CHARACTERIZATION OF TWO GENES INVOLVED IN *NEOTYPHODIUM LOLII* GROWTH.

A thesis in partial fulfillment of the requirements for the degree of Master of Science in Biochemistry at Massey University, Palmerston North, New Zealand

Duncan George Glenn McMillan

2004

ABSTRACT

Neotyphodium lolii is a filamentous fungus that forms symbiotic associations with *Lolium perenne*, growing in its intercellular spaces. It is a feature of the symbiosis that growth of the fungus and the plant is synchronized. When the grass leaf-blade grows, the fungus grows at the same rate, hence when the blade ceases extension the hyphae do likewise. In addition, *in planta* there is little hyphal branching, where as in culture hyphae branch at regular intervals. This suggests the existence of a regulatory mechanism *in planta* that partially dictates hyphal morphology and growth.

The criteria for choosing possible candidate genes relied on whether the gene had a function relating to hyphal branching and/or regulation of hyphal extension in several organisms. Three candidate genes were selected. Protein elongation factor 2 (*EF-2*; an elongation factor associated with the ribosome) was targeted to add more direct evidence to the high metabolic rate observed *in planta* using the GUS reporter gene by Tan *et al* (2001). Cell division control protein 12 (*CDC-12*); a septin which is involved in the construction of the 10 nm ring structure associated with cell division and whose mutation is lethal in yeast was chosen to help distinguish the growth mode of *N. lolii in planta*. A Stretch-activated Calcium Channel (*SACC*) which allows exogenous calcium into the cell upon application of lateral pressure on the membrane was targeted to help distinguish the possible recognition signal the hyphae make to elucidate when the host tissue is growing.

This project was then divided into four parts, one part per gene and a final part looking at the *in vitro* and *in vivo* expression of these genes. For the first three parts degenerate PCR was performed and appropriate-sized fragments cloned, sequenced and restriction mapped for *EF-2* and *CDC-12* (2066 bp and 514 bp respectively). Database searches were used to identify the sequences as potentially being the target genes. Degenerate PCR was unsuccessful for the *SACC*.

Southern blots were used to identify restriction enzymes for Inverse PCR; and this was used to obtain the remaining 5' and 3' regions of each target gene. Gene prediction software was used to predict gene structure; 5' and 3' RACE to confirm the length, introns and start/stop points of *EF-2* and *CDC-12* full gene transcripts (2,900 and 1,612 bp respectively). Internet-based sequence analysis tools subsequently were used to identify sequence features.

For the second part, expression of *EF-2* and *CDC-12* are investigated during various states of hyphal growth. Growth curves were constructed and *in vitro* expression analysis was achieved by Northern blot. The expression patterns of *EF-2* and *CDC-12* followed the growth state of *N. lolii*. RT PCR was used to confirm *in planta* expression of both genes and validate their uses for future studies.

ACKNOWLEDGEMENTS

I would like to convey my thanks to my supervisor, Jan Schmid for his patience and guidance in times of trial. I would also like to thank Ningxin Zhang for her help in the lab, without whose assistance I would have struggled to achieve results at all. Thank you both for taking time out of your busy schedules to help me with my project.

I would like to thank Mark Patchett, Mike Christensen and Errol Kwan for taking the time out to have all those philosophical chats with me about many aspects of research and ideas for further research. I have appreciated your friendship and advice. I would like to thank Rosie Bradshaw for acting as an interim supervisor and helping last-minute editing when she was very busy with other aspects of her work. I would like to acknowledge the members of Scott base, particularly Andrea Bryant, Michelle Bryant, Barry Scott and Carolyn Young for useful advice as well as everyone at IMBS for giving me tidbits of guidance throughout the years.

I would like to thank my family for their support and Kathryn Stowell for her understanding and advice for so many years. Thanks to the lads, I shall miss the chronic partying and science-free laughs. Lastly I would like to thank Mazzie who put up with my late-night swearing/ writing episodes when I was under pressure, and who offered encouragement when I needed it.

TABLE OF CONTENTS

FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 4

TABLES

ABBREVIATIONS

ATP	Adenosine 5'-triphosphate	h	Hour
GTP	Guanine 5'-triphosphate	mm	Millimeter
dNTP	Nucleotide 5'-triphosphate	cm	Centimeter
CDC	Cell cycle division protein	U	Units of Enzyme
EF	Protein Elongation factor		
SACC	Stretch-activated calcium channel		
mRNA	Messenger RNA		
MCS	Multiple cloning site		
GUS	β-glucuronidase		
Gd^{3+}	Galadeninum ion		
GST	Glutathione S-transferase		
kb	Nucleotides (Kilobases)		
b	Nucleotide (base)		
MOPs	3-N-Morpholinepropanesulfonic acid		
°C	Degrees centigrade		
µL	Micro-liter		
mL	Milli-liter		
L	Liter		
pM	Pico-mole		
µM	Micro-mole		
mM	Milli-mole		
M	Mole		
ng	Nanogram		
µg	Microgram		
mg	Milligram		
g	Gram		
s	Second		

Chapter 1

INTRODUCTION

1.1 *Neotyphodium lolii* Growth and Host Interactions

1.1.1 The *N. lolii* Life Cycle

The filamentous ascomycete fungus *Neotyphodium lolii* (Clavicipitaceae) and several grasses of the sub-family *Pooideae* form a symbiotic relationship. *Neotyphodium* endophytes are anamorphic (asexual), the life cycle occurs solely within the organs of the host grass and reproduction is by seed dissemination (vertical transmission; (Schardl and An, 1993)). *N. lolii* consists of elongated filaments (hyphae). These hyphae are divided into compartments by septa; compartments are connected by pores through which cytoplasm and organelles may pass. Nutrients are absorbed through the hyphal wall. Hyphae are present within the embryo that becomes infected early in its development. Each new host tiller starts as an axillary bud that develops on the stem apex, along with the new leaves. The *N. lolii* hyphae grow into the axillary bud, thus colonising each new tiller and axillary bud (Christensen and Bennett, 2004a). Hyphae only grow within the intercellular spaces of actively growing host tissue, localizing around vascular bundles and seldom branching (Figures 1.1 A) and B); (Schmid and Christensen, 1999). When a tiller becomes reproductive, *N. lolii* grows into the reproductive structures and into both the developing seed and embryo (Christensen and Bennett, 2004a). Embryonic hyphae grow with the embryo upon germination, while seed hyphae remain static (Philipson, 1986). The endophyte biomass is estimated to be less than 0.2% of the total infected plant biomass (Tan *et al.*, 2001).

1.1.2 Host Compatibility

The *Neotyphodium*/grass association has narrow limits of compatibility and if the fungus deviates from this, the results are hyphal death or plant-tissue death hence they are asymptomatic (Christensen, 1995; Schmid and Christensen, 1999). There is no evidence from studies that host grasses activate their defense mechanisms to the endophyte presence, however

further investigations have indicated that the host is also capable of responding to an incompatible association. Host specificity and compatibility studies have been carried out using seedling inoculation techniques. *N. lolii* can form novel associations with grasses closely related to their natural host species *Lolium perenne* (Christensen, 1995). Although some artificial combinations are fully compatible and stable most are not, resulting in the death of host tissue (hyper-sensitivity), stunted hosts or premature death of hyphae (Koga *et al.*, 1993; Christensen, 1995; Christensen *et al.*, 1997). It is unknown why attempts to form associations between *N. lolii* and grasses distantly related to *L. perenne* are unsuccessful.

1.1.3 Biological Benefits of Symbiosis

The presence of *N. lolii* in *L. perenne* enhances host physiological characteristics and in return the endophyte has a highly efficient means of dispersion through the seed and an inexhaustible supply of nutrients while the host lives. The intercellular spaces within *L. perenne* provide *N. lolii* with a stable environment in which they are free from the influence of competing microorganisms (Christensen, 1995; Schmid and Christensen, 1999; Tan *et al.*, 2001).

1.1.3.1 General Benefits

N. lolii infection is thought to enable greater vegetative vigour (tillering and shoot growth) and greater drought tolerance than non-infected plants (Clay, 1990). Endophyte-infected tall fescue (*Festuca arundinacea*) has been observed to maintain higher growth rates than non-infected tall fescue under drought conditions. It has been suggested by Clay (1990) that this is due to greater stomatal resistance and possibly higher photosynthetic rates. However these rates have been shown though glass house studies to be inconsistent and been unable to show any effect on growth or stress tolerance (Barker *et al.*, 1997). Endophyte infection has been shown to result in more reproductive organs and seed production, as well as improved germination success rate and faster seedling growth. *N. lolii* is also present between the seed coat and the aleurone layer, synthesizing alkaloids to provide protection for the seed and developing seedling from insects and parasitic nematodes (Clay, 1990; White, 1991).

1.1.3.2 Alkaloids and Bioprotection

There are three classes of secondary metabolite alkaloids that have been shown to protect the host from a variety of herbivorous grazers and parasites; Indole diterpenes (i.e. lolitrem B), ergot alkaloids (i.e. ergovaline) and pyrrolopyrazines (i.e. peramine; (Christensen *et al.*, 1993;

Scott *et al.*, 1999)). Lolitrem B is a potent tremorgen and was shown to be the cause of a neurological disorder in cattle and sheep known as 'ryegrass staggers'. Lolitrem B has also been field trial evaluated in a *Neotyphodium* sp. (strain AR6) that does not produce lolitrem. This demonstrated lolitrem B as a larval feeding deterrent to insects such as cutworm (*Graphania mutans*) and argentine stem weevil (*Listronotus bonariensis*), the latter of which is a primary agricultural pest in New Zealand pastures (Dymock *et al.*, 1989; Popay *et al.*, 1995). Ergovaline is the most abundant of the alkaloids produced *in planta* by *N. lolii*. Tissue culture and whole animal studies have identified ergovaline as a major causative agent of tall fescue (*F. arundinacea*) toxicosis and heat stress in cattle and sheep (Dyer, 1993; McLeay *et al.*, 2002). Alkaloid diet studies also demonstrated ergovaline to confer resistance against black beetle (*Heteronychus aratur*; (Ball *et al.*, 1997a)).

Ball *et al* (1997b) also showed peramine to be the major deterrent to feeding argentine stem weevil through similar studies. Distribution of peramine was higher in the lower blade portions than elsewhere, peaking adjacent to the ligule. This coincides with the observation by Keogh *et al* (1996) of argentine stem weevil feeding on only the tips of *L. perenne* lamina suggesting avoidance of the toxin (Keogh *et al.*, 1996; Ball *et al.*, 1997b). Interestingly cultures of *N. lolii* produce a reduced amount of lolitreme B than those *in planta* and no detectable peramine, suggesting that the production of these toxins is partly a plant-elicited response, thus the host promoting its own protection (Ball *et al.*, 1995).

1.1.4 The Phenomenon of *N. lolii* Growth

1.1.4.1 Hyphal Growth Dogma

The hypha is the characteristic growth form of fungal cells, a tube of approximately constant diameter ending in a curved tip. This filament may or may not be subdivided by septa into individual cells (Harold, 1990). Filamentous fungi are known to grow by two modes; hyphal tip extension and by the construction of new hyphal tips (branching).

With a few exceptions (i.e. *Coprinus cinereus,* in which chitin is synthesized uniformly throughout the length of the hyphae) hyphae elongate strictly by apical extension *via* the deposition of cell wall polysaccharides, in particular chitin and β-glucans (Harold, 1990). Confirmation of this specific deposition comes from microscopic autoradiography with tritiated *N*-acetylglucosamine and glucose (Gooday, 1971), and concurrently by rapid bursting

of growing hyphal tips by specific inhibitors of chitin and glucan synthesis respectively (Gooday, 1990). For apical extension to occur, the wall at the apex must be flexible (plastic) enough to allow its extension by the insertion of new material, and become rigid as it progresses to become the lateral wall of the growing hypha. A major part of apical growth is the forward transport of vesicles, which provide new cell wall and membrane material (Bartnicki-Garcia *et al.*, 1979). These vesicles accumulate in the Spitzenkorper ('apical body') before moving to the apical dome (Gooday, 1995). Wessels *et al* (1990) suggest that some of these vesicles contain chitin and glucan, and are deposited separately at the apex, becoming cross-linked into an insoluble matrix (Wessels, 1990). This steady forward movement of the vesicle supply center (VSC; i.e. soft endoplasmic reticulum) provides further evidence that there is an apically extending hypha. There is also evidence that cytoskeletal proteins have roles in this polarity by transporting the aforementioned vesicles, which center around the active roles of actin micofilaments and microtubules (Heath, 1994).

When conditions are favorable, hyphae branch, putting forth new tips from regions that were previously quiescent (Watters *et al.*, 2000). Branching can take place either apically, i.e. by splitting of the growing apex into 2 or more apices (induced), or laterally (native), i.e. the emergence of new apicies along hyphal cell walls. There is evidence to suggest that when a new tip (branch) is formed, it behaves exactly like new hyphae and extends by apical extension (Wessels, 1990). There are two schools of thought about the origins of branch initiation. The first school suggests branches are initiated by factors originating proximal to the new tip event itself. Bartnicki-Garcia *et al* (1990) suggested that the VSC at the tip is important in the branch determining process, and that there might be a hyphal-based pushing mechanism (i.e. a slight change in direction of vesicle flow so that a new spitzenkorper would form at a point on the rigid hyphal wall) responsible for the shift in which the cytoskeleton is a prime candidate.

The second school focuses on events which are controlled independently at the new tip. Bartnicki-Garcia *et al* (1990) also suggested that there might be a tip-based pulling mechanism that displaces the VSC for branching. Possible candidates for this role are actin microfilaments and integrin (Bartnicki-Garcia *et al.*, 1990). However it is generally conceded that some local dissolution of normally rigid cell wall is a prerequisite to new tip initiation. Mullins (1973) have shown prior to branch initiation there is a localization of vesicles, and have collected circumstantial evidence linking the secretion of cellulase in *Achlya biseualis* cells (an Oomycete) to the emergence of a branch (Mullins, 1973).

In addition calcium has been invoked as an important factor in both tip growth and hyphal branching (Reissig and Kinney, 1983). Ca^{2+} ions are well suited to such a role, given their ubiquitous functions as second messengers in eukaryotic cells (Stryer, 1995). In the hypha, Ca^{2+} is thought to play an important role in regulating apical extension through polarized growth (Takeuchi *et al.*, 1988; Harold, 1990). Levina *et al* found that the Ca^{2+} gradient peaked 3 μm behind the tip (Levina *et al.*, 1995). Further experiments showed that this apical Ca^{2+} gradient is present only in growing hyphae (Levina *et al.*, 1995). This provides evidence to support the theory that a Ca^{2+} gradient is required for hyphal growth

Most filamentous fungi have the following growth characteristics; i) Exponential growth (a constant elongation rate where hyphal tips are added to on a regular basis), ii) growth is limited by nutrient availability, and iii) hyphae branch at regular intervals (Gooday, 1995).

1.1.4.2 A Synchronous Growth Pattern

Observations of *N. lolii in planta* showed a complex interaction that may differ from the dogma in more than one way.

Christensen *et al* (2004b) observed hyphae within the tiller meristematic zone branch regularly, laterally filling the intercellular space they inhabit. It was also observed as hyphae extend into the leaf expansion zone they straighten, displaying a linear growth pattern within the leaves of its host, as few branches are formed (Figure 1.1 B); (Christensen and Bennett, 2004a). This concurred with findings by Tan *et al* (2001) that leaf blades grew in a linear fashion, and the number of hyphal strands in a given section of a leaf remained constant throughout the life of a leaf. This suggests that it is a feature of the symbiosis that growth of the fungus and the plant is synchronized (Christensen and Schmid, 1999). When the grass leaf-blade grows, the fungus grows at the same rate, hence when the blade ceases extension the hyphae do likewise.

As with growth observed in meristematic tissue, *N. lolii* hyphae *in culture* branch at regular intervals (regular tip addition) and growth is limited by nutrient availability in liquid culture, both features of the hyphal dogma (Tan *et al.*, 2001). This suggests growth of the hyphae is strongly regulated *in planta* as few branches form in the leaf and the hyphae cease growth in the presence of a constant nutrient supply. Interestingly although *N. lolii* growth ceases, metabolic activity remains high *in planta*. Reporter gene studies by Tan *et al* (2001) using β-

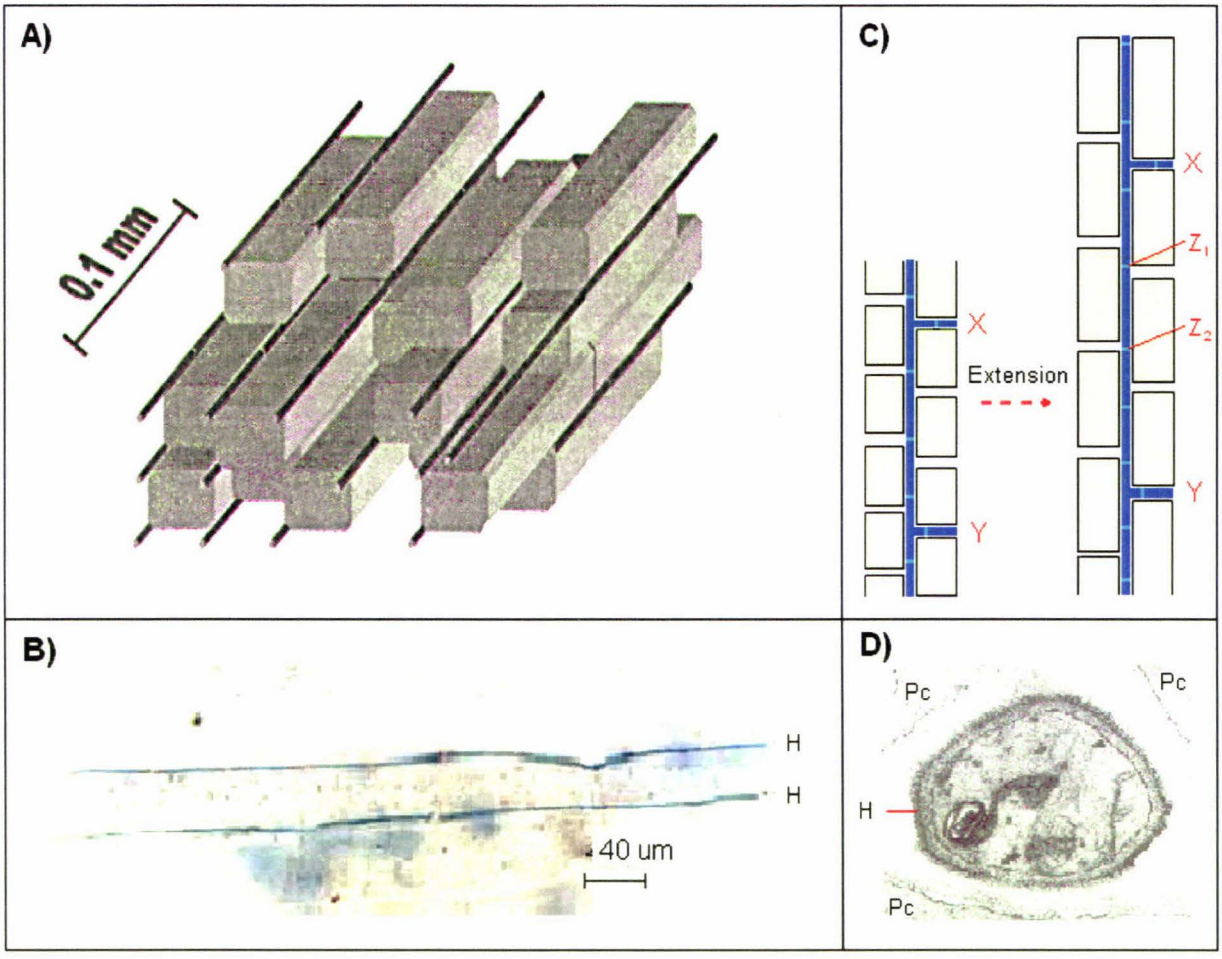
Figure 1.1: ***N. lolii in planta***

A) *N. lolii* hyphae (black rods) grow in the intercellular spaces of *L. perenne* cells (grey rectangles; (Schmid and Christensen, 1999)).

B) Longitudinal section of *N. lolii* non-branching hyphae (H) growing between *L. perenne* leaf sheath cells. Hyphae are stained with stained blue with 0.1% aniline blue (Christensen and Bennett, 2004a).

C) Schematic diagram of the possible phenomenon of Intercalary extension. Hyphae are represented by dark blue, septa by light blue and plant cells by cream. Points 'X' and 'Y' denote lateral hyphae branching off, the proposed intercalary extension between the two. Points Z_1 and Z_2 are septa that would be laid if this extension occurred (Christensen 2004b).

D) Transverse section of *N. lolii in planta*. Plant cells are labeled Pc, the hyphae indicated by H (Christensen, 2000).


A)
0.1 mm
C)
X
Extension
Y
X
Z1
Z2
Y
B)
H
H
40 um
D)
Pc
Pc
H
Pc

glucuronidase (GUS) gene under the control of a heterologous constitutive promoter (*A. nidulans gpdA*) demonstrated that metabolic activity remains at a high rate after growth has ceased, and conversely *in culture* that it does not. Observations by Christensen *et al* (2002) that hyphae increase in structural complexity, wall and septum thickness, and diameter support the conclusion of ongoing metabolic activity.

In order for *N. lolii* to successfully grow within its host exclusively by tip extension, the hypha must slide between the plant cells as the leaf expansion zone is just above its base (Schmid and Christensen, 1999), partially explaining why little hyphal branching is observed during plant growth stages. However this does not explain why when growth ceases the hyphae of *N. lolii* does not grow further or branch more.

Recent microscopy observations by Christensen *et al* (2004) query this theory of extension by exclusive tip extension. *N. lolii* hyphae appear to be in close contact with the host cells, transverse sections showing altered physical shape on a lateral plane (Figure 1.1 D). However this close contact may be mediated by a thick mucin layer, large heavily glycosylated proteins which are secreted on mucosal surfaces and as a lubricant to minimize shear stresses (Lodish *et al.*, 1999). This explanation is not entirely consistent with the shape alteration of the hyphae when only attached to one plant cell (Christensen, 2000).

In order for the exclusive tip growth model to be correct the oldest part of the hyphae would be located within the lower leaf sheath. Transverse section microscopy observations of the older upper zones of the leaf show hyphae increase in structural complexity, wall thickness, the number and thickness of septa and diameter, while basal sections show less complex hyphae. Christensen *et al* (2004) stated that these observations and longitudinal sections showing right-angled lateral branches in the leaf expansion zone are not consistent with an exclusive tip growth model. Christensen *et al* (2004) propose that although initial colonization is most likely by tip growth, growth within the leaf expansion zone may be due to intercalary extension, addition of material along the hyphal axis (Figure 1.1 C); (Christensen and Bennett, 2004b).

1.1.5 Molecular Investigation

Prior to the undertaking of this masterate project a table was constructed of genes (from several other organisms) that have putative functions in hyphal growth of *N. lolii* (Table A1.1). The specific criteria for choosing candidate genes relied on whether the gene had a function to do

with hyphal branching and/or regulation of hyphal extension in several organisms. Three candidate genes were selected:

- Protein Elongation Factor 2 (*EF-2*)
- Cell Division Control Protein 12 – a septin (*CDC-12*)
- Stretch-Activated Calcium Channel (SACC)

The *EF-2* gene was selected as a loss of function mutation resulted in hyper-branching (Propheta *et al.*, 2001) and it could be confirm GUS measurements of Lp19 metabolic activity in planta (Tan *et al.*, 2001). The *CDC-12* gene was selected due to septin intergral function in cell division (Momany *et al.*, 2001; Casamayor and Snyder, 2003) and its possible utility in identifying the Lp19 mode of hyphal extension *in planta* (Christensen and Bennett, 2004b). For both of these genes there were several fungal homologues identified with strong sequence similarity, hence the chance of isolating them using degenerate PCR was quite high. Although few fungal homologues were identified, and amoung those poor sequence similarity, the SACC gene was selected due to its function in transducing mechanical stimuli into growth signals (Garrill *et al.*, 1992).

1.2 *N. lolii* and Protein Elongation Factor 2

As fore-mentioned in Section 1.1.4.2 Tan *et al* (2001) demonstrated a constantly high metabolism *in planta* but conversely a high metabolism only in log-phase *in culture*. These studies were conducted using an *A. nidulans* (foreign) constitutive promoter, attached to a GUS reporter gene. The transformation of a vector containing this construct into *N. lolii* and subsequent plant re-infection is a lengthy process and not always successful (Tan *et al.*, 2001). To offer a more direct method of analyzing metabolism *in culture* and *in planta* the *N. lolii* Protein Elongation Factor 2 (EF-2p; plays an integral part in protein synthesis) was targeted for cloning. As *EF-2* mRNA can be isolated and frozen to allow direct measurement of metabolism in an instant in time, this would be an advantage over the current system in which GUS half-life influences results. If the GUS results are accurate a similar pattern of *EF-2* gene expression should be observed.

1.2.1 EF-2 Structure, Regulation and Modification

EF-2 is a 5-domain protein whose N-terminal portion is a guanine nucleotide binding domain (G-domain) consisting of two parts, a core and insert region which contain the EF-2 GTPase

activity (Kohno *et al.*, 1986). The remaining domains, designated from 2-5 from the N-to the C-terminus, contain various motifs that have functions in toxin binding, interactions with the ribosome and possible RNA binding (Ryazanov and Spirin, 1990; Perentesis *et al.*, 1992; Uchiumi and Kominami, 1994; Rodnina *et al.*, 1999; Shastry *et al.*, 2001; Jorgensen *et al.*, 2003). EF-2 proteins are structurally similar, although slightly larger than *E. coli* EF-G (Jorgensen *et al.*, 2003).

The G-domain of all known EF-2 contains the 6 highly conserved regions: G1-5 and the effector (E) motif. The G regions are highly conserved across all GTPase proteins (Bourne *et al.*, 1991). When an amino acid substitution was made of a glycine for a valine in the G1 region (also known as the P-loop) of c-Ha-ras1, binding of GTP was uninhibited. However GTPase activity was reduced suggesting that this region has important functions in GTP hydrolysis (Seeburg *et al.*, 1984). A number of other studies using c-Ha-ras1 involving substitutions of alanine for threonine in G2, and thiol-modification of cysteine in G5, showed both regions to have functions in GTP-binding (Kohno *et al.*, 1986). Studies in *E. coli* have shown the GTPase function of EF-G to drive the translocation reaction as hydrolysis occurs before the translocation event (Rodnina *et al.*, 1999).

The E motif is a conserved element across all EF-2. Other families of G-proteins have a similar motif that is only conserved within each family and this lead Kohno *et al* (1986) to suggest that this domain may interact with the ribosome. This is a likely scenario as most fungal EF-2 including those from *S. cerevisiae, M. grisea* and *N. crassa* (Perentesis *et al.*, 1992; Propheta *et al.*, 2001; Birren *et al.*, 2003) contain a conserved threonine residue which has been shown by Nairn and Palfrey (1987) to be the specific target of a mitogen activated CaM-dependent protein kinase III. Using elongation assays Nairn and Palfrey observed in rat pancreas cells that when this residue was phosphorylated EF-2 was inactive and removed from the protein synthesis pool (Nairn and Palfrey, 1987; Ryazanov and Spirin, 1990). Studies by Gschwendt *et al* (1989) showed that this inactivation by phosphorylation can be reversed by a type 2A protein phosphatase (Gschwendt *et al.*, 1989). Through this mechanism EF-2 activity can be tightly regulated, interaction of this motif would explain how EF-2 can be inactivated by a single phosphorylation event. However this may not be the sole method regulation, as not all fungal EF-2 contain this threonine residue. It is substituted with a methionine in *C. albicans* and *A. nidulans* (Mendoza *et al.*, 1999; Birren *et al.*, 2003).

Domains 2-5 are less well characterized in function, however they facilitate important interactions to form the active site and have several interactions with inhibitors. The most well characterized form of EF-2 inhibition is the ADP-ribosylation of *S. cerevisiae* and *H. sapiens* diphthamide residue which is unable to perform translocation reactions (Nygard and Nilsson, 1990). Diphthamide is a post-translationally modified histidine 715 at the tip of domain 4; it is unknown why this is modified as yeast mutational studies showed no difference in EF-2 translational properties with the modification, but required the conserved histidine presence for activity. It is suggested that this residue lies within the EF-2 active site, or may be part of the rRNA interaction (Omura *et al.*, 1989).

Observations from structural studies of EF-G demonstrated that domain 4 matches the tRNA anti-codon domain of the 30S subunit of the ribosome. This is consistent with findings that autoimmune antibodies specific for G1959 position of 26S rRNA inactivated EF-2 activity (Uchiumi and Kominami, 1994). This sheds light on the inhibition of ADP ribosylation and implicates domain 4 as interacting directly with 26S rRNA and acting in translocation catalysis.

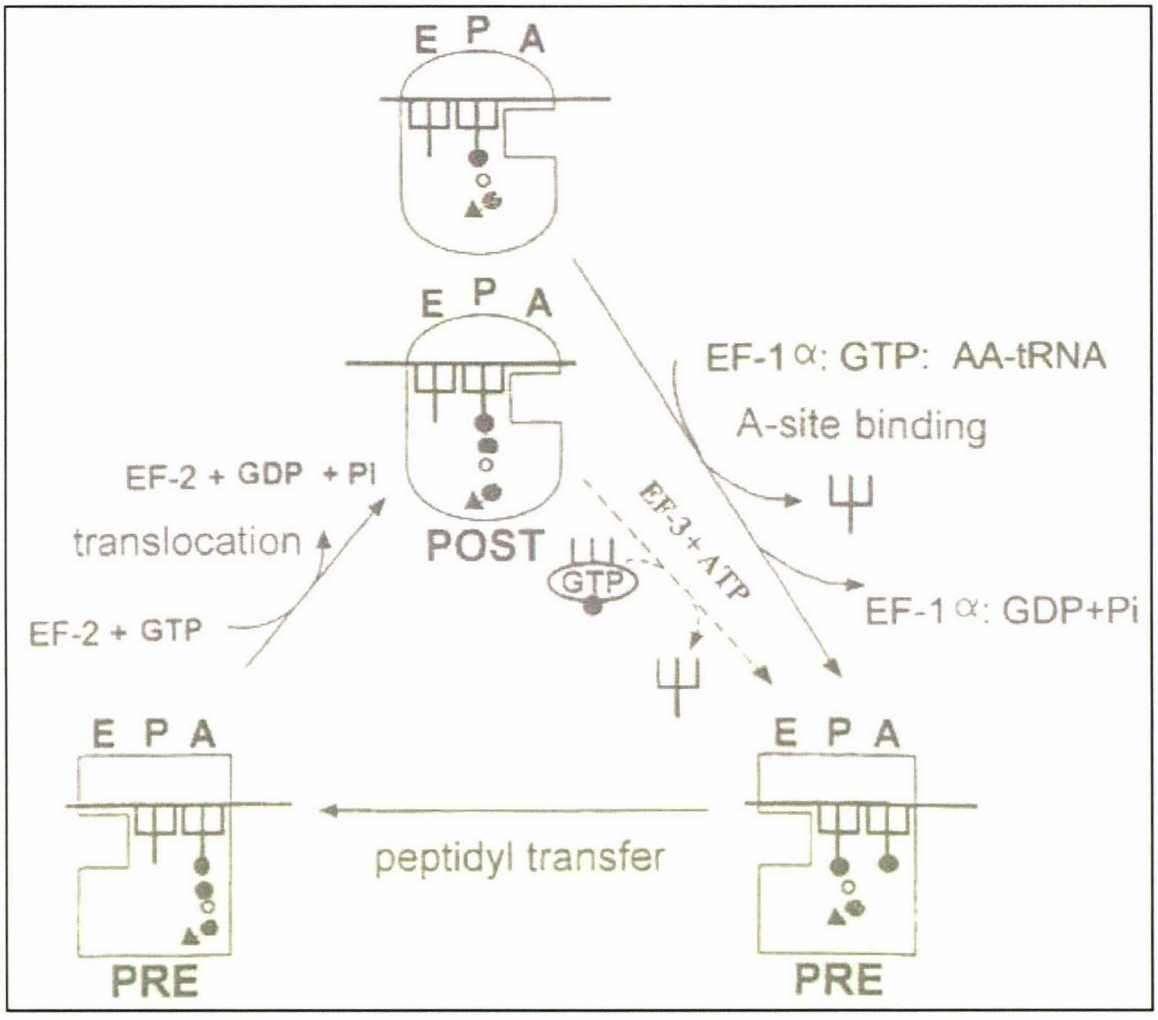
Recently EF-2 has been targeted for novel anti-fungal treatments. This identified a species-specific EF-2 inhibitor, sordarin. Shastry *et al* (2001) identified by mutagenesis three key residues involved with sordarin inhibition of EF-2: tyrosine 521, serine 523 and glutamic acid 524 (all located in domain 3). Substitution of serine resulted in insensitivity where as individual substitution of the tyrosine and glutamic acid resulted in partial sensitivity (Shastry *et al.*, 2001). Jorgensen *et al* (2003) showed that sordarin binding caused a large re-arrangement of domains 3, 4 and 5, explaining the inactivation (Jorgensen *et al.*, 2003).

1.2.2 Protein Synthesis and EF-2

Protein synthesis is a well-documented field and has been under intensive investigation through structural studies for several years. The *S. cerevisiae* 80S ribosome has been crystalized confirming the presence of three transfer RNA (tRNA) binding sites; the aminoacyl (A), peptidyl (P) and exit (E) sites (Triana *et al.*, 1994). Post initiation and following peptidyl transfer in the elongation cycle of protein synthesis, deactylated tRNA is bound to the P site and peptidyl-tRNA in the A site of the ribosome. For the protein elongation cycle to continue the A site needs to be vacated for the next aminoacyl-tRNA to bind to the ribosome (Riis *et al.*, 1990). This pre-translocational intermediate has been demonstrated using kinetic studies in

Figure 1.2: Protein Elongation in Fungi

Schematic representation of the reactions involved in protein synthesis. A, P and E indicate the three tRNA binding sites on the yeast ribosome described by Triana *et al* (1994), pre and post indicate the pre-translocational and post-translocational states of the ribosome. The fork-like symbols indicate tRNAs and all open and closed small circles represent amino acid residues attached to either an aminoacyl-tRNA (single symbol) or peptidyl-tRNA (multiple symbols); diagram from Chakraburtty, 1999.


E P A
E P A
EF-1α: GTP: AA-tRNA
A-site binding
EF-2 + GDP + PI
translocation
POST
GTP
EF-3+ATP
EF-1α: GDP+Pi
EF-2 + GTP
E P A
E P A
peptidyl transfer
PRE
PRE

E. coli to be the substrate of the translocase EF-G (EF-2 in eukaryotes), which catalyses the transfer of pepdityl-tRNA from the A site to the P site in a GTP driven reaction (Rodnina *et al.*, 1999). In most eukaryotes EF-2 is also thought to simultaneously eject the deacetylated-tRNA from the E site upon binding of a new aminoacyl-tRNA to the A site (Riis *et al.*, 1990). However in fungi this rate is slow without the recently discovered EF-3, which is unique to fungi and has been shown to catalyze this reaction using ATP in binding studies (Figure 1.2 for whole process; (Kamath and Chakraburtty, 1989).

1.3 *N. lolii* and Cell Division Control Protein 12

For the intercalary extension model suggested by Christensen *et al* (2004a) to be feasible (as seen in Figure 1.1 C) additional septa would be laid as the hyphae increases in length. Therefore CDC-12 (a septin) was targeted for cloning. Tracking the levels of *CDC-12* expression could enable the identification of areas of active cell division and hyphal thickening.

1.3.1 CDC-12 Structure

CDC-12 is a three-domain protein consisting of an N-terminal, central GTPase domain and a C-terminal coiled coil domain. Few studies have involved CDC-12 mutation as several studies have shown total loss of CDC-12 is lethal (Flescher *et al.*, 1993; Warenda and Konopka, 2002). The CDC-11 protein in *S. cerevisiae* is the most well characterized septin (Casamayor and Snyder, 2003). CDC-12 is part of a family of cell division control proteins found in animal and fungal cells known as septins. They have not been identified in plant cells. Septins are a family of cytoskeletal proteins first identified in *S. cerevisiae*. *S. cerevisiae* septin point or single amino acid mutants in CDC-3, CDC-10, CDC-11, and CDC-12 demonstrated defects in the cell division cycle whose phenotype consisted of defects in bud morphogenesis, and cytokinesis. Since their initial discovery septins have been implicated in a variety of other roles in morphogenesis, interacting with a number of different factors (Longtine *et al.*, 1995).

Deletion of the N-terminal portion of *S. cerevisiae* CDC-11 resulted in elongated cells, defective in cytokinesis and displaying morphology defects. Fusion protein studies showed that CDC-3, 10, 11 and 12 in *S. cerevisiae* bind phosphatidylinositol (3,4,5)-triphosphate, and that

deletion of the basic region in CDC-11 caused loss of binding. Taken together this suggests that the CDC-11 basic region is involved in either vesicle association or membrane interactions. Interestingly in *S. cerevisiae* CDC-11 this is a poly-basic region, while CDC-12 proteins only have a few basic residues (Trimble, 1999; Casamayor and Snyder, 2003).

The G-domain of CDC-12 contains the highly conserved regions called G1-G4, these G regions are highly conserved across all GTPase proteins (Bourne *et al.*, 1991). CDC-12 has not as yet been confirmed as a GTPase, but contains the similar sequence motifs to septins that have. A study with *S. cerevisiae* CDC-11 involving substitutions in the P-loop (G1) of glycine 29, 32 and 34 to alanine resulted in GTP binding but not hydrolysis, while other substitutions (i.e. glycine 230 to glutamatic acid) in regions G2-4 resulted in loss of GTP binding. These types of substitutions with other GTPases have caused identical effects (Section 1.2.1) (Casamayor and Snyder, 2003). The GTPase properties of the *Drosophila melanogaster* septin gene *pnut* was examined by Field *et al* (1996). Studies using [α-^{32}P] GTP showed that 75 % of bound GTP was recovered as GDP, strongly suggesting a GTPase activity (Field *et al.*, 1996).

Co-immunoprecipitation studies in *S. cerevisiae* CDC-11 demonstrated that the C-terminal coiled coil domain interacts with CDC-3 (another septin) and Bem4. Deletion of the CDC-11 coiled coil showed that it is partly responsible for proper localization at the *S. cerevisiae* mother-bud neck (Casamayor and Snyder, 2003).This suggests that septin coiled-coil domains may be involved in both homotypic and heterotypic protein-protein interactions. Interestingly while most septins, including CDC-12, possess a coiled-coil domain, CDC-10 lacks this feature (Momany *et al.*, 2001).

1.3.2 CDC-12 Function

Septins are broadly conserved in sequence. CDC-12 shows a ~26 % amino acid sequence identity with other septins, indicating that septins may not have specialized functions (Longtine *et al.*, 1995), however curiously both CDC-12 and CDC-3 have been shown to be essential in *S. cerevisiae*, *C. albicans* and *A. nidulans* (Flescher *et al.*, 1993; Momany *et al.*, 2001; Warenda and Konopka, 2002).

Septins have been long implicated as key proteins involved in cytokinesis. As the fungal cell divides a process known as septation occurs, a division of mother/daughter cells by formation

of a 10 nm ring complex called a 'septum'. These septa are typically every 10-15 μm in yeast (Longtine *et al.*, 1995) and 20 μm in *A. nidulans* (Momany *et al.*, 2001).

Haarer and Pringle (1987) using immunofluorescence demonstrated CDC-12 to be in the vicinity of the 10 nm ring of *S. cerevisiae* mother-bud neck (Haarer and Pringle, 1987), implicating CDC-12 involvement in the formation or a integral part of the septum complex. Field *et al* (1996) further confirmed septin function in *D. melanogaster* embryos using affinity chromotograpy with the localization of the septin *pnut* in septal rings (Field *et al.*, 1996). Mutation analysis in any one of the four *S. cerevisiae* septins causes a loss of septa localization at the neck (Flescher *et al.*, 1993), further reinforcing CDC-12 function as a component of the 10 nm ring. Recently, Westfall *et al* (2002) demonstrated the involvement of septins in the septation of the hyphal fungus *A. nidulans* using immunofluorescence. The CDC-3 (AspB gene) was fused to a GST-tag and shown to co-localise postmitotically with F-actin, showing CDC-3 was part of the septin ring. The pattern of septation was shown to be in a parasynchronous wave (septum forms from the most-to-least mature part of the growing hypha), and AspB was shown to localise premitotically to branch points, making it the only known branching site marker (Westfall and Momany, 2002). Taken together with immunofluorescence studies by Momany *et al* (2001) that demonstrated that CDC-12, -10 and -11 co-localise with CDC-3, these findings support the hypothesis that all four septins may play an integral part in branch formation and recruiting material to the new insertion point. Septins were also demonstrated to polymerise *in vitro.* Using immunofluorescence Field *et al* (1996) also proved *D. melanogaster* septins *pnut, Sep1* and *Sep2* associate in a complex forming 7 nm wide by 26 nm long filaments *in vitro*. This is suggested to correlate with the 10 nm wide by 28 nm long filaments observed *in vivo*, further implicating that the 10 nm ring is at least in part composed of a polymerized septin complex (Field *et al.*, 1996).

After formation the septin ring is assembled. *In vitro* mutagenesis studies of Cdc42p indicated that upon Cdc42p loss of GTPase function, septin 10 nm rings were wider and less organized than those found in wild-type in *S. cerevisiae*, suggesting that filament assembly into the characteristic septa is partially conducted by cdc42p (Gladfelter *et al.*, 2002). Similar observations were made in Gin4 kinase mutant cells, where instead of a 10 nm ring complex, they form 6-8 parallel bars lacking the uniform array and implicating Gin4 kinase as interacting with septin filaments (Carrol *et al.*, 1998).

Septins have also been shown to interact with proteins other than septins, generating speculation about septin function other than cytokinesis. Interestingly most of these interactions are with the essential CDC-12 and CDC-3. Recently GFP fluorescence has shown CDC-12 and CDC-3 to co-localize and interact with Int1p in *C. albicans* (Gale *et al.*, 2001). Int1p presence has been demonstrated through mutagenesis to cause a morphological switch from isotropic growth to highly polarized cells in *S. cerevisiae* (Gale *et al.*, 1996).

Septins have also been implicated in vesicle transport. Hsu *et al* (1998) found septins in Rat brain co-immunoprecipitated with the rSEC 6/8 complex (exocyst in *S. cerevisiae*; (Hsu *et al.*, 1998)), implicating a role in vesicle transport, this interaction could possibly be through the septin N-terminal basic domain. This also supports the suggestion by Westfall *et al* (2002) that septins may be involved in recruiting proteins and directing vesicles to pre-determined branch sites in *A. nidulans*.

1.4 *N. lolii* and Stretch Activated Calcium Channels

Both Christensen *et al* (2004a) and Tan *et al* (2001) hypothesize that in order for *N. lolii* hyphae to sense when the host is growing mechanosensitive channels may be involved in sensing surrounding cellular friction. These would be located at the tip of the growing hyphae sensing friction if either the hyphae is growing too fast, or if the host is growing faster and the hyphae need to increase speed of cell division; hence as Tan *et al* (2001) stipulates "-no net movement of tip versus surrounding tissue, and no friction would occur." An attractive proposition, if intercalary extension is true, is that the hyphae could be sensing stretch via mechanosensitive channels from attachments to plant cells. As previously mentioned in Section 1.1.5.1 calcium is an important signaling factor in hyphal growth, therefore a SACC was targeted for cloning. The existence of these mechanosensitive channels is well documented using patch-clamp experiments in fungi (Garrill *et al.*, 1992; Watts *et al.*, 1998), however comparatively few been isolated.

1.4.1 SACC Structure

Stretch-activated calcium channel structure is not well characterized in eukaryotes, so current information is drawn from studies of the large mechanosensitive channel MscL which has been

cloned and crystallized in *Mycobacterium tuberculosis* and *Esherichia coli* (Blount *et al.*, 1996; Sukharev *et al.*, 2001).

MscL is an L-type calcium channel consisting of five identical subunits which each contain a transmembrane and an amino terminal cytoplasmic helix. Each subunit is arranged as part of a ring, with the cytoplasmic helicies forming a 'gate' in the cell cytoplasm. Mutagenesis analysis of cysteine residues revealed that a cytoplasmic helix S1 crosslinks through phenylalanine 7 and 10 in S1 when the channel is closed. When the membrane is under stress (lateral tension) S1 residue isoleucine 3 interacts with isoleucine 96 of transmembrane helix M2 in an open conformation, this is stabilized by a cluster of phenylalanine residues in S1 (see Figure 1.3 (Sukharev *et al.*, 2001)).

Gd^{3+} was demonstrated to be a potent inhibitor of SACCs *N. crassa* hyphae. A decay of channel activity was seen in the presence of inhibitor, this was due to the effect of Gd^{3+} and not channel rundown, as Gd^{3+} had no effect on a spontaneous K^+ channel tested concurrently (Levina *et al.*, 1995).

1.4.2 SACC Function

Stretch-activated calcium channels are a key responsive element for hyphae in response to membrane stress (Garrill *et al.*, 1992; Watts *et al.*, 1998). Known *SACC* functions in fungi are regulation of cell growth and mating in *S. cerevisiae* and *S. pombe*. Patch-clamp studies using *Saprolegnia ferax* demonstrated that *SACCs* allow calcium into the cell upon application of membrane tension (Garrill *et al.*, 1992). Calcium has been invoked as an important factor in both tip growth and hyphal branching (Reissig and Kinney, 1983). Ca^{2+} ions are well suited to such a role, given their ubiquitous functions as second messengers in eukaryotic cells (Stryer, 1995). In the hypha, Ca^{2+} is thought to play an important role in regulating apical extension through polarized growth (Takeuchi *et al.*, 1988; Harold, 1990). However recently it has been shown in voltage clamping experiments on *Neurospora crassa* hyphae that there is no correlation between the size of transmembrane fluxes of Ca^{2+} and hyphal growth rates (Silverman-Gavrila and Lew, 2000). This leads to examination of intracellular Ca^{2+} for an obligatory role in tip growth. Schmid and Harold (1988) demonstrated that the growth of *N. crassa* hyphae is strongly dependent on the concentration of exogenous Ca^{2+} present in the media. It was observed that as Ca^{2+} concentrations approached 0.1 μmol/L, a severe reduction

in hyphal elongation occurred. At 0.1 μmol/L 20% of the population lost capacity for polarized growth, producing spherical cells up to 20 μm in diameter (Schmid and Harold, 1988). Similar data were obtained in *Saprolegnia ferax* by Jackson and Heath (Jackson and Heath, 1989). Thus *SACCs* may be required to supply exogenous calcium to maintain intracellular Ca^{2+} levels at a sufficient level to allow continued apical extension.

C. albicans represents a good example of sensory growth using *SACCs*. Using the *SACC* specific inhibitor Gd^{3+}, *C. albicans* hyphae were unable to sense topographical changes (Watts *et al.*, 1998). This suggests that SACCs may be required for *C. albicans* in a thigmotrophic response.

1.5 Aims of this study

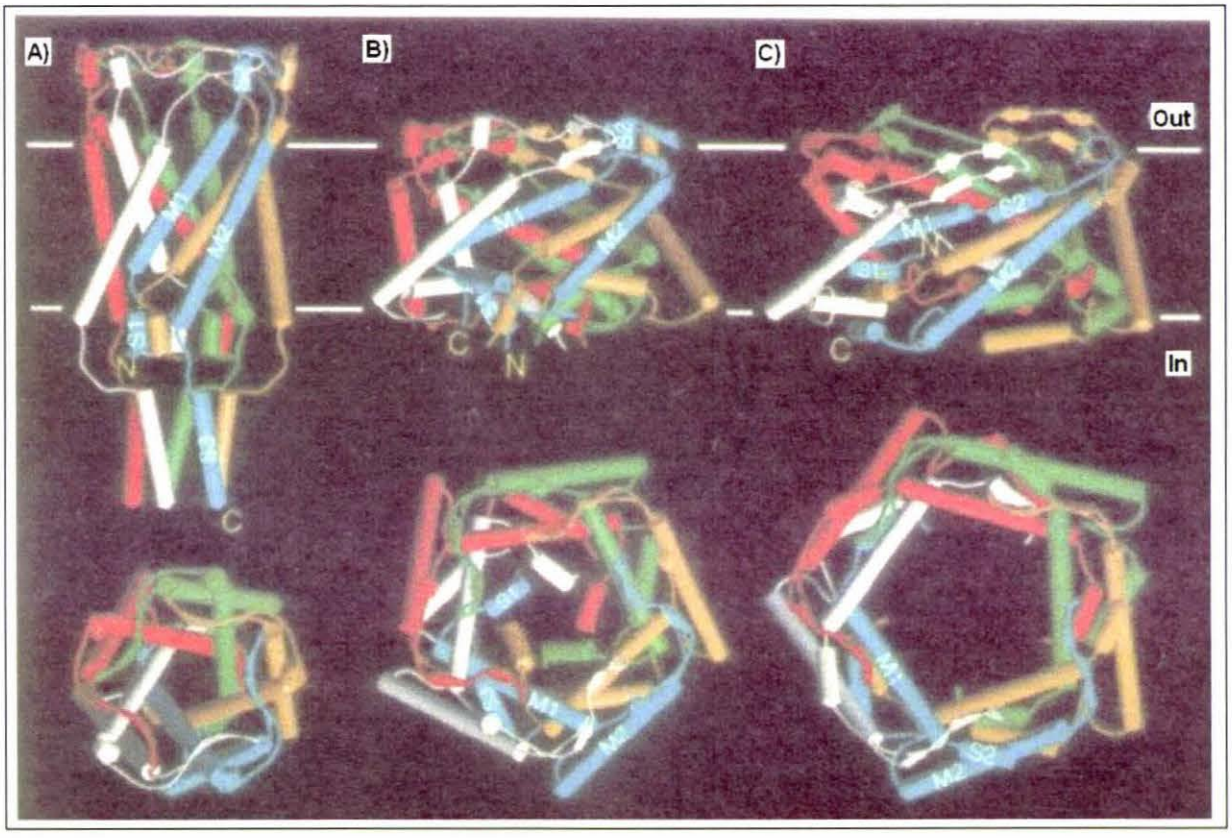
Due to unique growth characteristics and large effect on agriculture in New Zealand the *N. lolii* and *L. perenne* symbiosis is a dynamic field of study with a host of possible commercial applications not to mention greater understanding of fungal growth. Identification of the *N. lolii* signaling and growth mechanisms that regulate its growth and morphology *in planta* will extend our knowledge about filamentous fungi currently used in biotechnology. Advantages to be gained from this may be manipulation of hyphal extension in pathogenic fungi to slow or cease growth, development of antifungal agents and the ability to manipulate a fungus to produce secondary metabolites without resorting to adverse conditions. The aim of this project was to clone *EF-2*, *CDC-12* and *SACC* genes from the haploid *N. lolii* isolate Lp19 and to determine their expression *in culture*. This will allow expression of these genes to be determined *in planta* and offer some explanation about the *N. lolii* growth phenomenon *in planta*.

Figure 1.3: Molecular Model of SACC Opening

A) Resting conformation; Channel is closed, S1 phenylalanines closely associated.

B) Closed/Open conformation; Membrane is under low tension pore is slightly open

C) Open conformation; isoleucines interact from S1 and M1, pore is open. Diagram is taken from Sukharev *et al* 2001.


A)
B)
C)
Out
In
M1
M2
S1
S2
N
C

Chapter 2

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Origin of Materials

Yeast extract and Agar were purchased from Oxoid, Australia; all other organism growth media bases and components were purchased from Difco, BD Biosystems, NJ, USA.

Tris, Triton X-100, Lysozyme, MOPS, EDTA, SDS, Hoechst 33258 dye and calibration standards, Tween-80, Tween-20, dNTPs, X-Gal, IPTG, Ethidium Bromide, Ampicillin, GenElute mRNA preparation kits were purchased from Sigma Chemical Company, St Louis, USA.

Tri-reagent® for total RNA isolation (MRC) from ProGENZ Limited, Auckland, New Zealand.

Cellophane circles and 3MM paper was purchased from Whatman, Maidstone, England.

Restriction endonucleases and buffers and T4 DNA ligase were from either: New England Biolabs, MA, USA, or Roche Molecular Biochemicals, IN, USA.

Positively charged nitrocellulose filters, The DIG labeling System and labeled RNA quantification standards, DNA quantification standard, Expand Reverse Transcriptase kit for cDNA synthesis and High-pure Plasmid Isolation kits were purchased from Roche Molecular Biochemicals, IN, USA.

Taq-polymerase, PCR Purification and Gel Isolation (QIAquick) kits were purchased from Qiagen New Zealand distributors: Biolab Scientific Ltd, Albany, Auckland, New Zealand.

The cloning vector pGEM-T Easy was purchased from Promega Corporation, WI, USA

DNase, Random primers for RT PCR, Agarose powder for electrophoresis, GeneRacer kits for 5' and 3' RACE and all custom Oligonucleotides were ordered from Invitrogen Corporation, Invitrogen NZ limited, Penrose, Auckland, New Zealand.

Sucrose, NaCl, Glucose and all other general chemicals were purchased from BDH Laboratory Supplies, Poole, England

Microcentrifuge tubes were purchased from Axygen, Global Science and Technology Ltd, Glenfield, Auckland New Zealand; Cryotubes were purchased from Nunc Inc, Naperville, IL, USA. All other disposable plastic-tubes were purchase from Greiner Vacuette North America Inc, NC, USA.

2.1.2 Organism Strains and Plasmids

Organism strains, plasmids and donated DNA used in this study are listed in Table 2.1

2.1.3 Water Supply and Sterilization

Water (dH_2O) used in the preparation of all solutions was purified using a Millipore MilliQ Reagent water system. All equipment and solutions (unless otherwise specified) were sterilized at 121°C for 15 minutes in commercial autoclaves by various Technicians at IMBS; Massey University.

2.1.4 Media

Protocols for bacterial media preparation were obtained from Current Protocols In Molecular Biology (Ausubel *et al.*, 1997). *Neotyphodium lolii* protocols were obtained from Mike Christensen; AgResearch Grasslands Palmerston North (Christensen *et al.*, 1993) and Ningxin Zhang; IMBS Massey University Palmerston North. *Aspergillus nidulans* protocols were obtained from Pontecorvo *et al* (Pontecorvo *et al.*, 1953).

Table 2.1: Strains and Plasmids

Strain or Plasmid	Relevant Characteristics	Source or Reference
Fungal Strains		
***Neotyphodium* sp. (LpTG-1)**		
Lp19	*N. lolii* from *Lolium perenne*	Christensen *et al* (1993)
Aspergillus nidulans		
A57	*wild type*	Bradshaw *et al* (2001)
Bacterial Strain		
Escherichia coli		
DH5-α	*supE44, lacU169, 80 lacZ, hsdR17, recA1, endA1 gyrA96, thi-1, relA1*	Dr Max Scott (IMBS); Gibco
Plant Cultivars		
Lolium perenne		
NuiLp19	*N. lolii* (Lp19) infected tissue; *wild type*	Christensen *et al* (1993)
Nui	Non-infected tissue; *wild type*	Christensen *et al* (1993)
Plasmids		
pBR322	Ampicillin resistance	Ningxin Zhang (IMBS); Gibco
pUC118	Ampicillin resistance, *lacZ'*	Dr Barry Scott (IMBS); Gibco
pGEM-T Easy	Ampicillin resistance, interruptable *lacZ'*	Promega
pMM16	Contains *N. loii* (Lp19) β-tub full cDNA	Bryant (2004)
pDEF-1	*NEF-2* Degenerate PCR product (2066 bp) ligated into pGEM-T Easy	This study
pDEF-2	*NEF-2* Initial IPCR product (2042 bp) ligated into pGEM-T Easy	This study
pDEF-3	*NEF-2* Second IPCR product (3056 bp) ligated into pGEM-T Easy	This study
pDEF-4	*NEF-2* 5' RACE product (1083 bp) ligated into pGEM-T Easy	This study
pDEF-5	*NEF-2* 3' RACE product (2073 bp) ligated into pGEM-T Easy	This study
pDSEP-1	*nCDC-12* Degenerate PCR product (514 bp) ligated into pGEM-T Easy	This study
pDSEP-2	*nCDC-12* IPCR product (1994 bp) ligated into pGEM-T Easy	This study
pDSEP-3	*nCDC-12* 5' RACE product (736 bp) ligated into pGEM-T Easy	This study
pDSEP-4	*nCDC-12* 3' RACE product (1080 bp) ligated into pGEM-T Easy	This study
pDSACC-1	*A. nidulans* SACC probe (156 bp) ligated into pGEM-T Easy	This study

2.1.4.1 Bacterial Media

2.1.4.1.1 Luria Broth (LB) Medium

Contained (g/L): Tryptone, 10.0; Yeast extract, 5.0; NaCl, 5.0; NaOH 1 ml (1 mol/L). The pH was adjusted to 7.5 prior to autoclaving. Solid media were prepared by addition of agar (20 g/L). Where required media was supplemented with 100 µg/ml Ampicillin, 40 µg/ml isopropyl-β-D-galactoside (ITPG) and 60 µg/ml 5-bromo-4-chloro-3-indoyl-β-D-galactoside (X-Gal). Molten, sterilized solid media was cooled to 55°C before addition of the fore-mentioned supplements.

2.1.4.1.2 SOC Medium

Contained (g/L): Tryptone, 20.0; Yeast extract, 5.0; NaCl, 10 ml (1 mol/L); KCl, 2.5 ml (1 M); $MgCl_2.6H_20$, 0.95; $MgSO_4.7H_20$, 2.5; Glucose 3.6.

2.1.4.1.3 SOB Medium

Contained (g/L): Tryptone, 20.0; Yeast extract, 5.0; NaCl, 0.5; KCl 2.5 ml (1 M); $MgCl_2$, 5.0 ml (2 M) added just before use

2.1.4.2 Fungal Media

2.1.4.2.1 Potato Dextrose (PD) Broth Medium (*N. lolii*)

Potato dextrose broth (24 g) rehydrated in 1 L of water. The pH was adjusted to 6.2 using NaOH prior to autoclaving. Solid media were prepared by addition of agar (20 g/L).

2.1.4.2.2 MYG Medium (*A. nidulans*)

Contained (g/L): Malt extract, 5.0; Yeast extract, 2.5; D-Glucose, 10.0. The pH was adjusted to 6.5 prior to autoclaving. Solid media were prepared by addition of agar (20 g/L).

2.1.5 Buffers and Solutions

Protocols for bacterial media preparation were obtained from Current Protocols In Molecular Biology (Ausubel *et al.*, 1997).

2.1.5.1 Electrophoresis Buffers

2.1.5.1.1 1 x TAE Buffer

Contained 40 mM Tris-HCl, 2 mM $Na_2EDTA.2H_2O$ and 20 mM acetic acid

2.1.5.1.2 50 x TAE Buffer

Contained (g/L): Tris base, 242.0; Acetic acid, 57.1 ml; Na_2EDTA, 37.2.

2.1.5.1.3 1 x TBE Buffer

Contained 89 mM Tris-HCl, 2 mM Na_2EDTA and 89 mM Boric acid.

2.1.5.1.4 10 x TBE Buffer

Contained (g/L): Tris base, 162.0; Na_2EDTA, 9.5; Boric acid, 27.5.

2.1.5.1.5 SDS Loading Buffer

Contained 1 % (w/v) sodium dodecyl sulphate (SDS), 0.02 % Bromophenol blue, 20 % (w/v) sucrose and 5 mM Na_2EDTA (pH 8.0).

2.1.5.1.6 Ethidium Bromide (EtBr)

The solution prepared for use in staining agarose gels consisted of 1 µl of a 10 mg/ml EtBr stock solution per 10 ml of MilliQ H_2O (final concentration of 1 µg/ml).

