CHARACTERIZATION OF PRD-1 MUTATION IN NEUROSPORA CRASSA

GHAZALEH FIROOZI

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

December 2013

Abstract

In *Neurospora crassa*, rhythmic conidiation is controlled by several oscillators such as FRQ/WCC and FLO. The frequency (frq) gene and white collar genes (WC-1 and WC-2) are the most important components of the FRQ/WCC oscillator and *prd-1* and *prd-2* genes could be important components of the FLO. This project aims to characterize the *prd-1* mutation. This involved mapping the mutation using PCR analyses based on single nucleotide polymorphism (SNP) markers. After determining a minimal interval, candidate genes were sequenced, and knockout mutants were screened in this interval for prd-1 phenotype. A candidate gene was found to have a mutation that affects the splicing of the mRNA. The identity of the gene was confirmed by complementing the *prd-1* mutant with a wild type copy of the identified candidate gene. The *prd-1* gene is identified as an RNA helicase.

Acknowledgment

I would like to express my sincere gratitude to my supervisor Dr. Patricia Lakin-Thomas for the continuous support of my master study and research, for her patience, motivation, enthusiasm, and immense knowledge. Without her supervision and constant help this thesis would not have been possible.

Besides my advisor, I would like to thank the rest of my thesis committee Dr. Ron Pearlman, Dr. Joel Shore and Dr. Gerald. Audette for their encouragement and insightful comments.

My sincere thanks also goes to Dr. Keyur Advaharyu for extending help to carry out this research work.

Finally, the moral support and genuine affection rendered to me by my parents and my brothers are always remembered with deep gratitude.

Table of contents

List of Tables

List of Figures

List of Abbreviations

bd: Band

BLAST: Basic Local Alignment Search Tool

BMAL1: brain and muscle Arnt-like protein

BSA: Bovine Serum Albumin

CAPS: Cleaved Amplified Polymorphic Sequences

CCA1: Circadian Clock Associated 1

ccg: Clock Control Genes

chol-1: Choline mutatnt

CK2: Casein Kinase 2

Clk: Clock

CLK: CLOCK

CO: Constants

Cry: Cryptochrome

csp-1: Conidial separation 1

CWO: Clock Work Orange

CYC: CYCLE

DBT: DOUBLETIME

ELF3: Early Flowering 3

ELF4: early Flowering 4

FFC: FRQ-FRH Complex

FGSC: Fungal Genetic Stock Centre

FLO: FRQ-less Oscillator

FRH: FRQ-interacting RNA helicase

FRQ: Frequency

FT: Flowring Locus T

FWO: FRQ/White collar Oscillator

GI: Gigantea

hyg: Hygromycin

ipRGC: Intrinsically Photosensitive Retinal Ganglion Cells

LNs: Lateral Neurons

LUX: LUX Arrhythmo

LWD1: Light-Regulated WD1

LYH: Late Elongated Hypocotyl

PAS: PER-ARNT-SIM

PCR: Polymerase Chain Reactio

per: Period

prd: Period

PRR: Pseudo Response Regulator

PTO: Post-translational Oscillator

RNPs: Ribonucleoproteins

SCN: Superchiasmatic Nucleus

SNPs: Single Nucleotide polymorphisms

TIM: Timeless

TOC1: Timing of CAB2 Expression

TTFL: Transcriptional/Translational feedback loop

TTO: Transcription/Translation feedback oscillator

WC-1: White Collar 1

WC-2: White Collar 2

WCC: White Collar complex

ZTL: ZEITLUPE

1. Introduction

1.1 Circadian rhythms

The elements of temperature and light in our natural surroundings are inevitably altered by the earth rotating. Circadian is derived from the Latin word ‘circa’ which means about and ‘dies’ which means day. The evolution of circadian clocks is based on natural selection. Biological clocks are the endogenous cellular devices for measuring time. The clocks instil a benefit of survival by making it possible for a living being to anticipate the environmental changes that take place daily and therefore making its behaviour and physiology adapt to the daily changes at appropriate times. These clocks are synchronised by night-day cycles and enable a living form to synchronize with the daily light and dark cycles that are a result of the rotating earth (Panda et al. 2002).

Evolution created an accurate biological clock in many studied organisms. Studying and analysis of how such a precise biological clock is assembled and works have not been trivial tasks. Advanced development of genetic tools and the study of the genetics of circadian rhythms in model organisms like flies has opened up new avenues in the understanding of the daily regulation of circadian behaviours at the molecular level.

Circadian rhythms are a fundamental property of living cells. All the organisms from unicells to humans show behavioural and physiological rhythms that continue with a period of approximately 24 h (Dunlap 1999). These rhythms are driven by biological clocks that have three important criteria.

First, the rhythms have a 24-hour period. Second, the free-running period of 24 h is temperature compensated, which means that clocks do not run slower at lower temperatures or speed up at higher temperature. Third, they can also be synchronised (entrained) by external stimuli such as light, temperature or feeding schedules (Takahashi et al. 2008).

The importance of circadian clocks has been shown in prokaryotes and higher organisms (Woelfle et al. 2004, Dodd et al. 2005). Disruption of circadian rhythm in humans, as seen in rotational shift workers for example, causes important long-term health problems such as disorders of sleep, diabetes and obesity (Barger et al. 2009). In addition to sleep and feeding behaviour, the circadian clocks of animals also control rhythmicity in gene expression, cell cycle and metabolism, body temperature, hormone activity and blood pressure (de Paula et al., 2008). In plants, germination, flowering, fragrance emission and photosynthesis show circadian rhythms (McClung 2006).

Observations of oscillatory systems date as far back as 1729 when de Marian described circadian rhythms in plant leaves (de Marian 1729). Initially he set out to investigate the leaf movements of *Mimosa pudica* in the dark in the absence of a stimulus and cyclic environmental information. He found that the leaf rhythm continued, therefore there should be an endogenous clock (Dunlap et al. 2004). In the 1970s, the first mutation in *Drosophila* that disrupted the clock system was identified (Konopka et al. 1971). Decades later in the 1990s, the first mammalian clock mutants where described (Vitatema et al. 1994).

The circadian clock contains three elements: the oscillator that has a positive and negative feedback loop, the input such as temperature or light that can reset the oscillator and the output pathways that regulate the observed rhythmicity.

The proposed model for the oscillator is based on a transcription / translation feedback loop. In this loop transcription of clock genes is rhythmic and clock proteins feedback to regulate their own transcription negatively. To form a loop, rhythmic transcription leads to rhythmic levels of RNA and rhythmic mRNA levels cause rhythmic levels of proteins. The most important component for maintaining rhythmicity in organisms is rhythmic transcription of clock genes. In the transcription/translation oscillator (TTO) model Clock gene 2 is transcribed and translated into clock protein 2 and this protein positively regulates the transcription of clock gene 1. On the other hand clock protein 1 negatively regulates its own transcription and inhibits the positive effect of clock protein 2. Clock protein 1 also positively regulates the transcription and translation of clock gene 2. (Fig.1.1)

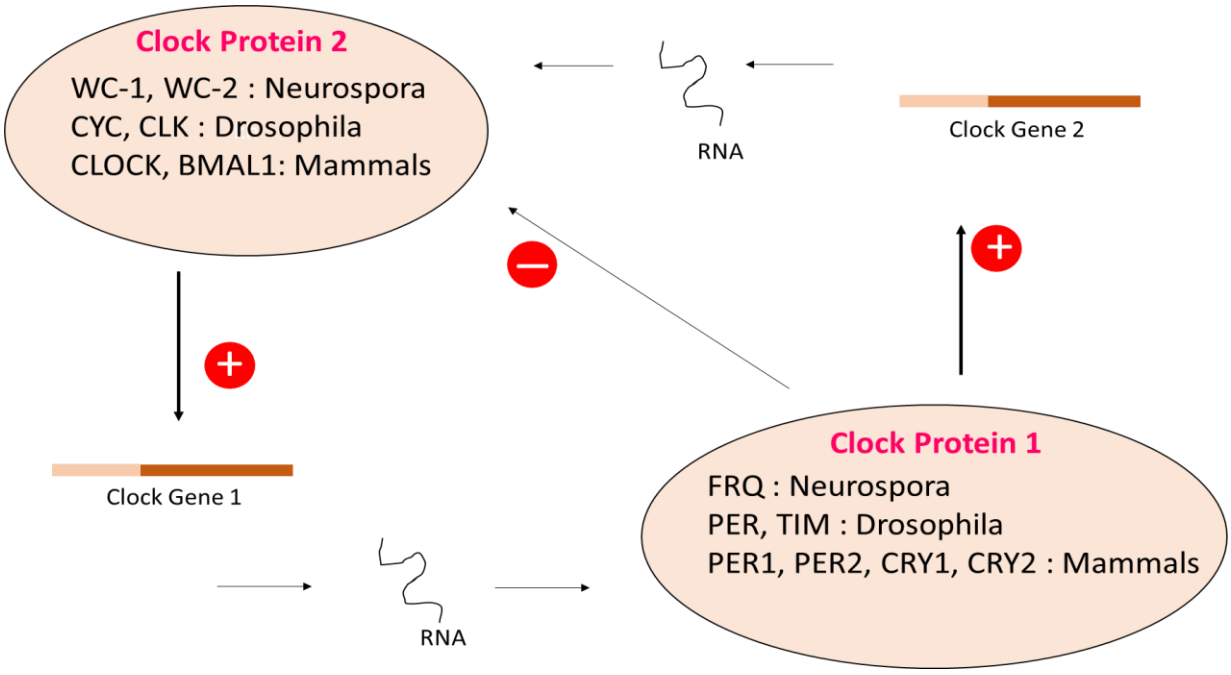


Fig 1.1. Transcription/translation feedback oscillator (TTO) model for circadian rhythms. Two interlocked loops are interacting with each other. Clock protein 1 positively regulates transcription of clock gene 2. Clock protein 2 positively regulates transcription of clock gene 1 (Lakin-Thomas, 2006 & de Paula et al. 2008).

1.2. Molecular rhythms in *Cyanobacteria*

The circadian rhythm in *Cyanobacterium Synechococcus* can be seen in the nitrogen fixation process. This process is regulated by the circadian clock to occur at night time because of the nitrogenase enzyme's oxygen sensitivity. Therefore, photosynthesis which produces oxygen occurs in the day and nitrogen fixation happens at night.

Essential clock genes in *Cyanobacteria* are *kaiA*, *kaiB* and *kaiC* that are found together in a cluster on the chromosome (Ishiura et al. 1998). Further studies of the kaiABC cluster suggested that the cynobacterial clockwork is similar to the mechanism that was proposed for the circadian system in eukaryotes, which is a

transcription/translation feedback loop (Hardin et al. 1990, Dunlap 1999) (Fig.1.2). The level of *kaiB* and *kaiC* mRNA and protein oscillate and overexpression of KaiC protein can repress the *kaiC* gene and reset the clock (Xu et al. 2000, Ishiura et al. 1998). It was also shown that KaiC can be phosphorylated rhythmically (Iwasaki et al. 2002). Deletion of any of these *kai* genes can abolish the rhythm and mutation in *kaiC, kaiB* or *kaiA* can change the period. *kaiC* has autophosphorylation and autodephosphorylation activity and it was shown that inhibition of a *kai* gene does not require a *kai* specific promoter (Ditty et al. 2003, Xu et al. 2003, Nakahira et al. 2004). Experiments from the Kondo team demonstrated reconstitution of the KaiC rhythm in vitro by mixing purified KaiA, KaiB and KaiC and also ATP in a test tube. It was found that KaiA promotes autophosphorylation of KaiC, however KaiB hinders the positive effect of KaiA and interaction between these three Kai proteins can create the oscillation of KaiC phosphorylation.

It was shown that oscillation in KaiC phosphorylation is the molecular time keeper in *Synechoccus* (Nakajima et al. 2005). Therefore, the underlying mechanism for *Cyanobacteria* is based on a post-translational oscillator (PTO) and also a transcriptional/translational feedback loop (TTFL). In the absence of KaiC phosphorylation the gene expression rhythmicity is gone so the KaiC oscillation is the master pacemaker in *Cyanobacteria*. Interaction of the TTFL and the protein oscillator can improve the robustness of the circadian clock (Johnson et al. 2008).

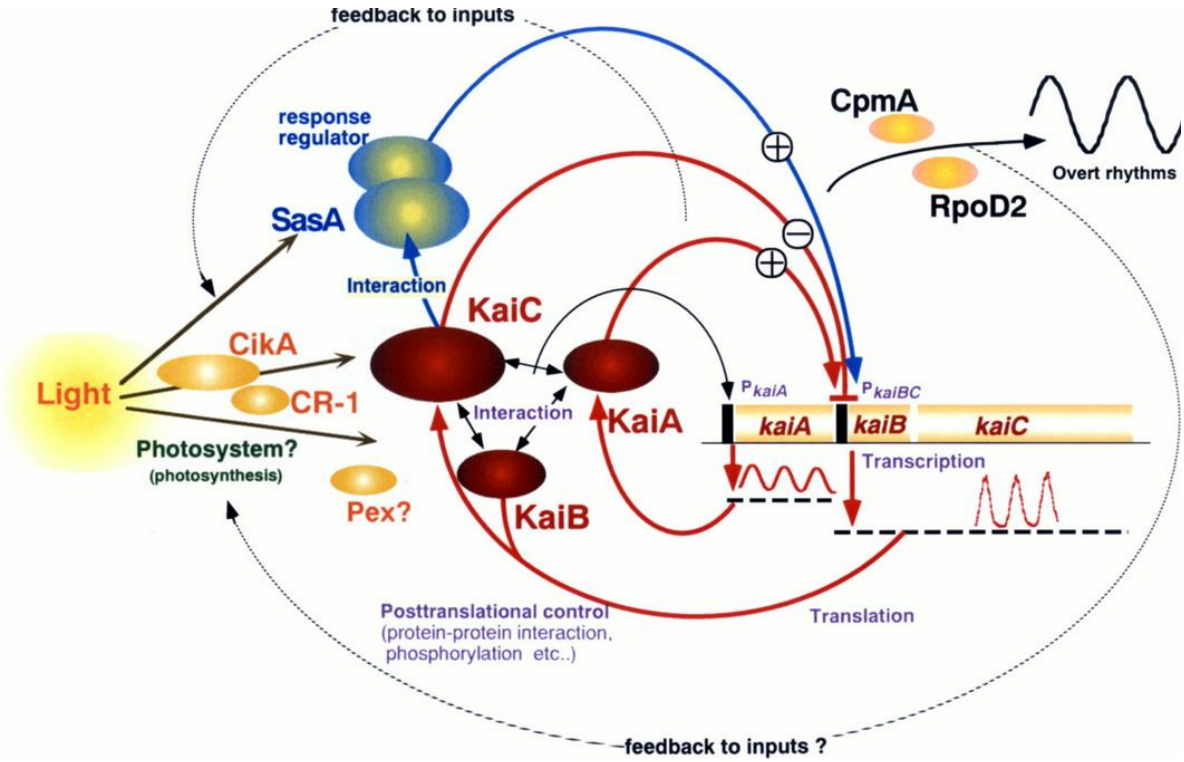


Fig 1.2. Proposed model for circadian system in *Cyanobacteria* (red loop). KaiC negatively regulates its own (kaiBC) expression to generate a molecular feedback loop, whereas KaiA activates kaiBC expression as a positive element to make the loop oscillate. Furthermore, posttranslational events like protein-protein interactions and phosphorylation may be required for regulation of their functions. SasA which is a KaiC-interacting sensory histidine kinase improves expression of *kaiBC* genes. SasA also likely adjusts sensitivity of robustness of the oscillator to light. CikA is a phytochrome-related histidine kinase that would be a circadian photoreceptor. The CR-1 and PEX proteins may also associate with some circadian photoreceptive pathways, while CpmA and RpoD2 are components of output pathways (Iwasaki and Kondo 2000).

1.3. Molecular rhythms in *Drosophila*

Eukaryotic systems share similar oscillators based on negative feedback loops that control the robust rhythm (Alabadi et al. 2001, Boxall et al. 2005). In addition to transcription-based interaction between clock components, post-transcriptional and post-translational regulation coupled with chromatin modifications are involved in regulating and controlling the oscillators (Covington et al. 2001, Daniel et al. 2004).

In the fruit fly *Drosophila melanogaster* lateral neurons (LNs) located in the central brain are considered as the central circadian pacemaker. LNs are able to control the rest, activity rhythms and timing of eclosion.

Molecular and genetic studies in *Drosophila* have revealed that the circadian rhythm mechanism is based on a TTO (Fig.1.3). The first known clock gene in *Drosophila* was the period (per) (Konopka et al. 1971) gene. Later, by identifying the other genes involved in this autoregulatry feedback loop, it was proposed that there are two interlocked feedback loops in *Drosophila*. Gene expression in the period/timeless (per/tim) and Clock (Clk) feedback loops is regulated primarily at the transcriptional and post-translational levels (Allada et al. 2003, Glossop et al. 1999, Young et al.2001). Because of this post-translational regulation, delays in abundance and activity of transcriptional inhibitors and activators can be seen. This delay regulates the timing of clock gene transcription during the circadian cycle.

For the circadian oscillator to function normally, it needs both types of regulation, transcriptional and post-translational. It was shown that in the absence of DOUBLETIME (DBT) kinase activity, or by altering the phosphorylation site of PER, the circadian period can be changed considerably and by combining these two the oscillator function

will be eliminated (Chiu et al. 2008, Rothenfluh et al. 2000). This means post-translational control is essential for rhythmicity. Similarly, circadian oscillator function is diminished or abolished by eliminating per and tim mRNA cycling (Hall et al. 2007, Yang et al. 2001), showing that transcriptional control is also necessary.

The feedback loops in *Drosophila* can be separated into different sequential phases including transcriptional activation, repression and reactivation. Hypophosphorylated CLOCK (CLK) makes a heterodimeric complex with CYCLE (CYC) and initiates the feedback loop. This complex binds to E-boxes and activates transcription of clock genes including period (per), timeless (tim), vrille (vri), PAR domain protein 1e (Pdp1e), and clockwork orange (cwo). A decrease in activation of CLK-CYC-dependent transcription and consequently a long period of behavioral rhythm can be seen in cwo mutants and CWO knockdown strains (Kadener et al. 2007, Lim et al. 2007, Richier et al. 2008). It is not known how CWO affects the activation of CLK-CYC-dependent transcription but studies showed that CWO is one of the important components of the clock (McDonald et al. 2001, Lin et al. 2002). It was also determined that CWO competes with CLK-CYC to bind to E1 E-boxes and acts to repress CLK–CYC transcription (Kadener et al. 2007). During the night, PER accumulates to high levels. TIM is needed for PER accumulation (Price et al. 1995) but it is not essential for repression. Therefore, a complex of PER and DOUBLETIME kinase (DBT) is sufficient to mediate repression (Chang et al. 2003). A complex of PER-DBT can interact with CLK-CYC, and PER-CLK binding is required for this repression by mediating hyperphosphorylation of CLK. This suggests that CLK hyperphophorylation can remove CLK-CYC from E-boxes to repress transcription (Kim et al. 2007).

At lights-on, transcriptional reactivation can be initiated when DBT- dependent phosphorylation encourages PER degradation (Nawathean et al. 2004). Hypophosphorylated CLK begins to accumulate and the next cycle of transcription will commence (Yu et al. 2006).

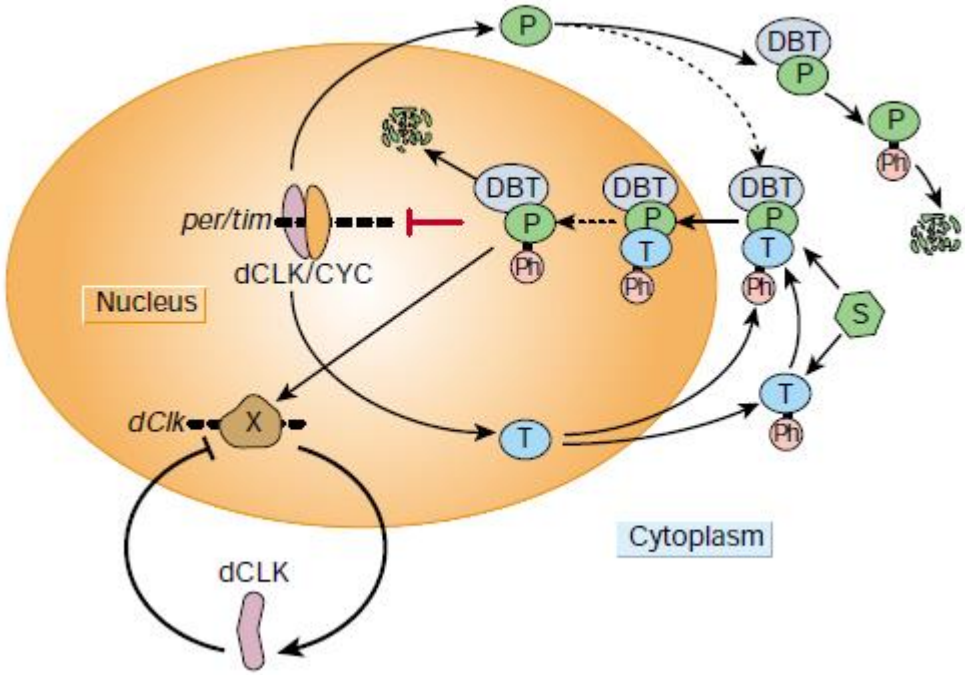


Fig 1.3. Model for *Drosophila* TTO. In *Drosophila* a heterodimer of two PAS-domain containing transcription factors dCLK/CYC binds to E-box of per/tim promoter and promotes their transcription. DBT phosphorylates cytoplasmic PER protein (P) and promotes its degradation. When TIM accumulates it can inhibit DBT activity and stabilize PER. Shaggy (S) phosphorylates TIM and causes the entry of PER/DBT/TIM into the nucleus. Later TIM is released from the complex and represses dCLK/CYC function. In the absence of TIM, PER is phosphorylated and degraded and dCLK/CYC is derepressed and a new cycle is starting. dCLK creates another feedback loop that inhibits its own transcription and PER promotes dCLK transcription (Adapted from Panda et al. 2002).

1.4. Molecular circadian network in plants

Circadian clocks in higher plants affect several regulatory pathways including metabolism, photosynthesis, hormone levels and growth regulation, nutrient uptake and flowering. Many studies were done to discover the transcription-based oscillator in *Arabidopsis*.

In *Arabidopsis*, the core oscillator model (Fig.1.4) was based on three components and loss of function of any of these clock genes results in a short period clock. These components are two MYB domain-containing transcription factors CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY), and a member of the PSEUDO RESPONSE REGULATOR (PRR) family, TIMING OF CAB2 EXPRESSION 1 (TOC1) (Edwards et al. 2006).

Several studies showed that TOC1 is a DNA-binding transcriptional repressor of CCA1 and LHY (Gendron et al. 2012). Along with CCA1 and LHY, TOC1 also binds the promoters and hinders the expression of the other oscillator components such as PRR5, PRR7, PRR9, LUX ARRHYTHMO (LUX), GIGANTEA (GI), and EARLY FLOWERING 4 (ELF4). In a morning-specific loop, CCA1 and LHY promote the expression of two TOC1 family members, PRR9 and PPR7, by directly binding to their promoters (Hanano et al. 2006, 2008). In return, PRR9, PRR7, and PRR5 function as transcriptional repressors to coregulate the expression of CCA1 and LHY. It was found that LIGHT-REGULATED WD1 (LWD1) contributes in a positive feedback loop with PRR9 and also indirectly increases expression of CCA1, LHY, PRR5 and TOC1. These results suggest that LWD1 functions as a transcriptional activator. It is known that in eukaryotic cells, post-translational regulation participates in keeping robust circadian

rhythms (McWatters et al. 2000, 2007). For instance, Casein Kinase 2 (CK2) is one of the conserved components in mammals, *Drosophila*, *Neurospora* and *Arabidopsis* that mediates regulation of clock genes (Michael et al. 2003, Millar et al. 1995). In *Arabidopsis*, CK2 phosphorylates CCA1 and LHY, and this process is considered to be important for CCA1 function.

In *Arabidopsis*, it was known that hypocotyl growth can be regulated by circadian clocks as well as several external cues (Park et al. 1999). Two key clock-related components in hypocotyl growth in *Arabidopsis* are PIF4 and PIF5. LUX which is one component of the evening complex (LUX, ELF3 and ELF4) binds to PIF4 and PIF5 promoters and represses them (Park et al. 1999, Kevei et al. 2006). Therefore hypocotyl growth is inhibited in early evenings and later on when the repression is released hypocotyl growth occurs.

Further studies demonstrated that molecular clock interactions can regulate the photoperiodic flowering. The known components that connect circadian oscillator to photoperiodic flowering are GIGANTEA (GI), flowering regulators CONSTANS (CO) and FLOWERING LOCUS T (FT) (Roenneberg et al. 1993, Sai et al. 1999).

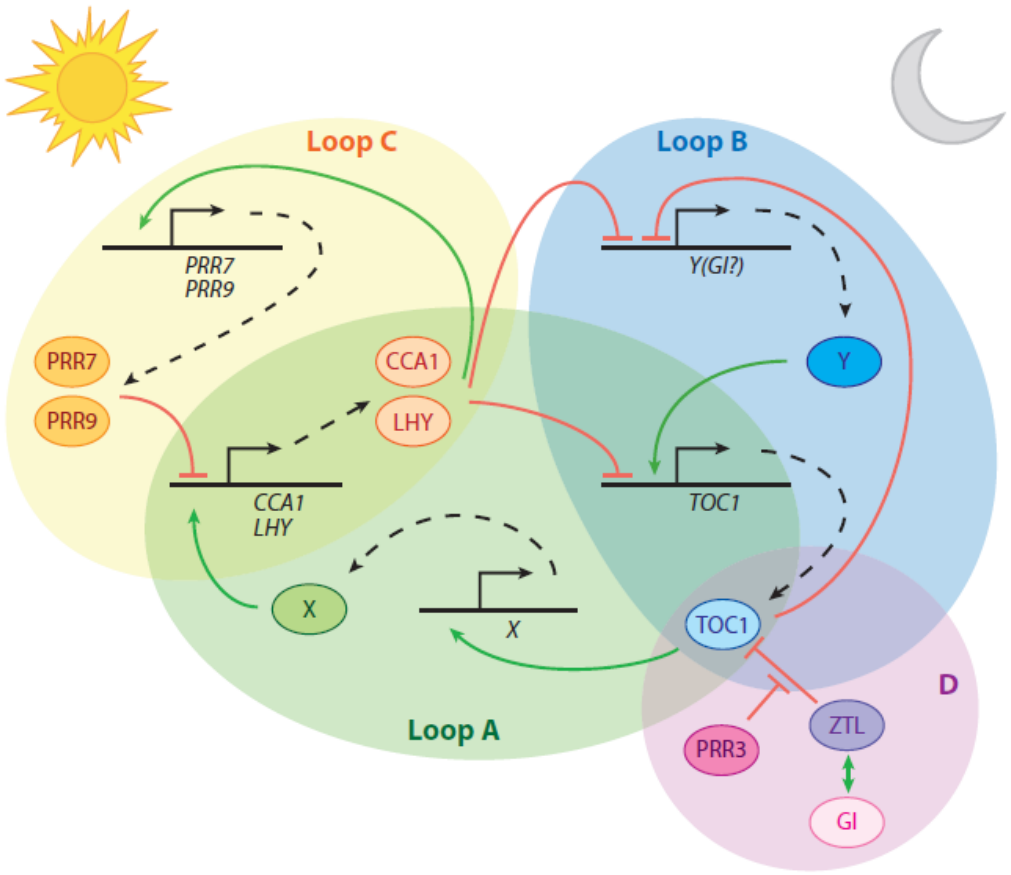


Fig 1.4. Model proposed for the plant clock. Loop A is the first identified feedback loop that consists of CCA1 and LHY that negatively regulate expression of TOC1. TOC1 can directly or indirectly activate the expression of CCA1 and LHY. Loop B is composed of two or more evening phase genes and loop C consists of morning phase genes like PRR7, PRR9, CCA1 and LHY. In loop D ZTL negatively regulate TOC1 protein abundance and its activity is regulated by PRR3 and GI (Harmer 2009).

1.5. Circadian rhythms in mammals

In mammalian cells, like other organisms, circadian clocks are driven based on autoregulatory negative-feedback transcriptional loops (Lowrey & Takahashi 2004,Takahashi et al. 2008). The core components of this network are CLOCK and BMAL1 which positively regulate the expression of Period (Per1 and Per2) and Cryptochrome (Cry1 and Cry2). Per and Cry proteins form a complex and after translocation into the nucleus they interact with CLOCK and BMAL1 and repress their own transcription. Another candidate gene that plays a role in the circadian gene network is Rev-erbα (Fig.1.5).

