Figure 4.19. 5'- and 3'-tagged region of insertion 85-1 of the promoter-tagged line ET2-85. Bold uppercase italics represent the core sequence of the ICEr1 *cis*-acting element. Underlined uppercase indicate an ABRE-like element. Bold uppercase represents a putative TATA box. Closed triangle represents T-DNA insertion site. Bold lowercase italics indicate sequences homologous to a *Persea americana* EST (CK758727). Boxed are nucleotides removed during T-DNA integration.

For tagged line ET2-179, four 5'-tagged sequences were identified (Figure 4.14). Isolation of 3'-tagged sequences revealed only non-specific or vector backbone sequences (Annex). The sequences 179-1, 179-2, 179-3 and 179-4 contain one, three, one and two ABRE-like elements, respectively. Only the sequence 179-3 possesses a candidate TATA box 474 bp upstream from the T-DNA insertion site. In addition, the sequences 179-2 and 179-4 contain one DRE-like element each, and two and one ICEr1-like elements, respectively. The sequences 179-2 and 179-3 show homology to known sequences. The sequence 179-2 contains homology over a length of 443 bp to two repetitive regions of the banana BAC clone MuG9 (AY4845883, 88% and 87% identity at $E < 5 \times 10^{-121}$ and 7×10^{-114}, respectively), and over 443 bp and 418 bp to the banana BAC clone MA4_54B05 (AC186753, 86% and 85% identity at $E < 4 \times 10^{-106}$ and 8×10^{-89}, respectively) of the wild banana 'Calcutta 4'. In addition, the 443 bp region starting 111 bp upstream of the T-DNA insertion site shows homology to the 3'-tagged repetitive region 34-1 of line ET2-34 (80% and 80% identity at $E < 5 \times 10^{-85}$ and $E < 3 \times 10^{-83}$, respectively) suggesting that the insertions 179-2 and 34-1 of two independent lines occurred in homologous repetitive regions. Consequently, homologies of the 443 bp region of 179-2 to the *Musa balbisiana* BAC clone MBP_31O07 of 'Pisang Klutuk Wulung' (AP009334, 91% identity at $E < 3 \times 10^{-73}$) and to the cDNA clone (DN238884, 90% identity at $E < 1 \times 10^{-81}$) isolated from leaves of *Musa acuminata* subsp. burmannicoides subjected to heat stress were also found. Moreover, homology to three banana ESTs including the EST accession 600170679T1 detected in the 3'-tagged sequence of insertion 34-1 was also obvious in the 5'-tagged 179-2 sequence (600178185T1, $E < 4 \times 10^{-67}$; 600098103T1, $E < 5 \times 10^{-51}$; 600170679T1, $E < 2 \times 10^{-50}$). The 5'-tagged sequence 179-3 showed homology to the banana BAC clone MuG9 (AY4845883, 82% identity at $E < 1 \times 10^{-28}$) in a region 2.2 kb apart from the homologous sequence of 179-2 in MuG9.

A summary of all the identified T-DNA flanking sequences and *in silico* analyses for the promoter-tagged lines ET2-17, ET2-34, ET2-85, ET2-156 and ET2-179 is shown in Annex. Homology searches of tagged sequences in promoter-tagged line ET2-156 revealed

the presence of DNA of prokaryote origin in 3'-tagged sequences suggesting the integration of *Agrobacterium* DNA not comprised in the pETKUL2 vector and the presence of *Rhizobium* DNA (Annex).

4.5 Discussion

T-DNA promoter and gene tagging studies in model plants usually yield on average one to two T-DNA copies per transgenic line (tobacco: Mudge and Birch 1998; Arabidopsis: Alonso *et al.* 2003; rice: Jeon *et al.* 2000). In this study, Southern hybridization analysis of 10 promoter-tagged banana lines showed an average of 3.6 T-DNA copies per line. This makes the isolation and identification of the activating 5'-tagged region(s) laborious and time consuming, and explains why researchers preferentially analyze mutants with a single T-DNA copy integration. Promoters of the Arabidopsis genes *EXORDIUM* (Farrar *et al.* 2003) and *HVT1* nucleic acid helicase (Wei *et al.* 1997) and of the *Lotus japonicus* gene *Cbp*1 (Webb *et al.* 2000), a tobacco meristem-specific promoter (Mudge and Birch 1998) and Arabidopsis root-specific cryptic promoters (Ökrész *et al.* 1998; Mollier *et al*. 2000) were all tagged in single T-DNA copy lines. Because of the low incidence of single copy T-DNA insertion in banana tagged lines one aim of this study is to demonstrate that the activated T-DNA insertion(s) can be identified in multicopy T-DNA mutants like line ET2-17.

Although few flanking sequences isolated by TAIL-PCR and I-PCR were not specific, the combination of both PCR-based methods yielded in the majority of the lines the expected number of RB flanking sequences (based on the T-DNA copy number) demonstrating the usefulness of different walking techniques with different modes of operation for retrieval of the flanking sequences in multicopy T-DNA lines. The physical linkage between the isolated 5'-tagged sequences and the *luc*$^+$ gene was proven by Southern hybridization.

Due to the relatively low number of retrieved 3'-tagged sequences, linking RB with LB flanks by PCR was incomplete. For instance in line ET2-17, five and three different 5'- and 3'-tagged sequences were isolated, respectively, of which only two 5'- and 3'-tagged sequences isolated from the T-DNA were linked. Isolation of the remaining 3'-tagged sequences was performed through TAIL-PCR with primers specific to the 5'-tagged regions that could not be linked (17-1 and 17-3). Subsequent linking-PCR confirmed the linkage between these remaining 5'- and 3'-tagged sequences. On the other hand, in the promoter-tagged line ET2-85 three and four 5'- and 3'-tagged different sequences were isolated, respectively, of which only two 5'- and 3'-tagged sequences were linked suggesting the presence of five 5'- and 3'-tagged sequences in this line. Truncations during T-DNA integration occur preferentially at the LB (Gheysen *et al.* 1991; Mayerhofer *et al.* 1991). However, Pérez-Hernández *et al.* (2006a) found that LB regions were more conserved than

RB in transgenic banana lines as analyzed by an anchored-PCR (A-PCR) based method. The authors suggest that the difference between the borders might be influenced by the shorter distance between the primer annealing site used for the A-PCR and the nicking site in the RB than in the LB (54 bp *vs*. 217 bp; Pérez-Hernández *et al.* 2006a). In the present work, the distance of the annealing sites for the primary, secondary and tertiary TAIL-PCR and I-PCR specific primers relative to the RB nicking site are 236 bp, 185 bp and 92 bp; and 339 bp, 218 bp, and 105 bp or 68 bp relative to the LB nicking site. Distances between primers and border nicking sites were thus comparable for both T-DNA borders making a difference in the detection efficiency of truncated insertions between the T-DNA borders less likely. Therefore, truncation at the LB region might have occurred in the promoter-tagged line ET2-17 at insertions 17-1 and 17-3 avoiding the annealing of specific primers for the direct isolation of 3'-tagged sequences, whereas in the promoter-tagged line ET2-85 both borders might have been truncated preventing the annealing of the specific primers used in TAIL- and I-PCR.

Sequence analysis revealed T-DNA tandem repeats in two (ET2-49 and ET2-156) out of nine (22.2%; Table 4.3) promoter-tagged lines analyzed, which correlates with frequencies obtained in rice (~29%, Zhu *et al.* 2006) and aspen (~21%, Kumar and Fladung *et al.* 2000). However, the only T-DNA tandem observed here was a direct repeat configuration, but with a different T-DNA/T-DNA junction region in each line. Furthermore, these two promoter-tagged lines contained an additional T-DNA tandem direct repeat with a truncated T-DNA in which the enhanced 35S promoter with a truncated neomycin phosphotransferase II (*neo*) was located upstream of the *luc*$^+$. Two other promoter-tagged lines (ET2-89 and ET2-111) showed similar T-DNA truncation patterns. In total, 5/9 or 55.6% of the lines showed multiple T-DNAs either in direct repeat and/or truncated conformation. Furthermore, 14 out of 16 promoter-tagged lines which were PCR-screened for the specific T-DNA truncation, whereby part of the enhanced 35S promoter and truncated *neo* are located upstream of the *luc*$^+$, revealed the presence of a T-DNA direct tandem repeat (9/16) and/or truncated T-DNA direct repeats (10/16). Besides their LUC activity profile under LT stress throughout development this finding supported the choice to analyze among others the lines ET2-17 and ET2-34 in greater detail which both lacked an enhanced 35S promoter and truncated *neo* upstream of the *luc*$^+$. T-DNA truncation in T-DNA tandem integration patterns are not uncommon and have been observed in Arabidopsis (Meza *et al.* 2002), rice (Zhu *et al.* 2006) and bentgrass (Fu *et al.* 2006). Furthermore, T-DNA truncation was detected in promoter-tagged lines of Arabidopsis (Kertbundit *et al.* 1998) and *Brassica napus* (Bade *et al.* 2003), in which the promoter driving the selectable marker gene was located upstream of the promoterless *uid*A and *uid*A::*neo* fusion gene, respectively. Bade *et al.* (2003) obtained two tagged lines with the rearranged 35S promoter upstream of the *uid*A::*neo* fusion gene and that showed constitutive GUS expression. The banana tagged lines ET2-49, ET2-111, and ET2-156 that

contained the enhanced 35S and truncated *neo* upstream of the *luc*$^+$ showed constitutive and strong or very strong LUC activity during the different developmental stages, except for line ET2-111 in early developmental stages (Table 3.3). The presence of an in-frame stop codon (TAA) 18 bp upstream of the *luc*$^+$ start codon may abolish translational fusion between the *neo* and *luc*$^+$, but at least in line ET2-49 a translational fusion might have occurred since it contained a truncation in the RB T-DNA region that resulted in the loss of the stop codon but the truncated *neo* and *luc*$^+$ remained in-frame. The possibility can also not be ruled out that the enhancer properties of the 35S promoter (Benfey *et al.* 1990) at one or more developmental stage(s) are to a certain extent responsible for LUC expression not only in such rearranged T-DNA insertions but in all the promoter-tagged lines and alter the true activity profile of the tagged promoter. A higher reporter gene activation frequency was obtained in transgenic rice when using a gene tagging vector containing the CaMV 35S enhancers than in a vector without the enhancers (Jeong *et al.* 2002). Nevertheless, the authors concluded that the sensitivity of the experiment was increased and that the enhancer elements do not affect the expression pattern of nearby rice promoters. On the other hand, Yoo *et al.* (2005) state that the 35S promoter driving the selectable marker gene is responsible for the altered expression pattern of transgenes and this effect might be governed by the presence of the 35S enhancer. Thus, caution in the interpretation of reporter gene activity should be taken when using a plant transformation vector carrying the CaMV 35S promoter.

In order to gain more understanding of the T-DNA integration process in banana, analysis of the T-DNA/banana DNA, T-DNA/T-DNA (tandem repeat) and T-DNA/vector backbone junctions was performed in nine promoter-tagged lines. Sequence analysis of the T-DNA/banana DNA junction regions demonstrated that the RB was more precisely nicked than the LB similar to other results in banana observed by Khanna *et al.* (2004) but in contrast to the data obtained by Pérez-Hernández *et al.* (2006a). The outer and inner surrounding regions of the LB affect the correct recognition of the LB and influence the presence of vector backbone sequences in the resulting transgenic plants (Podevin *et al.* 2006). Similarly, the surroundings regions of the RB are important for correct function during the nicking of the T-DNA (De Buck *et al.* 2000; Meza *et al.* 2002). Thus, the plasmid carrying the T-DNA sequence may play an important role in the correct nicking of the T-DNA and might be an explanation for the difference in correct nicking of the T-DNA borders described in the present work and transgenic banana lines analyzed by Pérez-Hernández *et al*. (2006a). In this study, 62.5% (10/16) of the RBs the cleavage site was between the third and fourth base as theoretically expected, which is in agreement with the results of Kim *et al.* (2003) in rice where 55% of the RBs examined (29/53) were nicked at the theoretical nicking site. The nicking of the LB was less accurately, only 11% (1/9) of the LBs in banana promoter-tagged lines was nicked at the expected position between the third and fourth base. This result again agrees well with transgenic rice lines where none of

the LBs examined was nicked at the theoretical cleavage site (Kim *et al.* 2003). Similar analysis of T-DNA border junction regions in rice revealed that the conserved cleavage occurred in 6% of the LBs and 43% of the RBs (Zhu *et al.* 2006). Reports in other plant species indicate similar frequencies of correct nicking of the RB including Arabidopsis and tobacco (7/15, Tinland 1996), aspen (18/27, Kumar and Fladung 2002) and bentgrass (13/26, Fu *et al.* 2006). The less accurate nicking at the LB might be a consequence of the low efficiency of nicking at this border. Probably another factor involved is the lower protection of the LB region of the single-strand T-DNA, unlike the RB region, resulting in a more random nicking.

By merging the data from 7 RB and 5 LB T-DNA/banana DNA junction regions, three classes of T-DNA integration were distinguished. Two out of 12 junction regions (16.7%) showed no filler DNA and no microhomology between the T-DNA and banana DNA ends, 25.0% (3/12) showed filler DNA and 58.3% (7/12) showed microhomology of maximum 7 bp. These data correlate well with the frequencies obtained in T-DNA/rice DNA junction regions comprising 9.8% with no filler DNA and no microhomology, 32.1% with filler DNA and 58.3% with microhomology (Zhu *et al.* 2006).

Removal of short fragments of plant DNA is common during T-DNA integration. In rice and in dicotyledonous plants the size of the deleted plant DNA at the site of T-DNA integration ranged from 0 to 70 bp (Kim *et al.* 2003) and from 13 to 73 bp (Gheysen *et al.* 1991; Mayerhofer *et al.* 1991), respectively. Linking-PCR was successful for five T-DNA integrations in three promoter-tagged lines which allowed, after sequencing, the verification of removed banana DNA during integration. Four (17-4, 34-1, 85-1 and 85-2) out of five T-DNA integration events showed deletion of banana DNA and three deleted regions fell into the range observed in the literature (0 -73 bp), but in one T-DNA integration event (85-2) 347 bp banana DNA was deleted. Similarly, Scholte *et al.* (2002) found that deletions comprising 5 to 404 bp occurred in T-DNA tagged *Medicago truncatula* lines.

In silico analysis of the four 5'-tagged candidate banana promoter sequences in line ET2-17 suggested that two promoters were tagged. The first T-DNA insertion most likely tagged a cryptic promoter although the functionality of the two candidate TATA boxes remains to be investigated, and contains relevant promoter elements to explain the LT up-regulated LUC activity in undifferentiated cultures. Tagging of cryptic promoters is not uncommon in plants and most of them are even tissue-specific (tobacco: seed coat, Fobert *et al.* 1994; *Arabidopsis*: callus and roots, Ökrész *et al.* 1998, guard cells, Plesch *et al.* 2000, and roots, Mollier *et al.* 2000 and Sivanandan *et al.* 2005). High homology of part of the 3'-tagged sequence to a banana EST and the last 90 amino acids of an unknown rice protein suggest that the second T-DNA insertion occurred in a coding region. Similarly, numerous reports describe preferential T-DNA integration in transcribed regions (Szabados *et al.* 2002; Alonso *et al.* 2003; An *et al.* 2003; Chen *et al.* 2003; Sallaud *et al.* 2003; Ryu *et al.* 2004).

T-DNA integration in promoter-tagged line ET2-34 occurred in a repetitive region. Southern hybridization using a *luc*$^+$ probe revealed the integration of at least one T-DNA copy. Although at first sight the repetitive regions in the tagged regions of insertion 34-1 indicated the absence of transcriptional activity in this region, searches in databases inferred that such repetitive sequences might be expressed due to the homology to banana ESTs. However, it can not be ruled out that the insertion of the T-DNA triggered the activity of surrounding sequences, which is a distinctive feature of cryptic promoters isolated by T-DNA tagging (Sivanandan *et al.* 2005).

For promoter-tagged line ET2-85, one out of three 5'-tagged sequences is expected to be a banana promoter due to the identification of a near-canonical TATA box and by the presence of a putative coding region in the corresponding 3'-tagged sequence. Several 5'-tagged sequences were obtained in line ET2-179 and identification of a candidate promoter sequence was not straightforward. One 5'-tagged sequence might be discarded due to the absence of representative *cis*-acting elements and TATA boxes and thus, one or even two of the remaining sequences could be the candidate promoter(s).

In summary, the number of T-DNA inserts averaged 3.6 in 10 independent promoter-tagged lines. Isolation of T-DNA flanking sequences was accomplished through TAIL-PCR and I-PCR. Sequence analysis revealed the presence of direct tandem repeats (lines ET2-49 and ET2-156) and vector backbone/T-DNA rearrangements (lines ET2-17, ET2-49, ET2-123, ET2-156, ET2-179; and ET2-17, ET2-49, ET2-89, ET2-111, ET2-156, respectively). The lines ET2-49, ET2-89, ET2-111 and ET2-156 contain in a 5'-tagged sequence the enhanced CaMV 35S promoter with part of the selectable marker gene *neo* from the tagging construct. Hence, most of the T-DNA tagged lines showing LUC activity contained multiple T-DNA copies and analysis of T-DNA flanking sequences revealed a high proportion of vector backbone and T-DNA rearrangement in the tagged lines. *In silico* analysis of the four 5'-tagged banana sequences in line ET2-17 suggest that two are candidates for driving LUC expression. The 5'-tagged sequence in line ET2-34 contained a repetitive region. In the line ET2-85 three candidate promoter sequences were retrieved with one containing a near-canonical TATA box and probably linked to a gene. In the following chapter the transcriptional fusion between the 5'-tagged sequences and the *luc*$^+$ gene is explored to discover, in the multicopy T-DNA lines ET2-17, ET2-85 and ET2-156, or confirm, in the single T-DNA copy line ET2-34, the tagged sequence that is responsible for the observed LUC activity.

Chapter 5 Identification of activated T-DNA insertion by RT-PCR

5.1 Introduction

The identification of promoters or genes in T-DNA tagging experiments is greatly facilitated when the sequence of the entire genome of the host plant is available. T-DNA flanking sequences from tagged lines can be easily mapped and analyzed based on the prior annotation of the sequences (Yamamoto *et al.* 2003). Moreover, in model plants the expression of the reporter gene in a T-DNA tagged line can often be linked to the known expression of the annotated gene and sometimes the promoter has already been identified and characterized in transgenic lines. However, expression can also be triggered by cryptic promoters when the T-DNA is integrated in antisense orientation of an annotated gene or in intergenic regions (Yamamoto *et al.* 2003). Cryptic promoters are frequently discovered in tagging experiments (Ökresz *et al.* 1998; De Greve *et al.* 2001; Fobert *et al.* 1994; Foster *et al.* 1999; Mollier *et al.* 2000; Sivanandan *et al.* 2005; Stangeland *et al.* 2005), and are found in T-DNA flanking sequences causing even silencing of transgenic promoters by aberrant transcription (Eike *et al.* 2005). In case of the non-model crop banana two additional complications exist that, until now, make the identification of tagged genes and promoters difficult. Firstly, only a fraction of the banana genome has been sequenced (~1% based on BAC sequencing, bioinfo.inibap.org/statusdb/stats.php?page=clone, 21 March 2008) excluding the possibility of mapping the banana T-DNA flanking sequences. However, additional BAC (end)-sequences (Cheung and Town 2007) and an increasing number of banana EST sequences from collections developed by EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária, Santos *et al.* 2005) and Syngenta (accessible under MTA, www.musagenomics.org/index.php?id=128) are annotated and become available increasing the probability of identifying the flanking sequences in the T-DNA tagged lines. Secondly, single T-DNA copy number lines are preferred in T-DNA tagging experiments because it facilitates the identification of the tagged sequence (paragraph 4.1). Independent data show that multiple T-DNAs occur in more than 50% (Khanna *et al.* 2004; Pérez Hernández *et al.* 2006a; Huang *et al.* 2007; this work, Figure 4.1) of transgenic banana lines. Hence, the majority of the tagged lines with an interesting LUC expression profile will carry more than one T-DNA copy. *In silico* analysis with bioinformatic tools (PlantCARE, PLACE, TSSP; paragraph 4.4.2) to identify candidate promoter elements may facilitate the selection of the banana promoter sequence responsible for the LUC expression in the promoter-tagged lines, but does not constitute an experimental proof. Obviously, another approach is needed to discriminate the activated from the non-activated T-DNA copy(ies).

The occurrence of transcriptional and translational fusions in which the T-DNA integrates downstream of a functional transcription start site in the 5’ untranslated region

and/or coding region is well documented in promoter and gene tagging experiments (André *et al.* 1986; Teeri *et al.* 1986; Kertbundit *et al.* 1991; Calderon-Villalobos *et al.* 2006). Thus, a truncated mRNA originating from the T-DNA RB flank should be fused to the *luc*$^+$ mRNA in banana lines showing LUC expression. Therefore, in promoter-tagged lines with multiple T-DNA copies the identification of the activated promoter(s) sequence(s) can be performed by an RT-PCR approach provided that the T-DNA integrated at a distance downstream of the functional transcription start site that allows the design of at least one T-DNA RB flank-specific primer. A similar approach was used for the detection of a tagged gene-reporter gene fusion transcript in T-DNA tagged rice (Ryu *et al.* 2004) and banana (Remy *et al.* 2005) lines.

In this chapter, the RT-PCR analysis of single (ET2-34) and multiple (ET2-17, ET2-85 and ET2-156) T-DNA copy lines to identify the RB T-DNA flanking sequence(s) transcriptionally fused to the *luc*$^+$ is discussed.

5.2 Identification of sequences with promoter activity by RT-PCR

An RT-PCR approach (Figure 5.1A) was followed for the analysis of transcriptional fusions between RB T-DNA flanking sequences (5'-tagged sequences) and the *luc*$^+$ transgene in four promoter-tagged lines (ET2-17, ET2-34, ET2-85 and ET2-156). Total RNA was isolated from different tissues of *in vitro* plant maintained at 26°C and first strand cDNA synthesis was performed using the gene-specific primer (GSP) Luc+R, which anneals at the positions +639 to +658 bp relative to the translation start codon of the *luc*$^+$ in the reverse orientation.