2.1.5.2 DNA Isolation

2.1.5.2.1 Lysis Buffer

Contained (g/L): Tris, 4.846; Sodium acetate, 1.64; Na_2EDTA, 0.37; SDS, 5.0. The pH was adjusted to 7.8 prior to autoclaving.

2.1.5.2.2 TE Buffer (10/0.1)

Contained 10 mM Tris-HCl and 0.1 mM Na_2EDTA. The pH was adjusted to 8.0 prior to autoclaving.

2.1.5.3 Southern Blotting (SB) Solutions

2.1.5.3.1 Solution 1 (SB 1)

Contained 0.25 M HCl

2.1.5.3.2 Solution 2 (SB 2)

Contained 0.5 M NaCl and 0.5 M NaOH

2.1.5.3.3 Solution 3 (SB 3)

Contained 0.5 M Tris-HCl and 2.0 M NaCl

2.1.5.3.4 20 x SSC

Contained 3.0 M NaCl and 0.2 M Sodium citrate. 2 x SSC was prepared by a 1:10 dilution of 20 x SSC in MilliQ dH_2O.

2.1.5.4 RNA Treatment

2.1.5.4.1 DNase I

DNase I was prepared in a storage buffer containing 20 mM Sodium acetate, 5 mM Calcium chloride, 0.1 mM Phenylmethanesulfonyl fluoride (PMSF) and 50% (v/v) Glycerol.

2.1.5.4.2 10 x MOPS Buffer

Contained (g/L): MOPS (3-[N-morpholino] propanesulphonic acid), 41.2; Sodium acetate 3-hydrate, 10.9; Na_2EDTA, 3.7. The pH was adjusted to 7.0 with NaOH, stored at room temperature and protected from all light sources.

2.1.5.5 DIG Detection/Hybridization Solutions (DS)

2.1.5.5.1 Blocking Solution (DS 1)

Contained 10.0 % blocking reagent in DS 2

2.1.5.5.2 Hybridization Solution (Standard)

Contained 5 x SSC, 0.1 % (w/v) Sodium lauroylsarcosine, 0.02 % (w/v) SDS and 1.0 % Blocking solution.

2.1.5.5.3 DIG Buffer 1 (DS 2)

Contained 150.0 mM NaCl and 100 mM Maleic acid at pH 7.5

2.1.5.5.4 DIG Buffer 2 (DS 3)

Contained 1.0 % (v/v) DS 1 in DS 2

2.1.5.5.5 Antibody Solution (DS 4)

Contained a 1:10,000 fold dilution of Anti-digoxigenin-AP Fab fragments in DS 3

2.1.5.5.6 DIG Buffer 3 (DS 5)

Contained 100 mM Tris-HCl and 100 mM NaCl, the pH was adjusted to 9.5 with NaOH

2.2 METHODS

Protocols were obtained from Current Protocols In Molecular Biology (Ausubel *et al.*, 1997) unless otherwise stated.

2.2.1 Growth of Cultures

Protocols for bacterial growth were obtained from Current Protocols In Molecular Biology (Ausubel *et al.*, 1997). *Neotyphodium lolii* protocols were obtained from Mike Christensen (Christensen *et al.*, 1993) and Ningxin Zhang. *Aspergillus nidulans* protocols were obtained from Pontecorvo *et al* (Pontecorvo *et al.*, 1953). Liquid cultures were incubated at 200 rpm on an orbital shaker (Lab-Line).

2.2.1.1 *E. coli* cultures (DH5-alpha)

E. coli DH5-alpha cultures were grown at 37°C overnight on LB plates or in LB broth with appropriate selection supplements (Section 2.1.4.1.1). When necessary, for short-term preservation (1-2 days) plates were sealed with parafilm and stored at 4°C. For long-term preservation, glycerol stocks were prepared by pelleting cells from an overnight liquid culture of a single clone by a 1 min centrifugation in 1.5 ml microcentrifuge tubes. The pellets were pooled and resuspended in 1 ml of 20 % glycerol and stored at -70°C.

2.2.1.1 *N. lolii* cultures (Lp19)

2.2.1.1.1 Growth on Solid Media by Subculture

Lp19 was isolated by Mike Christensen (AgResearch; Palmerston North) from its host *L. perenne* and grown at 22°C for 4 weeks on PD plates (Section 2.1.4.2.1). Approximately 5 mm x 5 mm sections of established colonies were excised for subsequent cultures on PD. These plate cultures were sealed with parafilm and stored at 4°C

2.2.1.1.2 Growth in Liquid Culture for DNA Isolation

A single 4 week old colony (Section 2.2.1.2.1) was macerated using a sterile scalpel blade and grown at 22°C for 10 days in 125 ml PD broth. This is also a pre-culture for growth curve cultures.

2.2.1.1.3 Growth Curves 1 (Zhuojian Liu at IMBS; Massey University Palmerston North)

A) Mycelium was harvested from a pre-culture (Section 2.2.1.2.2) by a 1 minute centrifugation using a Biofuge 13 (Heraeus) in 5 microcetrifuge tubes. The pellet was homogenized with 0.1 ml saline solution (NaCl 8.0 g/L), and re-centrifuged. The pellet was subsequently re-suspended, washed and centrifuged three times with 1 ml saline solution. After the final centrifugation 2 microcentrifuge tubes of mycelium were transferred to 10 ml tubes, individually re-suspended in potassium phosphate buffer (20 mM potassium phosphate; pH 5.2) and starved for 5 days at room temperature. The remaining 3 tubes of mycelium were freeze-dried overnight to measure dry weight.

B) The two tubes of starved mycelium from A) were independently washed with saline solution (as in A)). An estimated 1.0 g of mycelium was used to inoculate 500 ml PD broth; this culture was grown at 22°C until no more growth was seen.

C) Samples of the culture in B) were taken every 4 days during lag phase (30 ml) and every 2 days from early exponential phase to stationary phase (20 to 6 ml progressively). Samples were taken until 6 days into stationary phase.The sample volume was halved; one half was harvested by centrifugation using a Labofuge (Heraeus) and freeze-dried overnight to measure dry weight to construct a growth curve. The remaining half was filtered using nappy liners (Johnson and Johnson), washed with MilliQ DEPC-treated dH_2O and freeze-dried for RNA isolation (Section 2.2.16). All growth curves were carried out in duplicate.

2.2.1.1.4 Growth Curves 2

Two changes were made to the procedure outlined in Section 2.2.1.2.3. Firstly, the larger colonies were avoided from the pre-culture (Section 2.2.1.2.2) and secondly, the starvation in Section 2.2.1.2.3 B) was omitted.

2.2.1.2 *A. nidulans* cultures (A57)

2.2.1.2.1 Preparation of Spore Suspension

Spores were isolated from a sporulating culture by scraping the surface with a sterile inoculation loop and suspended in a Tween 80/Saline solution (containing (g/L): NaCl, 8.0; Tween 80, 0.025 % (v/v)). Spores were harvested by a 5-minute centrifugation at 6000 g, washed twice with sterile MilliQ and resuspended in Saline solution (NaCl 8.0 g/L). The concentration was estimated using a haemocytometer (Nuybauer) and stored at -20°C.

2.2.1.2.2 Culture

A spore suspension was spread onto sterile cellophane disks overlaid on MYG plates at a concentration of approximately 4.92 x 10^6 spores per plate (Section 2.1.4.2.2). The inoculated plates were incubated at 37°C overnight and then transferred to 30°C for the following day. The transfer to 30°C is a tempory inhibition of sporulation and encourages further hyphal growth for DNA isolation (Pontecorvo *et al.*, 1953).

2.2.2 DNA and Plasmid Isolation

2.2.2.1 DNA Isolation from Fungal Cultures by the Al-Samarrai *et al* (2000) Method

In a pre-cooled mortal and pestle approximately 30 mg of freeze-dried mycelia was ground to a fine powder under liquid nitrogen and resuspended in 500 µL lysis buffer (Section 2.1.5.2.1). Cell debris was precipitated by adding 165 µL NaCl (5 M) and the solution was inverted 10 times. This was centrifuged at 4°C for 20 minutes (all centrifugations were carried out at 13,000 g), the resulting supernatant transferred to a new tube and 400 µL chloroform added. The solution was inverted 50 times and centrifuged at room temperature for 20 minutes. The aqueous phase was transferred to a new tube, and the addition of chloroform repeated. Two volumes of 95 % ethanol were added, mixed, and centrifuged at room temperature for 5 minutes. The supernatant was discarded, precipitate resuspended in 500 µL lysis buffer and mixed. An equal volume of chloroform was added, mixed and centrifuged at room temperature for 10 minutes. The aqueous phase was transferred to a new tube, two volumes of cold (-20°C) 95 % ethanol added, mixed and centrifuged at room temperature for 2 minutes. The supernatant was discarded and this process was repeated three times with cold (-20°C) 70 % ethanol (Al-Samarrai and Schmid, 2000). The precipitated DNA was allowed to dry at room temperature before resuspension in 50 µL TE buffer (Section 2.1.5.2.2).

2.2.2.2 DNA Isolation from Infected Plant Tissue by an Adaptation of the Al-Samarrai *et al* (2000) Method

In essence this process is identical to the protocol described in Section 2.2.2.1, with two adjustments. Firstly, the organic phase from the chloroform treatment in RNA extraction from infected plant tissue (Section 2.2.16.1) was used and in this RNA was extracted from 1.5 g of

tissue, thus the quantities of solutions used were up-scaled to suit. Secondly, the process follows on from cell debris precipitation (addition of 5 M NaCl).

2.2.2.3 Large Scale Plasmid Isolation by Rapid Boil Method by Holmes and Quigley (1981)

Rapid boil plasmid preparations generate impure DNA extractions, but this method is sufficient to screen clones. Selected transformant cells (Section 2.2.12.3) from a 5 mL overnight *E. coli* culture (Section 2.2.1.1) were pelleted by centrifugation at 12,000 g for 1 minute. The supernatant was removed and the pellet was resuspended in 350 µL of STET (8% (w/v) sucrose, 50 mM EDTA, 50 mM Tris pH 8.0, 5% (v/v) Triton X-100) with 25 µL of freshly prepared lysozyme (10 mg/mL). The tubes were then placed into a boiling water bath for 40 seconds, followed by a 10-minute centrifugation at 12,000 g. The resulting gelatinous pellet was then removed and an equal volume of isopropanol was added to the supernatant, which was mixed and placed at -70°C for at least 30 minutes, to precipitate the DNA. The DNA was pelleted by centrifugation for 20 minutes at 12,000 g at 4°C. The supernatant was removed by aspiration and the pellet washed with 500 µL cold (-20°C) 95% ethanol, and centrifuged for 10 minutes (12,000 g at 4°C). Finally, all the supernatant was removed and the pellet dried and resuspended in 50 µL TE (Section 2.1.5.2.2). Samples were stored at -20°C until required (Holmes and Quigley, 1981).

2.2.2.4 Small Scale Plasmid Isolation

Small-scale plasmid DNA preparation was used when a small quantity of high quality plasmid DNA was required for PCR or sequencing. A 5 mL overnight culture was pelleted and plasmid DNA isolated using a High Pure® Plasmid Isolation Kit (Roche) according to the manufacturer's instructions. This system uses an alkaline lysis (containing SDS) to release plasmid DNA from the cell. Purification is achieved by plasmid DNA selectively binding to glass fibre matrix within the filter tubes in the presence of the chaotrophic salt guanidine-HCl. Samples were stored at -20°C until required.

2.2.3 Purification of DNA

2.2.3.1 Phenol/Chloroform Extraction

Phenol-chloroform extraction was used to purify DNA from proteins in solution. When chloroform is used in conjunction with phenol it improves the effectiveness of nucleic acid

extractions, due to its ability to denature proteins, remove lipids and enhance separation of the two phases (due to its high density; (Ausubel *et al.*, 1997)).An equal volume of Tris-equilibrated phenol/chloroform/isoamyl alcohol (25:24:1) was added to a DNA sample, mixed thoroughly and centrifuged at 13,000 g for 10 seconds. The DNA was precipitated with ethanol as described in section 2.2.3.2; the pellet dried and resuspended in 25 µL MilliQ dH_2O.

2.2.3.2 Ethanol Precipitation

To concentrate DNA in solution, one-tenth volume of 3 M sodium acetate (pH 5.2) and two volumes cold (-20°C) 95 % ethanol were added, mixed and left on ice for 30 minutes. The DNA was pelleted by centrifugation at 13, 000 g for 5 minutes and the supernatant discarded. The pellet was washed once with 70 % ethanol and allowed to dry before resuspension in 25 µL TE buffer (Section 2.1.5.2.2) or MilliQ dH_2O.

2.2.3.3 Agarose Gel Purification

Purification of DNA fragments was achieved by using agarose gel electrophoresis. Purification of small fragments (0-1 kb) was conducted using a purification kit, however this resulted in inadequate yields for larger fragments. In light of this, the Freeze-Squeeze technique was employed to purify larger fragments (Thuring *et al.*, 1975).

2.2.3.3.1 Purification of Small Fragments (0-1.5 kb)

The required DNA was excised from the gel and extracted using QIAquick® gel purification kit (Qiagen) as described by the manufacturers. This kit uses a system of spin cartridges containing a silica membrane. Initially the agarose gel is dissolved by a sodium perchlorate solution provided by the manufacturer, freeing DNA to adhere to the silica support. After washes with ethanol-based buffers, the DNA is released upon elution with TE buffer.

2.2.3.3.2 Purification of Large fragments (1.5 kb plus)

This was achieved using the Freeze-squeeze technique by Thuring *et al* (1975). The required DNA was excised from the gel and placed between parafilm at -20°C overnight. DNA from the gel was extracted by squeezing, the liquid collected and 50 µL 1 x TAE buffer (2.1.5.1.1) added to the agarose. The gel was re-frozen for 30 minutes, squeezed and the liquid pooled with the first extraction. The extract was centrifuged at 7,000 g for 10 seconds and supernatant in a new tube. This extraction was then ethanol precipitated (Section 2.2.3.2).

2.2.3.4 Purification of DNA from a PCR Reaction

PCR products were purified from PCR components before they could be used for probes, in ligations, or direct sequenced. This was achieved using QIAquick PCR purification kit (Qiagen). DNA binds to a silica resin filter within cartridge housing and eluted with 25 µL TE buffer (Section 2.1.5.2.2).

2.2.4 Nucleic Acid Quantification

2.2.4.1 Determination of DNA Concentration by Fluorometric Assay

Fluorometry is an accurate way to quantitify DNA and is more sensitive than most Spectrophotometric procedures. The Hoechst 33258 fluorochrome specifically binds DNA and has little affinity for RNA.

All DNA samples prior and subsequent to experimentation were quantitated using a TKO 100 fluorometer (Hoefer) according to manufacturer's protocol. The exitation and emission wavelengths were 365 and 460 nm respectively; the machine zeroed until a steady state was reached using 2 mL of a dye solution containing 1 x TNE buffer (10 mM Tris-HCl, 1 mM Na_2EDTA and 100 mM NaCl, pH 7.4) and 0.1 µg/mL Hoechst 33258 dye. The scale was set to 100 using 2 µL of 100 µg/mL calf thymus DNA added to 2 mL of dye solution. Once the scale was reliably set, 2 µL of sample DNA was added to 2 mL of dye solution and the resulting value recorded as concentration of the sample in ng/µL.

2.2.4.2 Determination of RNA Concentration by Spectrophotometric Assay

All RNA samples prior and subsequent to experimentation were quantitated using a ND-1000 UV-Vis Spectrophotometer (NanoDrop®) according to manufacturer's protocols. The instrument was zeroed by applying 2 µL DEPC MilliQ dH_2O to the mounting pedestal. RNA sample absorbance was measured in the same way at both 260 and 280 nm. The Nanodrop 2.5.0 software program indicated the purity by calculating the 260/280 nm ratio (where pure RNA is 2.0), and concentration of RNA to be calculated, using the relationship: A_{260} x 50 µg/mL. RNA preparations that resulted in a 260/280 ratio of 1.9-2.1 were considered useful for future experiments.

2.2.5 Agarose Gel Electrophoresis

Agarose gel electrophoresis is a method that separates DNA by charge and size. Linear DNA is negatively charged in an electrical field near neutrality (pH 8.0), and migrates toward the positive electrode. Mobility for linear DNA is primarily dependent on fragment size. After electrophoresis ethidium bromide is used to help visualize DNA as it intercalates forming a complex, which absorbs UV irradiation at 260 nm and fluorescence can be detected (Ausubel *et al.*, 1997).

All DNA fragments from PCR reactions or otherwise were size-fractionated by electrophoresis through 1 -2 % agarose (Invitrogen) dissolved in 1 x TAE (Section 2.1.5.1.1) or TBE (Section 2.1.5.1.3) buffer at 20-100 volts. The large Southern Gels (Section 2.2.13) were performed overnight at 20 volts using a HE99, while most other gels were run at 100 volts using a HE33 (Hoefer). The agarose gel was melted in a microwave and cooled to 55°C in a water bath, poured into a gel-casting tray and allowed to set. DNA samples were loaded into agarose wells by mixing with approximately 10% loading dye (Section 2.1.5.1.5). The loading dye results in the DNA solution being heavier than the buffer causing it to sink to the bottom of the wells. The bromophenol blue enables the visualization of the travel front. Electrophoresis was carried out in the same buffer used to make the gel. After electrophoresis, agarose gels were stained in ethidium bromide (Section 2.1.5.1.6) for 30 minutes, destained in water and DNA visualized by exposure to UV light, using a UV transilluminator (Alpha-Innotech). Post-staining gels were photographed using AlphaImager 2000 software (version 3.3d); sizes of DNA bands were determined by comparing them to a known DNA size standard marker (1 kb plus ladder, Invitrogen).

2.2.6 Restriction Endonuclease Digestion of DNA

Restriction Endonuclease digestion was performed at 37°C in the buffer stipulated by the manufacturer for 2 hours (plasmid digest) or overnight (Southern; Section 2.2.13 and Inverse PCR digests; Section 2.2.14).

Each digest consisted of 10 µL of DNA (diluted appropriately maximum 2 µg), 2 units of restriction enzyme along with 2 µL of appropriate 10 x buffer (as specified by the manufacturer) per µg of DNA in a total volume of 20 µL. Larger digests for Southern blots up-scaled the volumes described appropriately. Digests for Southern blot positive controls used 1 ng *NcoI* digested plasmid from which the probe was derived, this also to demonstrate that the level of fluorescence from a single copy gene in 1.0 µg genomic DNA is within detectable

limits. One tenth of each digest was checked for completion by comparison to undigested template using agarose gel electrophoresis (Section 2.2.5); incomplete digests were further incubated with 1 extra unit of restriction enzyme.

2.2.7 Ligation of DNA

2.2.7.1 Ligation into pGEM-T Easy vector

The pGEM-T Easy vector was selected for these studies because it has been shown to produce high copy number in *E. coli* and it is easily used for cloning in conjunction with *Taq* polymerase. pGEM-T Easy® vectors (appendix 2, Promega) are supplied linearized and ready for ligation as they have already been digested with *EcoRV*. They also contain a 3' terminal thymidine attached to either cut end, to facilitate efficient ligation of PCR products from a Taq polymerase reaction that adds 5' adenosine overhangs.

Ligation efficiency is dependent on the concentration of DNA ends in the reaction. To ensure complementary ends of vector and insert are joined, the concentration of insert must be higher than that of the vector (Berger and Kimmel, 1987).

The amount of vector DNA used in ligations was 50 ng, to which the insert DNA was added at either a 3:1 or 1:1 (insert: vector) molar ratio of vector. The amount of insert DNA that was added to ligations was calculated using the following formula.

$$\text{Amount of insert DNA (ng)} = \frac{\text{Amount of vector DNA (ng) x size of insert (bp)}}{\text{Size of vector (bp)}} \times \frac{\text{insert}}{\text{vector}}_{\text{molar ratio}}$$

Ligation mixes also contained 10 µl (2x) T4 DNA ligase buffer (100 mM Tris-HCl pH 7.6, 20 mM $MgCl_2$, 2 mM ATP, 2 mM DTT, 25% w/v PEG), 2 µL T4 DNA ligase (1 unit/µL, Roche) and sterile MilliQ dH_2O to make the final volume up to 10 µL. Ligation mixes were prepared, mixed thoroughly and incubated overnight at 16°C. To ensure each ligation was efficient, 2 µL of the ligation mix was removed prior to T4 DNA ligase addition and post reaction completion. These two samples were compared using agarose gel electrophoresis (Section 2.2.5). This procedure was preliminary to transformation protocol described in Section 2.2.12.

2.2.7.2 Intramolecular (Self) Ligation

There are two parameters in Self-ligation; if the concentration of DNA is low in a given self-ligation reaction there is a high chance that these two termini will meet, however as the

molecule increases in size the ends are less likely to meet (Ausubel *et al.*, 1997). Therefore for each reaction the optimal concentration was estimated using a size: concentration ratio graph (Ausubel *et al.*, 1997) and in both Inverse PCR scenarios 15 ng/μL was used. Ligation mixes contained 1.4 μg digested genomic DNA fragments, 10 μL (10 x) T4 DNA ligase buffer (500 mM Tris-HCl pH 7.5, 100 mM $MgCl_2$, 10 mM ATP, 100 mM DTT, 25 μg/ml BSA), 4 μL T4 DNA ligase (40 units/mL, NEB) and sterile MilliQ dH_2O to make the final volume up to 100 μL. Ligation mixes were prepared, mixed thoroughly and incubated over-night at 4°C. Ligation efficiency was tested as in Section 2.2.7.1. This procedure was a critical step for Inverse PCR (Section 2.2.14).

2.2.8 Degenerate and General Oligonucleotide Design

Degenerate primers were designed by identifying regions of homology within nucleotide sequence alignments and resolving ambiguities using degenerate code (Table 2.2). Primers were 20-25 bp in length and it was a requirement that both had similar annealing temperatures to increase specificity (Innis and Gelfand, 1994; Dieffenbach *et al.*, 1995). Poly- pyrimidine/ purine stretches and self-complementarity were avoided; guanine and cytosine bases were required toward the 3' end.

Primers (Table 2.2) were synthesized by the Invitrogen Corporation and provided as a dried stock. Each was rehydrated in TE buffer (Section 2.1.5.2.2) to a concentration of 100 pmol/μL and stored at -20°C. Primers were further diluted when required in TE buffer to a concentration of 10 pmol/μL for PCR reactions.

2.2.9 General and Degenerate PCR

PCR reactions contained 2.0 μL (10 pmol/μL) of each primer, 2.5 μL (2 mM) dNTPs (a mixture of dATP, dTTP, dCTP and dGTP), 2 μL 10 x PCR buffer (100 mM Tris-HCl, 500 mM KCl pH 8.3), 4 μL Q-solution, 0.20 μL *Taq* polymerase (5 U/μL), 1 μL template (approximately 10 ng/μL) and MilliQ dH_2O to make the total volume up to 20 μL. The *Taq* polymerase was always added last. Reactions were mixed thoroughly prior to a centrifugation at 13, 000 g in a Biofuge 13 (Heraeus) for 5 seconds and PCR reactions were carried out in one of two thermal cyclers, a Mastercycler (Eppendorf) was used for most reactions while a T-gradient (Biometra) was used for optimizing degenerate PCR. Both a positive and a negative control (no DNA template) were also included with each PCR reaction. For degenerate PCR

the positive control was DNA from an organism used to design the primers. This was to ensure that not only were the primers working, but also to give a rough size indication on the expected product. Other positive controls were specific for respective reactions and were to ensure the PCR reaction was working. The negative control contained no template DNA (appropriate volume replaced with MilliQ dH_2O); this was to check for the possibility of contamination.

General Thermal-cycling conditions for PCR

Initial Denaturation	94°C for 5 minutes.	
Denaturing	94°C for 1 minute.	30 x cycles.
Annealing	50 °C for 1 minute.	
Extension	72°C for 1 minute.	

The specifics of all PCR protocols are listed in Table A2.1; annealing temperature is dependent on the Tm of primers. Half of each PCR reaction was analyzed by agarose gel electrophoresis. Successful PCR reactions were repeated to obtain enough product for sequencing.

2.2.10 Sequencing

DNA sequencing was carried out by Lorraine Berry (MUSeq facility, Institute of Molecular BioSciences, Massey University) using dideoxy cycle sequencing with Big Dye terminators (Version 3.1, PE Biosystems, Foster City, CA, USA) and analyzed on an ABI 3730 automated DNA sequencer according to the manufacturers' instructions. Primers were diluted in TE buffer (Section 2.1.5.2.2) to 3.2 picomol/μL. Template concentration varied according to sample type, 1 ng per 100 bp and 100 ng/μL was used for PCR products and plasmid sequencing respectively. Using this protocol, direct sequencing was carried out by purifying a product from agarose gel (Section 2.2.3.3) and sequencing without cloning. Ambiguous sequence was manually solved using electropherogram data, and if necessary re-sequenced over the region.

When one set of primers was insufficient to cover an entire sequence, another set was designed using the 3' ends. These primers were then used to cover the remaining sequence; this technique is referred to as primer walking. All sequences obtained were sequenced three times using separate PCR reactions to eliminate the potential inclusion of PCR induced errors in sequence data.

2.2.11 Computer Analysis and Tools

A number of computer programs and web-based tools were utilized and unless otherwise specified, sequence data was entered and subsequently analyzed using default parameters for all these electronic resources and programs. Methods described by the program writers were followed. Two programs within the sequence analysis software provided by GCG (Version 9.1; Wisconsin Genetics Computer Group, USA), were used to construct restriction maps (map) and generate contigs of individual nucleotide sequences (FAS).

Several programs within the NCBI website (http://www.ncbi.nlm.nih.gov/) such as the Basic Local Alignment Search Tool; BLASTn for nucleotide, BLASTp for protein and Short/or exact sequence search were used for sequence similarity searches (Altschul *et al.*, 1990; Altschul *et al.*, 1997)

All PCR and construct sequences were directly compared with other sequences using ClustalX (http://www.es.embnet.org /Doc/clustalw/clustalx,html).

Filtered string-based queries for promoter elements were performed on Transcription Element Search System (TESS) (http://www.cbil.upenn.edu/tess), filter parameters used were "org Classification" – Eukaryota; Fungi (Schug and Overton, 1997).

Two programs within the European Bioinformatics Institute (http://www.ebi.ac.uk) were used to predict introns within (Genewise) and predict proteins from (Transeq) individual nucleotide sequences.

Codon usage tables were constructed from individual nucleotide sequences using version 4 of the Countcodon program (http://www.kazusa.or.jp/codon/countcodon.html; (Nakamura, 2004)).

Scans for putative motifs and domains, mass and pI of predicted proteins were conducted using the ScanProsite tool within PROSITE database (http://us.expasy.org/prosite/; (Gattiker *et al.*, 2002)).

The Coils program was used to search for possible coiled-coil motifs within protein sequences (http://us.expasy.org/tools/#proteome (Lupas *et al.*, 1991)).

Table 2.2: PCR and Sequencing Primers

A)

Primer	Size	Tm°C	Sequence (5' to 3')	Source
dEF2-Fwd	21 mer	66	GTCATYGCBCAYGTYGAYCAC	This Study
exEF2-Fwd	21 mer	64	GTTATTGCCCACGTCGATCAC	This Study
dEF2-Rev	21 mer	62	ACCWCKGTGRATRGAATCRGC	This Study
exEF2-Rev	21 mer	64	ACCACGGTGAATAGCATCAGC	This Study
EF2-IntF1	21 mer	62	GCCTTGATGAAGAGATCCTTCC	This Study
EF2-IntF2	21 mer	62	CATCGAACATGTGCTAACTGC	This Study
EF2-IntR1	21 mer	66	AGATGATGGAGCGTCTCTGGG	This Study
EF2-IntR2	21 mer	62	CCGAAAGACATCATCGACTGA	This Study
EF2-IntR3	21 mer	64	CTCGTGTCGGATTTCATGCGA	This Study
EF2-IntR4	21 mer	66	AGATGGGAGCAGCACAAACCC	This Study
EF2-IntRA	21 mer	64	GAAACTGCCCACAACCTCAAG	This Study
EF2-IPCR 5'	21 mer	66	GGAGAGCGGCAGTAACTTCAG	This Study
EF2-IPCR 3'	22 mer	68	ATCCATCCGCTTCAACGTCCTC	This Study
reEF2-IPCR 3'	21 mer	70	CGTGAGGGTCCCGTTGCTGAG	This Study
EF2 5' RACE	24 mer	74	GGAGGTTGGTGATCTCGTCGTTCT	This Study
EF2 3' RACE	25 mer	78	AGGGTAAGCAGCTCGAGCGTGCTTT	This Study
EF2-RACEc	24 mer	76	CCCTGACAAGGGCACCATTGCTTT	This Study
dSep-Fwd	20 mer	61	ATYCCHGTYATTGCCAAGGC	This Study
exSep-Fwd	20 mer	58	ATTCCTGTTATTGCCAAGGC	This Study
dSep-Rev	20 mer	58	TCHTCYTCCTTGAAYTTVGG	This Study
exSep-Rev	20 mer	56	TCATCTTCCTTGAATTTCGG	This Study
Sep-IPCR 5'	21 mer	62	CTTGAACTTGGCCAAATCAGC	This Study
Sep-IPCR 3'	20 mer	66	AAGTTCGGCGAAGCCCGTCC	This Study
Sep 5' RACE	24 mer	72	CTCGATATCGAGAGGCTTCAAGGT	This Study
Sep 3' RACE	25 mer	78	CTCGACGACCAGCACGAATCCTACA	This Study
Sep-RACEc	23 mer	72	CACAAGCGTCGACACCAGAAGCA	This Study
npSep-Fwd	21 mer	64	ATTGGCATTGCCAATCTGCCC	This Study
npSep-Rev	21 mer	66	AGGTCCTTGTTGAGACGGTCG	This Study
dSACC-Fwd	18 mer	46	TAYAARYVYTGGYTITGY	This Study
dSACC-Rev	18 mer	55	RTCYTTYTAIGGICCIGG	This Study
T1.1	20 mer	56	GAGAAAATGCGTGAGATTGT	Bryant (2003)
T1.2	20 mer	64	TGGTCAACCAGCTCAGCACC	Bryant (2003)
T1.1 IP-Fwd	26 mer	74	GTAACCGAGAAAATGCGTGAGATTGT	This Study

T1.2 IP-Rev	23 mer	74	ACCTGGTCAACCAGCTCAGCACC	This Study
pUC/M13F	22 mer	70	GCCAGGGTTTTCCCAGTCACGA	Promega
pUC/M13R	22 mer	62	TCACACAGGAAACAGCTATGAC	Promega
Generacer 5'	23 mer	74	CGACTGGAGCACGAGGACACTGA	Invitrogen
Generacer 3'	25 mer	76	GCTGTCAACGATACGCTACGTAACG	Invitrogen
Generacer 5' nested	26 mer	78	GGACACTGACATGGACTGAAGGAGTA	Invitrogen
Generacer 3' nested	23 mer	72	CGCTACGTAACGGCATGACAGTG	Invitrogen

B)

Degenerate Code For Primers

Letter	Represents	Letter	Represents
A	Adenosine	Y	Pyrimidine (C or T)
C	Cytosine	M	Amino (A or C)
G	Guanine	K	Keto group (G or T)
T	Thymidine (DNA)	B	Not A (C, G or T)
U	Uracil (RNA)	D	Not C (A, G or T)
S	Strong (C or G)	H	Not G (A, C or T)
W	Weak (A or T)	V	Not T (A, C or G)
R	Purine (A or G)	N or I	Any base or Inosine (A, C, G or T)

2.2.12 Cloning (Bacterial Transformation)

2.2.12.1 Preparation of Competent Cells

E.coli ultra-competent DH5 alpha cells were prepared using the simple and efficient method of Inoue *et al* (1990). An inoculation loop of DH5 alpha cells from a frozen stock was grown overnight on an LB plate (Section 2.2.1.1). Ten large colonies (approximately 2 mm in diameter) were isolated and used to inoculate 300 mL of SOB broth (Section 2.1.4.1.3); this was grown to an A_{600} of 0.6 at 18°C on an orbital shaker at 200 rpm (Lab-Line). The culture was incubated on ice for 10 minutes and pelleted in a RC-SB (GSA rotor; Sorvall) centrifuge at 3,000 g for 10 minutes at 4°C. The pellet was resuspended in 80 mL of cold (-20°C) TB (10 mM piperazine -N, N'-bis[2-ethanesulphonic acid] (PIPES), 15 mM $CaCl_2$, 250 mM KCl pH 6.7, 55 mM $MnCl_2.4H_2O$), the solution incubated on ice for 10 minutes, and spun down as above. The pellet was resuspended in 20 mL cold TB, and 1.4 mL of cold (-20°C) dimethyl sulfoxide (DMSO) was added. After the cell suspension was incubated for 10 minutes on ice, 500 µL aliquots were dispensed into microcentrifuge tubes, snap-frozen using liquid nitrogen and stored at -80°C (Inoue *et al.*, 1990). Cells were tested for competency by transformation with pBR322 before use (Section 2.2.12.2).

2.2.12.2 Transformation of *E.coli*

Half of a ligation reaction (Section 2.2.7.1) was added to 100 µL of DH5 alpha competent cells in a microfuge tube and incubated on ice for 20 minutes. The samples were then heat-shocked for 45 seconds in a water bath at 42°C and incubated on ice for 2 minutes. After the addition of 900 µl SOC medium, (Section 2.1.4.1.1) the samples were incubated for 1.5 hours at 37°C on an orbital shaker at 150 rpm (Lab-Line). The samples were then diluted appropriately and plated on either LB or LB/ Ampicillin/ IPTG/ X-Gal plates according to the nature of the sample and grown overnight (Section 2.2.1.1). This distinguished which plasmids had insert DNA. DH5 alpha colonies that appear blue are those containing plasmid without insert, whereas the white colonies carry plasmids with an insert in the correct location. The variation in color is generated as the β-galactosidase coding sequence is disrupted by the presence of the insert DNA and α-complementation cannot occur. The control plasmid pUC-118 which was used also confers ampicillin resistance and not only acts as a transformation control in this experiment, but also as a control to ensure IPTG and X-gal are functioning as selective markers. Ligation reactions with the control DNA assess the efficiency of the ligation. Typically, about 10-40% of colonies should be blue and no less than 60% of colonies should be

white. If no white colonies grew it could have been due to a failed ligation or transformation. In addition, if fewer than 50% of colonies were white then the ligations were likely to have been carried out using suboptimal conditions (Promega, 1999). Table A2.2 is an example of transformation results from this study.

2.2.12.3 Diagnostic PCR Screening of Transformants

Following the identification of a potential transformant by blue/white selection, white transformant colonies were screened using PCR (Section 2.2.9) with pUC/M13 primers (Table 2.2). A sterile toothpick was used to pick up a colony and dip this into a prepared PCR reaction mix (1 colony per reaction); following this the toothpick was then used to make a smear on a quadrant of a LB-amp plate and incubated overnight (Section 2.2.1.1).

The pUC/M13 primer PCR with non-transformants did not result in a product as if the vector is not religated, as they cannot amplify across the pGEM-T Easy MCS. If the vector was religated to itself (which it should not be due to 3' thymidine overhangs), but without an insert, a PCR product of ~300 bp results. If the insert was in the MCS then the PCR product that is amplified is ~300 bp plus the size of the insert. Each screen is specific due to the primers, and each agarose gel electrophoresis (Section 2.2.5) consisted of a no DNA lane, a positive control (positive control insert in plasmid resulting in a 542 bp + ~300) and a negative control (unligated plasmid) as well as plasmids to be screened. Following positive identification of a transformant, the cells were then cultured in 5 mL LB amp overnight (Section 2.2.1.1) for plasmid preparation and to retain the plasmids selectivity before storage.

2.2.12.4 Storage of clones

Glycerol stocks were made of selected transformed DH5 alpha lines by mixing 500 µL of an overnight culture with 500 µL sterile 40% glycerol in 1 mL sterile cryotubes and storing at -80°C.

2.2.13 Southern Blotting and Hybridization

2.2.13.1 Southern Blotting

Southern Blotting was based on protocol described by Southern (1975) and the DIG-system user guide (Southern, 1975; Van Miltenburg *et al.*, 1995). DNA samples (10 µg) were digested with the appropriate enzymes (Section 2.2.6) and electrophoresed through a 152 x 142 mm

0.7% agarose TBE gel for 16 hours at 20 volts (Section 2.2.5). This was photographed with a ruler to define migration distances of molecular weight markers. The gel was then submerged for depurination in SB 1 (Section 2.1.5.3.1) for 10 minutes, subsequently rinsed with MilliQ dH_2O before immersion in denaturation buffer SB 2 (Section 2.1.5.3.2) for 30 minutes. Following this the gel was rinsed with MilliQ dH_2O and submerged for neutralization in SB 3 (Section 2.1.5.3.3) for 30 minutes, all steps up to and including this the gel has been gently agitated room temperature.

Concurrent to gel treatments the blotting apparatus was assembled. Two sheets of Whatman 3MM paper was pre-soaked in 20 x SSC (Section 2.1.5.3.4) and placed on a plastic trough with wells at opposite ends. The wells were filled with 20 x SSC and either end of the paper allowed to rest in the liquid. A sheet of gladwrap was overlaid flush onto the soaked paper and a section 2 mm smaller than the gel area was excised. The gel was then placed over the excised area after neutralization treatment and overlaid with a piece of positively charged nylon membrane (Roche). Two pieces of Whatman 3MM paper (cut 2 mm less than the gel size) were pre-soaked in 2 x SSC and overlaid onto the membrane, followed by two additional pieces of dry paper of the same size. An arbitrary stack of dry paper towels was placed on top of the whole system, followed by a weight to ensure even DNA transfer.

After overnight blotting, the apparatus was disassembled and the DNA on wet membrane was cross-linked by ultraviolet light (energy 12,000 $\mu J/cm^2$; Ultra LUM) and allowed to air-dry.

2.2.13.2 Preparation of Digoxigenin-11-dUTP (DIG) Labeled DNA Probe

DNA probes were random prime DIG-labeled using DIG-High Prime (Van Miltenburg *et al.*, 1995). DNA to be labeled (500 µg per 16 µL) was heat-denatured in a boiling water bath (JB1; Grant Instruments) for 10 minutes and immediately put on ice. DIG-High prime (4 µL) was added to make a total volume of 20 µL, mixed and centrifuged in a Biofuge 13 (Hoefer) at 13,000 g for 3 seconds. The solution was incubated at 37°C in a JB1 water bath, and after 20 hours 2 µL EDTA (200 mM, pH 8.0) was added to stop the reaction. Estimation of probe yield was achieved by comparison to serial dilutions from an estimated 1ng/µL-0.01pg/µL (estimations were made using the DIG system users guide). Each dilution was spotted onto a positively charged nylon membrane (Roche) alongside DIG-labeled control DNA of the same known concentrations. The membrane was cross-linked by ultraviolet light (energy 12,000 $\mu J/cm^2$; Ultra LUM) and allowed to air-dry before chemiluminescent detection. Probes were only used if the actual concentration was equal to or greater than the control probes concentration.

2.2.13.3 Hybridization using DIG DNA Probes

Membranes for hybridization were pre-hybridized in a sealed glass tube with DIG standard hybridization solution (Section 2.1.5.5.2) for 2 hours at 55°C. Concurrent to this DIG-labeled probe was heat-denatured in a boiling water bath (JB1; Grant Instruments) for 10 minutes and immediately put on ice. The probe was diluted appropriately, added directly to the solution surrounding pre-hybridized membrane and hybridization commenced overnight at 55°C. After the hybridization the membrane was removed, washed twice in 2 x SSC (Section 2.1.5.3.4) containing 0.1% (w/v) SDS for 5 minutes at room temperature with gentle agitation (Heto Lab equipment). This was followed by two washes in 0.5 x SSC containing 0.1% (w/v) SDS for 15 minutes at 68°C with gentle agitation.

2.2.13.4 Detection of Hybridization using Chemiluminescence

Hybridized membranes (Section 2.2.13.3) were equilibrated in DS 2 (Section 2.1.5.5.3) for 1 minute and gently agitated (Heto Lab equipment) in DS 3 (Section 2.1.5.5.4) for 60 minutes before addition of Anti-DIG Fab fragments (Section 2.1.5.5.5). The membrane was then agitated for 30 minutes before the membrane was removed and washed twice in DS 2 containing 0.3% (v/v) Tween 20 for 15 minutes, before equilibration in detection buffer (DS 5; Section 2.1.5.5.6) for 2 minutes. The membrane was removed, quickly placed in an A4 copysafe pocket (Filemaster) while remaining wet and 500 µl of CSPD (1:100 dilution in DS 5) pipetted directly onto it. The pocket was sealed and the CSPD spread evenly using a wet tissue from outside the pocket, creating a liquid seal around the membrane. The membrane was sealed inside the copysafe pocket using a vacuum-packer and incubated at 37°C for 15 minutes. This was then exposed to X-ray film (Kodak) for 20 minutes; the exposure developed using an Auto-developer (Kodak).

2.2.14 Inverse PCR (IPCR)

IPCR is a technique to amplify DNA sequences that flank a region of known sequence. The methodology was based on protocol described by Ochman *et al* (1988). Genomic DNA was digested with appropriate enzymes and separated by gel electrophoresis as in Section 2.2.13.1. Using the markers from the molecular weight ladder, the region that contained the product from Southern hybridization was excised and purified (Section 2.2.3.3.2). This was self-ligated using T4 DNA ligase (Section 2.2.7.2), purified from ligase using phenol/chloroform (Section

2.2.3.1). To ensure a successful ligation took place ligated template was compared to unligated template using gel electrophoresis (Section 2.2.5).

Initially standard PCR technique (Section 2.2.9) was applied, however due to lack of success a modified protocol of the PCR technique was adopted. The Expand Long Template PCR system (Roche) was adopted; this follows the PCR technique closely with an additional step in the thermal cycling process and an increased quantity of template (100 ng). To improve the specificity of the IPCR reaction 1.25 % (v/v) formamide was also added to the reaction (Sarkar *et al.*, 1990). See Figure A2.2 for a schematic of the IPCR process.

General Thermalcycling conditions for IPCR

Initial Denaturation	93°C for 5 minutes	
Denaturing	93°C for 30 seconds	10 x cycles.
Annealing	50 °C for 30 seconds	
Extension	68°C for 45 seconds per kb	
Denaturing	93°C for 30 seconds	24 x cycles.
Annealing	50 °C for 30 seconds	
Extension	68°C for 45 seconds per kb	
	Plus 20 seconds/cycle	

The specifics of all PCR protocols are listed in Table A2.1 and annealing temperature is dependent on the Tm of primers. Half of each PCR reaction was analyzed by agarose gel electrophoresis

2.2.15 RNA Technique Precautions

To treat for RNA use, all glassware was soaked in chromic acid bath overnight, rinsed, autoclaved and baked for 4 hours at 180°C. Aerosol tips and previously unopened bags of disposable microcentrifuge tubes (Axygen) were autoclaved in RNA-treated glassware. Fresh disposable latex gloves were used to handle all glassware and during all experimental procedures. One set of Gilson pipettes were used exclusively for RNA; all reagents and solutions from RNA dedicated stocks. All solutions were treated with 0.01 % (v/v) DEPC with exception of solutions containing Na_2EDTA, Tris-HCl or MOPS buffer (these were made using 0.01 % (v/v) DEPC). Treatment of solutions made with DEPC was an incubated overnight at

37°C and autoclaved twice for DEPC removal. These precautions were observed during all RNA handling.

2.2.16 RNA Isolation

2.2.16.1 Total RNA Isolation

RNA isolation from fungal, infected plant tissue and plant tissue using Tri-reagent (TRIzol®) was based on protocol described by its manufacturers (MRC).

Generally, 1.0 g of tissue was ground to a fine powder with a mortar and pestle using liquid nitrogen and mixed with 10 mL TRIzol® to form a smooth paste. All subsequent activities were carried out at room temperature unless otherwise stated. The sample was allowed to thaw, transferred to a 14 mL Falcon tube (Greiner), centrifuged at 9,000 g using a SS34 (Soval) at 4°C for 10 minutes and the resulting supernatant transferred to a new falcon tube. Chloroform was added at 17 % (v/v), the solution mixed, incubated for 12 minutes and centrifuged at 9,000 rpm using an SS34 at 4°C for 15 minutes. The aqueous phase was transferred to a falcon tube containing 5.0 mL isopropanol, incubated for 10 minutes to precipitate RNA and centrifuged at 9,000 g using an SS34 at 4°C for 10 minutes. The supernatant was decanted, the RNA precipitate washed with 10 ml of 70 % ethanol and centrifuged at 6,700 g using an SS34 at 4°C for 10 minutes. The resulting pellet was air-dried and resuspended in 200 µL MilliQ dH_2O before DNase treatment (Section 2.2.16.2) and quantification (Section 2.2.4.2).

2.2.16.2 DNase Treatment of RNA

All isolated total RNA and mRNA was treated with RNase I free DNase to remove any contaminating DNA. Every 40 µg was mixed in a microcentrifuge tube with: 20 units DNase I (Invitrogen; Section 2.1.5.4.1); DNase I buffer (100 mM sodium acetate, 5 mM $MgSO_4$ at pH 5.0); 0.5 mM 1,4-dithiothreitol (DTT) and 50 units of RNase inhibitor (BRL) to a final volume of 800 µL. This was mixed and incubated at 37°C for 30 minutes. The reaction was then phenol/chloroform extracted (section 2.2.3.1) and ethanol precipitated (Section 2.2.3.2). The RNA was resuspended in 40 µL MilliQ dH_2O, spectophotometrically quantified (Section 2.2.4.2) and checked for RNA degradation and DNA contamination. Overly degraded RNA was not used in these studies and RNA was stored at -80°C.

2.2.16.3 mRNA Isolation

Messenger RNA (mRNA) isolation from total RNA using the GenElute™ mRNA minprep kit was based on a protocol described by its manufacturers (Sigma). This technique was used solely for enrichment to make cDNA.

Generally, 300 µg total RNA isolated in Section 2.2.16.1 was diluted into a total volume of 250 µL MilliQ dH_2O, 250 µL 2 x binding solution, 15 µL oligo dT_{30} polystyrene beads and mixed thoroughly. The solution was incubated at 70°C in a heating block (Multi-Blok) for 3 minutes, incubated at room temperature for 10 minutes and centrifuged at 13,000 g using a Biofuge 13 (Heraeus) for 2 minutes. The resulting pellet was washed twice with 500 µL wash solution, applied to the supplied filter-column and centrifuged as before. The mRNA was eluted by addition of 50 µL heated (70°C) elution solution (10 mM Tris-HCl, pH 7.5) and centrifugation at 13,000 g using a Biofuge 13 for 5 minutes. The mRNA was checked for rRNA contamination and overly contaminated samples were not used.

2.2.17 Northern Blotting and Hybridization

Northern Blotting was based on protocol described by Southern (1975) and the DIG-system user guide (Southern, 1975; Van Miltenburg *et al.*, 1995). When handling RNA samples the precautions outlined in Section 2.2.16 were observed.

2.2.17.1 Formaldehyde Gel Electrophoresis

An agarose gel (1 %) was prepared using 30 ml MOPS buffer (Section 2.1.5.4.2), 17 mL formaldehyde and 250 mL MilliQ dH_2O. The solution was poured into a 152 x 142 mm gel tray as in Section 2.2.5. Prior to sample loading the gel was electrophoresed at 70 volts for 10 mins.

Concurrent to gel setting, total RNA samples (10-50 µg) were mixed with 2.2 M formaldehyde, 50 % (v/v) formamide and 0.5 x MOPS buffer. Gel loading buffer was added to give a final concentration of 1x after samples were denatured in a JB1 water bath (Grant Instruments) at 65°C for 10 minutes. The prepared samples were loaded into the formaldehyde gel along with DIG-labeled RNA ladder (Roche MWM II; 1.5- 6.9 kb) and electrophoresed using a HE99 (Hoefer) at 70 volts for 5 hours at 4°C. The gel was stained and visualized as in Section 2.2.5 however DEPC MilliQ dH_2O was used.

2.2.17.2 Northern Blotting

RNA from the formaldehyde gel (Section 2.2.17.1) was blotted onto positively charged nylon membrane (Roche) by similar methods to those described in Section 2.2.13.1, however the formaldehyde gel was equilibrated for 30 minutes in 20 x SSC (Section 2.1.5.3.4). DIG labeling of probes, hybridization and chemiluminescent detection were performed as described in Sections 2.2.13.2, 2.2.13.3 and 2.2.13.4. The detection was exposed in a time-process (every 5 minutes) using a LAS-1000 (Fujifilm) instead of exposing to X-ray film.

2.2.17.3 Probe Stripping

The membrane to be stripped was washed in MilliQ dH_2O for 5 minutes and incubated in Northern probe stripping solution (50 % formamide, 50 mM Tris-HCl and 0.1 % (w/v) SDS at pH 8.0) in a water bath (Grant Instruments) at 68°C. The membrane was rinsed briefly twice, firstly in MilliQ dH_2O and secondly in 2 x SSC (Section 2.1.5.3.4). Membrane checked by chemiluminescence for DIG-labeled probe (Section 2.2.13.2) to ensure successful stripping had occurred. If stripping was unsuccessful the process was repeated.

2.2.18 RT PCR Techniques

2.2.18.1 RT PCR for *In planta* Analysis

Reverse transcription (RT) PCR is a two-step process in which cDNA is synthesized using mRNA and probed using gene-specific primers. The protocol used was adapted from the method outlined by Kawasaki (1990). As mRNA is handled during the cDNA synthesis the precautions outlined in Section 2.2.16 (Kawasaki, 1990).

2.2.18.1.1 Reverse Transcriptase Synthesis of cDNA using the Expand Kit

The Expand reverse transcriptase kit (Roche) was used according to manufacturer's instructions. In a microcentrifuge tube an appropriate dilution of DNase treated mRNA (1 µg; Section 2.2.16.3) was combined with 1.0 µL of random primers (50 pM; Invitrogen) and MilliQ dH_2O to a final volume of 11.7 µL. The solution was denatured at 95°C for 5 minutes using a Multi-blok heater (Lab-Line) and placed on ice. A cocktail of 1 µL Expand Reverse Transcriptase (50 units), 4.0 µL Expand Reverse Transcriptase buffer (5x), 0.8 µL dNTP's (1.25 mM each), 2.0 µl DTT (100 mM) and MilliQ dH_2O to a final volume of 20 µL were added, mixed and incubated at room temperature for 10 minutes. Following this the solution was incubated at 42°C in a JB1 water bath (Grant Instruments) for 45 minutes. The resulting

cDNA was stored at -20°C. Each time this reaction was performed a negative control containing no RNA was also prepared.

2.2.18.1.2 PCR probing of cDNA

Using cDNA synthesized in Section 2.2.18.2.1 a variation of general PCR (Section 2.2.9; PR9 Table A2.1) was carried out with selected GSPs (Table 2.2) designed the same way as RACE primers (Section 2.2.18.2.2) to improve specificity. As in Section 2.2.14, 1.25 % (v/v) formamide was also added to improve the specificity of the PCR reactions (Sarkar *et al.*, 1990). Half of each reaction was separated by gel electrophoresis (Section 2.2.5).

2.2.18.2 RACE

RACE is a variation of RT PCR, in that mRNA is synthesized into cDNA; however the method of PCR and purpose differs. RACE is a method of obtaining full-length cDNA from target genes to define transcription start/stop sites and intronic regions. RACE-Ready cDNA was kindly prepared by Ningxin Zhang (IMBS; Massey University; New Zealand).

2.2.18.2.1 Synthesis of RACE-Ready cDNA performed by Ningxin Zhang

The GeneRacer kit (Invitrogen) was used according to manufacturer's instructions. Messenger DNA free RNA was dephosphorylated and decapped before ligation of GeneRacer RNA oligo nucleotides (which contain the 5' and 3' Generacer primer sites). From this template RACE-Ready cDNA was synthesized using a SuperScript III RT reaction.

2.2.18.2.2 Design of RACE Primers and PCR

RACE Primers were designed in accordance with the GeneRacer kit (Invitrogen) manufacturer's instructions. This closely followed general PCR primer design (Section 2.2.9) with the following changes to improve specificity; the partnering primer to the 5' and 3' RACE primers were supplied by the manufacturer, the primers were generally longer (25 bp) to use higher annealing temperatures and no more than 2-3 guanine or cytosine nucleotides were included in the last 5 base-pairs. The initial RACE PCR was carried out with general PCR methods (Section 2.2.9) and the appropriate manufacturers PCR primers with a 5' or 3' GSP. This used RACE-ready cDNA prepared in Section 2.2.18.2.1 as a template. The nested RACE PCR was carried out using the appropriate manufacturer's nested PCR primers with a 5' or 3' GSP; identical conditions and a dilution of the initial RACE reaction as a template (PR6 Table 2.2). As in Section 2.2.14, 1.25 % (v/v) formamide was also added to improve the specificity of the PCR reactions (Sarkar *et al.*, 1990).

Chapter 3

nlEF-2 CLONING

3.1 Molecular Cloning Of A Protein Elongation Factor 2 Gene From *Neotyphodium lolii* (Lp19)

As described in Section 1.2 the *EF-2* gene was chosen for further investigations, because of its possible role in morphogenesis, its possible utility as a marker of endophyte metabolic activity and because it appeared to possess conserved regions that would facilitate heterologous cloning

3.1.1 Degenerate Primer Design

Database searches identified 20 fungal *EF-2* protein sequences (of these 6 were identified as *EF-2* homologues only by automatic annotation in genome projects and were excluded from further analysis). A number of ascomycete species sequences were selected for primer design on the basis that the species share a common ancestor with *N. lolii* (Spataford and Blackwell, 1993). Amino acid and nucleotide sequences for *Neurospora crassa* (Q96X45), *Candida parasilosis* (AF107291), *Candida albicans* (Y09664), *Candida glabrata* (AF107287), *Clavispora lusitaniae* (AF107290) and *Filobasidiella. neoformans* (AAG09782) from the NCBI website (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/) were downloaded and aligned with default settings using ClustalX (Section 2.2.11). Within the resulting alignment degenerate PCR primers dEF2-Fwd and dEF2-Rev were designed from two conserved regions (Section 2.2.8 and Figures 3.1 A and B). These sites were chosen firstly because they would amplify most of the gene and secondly, each site had similar base pair composition and therefore similar annealing temperature.

3.1.2 PCR amplification of a putative Lp19 *EF-2 (nlEF-2)* gene fragment

Optimization was carried out using a range of annealing temperatures between 50 and 60°C and varying amounts of Lp19 and *C. albicans* genomic DNA as templates (Section 2.2.9). A

Figure 3.1

A): Partial Alignment of Genes used to Design *EF-2* Degenerate PCR Primers.

Partial alignment *EF-2* sequences of: *C. albicans* (Mendoza *et al.* 1999), *C. parapsilosis* (Shastry *et al* 1998; direct submission), *C. glabrata* (Shastry *et al* 1998; direct submission), *C. lusitania* (Shastry *et al* 1998; direct submission), *F. neoformans* (Shastry *et al* 2001) and *N. crassa* (Propheta *et al* 2001). Each box around a sequence indicates the region the primer labeled was designed, and the stars above each column show a consensus for that base. The numbers annotate the locations of these sites in each of the sequences used and arrows show directionality. The septated line indicates there is more alignment between the two figures at the start and ends of these; however it is not shown for simplicity. The double backslash indicates that the 5`/3` (upper/lower) are ends of the same sequence.

B): Design of *EF-2* Degenerate PCR Primers.

The two regions identified in Figure 3.1 A) were selected for primer design, and the degenerate nucleotide sequence analyzed. Using degenerate code ambiguities were solved and dEF2-Fwd + dEF2-Rev were designed (degenerate nucleotide locations in red; Section 2.2.8 and Table 2.2). The region to be amplified in the *EF-2* genes between the sites used to design primers varied between *N. crassa*, 2091 bp and *C. albicans, C. parapsilosis*, 2035 bp. Taking this into consideration along with the suggestion by Shastry *et al* 2001 that *EF-2* genes are highly conserved; the estimated PCR product size was 2000-2100 bp.