The hypothalamic suprachiasmatic nucleus (SCN) was identified as the main circadian pacemaker for generation of circadian behavioral rhythms in mammals (Welsh et al. 2010). After the discovery of clock genes and by using gene reporter methods it was found that most peripheral organs and tissues can express circadian oscillations and this circadian expression is extensive throughout the body (Dibner et al. 2010). In order to have an accurate circadian system, the SCN controls the peripheral oscillators and loss of SCN can cause desynchronization of peripheral circadian clocks (Yoo et al. 2004). Therefore, tissue specific gene expressions are regulated by both “local” as well as central mechanisms. The SCN receives the input from retina photoreceptor cells called intrinsically photosensitive Retinal Ganglion Cells (ipRGC). In addition to ipRGC that seem to be important circadian photoreceptors, rod and cone photoreceptors also participate to provide photic inputs to SCN (Chen et al. 2011, Guler et al. 2008).

For many years it has been known that there is a relationship between the cell cycle and the clock work. DNA synthesis and replication are performed at night; therefore they are not exposed to harmful UV radiation. Clock components per-1 and per-2 are recognized as tumour suppressors and any defects in these genes can result in tumour formation *in vivo.* On the other hand cryptochrome (cry) is an oncogene and loss of cry can decrease the risk of cancer in p53 mutant mice (Yi et al. 2010). BMAL/CLOCK complexes can control the expression of cell cycle and cell proliferation genes such as c-Myc and wee1 (Reddy et al. 2005).

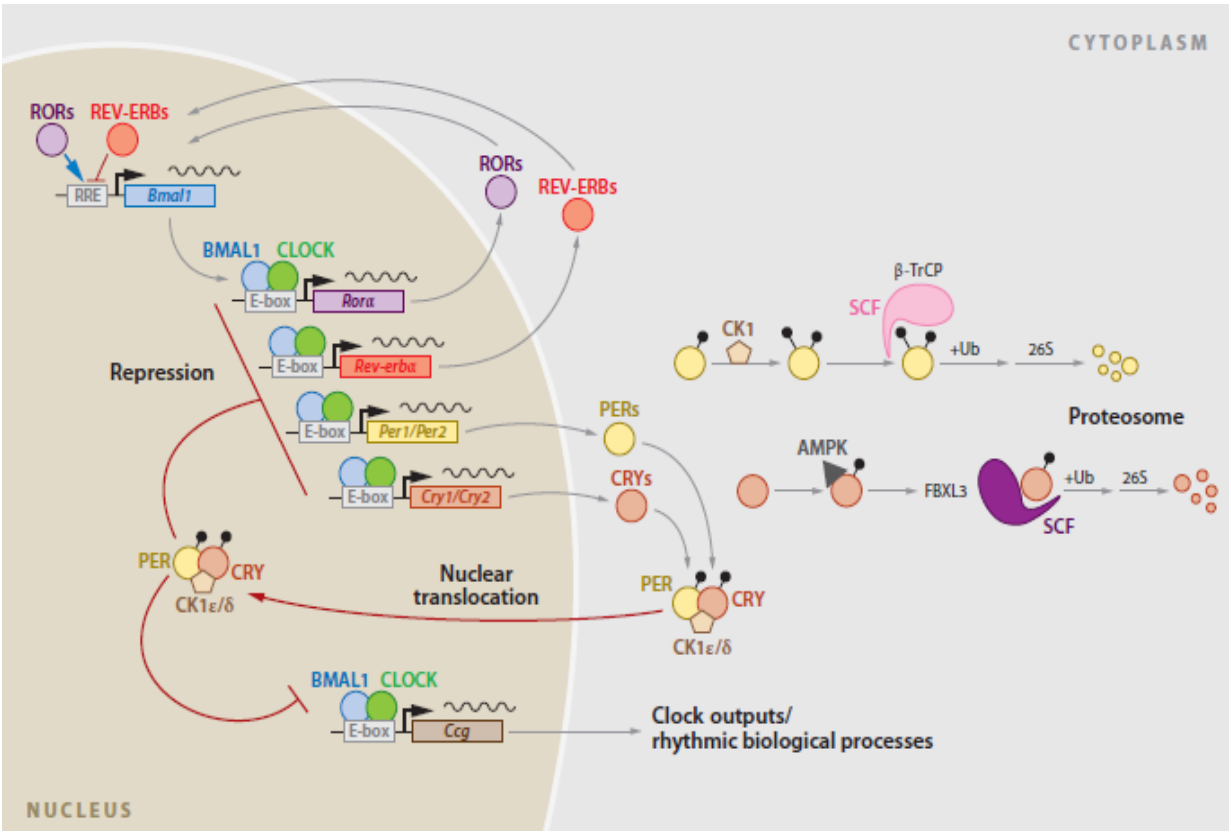


Fig 1.5. Proposed model for the circadian system in mammalian cells. In the nucleus the Clock/Bmal1 dimer binds to a E-box of several genes to activate their expressions such as Per1, Per2, Cry1, and Cry2. The Per/Cry protein dimers are phosphorylated in the cytoplasm by kinases such as casein kinase Iε /δ (CKIε /δ). Then the Per/Cry dimers translocate to the nucleus in a phosphorylation-regulated manner where they interact with the Clock/Bmal1 complex to repress their own activators. In addition to this core transcriptional feedback loop, there are other feedback loops involving Rev-erbα and Rorα that repress Bmal1 transcription. CLOCK/BMAL1 also regulates many downstream target genes known as clock-controlled genes (*ccgs*) (Mohawk et al. 2012).

1.6. Circadian rhythms and human health

Circadian variation can be observed in the human body. Circadian regulation can be seen in behavioral, neural, physiological and endocrine functions. Blood pressure, body temperature, respiration, awake-sleep cycles and hormone levels (cortisol and melatonin) are some of the circadian rhythms that are manifest through numerous physiological actions. These circadian rhythms have an impact on the health of a person (Levi et al. 2007, Hu et al. 2008).

Jet lag is a common effect of these rhythms and is a result of the variance between the genetic and the environment time. Stroke, sudden cardiac arrest that leads to death, cardiac arrhythmia and severe myocardial infarction are all cardiovascular illnesses that are linked to circadian rhythms and they mostly take place in the morning but rarely they happen during sleep time. Shift-work is another example whereby the chance of cardiovascular related deaths and cancer cases in human and hamsters is significantly increased by the desynchronization of the circadian clock with environmental shifting (Penev et al. 1998). Diseases such as ischaemic stroke, that shares risk factors with cardiovascular disease, are known to happen more in long-term female shift workers (Brown et al. 2009). There is also a correlation between breast cancer risk and long-term shift working (Schernhammer et al. 2001, 2006).

1.7. *Neurospora*

Neurospora is a filamentous fungus, a heterothallic species that produces numerous nuclei in distinct hyphae. It has two unique cycles in its life, the diploid sexual cycle and the haploid asexual cycle. In the asexual part of the cycle, the macroconidia germinate in favourable growth conditions to create the first germ tube. It later continues to develop by the tip extending and branching out to produce the threads of hyphae that do not have any cross walls and contain many haploid nuclei. A colony is able to generate millions of multinucleate conidia from aerial hyphae, and dispersal of these conidia can repeat the asexual cycle if they are located in suitable conditions (Davis, 2000).

The two mating types are determined by alternative DNA sequences at one chromosomal locus; these are called MAT A and MAT a. When colonies of different mating type come into contact, the sexual phase occurs. After fusion of cell walls and nuclei a diploid nucleus forms in a fruiting body called perithecia. These diploid nuclei undergo meiosis and four haploid nuclei are generated and they undergo mitosis and make eight ascospores that are placed in an ascus. Ascospores germinate and produce hyphae resulting in colonies exactly like those produced by asexual spores (Davis, 2000).

Neurospora is one of the best-studied microorganisms in circadian biology. Because it is a haploid organism, it is easy to isolate new mutations. The genome has been sequenced (Galagan et al. 2003), and annotation can be found online at Broad Institute Database (http://www.broadinstitute.org). Knockout strains can be generated (Colot et al. 2006), and they are available from the Fungal Genetic Stock Center (www.fgsc.net).

After *Drosophila*, which was the first organism in which a circadian clock mutation was identified (Konopka and Benzer, 1971), *Neurospora* was the second organism (Feldman and Hoyle, 1973). The circadian output in *Neurospora* can be observed as rhythmic conidia formation that can be visually seen as bands of conidia and interbands.

The first identified clock gene in *Neurospora* was frequency (*frq*). There are seven *frq* mutations that influence the rhythm in *Neurospora* which are frq^{1}, frq^{2}, frq^{3}, frq^{7}, frq^{9}, frq^{10} and frq^{11}. When strains are carrying these mutations, the period is changed or they are arrhythmic or have no temperature compensation (Morgan et al. 2001). There are many other clock-affecting mutations that are involved in the clock or can affect the clocks such as period (*prd-1*, *prd-2*, *prd-3*, *prd-4* and *prd-6*).

1.8. *Neurospora* FRQ/WCC feedback loop

Similar to the circadian oscillator in *Drosophila* and mammals, the core oscillator of *Neurospora* consists of a negative feedback loop (Aronson et al. 1994), in which FRQ (frequency), FRH (FRQ-interacting RNA helicase), WC-1 (white collar 1) and WC-2 are the main components (Fig.1.6). This oscillator is called the FRQ-WC oscillator (FWO). WC-1 and WC-2 are PAS (PER-ARNT-SIM) containing transcription factors that have GATA Zn finger DNA binding domains and are localized in the nucleus (Ballario et al. 1996, Linden et al. 1997). A complex of FRQ and FRH is the negative element, and WC-1 and WC-2 are positive elements (Aronson et al. 1994, Cheng et al. 2005). WC-1 and WC-2 form a heterodimeric complex (WCC) that binds to the Clock box (C box) in the *frq* promoter. After the FRQ protein is synthesized, it dimerizes with itself and forms a complex with FRH (Cheng et al. 2005, 2001).

The FRQ-FRH complex (FFC) inhibits WC complex activity by interacting with the WC proteins. FRQ stability is important for period length. When FRQ is synthesized it is progressively phosphorylated by kinases like CKI and CKII and when it is hyperphosphorylated it goes through the ubiquitination pathway and degradation. When the level of FRQ is low, WCC is no longer inhibited by FRQ-FRH complex and a new cycle starts (Liu et al. 2006). Therefore, three essential phases in the circadian rhythm in *Neurospora* are activation of *frq* transcription by WCC, inhibition of WCC by FFC and degradation and phosphorylation of FRQ.

In *Neurospora*, light acts rapidly through the WCC to induce high levels of *frq* expression. Light induction of *frq* requires two cis-acting elements (LREs) in the frq promoter. The LREs are bound by a WC-1/WC-2 complex (WCC), and light causes WCC to bind to the LREs. WC-1 is the blue light photoreceptor that with its cofactor FAD (flavin adenine dinucleotide) and WC-2 mediates light input to the circadian system (Dunlap et al. 2004).

There are some highly conserved components from *Neurospora* to mammals that are essential for post-translational regulation. Phosphorylation of FRQ and PER (although they are not homologs) occurs with kinases CKI and CKII and they are dephosphorylated by phosphatases PP2A (Akten et al. 2003, Gorl et al. 2001, Kloss et al. 1998).

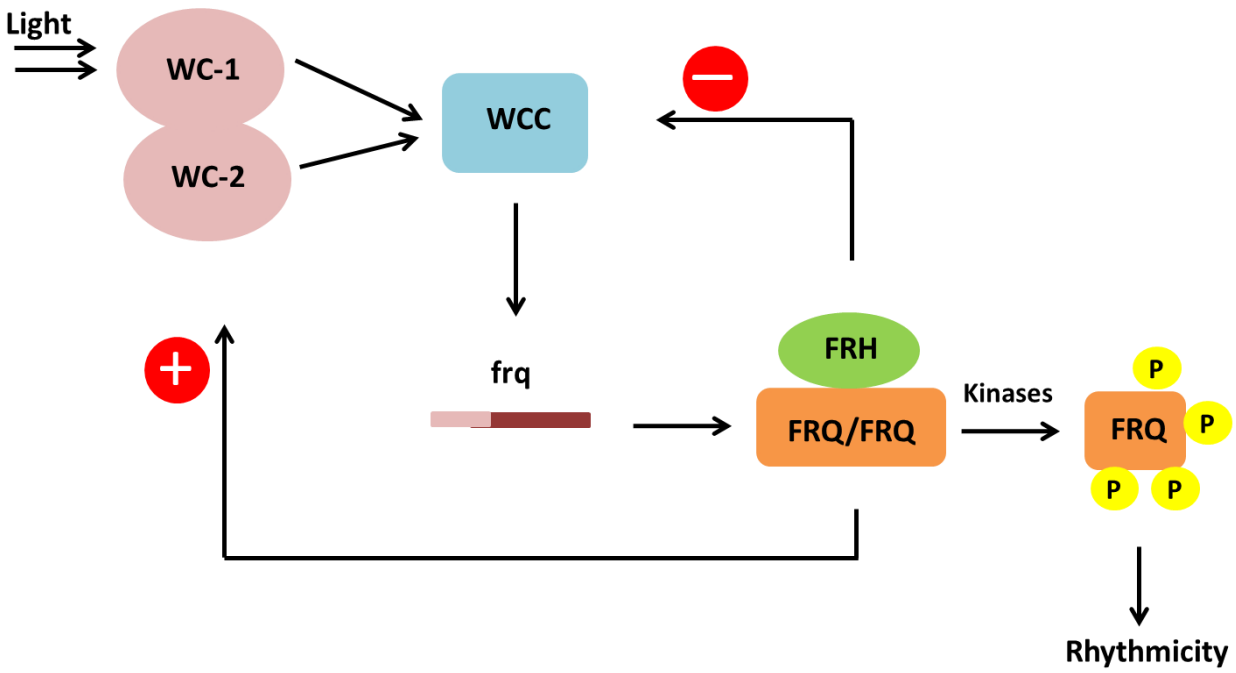


Fig 1.6. FRQ/WCC negative feedback oscillator in *Neurospora* (Liu & Bell-Pederson, 2006)

1.9. FRQ-less oscillator

Numerous studies demonstrated that rhythmic conidiation can be seen in the absence of clock gene function. The first report of rhythmicity in *frq* and wc null strains was observed when they were grown in extra-long race tubes (Aronson et al. 1994; Loros and Feldman, 1986). This means that in addition to the FRQ-based oscillator, there may be one or more FRQ-less oscillators that are involved in driving rhythmic conidiation (Froehlich et al. 2003; de Paula et al. 2006). The FRQ-less oscillator (FLO) does not have

circadian properties including temperature compensation and light entrainment and it also shows more variable periods compared to wild type (Dunlap and Loros, 2004).

The presence of a FLO (or FLOs) was also suggested by finding that *frq*, *wc-1* and *wc-2* null mutants that are arrhythmic in standard conditions, showed rhythmicity when they are grown on media enriched with farnesol or geraniol (Granshaw et al. 2003). Rhythmicity was restored when double mutants are constructed of clock genes with mutations that affect lipid synthesis like *cel* or *chol* mutations (Lakin-Thomas and Brody 2000). The *chol-1* (choline) mutation has effects on lipid synthesis and the *cho-1* mutant needs choline supplement for normal growth. Choline depletion inhibits growth and lengthens the period and also abolishes the temperature compensation of the conidiation rhythm (Lakin-Thomas, 1996, 1998).

Mutations in clock genes can be separated into two groups: the ones that do not have any effect on rhythmicity in the absence of the FWO oscillator and the ones that do have an effect. The second group may be the genes that are involved in FLOs or can affect the interaction between FWO and FLOs. Some of these mutations that have clock effects in the absence of FWO, are *ult* (Lombardi et al. 2007), *cel* and *chol* (Lakin-Thomas and Brody 2000), *UV90* and *prd-1*, *prd-2* and *prd-6*.

1.10. Effects of *prd* mutations

The effects of *prd-1*, *prd-2* and *prd-6* on rhythmicity in the absence of a functional FWO indicates that they have effects on the function of FLO. Other prd mutations have been shown to have no effect.

Previously it was shown that *prd-1* mutants can change the fatty acid composition (Cote and Brody, 1987) and it is likely that has an effect on lipid oscillations. The *prd-1* rhythm phenotype of the mutant is a long period in *frq*+ strains (approximately 25.8 hours) and also a long period of the FRQ-less rhythm with geraniol, and loss of the FRQ-less rhythm in *chol-1*. *Prd-1* also affects the FLO in *cel* (Lakin-Thomas and Brody 1985) and affects the heat entrainable FLO (Li and Lakin-Thomas 2010). The gene was not identified previously (Li and Lakin-Thomas 2010, Lombardi et al. 2007). It is described as a recessive mutation on chromosome III (Perkins et al. 2001).

The *prd-2* gene has not been identified yet but the *prd-2* mutation has the effects of a long period of conidiation rhythm in frq^{+} strains, a long period of FRQ-less rhythms with geraniol, loss of the FRQ-less rhythm in *chol-1*, and severe effects on the heat-entrainable FRQ-less rhythm (Li and Lakin-Thomas, 2010, Lombardi et al. 2007).

The *prd-3* gene was identified as the catalytic subunit of casein kinase 2 (CKII). Effects of this mutation are a long period of the conidiation rhythm, altered temperature compensation, loss of FRQ protein rhythm, and RNA rhythms of *frq*, *ccg-1*, and *ccg-2* (Mehra et al. 2009), (Yang et al. 2002).

The *prd-4* gene was identified as checkpoint kinase 2 (CHK2). The *prd-4* mutation can increase phosphorylation of FRQ (Pregueiro et al. 2006). It has a period-shortening

effect in *frq*$^{+}$ strains. It is known that *prd-3* and *prd-4* have slight effects on the free-running long period conidiation rhythm in *frq*^{10} *chol-1* strains on low choline and the entrained conidiation rhythm in the *frq*^{10} strain.

The *prd-6* mutation has been reported to be a mutation in a gene coding for a type 1 RNA helicase involved in translational termination and RNA metabolism (Compton and Feldman, 2004) and it shortens the period of the FRQ-less rhythm with geraniol (Lombardi et al. 2007).

Using microarrays, three clock-controlled genes (ccgs) were identified that display rhythmic mRNA levels in FRQ null strains. It is suggested that these ccgs are regulated by FLO (Correa et al. 2003). It was also shown that *ccg-16* (which is the most robustly rhythmic gene among ccgs) is rhythmic when FRQ levels are high; suggesting that the FLO oscillator that drives *ccg-16* rhythms is not inhibited by high or low FRQ levels. As was mentioned the FRQ-based oscillator required the activity of WC-1 and WC-2 (Loros and Dunlap, 2001; Dunlap and Loros, 2004). However, *ccg-16* rhythms do not require FRQ but do require WC-1 and WC-2 (de Paula et al. 2006). According to several studies it is proposed that in *Neurospora* at least three oscillators are running rhythmicity: the FRQ-based oscillator (Loros and Dunlap, 2001; Dunlap and Loros, 2004), FRQ/WC-independent or FLO (Merrow et al., 1999; Lakin-Thomas and Brody, 2000; Granshaw et al., 2003; Christensen et al., 2004), and also WC-FLO (de Paula et al. 2006).

Studies indicated that four *prd* mutations can have effects in FRQ-less rhythms in two conditions (*prd-1* and *prd-2* have the greatest effect, *prd-3* and *prd-4* have less effect): the free-running long-period conidiation in *frq*^{10} *chol-1* on low choline and the heat-entrained conidiation rhythm in *frq*^{10} strains. Therefore, *prd* mutations influence FLO with and

without choline starvation (Li et al. 2010). Results were obtained from Lombardi et al that *prd-1* and *prd-2* altered the period in geraniol and farnesol supplementation but they did not abolish the FRQ-less rhythmicity (Lombardi et al. 2007).

Another mutation that affects two FRQ-less rhythms similar to *prd-1* and *prd-2*, is *UV90*. It can influence the free-running rhythm under low-choline conditions and the heat-entrained rhythm in the choline-sufficient condition. It is not known if there are multiple FLOs or not, and each of the products of *UV90*, *prd-1* and *prd-2* may have different targets that affect the function of independent FLOs. In the simplest model there is a single FLO, and it is assumed that *UV90* and *prd-1* and *prd-2* gene products have one major target and function (Li et al. 2011).

In the proposed model for *Neurospora* oscillators, FRQ/WCC and FLO independently affect the rhythmicity. FLO is always functioning but its output and low amplitude are masked by FRQ/WCC. In the absence of FRQ/WCC and addition of different conditions like farnesol and *chol-1*, FLO exhibits higher amplitude and can drive the conidiation rhythm or other rhythms like nitrate reductase (NR) (Christensen et al. 2004) and *ccg-16*, shown in figure 1.7 (de Paula et al. 2006).

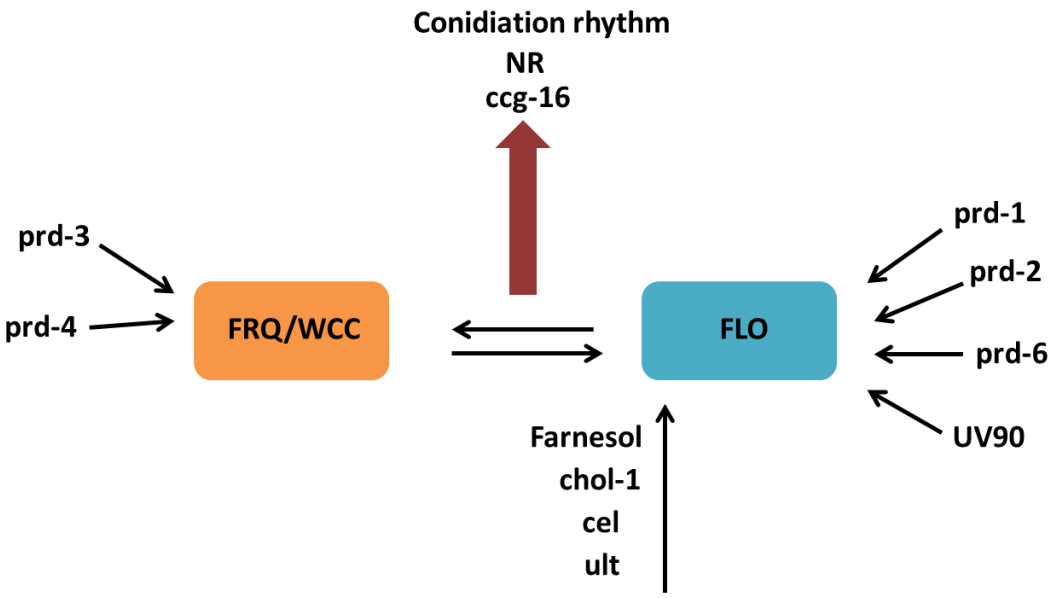


Fig 1.7. The proposed model for circadian oscillators in *Neurospora*. This model is based on a single FRQ-less oscillator that is in interaction with FRQ/WCC oscillator.

1.11. Objective of my project

The *prd-1* gene is known as part of FLO that affects both the period and temperature compensation. Previously mapping the *prd-1* gene was started using rapid genetic mapping techniques described by Jin et al (2007). Experiments showed that the *prd-1* gene is located on the right of the AluI(a) marker on chromosome III with a recombination rate of 2.2%.

My project aimed to confirm the location of the gene and also identify and characterize it. In order to do that, PCR and digestion by different enzymes was done to confirm the location of the gene. Genes in this interval were screened to check their phenotype. Candidate genes were sequenced to find the mutation. Several *prd-1* and prd^{+} strains were sequenced to confirm that the mutation exists in all *prd-1* strains and not in prd^{+} strains. cDNAs of *prd-1* and prd^{+} were compared. The candidate gene for *prd-1* was cloned in the *his-3* targeting plasmid and then complementation was carried out. BLAST search showed that the *prd-1* gene is an RNA helicase that has orthologs in *Saccharomyces.*

2. Methods and Materials

2.1. *Neurospora* strains

All *Neurospora* strains used in this study carry the backbone genotype *bd*; *csp*-1 as described previously (Lakin-Thomas, 1992). The genotype of *bd*; *csp*-1 will be referred to as WT (wild type) throughout this thesis. The *bd* mutation helps to express the conidial banding pattern when cultured in growth tubes or other closed culture vessels by reducing the sensitivity to CO_2 buildup (Davis, 2000). This mutation does not have any effect on the circadian clock. The *csp*-1 (conidial separation) mutation also has no effect on clocks and it is used to keep the conidiospores from detaching to limit contamination (Lakin-Thomas 2006). The *chol-1* mutation is deficient in an enzyme required for phosphatidylcholine biosynthesis. For normal growth and development it requires choline supplement in the growth medium. Choline starvation reduces its growth rate and lengthens the period (from 21.5 hours for WT to 60 hours).

2.2. Growth and Harvesting of Strains

In order to obtain DNA, strains were grown in liquid minimal media which is 50X vogels and glucose and water (Appendix-1). One ml of media was pipetted into microtiter wells and a small sample of conidia of each strain was inoculated using a sterile loop. The samples were placed at 30^{o}C for 2-3 days and they were harvested by vacuum filtration, rinsed with distilled water and covered with foil and stored at -80^{o}C.

2.3. Isolation of DNA

The method was adapted from Jin et al (2007). To isolate DNA for PCR, harvested fungi were placed in liquid nitrogen, and a cooled mortar and pestle were used to grind the frozen mycelia to fine powder. To each sample, 0.6 ml of extraction buffer (Appendix-1) and 3µl of proteinase K were added. After incubation in a water bath at 65^{o}C for 1 hour, 0.2 ml of 7.5M ammonium acetate was added and samples were incubated on ice for 5 minutes. Samples were centrifuged for 3 minutes at 16000 g and supernatants were transferred to fresh tubes. RNAase A (5mg/ml) was added and they were incubated at 37^{o}C for 1 hour, then 0.5ml of chloroform was added. After centrifuging the samples at 13250 g for 5 minutes the top aqueous layer was transferred to a fresh tube and 0.65 ml isopropanol was added. Samples were centrifuged again for 15 min at 13250 g. Isopropanol was removed and pellets were rinsed with 1ml of 70% ethanol. The pellets were dissolved in 0.1ml of TE buffer (Appendix-1). DNA concentration was measured by nanodrop spectrophotometer.

2.4. Primer preparation

Primers were ordered from Invitrogen and IDT (Integrated DNA Technologies). Primer tubes were centrifuged several times to collect primers. The “nmole” amount on the information that comes with the primers was used and the appropriate amount in µl of sterile TE was added to the primer tube. It was mixed and centrifuged briefly to collect the primer. The stock solutions of primer were 1 nmol/1µl.

For PCR a 10 µM concentration is needed, therefore each primer was diluted using 99 µl of sterile water + 1µl of primer to get 10 µM concentration.

2.5. Preparation of PCR samples

PCR recipe	Amount
Taq 10X Buffer	2 µl
Water (sterile,high quality)	14.52 µl
dNTP (10 mM)	0.4 µl
Taq polymerase (400 units)	0.08 µl
Forward Primer (10 µM)	1 µl
Reverse Primer (10 µM)	1 µl
DNA template (50 ng/µl concentration)	1 µl
	20 µl total

For CAPS markers, 50ng of DNA was used as a template. The first 4 ingredients were mixed in order to make a master mix. 17 µl of master mix was added to each tube and then 1 µl of Forward and Reverse primers and DNA template were added to each tube and mixed gently. The PCR program was based on New England Biolabs "*Taq* polymerase" PCR manual.

	Step	Temperature	Time
Initial DNA denaturing	Step 1:	94°C	3 min (30 seconds is typical, but 2-4 min is used to ensure separation of GC-rich sequence)
DNA denature (26 cycles)	Step 2:	94°C	15 sec for a DNA fragment less that 8-900 bp, for a little more that 1 kb use 30 sec
Primer annealing (26 cycle)	ii)	Varies*	15 sec for fragment less than 8-900 bp, for around 1 kb use 30 sec
Strand extension	iii)	72°C	1min (typically 1 min per 1 kb, use 1 min for smaller fragments)
Final extention	Step 3:	72°C	5 min (for less than 1000bp, for more use 10 min)

•The optimal annealing temperature is 3-5°C less than average Tm of primers that can be found from the manufacturer's information.

To get the PCR product for cloning, LA Taq (from TaKaRa) enzyme was used. LA Taq combines Taq DNA Polymerase and a DNA proofreading polymerase with 3' to 5'

exonuclease activity that is optimized for PCR amplification of very long DNA templates (long range PCR). The PCR recipe and program are presented below.

PCR recipe	Amount
LA Taq 10X Buffer	10 µl
Water (sterile,high quality)	54 µl
dNTP (2.5 mM)	16 µl
LA Taq polymerase (125 units)	2 µl
Forward Primer (10 µM)	8 µl
Reverse Primer (10 µM)	8 µl
DNA template (50 ng/ µl concentration)	2 µl
	50 µl total

Steps	Temperature	Time
Initial denaturation	94°C	5 min
Amplification (35 cycles)	94°C	40 sec
	60°C	40 sec
	72°C	3 min
Elongation	72°C	10 min

2.6. Restriction Enzyme Digestion

Using the New England Biolabs website (https://www.neb.com), it was determined if the enzyme needs Bovine Serum Albumin (BSA) and what temperature is needed for incubation and which buffer should be used. Different restriction enzymes from NEB were used.

