



CHAPTER ONE

INTRODUCTION

1.1 The banana plant

Banana (*Musa* spp.) belongs to the order Zingiberales, family Musaceae, section Eumusa and genus *Musa* (Simmonds, 1962; Purseglove, 1972; Stover and Simmonds, 1987). A banana plant is a tree-like, giant perennial tropical monocotyledonous herb of two to nine meters tall (Fig. 1.1). The plant has a pseudostem composed of compact overlapping leaf sheaths. The basal stem is condensed (corm) and only elongates through the pseudostem at flowering to eject the inflorescence. Cultivated bananas have descended from two wild seeded forms, *Musa acuminata* Colla (AA; $2n = 22$) and *M. balbisiana* Colla (BB; $2n = 22$) to give rise to AA, AB, AAA, AAB and ABB genomic composition (Simmonds, 1962; Purseglove, 1972; Stover and Simmonds, 1987). Triploids resulted from fertilization of egg cells with unreduced chromosome, caused by chromosome restitution, in one diploid parent by haploid pollen from another parent. With the exception of a few wild and fertile diploids and in conventional breeding work where seed set and germination can occur, banana plants propagate vegetatively through suckers that arise from the corm of the mother plant (Purseglove, 1972; Stover and Simmonds, 1987).

1.2 Banana breeding

1.2.1 Conventional breeding

Along with diseases, such as black Sigatoka and Fusarium wilt resistance and yield, early maturity and root system development have been traits of interest to classical banana breeders for improving banana productivity (Stover and Buddenhagen, 1986; Vuylsteke *et al.*, 1995; Gowen, 1996; Tenkouano *et al.*, 1998). However, no conventional breeding scheme has been developed to address these traits. Instead, fast maturity has been a secondary selection parameter in hybrid development (Vuylsteke *et al.*, 1993; Tenkouano *et al.*, 1998; Ortiz and Vuylsteke, 1998a; 1998b; Ortiz *et al.*, 1998). Differences in maturity time due to ploidy levels has been reported where diploids have a short fruit filling duration than triploid (Swennen and Vuylsteke, 2001). However, diploids are unattractive for large scale and commercial cultivation due to their low yields and growth vigour (Swennen and

Vuylsteke, 2001). In addition, diploids have been developed through human selection for their parthenocarpic fruits and seed sterility (Simmonds, 1962). Therefore, their biology restricts the introgression of their fast maturity trait into other banana variety through crossing breeding. The majority of the cultivated genotypes that constitute the popular varieties are triploids and highly sterile which limits their improvement through conventional breeding (Crouch *et al.*, 1989). In the few cases where seeds have been obtained, the low germinations limit the recovery of hybrids (Vuylsteke *et al.*, 1993; Ssebuliba *et al.*, 2005). Among the fertile banana cultivars (Ssebuliba *et al.*, 2005), the variability in their maturity time is not significant (Makumbi, 1995).

Application of cross breeding in banana is still limited by factors such as the low seed and pollen fertility, triploidy levels of the economically important cultivars, slow propagation, and long generation time from one generation to the next (Vuylsteke *et al.*, 1993; Crouch *et al.*, 1998; Sweenen and Vuylsteke, 2001; Ssebuliba *et al.*, 2005). Where hybrids have been recovered failure of the fruits to retain the desirable quality has been a drawback (Purseglove 1972; Pillay *et al.*, 2004).

1.2.2 Genetic engineering of banana

Genetic transformation has been used to introgress important agronomic traits in a number of monocotyledonous crops (Cheng *et al.*, 2004). In banana, this strategy has a potential for traits lacking in the *Musa* germplasm and for seed sterile cultivars (Crouch *et al.*, 1998) and the potential has been demonstrated for resistance against bacterial wilt (Tripathi *et al.*, 2010). However, employing the technique is limited by the complex gene interaction that influences quantitative traits (Crouch *et al.*, 1998). Gene transfer approaches and explants that have been used in banana include electroporation of protoplasts (Sagi *et al.*, 1994), particle bombardment of embryogenic cells (Sagi *et al.*, 1995; Becker *et al.*, 2001) and co-cultivation of *Agrobacterium tumefaciens* with embryogenic cells (Ganapathi *et al.*, 2001; Khanna *et al.*, 2004; Talengera *et al.*, 2010; Tripathi *et al.*, 2010). *Agrobacterium* has also been used on wounded corm discs (May *et al.*, 1995; Tripathi *et al.*, 2008). Of the gene delivery systems, *Agrobacterium* is prominent for its simplicity and delivery of intact

transgenes and low copy numbers and low incidences of transgene silencing (Dai *et al.*, 2001; Hu *et al.*, 2003; Veluthambi *et al.*, 2003). In banana, *Agrobacterium* is reported to give higher transformation frequencies and lower number of integration sites compared to particle bombardment (Arinaitwe, 2008). The use of corm discs as transformation explants can give results in a shorter time when compared to embryogenic cells (Tripathi *et al.*, 2008), but it carries the risk of generating chimeric transgenics (Khalil *et al.*, 2002; Jain, 2011). Strosse *et al.* (2004) noted that plants regenerated from embryogenic cell suspension (ECS) are from single cells. However, generating banana ECS is a relatively long process and is highly genotype and cultivar dependent (Strosse *et al.*, 2003 and 2004).

1.2.3 Factors influencing transgene expression

Expression of a transgene can be regulated by the strength and tissue specificity of a promoter used in the transgene construct (Potenza *et al.*, 2004; Dietz-Pfeilstetter, 2010) and by inclusion of an intron into the gene construct (Rose and Baliakoff, 2000; Cheng *et al.*, 2004). However, an integrated transgene can be silenced following failure at the transcription level caused by DNA-methylation commonly induced by promoter homology (Park *et al.*, 1996; Stam *et al.*, 1997). Similarly, silencing can be post-transcriptional where transcription occurs at a low level or its mRNA fails to accumulate as a result of high homology of the transgene with a native gene (Matzke and Matzke, 1995; Stam *et al.*, 1997; Kanno *et al.*, 2000). High homology of the transgene and the endogenous gene can lead to the silencing of both genes (Stam *et al.*, 1997). Strong promoters and multiple transgene copies lead to high transcript levels that result into gene silencing when a threshold level is exceeded (Stam *et al.*, 1997). A transgene integrated in a head to head and tail to tail can further induce methylation or cause DNA to DNA or RNA to DNA pairing that induces their degradation in the cell (Stam *et al.*, 1997; Lakshminarayan *et al.*, 2000). Due to the random integration of transgenes, the site of integration influences the extent of transcription (Park *et al.*, 1996; Stam *et al.*, 1997; Muskens *et al.*, 2000; Yoshida and Shinmyo, 2000; Gelvin, 2003; Filipecki and Malepszy, 2006; Qi and John, 2007). Transgenes integrated within a methylated chromosome area are more likely to get methylated and silenced. Genes inserted within a CG rich chromosome region or high gene

density regions will be easily transcribed while those in A/T rich areas are prone to silencing. According to Meyer (1995) and Down *et al.* (2001), transgene expression is also influenced by prevailing growth environment of the transgenic plants.

1.3 Banana growth and development

Where soil fertility, disease and pests are not critical constraints, fast maturity has been identified as the most important aerial growth trait, next to bunch size, in determining banana productivity (Stover, 2000). Maturity time determines the duration of the production cycle that comprise the vegetative growth phase and reproductive phases.

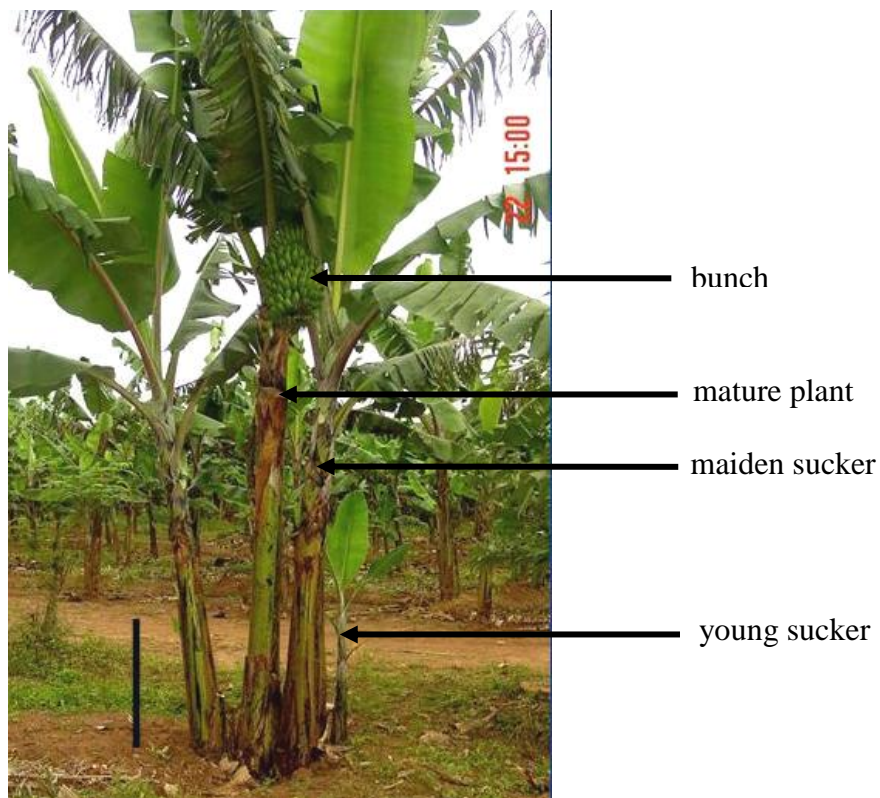


Fig.1.1 Growth cycle of a banana plant. Numbers show the successive vegetative phases before the plant flowers. Bar on the left = 1m