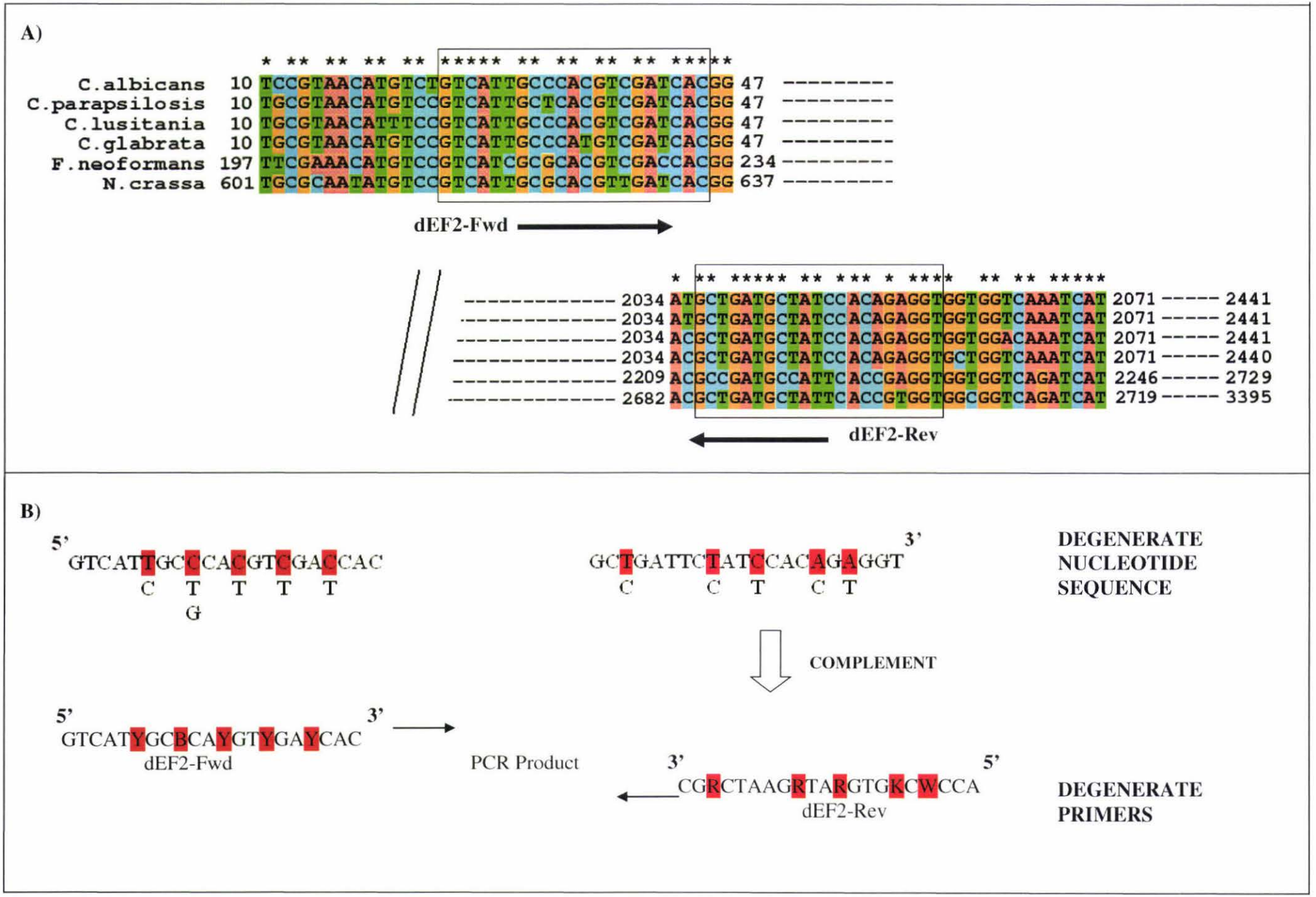
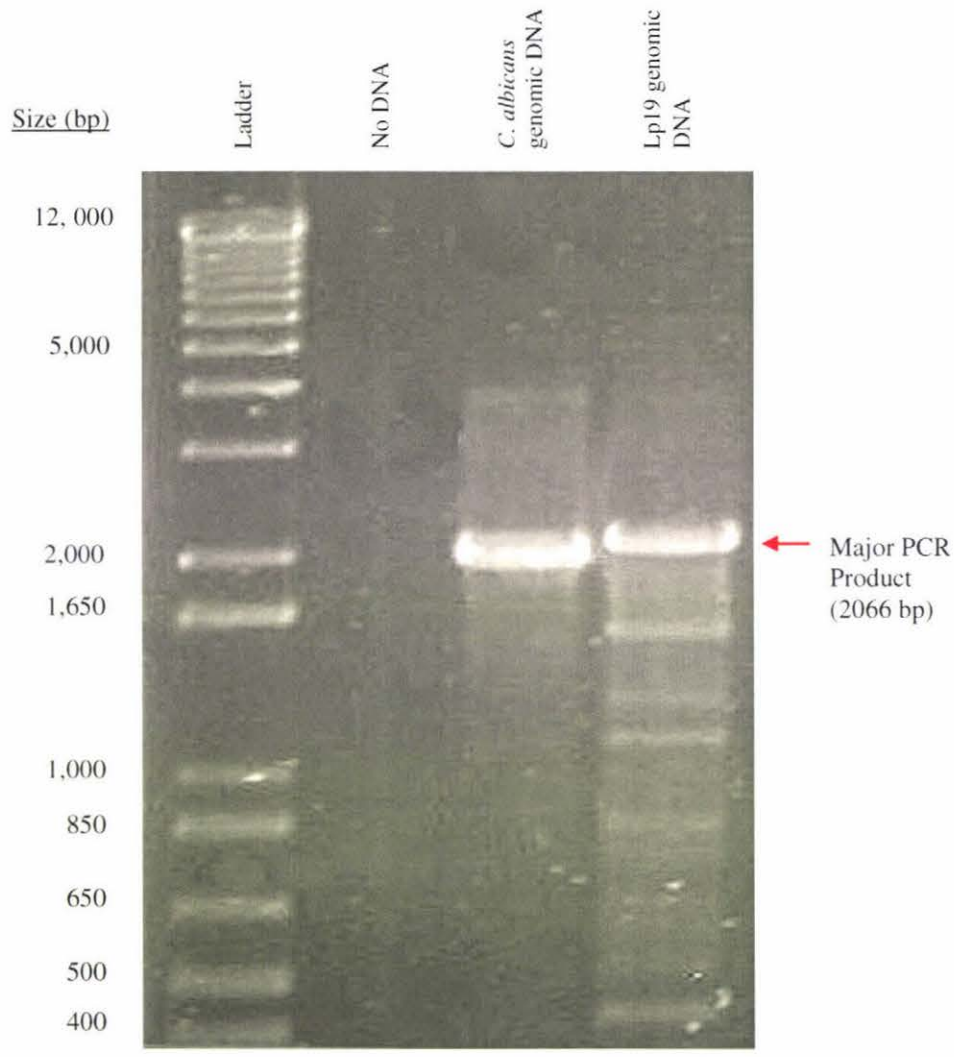

A)
C.albicans 10 TCCGTAACATGTCTGTCATTGCCCACGTCGATCACGG 47 ---------
C.parapsilosis 10 TGCGTAACATGTCCGTCATTGCTCACGTCGATCACGG 47 ---------
C.lusitania 10 TGCGTAACATTTCCGTCATTGCCCACGTCGATCACGG 47 ---------
C.glabrata 10 TGCGTAACATGTCCGTCATTGCCCATGTCGATCACGG 47 ---------
F.neoformans 197 TTCGAAACATGTCCGTCATCGCGCACGTCGACCACGG 234 ---------
N.crassa 601 TGCGCAATATGTCCGTCATTGCGCACGTTGATCACGG 637 ---------
dEF2-Fwd
---------- 2034 ATGCTGATGCTATCCACAGAGGTGGTGGTCAAATCAT 2071 ----- 2441
---------- 2034 ATGCTGATGCTATCCACAGAGGTGGTGGTCAAATCAT 2071 ----- 2441
---------- 2034 ACGCTGATGCTATCCACAGAGGTGGTGGACAAATCAT 2071 ----- 2441
---------- 2034 ACGCTGATGCTATCCACAGAGGTGCTGGTCAAATCAT 2071 ----- 2440
---------- 2209 ACGCCGATGCCATTCACCGAGGTGGTGGTCAGATCAT 2246 ----- 2729
---------- 2682 ACGCTGATGCTATTCACCGTGGTGGCGGTCAGATCAT 2719 ----- 3395
dEF2-Rev
B)
5' GTCATTGCCCACGTCGACCAC 3'
C T T T T
G
5' GCTGATTCTATCCACAGAGGT 3'
C C T C T
DEGENERATE NUCLEOTIDE SEQUENCE
COMPLEMENT
5' GTCATYGCBCAYGTYGAYCAC 3'
dEF2-Fwd
PCR Product
3' CGRCTAAGRTARGTGKCWCCA 5'
dEF2-Rev
DEGENERATE PRIMERS

Figure 3.2: PCR amplification of putative *nlEF-2* gene fragments

Both *C. albicans* and Lp19 PCR reactions were performed using 10 ng genomic DNA. 20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was loaded in lane 1.


Size (bp)
Ladder
No DNA
C. albicans genomic DNA
Lp19 genomic DNA
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
Major PCR Product (2066 bp)

temperature of 54°C using 10 ng of template resulted in both a good yield of a product of the expected size (a 2035 bp product was expected using *C. albicans* DNA as a template, and a product of comparable size from Lp19 DNA) and low levels of other products for both *C. albicans* and Lp19 templates (Figure 3.2; section 2.2.9 using PCR protocol 7 (PR7) in Table A2.1)

3.1.3 Cloning and sequencing of the putative *nlEF-2* fragment

The amplified putative *nlEF-2* fragment was purified using the freeze-squeeze technique and cloned into a pGEM-T Easy vector® (Section 2.2.12.2). Sixteen putative ligation product-containing transformants were selected. The presence of the putative *nlEF-2* fragment in the 16 clones was tested by PCR, using pUC/M13 universal primers (Section 2.2.12.3 and Table 2.2). All 16 clones generated the expected 2.4 kb product and 8 clones were stored on glycerol at -80°C. Plasmid containing the putative *nlEF-2* fragment (pDEF-1) was extracted from three clones. Inserts were partially sequenced using pUC/M13 universal primers resulting in 700 bp of reliable sequence from the 5' end and 700 bp of sequence from the 3' end of the fragment. Two new primers, EF2-IntF1 and EF2-IntR1 (Table 2.2) were then designed to these sequences and used to sequence the central remaining part of the fragment. Sequence assembly yielded a 2066 bp contig. A ClustalX alignment (Section 2.2.11) for all sequences showed the three inserts were identical in sequence.

3.1.4 Initial verification of the identity of the putative *nlEF-2* fragment.

The 2066 bp sequence was subjected to a BLASTn (Basic Local Alignment Search Tool) search through the NCBI website (Section 2.2.11). Sixteen of the 20 most significant matches were fungal *EF-2* genes (data not shown), with the Cot3 gene from *N. crassa* (Q96X45) being the most significant (0.0 E-value; 86 % identity). A putative intron site was identified from 673 to 701 by the signature sequences $^{5'}$GTANGT and YAG$^{3'}$ (Gurr *et al.*, 1987); this site agreed with the Genewise program intron predictions (Section 2.2.11). The Transeq program (Section 2.2.11) was used to predict a derived amino acid sequence from the 2066 bp sequence lacking the intron; this was subjected to a BLASTp search using default settings (Altschul *et al.*, 1990; Altschul *et al.*, 1997). The twenty most significant matches were fungal *EF-2*

protein sequences (data not shown) and as with nucleotide sequence Cot3 from *N. crassa* (EAA77131.1) was the most significant (0.0 E-value; 89% identity).

Together this evidence strongly suggested that the cloned 2066 bp PCR product was part of the *nlEF-2* gene.

3.1.5 Obtaining the *nlEF-2* 2066 bp flanking regions.

Inverse PCR was used to obtain the 5' and 3' flanking regions of the *nlEF-2* gene. IPCR is a method of genome walking in which a core sequence is known and primers are designed in the inverted orientation from conventional PCR primers (Triglia *et al.*, 1988). Restriction endonuclease digestion is required to manipulate genomic DNA into manageable sized fragments for DNA polymerase, thus creating a small fragment for self-religation and subsequent PCR (Figure A2.2). This follows a suggestion by Ochman *et al* that optimally the fragment of interest should no more than 2-3 kb longer than the core sequence due to Taq polymerase limitations (Ochman *et al.*, 1988). In addition Ausubel *et al* (1997) comments that smaller fragments are more easily self-ligated; this holds true for fragments to the lower size limit of 500 base pairs (Ausubel *et al.*, 1997). Selection of appropriate restriction enzymes can be determined by Southern blotting and hybridization procedures using all or part of the core sequence as a probe (Ochman *et al.*, 1988). Self-religation of the digested fragment is an essential element for IPCR in order for the PCR of opposite oriented primers to result in a product. It is also required that IPCR primers to be designed at least 50 base-pairs into known sequence for positive identification of IPCR product (Ausubel *et al.*, 1997).

There are two IPCR strategies that can be applied; the method used by Ochman *et al* uses restriction enzymes that do not cut within the core sequence and primers extending outward from its ends (Figure A2.2 A); (Ochman *et al.*, 1988)).. This may lead to large target amplicons for IPCR and most readily available restriction enzymes are infrequent cutters in high GC genomes (Huang *et al.*, 2000). The method of Siebert *et al* uses a restriction enzyme that cuts within the core sequence; so although the primers are extending in opposite directions they may be designed anywhere in the core region (Figure A2.2 B); (Siebert *et al.*, 1995). This potentially results in smaller amplicons for IPCR, however two IPCR reactions need to be designed to cover both the 5' and 3' flanking region of the core sequence. It was decided to use the Ochman *et al* (1988) method because it can amplify both flanking regions concurrently. If this strategy proved unsuccessful due to sub-optimal large fragments and inefficient self-

religation, the Siebert *et al* method would be applied. The acceptable range of potential digest products for self-ligation was from 4-5 kb, this size includes the 2066 bp product and as stipulated by Ochman *et al*, 2-3 kb of flanking sequence.

3.1.5.1 Enzyme Identification; Southern Blotting

GCG was used to generate a restriction map (Section 2.2.11) of the 2066 bp *nlEF-2* fragment and to identify restriction enzymes that did not cut within the fragment (Figure A3.1 A)). The *N. crassa*, *C. parasilosis*, *C. albicans EF-2* genes were also restriction mapped (data not shown) to find conserved restriction sites in the *nlEF-2* gene outside the cloned 2066 bp fragment which, if present in Lp19, would interfere with the IPCR. To minimize the size of the amplicons but still retain significant chance of amplifying the remaining *nlEF-2* gene six bp cutters were used (*EcoRV*, *Pvu I*, *EcoRI*, *Pvu II*, *BamHI*, *Mlu I*, *Sca I*, *Nde I*, *Sac I*, *Sma I* and *Nar I*).

The 2066 bp *nlEF-2* was amplified for a probe (PrEF2.1) by PCR using EF2-IntF1 and EF2-IntR1 primers with pDEF-1 as a template (Section 2.2.9 using PR1, Table A2.1). It was verified that PrEF2.1 was part of the *nlEF-2* gene by ClustalX alignment to the pDEF-1 sequence from which it was derived (Section 2.2.11). As more enzymes were tested than fit onto one blot, labeled PrEF2.1 was hybridized to two Southern blots (Figures 3.3 A) and B)) containing 1.0 µg digested Lp19 genomic DNA as well as 1.0 ng *Nco I* digested pDEF-1 (positive control; Section 2.2.13). PrEF2.1 hybridized to single Lp19 genomic DNA fragments of approximately 6.0 kb (*EcoRV*), 6.2 kb (*Pvu I*), 8.7 kb (*EcoRI*), 11.5 kb (*Pvu II*), slightly in excess of 12 kb (*BamHI*) in Figure 3.3 A) and 4.8 kb (*Mlu I*), 8.0 kb (*Sca I*), 9.5 kb (*Nde I*), 11 kb (*Sac I*) and slightly in excess of 12 kb (*Sma I* and *Nar I*) in Figure 3.3 B). Both *EcoRV* and *Pvu I* IPCR reactions would result in close to 4 kb sized products. As expected the PCR product hybridized to 696, 1066 and 3320 bp fragments of digested pDEF-1 in both blots. This is consistent with the location of the two *Nco I* sites on the restriction map of the putative *nlEF-2* fragment and the single *Nco I* site in the plasmid in relation to the insertion point in the pGEM-T easy vector (Figure A2.1). Using these results it is deduced that *Mlu I* is the optimal enzyme for use in following IPCR reactions and a 3 kb product would be expected. *Pvu II* and *EcoRV* were also identified as potential enzymes for use in IPCR.

3.1.5.2 Inverse PCR (IPCR)

Concurrent to the Southern analysis IPCR primers, EF2-IPCR 5' and EF2-IPCR 3' (Table 2.2) were designed 295 bp and 57 bp respectively from the 5' and 3' ends of the *nlEF-2* fragment

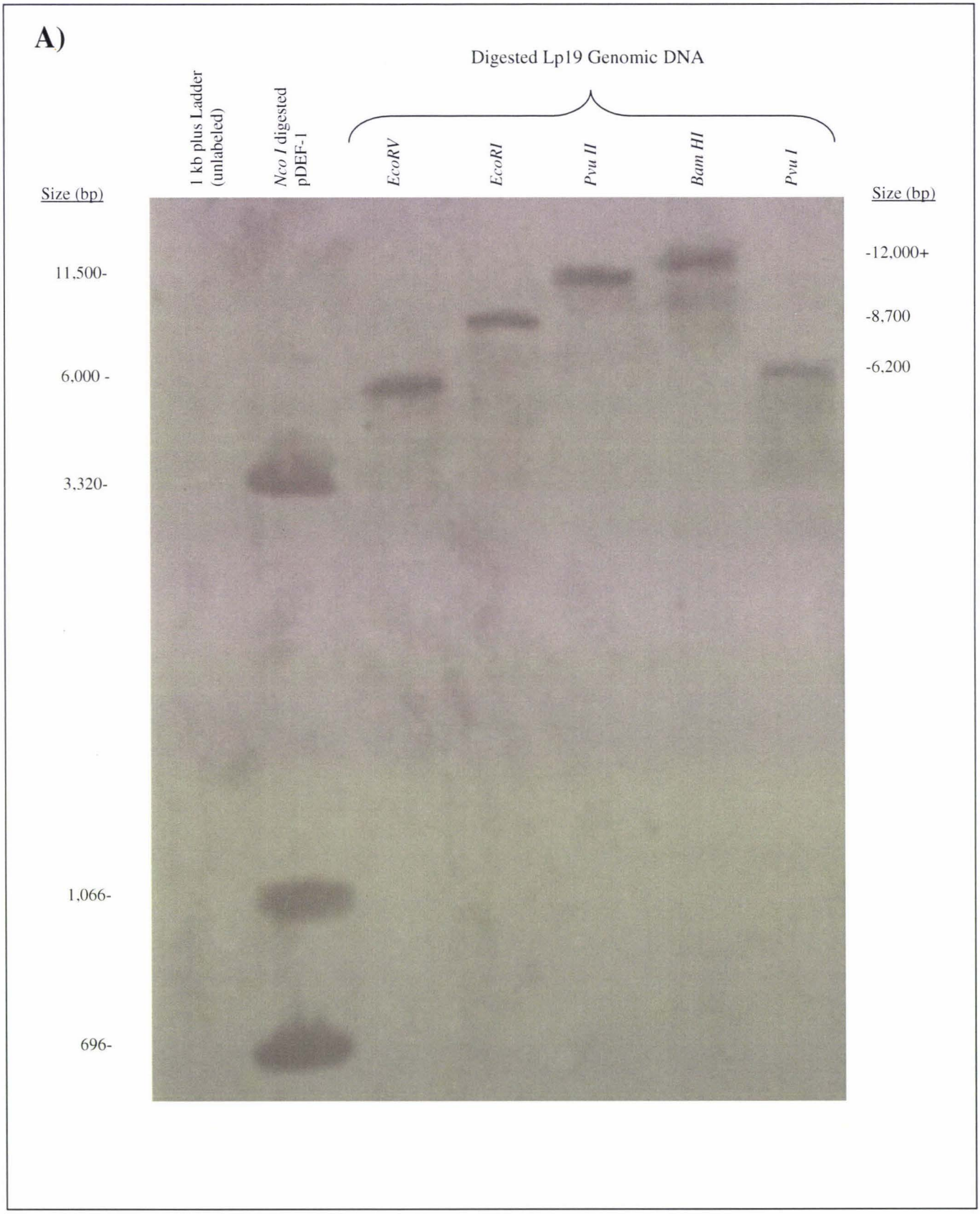
within pDEF-1 respectively (the *nlEF-2* 'core region'; Figure 3.4). Both primers were designed to prime in the opposite orientation from one another, therefore a product would only result if the DNA was cut and religated creating circular fragment. These sites were chosen for several reasons; firstly an overlap in sequence is required to positively identify sequencing from both ends as being part of the putative *nlEF-2* sequence already cloned (Figure 3.4). Secondly, EF2-IPCR 5' was designed further into the *nlEF-2* core region to act as a positive control to confirm the presence of the appropriate fragment in the template for IPCR reactions and to screen fragments resulting from IPCR reactions. Furthermore, exEF2-Fwd was designed from the region containing dEF2-Fwd for control use with EF2-IPCR 5' (Figure 3.4). *Mlu I* digests of 10 µg Lp19 genomic DNA were electrophoretically separated, and a region between 4.5 and 5.5 kb was excised and purified.

Purified fragment pooled from several reactions were ligated at 4°C and ligation products were purified using a phenol/chloroform extraction (Sections 2.2.7.2 and 2.2.3.1). The success of the ligation was monitored by electrophoresis. A preliminary control PCR reaction was performed using exEF2-Fwd and EF2-IPCR 5' and as expected this resulted in a 295 bp band (Figure 3.4), confirming that the fragment of interest was present in the IPCR template prepared and deeming it IPCR-ready (results not shown). Subsequent to these preparation procedures and control reactions, IPCR was executed using a long-range PCR protocol (Section 2.2.14).

Initial IPCR reactions with EF2-IPCR 5' and EF2-IPCR 3' were optimised using a range of annealing temperatures between 52 and 62°C, and varying amounts of IPCR-ready DNA as a template. Control templates included uncut Lp19 genomic DNA and *Mlu I* digested (unligated) Lp19 genomic DNA (Section 2.2.5). Single primer and water IPCR reactions were also used as controls in both sets of reactions. A temperature of 56°C using 100 ng of IPCR-ready DNA template resulted in a good yield of both ~2000 bp and ~1000 bp products; however neither was of expected size (Figure 3.5; PR3 Table A2.1). IPCR reactions using the Lp19 genomic DNA and *Mlu I* digested Lp19 genomic DNA did not result in any products. The IPCR-ready reactions resulted in two products of ~2000 bp and ~1000 bp respectively, validating the two products as the result of the IPCR treatments (results not shown). Single primer reactions showed the ~2000 bp band to be a product of EF2-IPCR 3' (Figure 3.5). Although several more IPCR reactions were performed, no 2800 bp fragment resulted using EF2-IPCR 5' and EF2-IPCR 3'. The ~2000 bp product was gel purified and a preliminary control PCR reaction was performed with the extracted product using EF2-Fwd and EF2-IPCR 5'. Unexpectedly a 295 bp resulted (Figure 3.4; Section 2.2.9 using PR4 Table A2.1), confirming that the fragment of

Figure 3.3: Southern Analysis for *nlEF-2* Inverse PCR

A) and B): Blots contained 1.0 µg Lp19 genomic DNA digested separately with enzymes indicated at top of wells and were hybridized with Digoxigenin-11-dUTP (Dig) labeled PrEF2.1. 1.0 µg unlabeled 1kb plus ladder (Invitrogen) serves as a negative control to show that DIG labeling is the source of fluorescence. Both autoradiographs were 30 min exposures using a 100 plus Automatic X-Ray Film Processor (All-Pro Imaging). Dig high-prime was used to label PrEF2.1 and the concentration of labeled probe estimated using a spot test. Hybridizations were performed at 55°C and the blots were washed at 55°C (Section 2.2.13.3).


A)
Digested Lp19 Genomic DNA
1 kb plus Ladder (unlabeled)
Nco I digested pDEF-1
EcoRV
EcoRI
Pvu II
Bam HI
Pvu I
Size (bp)
11,500-
6,000 -
3,320-
1,066-
696-
Size (bp)
-12,000+
-8,700
-6,200

B)

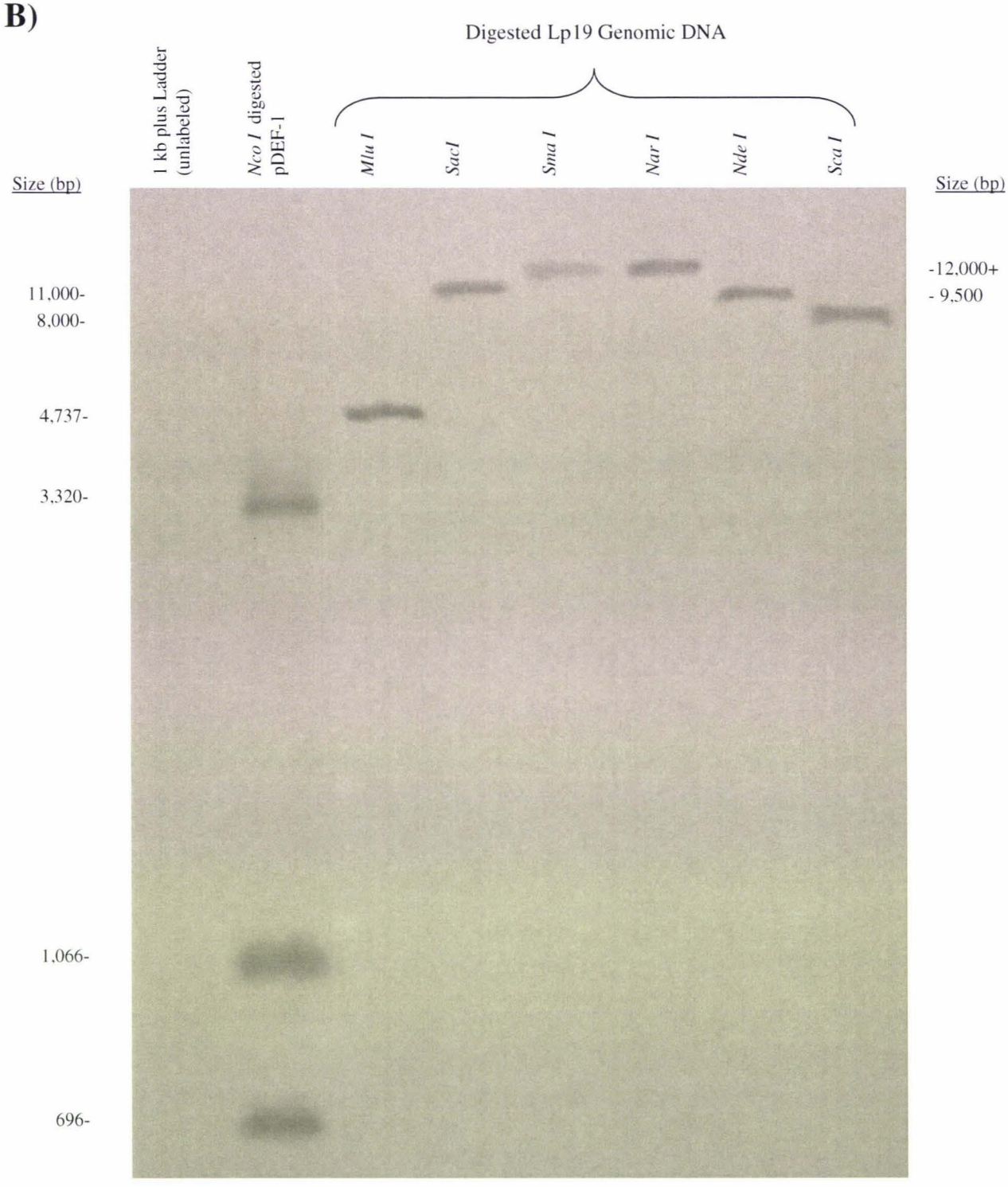
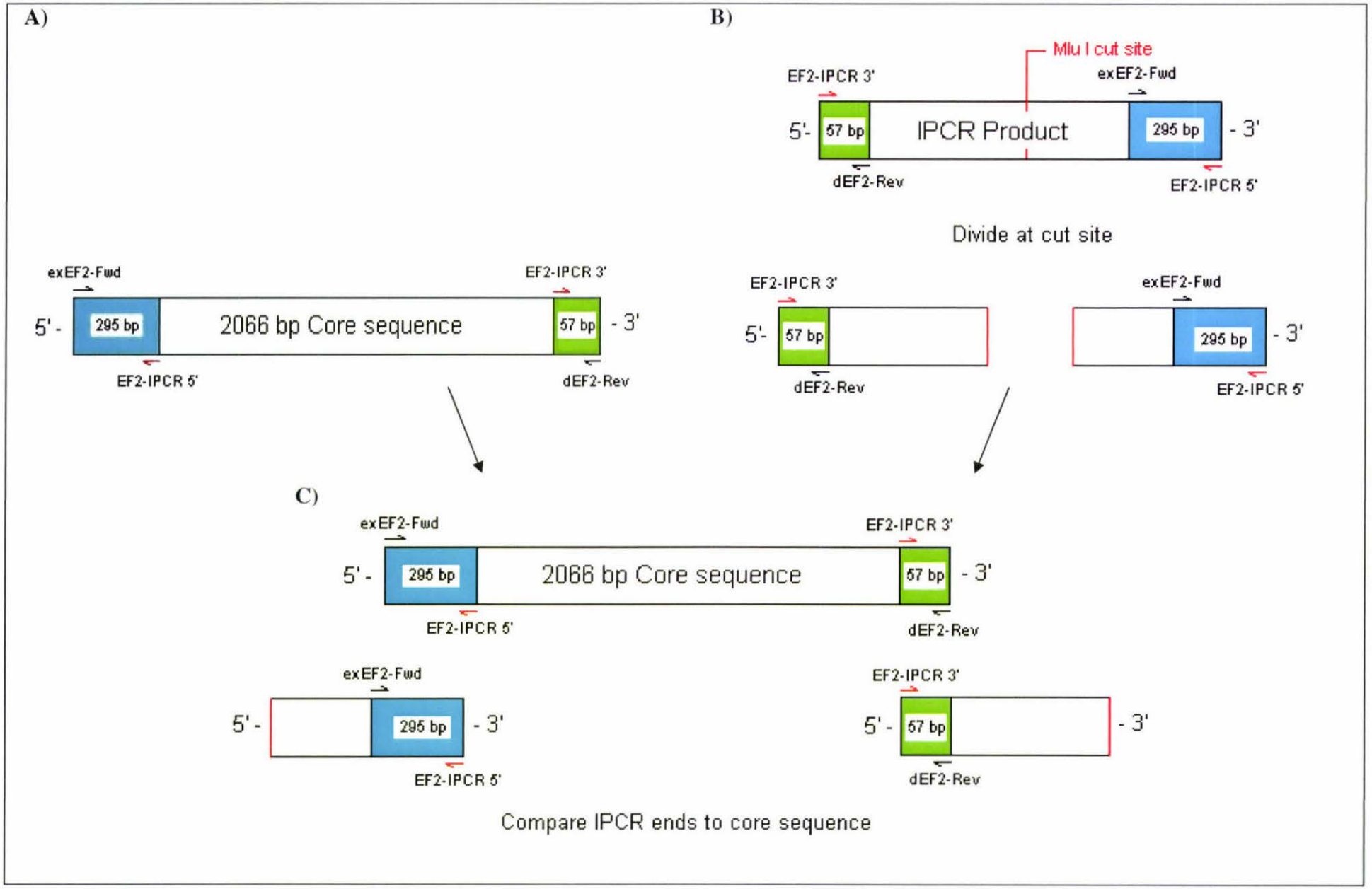

Digested Lp19 Genomic DNA
1 kb plus Ladder (unlabeled)
Nco I digested pDEF-1
Mlu I
Sac I
Sma I
Nar I
Nde I
Sca I
Size (bp)
11,000-
8,000-
4,737-
3,320-
1,066-
696-
Size (bp)
-12,000+
- 9,500

Figure 3.4: A Theoretical construction of a Complete *nlEF-2* gene from 2066 bp PCR product and the expected Inverse PCR product.

A) A schematic of the original 2066 bp degenerate PCR product. The relative location of the original PCR primers ex/dEF2-Fwd and dEF2-Rev (black arrows), and inverse PCR primers EF2-IPCR 5' and EF2-IPCR 3' (red arrows). The locations of expected sequence overlap with the inverse PCR product are shown in bright green and light blue, the latter being the site of the 295 bp internal control fragment.

B) A schematic of the expected IPCR product showing the expected overlapping sequence regions (colour-coded to the 2066 bp product A)) and the relative location of all PCR primers are shown as in A). A single *Mlu I* site is expected to be located somewhere between the 5' end of ex/dEF2-Fwd and the 3' end of dEF2-Rev (depicted as a red line). Subsequent to sequencing and identification of the *Mlu I* site, the IPCR sequence is divided at the *Mlu I* site and compared to the 2066 bp PCR product in C).

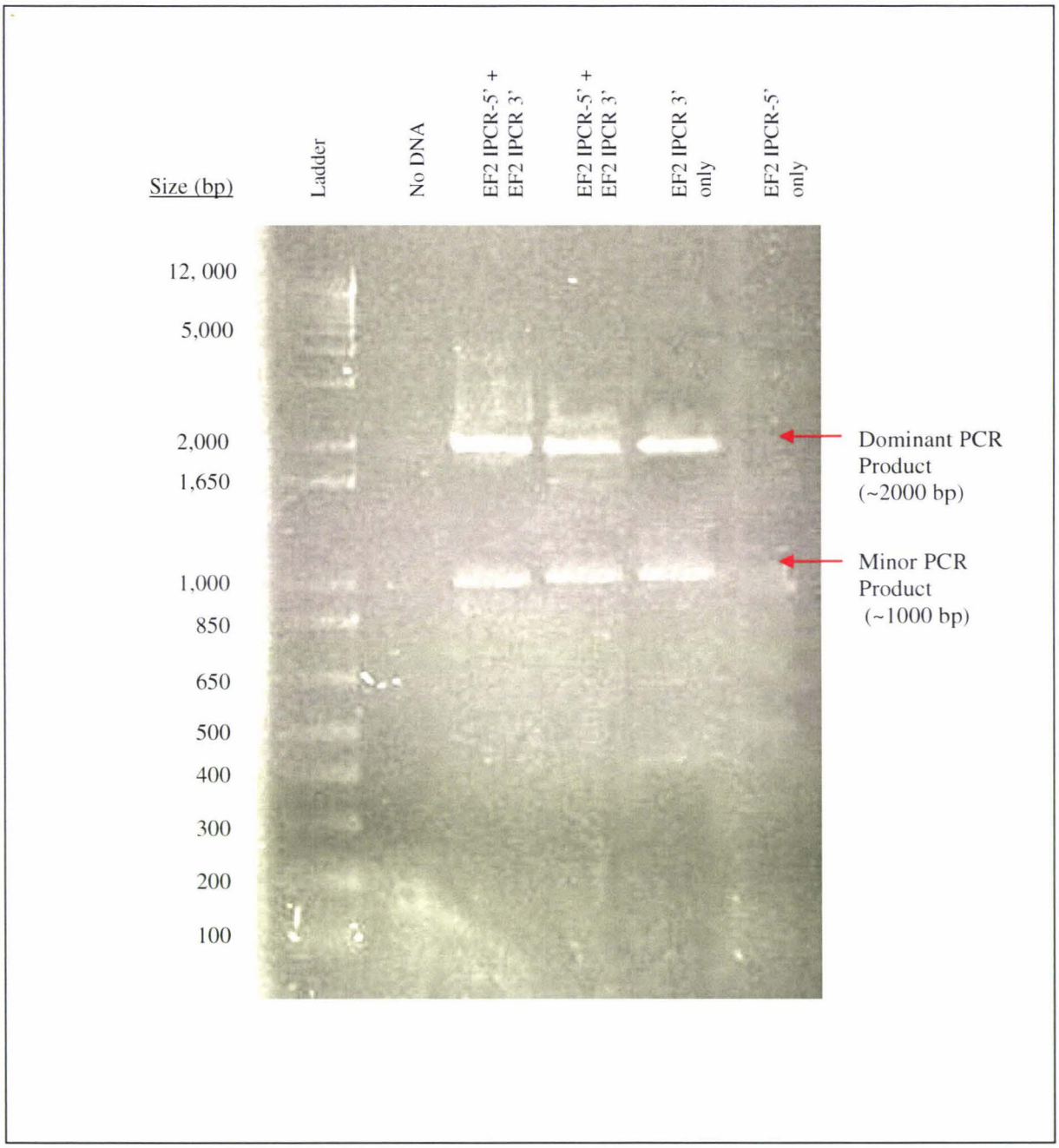

A)
B)
Mlu I cut site
EF2-IPCR 3'
exEF2-Fwd
5'-
57 bp
IPCR Product
295 bp
- 3'
dEF2-Rev
EF2-IPCR 5'
Divide at cut site
exEF2-Fwd
EF2-IPCR 3'
5' -
295 bp
2066 bp Core sequence
57 bp
- 3'
EF2-IPCR 5'
dEF2-Rev
C)
Compare IPCR ends to core sequence

interest was present in the IPCR extracted band and raising the possibility of two partial sequences of the *nlEF-2* gene (results not shown).

Initially an attempt was made to directly sequence (Section 2.2.10) using EF2-IPCR 5' and EF2-IPCR 3', however the electropherogram showed two sequences on top of one another. Taken together with the preliminary control PCR this strengthened the possibility of two very similar sized products. The purified fragments were then cloned into a pGEM-T easy vector. An additional PCR screen was performed with individual clones; bulk growth and screening methods (Sections 2.2.2.3 and 2.2.9 using PR4 Table A2.1) using exEF2-Fwd and EF2-IPCR 5' as a method to discern between the possible plasmid populations. A 295 bp product was expected from a plasmid containing the *nlEF-2* 5' end. However none of the 250 clones screened contained a plasmid which produced this band.
These clones yielded the expected ~2.3 kb fragment, and plasmid containing the putative *nlEF-2* IPCR fragment (pDEF-2) was extracted from the clones (Section 2.2.2.4). Three pDEF-2 inserts were partially sequenced using pUC/M13 universal primers. This resulted in 600 bp of reliable sequence from the 5' end and 750 bp of sequence from the 3' end of the fragment.

Two new primers, EF2-IntF2 and EF2-IntR2 (Table 2.2) were then designed to these sequences and used to sequence the central remaining part of the fragment. Sequence assembly yielded a 2042 bp contig. The three independent inserts were of identical sequence and as expected a solitary *Mlu I* site was found. The sequence was then divided at this restriction site to yield a 1017 bp 5' fragment and a 1025 bp 3' fragment and this screened for homology against the core region. It was established that 57 bp from the 5' fragment exactly matched the expected overlap in the 3' end of the core region (Figure A3.2), and when subjected to a BLASTn search the *N. crassa* Cot3 gene was the most significant match (7×10^{-79} E-value; 83 % match) for the 5' 476 bp of this sequence including the overlap. However, the 3' fragment showed no homology to any region in the core fragment and furthermore showed no homology to other *EF-2* sequences in a BLASTn search. The short or exact sequence matches tool on the NCBI website was employed as a tool to identify primer sites in all orientations (Section 2.2.11). This showed the 5' fragment had the expected sequence of EF2-IPCR 3' at its 5' end and the 3' fragment had the unexpected exact sequence of EF2-IPCR 3' at its 3' end in reverse (uxIPCR 3' site; Figures 3.8 and 3.9), therefore describing how this 2042 bp fragment resulted from the IPCR. Furthermore a third potential degenerate EF2-IPCR 3' site was found 23 bp upstream of the *Mlu I* cut site (deuxIPCR 3' site; Figures 3.8 and 3.11).

Figure 3.5: First Attempt of IPCR amplification of putative *nlEF-2* gene fragments
20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. 100ng of Lp19 IPCR-ready DNA was used in all reactions. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1 kb plus ladder (Invitrogen) was loaded in lane 1.

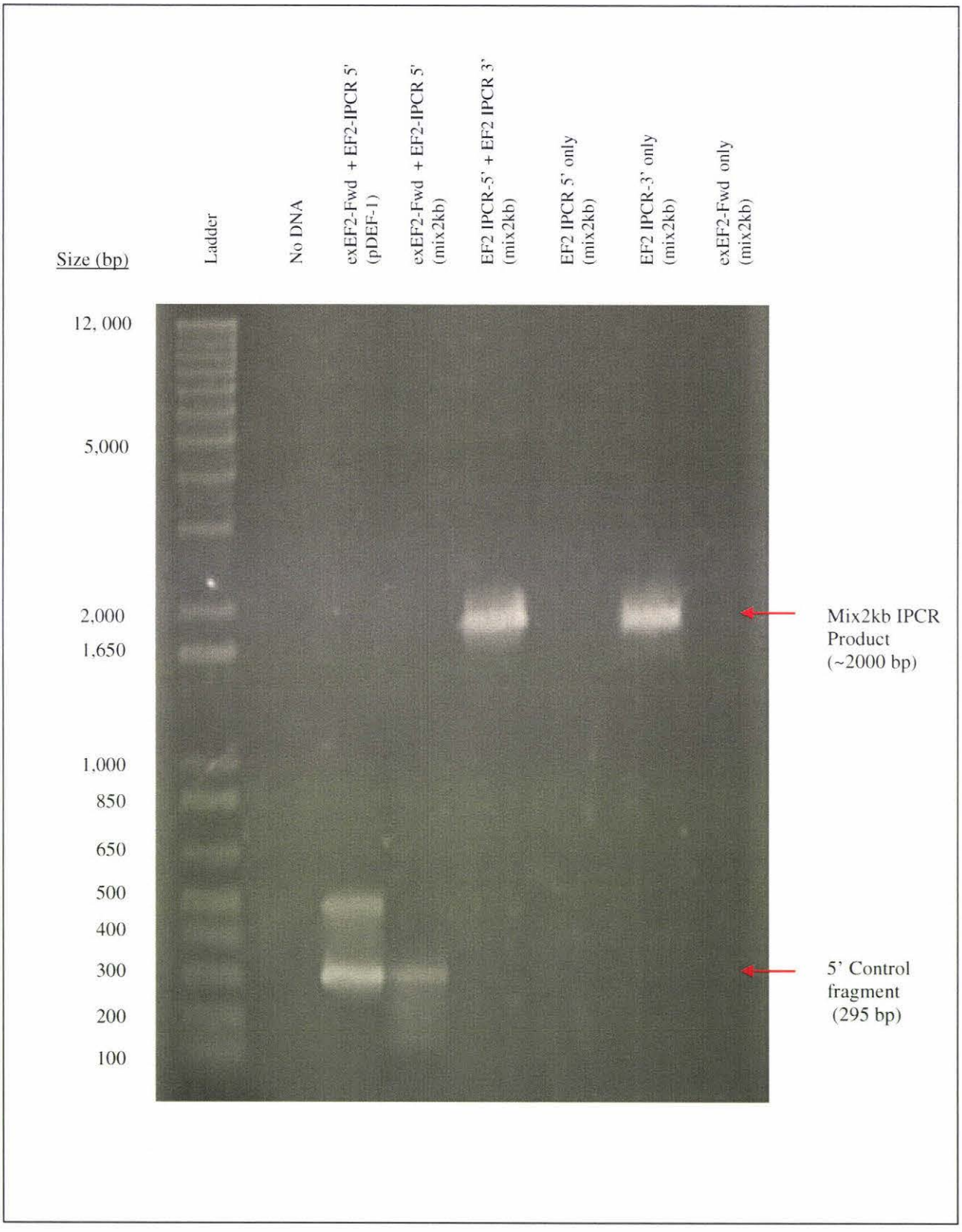

Size (bp)
Ladder
No DNA
EF2 IPCR-5' + EF2 IPCR 3'
EF2 IPCR-5' + EF2 IPCR 3'
EF2 IPCR 3' only
EF2 IPCR-5' only
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
Dominant PCR Product (~2000 bp)
Minor PCR Product (~1000 bp)

Taking this information together with the BLASTn results and the knowledge of the directionality of the priming; the 5' fragment was assembled into a contig using FAS within GCG and the EF2 core region (EF2-con1; Figure A3.2; Section 2.2.11) and the 3' fragment was hypothesized to be an upstream fragment (EF2-ups; Figure A3.2). In addition the EF2-IPCR 3' degenerate site indicated that the afore-mentioned hypothetical second product may be a product of EF2-IPCR 3' and EF2-IPCR 5' as a second set of direct sequencing reactions using EF2-IPCR 3' with a product (Figure 3.6; purified from the single primer product reactions) gave a clean single sequence (Section 2.2.10) that was an exact match (using ClustalX; Section 2.2.11) to the pDEF-2 sequences. The ~1000 bp product shown in figure 3.5 is a product of EF2-IPCR 3' alone, and upon sequence analysis (Figure 3.9) the uxIPCR 3' and deuxIPCR 3' sites are 1053 bp apart (Figures 3.8 and 3.11). Taking this information together is most likely the result of EF2-IPCR 3' priming at the uxIPCR 3' and deuxIPCR 3' sites (Figure 3.8). There are a number of possible reasons why the first IPCR failed to yield the expected 3 kb fragment, however the most likely is interference from more proximal primer sites. The lack of success in obtaining 5' 2030 bp clone may be due to less product than the 3' 2042 bp produced from the IPCR reaction (deuxEF2-IPCR 3' is a degenerate site and therefore less specific than the uxEF2-IPCR 3')

In summary only the 3' flanking region was cloned due to primer interference and single primer products. The 5' flanking region was yet to be cloned.

Figure 3.6: Identity verification of IPCR products

Each reaction used 10 ng of template. 20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was loaded in lane 1. The 500 bp product generated from the pDEF-1 is a pGEM-T vector product. This reaction is a positive control for the 5' verification reaction. Other reactions are to verify the identity of the template (recovery).

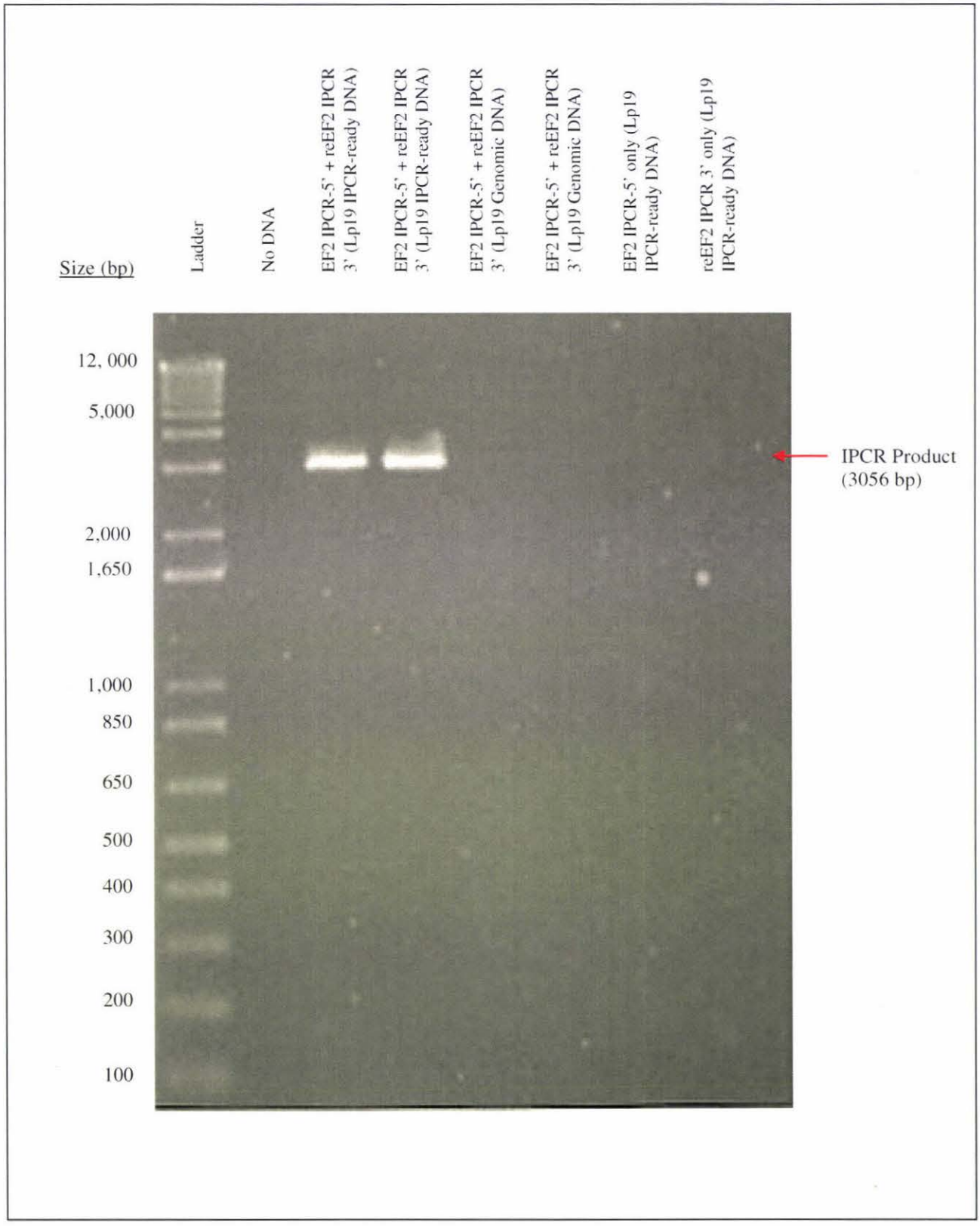

Size (bp)
Ladder
No DNA
exEF2-Fwd + EF2-IPCR 5' (pDEF-1)
exEF2-Fwd + EF2-IPCR 5' (mix2kb)
EF2 IPCR-5' + EF2 IPCR 3' (mix2kb)
EF2 IPCR 5' only (mix2kb)
EF2 IPCR-3' only (mix2kb)
exEF2-Fwd only (mix2kb)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
Mix2kb IPCR Product (~2000 bp)
5' Control fragment (295 bp)

Due to lack of success finding the 5' hypothetical PCR product amongst the IPCR clones, a new primer reEF2-IPCR 3' was designed from the core region 11 bp upstream of EF2-IPCR 3' (Table 2.2). IPCR was performed with EF2-IPCR 5' and reEF2-IPCR 3' using a long-range PCR protocol (Section 2.2.14 using PR5 Table A2.1). The identical conditions and initial reactions were followed as with EF2-IPCR 5' and reEF2-IPCR 3' however the estimated product size was ~3000 bp. A temperature of 58°C using 100 ng of IPCR ready DNA template resulted in a good yield of a ~3000 bp product (Figure 3.7). The ~ 3000 bp product was gel purified and a preliminary control PCR reaction was performed with the extracted product using EF2-Fwd and EF2-IPCR 5'. As expected a 295 bp resulted, confirming that the fragment of interest was present in the IPCR the extracted band.

The purified fragment was then cloned into a pGEM-T easy vector and clones were stored as previously described in section 3.1.3. These clones yielded the expected ~3.3 kb fragment, and plasmid containing the putative *nlEF-2* IPCR fragment (pDEF-3) was extracted from the clones (Section 2.2.2.4). The three pDEF-3 inserts were partially sequenced using pUC/M13 universal primers. This resulted in 800 bp of reliable sequence from the 5' end and 700 bp of sequence from the 3' end of the fragment. Two new primers EF2-IntR3 and EF2-IntR4 (Table 2.2) were then designed to these sequences and used along with EF2-IntF2 and EF2-IntR2 to sequence the central remaining part of the fragment. Sequence yielded a 3056 bp contig. The three independent inserts were of identical sequence and as expected a solitary *Mlu I* site was found. The sequence was divided at this restriction site to yield a 1049 bp 5' fragment and a 2007 bp 3' fragment and these screened for homology against the core region.

It was established that 57 bp from the 5' fragment and 299 bp from the 3' fragment exactly matched the expected overlap in the 3' end of the core region and 5' end of the core region respectively (Figure 3.8 and Figure A3.2). In order to independently verify the identity of both fragments, each was individually subjected to a BLASTn search (Section 2.2.11). As expected the *N. crassa* Cot3 gene was most significant match for both the 5' and 3' fragments and the 5' fragment had a single alignment site from 1-476 (476 bp) with an E-value of 7×10^{-79} and 83 % identity. The 3' fragment had 3 alignment sites; these are from positions 1285 – 1311 (27 bp; 0.001 E-value; 100 % identity), 1649 – 1873 (225 bp; 7×10^{-39} E-value; 84 % identity) and 1934 – 2007 (71 bp; 8×10^{-14} E-value; 90 % identity). Putative Lp19 *nlEF-2* sequence was present in EF2-con2. To confirm hypotheses about the origin of the first IPCR product ClustalX alignments of EF2-con2 with EF2-ups and EF2-con1 were performed (Figures 3.8 and A3.2; section 2.2.11). The exact homology of both EF2-ups and EF2-con1 to EF2-con2

Figure 3.7: Second Attempt of IPCR amplification of putative *nlEF-2* gene fragments
Each reaction used 100 ng of template. 20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was


Size (bp)
Ladder
No DNA
EF2 IPCR-5' + reEF2 IPCR 3' (Lp19 IPCR-ready DNA)
EF2 IPCR-5' + reEF2 IPCR 3' (Lp19 IPCR-ready DNA)
EF2 IPCR-5' + reEF2 IPCR 3' (Lp19 Genomic DNA)
EF2 IPCR-5' + reEF2 IPCR 3' (Lp19 Genomic DNA)
EF2 IPCR-5' only (Lp19 IPCR-ready DNA)
reEF2 IPCR 3' only (Lp19 IPCR-ready DNA)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
IPCR Product (3056 bp)

further supported hypothesis of how errors occurred and offered confirming evidence for size estimates for the first IPCR products in Figure 3.5.

Each position is given relative to the total size of the *Mlu I* fragment (5' and 3' fragments collectively) These results strongly suggest that the IPCR fragment contains the flanking regions of the 2066 bp fragment and therefore the remaining regions of the putative *nlEF-2* gene sequence. A full-length contig was assembled to create a 4737 bp sequence named EF2-con2 (Figure A3.2)

To confirm the full putative Lp19 *nlEF-2* sequence was within EF2-con2, a BLASTn search was conducted. Between nucleotides 1 and 2971 (Figure 3.11) sixteen of the 20 most significant matches were fungal *EF-2* genes (data not shown), with the *Cot3* gene from *N. crassa* (Q96X45) being the most significant (0.0 E-value; 87 % identity). No other regions in EF2-con2 had any homology to any other sequences, strongly suggested that the full gene sequence was obtained.

The Transeq program was used to predict an 844 residue amino acid sequence from EF2-con2 (Section 2.2.11) that showed strong sequence similarity when aligned *N. crassa Cot3* (EAA77131.1). The frame shift between ORFs and presence of stop codons between indicated the presence of three introns.

3.2 Intron confirmation and identification of transcription start/stop sites by Rapid Amplification of cDNA Ends (RACE)

To confirm the identification of putative introns and find the transcription start/stop sites, total RNA was isolated from actively growing Lp19 (Section 2.2.1.2.2) and analysed by RACE. In RACE technique, PCR is used to amplify cDNAs representing the region between a single point in an mRNA transcript and its 5' or 3' end. An internal stretch of sequence must already be known from the mRNA of interest. From this sequence, gene-specific primers are chosen that are oriented in the direction of the 5'- and 3'- ends. Like Inverse PCR this can be used to obtain a gene sequence, however this gives no information about flanking region gene structure. Gene-specific primers were designed to ensure sequence overlap so that the entire cDNA could be sequenced (see Figure 3.11; Table 2.2).

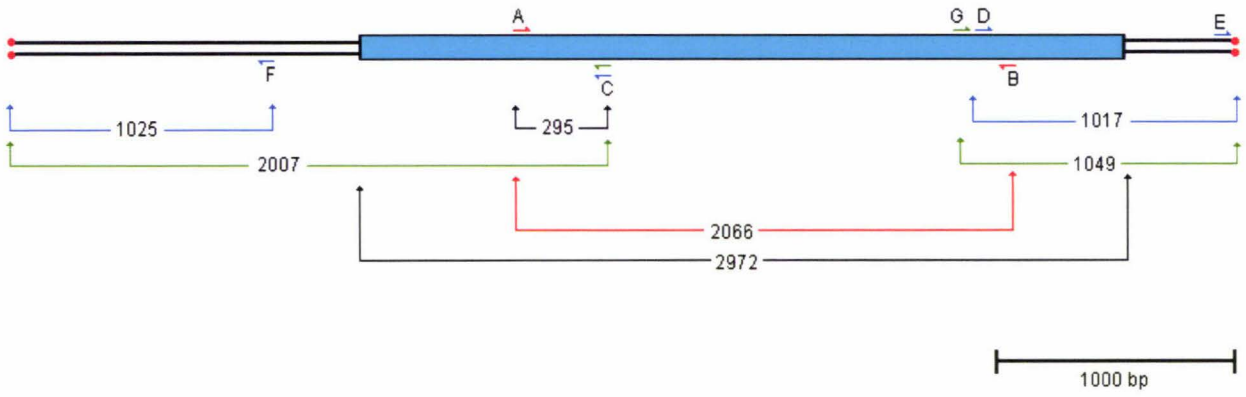
Figure 3.8: Schematic of Mapped EF2-con2 (Linear A) and Circular B))

The Lp19 *nlEF-2* gene coding region is identified by a light-blue enclosed box, the flanking regions by thick black lines and the Mlu I sites by red dots at the ends of these. Relative primer sites are labeled with letters (see key), and the sizes of products between each primer indicated in base-pairs. Each product is colour-coded to the primers used to obtain it; a blue line represents the 2042 bp IPCR Product (Figure 3.5); the second 3056 IPCR product by green (Figure 3.7), and the original 2066 bp PCR product (Figure 3.2) by red. The 295 bp internal control primer region and the total length of the gene are indicated in solid black. A broken line indicates the 2030 bp region (i-iii) where the second IPCR product from the first IPCR would be located and between points ii) – iii) on this is the 1053 bp product seen in Figure 3.5. Both figures compare the use of genomic DNA (A) to IPCR-ready self-ligated fragment (B) for IPCR. Shown together they demonstrate why no products are seen using A) as no primer sites meet where in (B) they do. The total length of the construct is 4737 bp.

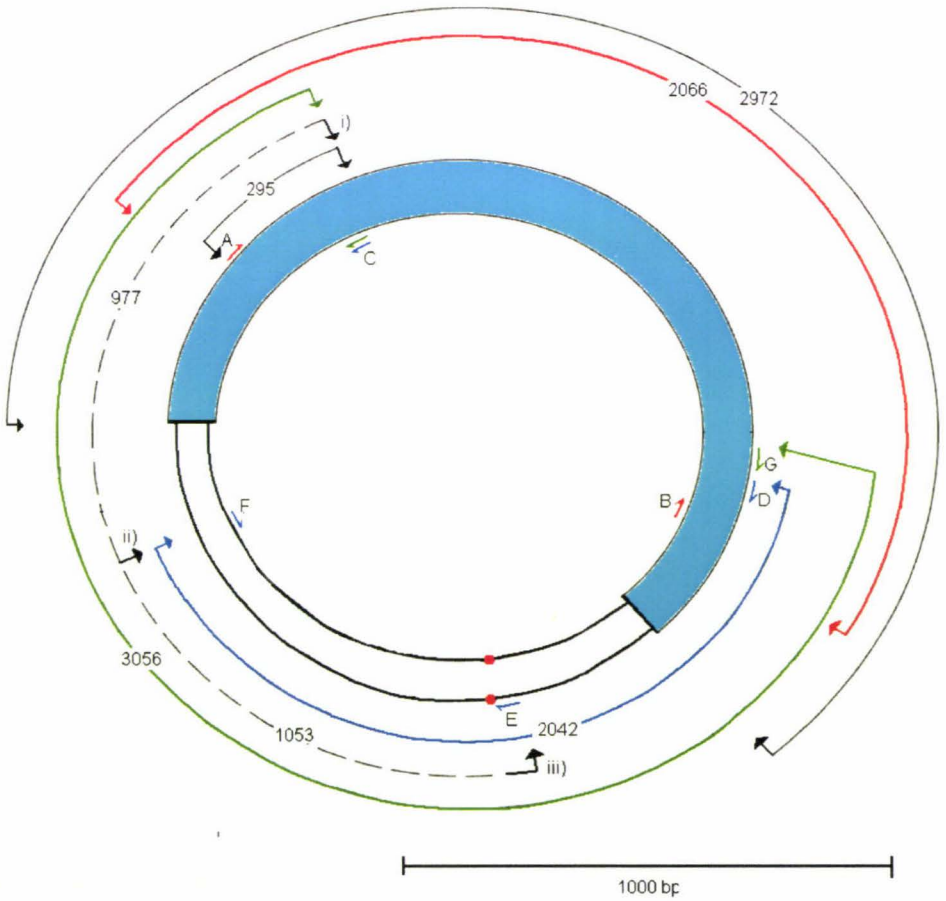
Key

A)	ex/dEF2-fwd	B)	ex/dEF2-rev
C)	EF2 IPCR-5'	D)	EF2 IPCR-3'
E)	deuxEF2 IPCR-3'	F)	uxEF2 IPCR 3'
G)	reEF2-IPCR-3'		

A)


A
G D
E
F
C
B
1025
295
1017
2007
1049
2066
2972
1000 bp

B)


2066
2972
i)
295
A
C
977
F
B
G
D
ii)
3056
E
2042
1053
iii)
1000 bp

To ensure no DNA was contaminating the cDNA template a control primer EF2-RACEc was designed from positions 962-985 within EF2-con2 (Table 2.2). This primer when used with EF2 5' RACE is expected to result in 287 bp and 316 bp products when used with uncontaminated Lp19 cDNA and genomic DNA templates respectively (between these two primers is a predicted 29 bp intron). Initially Lp19 total RNA was treated in two groups of procedures; firstly the 5' phosphate and cap structures were removed from mature mRNAs and an RNA oligo ligated as a substitute, this RNA oligo contains sites for GeneRacer 5' primer and 5' Nested primer.

Secondly, using this template RT PCR was conducted with an Oligo dT primer + another RNA oligo ligated at its 3' end. This 3' RNA oligo contains sites for GeneRacer 3' primer and 3' Nested primer, creating RACE-ready first-strand cDNA with known priming sites at the 5'- and 3'- ends. This preparation of RACE template was performed by Ningxin Zhang (IMBS, Massey University, Palmerston North; Section 2.2.18.2.1).

Initially two sets of PCR reactions for 5' and 3' RACE were performed using an annealing temperature of 65°C and 5 µl of three concentrations of serial diluted RACE-ready cDNA as template (1:20, 1:10 and 1:5) in a series of separate reactions (Section 2.2.18.2.2 using PR6 Table A2.1). Primers GeneRacer 5' and EF2 5' RACE were used for 5'- RACE while GeneRacer 3' and EF2 3' RACE were used for 3'- RACE. Following this initial reaction, the resulting product was diluted 1:100 and another round of PCR commenced using this template with primers GeneRacer 5' nested and EF2 5' RACE for 5'- RACE and GeneRacer 3'nested and EF2 3' RACE for 3'- RACE ((Templeton *et al.*, 1993); table 2.2). This step was taken to reduce background products and increase the specificity of the reaction.