Restriction Enzyme Recipe	with BSA	without BSA
Water (sterile, high quality)	Varies* (14.3 µl or 13.8 µl)	Varies* (14.5 µl or 14 µl)
Restriction Enzyme Buffer	2 µl	2 µl
restriction Enzyme	Varies* (0.5 µl or 1 µl)	Varies* (0.5 µl or 1 µl)
BSA (optional)	0.2 µl	0
DNA- PCR product	3 µl	3 µl
	20 µl	**20 µl**

The first 4 ingredients were mixed in this order: water, restriction enzyme buffer, restriction enzyme and BSA. When the master mix was ready 17 µl of master mix was added to 3 µl of PCR product and mixed gently. Samples were incubated in a water bath at the desired temperature for 1-1.5 hours.

2.7. Gel Electrophoresis

Small Gel (30 ml)		Large gel (150 ml)	
1X TAE Buffer	3 ml 10X TAE buffer + 27 ml water	1X TAE Buffer	15 ml 10X TAE buffer + 135 ml water
Agarose (1.5%)	0.45 g	Agarose (1.5%)	2.25 g
Ethidium Bromide (3%) or Red safe	EB: 0.9 µl Red safe: 3 µl	Ethidium Bromide (3%) or Red safe	EB: 4.5 µl Red safe: 8 µl

The appropriate amount of 10X TAE buffer (Appendix-1) and water were mixed to get 1X TAE, then agarose was added and heated in a microwave until the agarose was well dissolved. After cooling to about 55^{o}C, ethidium bromide or red safe were added and it was poured into the gel mold. After 20-30 minutes it was ready to use.

1X TAE was poured into the electrophoresis chamber and the gel was covered with buffer. Each sample was mixed with Loading Dye and loaded into wells. The gel was electrophoresed until blue dye had run about halfway down (small gel: 70 volts, ~ 45 min; large gel: 100 volts, ~ 45-60 min).

Undigested samples		Digested samples		Ladder	
6 X Loading Dye	2 µl	6 X Loading Dye	3 µl	6 X Loading Dye	1 µl
PCR product	10 µl	PCR product	15 µl	Ladder	1 µl
				Water	4 µl
	12 µl		18 µl		6 µl

2.8. Knockout screening in race tubes

Knockout strains were ordered from the Fungal Genetic Stock Centre (http://www.fgsc.net/) and were inoculated in race tubes, which are glass 30 cm tubes that are open at both ends. MA media was used for race tubes (Appendix-1). The race tubes were placed in light and 30^{o}C for one day and then they were transferred to constant dark and 22^{o}C. The growth front was marked every day and the exact time and date of marking were recorded. After the growth was finished the tubes were scanned and growth rate and period were calculated using lab generated software called Mac Tau.

2.9. Gene sequencing

In order to sequence a gene, the sequence and annotation were found in the Broad Institute Database (http://www.broadinstitute.org). For sequencing, upstream and downstream sequences also were used and were found from the same database. The primers were designed using ApE software (http://biologylabs.utah.edu/jorgensen/wayned/ape/). Primers with GC content almost 50%, Tm less that 60^{o}C and length of 18-22 nt were chosen. Primers Forward-1 and Reverse-1 were used for PCR and several primers were designed (every 500nt-800nt) to cover the whole gene.

PCR was carried out to amplify the gene using LA taq (as described in section 2.5). The PCR product was electrophoresed through a 1% agarose gel to check if the product size was what was expected. The PCR product was gel purified using the QIAGEN Gel Extraction kit according the manufacturer instructions. PCR product (5 µl) and 4µl of each primer were sent to the York Biology Molecular Core facility for sequencing. Sequencing results were aligned with the original sequence that was found from Broad

database using ApE software. According to Lee Wang in the core facility, the sequence for the first 20 and last 20 bases are not reliable. Therefore, if a missing base was seen at the beginning or at the end, sequencing was repeated or overlapped by the next primers to get more reliable results. Table 2.1 presents all the primers for PCR and sequencing the candidate *prd*-1 genes and *msp-8* gene.

Name	Sequence 5' ----> 3'	Length
5703-Forward 1	GCGGATCCGGTCTCCCCCTGGTCCCG	26 bp
5703-Forward 2	CATCCAGCCATCTGGGTCATC	21 bp
5703-Forward 3	CTCATCTCAGTCACTCAGTCA	21 bp
5703-Reverse 1	TCCGGGCCCTCTCAGAAACATTGCTTA	27 bp
5703-Reverse 2	GATGACCCAGATGGCTGGATG	21bp
7845 F-1	TGTCTCCATTACTCCCTTGAAC	22 bp
7845 R-1	GGGAAGCAGATGGACTATGTATG	23 bp
7824 F-1	CGAAGTCAACTGAAGCGATAGC	22 bp
7824 F-2	GGATGGAGGGAATACAGATCAG	22 bp
7824 R-1	TACGGCCCAGTAAGATTAGACG	22 bp
7817 F-1	ACGCAACTCAACCACCATTCC	21 bp
7817 F-2	TCGGATAAGCTGGTGTGGTCG	21 bp
7817 F-3	CGTAAACGCCTTGAGTAGCTGTG	23 bp
7817 F-4	GGTGGGAAGAAACGAAACGAACG	23 bp
7817 R-1	AAGGCGCATTGCTTTACACG	20 bp
7817 R-2	TTAGCGCCTTCTCCAGGTTTC	21 bp
7839 New F-1	TGCCTTATCGGCGCTTTATC	20 bp
7839 F-1	CTCAAATCCGCCTTCGCTGTAAC	23 bp
7839 F-2	ATCCGCATCTCACATCCACTAGC	23 bp
7839 New F-2	CACCAATCCTCCATCGACATCTC	23 bp
7839 F-3	GCGGCTACTCTAACAACTACTCC	23 bp
7839 New F-3	CGAACCTGCTCATTTCCTCTCTC	23 bp
7839 F-4	CGTGAGTTGGCTGTCCAGATTC	21 bp
7839 New F-4	TTCACATCAACGCTCAGCCTC	21 bp
7839 F-5	TCAAGACTGGCAAGAGCCCTATC	23 bp
7839 New F-5	TTGTCACCGAGAGAGGGTAGTAG	23 bp
7839 F-6	GTGTCACCTGAATCTGCCCATC	22 bp
7839 F-7	ATTGCTTTCAGACGCTCACTGG	22 bp
7839 F-10	CCCAACAACTCCGAGGATTAC	21 bp
7839 R-1	ACCTTGGAACGCAGCACAAATC	22 bp
7839 R-2	CTCATCGACGTGAAACCTACAGC	23 bp
7839 R-3	GCCGAACTTGGAGATTTCTTGC	22 bp
7839 R-4	ACCTCCACCACCACTGTAGC	20 bp
7839 New R-1	TGGCGATACAGACCTCGACTC	21 bp
7839 New R-2	GCCGAACTTGGAGATTTCTTGC	22 bp
7839 New R-3	AGCCTTGACCTCGTCCATAAC	21 bp
7839 New R-4	ATGTCCCAGTTCTGCTTCTGC	21 bp
pBM61 F-1	ATTCAGACCCCATTAGCCGTCCACGCC	27 bp
pBM61 R-1	GACGGGGTAGCTTGGCCCTAATTAACC	27 bp
1143 F-1	ACCTCTAGGTACTGATGAGCCAC	23 bp
1143 F-2	CTCACGGTGCTTGTCATCAAAC	22 bp
1143 F-3	TCTGCAACGCTTGAGCAACC	20 bp
1143 F-4	GACCCGACTACAACACCATCAAG	23 bp
1143 F-5	CGGTGTTTCCTTGGCATTCAG	21 bp
1143 F-6	TATTGACCTCTGTCGCAAGCTC	22 bp
1143 F-7	GTGTCATCAAGCAGGCATTGG	21 bp
1143 F-8	AGGACTTGGAAAGAGCGCAAAC	22 bp
1143 F-9	ACATGACGGGCTTCAGCATTG	21 bp
1143 F-10	AGAGCAATGGGCATCAAGGAC	21 bp
1143 F-11	CCCGACAAGAAGTTTGCAGCATC	23 bp
1143 F-12	CTTGCATATCAACACGGCATGG	22 bp
1143 F-13	TAGTGTCTAGGAGCAGCAACAGG	23 bp
1143 R-1	GATAAACAGGCGACACCAAAGC	22 bp
1143 R-2	ACTTCGCCAAGCACGATTACG	21 bp
1143 R-3	AACTTGATGGTCTTGGCCTCC	21 bp
1143 R-4	ATGCGGCTGCCCATTCTTTG	20 bp
1143 R-5	ATCTTCATCTCCTCGCCTTCCAG	23 bp
1143 R-6	GTCTTTGGGTTCTTTGCGCTGAG	23 bp

Table 2.1. List of primers used for sequencing candidate genes for *prd-1* and *msp-8.*

2.10. RNA Extraction

RNA of SG#26 (*bd csp*-1) and SG#260 (*bd csp*-1 *prd*-1) was extracted. To extract RNA, the strains were grown on minimal liquid media, harvested and turned into powder by grinding in liquid nitrogen as described earlier for DNA extraction. 1 ml of Trizol was added to screw-cap tubes of each sample. After vortexing, 0.2 ml choloroform was added to each tube and tubes were shaken for 15 seconds. They were incubated at room temperature for 5-12 minutes and centrifuged at 15770 g for 15 minutes at 4^{o}C. The top aqueous layer was transferred to a fresh tube, 0.5 ml of isopropanol was added to precipitate RNA and tubes were centrifuged at 15770 g for 8 minutes at 4 ^{o}C. The pellet of RNA was washed with 1 ml of 75% ethanol and centrifuged at 6160 g for 5 minutes at 4 ^{o}C. Ethanol was removed and the dried pellet was dissolved in DEPC (Diethylpyrocarbonate) water. Samples were incubated in a water bath for 10-15 minutes at 55-60^{o}C and stored at -80^{o}C. The RNA concentration was measured using the nanodrop spectrophotometer. RNA concentrations for SG#26 and Sg#260 were 3.06 µg/µl and 3.44 µg/µl, respectively.

2.11. RNA to cDNA

To make cDNA, a high capacity RNA to cDNA kit from Applied Biosystems was used. All procedures were conducted on ice. A master mix of 2X RT buffer (10 µl for each sample) and 20X RT enzyme (1 µl for each sample) was made. 9 µl of nuclease-free water and RNA (8 µl of nuclease-free water + 1µl of RNA sample) were mixed with 11 µl of master mix to make 20 µl of reaction for each sample. The PCR reaction condition is shown below.

	Step 1	Step 2	Step 3
Temperature (^{o}C)	37	95	4
Time (min)	60	5	∞

After the reverse transcription was finished, PCR with cDNA samples was done. Samples were mixed with loading dye and loaded on a 3% agarose gel.

2.12. PCR product preparation for cloning

PCR was carried out using LA taq as in section 2.5, with primers designed to have restriction enzyme sites that are not found inside the gene so the enzyme will not cut the gene. The PCR product that was carrying EcoRI restriction site was used for cloning.

2.13. Cloning the gene into pSTblue-1 plasmid

A mix was made of 2 µl PCR product plus 5.0 µl End Conversion Mix plus 7.5 µl Nuclease-free Water to a total of 10 µl.

The reaction was incubated at 22°C for 15 min. The reaction was inactivated by heating at 75°C for 5 min.

1 µl (50 ng) Blunt Vector and 1 µl (4 U) T4 DNA Ligase were added directly to the end-conversion reaction, to a total volume of 12 µl and they were incubated at room temperature for 15 min. After ligation, 1 µl of the ligation reaction and chemically competent cells were used for transformation and cells were plated on kanamycin plates.

2.14. Electroporation using One Shot TOP10 Electrocompetent *E.coli*

TOP10 *E.coli* competent cells from Invitrogen were used for transformation. These cells have high efficiency for cloning. Along with TOP10 Electrocomp *E.coli* cells, pUC19 plasmid (10 pg/µl) was used as a control. S.O.C (Super Optimal broth with Catabolite repression) medium and LB + ampicillin plates (Appendix-1) were placed at

37°C to be warm for transformation. A vial of One Shot TOP10 Electrocomp cells was used for each transformation. 1-2 µl of the DNA (10 pg-100ng) was added to a vial of One Shot cells and mixed gently. 1 µl of DNA (10pg for pUC19 and 5ng for pBM61) was used. The cells were transferred to a chilled electroporation cuvette on ice. The electroporator was set for bacterial transformation (600Ω, 25 uF, 2.5V) and after electroporation 250µl of pre-warmed S.O.C medium was added to each vial. The solutions were placed on a shaker for 1 hour at 37°C to allow expression of the antibiotic resistance gene. 10-150 µl from each transformation was spread on a pre-warmed selective plate and inverted plates were incubated overnight at 37°C.

2.15. Transformation using chemically competent *E.coli*

E.coli cells DH5α from glycerol stocks were streaked onto LB agar plates and incubated at 37°C overnight to get isolated single colonies. A 50 ml overnight culture of *E.coli* in LB broth was grown (Appendix-1) at 37°C, diluted in 500 ml LB and grown at 37°C to OD_{600} of 0.3-0.4. The cells were centrifuged at 1000 g and 4°C, pellets were kept ice-cold from this point and they were resuspended in one tenth volume (50 ml) ice-cold TSS (Appendix-1) and stored at -80°C.

For each transformation 50 µl of chemically competent *E.coli* cells was used. Plasmid DNA (1 µl, 10 pg-100ng) was added to *E.coli* cells and incubated on ice for 20-30 min. Cells were heat shocked for 45 seconds in 42°C water bath and quickly put on ice for 2 minutes. One ml of LB (without antibiotics) was added and shaken at 37°C and 250 rpm (revolutions per minute) for 1 hour. The cells were plated on selective plates and incubated overnight at 37°C.

2.16. Plasmid purification

Colonies that were grown on plates were picked with sterile toothpicks and each colony was grown in 3 ml LB (with antibiotics) overnight on 37^{o}C. Using the mini prep QIAGEN kit, the plasmid DNA was purified according to the manufacturer's instructions.

2.17. Transformation of *Neurospora*

2.17.1. Preparation of plasmid DNA

The plasmid DNA was linearized with an enzyme that cuts once outside of the insert and *his-3* gene. The restriction product was cleaned up using the PCR purification QIAGEN kit according to the manufacturer's instructions. The concentration of DNA was measured by nanodrop spectrophotometer. Plasmid DNA (10-15µg) in 15µl water was needed for transformation, but the DNA was not concentrated enough so it was precipitated to concentrate it. Plasmid DNA was diluted with water to 100µl. Sodium acetate (10 µl, 3M, pH 8) and 200µl absolute ethanol were added and the solution was kept at -80^{o}C overnight. It was centrifuged at 18510 g, 4^{o}C for 5 minutes and the supernatant was removed. The pellet was washed with 75% ethanol, centrifuged and dried at room temperature 10-20 minutes and dissolved in 15µl sterile water.

2.17.2. Growth and harvest of conidia

The *his-3* strain was grown in 4-5 slants at 30^{o}C for 4-7 days. Five 1000ml flasks containing 100 ml medium for growing conidia (Appendix-1) were made and the contents of slants were inoculated by sterile loop. Cultures were grown in constant dark at 30^{o}C for 3 days and then constant light at room temperature for 5-8 days.

Water (50 ml) was added to each flask and using a sterile loop, the conidia were scraped off and poured into conical tubes. This was repeated with the other flasks and tubes were vortexed vigorously and were poured through autoclaved muslin in a funnel into fresh tubes. Tubes were centrifuged at 2730 g for 5 minutes, 25^{o}C. The supernatant was discarded and all pellets were pooled into one tube. The pellet was washed three times with 40 ml water and resuspended in residual water and kept on ice.

2.17.3. Transformation of spores by electroporation

Spores were centrifuged at 2730 g for 5 minutes at 4^{o}C and the supernatant was removed. One ml 1M sorbitol was added and after vortex and centrifuging it was removed. Sorbitol (40µl 1M) (Appendix-1) was added and spores were mixed and kept on ice. The results of transformation showed that 40µl 1M sorbitol was too viscous and transformation was not successful. Adding 2ml of sorbitol (1M) produced better results because the spore concentration was less.

Plasmid DNA (15µl) was added to 250 µl of conidia, mixed gently and samples were put on ice for 5 minutes. Conidia and plasmid DNA were transferred to a chilled electroporation cuvette using a pasture pipette. The electroporator was set for fungal transformation (voltage: 1.5V, capacitance: 25 uF, resistance: 600 Ω). After electroporation 1ml recovery medium (Appendix-1) was added to the cuvette and gently mixed by pipetting and the mixture was transferred to a 1.5ml tube. Tubes were put on a shaker at room temperature for 1 hour.

2.17.4. Growth of transformants

Two 1 ml dilutions of 1/10 and 1/100 of spores were made. Recovered spores (100 µl) were removed and 900 µl recovery medium was added to make a 1/10 dilution. From the 1/10 dilution, 100 µl was removed and 900 µl recovery medium was added to that to make a 1/100 dilution. Each dilution was mixed well with 60 ml plating medium (Appendix-1) and poured into three 10 cm petri plates (20 ml per plate). Plates were incubated face up at 30^{o}C for 3-7 days for colonies to grow.

Individual colonies were punched out with the narrow tip of a pasteur pipette into slant tubes with minimal medium (Appendix-1) (no histidine) and incubated at 30^{o}C to grow. To observe the phenotype of transformants they were inoculated in race tubes from slants and placed in constant light at 30^{o}C for a day and transferred to constant dark at 22^{o}C. Growth rate and period were calculated using Mac Tau software.

2.18. Microconidia preparation

The medium for microconidia (Appendix-1) was poured into 10 cm petri plates, 20 ml per culture. Plates were covered with sterile cellophane. The center of the cellophane was pierced by a sterile needle and inoculated with a few spores near the hole. Plates were incubated upside down at room temperature in a humid atmosphere for 7-12 days. The cellophane was removed and 2 ml of sterile water was poured into the plates to wash the surface; after a few minutes the water was removed with a sterile pasture pipette into sterile 1.5 ml conical tubes. Tubes were centrifuged at 21470 g for 1 minute at room temperature. Supernatant was removed and spores were resuspended in the residual water. Spores were counted using a hematocytometer and microscope. A volume that

contains 100 spores was calculated. Spores were plated in three different plates of 20, 100 and 500 spores per plate. Colonies were picked after 3-5 days of growth at 30^{o}C.

2.19. Screening the progeny from a cross

Crosses were set up by Dr. Lakin-Thomas and spores were picked from these crosses. From a plate of 4% agar, a slice was cut as big as a slide and transferred to a clean glass slide. A few drops of 1:10 diluted bleach was poured on the agar. Agar was cut in the shape of a grid that has 100 small squares. To pick spores, a flamed loop was used to pick up perithecia and they were placed on the agar slide under a dissecting scope. Using the loop, a perithecuim was broken and spores were separated and each spore placed on a small square of agar. Each square was placed into a baby tube containing minimal medium and choline (~ 170 spores were picked). Baby tubes were placed in a water bath at 55-60^{o}C for 30 minutes to heat shock to induce germination. Baby tubes were kept at room temperature for 7-10 days to grow.

In order to test baby tubes, the first characteristic to select is *csp* $^{+}$ or *csp*-1. By tapping the tube *csp*$^{+}$ and *csp*-1 progeny were separated. Tubes with flying spores are *csp*$^{+}$ and clumpy ones are *csp*-1. Each *csp*-1 progeny was inoculated in 2 race tubes containing MA medium (one with choline and one without choline). The *chol*$^{+}$ and *chol-1* progeny can be separated by growth rate with and without choline. Using a hygromycin resistance test, it can be determined which progeny have the knockout background. Each progeny was inoculated in two stock tubes with 1ml minimal liquid media (one with hygromycin and one without hygromycin). They were placed on a shaker for 2-3 days at room temperature. Progeny that grew in hygromycin media (*hyg*^{r}) have the background of the knockout parent. To make a knockout strain, the knockout gene is replaced by the

hph cassette which confers hygromycin resistance (Colot et al., 2006). Observation on race tubes of growth rate and period were used to identify *bd* or *prd* mutations.

3. Results

3.1. CAPS markers and mapping

In order to map the prd-1 mutation, mapping was performed based on single nucleotide polymorphisms (SNPs). In this technique co-segregation analysis of progeny from a single cross between a wild type and mutant strain from different backgrounds was done. The Oakridge and Mauriceville strains have many SNPs (Jin et al. 2007).

The wild type strain Mauriceville (ras^{bd+}, $csp\text{-}1^{+}$, $prd\text{-}1^{+}$) and Oakridge strain (ras^{bd}, csp-1, prd-1) were previously crossed by Kamyar Motavaze and progeny were examined with Cleaved Amplified Polymorphic Sequences (CAPS) markers. The markers were designed to have SNPs that also act as a restriction enzyme site (Jin et al. 2007). The progeny with Oakridge background have the restriction site and a PCR product that contains the SNP will be cut while Mauriceville stays uncut. Markers in different locations were chosen and DNA of Mauriceville and Oakridge parents as well as prd-1 progeny and prd+ progeny that showed recombination in this region were examined.

If progeny with a prd^{+} Mauriceville phenotype are digested it means that they have the SNP from the Oakridge parent and recombination had occurred during meiosis. Similarly, a strain with prd-1 Oakridge phenotype that remains uncut after restriction enzyme digestion means that it has DNA of Mauriceville background. To map the gene, the recombination frequency was calculated to find the approximate distance between markers. The higher the recombination rate, the further the marker is from the mutant gene.

From previous experiments done by Kamyar Motavaze it was found that the prd-1 gene should be located between AluI (a) marker and the centromere (Fig.3.1). The purpose of this experiment was to confirm the location of prd-1. Therefore, based on previous experiments, primers were chosen to do PCRs and digestion with different restriction enzymes (Table.3.1). Progeny were chosen that should have a recombination event in this region, based on previous results with CAPS markers mapping. A summary of digestions is shown in Table 3.2. My result confirmed the location of the gene. The results with the TSP 509 I enzyme were not consistent with Kamyar's result. The TaqI(a) and TaqI(b) results were also not consistent with Kamyar's results probably due to using old enzymes. The TaqI(a) and TaqI(b) markers also did not give consistent results with the control strains: all Oakridge controls should be cut but some were uncut. Therefore the results with TSP 509I and TaqI(a) and TaqI(b) were disregarded. The rest of the digestions showed that prd-1 should be somewhere between HaeIII and MspI markers. The PCR results and digestions with different restriction enzymes are shown in figures 3.3 - 3.11.

Marker	SNP Detail	Position	Primer sequence 5' --->3'	Primer's Name	PCR size	Enzyme site
MseI	3_641804_A → G_MseI +	640804-642804 +	F- GCTAGTCTCCTTGAGATCGTCCC R- GTGACGATCAGCGCAAACT	641804-F 641804-R	359 bp	TTAA
TaqI (b)	3_1212329_C →T_TaqI +	1211329 - 12133329 +	F- CATTGCAAGTCCTCTCCAAA R- GTTCTGTGCGTGGAGTACAAC	1211329-F 1211329-R	358 bp	TCGA
AluI (b)	3_1502999_C →T_AluI +	1501999 - 1503999+	F- CATATGTATCCGTCTACGCCAC R- GGCTGGGTCTCGTTTTCTCT	1501999-F 1501999-R	358 bp	AGCT
HealII	3_1083480_C →T_HaeIII +	1082480 - 1084480+	F- GACACTGGATTGTAACTTCGACG R- CATCCGAGATTTTCTCACAAGAA	108348-F 108348-R	354 bp	GCCC/CCGG
MspI	3_545271_G →A_MspI +	545000-546000 +	F- AACACGCTATTCGCTCATTTG R- CGTTTTCGCTGTCTTGTTCTGTG	545000-F 545000-R	957 bp	CCGG
TaqI (a)	3_578038_C →T_TaqI +	577600-579000 +	F- CAGTACTTGTAGATAGTCTTGCATCG R- GCACTCCTCGGTGTAGTACACACTT	577600-F 579000-R	1350 bp	TCGA
BstUI	3_597381_C →G_BstUI +	597000-598400 +	F- TGTAAGGAAAGCTGAGATATCCTAG R- CGATTTTCAGAGAGAGCTTGTACGG	597000-F 598400-R	1351 bp	CGCG
Tsp509I	3_515957_C →T_Tsp509I +	515731-516168 +	F- GTGAGCCAGATGACGTTGTG R- GTTCAGCCCAGCTTCCCTAT	P-21 Q-21	398 bp	AATT
AluI (a)	3_639318_C →T_AluI +	639225-639524 +	F- TGGGCGGTCAACTCCAAAGG R- CAGGGCAACTTGTTTGGGCT	3-17-AluI-F 3-17-AluI-R	260 bp	AGCT

Table 3.1. Markers with SNP details and primer sequence. Numbers under SNP Detail identify the position of the SNP on linkage group 3.

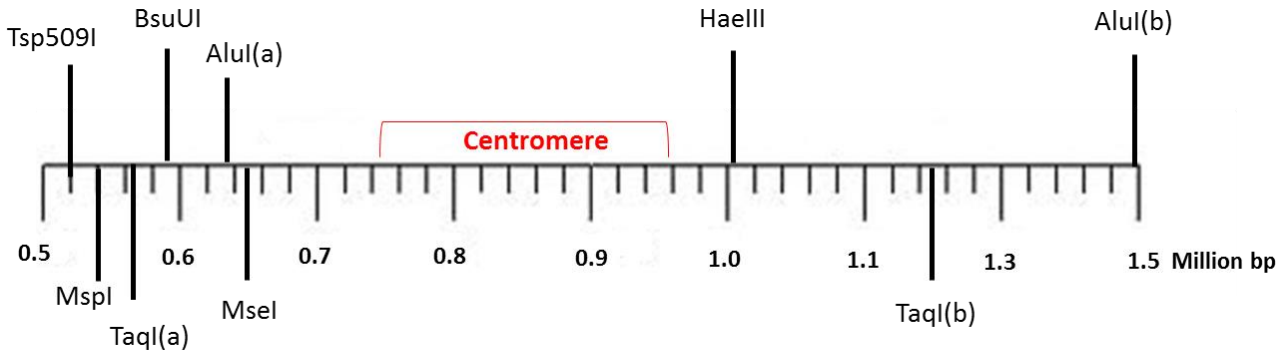


Fig 3.1. Marker locations on linkage group III of *Neurospora* crassa. It was predicted that the prd-1 gene is between HaeIII marker and MspI marker.

			Restriction Enzyme						
Progeny #		**Phenotype**	**Tsp509I**	**MspI**	**BstUI**	**AluI (a)**	**MseI**	**HaeIII**	**AluI(b)**
17	**Progeny from the cross**	prd-1	Cut	Uncut	Cut	Cut	Cut	Cut	Cut
44		prd-1	Cut	Uncut	Cut	Cut	Cut	Cut	Cut
46		prd-1	Cut	Uncut	Cut	Cut	Cut	Cut	Cut
A7		prd-1	Uncut	Uncut	Cut	Cut	Cut	Cut	Cut
A57		prd+	Uncut	Uncut	Uncut	Uncut	Uncut	Cut	Cut
SG#258	**Controls**	Mauriceville	Uncut	Uncut	Uncut	Uncut	Uncut	Uncut	Uncut
SG#259		Oakridge (OR)	Cut		Cut	Cut	Cut	Cut	Cut
SG#1		prd+ (Oakridge background)	Cut	Cut	Cut	Cut	Cut	Cut	Cut
SG#250(a)		prd -1(Oakridge background)	Cut	Cut	Cut	Cut	Cut	Cut	Cut
SG#250(b)		prd-1 (Oakridge background)	Cut		Cut	Cut	Cut	Cut	Cut
SG#80		prd+ (Oakridge background)			Cut	Cut		Cut	
SG#213		prd+ (Oakridge background)			Cut	Cut		Cut	

Table 3.2. PCR and digestion results of different CAPS markers. Yellow is Mauriceville sequence and purple is Oakridge sequence. TaqI(a) and TaqI(b) column were not included in the table because the control results were not consistent. Orange is the inconsistent result with previous result.

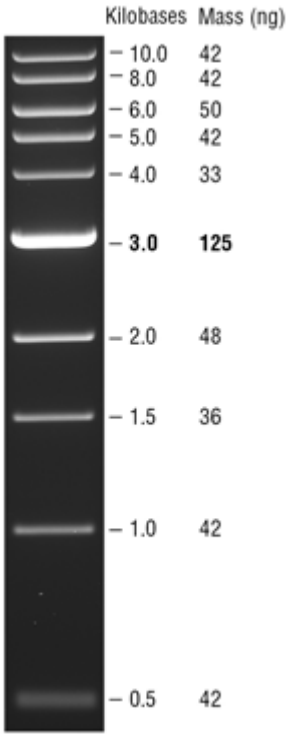


Fig 3.2. 1 kb DNA ladder from New England BioLabs was used in all experiments.