The vegetative phase covers the duration a plant takes to flower and shooting interval between ratoon crops while the reproductive phase covers the fruit filling time (Tenkouano *et al.*, 1998). From a field germplasm evaluation study conducted in Uganda (Makumbi, 1995), bananas take 13-17 months from sucker emergence to flowering, and 4-5 months for the fruits to mature. Surveys in East Africa (Karamura *et al.*, 2004; Gaidashova *et al.*, 2005) and West Africa (Efenden *et al.*, 2003) established maturation rates as one of the criteria used by farmers to select cultivars to grow. The relatively long cropping duration of banana was identified as disincentive to farmers' adoption rate of improved production technologies, such as fertilizer, when they consider banana growing as a long-term investment (Van Asten *et al.*, 2004). Similarly, long cropping duration of banana has also been a drawback to conventional breeding efforts due to the more than 2 years required generating a crop cycle from a seed (Crouch *et al.*, 1998; Pillay *et al.*, 2002).

1.3.1 Banana leaf and root growth and productivity

Physiological studies of banana have correlated leaf production with the plant development, architecture and yield. Depending on the cultivar, floral initiation in banana is induced after the plant has emitted a specific number of leaves (Stover and Simmonds, 1987; Swennen and De Langhe, 1985, Swennen and Vuylsteke, 2001). Further, growth studies in banana have positively correlated leaf emergence and growth rates to faster growth and bunch weight (Swennen and De Langhe, 1985). In addition, Swennen and De Langhe (1985) observed a positive correlation between leaf emergence rates, growth rates and early flowering in plantains (AAB). Because the banana pseudostem consists of overlapping leaf sheaths, foliar growth has a direct influence plant height and circumference of the pseudostem (Stover and Simmonds, 1987).

The root system determines the plant's ability to obtain both water and mineral nutrients from the soil (Taiz and Zeiger, 2006). In banana, a positive relationship between root development and aerial growth has been recorded (Blomme *et al.*, 2001; Lecompte *et al.*, 2002). Because banana plants are characterized by a shallow root system (Purseglove, 1972; Stover and Simmonds 1987), an improved root system is a desirable trait against

toppling and nematode damage (Tenkouano *et al.*, 1998). However, breeding for root growth is not feasible with the available conventional approach.

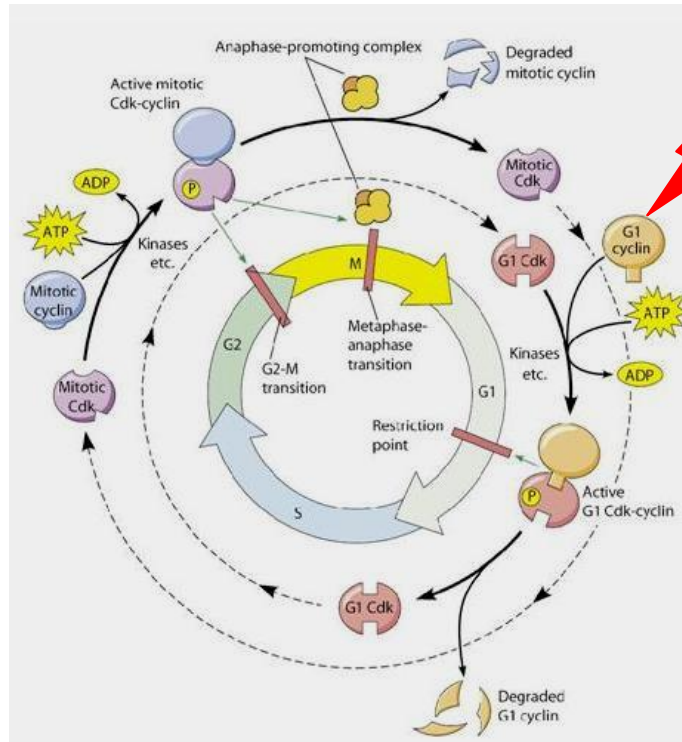
1.3.2 Biotechnology and banana plant architecture

Mutagens, such as gamma rays, can induce genetic variability in plants from which rare traits can be selected for (Van Harten, 1998). This approach has generated a fast maturing Grand Naine banana cultivar “Novaria” (Roux, 2004). However, useful mutation occurs at a very low frequency and can affect other non targeted genes (Van Harten, 1998). In light of these limitations, genetic transformation that is considered to be precise in modifying traits of interest without altering other plant characteristics (Sharma and Ortiz, 2000) remains a desirable option. In particular, over-expression of Arabidopsis *CyclinD2;1* gene in tobacco gave transgenic plants with faster shoot and root growth, faster leaf emission and earlier flowering (Cockroft *et al.*, 2000; Boucheron *et al.*, 2005) and this cyclin type has also been used in this study to possibly change the banana architecture. *CyclinD*-type is a protein sub unit that plays a major regulatory role of the cell cycle progression at the G1/S transition (Meijer and Murray, 2000; Shen, 2002; Inzé and De Veylder, 2006; Brooker, 2009; Fig. 1.2).

1.4 Plant cell cycle

Cells reproduce themselves through the process of cell division cycle or cell cycle (Taiz and Zeiger, 2006). In plants, growth and development is achieved through regulated cell division, cell expansion and differentiation (De Veylder *et al.*, 2003; Taiz and Zeiger, 2006; Brooker, 2009). The phases of the cell cycle and their molecular control mechanism are widely documented (Sorrell *et al.*, 1999; Black and Azizkhan-Clifford, 1999; Meijer and Murray, 2000; Shen, 2002; Dewitte and Murray, 2003; Inzé and De Veylder, 2006; Brooker, 2009; Wayne *et al.*, 2009; Fig. 1.2). According to these authors, decision for cells to resume division from a resting state (G₀) or to continue division or to differentiate is made at the transition of Gap1 (G₁) and synthesis (S) phase of the cell cycle. The initiation of the cell cycle at these points is regulated by the Cyclin Dependent Kinase (CDK)

proteins whose activity is modulated by the cyclin protein sub units. In the presence of CDK-activating kinases (CAKs), cyclins complexes with Cyclin Dependent Kinase (CDK) catalytic subunits to activate the phosphorylation process of the E2F transcription factor-retinoblastoma protein (E2F/Rb) dimers. Phosphorylated RbR is demobilized from promoter-bound Rb/E2Fa-DPa hetero-dimer transcription factor complex allowing the free E2Fa transcription factor to bind to the promoters of genes that are responsible for the S-phase entry phase. Among the transcribed genes are topo-isomerase that relaxes chromosomal strands, acetyltransferase responsible for the acetylation of the histone protein that causes chromatin relaxation and DNA polymerase (Vlieghe *et al.*, 2003; Brooker, 2009). By contrast, the E2F-RbR pathway is inhibited by E2Fc through its competitive inhibiting the DPa subunit (Vannesta *et al.*, 2005). At the late S-phase and G₂-to M-phase cyclins A- and B- types form complexes with CDKB, respectively (Soni *et al.*, 1995; Mironov *et al.*, 1999). The activated CDKB directly phosphorylate the histone proteins leading to chromosome condensation and the lamins that lead to disintegration of the nuclear membrane (Brooker, 2009). Cyclin D-type protein sub units effect intracellular transduction of signals as they pertain to the environment, mitogenic and nutrient status that influences the cell cycle progression (Weinert, 1998; Stals and Inzé 2001; Inzé, 2005; Inzé and De Veylder, 2006; Menges *et al.*, 2007). Progression of the cell cycle is regulated so as to match cell division with the appropriate environment, to prevent replication of cells with damaged genomes and to regulate growth and development (Dewitte and Murray, 2003).



Overexpression
of *CyclinD2;1*

Fig. 1.2 Eukaryote cell cycle. G1: Gap 1 phase; S: Synthesis phase; G2: Gap 2 phase; M: Mitosis phase. Cdk: Cyclin dependent kinase; ATP: Adenosine triphosphate; ADP: Adenosine monophosphate. Arrow indicates the intervention of this study. (Source: Wayne *et al.*, 2009).

1.4.1 Classification of cyclins

From genome-wide analysis of Arabidopsis, rice and maize Cyclin A, B, C, D, F, H, J18, L, P, SDS, T, and U have been identified, with cyclins A, B, C, H and L being shared between the plants and animals (Wang *et al.*, 2004; Hu *et al.*, 2010). Cyclins A, B and D are the most important in mitotic cell division of plants (Reanudin *et al.*, 1996; Wang *et al.*, 2004). Cyclins contain a cyclin_N and cyclin_C domain; the cyclin_N domain that comprise the CDK binding site is more conserved (Wang *et al.*, 2004; Hu *et al.*, 2010).