When performing 5' RACE, a 1:5 dilution of template resulted in a 1023 bp product (Figure 3.9). For 3' RACE a 1:10 dilution of template resulted in a 2031 bp product (Figure 3.10). Several more reactions were carried out under the same conditions to obtain an adequate amount of product for purification. Single primer and water RACE reactions were also used as negative controls in both sets of reactions. The 1023 and 2031 bp products were gel purified and cloned into pGEM-T easy vectors. Using the previously described screen in Section 3.1.3, these clones yielded the expected ~1.4 and ~2.3 kb fragments for 5' and 3' RACE respectively, and plasmid containing each of the RACE products (pDEF-4; 5' RACE and pDEF-5; 3' RACE) was extracted from the clones (Section 2.2.2.4). The three independent pDEF-4 inserts were sequenced using pUC/M13 universal primers and the sequences assembled into a contig (EF2-con3). The three independent pDEF-5 inserts were partially sequenced using pUC/M13

Intron	5' Donor Site		Branch site		3' Acceptor site		Total Length (bp)
	Sequence	Start Point	Sequence	Start Point	Sequence	End point	
1	A-GTAAGT	9	ACTAAC	296	GTAG-C	355	346
2	G-GTACGT	1088	GCTAAC	1099	CTAG-G	1115	29
3	C-GTAAGT	2891	GCTAAC	2937	ACAG-T	2951	60
N. crassa Consensus	G-GTRMGY		RCTRAC		WYAG-G		

Table 3.1; Comparison of intron consensus sequences found within the putative *nlEF-2* gene to *N. crassa* consensus sequences. All sequences are listed 5' to 3'; left to right. Within the *N. crassa* consensus sequences M = C or A; R = A or G; W = A or T and Y = C or the 'Start/End points' are positions relative to the translation start point in figure 3.11 and the total length of each intron is indicated on the far left column in base pairs.

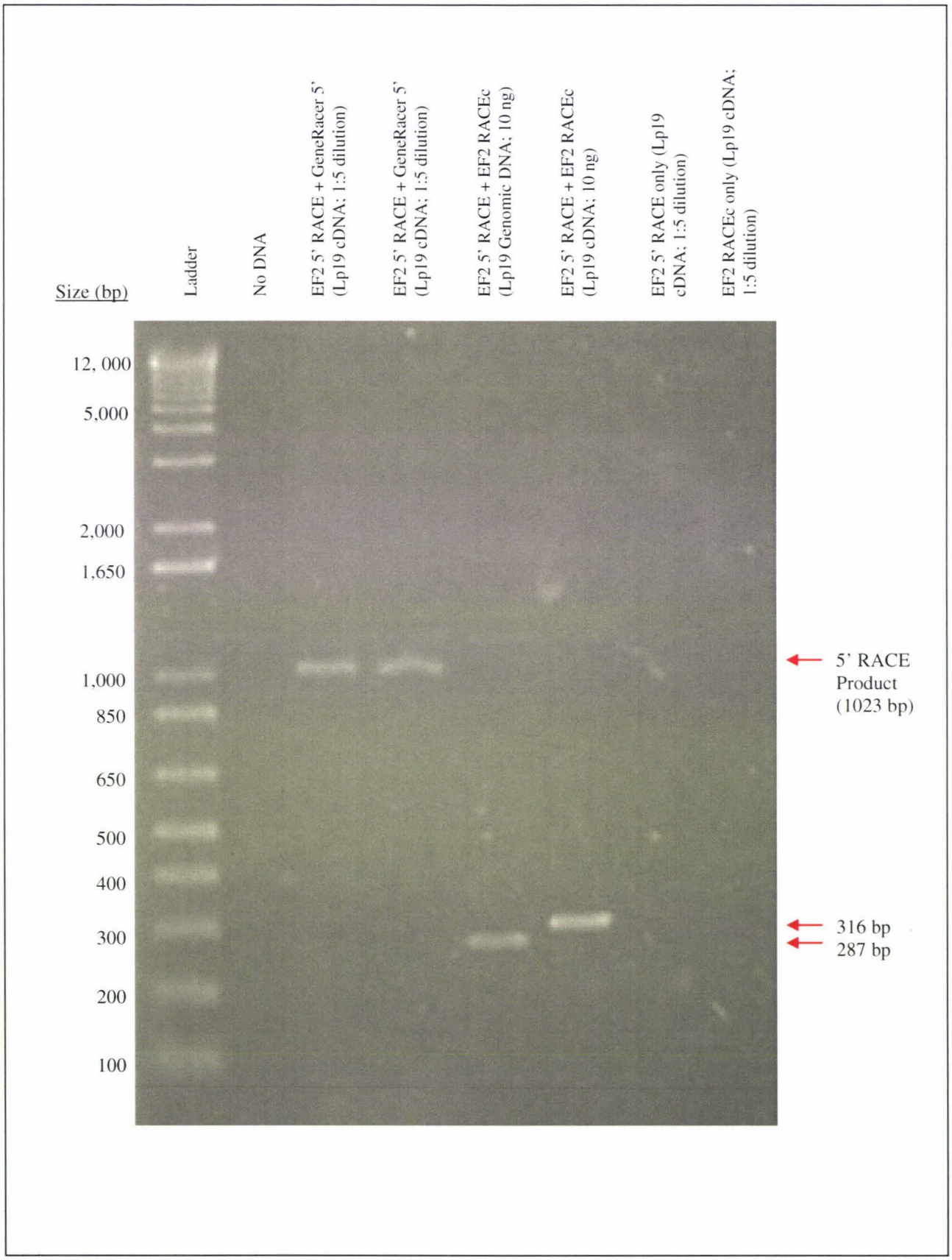
universal primers resulted in 700 bp of reliable sequence from the 5' end and 700 bp of sequence from the 3' end of the fragment. One new primer EF2-IntRa (Table 2.2) was designed to the 3' end of the pUC/M13L sequence and used along with exEF2-Rev to sequence the central remaining part of the fragment. Sequence assembly yielded a 2,949 bp contig (EF2-con4).

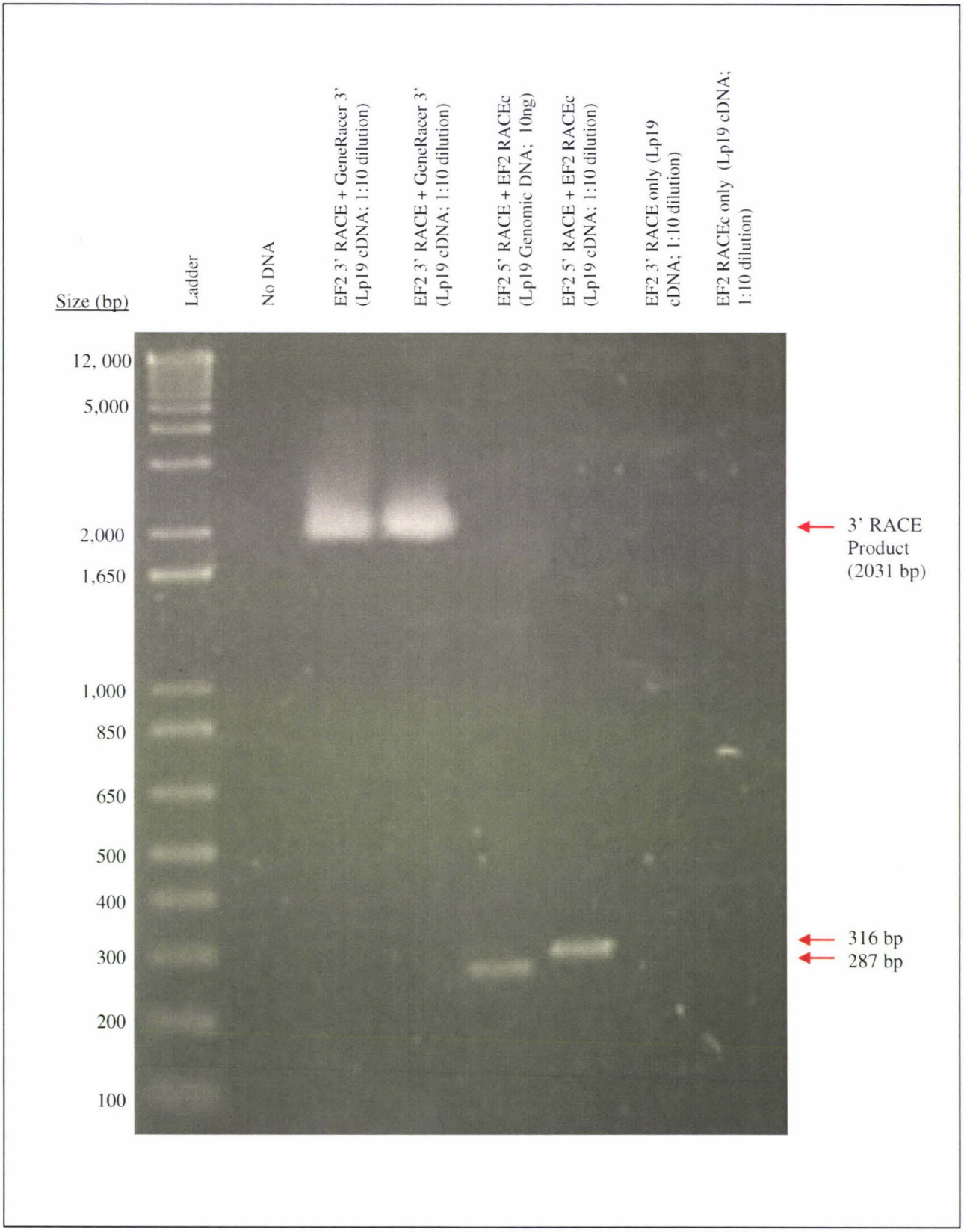
ClustalX alignments (Section 2.2.11) for each of the assembled sequences in both 5' and 3' RACE showed that the three contigs were identical. As additional controls the 316 and 287 bp products were extracted separately from the gels in figures 3.10 and 3.11 and direct-sequenced using EF2 5' RACE and EF2 RACEc. EF2-con2 (~160-3085), EF2-con5, and the two controls 316 + 287 bp products were subsequently aligned using ClustalX (Section 2.2.11; Figure A3.3).

Figures 3.9: *nlEF-2* 5' RACE

3.10: *nlEF-2* 3' RACE (following page)

20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1 kb plus ladder (Invitrogen) was loaded in lane 1. As expected EF2-RACEc and EF2 5' RACE control PCR reactions resulted in single 287 bp and 316 bp bands with Lp19 cDNA and genomic DNA templates respectively, validating the RACE result.


Size (bp)
Ladder
No DNA
EF2 5' RACE + GeneRacer 5' (Lp19 cDNA; 1:5 dilution)
EF2 5' RACE + GeneRacer 5' (Lp19 cDNA; 1:5 dilution)
EF2 5' RACE + EF2 RACEc (Lp19 Genomic DNA; 10 ng)
EF2 5' RACE + EF2 RACEc (Lp19 cDNA; 10 ng)
EF2 5' RACE only (Lp19 cDNA; 1:5 dilution)
EF2 RACEc only (Lp19 cDNA; 1:5 dilution)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
5' RACE Product (1023 bp)
316 bp
287 bp


Size (bp)
Ladder
No DNA
EF2 3' RACE + GeneRacer 3' (Lp19 cDNA; 1:10 dilution)
EF2 3' RACE + GeneRacer 3' (Lp19 cDNA; 1:10 dilution)
EF2 5' RACE + EF2 RACEc (Lp19 Genomic DNA; 10ng)
EF2 5' RACE + EF2 RACEc (Lp19 cDNA; 1:10 dilution)
EF2 3' RACE only (Lp19 cDNA; 1:10 dilution)
EF2 RACEc only (Lp19 cDNA; 1:10 dilution)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
3' RACE Product (2031 bp)
316 bp
287 bp

3.3 Analysis and Discussion of Sequence Results

3.3.1 *N. lolii* (Lp19) Protein Elongation Factor 2 (*nlEF-2*)

The *nlEF-2* gene was cloned from Lp19 by PCR, allowing sequencing and characterization of the entire coding region as well as the 5' and 3' non-translated and non-transcribed regions.

The *nlEF-2* gene of Lp19 was isolated by degenerate PCR using the primers dEF2-Fwd and dEF2-Rev. Although a range of products were amplified the most abundant product was a 2066 bp band, whose sequence suggested it is part of a putative *EF-2* gene. IPCR was used to obtain the flanking regions. Initially IPCR was only partially successful using the primers EF2-IPCR 5' and EF2-IPCR 3', which resulted in a 2042 bp fragment where 2800 bp was expected (Figure 3.5). Two additional EF2-IPCR 3' sites were identified using the sequence information (Figure 3.11), and their relative locations suggested why the initial PCR failed because of interference (Figure 3.8). A new IPCR 3' primer was designed, the process repeated and this resulted in a 3056 bp product containing the full flanking regions (Figure A3.2). The flanking regions and 2066 bp degenerate PCR product were constructed into the full *nlEF-2* sequence.

To identify the *nlEF-2* transcribed region RACE was performed using an overlapping set of primers for the 5' and 3' ends. After sequencing both products were assembled into a contig containing the full mRNA sequence, identifying the transcription start site, polyadenylation site and three intronic regions. This revealed that the *nlEF-2* transcripts are approximately 2,949 bp in length.

Lp19 has a haploid genome and has previously been shown by Southern analysis to possess a single *HMG CoA reductase* gene (Dobson, 1997). A similar pattern was observed in EF-2 homologues as *nlEF-2* hybridized to a single band per digestion, indicating a single copy was present, where two copies of both genes were present in *S. cerevisiae* and *U. maydis*.

Figure 3.11: Full Nucleotide sequence of EF2-con2

The 4737 bp EF2-con2 sequence contains the full *nlEF-2* gene (1-2972), a portion of the 5' non-coding region (-1296-1) and 3' non-coding sequence (2973-3340). Numbering is relative to the putative translation start site. The *Mlu I* cut site is indicated by a backslash, the sequence underlined.

- Primer locations and *nlEF-2* Coding region features:

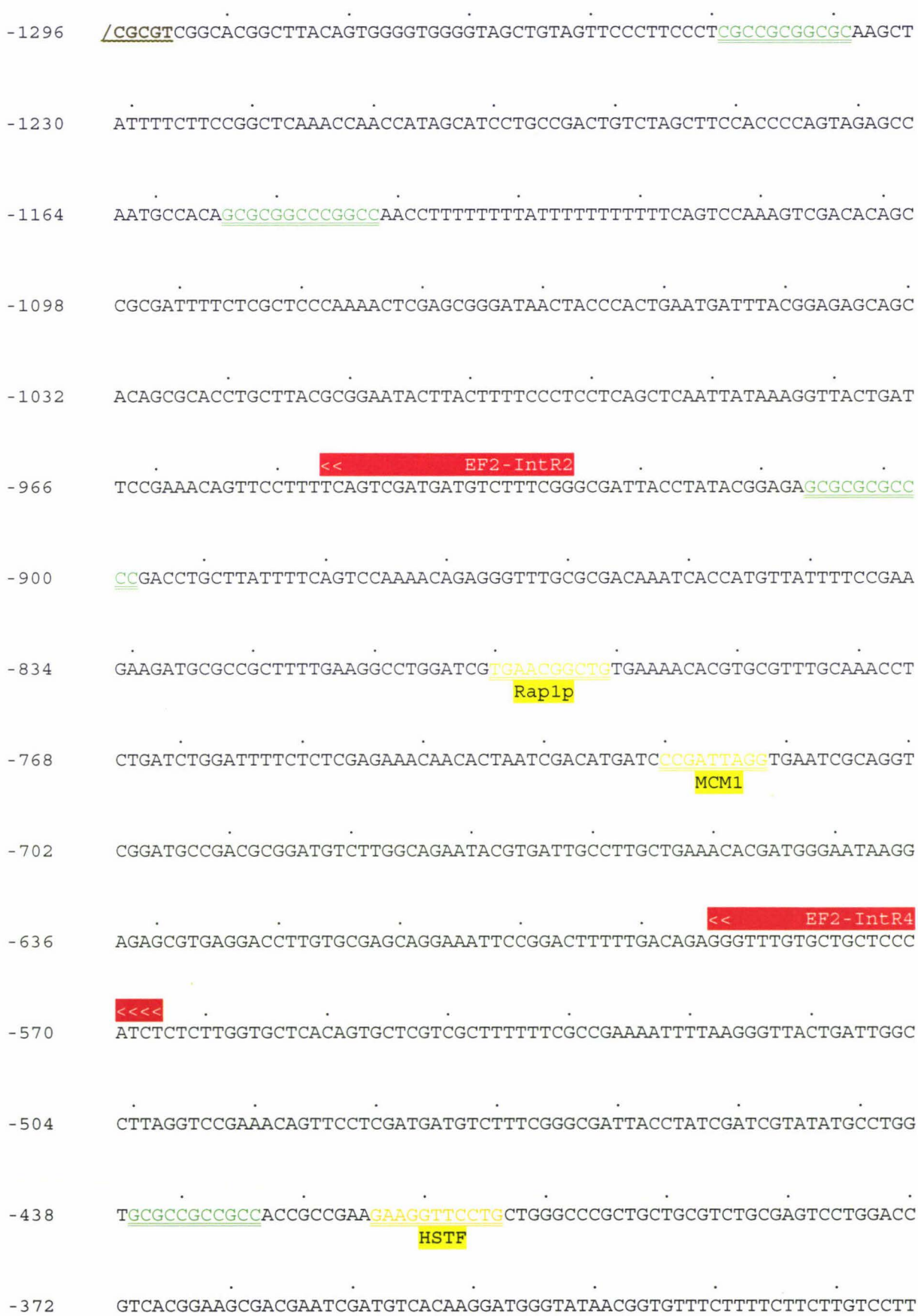
The amino acid residues encoded are positioned beneath the corresponding nucleotide sequence, with putative translational start and stop sites shown in red text. Blue lettering indicate the introns present, intron consensus sequences are boxed. Various classes of primer sites are indicated by highlighted labeling beneath corresponding sequence. Red primers indicate the sites for original degenerate PCR and sequencing primers. Dark green and dark blue indicate IPCR primer sites and unexpected primer sites respectively. Pink primers indicate RACE primer sites.

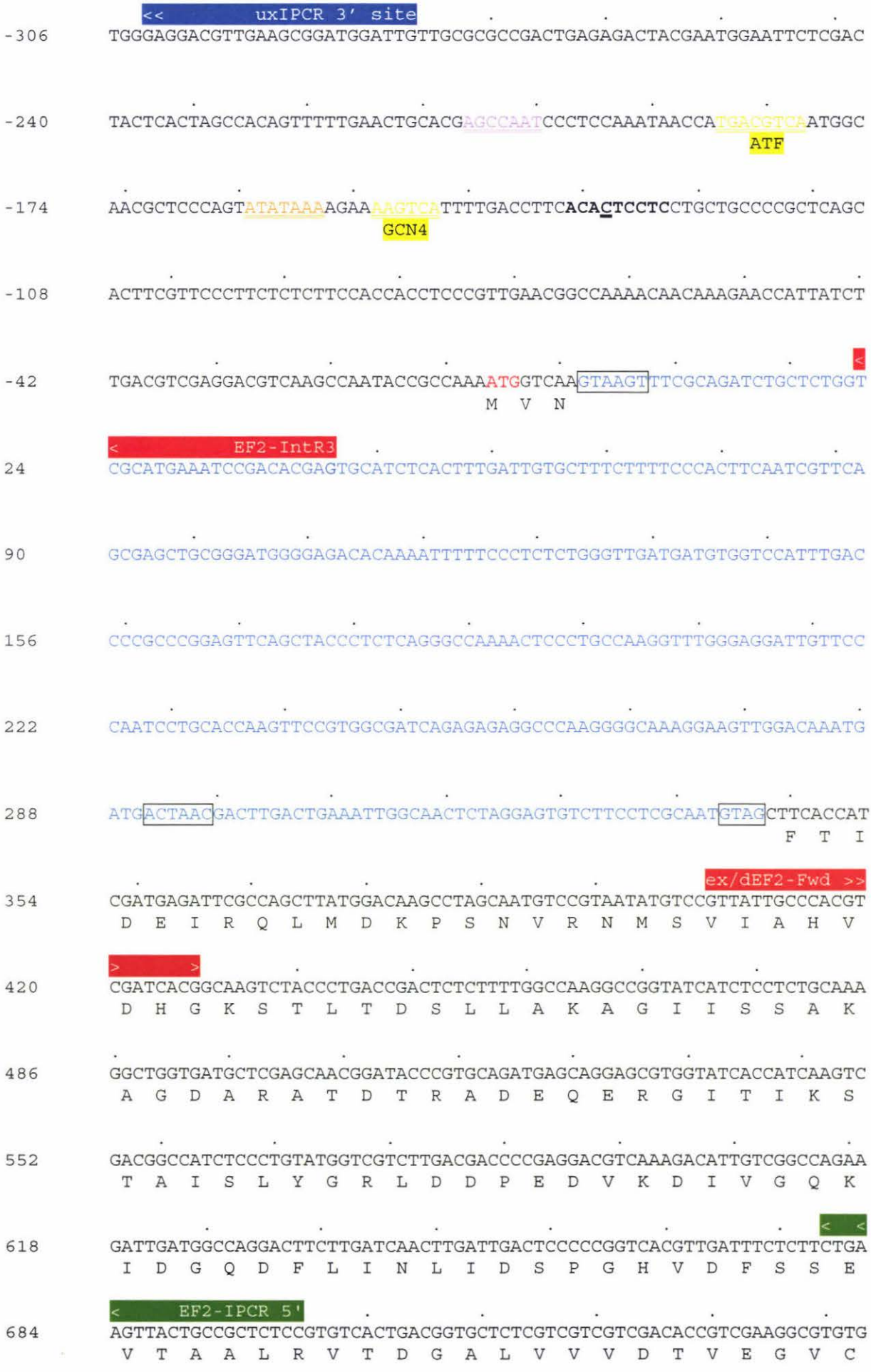
- 5' Non-Coding region

The sequence surrounding the transcriptional start site is in bold, actual site is also underlined; all other features are double underlined. Putative TATA, CCAAT and GC boxes are purple, orange and dark green lettering respectively. Possible promoter binding sites are in yellow, the type of binding site is highlighted below each site.

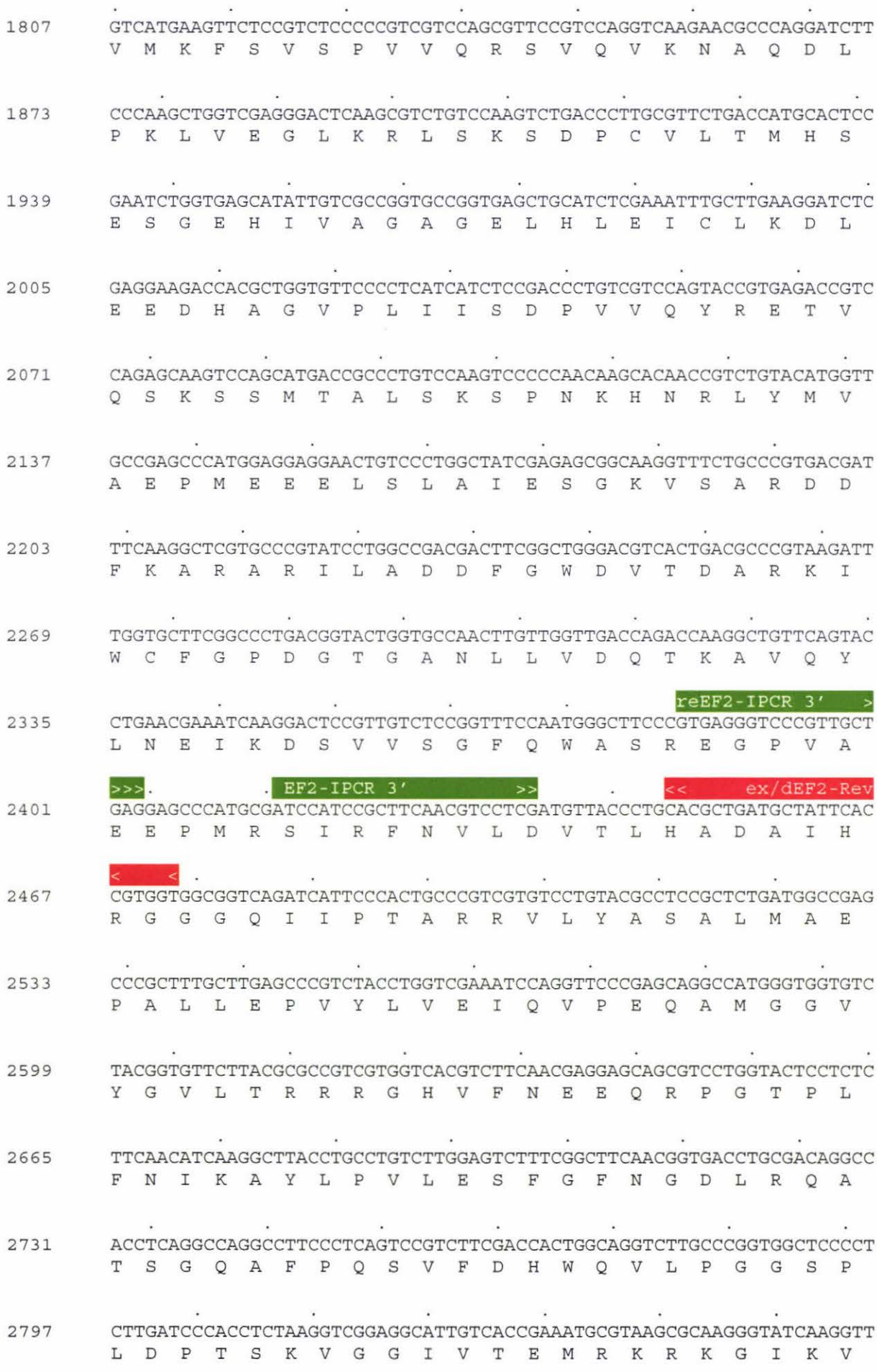
- 3' Non-coding region

The transcriptional stop site is in bold and underlined. The putative efficiency element and positioning elements are pink text, double underlined and labeled 1 and 2 by highlighted numbering below each site respectively


-1296 /CGCGTCGGCACGGCTTACAGTGGGGTGGGGTAGCTGTAGTTCCCTTCCCTCGCCGCGGCGCAAGCT
-1230 ATTTTCTTCCGGCTCAAACCAACCATAGCATCCTGCCGACTGTCTAGCTTCCACCCCAGTAGAGCC
-1164 AATGCCACAGCGCGGCCCGGCCAACCTTTTTTTTATTTTTTTTTTTCAGTCCAAAGTCGACACAGC
-1098 CGCGATTTTCTCGCTCCCAAAACTCGAGCGGGATAACTACCCACTGAATGATTTACGGAGAGCAGC
-1032 ACAGCGCACCTGCTTACGCGGAATACTTACTTTTCCCTCCTCAGCTCAATTATAAAGGTTACTGAT
<< EF2-IntR2
-966 TCCGAAACAGTTCCTTTTCAGTCGATGATGTCTTTCGGGCGATTACCTATACGGAGAGCGCGCGCC
-900 CCGACCTGCTTATTTTCAGTCCAAAACAGAGGGTTTGCGCGACAAATCACCATGTTATTTTCCGAA
-834 GAAGATGCGCCGCTTTTGAAGGCCTGGATCGTGAACGGCTGTGAAAACACGTGCGTTTGCAAACCT
Rap1p
-768 CTGATCTGGATTTTCTCTCGAGAAACAACACTAATCGACATGATCCCGATTAGGTGAATCGCAGGT
MCM1
-702 CGGATGCCGACGCGGATGTCTTGGCAGAATACGTGATTGCCTTGCTGAAACACGATGGGAATAAGG
<< EF2-IntR4
-636 AGAGCGTGAGGACCTTGTGCGAGCAGGAAATTCCGGACTTTTTGACAGAGGGTTTGTGCTGCTCCC
<<<<
-570 ATCTCTCTTGGTGCTCACAGTGCTCGTCGCTTTTTTCGCCGAAAATTTTAAGGGTTACTGATTGGC
-504 CTTAGGTCCGAAACAGTTCCTCGATGATGTCTTTCGGGCGATTACCTATCGATCGTATATGCCTGG
-438 TGCGCCGCCGCCACCGCCGAAGAAGGTTCCTGCTGGGCCCGCTGCTGCGTCTGCGAGTCCTGGACC
HSTF
-372 GTCACGGAAGCGACGAATCGATGTCACAAGGATGGGTATAACGGTGTTTCTTTTCTTCTTGTCCTT


<< uxIPCR 3' site
-306 TGGGAGGACGTTGAAGCGGATGGATTGTTGCGCGCCGACTGAGAGACTACGAATGGAATTCTCGAC
-240 TACTCACTAGCCACAGTTTTTGAACTGCACGAGCCAATCCCTCCAAATAACCATGACGTCAATGGC
ATF
-174 AACGCTCCCAGTATATAAAAGAAAAGTCATTTTGACCTTCACACTCCTCCTGCTGCCCCGCTCAGC
GCN4
-108 ACTTCGTTCCCTTCTCTCTTCCACCACCTCCCGTTGAACGGCCAAAACAACAAAGAACCATTATCT
-42 TGACGTCGAGGACGTCAAGCCAATACCGCCAAAATGGTCAAGTAAGTTTCGCAGATCTGCTCTGGT
M V N
< EF2-IntR3
24 CGCATGAAATCCGACACGAGTGCATCTCACTTTGATTGTGCTTTCTTTTCCCACTTCAATCGTTCA
90 GCGAGCTGCGGGATGGGGAGACACAAAATTTTTCCCTCTCTGGGTTGATGATGTGGTCCATTTGAC
156 CCCGCCCGGAGTTCAGCTACCCTCTCAGGGCCAAAACTCCCTGCCAAGGTTTGGGAGGATTGTTCC
222 CAATCCTGCACCAAGTTCCGTGGCGATCAGAGAGAGGCCCAAGGGGCAAAGGAAGTTGGACAAATG
288 ATGACTAACGACTTGACTGAAATTGGCAACTCTAGGAGTGTCTTCCTCGCAATGTAGCTTCACCAT
F T I
ex/dEF2-Fwd >>
354 CGATGAGATTCGCCAGCTTATGGACAAGCCTAGCAATGTCCGTAATATGTCCGTTATTGCCCACGT
D E I R Q L M D K P S N V R N M S V I A H V
420 CGATCACGGCAAGTCTACCCTGACCGACTCTCTTTTGGCCAAGGCCGGTATCATCTCCTCTGCAAA
D H G K S T L T D S L L A K A G I I S S A K
486 GGCTGGTGATGCTCGAGCAACGGATACCCGTGCAGATGAGCAGGAGCGTGGTATCACCATCAAGTC
A G D A R A T D T R A D E Q E R G I T I K S
552 GACGGCCATCTCCCTGTATGGTCGTCTTGACGACCCCGAGGACGTCAAAGACATTGTCGGCCAGAA
T A I S L Y G R L D D P E D V K D I V G Q K
618 GATTGATGGCCAGGACTTCTTGATCAACTTGATTGACTCCCCCGGTCACGTTGATTTCTCTTCTGA
I D G Q D F L I N L I D S P G H V D F S S E
EF2-IPCR 5'
684 AGTTACTGCCGCTCTCCGTGTCACTGACGGTGCTCTCGTCGTCGTCGACACCGTCGAAGGCGTGTG
V T A A L R V T D G A L V V V D T V E G V C


750
CGTCCAGACCGAGACTGTGCTCCGACAGGCTCTTGGTGAGCGTATCAAGCCCGTCATTATTATCAA
V Q T E T V L R Q A L G E R I K P V I I I N
816
CAAGGTCGACCGTGCCCTTCTCGAACTTCAGGTCTCCAAGGAGGATCTGTACCAATCTTTCTCTCG
K V D R A L L E L Q V S K E D L Y Q S F S R
882
AACCATCGAGTCCGTCAACGTCATCATCTCCACCTACCTGGACAAGGCTCTGGGTGACGTTCAGGT
T I E S V N V I I S T Y L D K A L G D V Q V
EF2-RACEc >>
948
CTACCCTGACAAGGGCACCATTGCTTTTGGTTCCGGTCTGCACGGCTGGGCCTTCACTATTCGTCA
Y P D K G T I A F G S G L H G W A F T I R Q
EF2-IntR1 >>
1014
GTTCGCTGTTCGATACGCCAAGAAGTTTGGTGTCGACAAGAACAAGATGATGGAGCGTCTCTGGGT
F A V R Y A K K F G V D K N K M M E R L W
1080
ACGTATTTGCTAACTGGCCCATTCTAGGGTGACAACTACTTCAACCCTCACACCAAGAAGTGGACC
G D N Y F N P H T K K W T
EF2 3' RACE >>
1147
AAAACGGCACGTTACGAGGGTAAGCAGCTCGAGCGTGCTTTCAATCAGTTCATCTTGGACCCCATT
K T A R Y E G K Q L E R A F N Q F I L D P I
<< EF2 5' RACE
1213
TTCAAGATCTTCAATGCCGTCATGAACTTCAAGAACGACGAGATCACCAACCTCCTCGAGAAGCTG
F K I F N A V M N F K N D E I T N L L E K L
1279
TCCCTTAAGCTTAGCGTTGATGACCGTGCCAAGGAGGGCAAGCAGCTACTCAAGGTTGTCATGCGA
S L K L S V D D R A K E G K Q L L K V V M R
1345
ACTTTCTTGCCTGCAGCCGACTCCTTGTTGGAGATGATGATCCTTCACCTTCCTTCTCCTGTCACT
T F L P A A D S L L E M M I L H L P S P V T
1411
GCTCAGACCTACCGTGTCGAGACTTTGTACGAAGGTCCCATGGACGACGAAGCCGCCATTGGTATC
A Q T Y R V E T L Y E G P M D D E A A I G I
1477
CGTGACTGCGACCCCAAGGGTCCTCTCATGTTGTACGTCTCCAAGATGGTTCCCACCCTTGACAAG
R D C D P K G P L M L Y V S K M V P T L D K
1543
GGCCGATTCTTCGCCATCGGTCGTGTCTTCTCCGGTACTGTCCGCTCTGGTCTCAAGGTCTGCATC
G R F F A I G R V F S G T V R S G L K V C I
<< EF2-IntF1
1609
AGGGGCCCCAATTACACCCCTGGCAAGAAGGAAGATCTCTTCATCAAGGCTATCCAGCGAACCGTT
R G P N Y T P G K K E D L F I K A I Q R T V
1675
CTCATGATGGGCGGCAAGGTCGAGCCCATCGATGACATGCCTGCCGGCAACATTGTGGGTCTTGTT
L M M G G K V E P I D D M P A G N I V G L V
EF2-IntRa >>
1741
GGTATCGATCAGTTCTTGCTCAAGTCTGGTACTCTCACCACCTCCGAAACTGCCCACAACCTCAAG
G I D Q F L L K S G T L T T S E T A H N L K


1807 GTCATGAAGTTCTCCGTCTCCCCCGTCGTCCAGCGTTCCGTCCAGGTCAAGAACGCCCAGGATCTT
V M K F S V S P V V Q R S V Q V K N A Q D L
1873 CCCAAGCTGGTCGAGGGACTCAAGCGTCTGTCCAAGTCTGACCCTTGCGTTCTGACCATGCACTCC
P K L V E G L K R L S K S D P C V L T M H S
1939 GAATCTGGTGAGCATATTGTCGCCGGTGCCGGTGAGCTGCATCTCGAAATTTGCTTGAAGGATCTC
E S G E H I V A G A G E L H L E I C L K D L
2005 GAGGAAGACCACGCTGGTGTTCCCCTCATCATCTCCGACCCTGTCGTCCAGTACCGTGAGACCGTC
E E D H A G V P L I I S D P V V Q Y R E T V
2071 CAGAGCAAGTCCAGCATGACCGCCCTGTCCAAGTCCCCCAACAAGCACAACCGTCTGTACATGGTT
Q S K S S M T A L S K S P N K H N R L Y M V
2137 GCCGAGCCCATGGAGGAGGAACTGTCCCTGGCTATCGAGAGCGGCAAGGTTTCTGCCCGTGACGAT
A E P M E E E L S L A I E S G K V S A R D D
2203 TTCAAGGCTCGTGCCCGTATCCTGGCCGACGACTTCGGCTGGGACGTCACTGACGCCCGTAAGATT
F K A R A R I L A D D F G W D V T D A R K I
2269 TGGTGCTTCGGCCCTGACGGTACTGGTGCCAACTTGTTGGTTGACCAGACCAAGGCTGTTCAGTAC
W C F G P D G T G A N L L V D Q T K A V Q Y
reEF2-IPCR 3'
2335 CTGAACGAAATCAAGGACTCCGTTGTCTCCGGTTTCCAATGGGCTTCCCGTGAGGGTCCCGTTGCT
L N E I K D S V V S G F Q W A S R E G P V A
>>> EF2-IPCR 3' >> << ex/dEF2-Rev
2401 GAGGAGCCCATGCGATCCATCCGCTTCAACGTCCTCGATGTTACCCTGCACGCTGATGCTATTCAC
E E P M R S I R F N V L D V T L H A D A I H
< <
2467 CGTGGTGGCGGTCAGATCATTCCCACTGCCCGTCGTGTCCTGTACGCCTCCGCTCTGATGGCCGAG
R G G G Q I I P T A R R V L Y A S A L M A E
2533 CCCGCTTTGCTTGAGCCCGTCTACCTGGTCGAAATCCAGGTTCCCGAGCAGGCCATGGGTGGTGTC
P A L L E P V Y L V E I Q V P E Q A M G G V
2599 TACGGTGTTCTTACGCGCCGTCGTGGTCACGTCTTCAACGAGGAGCAGCGTCCTGGTACTCCTCTC
Y G V L T R R R G H V F N E E Q R P G T P L
2665 TTCAACATCAAGGCTTACCTGCCTGTCTTGGAGTCTTTCGGCTTCAACGGTGACCTGCGACAGGCC
F N I K A Y L P V L E S F G F N G D L R Q A
2731 ACCTCAGGCCAGGCCTTCCCTCAGTCCGTCTTCGACCACTGGCAGGTCTTGCCCGGTGGCTCCCCT
T S G Q A F P Q S V F D H W Q V L P G G S P
2797 CTTGATCCCACCTCTAAGGTCGGAGGCATTGTCACCGAAATGCGTAAGCGCAAGGGTATCAAGGTT
L D P T S K V G G I V T E M R K R K G I K V


EF2-IntF2 >>
2863 GAAGTTCCTGGCGTTGAGAACGTAAGTTGACTGCATTTTCACTTTCCTAGCTCATCGAACATGTGC
E V P G V E N
2929 TAACTGCCTACACAGTACTATGACAAGCTGTAAATAGCTATTCCAAATGCCGATAGAATCATGTGT
Y Y D K L *
2995 AGTGCCAACAGAATGGATAAAAAAAAATCGATAGGCTGCTAGAATAAACTCTCCTCGCGCACACCC
-1-
-2-
3061 ACGAATGGGCACATCTCGTCGATAGGCGCCGTGGCAACTCCAGCGGATATCGAAAGCGTAGGTTAG
3127 TCTACGGGCAAGCCGATCGCGGTACAAACAAAGCAACCGAATGCAAGGTTGAGGCTATAATCAAGG
3193 GACACAGAGCAAGGTTCAGCTTTTGATACATGTAGTTATGGAATCAAGAACTGTGCTTTGAGTGAC
3259 TCGCGTGAATCAAGGTCTCTTAAATTCACAGACTATACGTGGTACAGCCCGAATGCAAGGTTGAAA
3325 ACTCAGCGACCTGATCGACAGCGAATGCAAGGTTGTCTCTTAAAAGACAATAGAATAGGAGGAATG
deuxIPCR 3' site >>
3391 ATCGCAGATTTACCTCATTCGTTTGCTCGACGCCCTCGCA/

3.3.2 Intron Sites

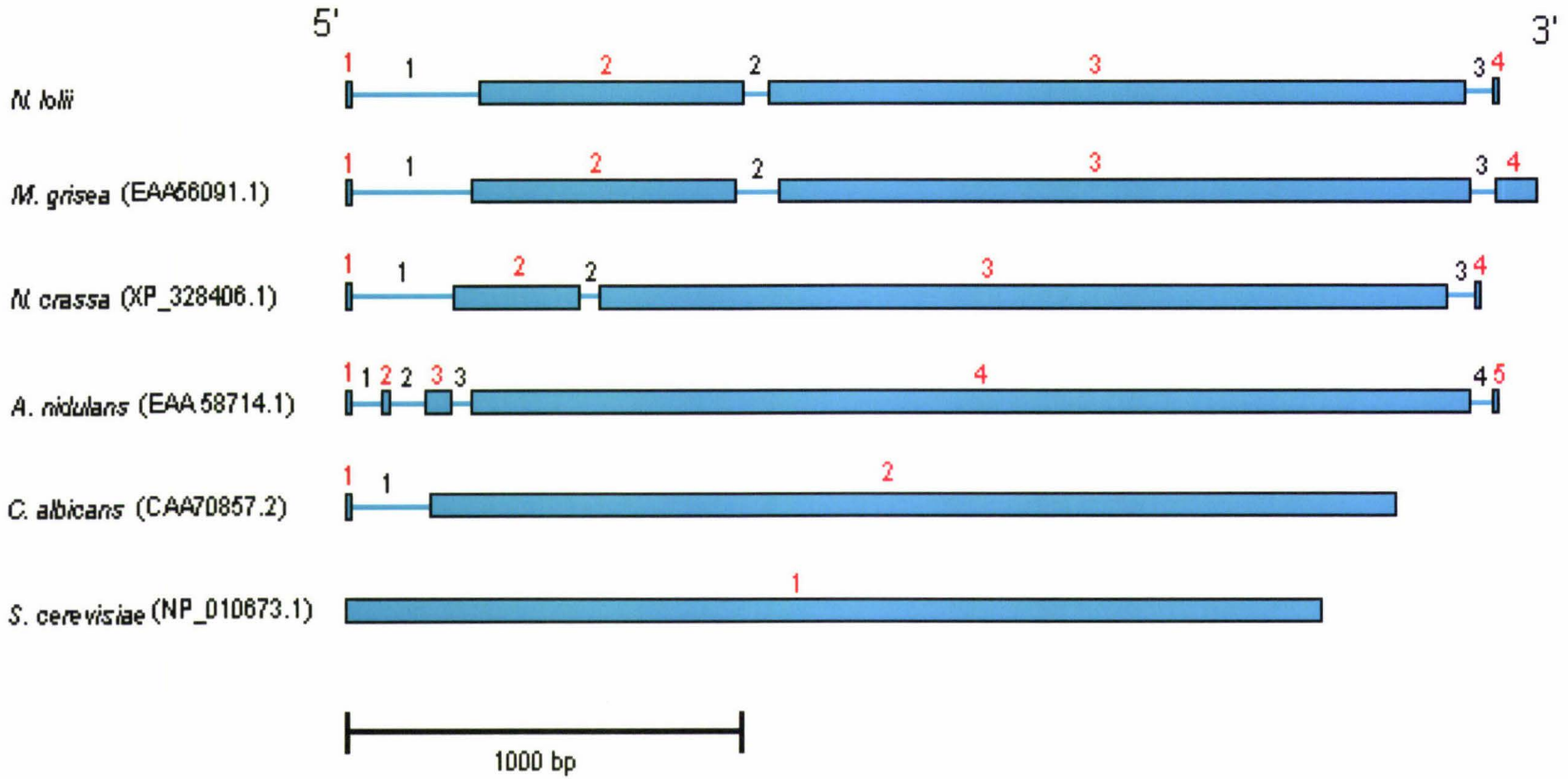
The three putative introns were confirmed by RACE. The intron sites are conserved in *EF-2* genes from more closely related organisms, such as *M. grisea* (Birren *et al.*, 2003) and *N. crassa* (Propheta *et al.*, 2001) both of which have similar length introns in approximately the same positions (Figure 3.12). In contrast the *C. albicans EF-2* gene (Mendoza *et al.*, 1999) contained only one intron at the same conserved position as intron 1 in Lp19 and both *S. cerevisiae* genes (Perentesis *et al.*, 1992) lacked introns entirely. These observations concur with current literature that filamentous fungal genes are often interrupted by introns where as introns are unusual in budding yeast (Fink, 1987; Gurr *et al.*, 1987). Gurr *et al* (1987) also suggested that intronic sequences in fungi are often less than 100 bp long and the lengths of introns 2 and 3 (29 and 62 bp respectively) comply with this statement. Intron 1 is much larger (346 bp) and has been shown in Figure 3.12 to be relatively conserved in other filamentous fungal *EF-2* genes from *N. crassa* (218 bp) and *M. grisea* (329 bp), to the yeast *C. albicans* (200 bp). In contrast this intronic region in *A. nidulans* (Birren *et al.*, 2003) is interrupted by two short exons. The similarity of *EF-2* gene structure between fungal genera is overall quite conserved, which leads to speculation about the origin of these sequences and concurs with phylogenetic relationships between these species (Spataford and Blackwell, 1993).

3.3.3 Identification Of Transcription Start/Stop Sites

Using 5' RACE results, the location of the major transcription start site was identified 122 bp upstream of the putative translation start (-122; Figure 3.11). This location is consistent with a suggestion by Gurr et al (1987) that transcription start points are between 80-400 bp upstream of translation start codon AUG (Gurr *et al.*, 1987). The sequence around the transcription start site (bold, underlined) ACA/**C**TCCTC showed similarity to several consensus sequences which included the *A. nidulans* consensus sequence predicted by Elder A/**G**CTGTTC (Elder, 1992) and general consensus KCW/**B**HYBY ((Bucher, 1990); Table 2.2). A poly-adenylation site was also identified in the EF2-con5/ EF2-con2 ClustalX alignment (Figure A3.3); 95 bp downstream of the translation stop site. The poly-adenosine tail was approximately 173 bases in length; this is typical of that found in higher eukaryotes and also comparable to mRNAs from *S. pombe* (Barker *et al.*, 1987), although in contrast to *S. cerevisiae* in which poly-A tails are 70-80 bp in length (Gallie, 2002). In fungi two poly-adenylation signals are commonly seen; the efficiency and positioning elements. The efficiency element may be found a variable distance upstream from the translation stop point and usually consist of UA or U-rich sequences (Zhao *et al.*, 1999).

Figure 3.12: Schematic comparison of *EF-2* gene structure

Exons are labeled with red numbers, and are represented by enclosed blue boxes. Introns are labeled with black numbers, and are represented by open blue lines. See table A3.1 A) for actual figures of exon and intron lengths.


5'
3'
N. lolii
M. grisea (EAA66091.1)
N. crassa (XP_328406.1)
A. nidulans (EAA58714.1)
C. albicans (CAA70857.2)
S. cerevisiae (NP_010673.1)
1000 bp

The most likely efficiency element sequence was AUAAAAAAAA; located 43 bp upstream of the polyadenylation site. The positioning element is also a UA-rich sequence located 11-30 bp upstream of the polyadenylation site, AAUAAA and AAAAAA are the most efficient (Guo and Sherman, 1995). This is required for both cleavage at the 3' end of most mRNAs and directing polyadenylation to a specific region or site (Zhao *et al.*, 1999). The closest match to a positioning element was AAUAAA, identified 18 bp upstream of the polyadenylation site.

3.3.4 Analysis of the *nlEF-2* gene promoter region

The 5' untranscribed region of *nlEF-2* was searched using a filtered string-based query on Transcription Element Search System (TESS) for promoter elements (Section 2.2.11). This was a general search to identify any feature that might further implicate *nlEF-2* function in protein biosynthesis.

In higher eukaryotes TATA and inverted Y-boxes (CCAAT) are often found 20-40 bp and 70-90 bp upstream of transcription start point respectively. Both are common and have been shown to be important for transcription initiation (Ballance, 1986). Although it is uncertain how these function, it has been suggested that TATA boxes are involved in determining the transcriptional efficiency and are usually found on the promoters of highly expressed genes. They are a target for various TATA-box binding proteins (TBPs) and are a crucial positioning component of the core promoter (Weaver, 1999). The inverted Y-box (CCAAT) does not appear to have a direct role in promoter specificity, but its inclusion increases promoter strength and influences basal-level transcription (Yoon *et al.*, 1999). Possible TATA and inverted Y-boxes were identified 31 bp and 78 bp upstream of the transcription start point respectively (-152 and -199; Figure 3.11), the location of these provides further evidence of the transcription start location and suggests that *nlEF-2* is a highly expressed gene.

Another common upstream element is the GC box, which is often present as an enhancing element in the promoters of house-keeping genes in fungi (Nakanishi *et al.*, 1988). Four GC boxes were identified 305, 777, 1033 and 1124 bp from the transcriptional start point (-427, -899, -1155 and –1246; figure 3.11).

In many other highly expressed fungal genes a pyrimidine-rich tract is upstream of the transcription start, the extent of this has been speculated to be important in determining the level of gene expression (Ballance, 1986). *nlEF-2* lacked an extended pyrimidine tract, however a short T-rich tract is seen between the TATA box and the transcriptional start (Figure 3.11).

S. cerevisiae sequences in this region have been suggested to have a greater role in determining transcriptional levels than other upstream elements (Chen and Struhl, 1985). The presence of these strong transcription elements is not unexpected as in *S. cerevisiae EF-2* is a highly expressed gene (Perentesis *et al.*, 1992).

In addition to these common motifs, five other putative elements were identified. Three of these are sites expected for a gene involved in translation. A potential GCN4 site was located at position -141 (Figure 3.11), in *S. cerevisiae* this has been shown by mutagenesis studies as an activator of genes involved in protein biosynthesis and a positive transcription factor in general control promoters (Arndt and Fink, 1986). This may be the most important element in the promoter region taking Chen and Struhl's (1985) suggestion into account and GCN4 can directly interact with RNA polymerase II. The MCM1 binding site that was discovered in *S. cerevisiae* was identified at position -714 (Figure 3.11). This is a transcriptional activator often found in genes involved in the cell cycle (Kuo and Grayhack, 1994) and therefore not unexpected in a gene whose function is closely associated with growth. As homologues of nlEF-2p are involved in protein synthesis it was not unexpected to find a potential Rap1p site at position -794 (Figure 3.11). Rap1 has been demonstated to be an activator of ribosomal genes, or repressor in response to a secretory defect (Mizuta *et al.*, 1998).

The function of the two other transcriptional elements identified is unclear. *S. cerevisiae* and Hamster *EF*-2 gene promoters contain an ATF site which is involved in the expression of cAMP inducible genes (Nakanishi *et al.*, 1988; Lin and Green, 1989; Perentesis *et al.*, 1992). This was also present in the *nlEF-2* promoter at position -166 (Figure 3.11). A HSTF site was identified at position -407; this is a potential activator in response to elevated temperatures (Wiederrecht *et al.*, 1988).

3.3.5 Analysis of sequences resulting from *nlEF-2* gene

3.3.5.1 Identification of Putative Translation Start/Stop Sites

The proposed translation start site of *nlEF-2* ORF 1 was identified by a number of factors. Firstly the BLASTp prediction in section 3.1.5.2 identified homology to the *N. crassa EF-2* gene and secondly the Transeq program was used to predict a derived amino acid sequence from EF2-con2 (Section 2.2.11). Lastly the suggestion of translation start site location was re-enforced by an upstream discontinuity of coding sequence and by the presence of specific nucleotides at the start as predicted by Kozak (Figure 3.11; (Kozak, 1981)). The sequence

around the putative start site GCCAAA**AUG**G showed similarity to several consensus sequences which included the vertebrate consensus sequence predicted by Kozak, GCCRMC**AUG**G (Kozak, 1981) and two proposed consensus sequences for *N. crassa*: CNNNCAMU**AUG**GC (Bruchez *et al.*, 1993) and CAMM**AUG**GCU ((Edelmann and Staben, 1994); Table 2.2). In addition, this sequence showed homology to two prior-characterized genes; it was an exact match to the sequence surrounding the *N. crassa EF-2* gene and a close match to the Lp19 Beta-Tubulin gene GAGAAA**AUG**C (Bryant, 2003).

The translational stop site was predicted by the Transeq program and is encoded by the Ochre codon **TAA** (Weaver, 1999) 2969 bp downstream of the translation start site (Figure 3.11). The suggestion of translation stop site location was re-enforced by a downstream discontinuity of coding sequence.

3.3.5.2 *nlEF-2* putative amino acid sequence: homology to elongation factor 2

The predicted 844 residue polypeptide (nlEF-2p) had a calculated mass of 93.44 kDa and a pI of 6.30 ((Wilkins *et al.*, 1998); Section 2.2.11). A BLASTp search using default parameters revealed that nlEF-2p is homologous to the eukaryotic EF-2 protein and suggests that EF-2 sequences are conserved through evolution. The degree of predicted amino acid sequence identity ranges from 69% relative to EF-2 polypeptides found in *Homo sapiens* (Rapp *et al.*, 1989) and *Drosophila* (Grinblat *et al.*, 1989) to over 76% in fungal species (Figure 3.13). nlEF-2p and its prokaryotic homologue, *E. coli* elongation factor G (Ovchinnikov *et al.*, 1983), are overall 34% identical to one another, however several regions in both the amino and carboxyl termini show up to 75% identity.

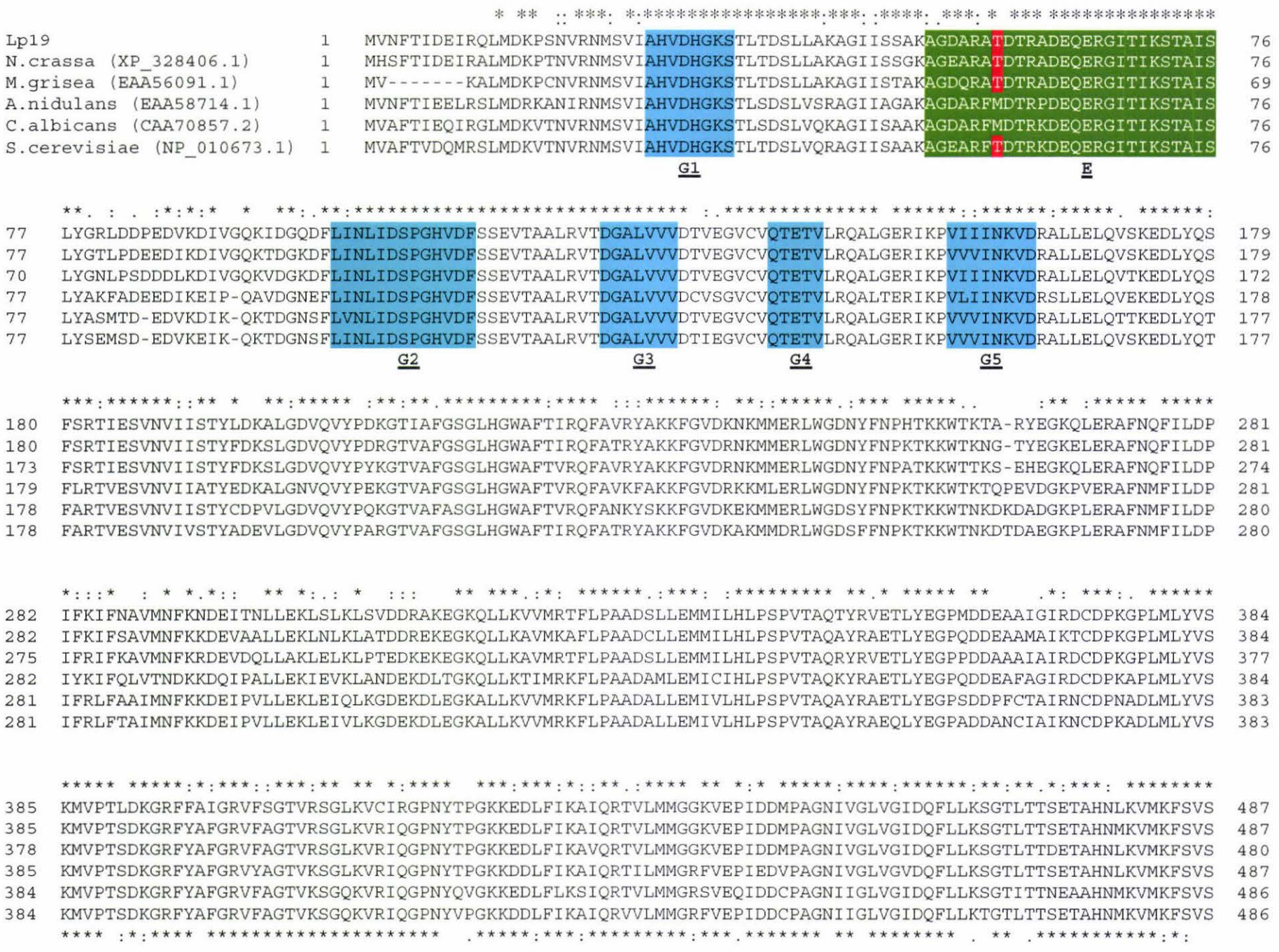
As previously discussed in Section 1.2.1 EF-2 catalyzes the translocation of peptidyl-tRNA from the A to the P site on the ribosome in a GTP-driven reaction. Like other EF-2 (Section 1.2.2), nlEF-2p contains 5 putative structural domains: residues 2-220 and 330-346 (G-domain core), 221-329 (G-domain insert), 347-482 (domain 2), 483-559 (domain 3), 560-727 and 802-844 (domain 4), 728-801 (domain 5). The positions of nlEF-2p G-domain and domains 2-5 were identified by comparison (Figure 3.13) to *S. cerevisiae* EF-2 for which a crystal structure has been solved (Jorgensen *et al.*, 2003). PROSITE was scanned using default settings; this revealed a G-domain and a putative helix-turn-helix motif (HLH; section 2.2.11).

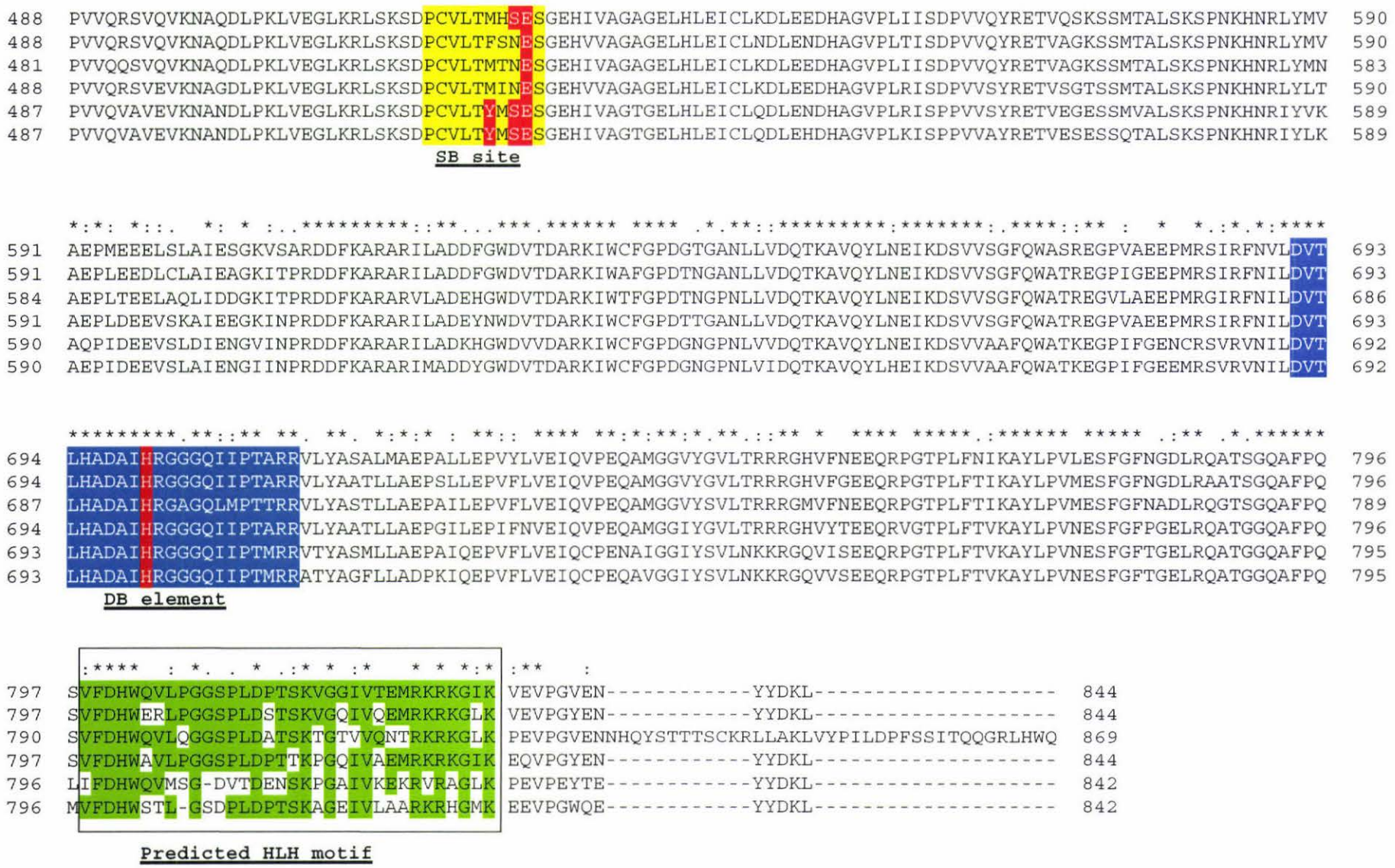
Firstly, the G-domain of nlEF-2p possesses the canonical GTP-binding motifs (G1-G5) found in all other GTP-binding proteins, as well as an effector (E) motif which is conserved among all prokaryotic and eukaryotic elongation factors (Figure 3.13; also see Section 1.2.2).