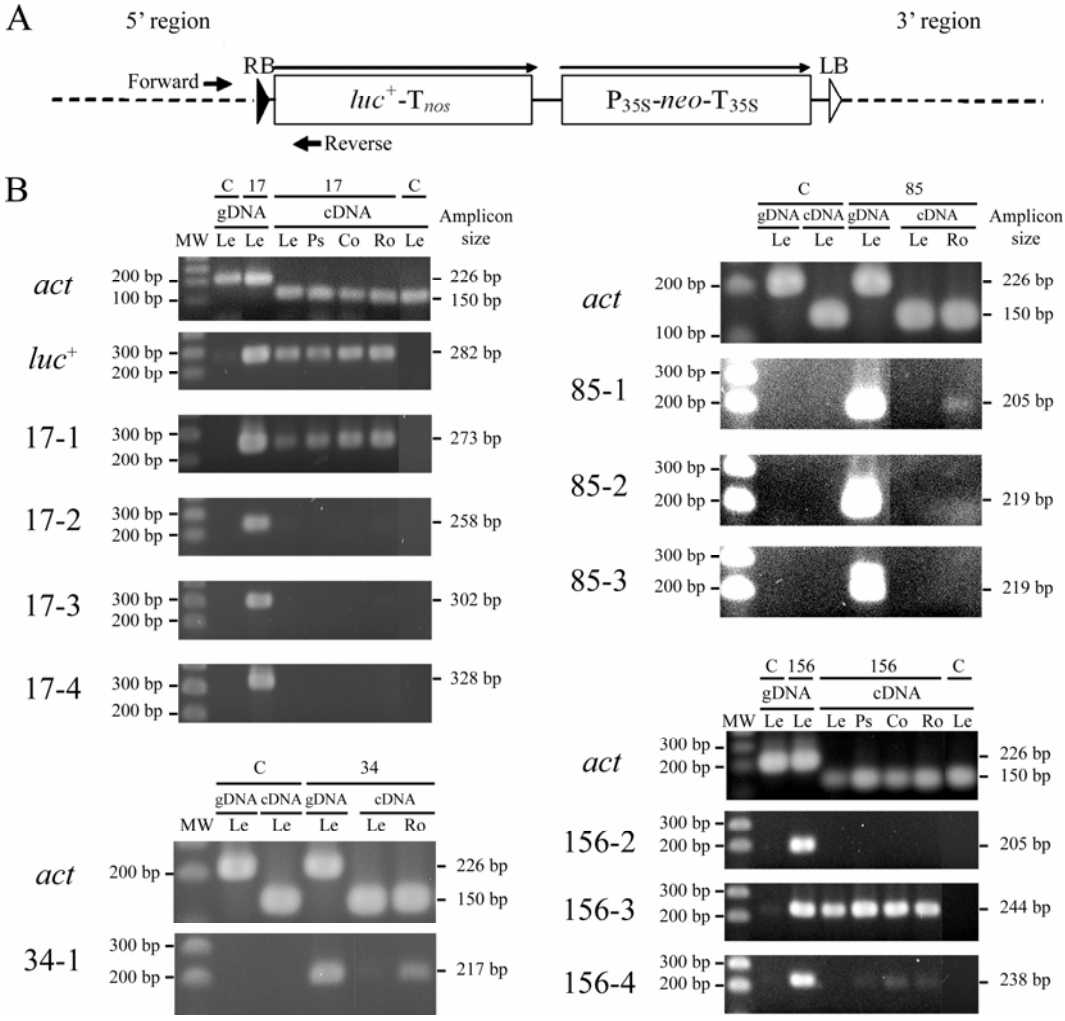

Figure 5.1. RT-PCR analysis of *in vitro* plant tissues at 26°C for transcriptional fusion between RB T-DNA flanking sequences (5'-tagged sequences) and *luc*$^+$ in the promoter-tagged lines ET2-17, ET2-34, ET2-85 and ET2-156. (A) Schematic representation of the annealing sites for the primers used in the RT-PCR. Dotted lines indicate T-DNA flanking sequences (5' and 3' region). The forward and reverse primers are illustrated as small arrows. The forward primers annealed within a distance of 70 to 1 bp from the RB, while the reverse primers anneal at the positions +200 (TAILRBluc1) and +149 (TAILRBluc2) relative to the start codon of the *luc*$^+$. Closed and open triangles indicate right (RB) and left (LB) T-DNA borders, respectively. Long arrows indicate direction of the transcription for the *luc*$^+$ and *neo* cassettes. The diagram is not precisely according to scale. (B) All RT-PCR reactions were stopped after 35 amplification cycles and the reaction of the codon-optimized luciferase gene (*luc*$^+$) is illustrated for line ET2-17 only while those for the housekeeping actin gene (*act*) is depicted for all the lines. RT-PCR was performed on cDNA synthesized using an oligo(dT)$_{18}$ primer for the *act* gene. Actin primers ActinF3/ActinR2 flank an intron allowing discrimination between a genomic DNA (gDNA) product (226 bp) and the reverse transcribed DNA (cDNA) product (150 bp). For all other RT-PCR reactions the cDNA template was produced by the *luc*$^+$ gene specific primer Luc+R. Control primers for the *luc*$^+$ transgene were LucL2/LucR5. Forward primers annealing at the different 5'-tagged sequences were used in combination with the reverse primer TAILRBLUC1 (line ET2-17) or TAILRBLUC2 (lines ET2-34, ET2-85 and ET2-156) which anneal in the *luc*$^+$ gene. Transcription of the 5'-tagged sequences 17-1, 17-2, 17-3 and 17-4 of line ET2-17 was verified employing the sequence-specific forward primer 17-RT-1, 17-RT-2, 17-RT-3 and 17-RT-4, respectively. For the 5'-tagged sequence 34-1 of line ET2-34 the forward primer employed was 34-RT-1, while for the lines ET2-85 and ET2-156 the corresponding forward primers were 85-RT-1, 85-RT-2, 85-RT-3, and 156-RT-2, 156-RT-3, 156-RT-4, respectively. C: non-transformed control plant. 17: tagged line ET2-17. 34: tagged line ET2-34. 85: tagged line ET2-85. 156: tagged line ET2-156. Le: leaf; Ps: pseudostem; Co: corm; Ro: root, all from *in vitro* plants. The expected amplicon sizes are indicated at the right of the images. MW: Smart Ladder SF (100 bp, Eurogentec, Seraing, Belgium).

The RT-PCR for the *luc*$^+$ transgene is illustrated only for line ET2-17, while for all four tagged lines the RT-PCR for the housekeeping actin gene (*act*) is shown (Figure 5.1B). Since the actin primers span an intron, the smaller size of the cDNA (150 bp) *versus* the gDNA (226 bp) PCR product demonstrated the absence of genomic DNA in the cDNA preparations in the four promoter-tagged lines analyzed. The expected *luc*$^+$ signal of 282 bp was obtained using either gDNA or cDNA as template demonstrating the presence and transcription of this transgene, respectively, in all tested *in vitro* tissues (leaf, pseudostem, corm and root) of line ET2-17. These results are also in agreement with the real-time measurements at 26°C and *in vitro* plant stage in this promoter-tagged line (Figure 3.4). Similar results for the *luc*$^+$ gene were observed for the other three promoter-tagged lines (data not shown).

RT-PCR reactions for candidate promoters were performed with a primer binding within a distance of 70 to 1 bp from the RB T-DNA junction of each of the different 5'-tagged sequences in combination with a *luc*$^+$-specific primer. In line ET2-17, a product of the expected size (273 bp) was only obtained for sequence 17-1 in all tissues tested (Figure 5.1B) demonstrating that this sequence is transcriptionally fused to the *luc*$^+$ gene. The difference in amplicon intensities between the different tissues with the strongest and weakest signal in root and leaf tissue, respectively, suggested a differential LUC expression in these tissues. However, RNA samples isolated from root and leaf showed in general a higher A_{260}/A_{230} absorbance ratio, and therefore a higher RNA quality, than that from corm and pseudostem tissue (1.47-1.72 and 0.96-1.37, respectively). RNA quality also varied between different RNA isolation events (data not shown) and may have an effect on cDNA synthesis.

Nevertheless, to gain more insight in the possible differences in *luc*$^+$ activation in the different tissues semi-quantitative PCR was performed for the four candidate promoter sequences of line ET2-17 in leaf and root tissue (Figure 5.2). Semi-quantitative RT-PCR analysis as presented in this work refers to non-competitive RT-PCR. Synthesis of cDNA was done using the GSP Luc+R or the oligo$(dT)_{18}$ primer that should bind to all mRNA molecules. To our surprise, the GSP Luc+R primer was able to anneal to other mRNA molecules than the *luc*$^+$ mRNA molecules since the *act* RT-PCR product was amplified in leaf and root tissue albeit at a very low level (Figure 5.2). Transcription of the *luc*$^+$ transgene in line ET2-17 clearly occurred in root tissue as detected with both cDNA preparations, but it was hardly detectable in leaf tissue (Figure 5.2). A strong RT-PCR signal was obtained before in the latter tissue with the GSP Luc+R cDNA preparation primer (Figure 5.1B). Different RNA isolations were performed for the samples shown in Figures 5.1B and 5.2, yet similar A_{260}/A_{230} ratios were obtained with leaf tissue (1.47 and 1.30, respectively) excluding RNA quality in this case as the source for the discrepancy in signal intensities in leaf tissue. GSP Luc+R-produced cDNA yielded the expected amplicon for sequence 17-1 already at 30 cycles in leaf and root tissue although the signal became

only clearly visible at 35 cycles. For the other candidate promoter sequences of line ET-17 35 cycles did not yield any product irrespective of the tissue and primer used for the cDNA synthesis (Figure 5.2).

The *luc*$^+$ activation by the flanking sequence 17-1 in leaf and root tissue of *in vitro* plants cultured at 26°C was detected with GSP produced cDNA at 35 RT-PCR cycles in two independent cDNA synthesis events (Figures 5.1B and 5.2). In addition, the results obtained for sequence 17-1 at 26°C were confirmed in freshly isolated RNA samples from leaf and root tissues of *in vitro* plants maintained at 26°C and 8°C (Figure 5.3) demonstrating not only that it is transcriptionally fused to the *luc*$^+$ transgene but also its activity at LT. Expression at 8°C appeared comparable to that at 26°C in root tissue, but the intensity of the amplicons obtained at 8°C seemed to be slightly stronger than those at 26°C in leaf tissue (Figure 5.3). For sequence 17-2 transcription was detected in leaf tissue only when sampling at 8°C and at 40 RT-PCR cycles, whereas in root tissue the expected RT-PCR signal was already detectable at 35 cycles at 26°C and 8°C which contradicts the results obtained in Figure 5.2 for samples at 26°C. However, the root RNA quality was much higher for the samples shown in Figure 5.3 than in Figure 5.2 (A_{260}/A_{230} ratio of 2.08-2.15 and 1.12, respectively) and might result in differences in efficiency of cDNA synthesis. Taken together, the RT-PCR data suggest that besides sequence 17-1 also sequence 17-2 weakly activates the *luc*$^+$ gene.

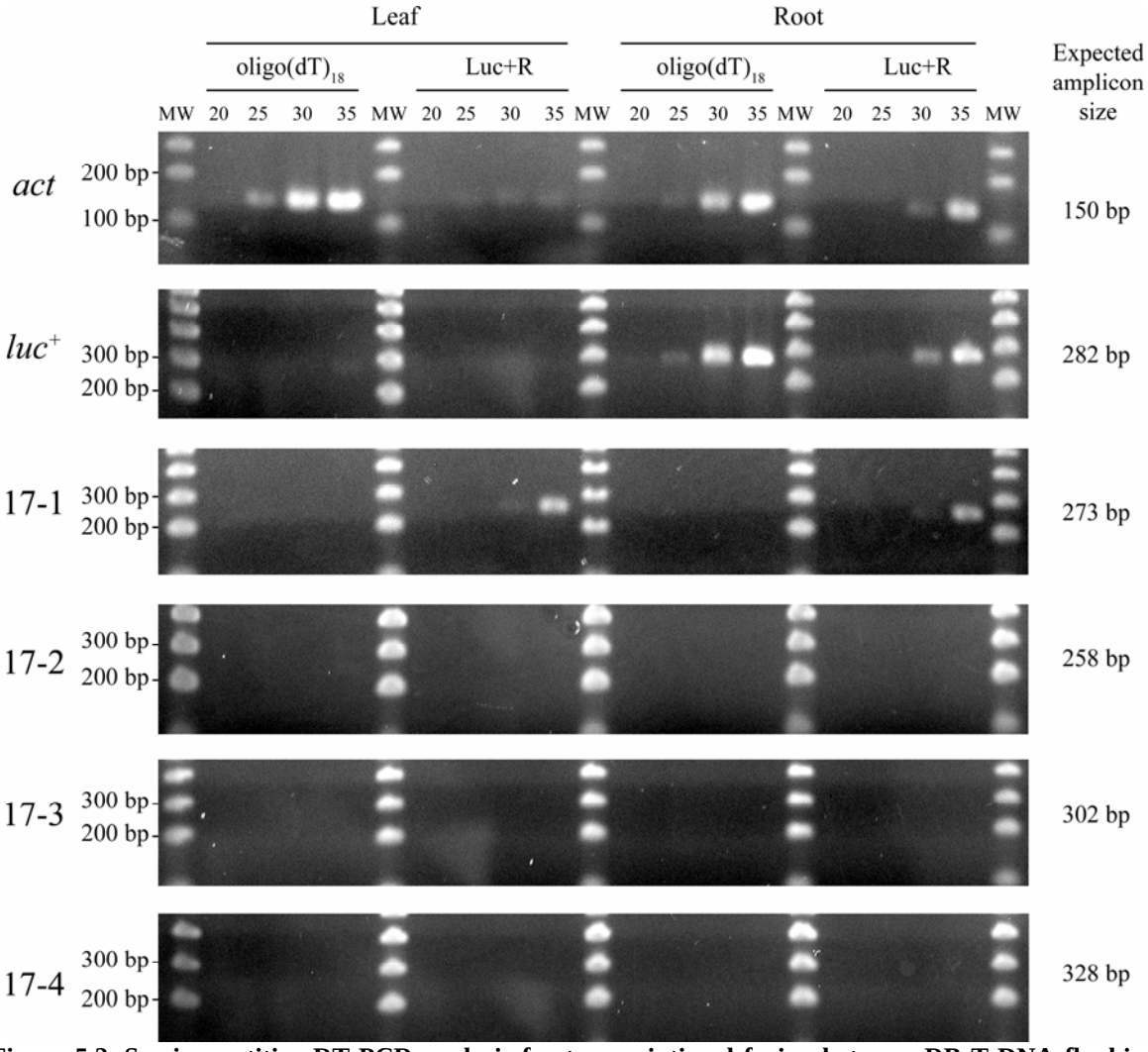

Figure 5.2. Semi-quantitive RT-PCR analysis for transcriptional fusion between RB T-DNA flanking sequences (5'-tagged sequences) and *luc*$^+$ in the promoter-tagged line ET2-17 in leaf and root tissues from *in vitro* plants maintained at 26°C. Oligo(dT)$_{18}$ and Luc+R indicate cDNA synthesized using the oligo(dt)$_{18}$ or the *luc*$^+$ gene specific Luc+R primer, respectively. Fresh RNA was isolated independently from the samples shown in Figure 5.1B. The numbers 20, 25, 30 and 35 indicate the number of amplification cycles. The RT-PCR amplicons were produced as described in Figure 5.1 and the expected amplicon sizes are indicated on the right of the images. MW: Smart Ladder SF (100 bp, Eurogentec, Seraing, Belgium).

For line ET2-34, an amplicon of the expected size (217 bp) in the root was obtained by RT-PCR confirming that the 5'-tagged sequence 34-1 can be associated with the transcription of the activated *luc*$^+$ gene (Figure 5.1B).

An amplicon of the expected size was weakly detected in the root but not in the leaf tissue of line ET2-85 for the flanking sequence 85-1 (Figure 5.1B) indicating the transcriptional fusion between this sequence and the *luc*$^+$ gene at least in the root tissue. The RNA quality of leaf and root samples was comparable (A_{260}/A_{230} of 1.17 and 1.35, respectively) though low and might prevented efficient cDNA synthesis. The remaining two sequences of line ET2-85 showed no detectable amplicon of the expected size in either tissue (Figure 5.1B).

In none of the four tested tissues of line ET2-156 an RT-PCR product was produced by the flanking sequence 156-2 (Figure 5.1B). By contrast, the flanking sequence 156-3 was clearly transcriptionally fused to the *luc*$^+$ as demonstrated by the presence of a high intensity amplicon of the expected size in all tissues (Figure 5.1B). Very weak *luc*$^+$ activation by the flanking sequence 156-4 was also detected in all tested tissues. RT-PCR analysis was not performed for sequence 156-1 because it comprised a T-DNA tandem repeat (Annex).

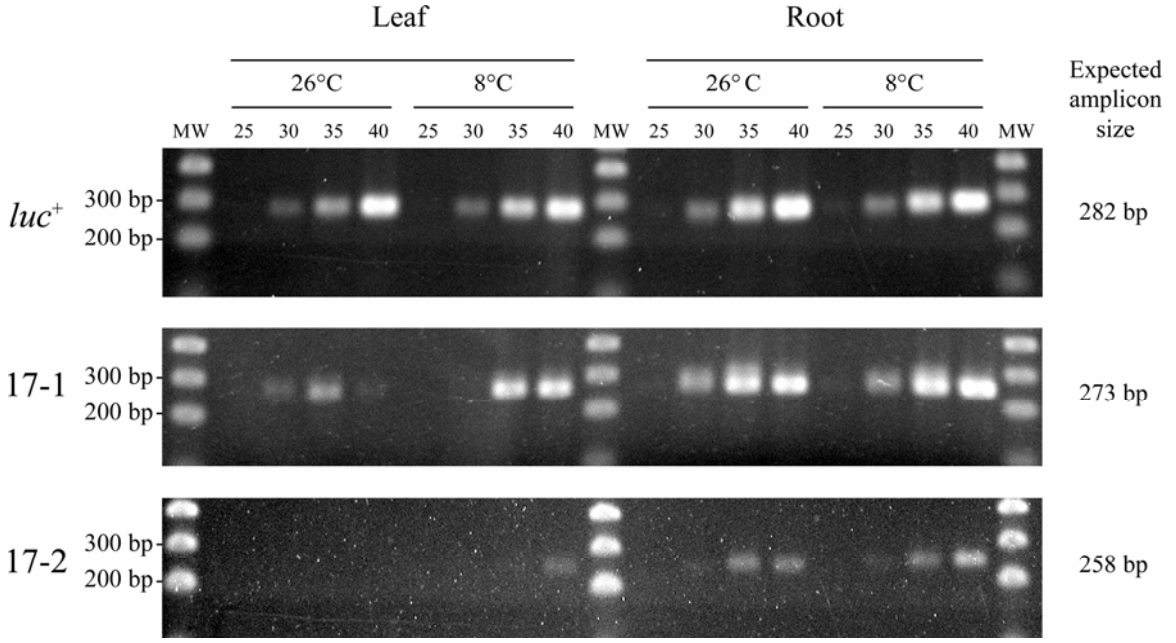

Figure 5.3. Semi-quantitive RT-PCR analysis for transcriptional fusion between RB T-DNA flanking sequences (5'-tagged sequences) and *luc*$^+$ in leaf and root tissues of *in vitro* plants of the promoter-tagged line ET2-17 maintained at different temperatures. Plant tissues were frozen for RNA isolation after 6 h at 26°C and 16 h at 8°C for the 26°C and 8°C treatments, respectively. Synthesis of cDNA was performed using the *luc*$^+$ gene specific primer Luc+R. The numbers 25, 30, 35 and 40 indicate the number of amplification cycles. The RT-PCR reaction for the *luc*$^+$ transgene was performed with the primer pair LucL2/LucR5. The RT-PCR amplicons were produced as described in Figure 5.1. The expected amplicon sizes are indicated on the right of the images. MW: Smart Ladder SF (100 bp, Eurogentec, Seraing, Belgium).

5.3 Discussion

Under the assumption that the in-frame stop codon 18 bp upstream of the *luc*$^+$ gene in pETKUL2 (Figure 4.2) is not deleted during T-DNA integration it should reduce the likelihood of translational fusions. Hence, taking advantage of the transcriptional fusion between the 5'-tagged sequences with the *luc*$^+$ in promoter-tagged banana lines, the identification of the activated T-DNA copy in multiple T-DNA copy lines was performed *via* RT-PCR.

Two 5'-tagged sequences, 17-1 and 17-2, are transcriptionally fused to the *luc*$^+$ in line ET2-17, albeit to a different level with 17-1 being the most transcriptionally active sequence. Transcriptional activity was expected for sequence 17-2 due to the integration of

the T-DNA in a transcribed region (Figure 4.14A). Whereas the transcriptional activity of the 5'-tagged sequence 17-1 which is neither homologous to any known coding region nor the corresponding 3'-tagged region (Figure 4.14A), the hypothesis is strengthened that a cryptic promoter has been tagged in line ET2-17. Verification of the absence of a transcript at the 17-1 integration site in a non-transformed plant should yield additional proof that 17-1 harbours a cryptic promoter.

Besides the starting amount of RNA which was kept constant throughout all our analyses at 50 ng per reaction, non-competitive semi-quantitative RT-PCR analysis is influenced by several factors including cDNA synthesis efficiency, which in turn, is affected by RNA quality. The latter seems not only linked to the type of *in vitro* tissue with in general higher RNA quality from root and leaf tissue than from corm and pseudostem, but also to the quality of the tissue. For instance, a specific amplicon for the *luc*$^+$ gene was detected in leaf tissue of line ET2-17, in which the GSP primer was used for cDNA synthesis (Figure 5.1, RNA A_{260}/A_{230} ratio of 1.47). However, an RT-PCR performed on GSP produced cDNA from an independent RNA isolation of leaf tissue (RNA A_{260}/A_{230} ratio of 1.3) revealed a very weak signal at the same number of 35 PCR cycles (Figure 5.2). Most likely, the level of tissue blackening, which is an indication of the amount of polyphenolic compounds, is directly correlated with RNA quality. As these compounds are known to cause difficulties in nucleic acid purification in many plant species and banana is known to contain high amounts of their precursors as well as of polyphenol oxidase (Wuyts *et al.* 2006), it is not surprising that independent RNA isolations from the same type of tissue can yield very different RNA qualities. Thus, care in the interpretation of the semi-quantitative RT-PCR analysis for the LUC expression in the different tissues tested is warranted. For instance, the differences in amplicon intensities between different tissues of *in vitro* plants (Figure 5.1) suggest a tissue specific level of LUC expression in line ET2-17, but in view of the aforementioned remarks we only conclude that 17-1 is transcriptionally active in all tested *in vitro* tissues and hence, has constitutive activity at the plant level. Furthermore, the presence of the specific amplicon for the tagged sequence 17-1 using the GSP but its absence using the oligo$(dT)_{18}$ primer for cDNA synthesis may be due to the low efficiency of full-length cDNA synthesis with the latter primer which will abolish the annealing site of the 5'-tagged specific primer as suggested by Zhang and Byrne (1999).

Although cDNA synthesis might affect semi-quantitative RT-PCR results, a general comparison within the same tissue indicated that the LUC expression directed by the two transcriptionally fused sequences 17-1 and 17-2 in line ET2-17 is not lower at 8°C than at 26°C at *in vi*tro plant stage (Figure 5.3). Thus, the decrease in LUC activity observed at the *in vitro* plant stage at LT (Figure 3.4) may reflect a post-transcriptional regulation and not the activity of the promoter.

RT-PCR analyses revealed transcriptional fusions between the 5'-tagged sequences and the *luc*$^+$ in the tagged lines ET2-34 and ET2-85 (Figure 5.1B). Since the T-DNA

integrated in a repetitive region (Figure 4.14B), most likely a cryptic promoter is tagged in line ET2-34. The transcriptional fusion of the 5'-tagged sequence 85-1 of line ET2-85 with *luc*$^+$ confirmed that this T-DNA copy was inserted between a promoter and a coding region as observed in the *in silico* analysis (Figure 4.14C). The lack of the expected RT-PCR amplicons suggested that 5'-tagged sequences 85-2 and 85-3 were not transcriptionally fused to *luc*$^+$. RNA quality of leaf and root tissues expressed as A_{260}/A_{230} value was below the optimal value of 1.8 or higher in the RNA preparations of lines ET2-34 and ET2-85 which suggests suboptimal cDNA synthesis for both tissue samples. Therefore, a tissue specific expression could not be reliable determined by RT-PCR.

In the tagged line ET2-156 expression of the *luc*$^+$ transgene is strongly driven by the truncated 35S promoter and *neo* gene present in sequence 156-3 (Table 4.4, Annex) as demonstrated by the intense RT-PCR signal (Figure 5.1B). No transcript was detected for sequence 156-2, whereas a very weak signal occurred for sequence 156-4. In conclusion, no native banana promoter was tagged in line ET2-156. From these data it is also clear that truncated promoters originating from the selectable marker gene cassette which integrate upstream of the promoterless reporter gene can lead to false positive LUC activation. PCR screening for such rearrangements using a forward primer for the promoter of the selectable marker gene in combination with a reverse primer for the *luc*$^+$ can identify such lines prior to the sequencing of the 5'-tagged regions (Table 4.4, Figure 4.11).