1.4.2 Regulation of cyclins

Cyclins are regulated through synthesis and destruction. The transcription of cyclin A and B is induced by auxins, cytokinins and gibberellin acid (Stals and Inzé, 2001). Synthesis of cyclins D1, D2, D3 and D4 is induced by sucrose while cyclin D1 and D3 are stimulated by both sucrose and cytokinins (De Veylder *et al.*, 1999; Riou-Khamlichi *et al.*, 1999; Riou-Khamlichi *et al.*, 2000; De Veylder *et al.*, 2003). CyclinD3 is induced by auxin, cytokinin, gibberellins and brassinosteroids (Rossi and Vitoro, 2001). CyclinD2 is sequentially activated earlier than cyclinD3 in the cell cycle (Riou-Khamlichi *et al.*, 2000; Masubelele *et al.*, 2005). For their deregulation, cyclin A- and B- type contain a destruction box while some D- type cyclins have PEST motifs in their sequences that facilitate their degradation by ubiquitin-mediated proteolysis (Genschik *et al.*, 1998; DeSalle and Pagano, 2001; Vandepoele *et al.*, 2002).

Cell cycle progression can also be modulated by proteins that inhibit cyclin/CDK association or direct kinase activity. Inhibitor of cdc2 kinase (ICK1) or KIP related proteins (KRP) interfere with the formation of CDK/Cyclin complex (Oakafull *et al.*, 2002, Stals *et al.*, 2000). This inhibits the CDK kinase activity with a resultant reduction of cell division and decline in plant organ development (Wang *et al.*, 2000; Barroco *et al.*, 2006). In Arabidopsis, the two versions of CKI interact with CDKA (Wang *et al.*, 1998) while the monocot ICKs/KPR inhibits both A- and D-type cyclin/CDK complexes (Coelho *et al.*, 2005). Abscisic acid (ABA) is reported to inhibit cell proliferation through upregulating

ICK1 transcription (Wang *et al.*, 1998). Alternatively, Wee1 kinase has an inhibitory phosphorylation effect on the mitotic CDK activity that results into endoreduplication (Sun *et al.*, 1999).

1.4.3 Role of the cell cycle in plant development

Plant developmental studies have correlated organ growth rate to the proliferation rate of the cells, their rate and duration of the expansion (MacAdam *et al.*, 1989; Bultynck *et al.*, 2003; Beemster *et al.*, 2003; Beemster *et al.*, 2005). The rate of cell production is determined by the meristem size and rate of cell division (Beemster and Baskin, 1998; Fiorani *et al.*, 2000). Environmental stress factors have been shown to reduce plant growth through shortening the meristem size and prolonging the cell cycle duration (Granier *et al.*, 2000; West *et al.*, 2004), and by reducing cell division and expansion (Kavanova *et al.*, 2006). The cell duration is prolonged by arresting the cell cycle at the G1 phase (Granier and Tardieu, 1999). On the contrary, the basis of enhanced growth of cyclinD transgenics has been attributed to shortened cell cycle duration and resultant increased rate of cell division (Cockroft *et al.*, 2000).

1.4.4 Cyclin expression studies

Several over-expression studies of cyclins have been conducted to evaluate their effect on cell proliferation and plant development. Cases where the transgenes contributed to the phenotypes are summarized in Table 1.1. Over-expression of *CyclinB1;1* and *B1;2* enhanced cell proliferation that resulted into faster root growth in *Brassica* (Doerner *et al.*, 1996) and in rice (Lee *et al.*, 2003). In Arabidopsis, over-expression of *Arath;CyclinB1;2* induced cell division with a reduction in ploidy in trichomes (Schnittger *et al.*, 2002). Tobacco BY-2 cells transformed with *CyclinD1* from *Antirrhinum majus* exhibited accelerated cell cycle (Koroleva *et al.*, 2004). Upregulation of *Arath;CyclinD2;1* and *CyclinD3;1* in Arabidopsis increased cell proliferation but at the expense of differentiation (Dewitte *et al.*, 2003). *Arath;CyclinD1;1* and *D2:1* increased cell proliferation in germinating Arabidopsis seeds with concomitant faster germination (Masubelele *et al.*,

2005). Similarly, over-expression of wheat *CycD2;1* in Arabidopsis resulted in increased cell proliferation and taller plants but the leaves were curled (Wang *et al.*, 2006). In contrast, over-expression of Arabidopsis and tobacco *CycD2;1* resulted into plants with faster root and shoot growth, faster leaf emission and early flowering (Cockroft *et al.*, 2000; Boucheron *et al.*, 2005; Guo and Wang, 2008). Although the same cyclin enhanced shoot and root growth in rice *in vitro* (Oh *et al.*, 2008), the effect was not evident in potted plants. Due to these contrasting reports, it is currently unpredictable if a particular overexpression construct of *cyclinD2* will yield growth stimulation in banana. Therefore, this study was conducted to evaluate the growth stimulating role by *cyclinD2* over-expression.

Banana plants reproduce continuous through suckering consequently allowing flowering to happen on each mat throughout the year. This growth characteristic provides a continuous and stable source of food when the annual crops are out of production. And on the basis that productivity in banana is measured in terms of fruit weight per unit area per unit time (Stover, 2000; Hauser and Van Asten, 2010), accelerating banana plant growth and development will reduce the time a plant takes to flower, increase the rate of ratooning and offer more harvestable bunches in a year. Sweenen and De Langhe (1985) observed a positive relationship between growth, leaf emergence rates and early flowering in bananas. More practical, surveys carried out on banana production in Uganda (Karamura *et al.*, 2004), Rwanda (Gaidashova *et al.*, 2005) and Camerron (Efenden *et al.*, 2003) identified the desire by farmers for early maturing banana cultivars. However, lack of a breeding scheme for faster growing bananas coupled with the low seed fertility and transfer of undesired characteristics of parental clones to hybrids preclude the application of the conventional crossing method to breed for quick maturing bananas. Given these limitations, it is imperative to evaluate genetic engineering of banana plants where a single trait can be introgressed while preserving other plant characteristic.

Table 1.1 Over-expression studies of cyclins and resultant phenotypes

Component	Donor plant	Expressing plant	Phenotype modification	References
<i>CyclinB;1</i>	<i>Arabidopsis thaliana</i>	<i>A. thaliana</i>	Enhanced root growth	Doerner <i>et al.</i> , 1996.
<i>CyclinB2;2</i>	Rice	Rice	Accelerated root growth	Lee <i>et al.</i> , 2003
<i>CylinD2;1</i>	<i>A. thaliana</i>	<i>Nicotiana tabacum</i>	Accelerated shoot and root growth <i>ex vitro</i>	Cockcroft <i>et al.</i> , 2000
<i>CylinD2;1</i> and <i>CycclinD3;1</i>	<i>A. thaliana</i>	<i>N. tabacum</i>	Enhanced leaf production rate and plant height	Boucheron <i>et al.</i> , 2005
<i>CylinD2;1</i>	<i>Triticum aestivum</i>	<i>A. thaliana</i>	Retarded plant growth and curled plants	Wang <i>et al.</i> , 2006
<i>CylinD2;1</i>	<i>A. thaliana</i>	<i>Oryza sativa</i>	Enhanced shoot and root growth <i>in vitro</i> .	Oh <i>et al.</i> , 2008
<i>CylinD3;1</i>	<i>Arabidopsis thaliana</i>	<i>Arabidopsis thaliana</i>	Accelerated cell proliferation but low differentiation leading to retarded growth	Dewitte <i>et al.</i> , 2003
<i>CylinD3;4</i>	<i>N. tabacum</i>	<i>N. tabacum</i>	Accelerated growth <i>ex vitro</i>	Guo and Wang, 2008

1.5 Working hypothesis and objectives

Motivation to carry out the following study was based on results reported by Cockroft *et al.* (2000) that transformed tobacco plants over-expressing *CyclinD2;1* had accelerated rate of leaf initiation and stem growth rate attributed to shortened cell cycle duration and increased rate of cell division. It was therefore hypothesized that transformation of banana plants with a *Cyclin D2;1* gene would accelerate the cell cycle resulting in accelerated plant growth, and subsequently translate into shortened growth cycle of the banana crop. This would allow earlier fruit development and consequently earlier fruit harvesting that could offer more harvestable bunches in a year resulting in a significant economic benefit.

To find support for the set working hypothesis the following objectives were set to:

1. Isolate a *CyclinD2* gene from banana for over-expression of the coding sequence in a homologous host plant and characterize the isolated *CyclinD2* coding sequence to determine any homology with other *CyclinD*-type genes.
2. Transform banana plants with either the *Arabidopsis* or the *Musac;CyclinD2;1* transgene to generate transformed plants for phenotypic evaluation.
3. Determine the relative expression of the *Arabidopsis* and *Musac;CyclinD2;1* coding sequence in transformed banana plants allowing to relate transcription with phenotypic changes.
4. Evaluate growth and development of the transformed plants to determine if *CyclinD2;1* over-expression results in faster plant growth.

From this study, the first *CyclinD2;1* coding sequence was isolated from banana thus contribution towards assembling the banana genome information that is lacking. In addition, the structural characterization of the isolated sequence shed light on the genomic relationship of banana, other monocots and the model plants such as *Arabidopsis*. The thesis work also demonstrated the feasibility of transforming and overexpressing a *CyclinD2;1* gene in a more complex monocot species, banana, and provided support on the growth enhancement resulting from the *CyclinD2;1* transgenes in this crop.