Figure 3.13: Alignment Comparing nlEF-2p and Fungal EF-2 Protein Sequences

ClustalX protein alignment of EF-2 from: Lp19 (this study), *N. crassa* (Propheta *et al* 2001), *M. grisea*, *A. nidulans* (Birren *et al.*, 2003), *C. albicans* (Mendoza *et al.* 1999) and *S. cerevisiae* (Perentesis *et al.*, 1992); each showed 90, 87, 82, 77 and 77% identity to nlEF-2p. Each of the G-domain motifs is highlighted light blue and labeled G1-G5. The elongation factor effector motif is highlighted in dark green; within this the relative location of EF-2 kinase target threonine highlighted in red (if present at this site). The Sordarin binding (SB) site is highlighted in yellow; within this the relative location of critical residues highlighted in red (if present at this site). A dipthamide binding element is highlighted in dark blue; the conserved histidine residue highlighted in red. A potential helix-loop-helix motif is boxed and residues identical to nlEF-2p highlighted in light green. Each potential site is labeled appropriately (bold, underlined).

Identical residues are indicated by a * above each column, highly and slightly similar residues by : and . respectively. Numbers annotate the locations of sites in each of the sequences.


Lp19
N.crassa (XP_328406.1)
M.grisea (EAA56091.1)
A.nidulans (EAA58714.1)
C.albicans (CAA70857.2)
S.cerevisiae (NP_010673.1)
G1
E
G2
G3
G4
G5


SB site
DB element
Predicted HLH motif

When examined by eye for the location and sequence of these GTPase domains predicted by PROSITE concurred with motifs using consensus sequences; however the effector motif tends to be highly conserved within, but not between families of GTP-binding proteins (Kohno *et al.*, 1986; Bourne *et al.*, 1991). The nlEF-2p effector motif also contains the EF-2 kinase target threonine residue (Figure 3.13). As discussed in Section 1.2.2 EF-2 kinase has been shown to inactivate EF-2 by phosphorylation of this residue, offering a possible regulatory mechanism for nlEF-2p (Ryazanov and Spirin, 1990). As this single phophorylation event inactivates EF-2, and GTP to GDP hydrolysis causes conformational changes within EF-2 to allow for translocation, it is not inconceivable that the third phosphate from hydrolysis EF-2 conformation might phosphorylate this threonine. However since EF-2 is in an active conformation the phophorylated residue may be dephosphorylated during the translocation reaction. This would explain why phophorylating this specific threonine, while EF-2 remains in an inactive conformation, would inhibit translocation as this would then be unable to undergo phophorylation by ATP.

Secondly, nlEF-2p positions 798-831 fit, with one mismatch, the PROSITE consensus for the HLH motif of homeo-box proteins (Figure 3.13). This is similar to the HLH site identified in *C. albicans* EF-2 by Mendoza *et al* (1999) who predicted this region to interact with rRNA, based on the evidence of ribosomal protein R11 binding RNA using an HLH-like motif (Xing *et al.*, 1997; Mendoza *et al.*, 1999). This observation is supported by a structure of *S. cerevisiae* EF-2 showing a helix at the 3' end of domain 5 and a loop to domain 4 C-terminal helix (Jorgensen *et al.*, 2003) and further validated as this region within *E. coli* EF-G has been shown to reach into the decoding region of the 30S subunit (Joseph, 2003), which through deletion studies caused a 1000-fold reduction in translocation (Rodnina *et al.*, 1999). Previously discussed in Section 1.2.2, *H. sapiens* and *S. cerevisiae* EF-2 proteins have been shown to have a dipthamide element. The dipthamide motif contains a post-translationally modified histidine residue which is specifically targeted by diphtheria toxin for ADP-ribosylation, inactivating EF-2 (Rapp *et al.*, 1989; Perentesis *et al.*, 1992). In this motif nlEF-2p and the other fungal EF-2 show almost exact sequence homology to *S. cerevisiae* EF-2, with the histidine conserved in all (Figure 3.13). Fungal EF-2 proteins were recently identified as the target of the novel antifungal sordarin (Section 1.2.2). The Sordarin Binding site (SB site) has been previously characterized in several *Candida* sp. and three residues were stipulated to be defining sordarin sensitivity (Shastry *et al.*, 2001). The sordarin binding site is present in nlEF-2p; serine (S524) and glutamic acid (E525) residues in the positions corresponding to sordarin sensitive *S. cerevisiae* S523 and E524; however methionine (M522) replaces tyrosine

(Y521). This raises the possibility that Lp19 may be partially sordarin sensitive, and therefore the closely related choke pathogen *E. typhina* may also be partially sordarin sensitive. This suggests that nlEF-2p may be a potential target for antifungal agents in commercial applications.

3.3.5.3 Condon Usage of *nlEF-2*

Condon usage was analysed by the Countcodon program (Section 2.2.11; Table A3.2 A)). Like many other highly expressed fungal genes, *nlEF-2* showed a strong codon bias. A preference was found in the third nucleotide position for a pyrimidine, and this was usually cytosine. Where a purine was in the third position, guanine was favored highly over adenosine. This was expected as highly expressed genes often show a codon bias, a trend which is though to be a result of highly expressed genes requiring more common isoaccepting transfer RNAs (Gurr *et al.*, 1987).

Chapter 4

nlCDC-12 CLONING

4.1 Molecular Cloning Of A Cell Cycle Division Protein 12 Gene From *Neotyphodium lolii* (Lp19)

As described in Section 1.3 the *CDC-12* gene was chosen for further investigations, because of its possible role in morphogenesis and because it appeared to possess conserved regions that would facilitate heterologous cloning. The methods used are almost identical to those in Chapter 3.

4.1.1 Degenerate Primer Design

Database searches identified 9 fungal *CDC-12* protein sequences (of these, 2 were identified as *CDC-12* homologues only by automatic annotation in genome projects and were excluded from further analysis). A number of ascomycete species sequences were selected for primer design. Amino acid and nucleotide sequences for *Aspergillus nidulans* (AAK21000.2), *Neurospora crassa* (EAA53189.1), *Candida albicans* (AAM51627.1), *Mucor circinelloides* (CAB61437.1) from the NCBI website were downloaded and aligned with default settings using ClustalX (Section 2.2.11). Within the resulting alignment, degenerate PCR primers dSep-Fwd and dSep-Rev were designed from two conserved regions (Table 2.2 and Figures 4.1 A and B). These sites were chosen firstly because they would amplify most of the gene and secondly, each site had similar base pair composition and therefore similar annealing temperature.

4.1.2 PCR amplification of a putative Lp19 *CDC-12* (*nlCDC-12*) gene fragment

PCR optimization was carried out using a range of annealing temperatures between 50 and 60°C and varying amounts of Lp19 and *A. nidulans* genomic DNA as templates (Section 2.2.9). A temperature of 52°C using 10 ng of template resulted in both a good yield of a

Figure 4.1

A): Partial Alignment of Genes used to Design *CDC-12* Degenerate PCR Primers.

Partial alignment *CDC-12* sequences of: *N. crassa* (Galagan *et al.*, 2003), *A. nidulans* (Momany *et al.*, 2001), *M. circinelloides* (Iturriaga and Eslava, 1999) and *C. albicans* (Warenda and Konopka, 2002). Each box around a sequence indicates the region the primer labeled was designed, and the stars above each column show a consensus for that base. The numbers annotate the locations of these sites in each of the sequences used and arrows show directionality. The septated line indicates there is more alignment between the two figures at the start and ends of these; however it is not shown for simplicity. The double backslash indicates that the 5'/3' (upper/lower) are ends of the same sequence.

B): Design of *CDC-12* Degenerate PCR Primers.

The two regions identified in Figure 4.1 A) were selected for primer design, and the degenerate nucleotide sequence analyzed. Primers dSep-Fwd + dSep-Rev were designed (degenerate nucleotide locations in red; Section 2.2.8 and Table 2.2). The region to be amplified in the *CDC-12* genes between the sites used to design primers varied between *N. crassa*, 494 bp and *M. circinelloides* 512 bp. Taking this into consideration along with the suggestion by Momany *et al* (2000) that *CDC-12* gene structure is conserved, the estimated PCR product size was 490-515 bp.

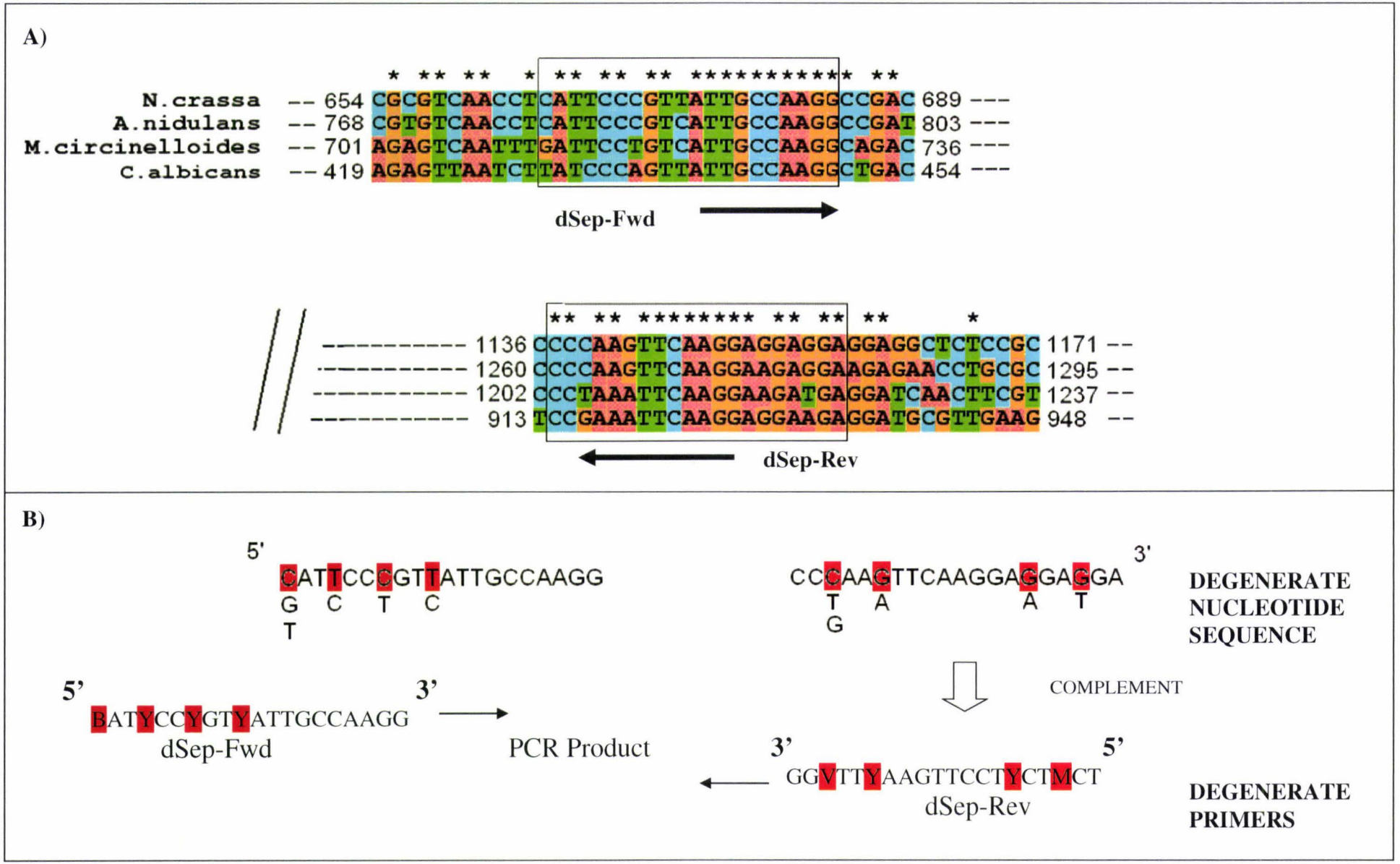
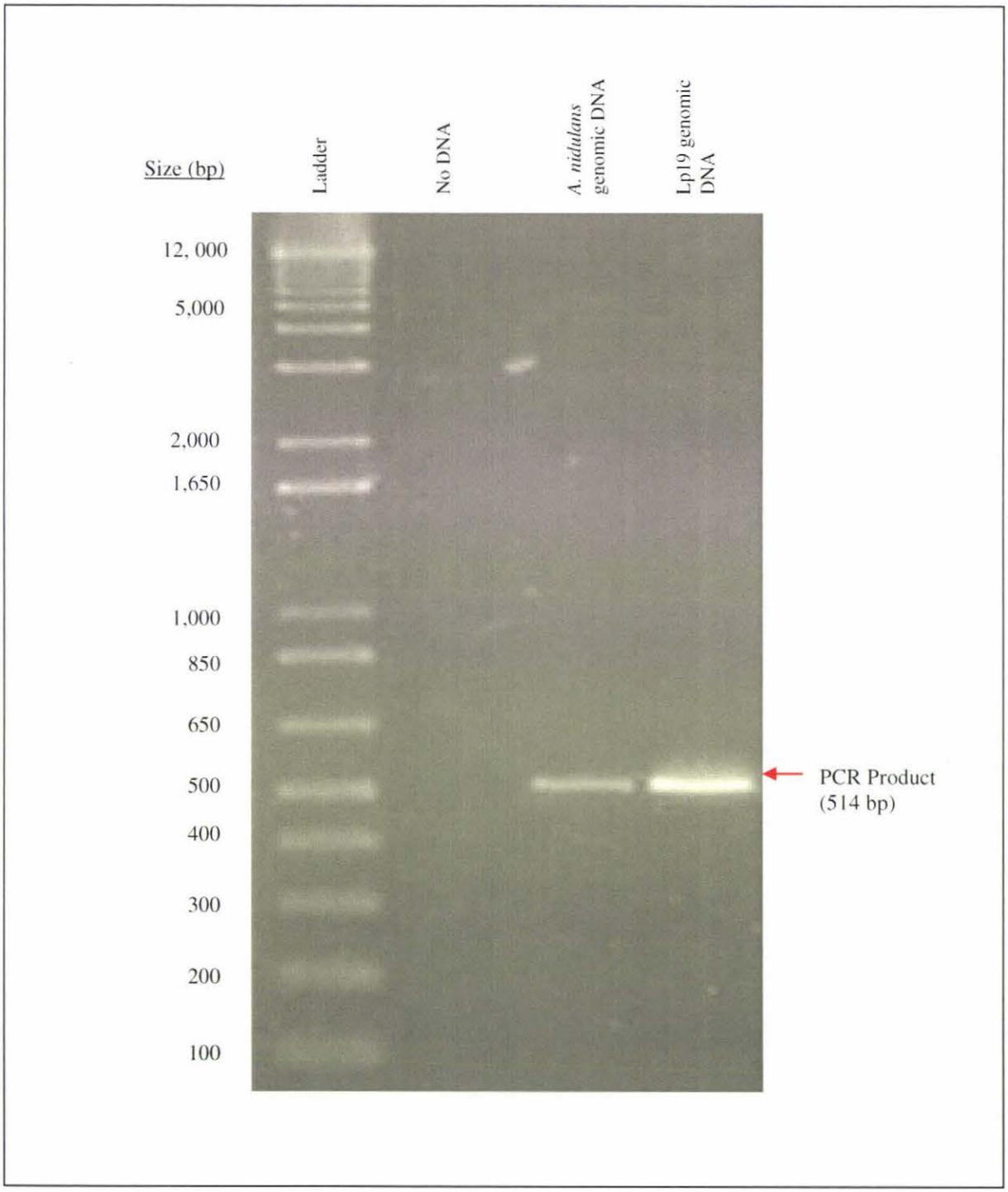

A)
N.crassa -- 654 CGCGTCAACCTCATTCCCGTTATTGCCAAGGCCGAC 689 ---
A.nidulans -- 768 CGTGTCAACCTCATTCCCGTCATTGCCAAGGCCGAT 803 ---
M.circinelloides -- 701 AGAGTCAATTTGATTCCTGTCATTGCCAAGGCAGAC 736 ---
C.albicans -- 419 AGAGTTAATCTTATCCCAGTTATTGCCAAGGCTGAC 454 ---
dSep-Fwd
1136 CCCCAAGTTCAAGGAGGAGGAGGAGGCTCTCCGC 1171 --
1260 CCCCAAGTTCAAGGAAGAGGAAGAGAACCTGCGC 1295 --
1202 CCCTAAATTCAAGGAAGATGAGGATCAACTTCGT 1237 --
913 TCCGAAATTCAAGGAGGAAGAGGATGCGTTGAAG 948 --
dSep-Rev
B)
5'
CATTCCCGTTATTGCCAAGG
G C T C
T
CCCAAGTTCAAGGAGGAGGA 3'
T A A T
G
DEGENERATE NUCLEOTIDE SEQUENCE
COMPLEMENT
5' 3'
BATYCCYGTYATTGCCAAGG
dSep-Fwd
PCR Product
3' 5'
GGVTTYAAGTTCCTYCTMCT
dSep-Rev
DEGENERATE PRIMERS

Figure 4.2: PCR amplification of putative *nlCDC-12* gene fragments

Both *A. nidulans* and Lp19 PCR reactions were performed using 10 ng genomic DNA. 20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was loaded in lane 1.


Size (bp)
Ladder
No DNA
A. nidulans genomic DNA
Lp19 genomic DNA
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
PCR Product
(514 bp)

product of the expected size (a 504 bp product was expected using *A. nidulans* DNA as a template, and a product of comparable size from Lp19 DNA) and low levels of other products for both *A. nidulans* and Lp19 templates (Figure 4.2; Section 2.2.9 using PR7 Table A2.1).

4.1.3 Cloning and sequencing of the putative *nlCDC-12* fragment

The amplified putative *nlCDC-12* fragment was purified using a Qiagen gel extraction kit and cloned into a pGEM-T Easy vector®. Sixteen putative recombinant transformants were selected. The presence of the putative *nlCDC-12* fragment in the 16 clones was tested by PCR using pUC/M13 universal primers. All 16 clones generated the expected 700 bp product and 8 clones were stored on glycerol at -80°C. Plasmid containing the putative *nlCDC-12* fragment (pDSEP-1) was extracted from three clones. Inserts were sequenced using pUC/M13 universal primers resulting in reliable sequence data for both strands of the 514 bp insert. A ClustalX alignment for each end of the sequences showed that the three inserts were identical in sequence.

4.1.4 Initial verification of the identity of the putative *nlCDC-12* fragment.

The 514 bp sequence was subjected to a BLASTn search through the NCBI website. Eighteen of the 20 most significant matches were fungal septin genes (data not shown), with the nine *CDC-12* genes being the most significant. Of these the *CDC-12* gene from *N. crassa* (EAA53189.1) was the most significant (1×10^{-61} E-value; 84 % identity). A putative intron site was identified from 65 to 131 by the signature sequences $^{5'}$GTANGT and YAG$^{3'}$ (Gurr *et al.*, 1987); this site agreed with the Genewise program intron predictions (Section 2.2.11). The Transeq program was used to predict a derived amino acid sequence from the 514 bp sequence lacking the intron; this was subjected to a BLASTp search using default settings (Section 2.2.11). Ten of the twenty most significant matches were fungal *CDC-12* protein sequences (data not shown) and as with nucleotide sequence the *CDC-12* gene from *N. crassa* (XP_323096) being the most significant (2×10^{-71} E-value; 94% identity).

Together this evidence strongly suggested that the cloned 514 bp PCR product was part of the Lp19 *CDC-12* gene.

4.1.5 Obtaining the *nlCDC-12* 514 bp flanking regions

As in Chapter 3 Inverse PCR (IPCR) to obtain the 5' and 3' flanking regions of the *nlCDC-12* gene.

As with the *nlEF-2* gene it was decided to use the Ochman *et al* method because it can amplify both flanking regions concurrently (Ochman *et al.*, 1988). If this strategy proved unsuccessful due to sub-optimal large fragments and inefficient self-religation, the Siebert *et al* method would be applied (Siebert *et al.*, 1995). However this would be more unlikely for the *nlCDC-12* gene as the core sequence is smaller (514 bp) than the core sequence for the *nlEF-2* gene (2066 bp).

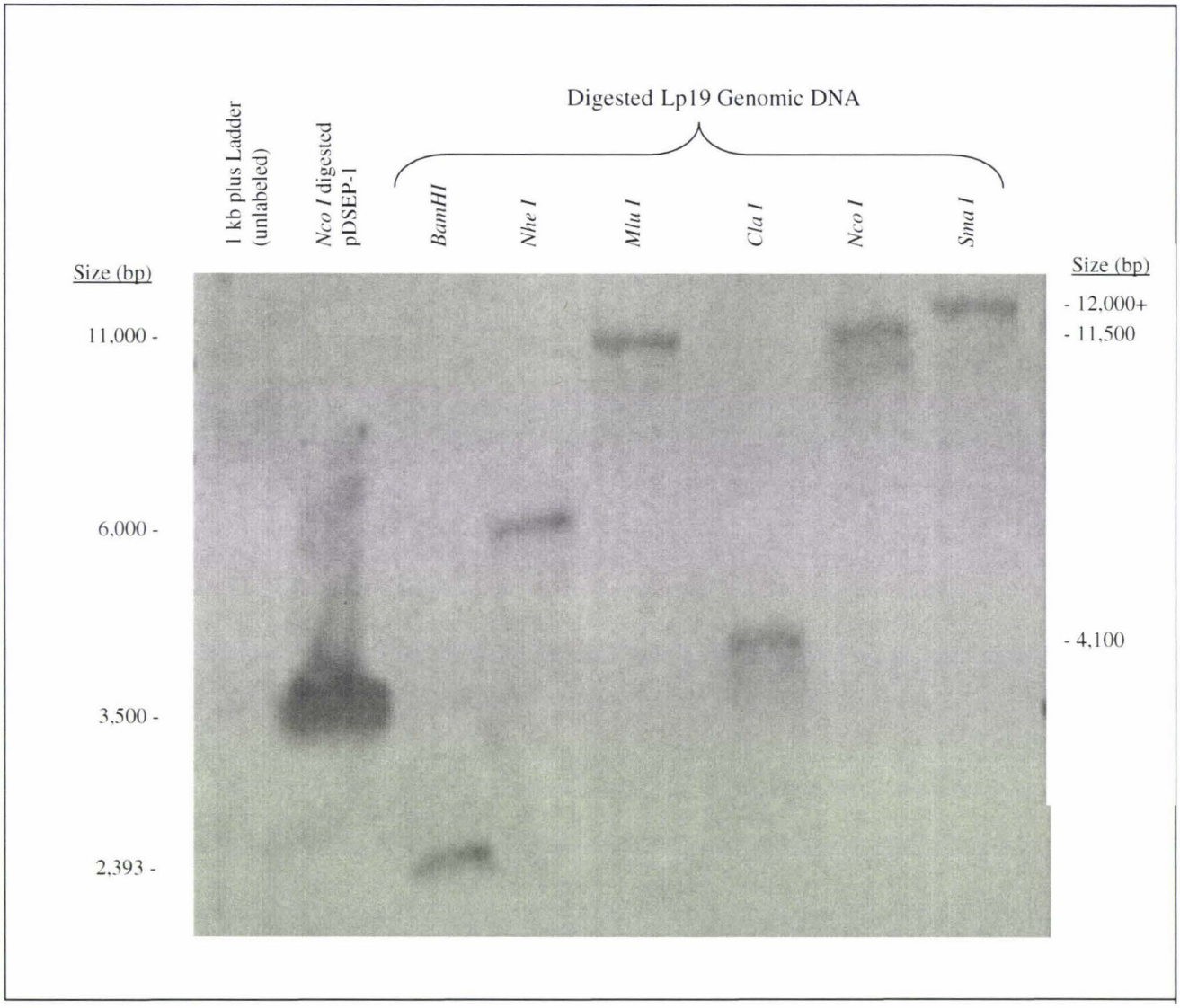
4.1.5.1 Enzyme Identification; Southern Blotting

GCG was used to generate a restriction map of the 514 bp *nlCDC-12* fragment and to identify restriction enzymes that did not cut within the fragment (Figure A3.1 B)). The *N. crassa*, *M. circinelloides* and *Aspergillus nidulans CDC-12* genes were also restriction mapped (data not shown) to find conserved restriction sites in the *CDC-12* gene outside the cloned 514 bp fragment region which, if present in Lp19 would interfere with the IPCR. To minimize the size of the amplicons but still retain significant chance of amplifying the remaining *nlCDC-12* gene six bp cutters were used (*BamHI*, *NheI*, *MluI*, *Cla I*, *Nco I* and *Sma I*).

The 514 bp fragment was generated for a probe (PrSep1) by PCR using dSep-Fwd and dSep-Rev primers with pDSEP-1 as a template (Section 2.2.9; using PR7 Table A2.1). It was verified that PrSep.1 was part of the *nlCDC-12* gene by ClustalX alignment to the pDSEP-1 sequence from which it was derived (Section 2.2.11). Labeled PrSep.1 was hybridized to a Southern blot (Figure 4.3) containing 1.0 µg digested *N. lolii* genomic DNA as well as 1.0 ng *NcoI* digested pDSep-1 (positive control). PrSep1 hybridized to single Lp19 genomic DNA fragments of approximately 2.4 kb (*BamHI*), 4.1 kb (*Cla I*), 6.0 kb (*Nhe I*), 11.0 kb (*Mlu I*), 11.5 kb (*Nco I*), and slightly in excess of 12 kb (*Sma I*). As expected the PCR product hybridized to a 3.5 kb fragment of digested pDSEP-1 in both blots. This is consistent with the location of the single NcoI site in the pGEM-T easy vector (Figure A2.1). Using these results it was deduced that *BamHI* is the optimal enzyme for use in following IPCR reactions and a 3 kb product would be expected. *Cla I* was also identified as a potential enzyme for use in IPCR.

Figure 4.3: Southern Analysis for *nlCDC-12* Inverse PCR

Blot contained 1.0 µg Lp19 genomic DNA digested separately with enzymes indicated at top of wells and were hybridized with Digoxigenin-11-dUTP (Dig) labeled PrSep1. 1.0 µg unlabeled 1kb plus ladder (Invitrogen) serves as a negative control to show that DIG labeling is the source of fluorescence. Autoradiograph was a 35 min exposure using a 100 plus Automatic X-Ray Film Processor (All-Pro Imaging). Dig high-prime was used to label PrSep1 and the concentration of labeled probe estimated using a spot test. Hybridizations were performed at 55°C and the blots were washed at 55°C (Section 2.2.13.3).


Digested Lp19 Genomic DNA
1 kb plus Ladder (unlabeled)
Nco I digested pDSEP-1
BamHI
Nhe I
Mlu I
Cla I
Nco I
Sma I
Size (bp)
11,000 -
6,000 -
3,500 -
2,393 -
Size (bp)
- 12,000+
- 11,500
- 4,100

4.1.5.2 Inverse PCR (IPCR)

Using the same principles as outlined in Chapter 3, IPCR primers, Sep-IPCR 5' and Sep-IPCR 3' (Table 2.2) were designed 57 bp and 56 bp respectively from the 5' and 3' ends of the *nlCDC-12* fragment within pDSEP-1 (the *nlCDC-12* 'core region'; Figure 4.4). *BamHI* digests of 10 µg Lp19 genomic DNA were electrophoretically separated and a region between 2.0 and 3.0 kb was excised and purified. Pooled purified fragment from several reactions were ligated at 4°C and ligation products cleaned by phenol/chloroform protocol extraction. The success of ligation was monitored by gel electrophoresis (successful ligations generated IPCR-ready DNA) and IPCR was subsequently carried out using a long-range PCR protocol (Section 2.2.14). Initial IPCR reactions with Sep-IPCR 5' and Sep-IPCR 3' were optimised using a range of annealing temperatures between 50 and 60°C, using varying amounts of IPCR-ready DNA as a template. Control templates included uncut Lp19 genomic DNA and *BamHI* digested (unligated) Lp19 genomic DNA. Single primer and water IPCR reactions were also used as controls in both sets of reactions. A temperature of 55°C using 100 ng of IPCR ready DNA template resulted in a good yield of a ~2000 bp product (Figure 4.5; Section 2.2.14 using PR8 Table A2.1). No PCR products were obtained with any of the controls.

The ~2000 bp product was gel purified and cloned into a pGEM-T easy vector and clones were stored as previously described in Section 3.1.3. These clones yielded the expected ~2.3 kb fragment, and plasmid containing the putative *nlCDC-12* IPCR fragment (pDSEP-2) was extracted from the clones (Section 2.2.3.3.1). Three pDSEP-2 inserts were partially sequenced using pUC/M13 universal primers. This resulted in 560 bp of reliable sequence from the 5' end and 670 bp of sequence from the 3' end of the fragment. Two new primers, Sep-IntF1 and Sep-IntR1 (Table 2.2) were then designed to these sequences and used to sequence the central remaining part of the fragment. Sequence yielded a 1994 bp contig. The three independent inserts were of identical sequence and as expected a solitary *BamHI* site was found. The sequence was then divided at this restriction site to yield a 612 bp 5' fragment and a 1382 bp 3' fragment and these screened for homology against the core region. It was established that 56 bp from the 5' fragment and 57 bp from the 3' fragment exactly matched the expected overlaps (as shown in Figure 4.4). Each fragment was individually subjected to a BLASTn search (Section 2.2.11). As expected the *N. crassa CDC-12* gene was most significant match for both the 5' and 3' fragments and the 5' fragment had a single alignment site from 31- 255 (226 bp) with an E-value of 7×10^{-24} and 82 % identity. The 3' fragment had 3 alignment sites; these are from positions 1287 – 1354 (68 bp; 4×10^{-21} E-value; 95 % identity), 1458 – 1621 (164 bp; 9×10^{-50} E-value; 90 % identity) and 1707 – 1957 (251 bp; 1×10^{-70} E-value; 88 %

Figure 4.4: A Theoretical construction of a Complete *nlCDC-12* gene from 514 bp PCR product and the expected Inverse PCR product.

A) A schematic of the original 514 bp degenerate PCR product. The relative location of the original PCR primers ex/dSep-Fwd and ex/dSep-Rev (black arrows), and inverse PCR primers Sep-IPCR 5' and Sep-IPCR 3' (red arrows). The locations of expected sequence overlap with the inverse PCR product are shown in bright green and light blue.

B) A schematic of the expected IPCR product showing the expected overlapping sequence regions (colour-coded to the 514 bp product A)) and the relative location of all PCR primers are shown as in A). A single *BamHI* site is expected to be located somewhere between the 5' end of ex/dSep-Fwd and the 3' end of dSep-Rev (depicted as a red line). Subsequent to sequencing and identification of the *BamHI* site, the IPCR sequence is divided at the *BamHI* site and compared to the 514 bp PCR product in C).

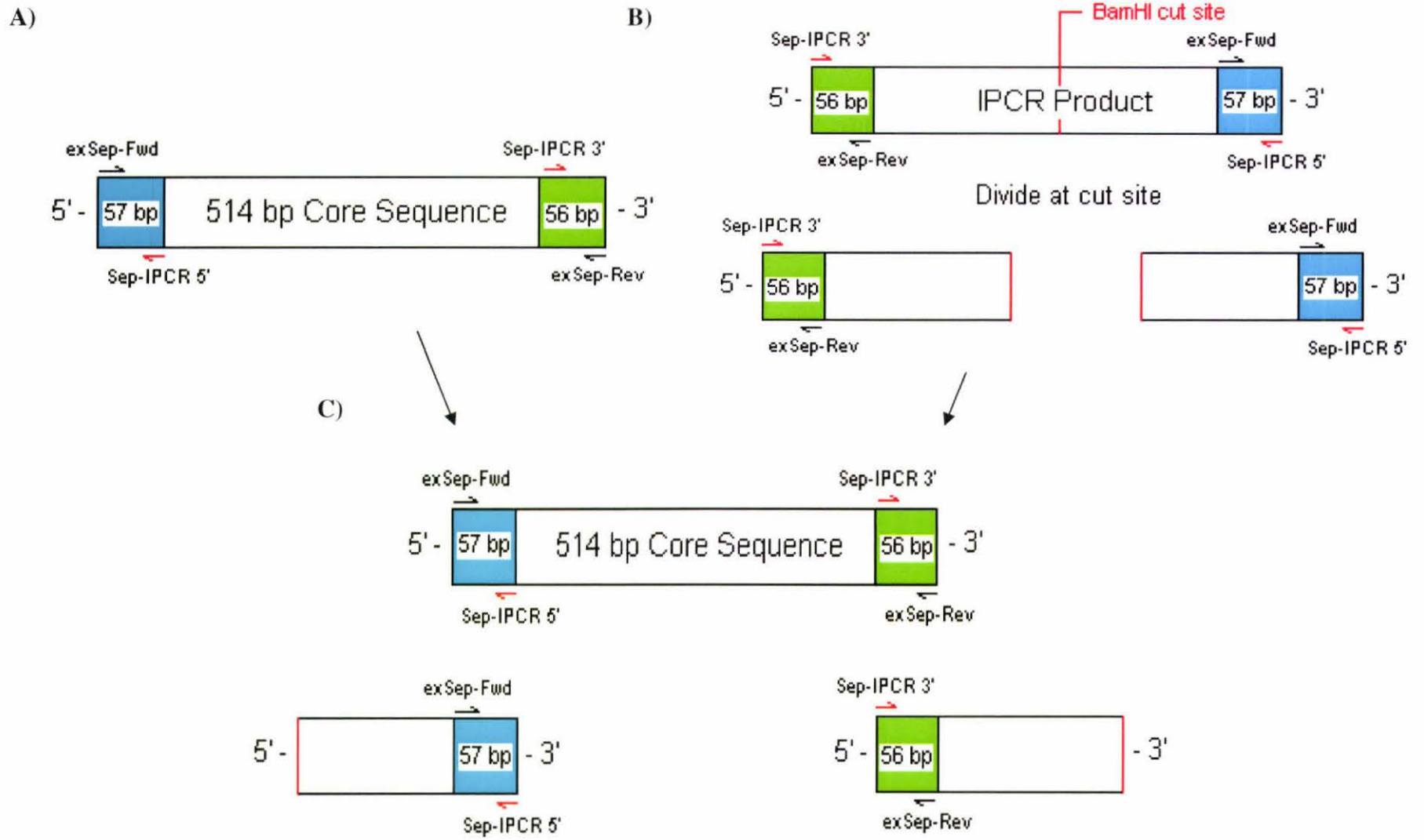
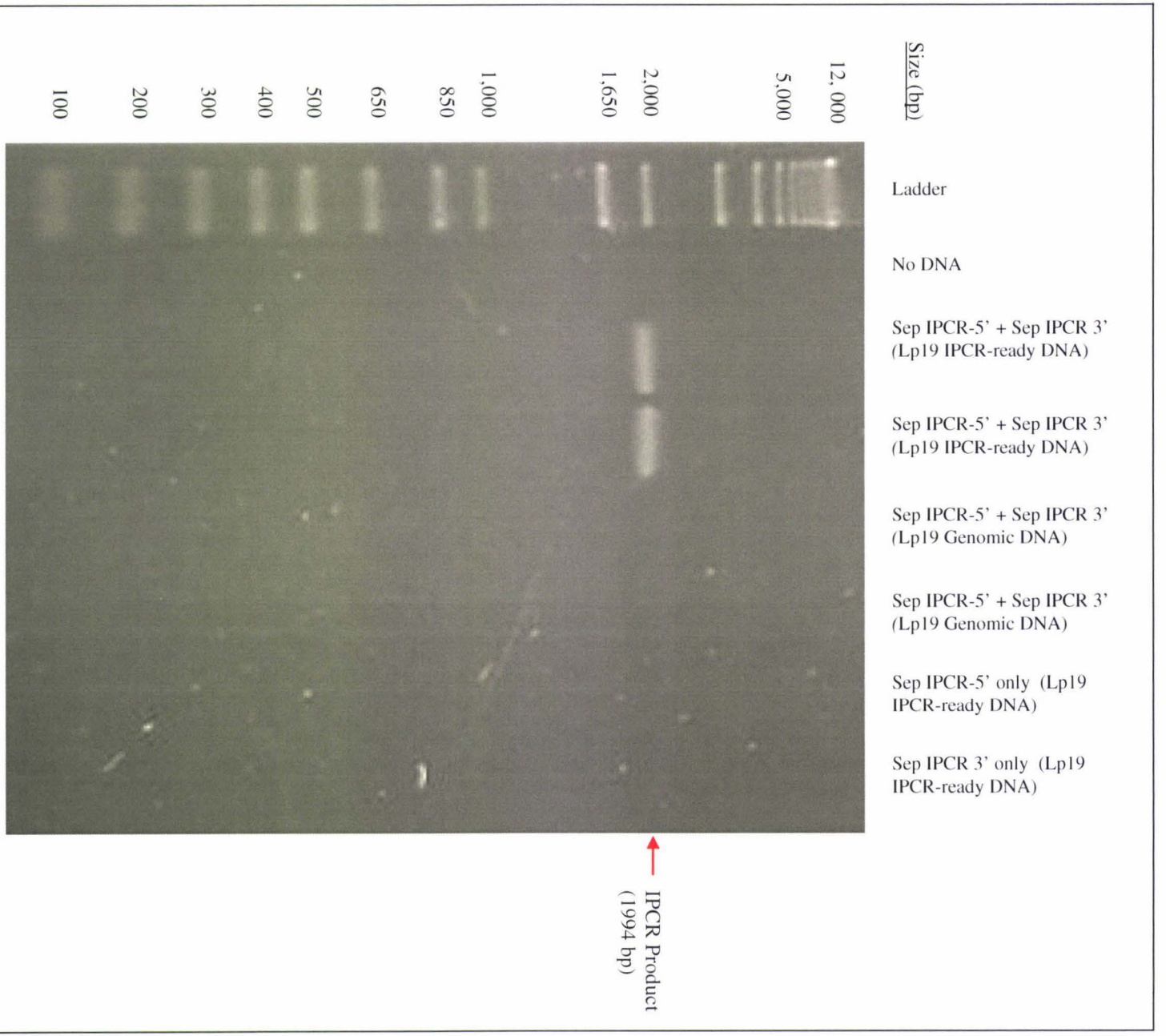

A)
exSep-Fwd
Sep-IPCR 3'
5' -
57 bp
514 bp Core Sequence
56 bp
- 3'
Sep-IPCR 5'
exSep-Rev
B)
BamHI cut site
Sep-IPCR 3'
exSep-Fwd
5' -
56 bp
IPCR Product
57 bp
- 3'
exSep-Rev
Sep-IPCR 5'
Divide at cut site
C)
Compare IPCR ends to core sequence

Figure 4.5: IPCR amplification of putative *nlCDC-12* gene fragments

Each reaction used 100 ng of template. 20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was loaded in lane 1.


Size (bp)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
Ladder
No DNA
Sep IPCR-5' + Sep IPCR 3' (Lp19 IPCR-ready DNA)
Sep IPCR-5' + Sep IPCR 3' (Lp19 IPCR-ready DNA)
Sep IPCR-5' + Sep IPCR 3' (Lp19 Genomic DNA)
Sep IPCR-5' + Sep IPCR 3' (Lp19 Genomic DNA)
Sep IPCR-5' only (Lp19 IPCR-ready DNA)
Sep IPCR 3' only (Lp19 IPCR-ready DNA)
IPCR Product (1994 bp)

identity). Each position is given relative to the total size of the *BamHI* fragment (5' and 3' fragments collectively). These results strongly suggest that the IPCR fragment contains the flanking regions of the 514 bp fragment and therefore the remaining regions of the putative *nlCDC-12* gene sequence.

A full-length contig was assembled to create a 2393 bp sequence named Sep-con1 (as in Figure 4.4). To confirm the full putative Lp19 *nlCDC-12* sequence was within Sep-con1, a BLASTn search was conducted. Between nucleotides 1 and 1377 (Figure 4.8) there were 4 significant matches, all were fungal *CDC-12* genes (data not shown) with the *CDC-12* gene from *N. crassa* (XM_323095) being the most significant (4×10^{-89} E-value; 86 % identity). No other regions in Sep-con1 had any homology to any other sequences, strongly suggesting that the full putative Lp19 *nlCDC-12* sequence was present in Sep-con1. The Transeq program was used to predict a 386 residue amino acid sequence from Sep-con1 that showed strong sequence similarity to *N. crassa CDC-12* (XP_323096). The frame shift between ORFs and presence of stop codons between indicated the presence of three introns.

4.2 Intron confirmation and Identification of Transcription start/stop sites by Rapid Amplification of cDNA Ends (RACE)

Three putative introns were located by discontinuity of ORFs. All three contain 5' donor site and 3' acceptor site sequences which show similarities to consensus sequences determined for splicing sites in *N. crassa* genes (Bruchez *et al.*, 1993); in addition each contained a universally conserved branch point sequence within 55 bp of the 3' splice site (Langford *et al.*, 1984); Table 4.1). To confirm the identification of putative introns and find the transcription start/stop sites, total RNA was isolated from actively growing Lp19 (Section 2.2.1.2.2) and analyzed by RACE (Section 3.2). Gene-specific primers were designed to ensure sequence overlap so that entire cDNA could be sequenced (see Figure 4.8; Table 2.2).

To ensure no DNA was contaminating the cDNA template a control primer Sep-RACEc was designed from positions 265-287 within Sep-con1 (Table 2.2). This primer when used with Sep 5' RACE is expected to result in 300 bp and 368 bp products when used with uncontaminated

Intron	5' Donor Site		Branch site		3' Acceptor site		Total Length (bp)
	Sequence	Start Point	Sequence	Start Point	Sequence	End point	
1	G-GTATAT	106	ACTAAC	133	TCAG-G	187	82
2	G-GTAAGC	351	CCTACC	385	CTAG-T	418	68
3	A-GTAAGT	730	GCTAAC	781	CCAG-A	795	66
N. crassa Consensus	G-GTRMGY		RCTRAC		WYAG-G		

Table 4.1; Comparison of intron consensus sequences found within the putative *nlCDC-12* gene to *N. crassa* consensus sequences. All sequences are listed 5' to 3'; left to right. Within the *N. crassa* consensus sequences M = C or A; R = A or G; W = A or T and Y = C or the 'Start/End points' are positions relative to the translation start point in figure 4.8 and the total length of each intron is indicated on the far left column in base pairs.

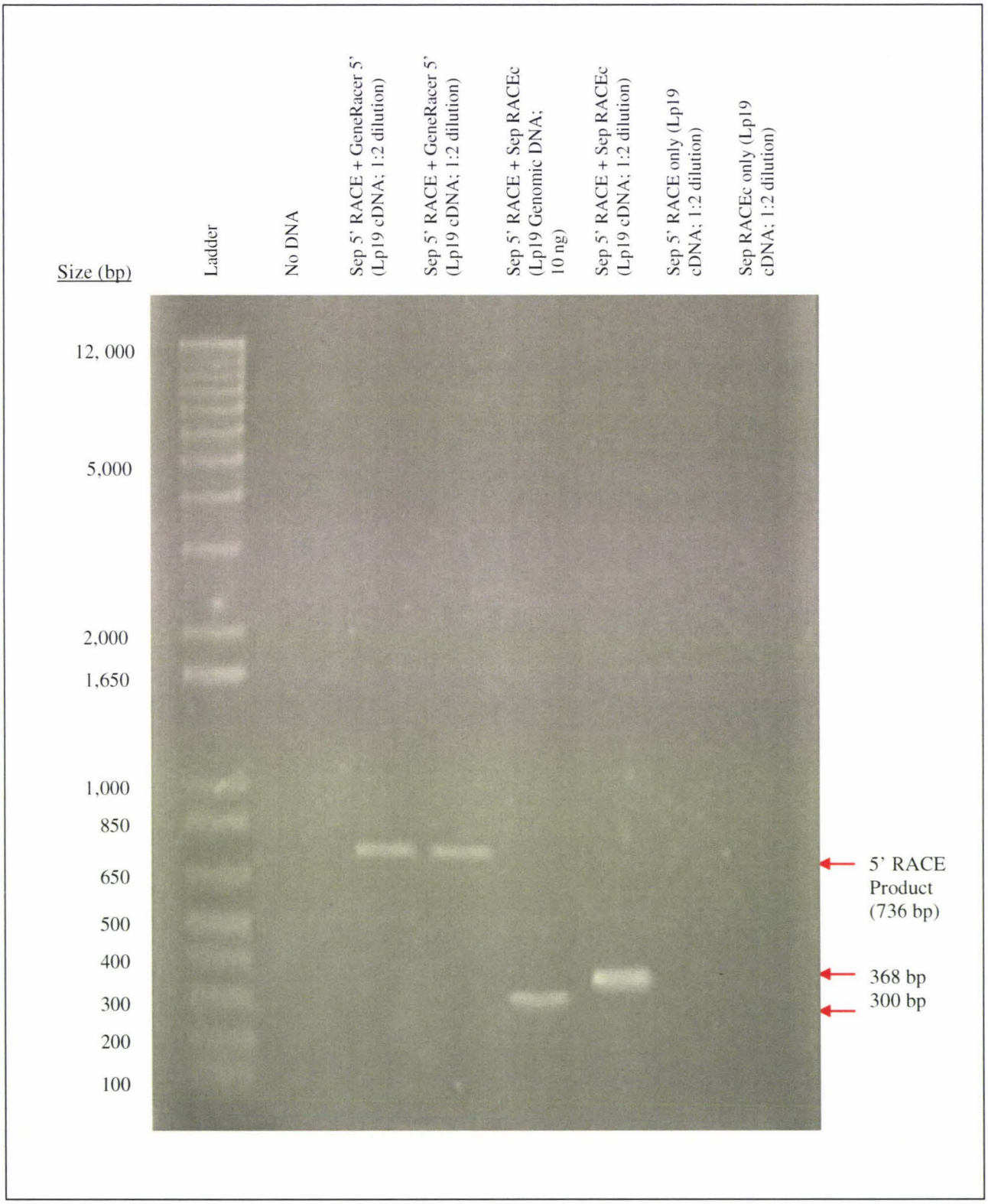
Lp19 cDNA and genomic DNA templates respectively (between these two primers is a predicted 68 bp intron). RACE-ready cDNA template was prepared by Ningxin Zhang (IMBS, Massey University, Palmerston North; see Section 3.2.1) and RACE was performed using identical protocol as in section 3.2. Primers GeneRacer 5' and Sep 5' RACE were used for 5'-RACE while GeneRacer 3' and Sep 3' RACE were used for 3'- RACE.

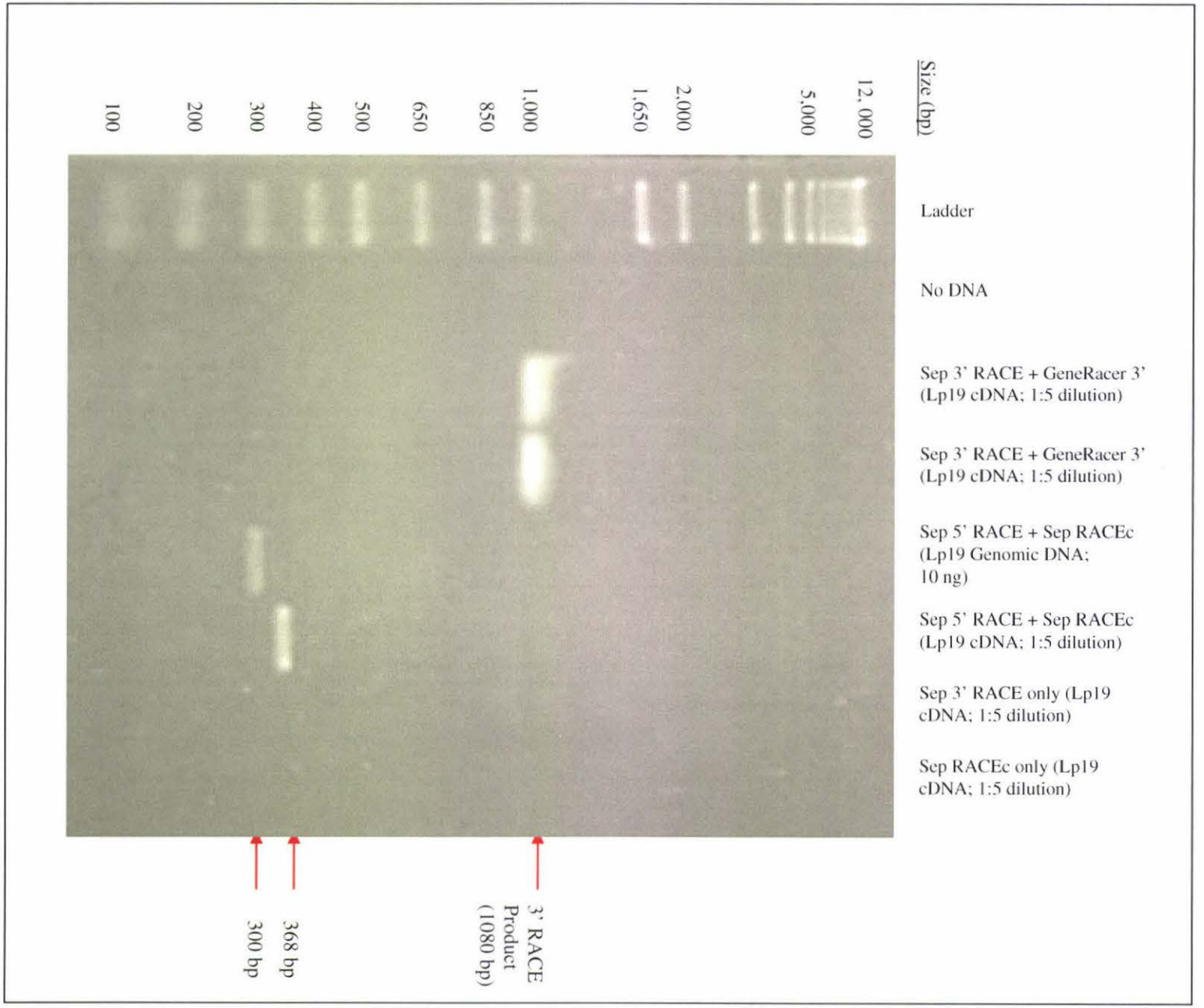
Initially the reaction was unsuccessful for 5' RACE, however after increasing the template concentration to a 1:2 dilution, a 736 bp product was obtained (Figure 4.6). For 3' RACE a 1:5 dilution of template resulted in a 1080 bp product (Figure 4.7). The 736 and 1080 bp products were gel purified and cloned into pGEM-T easy vectors. Plasmid containing each of the RACE products (pDSEP-3; 5' RACE and pDSEP-4; 3' RACE) was extracted from recombinant clones. Three independent inserts were sequenced using pUC/M13 universal primers for both pDSEP-3 and pDSEP-4 and the sequences assembled into contigs (Sep-con2 for pDSEP-3 and Sep-con3 for pDSEP-4). ClustalX alignments for each of the assembled sequences in both 5' and 3' RACE showed that the three inserts were identical. The overlapping 5' RACE and 3' RACE sequences were then assembled together into a contig (Sep-con4). As additional controls the genomic and cDNA PCR products (368 and 300 bp) shown in Figures 4.6 and 4.7 were direct sequenced and shown to match expected sections of Sep-con4 sequence (Figure A3.4).

Figures 4.6: *nlCDC-12* 5' RACE

4.7: *nlCDC-12* 3' RACE (following page)

20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was loaded in lane 1. As expected Sep-RACEc and Sep 5' RACE control PCR reactions resulted in single 300 bp and 368 bp bands with Lp19 cDNA and genomic DNA templates respectively, validating the RACE result.


Size (bp)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
Ladder
No DNA
Sep 5' RACE + GeneRacer 5' (Lp19 cDNA; 1:2 dilution)
Sep 5' RACE + GeneRacer 5' (Lp19 cDNA; 1:2 dilution)
Sep 5' RACE + Sep RACEc (Lp19 Genomic DNA; 10 ng)
Sep 5' RACE + Sep RACEc (Lp19 cDNA; 1:2 dilution)
Sep 5' RACE only (Lp19 cDNA; 1:2 dilution)
Sep RACEc only (Lp19 cDNA; 1:2 dilution)
5' RACE Product (736 bp)
368 bp
300 bp


Size (bp)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
Ladder
No DNA
Sep 3' RACE + GeneRacer 3' (Lp19 cDNA; 1:5 dilution)
Sep 3' RACE + GeneRacer 3' (Lp19 cDNA; 1:5 dilution)
Sep 5' RACE + Sep RACEc (Lp19 Genomic DNA; 10 ng)
Sep 5' RACE + Sep RACEc (Lp19 cDNA; 1:5 dilution)
Sep 3' RACE only (Lp19 cDNA; 1:5 dilution)
Sep RACEc only (Lp19 cDNA; 1:5 dilution)
3' RACE Product (1080 bp)
368 bp
300 bp

4.3 Analysis and Discussion of Sequence Results

4.3.2 *N. lolii* (Lp19) Cell Division Control Protein 12 (*nlCDC-12*)

The *nlCDC-12* gene was cloned from Lp19 by PCR, allowing sequencing and characterization of the entire coding region as well as the 5' and 3' non-translated and non-transcribed regions. The *nlCDC-12* gene of Lp19 was isolated by degenerate PCR using the primers dSep-Fwd and dSep-Rev. This resulted in a 514 bp band; whose sequence suggested it is part of a putative *CDC-12* gene. IPCR was used to obtain flanking regions resulting in a 2,393 fragment containing the full *nlCDC-12* sequence. To identify the *nlCDC-12* transcribed region RACE was performed using an overlapping set of primers for the 5' and 3' ends. After sequencing the products were assembled into a contig containing the full mRNA sequence, identifying the transcription start site, polyadenylation site and three intronic regions. This revealed that *nlCDC-12* transcripts are approximately 1,612 bp in length.

4.3.2 Intron Sites

The three putative introns were confirmed by RACE. The intron sites are conserved in *nlCDC-12* genes from more closely related organisms, such as *N. crassa* (Galagan *et al.*, 2003) which has similar length introns in approximately the same positions (Figure 4.9). In contrast both the *S. pombe* or *S. cerevisiae CDC-12* gene lacked introns (Goffeau *et al.*, 1996; Wood *et al.*, 2002). The *A. nidulans CDC-12* gene (Momany *et al.*, 2001) has a structure similar to *nlCDC-12*, with the exception of the first exon which is split into three exons in *A. nidulans*. This 5' small exon pattern is also shown in the *A. nidulans EF-2* gene. *M. circinelloides CDC-12* (Iturriaga and Eslava, 1999) gene structure has regularly spaced intronic regions which are not observed in other *CDC-12* genes. The similarity of *nlCDC-12* gene structure between fungal genera is most conserved among more closely related species, however where spacing and quantity of intronic regions differ, when present they are all of similar length. This leads to speculation about the origin of these sequences and concurs with phylogenetic relationships between these species (Spataford and Blackwell, 1993).

Figure 4.8: Full Nucleotide sequence of Sep-con1

The 2393 bp Sep-con1 sequence contains the full *nlCDC-12* gene (1-1377), a portion of the 5' non-coding region (-658-1) and 3' non-coding sequence (1378-1679). Numbering is relative to the putative translation start site. The *BamHI* cut site is indicated by a backslash, the sequence wave underlined.

- Primer locations and *nlCDC-12* Coding region features:

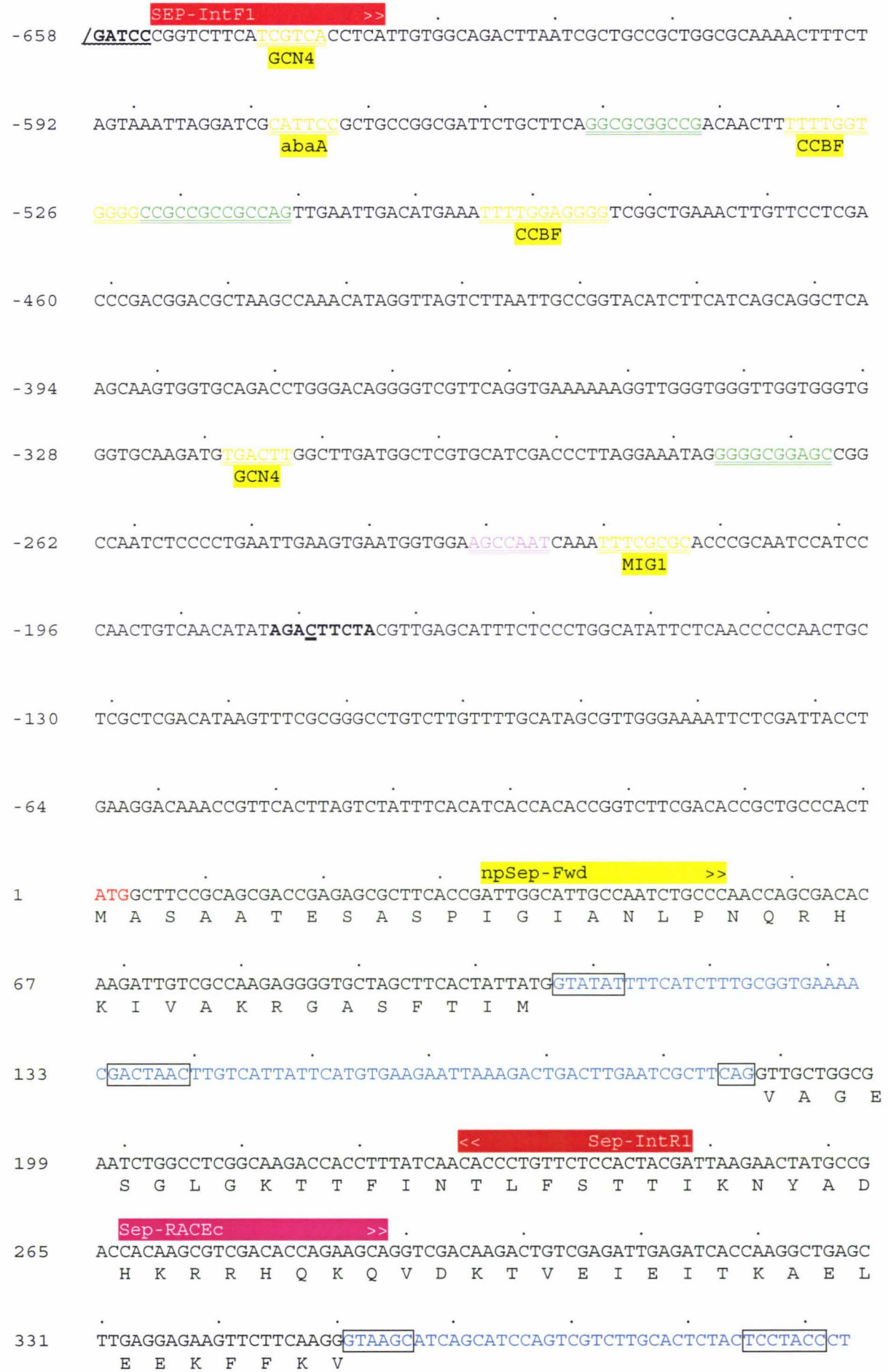
The amino acid residues encoded are positioned beneath the corresponding nucleotide sequence, with putative translational start and stop sites in red text. Blue lettering indicate the introns present, intron consensus sequences are boxed. Various classes of primer sites are indicated by highlighted labeling beneath corresponding sequence. Red primers indicate the sites for original degenerate PCR and sequencing primers. Dark green and dark blue indicate IPCR primer sites and unexpected primer sites respectively. Pink and yellow primers indicate RACE primer sites and Northern probe primers respectively.