In summary, transcriptional fusion between 5'-tagged sequences and the *luc*$^+$ gene was detected and confirmed banana promoter activity in tagged lines ET2-17, ET2-34 and ET2-85. After cDNA synthesis a forward primer annealing to the 5'-tagged sequence near the RB T-DNA sequence was employed in combination with a reverse primer which anneals to the *luc*$^+$. The RT-PCR approach described in this chapter facilitated the identification of the 5'-tagged sequence responsible for the LUC expression observed in the promoter-tagged lines, especially in multicopy T-DNA lines. However, a final validation of the 5'-tagged sequence is required to confirm the promoter activity by cloning the candidate promoter sequences upstream of a reporter gene and then integrate the different chimeric cassettes into the banana genome. Therefore, in the following chapter further analysis of promoter activity was performed by cloning the 5'-tagged sequences of the promoter tagged lines ET2-17, ET2-34 and ET2-85 upstream of the *uid*A reporter gene and by back-transforming into banana ECS.

Chapter 6 Back-transformation of candidate promoter sequences

6.1 Introduction

T-DNA tagging allows the isolation and identification of promoters and genes without information on the genome. Back-transformation of the tagged promoter fused to a reporter gene in non-transformed control plants has yielded for many promoters an activity pattern similar to that of the original tagged line (Kertbundit *et al.* 1991; Ökrész *et al.* 1998; Foster *et al.* 1999; Custers *et al.* 2002; Sivanandan *et al.* 2005 and others). However, other research groups found differences in reporter gene expression patterns between the original T-DNA tagged line and back-transformed lines (Mollier *et al.* 2000; Farrar *et al.* 2003; Bade *et al.* 2003; Stangeland *et al.* 2005). Besides to proof that a 5'-tagged sequence has promoter activity, back-transformation is clearly also useful to study the activity pattern of the cloned 5'-tagged sequence and compare it to that of the tagged version (in the original tagged line).

Most T-DNA tagging and the corresponding back-transformation experiments rely on the use of the *β*–glucuronidase or GUS encoding *uid*A reporter gene (Table 1.2). The GUS enzyme is stable, its activity can be quantitatively assayed in a reliable manner by fluorometric measurement and precise localization of GUS activity is possible *via* histochemical staining (Jefferson *et al.* 1987). The latter does not require expensive equipment, but GUS analyses can nevertheless become expensive when analyzing larger and/or many plants because of the substrate costs (Springer 2000). In contrast to the detection of LUC expression, GUS expression can only be studied *in planta* at specific time points and not in real-time because of the lethality of the assays. On the other hand, a wide variety of promoter cloning vectors exists that contain the *uid*A reporter gene (e.g. pCAMBIA vectors, www.cambia.org). For a first screening for promoter activity of the isolated candidate promoter sequences from the promoter tagged lines ET2-17, ET2-34 and ET2-85 it was therefore decided to back-transform the different 5'-tagged sequences into banana by transcriptionally fusing them to the *uid*A reporter gene.

Hence, in this chapter, the promoter activity of transcriptionally active 5'-tagged sequences of promoter-tagged lines ET2-17, ET2-34 and ET2-85 was analyzed in back-transformed banana lines by qualitative and quantitative GUS assays.

6.2 Construction of vectors

To verify the promoter activity of candidate promoter sequences identified and isolated by T-DNA tagging, 5'-tagged sequences showing transcriptional fusion with the *luc*$^+$ gene in lines ET2-17, ET2-34 and ET2-85 (Chapter 5) were cloned into the multiple cloning site (MCS) of the pCAMBIA 1391Z vector (www.cambia.org). The 5'-tagged sequences 17-1, 17-2, 17-3, 34-1, and 85-1 were PCR amplified from the original promoter-tagged lines using specific primers (Figure 2.5), and the distances of these primers relative to the RB T-DNA junction are given in Table 6.1.

PCR was performed using the Extensor Hi-Fidelity PCR Enzyme Mix (ABgene, Epsom, UK) and primers employed were designed either with or without different restriction sites at their 5' end (Table 6.1). The restriction sites were designed for unidirectional cloning of the candidate promoter sequence and/or as an alternative to *Eco*RI cloning because the recognition site of this enzyme was present in the candidate promoter sequence. Despite the presence of non-specific products, amplicons of the expected size were obtained for each candidate promoter sequence (Figure 6.1). For the tagged sequence 34-1, the PCR product revealed the presence of more than one amplicon (Figure 6.1) due to the annealing of the forward primer in the repetitive region of the 5'-tagged sequence (Figure 4.18A).

Following PCR amplification the candidate promoter sequences were cloned in the pCR®4-TOPO® vector and introduced in chemically competent *E. coli* cells (Sambrook *et al.* 1989) for vector multiplication. Bacterial colonies containing the cloned amplicon as verified by colony PCR (data not shown) were cultured for plasmid purification.

Table 6.1. Cloning of tagged candidate promoter sequences in pCAMBIA 1391Z

		Primers used for cloning[a]							Sequence comparison[b]		
		Forward		Reverse							
Line-promoter sequence[c]	**Resulting vector**	**ID**	**5' Modif.**	**ID**	**5' Modif.**	**Expected promoter size (bp)**	**Obtained promoter size (bp)**	**Distance from RB[d]**	**No. of mis-matches**	**No. of deleted bp**	**No. of additional bp**
17-1	pESKUL1	17-1F	None	17-1R1	None	1738[e]	1742	-1742 to -5	4	0	4
17-1	pESKUL2	17-1FSacI	*Sac*I	17-1R4AscI	*Asc*I	1350	1352	-1742 to -393	4	0	2
17-2	pESKUL3	17-2FHindIII	*Hind*III	17-2R1BamHI	*Bam*HI	1198	1198	-1210 to -13	3	0	0
17-3	pESKUL4	17-3FHindIII	*Hind*III	17-3R1BamHI	*Bam*HI	681	681	-686 to -6	3	0	0
34-1	pESKUL5	34-1F	None	34-1R	None	828	828	-856 to -29[f]	4	0	0
34-1	pESKUL6	34-1F	None	34-1R	None	1216	1318[g]	-1244 to -29[f]	37	8	1 and 109[g]
85-1	pESKUL7	85-1F2PstI	*Pst*I	85-1R2	None	599	599	-604 to -6	4	0	0

[a]Candidate promoter sequences were PCR amplified from the original tagged lines using primers with or without a 5' modification consisting of a restriction enzyme site.
[b]Sequence comparison between the 5'-tagged sequences amplified by the Extensor Hi-Fidelity PCR Enzyme Mix (Abgene) (obtained promoter) and by the *Taq* DNA polymerase (New England Biolabs) in the original TAIL-PCR or I-PCR (expected promoter). The number of mismatches and deleted or additional nucleotides are shown.
[c] pETKUL2 tagged line number-candidate promoter sequence number.
[d]Distances from the RB junction of the amplified candidate promoter sequences, with negative (-) positions upstream of the insertion site in the original ET2-line.
[e]5'-tagged sequence isolated by I-PCR employing the Expand High Fidelity PCR System from Roche.
[f]The primer 34-1F anneals twice due to the presence of a direct repeat.
[g]An addition of a 109 bp fragment at 125 bp upstream of the 3' end.
bp refers to base pair.

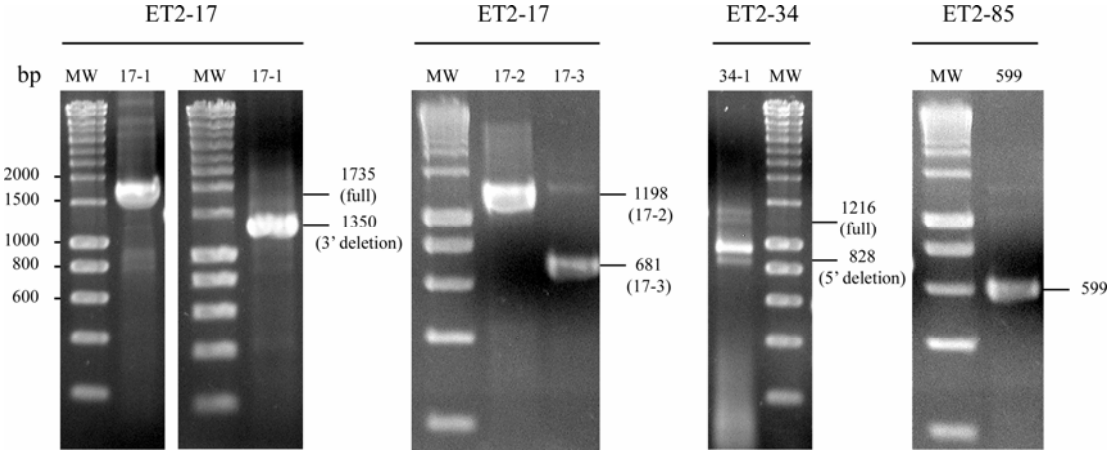

Figure 6.1. PCR amplification of candidate promoter sequences from tagged lines ET2-17, ET2-34 and ET2-85 using a polymerase mix with proofreading activity. Following addition of genomic DNA (2-4 ng μL^{-1}), promoter specific primers (200 µM each; Table 6.1) and the Extensor Hi-Fidelity PCR Enzyme Mix (0.03 U μL^{-1}; ABgene, Epsom, UK) to the master mix the candidate promoter sequences were amplified as follows: 94.0°C for 2 min, then 35 cycles of denaturation at 94.0°C for 30 s, annealing of primers at 60.0°C for 30 s, elongation at 68.0°C for 2 min, and ending with a final elongation at 68.0°C for 10 min. Above the figures the pETKUL2-tagged line number is given and above the lanes the candidate tagged promoter sequence is indicated (tagged line-corresponding 5'-tagged sequence). The expected size(s) (in bp) are indicated on the right of each image. MW: Smart Ladder (1 kb, Eurogentec, Seraing, Belgium). bp: base pair.

Following plasmid purification, excision of the candidate promoter sequence was performed using the *Eco*RI enzyme unless the primers were modified with other restriction sites (Table 6.1), the restricted fragment was then isolated and cloned in the *Eco*RI site (or the other corresponding restriction sites) linearized pCAMBIA 1391Z vector. The resulting pESKUL vectors (Table 6.1) were transferred to chemically competent *E. coli* cells by heat shock. The correct orientation of the candidate promoters as identified in the original tagged lines was verified by colony PCR employing a forward and a reverse primer annealing in the candidate promoter sequence and the *uid*A^{INT} reporter gene, respectively (data not shown).

Bacterial colonies harbouring the cloned promoter in the correct orientation as shown by the presence of an amplicon were cultured and the pESKUL vectors (Table 6.1) were isolated. The cloned candidate promoters were sequenced (obtained promoter sequence in Table 6.1) and compared to the original 5'-tagged sequences isolated by TAIL- or I-PCR from the tagged lines (expected promoter sequence in Table 6.1). Not taking the obtained 1318 bp size candidate promoter 34-1 into account, this comparison revealed three to four mismatches and up to four additional nucleotides in the cloned candidate promoters (Table 6.1). More specifically, the additional nucleotides obtained when amplifying with a high fidelity enzyme mix lead to minor differences in promoter length between the expected and the cloned promoter in the pCAMBIA 1391Z vector for both variants of sequence 17-1 (Table 6.1). The obtained sequence amplified with a HF polymerase mix was submitted to the GenBank under the accession number EU161097. Sequence analysis

of the amplicon obtained for the full-length 5'-tagged sequence 34-1 (resulting vector pESKUL6) revealed the presence of 37 mismatches, the deletion of 8 bp in total, an additional single bp and an additional fragment of 109 bp starting at 125 bp from the RB T-DNA junction site (Table 6.1). In addition, the expected amplicon of 828 bp with four mismatches but without any additional bp was obtained corresponding to the short version of the 5'-tagged sequence 34-1 (resulting vector pESKUL5, Table 6.1).

The pESKUL vectors were introduced into *A. tumefaciens* strain EHA105 by electroporation. Additionally, restriction digestion was performed on the different pESKUL vectors isolated from *E. coli* and *A. tumefaciens* confirming the correct orientation of the candidate promoter sequences upstream of the *uid*A^{INT} reporter gene (data not shown).

6.3 Back-transformation

6.3.1 Baseline promoter activity

Embryogenic cell suspension (ECS) from the dessert banana 'Grand naine' (GN) and the plantain 'Three Hand Planty' (THP) were cocultivated with *A. tumefaciens* harboring the pESKUL vectors. After six days of cocultivation, the embryogenic cells were histochemically stained for detection of transient GUS activity. All the pESKUL vectors shown in Table 6.1 were used to transform GN, while THP cells were only transformed with the pESKUL1 and pESKUL5 vectors. As negative control to determine the background, ECS were transformed with the pCAMBIA 1391Z cloning vector lacking a promoter to drive the *uid*A^{INT} gene. ECS were also transformed with the positive control vectors pLSKUL23 and pLSKUL25, in which the enhanced 35S CaMV promoter and maize *Ubi*1 promoter, respectively, drive the *uid*A^{INT} reporter gene. GN cells transformed with the pESKUL1 and pESKUL2 revealed up to 10 and 8 blue foci per sample of approximately 50 mg of embryogenic cells, respectively. In the pESKUL3, pESKUL5 and pESKUL6 transformed samples up to 3, 2 and 2 blue foci were detected, respectively, whereas the pESKUL7 transformed cells showed up to 35 blue foci per sample. No detectable background staining was observed with pCAMBIA 1391Z embryogenic cells. As expected for the positive control vectors pLSKUL23 and pLSKUL25, around 1000 blue foci were obtained per GN sample. On the other hand, only the THP embryogenic cells transformed with the pLSKUL25 vector showed detectable blue foci, but at a lower number than the GN pLSKUL25 transformed cells (around 500 blue foci per sample) and GUS activity was absent in pESKUL1 and pESKUL5 transformed THP cells.

After two and three months of selection for transgenic THP and GN lines, respectively, histochemical GUS staining of antibiotic-resistant cell colonies maintained at 26°C was performed (Figure 6.2). THP pESKUL1 transformed cell colonies showed clear GUS activity but at a lower intensity and extent than the positive control cell colonies that

contain the maize *Ubi*1 promoter (pLSKUL25). Thus, the promoter activity of the THP 17-1 sequence was confirmed in THP following back-transformation. On the other hand, the pCAMBIA 1391Z and the pESKUL5 transformed colonies showed no detectable GUS activity. In these samples a blue-brown precipitate or color was observed in small cell clusters surrounding the cell colonies, which is most likely due to the presence of phenolic compounds in the dead or dying cells and does not reflect GUS activity. In GN, only the cell colonies harboring the pESKUL1, pESKUL2 and pESKUL7 expressed the *uid*A^{INT} transgene demonstrating that the THP tagged sequences 17-1 and 85-1 possess promoter activity in GN. Furthermore, GUS activity was detected at a higher intensity in the pESKUL1 than in the pESKUL2 lines. However, neither version of the 5'-tagged sequence 34-1, nor the full-length (pESKUL6) nor the 5' deletion variant (pESKUL5), was able to drive GUS expression (Figure 6.2).

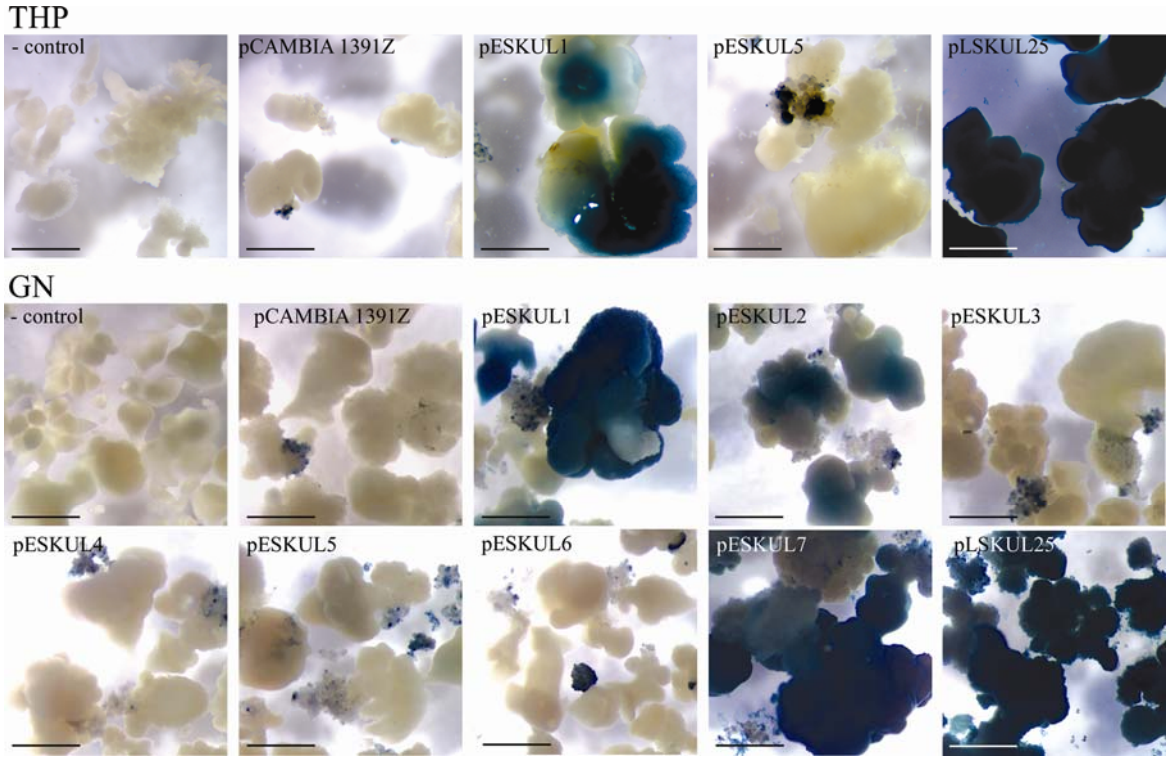

Figure 6.2. Histochemical GUS staining of transgenic cell colonies of the plantain 'Three Hand Planty' (THP) and the dessert banana 'Grand naine' (GN) maintained at 26°C two and three months after back-transformation, respectively, with the candidate promoter sequences fused to the *uid*A^{INT} gene. Non-transformed cells and cells transformed with the pCAMBIA 1391Z cloning vector are indicated as '-control' and 'pCAMBIA 1391Z', respectively. Only the pESKUL1 and pESKUL5 vectors were used to transform THP embryogenic suspension cells, while all pESKUL vectors were employed for transformation of GN embryogenic suspension cells. The maize ubiquitin promoter drives the *uid*A^{INT} reporter gene in vector pLSKUL25. Bars indicate 1 mm.

Complementary to the histochemical analysis, quantitative fluorometric GUS assays were performed on extracts of pooled transgenic cell colonies maintained on ZZ medium at 26°C three to four months after back-transformation (Figure 6.3). As observed in the histochemical analysis (Figure 6.2), GUS enzymatic activity above background was clearly detected in the pESKUL1 THP cell colonies (Figure 6.3A). The positive control lines (pLSKUL23 and pLSKUL25) revealed a 20 (pLSKUL23)- and 21 (pLSKUL25)-fold higher GUS activity than the pESKUL1 lines (518 ± 40 and 539 ± 60 *vs.* 26 ± 6 pmol MU h^{-1} μg^{-1} protein, respectively; Figure 6.3A). GUS enzymatic activity was not detected in non-transformed pCAMBIA 1391Z and pESKUL5 transformed cell colonies (Figure 6.3A). The GUS enzymatic activity measured in the pESKUL1 GN cell colonies (105 ± 36 pmol MU h^{-1} μg^{-1} protein; Figure 6.3A) is in agreement with the histochemical GUS staining (Figure 6.2), but the weaker histochemical GUS activity of the pESKUL2 cell colonies (Figure 6.2) was surprisingly, not detected in the fluorometric analysis (Figure 6.3B).

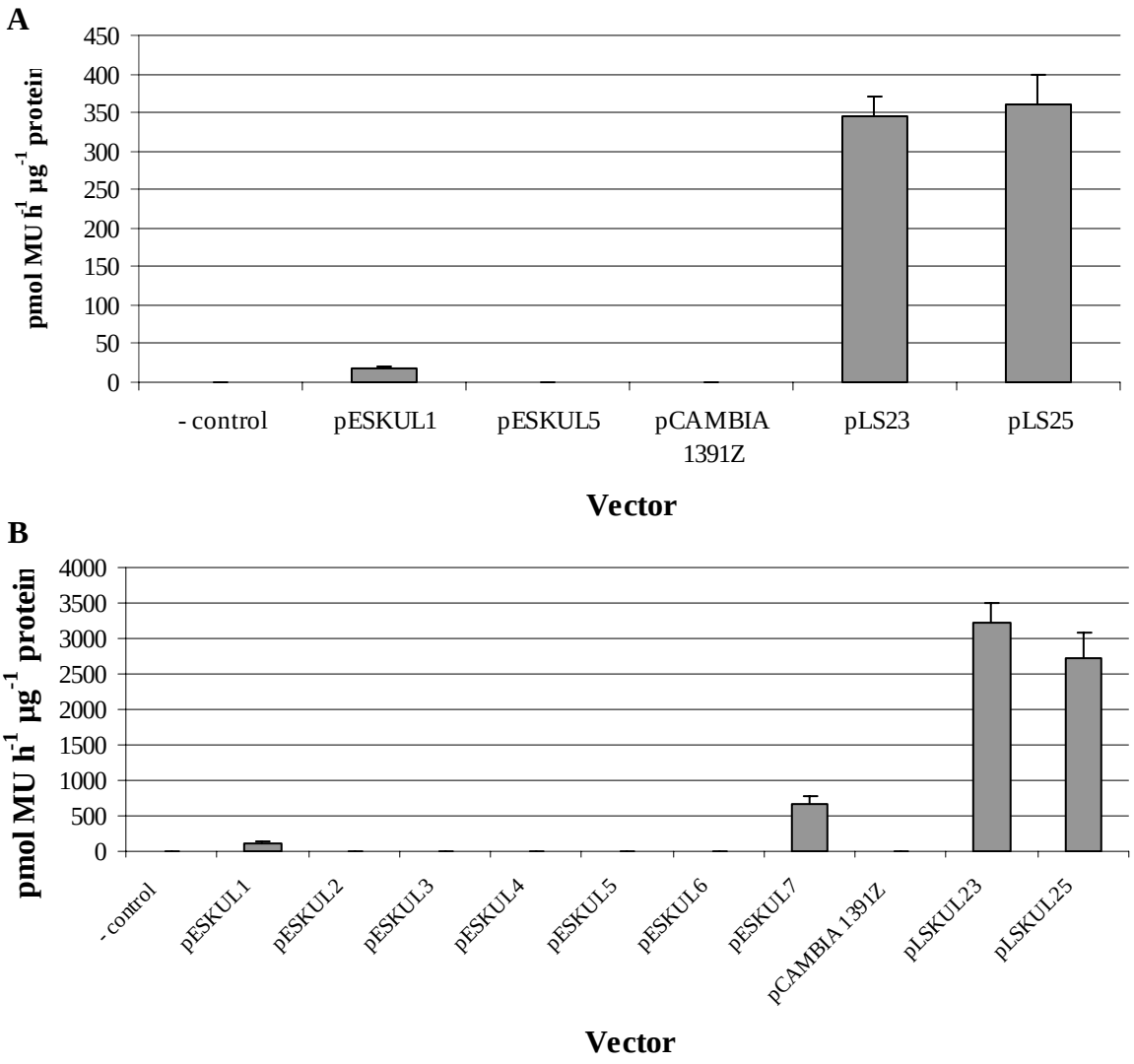

Figure 6.3. Enzymatic GUS activity in cell colonies maintained on ZZ medium and at 26°C three to four months after *Agrobacterium* back-transformation with candidate promoter sequences fused to the *uid*A^{INT} gene. (A) Approximately 300 mg fresh weight of cell colonies of the plantain 'Three Hand Planty' (THP) were screened for GUS activity after three months of back-transformation with the vectors pESKUL1 and pESKUL5. (B) Approximately 50 mg fresh weight of cell colonies of the dessert banana 'Grand naine' (GN) were screened for GUS activity after four months of back-transformation with the vectors pESKUL1, pESKUL2, pESKUL3, pESKUL4, pESKUL5, pESKUL6 and pESKUL7. In addition, the fluorometric GUS assay was performed on THP and GN cell colonies that were transformed with the control vectors pCAMBIA 1391Z, pLSKUL23 (enhanced 35S CaMV promoter) and pLSKUL25 (maize ubiquitin promoter). Non-transformed cells are indicated as '- control'. GUS activity is expressed as pmol MU h^{-1} μg^{-1} of total protein. Each entry is the average of three independent measurements and the standard deviations are indicated.