CHAPTER TWO

ISOLATION OF A BANANA *CYCLIN D2; 1* GENE HOMOLOG

2.1 Introduction

Expression analyses of cyclins have shown their transcription abundance to coincide with active cell multiplication in cell suspension and plant organs (Soni *et al.*, 1995; Dewitte *et al.*, 2003; Jang *et al.*, 2005). Subsequently, these stages have been targeted to isolate the respective genes (Hirt *et al.*, 1992; Dahl *et al.*, 1995; Magyar *et al.*, 1997; Freeman *et al.*, 2002). Since cyclins have a high conservation in their functional domains, such as the cyclin boxes (Renaudin *et al.*, 1996; Wang *et al.*, 2004; Menges *et al.*, 2007), these functional domains have been used for isolating new plant cyclins. Initially, cyclin isolation relied on cDNA libraries that were screened using radioactively labeled probes generated with degenerated primers. These primers were designed within the conserved regions of characterized cyclins (Hata *et al.*, 1991; Dhahl *et al.*, 1995; Sorrel *et al.*, 1999; Kvarnheden, *et al.*, 2000; Freeman *et al.*, 2003). Polymerase chain reaction (PCR) using degenerated primers have been recently used to amplify conserved regions of genes from reverse transcribed RNA sequences (Wang, *et al.*, 2006; Primrose and Twyman, 2006). This method does not require construction of libraries or radioactive material and is also faster. Since all isolation techniques rely on conserved regions of the genes, they rarely generate full length sequences. As such, the flanking regions of the gene can be generated using the Rapid Amplification of cDNA Ends (RACE) technique (Frohman *et al.*, 1988). Also when no introns exist in the targeted region, genome walking of uncloned genomic libraries has been used (Siebert *et al.*, 1995; Rishi *et al.*, 2004).

The objective of this study was to isolate a D-type cyclin from banana to carry out a sequence comparison with existing *cyclin* genes. Also the gene was isolated to be over-expressed in banana and to evaluate if over-expression changes the morphology of the banana plant. In general, *Musa* genomic information is still limited and existing data are in form of expressed sequence tags (ESTs) that are not related to plant development. In this part of the study, sequence information from other monocots was used to isolate a cyclinD from banana. Applying a combination of RT-PCR, RACE and genome walking a full genomic and cDNA of *cyclinD2;1* was isolated from an East African highland banana

(AAA) cultivar ‘Nakasabira’, a dessert (AAB) cultivar ‘Sukalindiizi’ and one of their progenitors, ‘Calcutta 4’ (*Musa acuminata* spp. *burmannicoides*).

2.2 Materials and Methods

2.2.1 RNA isolation

Due to the association of high transcription of *cyclins* with actively dividing cells, a highly proliferating embryogenic cell suspension of a Highland cooking banana (‘Nakasabira’; *Musa* spp. AAA group) and of a dessert banana (‘Sukalindiizi’; AAB group) were used as a source for cyclin cDNA. The cell suspension was established from callus that was initiated from young male flowers and was harvested six days after sub-culturing into a fresh MA2 liquid medium formulation according to Côte *et al.* (1996) supplemented with 1 mg/L 2,4-D. Two hundred micro liters of packed cell volume were transferred into a 2 ml Eppendorf tube and washed twice with sterile distilled water. The cells were frozen and ground using liquid nitrogen in a mortar with a pestle. Total RNA was isolated from 50 mg of the powder using the Nucleospin® plant kit (BD Biosciences Clontech) following the recommended protocol of the supplier. Similarly, total RNA was isolated from the shoot meristem of *in vitro* proliferating shoot tip cultures of one of the progenitors of bananas ‘Calcutta-4’ (*Musa acuminata*, AA, subspecies *burmannicoides*). The quality and concentration of the RNA was assessed on a 1% agarose gel and a spectrophotometer (Nanodrop®, ND 1000), respectively, hereafter the RNA was stored at -80°C.

2.2.2 cDNA synthesis

The first strand cDNA was synthesized from 0.2 µg of the total RNA using the Improm-II™ Reverse Transcription system and by applying the recommended protocol of the supplier (Promega). Oligo dT15 (0.5µg) was annealed to the RNA at 70°C for 5 min followed by chilling on ice for 5 min. Reverse transcription was performed in a 20µl reaction mixture containing 7.5 mM MgCl₂, 0.215 mM dNTPs, 2.5 Units of reverse transcriptase and 10 Units of RNase inhibitor (RNasin). A PCR reaction was run for 5 min

at 25°C, 1 hr at 42°C, 15 min at 70°C and then the reaction was cooled at 4°C. The quality of cDNA was checked by using 1 µl of cDNA in a PCR reaction with banana *Actin* specific primers BanActFw (5'-CTGGTGATGGTGTGAGCCAC-3') and BanActRv (5'-CAGGGCAACGTAGGCAAGCT-3') designed from *Musa Actin* (Genbank accessions AF285176 and AY904067) to give a 200 bp amplicon. PCR amplification was carried out for 3 min at 96°C followed by 35 cycles of 20 sec at 94°C, 20 sec at 58°C, 30 sec at 72°C and a final extension of DNA strands for 7 min at 72°C.

2.2.3 PCR reactions and purification

All PCR products were fractionated on 1% agarose gel using 1X TAE buffer. To enable viewing the amplification products, ethidium bromide was incorporated into the gel at 0.1 µg/ml prior to casting. The stained gels were viewed under U.V.-light to identify amplification products and to isolate separated PCR products from the gel. Gel pictures were acquired with a photo documentation system.

2.2.4 Degenerated primer design

A pair of degenerated primers was designed using the consensus-degenerate hybrid oligonucleotide primers (CODEHOP) program (Rose *et al.*, 1998) and based on the nucleotide sequence within the conserved region of the *cyclin* genes (Wang *et al.*, 2004). The conserved region was obtained from a multiple alignment of monocot CycD2:1 amino acid sequences of *Zea mays* cultivar Mo17 (accession number AAL83926), *Triticum aestivum* (AAQ08041), *Oryza sativa* (EAZ43092) and *Saccharum officinarum* (AAV28532). Degenerated primers (5'-CGCCTCCATCCTGCTGtgygcngarga-3') and (5'-GAAGGAGCAGGGGTCacngentgcat-3') produced a 600 bp amplicon. PCR was conducted with 3 µl cDNA in a 50 µl reaction mixture containing 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.3 µM each of primer and 1 unit of Bio-Taq DNA polymerase enzyme. To increase the binding specificity of the primer, a touchdown PCR (Don *et al.*, 1991) was applied. The PCR cycle program was 3 min at 95°C; 15 cycles of 45 sec at 94°C, 45 sec at

70°C with a reduction of 1°C per cycle, 1 min at 72°C; 25 cycles of 45 sec at 94°C, 45 sec at 58°C, 1 min at 72°C and a final extension of DNA strands for 7 min at 72°C.

2.2.5 Genome walking

To extend the flanking regions of the first fragment, primers were designed towards the ends of the first gene fragment. To obtain the 3' downstream region, forward gene specific primer MagwRv1 (5'-TCTGTCCCTGGATTTACAGGTTGGTG-3') and a gene specific nested primer MagwRv2 (5'-GAGAATGGAGCTTCTGGTGATGAG-3') were designed and used together with GeneRacer™ 3' primer (5'-GCTGTCAACGATACGCTACGTAA CG-3') and GeneRacer™ 3' nested primer (5'-CGCTACGTAACGGCATGACACAGTG-3'). The PCR reaction was conducted in a 25 µl reaction mixture with 3 µl of single strand cDNA and the PCR reaction components mentioned above. The only difference was that the concentration of the Generacer™ 3' primer was doubled (0.6 µM) but the gene specific primer maintained at 0.3 µM. The nested PCR reaction was conducted in a 50 µl reaction volume using 2 µl of the primary PCR product and the gene specific and Generacer™ 3' nested primer used at the same concentration of 0.3 µM. The PCR was performed for 4 min at 94°C followed by 5 cycles of 30 sec at 94°C, 30 sec at 70°C, 1.5 min at 72°C; 5 cycles of 30 sec at 94°C, 30 sec at 65°C, 1.5 min at 72°C; 25 cycles of 30 sec at 94°C, 30 sec at 60°C, 1.5 min at 72°C and a final extension of DNA strands for 10 min at 72°C.

A Blast search (Altschul *et al.*, 1997) for the 5'-end of the *cyclin* gene indicated that no intron was to be expected. Thus, the upstream of the banana *cyclin* gene was generated by genome walking using restriction enzyme treated genomic DNA of the banana cultivar 'Nakasabira' (Section 1.3.6). The primary PCR was performed with gene specific primer MagwFw1 (5'-ATCTCTCCTGATGGAGGAGTCCAACG-3') and adapter specific primer AP1 (5'-GGATCCTAATACGACTCACTATAGGGC-3'). A nested PCR reaction was carried out with a gene specific primer MagwFw2 (GCTCTCTCGACCAACAAGCTCAAC) and a nested adapter specific primer AP2 (5'-AATAGGGCTCGAGCGGC-3'). The primary PCR reaction was conducted in a 25µl reaction mixture using gene specific primer1 and adapter primer1 with 2 µl of cut genomic DNA. The secondary PCR reaction

was carried out with 2 µl of the 100-times diluted first PCR amplification product using gene specific primer² and adapter primer². The PCR reaction included denaturing for 3 min at 94°C; 10 cycles for 30 sec at 94°C, 5 min at 68°C, followed by 25 cycles of 20 sec at 94°C, 5 min at 68°C with 10 sec increment per cycle and a final DNA strand extension of 10 min at 68°C.