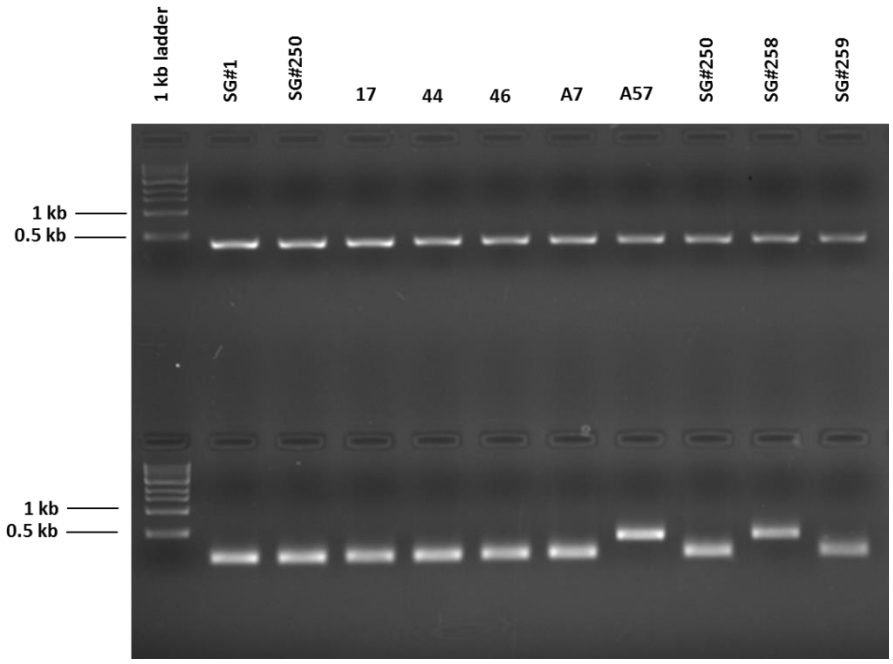


Fig 3.3. PCR for marker MseI. Top lanes are PCR product before digestion, bottom lanes are after 1 hour digestion by MseI enzyme. A57 and SG#258 stay uncut and this result is consistent with previous results obtained by Kamyar Motavaze.

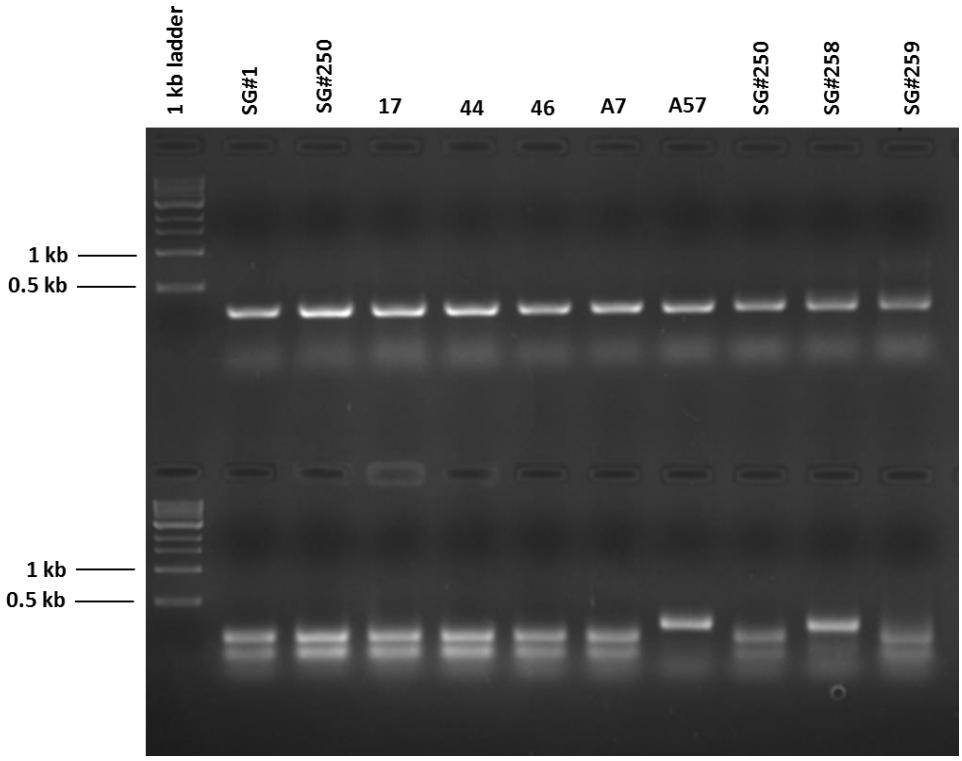


Fig 3.4. PCR with primers 3-17 AluI-F and 3-17 AluI-R for marker AluI(a). Top lanes are PCR product before digestion and bottom lanes are after 1 hour digestion by AluI. A57 and SG#258 are uncut.

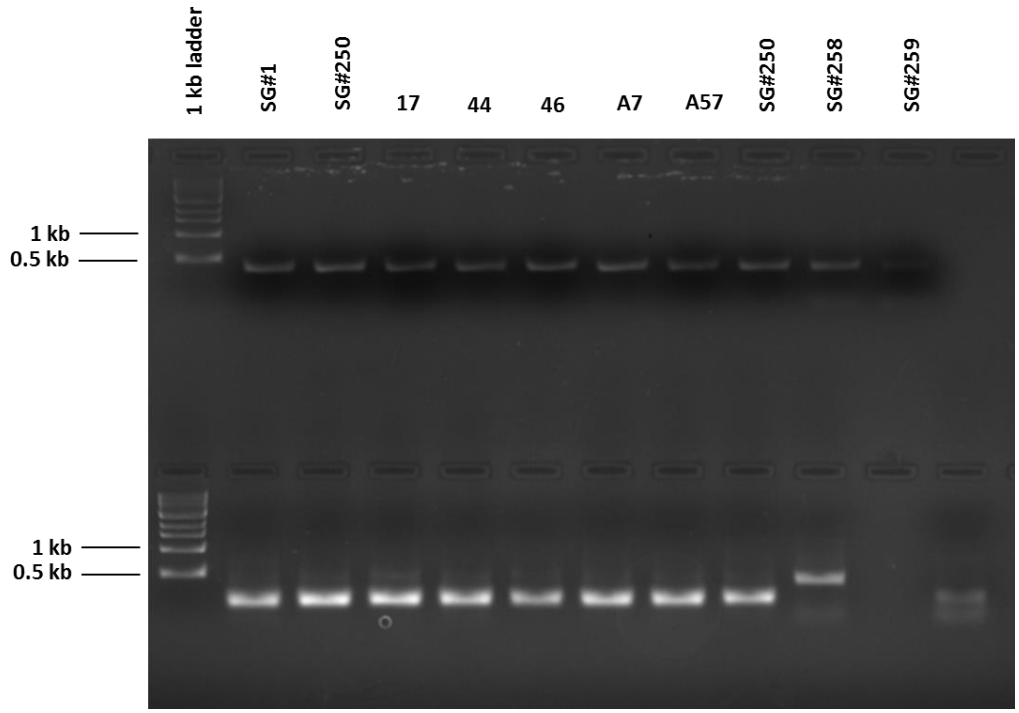


Fig 3.5. PCR with primers 1501999-F and 1501999-R for marker AluI(b). Top lanes are PCR product and bottom lanes are after digestion by AluI. SG#258 stays uncut.

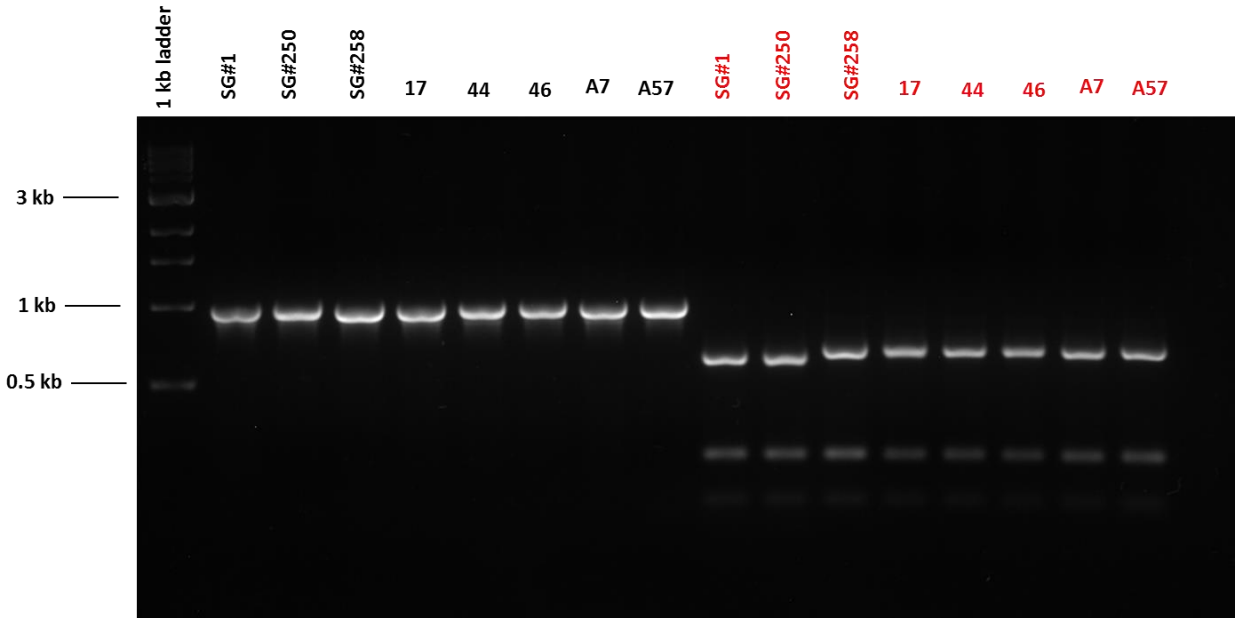


Fig 3.6. PCR with primers 545000-F and 546000-R for marker MspI. PCR products are shown in black and after digestion by MspI they are shown in red. SG#1 and SG#250 are cut.

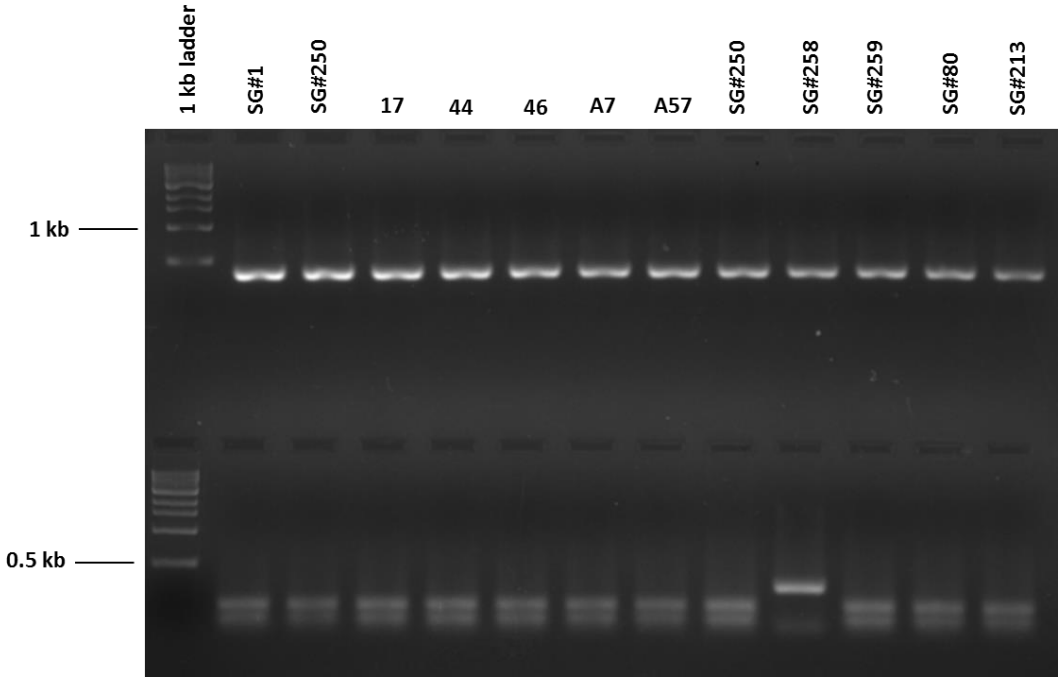


Fig 3.7. PCR with primers108348-R and 108348-F for marker HaeIII. Top lanes are PCR product and bottom lanes are after digestion by HaeIII. SG#258 stays uncut.

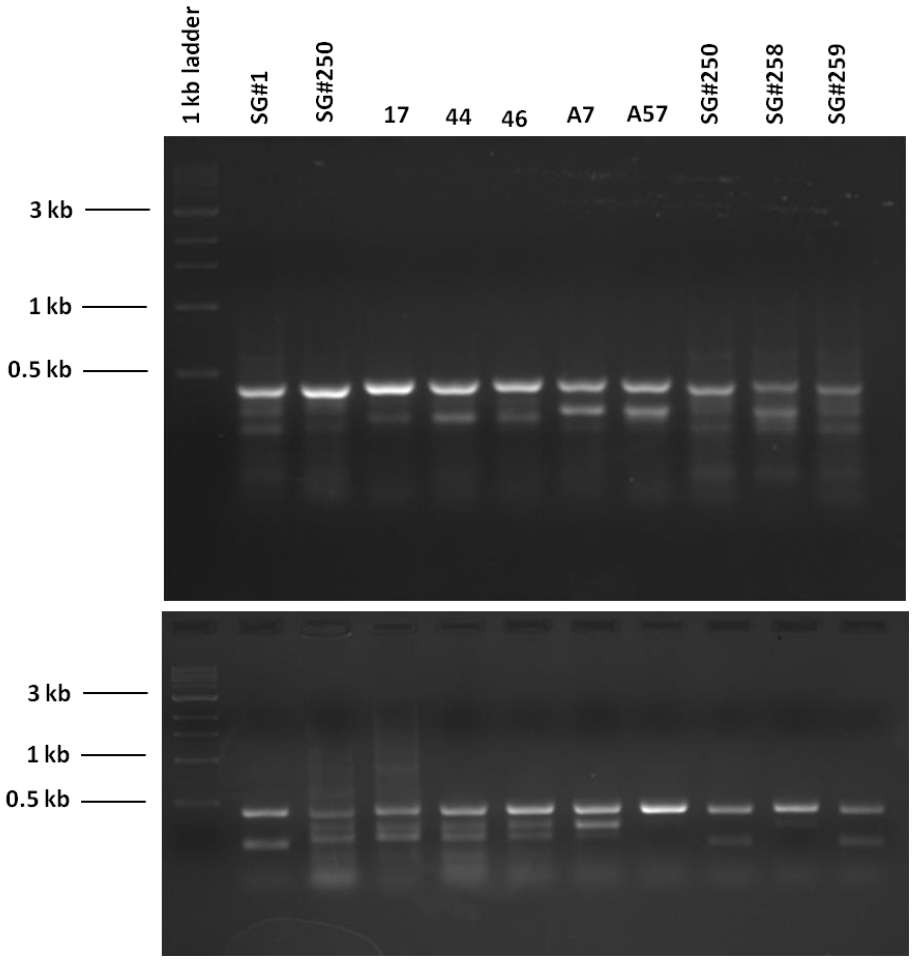


Fig 3.8. PCR with primers P21 and Q21 for marker TSP509I. Top lanes are PCR product and bottom lanes are after digestion by TSP509I. prd-1 progeny 17, 44, 46 are cut which shows inconsistent result. Non-specific bands can be seen in this gel picture.

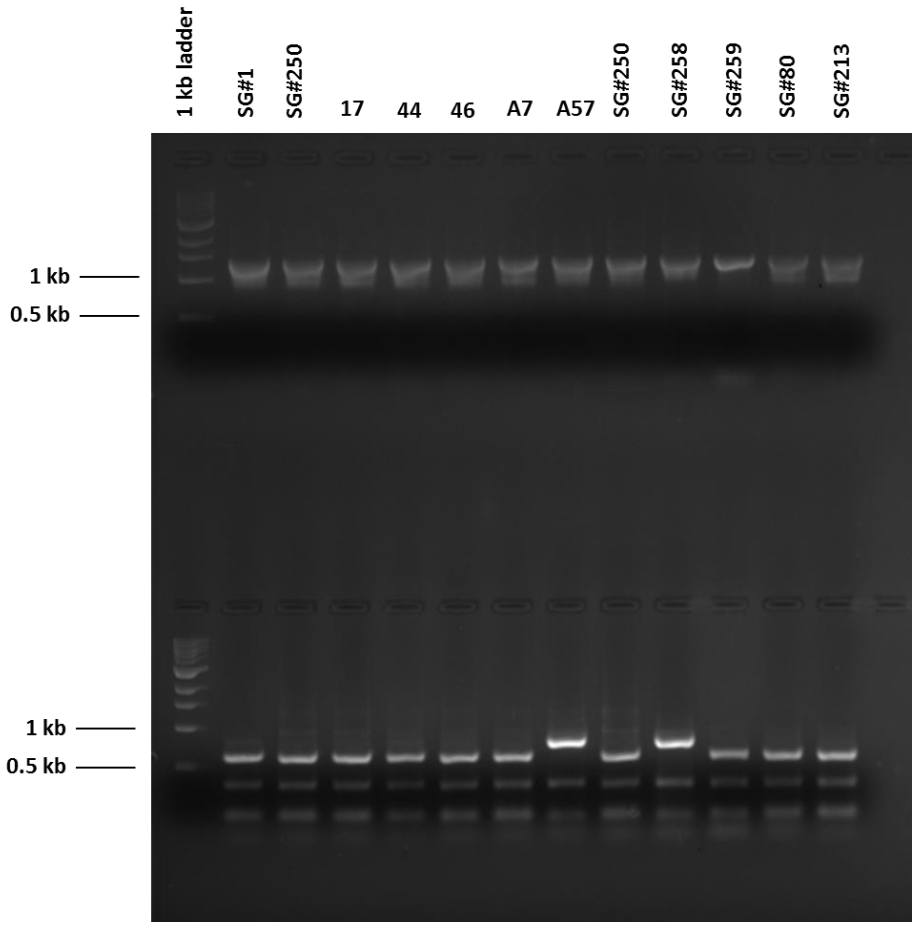


Fig 3.9. PCR with primers 597000F and 598400R for marker BstUI. Top lanes are PCR product and bottom lanes are after digestion by BstUI. A57 and SG#258 are uncut.

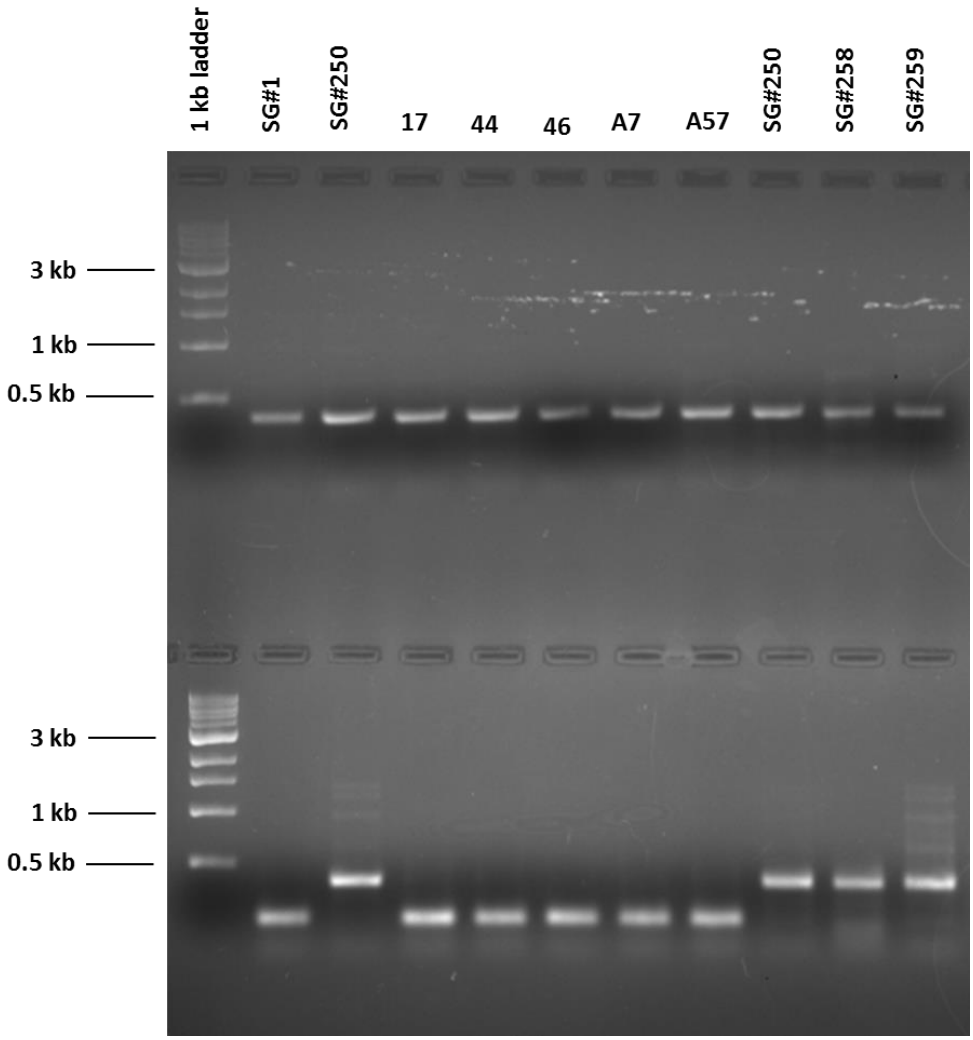


Fig. 3.10. PCR with primers 1211329F and 1211329R for marker TaqI(b). Top lanes are PCR product and bottom lanes are after digestion by TaqI. SG#250, SG#259 and SG#258 are uncut.

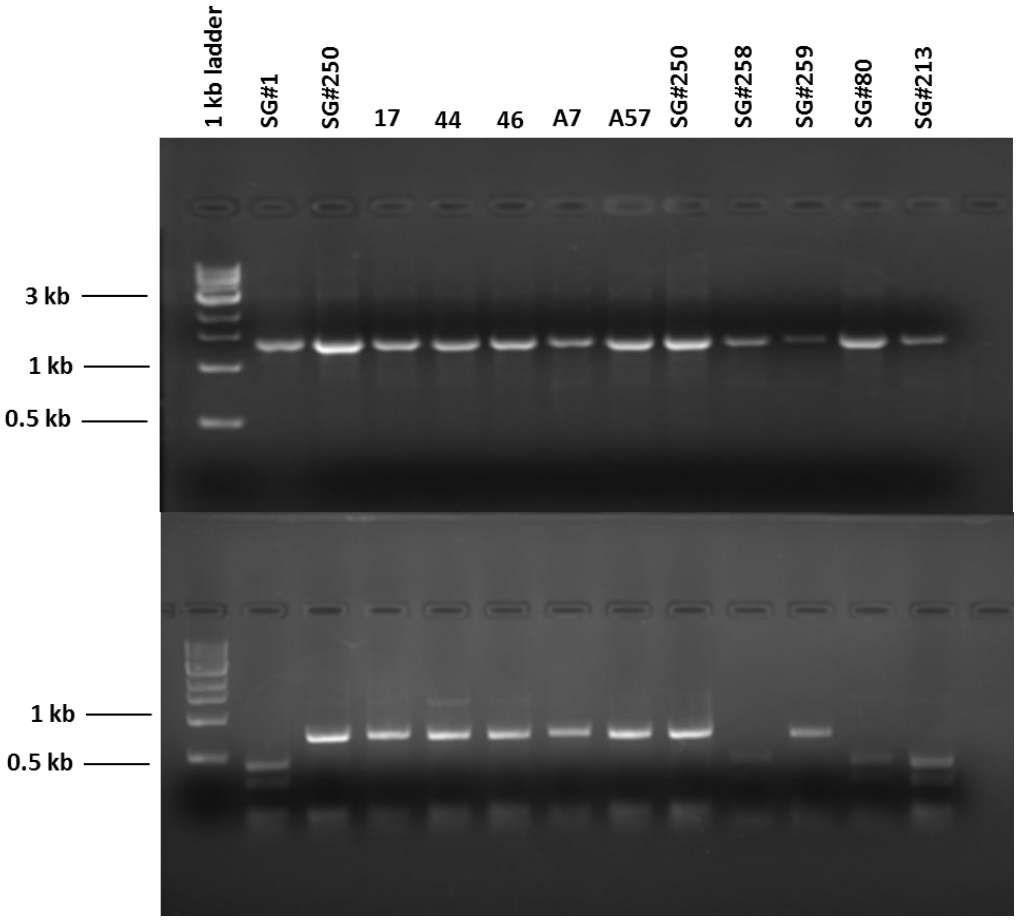


Fig. 3.11. PCR with primers 577600-F and 577600-R for TaqI(a). Top lanes are PCR product and bottom lanes are after digestion by TaqI. SG#1, SG#258, SG#80 and SG#213 are cut.

3.2. Screening knockouts and sequencing

After determining the approximate region of the prd-1 gene, screening knockout mutants in this interval was the next step toward identifying the gene. To make knockout genes, genes were replaced by the hph cassette which confers hygromycin resistance (Colot et al. 2006). Using the Broad Institute database website (www.broadinstitute.org) and genome annotations in the database, the genes in that specific interval (between HaeIII and MspI markers) were found. A number of knockouts in this region (left of centromere) from the Fungal Genetics Stock Center were already ordered by Keyur Adhvaryu and they were available in the lab and they were screened in race tubes. Period and growth rate for all the tubes were calculated using Mac Tau software.

Table 3.3 and 3.4 present lists of knockout genes screened in two experiments. Among 33 knockouts (Table 3.3 and 3.4) only one gene (NCU05703) had a slow growth rate (~ 1.6 mm/hr), which is a characteristic of the prd-1 mutation in addition to a long period. The knockouts are not carrying the *bd* mutation therefore they do not show any bands and period. The phenotype of NCU05703 is shown in figure 3.12 in comparison to controls.

On the right side of the centromere there are 32 genes which are listed in Table 3.5 (data collected from Broad Institute Database). From this list some of them were chosen based on their function and the knockout genes were ordered from FGSC and screened for growth rate on race tubes. Table 3.6 presents the phenotype of the knockout genes from FGSC. Among these knockouts FGSC#11688 (NCU07817) exhibits a slow growth rate and therefore it could be one of the candidates for the *prd-1* gene.

In order to sequence the candidate genes, primers were designed for PCR and sequencing (each interval covers 500bp-800bp). Around 200-400bp was included for upstream and downstream regions. PCR was carried out using genomic DNA from strain SG#250 (*bd csp prd-1*) to amplify the gene and the PCR product was electrophoresised through a gel and gel purified (Fig.3.13). The DNA samples and primers were sent to York Biology Molecular Core facility for sequencing. After alignment of the sequencing results, no mutations were found in these two genes, therefore they are not the *prd-1* gene.

Some of the genes from table 3.5 were chosen according to their function, for which knockouts were not available in FGSC and these genes were sequenced directly without screening. Primers were designed and PCR with SG#250 DNA was carried out and the samples were sent to York Biology Molecular Core facility for sequencing. Table 3.7 lists the sequenced genes with the number of upstream and downstream base pairs, location of the genes and forward and reverse primers. NCU07839 is the only gene that showed a mutation. To check whether this mutation exists in all *prd-1* strains, DNA was extracted from three *prd-1* progeny (17, 44, and 46) and the sequencing was repeated for them. DNA samples from SG#1 and SG#170 were also sequenced. In all *prd-1* strains this particular mutation could be seen, however in the wild type and *prd+* progeny this mutation was absent (table 3.8). Figure 3.14 shows the alignment of wild type NCU07839 sequence and mutant.

FGSC #	Mating Type	Genotype	Growth Rate	Period
15744	a	NCU05730.2	2.49	0
17654	A	NCU05699.2	2.72	0
17655	A	NCU05701.2	2.84	0
17657	a	NCU05702.2	2.66	0
17658	a	NCU05704.2	2.68	0
17659	A	NCU05711.2	2.78	0
17660	a	NCU05718.2	2.67	0
17717	A	NCU05698.2	2.70	0
18004	A	NCU05703.2	1.60	0
18005	A	NCU05707.2	2.71	0
18007	A	NCU05724.2	2.70	0
18008	A	NCU05727.2	2.73	0
18009	A	NCU05729.2	2.63	0
18075	A	NCU05705.2	2.82	0
18076	A	NCU05712.2	2.64	0
18077	a	NCU05713.2	2.67	0
18079	A	NCU05722.2	2.59	0
18080	A	NCU05728.2	2.49	0
18786	A	NCU05697.2	2.43	0
20883	A	NCU05723.2	2.24	0
21583	a	NCU10596.2	2.66	0
SG#259	a	O.R	2.46	0
SG#170	a	bd prd-1	0.95	29.475
SG#4	A	bd	1.12	26.71

Table 3.3. List of knockout genes screened in the first experiment, grown on race tubes. Growth rate and period were calculated by Mac Tau software. The bd mutation slows the growth rate and the knockouts grow fast because they are bd^{+}.