Noteworthy is the fact that the THP full-length 17-1 promoter was four times more active in GN than in THP back-transformed cell colonies (105 ± 36 and 26 ± 6 pmol MU h^{-1} μg^{-1} protein, respectively), although the enhanced 35S and maize *Ubi*1 promoters were also more active in GN than in THP cell colonies [6- (3211 ± 280 *vs.* 518 ± 40 pmol MU h^{-1} μg^{-1} protein) and 5-fold (2720 ± 369 *vs.* 539 ± 60 pmol MU h^{-1} μg^{-1} protein), respectively]. The GN cells harboring the pCAMBIA 1391Z, pESKUL3, pESKUL4, pESKUL5 and pESKUL6 were not showing GUS enzymatic activity, similar to protein extracts of the non-

transformed cells (Figure 6.3B). Similar to the data obtained in the THP transgenic cell colonies, in GN the positive control pLSKUL25 colonies conferred a 26-fold higher GUS enzymatic activity than the pESKUL1 colonies (2720 ± 369 *vs.* 105 ± 36 pmol h^{-1} μg^{-1} protein, respectively; Figure 6.3B). However, this difference was reduced to 4-fold in the pESKUL7 colonies (2720 ± 369 *vs.* 656 ± 126 pmol MU h^{-1} μg^{-1} protein, Figure 6.3B). In conclusion, the THP 85-1 promoter has a 6-fold higher activity in transgenic GN cell colonies than the THP full-length 17-1 promoter.

6.3.2 Promoter activity at low temperature

The back-transformed cell cultures were LT treated using the temperature controlled plate (paragraph 2.2.4.2) to compare the activity of the candidate promoter sequences to that in the original tagged lines under identical temperature conditions. Two (THP) and six (GN) months after back-transformation, transgenic cell colonies maintained on ZZ medium were histochemically stained for GUS activity after 17 h at 26°C and after a temperature regime of 6 h at 26°C followed by 10 h at 8°C (26°C and 8°C treatments, respectively; Figure 6.4). Histochemical GUS activity was comparable between the 26°C and the 8°C treatments in both cultivars for cell colonies transformed with the pESKUL1 vector. Similar results were obtained in transgenic cell colonies transformed with the positive control vectors (pLSKUL23 and pLSKUL25). On the other hand, GN pESKUL7 transformed colonies showed an apparent increase in GUS activity by the LT treatment of 8°C (Figure 6.4). In contrast to previous assays at 26°C (Figure 6.2), the GUS intensity of pESKUL1 transformed cell colonies was lower in GN than in THP (Figure 6.4). GUS intensity of pLSKUL25 cell colonies showed a similar pattern with lower activity detected in GN than in THP. The lack of GUS activity in pESKUL5 lines of both cultivars at 26°C was confirmed (Figures 6.2 and 6.4) and was also absent at 8°C (Figure 6.4). This observation holds true for the cloning vector pCAMBIA 1391Z demonstrating the lack of 35S enhancer induced GUS activity. The negative control colonies were also free of background staining at LT.

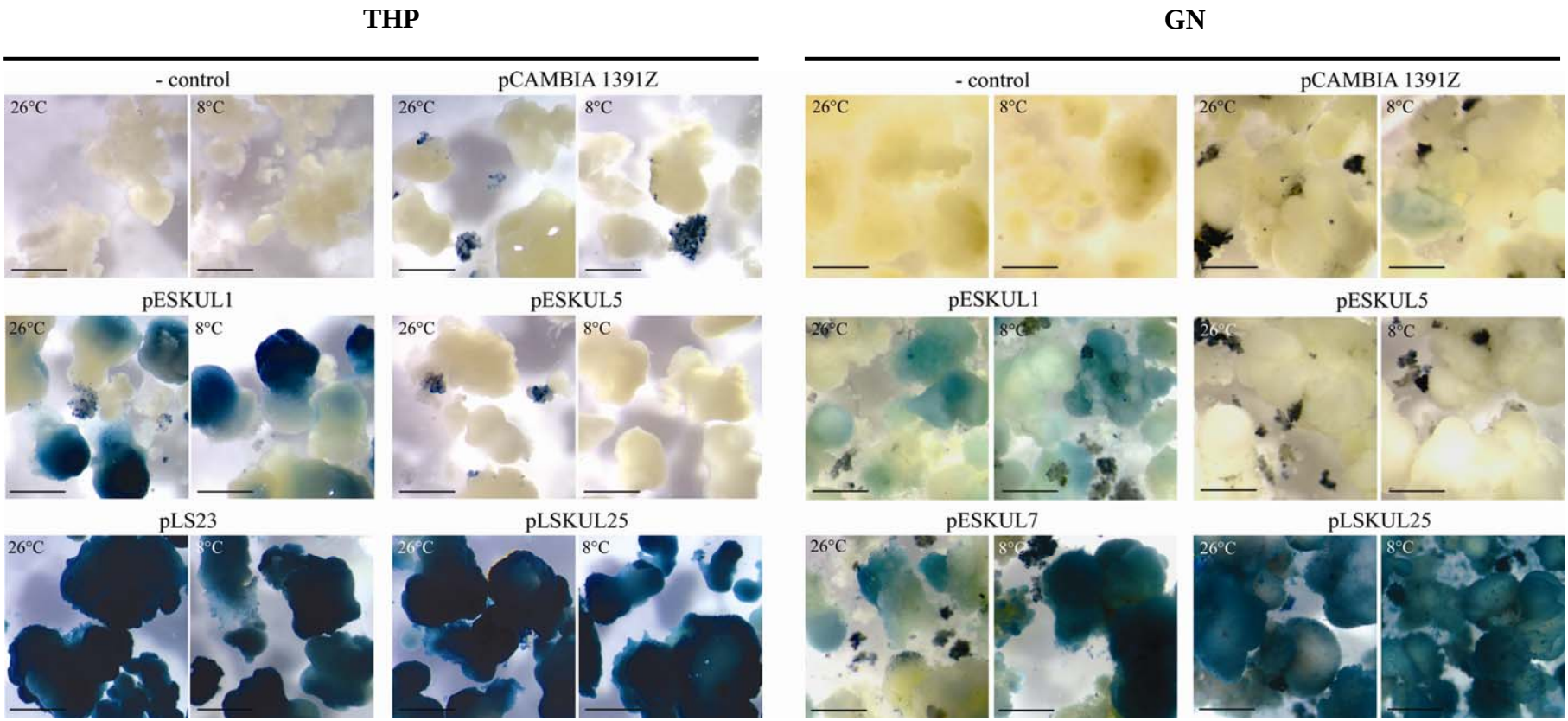

Figure 6.4. Histochemical GUS assay of transgenic cell colonies of the plantain 'Three Hand Planty' (THP) and dessert banana 'Grand naine' (GN) maintained on ZZ medium and subjected to a low temperature treatment two and six months after *Agrobacterium* back-transformation with candidate promoter sequences fused to the *uid*A^{INT} gene, respectively. Cell colonies for the 26°C treatment were incubated at 26°C for 17 h, while cells for the 8°C treatment were first incubated at 26°C for 6 h and then subjected to 8°C for 10 h including an one hour period of temperature transition. Non-transformed cells and cells transformed with the cloning vector pCAMBIA 1391Z are indicated as '- control' and 'pCAMBIA 1391Z', respectively. The vectors pESKUL1 and pESKUL5 were used for transformation of THP embryogenic cell suspensions, while the vectors pESKUL1, pESKUL5 and pESKUL7 were used to transform GN. The enhanced 35S CaMV promoter and maize ubiquitin promoter drive the *uid*A^{INT} reporter gene in vectors pLSKUL23 and pLSKUL25, respectively. Bars indicate 1 mm.

The original promoter-tagged line ET2-17 showed an immediate increase in LUC expression upon LT (8°C) treatment and the expression remained elevated until at least 17 h at 8°C three months after transformation (Figure 6.5, partly copied from Figure 3.4). Similarly, but in contrast to the histochemical GUS assay, a significant induction of enzymatic GUS activity upon the temperature decrease to 8°C was observed in back-transformed pESKUL1 cell colonies six months after transformation (Figure 6.5). However, the level of reporter gene induction was lower in the back-transformed cell colonies than in the original tagged cell colony. Induction of reporter gene activity by the 8°C treatment was 11.2-fold in the original tagged ET2-17 line (749 *vs.* 8434 RLU), and 2.2-fold in the pESKUL1 back-transformed lines (50 ± 0 *vs.* 111 ± 28 pmol MU h^{-1} μg^{-1} protein). Obviously, the RLU values of the same time points as assayed in the fluorometric GUS assay, *i.e.* 0 h and 11 h from the start of the temperature transition, are given here.

After three months of back-transformation part of the cell colonies were subcultured on RD1 medium while others were maintained on ZZ medium. Back-transformed lines of THP (data not shown) and GN (Figure 6.6) were histochemically stained for GUS activity six months after transformation when maintained on RD1 medium. They were again subjected to the 26°C treatment (26°C for 17 h) or to the 8°C treatment (8°C for 10 h after an incubation of 6 h at 26°C). Non-transformed cell cultures and cell cultures transformed with pESKUL1, pESKUL5 or pLSKUL25 were divided into two, one part for each temperature treatment. As observed in the screening at the ZZ stage (Figure 6.4), the pESKUL1 and pLSKUL25 cell cultures displayed similar GUS staining at 26°C and at 8°C, while GUS activity was not detected in pESKUL5 cell cultures at either temperature (Figure 6.6A). At 26°C cell cultures transformed with the pESKUL2 vector (3' deletion variant of the 17-1 promoter sequence) displayed weaker GUS activity than pESKUL1 (full-length 17-1 promoter sequence) transformed cell cultures, while pESKUL7 lines showed the strongest GUS activity (Figure 6.6B). Together with the absence of activity for the remaining candidate promoters tested (pESKUL3, pESKUL4 and pESKUL6) these findings confirm the results at the ZZ stage.

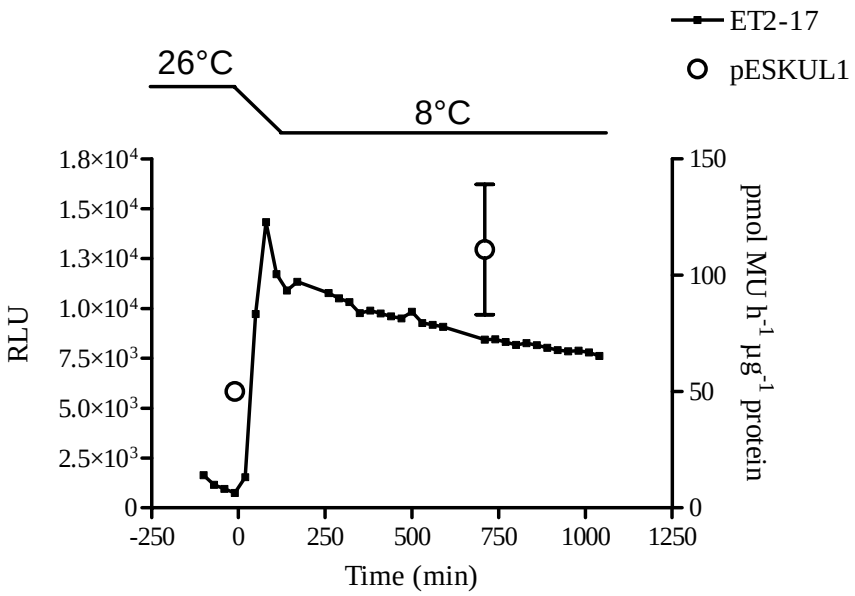

Figure 6.5. Quantitative LUC and GUS expression analysis under low temperature stress of the original promoter-tagged line ET2-17 and the back-transformed pESKUL1 lines, respectively. The 'Three Hand Planty' cell colony line ET2-17 transformed with the promoter tagging vector pETKUL2 was identified by LUC screening of transgenic cell colonies maintained on ZZ medium three months after transformation by the application of 0.1 mM luciferin and monitoring in real-time at 26°C and 8°C (Table 3.4 and Figures 3.4 and 3.5). LUC activity was monitored for 2 h at 26°C, at time point zero the temperature was set to 8°C which was reached 1 h later, and then maintained for 17 h (solid line above the graphs). One LUC image of 20 min exposure time was recorded every 30 min. LUC activity expressed in relative light units (RLU) was corrected for the background measured in a negative, non-transformed control line and is indicated in the left Y axis (closed squares). The region of interest for quantification of LUC activity was 0.34 cm^2. Enzymatic GUS activity was measured in cell extracts of 'Grand naine' cell colonies maintained on ZZ medium six months after *Agrobacterium* back-transformation with the pESKUL1 vector. This vector carries the candidate promoter sequence 17-1 identified in the promoter-tagged line ET2-17 fused to the *uid*A^{INT} transgene. Back-transformed colonies were screened independently at 26°C for 17 h or first kept at 26°C for 6 h and then subjected to 8°C for 10 h including a 1 h period of temperature transition (open circles). Approximately 170 to 180 mg fresh weight of transgenic cell colonies were processed for preparation of protein extracts. GUS activity expressed as pmol MU h^{-1} μg^{-1} of total protein is indicated in the right Y axis. Each value is the average of three independent measurements and the standard deviations are indicated. Note that the standard deviation was near 0 and therefore not distinguishable for the extracts of the colonies at 26°C for 17 h.

To confirm and quantify the GUS activity under LT stress in the back-transformed lines, fluorometric GUS assays were performed for the lines maintained on RD1 and ZZ medium six months after back-transformation (Figure 6.7). Eight lines were randomly chosen for each vector employed in the back-transformation. Cell cultures on RD1 medium were divided in half and each part went through different temperature treatments, while lines maintained on ZZ medium were not divided due to their small size. The pESKUL1 lines displayed similar GUS activity at 26°C and at 8°C on RD1 medium independent of the cultivar (Figure 6.7A and 6.7B). However, the THP pESKUL1 lines showed no GUS activity in cell colonies maintained at ZZ at either temperature.

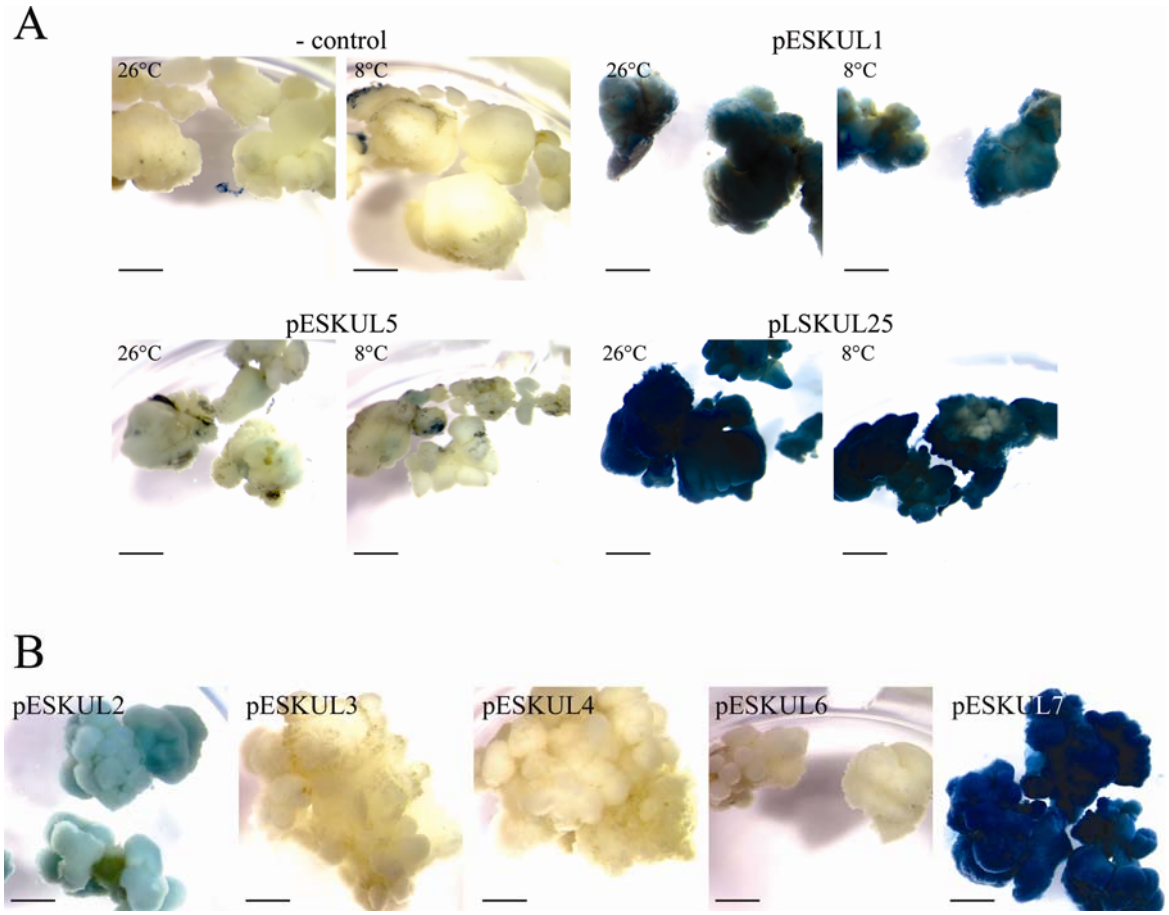

Figure 6.6. Histochemical GUS analysis of transgenic cell cultures of the dessert banana 'Grand naine' maintained on RD1 medium six months after back-transformation with candidate promoter sequences fused to the *uid*A^{INT} gene. (A) Eleven days before the temperature treatment each cell culture line assayed was divided in two equal parts, one for screening at 26°C and the other for the screening at 8°C. Non-transformed (- control), pESKUL1, pESKUL5 and pLSKUL25 (maize ubiquitin promoter) transformed cell cultures were incubated at 26°C for 17 h for the 26°C treatment, whereas cultures for the 8°C treatment were first incubated at 26°C for 6 h and then subjected to 8°C for 10 h including an one hour period of temperature transition. (B) Cells transformed with the pESKUL2, pESKUL3, pESKUL4, pESKUL6 and pESKUL7 were maintained at 26°C before histochemical GUS staining. Bars indicate 1 mm.

As already indicated in Figure 6.5 the enzymatic GUS activity was 2.2-fold induced by the 8°C temperature stress in GN pESKUL1 lines at the ZZ stage, but this induction disappeared at the RD1 stage (Figure 6.7B). GUS activity in extracts isolated from pESKUL5, pCAMBIA 1391Z and non-transformed lines was not detected for the different cultivar and media. Except for GN cell cultures on RD1 medium, significant differences in enzymatic GUS activity were recorded between both temperatures for all positive control pLSKUL25 lines. Moreover, whereas a significant decrease in GUS activity occurred at 8°C in these positive control cultures of both cultivars at the ZZ stage, this low temperature treatment caused a significant increase in GUS activity at the RD1 stage in THP.

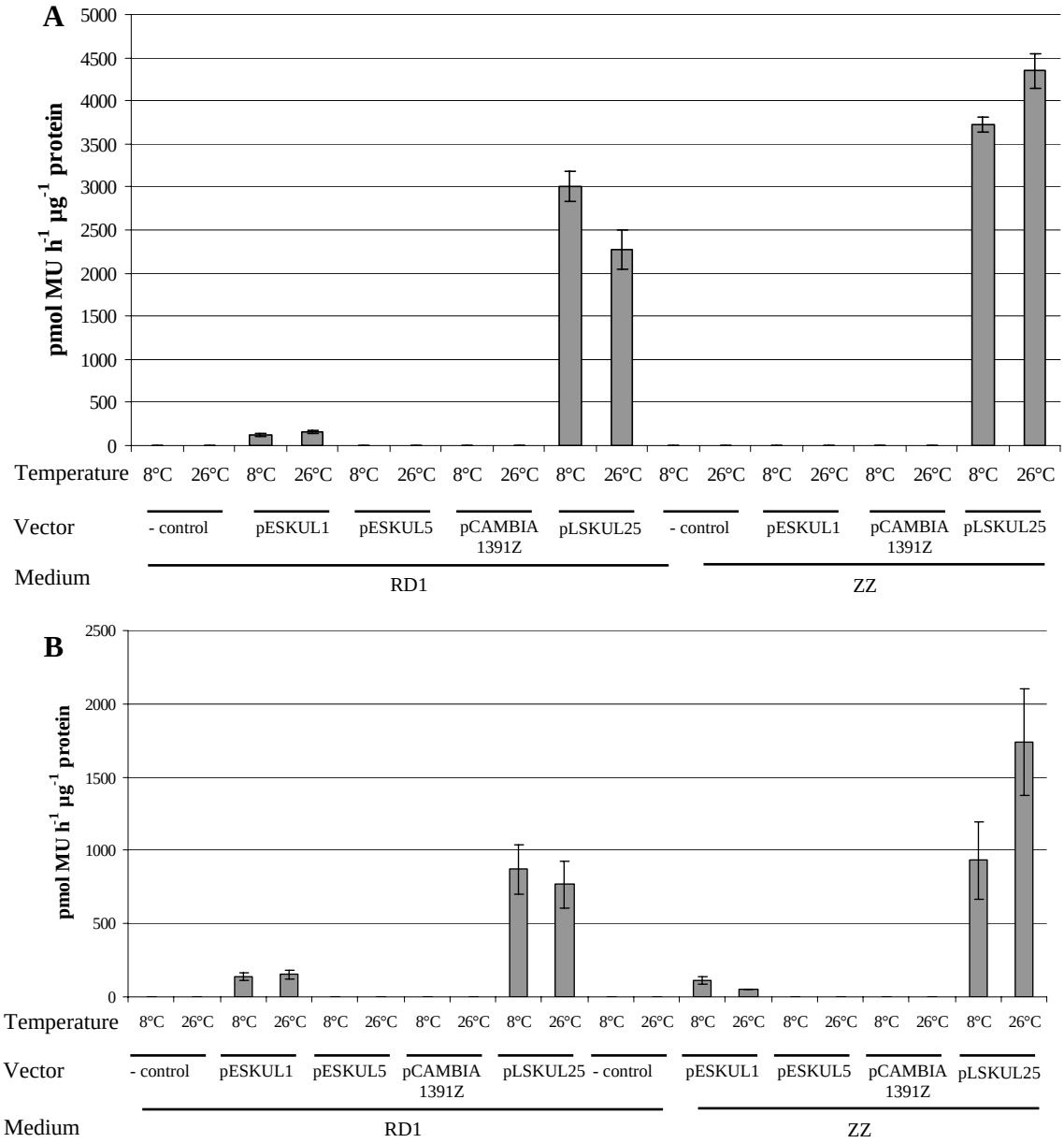

Figure 6.7. Enzymatic GUS activity in cell extracts of 'Three Hand Planty' (A) and 'Grand naine' (B) cultures maintained on ZZ or RD1 medium at two different temperatures six months after *Agrobacterium* back-transformation with candidate promoter sequences fused to the *uid*A^{INT} gene. Eleven days prior to the temperature treatments each cell culture assayed was divided in two, one part for each temperature screening for the lines maintained on RD1 medium. Cell cultures maintained on ZZ medium were incubated at 26°C or 8°C. The temperature treatments consisted of 17 h at 26°C for the 26°C treatment or 6 h at 26°C followed by 10 h at 8°C including a 1 h period of temperature transition. E ight independent lines were randomly chosen per vector and pooled for the analysis. Approximately 180 to 530 mg fresh weight of transgenic cell cultures maintained on RD1 medium was processed per vector, while for lines maintained on ZZ medium this amounted up to 60 to 220 mg fresh weight of transgenic cell colonies. Non-transformed cell cultures and cell cultures transformed with the cloning vector pCAMBIA 1391Z are indicated as '- control' and 'pCAMBIA 1391Z', respectively. Only the banana promoter containing vectors pESKUL1 and pESKUL5 were used besides the positive control vectors pLSKUL23 (pe35S::*uid*A^{INT}) and pLSKUL25 (p*Ubi*1::*uid*A^{INT}). GUS activity is expressed as pmol MU h^{-1} μg^{-1} of total protein. Each entry is the average of three independent measurements, the standard deviation is indicated.