2.2.6 Preparation of restriction enzyme-digested genomic DNA

Genomic DNA was extracted from the cigar leaf of the field-grown plants of the Highland banana cultivar 'Nakasabira' using the CTAB method (Ude *et al.*, 2002). DNA was purified with a CFXTM DNA and gel band purification kit (Amersham Biosciences). Digestion and ligation of adapters were performed on 3.5 µg DNA following the protocol described by Siebert *et al* (1995). Briefly, the process involves digestion of the DNA for 18 hrs in separate 20 µl reaction mixtures containing either restriction enzymes *PvuII*, *ScaI* or *StuI*. Reactions were terminated by heating for 20 min at 65°C for enzyme *ScaI* and at 80°C for *PvuII* and *StuI*.

The digested DNA was precipitated by adding 20 µl of phenol/chloroform/isoamyl alcohol (25:24:1). This was followed by gentle vortexing of the mixture and centrifugation for 3 min at 13,000 rpm at room temperature. The top phase was transferred into a new Eppendorff tube to which 1/10 volume 3M NaOAc and 20 µg glycogen (Fermentas) were added. Two and half volume 95% ethanol was added, mixed by inverting the tube 6 times and the mixture was kept at -80°C for 1 hr. The mixture was centrifuged for 15 min at 13,000 rpm and 4°C. The recovered pellet was rinsed with 200µl of 70% ethanol and further precipitated by centrifugation for 5 min at 10,000 rpm and 4°C. Ethanol was pipetted off and the pellet was vacuum-dried and suspended in 20 µl sterile distilled water. Adapter 1 (CTAATACGACTCACTATAGGGCTCGAGCGGCCCGCCCGGGCAGGT) and adapter 2 (ACCTGCCC-NH₂) (Siebert *et al.*, 1995) were ligated to the blunt-ended digested DNA fragments. This involved mixing 5 µM of each adapter, heating the mixture at 95°C for 3 min and allowing it to cool at room temperature. The adapter mixture was added to the DNA that had been relaxed by heating for 2 min at 50°C and cooled on ice.

Three micro liters of ligase 10x buffer and 2 µg T4 DNA ligase (Roche Applied Science) were added to the mix and the final volume made to 30 µl with sterile distilled water. The ligation reaction was conducted at 20°C for 17 hrs and terminated by heating at 65°C for 10 min. The final volume was made to 100 µl and the adapter-ligated product stored at -20°C.

2.2.7 DNA cloning

For cloning, PCR products were run on an agarose gel and the bands of interest were extracted from the gel, purified and cloned into the plasmid pCR®4-TOPO (Invitrogen). This involved ligating 30 ng DNA extracted from the gel with 10 ng of plasmid DNA in the presence of 1 µl NaCl in a 10 µl reaction mixture for 30 min at 25°C. Ligated plasmid DNA was subsequently transformed into competent *Escherichia coli* cells using the heat shock method. This involved placing a 1.5 ml Eppendorf tube containing 50 µl of *E. coli* cells on ice for 30 min after which the TOPO vector was added, mixed gently and the mixture was maintained on ice for another 30 min. The Eppendorf tube containing the bacteria-plasmid mixture was immersed for 45 sec in a 42°C water bath and immediately cooled on ice for 2 min. SOC medium (250 µl) was added to the cells and the mixture was incubated for 1 hr at 37°C on a shaker (150 rpm). To select for transformed *E. coli* cells, 50 µl of the cell suspension was spread on a plate containing 25 ml Luria Bertani (LB) medium (g/liter: 10 tryptone, 5 yeast extract, 10 NaCl₂, 15 agar) supplemented with 50 µg/ml kanamycin. The plates were incubated for 17 hrs at 37°C. Five bacterium colonies were selected and each inoculated in 5 ml of liquid LB in the presence of kanamycin and grown for 17 hrs under shaking at 200 rpm.

TOPO plasmid containing the cloned DNA insert was isolated from *E. coli* using the GeneJET™ Plasmid miniprep kit (Fermentas Life Science). The presence of the insert was confirmed by restriction enzyme analysis digesting 1 µg of the plasmid with 1 unit of *EcoR*I in a 20µl reaction mixture for 1 h at 37°C.

2.2.8 DNA sequencing

After extraction from the gel, the PCR products were prepared for sequencing by running a sequencing PCR reaction in a 10 µl reaction mixture containing 400 ng TOPO plasmid DNA, 2 µl 5x Big-Dye buffer, 2 µl BigDye and 1 µM M13 forward or reverse primer. The PCR program comprised of 95°C for 1 min, 30 cycles of 95°C (10 sec), 52°C (5 sec) and 60°C (4 min). PCR products were purified with a 6.5% Sephadex® G-50 (Sigma cat # G5050-100G) column. PCR samples were loaded onto the centre of the Sephadex column and centrifuged at 2800 rpm for 2 min. The eluted products were dried in a concentrator and sequenced on an ABI377 DNA sequencer.

2.2.9 Isolation of cyclin sequences

Obtained sequences were used to search the nucleotide database of the National Center for Biotechnology (www.ncbi.nlm.nih.gov). The sequences that gave positive *cyclin* hits were identified and used in the ExPASy online program for translation into amino acid sequences and checking for an uninterrupted open reading frame. The 5' end of the gene was determined by the ATG start codon while the TGA codon indicated the 3' termination. Based on the identified full nucleotide sequence, a forward primer (5'-ATGAGTCCAAGT TGTGACTGCG-3') and reverse primer (5'-TCATCTGGTTGTTTTCTCTCT-3') were designed to amplify the full cyclin coding region from a cDNA isolated from plants of cultivar 'Nakasabira'. This involved using 0.3 µM of primers and 1 Unit proof reading *pfu* Taq polymerase in a 50 µl reaction mixture containing 1.5 mM MgCl₂ and 0.2 mM dNTPs. The PCR program consisted of 94°C (4 min), 35 cycles of 94°C (30 sec), 60°C (30 sec), 72°C (1.5 min) and final extension of DNA strands 72°C (10 min). The amplified PCR products were run on and isolated from a 1% agarose gel, cloned into the TOPO plasmid and *E. coli* cells were transformed with the plasmid. Because *pfu* Taq polymerase does not incorporate the 3'-terminal deoxyadenosine overhangs that facilitate ligation into the TOPO plasmid, an A-tailing reaction was performed on the PCR products. A 20 µl tailing reaction was conducted using 11 µl of gel extract to which 2 µl reaction buffer containing 1.5 mM MgCl₂, 0.1 mM dNTPs and 1.5 Units of Bio-Taq DNA polymerase enzyme was added. The

mixture was incubated for 10 min at 72°C, immediately placed on ice and 4 µl of the reaction was ligated into the TOPO plasmid. The TOPO plasmid was purified from transformed *E. coli* cells and sequenced. The resultant sequences were aligned to obtain the consensus sequence. The same primers were used to amplify the corresponding coding sequences from cDNAs of plants of the banana cultivar ‘Sukalindiizi’ and from a diploid wild type ‘Calcutta-4’. Furthermore, these primers were used to amplify the full genomic cyclin from genomic DNA of cultivar ‘Nakasabira’.

2.2.10 Bioinformatics

Arabidopsis cyclin sequences were retrieved from the NCBI website (<http://www.ncbi.nlm.nih.gov>, Altschul *et al.*, 1997). Rice cyclins were retrieved from the rice genome database (<http://www.gramene.org/protein> and <http://www.iniprot.org>), while the maize orthologs were downloaded from <http://www.archieve.maizesequence.org>. For tobacco and wheat cyclins and to verify the retrieved sequences, BLAST searches were performed on the GenBank of the NCBI website. PEST motifs in the protein were determined using the ePESTfind online program (<http://www.EMBOSS.bioinformatics.nl/cgi-bin/emboss/epestfind>). The program was run at the default setting of five for the minimum number of amino acids between positive franks. To classify the isolated banana cyclin, phylogenetic analysis was performed using representative cyclin sequences from *Arabidopsis*, maize, rice, sugarcane, wheat, and a phylogram was constructed using Neighbor-Joining method with MEGA program version 3.1 (Kumar *et al.*, 2004). The same MEGA program was employed to analyze the introns in the genes by aligning the genomic and cDNA sequences of the gene. Relatedness between the banana cyclin and *cyclinD2* of monocots and *Arabidopsis* was determined with ClustalW program (<http://www.ebi.ac.uk/Tools/clustalw2/>). Specificity of the isolated cyclin to banana was determined by performing a PCR reaction on genomic DNA of selected banana accessions using primers designed at the flanking regions of the cDNA. The accessions were Yalim (AA), Kahuti-Mshale (AA), Sukalindiizi (AAB), Ibwi (AAA), Mbwazirume (AAA), Gros Michel (AAA), their wild parents ‘Calcutta-4’ (*Musa acuminata* spp. *Burmannicoides*, AA), *M. balbisiana* (BB), *M. zebrina*, *M. malacensis* (AA), and wild relative *Ensete*.