FGSC #	Mating Type	Genotype	Growth Rate	Period
12095	A	NCU05685.2	2.69	0
12463	a	NCU05710.2	2.89	0
13496	a	NCU05693.2	2.24	0
13679	A	NCU05716.2	2.74	0
13682	A	NCU05735.2	2.82	0
13808	a	NCU05715.2	2.65	0
13982	A	NCU05706.2	2.93	0
17714	a	NCU05691.2	2.9	0
18010	A	NCU05732.2	2.89	0
18875	A	NCU05731.2	2.89	0
18924	A	NCU05733.2	2.97	0
19933	a	NCU05694.2	2.88	0
19937	A	NCU05721.2	2.81	0
SG#4	A	bd	1.11	28.39
SG#170	a	bd prd-1	0.87	30.79
SG#259	a	WT	2.89	0
SG#1		bd csp-1	1.21	21.55
SG#250		bd csp-1 prd-1	0.74	23.68
17		bd csp-1 prd-1	0.95	23.43
41		bd csp-1 prd-1	0.54	24.87
A7		bd csp-1 prd-1	0.55	27.73
A57		bd csp-1	2.57	19.97

Table 3.4. List of knockout genes in the second experiment, grown on race tubes. Growth rate and period were calculated by Mac Tau software.

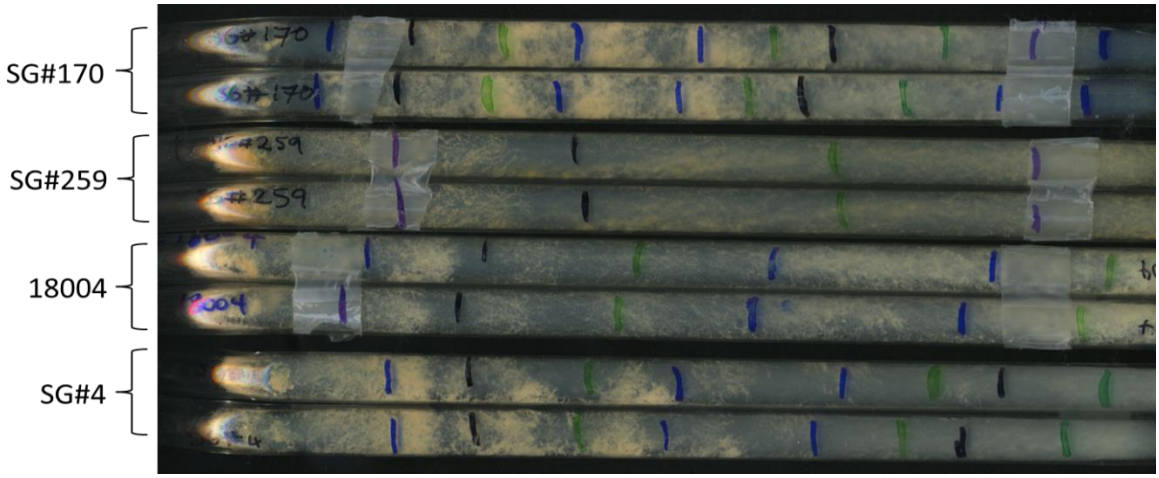


Fig 3.12. Screening of NCU05703 (FGSC#18004) on race tubes compared to controls. They were grown from left to right. Pen marks indicate 24 hour growth intervals.

Locus	Gene Name	Location	Length		Locus	Gene Name	Location	Length
NCU07836.7	pantotheic acid-3	N. crassa OR74A (NC12): Supercontig 3: 1029521-1031296 -	1776		NCU07834.7	MYB DNA-binding domain-containing protein	N. crassa OR74A (NC12): Supercontig 3: 1020153-1021420 +	1268
NCU07826.7	40S ribosomal protein S19	N. crassa OR74A (NC12): Supercontig 3: 991635-993299 -	1665		NCU07837.7	Cockayne syndrome B ortholog	N. crassa OR74A (NC12): Supercontig 3: 1031459-1035777 +	4319
NCU07846.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1077781-1080168 -	2388		NCU07845.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1073692-1074469 +	778
NCU07844.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1071357-1072873 -	1517		NCU07820.7	pantothenate transporter	N. crassa OR74A (NC12): Supercontig 3: 979196-980891 +	1696
NCU07818.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 975361-976669 +	1309		NCU07831.7	eukaryotic translation initiation factor 3	N. crassa OR74A (NC12): Supercontig 3: 1003025-1006515 -	3491
NCU07833.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1016388-1017001 -	614		NCU07828.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 997573-999548 +	1976
NCU07842.7	Woronin sorting complex	N. crassa OR74A (NC12): Supercontig 3: 1066817-1069356 +	2540		NCU07827.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 993916-997075 +	3160
NCU07815.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 965451-967346 +	1896		NCU07822.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 983811-985578 +	1768

Locus	Gene Name	Location	Length		Locus	Gene Name	Location	Length
NCU05742.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 813119-813977 -	859		**NCU07839.7**	ATP-dependent RNA helicase dbp-2	N. crassa OR74A (NC12): Supercontig 3: 1037469-1041521 +	4053
NCU07843.7	ELL complex subunit Eap30	N. crassa OR74A (NC12): Supercontig 3: 1069160-1070961 -	1802		**NCU07823.7**	scytalone dehydratase	N. crassa OR74A (NC12): Supercontig 3: 985729-986672 +	944
NCU07830.7	cytoplasmic ribosomal protein-2	N. crassa OR74A (NC12): Supercontig 3: 1000681-1002117 +	1437		**NCU07841.7**	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1063208-1065157 -	1950
NCU07824.7	MDM complex subunit 10	N. crassa OR74A (NC12): Supercontig 3: 986903-989037 +	2135		**NCU07829.7**	60S ribosomal protein L7	N. crassa OR74A (NC12): Supercontig 3: 998601-1000625 -	2025
NCU07825.7	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 988745-991445 -	2701		**NCU07821.7**	dimethylaniline monooxygenase	N. crassa OR74A (NC12): Supercontig 3: 981619-983387 +	1769
NCU07832.7	pre-mRNA processing splicing factor 8	N. crassa OR74A (NC12): Supercontig 3: 1007087-1015086 +	8000		**NCU07816.7**	magnesium and cobalt transporter CorA	N. crassa OR74A (NC12): Supercontig 3: 967898-970816 +	2919
NCU07819.7	alpha-ketoglutarate-dependent taurine dioxygenase	N. crassa OR74A (NC12): Supercontig 3: 976728-978170 -	1443		**NCU07838.7**	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1036269-1037139 +	871
NCU07817.7	non-anchored cell wall protein-3	N. crassa OR74A (NC12): Supercontig 3: 972480-974377 +	1898		**NCU07840.7**	hypothetical protein	N. crassa OR74A (NC12): Supercontig 3: 1059964-1062628 -	2665

Table 3.5. List of the genes on the right side of centromere between centromere and HaeIII marker collected from Broad Institute database.

FGSC#	Mating Type	Genotype	Growth Rate	Period
18136	A	NCU07844.2	1.71	0
19588	A	NCU07819.2	1.66	0
19591	A	NCU07820.2	1.62	0
19593	A	NCU07822.2	1.45	0
20493	A	NCU07833.2	1.55	0
13786	a	NCU07836.2	1.45	0
14353	a	NCU07823.2	1.55	0
14940	A	NCU07818.2	1.64	0
18132	a	NCU07827.2	1.76	0
18133	A	NCU07841.2	1.61	0
11688	A	NCU07817.2	0.89	0
11969	A	NCU07816.2	1.64	0
12027	a	NCU07840.2	1.59	0
12075	a	NCU07834.2	1.47	0
13121	a	NCU07837.2	1.72	0
21085	a	NCU07821.2	1.53	0
SG#170	a	bd prd-1	0.83	26.5
SG#259	a	WT	1.56	0
SG#4	A	bd	1.05	26.062

Table 3.6. List of knockout genes ordered from Fungal Genetic Stock Centre (FGSC). Growth rate and period were calculated by McTau software.

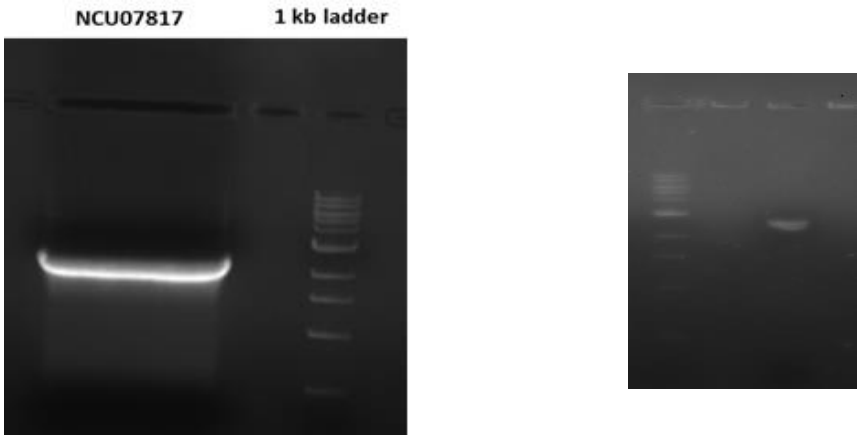


Fig 3.13. PCR of NCU07817 with SG#250 DNA as template (left). Second gel (right) is after gel purification with QIAGEN kit to show the concentration for sequencing. Similar gels were run for other sequenced genes.

Locus	Gene Location	PCR	Forward and reverse primers	Mutation
NCU05703	569539-570430	568977-570929	F-1 5' GGTCTCCCCCTGGTCCCG 3'	NO
		562nt up, 499nt down	R-1 5' TCTCAGAAACATTGCTTA 3'	
NCU07817	973184-974249	973002-974293	F-1 5' ACGCAACTCAACCACCATTCC 3'	NO
		182nt up, 44nt down	R-1 5' AAGGCGCATTGCTTTACACG 3'	
NCU07824	986903-989037	986420-989493	F-1 5' CGAAGTCAACTGAAGCGATAGC 3'	NO
		483nt up, 456nt down	R-1 5' TACGGCCCAGTAAGATTAGACG 3'	
NCU07845	1073692-1074469	1073379-1074790	F-1 5' TGTCTCCATTACTCCCTTGAAC 3'	NO
		313nt up, 321nt down	R-1 5' GGGAAGCAGATGGACTATGTATG 3'	
NCU07839	1037469-1041521	1037113-1041895	F-1 5' CTCAAATCCGCCTTCGCTGTAAC 3'	YES
		356nt up, 374nt down	R-1 5' ACCTTGGAACGCAGCACAAATC 3'	

Table 3.7. List of sequenced genes, their location and forward, reverse primers.

Strain	Genotype	Observed mutation
SG#1	w.t	NO
SG#170	bd prd-1	YES
SG#250	bd csp-1 prd-1	YES
17	prd-1	YES
44		
46		
A57	prd+	NO
56		
62		

Table 3.8. List of sequenced prd-1 and prd+ progeny to confirm the mutation can be seen in all prd-1 progeny

Fig 3.14. Alignment of w.t NCU07839 sequence (top) and mutant DNA sequence (bottom) (mutation of G → A is show by #). Only alignment with primer F-5 is shown. Every 10 nucleotides are scaled by *.

3.3. Characterizing the mutation

Annotation of the gene using the Broad Institute Database shows the gene has four introns that are 138 bp, 520 bp, 63 bp and 57 bp respectively that should be removed by RNA splicing to generate the final mature RNA product of the gene. After aligning the sequences, it was found that the mutation is located on the first base of the third intron (G→A) and may prevent the third intron from being spliced (Fig.3.15)

In order to check the effect of this mutation on splicing, RNA was extracted and using an RNA to cDNA kit, cDNA of wild type and mutant was made. As shown in figure 3.15 primers were designed to cover the specific intron. After PCR and electrophoresing the product on a gel it can be seen that the size of the band in WT and mutant does not show a difference in the size of the intron (which is 63 bp), therefore the mutation does not disrupt splicing completely (Fig.3.16). Since there is a difference in brightness of bands between WT and mutant cDNA, the bands were cut and sent for sequencing. Alignment of WT and mutant cDNA showed that this mutation has an effect on splicing and splicing does not occur properly: there is an extra G and A in the mutant cDNA (Fig.3.17). This will change the reading frame and influence the amino acid sequence. This gene is an RNA helicase and the amino acid sequence of wild type shows that it contains a DEXDc domain and a HELCc domain. Based on cDNA alignment it was found that this mutation affects the amino acid sequence as shown in figure 3.18 and 3.19.

```
CTTCCTTCCTTCCACCCACCCTCCATCTCCAATATCCCTTCCCGCCTACTTCTTGCGATTCTGCAAGATT  < 70
         10        20        30        40        50        60

CTTCATCCGCGATCCTTCTTTAACCAGTCGGACTGCTTACTTGTAACCACTTCACGCGTGTCGTTTGGTT  < 140
         80        90       100       110       120       130

CTTGTCTCTAGACGTCTGTCTTTTCCGACGTCTTCTTTTCCCGACAATCGATACCCTTGCGCCTCTAGGC  < 210
        150       160       170       180       190       200

CTTGTTTGCTCGTCGTTTCGGTCTACCCTTCAGCAACCCCAATAACAACCGATCTTCAATTTTCAACCTT  < 280
        220       230       240       250       260       270

TTCTTAAACAACCAGCAAACATGTCCGGATCTTACGGCGGCGGCGGCTACGGTGGCCGCGGTGGCGGTGG  < 350
        290       300       310       320       330       340

TGGTGGATATTCTAACGGGTACGATCGCAATGGCGGCGGCTACTCTAACAACTACTCCTCACACGGGTAT  < 420
        360       370       380       390       400       410

GATGTTTTAGCGCGAACCTGCTCATTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT  < 490
        430       440       450       460       470       480

CTCTCTCTCTCTCTCTCTCTCTCTTTGCATTTCAAGATCAAATTGCTAACATCTGGTTCAACAGTGGCTC  < 560
        500       510       520       530       540       550

GAATGGCTACGGAGGCGGCGGCGGAGGATATGGCGGCGGTGGTGGTGGCTACGGAGGCGGCGGCTACGGT  < 630
        570       580       590       600       610       620

GGTGGTGGCGGCGGTGATAGGATGTCTGCTCTCGGTGCTGGTCTGCAGAAGCAGAACTGGGACATGAGCG  < 700
        640       650       660       670       680       690

CTCTTCCCAAGTTCGAGAAGTCCTTCTACCAAGAGCATCCTAGCGTCGCCAACCGATCTCCTGCTGAGGT  < 770
        710       720       730       740       750       760

CGACAAGTTCCGCGCAGACCACTCCATTGCCGTCTTCGGTAACAACGTTCCCAAGCCTGTCGAGACCTTC  < 840
        780       790       800       810       820       830

GACGAGGCTGGTTTCCCTCGCTACGTTATGGACGAGGTCAAGGCTCAGGGTTTCCCTGCTCCTACTGCCA  < 910
        850       860       870       880       890       900
```

```
TTCAGTCGCAGGGCTGGCCCATGGCCCTTTCTGGTCGCGATGTCGTCGGTATTGCCGAGACCGGTTCCGG  < 980
         920       930       940       950       960       970

AAAGACGCTCACCTACTGCCTTCCCGCCATCGTTCACATCAACGCTCAGCCTCTCCTCGCTCCCGGCGAT  < 1050
         990       1000      1010      1020      1030      1040

GGCCCTATCGTCCTCATCCTCGCCCCTACCCGTGAGTTGGCTGTCCAGATTCAGCAAGAAATCTCCAAGT  < 1120
         1060      1070      1080      1090      1100      1110

TCGGCAAGTCGTCCCGCATTCGCAACACCTGCGTCTACGGTGGTGTCCCCAAGGGTCCCCAGATTCGCGA  < 1190
         1130      1140      1150      1160      1170      1180

TCTTTCCAGGGGAGTCGAGGTCTGTATCGCCACCCCCGGCCGTTTGATCGATATGCTCGAGTCTGGCAAG  < 1260
         1200      1210      1220      1230      1240      1250

ACCAACCTCCGTCGTGTCACCTACCTTGTCCTCGATGAGGCTGATCGCATGTTGGACATGGGTTTCGAGC  < 1330
         1270      1280      1290      1300      1310      1320

CCCAAATTCGCAAGATCATCGGCCAGATTCGCCCTGATCGTCAGACTCTCATGTGGTCTGCTACCTGGCC  < 1400
         1340      1350      1360      1370      1380      1390

CAAGGAGGTCCGCAACTTGGCCGCCGACTTCTTGACCGACTTCATCCAGGTCAACATCGGTTCCATGGAT  < 1470
         1410      1420      1430      1440      1450      1460

TTGGCTGCCAACCACCGTATCACCCAGATCGTCGAGGTTGTTTCCGAGTCCGAGAAGCGTGATCGCATGA  < 1540
         1480      1490      1500      1510      1520      1530

TCAAGCATCTTGAGAAGATTATGGAGGGCCGCGAGAACCAGAACAAGATCCTCATCTTCACCGGCACCAA  < 1610
         1550      1560      1570      1580      1590      1600

GCGTGTCGCCGACGACATCACCCGCTTCCTCCGCCAGGACGGCTGGCCCGCTCTTTCCATCCACGGCGAC  < 1680
         1620      1630      1640      1650      1660      1670

AAGCAACAGAACGAGCGTGACTGGGTTTTGGACCAGTTCAAGACTGGCAAGAGCCCTATCATGGTTGCCA  < 1750
         1690      1700      1710      1720      1730      1740

CCGACGTGGCTTCTCGTGGTATCGGTATGTATCCATCTTACCCCTCCGCCCTGCCCAAACCCCTCCCCCC  < 1820
```

```
         1760      1770      1780      1790      1800      1810

TTCTTCTCTCCCCCATCTCAGGCGACTTCCAGCATACGTACGTGCTTGTGCTTTACTCTTATGTCTAGGG  < 1890
         1830      1840      1850      1860      1870      1880

GCTCTCTTGATTGCGCGAACTTGCCTGACCTCGTGTCGACTTGGTCTATCTGAGTGGAGTCTTGGGTTCG  < 1960
         1900      1910      1920      1930      1940      1950

TGACGTAGCAGAAACAGTGGACTATGCTACGCCCGCGTGATTCTTCAATTGTACCTTTTTATCCCAGTTG  < 2030
         1970      1980      1990      2000      2010      2020

TGGTACAAGCGAGGAGTAGGCATTTCTTGATTCCCGTCAAGCTTGTCTCCAAGATTCCCAAGGGGATTTG  < 2100
         2040      2050      2060      2070      2080      2090

CATGCATGGCCCAAGGATCTCTGGAACCTTGGATCCCAACCCGGCATATTTTTCGCAATGGACAGCCCAC  < 2170
         2110      2120      2130      2140      2150      2160

CTGCCAAGGAGAGAAGTACTACCAGAGGATCACAGAACATACCTAGATATTCGTGTTCTCCGTCCGTTCT  < 2240
         2180      2190      2200      2210      2220      2230

GGAATGTCGTTATTCCAAACCCTGCTATTGCTAACCACGGTCCTTTAAAAAAAGATGTGCGCAACATTAC  < 2310
         2250      2260      2270      2280      2290      2300

TCACGTGCTCAACTATGACTACCCCAACAACTCCGAGGATTACATCCATCGTATCGGCCGAACTGGTCGT  < 2380
         2320      2330      2340      2350      2360      2370

GCCGGTGCGAAGGGTACTGCCATTACTTTCTTCACCACTGACAGTAAGTGTCACCTGAATCTGCCCATCC  < 2450
         2390      2400      2410      2420      2430      2440

TTAGCCGAACGAAATTCTAATGTTATCTCTCCACAGACTCCAAGCAGGCTCGTGAGCTCGTCGGTGTGCT  < 2520
         2460      2470      2480      2490      2500      2510

TCAAGAAGCTAAGCAGCAGATTGACCCTAGACTTGCCGAGATGGCTCGCTACAGTGGTGGTGGAGGTGGC  < 2590
         2530      2540      2550      2560      2570      2580

CGCTTCGGTGGCTATCGCGGCCGTGGCGGCGGTGGATGGCGCGGTGGTAAGTCATTTTCACGCAGGTGAA  < 2660
         2600      2610      2620      2630      2640      2650
```

```
TTTGGCAGTTCCAATCCTAATAACCCTCATTAGGCCGTGGCGGCGGTGGTGGCGGCGGCTCCGTCGGTGG  < 2730
         2670      2680      2690      2700      2710      2720

TGCTAATGCTCTTCCGCTTAACAACCGCCGGTGGTAAAGCACCACTTGAAGCGAATCAGCACATATCAAT  < 2800
         2740      2750      2760      2770      2780      2790

TGACCTTCCTGGTCCAATCCATTGTCGCCATGGTCATTCCACACGACGAATGGCCCATTACAGCGACAAT  < 2870
         2810      2820      2830      2840      2850      2860

TTCGACCTCGAACATGATCACGGCGTTTTAGTAGAGGATACGGTCTAGAGGTTGGTTTATTGCTTTCAGA  < 2940
         2880      2890      2900      2910      2920      2930

CGCTCACTGGTGTCTGTTTTCACCTTGACGTCTAGGTGGTATGGGTGTGGCTTTGTTCTTTTTTTCTCGT  < 3010
         2950      2960      2970      2980      2990      3000

ATTCCCTTATTGTCATGCTTGCGGGTTACTATTTTCACTGAGGTCAGCCGGGGTTCCAAAATTTTCCTTC  < 3080
         3020      3030      3040      3050      3060      3070

ATTTTTTCGGAAAGTCTCATCTACAGGTCATGTTGCACCACAACTGCACTTGCAGATCATCAAACTGTCC  < 3150
         3090      3100      3110      3120      3130      3140

AGGTCAAGATGGCCGGGTTTTATACCCTTCACCATATCTCGTCATGCCATTGCTCCGTCAAGGCATTATA  < 3220
         3160      3170      3180      3190      3200      3210

TGACTTATGACGACGGCTTATCAGTCGAGGTGTGAACAAGAGATTTCAACTCTCTTGCTTCCCTAACAGG  < 3290
         3230      3240      3250      3260      3270      3280

ACGGGCGCAGGCTTCTCGCCTCTCATCCTCGTTGATTAGTGATGATCTTACCGCAGACTTGTGGTAGCAG  < 3360
         3300      3310      3320      3330      3340      3350

TCGATAGGTACTATCCTCTGCTTTTTGTGTTGTACAATAGGAAGAGGATCCTTTTCAAAGCATCACGGCT  < 3430
         3370      3380      3390      3400      3410      3420

TGAAAGATTTTCGAAGGATCTGTCTAGTCACCAAGAGAGTCAGAGATAAACACGCGCTTTGTTTGCGCCC  < 3500
         3440      3450      3460      3470      3480      3490

TTGGGCGCTGCAAGACTACTGCTCACCAGCTGGGTGGGATGTTTACGGTAATTGTCACCGAGAGAGGGTA  < 3570
         3510      3520      3530      3540      3550      3560
```

```
GTAGAAATGAACTGGGTAGGTCTCCAAAGAAAGATACAGATACCCTTTATGCTTAGATGGAAAAGAACAT  < 3640
         3580       3590       3600       3610       3620       3630

TTACGCTAG  < 3649
```

Fig 3.15. NCU07839 gene. Start and stop codons are shown in turquoise color. Introns are shown in red, the mutation at the beginning of the intron shown in green. Primer F-10 and R-4 (shown in purple) were used for amplifying the region of DNA which contains this intron.

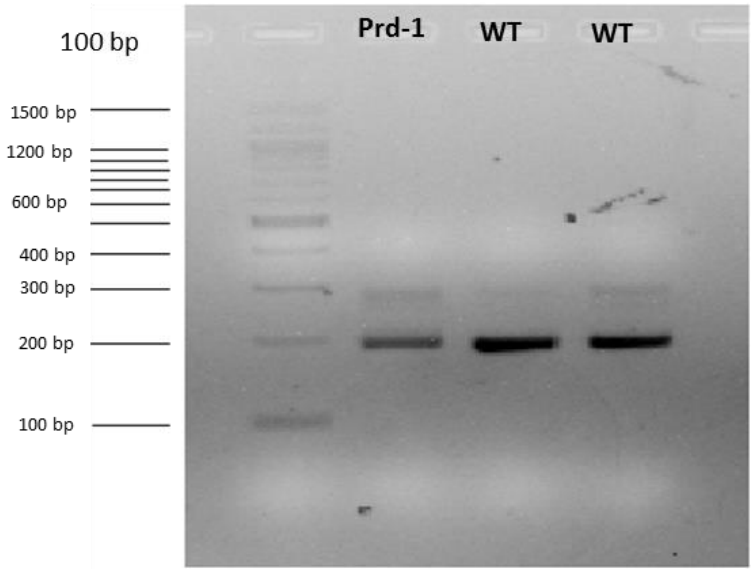


Fig 3.16. PCR of cDNA from wild type and prd-1 on 3% agarose gel

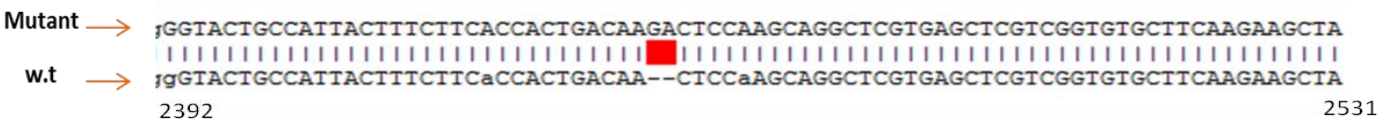


Fig 3.17. Alignment of part of wild type and prd-1 cDNA after sequencing.

Fig 3.18. Wild type amino acid sequence and protein domains. Purple and green are low complexity regions, red is DEXDc domain and blue is HELICc domain.

Fig 3.19. *prd-1* amino acid sequence and protein domains. Purple and green are low complexity regions; red is DEXDc domain and blue is HELICc domain. From amino acid 489 the sequence is different in wild type and mutant.

3.4. Cloning the candidate gene

To confirm the identity of the candidate gene it is necessary to do a complementation test. The first step for complementation is to clone the gene into the *his-3* targeting plasmid (pBM61) (Fig.3.20). The next step is to transform the plasmid into a *prd-1*, *his-3* strain and assay the rescued phenotype. This integration into the chromosome of pBM61 plasmid with its insert occurs by double crossover (Fig.3.21). Integration of an insert at the *his-3* locus happens without duplication of the *his-3* sequence and reconstitutes the *his-3*$^{+}$ phenotype. Selection is applied for His^{+} colonies after transformation.

Cloning was done first into the pSTblue-1 plasmid (Fig.3.26) and then after cloning to that plasmid the insert was subcloned into pBM61. Cloning into the pSTblue-1 plasmid was done using the EcoRI construct. This cloning was tried twice. In the first trial colonies were found on the plates and they were picked and tested using HindIII (2 bands of 6 kb and 2.5 kb were expected) and BamHI (4 bands were expected), but the DNA did not show the expected size after digestions (Fig. 3.27 and 3.28). In the second trial with pSTblue-1 plasmid the DNA was tested with BamHI and colonies 10 and 16 showed the expected size after digestion (Fig.3.29). These two were tested with HindIII and DNA from colony 10 showed the expected size and therefore the insert seemed to be in the right place (Fig.3.30).

This cloned plasmid was used for subcloning into pBM61. The cloned plasmid was digested with EcoRI to get the insert, it was ligated with pBM61 plasmid digested with EcoRI, and transformed into DH5 alpha chemical competent cells. Many colonies were seen in experimental plates and they were picked and DNA tested using BamHI, HindIII

and EcoRI (Fig.3.31, 3.32, 3.33 respectively). As can be seen there were some colonies (number 1-5) with the expected size. Correct cloning was confirmed by sequencing the plasmid and colony 3 was chosen for complementation.