In summary, and despite the fact that a different reporter gene was used than for the tagging, recapitulation of the enhanced promoter activity upon LT at the undifferentiated cell colony stage in the original promoter-tagged ET2-17 line was obtained in the back-transformed cell colonies for the promoter sequence 17-1. In two (ET2-17 and ET2-85) out of three promoter-tagged lines the promoters have been identified and their activity confirmed *via* back-transformation.

6.4 Discussion

Candidate promoters which were transcriptionally active in the tagged lines ET2-17, ET2-34 and ET2-85 were fused to the *uid*A^{INT} reporter gene in the pCAMBIA 1391Z vector for verification of promoter activity into two different banana varieties, the plantain THP and the dessert banana GN. Transient GUS activity (TGA) analyses revealed a very weak promoter activity for all the candidate promoter sequences in GN and THP embryogenic cells. TGA frequencies varied from less than 1% up to 3-4% of that reached by the positive control enhanced 35S and maize *Ubi*1 promoters. The TGA frequencies obtained with the latter two promoters (ca. 1000 blue foci per 50 mg fresh weight of GN suspension cells) confirm previous findings (Sagi *et al.* 1995a; Arinaitwe 2008) and demonstrate efficient T-DNA transfer. Two to six months after transformation transgenic cell colonies containing the THP promoter sequences 17-1 (pESKUL1: full-length, and pESKUL2: 3' deletion variant) or 85-1 (pESKUL7) showed stable baseline GUS activity as detected by histochemical and enzymatic analyses. Thus, the sequences 17-1 and 85-1 are responsible for the LUC expression observed in the tagged lines ET2-17 and ET2-85, respectively. Based on fluorometric analyses, the full-length 17-1 promoter sequence is 21 (THP)- to 26 (GN)-fold less active at 26°C than the maize *Ubi*1 promoter, which is among the strongest known plant promoters in banana (Dugdale *et al.* 1998; Hermann *et al.* 2001a; Atkinson *et al.* 2004). Similar differences existed between the 17-1 and enhanced 35S promoters. These results also demonstrate that promoter sequences isolated from a plantain are active in a dessert banana. Although the histochemical GUS staining suggested that the 3' deletion variant of the 17-1 promoter was weakly active in GN cell colonies, enzymatic GUS activity was not detected. Based on the fluorometric GUS data, all elements necessary for the recruitment of the RNA polymerase might be located downstream of position -393 and the candidate TATA identified within the region -392 to -1 relative to the T-DNA integration site in the original promoter-tagged line (Figure 4.14A) might be functional. However, further expression analyses need to be performed with 5' and 3' deletion variants of sequence 17-1 to confirm the activity of the TATA boxes within the -392 to -1 region. The histochemical GUS staining did not reveal any major difference in activity between the 17-1 and 85-1 promoters, but the quantitative fluorometric GUS data clearly show that the latter promoter is four times more active than the former one in GN cell colonies.

Therefore, the maize *Ubi*1 promoter was only 6 times more active than the THP 85-1 promoter in GN cell colonies suggesting that the 85-1 promoter can be considered a strong promoter in banana. This promoter activity is fully in agreement with the transcriptional fusion detected between the 85-1 sequence and the *luc*$^+$ in the original tagged line (Figure 5.1B). Finally, all these data correlate well with the LUC expression at 26°C detected in the original tagged lines ET2-17 and ET2-85 at cell colony stage three months after transformation (Table 3.3 and Figure 3.4).

Although tagged sequence 17-2 weakly activated *luc*$^+$ as detected by RT-PCR analysis (Figure 5.3), it did not have promoter activity in back-transformed cell colonies. Perhaps this sequence might still be part of a coding region (Figure 4.14A) and the more upstream sequences contain promoter elements. As expected since it was not transcriptionally fused to *luc*$^+$ (Figure 5.1B), sequence 17-3 was not able to drive transcription of the *uid*A^{INT} reporter gene in back-transformed cell colonies. The transcriptionally active sequence 34-1 (Figure 5.1B) could also not drive GUS expression in back-transformed cell colonies, neither as the full-length (with a fragment containing a 109 additional base pairs; Table 6.1) nor the 5' deletion variant. Again, the lack of promoter activity might be due to the absence of the necessary elements for transcription activation, or an additional 5'-tagged sequence might be responsible for the LUC expression that was not identified in the tagged line ET2-34. Major differences in gene expression between the original tagged line and the back-transformed lines might be due to the exclusion of regulatory DNA elements in the cloned promoter-reporter gene fusion (Ökrész *et al.* 1998; Farrar *et al.* 2003; Mollier *et al.* 2000; Stangeland *et al.* 2005). Expression analysis of the original T-DNA tagged line, therefore, may reflect promoter activity better than cloned promoter-gene fusions because all regulatory elements are present in their natural context. This conclusion is of course only valid under the assumption that the disruption of the promoter or gene by the T-DNA does not alter the natural activity pattern of the promoter (Springer 2000). Accurate control of reporter gene expression might be also dependent on sequences that are not strictly part of the promoter like the 5' and/or 3' untranslated regions of a gene (Taylor 1997). This possibility can also be explored by back-transformation, but only if the structure of the corresponding tagged gene is known.

Due to the long half-life of the GUS enzyme (~60 h, Jefferson *et al.* 1986) and the destructive properties of the GUS assays, the *uid*A reporter gene is not suited for real-time gene expression studies (De Ruijter *et al.* 2003). On the other hand, for the analysis of promoter activity under two or more different conditions (i.e. untreated vs. treated) using the *uid*A reporter gene one or more distinct time points can be assayed. On the contrary, the *luc* reporter gene allows real-time detection of LUC enzymatic activity in a non-invasive and non-destructive manner combined with high sensitivity (Ow *et al.* 1986) due to its short half-life (Van Leeuwen *et al.* 2000). Promoter activity at 26°C and 8°C was analyzed in real-time in the original *luc*$^+$-tagged lines, whereas in back-transformed lines it was

assessed by GUS enzymatic activity measurements in protein extracts of cell colony lines incubated at 26°C for 17 h or at 26°C for 6 h followed by 10 h at 8°C. Although a direct comparison between these two reporter gene systems is not appropriate because of the aforementioned inherent differences, both showed an up-regulation of 17-1 promoter activity by LT treatment at cell colony stage (ZZ medium). Apparently a 2-fold difference of GUS activity could not be differentiated by histochemical assays as observed in Figure 6.4. The lower level of promoter induction by 8°C in back-transformed lines than in the original tagged ET2-17 line (2.2-fold vs. 11.2-fold, respectively) may be caused by the stability of the GUS enzyme. Low temperature up-regulated promoters have also been identified using the *uid*A reporter gene in Arabidopsis plants (4°C; Mandal *et al.* 1995) and in rice shoots (5°C; Lee *et al.* 2004a).

The absence of promoter activity was confirmed for the sequence 34-1 (pESKUL5) at LT similar to pCAMBIA 1391Z and non-transformed lines. On the other hand, the promoter sequence 85-1 (pESKUL7) showed an apparent increase of the activity at LT as revealed by the histochemical analysis. Fluorometric analyses were not performed in the lines transformed with the pESKUL7 vector and therefore the increase in GUS activity by LT could not be verified quantitatively. However, the original tagged line ET2-85 showed a decrease of LUC activity at LT contradicting the histochemical data, and suggesting that further analysis should be performed to confirm the activity pattern of the cloned promoter 85-1 at LT. Developmental regulation of the identified promoters 17-1 and 85-1 needs to be further investigated because only early undifferentiated stages have been analyzed. Preliminary results in back-transformed lines at later developmental stages revealed that these sequences resemble the promoter activity in the original promoter-tagged lines ET2-17 and ET2-85, respectively (data not shown). Moreover, these results indicate that the promoter activity of sequence 85-1 increased to the levels of the enhanced 35S and *Ubi*1 promoters (data not shown).

The decrease in GUS activity in the pLSKUL25 transformed cell colonies under 8°C stress might be attributed to an inhibiting effect of LT on the maize *Ubi*1 promoter activity. A consistent 2- to 4-fold lower GUS activity at 4°C than at 24°C in wheat leaves controlled by the maize *Ubi*1 promoter has been reported by Ouellet *et al.* (1998). In contrast, exposure of two stably transformed (*UBI1:GUS*) rice callus lines to freezing (-20°C for 1 h) resulted in an approximately 3-fold increased GUS expression (Perales *et al.* 2007). Furthermore, pLSKUL25 banana cell cultures maintained on RD1 medium showed an increase in GUS activity at 8°C compared to the 26°C treatment. Further fluorometric GUS assays should be performed in the positive control lines to confirm the different GUS activity patterns observed at 26°C and 8°C at these developmental stages. Apparently, LT conditions play an important role in the induction or suppression of the maize *Ubi*1 promoter, which also warrants further investigation on the activation of promoter sequence 17-1 under different LT conditions.

In conclusion, two novel promoters have been identified and characterized in banana, one of which showed LT up-regulation at undifferentiated cell colony stage while both posses a baseline constitutive activity during the *in vitro* regeneration process.

Chapter 7 General discussion and perspectives

Promoters are essential to create transgenic plants *via* genetic transformation to drive expression of desired genes. However, only a limited number of promoters from banana have been isolated, characterized and used in transgenic bananas. In addition, banana promoters fused to banana genes should provide more public acceptance than the use of heterologous sequences. A high proportion of cultivated bananas are sterile and thus, genetic transformation is a useful tool to increase resistance to biotic and abiotic stresses. In this work, novel banana promoters were tagged in the banana cultivars 'Three Hand Planty' (THP), 'Williams' and 'Cacambou'. Embryogenic cell suspension (ECS) samples of each cultivar were co-cultivated with *A. tumefaciens* strain EHA105 harboring the promoterless *luc*$^{+}$ (pETKUL2) or *luc*$^{+INT}$ (pKCKUL1) near the RB of the T-DNA. Two to three months after transformation, a liquid nitrogen-cooled CCD camera allowed high-throughput screening of approximately 89,000 independent transgenic cell colonies for LUC activation. In total, 697 light emitting cell colonies were identified and 96 THP lines were re-screened throughout the *in vitro* regeneration process. The main reason for this early screening is the required scale of the work, because banana T-DNA tagging efficiency, like in many other plant species, is less than 5% (Remy *et al.* 2005). In banana, only ca. 20% of transgenic lines carry 1 copy of the transgene. Therefore, early determination of the T-DNA copy number could greatly reduce the number of candidate lines for re-screening at later developmental stages because lines with multiple T-DNA copies could be discarded (see below). Due to the limited amount of material available at the early cell colony stage, Southern blot analysis was excluded, but real-time quantitative PCR could circumvent this problem (Yang *et al.* 2005).

The presence of an intron in the *luc*$^{+}$ transgene increased the baseline tagging frequency from 0.89% to 1.38%, probably due to an intron enhancement effect on LUC expression allowing the detection of weaker promoters. Alternatively, the use of recently improved *luc* versions might also increase the sensitivity of the screening. Such improved luciferase versions (e.g. the *luc2* and the destabilized luciferases are already commercially available (www.promega.com/tbs/tm259/tm259.pdf). The latter have a reduced half-life allowing an even more real-time or closer monitoring of the response in transcriptional rate during stress treatments reflecting the true promoter activity. Not only may these improved *luc* versions yield a greater number of tagged promoters, but also the integration time for monitoring LUC activity by CCD camera observation might be reduced from 20 min to 5 min or less, thereby reducing the background noise of the acquired images.

Detailed analysis of reporter gene activity during *in vitro* regeneration and LT stress was performed in a core group of 23 promoter-tagged THP lines. Two lines showed an increase in LUC activity at the early developmental stages by LT (8°C) treatment (Table

3.3, lines ET2-17 and ET2-42). However, the up-regulation by 8°C was absent at subsequent developmental stages for these two lines (Table 3.3 and Figure 3.4) indicating that the LT response was modulated during the *in vitro* regeneration process. In addition, three promoter-tagged lines which showed LUC activity at the undifferentiated cell colony stage were not preserving that activity at later regeneration stages (ET2-28, ET2-37 and ET2-62) implying that the promoters tagged in these lines are only active at undifferentiated stages. Thus, the tagging system also enabled the identification of banana promoters that are restricted in their activity to the selection period of 2 to 3 months (Pérez Hernández *et al.* 2006b). Such promoters are thus suitable to drive the expression of the selectable marker gene only when needed.

To verify the promoter activity pattern at different temperatures, LUC enzymatic activity should be measured in extracts of samples taken at the different temperatures. This assay will allow discrimination between a promoter effect and a more general effect of the temperature at the post-translational level, because it is known that LUC enzymatic activity decreases when lowering the temperature at least in solution (McElroy and Seliger 1961; Dickinson *et al.* 1993; Ueda *et al.* 1994). Despite several attempts, however, researchers in our laboratory have not been able yet to set up a reliable LUC enzymatic activity assay for banana tissues including cell colonies (Serge Remy, personal communication).

One of the most important factors in T-DNA tagging experiments is the number of integrated T-DNA copies in the tagged lines. One T-DNA copy is preferred since it greatly simplifies the isolation of the activated promoter. Possible causes for multiple T-DNA copies and complex rearrangements integrated into the plant genome are the vector harboring the T-DNA (Scholte *et al.* 2002) and/or the *A. tumefaciens* strain used (Jorgensen *et al.* 1987). Multiple T-DNA integration events were detected in T-DNA tagging experiments where pBin19 derived vectors were employed (*Medicago trunculata*: Trinh *et al.* 1998, Scholte *et al.* 2002; *Lotus japonicus*: Martirani *et al.* 1999; *Nicotiana tabacum*: Fobert *et al.* 1991). Most of the banana promoter-tagged lines analyzed in this work were transformed with the pETKUL2 tagging vector (Remy *et al.* 2005) and contained multiple and complex T-DNA integration events. The pETKUL2 vector is a derivate of the pCAMBIA 2300 vector and thus from the pPZP binary plasmid family (www.cambia.org). Similar complex T-DNA integration events were observed in transgenic Creeping bentgrass (*Agrostis palustris* Huds.) when using pPZP derived vectors (Fu *et al.* 2006). However, an alternative explanation for the presence of multiple T-DNA copies in transgenic lines is the presence of the selectable marker gene *neo*. Expression of a single *neo* might not be enough to confer resistance to antibiotics during the selection period, resulting in the selection for transgenic lines with multiple T-DNA copies (Scholte *et al.* 2002). The type of explant used for transformation might also influence the number of T-DNA copies (Grevelding *et al.* 1993). Most likely a combination of several factors promotes multiples T-DNA copies. Especially in the case of tandem repeats, the type and physiology of bacterial *inoculum*, the

copy number of the T-DNA replicon, and the physiology and competence of the explant to be transformed play jointly a role in the integration process (De Neve *et al.* 1997). Due to the number and complexity of the factors involved, it is not yet possible to control the number of integrated T-DNA copies and thus selection for single T-DNA copy lines remains necessary.

Identification of a specific T-DNA rearrangement in which the enhanced CaMV 35S promoter with part of the *neo* selectable marker gene integrated upstream of the *luc*$^+$ was observed in some promoter-tagged banana lines. Hence, the observed LUC activity might be (partly) controlled by the enhanced CaMV 35S promoter. To detect this complex T-DNA integration event and discard the positive promoter-tagged lines, a PCR screening was performed prior to the isolation of the T-DNA flanking sequences. Moreover, the presence of CaMV 35S promoter sequences in the T-DNA could affect the level and/or pattern of promoter activity in the vicinity, including the putative promoters tagged, due to its enhancer properties. The presence of promoters lacking enhancer elements in the T-DNA of the tagging vector might therefore, result in a more reliable evaluation of the true (native) activity of the tagged promoters. Perhaps the banana promoters isolated in this and previous work (Remy *et al.* 2005) could drive the selectable marker gene.

The time consuming RT-PCR analysis to identify the LUC activating flanking sequence could be avoided if single T-DNA integration is obtained. On the other hand, even for lines carrying a single T-DNA RT-PCR is useful for early confirmation of promoter activity. When multiple T-DNA flanking sequences are identified, the RT-PCR approach proved to be a necessary and powerful tool to identify the T-DNA flanking sequence with transcriptional activity.

In conclusion, two banana promoters (17-1 and 85-1) were isolated and their activities were confirmed in back-transformed lines at the cell colony stage. The activity of the novel promoters was lower than the constitutive enhanced 35S and maize *Ubi*1 promoters at early undifferentiated stages (Figure 6.3B). Yet they might be useful for applications requiring lower gene expression levels and thus with less energy input of the cells and less interference with essential pathways. Additional characterization of these banana promoters in back-transformed banana and other plant species is therefore advisable. As only early undifferentiated stages of back-transformed lines in combination with LT stress were analyzed, determination of the developmental regulation of the tagged promoters is ongoing. Finally, analysis of promoter activity should be performed under greenhouse and field conditions as well to assess their possible use for transgenic banana production.

References

Acereto-Escoffié POM, Chi-Manzanero BH, Echeverría-Echeverría S, Grijalva R, Kay AJ, González-Estrada T, Castaño E, Rodríguez-Zapata LC (2005) *Agrobacterium*-mediated transformation of *Musa acuminata* cv. "Grand Nain" scalps by vacuum infiltration. Scientia Horticulturae 105: 359-371

Aflalo C (1991) Biologically localized firefly luciferase: a tool to study cellular processes. International Review of Cytology 130: 269-323

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research 25: 4692-4693

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen HM, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301: 653-657

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25: 3389-3402

Alvarado MC, Zsigmond LM, Kovacs I, Cseplo A, Koncz C, Szabados LM (2004) Gene trapping with firefly luciferase in Arabidopsis. Tagging of stress-responsive genes. Plant Physiology 134: 18-27

An SY, Park S, Jeong DH, Lee DY, Kang HG, Yu JH, Hur J, Kim SR, Kim YH, Lee M, Han SK, Kim SJ, Yang JW, Kim E, Wi SJ, Chung HS, Hong JP, Choe V, Lee HK, Choi JH, Nam JM, Kim SR, Park PB, Park KY, Kim WT, Choe S, Lee CB, An GH (2003) Generation and analysis of end sequence database for T-DNA tagging lines in rice. Plant Physiology 133: 2040-2047

André D, Colau D, Schell J, Montagu M, Hernalsteens J-P (1986) Gene tagging in plants by a T-DNA insertion mutagen that generates APH(3')II-plant gene fusions. Molecular and General Genetics 204: 512-518

Anuradha TS, Jami SK, Datla RS, Kirti PB (2006) Genetic transformation of peanut (*Arachis hypogaea* L.) using cotyledonary node as explant and a promoterless *gus::npt*II fusion gene based vector. Journal of Biosciences 31: 235-246

Arinaitwe G (2008) An improved *Agrobacterium*-mediated transformation method for banana and plantain (*Musa* spp.). Dissertationes de agricultura. Ph.D. thesis 781. Katholieke Universiteit Leuven, Belgium. Faculteit Bio-Ingenieurswetenschappen. 156 p.