2.3 Results

2.3.1 Cyclin isolation

Polymerase chain reaction (PCR) with degenerated primers generated a 586 bp amplicon (Fig. 2.1C-i). When the derived sequence was blasted against the NCBI Genbank, the sequence showed homology to N-termini of cyclin D2 type proteins. Thus, the 5' end of the gene was obtained by genome walking using digested genomic DNA generated by restriction with enzymes *ScaI* and *PvuII*. Application of a nested PCR reaction using primers MagwFw2 and adapter specific primer AP2 yielded in 400 bp fragments from cut DNA (Fig. 2.1C-ii) that comprised of 196 bases upstream the ATG start codon. For nested PCR, products from a primary PCR reaction using primers MagwFw1 and adapter specific primer AP1, which gave a smear on the gel, were diluted 100-times for the nested reaction. When this sequence was translated with the Expasy program it yielded a full N-terminus of the gene. The 3'-RACE PCR generated a 531 bp fragment (Fig. 2.1C-iii). Blast search with this sequence gave very strong hits of the cyclinD C-terminus. When the three fragments were joined (Fig. 2.1B) and submitted to the Expasy program for transformation, an open reading frame (ORF) with ATG and TGA as start and stop codon, respectively, was generated. Finally, primers that were designed at the 5' and 3' termini picked a full-length ORF of 1035 bp from cDNA (Fig. 2.2). The nucleotide sequence had a CG and AT content of 49 and 51%, respectively, an open reading frame of 345 amino acids and a protein size of 629,017 Daltons. A similar product was amplified using the same primer pair on cDNA from proliferating cell suspension of banana cultivar 'Sukalindiizi' (AAB) and shoot meristematic tissue of a wild diploid banana 'Calcutta-4' (AA).

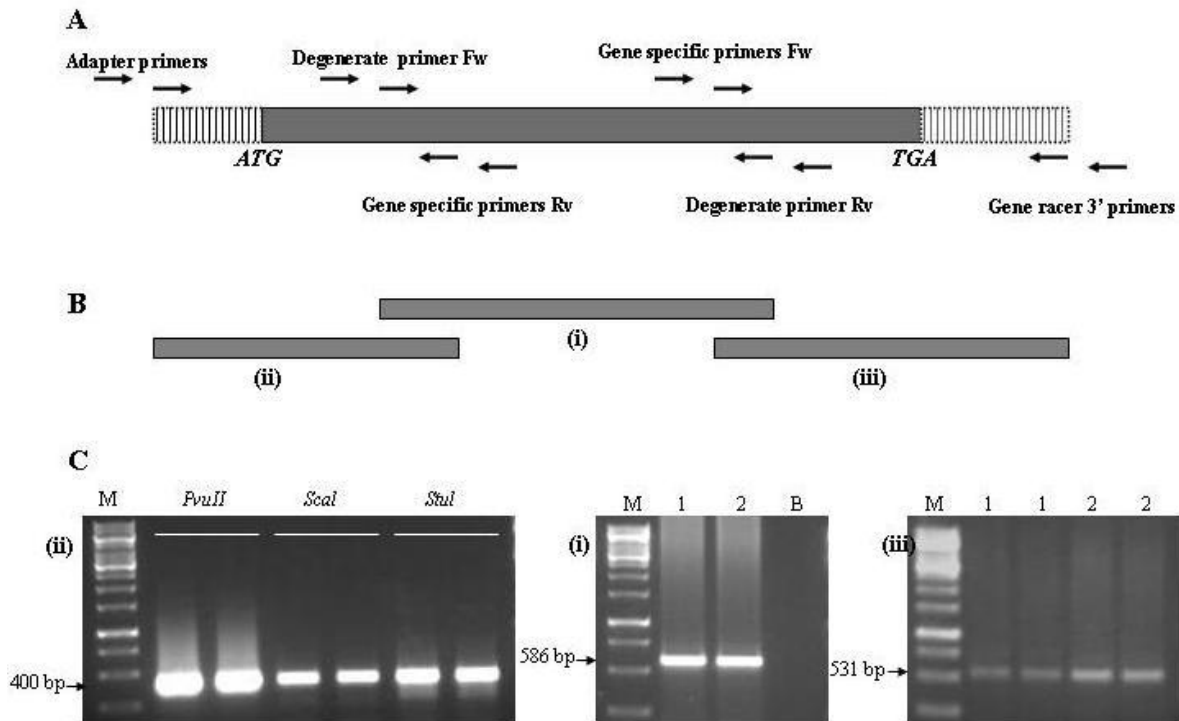


Fig. 2.1 Isolation of *CyclinD2;1* gene homolog from banana using PCR, RACE and genome walking. (A) Schematic overview of the primer positions. (B): (i) sequence of gene isolation where gene segments were amplified from the conserved region of the gene using degenerated primers and cDNA; (ii) 3' end of the gene generated by 3' RACE using forward primers designed in the first fragment together with Generacer 3' primers; (iii) 5' extension of the gene amplified through genome walking of genomic DNA using reverse gene specific and adapter primers. (C) Representative agarose gels for the respective steps: PCR products for (i), RACE secondary PCR reaction for (iii) and genome walk secondary PCR for (ii). M: 1Kb DNA ladder (GeneRuler™), 1 and are PCR samples, genome libraries were prepared with restriction enzymes *PvuII*, *ScaI* and *StuI*.

2.3.2 Molecular characterization of the coding sequence

Predicted amino acid residues of the three banana genotypes were 100% similar. Domain analysis of the protein established 134 and 87 amino acid of the cyclin_N and cyclin_C domains (Wang *et al.*, 2004; Menges *et al.*, 2007), respectively (Fig. 2.2). Furthermore, the banana *cyclin* gene was characterized by LxCxE motif (where x is any amino acid residue, i.e. LLCAE) and a retinoblastoma (RB) related protein binding motif (Fig. 2.3; Wang *et al.*, 2004; Menges *et al.*, 2007). The LxCxE motif was similar to that of maize, sugarcane and wheat but different from the *Arabidopsis* (LACGE) and tobacco (LLCTE). A unique motif, IWKVHAHY, was also identified within the cyclin N_domain (Fig. 2.2). In addition, the N_domain had the D, K, L and E residues that are highly conserved and required for cyclin function (Menges *et al.*, 2007). One PEST motif was identified at amino acid position 1 to 33 (PESTFIND score of +2.13; Fig 2.2). PEST motifs are signals for rapid protein degradation and are an integral part of some cyclins (Renaudin *et al.*, 1996; Wang *et al.*, 2004). PEST motifs are characterized by a stretch of at least twelve negatively charged amino acids mainly proline (P), glutamic acid (E), serine (S), threonine (T), and aspartic acid (D) relative to the positively charged ones (Rechsteiner and Rogers, 1996). Based on the proportion of these amino acids to the positively charged residues, an algorithm computes the PEST score with a threshold of 5.0, below which the score is considered poor.

Phylogenic analysis between the full *Musa* cDNA amino acids sequences and the D-type cyclins from other plant species placed the *Musa cyclin* gene in a monocot specific clade (Fig. 2.4). Within this clade, the *Musa cyclin* had a 17% bootstrap support. At amino acid sequence levels, *Musac;CYCD2;1* was 40% identical to *Arabidopsis*, 54% to rice, 57% to maize, 54% to sugar cane, 57% to wheat cyclinD2.