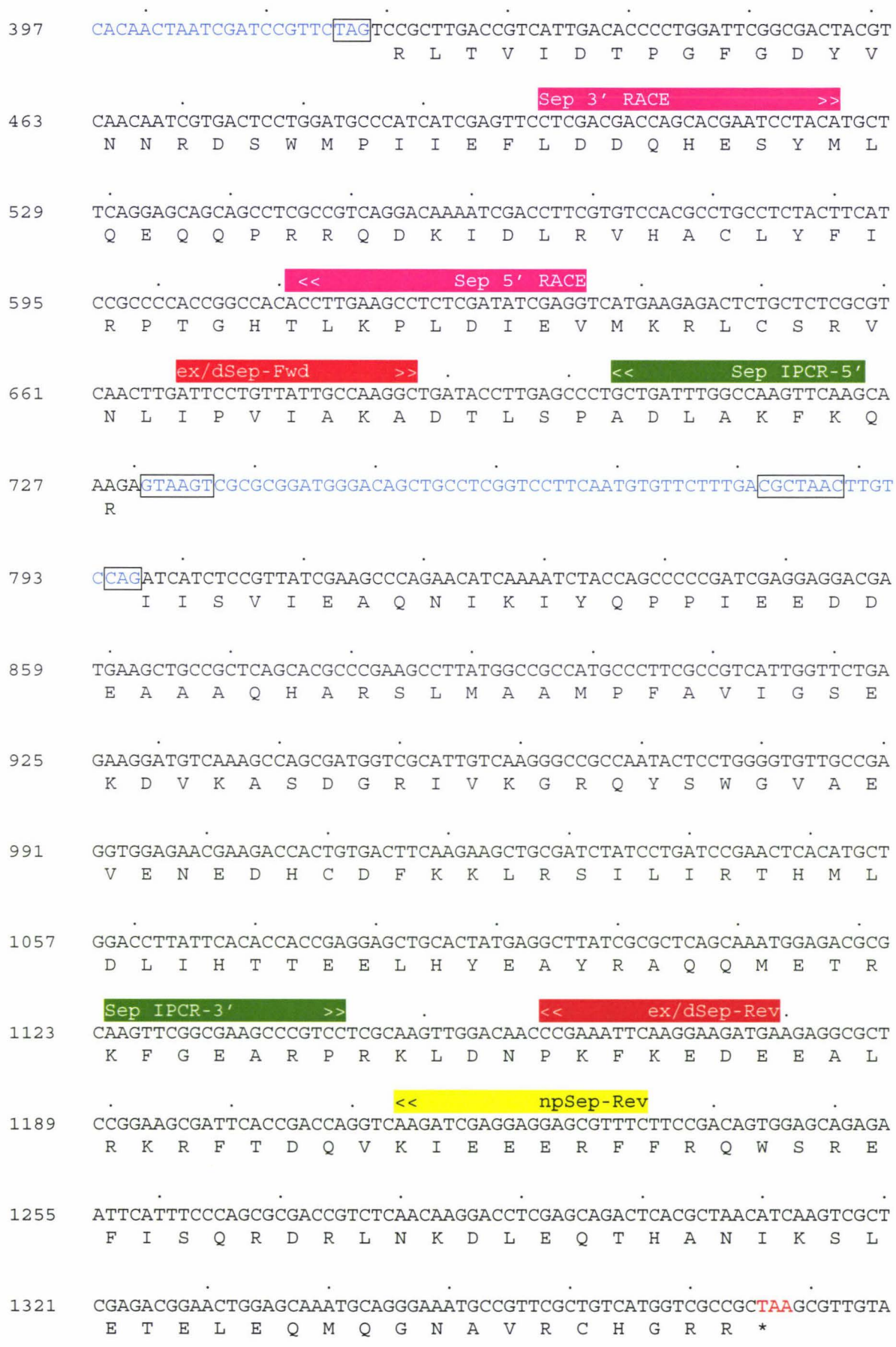
- 5' Non-Coding region

The sequence surrounding the transcriptional start site is in bold, actual site is also underlined; all other features are double underlined. Putative CCAAT and GC boxes are purple and dark green lettering respectively. Possible promoter binding sites are in yellow, the type of binding site labeled and highlighted.

- 3' Non-coding region

The transcriptional stop site is in bold and underlined. The putative efficiency element and positioning elements are pink text, double underlined and labeled 1 and 2 by highlighted numbering below each site respectively.


SEP-IntF1 >>
-658 /GATCCCGGTCTTCATCGTCACCTCATTGTGGCAGACTTAATCGCTGCCGCTGGCGCAAAACTTTCT
GCN4
-592 AGTAAATTAGGATCGCATTCCGCTGCCGGCGATTCTGCTTCAGGCGCGGCCGACAACTTTTTTGGT
abaA
CCBF
-526 GGGGCCGCCGCCGCCAGTTGAATTGACATGAAATTTTGGAGGGGTCGGCTGAAACTTGTTCCTCGA
CCBF
-460 CCCGACGGACGCTAAGCCAAACATAGGTTAGTCTTAATTGCCGGTACATCTTCATCAGCAGGCTCA
-394 AGCAAGTGGTGCAGACCTGGGACAGGGGTCGTTCAGGTGAAAAAAGGTTGGGTGGGTTGGTGGGTG
-328 GGTGCAAGATGTGACTTGGCTTGATGGCTCGTGCATCGACCCTTAGGAAATAGGGGGCGGAGCCGG
GCN4
-262 CCAATCTCCCCTGAATTGAAGTGAATGGTGGAAGCCAATCAAATTTCGCGCACCCGCAATCCATCC
MIG1
-196 CAACTGTCAACATATAGACTTCTACGTTGAGCATTTCTCCCTGGCATATTCTCAACCCCCAACTGC
-130 TCGCTCGACATAAGTTTCGCGGGCCTGTCTTGTTTTGCATAGCGTTGGGAAAATTCTCGATTACCT
-64 GAAGGACAAACCGTTCACTTAGTCTATTTCACATCACCACACCGGTCTTCGACACCGCTGCCCACT
npSep-Fwd >>
1 ATGGCTTCCGCAGCGACCGAGAGCGCTTCACCGATTGGCATTGCCAATCTGCCCAACCAGCGACAC
M A S A A T E S A S P I G I A N L P N Q R H
67 AAGATTGTCGCCAAGAGGGGTGCTAGCTTCACTATTATGGTATATTTTCATCTTTGCGGTGAAAA
K I V A K R G A S F T I M
133 CGACTAACTTGTCATTATTCATGTGAAGAATTAAAGACTGACTTGAATCGCTTCAGGTTGCTGGCG
V A G E
<< Sep-IntR1
199 AATCTGGCCTCGGCAAGACCACCTTTATCAACACCCTGTTCTCCACTACGATTAAGAACTATGCCG
S G L G K T T F I N T L F S T T I K N Y A D
Sep-RACEc >>
265 ACCACAAGCGTCGACACCAGAAGCAGGTCGACAAGACTGTCGAGATTGAGATCACCAAGGCTGAGC
H K R R H Q K Q V D K T V E I E I T K A E L
331 TTGAGGAGAAGTTCTTCAAGGGTAAGCATCAGCATCCAGTCGTCTTGCACTCTACTCCTACCCCT
E E K F F K V


397 CACAACTAATCGATCCGTTCTAGTCCGCTTGACCGTCATTGACACCCCTGGATTCGGCGACTACGT
R L T V I D T P G F G D Y V
Sep 3′ RACE >>
463 CAACAATCGTGACTCCTGGATGCCCATCATCGAGTTCCTCGACGACCAGCACGAATCCTACATGCT
N N R D S W M P I I E F L D D Q H E S Y M L
529 TCAGGAGCAGCAGCCTCGCCGTCAGGACAAAATCGACCTTCGTGTCCACGCCTGCCTCTACTTCAT
Q E Q Q P R R Q D K I D L R V H A C L Y F I
<< Sep 5′ RACE
595 CCGCCCCACCGGCCACACCTTGAAGCCTCTCGATATCGAGGTCATGAAGAGACTCTGCTCTCGCGT
R P T G H T L K P L D I E V M K R L C S R V
ex/dSep-Fwd >>
<< Sep IPCR-5′
661 CAACTTGATTCCTGTTATTGCCAAGGCTGATACCTTGAGCCCTGCTGATTTGGCCAAGTTCAAGCA
N L I P V I A K A D T L S P A D L A K F K Q
727 AAGAGTAAGTCGCGCGGATGGGACAGCTGCCTCGGTCCTTCAATGTGTTCTTTGACGCTAACTTGT
R
793 CCAGATCATCTCCGTTATCGAAGCCCAGAACATCAAAATCTACCAGCCCCCGATCGAGGAGGACGA
I I S V I E A Q N I K I Y Q P P I E E D D
859 TGAAGCTGCCGCTCAGCACGCCCGAAGCCTTATGGCCGCCATGCCCTTCGCCGTCATTGGTTCTGA
E A A A Q H A R S L M A A M P F A V I G S E
925 GAAGGATGTCAAAGCCAGCGATGGTCGCATTGTCAAGGGCCGCCAATACTCCTGGGGTGTTGCCGA
K D V K A S D G R I V K G R Q Y S W G V A E
991 GGTGGAGAACGAAGACCACTGTGACTTCAAGAAGCTGCGATCTATCCTGATCCGAACTCACATGCT
V E N E D H C D F K K L R S I L I R T H M L
1057 GGACCTTATTCACACCACCGAGGAGCTGCACTATGAGGCTTATCGCGCTCAGCAAATGGAGACGCG
D L I H T T E E L H Y E A Y R A Q Q M E T R
Sep IPCR-3′ >>
<< ex/dSep-Rev
1123 CAAGTTCGGCGAAGCCCGTCCTCGCAAGTTGGACAACCCGAAATTCAAGGAAGATGAAGAGGCGCT
K F G E A R P R K L D N P K F K E D E E A L
<< npSep-Rev
1189 CCGGAAGCGATTCACCGACCAGGTCAAGATCGAGGAGGAGCGTTTCTTCCGACAGTGGAGCAGAGA
R K R F T D Q V K I E E E R F F R Q W S R E
1255 ATTCATTTCCCAGCGCGACCGTCTCAACAAGGACCTCGAGCAGACTCACGCTAACATCAAGTCGCT
F I S Q R D R L N K D L E Q T H A N I K S L
1321 CGAGACGGAACTGGAGCAAATGCAGGGAAATGCCGTTCGCTGTCATGGTCGCCGCTAAGCGTTGTA
E T E L E Q M Q G N A V R C H G R R *

1387 TCGATTCTTTATATCAATTAAAAAATGTTTGGATCGATAGTTTTATAATATACATCTGCTAACATA

-1- -2

1453 TATACCATTTTTTTCCGATAGCTACTAGAAAAAAAGCATAGACATAAAAAATTTTTACATATAAAT

1520 ATTTTCGCGTTGTTATCGATATATTTTCAAAGCTGATCGACTTTATTCGTATATAGCGCCTCTCAG

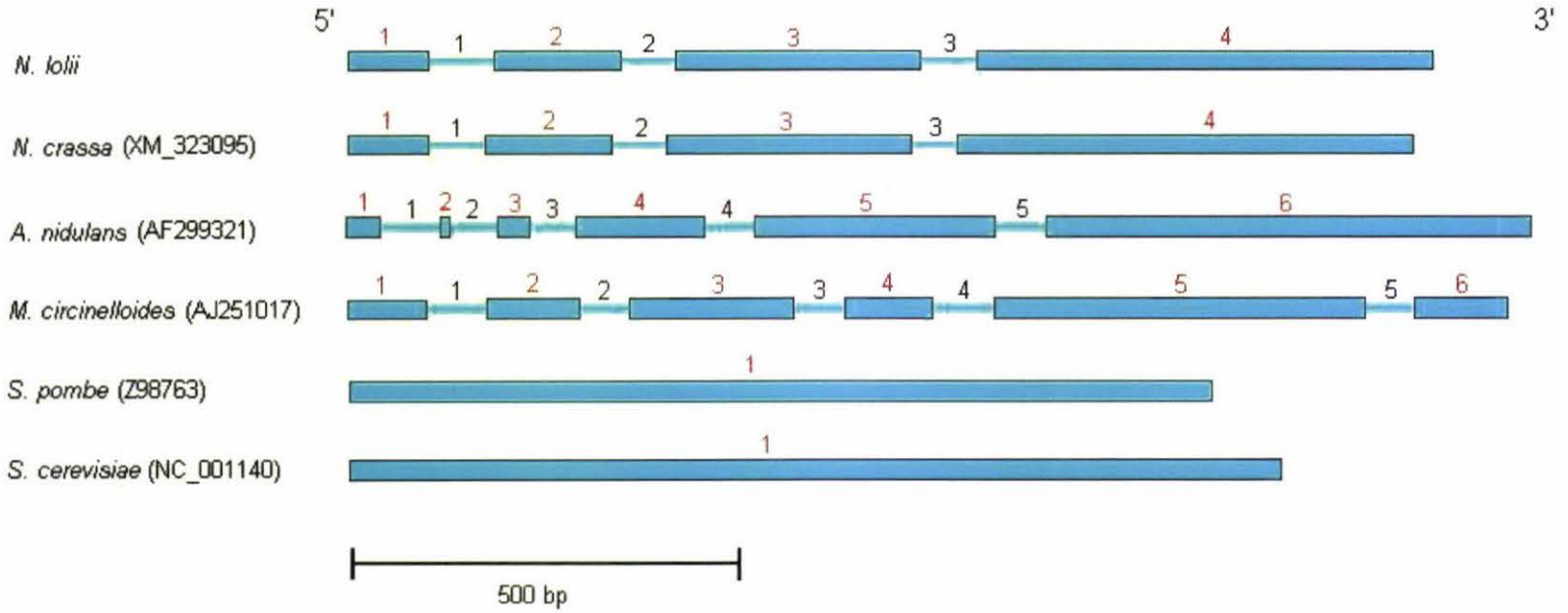
1586 TAACTGATTTTTTTGGTTCCTCATACCTTGAAATCTACTCATCAGCTTTTTGACTTTGGTTCCTTC

1652 TTCGTCGCTTTCAAATCCAGCGATAAAGACAGCTCTTATAGAATCAGCTTCGTTATCACATAATGG

1718 CTTTTACCACATTTTGGTG/

Figure 4.9: Schematic comparison of *CDC-12* gene structure

Exons are labeled with red numbers, and are represented by enclosed blue boxes. Introns are labeled with black numbers, and are represented by open blue lines. See Table A3.1 for actual figures of exon and intron lengths.


5'
3'
N. lolii
N. crassa (XM_323095)
A. nidulans (AF299321)
M. circinelloides (AJ251017)
S. pombe (Z98763)
S. cerevisiae (NC_001140)
500 bp

4.3.3 Identification Of Transcription Start/Stop Sites

Using 5' RACE results the location of the major transcription start site was identified 179 bp upstream of the putative translation start (Figure 4.8). The sequence around the transcription start site (bold, underlined) AGA/**C**TTCTA showed similarity to several consensus sequences which included the *A. nidulans* consensus sequence A/**G**CTGTTC (Elder, 1992) and general consensus KCW/**B**HYBY ((Bucher, 1990); Table 2.2). A poly-adenylation site was also identified 71 bp downstream of the translation stop site (Figure 4.8). The Poly-A tail was approximately 197 bases in length; this is typical of that found in higher eukaryotes and also comparable to mRNAs from *S. pombe* (Barker *et al.*, 1987) and *nlEF-2* (this study). As mentioned previously (Section 3.3.3), in fungi two poly-adenylation signals are commonly seen; the efficiency and positioning elements. The most likely efficiency element (Zhao *et al.*, 1999) sequence was AAUUAAAAAA; located 47 bases upstream of the polyadenylation site. The closest match to a positioning element (Guo and Sherman, 1995) was AAUAUA, identified 16 bp upstream of the polyadenylation site.

4.3.4 Analysis of the *nlCDC-12 gene* promoter region

The 5' untranscribed region of *nlCDC-12* was searched in the same way as that described for the *nlEF-2* gene in Section 3.3.4.

Using a filtered string-based query on Transcription Element Search System (TESS) for promoter elements (Section 2.2.11). No TATA boxes were identified, however an inverted Y-box was identified 53 bp upstream of the transcription start point (-231; Figure 4.8). Three G/C boxes were identified 98, 344 and 372 bp from the transcriptional start point (-276, -522 and–550; Figure 4.8). These findings correlate with reports that G/C boxes are required for the efficient transcription of TATA-less promoters (Smale, 1997).

In addition to these common motifs, four other putative elements were identified. Three of these sites are expected for a gene involved in the cell cycle and development. A potential E2F site was located at position -220 (Figure 4.8); in *S. pombe* this has been shown by reporter-gene studies to activate several promoters of genes whose products are implicated in the control of cell proliferation (Malhotra *et al.*, 1993). Two potential CCBF binding sites were identified at positions -493 and -533 in the *nCDC-12* promoter. In *S. cerevisiae* foot-printing experiments showed this factor was required for cell-cycle dependent transcriptional activation of the HO gene (Andrews and Herskowitz, 1989). An abaA binding site was identified at

position –577 (Figure 4.8); in *A. nidulans* this has been demonstrated using mobility-shift assays to be a transcriptional activator that acts as a genetic switch in developmental control (Andrianopoulos and Timberlake, 1994).

The function of the other transcriptional element identified is unclear. Two potential GCN4 sites were located at positions -318 and -644 (Figure 4.8). As discussed in Section 3.3.4 this factor has been shown to be involved in protein biosynthesis and a positive transcription factor in general control promoters, its function if any may be the latter as CDC-12 proteins have not been implicated to be involved in protein biosynthesis.

4.3.5 Analysis of sequences resulting from the *nlCDC-12* gene

4.3.5.1 Identification of Putative Translation Start/Stop Sites

The proposed translation start site of *nlCDC-12* ORF 1 was identified by a number of factors. Firstly the BLASTp prediction in Section 4.1.5.2 identified homology to the *N. crassa CDC-12* gene and secondly the Transeq program was used to predict a derived amino acid sequence from Sep-con1 (Section 2.2.11). Lastly the suggestion of translation start site location was re-enforced by an upstream discontinuity of coding sequence and by the presence of specific nucleotides at the start as predicted by Kozak (Figure 4.8). The sequence around the putative start site CCCACT**AUG**G showed similarity to several consensus sequences as discussed for *nlEF-2* (Kozak, 1981; Bruchez *et al.*, 1993; Edelmann and Staben, 1994). The translational stop site is 1161 bp downstream of the translation start site (Figure 4.8). The stop site location was re-enforced by a downstream discontinuity of coding sequence.

4.3.5.2 *nlCDC-12* putative amino acid sequence: homology to CDC-12

The predicted 386 residue polypeptide (nlCDC-12p) had a calculated mass of 44.72 kDa and a pI of 8.49 ((Wilkins *et al.*, 1998); section 2.2.11). A BLASTp search using default parameters revealed that nlCDC-12p is homologous to the eukaryotic CDC-12 proteins (Section 2.2.11). The degree of predicted amino acid sequence identity ranges from 93% relative to CDC-12 polypeptides found in *N. crassa* (Galagan *et al.*, 2003) to 56% in yeast species (Figure 4.10).

As previously discussed in section 1.3.1 CDC-12p is a G-protein critical for cell division and whose ability to hydrolyze GTP is stipulated to be involved in the formation of the 10 nm filament during septation (Kinoshita *et al.*, 1997). The N-terminal portion of CDC-12 contains

several conserved motifs typical of G-proteins which contain the GTPase activity, while the C-terminus contains putative domains that interact with other septins (Momany *et al.*, 2001).

A PROSITE scan (Gattiker *et al.*, 2002) showed that nlCDC-12p possesses the canonical GTP-binding motifs (G1, 2, 3, and 5) found in all other GTP-binding proteins, (Figure 4.10; also see Section 1.3.1). The location and sequence of these GTPase domains concurred with those of motifs using consensus sequences (Bourne *et al.*, 1991). The GTPase domain has been shown to be involved in proper septin polymerization and localization in mice (Kinoshita *et al.*, 1997), it may also have other functions involved with signaling (Section 1.3.1).

A coiled-coil motif was predicted using a (using Coils Program; probability of 0.769; section 2.2.11) between residues 348-385.

However relatively little homology is present towards the C-terminus of this domain considering its presence among all CDC-12 homologues. Coiled-coil domains have been previously characterized to be protein-protein interaction domains, therefore providing a possible mechanism for nlCDC-12p to interact with other septins to form the 10 nm ring (Momany *et al.*, 2001).

4.3.5.3 Codon Usage of *nlCDC-12*

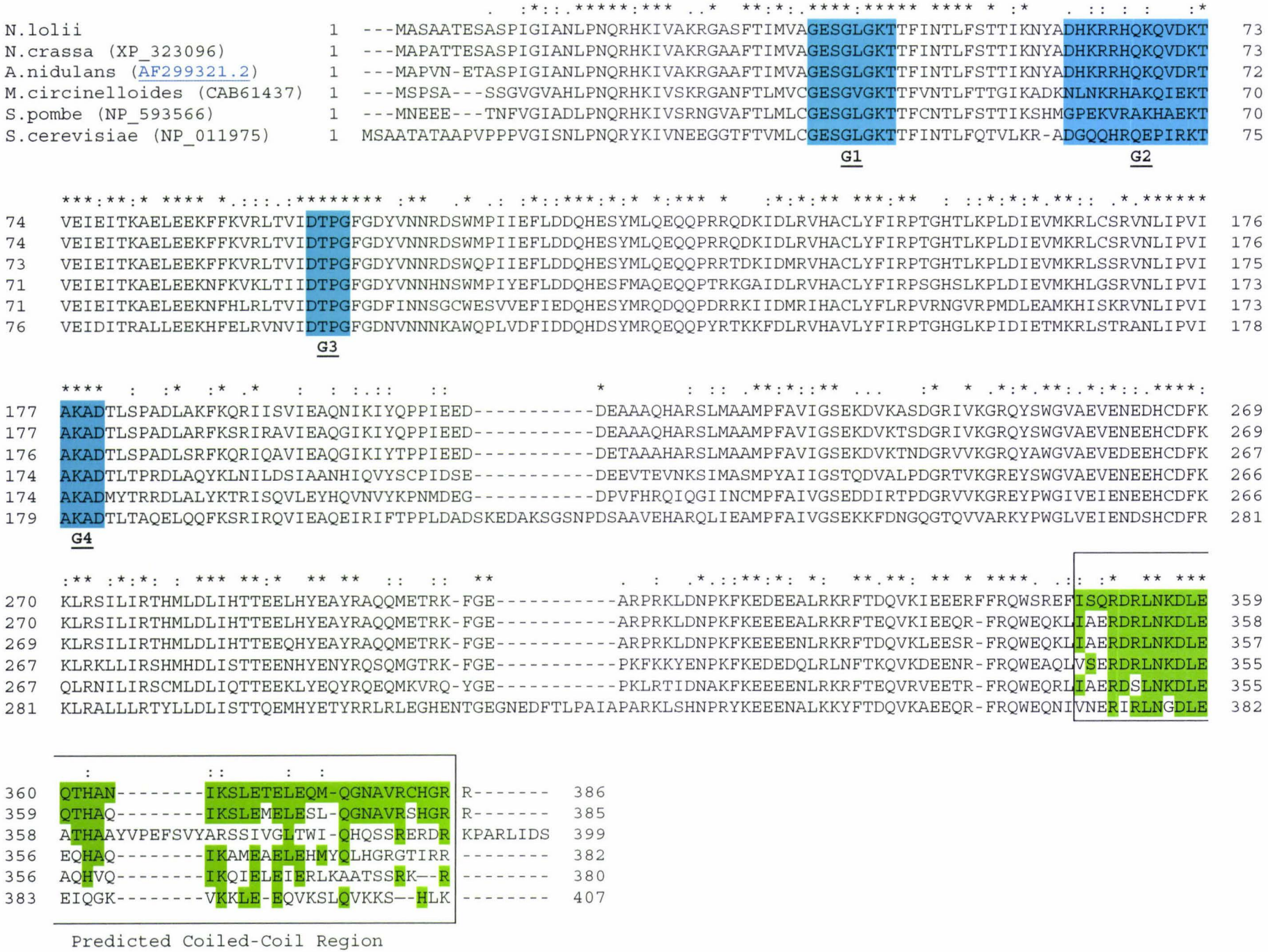
Codon usage was analyzed by the Countcodon program (Section 2.2.11; Table A3.2 A)). Surprisingly like *nlEF-2, nlCDC-12* showed a strong codon bias. A preference was found in the third nucleotide position for a pyrimidine, and this was usually cytosine. Where a purine was in the third position, guanine was favored highly over adenosine. This was unexpected as *nCDC-12* has not been shown in other organisms to be highly expressed (Momany *et al.*, 2001; Warenda and Konopka, 2002) and codon bias is thought to be more pronounced in highly expressed genes. However, although in general *CDC-12* is not highly expressed, high transient levels of expression have been identified in *A. nidulans* (Momany *et al.*, 2001). This suggests that although *CDC-12* is not a constitutively highly expressed gene such as *EF-2*; that during the appropriate periods of the cell cycle *CDC-12* expression may occur at high levels. Another possible explanation is discussed in Section 7.1.

Figure 4.10: Alignment Comparing nlCDC-12p and Fungal CDC-12 Protein Sequences

ClustalX protein alignment of CDC-12 from: Lp19 (this study), *N. crassa* (Galagan *et al.*, 2003), *A. nidulans* (Momany *et al.*, 2001), *M. circinelloides* (Iturriaga and Eslava, 1999), and *S. cerevisiae* (Goffeau *et al.*, 1996); each showed 93, 87, 65, 58 and 56% identity to nlCDC-12p.

Each of the G-domain motifs are highlighted light blue and labeled G1-G5. A potential coiled coil motif is boxed and residues identical to nlCDC-12p highlighted in light green. Each potential site is labeled appropriately (bold, underlined).

Identical residues are indicated by a * above each column, highly and slightly similar residues by : and . respectively. Numbers annotate the locations of sites in each of the sequences.


N.lolii
N.crassa (XP_323096)
A.nidulans (AF299321.2)
M.circinelloides (CAB61437)
S.pombe (NP_593566)
S.cerevisiae (NP_011975)
G1
G2
G3
G4
Predicted Coiled-Coil Region

Chapter 5

SACC CLONING

5.1 Attempted Cloning Of A Stretch-Activated Calcium Channel Gene From *Neotyphodium lolii* (Lp19)

As described in Section 1.4 the *SACC* gene was chosen for further investigation because of its possible role in morphogenesis, although it contained few conserved regions that would facilitate heterologous cloning. The methods used are almost identical to those in Chapter 3.

5.1.1 Degenerate Primer Design

Database searches identified 7 fungal *SACC* protein sequences (of these, 2 were identified as *SACC* homologues only by automatic annotation in genome projects and were excluded from further analysis). A number of ascomycete species sequences were selected for primer design. Initially amino acid and nucleotide sequences for *Aspergillus nidulans* (AAM47511), *Neurospora crassa* (CAB91331), *Schizosaccharomyces pombe* (JC7212), *Candida albicans* (EAK91710) and *Saccharomyces cerevisiae* (BAA06895) from the NCBI website were aligned using ClustalX (Section 2.2.11). Due to lack of sequence identity within the nucleotide alignments sites for degenerate PCR primer design were identified within the amino-acid sequence alignment. For this only *N. crassa, A. nidulans and S. pombe* proteins were used, the other two yeast proteins being more evolutionarily distant (Gurr *et al.*, 1987). The degenerate primers dSACC-Fwd and dSACC-Rev were designed from the most conserved regions in the alignment (Figure 5.1; Table 2.2).

5.1.2 Attempted PCR amplification of a putative Lp19 *SACC* gene fragment

A set of initial PCR reactions were performed using a range of annealing temperatures between 50 and 60°C and varying amounts of Lp19 and *A. nidulans* genomic DNA as templates in a series of separate reactions (Section 2.2.9). The aim was to find conditions at which products

of the expected size (approximately 156 bp) were formed by PCR. A range of products were amplified by PCR between 50-55°C but very little at higher temperatures. There were no products of the expected size using Lp19 DNA; however a product of 156 bp was amplified using *A. nidulans* DNA. A range of differing conditions were applied; 5-100 ng DNA template (increments of 5 ng), 0.0- 5.0 mM magnesium concentrations (increments of 0.5 mM), and 1.0-5.5 % formamide which has been demonstrated to increase specificity of priming (Sarkar *et al.*, 1990). No products of the expected size were obtained with Lp19 template for any of these conditions.

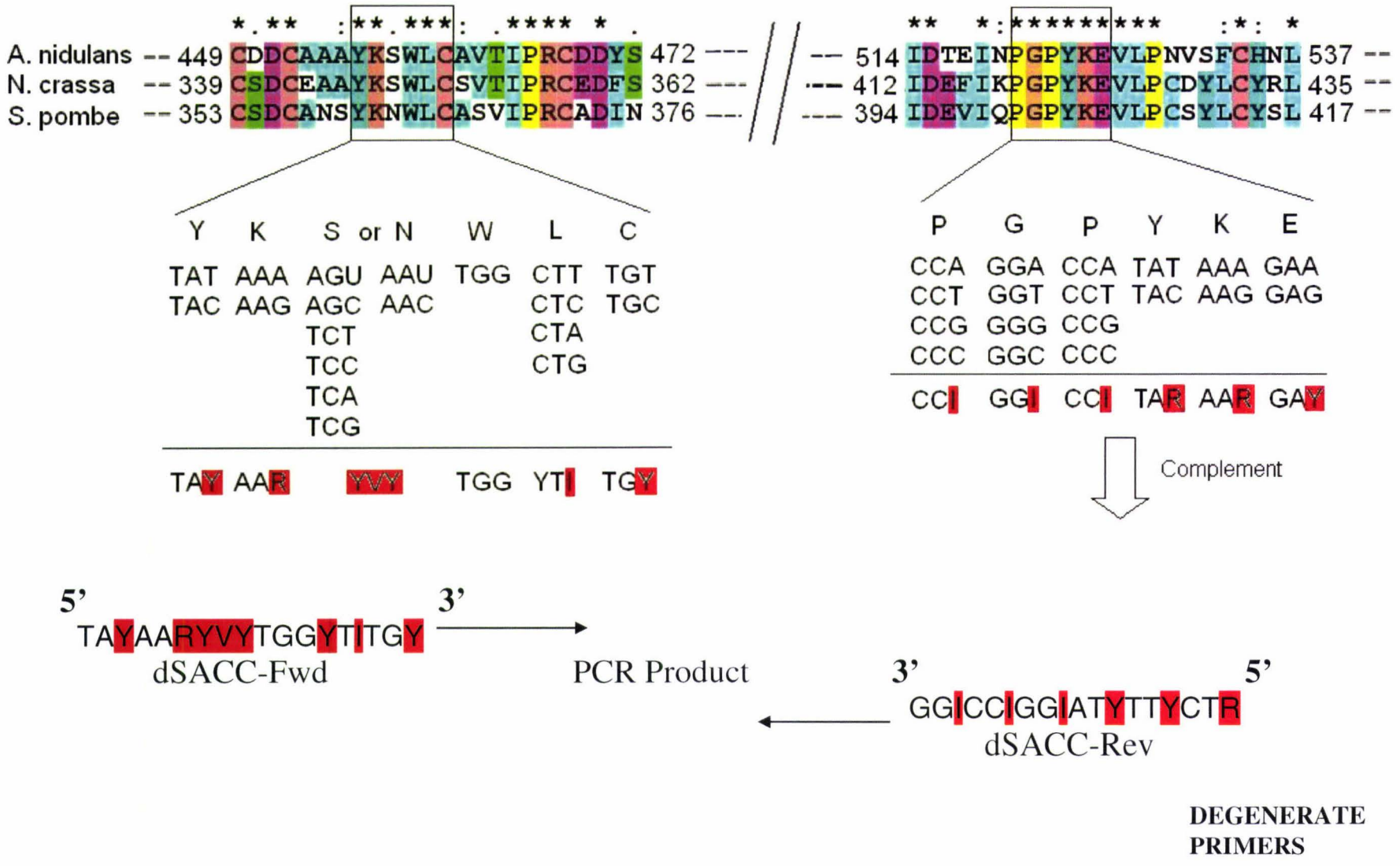
5.1.3 Attempted Detection of a putative Lp19 *SACC* using an *A. nidulans* probe on a Southern Blot

To enrich the template for further PCR reactions and possibly omit non-specific products from out-competing the target it was decided to attempt to detect the Lp19 *SACC* gene using a section of the *A. nidulans SACC* gene as a probe. The 156 bp *A. nidulans* product was cloned and sequenced. PrSACC1 was generated by PCR using dSACC-Fwd and dSACC-Rev primers. DIG-labeled PrSACC was hybridized to a Southern blot containing 1.0 µg Lp19 genomic DNA digested with *BamHI*, *NheI*, *MluI*, *ClaI*, *NcoI* and *SmaI* as well as 1.0 ng *NcoI* digested pDSACC (positive control). No hybridization was detected over a series of hybridization temperatures (45-55°C in increments of 5°C) to any fragment in the Lp19 genome although the probe hybridized to the positive control.

In summary, degenerate PCR for the Lp19 *SACC* gene was unsuccessful, as was using a probe from a related species.

Figure 5.1

A): Partial Alignment of Proteins used to Design *SACC* Degenerate PCR Primers.
Partial alignment of *SACC* amino acid sequences from: *N. crassa* (Galagan *et al.*, 2003), *A. nidulans* (Jackson, 2002) and *S. pombe* (Tasaka *et al.*, 2000). Each box around a sequence indicates the region the primer labeled was designed, and the stars above each column show a consensus for that residue. The numbers annotate the locations of these sites in each of the sequences used and arrows show directionality. The septated line indicates there is more alignment between the two figures at the start and ends of these. The double backslash indicates that the upper portion and lower portion of sequence are of the same alignment only a number of residues apart.
The two regions identified in were selected for primer design, and all possible combinations of potential codons analyzed. Using the degenerate code ambiguities were solved and dSACC-Fwd + dSACC-Rev were designed (degenerate nucleotide locations in red; Section 2.2.8 and Table 2.2). The estimated PCR product size was approximately 150 base pairs.


A. nidulans -- 449 CDDCAAAYKSWLCAVTIPRCDDYS 472 ---
N. crassa -- 339 CSDCEAAYKSWLCSVTIPRCEDFS 362 ---
S. pombe -- 353 CSDCANSYKNWLCASVIPRCADIN 376 ---
--- 514 IDTEINPGPYKEVLPNVSFCHNL 537 --
--- 412 IDEFIKPGPYKEVLPCDYLCYRL 435 --
--- 394 IDEVIQPGPYKEVLPCSYLCYSL 417 --
Y K S or N W L C
TAT AAA AGU AAU TGG CTT TGT
TAC AAG AGC AAC CTC TGC
TCT CTA
TCC CTG
TCA
TCG
TAY AAR YVY TGG YTI TGY
P G P Y K E
CCA GGA CCA TAT AAA GAA
CCT GGT CCT TAC AAG GAG
CCG GGG CCG
CCC GGC CCC
CCI GGI CCI TAR AAR GAY
Complement
5'
TAYAARYVYTGGYTITGY
3'
dSACC-Fwd
PCR Product
3'
5'
GGICCIGGIATYTTYCTR
dSACC-Rev
DEGENERATE PRIMERS

Chapter 6

nlEF-2 AND *nlCDC-12* EXPRESSION STUDIES

6.1 Expression of *nlEF-2* and *nlCDC-12 in culture.*

To characterize the expression pattern of *nlEF-2* and *nlCDC-12* in culture a decision was made to perform Northern analysis. Northern analysis has many advantages over more contempary techniques such as RT-PCR in that semi-quantitative data can be obtained, transcript size estimated, and it is recyclable although an order of magnitude more template is required due to lower sensitivity (Ausubel *et al.*, 1997).

Initially growth curves for RNA extraction were prepared according to protocol in section 2.2.1.2.3 (Figure 6.1). Due to problems the standard procedures were altered from the first growth curve onward to omit the starvation step and a modification to the harvesting from pre-cultures that resulted in a shorter lag phase and a more even culture (Figure 6.1; Section 2.2.1.2.4). To compare expression of both genes over *N. lolii* growth phases on the Northern blot, total RNA was isolated from samples taken from three areas within the growth curve. The growth phases selected were early exponential, mid-exponential and stationary phases. To identify transcripts of interest a 295 bp *nlEF-2* probe (PrEF2.2) and a 1039 bp *nlCDC-12* probe (PrSep2) were generated by PCR using exEF2-Fwd and EF2-IPCR 5'; npSep-Fwd and npSep-Rev respectively (Table 2.2; PR2 and 10 Table A2.1). Both PCR reactions used RACE-ready cDNA prepared in Section 3.2.1 as a template. To check if RNA transfer was efficient it was decided to use the constitutively expressed *N. lolii β-tubulin2* gene as a control (Bryant, 2003). A 351 bp *β-Tubulin2* cDNA probe (PrbTUB) was generated by PCR using T1.1 and T1.2 primers with pMM16 as a template (Section 2.2.9 using PR11 Table A2.1 and Table 2.2). Northern blots were prepared in duplicate containing varying amounts of Lp19 total RNA and hybridized with PrbTUB (Section 2.2.17.2). To allow semi-quantification by comparison for expression change, Mexp2 was loaded multiple times and the band density estimated over a range of concentrations. A band of the expected size (about 2.0 kb) was observed in most lanes, with exception of Mexp2 0.5 μg where it is assumed to be below detectable limits (Figure 6.2). As expected the level of *β-tubulin2* expression reflected the quantity of total RNA loaded, and was relatively consistent across growth phases although a

Figure 6.1: Lp19 Growth Curve

A log graph depicting the typical sigmoidal growth pattern associated with exponentially growing fungi *in culture*. Each point on both graphs indicated a point during growth where a sample of the culture was taken, (Table A3.3 for actual figures). The selected points from which samples were taken for total RNA extraction and subsequent Northern analysis are circled in gray and numbered. The early exponential phase (Eexp1) sample is labeled 1 and mid-exponential phase samples (Mexp1 and Mexp2) labeled 2 and 3 respectively. The stationary phase samples (Sta1 and Sta2) are labeled 4 and 5 respectively. Curve A was obtained by growing Lp19 according to standard procedures (Section 2.2.1.2.3) and curve B by altered procedures (Section 2.2.1.2.4).

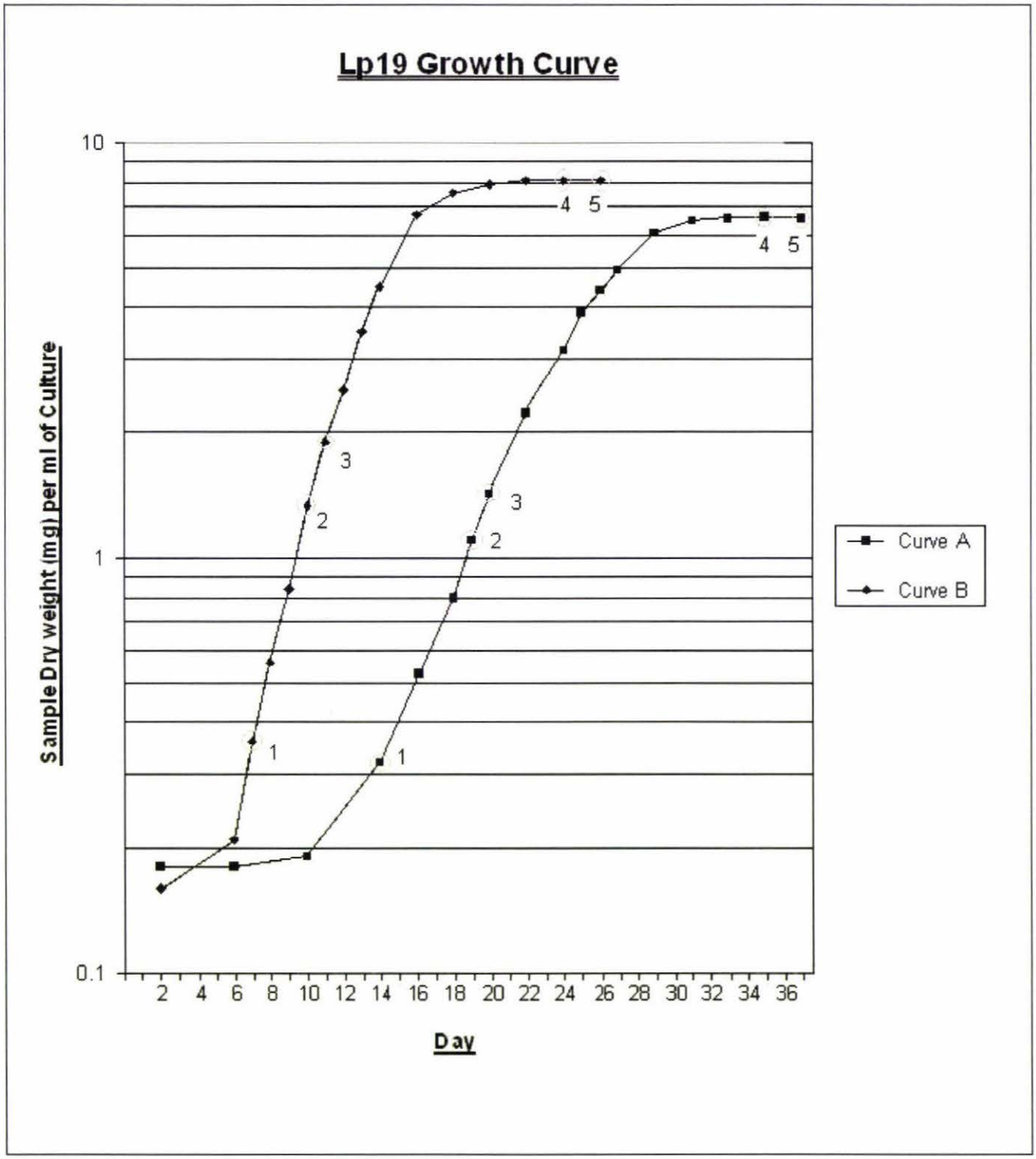
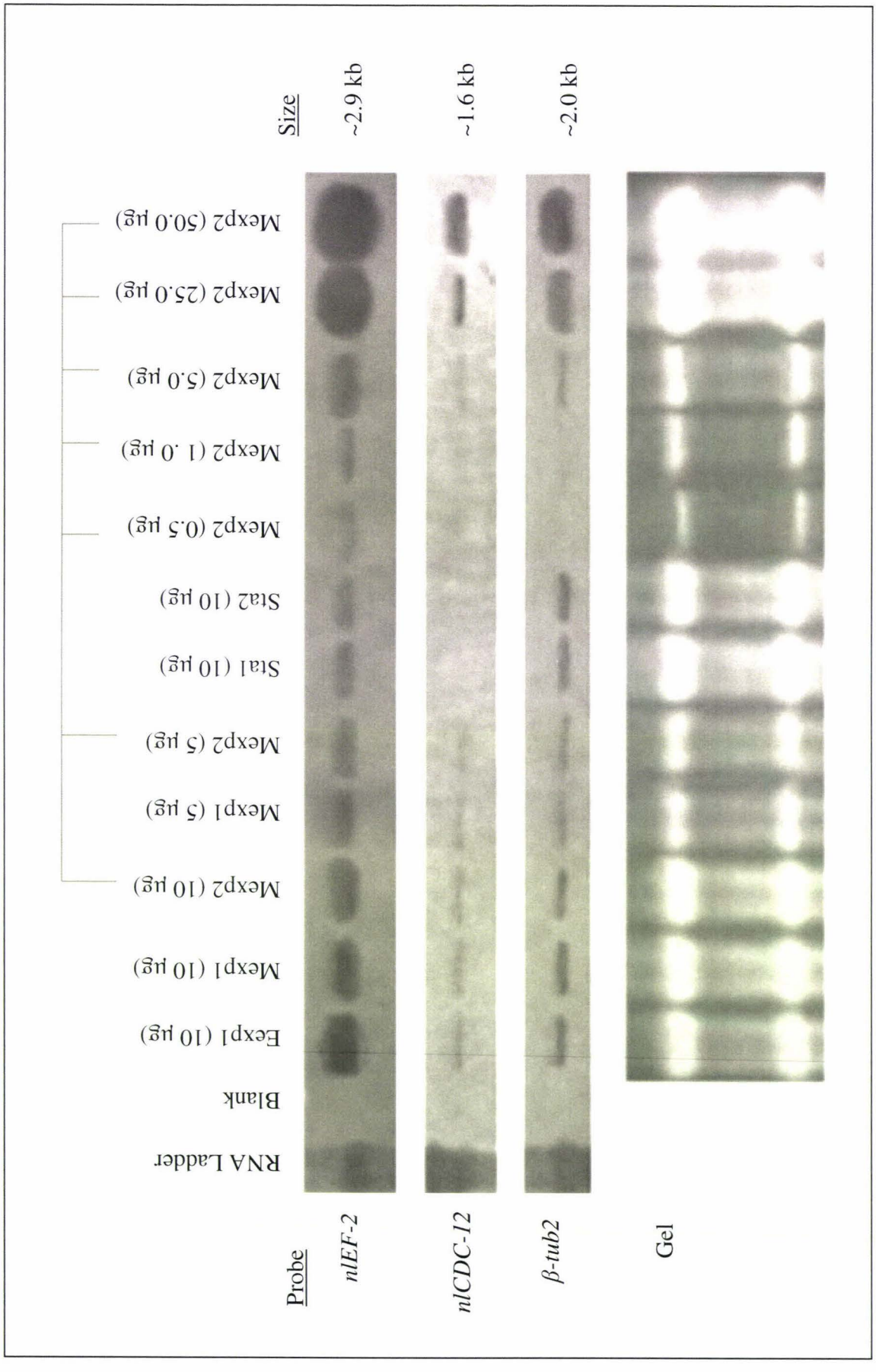

Lp19 Growth Curve
10
1
0.1
Sample Dry weight (mg) per ml of Culture
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36
Day
1
2
3
4 5
1
2
3
4 5
Curve A
Curve B

Figure 6.2: Northern Analysis of *nlEF-2* and *nlCDC-12* Expression

Total Lp19 RNA was used from early exponential phase (Eexp1), two different time points in mid exponential phase (Mexp1 and Mexp2) and two different time points in stationary phase (Sta1 and Sta2); times of sampling are marked in figure 5.1. Varying amounts of RNA were loaded on a gel (indicated at the top of each lane) and electrophoretically separated. A Northern blot was probed successively with Dig high-prime labeled probes directed against the *nlEF-2*, *nlCDC-12* and *β-tub2* gene transcripts (probe sizes were 245, 1039 and 351 bp respectively), the concentration of labeled probe estimated using a spot test. Hybridization was detected using a LAS-1000 (Fujifilm Intelligent dark box II) at the expected sizes (indicated at the right hand side of the figure); these were determined using a Dig labeled RNA ladder (Roche). Detections were 35 min exposures. To allow semi-quantitative assessment of expression levels, one of samples (Mexp2) was loaded in a range of concentrations (marked by grey bracket). The RNA gel and *β-tub2* results show that RNA loading and subsequent transfer from the gel to the membrane was consistent. Hybridizations were performed at 55°C and the blots were washed at 55°C (Section 2.2.17.2).


Probe
nlEF-2
nlCDC-12
β-tub2
Gel
Size
~2.9 kb
~1.6 kb
~2.0 kb
RNA Ladder
Blank
Eexp1 (10 µg)
Mexp1 (10 µg)
Mexp2 (10 µg)
Mexp1 (5 µg)
Mexp2 (5 µg)
Sta1 (10 µg)
Sta2 (10 µg)
Mexp2 (0.5 µg)
Mexp2 (1 .0 µg)
Mexp2 (5.0 µg)
Mexp2 (25.0 µg)
Mexp2 (50.0 µg)

slight decrease was observed in stationary phase. The blot was then stripped and hybridized with PrEF2.2 (Sections 2.2.17.3 and 2.2.17.2). This concurred with RACE results (Section 3.2.1) that *nlEF-2* transcripts are approximately 2.9 kb in length. Peak expression was observed at early exponential phase, and then paralleled the growth phase of the culture. Message declined by approximately 10-fold during stationary phase cultures (Figure 6.2). The blot was stripped again and hybridized with PrSep2 (Sections 2.2.17.3 and 2.2.17.2). This concurred with RACE result (Section 4.2.1) that *nlCDC-12* transcripts are approximately 1.6 kb in length. Expression was observed at early and mid exponential stages, no signals were detected at stationary phase suggesting that there was at least a 4-fold decrease in expression (within detectable levels). These results were repeated in the Northern prepared with samples from the replicate growth curve (results not shown).

6.2 Expression of *nlEF-2* and *nlCDC-12 in planta*

As discussed in Section 1.1.1, in planta Lp19 hyphae have a low biomass (less than of 0.2% of total infected plant tissue). In consideration of this it would not be feasible to conduct a Northern analysis of *nlEF-2* or *nlCDC-12* expression using this technique. Therefore it was decided to conduct a pilot experiment to determine the presence of *nlEF-2* and *nlCDC-12* mRNA transcripts *in planta* by RT-PCR.

According to studies by Tan *et al* (2001) metabolism is constantly high when hyphal extension ceases and observations of hyphal wall and septum thickening by Christensen *et al* (2000) reinforce this. If both statements were accurate then both *nlEF-2* and *nlCDC-12* would be expressed to varying degrees in mature leaves. As Tan *et al* (2001) observed the whole leaf sheath section of Lp19 endophyte infected *L. perenne* tissue (Nui cultivar; Table 2.1) is richer in endophyte than other sections (0.2% peaking in older tissue). As only the youngest sheath region is still growing and the outer sheath is starting to degrade as the tissue becomes less dense (Christensen *et al.*, 2001), it was decided to test both genes expression using the 2nd youngest leaf sheath region (Figure A1.1). Lp19 endophyte infected and non-infected *L. perenne* tissue (Nui cultivar; Table 2.1) was used to isolate total RNA. Non-infected plant tissue was extracted as a negative control and a portion of each was also used to isolate genomic DNA for control reactions. To enhance transcript yield and reduce any potential interference, total RNA was separated using GenElute columns (Sigma) to isolate mRNA

(Section 2.2.16.3). cDNA was synthesized using the mRNA template with random primers (Invitrogen, Sections 2.2.4.2 and 2.2.18.1.1).

PCR amplification was then carried out with primers EF2-RACEc and EF2 5' RACE for *nlEF-2*; Sep-RACEc and Sep 5' RACE for *nlCDC-12*; Sections 3.2.1 and 4.2.1 respectively.

These primers are also those used in RACE controls and were selected for PCR reactions as they amplify across an intron in their respective targets. To maintain consistency new *β-tubulin2* primers (T1.1 IP-Fwd and T1.2 IP-Rev) were designed using a ClustalX alignment of *Epichloe typhina* and Lp19 *β-tubulin2* genes to act as a positive control. These new *β-tubulin2* primers were extended forms of T1.1 and T1.2. All sets of primers were expected to result in similar-sized products (287-368 and 736 bp) and all have similar Tm, this enabled PCR reactions for both genes to be used during the same protocol. The four template types were compared using each set of primers to ensure the validity of results (non-infected genomic DNA and cDNA, infected genomic DNA and cDNA). Using infected tissue genomic DNA reactions were expected to result in 316 (*nlEF-2*), 368 (*nlCDC-12*) and 736 (*β-tubulin2*) bp products, while cDNA reactions are expected to result in 287 (*nlEF-2*), 300 (*nlCDC-12*) and 360 (*β-tubulin2*) bp products. PCR reactions using non-infected tissue were not expected to result in products, and function as negative controls.

Primarily two sets of PCR reactions were performed at an annealing temperature of 65°C using 10 µl cDNA or 100 ng genomic DNA in 12 separate reactions (Section 2.2.18.2.2 using PR9 Table A2.1). Following this initial reaction, the resulting product was diluted 1:100 and another round of PCR commenced using this template, this step was taken to reduce background products and increase the specificity of the reaction. All products of expected sizes were identified (Figure 6.3). These were direct-sequenced and the resulting sequences for each gene were verified by individual comparison to one another (cDNA v.s DNA product) and the respective gene sequences using ClustalX (Section 2.2.11). All sequences derived from cDNA lacked predicted intronic sequence, confirming the expression of all three genes *in planta*. This shows that both genes are detectable using Real-time RT-PCR and that they can be used to monitour endophyte *in planta*.

Figure 6.3: *In planta* Analysis of *nlEF-2* and *nlCDC 12* Expression

20 µl of each PCR reaction was electrophoretically separated on a 1% agarose gel in 1xTAE buffer for about 1 hour at 100 V. Molecular sizes of bands are indicated on the left in base pairs (bp), (0.5 µg of 1Kb plus ladder (Invitrogen) was loaded in lane 1. As expected genomic DNA (gDNA) and cDNA derived from infected (I.F) *L. perenne* resulted in products while non-infected (N.I) did not, validating that the products were a consequence of *N. lolii* presence in the tissue. EF2-RACEc and EF2 5' RACE PCR reactions resulted in single 300 bp and 368 bp bands with cDNA and genomic DNA templates respectively. Sep-RACEc and Sep 5' RACE PCR reactions resulted in single 300 bp and 368 bp bands with cDNA and genomic DNA templates respectively. These results are indicative that both genes are expressed *in planta*.

T1.1 IP-Fwd and T1.2 IP-Rev PCR positive control reactions resulted in single 360 bp and 736 bp bands with cDNA and genomic DNA templates respectively.

Size (bp)
12, 000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100
Ladder
Blank
β-tub2 (N.I; gDNA)
β-tub2 (N.I; cDNA)
β-tub2 (If; gDNA)
β-tub2 (If; cDNA)
nlEF-2 (N.I; gDNA)
nlEF-2 (N.I; cDNA)
nlEF-2 (If; gDNA)
nlEF-2 (If; cDNA)
nlCDC-12 (N.I; gDNA)
nlCDC-12 (N.I; cDNA)
nlCDC-12 (If; gDNA)
nlCDC-12 (If; cDNA)

Chapter 7

Discussion and Future Directions

7.1 Summary of *nlEF-2* and *nlCDC-12* General Features

N. lolii (Lp19) was isolated by Christensen *et al* (1993) from perennial ryegrass *Lolium perenne*. The *EF-2* and *CDC-12* genes were isolated, cloned and sequenced (Sections 3.3.1 and 4.3.1 respectively). Transcript and flanking regions for both genes were identified using IPCR and RACE techniques. The use of formamide (Sarkar *et al.*, 1990) during IPCR and RACE PCR protocols was invaluable for retaining primer specificity.

The overall structure of each gene is very similar to those of *N. crassa* and other phylogenetically close relatives, yet contrasts markedly with those of more distant species such as *S. cerevisiae* whose genes generally lack introns (Fink, 1987). *nlEF-2* and *nlCDC-12* displayed conserved motifs for introns (Gurr *et al.*, 1987), transcription start and stop sites (Bucher, 1990; Elder, 1992) and translation start sites (Kozak, 1981; Bruchez *et al.*, 1993; Edelmann and Staben, 1994) and stop sites (Weaver, 1999) from a variety of related fungi.

Interestingly there is a general third codon bias in Lp19 genes and those from several close relatives. Lp19 *β-tub2* (Bryant, 2003), *hmg CoA reductase* (Dobson, 1997), peptide synthetase (Panaccione *et al.*, 2001), *pyr4* genes (Collett *et al.*, 1995) and *E. typhina β-tub2* gene (Byrd *et al.*, 1990) all show a codon bias away from using adenosine in this position and not all are highly expressed genes, suggesting that a bias away from using adenosine in the third codon may be a more extensive phenomenon in endophytic genes, aquired from a common ancestor.

7.2 Summary of *nlEF-2* Specific Features

7.2.1 Promoter Elements

Analysis of the *nlEF-2* gene promoter revealed the presence of a number of putative regulatory elements that may influence high expression including a TATA box, an inverted Y-box and four GC boxes, the later which is also common to house-keeping genes (Nakanishi *et al.*, 1988;

Weaver, 1999). A putative GCN4 element identified within the promoter may be an activator of *nlEF-2* as its function is involved in activation of protein biosynthesis (Arndt and Fink, 1986). Two other elements found in the promoter, Rap1p and MCM1 may also be activators of *nlEF-2* as their functions in *S. cerevisiae* have been associated with activation of genes within the cell cycle and ribosomal genes respectively (Kuo and Grayhack, 1994; Mizuta *et al.*, 1998). This suggests that the mechanisms used to control the expression of *nlEF-2* may be common to other genes involved in protein biosynthesis.

7.2.2 Putative Protein Elements

nlEF-2 encodes a predicted protein of 844 amino acids and exhibits between 77 (*S. cerevisiae*) and 90% (*N. crassa*) sequence identity within EF-2 proteins from other fungal species. This high level of identity underlines the functional importance of EF-2 by sequence conservation through evolution. Like other EF-2, nlEF-2p was found to be a five-domain protein (Section 3.3.5.2). The N-terminal G-domain (residues 2-346; Figure 3.13) contained the highly conserved GTP-binding motifs G1-5, all of which have been demonstrated to be critical for the GTPase catalytic activity of EF-2 in *H. sapiens* (Rapp *et al.*, 1989), *S. cerevisiae* (Perentesis *et al.*, 1992) and *E. coli* (Ovchinnikov *et al.*, 1983). The regulation of EF-2 activity may be by phosphorylation/dephosphorylation of the conserved Threonine^{57} residue which is present in signature EF-2 effector motif and also located in the G-domain (Ryazanov and Spirin, 1990; Bourne *et al.*, 1991). The other 4 domains are involved in ribosomal and inhibitor interactions. The C-terminal end of domain 4 (residues 802-844) contain a predicted helix-turn-helix motif which may interact with RNA in the decoding region of the 30S ribosomal subunit and enhance translation (Joseph, 2003). Possible inhibitor interaction domains identified include the dipthamide binding domain which contained the conserved Histidine^{700} residue which is the specific target of Diptheria toxin (Perentesis *et al.*, 1992), and the Sordarin binding domain which contained Serine^{524} and Glutamic acid^{525}, two of the three critical residues that confer sordarin sensitivity possibly rendering Lp19 partially sensitive (Shastry *et al.*, 2001). Should this the case, there is a strong chance that the closely related ryegrass 'choke' pathogens *Epichloe* spp. may have the same sequence features. This could have implications for Sordarin use in control of this pathogen.