Arinaitwe G, Remy S, Strosse H, Swennen R, Sági L (2004) *Agrobacterium*- and particle bombardment-mediated transformation of a wide range of banana cultivars. In: Banana Improvement: Cellular, Molecular Biology, and Induced Mutations. S. Mohan Jain and R. Swennen, editors. Science Publishers Inc., Enfield, NH, USA. pp. 351-357

Atkinson HJ, Grimwood S, Johnston K, Green J (2004) Prototype demonstration of transgenic resistance to the nematode *Radopholus similis* conferred on banana by a cystatin. Transgenic Research 13: 135-142

Bade J, van Grinsven E, Custers J, Hoekstra S, Ponstein A (2003) T-DNA tagging in *Brassica napus* as an efficient tool for the isolation of new promoters for selectable marker genes. Plant Molecular Biology 52: 53-68

Bajic VB, Seah SH, Chong A, Zhang GL, Koh JLY, Brusic V (2002) Dragon Promoter Finder: recognition of vertebrate RNA polymerase II promoters. Bioinformatics 18: 198-199

Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana COR15a* has *cis*-acting elements that confer cold-regulated, drought-regulated and ABA-regulated gene-expression. Plant Molecular Biology 24: 701-713

Barthels N, Van Der Lee FM, Klap J, Goddijn OJM, Karimi M, Puzio P, Grundler FMW, Ohl SA, Lindsey K, Robertson L, Robertson WM, Van Montagu M, Gheysen G, Sijmons PC (1997) Regulatory sequences of Arabidopsis drive reporter gene expression in nematode feeding structures. Plant Cell 9: 2119-2134

Bechtold N, Jaudeau B, Jolivet S, Maba B, Vezon D, Voisin R, Pelletier G (2000) The maternal chromosome set is the target of the T-DNA in the *in planta* transformation of *Arabidopsis thaliana*. Genetics 155: 1875-1887

Becker DK, Dugdale B, Smith MK, Harding RM, Dale JL (2000) Genetic transformation of Cavendish banana (*Musa* spp. AAA group) cv 'Grand Nain' via microprojectile bombardment. Plant Cell Reports 19: 229-234

Benfey PN, Chua NH (1990) The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. Science 250: 959-966

Berk AJ (1999) Activation of RNA polymerase II transcription. Current Opinion in Cell Biology 11: 330-335

Biddington NL (1992) The influence of ethylene in plant-tissue culture. Plant Growth Regulation 11: 173-187
Bosque-Pérez NA, May GD, Arntzen CJ (2000) Applicability of an *Agrobacterium*-based system for the transformation of *Musa* species with diverse genomic constitution and ploidy level. Acta Horticulturae 540: 193-201
Bouchez D, Camilleri C, Caboche M (1993) A binary vector based on Basta resistance for *in planta* transformation of *Arabidopsis thaliana*. Comptes Rendus De L' Académie Des Sciences Serie III-Sciences De La Vie-Life Sciences 316: 1188-1193
Bourdon V, Harvey A, Lonsdale DM (2001) Introns and their positions affect the translational activity of mRNA in plant cells. EMBO Reports 2: 394-398
Breathnach R, Chambon P (1981) Organization and expression of eukaryotic split genes-coding for proteins. Annual Review of Biochemistry 50: 349-383
Bresolin G, Neuhaus K, Scherer S, Fuchs TM (2006) Transcriptional analysis of long-term adaptation of *Yersinia enterocolitica* to low-temperature growth. Journal of Bacteriology 188: 2945-2958
Brody JR, Kern SE (2004) Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. Biotechniques 36: 214-216
Brown APC, Dunn MA, Goddard NJ, Hughes MA (2001) Identification of a novel low-temperature-response element in the promoter of the barley (*Hordeum vulgare* L) gene *blt101.1*. Planta 213: 770-780
Bustos MM, Begum D, Kalkan FA, Battraw MJ, Hall TC (1991) Positive and negative *cis*-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. EMBO Journal 10: 1469-1479
Buzas DM, Lohar D, Sato S, Nakamura Y, Tabata S, Vickers CE, Stiller J, Gresshoff PM (2005) Promoter trapping in *Lotus japonicus* reveals novel root and nodule GUS expression domains. Plant and Cell Physiology 46: 1202-1212
Calderon-Villalobos LIA, Kuhnle C, Li HB, Rosso M, Weisshaar B, Schwechheimer C (2006) LucTrap vectors are tools to generate luciferase fusions for the quantification of transcript and protein abundance *in vivo*. Plant Physiology 141: 3-14
Campisi L, Yang YZ, Yi Y, Heilig E, Herman B, Cassista AJ, Allen DW, Xiang HJ, Jack T (1999) Generation of enhancer trap lines in Arabidopsis and characterization of expression patterns in the inflorescence. Plant Journal 17: 699-707
Carpentier SC, Witters E, Laukens K, Van Onckelen H, Swennen R, Panis B (2007) Banana (*Musa* spp.) as a model to study the meristem proteome: Acclimation to osmotic stress. Proteomics 7: 92-105
Castle J, Szekeres M, Jenkins G, Bishop GJ (2005) Unique and overlapping expression patterns of Arabidopsis *CYP85* genes involved in brassinosteroid C-6 oxidation. Plant Molecular Biology 57: 129-140
Cazzonelli CI, Mccallum EJ, Lee R, Botella JR (2005) Characterization of a strong, constitutive mung bean (*Vigna radiata* L.) promoter with a complex mode of regulation in planta. Transgenic Research 14: 941-967
Chakrabarti A, Ganapathi TR, Mukherjee PK, Bapat VA (2003) MSI-99, a magainin analogue, imparts enhanced disease resistance in transgenic tobacco and banana. Planta 216: 587-596
Chandrasekharan MB, Bishop KJ, Hall TC (2003) Module-specific regulation of the β-phaseolin promoter during embryogenesis. Plant Journal 33: 853-866
Chen SY, Jin WZ, Wang MY, Zhang F, Zhou J, Jia OJ, Wu YR, Liu FY, Wu P (2003) Distribution and characterization of over 1000 T-DNA tags in rice genome. Plant Journal 36: 105-113
Chen WQ, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou GZ, Whitham SA, Budworth PR, Tao Y, Xie ZY, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T (2002) Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell 14: 559-574
Cheung F, Town CD (2007) A BAC end view of the *Musa acuminata* genome. BMC Plant Biology 7: 29
Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong XH, Agarwal M, Zhu JK (2003) ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Genes & Development 17: 1043-1054
Chinnusamy V, Zhu J, Zhu JK (2006) Gene regulation during cold acclimation in plants. Physiologia Plantarum 126: 52-61
Chiu WL, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J (1996) Engineered GFP as a vital reporter in plants. Current Biology 6: 325-330
Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Molecular Biology 18: 675-689
Custers JHHV, Melchers LS, Tigelaar H, Bade JB, Spiegeler JJM, Van Der Meijs PJ, Simons BH, Stuiver MH (2002) T-DNA tagging of a pathogen inducible promoter in *Arabidopsis thaliana*. Molecular Plant

Pathology 3: 239-249
De Buck S, De Wilde C, Van Montagu M, Depicker A (2000) T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. Molecular Breeding 6: 459-468
De Buck S, Jacobs A, Van Montagu M, Depicker A (1999) The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. Plant Journal 20: 295-304
De Greve H, Nguyen VK, Deboeck F, Lin TT, Karimi M, Hernalsteens JP (2001) T-DNA tagging of the translation initiation factor eIF-4A1 of *Arabidopsis thaliana*. Plant Science 161: 685-693
De Neve M, Debuck S, Jacobs A, Vanmontagu M, Depicker A (1997) T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. Plant Journal 11: 15-29
De Ruijter NCA, Verhees J, Van Leeuwen W, Van Der Krol AR (2003) Evaluation and comparison of the GUS, LUC and GFP reporter system for gene expression studies in plants. Plant Biology 5: 103-115
Debondt A, Eggermont K, Druart P, Devil M, Goderis I, Vanderleyden J, Broekaert WE (1994) *Agrobacterium*-mediated transformation of apple (*Malus* X *domestica* Borkh): an assessment of factors affecting gene-transfer efficiency during early transformation steps. Plant Cell Reports 13: 587-593
Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: version II. Plant Molecular Biology Reports 1: 19-21
DeLuca M, McElroy WD (1974) Kinetics of firefly luciferase catalyzed reactions. Biochemistry 13: 921-925
Desfeux C, Clough SJ, Bent AF (2000) Female reproductive tissues are the primary target of *Agrobacterium*-mediated transformation by the Arabidopsis floral-dip method. Plant Physiology 123: 895-904
Dhed'a D, Dumortier F, Panis B, Vuylsteke D, De Langhe E (1991) Plant regeneration in cell suspension cultures of the coocking banana cv. 'Bluggoe' (*Musa* spp. AAB group). Fruits 46: 125-135
Dickinson R, Franks NP, Lieb WR (1993) Thermodynamics of anesthetic protein interactions. Temperature studies on firefly luciferase. Biophysical Journal 64: 1264-1271
Dugdale B, Becker DK, Beetham PR, Harding RM, Dale JL (2000) Promoters derived from banana bunchy top virus DNA-1 to -5 direct vascular-associated expression in transgenic banana (*Musa* spp.). Plant Cell Reports 19: 810-814
Dugdale B, Becker DK, Harding RM, Dale JL (2001) Intron-mediated enhancement of the banana bunchy top virus DNA-6 promoter in banana (*Musa* spp.) embryogenic cells and plants. Plant Cell Reports 20: 220-226
Dugdale B, Beetham PR, Becker DK, Harding RM, Dale JL (1998) Promoter activity associated with the intergenic regions of banana bunchy top virus DNA-1 to-6 in transgenic tobacco and banana cells. Journal of General Virology 79: 2301-2311
Dunn MA, White AJ, Vural S, Hughes MA (1998) Identification of promoter elements in a low-temperature-responsive gene (*blt*4.9) from barley (*Hordeum vulgare* L.). Plant Molecular Biology 38: 551-64
Dynan WS (1986) Promoters for housekeeping genes. Trends in Genetics 2: 196-197
Eastmond PJ, Hooks MA, Williams D, Lange P, Bechtold N, Sarrobert C, Nussaume L, Graham IA (2000) Promoter trapping of a novel medium-chain acyl-CoA oxidase, which is induced transcriptionally during Arabidopsis seed germination. Journal of Biological Chemistry 275: 34375-34381
Eike MC, Mercy IS, Aalen RB (2005) Transgene silencing may be mediated by aberrant sense promoter sequence transcripts generated from cryptic promoters. Cellular and Molecular Life Sciences 62: 3080-3091
Elliott AR, Campbell JA, Dugdale B, Brettell RIS, Grof CPL (1999) Green fluorescent protein facilitates rapid in vivo detection of genetically transformed plant cells. Plant Cell Reports 18: 707-714
Ermolenko DN, Makhatadze GI (2002) Bacterial cold-shock proteins. Cellular and Molecular Life Sciences 59: 1902-1913
Espelund M, Jakobsen KS (1992) Cloning and direct sequencing of plant promoters using primer-adapter mediated PCR on DNA coupled to a magnetic solid phase. Biotechniques 13: 74-81
FAOSTAT (2006) Food and Agricultural Organization of the United Nations Statistical Databases, , ProdSTAT, Crops - Bananas, plantains - World Production <http://faostat.fao.org/> Accessed 08 December 2007
Farewell A, Neidhardt FC (1998) Effect of temperature on *in vivo* protein synthetic capacity in *Escherichia coli*. Journal of Bacteriology 180: 4704-4710
Farrar K, Evans IM, Topping JF, Souter MA, Nielsen JE, Lindsey K (2003) *EXORDIUM* - a gene expressed in proliferating cells and with a role in meristem function, identified by promoter trapping in Arabidopsis. Plant Journal 33: 61-73
Favery B, Lecomte P, Gil N, Bechtold N, Bouchez D, Dalmasso A, Abad P (1998) RPE, a plant gene involved in early developmental steps of nematode feeding cells. EMBO Journal 17: 6799-6811

Feldmann KA (1991) T-DNA insertion mutagenesis in *Arabidopsis*: mutational spectrum. Plant Journal 1: 71-82

Feldmann KA, Marks MD (1987) *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*: a non-tissue culture approach. Molecular and General Genetics 208: 1-9

Fickett JW, Hatzigeorgiou AC (1997) Eukaryotic promoter recognition. Genome Research 7: 861-878

Fobert PR, Labbe H, Cosmopoulos J, Gottlob-McHugh S, Ouellet T, Hattori J, Sunohara G, Iyer VN, Miki BL (1994) T-DNA tagging of a seed coat-specific cryptic promoter in tobacco. Plant Journal 6: 567-577

Fobert PR, Miki BL, Iyer VN (1991) Detection of gene regulatory signals in plants revealed by T-DNA-mediated fusions. Plant Molecular Biology 17: 837-851

Foster E, Hattori J, Labbe H, Ouellet T, Fobert PR, James LE, Iyer VN, Miki BL (1999) A tobacco cryptic constitutive promoter, tCUP, revealed by T-DNA tagging. Plant Molecular Biology 41: 45-55

Fowler S, Thomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-1690

Fu DL, St Amand PC, Xiao YM, Muthukrishnan S, Liang GH (2006) Characterization of T-DNA integration in creeping bentgrass. Plant Science 170: 225-237

Gallie DR (1993) Posttranscriptional regulation of gene-expression in plants. Annual Review of Plant Physiology and Plant Molecular Biology 44: 77-105

Galán Saúco V, Cabrera Cabrera J, Hernández Delgado PM (1992) Phenological and production differences between greenhouse and open-air bananas (*Musa acuminata* Colla AAA cv. Dwarf Cavendish) in Canary Islands. Acta Horticulturae 296: 97-111

Ganapathi TR, Higgs NS, Balint-Kurti PJ, Arntzen CJ, May GD, Van Eck JM (2001) *Agrobacterium*-mediated transformation of embryogenic cell suspensions of the banana cultivar Rasthali (AAB). Plant Cell Reports 20: 157-162

Ganry J (1973) Étude du développement du système foliaire du bananier en fonction de la température. Fruits 28: 499-516

Gheysen G, Villarroel R, Vanmontagu M (1991) Illegitimate recombination in plants: a model for T-DNA integration. Genes & Development 5: 287-297

Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression. Plant Journal 16: 433-442

Goddijn OJM, Lindsey K, Vanderlee FM, Klap JC, Sijmons PC (1993) Differential gene expression in nematode-induced feeding structures of transgenic plants harboring promoter-*gus*A fusion constructs. Plant Journal 4: 863-873

Goldsbrough AP, Albrecht H, Stratford R (1993) Salicylic acid-inducible binding of a tobacco nuclear-protein to a 10-bp sequence which is highly conserved amongst stress-inducible genes. Plant Journal 3: 563-571

Gong ZZ, Lee H, Xiong LM, Jagendorf A, Stevenson B, Zhu JK (2002) RNA helicase-like protein as an early regulator of transcription factors for plant chilling and freezing tolerance. Proceedings of the National Academy of Sciences USA 99: 11507-11512

Grace ML, Chandrasekharan MB, Hall TC, Crowe AJ (2004) Sequence and spacing of TATA box elements are critical for accurate initiation from the β-phaseolin promoter. Journal of Biological Chemistry 279: 8102-8110

Grevelding C, Fantes V, Kemper E, Schell J, Masterson R (1993) Single-copy T-DNA insertions in *Arabidopsis* are the predominant form of integration in root-derived transgenics, whereas multiple insertions are found in leaf disks. Plant Molecular Biology 23: 847-860

Grundemann D, Schomig E (1996) Protection of DNA during preparative agarose gel electrophoresis against damage induced by ultraviolet light. Biotechniques 21: 898-903

Gu RL, Zhao L, Zhang Y, Chen XP, Bao J, Zhao JF, Wang ZY, Fu JJ, Liu TS, Wang JH, Wang GY (2006) Isolation of a maize beta-glucosidase gene promoter and characterization of its activity in transgenic tobacco. Plant Cell Reports 25: 1157-1165

Guo Y, Xiong L, Ishitani M, Zhu JK (2002) An Arabidopsis mutation in translation elongation factor 2 causes superinduction of CBF/DREB1 transcription factor genes but blocks the induction of their downstream targets under low temperatures. Proceedings of the National Academy of Sciences USA 99: 7786-7791

Haake V, Cook D, Riechmann JL, Pineda O, Thomashow MF, Zhang JZ (2002) Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. Plant Physiology 130: 639-648

Halfon MS (2006) (Re)modeling the transcriptional enhancer. Nature Genetics 38: 1102-1103

Hall TC, Chandrasekharan MB, Li GF (1999) Phaseolin: its past, properties, regulation and future. In: Seed Proteins. P.R. Shewry and R. Casey, editors. Kluwer Academic Publ., Dordrecht, Netherlands. pp.209-240

Heim R, Cubitt AB, Tsien RY (1995) Improved green fluorescence. Nature 373: 663-664

Heino P, Palva ET (2003) Signal transduction in plant cold acclimation. In: Plant Responses to Abiotic Stress.

Topics in Current Genetics, 4. H. Hirt and K Shinozaki, editors. Springer-Verlag Berlin Heidelberg, Germany. pp. 151-186

Herman L, Jacobs A, Van Montagu M, Depicker A (1990) Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA integration events. Molecular and General Genetics 224: 248-256

Hermann SR, Becker DK, Harding RM, Dale JL (2001b) Promoters derived from banana bunchy top virus-associated components S1 and S2 drive transgene expression in both tobacco and banana. Plant Cell Reports 20: 642-646

Hermann SR, Harding RM, Dale JL (2001a) The banana actin 1 promoter drives near-constitutive transgene expression in vegetative tissues of banana (*Musa* spp.). Plant Cell Reports 20: 525-530

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Research 27: 297-300

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene-transfer to plants. Transgenic Research 2: 208-218

Hraska M, Rakousky S, Curn V (2006) Green fluorescent protein as a vital marker for non-destructive detection of transformation events in transgenic plants. Plant Cell Tissue and Organ Culture 86: 303-318

Huang X, Huang XL, Xiao W, Zhao JT, Dai XM, Chen YF, Li XJ (2007) Highly efficient *Agrobacterium*-mediated transformation of embryogenic cell suspensions of *Musa acuminata* cv. Mas (AA) via a liquid co-cultivation system. Plant Cell Reports 26: 1755-1762

Hur J, Jung KH, Lee CH, An GH (2004) Stress-inducible *OsP5CS2* gene is essential for salt and cold tolerance in rice. Plant Science 167: 417-426

Ishitani M, Xiong LM, Lee HJ, Stevenson B, Zhu JK (1998) *HOS1*, a genetic locus involved in cold-responsive gene expression in Arabidopsis. Plant Cell 10: 1151-1161

Ishitani M, Xiong LM, Stevenson B, Zhu JK (1997) Genetic analysis of osmotic and cold stress signal transduction in Arabidopsis: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways. Plant Cell 9: 1935-1949

Israeli Y, Lahav E (2000) Injuries to banana caused by adverse climate and weather. In: Diseases of Banana, Abacá and Enset. D.R. Jones, editor. CAB International, Wallingford, UK. pp. 351-374

Jacobsen E, Schouten HJ (2007) Cisgenesis strongly improves introgression breeding and induced translocation breeding of plants. Trends in Biotechnology 25: 219-223

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF (1998) Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. Science 280: 104-106

Janssens H, Hou SL, Jaeger J, Kim AR, Myasnikova E, Sharp D, Reinitz J (2006) Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster even skipped* gene. Nature Genetics 38: 1159-1165

Jefferson RA, Burgess SM, Hirsh D (1986) *β*-Glucuronidase from *Escherichia coli* as a gene-fusion marker. Proceedings of the National Academy of Sciences USA 83: 8447-8451

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: *β*-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO Journal 6: 3901-3907

Jeon JS, An GH (2001) Gene tagging in rice: a high throughput system for functional genomics. Plant Science 161: 211-219

Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Yang K, Nam J, An K, Han MJ, Sung RJ, Choi HS, Yu JH, Choi JH, Cho SY, Cha SS, Kim SI, An G (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant Journal 22: 561-570

Jeong DH, An SY, Kang HG, Moon S, Han JJ, Park S, Lee HS, An KS, An GH (2002) T-DNA insertional mutagenesis for activation tagging in rice. Plant Physiology 130: 1636-1644

Jeong DH, An SY, Park S, Kang HG, Park GG, Kim SR, Sim J, Kim YO, Kim MK, Kim SR, Kim J, Shin M, Jung M, An G (2006) Generation of a flanking sequence-tag database for activation-tagging lines in japonica rice. Plant Journal 45: 123-132

Jorgensen R, Snyder C, Jones JDG (1987) T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. Molecular and General Genetics 207: 471-477

Kang GZ, Wang CH, Sun GC, Wang ZX (2003) Salicylic acid changes activities of H_2O_2-metabolizing enzymes and increases the chilling tolerance of banana seedlings. Environmental and Experimental Botany 50: 9-15

Karimi M, Van Montagu M, Gheysen G (1999) Hairy root production in *Arabidopsis thaliana*: cotransformation with a promoter-trap vector results in complex T-DNA integration patterns. Plant Cell Reports 19: 133-142

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nature Biotechnology 17: 287-291

Kertbundit S, De Greve H, Deboeck F, Van Montagu M, Hernalsteens JP (1991) *In vivo* random *β*-glucuronidase gene fusions in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences USA 88: 5212-5216

Kertbundit S, Linacero R, Rouze P, Galis I, Macas J, Deboeck F, Renckens S, Hernalsteens JP, De Greve H (1998) Analysis of T-DNA-mediated translational *β*-glucuronidase gene fusions. Plant Molecular Biology 36: 205-217

Khanna H, Becker D, Kleidon J, Dale J (2004) Centrifugation Assisted *Agrobacterium tumefaciens*-mediated Transformation (CAAT) of embryogenic cell suspensions of banana (*Musa* spp. Cavendish AAA and Lady finger AAB). Molecular Breeding 14: 239-252

Kim HJ, Kim YK, Park JY, Kim J (2002) Light signalling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in *Arabidopsis thaliana*. Plant Journal 29: 693-704

Kim JC, Jeong JC, Park HC, Yoo JH, Koo YD, Yoon HW, Koo SC, Lee SH, Bahk JD, Cho MJ (2001) Cold accumulation of SCOF-1 transcripts is associated with transcriptional activation and mRNA stability. Molecules and Cells 12: 204-208

Kim SR, Lee J, Jun SH, Park S, Kang HG, Kwon S, An G (2003) Transgene structures in T-DNA-inserted rice plants. Plant Molecular Biology 52: 761-773

Knight H, Trewavas AJ, Knight MR (1996) Cold calcium signaling in Arabidopsis involves two cellular pools and a change in calcium signature after acclimation. Plant Cell 8: 489-503

Knight H, Zarka DG, Okamoto H, Thomashow ME, Knight MR (2004) Abscisic acid induces *CBF* gene transcription and subsequent induction of cold-regulated genes via the CRT promoter element. Plant Physiology 135: 1710-1717

Knudsen S (1999) Promoter2.0: for the recognition of PolII promoter sequences. Bioinformatics 15: 356-361

Koncz C, Martini N, Mayerhofer R, Konczkalman Z, Korber H, Redei GP, Schell J (1989) High-frequency T-DNA-mediated gene tagging in plants. Proceedings of the National Academy of Sciences USA 86: 8467-8471

Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: Evidence for multiple complex patterns of integration. Plant Journal 11: 945-957

Kreps JA, Wu YJ, Chang HS, Zhu T, Wang X, Harper JF (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. Plant Physiology 130: 2129-2141

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis. Plant Cell 11: 2283-2290

Kumar S, Fladung M (2000) Transgene repeats in aspen: molecular characterisation suggests simultaneous integration of independent T-DNAs into receptive hotspots in the host genome. Molecular and General Genetics 264: 20-28

Kumar S, Fladung M (2002) Transgene integration in aspen: structures of integration sites and mechanism of T-DNA integration. Plant Journal 31: 543-551

Kutach AK, Kadonaga JT (2000) The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. Molecular and Cellular Biology 20: 4754-4764

Last DI, Brettell RIS, Chamberlain DA, Chaudhury AM, Larkin PJ, Marsh EL, Peacock WJ, Dennis ES (1991) pEmu: an improved promoter for gene expression in cereal cells. Theoretical and Applied Genetics 81: 581-588

Lee H, Guo Y, Ohta M, Xiong LM, Stevenson B, Zhu JK (2002) *LOS2*, a genetic locus required for cold-responsive gene transcription encodes a bi-functional enolase. EMBO Journal 21: 2692-2702

Lee HJ, Xiong LM, Gong ZZ, Ishitani M, Stevenson B, Zhu JK (2001) The *Arabidopsis HOS1* gene negatively regulates cold signal transduction and encodes a RING finger protein that displays cold-regulated nucleo-cytoplasmic partitioning. Genes & Development 15: 912-924

Lee S, Jung KH, An G, Chung YY (2004b) Isolation and characterization of a rice cysteine protease gene, *OsCP1*, using T-DNA gene-trap system. Plant Molecular Biology 54: 755-765

Lee SC, Kim JY, Kim SH, Kim SJ, Lee K, Han SK, Choi HS, Jeong DH, An GH, Kim SR (2004a) Trapping and characterization of cold-responsive genes from T-DNA tagging lines in rice. Plant Science 166: 69-79

Lee TI, Young RA (2000) Transcription of eukaryotic protein-coding genes. Annual Review of Genetics 34: 77-137

Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y , Van De Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. Nucleic Acids Research 30: 325-327

Li GF, Chandrasekharan MB, Wolffe AP, Hall TC (2001) Chromatin structure and phaseolin gene regulation. Plant Molecular Biology 46: 121-129

Li GF, Hall TC (1999) Footprinting *in vivo* reveals changing profiles of multiple factor interactions with the *β*-phaseolin promoter during embryogenesis. Plant Journal 18: 633-641

Lima VM, Magioli C, Gerhardt LBD, Tarre E, Menezes RMG, Sachetto-Martins G, Margis-Pinheiro M (2002) Bean class IV chitinase promoter is modulated during plant development and under abiotic stress. Physiologia Plantarum 116: 512-521

Lindsey K, Wei W, Clarke MC, McArdle HF, Rooke LM, Topping JF (1993) Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. Transgenic Research 2: 33-47

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell 10: 1391-1406

Liu YG, Mitsukawa N, Oosumi T, Whittier RF (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant Journal 8: 457-463

Mandal A, Sandgren M, Holmstrom KO, Gallois P, Palva ET (1995) Identification of *Arabidopsis thaliana* sequences responsive to low temperature and abscisic acid by T-DNA tagging and *in-vivo* gene fusion. Plant Molecular Biology Reporter 13: 243-254

Maoz A, Mayr R, Bresolin G, Neuhaus K, Francis KP, Scherer S (2002) Sensitive *in situ* monitoring of a recombinant bioluminescent *Yersinia enterocolitica* reporter mutant in real time on camembert cheese. Applied and Environmental Microbiology 68: 5737-5740