1	ctc gcc cgt ctt ttg tct cac tga gca	ATG AGT CCA AGT TGT GAC	6
		<i>M S P S C D</i>	
46	TGC GCC TCC TCT ATC	CTC CTG TGC GCT GAG GAC AAC GAC AGC ATC	21
	<i>C A S S I</i>	L L C A E	
91	CTG GGC TTC GAC GAT	GGT GAG GAG GAG GGC GGG CAT AGG CCT GGA	36
	<i>L G F D D G E E E G</i>	G H R P G	
136	TGG GTT TCT GAA CCA AAA AGG TGC GAT TTT TAT GGG GAT ATC CTC		51
	<i>W V S E P K R C D F Y G D I L</i>		
181	GCG GGT TTT CCC CTG CAG TCG GAT GAG TTG TTG AGC TTG TTG GTC		66
	<i>A G F P L Q S D E L L S L L V</i>		
226	GAG AGA GAG CAG GAG CAT CTG CCG AGG GAG GAC TAC CGC GAG AGG		81
	<i>E R E Q E H L P R E D Y R E R</i>		
271	CTG TGT TCT GGG GCG TTG GAC TCC TCC ATC AGG AGA GAT GCT ATT		96
	<i>L C S G A L D S S I R R D A I</i>		
316	GAT TGG ATT TGG AAG GTT CAT GCT CAT TAC	AAT TTT GGA CCA CTG	111
	<i>D W I W K V H A H Y N F G P L</i>		
361	AGT GCC TAT TTA TCT GTA AAT TAC TTG GAT AGG TTC CTC TCT TCC		126
	<i>S A Y L S V N Y L D R F L S S</i>		
406	TAT GAT CTC CCA CAA GGC AAG GCT TGG ATG ACA CAA CTA TTA TCT		141
	<i>Y D L P Q G K A W M T Q L L S</i>		
451	GTG GCC TGC CTA TCT TTG GCT GCC AAG ATG GAG GAA ACT GAA GTT		156
	<i>V A C L S L A A K M E E T E V</i>		
496	CCT CTG TCC CTG GAT TTA CAG GTT GGT GAG GCA AAA TAT ATA TTT		171
	<i>P L S L D L Q V G E A K Y I F</i>		
541	GAA GGA AGG ACA ATT CAG AGA ATG GAG CTT CTG GTG ATG AGC ACC		186
	<i>E G R T I Q R M E L L V M S T</i>		
586	CTC AAA TGG AGG ATG CAA GCT GTG ACT CCT TTC TCA TTC ATA GAT		201
	<i>L K W R M Q</i>	A V T P F S F I D	
631	TTC TTC CTC CAC AAG TTC AAT GGT TGT GGT GCA CCT AGC AAG TTG		216
	<i>F F L H K F N G C G A P S K L</i>		
676	TCA CTT TCT CGA TCT GCC GAA CTC ATC CTG AGC ACA ATT AGA GGC		231
	<i>S L S R S A E L I L S T I R G</i>		
721	ATT GAT TTC CTA GCA TTC AGA CCC TCA GTA ATT TCT GCA GCC ATT		246
	<i>I D F L A F R P S V I S A A I</i>		
766	GCA CTG TTG GTT TTG GGA GAA ACC CAG ATT GTG GAT GTT GAG GAG		261
	<i>A L L V L G E T Q I V D V E E</i>		
811	GCC TTG TCT TGT TGC TGT CAT GTA GCT AAG GAG GGG GTG TTG GGA		276
	<i>A L S C C C H V A K E G V L G</i>		
856	TGT TAT GAA GTG ATT CAA GAC AAA GTA TTG ATG AGG AAG CAG TCA		291
	<i>C Y E V I Q D</i>	K V L M R K Q S	
901	GCC AAA GAC CTA GTC TCC TCG GTA TCC TGT GTG CCC CAA AGC CCT		306
	<i>A K D L V S S V S C V P Q S P</i>		
946	GTT GGG GTG TTG CAT GCT GCA TGC CTG AGC TAC AAG AGT GAT GAT		321
	<i>V G V L H A A C L S Y K S D D</i>		
991	GCA ACT GTT GTG TCA CAT GCA ACA TGT CTT GGT CTC TCT CCA GCT		336
	<i>A T V V S H A T C L G L S P A</i>		
1036	AGC AAG AGG AGG AAA ACA ACC AGA TGA tca ctg tca tgc cgt tac		345
	<i>S K R R K T T R R</i>		

Fig. 2.2 Nucleotide and deduced amino acid sequences of *Musac;CyclinD2;1* cDNA. The PEST motif (amino acids 1 to 33) is bolded and in italics; retinoblastoma (RB) related protein binding motif LxCxE is boxed in solid; cyclin N_domain (134 amino acids) is bolded and underlined; Cyclin C_domain (87 amino acids) is in italics and underlined; signature motif is boxed in dotted line. Nucleotides are numbered on left and amino acid residues on the right.

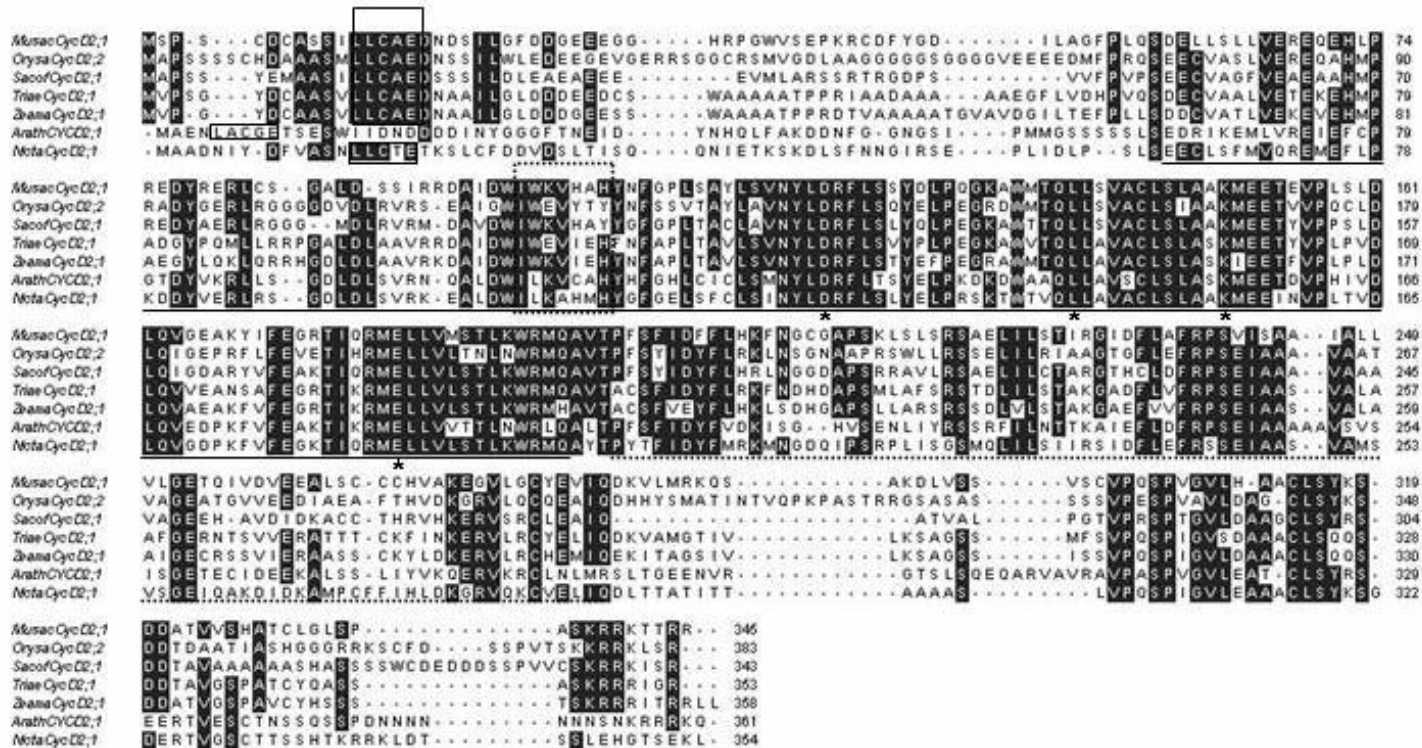


Fig. 2.3 Multiple alignment of *cyclinD2;1* amino acid sequence of banana with Arabidopsis (Arath, *Arabidopsis thaliana*), tobacco (Nicta, *Nicotiana tabacum*), rice (Orysa, *Oryza sativa*), sugarcane (Sacof, *Saccharum officinarum*), maize (Zeama, *Zea mays*) and wheat (Triae, *Triticum aestivum*). The LxCxE is presented in banana as LLCAE and similar to maize, sugarcane and wheat. The cyclin N_domain is underlined (solid line) and the C_domain with a dotted line; cyclin signature motifs are boxed (dotted line), where IWKVHAHY is specific to banana. Four conserved residues required for cyclin function are underscored with stars. Dashes indicate gaps required to maximally align the sequences.