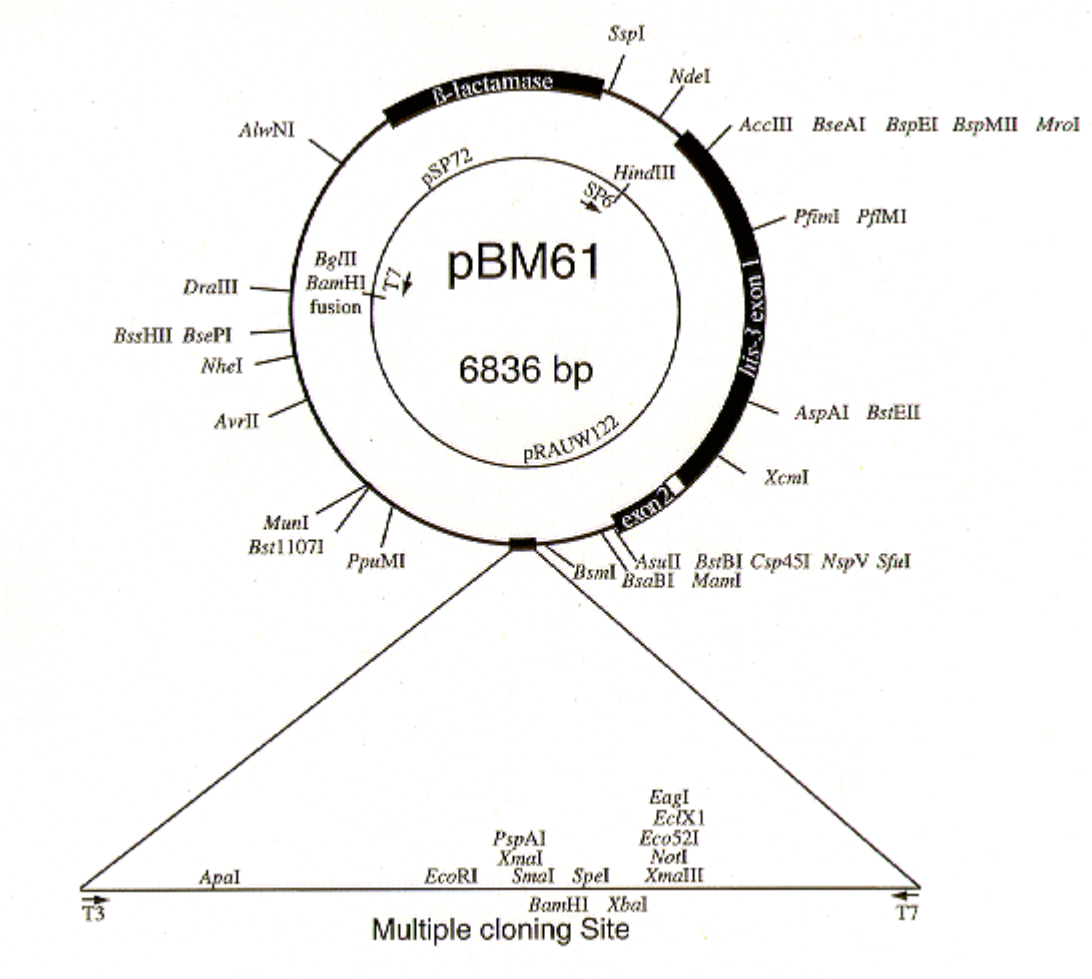


Fig 3.20. pBM61 *his-3* targeting plasmid. Adapted from www.fgsc.net.

Up and downstream of gene	Restriction Enzyme	Forward and reverse primers	Size of plasmid and insert	Insert Size
166 bp upstream	EcoRI	F-1 5' AATT GAATTC ATCCGCATCTCACATCCACTAGC 3'	10971 bp	3920
104 bp downstream		R-2 5' CTCATCGACGTGAAACCTACAGC 3'		
1 kb upstream	NotI	F-1 5' ATAAGAAT GCGGCCGC ACATGGAAGTGGTGGTTTGT 3'	12328 bp	5605
1 kb downstream	Apa I	R-1 5' TCC GGGCCC AACAAATATTAACGAGAATCTTCA 3'		
202 bp upstream	BsaBI	F-1 5' TCC GATGTATATC GGTCTGGCACAGGATCCGCATCTC 3'	10601 bp	4351
500 bp downstream	ApaI	R-1 5' TCC GGGCCC AAATCAACACCGGAATCATAAGAC 3'		
1 kb upstream	BsaBI	F-1 5' TCC GATGTATATC ACATGGAAGTGGTGGTTTGTGTTG 3'	11899 bp	5649
976 bp downstream	ApaI	R- 1 5' TCC GGGCCC AACAAATATTAACGAGAATCTTCA 3'		

Table 3.9. Cloning trials with different enzymes.

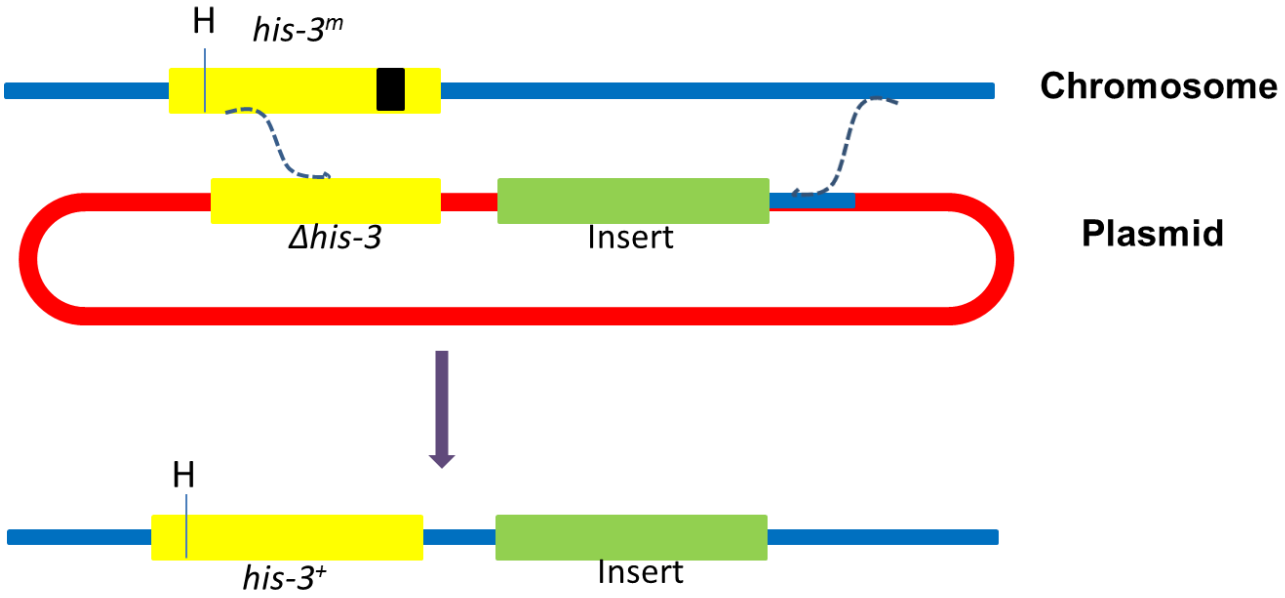


Fig 3.21. Integration of pBM61 plasmid and insert and reconstitution of *his-3*$^{+}$ genotype.

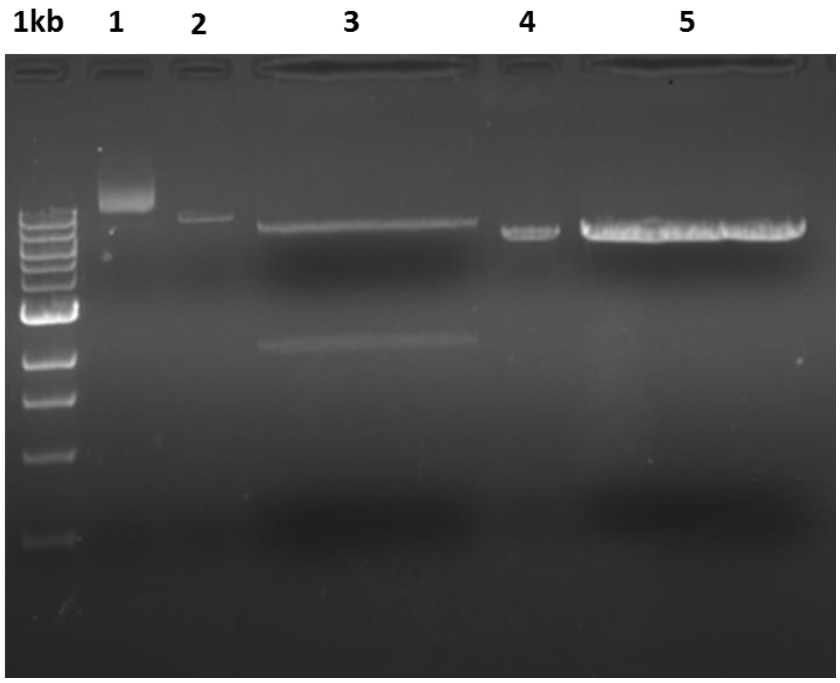


Fig 3.22. Digestions of plasmid and insert with NotI and ApaI restriction enzymes shown with 1kb ladder.

1. Uncut plasmid
2. Plasmid digested by NotI
3. Plasmid double digested by NotI and ApaI
4. Insert digested by NotI
5. Insert double digested by NotI and ApaI

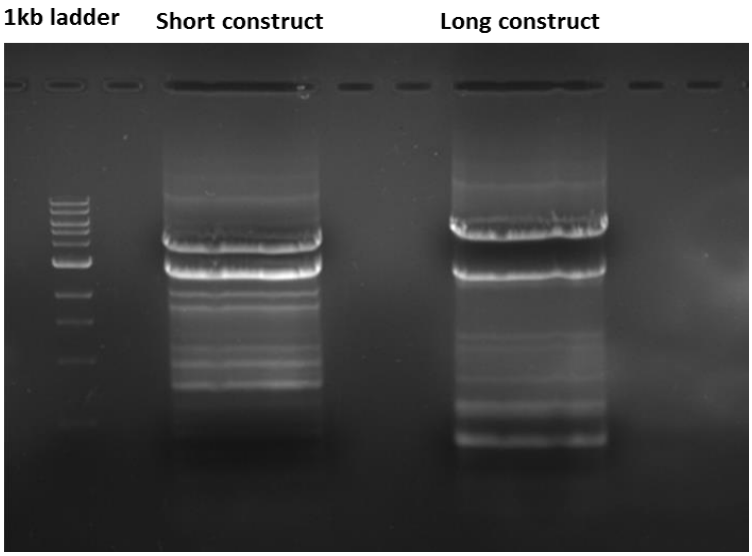


Fig 3.23. Short and long construct PCR products for cloning using BsaBI and ApaI. The first bands (short construct: ~ 4 kb and long construct ~ 5.5 kb) are inserts. The other bands are non-specific.

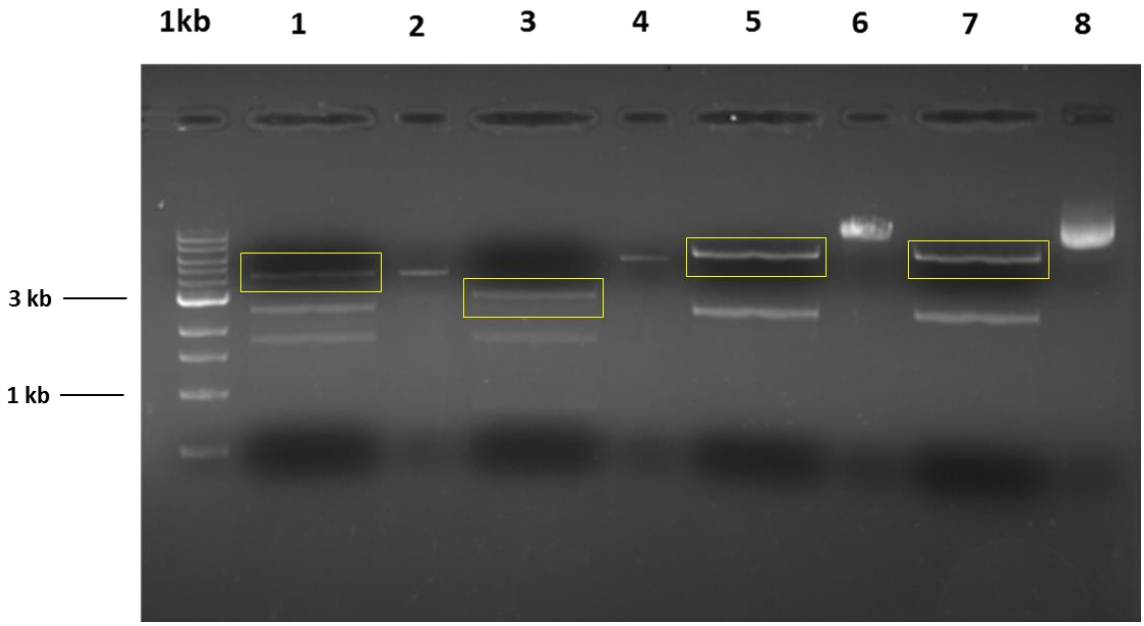


Fig 3.24. Digestion of insert and plasmid by BsaBI and ApaI. Marked bands with yellow rectangle are the correct bands.

1. Large PCR double digest (first band is the correct band)
2. Small PCR first digest
3. Small PCR double digest (first band is the correct band)
4. Large PCR first digest
5. Plasmid double digest (first band is the correct band)
6. Plasmid first digest
7. Plasmid double digest (first band is the correct band)
8. Uncut plasmid

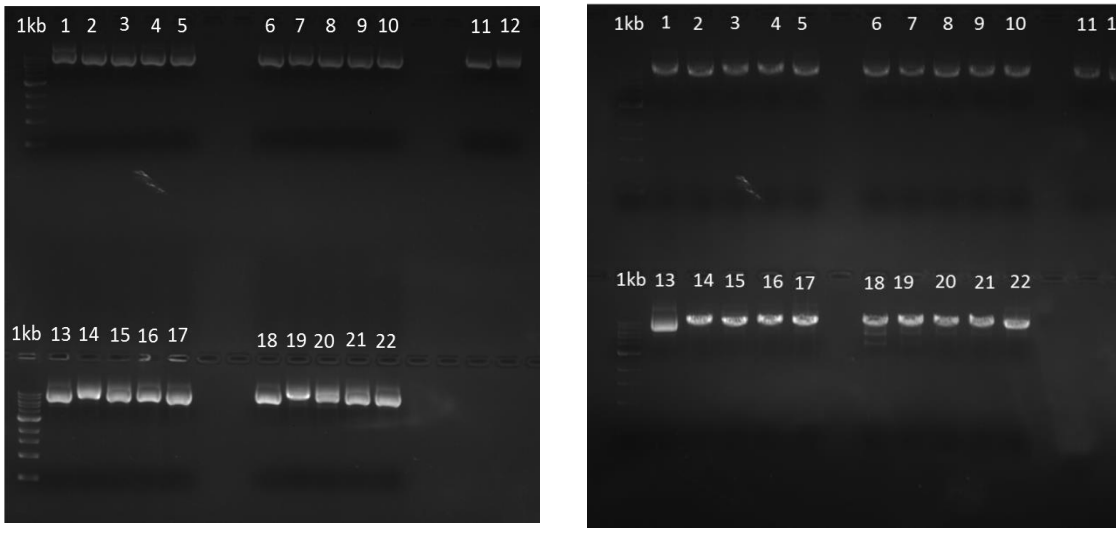


Fig.3.25. Cloning with EcoRI. Colonies were picked and plasmids were digested by EcoRI on the left picture (2 bands expected) and BamHI on the right picture (3 bands expected).

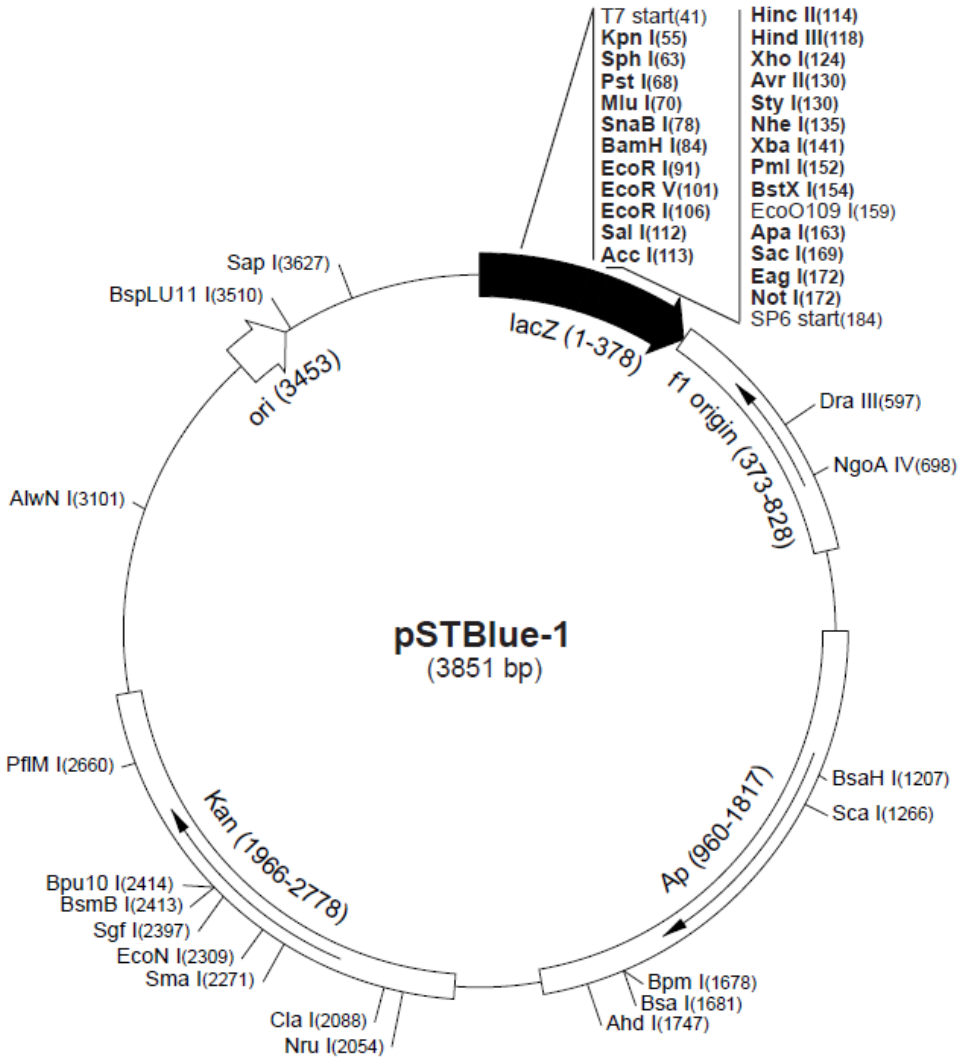


Fig 3.26. pSTblue-1 plasmid. Adopted from Novagen website (http://www.emdmillipore.ca).

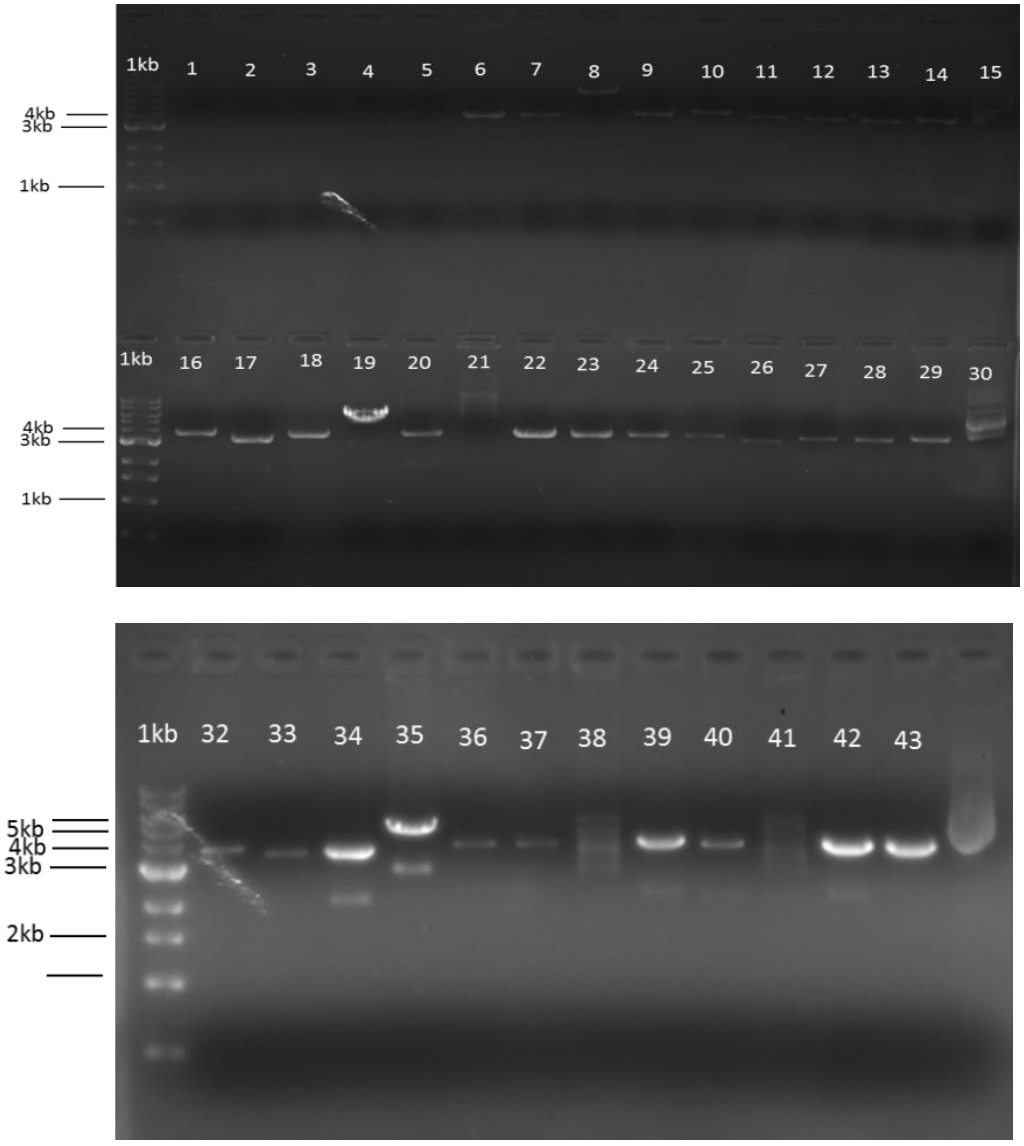


Fig 3.27. Digestion of colonies picked from the first trial of PSTblue cloning by *HindIII* (two bands of ~6kb and ~2.5 kb were expected).

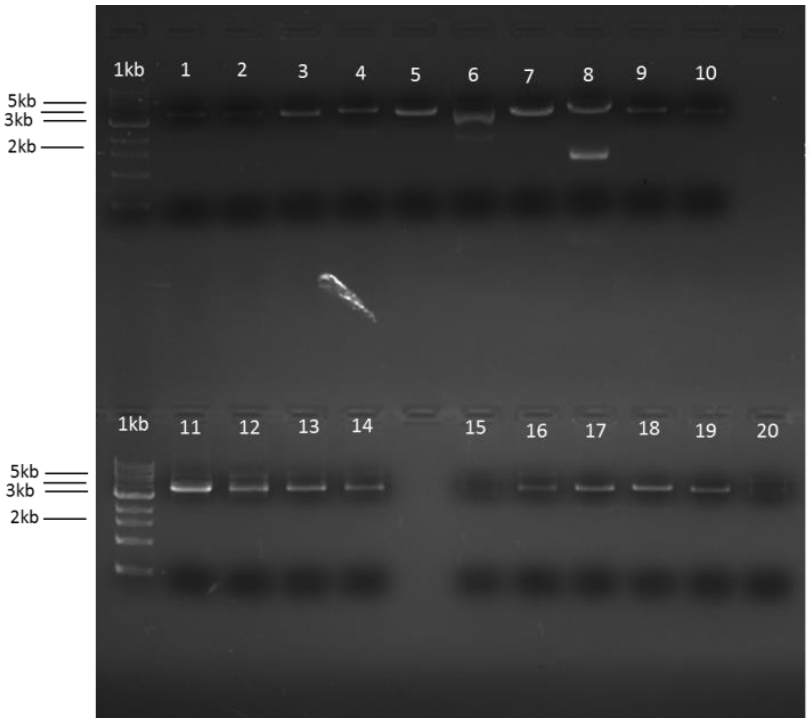


Fig 3.28. Digestion of colonies picked form the first trial of PSTblue cloning by *BamHI* (four bands were expected).

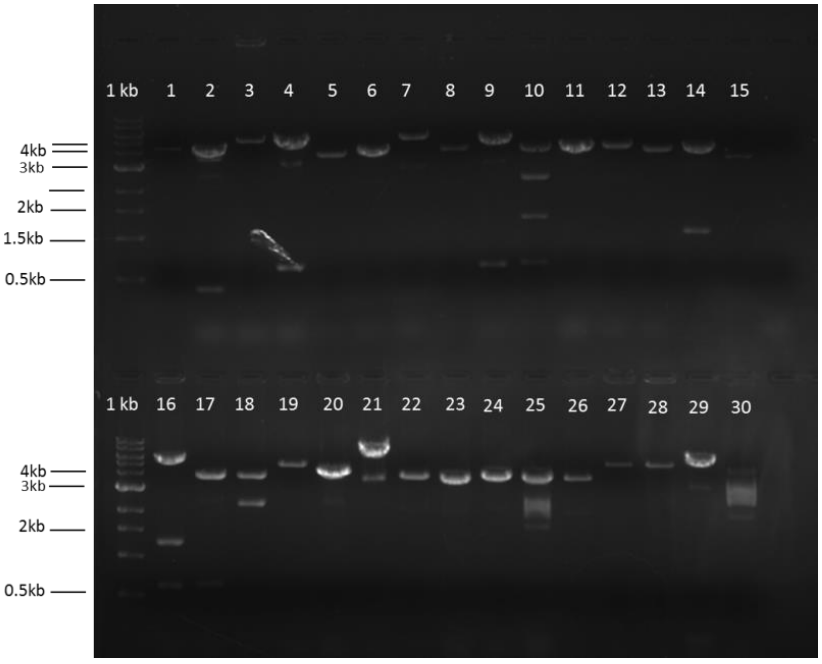


Fig 3.29. Digestion of colonies picked from the second trial of PSTblue cloning by *BamHI*

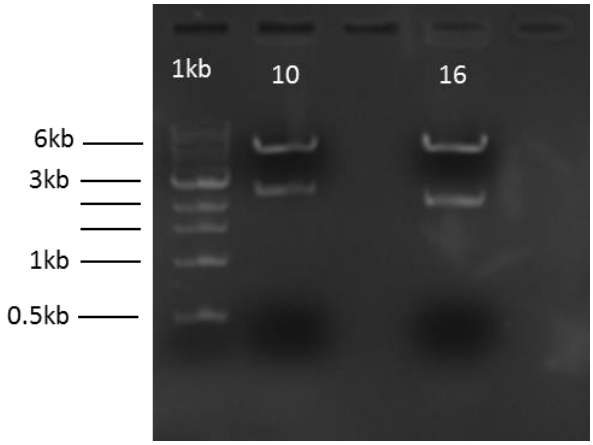


Fig 3.30. Digestion of colonies #10 and #16 picked from cloning 2 by *HindIII*

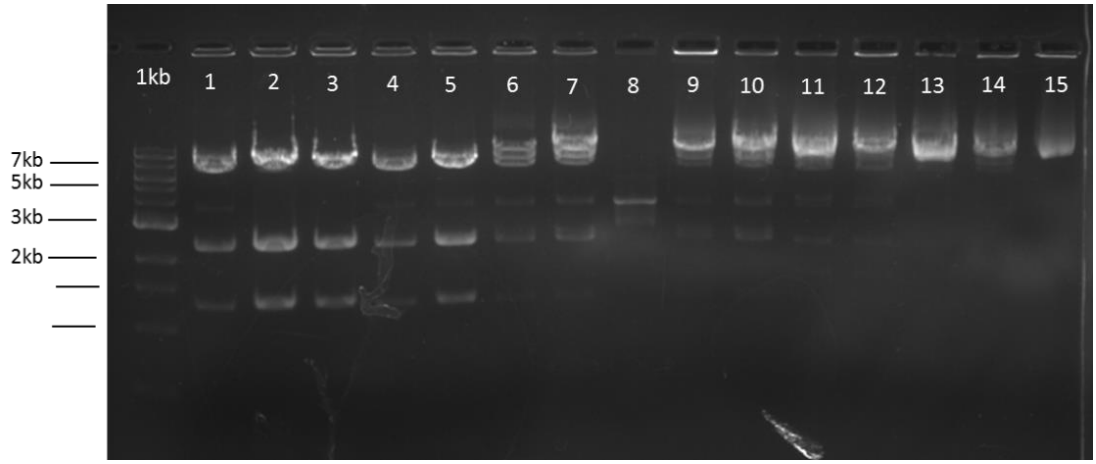


Fig 3.31. Colonies picked from sub-cloning to pBM61 digested by *BamHI*. 3 bands of 7.3 kb, 2.3 kb and 1.3 kb were expected. Colonies number 1, 2, 3, 4 and 5 show the expected sizes.