Marilley M, Pasero P (1996) Common DNA structural features exhibited by eukaryotic ribosomal gene promoters. Nucleic Acids Research 24: 2204-2211

Martienssen RA (1998) Functional genomics: Probing plant gene function and expression with transposons. Proceedings of the National Academy of Sciences USA 95: 2021-2026

Martirani L, Stiller J, Mirabella R, Alfano F, Lamberti A, Radutoiu SE, Iaccarino M, Gresshoff PM, Chiurazzi M (1999) T-DNA tagging of nodulation- and root-related genes in *Lotus japonicus*: expression patterns and potential for promoter trapping and insertional mutagenesis. Molecular Plant-Microbe Interactions 12: 275-284

Matsumoto K, Cunha NB, Morais LS, Aragão FJL (2007) Instability of transgenes in in vitro-maintained banana plants. Acta Horticulturae 738: 529-533

Matsumoto K, Morais LS, Vianna GR, Aragão FJL, Rech EL (2002) Genetic transformation of banana embryogenic cells through particle bombardment using a herbicide resistance gene as selectable marker. Acta Horticulturae 575: 61-67

Maximova SN, Dandekar AM, Guiltinan MJ (1998) Investigation of *Agrobacterium*-mediated transformation of apple using green fluorescent protein: high transient expression and low stable transformation suggest that factors other than T-DNA transfer are rate-limiting. Plant Molecular Biology 37: 549-559

May GD, Afza R, Mason HS, Wiecko A, Novak FJ, Arntzen CJ (1995) Generation of transgenic banana (*Musa acuminata*) plants via *Agrobacterium*-mediated transformation. Biotechnology 13: 486-492

Mayerhofer R, Konczkalman Z, Nawrath C, Bakkeren G, Crameri A, Angelis K, Redei GP, Schell J, Hohn B, Koncz C (1991) T-DNA integration: a mode of illegitimate recombination in plants. EMBO Journal 10: 697-704

Maziah M, Sariah M, Sreeramanan S (2007a) Transgenic banana Rastali (AAB) with β-1, 3-glucanase gene for tolerance to Fusarium wilt race 1 disease via *Agrobacterium*-mediated transformation system. Plant Pathology Journal 6: 271-282

Maziah M, Sreeramanan S, Puad A, Sariah M (2007b) Production of transgenic banana cultivar, Rastali (AAB) via *Agrobacterium*-mediated transformation with rice chitinase gene. Journal of Plant Sciences 2: 504-517

McElroy WD, Seliger HH (1961) Mechanisms of bioluminescent reactions. In: Light and Life. W.D. McElroy and B. Glass, editors. The Johns Hopkins Press, Baltimore, MD, USA. pp.219-257

Medina J, Bargues M, Terol J, Perez-Alonso M, Salinas J (1999) The Arabidopsis *CBF* gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. Plant Physiology 119: 463-469

Mellor J (2006) Dynamic nucleosomes and gene transcription. Trends in Genetics 22: 320-329

Mendel RR, Muller B, Schulze J, Kolesnikov V, Zelenin A (1989) Delivery of foreign genes to intact barley cells by high-velocity microprojectiles. Theoretical and Applied Genetics 78: 31-34

Meza TJ, Stangeland B, Mercy IS, Skarn M, Nymoen DA, Berg A, Butenko MA, Hakelien AM, Haslekas C, Meza-Zepeda LA, Aalen RB (2002) Analyses of single-copy *Arabidopsis* T-DNA-transformed lines show that the presence of vector backbone sequences, short inverted repeats and DNA methylation is not sufficient or necessary for the induction of transgene silencing. Nucleic Acids Research 30: 4556-4566

Michiels A, Van Laere A, Van Den Ende W, Tucker M (2004) Expression analysis of a chicory fructan 1-exohydrolase gene reveals complex regulation by cold. Journal of Experimental Botany 55: 1325-1333

Millar AJ, Short SR, Chua NH, Kay SA (1992) A novel circadian phenotype based on firefly luciferase expression in transgenic plants. Plant Cell 4: 1075-1087

Mollier P, Hoffmann B, Orsel M, Pelletier G (2000) Tagging of a cryptic promoter that confers root-specific *gus* expression in *Arabidopsis thaliana*. Plant Cell Reports 19: 1076-1083

Mollier P, Montoro P, Delarue M, Bechtold N, Bellini C, Pelletier G (1995) Promoterless *gusA* expression in a large number of *Arabidopsis thaliana* transformants obtained by the *in planta* infiltration method. Comptes Rendus De L' Académie Des Sciences Serie III-Sciences De La Vie-Life Sciences 318: 465-474

Moore NY (1994) Fusarium wilt of banana: pathogen variability and host pathogen interaction. Ph.D. thesis. The University of Queensland, Brisbane, Australia.

Moore NY, Pegg KG, Allen RN, Irwin JAG (1993) Vegetative compatibility and distribution of *Fusarium oxysporum* f. sp. *cubense* in Australia. Australian Journal of Experimental Agriculture 33: 797-802

Mudge SR, Birch RG (1998) T-DNA tagging and characterisation of a novel meristem-specific promoter from tobacco. Australian Journal of Plant Physiology 25: 637-643

Mueller PR, Wold B (1989) *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. Science 246: 780-786

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiologia Plantarum 15: 473-497

Muskens MWM, Vissers APA, Mol JNM, Kooter JM (2000) Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. Plant Molecular Biology 43: 243-260

Nakamura M, Tsunoda T, Obokata J (2002) Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. Plant Journal 29: 1-10

Nakashima K, Yamaguchi-Shinozaki K (2006) Regulons involved in osmotic stress-responsive and cold stress-responsive gene expression in plants. Physiologia Plantarum 126: 62-71

Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses. Plant Journal 34: 137-148

Nott A, Le Hir H, Moore MJ (2004) Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes & Development 18: 210-222

Novillo F, Alonso JM, Ecker JR, Salinas J (2004) CBF2/DREB1C is a negative regulator of CBF1/DREB1B and CBF3/DREB1A expression and plays a central role in stress tolerance in Arabidopsis. Proceedings of the National Academy of Sciences USA 101: 3985-3990

Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. Genetics 120: 621-623

Ohler U, Harbeck S, Niemann H, Noth M, Reese MG (1999) Interpolated Markov chains for eukaryotic promoter recognition. Bioinformatics 15: 362-369

Ohler U, Niemann H, Liao GC, Rubin GM (2001) Joint modeling of DNA sequence and physical properties to improve eukaryotic promoter recognition. Bioinformatics 17: S199-S206

Ouellet F, Vazquez-Tello A, Sarhan F (1998) The wheat *wcs120* promoter is cold-inducible in both monocotyledonous and dicotyledonous species. FEBS Letters 423: 324-328

Ow DW, Wood KW, DeLuca M, de Wet JR, Helinski DR, Howell SH (1986) Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants and transgenic plants. Science 234: 856-859

Ökrész L, Mathe C, Horvath E, Schell J, Koncz C, Szabados L (1998) T-DNA trapping of a cryptic promoter identifies an ortholog of highly conserved SNZ growth arrest response genes in Arabidopsis. Plant Science 138: 217-228

Pandey SP, Krishnamachari A (2006) Computational analysis of plant RNA Pol-II promoters. Biosystems 83: 38-50

Panis B, Helliot B, Strosse H, Remy S, Lepoivre P, Swennen R (2005b) Germplasm conservation, virus eradication and safe storage of transformation competent cultures in banana: The importance of cryopreservation. Acta Horticulturae 692: 51-59

Panis B, Piette B, Swennen R (2005a) Droplet vitrification of apical meristems: a cryopreservation protocol applicable to all *Musaceae*. Plant Science 168: 45-55

Panis B, Strosse H, Van Den Hende S, Swennen R (2002) Sucrose preculture to simplify cryopreservation of banana meristem cultures. Cryoletters 23: 375-384

Panis B, Totte N, Vannimmen K, Withers LA, Swennen R (1996) Cryopreservation of banana (*Musa* spp.) meristem cultures after preculture on sucrose. Plant Science 121: 95-106

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proceedings of the National Academy of Sciences USA 85: 2444-2448

Pegg KG, Moore NY, Bentley S (1996) Fusarium wilt of banana in Australia: a review. Australian Journal of Agricultural Research 47: 637-650

Pei XW, Chen SK, Wen RM, Ye S, Huang JQ, Zhang YQ, Wang BS, Wang ZX, Jia SR (2005) Creation of

transgenic bananas expressing human lysozyme gene for Panama wilt resistance. Journal of Integrative Plant Biology 47: 971-977

Pellegrineschi A, Reynolds M, Pacheco M, Brito RM, Almeraya R, Yamaguchi-Shinozaki K, Hoisington D (2004) Stress-induced expression in wheat of the Arabidopsis thaliana DREB1A gene delays water stress symptoms under greenhouse conditions. Genome 47: 493-500

Perales L, Peńarrubia L, Cornejo MJ (2007) Induction of a polyubiquitin gene promoter by dehydration stresses in transformed rice cells. Journal of Plant Physiology 165: 159-171

Pfeifer GP, Steigerwald SD, Mueller PR, Wold B, Riggs AD (1989) Genomic sequencing and methylation analysis by ligation mediated PCR. Science 246: 810-813

Phillips JR, Dunn MA, Hughes MA (1997) mRNA stability and localisation of the low-temperature-responsive barley gene family *blt14*. Plant Molecular Biology 33: 1013-1023

Pineda CR, Toro Perea N, Narvaez J, Orozco ML, Laignelet A, Cárdenas H (2002) Genetic transformation by *Agrobacterium tumefaciens* of embryogenic cell suspensions of plantain 'Dominico hartón' (*Musa* AAB Simmonds). Infomusa 11: 9-13

Plesch G, Kamann E, Mueller-Roeber B (2000) Cloning of regulatory sequences mediating guard-cell-specific gene expression. Gene 249: 83-89

Ploetz RC, Herbert J, Sebasigari K, Hernández JH, Pegg KG, Ventura JA, Mayato LS (1990) Importance of fusarium wilt in different banana-growing regions. In: Fusarium Wilt of Bananas. R.C. Ploetz, editor. APS Press, American Phytopathological Society. St. Paul, MN, USA. pp. 9-26

Podevin N, De Buck S, De Wilde C, Depicker A (2006) Insights into recognition of the T-DNA border repeats as termination sites for T-strand synthesis by *Agrobacterium tumefaciens*. Transgenic Research 15: 557-571

Prasad AM, Sivanandan C, Resminath R, Thakare DR, Bhat SR, Srinivasan (2005) Cloning and characterization of a pentatricopeptide protein encoding gene (*LOJ*) that is specifically expressed in lateral organ junctions in *Arabidopsis thaliana*. Gene 353: 67-79

Praz V, Perier R, Bonnard C, Bucher P (2002) The Eukaryotic Promoter Database, EPD: new entry types and links to gene expression data. Nucleic Acids Research 30: 322-324

Pérez-Hernández J, Swennen R, Sági L (2006a) Number and accuracy of T-DNA insertions in transgenic banana (*Musa* spp.) plants characterized by an improved anchored PCR technique. Transgenic Research 15: 139-150

Pérez Hernández JB (2000) Development and application of *Agrobacterium*-mediated genetic transformation to increase fungus-resistance in banana (*Musa* spp). Dissertationes de agricultura. Ph.D. thesis 442. Katholieke Universiteit Leuven, Belgium. Faculteit Bio-ingenieurswetenschappen. 187 p.

Pérez Hernández JB, Remy S, Galán Saúco V, Swennen R, Sági L (1998) Chemotaxis, attachment and transgene expression in the *Agrobacterium*-mediated banana transformation system. Mededelingen Faculteit Landbouwwetenschappen Universiteit Gent 63:1603-1606

Pérez Hernández JB, Remy S, Swennen R, Sági L (2006b) Banana (*Musa* sp.). In: Methods in Molecular Biology, 344: Agrobacterium Protocols, volume 2. K. Wang, editor. Humana Press Inc., Totowa, NJ, USA. pp. 167-175

Qin GJ, Kang DM, Dong YY, Shen YP, Zhang L, Deng XH, Zhang Y, Li S, Chen N, Niu WR, Chen C, Liu PC, Chen HD, Li JG, Ren YF, Gu HY, Deng XW, Qu LJ, Chen ZL (2003) Obtaining and analysis of flanking sequences from T-DNA transformants of *Arabidopsis*. Plant Science 165: 941-949

Radhamony RN, Prasad AM, Srinivasan R (2005) T-DNA insertional mutagenesis in *Arabidopsis*: a tool for functional genomics. Electronic Journal of Biotechnology 8: 82-106

Ramachandran S, Sundaresan V (2001) Transposons as tools for functional genomics. Plant Physiology and Biochemistry 39: 243-252

Reese MG (2001) Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. Computers & Chemistry 26: 51-56

Remans T (2002) Characterisation, optimisation and application of banana streak virus (BSV) promoters. Ph.D. thesis. Department of Biochemistry and Molecular Biology, The University of Queensland. Lucia, Australia. 257 p.

Remy S (2000) Genetic transformation of banana (*Musa* spp.) for disease resistance by expression of antimicrobial proteins. Dissertationes de agricultura. Ph.D. thesis 420. Katholieke Universiteit Leuven, Belgium. Faculteit Bio-ingenieurswetenschappen. 341 p.

Remy S, Buyens A, Cammue BPA, Swennen R, Sági L (1998b) Production of transgenic banana plants expressing antifungal proteins. Acta Horticulturae 490: 433-436

Remy S, De Weerdt G, Deconinck I, Swennen R, Sági L (2004) An ultrasensitive luminescent detection system in banana biotechnology: from promoter tagging to southern hybridization. In: Banana Improvement: Cellular, Molecular Biology, and Induced Mutations. S. Mohan Jain and R. Swennen, editors. Science Publishers Inc., Enfield, NH, USA. pp. 307-319

Remy S, François I, Cammue BPA, Swennen R, Sági L (1998a) Co-transformation as a potential tool to create multiple and durable resistance in banana (*Musa* spp.). Acta Horticulturae 461: 361-365

Remy S, Thiry E, Coemans B, Windelinckx S, Swennen R, Sági L (2005) Improved T-DNA vector for tagging plant promoters via high-throughput luciferase screening. Biotechniques 38: 763-770

Resminath R, Prasad AM, Thakare DR, Sivanandan C, Bhat SR, Srinivasan (2005) Promoter trapping in Arabidopsis using T-DNA insertional mutagenesis. Journal of Plant Biochemistry and Biotechnology 14: 1-8

Richardson K, Fowler S, Pullen C, Skelton C, Morris B, Putterill J (1998) T-DNA tagging of a flowering-time gene and improved gene transfer by in planta transformation of *Arabidopsis*. Australian Journal of Plant Physiology 25: 125-130

Rios G, Lossow A, Hertel B, Breuer F, Schaefer S, Broich M, Kleinow T, Jasik J, Winter J, Ferrando A, Farras R, Panicot M, Henriques R, Mariaux JB, Oberschall A, Molnar G, Berendzen K, Shukla V, Lafos M, Koncz Z, Redei GP, Schell J, Koncz C (2002) Rapid identification of Arabidopsis insertion mutants by non-radioactive detection of T-DNA tagged genes. Plant Journal 32: 243-253

Robinson JC (1996) Bananas and Plantains. Crop Production Science in Horticulture, 5. CAB International, Wallingford, UK. 238 p.

Robinson JC, Anderson T (1991) The influence of temperature on dry matter assimilation and distribution in young banana plants. Newsletter of the International Group on Horticultural Physiology of Banana 14: 37

Rombauts S, Florquin K, Lescot M, Marchal K, Rouze P, Van De Peer Y (2003) Computational approaches to identify promoters and cis-regulatory elements in plant genomes. Plant Physiology 132: 1162-1176

Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B (2003) An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Molecular Biology 53: 247-259

Roux N, Baurens FC, Doležel J, Hřibová E, Heslop-Harrison P, Town C, Sasaki T, Matsumoto T, Aert R, Remy S, Souza M, Lagoda P (2008) Genomics of banana and plantain (*Musa* spp.), major staple crops in the tropics. In: Genomics of Tropical Crop Plants. Plant Genetics and Genomics: Crops and Models, 1. P.H. Moore and R. Ming, editors. Springer Science+Business Media, New York, NY, USA. pp. 83-111

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics Methods and Protocols. Methods in Molecular Biology, 132. S. Krawetz and S. Misener, editors. Humana Press, Totowa, NJ, USA. pp. 365-386

Ryu CH, You JH, Kang HG, Hur JH, Kim YH, Han MJ, An KS, Chung BC, Lee CH, An GH (2004) Generation of T-DNA tagging lines with a bidirectional gene trap vector and the establishment of an insertion-site database. Plant Molecular Biology 54: 489-502

Sallaud C, Gay C, Larmande P, Bes M, Piffanelli P, Piegu B, Droc G, Regad F, Bourgeois E, Meynard D, Perin C, Sabau X, Ghesquiere A, Glaszmann JC, Delseny M, Guiderdoni E (2004) High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. Plant Journal 39: 450-464

Sallaud C, Meynard D, van Boxtel J, Gay C, Bes M, Brizard JP, Larmande P, Ortega D, Raynal M, Portefaix M, Ouwerkerk PB, Rueb S, Delseny M, Guiderdoni E (2003) Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. Theoretical and Applied Genetics 106: 1396-408

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, second edition. Cold Spring Harbor Laboratory Press, NY, USA.

Sangwan V, Foulds I, Singh J, Dhindsa RS (2001) Cold-activation of *Brassica napus BN115* promoter is mediated by structural changes in membranes and cytoskeleton, and requires Ca^{2+} influx. Plant Journal 27: 1-12

Santos CMR, Martins NF, Horberg HM, De Almeida ERP, Coelho MCF, Togawa RC, Da Silva FR, Caetano AR, Miller RNG, Souza MT (2005) Analysis of expressed sequence tags from *Musa acuminata* ssp. *burmannicoides*, var. *Calcutta 4* (AA) leaves submitted to temperature stresses. Theoretical and Applied Genetics 110: 1517-1522

Santos E, Remy S, Thiry E, Windelinckx S, Swennen R , Sági L (2007) Tagging novel promoters in banana using the luciferase rerporter gene. Acta Horticulturae 763: 99-107

Sági L, May GD, Remy S, Swennen R (1998) Recent developments in biotechnological research of banana (*Musa* spp.). In: Biotechnology and Genetic Engineering Reviews. M.P. Tombs, editor. Intercept Ltd, Andover, UK. pp. 313-327

Sági L, Panis B, Remy S, Cammue BPA, Swennen R (1995c) Stable and transient genetic transformation of banana (*Musa* spp. cv. 'Bluggoe', ABB-group) using cells and protoplasts. In: Proceedings of the 11th ACORBAT Meeting. V. Morales Soto, editor. San José, Costa Rica, 13-18 February 1994. pp. 85-102

Sági L, Panis B, Remy S, Schoofs H, Desmet K, Swennen R, Cammue BPA (1995a) Genetic transformation of banana and plantain (*Musa* spp.) via particle bombardment. Biotechnology 13: 481-485

Sági L, Remy S, Panis B, Swennen R, Volckaert G (1994) Transient gene expression in electroporated banana (*Musa* spp., cv. Bluggoe, ABB group) protoplasts isolated from regenerable embryogenetic cell suspensions. Plant Cell Reports 13: 262-266

Sági L, Remy S, Panis B, Volckaert G, Swennen R (1992) Transient gene expression in banana protoplasts. Banana Newsletter 15: 42

Sági L, Remy S, Verelst B, Panis B, Cammue BPA , Volckaert G, Swennen R (1995b) Transient gene expression in transformed banana (*Musa* cv. 'Bluggoe') protoplasts and embryogenic cell suspensions. Euphytica 85: 89-95

Schatz T, Langowski J (1997) Curvature and sequence analysis of eukaryotic promoters. Journal of Biomolecular Structure and Dynamics 15: 265-275

Schenk PM, Remans T, Sági L, Elliott AR, Dietzgen RG, Swennen R, Ebert PR, Grof CPL, Manners JM (2001) Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants. Plant Molecular Biology 47: 399-412

Schenk PM, Sági L, Remans T, Dietzgen RG, Bernard MJ, Graham MW, Manners JM (1999) A promoter from sugarcane bacilliform badnavirus drives transgene expression in banana and other monocot and dicot plants. Plant Molecular Biology 39: 1221-1230

Scherf M, Klingenhoff A, Werner T (2000) Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: A novel context analysis approach. Journal of Molecular Biology 297: 599-606

Schneeberger RG, Zhang K, Tatarinova T, Troukhan M, Kwok SF, Drais J, Klinger K, Orejudos F, Macy K, Bhakta A, Burns J, Subramanian G, Donson J, Flavell R, Feldmann KA (2005) Agrobacterium T-DNA integration in Arabidopsis is correlated with DNA sequence compositions that occur frequently in gene promoter regions. Functional & Integrative Genomics 5: 240-253

Scholte M (2002) T-DNA tagging in *Medicago truncatula*. Ph.D. thesis. Wageningen University. The Netherlands. 109 p.

Scholte M, D'erfurth I, Rippa S, Mondy S, Cosson V, Durand P, Breda C, Trinh H, Rodriguez-Llorente I, Kondorosi E, Schultze M, Kondorosi A, Ratet P (2002) T-DNA tagging in the model legume *Medicago truncatula* allows efficient gene discovery. Molecular Breeding 10: 203-215

Schouten HJ, Krens FA, Jacobsen E (2006) Cisgenic plants are similar to traditionally bred plants: International regulations for genetically modified organisms should be altered to exempt cisgenesis. EMBO Reports 7: 750-753

Schünmann PHD, Richardson AE, Vickers CE, Delhaize E (2004) Promoter analysis of the barley *Pht1;1* phosphate transporter gene identifies regions controlling root expression and responsiveness to phosphate deprivation. Plant Physiology 136: 4205-4214

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashizaki Y, Kawai J, Carninci P, Itoh M, Ishii Y, Arakawa T, Shibata K, Shinagawa A, Shinozaki K (2002) Functional annotation of a full-length Arabidopsis cDNA collection. Science 296: 141-145

Sessions A, Burke E, Presting G, Aux G, Mcelver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell 14: 2985-2994

Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV (2003) PlantProm: a database of plant promoter sequences. Nucleic Acids Research 31: 114-117

Shahmuradov IA, Solovyev VV, Gammerman AJ (2005) Plant promoter prediction with confidence estimation. Nucleic Acids Research 33: 1069-1076

Sherf BA, Wood KV (1994) Firefly luciferase engineered for improved genetic reporting. Promega Notes Magazine 49: 14-21

Shinwari ZK, Nakashima K, Miura S, Kasuga M, Seki M, Yamaguchi-Shinozaki K, Shinozaki K (1998) An Arabidopsis gene family encoding DRE/CRT binding proteins involved in low-temperature-responsive gene expression. Biochemical and Biophysical Research Communications 250: 161-170

Simpson GG, Filipowicz W (1996) Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organisation of the spliceosomal machinery. Plant Molecular Biology 32: 1-41

Sivanandan C, Sujatha TP, Prasad AM, Resminath R, Thakare DR, Bhat SR, Srinivasan (2005) T-DNA tagging and characterization of a cryptic root-specific promoter in Arabidopsis. Biochimica et Biophysica Acta-Gene Structure and Expression 1731: 202-208

Smale ST (1994) Core promoter architecture for eucaryotic protein-coding genes. In: Transcription: mechanisms and regulation. R. C. Conaway and J. W. Conaway, editors. Raven Press, Ltd., New York, NY, USA. pp. 63-81

Smale ST (1997) Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes.