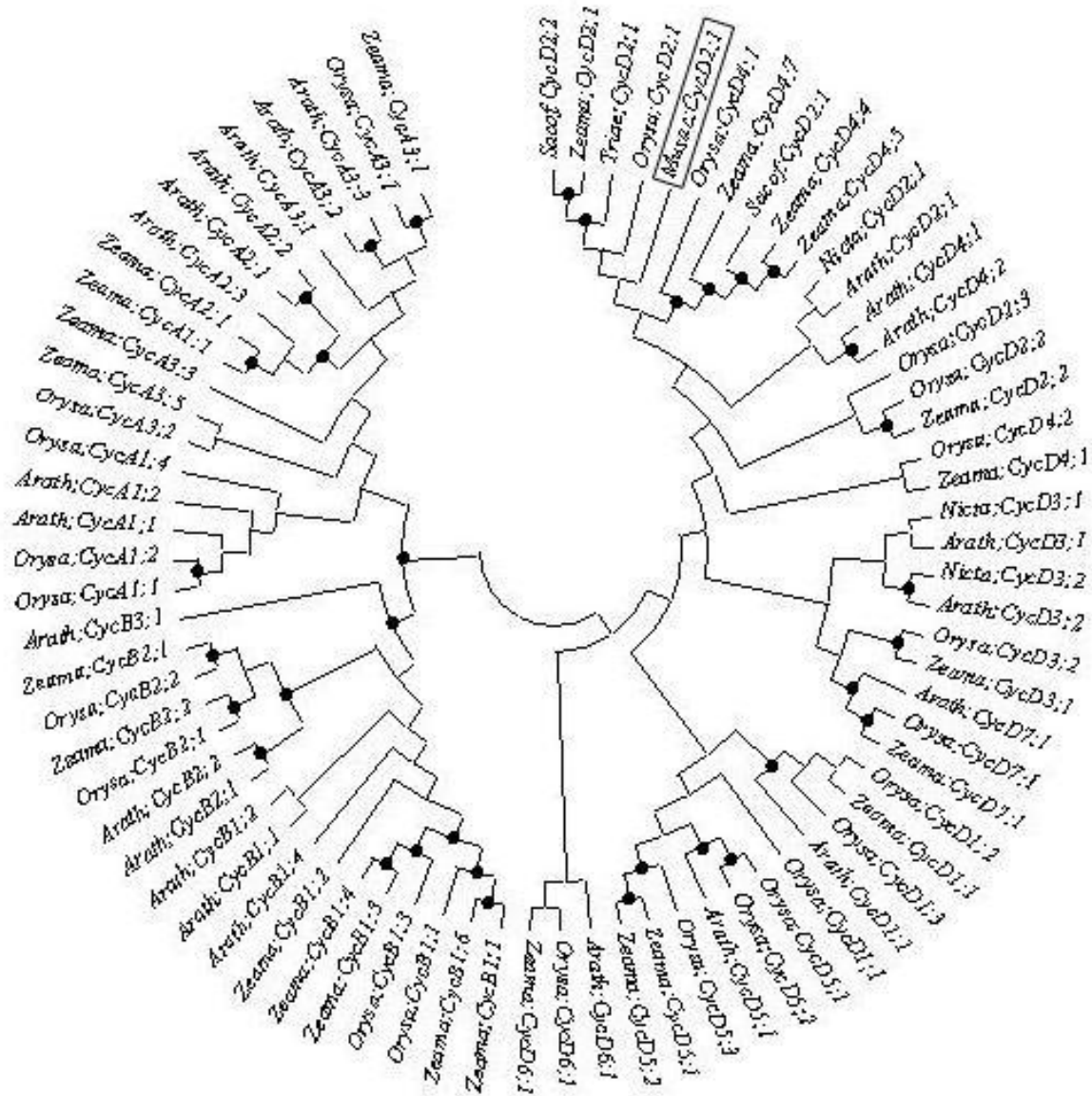


Fig. 2.4 Phylogenetic relationship between *Musac;CyclinD2;1* (boxed) and other plant cyclins. Amino acid sequences were aligned using ClustalW program and the phylogram was constructed using the neighbour-joining method. Names are presented by the first three letters of the genus and two of species: Arath, *Arabidopsis thaliana*; Musac, *Musa acuminata*; Nicta, *Nicotiana tabacum*; Orysa, *Oryza sativa*; Sacof, *Saccharum officinarum*; Triae, *Triticum aestivum*; Zeama, *Zea mays*. Nodes with bootstrap values above 70 are shown and are indicated as bullets.

2.3.3 Molecular characterization of the genomic sequence

To test the specificity of the coding sequence of this cyclin homolog to *Musa*, a PCR reaction was conducted with gene specific primers designed for the flanking ends of the ORF. Genomic DNA from eleven banana accessions, their wild relative Calcutta-4, and wild relative, Enset was used. An amplicon of 2033 bp was generated that represented a full-length genomic *CyclinD* gene (Fig. 2.5). Alignment of genomic DNA of cultivar ‘Nakasabira’ and its cDNA and the cDNA from ‘Sukalindiizi’ and ‘Calcutta-4’ revealed 6 exons of length: 1, 304 bp; 2, 86 bp; 3, 101 bp; 4, 201 bp; 5, 122 bp; 6, 221 bp (Table 2.1). The introns were respectively spliced at GG/CG-GG/TT, CA/GT-AG/CA, GT/GA-GT/TG, GG/TA-GG/CA and GG/TA-GG/AG sites.

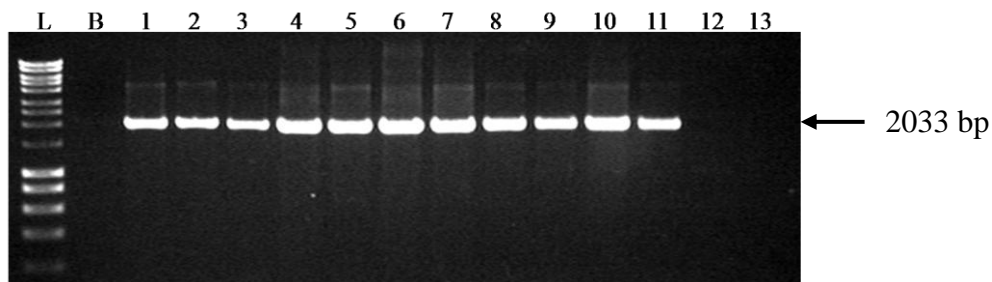


Fig. 2.5 Presence of *Musac;CyclinD2;1* coding region in selected *Musa* accessions after PCR reaction with genomic DNA. L: 200 bp HyperLadder™ (Bioline); B: Buffer; 1: Calcutta-4 (AA); 2: *Musa balbisiana* (BB); 3: *Zebrina*; 4: *M. ornata*; 5: *M. malaccensis* (AA); 6: Yalim (AA); 7: Kahuti (Mshale-AAA); 8: Ibwi (AAA); 9: Nakasabira (AAA); 10: Gros Michel (AAA); 11: Enset; 12: Maize; 13: Tobacco.

Table 2.1 Banana *CyclinD2;1* gene structure represented by size and composition of the exons and intron, and splicing sites.

Segment	Size (bp)	CG (%)	AT (%)	Splicing site
Exon 1	304	56	44	
Intron 1	534	33	67	GG/CG---GG/TT
Exon 2	123	40	60	
Intron 2	86	31	69	CA/GT---AG/CA
Exon 3	84	49	51	
Intron 3	101	36	64	GT/GA---GT/TG
Exon 4	99	44	56	
Intron 4	201	43	57	GG/TA---GC/CA
Exon 5	158	46	54	
Intron 5	122	32	68	GG/TA---GG/AG
Exon 6	221	49	51	

2.4 Discussion

A banana *CyclinD* gene was isolated through a combination of genome walking and the 3' RACE technique. To our knowledge, this is the first cell cycle gene to be isolated in banana. The gene sequence isolated has a cyclin core of N and C domains constituting the binding site for the cyclin-dependent kinases (CDK) (Sorrell *et al.*, 1999; Wang *et al.*, 2004; Menges *et al.*, 2007). The LxCxE protein sequence (in the form of LLCAE) at the N-terminal of the protein is essential for the cyclinD-associated kinase to bind and phosphorylate the retinoblastoma-related (RBR) protein. The LLCAE sequence is characteristic of cyclinD2 types (Sorrell *et al.*, 1999; Menges *et al.*, 2007) and this feature places the isolated banana cyclin into the D2-group. Further evidence for placing the banana cyclin into the D group is the six exons were identified in the isolated banana *cyclin*. This gene structure is in agreement with observations by Menges *et al.* (2007) on *cyclinD* genes in representative members of the angiosperm. Phylogenetic analysis also placed the banana cyclin close to monocot cyclin D2 and D4, although comparison of the overall banana genome structure and organization placed banana closer to Arabidopsis than to rice (Dickman, 2004).

PCR further confirmed the presence of the gene in eleven banana accessions and their wild Enset relative (Fig. 2.6). According to the classification of banana (Simmonds, 1962; Stover and Simmonds, 1987), *Musa ornata* belongs to the section Rhodochlamys, while the rest of the bananas tested here belong to the section Eumusa; the two sections are of the genus *Musa*. Likewise, Enset is a member of the genus Ensete, which together with the genus *Musa* belongs to the family Musaceae. Therefore, the presence of amplification from banana and Enset further suggests that the detected cyclin sequence is conserved across the Musaceae family. Future sequencing the product from the distant banana relatives will further confirm this.

The *cyclin* gene is the first to be isolated from banana. Since the sequenced DNA was extracted from bananas of AA, AAA and AAB genomes, where the acuminata genome dominates, it can be denoted *Musa acuminata; Cyclin D2;1* or *Musac; CyclinD2;1* according

to the guidelines of the Commission on Plant Gene Nomenclature (Renaudin *et al.*, 1996). The IWKVHAHY sequence found within the cyclin N-domain is so far unique for banana. This motif can be considered as a signature motif for the isolated banana cyclin, since it is variable and species-specific (Wang *et al.*, 2004). Further, the PEST region that was identified had a low score. This low score has been reported in the majority of the CyclinD-type proteins (Sorrell *et al.*, 1999; Wang *et al.*, 2004; Menges *et al.*, 2007). For the degradation of such proteins, Rechsteiner and Rogers (1996) proposed the phosphorylation of the serines and threonines as a possible mechanism for activating their latent PEST signals.

With regards to the splicing sites recognized in mRNA processing, only one *cyclin* intron out of the five had the conventional splicing border composition of AG---GU at the 5' and 3' ends reported for plants and vertebrates (Brown and Simpson, 1998; Reddy, 2001). A random sample in Genbank on genes that have been submitted as genomic DNA or cDNA identified this non-conformity to range between 50 and 100% in the Arabidopsis, popular and rice *cyclins* as well as banana *actins*.

In conclusion, a *Musac;CyclinD2;1* gene ortholog was isolated from banana, with higher sequence identity to monocot plants than *Arabidopsis*. This is the first cyclin to be isolated from banana and the sequence can contribute to the ongoing banana genome work (www.musagenomics.org). The sequence was deposited into a public database as *Musac;CYCD2;1* (GenBank accession number HQ839770). In chapter three an expression vector was constructed to facilitate the over-expression of this cyclin and in chapter five the resultant effect on banana plant growth is presented.