7.3 *nlEF-2* Expression and Future Directions

nlEF-2 was demonstrated to be expressed *in culture* and *in planta*. The results of *nlEF-2* expression during Lp19 growth curves support the *in culture* constitutive GUS-reporter gene assay results by Tan *et al* (2001). As the expression of GUS at exponential stage was at its peak so was *nlEF-2* expression. When GUS reporter gene expression dropped markedly in stationery phase, the *nlEF-2* expression dropped with, but not quite as low as the GUS reporter gene (Tan *et al.*, 2001). The possible explanation was although growth had ceased within detectable limits; perhaps the stationary phase samples were not taken deep enough into stationary phase. However *nlEF-2* expression was detected in mature leaves, further confirming metabolism continues while growth ceases *in planta*. Levels of expression in different plant tissue can be examined by Real-time PCR to determine areas and levels of metabolic activity when growth has stopped. This would be an easier and more meaningful test than any reporter gene construct (i.e. GUS or Green fluorescent protein; GFP) transformed into *N. lolii*. Reinfection is a difficult process (Tan *et al.*, 2001) and while GUS and GFP have comparatively shorter halve-lives (i.e. you can't tell whether expression happened now or up to 24 hours ago) RNA can be snap frozen from an instant in time.

7.4 *nlCDC-12* Sequence Features Summary

7.4.1 Promoter Elements

Analysis of the *nlCDC-12* gene promoter revealed the presence of a number of putative regulatory elements such as an inverted Y-box and three GC boxes, the latter of which is also common to house-keeping genes and is likely to be important for *nlCDC-12* expression as the promoter lacks a TATA box (Nakanishi *et al.*, 1988; Smale, 1997). A putative E2F element identified within the promoter may be an activator of *nlCDC-12* as its function is involved in control of cell proliferation (Malhotra *et al.*, 1993). Two other potentially important elements were found in the promoter. The CCBF site may be an activator of *nlCDC-12* as its function in *S. cerevisiae* has been associated with activation of genes within the cell cycle (Andrews and Herskowitz, 1989), and the abaA site has been shown in *A. nidulans* to act as a switch during cellular development (Andrianopoulos and Timberlake, 1994). This suggests that the

mechanisms used to control the expression of *nlCDC-12* may be common to other genes involved in cell division and proliferation.

7.4.2 Putative Protein Elements

nlCDC-12 encodes a predicted protein of 386 amino acids and exhibits between 56 (*S. cerevisiae*) and 93% (*N. crassa*) sequence identity with proteins from other fungal species. This level of sequence identity exhibits the functional importance CDC-12 by sequence conservation through evolution. Like other CDC-12, nlCDC-12p was found to contain an N-terminal GTPase domain (residues 26-180; figure 4.10) and contained the highly conserved GTP-binding motifs G1-4 (Bourne *et al.*, 1991), the later have been demonstrated using mutagenesis studies in mice to be involved in proper septin localization and polymerization and contain the GTPase catalytic activity of CDC-12 (Kinoshita *et al.*, 1997). A coiled-coil motif common to most septins was predicted in the C-terminal end of *nCDC-12* (residues 348-385), this may facilitate interactions with other septins in the formation of the 10 nm ring (Casamayor and Snyder, 2003). As an aside the naming system of CDC-12 proteins is in need of review. A group of formins are referred to as 'cdc12' although they show no homology to the CDC-12 proteins in this study (Chang *et al.*, 1997). This problem may have originated by arbitarily naming mutations cdc1, 2, 3 etc through genetic screens in the 1970s (CDC-12 mutations are lethal (Momany *et al.*, 2001; Warenda and Konopka, 2002)), the concept of naming genes from homology only occurred in later years (Chang, 2004).

7.5 *nlCDC-12* Expression and Future Directions

nlCDC-12 was expressed *in culture* and *in planta*. The results of *nlCDC-12* expression during growth of Lp19 showed that *nlCDC-12* is most highly expressed during stages of exponential growth, and below detectable limits of expression (at least a 4-fold decrease) in stationary phase. These results suggest that theoretically septa will only be laid down during periods of Lp19 growth. Interestingly this was not the case *in planta* as *nlCDC-12* was expressed in mature leaf tissue that according to current theory of endophytic growth (Section 1.1.4.2) is non-growing. This is likely be low-level expression involved in cell wall and septum thickening observed by Christensen *et al* (2004), more septal compartments being formed, nlCDC-12 functioning in processes other than septum formation or a combination (Section 1.3.2). Before further

experiments the level of this expression would need to be assessed using Real-time PCR. Post-testing regions of *N. lolii* growth can be examined *in planta* by age of septa and this would contribute to the *in planta* growth modeling of Lp19. If septa were younger in the leaf expansion zone than at hyphal tips this would support the theory of intercalary extension by Christensen *et al* (2004; (Christensen and Bennett, 2004b)). This type of analysis could be achieved by firstly transforming Lp19 with two constructs; a full cDNA copy of *nlCDC-12* with a specific epitope under the control of a heterologous constitutive promoter which is able to be inactivated, and a full cDNA copy of *nlCDC-12* with a different epitope under the control of a leak-free inducible promoter. Using *N. lolii* growing *in planta* under this system, the constitutive promoter *nlCDC-12* would be inactivated concurrent to the activation of the inducible promoter *nlCDC-12*. Older and younger/older septa location can be distinguished using ELISA with monoclonal antibodies (Stryer, 1995); ideally a third inducible promoter would also be used with a different epitope to give better pattern over leaf sections. To ensure successful transformations differing selective markers on each plasmid can be used (such as hygromycin and tetracycline). In addition this could be used in conjunction with co-immunoprecipitation reactions and could reveal other proteins involved in branch site selection and help explain irregular branching *in planta* compared with the regular branching phenotypes observed *in culture* (Schmid and Christensen, 1999).

7.6 *SACC* Future Directions

Although there was a lack of success with degenerate primers and probing using a section of *A. nidulans SACC*, the identification of new *SACC* genes would facilitate the design of new primers. A possible reason why probing for *SACC* was unsuccessful is that all fungal *SACC* genes cloned to date have functions integrated with mating (Iida *et al.*, 1994; Tasaka *et al.*, 2000), where *N. lolii* Lp19 is asexual (Section 1.1.1) and may have a modified version with differing function. Taking into account the similarity found between the *nlCDC-12* and *nlEF-2* genes with their respective homologues from *N. crassa* and *M. grisea*, and the 3rd codon bias away from using adenosine the potential for re-design of degenerate primers is promising. Unfortunately a great deal of research is required before *SACC* homologues from more closely related species such as *M. grisea*, *U. maydis* and *F. graminearium* are characterized beyond automatic annotation in genome projects. The potential uses of cloning and characterizing the

Lp19 *SACC* gene could possibly solve the question of the unique growth patterns exhibited *in planta.* This could be examined using molecular techniques such as labeling for distribution analysis, expression, mutagenesis and knockouts for functional studies. Cloning and expression studies may lead to the discovery of how Lp19 growth is controlled in planta.

Appendix 1

Table A1.1: Candidate Genes for Masters Studies

Organisms from which homologues have been isolated are listed in three-letter code (key at foot of table).

Contents

- *Cph1*

Organism:	*Candida albicans*
Accession number:	A54767
Homologues:	K.la; C.al; C.lu; M.gr; N.cr; E.ni; P.ma
Associated pathway:	Regulates enzyme activities of phospholipase and superoxide dismutase, regulates virulence genes (C.ne); control mating and yeast/hypha transitions (all); contributes to cell wall integrity, adhesion and invasion of epthelial cells, hyphal elongation (C.al).
Putative *N. lolli* function:	Regulation of hyphal development/morphogenesis
Notes:	In other organisms this gene is known as *Ste12*. Has a positive effect; transcription factor; may be MAPK target.

References: (Liu *et al.*, 1994; Kanzaki *et al.*, 1999; Tasaka *et al.*, 2000; Short *et al.*, 2002)

• *Cph2*

Organism: *Candida albicans*

Accession number: AF349507

Homologues: none identified

Associated pathway: Responsible for TEC1 transcription, regulates hyphal development

Putative *N. lolli* function: Regulation of hyphal development/morphogenesis

Notes: bHLH; regulates hyphal development partly via Tec1

References: (Lane *et al.*, 2001)

• *Tup1*

Organism: *Candida albicans*

Accession number: AF005741

Homologues: S.ce; K.la; Y.li; E.ni; N.cr; D.di; S.po; P.an

Associated pathway: Repression and histone binding functions are conserved across all species listed. These functions also include: repression of genes regulated by cell type; glucose; oxygen; DNA damage and other signals.

Putative *N. lolli* function: Possible hyphal repressor and/or regulator of hyphal branching

Notes: negative effect; transcription factor; general repressor

References: (Braun and Johnson, 1997; Gale *et al.*, 1998; Braun *et al.*, 2000; Braun and Johnson, 2000; Dieterich *et al.*, 2002)

• *Efg1*

Organism: *Candida albicans*

Accession number: P43064 / Z32687

Homologues:	S.ce; Y.li
Associated pathway:	Enhancing filament growth factor (C.al); Mycelial growth factor (Y.li); plays a general regulatory role in the PKA signal transduction pathway,a positive regulator of mitosis, a positive regulator of filamentation, function as negative regulators of meiosis (S.ce).
Putative *N. lolli* function:	Enhancing or regulating mycelial growth.
Notes:	Positive effect; transcription factor; may be MAPK targets
References:	(Stoldt *et al.*, 1997; Braun and Johnson, 2000; Leng *et al.*, 2001; Wartmann *et al.*, 2001; Dieterich *et al.*, 2002)

• *Cdc42*

Organism:	*Saccharomyces.cerevisiae*
Accession number:	NP_013330
Homologues:	E.go; C.al; Y.li; S.co; S.po; E.de; M.gr; S.bo; W.ba; C.tr
Associated pathway:	promotes assembly of a ring of septins at the site of future bud emergence (S.ce); Filament formation and invasive growth (Y.li); Cell polarization (all); Maintain actin cytoskeleton at hyphal tips (all); role in cell division and regulates endocytosis (all).
Putative *N. lolli* function:	Regulator of hyphal growth and morphogenesis
Notes:	
References:	(Boyce *et al.*, 2001)

• *Bmh1*

Organism:	*Saccharomyces.cerevisiae*
Accession number:	S30863
Homologues:	S.co; L.ed; P.ca; T.ha; H.je; C.al; E.ni; H.sa; X.la; D.me; A.th
Associated pathway:	essential for psuedohyphal induction and normal bud cell development,important in rapamycin-sensitive signalling

	(S.ce;D.di); cell cycle control post-DNA damage (S.po;C.al)
Putative *N. lolli* function:	Signalling protein involved in a number of cascades including hyphal development
Notes:	All of these can substitute for each other in functional aspects
References:	(Van Heusden *et al.*, 1992; Cognetti *et al.*, 2002)

• *Bmh2*

Organism:	*Saccharomyces.cerevisiae*
Accession number:	S51250
Homologues:	S.co; L.ed; P.ca; T.ha; H.je; C.al; E.ni; H.sa; X.la; D.me
Associated pathway:	essential for psedohyphal induction and normal bud cell development,important in rapamycin-sensitive signalling (S.ce;D.di); cell cycle control post-DNA damage (S.po;C.al)
Putative *N. lolli* function:	Signalling protein involved in a number of cascades including hyphal development
Notes:	All of these can substitute for each other in functional aspects.
References:	(Van Heusden *et al.*, 1992; Cognetti *et al.*, 2002)

• *DMC1*

Organism:	*Saccharomyces.cerevisiae*
Accession number:	A38214
Homologues:	C.al; S.po; P.os; C.ci; M.mu; C.py; H.sa; L.ma; B.mo
Associated pathway:	involved in meiosis through renaturation of ssDNA into supercoiled dsDNA (S.ce);
Putative *N. lolli* function:	Meiotic recombination

Notes:

References: (Fukushima *et al.*, 2000; Hong *et al.*, 2001)

• *Tec1*

Organism: *Saccharomyces.cerevisiae*

Accession number: A35667

Homologues: C.al

Associated pathway: Expressed in hyphal form, activates genes such as FLO11 for hyphal extension (S.ce;C.al).

Putative *N. lolli* function: Hyphal extension regulator

Notes: transcriptional activator

References: (Laloux *et al.*, 1990; Jahn *et al.*, 1993; Gavrias *et al.*, 1996; Oehlen and Cross, 1998; Schweizer *et al.*, 2000; La Valle and Wittenberg, 2001)

• *Dig1*

Organism: *Saccharomyces.cerevisiae*

Accession number: AAB68172

Homologues: None identified

Associated pathway: Co-factor in repression of filamentous response elements through Kss1 repressing ste12.

Putative *N. lolli* function: Repressor of hyphal extension

Notes: Involved in morphological regulation

References: (Bardwell *et al.*, 1998)

• *Dig2*

Organism: *Saccharomyces.cerevisiae*

Accession number: AAB64911

Homologues: None identified

Associated pathway:	Co-factor in repression of filamentous response elements through Kss1 repressing ste12.
Putative *N. lolli* function:	Repressor of hyphal extension
Notes:	Involved in morphological regulation
References:	(Bardwell *et al.*, 1998)

• *Rad6*

Organism:	*Candida albicans*
Accession number:	AAC24765
Homologues:	N.ha; N.cr; A.ni; S.ce; S.po; C.el; D.me; H.sa; M.mu;
Associated pathway:	Represses yeast/hypha morphogenesis (C.al); Ubiqutin binding protein required for DNA repair (S.ce).
Putative *N. lolli* function:	DNA repair protein and/or hyphal growth intiation regulator
Notes:	responds to DNA damage; also represses hyphal formation
References:	(Leng *et al.*, 2001)

• *CaSln1*

Organism:	*Candida albicans*
Accession number:	BAA24951
Homologues:	S.ce; K.la; V.ch
Associated pathway:	sensor to osmotic changes in environment (all), has a function in hyphal formation (C.al).
Putative *N. lolli* function:	Regulation of hyphal development/morphogenesis and/or initiator of hyphal growth.
Notes:	membrane sensor kinase
References:	(Nagahashi *et al.*, 1998)

• *CaNik1*

Organism:	*Candida albicans*
Accession number:	AAC72284
Homologues:	M.gr; N.cr; B.fu; N.ha; F.so; D.di; S.co; M.lo; Nost.sp
Associated pathway:	sensor to osmotic changes in environment (all), has a function in hyphal formation (C.al).
Putative *N. lolli* function:	Regulation of hyphal development/morphogenesis and/or initiator of hyphal growth.
Notes:	Cytoplasmic sensor kinase, deletion of this or any his kinases reduces hyphal formation
References:	(Nagahashi *et al.*, 1998)

• *Ssk1*

Organism:	*Candida albicans*
Accession number:	AAD5513
Homologues:	S.ce;C.al;S.po
Associated pathway:	sensor of oxidative stress (S.ce); has a role in hyphal development
Putative *N. lolli* function:	Role in hyphal development
Notes:	deletion of this reduces hyphal formation
References:	(Calera *et al.*, 2000)

• *Cot-3*

Organism:	*Neurospora crassa*
Accession number:	AAK49353
Homologues:	S.ce; S.po; C.al; C.gl; C.pa; C.lu; F.ne; C.tr; G.ga; C.el
Associated pathway:	hyphal elongation/branching (N.cr); general role in protein synthesis (all).
Putative *N. lolli* function:	Role in hyphal development
Notes:	A mutation results in abnormal hyphal elongation

	and branching
References:	(Propheta *et al.*, 2001)

• *Cot-1*

Organism:	*Neurospora crassa*
Accession number:	P38679
Homologues:	C.tr; N.cr; S.po; U.ma; H.sa; M.mu;
Associated pathway:	kinase is required for one or more events essential for hyphal elongation.
Putative *N. lolli* function:	Hyphal extension regulator
Notes:	Serine/Threonine-Protein Kinase; required for hyphal elongation
References:	(Yarden *et al.*, 1992; Lauter *et al.*, 1998)

• *CDC-12*

Organism:	*Neurospora crassa*
Accession number:	EAA53189.1
Homologues:	A.ni; N.cr; M.ci; C.al; S.po; S.ce
Associated pathway:	Part of a group of cell division proteins involved with cytokinesis (septation). May be involved in branch initiation with Int1p.
Putative *N. lolli* function:	Integral protein involved in 10 nm ring formation (septation)
Notes:	GTPase activity and polymerises
References:	(Gale *et al.*, 2001; Momany *et al.*, 2001; Warenda and Konopka, 2002; Westfall and Momany, 2002)

• *Yam8*

Organism:	*Schizosaccharomyces pombe*
Accession number:	JC7212
Homologues:	S.ce; C.al; M.gr
Associated pathway:	mediate cellular responses to mechanical stimuli
Putative *N. lolli* function:	Hyphal extension regulator through mechanical stress
Notes:	stretch-activated Ca2+-permeable channel protein homolog
References:	(Kanzaki *et al.*, 1999; Tasaka *et al.*, 2000; Short *et al.*, 2002)

• *Nrg1*

Organism:	*Candida albicans*
Accession number:	AAK09366
Homologues:	S.ce
Associated pathway:	represses a subset of Tup1-regulated genes, which includes known hypha-specific genes and other virulence factors (C.al); involved in glucose repression (S.ce).
Putative *N. lolli* function:	Repressor of hyphal extension
Notes:	
References:	(Murad *et al.*, 2001)

• *Int1*

Organism:	*Candida albicans*
Accession number:	AAA96019
Homologues:	None identified
Associated pathway:	hyphal growth, adhesion to epithelial cells
Putative *N. lolli* function:	Possible gene involved in hyphal adhesion to plant cells
Notes:	
References:	(Gale *et al.*, 1998)

• *Ece1*

Organism:	*Candida albicans*
Accession number:	Q07730
Homologues:	None identified
Associated pathway:	Expressed in association with cell elongation. May play a role in the process of hyphal growth of formation.
Putative *N. lolli* function:	Hyphal extension regulator
Notes:	Expressed in elongating hyphae, but not in budding yeast cells
References:	(Birse *et al.*, 1993)

• *Cho1*

Organism:	*Fusarium graminearum*
Accession number:	NP_010943
Homologues:	T.ae; S.ce; S.po
Associated pathway:	phosphatidylserine synthase
Putative *N. lolli* function:	Regulator of hyphal growth and branching
Notes:	
References:	(Atkinson *et al.*, 1980; Fecikova *et al.*, 1982; Wiebe *et al.*, 1989; Binks *et al.*, 1991; Markham *et al.*, 1993; Lakin-Thomas, 1998)

Key

A.fu = *A. fumigatus*
A.ni = *A. nidulans*
B.fu = *B. fukleliana*
B.mo = *B. mori*
C.a = *C. albicans*
C.ci = *C. cinerca*
C.el = *C. elegans*
C.gl = *C. glabrata*
C.im = *C. immitis*
C.lu = *C. lusitaniae*
C.pa = *C.parapsilosis*
C.py = *C. pyrrhogaster*
C.tr = *C. trifolii*
C.tr = *C. tropicalis*
D.di = *D. discoideum*
D.me = *D. melanogaster*
E.de = *E. dermatitidis*
E.go = *E. gossypii*
F.ne = *F. neoformans*
F.so = *F. solani*
G.ga = *G. gallus*
H.je = *H. jecorina*
H.sa = *H. sapiens*
K.la = *K. lactis*
L.bi = *L. bicolour*
L.ed = *L. edodes*
L.ma = *L. major*
M.gr = *M. grisea*
M.lo = *M. loti*
M.mu = *M. musculus*
M.ra = *M. raccmosus*
N.cr = *N. crassa*
N.ha = *N. haematococca*
Nost.sp= *Nostoc sp*
P.an = *P. angusta*
P.an = *P. anserine*
P.bl = *P. blakesleeanus*
P.br = *P. brasiliensis*
P.ca = *P. carinii*
Piso.sp= *Pisolithus sp.441*
P.ma = *P. marneffei*
P.no = *P. nodorum*
P.os = *P. ostreatus*
R.mi = *R. microsporus*
S.bo = *S. bovines*
S.ce = *S. cerevisiae*
S.co = *S. coelicolor*
S.co = *S. commune*
S.po = *S. pombe*
T.ha = *T. harzianum*
U.ma = *U. maydis*
V.ch = *V. cholerae*
W.ba = *W. bancrofti*
X.la = *X. laevius*
Y.li = *Y. lipolytica*

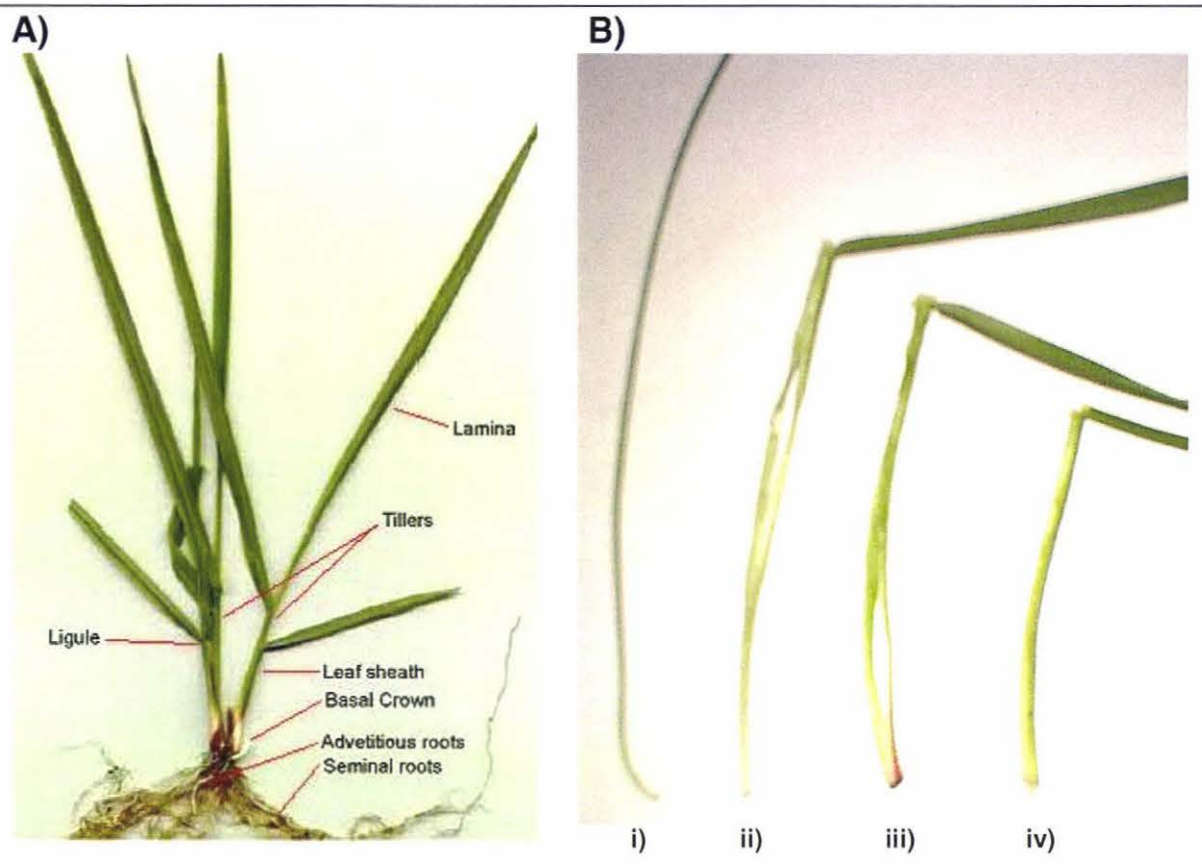


Figure A1.1: Host Grass

A) Example host grass (*F. arundinacea*) in which the basic structure is label appropriately. Tillers come up from the basal crown. Each tiller is composed of several leaves that are rolled together in the lower part of the tiller to form a pseudostem. Each leaf is composed of a lamina and a sheath. The sheath and lamina are separated by the ligule. The tip of the blade is the first part of the leaf to be formed and the base of the sheath the last. Adventitious and seminal roots arise from the base of tillers ((Langer, 1974); photo: (Christensen and Bennett, 2004a)) .

B) *Lolium perenne* tillers were dissected into leaves, i) being the innermost and iv) the outermost respectively. The leaf sheath from leaf ii) was selected for RNA and DNA isolation and to maintain consistency only tillers consisting of four leaves were used (i-iv).

Appendix 2

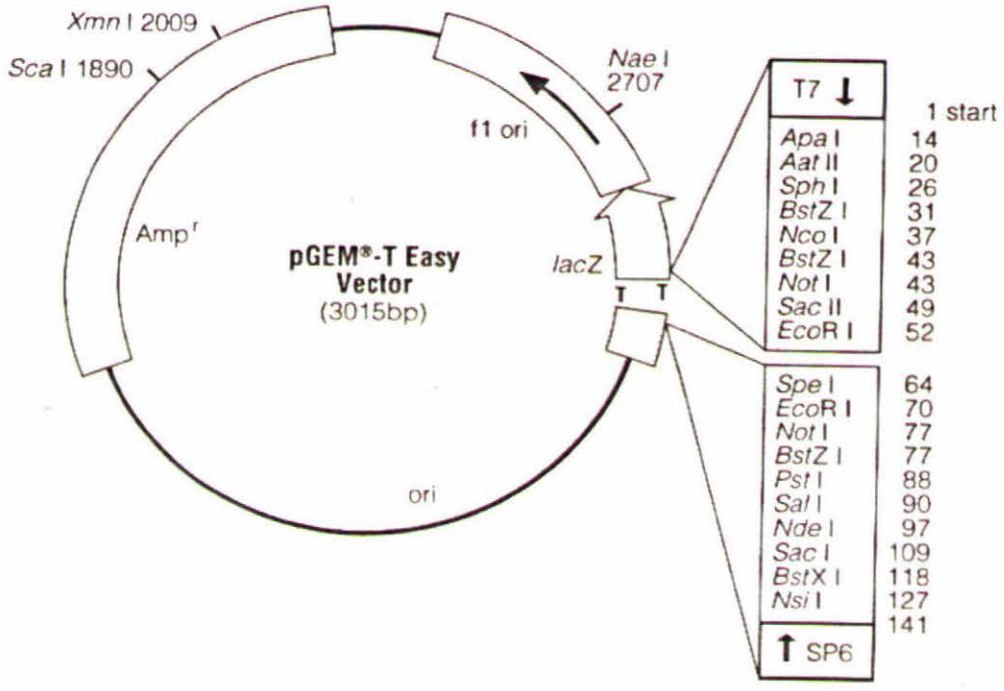


Figure A2.1: pGEM-T Easy vector Map

The pGEM-T Easy vector allows blue/white selection of clones by disruption of the *lacZ'* when a fragment of interest is ligated into the multiple cloning site. The thymine 3' over hangs allows easy cloning using a Taq polymerase derived PCR product. The full sequence for pGEM-T Easy is available on http://www.promega.com/vectors/pgemt.txt.

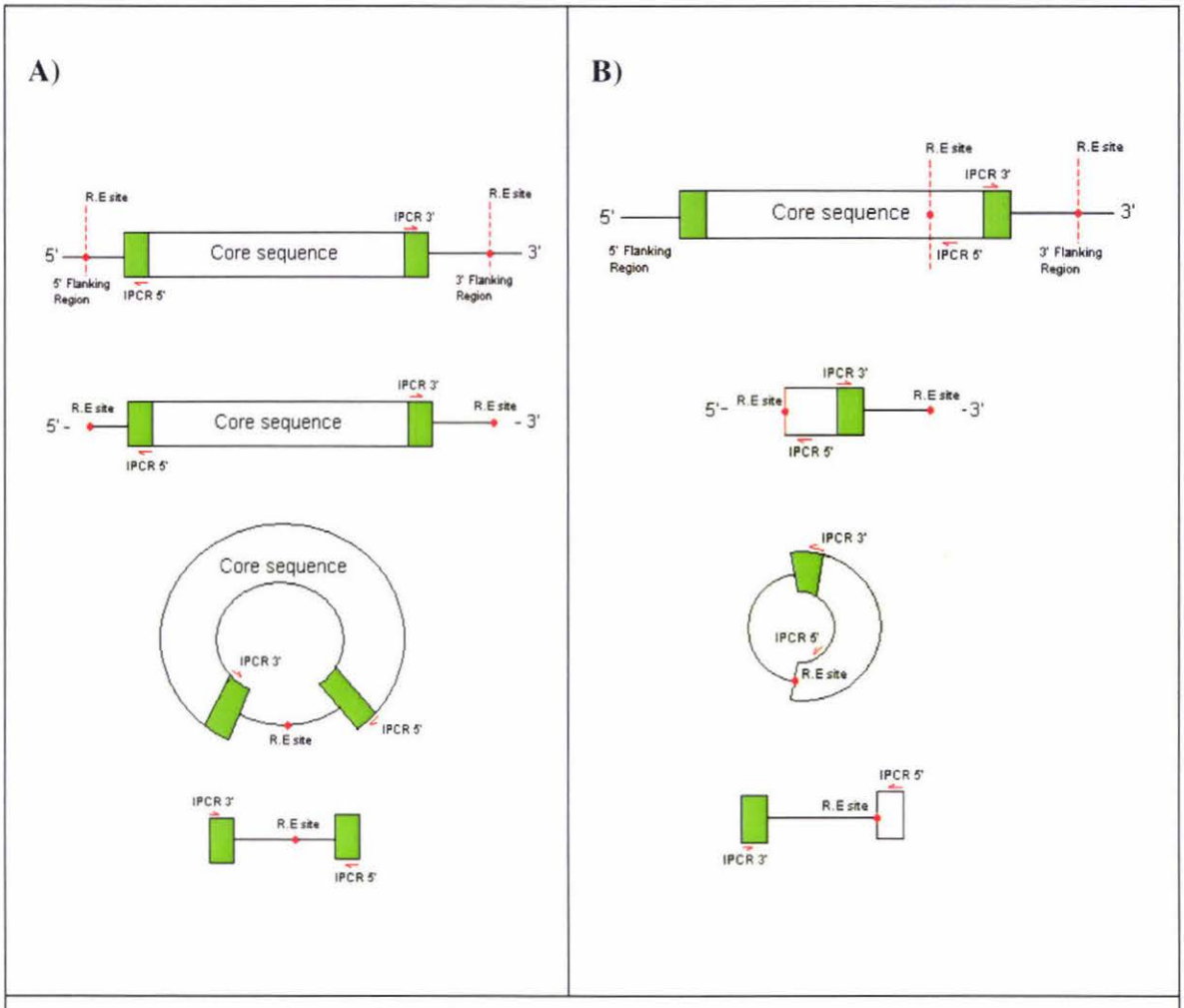


Figure A2.2 IPCR Strategies Schematic

Strategy A) uses restriction enzymes that do not cut within the core sequence and primers extending outward from its ends (Ochman *et al.*, 1988), while strategy B) uses a restriction enzyme that cuts within the core sequence; so although the primers are extending in opposite directions they may be designed anywhere in the core region (Siebert *et al.*, 1995)

Protocol (PR)	Denaturation (°C)	Time (seconds)	Annealing Temperature (°C)	Time (seconds)	Elongation (°C)	Time (seconds)	Final Elongation (°C)	Time (seconds)	# Cycles
PR 1	94	30	54	30	72	90	72	300	30
PR 2	94	60	55	40	72	50	72	420	30
PR 3	94	30	56	30	72	120	72	360	30
PR 4	94	30	55	30	72	50	72	300	30
PR 5	93	30	58	30	68	135			10
	93	30	58	30	68	155	68	480	24
PR 6	94	30	65	30	72	90	72	300	35
PR 7	94	30	52	30	72	90	72	300	30
PR 8	94	30	55	30	68	126			10
	94	30	55	30	68	146	68	480	24
PR 9	94	30	65	30	72	90	72	300	40
PR 10	94	30	55	30	72	90	72	300	30
PR 11	94	30	55	30	72	90	72	300	30

Protocol (PR)	Notes	PR 6	RACE PCR
PR 1	*nlEF-2* Degenerate PCR	PR 7	*nlCDC-12* Degenerate PCR
PR 2	Checking Clones for Insert	PR 8	*nlCDC-12* IPCR Part 1 (Long range)
PR 3	*nlEF-2* Initial IPCR		*nlCDC-12* IPCR Part 2 (Long range) elongation
PR 4	*nlEF-2* Control IPCR/ Northern Probe		increases by 20 seconds per cycle
PR 5	*nlEF-2* IPCR Part 1 (Long Range)	PR 9	RT-PCR
	nlEF-2 IPCR Part 2 (Long Range) elongation	PR 10	*CDC-12* Northern Probe
	increases by 20 seconds per cycle	PR 11	*Beta-Tubulin2* Northern Probe

Table A2.1: PCR Protocols

Reaction	Ligation mixture	DH5ά competent cells	Dilution	Plates Used	Number of colonies White	Blue
1) DH5ά competent cells (negative control)	- -	10 µL	10^6 10^7 10^6	 LB plates LB A/I/X	TMC 281 0	- - -
2) pUC 118 (transformation/reagent control)	-	100 µL	Undiluted	LB A/I/X	-	559
3) Vector - ligase (control)	10 µL	100 µL	Undiluted 1:10	LB A/I/X LB A/I/X	0 0	27 4
4) Vector + control DNA (positive control)	10 µL	100 µL	Undiluted 1:10	LB A/I/X LB A/I/X	693 71	84 10
5) Vector plus ligase (background control)	10 µL	100 µL	Undiluted 1:10	LB A/I/X LB A/I/X	0 0	14 2
6) Vector plus PCR DNA	10 µL	100 µL	Undiluted 1:10	LB A/I/X LB A/I/X	252 29	22 3

Table A2.2: Transformation Results for *nlEF-2* degenerate PCR

The results in Table A2.2 are an example of a pGEM-T easy transformation results.

This shows that ~11.5% of colonies in reaction 4 are blue and the majority of colonies were white, indicating that the ligation efficiency was acceptable. The blue colonies that arose were probably due to undigested pGEM-T vectors or the absence of a 3' terminal thymidine attached to ends of the vector (which meant the vector could then religate). The background blue colonies observed in control reactions 3 and 5 indicate that the majority of blue colonies were due to vector self-ligation. The ligation of the pGEM-T Easy vector and the 2066 bp PCR product (reaction 6) resulted in ~91.5% white colonies, each a potential transformant. Reaction 2 is a transformation control and shows the efficiency of the DH5 alpha competent cells (6.29×10^5 cells/µg) to be low, but adequate for the transformation.

Appendix 3

Figure A3.1: Endonuclease non-Cutters

A) There were 104 known restriction endonucleases that did not have a digestion site within the *nlEF-2* 2066 bp degenerate PCR product and B) there were 167 that did not have a digestion site within the *nlCDC-12* 514 bp degenerate PCR products. Enzymes highlighted in bold and underlined were used in this study.

Table A3.1: Intron and Exon Lengths and locations for Figures 3.12 (A) and 4.9 (B)
All lengths are in base pairs and location refers to the site of each intron/exon in relation to the start codon. Introns and Exons are numbered from the 5' end of the gene.

A)

AatI; AccIII; Acc65I; AflIII; AgeI; AhaIII; Alw44I; AocI; AosI; ApaLI; AscI; AseI; AsnI; Asp700; Asp718; AspEI; AsuII; AviII; AvrII; ***BamHI;*** *BbrPI; BglI; BinI; BpmI; BseAI; BsePI; BsiWI; BsmI; BspEI; BspLU11I; BspMI; BssHII; Bst1107I; BstBI; BstEII; BstUI; Bsu36I; CfoI; DraI; DraIII; EagI; Eam1105I; Ecl136II; EclXI; Eco47III;* ***EcoRI;*** ***EcoRV;*** *FnuDII; FspI; GsuI; HaeII; HhaI; HinPI; HpaI; KasI; KpnI; KspI; MfeI;* ***MluI;*** *MroI; MstI; MstII; MunI; MvnI;* ***NarI;*** ***NdeI;*** *NheI; NotI; NruI; NsiI; NspBII; NspV; PacI; PinAI; PmaCI; PmeI; PmlI; Psp1406I;* ***PvuI;*** ***PvuII;*** *RsrII;* ***SacI;*** *SacII; SauI;* ***ScaI;*** *SexAI; SfiI; SfuI;* ***SmaI;*** *SnoI; SpeI; SphI; SrfI; Sse8387I; SspI; SstI; SstII; StuI; SwaI; ThaI; XbaI; XmaI; XmaIII and XmaCI.*

B)

AatI; AatII; AccI; AccIII; Acc65I; AcyI; AflII; AflIII; AgeI; AhaII; AhaIII; Alw44I; AlwNI; AocI; AosI; ApaI; ApaLI; AscI; AseI; AsnI; AspI; Asp700; Asp718; AspEI; AspHI; AsuII; AvaI; AviII; AvrII; ***BamHI;*** *BanI; BbrPI; BclI; BfaI; BfrI; BglI; BglII; BinI; BpmI; BsaI; BsaAI; BsaBI; BsaHI; BseAI; BsePI; BsiWI; BsmI; Bsp1407I; BspDI; BspEI; BspHI; BspLU11I; BspMI; BsrI; BsrFI; BssGI; BssHII; Bst1107I; BstBI; BstEII; BstXI; BstYI; Bsu36I; Cfr10I;* ***ClaI;*** *DraI; DraII; DraIII; DrdI; DsaI; EagI; Eam1105I; EarI; Ecl136II; EclXI; Eco47III; Eco57I; EcoNI; EcoO109I; EcoRI; EcoRV; FokI; FspI; GsuI; HaeII; HgiAI; HincII; HindII; HindIII; HpaI; HpaII; HphI; KasI; KpnI; KspI; MaeI; MaeII; MamI; MfeI;* ***MluI;*** *MroI; MseI; MspI; MstI; MstII; MunI; NaeI; NarI; NciI;* ***NcoI;*** *NdeI; NgoMI;* ***NheI;*** *NlaIV; NotI; NruI; NsiI; NspV; PacI; PaeR7I; PflMI; PinAI; PleI; PmaCI; PmeI; PmlI; PpuMI; Psp1406I; PstI; RcaI; RmaI; RsaI; RsrII; SacI; SacII; SalI; SauI; ScaI; SexAI; SfaNI; SfcI; SfiI; SfuI; SgrAI;* ***SmaI;*** *SnaBI; SnoI; SpeI; SphI; SrfI; Sse8387I; SspI; SspBI; SstI; SstII; StuI; SwaI; Tru9I; Tth111I; Van91I; XbaI; XcmI; XhoI; XhoII; XmaI; XmaIII and XmaCI*

A)

Organism	Exons			Introns		
	Number	Length	Location	Number	Length	Location
N. lolii (Lp19)	1	8	1-8	1	346	9-355
	2	733	356-1087	2	29	1088-1115
	3	1776	1116-2890	3	62	2891-2951
	4	18	2952-2968			
M. grisea	1	8	1-8	1	329	9-336
	2	712	337-1047	2	125	1048-1171
	3	1776	1172-2946	3	79	2947-3024
	4	114	3025-3137			
N. crassa	1	8	1-8	1	218	9-225
	2	171	226-395	2	54	396-448
	3	2228	449-2675	3	57	2676-2731
	4	18	2732-2748			
A. nidulans	1	8	1-8	1	80	9-87
	2	16	88-102	2	92	103-193
	3	58	194-150	3	56	151-205
	4	2435	206-2639	4	56	2640-2694
	5	18	2695-2711			
C. albicans	1	4	1-4	1	200	5-203
	2	2525				
S. cerevisiae	1	2529	N/A	N/A	N/A	N/A

B)

Organism	Exon			Intron		
	Number	Length	Location	Number	Length	Location
N. lolii (Lp19)	1	105	1-105	1	82	106-187
	2	163	188-350	2	68	351-418
	3	311	419-729	3	66	730-795
	4	579	796-1374			
N. crassa	1	105	1-105	1	71	106-176
	2	163	177-339	2	68	340-407
	3	311	408-718	3	57	719-775
	4	579	776-1354			
A. nidulans	1	45	1-45	1	73	46-118
	2	14	119-132	2	64	133-196
	3	43	197-239	3	60	240-299
	4	163	300-462	4	59	463-521
	5	311	522-832	5	67	833-899
	6	624	900-1523			
M. circinelloides	1	103	1-103	1	73	104-176
	2	132	177-308	2	58	309-366
	3	217	367-583	3	64	584-647
	4	118	648-765	4	76	766-841
	5	476	842-1317	5	68	1318-1385
	6	123	1386-1508			
S. pombe	1	1143	N/A	N/A	N/A	N/A
S. cerevisiae	1	1224	N/A	N/A	N/A	N/A

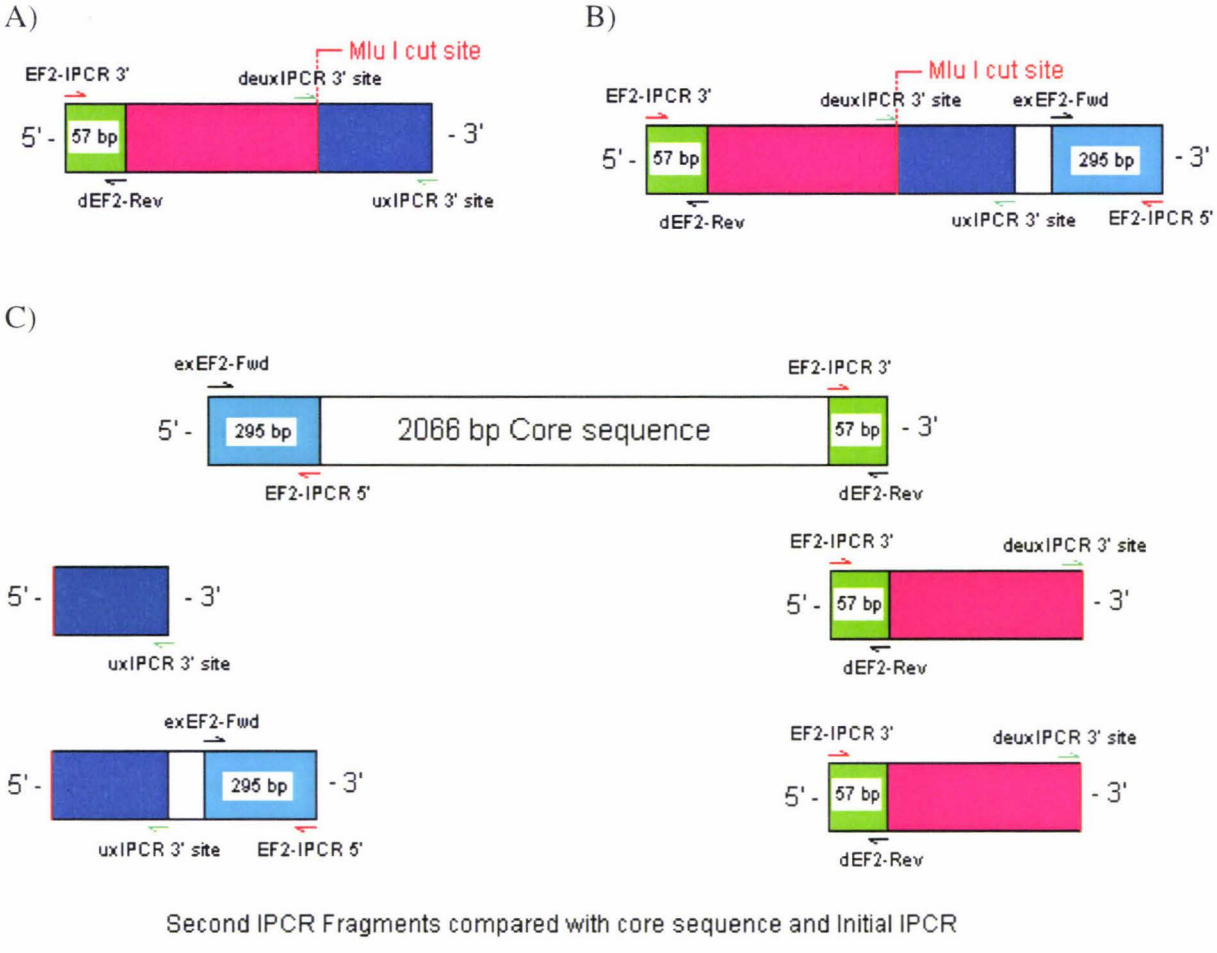


Figure A3.2: Schematic assembly of EF2-con1 and EF2-con2

Colored regions indicate identical sequences and relative location of primer sites and Mlu I sites are labeled.

A) Original IPCR product from figure 3.5. Upon division at the Mlu I site homology was found in the 5' fragment to the core region in C), however no homology was found in the 3' fragment. The core region and 5' fragment were subsequently assembled using FAS within GCG into EF2-con1 (section 2.2.11).

B) Second IPCR product from figure 3.7. Upon division at the Mlu I site homology was found in both the 5' and 3' fragments to the core region in C). Using a ClustalX on default settings the 5' end of the 3' fragment was found to be identical to the 3' fragment from A). The core region, 5' and 3' fragments were subsequently assembled using FAS within GCG into EF2-con2 (section 2.2.11)

Table A3.2: *nlEF-2* and *nlCDC-12* Codon Usage

Codon usage tables for A) the *nlEF-2* gene and B) the *nlCDC-12* gene.

A)

nlEF-2

Amino Acid	Codon	Number	/1000	Fraction
Gly	GGG	0	0	0
Gly	GGA	2	2.4	0.03
Gly	GGT	40	47.3	0.63
Gly	GGC	22	26.4	0.44
Glu	GAG	36	42.6	0.29
Glu	GAA	15	17.8	0.71
Asp	GAT	18	21.3	0.375
Asp	GAC	40	47.3	0.625
Val	GTG	3	3.6	0.04
Val	GTA	0	0	0
Val	GTT	24	28.4	0.31
Val	GTC	51	60.4	0.65
Ala	GCG	0	0	0
Ala	GCA	5	5.9	0.09
Ala	GCT	22	26	0.37
Ala	GCC	32	37.9	0.54
Arg	AGG	1	1.2	0.02
Arg	AGA	0	0	0
Ser	AGT	0	0	0
Ser	AGC	5	5.9	0.1
Lys	AAG	55	65.1	0.96
Lys	AAA	2	2.4	0.04
Asn	AAT	5	5.9	0.19
Asn	AAC	22	26	0.81
Met	ATG	24	28.4	1
Ile	ATA	0	0	0
Ile	ATT	18	21.3	0.35
Ile	ATC	33	39.1	0.65
Thr	ACG	4	4.7	0.09
Thr	ACA	0	0	0
Thr	ACT	14	16.6	0.31
Thr	ACC	27	32	0.6
Trp	TGG	7	8.3	1
End	TGA	0	0	0
Cys	TGT	0	0	0

B)

nlCDC-12

Amino Acid	Codon	Number	/1000	Fraction
Gly	GGG	0	0	0
Gly	GGA	2	5.2	0.13
Gly	GGT	5	12.9	0.33
Gly	GGC	8	20.7	0.54
Glu	GAG	25	64.6	0.71
Glu	GAA	10	25.8	0.29
Asp	GAT	7	18.1	0.29
Asp	GAC	17	42.9	0.71
Val	GTG	1	2.6	0.05
Val	GTA	0	0	0
Val	GTT	5	12.9	0.26
Val	GTC	13	33.6	0.69
Ala	GCG	2	5.2	0.06
Ala	GCA	1	2.6	0.03
Ala	GCT	12	31	0.39
Ala	GCC	16	41.3	0.52
Arg	AGG	1	2.6	0.03
Arg	AGA	3	7.8	0.1
Ser	AGT	0	0	0
Ser	AGC	6	15.5	0.32
Lys	AAG	26	67.2	0.87
Lys	AAA	4	10.3	0.13
Asn	AAT	3	7.8	0.23
Asn	AAC	10	25.8	0.77
Met	ATG	10	25.8	1
Ile	ATA	0	0	0
Ile	ATT	13	33.6	0.33
Ile	ATC	17	43.9	0.77
Thr	ACG	3	7.8	0.14
Thr	ACA	0	0	0
Thr	ACT	5	12.9	0.24
Thr	ACC	13	33.6	0.62
Trp	TGG	3	7.8	1
End	TGA	0	0	0
Cys	TGT	2	5.2	0.5

Cys	TGC	6	7.1	1	Cys	TGC	2	5.2	0.5
End	TAG	0	0	0	End	TAG	0	0	0
End	TAA	1	1.2	1	End	TAA	1	2.6	1
Tyr	TAT	2	2.4	0.1	Tyr	TAT	3	7.8	0.38
Tyr	TAC	18	21.3	0.9	Tyr	TAC	5	12.9	0.62
Leu	TTG	16	18.9	0.2	Leu	TTG	6	15.5	0.22
Leu	TTA	0	0	0	Leu	TTA	0	0	0
Phe	TTT	2	2.4	0.06	Phe	TTT	1	2.6	0.06
Phe	TTC	30	35.5	0.94	Phe	TTC	16	41.3	0.94
Ser	TCG	1	1.2	0.02	Ser	TCG	1	2.6	0.05
Ser	TCA	1	1.2	0.02	Ser	TCA	1	2.6	0.05
Ser	TCT	15	17.8	0.29	Ser	TCT	4	10.3	0.21
Ser	TCC	30	35.5	0.43	Ser	TCC	7	18.1	0.37
Arg	CGG	0	0	0	Arg	CGG	1	2.6	0.03
Arg	CGA	9	10.7	0.2	Arg	CGA	7	18.1	0.23
Arg	CGT	31	36.7	0.77	Arg	CGT	7	18.1	0.23
Arg	CGC	5	5.9	0.11	Arg	CGC	13	33.6	0.38
Gln	CAG	30	35.5	0.94	Gln	CAG	17	43.9	0.81
Gln	CAA	2	2.4	0.06	Gln	CAA	4	10.3	0.19
His	CAT	2	2.4	0.125	His	CAT	1	2.6	0.08
His	CAC	14	16.6	0.875	His	CAC	12	31	0.92
Leu	CTG	24	28.4	0.31	Leu	CTG	7	18.1	0.26
Leu	CTA	1	1.2	0.01	Leu	CTA	0	0	0
Leu	CTT	16	18.9	0.2	Leu	CTT	5	12.9	0.19
Leu	CTC	22	26	0.28	Leu	CTC	9	23.3	0.33
Pro	CCG	0	0	0	Pro	CCG	3	7.8	0.21
Pro	CCA	0	0	0	Pro	CCA	0	0	0
Pro	CCT	18	21.3	0.45	Pro	CCT	6	15.5	0.43
Pro	CCC	22	26	0.55	Pro	CCC	5	12.9	0.36

Day	Curve A (mg/mL)	Sample Size (mL)	Curve B (mg/mL)	Sample Size (mL)
2	0.18	50	0.16	50
6	0.18	50	0.21	50
7			0.36	50
8			0.56	50
9			0.84	50
10	0.19	50	1.33	50
11			1.9	25
12			2.53	15
13			3.5	15
14	0.32	50	4.5	15
15				
16	0.533	50	6.7	15
17				
18	0.796	50	7.54	10
19	1.104	25		
20	1.428	25	7.93	10
21				
22	2.224	15	8.1	10
23				
24	3.16	15	8.11	10
25	3.9	15		
26	4.39	15	8.08	10
27	4.9	15		
29	6.05	10		
31	6.44	10		
33	6.56	10		
34	6.58	10		
35	6.57	10		

Table A3.3: Dry Weight Measurements for Northern Samples

Figure A3.3: *nlEF-2* Sequence Characterization Alignment

Alignment of EF2-con2 (partial genomic DNA sequence: ¯160-3085), EF2-con5 (contig of 5' and 3' *nlEF-2* RACE products) and two RACE control products generated using primers EF2-RACEc and EF2 5' RACE with either cDNA or genomic DNA (gDNA). Stars label homologous sequence; putative intronic sequence is confirmed in grey text. As expected the control reaction using cDNA yielded a product lacking the 29 bp intronic region present in the control reaction using gDNA. Transcription start/stop sites are bold and underlined; putative translation start/stop codons are bold and double underlined.

```
EF2-con5        ----------------------------------------CTCCTCCTGCTGCCCCGCTCA
EF2-con2        CTCCCAGTATATAAAAGAAAAGTCATTTTGACCTTCACACTCCTCCTGCTGCCCCGCTCA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                                                ********************

EF2-con5        GCACTTCGTTCCCTTCTCTCTTCCACCAATCTTGACGTCGAGGACGTCAAGCCAATACCG
EF2-con2        GCACTTCGTTCCCTTCTCTCTTCCACCAATCTTGACGTCGAGGACGTCAAGCCAATACCG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        ATCTTGACGTCGAGGACGTCAAGCCAATACCGCCAAAATGGTCAA--------------
EF2-con2        ATCTTGACGTCGAGGACGTCAAGCCAATACCGCCAAAATGGTCAAGTAAGTTTCGCAGAT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                *********************************************

EF2-con5        ------------------------------------------------------------
EF2-con2        CTGCTCTGGTCGCATGAAATCGGACACGAGTGCATCTCACTTTGATTGTGCTTTCTTTTC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------

EF2-con5        ------------------------------------------------------------
EF2-con2        CCACTTCAATCGTTCAGCGAGCTGCGGGATGGGGAGACACAAAATTTTTCCCTCTCTGGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------

EF2-con5        ------------------------------------------------------------
EF2-con2        TTGATGATGTGGTCCATTTGACCCCGCCCGGAGTTCAGCTACCCTCTCAGGGCCAAAACT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------

EF2-con5        ------------------------------------------------------------
EF2-con2        CCCTGCCAAGGTTTGGGAGGATTGTTCCCAATCCTGCACCAAGTTCCGTGGCGATCAGAG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
```

```
EF2-con5        ------------------------------------------------------------
EF2-con2        AGAGGCCCAAGGGGCAAAGGAAGTTGGACAACATAACATGATGGACTTGACTGAAATTGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------

EF2-con5        --------------------------------CTTCACCATCGATGAGATTCGCCAGCTTA
EF2-con2        CAACTCTAGGAGTGTCTTCCTCGCAATGTAGCTTCACCATCGATGAGATTCGCCAGCTTA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                                               ****************************

EF2-con5        TGGACAAGCCTAGCAATGTCCGTAATATGTCCGTTATTGCCCACGTCGATCACGGCAAGT
EF2-con2        TGGACAAGCCTAGCAATGTCCGTAATATGTCCGTTATTGCCCACGTCGATCACGGCAAGT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CTACCCTGACCGACTCTCTTTTGGCCAAGGCCGGTATCATCTCCTCTGCAAAGGCTGGTG
EF2-con2        CTACCCTGACCGACTCTCTTTTGGCCAAGGCCGGTATCATCTCCTCTGCAAAGGCTGGTG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        ATGCTCGAGCAACGGATACCCGTGCAGATGAGCAGGAGCGTGGTATCACCATCAAGTCGA
EF2-con2        ATGCTCGAGCAACGGATACCCGTGCAGATGAGCAGGAGCGTGGTATCACCATCAAGTCGA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CGGCCATCTCCCTGTATGGTCGTCTTGACGACCCCGAGGACGTCAAAGACATTGTCGGCC
EF2-con2        CGGCCATCTCCCTGTATGGTCGTCTTGACGACCCCGAGGACGTCAAAGACATTGTCGGCC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        AGAAGATTGATGGCCAGGACTTCTTGATCAACTTGATTGACTCCCCCGGTCACGTTGATT
EF2-con2        AGAAGATTGATGGCCAGGACTTCTTGATCAACTTGATTGACTCCCCCGGTCACGTTGATT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TCTCTTCTGAAGTTACTGCCGCTCTCCGTGTCACTGACGGTGCTCTCGTCGTCGTCGACA
EF2-con2        TCTCTTCTGAAGTTACTGCCGCTCTCCGTGTCACTGACGGTGCTCTCGTCGTCGTCGACA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CCGTCGAAGGCGTGTGCGTCCAGACCGAGACTGTGCTCCGACAGGCTCTTGGTGAGCGTA
EF2-con2        CCGTCGAAGGCGTGTGCGTCCAGACCGAGACTGTGCTCCGACAGGCTCTTGGTGAGCGTA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TCAAGCCCGTCATTATTATCAACAAGGTCGACCGTGCCCTTCTCGAACTTCAGGTCTCCA
EF2-con2        TCAAGCCCGTCATTATTATCAACAAGGTCGACCGTGCCCTTCTCGAACTTCAGGTCTCCA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************
```

```
EF2-con5        AGGAGGATCTGTACCAATCTTTCTCTCGAACCATCGAGTCCGTCAACGTCATCATCTCCA
EF2-con2        AGGAGGATCTGTACCAATCTTTCTCTCGAACCATCGAGTCCGTCAACGTCATCATCTCCA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CCTACCTGGACAAGGCTCTGGGTGACGTTCAGGTCTACCCTGACAAGGGCACCATTGCTT
EF2-con2        CCTACCTGGACAAGGCTCTGGGTGACGTTCAGGTCTACCCTGACAAGGGCACCATTGCTT
Control_cDNA    ------------------------------------CCCTGACAAGGGCACCATTGCTT
Control_gDNA    ------------------------------------CCCTGACAAGGGCACCATTGCTT
                ************************************************************

EF2-con5        TTGGTTCCGGTCTGCACGGCTGGGCCTTCACTATTCGTCAGTTCGCTGTTCGATACGCCA
EF2-con2        TTGGTTCCGGTCTGCACGGCTGGGCCTTCACTATTCGTCAGTTCGCTGTTCGATACGCCA
Control_cDNA    TTGGTTCCGGTCTGCACGGCTGGGCCTTCACTATTCGTCAGTTCGCTGTTCGATACGCCA
Control_gDNA    TTGGTTCCGGTCTGCACGGCTGGGCCTTCACTATTCGTCAGTTCGCTGTTCGATACGCCA
                ************************************************************

EF2-con5        AGAAGTTTGGTGTCGACAAGAACAAGATGATGGAGCGTCTCTGGG--------------
EF2-con2        AGAAGTTTGGTGTCGACAAGAACAAGATGATGGAGCGTCTCTGGGTACGTATTTGCTAAC
Control_cDNA    AGAAGTTTGGTGTCGACAAGAACAAGATGATGGAGCGTCTCTGGG--------------
Control_gDNA    AGAAGTTTGGTGTCGACAAGAACAAGATGATGGAGCGTCTCTGGGTACGTATTTGCTAAC
                ************************************************************

EF2-con5        --------------GTGACAACTACTTCAACCCTCACACCAAGAAGTGGACCAAAACGGC
EF2-con2        TGGCCCATTCTAGGGTGACAACTACTTCAACCCTCACACCAAGAAGTGGACCAAAACGGC
Control_cDNA    --------------GTGACAACTACTTCAACCCTCACACCAAGAAGTGGACCAAAACGGC
Control_gDNA    TGGCCCATTCTAGGGTGACAACTACTTCAACCCTCACACCAAGAAGTGGACCAAAACGGC
                ************************************************************

EF2-con5        ACGTTACGAGGGTAAGCAGCTCGAGCGTGCTTTCAATCAGTTCATCTTGGACCCCATTTT
EF2-con2        ACGTTACGAGGGTAAGCAGCTCGAGCGTGCTTTCAATCAGTTCATCTTGGACCCCATTTT
Control_cDNA    ACGTTACGAGGGTAAGCAGCTCGAGCGTGCTTTCAATCAGTTCATCTTGGACCCCATTTT
Control_gDNA    ACGTTACGAGGGTAAGCAGCTCGAGCGTGCTTTCAATCAGTTCATCTTGGACCCCATTTT
                ************************************************************

EF2-con5        CAAGATCTTCAATGCCGTCATGAACTTCAAGAACGACGAGATCACCAACCTCCTCGAGAA
EF2-con2        CAAGATCTTCAATGCCGTCATGAACTTCAAGAACGACGAGATCACCAACCTCCTCGAGAA
Control_cDNA    CAAGATCTTCAATGCCGTCATGAACTTCAAGAACGACGAGATCACCAACCTCC-------
Control_gDNA    CAAGATCTTCAATGCCGTCATGAACTTCAAGAACGACGAGATCACCAACCTCC-------
                ************************************************************

EF2-con5        GCTGTCCCTTAAGCTTAGCGTTGATGACCGTGCCAAGGAGGGCAAGCAGCTACTCAAGGT
EF2-con2        GCTGTCCCTTAAGCTTAGCGTTGATGACCGTGCCAAGGAGGGCAAGCAGCTACTCAAGGT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TGTCATGCGAACTTTCTTGCCTGCAGCCGACTCCTTGTTGGAGATGATGATCCTTCACCT
EF2-con2        TGTCATGCGAACTTTCTTGCCTGCAGCCGACTCCTTGTTGGAGATGATGATCCTTCACCT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TCCTTCTCCTGTCACTGCTCAGACCTACCGTGTCGAGACTTTGTACGAAGGTCCCATGGA
EF2-con2        TCCTTCTCCTGTCACTGCTCAGACCTACCGTGTCGAGACTTTGTACGAAGGTCCCATGGA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************
```

```
EF2-con5        CGACGAAGCCGCCATTGGTATCCGTGACTGCGACCCCAAGGGTCCTCTCATGTTGTACGT
EF2-con2        CGACGAAGCCGCCATTGGTATCCGTGACTGCGACCCCAAGGGTCCTCTCATGTTGTACGT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CTCCAAGATGGTTCCCACCCTTGACAAGGGCCGATTCTTCGCCATCGGTCGTGTCTTCTC
EF2-con2        CTCCAAGATGGTTCCCACCCTTGACAAGGGCCGATTCTTCGCCATCGGTCGTGTCTTCTC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CGGTACTGTCCGCTCTGGTCTCAAGGTCTGCATCAGGGGCCCCAATTACACCCCTGGCAA
EF2-con2        CGGTACTGTCCGCTCTGGTCTCAAGGTCTGCATCAGGGGCCCCAATTACACCCCTGGCAA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        GAAGGAAGATCTCTTCATCAAGGCTATCCAGCGAACCGTTCTCATGATGGGCGGCAAGGT
EF2-con2        GAAGGAAGATCTCTTCATCAAGGCTATCCAGCGAACCGTTCTCATGATGGGCGGCAAGGT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CGAGCCCATCGATGACATGCCTGCCGGCAACATTGTGGGTCTTGTTGGTATCGATCAGTT
EF2-con2        CGAGCCCATCGATGACATGCCTGCCGGCAACATTGTGGGTCTTGTTGGTATCGATCAGTT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CTTGCTCAAGTCTGGTACTCTCACCACCTCCGAAACTGCCCACAACCTCAAGGTCATGAA
EF2-con2        CTTGCTCAAGTCTGGTACTCTCACCACCTCCGAAACTGCCCACAACCTCAAGGTCATGAA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        GTTCTCCGTCTCCCCCGTCGTCCAGCGTTCCGTCCAGGTCAAGAACGCCCAGGATCTTCC
EF2-con2        GTTCTCCGTCTCCCCCGTCGTCCAGCGTTCCGTCCAGGTCAAGAACGCCCAGGATCTTCC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CAAGCTGGTCGAGGGACTCAAGCGTCTGTCCAAGTCTGACCCTTGCGTTCTGACCATGCA
EF2-con2        CAAGCTGGTCGAGGGACTCAAGCGTCTGTCCAAGTCTGACCCTTGCGTTCTGACCATGCA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CTCCGAATCTGGTGAGCATATTGTCGCCGGTGCCGGTGAGCTGCATCTCGAAATTTGCTT
EF2-con2        CTCCGAATCTGGTGAGCATATTGTCGCCGGTGCCGGTGAGCTGCATCTCGAAATTTGCTT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        GAAGGATCTCGAGGAAGACCACGCTGGTGTTCCCCTCATCATCTCCGACCCTGTCGTCCA
EF2-con2        GAAGGATCTCGAGGAAGACCACGCTGGTGTTCCCCTCATCATCTCCGACCCTGTCGTCCA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************
```

```
EF2-con5        GTACCGTGAGACCGTCCAGAGCAAGTCCAGCATGACCGCCCTGTCCAAGTCCCCCAACAA
EF2-con2        GTACCGTGAGACCGTCCAGAGCAAGTCCAGCATGACCGCCCTGTCCAAGTCCCCCAACAA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        GCACAACCGTCTGTACATGGTTGCCGAGCCCATGGAGGAGGAACTGTCCCTGGCTATCGA
EF2-con2        GCACAACCGTCTGTACATGGTTGCCGAGCCCATGGAGGAGGAACTGTCCCTGGCTATCGA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        GAGCGGCAAGGTTTCTGCCCGTGACGATTTCAAGGCTCGTGCCCGTATCCTGGCCGACGA
EF2-con2        GAGCGGCAAGGTTTCTGCCCGTGACGATTTCAAGGCTCGTGCCCGTATCCTGGCCGACGA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CTTCGGCTGGGACGTCACTGACGCCCGTAAGATTTGGTGCTTCGGCCCTGACGGTACTGG
EF2-con2        CTTCGGCTGGGACGTCACTGACGCCCGTAAGATTTGGTGCTTCGGCCCTGACGGTACTGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TGCCAACTTGTTGGTTGACCAGACCAAGGCTGTTCAGTACCTGAACGAAATCAAGGACTC
EF2-con2        TGCCAACTTGTTGGTTGACCAGACCAAGGCTGTTCAGTACCTGAACGAAATCAAGGACTC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CGTTGTCTCCGGTTTCCAATGGGCTTCCCGTGAGGGTCCCGTTGCTGAGGAGCCCATGCG
EF2-con2        CGTTGTCTCCGGTTTCCAATGGGCTTCCCGTGAGGGTCCCGTTGCTGAGGAGCCCATGCG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        ATCCATCCGCTTCAACGTCCTCGATGTTACCCTGCACGCTGATGCTATTCACCGTGGTGG
EF2-con2        ATCCATCCGCTTCAACGTCCTCGATGTTACCCTGCACGCTGATGCTATTCACCGTGGTGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CGGTCAGATCATTCCCACTGCCCGTCGTGTCCTGTACGCCTCCGCTCTGATGGCCGAGCC
EF2-con2        CGGTCAGATCATTCCCACTGCCCGTCGTGTCCTGTACGCCTCCGCTCTGATGGCCGAGCC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        CGCTTTGCTTGAGCCCGTCTACCTGGTCGAAATCCAGGTTCCCGAGCAGGCCATGGGTGG
EF2-con2        CGCTTTGCTTGAGCCCGTCTACCTGGTCGAAATCCAGGTTCCCGAGCAGGCCATGGGTGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TGTCTACGGTGTTCTTACGCGCCGTCGTGGTCACGTCTTCAACGAGGAGCAGCGTCCTGG
EF2-con2        TGTCTACGGTGTTCTTACGCGCCGTCGTGGTCACGTCTTCAACGAGGAGCAGCGTCCTGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************
```

```
EF2-con5        TACTCCTCTCTTCAACATCAAGGCTTACCTGCCTGTCTTGGAGTCTTTCGGCTTCAACGG
EF2-con2        TACTCCTCTCTTCAACATCAAGGCTTACCTGCCTGTCTTGGAGTCTTTCGGCTTCAACGG
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        TGACCTGCGACAGGCCACCTCAGGCCAGGCCTTCCCTCAGTCCGTCTTCGACCACTGGCA
EF2-con2        TGACCTGCGACAGGCCACCTCAGGCCAGGCCTTCCCTCAGTCCGTCTTCGACCACTGGCA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        GGTCTTGCCCGGTGGCTCCCCTCTTGATCCCACCTCTAAGGTCGGAGGCATTGTCACCGA
EF2-con2        GGTCTTGCCCGGTGGCTCCCCTCTTGATCCCACCTCTAAGGTCGGAGGCATTGTCACCGA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        AATGCGTAAGCGCAAGGGTATCAAGGTTGAAGTTCCTGGCGTTGAGAAC-----------
EF2-con2        AATGCGTAAGCGCAAGGGTATCAAGGTTGAAGTTCCTGGCGTTGAGAACGTAAGTTGACT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                *************************************************

EF2-con5        ----------------------------------------------------TACTATGAC
EF2-con2        GCATTTTCACTTTCCTAGCTCACATCGAACATGTGCTAACTGCCTACACAGTACTATGAC
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                                                                 *********

EF2-con5        AAGCTGTAAATAGCTATTCCAAATGCCGATAGAATCATGTGTAGTGCCAACAGAATGGAT
EF2-con2        AAGCTGTAAATAGCTATTCCAAATGCCGATAGAATCATGTGTAGTGCCAACAGAATGGAT
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ************************************************************

EF2-con5        AAAAAAAATCGATAGGCTGCTAGAATAAACTCTCCTCGCGCAAAAAAAAAAAAAAAAAAA
EF2-con2        AAAAAAAATCGATAGGCTGCTAGAATAAACTCTCCTCGCGCACACCCACGAATGGGCACA
Control_cDNA    ------------------------------------------------------------
Control_gDNA    ------------------------------------------------------------
                ****************************************** *   *  **     * *
```