Biochimica et Biophysica Acta-Gene Structure and Expression 1351: 73-88
Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. Annual Review of Biochemistry 72: 449-479
Springer PS (2000) Gene traps: tools for plant development and genomics. Plant Cell 12: 1007-1020
Sreenivasulu N, Sopory SK, Kishor PBK (2007) Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. Gene 388: 1-13
Sreeramanan S, Maziah M, Abdullah MP, Sariah M, Xavier R (2006c) Transient expression of *gusA* and *gfp* gene in *Agrobacterium*-mediated banana transformation using single tiny meristematic bud. Asian Journal of Plant Sciences 5: 468-480
Sreeramanan S, Maziah M, Abdullah MP, Sariah M, Xavier R, Nor'Aini MF (2005) Physical and biological parameters affecting transient GUS and GFP expression in banana via particle bombardment. Asia Pacific Journal of Molecular Biology and Biotechnology 13: 35-57
Sreeramanan S, Maziah M, Rosli NM, Sariah M, Xavier R (2006a) Particle bombardment-mediated co-transformation of chitinase and β-1, 3 glucanase genes in banana. Biotechnology 5: 203-216
Sreeramanan S, Maziah M, Rosli NM, Sariah M, Xavier R (2006b) Enhanced tolerance against a fungal pathogen, *Fusarium oxysporum* f.sp. *cubense* (race 1) in transgenic silk banana. International Journal of Agricultural Research 1: 342-354
Stahl R, Horvath H, Van Fleet J, Voetz M, Von Wettstein D, Wolf N (2002) T-DNA integration into the barley genome from single and double cassette vectors. Proceedings of the National Academy of Sciences USA 99: 2146-2151
Stangeland B, Nestestog R, Grini PE, Skrbo N, Berg A, Salehian Z, Mandal A, Aalen RB (2005) Molecular analysis of Arabidopsis endosperm and embryo promoter trap lines: reporter-gene expression can result from T-DNA insertions in antisense orientation, in introns and in intergenic regions, in addition to sense insertion at the 5' end of genes. Journal of Experimental Botany 56: 2495-2505
Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a *cis*-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proceedings of the National Academy of Sciences USA 94: 1035-1040
Strosse H (2005) Optimisation from *in vitro* multiple meristem cultures and embryogenic cell suspensions in banana (*Musa* spp.). Ph.D. thesis 651. Katholieke Universiteit Leuven, Belgium. Faculteit Bio-ingenieurswetenschappen. 245 p.
Strosse H, Domergue R, Panis B, Escalant J-V, Côte F (2003) Banana and plantain embryogenic cell suspensions. INIBAP Technical Guidelines 8. The International Network for the Improvement of Banana and Plantain, Montpellier, France. 32 p.
Struhl K (2001) Gene regulation - A paradigm for precision. Science 293: 1054-1055
Sundaresan V, Springer P, Volpe T, Haward S, Jones JD, Dean C, Ma H, Martienssen R (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. Genes & Development 9: 1797-810
Sunil Kumar GB, Ganapathi TR, Revathi CJ, Srinivas L, Bapat VA (2005) Expression of hepatitis B surface antigen in transgenic banana plants. Planta 222: 484-493
Suzuki M, Ketterling MG, McCarty DR (2005) Quantitative statistical analysis of *cis*-regulatory sequences in ABA/VP1-and CBF/DREB1-regulated genes of Arabidopsis. Plant Physiology 139: 437-447
Swennen R, Arinaitwe G, Cammue BPA, François I, Panis B, Remy S, Sági L, Santos E, Strosse H, Van den Houwe I (2003) Transgenic approaches for resistance to *Mycosphaerella* leaf spot diseases in *Musa* spp. In: Mycosphaerella leaf spot diseases of bananas: present status and outlook. Proceedings of the 2nd International workshop on *Mycosphaerella* leaf spot diseases of bananas. L. Jacome, P. Lepoivre, D. Marin, R. Ortiz, R. Romero and J.V. Escalant, editors. San José, Costa Rica, 20-23 May 2002. INIBAP, Montpellier, France. pp. 209-238
Szabados L, Kovacs I, Oberschall A, Abraham E, Kerekes I, Zsigmond L, Nagy R, Alvarado M, Krasovskaja I, Gal M, Berente A, Redei GP, Ben Haim A, Koncz C (2002) Distribution of 1000 Sequenced T-DNA tags in the Arabidopsis genome. Plant Journal 32: 233-242
Taiz L, Zeiger E (2002) Plant Physiology, third edition. Sinauer Associates Inc. Publ. Sunderland, MA, USA. 690 p.
Tatusova TA, Madden TL (1999) BLAST 2 SEQUENCES, a new tool for comparing protein and nucleotide sequences. FEMS Microbiology Letters 174: 247-250
Taylor CB (1997) Promoter fusion analysis: An insufficient measure of gene expression. Plant Cell 9: 273-275
Teeri TH, Herrera-Estrella L, Depicker A, Van Montagu M, Palva ET (1986) Identification of plant promoters *in situ* by T-DNA-mediated transcriptional fusions to the *npt*-II gene. EMBO Journal 5: 1755-1760
The Global *Musa* Genomics Consortium (2002) A strategy for the Global *Musa* Genomics Consortium.

Report of a meeting held in Arlington, USA, 17-20 July 2001. International Network for the Improvement of Banana and Plantain, Montpellier, France. 44 p.

Thijs G, Marchal K, Lescot M, Rombauts S, De Moor B, Rouze P, Moreau Y (2002) A Gibbs sampling method to detect overrepresented motifs in the upstream regions of coexpressed genes. Journal of Computational Biology 9: 447-464

Thomashow MF (2001) So what's new in the field of plant cold acclimation? Lots! Plant Physiology 125: 89-93

Thompson JF, Hayes LS, Lloyd DB (1991) Modulation of firefly luciferase stability and impact on studies of gene regulation. Gene 103: 171-177

Tinland B (1996) The integration of T-DNA into plant genomes. Trends in Plant Science 1: 178-184

Topping JF, Agyeman F, Henricot B, Lindsey K (1994) Identification of molecular markers of embryogenesis in *Arabidopsis thaliana* by promoter trapping. Plant Journal 5: 895-903

Topping JF, Wei WB, Lindsey K (1991) Functional tagging of regulatory elements in the plant genome. Development 112: 1009-1019

Triglia T, Peterson MG, Kemp DJ (1988) A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Research 16: 8186

Trinh TH, Ratet P, Kondorosi E, Durand P, Kamate K, Bauer P, Kondorosi A (1998) Rapid and efficient transformation of diploid *Medicago truncatula* and *Medicago sativa* ssp. *falcata* lines improved in somatic embryogenesis. Plant Cell Reports 17: 345-355

Tripathi L, Tripathi JN, Hughes JD (2005) *Agrobacterium*-mediated transformation of plantain (*Musa* spp.) cultivar Agbagba. African Journal of Biotechnology 4: 1378-1383

Turner DW, Lahav E (1983) The growth of banana plants in relation to temperature. Australian Journal of Plant Physiology 10: 43-53

Ueda I, Shinoda F, Kamaya H (1994) Temperature-dependent effects of high-pressure on the bioluminescence of firefly luciferase. Biophysical Journal 66: 2107-2110

Van Buskirk HA, Thomashow MF (2006) Arabidopsis transcription factors regulating cold acclimation. Physiologia Plantarum 126: 72-80

Van Leeuwen W, Hagendoorn MJM, Ruttink T, Van Poecke R, Van der Plas LHW, Van der Krol AR (2000) The use of the luciferase reporter system for *in planta* gene expression studies. Plant Molecular Biology Reporter 18: 143a-143t

Van Leeuwen W, Ruttink T, Borst-Vrenssen AWM, Van Der Plas LHW, Van Der Krol AR (2001) Characterization of position-induced spatial and temporal regulation of transgene promoter activity in plants. Journal of Experimental Botany 52: 949-959

Venter M, Botha FC (2004) Promoter analysis and transcription profiling: Integration of genetic data enhances understanding of gene expression. Physiologia Plantarum 120: 74-83

Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF (2005) Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of Arabidopsis. Plant Journal 41: 195-211

Wang XL, Peng XX (2001a) Cloning of promoter of banana fruit-specific ACC synthase gene and primary study on its function. Chinese Journal of Biotechnology 17: 293-296

Wang XL, Peng XX (2001b) Cloning of promoter of banana fruit ripening-related ACO1 and primary study on its function. Chinese Journal of Biotechnology 17: 428-431

Warshawsky D, Miller L (1994) A rapid genomic walking technique based on ligation-mediated PCR and magnetic separation technology. Biotechniques 16: 792-794

Webb KJ, Skot L, Nicholson MN, Jorgensen B, Mizen S (2000) *Mesorhizobium loti* increases root-specific expression of a calcium-binding protein homologue identified by promoter tagging in *Lotus japonicus*. Molecular Plant-Microbe Interactions 13: 606-616

Wei WB, Twell D, Lindsey K (1997) A novel nucleic acid helicase gene identified by promoter trapping in Arabidopsis. Plant Journal 11: 1307-1314

Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia YJ, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in Arabidopsis. Plant Physiology 122: 1003-1013

Wenck A, Pugieux C, Turner M, Dunn M, Stacy C, Tiozzo A, Dunder E, Van Grinsven E, Khan R, Sigareva M, Wang WC, Reed J, Drayton P, Oliver D, Trafford H, Legris G, Rushton H, Tayab S, Launis K, Chang YF, Chen DF, Melchers L (2003) Reef-coral proteins as visual, non-destructive reporters for plant transformation. Plant Cell Reports 22: 244-251

Wood KV (1995) Marker proteins for gene expression. Current Opinion in Biotechnology 6: 50-8

Wu CH, Madabusi L, Nishioka H, Emanuel P, Sypes M, Arkhipova I, Gilmour DS (2001) Analysis of core promoter sequences located downstream from the TATA element in the hsp70 promoter from *Drosophila melanogaster*. Molecular and Cellular Biology 21: 1593-1602

Wuyts N, De Waele D, Swennen R (2006) Activity of phenylalanine ammonia-lyase, peroxidase and polyphenol oxidase in roots of banana (*Musa acuminata* AAA, cvs Grande Naine and Yangambi km5) before and after infection with *Radopholus similis*. Nematology 8: 201-209

Xiong LM, Lee BH, Ishitani M, Lee H, Zhang CQ, Zhu JK (2001) FIERY1 encoding an inositol polyphosphate 1-phosphatase is a negative regulator of abscisic acid and stress signaling in Arabidopsis. Genes & Development 15: 1971-1984

Xiong LM, Schumaker KS, Zhu JK (2002) Cell signaling during cold, drought, and salt stress. Plant Cell 14: S165-S183

Yamaguchi-Shinozaki K, Shinozaki K (1994) A novel *cis*-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt Stress. Plant Cell 6: 251-264

Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. Trends in Plant Science 10: 88-94

Yamamoto YY, Ichida H, Abe T, Suzuki Y, Sugano S, Obokata J (2007b) Differentiation of core promoter architecture between plants and mammals revealed by LDSS analysis. Nucleic Acids Research 35: 6219-6226

Yamamoto YY, Ichida H, Matsui M, Obokata J, Sakurai T, Satou M, Seki M, Shinozaki K, Abe T (2007a) Identification of plant promoter constituents by analysis of local distribution of short sequences. BMC Genomics 8: 67

Yamamoto YY, Obokata J (2007c) ppdb: a plant promoter database. Nucleic Acids Research 36: D977-D981

Yamamoto YY, Tsuhara Y, Gohda K, Suzuki K, Matsui M (2003) Gene trapping of the *Arabidopsis* genome with a firefly luciferase reporter. Plant Journal 35: 273-283

Yang IC, Iommarini JP, Becker DK, Hafner GJ, Dale JL, Harding RM (2003) A promoter derived from taro bacilliform badnavirus drives strong expression in transgenic banana and tobacco plants. Plant Cell Reports 21: 1199-1206

Yang LT, Ding JY, Zhang CM, Jia JW, Weng HB, Liu WX, Zhang DB (2005) Estimating the copy number of transgenes in transformed rice by real-time quantitative PCR. Plant Cell Reports 23: 759-763

Yanofsky MF, Porter SG, Young C, Albright LM, Gordon MP, Nester EW (1986) The virD operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. Cell 47: 471-477

Yoo SY, Bomblies K, Yoo SK, Yang JW, Choi MS, Lee JS, Weigel D, Ahn JH (2005) The 35S promoter used in a selectable marker gene of a plant transformation vector affects the expression of the transgene. Planta 221: 523-530

Zarka DG, Vogel JT, Cook D, Thomashow MF (2003) Cold induction of Arabidopsis *CBF* genes involves multiple ICE (Inducer of *CBF* expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. Plant Physiology 133: 910-918

Zhang JL, Byrne CD (1999) Differential priming of RNA templates during cDNA synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase PCR. Biochemical Journal 337: 231-241

Zhang MQ (1998) Identification of human gene core promoters *in silico*. Genome Research 8: 319-326

Zhou DX, Li YF, Rocipon M, Mache R (1992) Sequence-specific interaction between S1F, a spinach nuclear factor, and a negative *cis*-element conserved in plastid-related genes. Journal of Biological Chemistry 267: 23515-23519

Zhu Q, Dabi T, Lamb C (1995) TATA box and initiator functions in the accurate transcription of a plant minimal promoter *in vitro*. Plant Cell 7: 1681-1689

Zhu QH, Ramm K, Eamens AL, Dennis ES, Upadhyaya NM (2006) Transgene structures suggest that multiple mechanisms are involved in T-DNA integration in plants. Plant Science 171: 308-322

List of publications

Santos E, Remy S, Thiry E, Windelinckx S, Swennen R, Sági L (2007) Tagging low-temperature responsive promoters in banana using the luciferase reporter gene. Acta Horticulturae 763: 99-107

Swennen R, Arinaitwe G, Cammue BPA, François I, Panis B, Remy S, Sági L, **Santos E**, Strosse H, Van den houwe I (2003) Transgenic approaches for resistance to Mycosphaerella leaf spot diseases in *Musa* spp. L. Jacome, P. Lepoivre, D. Marin, R. Ortiz, R. Romero and J.V. Escalant, editors. Mycosphaerella leaf spot diseases of bananas: present status and outlook. Proceedings of the 2nd International workshop on Mycosphaerella leaf spot diseases of bananas. San José, Costa Rica, 20-23 May 2002. INIBAP, Montpellier, France: 209-238

Santos E, Remy S, Thiry E, Windelinckx S, Swennen R, Sági L (in revision) Characterization and isolation of a T-DNA tagged banana promoter active during *in vitro* regeneration and low temperature stress. Journal of Experimental Botany

Annex

Annex. *In silico* analysis of T-DNA flanking sequences in promoter-tagged lines. The accession with the highest gene homology is shown if the E value is below 1 x 10^{-18} and is indicated by the genus and species, gene/clone, E value and accession number. Analysis was performed in the period comprising July 2005 until February 2007

						Homology search				***cis*-acting elements/promoter search**[a]			
Line	**T-DNA copy number**[b]	**Border**	**ID**[c]	**Isolation method**	**Size (bp)**	**BlastN**[d]	**Fasta3**[e]	**BlastN EST**[f]	**Banana EST**[g]	**PlantCare**	**PLACE**	**TSSP**	**Remark**
17	5	RB	1	I-PCR BsrGI TAIL-PCR AD2-1	2438 1440	No	No	No	No	ABRE, DRE, TATA	ABRE, DRE, TATA	Yes (2)	One ICEr1-like element[h]
			2	TAIL-PCR AD2-1	1275	No	No	No	No	DRE, TATA	DRE, TATA	No	
			3	TAIL-PCR AD2-1 TAIL-PCR AD2 I-PCR BsrGI	731 647 699	No	No	No	No	No	No	No	
			4	TAIL-PCR AD2-1	266	No	No	No	No	TATA	No	No	
			5	TAIL-PCR AD2	859	NA	NA	NA	NA	NA	NA	NA	vector backbone
			A	TAIL-PCR AD2	~2500	*Musa x paradisiaca*, CONSTANT-like protein gene, E=2e-20, DQ153049	*Oryza sativa*, Hd1 gene, E=1.2e-28, BD094727	*Malus x domestica*, cDNA clone, E=4e-31, EB176239	No	NA	NA	No	Non-specific sequence[h]
			B	TAIL-PCR AD2-1	1235	No	No	No	No	ABRE, DRE, TATA	DRE, TATA	No	Non-specific sequence[i]
		LB	1	TAIL-PCR AD5 TAIL-PCR AD5 TAIL-PCR AD2	1683 1044 512	No	No	No	No	NA	NA	No	Sequence isolated from the 5'-tagged sequence 1 of line 17
			2	I-PCR BsrGI TAIL-PCR AD2 TAIL-PCR AD2 TAIL-PCR AD2-1 I-PCR BclI	1165 ~2000 1122 1086 291	No	No	No	E=1e-135, 600092615T1	NA	NA	No	
			3	TAIL-PCR AD	932	No	No	No	No	NA	NA	No	Sequence isolated from the 5'-tagged sequence 3 of line 17
			4	I-PCR BclI TAIL-PCR AD2-1	727 504	No	No	No	No	NA	NA	No	
			5	TAIL-PCR AD2	630	NA	NA	NA	NA	NA	NA	No	pe35S-*neo* sequence

Annex. (continued)

						Homology search				*cis*-acting elements/promoter search[a]			
Line	**T-DNA copy number[b]**	**Border**	**ID[c]**	**Isolation method**	**Size (bp)**	**BlastN[d]**	**Fasta3[e]**	**BlastN EST[f]**	**Banana EST[g]**	**PlantCare**	**PLACE**	**TSSP**	**Remark**
34	1	RB	1	TAIL-PCR AD2-1 TAIL-PCR AD2-5 TAIL-PCR AD2	1298 969 532	No	No	No	No	DRE, TATA	ABRE, DRE, TATA	No	
		LB	1	TAIL-PCR AD2-5 TAIL-PCR AD2-5	1494 698	*Musa acuminata*, MuG9, E=6e-103, AY484588	*Musa acuminata*, MuG9, E=2.1e-114, AY484588	*Musa acuminata*, HOT stress cDNA clone, E=3e-22, DN238884.	E= 2e-22, 600170679T1	NA	NA	NA	Homologous to partial region of the 5'-tagged 179-2
			A	IPCR-BclI	1144	*Musa acuminata* retrotransposon monkey, E=0.0, AF143332	*Musa acuminata* retrotransposon monkey, E=0.0, AF143332	No	NT	NA	NA	NA	Non-specific sequence[i]
85	3	RB	1	TAIL-PCR AD2	625	No	No	No	No	TATA	ABRE, TATA	No	One ICEr1-like element[h]
			2	I-PCR BsrGI	386	*Oryza sativa*, mRNA Auxin response factor 10, E=5e-19, AB071299	*Oryza sativa*, mRNA Auxin response factor 10, E=2.8e-27, AB071299	No	No	No	No	No	
			3	TAIL-PCR AD2	598	No	No	No	No	TATA	TATA	No	
		LB	1	TAIL-PCR AD2-1	561	No	*Medicago truncatula*, clone mth, E=5.6e-38, AC141115	*Persea americana*, cDNA clone, E=4e-22, CK758727	No	NA	NA	NA	
			2	TAIL-PCR AD2	1184	No	No	No	No	NA	NA	NA	
			4	TAIL-PCR AD2	1596	No	No	No	No	NA	NA	NA	
			5	TAIL-PCR AD2	1773	No	*Oryza sativa*, cDNA clone, E=6.6e-69, CT830291	No	No	NA	NA	NA	
			A	TAIL-PCR AD2	1185	No	No	No	NT	NA	NA	NA	Non-specific sequence[i]

Annex. (continued)

						Homology search				*cis*-acting elements/promoter search[a]			
Line	**T-DNA copy number[b]**	**Border**	**ID[c]**	**Isolation method**	**Size (bp)**	**BlastN[d]**	**Fasta3[e]**	**BlastN EST[f]**	**Banana EST[g]**	**PlantCare**	**PLACE**	**TSSP**	**Remark**
156	5	RB	1	TAIL-PCR AD2-1	355	NA	NA	NA	NA	NA	NA	NA	T-DNA tandem integration
			2	I-PCR BsrGI	718	No	No	No	NT	DRE, TATA	DRE, TATA	No	
			3	TAIL-PCR AD2-5	445	NA	NA	NA	NA	NA	NA	NA	pe35S-*neo* sequence
		LB	1	I-PCR BsrGI	1583	*Agrobacterium tumefaciens*, insertion element IS113, E=0.0, ATUIS22A	*Agrobacterium tumefaciens*, insertion element IS113, E=2.7e-207, M82888	No	No	NA	NA	NO	Tandem integration plus insertion of IS1131 element (715 bp) from *Agrobacterium tumefaciens*
			2	I-PCR BsrGI	1829	*Rhizobium leguminosarum*, DNA for repA, repB and repC genes, E= 3e-61 X89447	*Rhizobium leguminosarum*, DNA for repA, repB and repC genes, E=4.2e-208 X89447	No	No	NA	NA	NO	Not fully sequenced
			3	I-PCR BclI I-PCR BsrGI	803 700	NA	NA	NA	NA	NA	NA	NA	T-DNA tandem integration, same as the 5'-tagged seq. 156-1
179	4	RB	1	TAIL-PCR AD2-5	501	No	No	No	No	No	ABRE	NO	
			2	I-PCR BclI TAIL-PCR AD2-1 TAIL-PCR AD2 I-PCR BsrGI	1648 1168 880 186	*Musa acuminata* MuG9, E=5e-121, AY484588	*Musa acuminata*, MuG9, E=3.6-86, AY484588	*Musa acuminata*, E=1e-81, DN238884	4E-67, 600178185T1	ABRE, DRE	ABRE, DRE		Two ICEr1-like element[h] Homologous to partial region of the 5'-tagged seq. 34-1
			3	I-PCR BclI TAIL-PCR AD2-1	705 540	*Musa acuminata*, MuG9, E=1e-28, AY484588	*Musa acuminata*, MuG9, E=8.2e-66, AY484588	No	No	TATA	ABRE, TATA	NO	
			4	I-PCR BclI TAIL-PCR AD2-5	1061 659	No	No	No	No	No	ABRE, DRE	No	One ICEr1-like element[h]
		LB	1	I-PCR BclI	413	NT	NT	NT	NT	NT	NT	NT	vector backbone
			A	I-PCR BclI	~2000	NT	NT	NT	NT	NT	NT	NT	Non-specific sequence[i]

[a]T-DNA flanking sequences were queried for the presence of, ABRE- and DRE-like elements and TATA boxes using the PlantCARE and PLACE software, while the TSSP software was used for the identification of promoters.

[b]Minimun number of T-DNA copies determined by Southern blot hybridization using a DIG-labeled *luc*$^+$ probe.

[c]Sequence identification number of promoter-tagged lines. Non-specific sequences are identified with a letter.

[d]T-DNA flanking sequences were analyzed with blastn and blastx in the GeneBank database "nr" (all GenBank + RefSeq Nucleotides + EMBL + DDBJ + PDB sequences, excluding HTGS0,1,2, EST, GSS, STS, PAT, WGS) to all organisms.

[e]Fasta3 software was used for homology search for the T-DNA flanking sequences.

[f]T-DNA flanking sequences were analyzed using the blastn in the "EST" database (GenBank + EMBL + DDBJ sequences from EST Divisions).

[g]T-DNA flanking sequences were analyzed with blastn in a banana EST database donated by Syngenta to the Global *Musa* Genomics Consortium. [h]The core sequence of the ICEr1 element (CACATG) was located manually. [i]Non-specifc sequences were catalogued to T-DNA flanking sequences lacking the pETKUL2 specific primers annealing site.

NT refers to not tested and NA refers to not applicable, while pe35S-*neo* indicates the presence of the enhanced 35S promoter and truncated neomycin phosphotransferas II selectable marker gene.