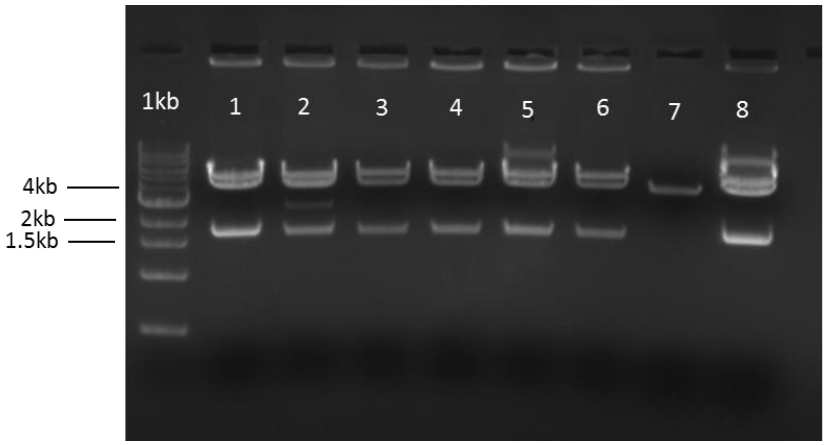


Fig 3.32. Colonies picked from sub-cloning to pBM61 digested by *HindIII*. 3 bands of 5.2 kb, 4 kb and 1.7 kb were expected. Colonies number 1, 2, 3, 4, 5 and 6 show the expected sizes. Colonies 7 and 8 do not look as predicted.

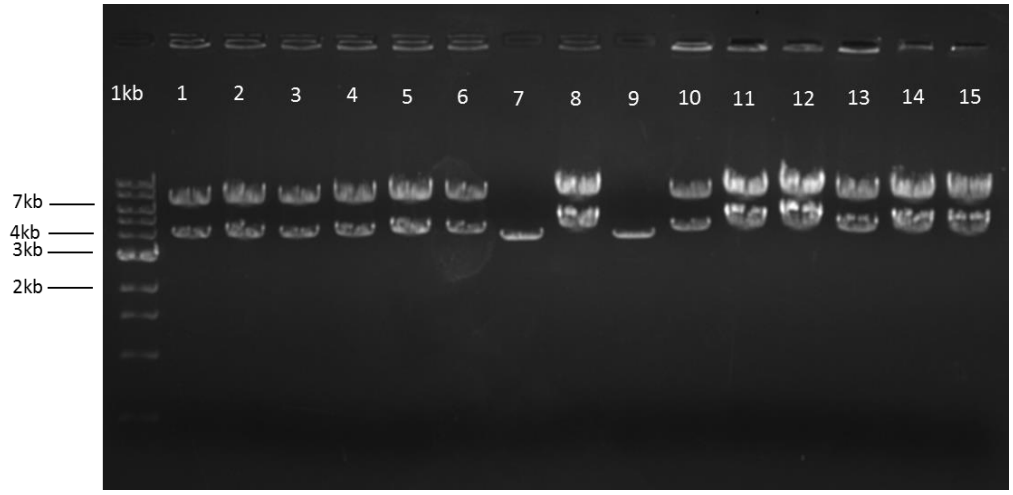


Fig 3.33. Colonies picked from sub-cloning to pBM61 digested by *EcoR1*. 2 bands of 7 kb and 4 kb were expected. All the colonies show the predicted size except colonies number 7 and 9.

3.5. Complementation

After cloning the gene into the *his-3* targeting plasmid, the plasmid was transformed into fungal spores. For complementation, strain SG#310 (*bd csp-1 chol-1 his-3 prd-1*) was used. Fresh spores were harvested and in the first trial they were suspended in 40 ul of sorbitol. The spore suspension was too viscous and transformation failed. In the second trial spores were suspended in 2ml of sorbitol and transformation was successful. After electroporation they were plated and after 3 days 33 colonies were picked and grown on choline media. All of the colonies grew on choline media and all of 33 transformants were placed on race tubes to check the growth rate and period compared to SG#26 (*bd csp-1*) and SG#260 (*bd csp-1 prd-1*).

The results from race tubes showed that transformants were complemented and showed growth rate and period similar to wild type (Fig.3.34 and Table 3.10). Because the transformed macroconidia may have two or three nuclei, the transformants may be heterokaryons with both *his-3* and his^{+} nuclei. Therefore, microconidia were made to purify single nuclei from strains 29, 4, 14, 34 and 22. In order to make microconidia, a medium that induces stress, prevents the fungus from making aerial hyphae and forces the fungus to make microconidia that have single nuclei was used. Microconidia were plated and colonies were picked for each strain. The growth rates and periods of microconidia colonies growing on race tubes are shown in Figure 3.35 and Table 3.11.

To check the transformation, DNA from these transformants was extracted. Primers were designed to amplify the fusion between the insert and the *his-3* downstream insertion site (Fig. 3.36) and PCR was done with DNA from the original transformants.

The expected bands were ~1.5kb and could be seen on the gel (Fig.3.37) which means that transformation was successful. This result indicates that the *prd-1* gene is identified as an RNA helicase that is including a DEAD box domain and helicase domain.

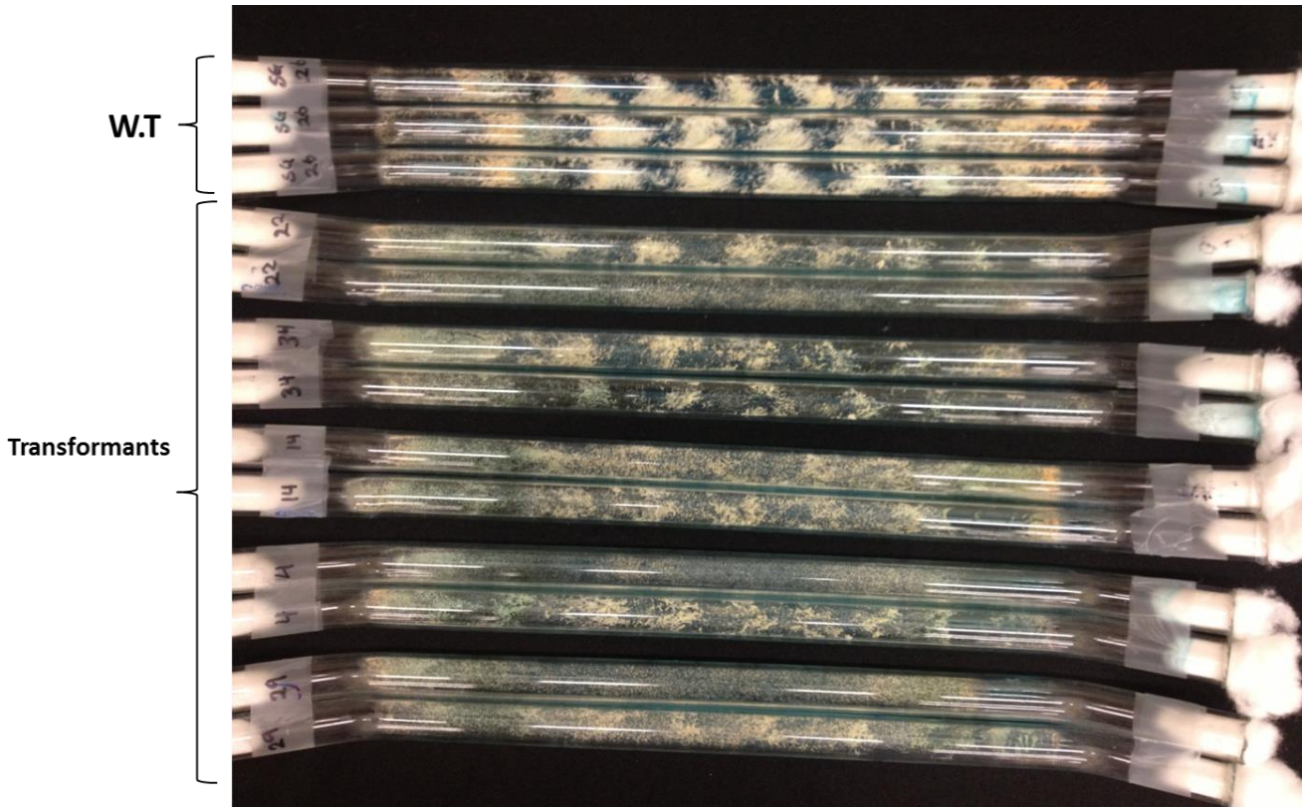


Fig 3.34. prd-1 transformants phenotype

Strain Number	Growth Rate	Period	G.R Average	Pr. Average
SG#26(bd,csp-1)	1.34	22.54	1.33	22.27
	1.34	22.55		
	1.32	21.73		
SG#260(bd,csp-1,prd-1)	0.78	24.77	0.78	24.62
	0.80	24.48		
	0.76	24.61		
Colony#34	1.36	21.33	1.31	22.09
	1.26	22.85		
Colony#14	1.27	22.56	1.23	23.51
	1.20	*24.47*		
Colony#4	1.26	20.14	1.26	20.14
	0.45			
Colony#29	1.20	23.21	1.20	23.21
	0.53			
Colony#22	1.35	22.77	1.35	22.77
	0.88			

Table 3.10. Growth rate and period of transformants after complementation. Two or three replicate race tubes were grown for each strain. For colonies # 4, 29 and 22 only one of the tubes showed bands and the growth rate was slow so they were not complemented.

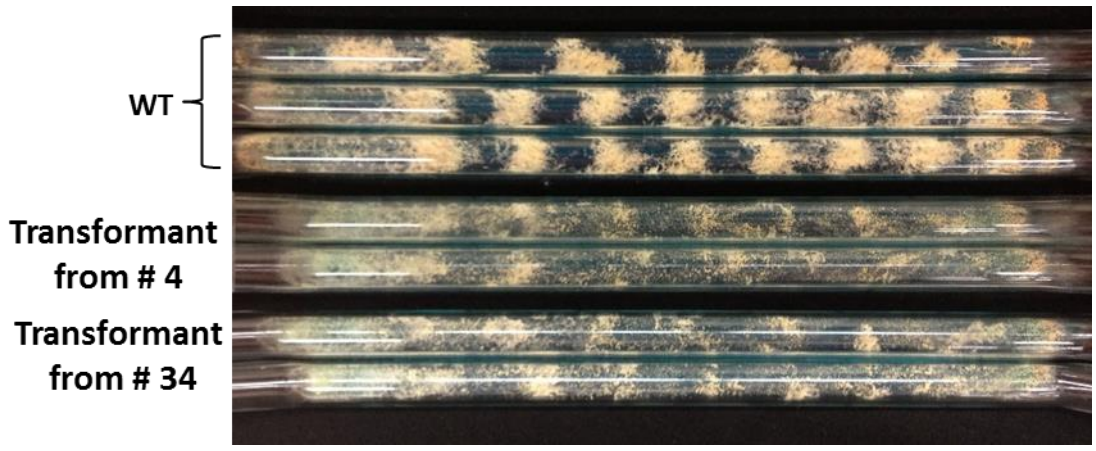


Fig 3.35. Phenotypes of prd-1 microconidia colonies.

Strains	Period (h)	Growth Rate (mm/h)
w.t	21.59	1.27
prd-1 chol-1	25	0.74
Transformant 1 (from # 4)	21.49	1.22
Transformant 2 (from #34)	22.83	1.19

Table 3.11. Average of growth rate and period of microconida colonies compared to w.t and mutant strains.

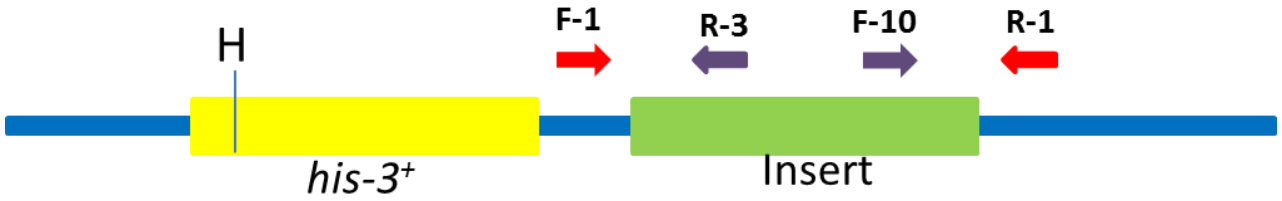


Fig 3.36. Designing the fusion primers. Fusion primers were designed (F-1 and R-1 are in the *his-3* downstream sequence) to check if the insert is in the right place.

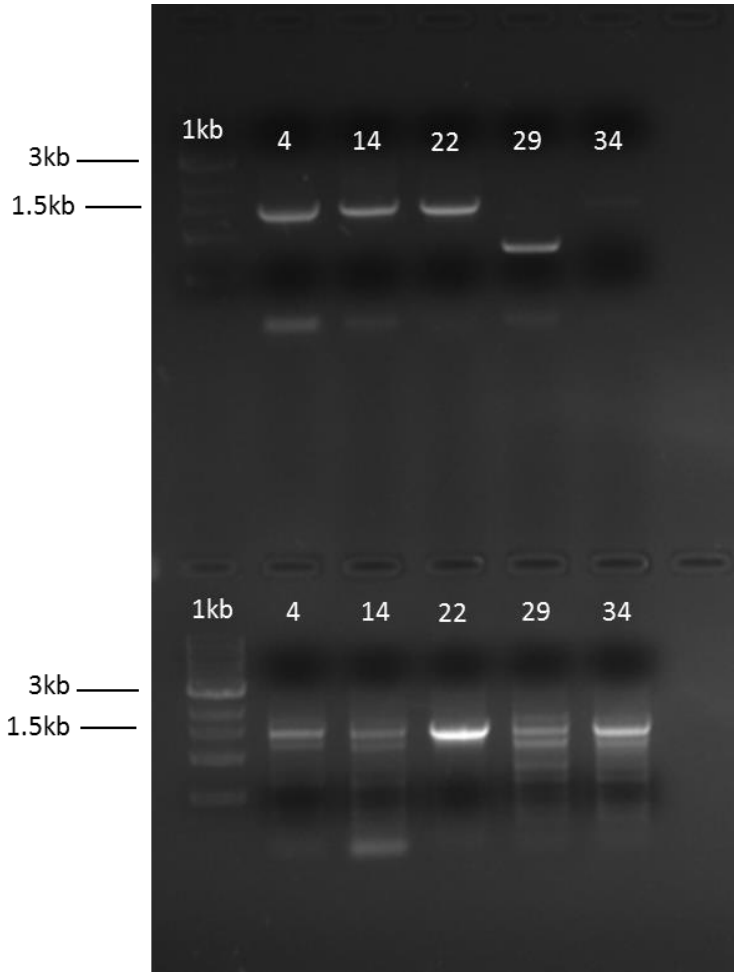


Fig 3.37. PCR with fusion primers of transformant DNA.

Top lanes are primers F-10 and R-1, bottom lanes are primers F-1 and R-3. Number 4, 14 and 22 showed the expected size.

3.6. Effect of *prd-6* on FLO

As mentioned in the introduction, five *prd* mutations affect the period length of the conidiation rhythm. From the experiments of Li et al. (2010) it was shown that *prd-1* and *prd-2* genes are important components of FLO. The *prd-6* gene was also shown to be involved in FLO by adding geraniol or farnesol to the media (Lombardi et al. 2007) and it has the role of an RNA helicase. Because I have identified *prd-1* as an RNA helicase, I wanted to see if other helicases are involved in FLO. In order to check if *prd-6* also affects the FLO in our lab assay, which is low and high choline concentration media, a cross between SG#127 (*bd csp-1 chol-1 frq*^{9}) and FGSC#11230 (*prd-6* ^{ko}) was set up to construct *chol-1 prd-6* double mutants.

187 spores were picked and were checked for *csp*$^{+}$ and *csp-1* by tapping the baby tubes: the ones with flying spores are *csp*$^{+}$, however progeny with clumpy spores are *csp-1*. 63 *csp-1* progeny were tested on race tubes (with minimal and choline media) The growth rate and period were calculated by Mac Tau software. *Chol-1* progeny grow slowly on minimal medium. Progeny that are *bd* show bands in race tubes and *bd*$^{+}$ progeny show fast growth rate and no bands. In *frq*^{9} progeny, mycelia are fluffy and do not show any bands. Progeny were also tested for hygromycin resistance; the *hyg*^{r} progeny have the *prd-6* knockout parent (the knockout gene is replaced by the *hph* cassette which confers hygromycin resistance) (Colot et al., 2006).

Results of race tubes are shown in Table 3.12. In order to get better results this experiment was repeated with more replicates. I repeated the race tubes for three progeny of each genotype, and for each strain five tubes of minimal and five tubes of choline media were used. The summary of results is shown in table 3.13. For statistical test, all 15

replicates of one genotype were pooled. A t-test between *frq*^{9} *prd*$^{+}$ and *frq*^{9} *prd-6* was calculated for growth rate in minimal media (p-value = $2.18*10^{-6}$, t =6, df = 11) and for growth rate in choline media (p-value = $2.22*10^{-9}$, t = 2.69, df = 11) and shows that the growth rates are significantly different. Periods in minimal media are not different (p = 0.39, t = 0.76, df = 11). A t-test between *frq*$^{+}$ *prd*$^{+}$ and *frq*$^{+}$ *prd-6* shows that in choline media growth rate and period are significantly different and period is shorter in *prd-6* strains (p = $2.37*10^{-8}$, t = 7.95, df = 11 and p = $1.56*10^{-8}$, t = 7.2, df = 11 respectively). In minimal media growth rate is not significantly different however period is significantly different (p = 0.55, t = 0.5, df = 11 and p = $6.71*10^{-5}$, t = 4.18, df = 11 respectively). It can be concluded that *prd-6* has an effect on FLO and lengthens the period but to confirm this conclusion more repetition of experiments needs to be done.

Genotye	Strain#	minimal		choline	
		Growth R.	Period	Growth R.	Period
bd csp-1 chol-1 frq9 prd-6+	#13	0.49	65.95	1.22	0
	#22	0.5	81.17	1.19	0
	#28	0.45	93.27	1.22	0
	#10	0.54	50.96	1.29	0
bd csp-1 chol-1 frq9 prd-6	#11	0.72	54.8	1.12	0
	#17	0.7	70.95	1.08	0
	#1	0.6	81.79	1.06	0
	#21	0.63	81.87	1.09	0
	#46	0.56	98.12	1.01	0
	#31	0.64	75.57	1.31	0
bd csp-1 chol-1 frq+ prd+	#25	0.53	0	1.17	25.13
	#48	0.59	0	1.18	25.54
	#61	0.65	0	1.28	25.42
	#56	0.51	65.88	1.25	24.3
	#30	0.65	0	1.12	25.74
bd csp-1 chol-1 frq+ prd-6	#35	0.42	78.2	0.88	25.65
	#4	0.47	77.3	0.97	19.15
	#43	0.48	80.29	0.89	19.23
	#47	0.52	72.14	0.97	21.88
	#58	0.53	65.98	1.07	19.44
	#38	0.51	78.45	1.02	19.92
	#32	0.44	51.23	0.72	22.74
	#63	0.54	55.75	0.89	21.43
bd csp-1	SG#1	1.51	20.51	1.46	19.28
bd csp-1 chol-1 frq^{9}	SG#127	0.44	86.93	1.55	38.41
bd csp-1 chol-1	SG#26	0.47	76.39	1.44	19.32

Table 3.12. Strains from cross between SG#127 and FGSC# 1123

Genotype	Progeny #	Minimal		Choline	
		Period	G.R	Period	G.R
bd csp-1 chol-1 frq9 prd+	#10	55.46± 4.1	0.56± 0.01	0.00	1.56± 0.04
	#22	54.87± 2.4	0.57± 0.01	0.00	1.64± 0.09
	#28	54.80± 1.7	0.53± 0.009	0.00	1.74± 0.07
bd csp-1 chol-1 frq+ prd+	#61	51.84± 3.13	0.50± 0.02	21.15± 0.32	1.54± 0.07
	#56	61.01± 4.36	0.50± 0.03	21.14± 0.17	1.76±0.08
	#25	54.31± 6.7	0.57± 0.03	21.31± 0.37	1.57± 0.03
bd csp-1 chol-1 frq9 prd-6	#17	51.79± 1.5	0.63± 0.01	0.00	1.27± 0.01
	#21	44.92± 4.2	0.74± 0.04	0.00	1.26± 0.009
	#46	60.51± 4.4	0.74± 0.03	0.00	1.15± 0.03
bd csp-1 chol-1 frq+ prd-6	#58	74.75± 4.5	0.53± 0.01	17.04± 0.07	1.25± 0.01
	#43	77.90± 1.7	0.47± 0.01	18.54± 1.15	1.13± 0.03
	#47	65.31± 2.66	0.54± 0.01	16.5± 0.6	1.32± 0.01
bd csp-1 chol-1	SG#26	55.48± 3.6	0.52± 0.02	21.32± 0.1	1.42± 0.02
bd csp-1 chol-1 frq9	SG#127	59.33± 3.2	0.52± 0.03	0.00	1.68± 0.02

Table 3.13. Summary of more replications of prd-6 experiment. Mean and SEM of each group were calculated.

3.7. Is prd-2 an RNA helicase?

According to data from *prd-2* mapping from our lab (*prd-2* has not been identified yet), there were no recombinants between the markers that were used in mapping and a region containing the gene NCU01143 known as *msp-8,* an RNA helicase. We hypothesized that *msp-8* could be *prd-2* and can interact with other helicases such as *prd-1* and *prd-6*, so it was decided to sequence the gene. In order to sequence the *msp-8* gene, I extracted DNA from SG#171 (*bd prd-2*) for PCR. Primer F-1 and R-1 were used for PCR (table 3.14). Thirteen forward and six reverse primers were designed to cover 519 bp upstream and 607 bp downstream of the gene. Sequencing was repeated for some regions but I didn't find any mutation in this gene, therefore it can be concluded that *msp-8* is not *prd-2*.

However, if RNA helicases are important in the FLO, *msp-8* could have an effect on FLO. Therefore based on this hypothesis a cross between SG#127 (*bd csp-1 chol-1 frq*^{9}) and FGSC#12359 (*msp-8* ^{ko}) was set up. From the cross 167 spores were picked, 61 *csp-1* progeny were tested in race tubes on minimal and choline media, 33 *chol-1* progeny were collected and the growth rate and period were calculated by Mac Tau software. Genotypes for *csp-1, bd, chol-1* and *frq9* were determined as for the *prd-6* cross. Progeny were also tested for hygromycin resistance; the *hyg*^{r} progeny have *msp-8* knockout. The preliminary results are shown in table 3.14, but to get more reliable results the experiment was repeated using a few progeny with more replications.

From results presented in Table 3.15, and t-test between *frq*$^{+}$ *msp*$^{+}$ and *frq*$^{+}$ *msp-8* it can be concluded that there is no significant difference in growth rate on minimal media ($p = 0.3$, $t = 0.8$, $df = 11$). There is no difference in growth rate and period in choline media ($p = 0.5$, $t = 0.51$, $df = 11$ and $p = 0.3$, $t = 0.8$, $df = 11$ respectively). T-test between *frq*^{9} *msp*$^{+}$ and *frq*^{9} *msp-8* shows that there is a significant difference in growth rate in minimal media ($p = 0.006$, $t = 3$, $df = 11$) but there is no difference in choline media ($p = 0.24$, $t = 1.29$, $df = 6$). *Msp-8* strains were arrhythmic on minimal unlike *msp*$^{+}$. Therefore, Msp-8 may be one of the components of the FLO but further experiments are being done to confirm these results.

Genotype	Strain #	Minimal		Choline	
		Growth Rate	Period	Growth Rate	Period
bd csp-1 chol-1 frq^{+} msp-8^{+}	#35	0.38	0	1.11	24.76
	#43	0.44	0	1.36	21.07
	#44	0.39	0	1.22	22.04
	#48	0.27	106.23	1.07	19.92
	#51	0.35	0	1.36	22.05
	#54	0.4	0	1.46	20.42
	#65	0.41	0	1.36	20.79
	#68	0.41	0	1.36	23.42
bd csp-1 chol-1 frq^{9} msp-8^{+}	#1	0.35	0	1.35	0
	#2	0.4	0	1.57	0
	#5	0.39	0	1.58	0
	#10	0.43	0	1.54	0
	#11	0.4	75.68	1.52	0
	#12	0.43	0	1.54	0
	#26	0.4	0	1.54	0
	#38	0.44	0	1.38	0
bd csp-1 chol-1 frq^{+} msp-8	#31	0.38	0	1.36	21.27
	#32	0.33	0	1.29	21.99
	#34	0.31	0	1.28	22.15
	#40	0.42	0	1.47	21.35
	#50	0.36	83.94	1.48	21.6
bd csp-1 chol-1 frq^{9} msp-8	#4	0.39	0	1.58	0
	#18	0.34	0	1.47	0
	#19	0.36	0	1.53	0
	#20	0.65	0	1.43	0
	#22	0.39	0	1.46	0
	#27	0.66	0	1.46	0
bd csp-1	SG#1	1.51	20.51	1.46	19.28
bd csp-1 chol-1 frq^{9}	SG#127	0.44	86.93	1.55	38.41
bd csp-1 chol-1	SG#26	0.47	76.39	1.44	19.32

Table 3.14. Choline requiring progeny from the msp-8 cross. Growth rate and period are shown.

Genotype	Progeny #	Minimal		Choline	
		Period	G.R	Period	G.R
bd csp-1 chol-1	#48	59.25± 3.9	0.43± 0.02	20.52± 0.5	1.59± 0.006
	#58	65.25± 6.1	0.43± 0.02	20.44± 0.2	1.47± 0.01
	#43	65.37± 2.1	0.43± 0.01	21.17± 0.4	1.43± 0.02
bd csp-1 chol-1 frq9	#11	60.85± 5.2	0.45 ±0.01	0.00	1.56± 0.01
	#12	62.74± 3.07	0.45±0.008	0.00	1.54± 0.02
	#26	56.92± 7.8	0.46± 0.008	0.00	1.48±0.05
bd csp-1 chol-1 msp-8	#50	0.00	0.38± 0.01	20.9± 0.2	1.39± 0.02
	#51	0.00	0.33± 0.008	21.15± 0.24	1.44± 0.06
	#40	0.00	0.51±0.027	20.79± 0.18	1.59± 0.03
bd csp-1 chol-1 msp-8 frq9	#4	0.00	0.4± 0	0.00	1.55± 0.02
	#18	0.00	0.43± 0.02	0.00	1.61± 0.03
	#22	0.00	0.42± 0.02	0.00	1.5± 0.04
bd csp-1 chol-1	SG#26	56.96± 3.08	0.52± 0.01	20.42± 0.2	1.45± 0.01
bd csp-1 chol-1 frq9	SG#127	62.29± 1.65	0.45± 0.01	0.00	1.55± 0.007

Table 3.15. Summary of more replications of msp-8 experiment. Mean and SEM of each group is shown. Five replicate tubes were averaged for each strain.

4. Discussion

4.1. Gene identification

The purpose of this thesis was to finish mapping the *prd-1* gene and identify the gene. CAPS markers were used to identify a narrow interval containing the gene. The mapping was continued by sequencing the genes in the narrowed interval and the mutation was found in one of the candidate genes. In order to identify the gene by complementing the mutation, the gene was cloned in the *his-3* targeting plasmid and transformed into the *his-3 prd-1* strain.

After complementation of the *prd-1* candidate gene, and rescue of the wild type phenotype (growth rate and period) it was confirmed that this gene (NCU07839) is *prd-1*. This gene is identified as an RNA helicase that contains a DEAD box domain and a helicase domain. The DEAD-box class of RNA helicases have fundamental roles in RNA and RNP structure in all the phases of RNA metabolism.

4.2. Amino acid annotation

The mutation is a substitution of G to A at the 5' splice junction of one of the introns and cDNA alignments showed that this mutation changes the reading frame and can change the amino acid sequence of the protein produced. The wild type amino acid sequence shows that the protein contains a DEXDc domain (DEAD-like helicases superfamily) which is located from amino acid 151 to 354, and HELICc domain (helicase superfamily c-terminal domain) which is located from amino acid 395 to 476. There are also two low complexity regions at the beginning and the end of the protein. The region

starts at position 2 and ends at position 69, and position 519 and ends at position 551. Domains are shown in figure 4.1.

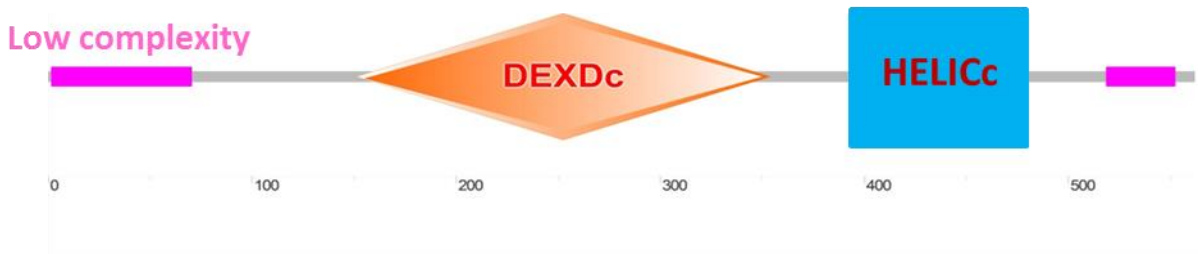


Fig. 4.1. Domains of RNA helicase that are the DEAD-box domain and helicase domain and two low complexity regions. The protein sequence from Broad Institute Database was introduced into SMART website (smart.embl-heidelberg.de) to generate this schematic.

The wild type amino acid sequence is shown in Figure 4.2. Different domains are shown by different colors. After changing the reading frame based on cDNA sequencing for the mutant, the stop codon was eliminated based on the alteration of the sequence (Fig 4.3). By adding nucleotides downstream to find the stop codon the amino acid sequence of the mutant was predicted (Fig.4.4). Wild type protein has 563 amino acids and mutant protein has 611 amino acids.