Figure A3.4: *nlCDC-12* Sequence Characterization Alignment

Alignment of Sep-con1 (partial genomic DNA sequence: ¯239-1500), Sep-con4 (contig of 5' and 3' *nlCDC-12* RACE products) and two RACE control products generated using primers Sep-RACEc and Sep 5' RACE with either cDNA or genomic DNA (gDNA). Stars label homologous sequence; putative intronic sequence is confirmed in grey text. As expected the control reaction using cDNA yielded a product lacking the 68 bp intronic region present in the control reaction using gDNA. Transcription start/stop sites are bold and underlined; putative translation start/stop codons are bold and double underlined.

```
control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        ------------------------------------------------------------
Sep-con1        TGGTGGAAGCCAATCAAATTTCGCGCACCCGCAATCCATCCCAACTGTCAACATATCAGC


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        CTTCGAGTTGAGCATTTCTCCCTGGCATATTCTCAACCCCCAACTGCTCGCTCGACATAA
Sep-con1        CTTCGAGTTGAGCATTTCTCCCTGGCATATTCTCAACCCCCAACTGCTCGCTCGACATAA
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        GTTTCGCGGGCCTGTCTTGTTTTGCATAGCGTTGGGAAAATTCTCGATTACCTGAAGGAC
Sep-con1        GTTTCGCGGGCCTGTCTTGTTTTGCATAGCGTTGGGAAAATTCTCGATTACCTGAAGGAC
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        AAACCGTTCACTTAGTCTATTTCACATCACCACACCGGTCTTCGACACCGCTGCCCGCTA
Sep-con1        AAACCGTTCACTTAGTCTATTTCACATCACCACACCGGTCTTCGACACCGCTGCCCGCTA
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        TGGCTTCCGCAGCGACCGAGAGCGCTTCACCGATTGGCATTGCCAATCTGCCCAACCAGC
Sep-con1        TGGCTTCCGCAGCGACCGAGAGCGCTTCACCGATTGGCATTGCCAATCTGCCCAACCAGC
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        GACACAAGATTGTCGCCAAGAGGGGTGCTAGCTTCACTATTATG----------------
Sep-con1        GACACAAGATTGTCGCCAAGAGGGGTGCTAGCTTCACTATTATGGTATATTTTCATCTTT
                ******************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        ------------------------------------------------------------
Sep-con1        GCGGTGAAAACACTGAACTTGTCATTATTCATGTGAAGAATTAAAGACTGACTTGAATCG
```

```
control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        ------GTTGCTGGCGAATCTGGCCTCGGCAAGACCACCTTTATCAACACCCTGTTCTCC
Sep-con1        CTTCAGGTTGCTGGCGAATCTGGCCTCGGCAAGACCACCTTTATCAACACCCTGTTCTCC
                      ************************************************

control_cDNA    ------------------------CACAAGCGTCGACACCAGAAGCAGGTCGACAAGACT
control_gDNA    ------------------------CACAAGCGTCGACACCAGAAGCAGGTCGACAAGACT
Sep-con4        ACTACGATTAAGAACTATGCCGACCACAAGCGTCGACACCAGAAGCAGGTCGACAAGACT
Sep-con1        ACTACGATTAAGAACTATGCCGACCACAAGCGTCGACACCAGAAGCAGGTCGACAAGACT
                ************************************************************

control_cDNA    GTCGAGATTGAGATCACCAAGGCTGAGCTTGAGGAGAAGTTCTTCAAGG-----------
control_gDNA    GTCGAGATTGAGATCACCAAGGCTGAGCTTGAGGAGAAGTTCTTCAAGGGTAAGCATCAG
Sep-con4        GTCGAGATTGAGATCACCAAGGCTGAGCTTGAGGAGAAGTTCTTCAAGG-----------
Sep-con1        GTCGAGATTGAGATCACCAAGGCTGAGCTTGAGGAGAAGTTCTTCAAGGGTAAGCATCAG
                ************************************************************

control_cDNA    ------------------------------------------------------------TCC
control_gDNA    CATCCAGTCGTCTTGCACTCTACTCCTACCCTGGCACAACTAATCGATCCGTTCTAGTCC
Sep-con4        ------------------------------------------------------------TCC
Sep-con1        CATCCAGTCGTCTTGCACTCTACTCCTACCCTGGCACAACTAATCGATCCGTTCTAGTCC
                ************************************************************

control_cDNA    GCTTGACCGTCATTGACACCCCTGGATTCGGCGACTACGTCAACAATCGTGACTCCTGGA
control_gDNA    GCTTGACCGTCATTGACACCCCTGGATTCGGCGACTACGTCAACAATCGTGACTCCTGGA
Sep-con4        GCTTGACCGTCATTGACACCCCTGGATTCGGCGACTACGTCAACAATCGTGACTCCTGGA
Sep-con1        GCTTGACCGTCATTGACACCCCTGGATTCGGCGACTACGTCAACAATCGTGACTCCTGGA
                ************************************************************

control_cDNA    TGCCCATCATCGAGTTCCTCGACGACCAGCACGAATCCTACATGCTTCAGGAGCAGCAGC
control_gDNA    TGCCCATCATCGAGTTCCTCGACGACCAGCACGAATCCTACATGCTTCAGGAGCAGCAGC
Sep-con4        TGCCCATCATCGAGTTCCTCGACGACCAGCACGAATCCTACATGCTTCAGGAGCAGCAGC
Sep-con1        TGCCCATCATCGAGTTCCTCGACGACCAGCACGAATCCTACATGCTTCAGGAGCAGCAGC
                ************************************************************

control_cDNA    CTCGCCGTCAGGACAAAATCGACCTTCGTGTCCACGCCTGCCTCTACTTCATCCGCCCCA
control_gDNA    CTCGCCGTCAGGACAAAATCGACCTTCGTGTCCACGCCTGCCTCTACTTCATCCGCCCCA
Sep-con4        CTCGCCGTCAGGACAAAATCGACCTTCGTGTCCACGCCTGCCTCTACTTCATCCGCCCCA
Sep-con1        CTCGCCGTCAGGACAAAATCGACCTTCGTGTCCACGCCTGCCTCTACTTCATCCGCCCCA
                ************************************************************

control_cDNA    CCGGCCACACCTTGAAGCCTCTCGATATCGAG---------------------------
control_gDNA    CCGGCCACACCTTGAAGCCTCTCGATATCGAG---------------------------
Sep-con4        CCGGCCACACCTTGAAGCCTCTCGATATCGAGGTCATGAAGAGACTCTGCTCTCGCGTCA
Sep-con1        CCGGCCACACCTTGAAGCCTCTCGATATCGAGGTCATGAAGAGACTCTGCTCTCGCGTCA
                ************************************************************

control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        ACTTGATTCCTGTTATTGCCAAGGCTGATACCTTGAGCCCTGCTGATTTGGCCAAGTTCA
Sep-con1        ACTTGATTCCTGTTATTGCCAAGGCTGATACCTTGAGCCCTGCTGATTTGGCCAAGTTCA
                ************************************************************

control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        AGCAAAGA---------------------------------------------------
Sep-con1        AGCAAAGAGTAAGTCGCGCGGATGGGACAGCTGCCTCGGTCCTTCAATGTGTTCTTTGAC
                ********
```

```
control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        --------------ATCATCTCCGTTATCGAAGCCCAGAACATCAAAATCTACCAGCCCC
Sep-con1        GCTAACTTGTCCAGATCATCTCCGTTATCGAAGCCCAGAACATCAAAATCTACCAGCCCC
                              ********************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        CGATCGAGGAGGACGATGAAGCTGCCGCTCAGCACGCCCGAAGCCTTATGGCCGCCATGC
Sep-con1        CGATCGAGGAGGACGATGAAGCTGCCGCTCAGCACGCCCGAAGCCTTATGGCCGCCATGC
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        CCTTCGCCGTCATTGGTTCTGAGAAGGATGTCAAAGCCAGCGATGGTCGCATTGTCAAGG
Sep-con1        CCTTCGCCGTCATTGGTTCTGAGAAGGATGTCAAAGCCAGCGATGGTCGCATTGTCAAGG
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        GCCGCCAATACTCCTGGGGTGTTGCCGAGGTGGAGAACGAAGACCACTGTGACTTCAAGA
Sep-con1        GCCGCCAATACTCCTGGGGTGTTGCCGAGGTGGAGAACGAAGACCACTGTGACTTCAAGA
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        AGCTGCGATCTATCCTGATCCGAACTCACATGCTGGACCTTATTCACACCACCGAGGAGC
Sep-con1        AGCTGCGATCTATCCTGATCCGAACTCACATGCTGGACCTTATTCACACCACCGAGGAGC
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        TGCACTATGAGGCTTATCGCGCTCAGCAAATGGAGACGCGCAAGTTCGGCGAAGCCCGTC
Sep-con1        TGCACTATGAGGCTTATCGCGCTCAGCAAATGGAGACGCGCAAGTTCGGCGAAGCCCGTC
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        CTCGCAAGTTGGACAACCCGAAATTCAAGGAAGATGAAGAGGCGCTCCGGAAGCGATTCA
Sep-con1        CTCGCAAGTTGGACAACCCGAAATTCAAGGAAGATGAAGAGGCGCTCCGGAAGCGATTCA
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        CCGACCAGGTCAAGATCGAGGAGGAGCGTTTCTTCCGACAGTGGAGCAGAGAATTCATTT
Sep-con1        CCGACCAGGTCAAGATCGAGGAGGAGCGTTTCTTCCGACAGTGGAGCAGAGAATTCATTT
                ************************************************************


control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        CCCAGCGCGACCGTCTCAACAAGGACCTCGAGCAGACTCACGCTAACATCAAGTCGCTCG
Sep-con1        CCCAGCGCGACCGTCTCAACAAGGACCTCGAGCAGACTCACGCTAACATCAAGTCGCTCG
                ************************************************************
control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        AGACGGAACTGGAGCAAATGCAGGGAAATGCCGTTCGCTGTCATGGTCGCCGCTAAGCGT
Sep-con1        AGACGGAACTGGAGCAAATGCAGGGAAATGCCGTTCGCTGTCATGGTCGCCGCTAAGCGT
                ************************************************************
```

```
control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        TGTATCGATTCTTTATATCTGGATCAATTAAAAAATGTTGATAGTTTTATAATATACATC
Sep-con1        TGTATCGATTCTTTATATCTGGATCAATTAAAAAATGTTGATAGTTTTATAATATACATC
                ************************************************************

control_cDNA    ------------------------------------------------------------
control_gDNA    ------------------------------------------------------------
Sep-con4        TGCTAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Sep-con1        TGCTAACATATATACCATTTTTTTCCGATAGCTACTAGAAAAAAAGCATAGACATAAAAA
                ******** * * *  *          * *   *  * *******  * * * * *****
```

A3.1 Miscellaneous Probes

β-tubulin2 Gene Probe (Northern Blots- 351 bp)

GAGAAAATGCGTGAGATTGTTCATCTTCAAACCGGTCAGTGCGGTAACCAAATTGGTGCTGCTTTCTGGCAGACCA
TCTCTGGCGAGCACGGCCTCGACAGCAATGGTGTGTACAATGGTACCTCCGAGCTCCAGCTCGAGCGTATGAGTGT
CTACTTCAACGAGGCTTCTGGCAACAAGTATGTTCCTCGCGCTGTCCTCGTCGATCTCGAGCCTGGTACCATGGAT
GCAGTCCGTGCCGGTCCCTTCGGTCAGCTTTTCCGTCCTGACAACTTCGTCTTCGGTCAGTCTGGTGCTGGCAACA
ACTGGGCCAAGGGTCACTACACTGAGGGTGCTGAGCTGGTTGACCA

A. nidulans SACC Gene Probe (264 bp)

GGCTCTGCGCCGTCACCATTCCTCGCTGCGACGACTACTCCAGCACATCCAATAGGACTGCTGTCATGGTGCGTAA
TGCGGCGCAACCTTTCCCTAATGGCACCGAGATCACAAACCAGACCCTCCGCGATAGCCCGATCACCAATCGTCCT
AGAAACAGCGGGCTGATTGATACGGAAATCAACCCCGGTCCATATAAGGAAGTGCTTCCCAATGTCAGTTTCTGCC
ATAACCTCGTGCGAAGCTGTCCTATGTCACTTGGGT

REFERENCES

Al-Samarrai, T.H., and Schmid, J. (2000) A simple method for extraction of fungal genomic DNA. *Letters in Applied Microbiology* **30**: 53-56.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.

Andrews, B.J., and Herskowitz, I. (1989) Identification of a DNA binding factor involved in cell-cycle control of the yeast HO gene. *Cell* **57**: 21-29.

Andrianopoulos, A., and Timberlake, W.E. (1994) The *Aspergillus nidulans* abaA gene encodes a transcriptional activator that acts as a genetic switch to control development. *Molecular and Cellular Biology* **14**: 2503-2515.

Arndt, K., and Fink, G.R. (1986) GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5' TGACTC 3' sequences. *Proceeding of the National Academy of Sciences (USA)* **83**: 8516-8520.

Atkinson, K.D., Jensen, B., Kolat, A.I., Storm, E.M., Henry, S.A., and Fogel, S. (1980) Yeast mutants auxotrophic for choline or ethanolamine. *Journal of Bacteriology* **141**: 558-564.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K.e. (1997) *Current Protocols in Molecular Biology*. New York: John Wiley and Sons, Inc.

Ball, O.J.P., Prestidge, R.A., and Spronsen, J.M. (1995) Interrelationships between *Acremonium lolii*, peramine and lolitrem B in perennial ryegrass. *Applied and Enviromental Microbiology* **61**: 1527-1533.

Ball, O.J.P., Miles, C.O., and Prestidge, R.A. (1997a) Ergopeptide alkaloids and *Neotyphodium lolli*-mediated resistance in perennial ryegrass agains adult *Heteronynchus arator* (Coleoptera: Scarabaeidae). *Journal of Economic Entomology* **90**: 1382-1391.

Ball, O.J.P., Barker, G.M., Prestidge, R.A., and Lauren, D.R. (1997b) Distribution and accumulation of the alkaloidperamine *Neotyphodium lolli* in infected perennial ryegrass. *Journal of Chemical Ecology* **23**: 1419-1434.

Ballance, D.J. (1986) Sequences important for gene expression in filamentous fungi. *Yeast* **2**: 229-236.

Bardwell, L., Cook, J.G., Zhu-Shimoni, J.X., Voora, D., and Thorner, J. (1998) Differential regulation of transcription: repression by unactivated mitogen-activated protein kinase Kss1 requires the Dig1 and Dig2 proteins. *Proceedings of the National Academy of Sciences (USA)* **95**: 15400-15405.

Barker, D.G., White, J.H., and Johnson, L.H. (1987) Molecular characterization of the DNA ligase gene, CDC17, from the fission yeast *Schizosaccharomyces pombe*. *European Journal of Biochemistry* **162**: 659-667.

Barker, D.J., Hume, D.E., and Quigley, P.E. (1997) Negligible physiological responses to water deficit of endophyte infected and uninfected ryegrass. In *Neotyphodium/grass Interactions*. Bacon, C.W., and Hill, N.S. (eds). New York and London: Plenum Press, pp. 137-139.

Bartnicki-Garcia, S., Ruiz-Herrera, J., and Bracker, C.E. (1979) *Chitosomes and Chitin synthesis*. Cambridge: Cambridge University Press.

Bartnicki-Garcia, S., Hegert, F., and Gierz, G. (1990) A novel computor model for generating cell shape: application to fungal morphogenesis. In *Biochemistry of Cell Walls and Membranes in Fungi*. (ed. P. J. Kuhn, A.P.J.T., M. J. Jung, M. W. Goosey and L. G. Copping) (ed). Berlin: Springer-Verlag, pp. pp. 43-60.

Binks, P.R., Robson, G.D., Goosey, M.W., and Trinci, A.P. (1991) Relationships between phosphatidylcholine content, chitin synthesis, growth, and morphology of *Aspergillus nidulans* choC. *FEMS Microbiology Letters* **67**: 159-164.

Birren, B., Nusbaum, C., and Aboluelleil, A. (2003). Direct Submission; GenBank

Birse, C.E., Irwin, M.Y., Fonzi, W.A., and Sypherd, P.S. (1993) Cloning and characterization of ECE1, a gene expressed in association with cell elongation of the dimorphic pathogen *Candida albicans*. *Infective Immunology* **61**: 3648-3655.

Blount, P., Sukharev, S.I., Moe, P.C., Nagle, S.K., and Kung, C. (1996) Towards an understanding of the structural and functional properties of MscL, a mechanosensitive channel in bacteria. *Biological Cell* **87**: 1-8.

Bourne, H.R., Sanders, D.A., and McCormick, F. (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**: 117-127.

Boyce, K.J., Hynes, M.J., and Andrianopoulos, A. (2001) The CDC42 homolog of the dimorphic fungus *Penicillium marneffei* is required for correct cell polarization during growth but not development. *Journal of Bacteriology* **183**: 3447-3457.

Braun, B.R., and Johnson, A.D. (1997) Control of filament formation in *Candida albicans* by the transcriptional repressor TUP1. *Science* **277**: 105-109.

Braun, B.R., and Johnson, A.D. (2000) TUP1, CPH1 and EFG1 make independent contributions to filamentation in *Candida albicans*. *Genetics* **155**: 57-67.

Braun, B.R., Head, W.S., Wang, M.X., and Johnson, A.D. (2000) Identification and characterization of TUP1-regulated genes in *Candida albicans*. *Genetics* **156**: 31-44.

Bruchez, J.J.P., Eberle, J., and Russo, V.A.E. (1993) Regulatory sequences in the transcription of *Neurospora crassa* genes: CAAT box, TATA box, introns, poly (A) tail formation sequences. *Fungal Genetics Newsletter* **40**: 89-96.

Bryant, M. (2003) Cloning and Characterization of Two Subtilisin-Like Protease Genes from *Neotyphodium lolii*. In *Molecular Biosciences*. Palmerston North: Massey University.

Bucher, P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoters. *Journal of Molecular Biology* **212**: 563-578.

Byrd, A.D., Schardl, C.L., Songlin, P.J., Mogen, K.L., and Siegel, M.R. (1990) The beta-tubulin gene of *Epichloe typhina* from perennial ryegrass (*Lolium perenne*). *Current Genetics* **18**: 347-354.

Calera, J.A., Zhao, X.J., and Calderone, R. (2000) Defective hyphal development and avirulence caused by a deletion of the SSK1 response regulator gene in *Candida albicans*. *Infective Immunology* **68**: 518-525.

Carrol, C.W., Altman, R., Schieltz, D., Yates, J.R., and Kellogg, D.E. (1998) The septins are required for the mitosis-specific activation of the Gin4 kinase. *Journal of Biological Chemistry* **143**: 709-717.

Casamayor, A., and Snyder, M. (2003) Molecular dissection of a yeast septin: distinct domains are required for septin interaction, localization and function. *Molecular and Cellular Biology* **23**: 2762-2777.

Chang, F. (2004) Personal Communication. In. McMillan, D. (ed). Dunedin.

Chang, F., Drubin, D., and Nurse, P. (1997) Cdc12p, a prtein required for cytokinesis in fission yeast, is a component of the cell division ring and interacts with profilin. *The Journal of Cell Biology* **137**: 169-182.

Chen, W., and Struhl, K. (1985) Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. *European Molecular Biology Organization Journal* **4**: 3273-3280.

Christensen, M.J. (1995) Variation in the ability of Acremonium endophytes of *Lolium perenne*, *Festuca arundinacea* and *F. pratensis* to form compatible associations in the grasses. *Mycological Research* **99**: 466-470.

Christensen, M.J. (2000) Growth of fungal endophytes in host grasses. a pictorial review. In. Palmerston North: AgResearch Grasslands, p. 12.

Christensen, M.J., and Schmid, J. (1999) Ryegrass endophyte: host/fungus interaction. In *Grassland Research and Practice Series*, pp. 101-106.

Christensen, M.J., and Bennett, R.J. (2004a) Tall Fescue-an Endophyte Symbiosis. In. Palmerston North,: AgResearch Limited, pp. 1-11.

Christensen, M.J., and Bennett, R.J. (2004b) Fungal endophytes: a challenge to fungal tip growth dogma. In. McMillan, D. (ed). Palmerston North.

Christensen, M.J., Bennett, R.J., and Schmid, J. (2001) Growth of *Epichloe/Neotyphodium* and p-endophytes in leaves of *Lolium* and *Festuca* grasses. *Mycological Research* **105**: 1-13.

Christensen, M.J., Leuchtmann, A., Rowan, D.D., and Tapper, B.A. (1993) Taxonomy of *Acremonium* endophytes of tall fescue (*Festuca arundinacea*), meadow fescue (*F. pratensis*) and perennial ryegrass (*Lolium perenne*). *Mycological Research* **97**: 1083-1092.

Christensen, M.J.O.J.-P., Ball, O.J.-P., Bennett, R.J., and Schardl, C.L. (1997) Fungal and host genotype effects on compatibility and vascular colonisation by *Epichloë festucae*. *Mycological Research* **101**: 493-501.

Clay, K. (1990) Fungal Endophytes of Grasses. *Annual Review of Ecology and Systematics* **21**: 275-297.

Cognetti, D., Davis, D., and Sturtevant, J. (2002) The *Candida albicans* 14-3-3 gene, BMH1, is essential for growth. *Yeast* **19**: 55-67.

Collett, M.A., Bradshaw, R.E., and Scott, D.B. (1995) A mutualistic fungal symbiont of perennial ryegrass contains two different pyr4 genes, both expressing orotidine-5'-monophosphate decarboxylase. *Gene* **158**: 31-39.

Dieffenbach, C.W., Lowe, T.M.J., and Dveksler, G.S. (1995) *General Concepts for PCR Primer Design, in PCR Primer, A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.

Dieterich, C., Schandar, M., Noll, M., Johannes, F.J., Brunner, H., Graeve, T., and Rupp, S. (2002) In vitro reconstructed human epithelia reveal contributions of *Candida albicans* EFG1 and CPH1 to adhesion and invasion. *Microbiology* **148**: 497-506.

Dobson, J.M. (1997) Molecular characterisation of the HMG CoA reductase gene from *Neotyphodium lolii*. In *Institute of Molecular Biosciences*. Palmerston North: Massey University, p. 100.

Dyer, D.C. (1993) Evidence that ergovaline acts on serotonin receptors. *Life Sciences* **53**: 223-228.

Dymock, J.J., Rowan, D.D., and McGee, I.R. (1989) Effects of endophyte-produced mycotoxins on Argentine stem weevil and cutworm, *Graphania mutans*. In *Proceedings of the 5th Australasian Conference on Grassland Invertebrate Ecology*, pp. 35-43.

Edelmann, S.E., and Staben, C. (1994) A statistical analysis of sequence features within genes from *Neurospora crassa*. *Experimental Mycology* **18**: 70-81.

Elder, D. (1992) Consensus sequences at transcriptional start sites. *Reviews in Biology* **85**: 103-108.

Fecikova, H., Gbelska, Y., Barath, Z., and Grmanova, M. (1982) Relationship between the composition of phospholipids and respiratory activity of choline-deficient mutants of *Neurospora crassa*. *Folia Microbiology (Praha)* **27**: 76-80.

Field, C.M., Al-Awar, O., Rosenblatt, J., Wong, M.L., Alberts, B., and Mitchison, T. (1996) A purified *Drosophila* septin complex forms filaments and exhibits GTPase activity. *The Journal of Cell Biology* **133**: 605-616.

Fink, G.R. (1987) Pseudogenes in Yeast? *Cell* **49**: 5-6.

Flescher, E.G., Madden, K., and Snyder, M. (1993) Components required for cytokinesis are important for bud site selection in yeast. *Journal of Cellular Biology* **122**: 373-386.

Fukushima, K., Tanaka, Y., Nabeshima, K., Yoneki, T., Tougan, T., Tanaka, S., and Nojima, H. (2000) Dmc1 of *Schizosaccharomyces pombe* plays a role in meiotic recombination. *Nucleic Acids Research* **28**: 2709-2716.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., and *al, e.* (2003) The genome sequence of the filamentous fungus Neurospora crassa. *Nature* **422**: 859-868.

Gale, C., Gerami-Nejad, M., McClellan, M., Vandoninck, S., Longtine, M.S., and Berman, J. (2001) *Candida albicans* Int1p interacts with the septin ring in yeast and hyphal cells. *Molecular Biology of the Cell* **12**: 3538-3549.

Gale, C., Finkel, D., Tao, N., Meinke, M., McClellan, M., Olson, J. et al. (1996) Cloning and Expression of a gene encoding an intergrin-like protein in *Candida albicans*. *Proceedings of the National Academy of Sciences (USA)* **93**: 357-361.

Gale, C.A., Bendel, C.M., McClellan, M., Hauser, M., Becker, J.M., Berman, J., and Hostetter, M.K. (1998) Linkage of adhesion, filamentous growth, and virulence in *Candida albicans* to a single gene, INT1. *Science* **279**: 1355-1358.

Gallie, D.R. (2002) Protein-Protein interactions required during translation. *Plant Molecular Biology* **50**: 949-970.

Garrill, A., Lew, R.R., and Heath, I.B. (1992) Stretch-activated Ca^{2+} and Ca^{2+} -activated K^{+} channels in the hyphal tip plasma membrane of the oomycete *Saprolegnia ferax*. *Journal of Cell Science* **101**: 721-730.

Gattiker, A., Gasteiger, E., and Bairoch, A. (2002) ScanProsite: a reference implementation of a PROSITE scanning tool. *Applied Bioinformatics* **1**: 107-108.

Gavrias, V., Andrianopoulos, A., Gimeno, C.J., and Timberlake, W.E. (1996) *Saccharomyces cerevisiae* TEC1 is required for pseudohyphal growth. *Molecular Microbiology* **19**: 1255-1263.

Gladfelter, A.S., Bose, I., Zyla, T.R., Bardes, E.S.G., and Lew, D.J. (2002) Septin ring assembly involves cycles of GTP loading and hydrolysis by Cdc42p. *Journal of Cellular Biology* **156**: 315-326.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., and Feldmann, H. (1996) Life with 6000 genes. *Science* **274**: 546-547.

Gooday, G.W. (1971) An autoradiographic study of hyphal growth of some fungi. *Journal of General Microbiology* **67**: 125-133.

Gooday, G.W. (1990) *Inhibition of chitin metabolism*. Berlin: Springer-Verlag.

Gooday, G.W. (1995) The dynamics of hyphal growth. *Mycological Research* **99**: 385-394.

Grinblat, Y., Brown, N.H., and Kafatos, F.C. (1989) Isolation and characterization of the *Drosophila* translational elongation factor 2 gene. *Nucleic Acids Research* **17**: 7303-7314.

Gschwendt, M., Kittstein, W., Mieakes, G., and Marks, F. (1989) A type 2A protein phosphatase dephosphorylates the elongation factor 2 and is stimulated by the phorbol ester TPA in mouse epidermis *in vivo*. *FEBS Letters* **257**: 357-360.

Guo, Z., and Sherman, F. (1995) 3'-end-forming signals of yeast mRNA. *Molecular and Cellular Biology* **15**: 5983-5990.

Gurr, S.J., Unkles, S.E., and Kinghorn, J.R. (1987) *Gene Structure In Eukaryotic Microbes*. Washington DC: IRL Press.

Haarer, B.K., and Pringle, J.R. (1987) Immunofluorescence localization of the *Saccharomyces cerevisiae CDC12* gene product to the vicinity of the 10-nm filaments in the mother-bud neck. *Molecular and Cellular Biology* **7**: 3678-3687.

Harold, F.M. (1990) To shape a cell: an inquiry. *Microbiological Reviews* **54**: 381-431.

Heath, I.B. (1994) *The Mycota, I, Growth, Differentiation and Sexuality*. Berlin: Springer-Verlag.

Holmes, D.S., and Quigley, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* **114**: 193-197.

Hong, E.L., Shinohara, A., and Bishop, D.K. (2001) *Saccharomyces cerevisiae* Dmc1 protein promotes renaturation of single-strand DNA (ssDNA) and assimilation of ssDNA into homologous super-coiled duplex DNA. *Journal of Biological Chemistry* **276**: 41906-41912.

Hsu, S.C., Hazuka, C., Roth, R., Foletti, D., Heuser, J., and Scheller, R. (1998) Subunit composition, protein interactions and structures of the mammalian brain sec6/8 complex and septin filaments. *Neuron* **20**: 1111-1122.

Huang, G., Zhang, L., and Birch, R.G. (2000) Rapid amplification and cloning of Tn5 flanking fragments by inverse PCR. *Letters in Applied Microbiology* **31**: 149-153.

Iida, H., Nakamura, H., Ono, T., Okmuro, M.S., and Anraku, Y. (1994) MID1, a Novel Saccharomyces cerevisiae Gene Encoding a Plasma Membrane Protein, Is Required for Ca21 Influx and Mating. *Molecular and Cellular Biology* **14**: 8259-8271.

Innis, M.A., and Gelfand, D.H. (1994) *Optimization of PCRs, in PCR protocols, A Guide to Methods and Applications*. London: CRC Press.

Inoue, H., H., N., and Okayama, H. (1990) High effeciency transformation of *Escherichia coli* with plasmids. *Gene* **96**: 23-28.

Iturriaga, E.A., and Eslava, A.P. (1999) Unpublished-Direct Submission. *GenBank*.

Jackson, S.L., and Heath, I.B. (1989) Effects of exogenous calcium ions on tip growth, intracellular Ca2+ concentration, and actin arrays in hyphae of the fungus *Saprolegnia ferax*. *Experimental Mycology* **13**: 1-12.

Jahn, C.L., Doktor, S.Z., Frels, J.S., Jaraczewski, J.W., and Krikau, M.F. (1993) Structures of the *Euplotes crassus* Tec1 and Tec2 elements: identification of putative transposase coding regions. *Gene* **133**: 71-78.

Jorgensen, R., Ortiz, P.A., Carr-Schmid, A., Nissen, P., Kinzy, T.G., and Anderson, G.R. (2003) Two crystal structures demonstrate large conformational changes in the eukaryotic ribosomal translocase. *Nature Structural Biology* **10**: 379-385.

Joseph, S. (2003) After the ribosome structure: How does translocation work? *RNA* **9**: 160-164.

Kamath, A., and Chakraburtty, K. (1989) Role of yeast elongation factor 3 in the elongation cycle. *Journal of Biological Chemistry* **264**: 15423-15428.

Kanzaki, M., Nagasawa, M., Kojima, I., Sato, C., Naruse, K., Sokabe, M., and Iida, H. (1999) Molecular identification of a eukaryotic, stretch-activated nonselective cation channel. *Science* **285**: 882-886.

Kawasaki, E.S. (1990) Amplification of RNA. In *PCR Protocols: A Guide to Methods and Applications*. Innis, M.A., Sninsky, J.J., and White, T.J. (eds). San Diego: Academic Press, Inc, pp. 21-27.

Keogh, R.G., Tapper, B.A., and Fletcher, R.H. (1996) Distributions of the fungal endophyte Acremonium lolii, and of the alkaloids lolitrem B and peramine, within perennial ryegrass. *New Zealand Journal of Agricultural Research* **39**: 121-127.

Kinoshita, M., Kumar, S., Mizoguchi, A., Ide, C., Kinoshita, A., Haraguchi, T. et al. (1997) Nedd5, a mammalian septin, is a novel cytoskeletal component interacting with actin-based structures. *Genetic Development* **11**: 1535-1547.

Koga, H.M.J., Christensen, M.J., and Bennett, R.J. (1993) Incompatibility of some grass/Acremonium endophyte associations. *Mycological Research* **97**: 1237-1244.

Kohno, K., Uchida, T., Ohkubo, H., Nakanishi, S., Nakanishi, T., Fukui, T. et al. (1986) Amino acid sequence of mammalian elongation factor 2 deduced from the cDNA sequence: Homology with GTP-binding proteins. *Proceedings of the National Academy of Sciences (USA)* **83**: 4978-4982.

Kozak, M. (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Research* **9**: 5233-5252.

Kuo, M.H., and Grayhack, E. (1994) A libary of yeast genomic MCM1 binding sites contains gene involved in cell cycle control, cell wall and membrane structure, and metabolism. *Molecular and Cellular Biology* **14**: 348-359.

La Valle, R., and Wittenberg, C. (2001) A role for the Swe1 checkpoint kinase during filamentous growth of *Saccharomyces cerevisiae*. *Genetics* **158**: 549-562.

Lakin-Thomas, P.L. (1998) Choline depletion, frq mutations, and temperature compensation of the circadian rhythm in *Neurospora crassa*. *Journal of Biological Rhythms* **13**: 268-277.

Laloux, I., Dubois, E., Dewerchin, M., and Jacobs, E. (1990) TEC1, a gene involved in the activation of Ty1 and Ty1-mediated gene expression in *Saccharomyces cerevisiae*: cloning and molecular analysis. *Molecular and Cellular Biology* **10**: 3541-3550.

Lane, S., Zhou, S., Pan, T., Dai, Q., and Liu, H. (2001) The basic helix-loop-helix transcription factor Cph2 regulates hyphal development in *Candida albicans* partly via TEC1. *Molecular and Cellular Biology* **21**: 6418-6428.

Langer, R.H.M. (1974) *How Grasses Grow*. Southampton: Camelot Press Ltd.

Langford, C.J., Klinz, F.J., Donath, C., and Gallwitz, D. (1984) Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast. *Cell* **36**: 645-653.

Lauter, F.R., Marchfelder, U., Russo, V.E.A., Yamashiro, C.T., Yatzkan, E., and Yarden, O. (1998) Photoregulation of cot-1, a Kinase-Encoding Gene Involved in Hyphal Growth in *Neurospora crassa*. *Fungal Genetics and Biology* **23**: 300-310.

Leng, P., Lee, P.R., Wu, H., and Brown, A.J. (2001) Efg1, a morphogenetic regulator in *Candida albicans*, is a sequence-specific DNA binding protein. *Journal of Bacteriology* **183**: 4090-4093.

Levina, N.N., Lew, R.R., Hyde, G.J., and Heath, I.B. (1995) The roles of Ca^{2+} and plasma membrane ion channels in hyphal tip growth of *Neurospora crassa*. *Journal of Cell Science* **108**: 3405-3417.

Lin, Y.S., and Green, M.R. (1989) Identification and Purification of a *Saccharomyces cerevisiae* protein with the DNA binding specificity of mammalian activating transcription factor. *Proceedings of the National Academy of Sciences (USA)* **86**: 109-113.

Liu, H., Kohler, J., and Fink, G.R. (1994) Suppression of hyphal formation in *Candida albicans* by mutation of a STE12 homolog. *Science* **266**: 1723-1726.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaria, P., Baltimore, D., and Darnell, J.E. (1999) *Molecular Cell Biology*. New York: W. H. Freeman and Company.

Longtine, M.S., DeMarini, D.J., Valencik, M.L., Al-Awar, O., Fares, H., DeVirgilio, C., and Pringle, J.R. (1995) The septins: roles in cytokinesis and other processes. *Current Opinion In Cell Biology* **8**: 106-119.

Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting Coiled Coils from Protein Sequences. *Science* **252**: 1162-1164.

Malhotra, P., Manohar, C.F., Swaminathan, S., Toyama, R., Dhar, R., Reichel, R., and Thimmapaya, B. (1993) E2F site activates transcription in fission yeast Schizosaccharomyces pombe and binds to a 30-kDa transcription factor. *Journal of Biological Chemistry* **268**: 20392-20401.

Markham, P., Robson, G.D., Bainbridge, B.W., and Trinci, A.P. (1993) Choline: its role in the growth of filamentous fungi and the regulation of mycelial morphology. *FEMS Microbiol Reviews* **10**: 287-300.

McLeay, L.M., Smith, B.L., and Reynolds, G.W. (2002) Cardiovascular, respiratory, and body temperature responses of sheep to the ergopeptides ergotamine and ergovaline. *Annual Journal of Vetenary Research* **63**: 387-393.

Mendoza, A., Serramia, M.J., Capa, L., and Garcia-Bustos, J.F. (1999) Translation elongation factor 2 is encoded by a single essential gene in Candida albicans. *Gene* **229**: 183-191.

Mizuta, K., Tsujii, R., Warner, J.R., and Nishiyama, M. (1998) The C-terminal silencing domain of Rap1 is essential for the repression of ribosomal protein genes in response to a defect in the secretory pathway. *Nucleic Acids Research* **26**: 1063-1069.

Momany, M., Zhao, J., Lindsey, R., and Westfall, P.J. (2001) Characterization of the Aspergillus nidulans septin (asp) gene family. *Genetics* **157**: 969-977.
Mullins, J.T. (1973) Lateral branch formation and cellulase production in the water moulds. *Mycologia* **65**: 1007-1014.

Murad, A.M., Leng, P., Straffon, M., Wishart, J., Macaskill, S., MacCallum, D. et al. (2001) NRG1 represses yeast-hypha morphogenesis and hypha-specific gene expression in *Candida albicans*. *EMBO Journal* **20**: 4742-4752.

Nagahashi, S., Mio, T., Ono, N., Yamada-Okabe, T., Arisawa, M., Bussey, H., and Yamada-Okabe, H. (1998) Isolation of CaSLN1 and CaNIK1, the genes for osmosensing histidine kinase homologues, from the pathogenic fungus *Candida albicans*. *Microbiology* **144 (Pt 2)**: 425-432.

Nairn, A.C., and Palfrey, H.C. (1987) Identification of the major Mr 100,000 substrate for calmodulin-dependent protein kinase III in mammalina cells as elongation factor-2. *The Journal of Biological Chemistry* **262**: 17299-17303.

Nakamura, Y.K. (2004). Codon Usage Database. URL
http://www.kazusa.or.jp/codon/countcodon.html

Nakanishi, T., Kohno, K., Ishiura, M., Ohashi, H., and Uchida, T. (1988) Complete Nucleotide Sequence and Characterization of the 5'-Flanking Region of Mammalian Elongation Factor 2 Gene. *The Journal of Biological Chemistry* **263**: 6384-6391.

Nygard, O., and Nilsson, L. (1990) Kinetic determination of the effects of ADP-ribosylation on the interaction of eukaryotic elongation factor 2 with ribosomes. *Journal of Biological Chemistry* **265**: 6030-6034.

Ochman, H., Gerber, A.S., and Hartl, D.L. (1988) Genetic Applications of an Inverse Polymerase Chain Reaction. *Genetics* **120**: 621-623.

Oehlen, L., and Cross, F.R. (1998) The mating factor response pathway regulates transcription of TEC1, a gene involved in pseudohyphal differentiation of *Saccharomyces cerevisiae*. *FEBS Letters* **429**: 83-88.

Omura, F., Kohno, K., and Uchida, T. (1989) The histidine residue of codon 715 is essential for function of elongation factor 2. *European Journal of Biochemistry* **180**: 1-8.

Ovchinnikov, I.A., Alakhov, I.B., Bundulis, I.P., Bundule, M.A., and Vinokurov, L.M. (1983) Primary structure of the elongation factor G from Escherichia coli IX. Structure of peptides generated by cyanogen bromide cleavage of the G-factor isolated on thiol-activated sepharose and of the products of the G-factor cleavage at Asp-Pro bonds. Complete primary structure. *Bioorganism Khimistry* **9**: 343-357.

Panaccione, D.G., Johnson, R.D., Wang, J., Young, C.A., Damrongkool, P., Scott, B., and Schardl, C.L. (2001) Elimination of ergovaline from a grass-Neotyphodium endophyte symbiosis by genetic modification of the endophyte. *Proceedings of the National Academy of Sciences (USA)* **98**: 12820-12825.

Perentesis, J.P., Phan, L.D., Gleason, W.B., and LaPorte, D.C. (1992) *Saccharomyces cerevisiae* Elongation Factor 2. *The Journal of Biological Chemistry* **267**: 1190-1197.

Philipson, M.a.M.C.C. (1986) The relationship of host and endophyte during flowering, seed formation, and germination of *Lolium perenne*. *New Zealand Journal of Botany* **24**: 135-134.

Pontecorvo, G., Roper, J.A., Hemmons, D.W., MacDonald, K.D., and Bufton, A.W.J. (1953) The genetics of *Aspergillus nidulans*. *Advances In Genetics* **5**: 141-238.

Popay, A.J., Hume, D.E., Mainland, R.A., and Saunders, C.J. (1995) Field resistance to Argentine stem weevil (Listronotus bonariensis) in different ryegrass cultivars infected with an endophyte deficient in lolitrem B. *New Zealand Journal of Agricultural Research* **38**: 519-528.

Promega (1999) pGEM-T and pGEM-T Easy Vector Systems. In. Madison, USA: Promega Corporation.

Propheta, O., Vierula, J., Toporowski, P., Gorovits, R., and Yarden, O. (2001) The *Neurospora crassa* colonial temperature-sensitive 3 (cot-3) gene encodes protein elongation factor 2. *Molecular Gene and Genetics* **264**: 894-901.

Rapp, G., Klaudiny, J., Hagendorff, G., Luck, M.R., and Scheit, K.H. (1989) Complete sequence of the coding region of human elongation factor 2 (EF-2) by enzymatic amplification of cDNA from human ovarian granulosa cells. *Biological Chemistry* **370**: 1071-1075.

Reissig, J.L., and Kinney, S.G. (1983) Calcium as a Branching Signal in *Neurospora crassa*. *Journal of Bacteriology* **154**: 1397-1402.

Riis, B., Rattan, S.I.S., Clark, B.F.C., and Merrick, W.C. (1990) Eukaryotic protein elongation factors. *TIBS* **15**: 420-424.

Rodnina, M.V., Savelsbergh, A., Katunin, V.I., Semenkov, Y.P., and Wintermeyer, W. (1999) Thiostrepton inhibits the turnover but not the GTPase of elongation factor G on the ribosome. *Proceedings of the National Academy of Sciences (USA)* **96**: 9586-9590.

Ryazanov, A.G., and Spirin, A.S. (1990) Phosphorylation of elongation factor 2 by EF-2 kinase affects rate of translation. *New Biology* **2**: 843-850.

Sarkar, G., Kapelner, S., and Sommer, S.S. (1990) Formamide can dramatically improve the specificity of PCR. *Nucleaic Acids Research* **18**: 7465.

Schardl, C.L., and An, Z. (1993) *Molecular biology and genetics of protective fungal endophytes of grasses*. New York: Plenum Press.

Schmid, J., and Harold, F.M. (1988) Dual roles for Calcium Ions in Apical Growth of *Neurospora crassa*. *Journal of General Microbiology* **134**: 2623-2631.

Schmid, J., and Christensen, M.J. (1999) Ryegrass endophyte: host/fungus interaction. In *Grassland Research and Practice Series*, pp. 101-106.

Schug, J., and Overton, C.G. (1997). TESS: Transcription Element Search Software on the WWW'. URL http://www.cbil.upenn.edu/tess

Schweizer, A., Rupp, S., Taylor, B.N., Rollinghoff, M., and Schroppel, K. (2000) The TEA/ATTS transcription factor CaTec1p regulates hyphal development and virulence in *Candida albicans*. *Molecular Microbiology* **38**: 435-445.

Scott, B., Young, C.A., and McMillan, L. (1999) Molecular biology of *Epichloe* endophyte toxin biosynthesis. In *Ryegrass Endophyte: An essential New Zealand symbiosis*. Woodfield, D.R., and Matthew, C. (eds). Palmerston North: Double M Publishing, pp. 77-83.

Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V., and Levinson, A.D. (1984) Biological properties of human c-Ha-ras1. *Nature* **312**: 71-75.

Shastry, M., Nielsen, J., Ku, T., Hsu, M.J., Liberator, P., Anderson, J. et al. (2001) Species-specific inhibition of fungal protein synthesis by sordarin: identification of a sordarin-specificity region in eukaryotic elongation factor 2. *Microbiology* **147**: 383-390.

Short, K.M., Hopwood, B., Yi, Z., and Cox, T.C. (2002) MID1 and MID2 homo- and heterodimerise to tether the rapamycin-sensitive PP2A regulatory subunit, Alpha 4, to microtubules: implications for the clinical variability of X-linked Opitz GBBB syndrome and other developmental disorders. *BMC Cellular Biology* **3**: 1.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Research*: 1087-1088.

Silverman-Gavrila, L.B., and Lew, R.R. (2000) Calcium and tip growth in *Neurospora crassa*. *Protoplasma* **213**: 203-217.

Smale, S.T. (1997) Transcription initiation from TATA-less promoters within eukaryotic protein-encoding genes. *Biochimica et Biophysica Acta* **1351**: 73-88.

Southern, E.M. (1975) Detection of specific sequences amoung DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503-517.

Spataford, J.W., and Blackwell, M. (1993) Molecular systematics of unitunicate perithecial ascomycetes: the clavicipitales-hypocreales connection. *Mycological Research* **85**: 912-922.

Stoldt, V.R., Sonneborn, A., Leuker, C.E., and Ernst, J.F. (1997) Efg1p, an essential regulator of morphogenesis of the human pathogen *Candida albicans*, is a member of a conserved class of bHLH proteins regulating morphogenetic processes in fungi. *EMBO Journal* **16**: 1982-1991.

Stryer, L. (1995) *Biochemistry*. New York: W. H. Freeman and Company.

Sukharev, S.I., Betanzos, M., Chiang, C.-S., and Guy, R.H. (2001) The gating mechanism of the large mechanosensitive channel MscL. *Nature* **409**: 720-724.

Takeuchi, Y., Schmid, J., Caldwell, J.H., and Harold, F.M. (1988) Transcellular Ion Currents and Extension of *Neurospora crassa* hyphae. *Journal of Membrane Biology* **101**: 33-41.

Tan, Y.Y., Spiering, M.J., Scott, V., Lane, G.A., Christensen, M.J., and Schmid, J. (2001) In planta regulation of an endophytic fungus and maintenance of high metabolic rates in its mycelium in the absence of apical extension. *Applied and Enviromental Microbiology* **67**: 5377-5383.

Tasaka, Y., Nakagawa, Y., Sato, C., Mino, M., Uozumi, N., Murata, N. et al. (2000) Yam8(+), a *Schizosaccharomyces pombe* gene, is a potential homologue of the *Saccharomyces cerevisiae* MID1 gene encoding a stretch-activated Ca(2+)-permeable channel. *Biochemical and Biophysical Research Communications* **269**: 265-269.

Templeton, N.S., Ureelay, E., and Safer, B. (1993) Reducing artifact and increasing the yield of specific DNA target fragments during PCR-RACE or anchor PCR. *BioTechniques* **15**: 48-50.

Thuring, R.W.J., Sanders, J.P.M., and Borst, P. (1975) A Freeze-Squeeze Method For Recovering Long DNA From Agarose Gels. *Analytical Biochemistry* **66**: 213-220.

Triana, F.J., Nierhaus, K.H., and Chakraburtty, K. (1994) Transfer RNA binding to 80s ribosomes from yeast. Evidence for three sites. *Biochemical and Molecular Interpretations* **32**: 336-342.

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

Trimble, W.S. (1999) Septins:A highly conserved family of membrane-associated GTPases with functions in Cell division and beyond. *Journal of Membrane Biology* **169**: 75-81.

Uchiumi, T., and Kominami, R. (1994) A functional site of the GTPase-associated center within 28S ribosomal RNA probed with an anti-RNA autoantibody. *EMBO Journal* **13**: 3389-3394.

Van Heusden, G.P., Wenzel, T.J., Lagendijk, E.L., de Steensma, H.Y., and van den Berg, J.A. (1992) Characterization of the yeast BMH1 gene encoding a putative protein homologous to mammalian protein kinase II activators and protein kinase C inhibitors. *FEBS Letters* **302**: 145-150.

Van Miltenburg, R., Ruger, B., Grunewald-Janho, S., and Schroder, C. (eds) (1995) *The DIG System User's Guide for Filter Hybridization*: Boehringer Mannheim.

Warenda, A.J., and Konopka, J.B. (2002) Septin function in *Candida albicans* morphogenesis. *Molecular Biology of the Cell* **13**: 2732-2746.

Wartmann, T., Gellissen, G., and Kunze, G. (2001) Regulation of the AEFG1 gene, a mitochondrial elongation factor G from the dimorphic yeast *Arxula adeninivorans* LS3. *Current Genetics* **40**: 172-178.

Watters, M.K., Virag, A., Haynes, J., and Griffiths, A.J.F. (2000) Branch initiation in *Neurospora* is influenced by events at the previous branch. *Mycological Research* **104**: 805-809.

Watts, H.J., Very, A.A., Perera, T.H.S., Davies, J.M., and Gow, N.A.R. (1998) Thigmotrophism and stretch-activated channels in the pathogenic fungus *Candida albicans*. *Microbiology* **144**: 689-695.

Weaver, R.F. (1999) *Molecular Biology*. Boston: WBC/McGraw-Hill.

Wessels, J.G.H. (1990) Role of cell wall architecture in fungal tip growth generation. In *Tip Growth in Plant and Fungal Cells*. Heath), e.I.B. (ed). New York: Academic Press, pp. pp. 1-29.

Westfall, P.J., and Momany, M. (2002) *Aspergillus nidulans* Septin AspB plays Pre- and Postmitotic roles in Septum, Branch and Conidiophore Development. *Molecular Biology of the Cell* **13**: 110-118.

White, J.F.J., A. C. Morrow, G. Morgan-Jones and D. A. Chambless (1991) Endophyte-host associations in forage grasses. XIV. Primary stromata formation and seed transmission in *Epichloë typhina:* developmental and regulatory aspects. *Mycologia* **83**: 72-81.

Wiebe, M.G., Robson, G.D., and Trinci, A.P. (1989) Effect of choline on the morphology, growth and phospholipid composition of *Fusarium graminearum*. *Journal of Genetics and Microbiology* **135** (**Pt 8**): 2155-2162.

Wiederrecht, G., Seto, D., and Parker, C. (1988) Isolation of the gene encoding the *S. cerevisiae* Heat shock transcription factor. *Cell* **54**: 841-853.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., and Hochstrasser, D.F. (1998) *Protein Identification and Analysis Tools in the ExPASy Server*. New Jersey: Humana Press.

Wood, V., Gwilliam, R., Rajandream, M.-A., and al, e. (2002) The genome sequence of Schizosaccharomyces pombe. *Nature* **415**: 871-880.

Xing, Y.Y., GuhaThakurta, D., and Draper, D.E. (1997) The RNA binding domain of ribosomal protein L11 is structurally similar to homeo-domains. *Nature Structural Biology* **4**: 24-27.

Yarden, O., Plamann, M., Ebbole, D.J., and Yanofsky, C. (1992) cot-1, a gene required for hyphal elongation in *Neurospora crassa*, encodes a protein kinase. *EMBO Journal* **11**: 2159-2166.

Yoon, J.H., Kim, J.K., Rha, G.B., Oh, M., Park, S., Seong, R.H. et al. (1999) Sp1 mediates cell proliferation-dependent regulation of rat DNA topoisomerase IIa gene promoter. *Biochemical Journal* **344**: 367–374.

Zhao, J., Hyman, L., and Moore, C. (1999) Formation of mRNA 3' Ends in Eukaryotes: Mechanism, Regulation, and Interrelationships with Other Steps in mRNA Synthesis. *Microbiology and Molecular Biology Reviews* **63**: 405-445.