As can be seen this mutation does not have any effect on the DEAD-box domain or helicase domain but it changes the low complexity region at the C-terminus. This region is enriched by Glycine in wild type; however it is enriched in Valine in the mutant (Fig 4.3 and 4.4). This mutation may have an effect on a binding site between this protein and other proteins. As can be seen in predicted tertiary structure of the wild type and mutant, in the mutant there is an extra turn in the last helix at the C-terminus (shown with a white

arrow) therefore there is a difference in arrangement and conformation of the wild type and mutant proteins (Fig 4.5).

Fig. 4.2. Wild type amino acid sequence. Purple and green are low complexity regions; red is DEXDc domain and blue is HELICc domain. Every 10 amino acid is scaled by *.

Fig 4.3. Mutant amino acid sequence without stop codon. Purple and green are low complexity regions; red is DEXDc domain and blue is HELICc domain. From amino acid 489 the sequence is different in wild type and mutant. Every 10 amino acid is scaled by *.

Fig 4.4. Mutant amino acid sequence with stop codon (mutation makes the protein larger than wild type). Every 10 amino acid is scaled by *.

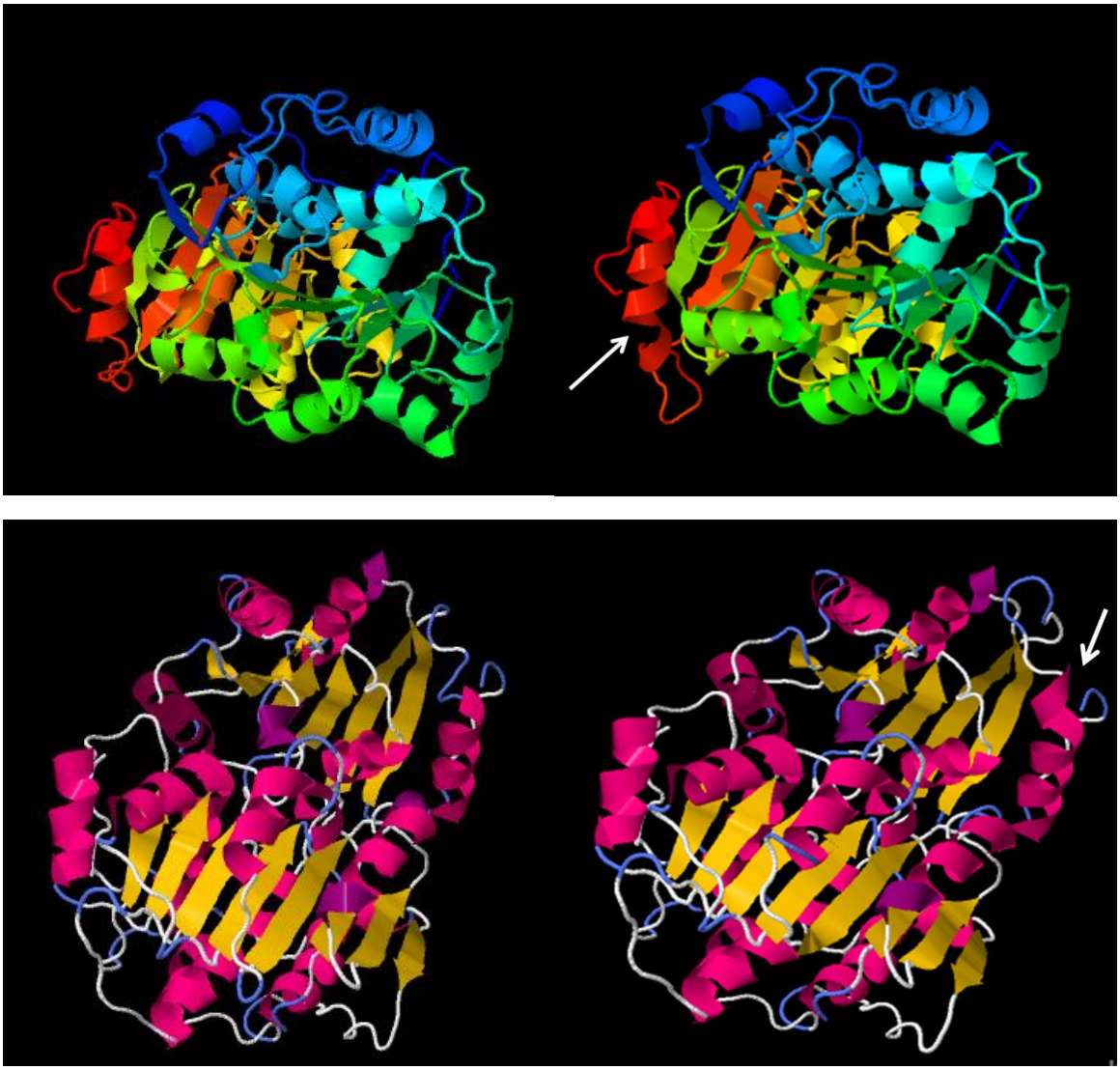


Fig 4.5. Wild type (left panels) and mutant (right panels) predicted tertiary structure from Phyre website (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index). White arrow shows the difference between wild type and mutant.

4.3. DEAD-box RNA helicases

While RNA is known for being single stranded, more often than not it has the tendency to fold on itself and form intramolecular helices important for RNA molecule function. The efficient binding and interaction of RNA with RNA-binding proteins is almost entirely dependent on the RNA folding (Zemora and Waldsich 2010; Woodson 2010). One class of enzymes that control and regulate cellular RNA structure are DEAD-box RNA helicases. These helicases belong to one of the largest classes of the RNA helicase superfamily. DEAD-box proteins have a wide range of functions some of which include: duplex unwinding, ATPase activity, displacement of RNA-binding protein from single-stranded RNA, RNA strand annealing, and acting as a platform for larger ribonucleoprotein complexes. Despite being classically defined as helicase enzymes, DEAD-box proteins can also function as RNA chaperones, and are involved in numerous biochemical processes to form and assemble RNPs.

DEAD-box helicases share the highly conserved SF2 helicase core that contains two nearly equal domains (Fig. 4.6), with each domain bearing similarity to the recombination protein RecA. The C- and N-terminal domains of DEAD-box helicases are often not conserved. Two helicase domains are connected with a flexible linker, allowing it to change orientation. This is particularly important as it facilitates the opening and closing of these two domains, which are essential for the function of the enzyme. For instance when the core domains are closed, the enzyme can have an ATPase function (Linder et al. 2011). The name DEAD box for this family is given because of motif II that contains the Asp-Glu-Ala-Asp motif.

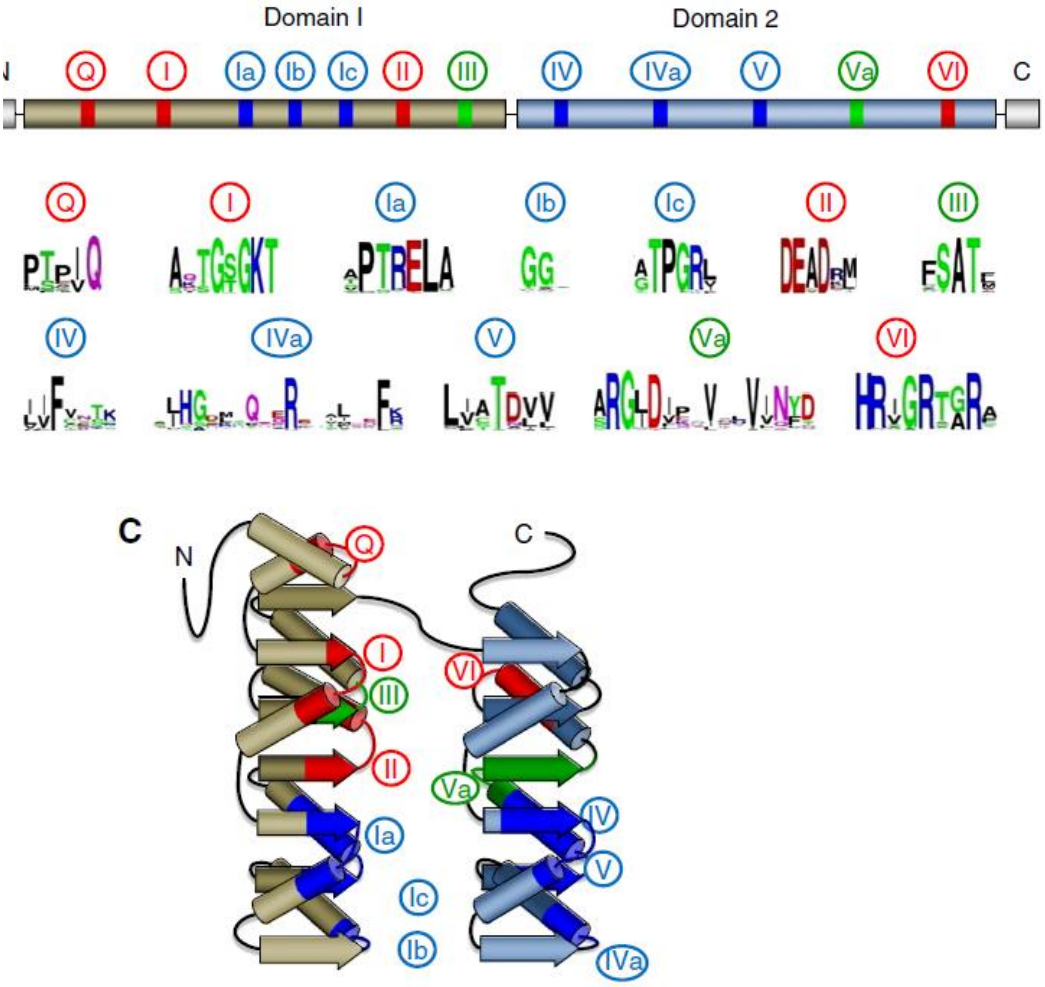


Fig 4.6. A) The primary structure of DEAD-box RNA helicase. Domain 1 (N-terminus) is shown in brown and domain 2 (C-terminus) is shown in blue. Different motifs are shown and color-coded based on their activity (blue is RNA binding, red is ATP binding and green is coordination between ATP and RNA binding). Circled numbers show the location of the sequenced motifs. B) The conserved sequences of motifs of DEAD-box helicases are shown. Amino acid conservation is represented by the height of the letter. C) Tertiary structure of two RecA like domains is shown, β-

strands are indicated by arrows, α-helices are indicated by cylinders. (Adapted from Putnam and Jankowsky 2013)

4.4. DEAD-box RNA helicase roles in the cell

4.4.1. The basic biochemical features

RNA binding and coupling to ATP binding: RNA is bound by the helicase core of probably all DEAD-box proteins through a set of conserved residues in motifs Ia, b, c and motifs IV, IVa, V, and Va (Fig.4.6). Mallam and colleagues (2012) have shown that single stranded RNA can bind to both helicase domains; they further conclude that in this way RNA stimulates the formation of the closed state of these domains thereby forming the active site for ATP hydrolysis. It is worth mentioning that RNA duplex can only bind to domain 2 (Mallam et al. 2012).

Nucleotide binding and metabolite sensing: According to Parvatiya and colleagues (2012) some DEAD-box helicases can bind other adenine-based and related nucleotides (i.e. AMP). In the case of these helicases, AMP binds exclusively to ATP binding residues in the helicase domain 1 (Hogbom et al. 2007). It is also known that DDX41 binds to cyclic di-GMP (c-di-GMP) and cyclic di-AMP (c-di-AMP) but it is not known whether they bind like ATP or ADP. Consequently, it can be concluded that DEAD-box helicases could play a role in providing energy for RNA metabolism in response to metabolic stresses that increase AMP levels (Parvatiyar et al. 2012)

Protein binding: Like most other enzymes, DEAD-box helicases require other proteins to carry out their functions. Such proteins bind a DEAD-box helicase at the auxiliary and core domains of the enzyme (Linder et al. 2011).

Oligomerization and posttranslational modifications: RNA, nucleotide and protein binding of DEAD-box helicases depends on the oligomeric state of the DEAD-box helicase, and its post-translational modifications define how RNA, nucleotides, and proteins bind the helicase. DEAD-box helicases can be functional in oligomeric, dimeric or even monomeric states. The post-translational modifications of DEAD-boxes include: phosphorylation, methylation, and sumo modification (Jacobs et al. 2007).

4.4.2. RNA-related activities of DEAD-box helicases

RNA-stimulated ATPase activity: DEAD-box helicases are often referred to as RNA dependent ATPases, as they are all capable of hydrolyzing ATP in the presence of RNA (Parvatiyar et al. 2012).

RNA unwinding: The ATP driven unwinding of RNA, requires coordination between ATP and RNA binding. RNA unwinding is associated with conformational changes of the enzyme (Jarmoskaite et al. 2011). Mutations in DEAD-box proteins can impair unwinding the RNA duplexes, therefore impairing cellular functions.

RNA clamping: RNA clamping refers to the capacity of a DEAD-box helicase to remain bound to RNA for a longer period of time. Therefore, it can make a platform for larger complexes.

Protein displacement: Remodeling of RNA–protein complexes (RNPs) by displacement of proteins is thought to be a main function of DEAD-box and other helicases. Thus it can help to unwind the duplex RNA.

Strand annealing: Several DEAD-box helicases are known to promote the formation of duplex structure. Strand annealing does not require ATP or ADP. Mallam et al. (2012) have shown that that DEAD-box helicases bind to a duplex in domain 2. It was proposed that the ability to promote duplex formation can be related to other aspects of the enzyme's function as RNA chaperones (Jarmoskaite et al. 2011).

RNA structure conversion: DEAD-box helicases have the function of converting one RNA structure to another. This function is also known as RNA chaperones.

4.5. RNA helicases and regulation of circadian rhythms in *Neurospora*

FRQ-interacting RNA helicase (FRH) is a member of the SK12 subfamily of RNA helicases, and an important DEAD box-containing RNA helicase in *Neurospora*. FRH contains a DEAD/DEAH box and helicase C domains. It makes a complex with FRQ (FFC) and is known as an important clock component. Chang et al. (2005) have used *frh* specific dsRNA that downregulates *frh* expression, to show that a complex of FRQ-FRH plays an important role in sustaining the steady level of FRQ.

When FRH is downregulated levels of FRQ reduce and levels of *frq* RNA increase, this can impair the negative feedback loop. Moreover, it is known that FRH moderates the interaction between FRQ and WC complexes. FRQ-FRH interaction is important for FRQ-WC interactions because once the former interaction is disrupted, the latter interaction will be eliminated (Cheng et al. 2005).

Dob1p/Mtr4p is known to be a homolog for FRH in yeast. Dob1p/Mtr4p is a RNA helicase that is the main component of the exosome complex and acts as a cofactor for the exosome to degrade or process RNA. (Liang et al. 1996; de la Cruz et al. 1998; Jacobs et al. 1998; Torchet et al. 2002).

A study from Shi and colleagues (2010) has shown that a mutation in *frh* (frh^{R806H}) can stop the rhythmicity but the interaction between FRQ and FRH remained intact indicating that FRQ-FRH interaction with WCC is disturbed. In the mutant strain phosphorylation of WC-1 is reduced and WC-1 stays hypophosphorylated. Overexpression of FRH leads to accumulation of WC-1 indicating that FRQ-FRH is important in WC-1 stability. In the FRH protein that has 1107 amino acids, the DEAD/DEAH domain is located in the N-terminal region and the helicase domain is placed in the middle of the protein. The R806H mutation was created close to the C-terminal DSHCT domain. The functional domains for RNA helicase stay intact in this mutant, therefore the interaction of FRH and FRQ was not disrupted. However, it was speculated in this study that this mutation impairs the binding region of WCC. This study showed that although FRH is essential and functional in the exosome, it has a specific role in clock-related responses (Shi et al. 2010).

4.6. Other DEAD-box helicases and RNA-binding proteins that are involved in circadian rhythms

Recently it was reported that RNA helicases DDX5 (DEAD-box polypeptide 5) and DHX9 (DEAH-box protein 9) are components of mouse PER complexes. Padamanbhan et al. (2012) showed that knockdown of DDX5 and DHX9 can lead to

increasing the per1 gene transcription and can alter the period. Moreover, FRH in *Neurospora* is related to DDX5 and DHX9.

Mouse LARK protein (mLARK) is an RNA-binding protein that binds to the 3' UTR of per1 and upregulates mPER1 protein expression. The homolog of mLARK in human, RNA-binding motif protein 4 (RBM4), has important roles in mRNA metabolism and translation control and alternative mRNA splicing. Knockdown of mLARK by siRNA resulted in shortening the period; however overexpression of mLARK lengthened the period. Thus, it is speculated that mLARKs have a role in mPER1 translation (Kojima et al. 2007).

In *Drosophila* it was known that decreased or increased levels of lark resulted in early- or late pupa eclosion phenotypes, respectively (Newby and Jackson 1993). In a recent study in *Drosophila*, using RNAi to knockdown LARK and decrease its levels in clock neurons, it was shown that knockdown of LARK in clock neurons and also non-clock cells resulted in arrythmicity in locomotor activity. This indicated that LARK is important for activity rhythms (Sundram et al. 2012).

In Arabidopsis, a nuclear RNA-binding protein (AtGRP7) has been found as a component of the circadian clock system.This protein is an output that may regulate clock-controlled transcripts but it is not part of the oscillator. Its effect on translation has not been well studied (Heintzen et al. 1997).

4.7. PRD-1 homolog

Using a BLAST search the *Prd-1* gene was found to be ortholog with Dbp2 (Appendix-2). Cloutier et al. (2012) have demonstrated that DEAD-box protein Dbp2 is associated with transcribing genes and has an important role in gene expression.

They have proposed that Dbp2 is a double- stranded RNA-specific ATPase that is required for accurate transcription and has a critical role in maintaining transcriptome and transcription fidelity (Cloutier et al. 2012).

Cotranscriptional and posttranscriptional processes such as splicing, 3’ end cleavage and polyadenylation require proper nuclear mRNP assembly. Thus, mRNP must assemble with RNA-binding proteins to be involved in gene expression. It was shown that Dbp2 has a crucial role in mRNP assembly in the nucleus (Ma et al. 2013). The human ortholog of Dbp2, named p68, has been associated with many transcriptional and post-transcriptional processes like transcriptional regulation, alternative splicing, and microRNA processing (Janknecht 2010). These proteins are conserved in all eukaryotes.

The *prd-1* mutation is known to be not in the helicase and DEAD-box domain. The knockout for this gene does not exist in the stock centre indicating that the knockout of this gene should be lethal. Therefore, it can be said that PRD-1 protein is essential for *Neurospora.*

4.8. Future Directions

Since this mutation is identified as an RNA helicase there are several questions that can be addressed.

It is important to know what protein complexes are interacting with wild type that are absent or altered in PRD-1. In order to address this question an epitope tag can be used and after complementation and the phenotype verification, by immunoprecipitation and mass spectrometry the complex can be identified. After identifying the protein complex, mutations can be made in these genes or if the knockout is available it can be ordered from the stock centre, assaying the phenotype on race tubes can confirm their associations with FLO.

Since the mutation does not have effects on helicase and DEAD-box domains it may have effects on binding sites with other complexes. Since the gene is an RNA helicase, it is important to know whether catalytic activity of PRD-1 is required for its function. To approach this question complementation with a construct that has a mutation in the catalytic domain, which can include mutation in the RNA binding and ATP hydrolysis domain, can be done. If the complementation fails, it means catalytic activity of PRD-1 can regulate the circadian rhythm. Therefore it will be interesting to know what the RNA putative substrates are by RNA sequencing of immunoprecipitated PRD-1.

As was mentioned PRD-1 is an important component of FLO and it is important to find out how the substrates (which are RNAs) are functional for the FRQ-less oscillator. It is critical to understand what these RNAs code for and what proteins are being regulated by this RNA helicase (like FRQ for FRH). After identifying transcripts,

knockouts for these genes can be used from the stock center and if they are not available in vitro mutagenesis can be done to check their phenotype in race tubes to see how important they are for FLO.

PRD-6 is also identified as an RNA helicase, therefore there is a hypothesis that PRD-1 and PRD-6 interact with each other or even with other PRDs. This can be tested by the immunoprecipitation methods described above.

Another important question is whether the gene and protein expression are rhythmic for *prd-1*, which can be explored using northern and western blots respectively. Using the same technique the pattern of expression of WCC or FRQ will be studied in the *prd-1* mutant.

References

Akten B, Jauch E, Genova GK, Kim EY, Edery I, Raabe T, Jackson FR (2003) A role for CK2 in the *Drosophila* circadian oscillator. *Nat Neurosci* 6:251–257.

Alabadi D, Oyama T, Yanovsky MJ, Harmon FG, Mas P, Kay SA (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the *Arabidopsis* circadian clock. *Science* 293:880–83.

Allada R (2003) Circadian clocks: a tale of two feedback loops. *Cell* 112: 284–6.

Aronson B, Johnson K, Loros JJ, Dunlap JC (1994). Negative feedback defining a circadian clock: autoregulation in the clock gene frequency. *Science* 263:1578–1584.

Ballario P, Vittorioso P, Magrelli A, Talora C, Cabibbo A, Macino G (1996) White collar-1, a central regulator of blue-light responses in *Neurospora*, is a zinc-finger protein. *EMBO J* 15:1650-1657.

Barger LK, Lockley SW, Rajaratnam SM, Landrigan CP (2009) Neurobehavioral, health, and safety consequences associated with shift work in safety-sensitive professions. *Curr Neurol Neurosci Rep* 9:155–164.

Boxall SF, Foster JM, Bohnert HJ, Cushman JC, Nimmo HG, Hartwell J (2005) Conservation and divergence of circadian clock operation in a stress-inducible Crassulacean acid metabolism species reveals clock compensation against stress. *Plant Physiol* 137:969–82.

Brown DL, Feskanich D, Sánchez BN, Rexrode KM, Schernhammer ES, Lisabeth LD (2009) Rotating night shift work and the risk ofischemic stroke. *Am J Epidemiol* 169:1370–1377

Chang DC, Reppert SM (2003) A novel C-terminal domain of *Drosophila* PERIOD inhibits dCLOCK : CYCLE mediated transcription. *Curr Biol* 13:758–62.

Chen SK, Badea TC, Hattar S (2011) Photoentrainment and pupillary light reflex are mediated by distinct populations of ipRGCs. *Nature* 476:92–95.

Cheng P, He Q, Wang L, Liu Y(2005). Regulation of the *Neurospora* circadian clock by an RNA helicase. *Genes Dev* 19:234–241.

Cheng P, Yang Y, Heintzen C, Liu Y (2001) Coiled-coil domain mediated FRQ-FRQ interaction is essential for its circadian clock function in *Neurospora EMBO J* 20:101–108.

Chiu JC, Vanselow JT, Kramer A, Edery I (2008) The phosphooccupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. *Genes Dev* 22:1758–72.

Christensen MK, Falkeid G, Loros JJ, Dunlap JC, Lillo C, and Ruoff P (2004) A nitrate-induced frq-less oscillator in *Neurospora crassa*. *J Biol Rhythms* 19:280-286.

Cloutier SC, Ma WK, Nguyen LT, Tran EJ (2012) The DEAD-box RNA helicase Dbp2 connects RNA quality control with repression of aberrant transcription. *J Biol Chem* 287:26155-26166.

Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovick KA, Dunlap JC (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc Natl Acad Sci* 103(27):10352-10357.

Compton JE, Feldman JF (2004) Many genetic loci contribute to circadian rhythms in *Neurosporacrassa*, pp. 48–66 in Circadian Clocks in Eukaryotic Microbes, edited by F. Kippert. *Landes BioScience Georgetown, TX*.

Correa A, Lewis ZA, Greene AV, March IJ, Gomer RH, and Bell-Pedersen D (2003) Multiple oscillators regulate circadian gene expression in *Neurospora*. *Proc Natl Acad Sci U S A* 100:13597-13602.

Cote GG, Brody S (1987) Circadian rhythms in *Neurospora crassa*: a clock mutant, *prd-1*, is altered in membrane fatty acid composition. *Biochim Biophys Acta* 904:131–139.

Covington MF, Panda S, Liu XL, Strayer CA,Wagner DR, Kay SA (2001) ELF3 modulates resetting of the circadian clock in *Arabidopsis*. *Plant Cell* 13:1305–15.

Daniel X, Sugano S, Tobin EM (2004) CK2 phosphorylation of CCA1 is necessary for its circadian oscillator function in *Arabidopsis*. *Proc Natl Acad Sci U S A* 101:3292–97.

Davis, R. H. (2000) *Neurospora*: Contributions of a Model Organism. *Oxford University Press, Inc New York.*

de la Cruz J, Kressler D, Tollervey D, Linder P (1998) Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3 end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J* 17:1128–1140.

de Mairan J (1729) Observation botanique. Histoire de L'Academie Royale des *Science*s, pp 35-36.

de Paula RM, Lewis ZA, Greene AV, Seo KS, Morgan LW, Vitalini MW, Bennett L, Gomer RH, Bell-Pedersen D (2006) Two circadian timing circuits in *Neurospora* crassa cells share components and regulate distinct rhythmic processes. *J Biol Rhythms* 21(3):159-68.

de Paula RM, Lamb TM, Bennet L, Bell-Pederson D (2008) A connection between MAPK pathways and circadian clock. *Cell Cycle* 7:2630-2634.

Dibner C, Schibler U, Albrecht U (2010) The mammalian circadian timing system: organization and coordination of central and peripheral clocks. *Annu Rev Physiol* 72:517–49.

Ditty JL, Williams SB, and Golden SS (2003) A *Cyanobacterial* circadian timing mechanism. *Annu Rev Genet* 37:513-543.

Dodd AN, Salathia N, Hall A, Kévei E, Tóth R, Nagy F, Hibberd JM, Millar AJ, Webb AAR (2005) Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* 309:630–633.

Dunlap JC (1999) Molecular bases for circadian clocks. *Cell* 96:271–290.

Dunlap JC, Loros JJ (2004) The *Neurospora c*ircadian system. *J Biol Rhythms* 19(5):414-24.

Dunlap JC, Loros JJ, DeCoursey PJ (2004) Chronobiology: BiologicalTimekeeping. *Sinauer Associates Inc Publishers Massac.*

Edwards KD, Anderson PE, Hall A, Salathia NS, Locke JC, et al (2006) FLOWERING LOCUS C mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock. *Plant Cell* 18:639–50.

Feldman JF, Hoyle MN (1973) Isolation of circadian clock mutants of *Neurospora crassa*. *Genetics* 75:605-13.

Froehlich AC, Pregueirio A, Lee K, Denault D, Colot H, Nowrousian M, Loros JJ, Dunlap JC (2003) The molecular workings of the *Neurospora* biological clock. *Novartis Found Symp* 253:184-98.

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W,Ma L, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D et al (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422:859-868.

Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay SA. (2012). *Arabidopsis* circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proc Natl Acad Sci USA* 109:3167–3172.

Glossop NR, Lyons LC, Hardin PE (1999) Interlocked feedback loops within the *Drosophila* circadian oscillator. *Science* 286(5440):766–8.

Gorl M, Merrow M, Huttner B, Johnson J, Roenneberg T, Brunner M (2001) A PEST-like element in FREQUENCY determines the length of the circadian period in *Neurospora crassa*. *EMBO J* 20:7074–7084.

Granshaw T, Tsukamoto M, Brody S (2003) Circadian rhythms in *Neurospora crassa*: farnesol or geraniol allow expression of rhythmicity in the otherwise arrhythmic strains frq10, wc-1, and wc-2. *J Biol Rhythms* 18: 287–296.

Guler AD, Ecker JL, Lall GS, Haq S, Altimus CM, et al (2008) Melanopsin cells are the principal conduits for rod-cone input to nonimage-forming vision. *Nature* 453:102–5.

Hall JC, Chang DC, Dolezelova E (2007) Principles and problems revolving around rhythm-related genetic variants. *Cold Spring Harb Symp Quant Biol* 72:215–32.

Hanano S, Domagalska MA, Nagy F, Davis SJ (2006) Multiple phytohormones influence distinct parameters of the plant circadian clock. *Genes Cells* 11:1381–92.

Hanano S, Stracke R, Jakoby M, Merkle T, Domagalska MA, et al (2008) A systematic survey in *Arabidopsis thaliana* of transcription factors that modulate circadian parameters. *BMC Genomics* 9:182.

Hardin PE, Hall JC, and Rosbash M (1990) Feedback of the *Drosophila* period gene product on circadian cycling of its messenger RNA levels. *Nature* 343:536–540.

Harmer L (2009) The circadian system in higher plants. *Annu Rev Plant Biol* 60:357–77.

Heintzen C, Nater M, Apel K, Staiger D (1997) AtGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 94:8515–8520.

Hogbom M, Collins R, van den Berg S, Jenvert RM, Karlberg T, Kotenyova T, Flores A, Karlsson Hedestam GB, Schiavone LH (2007) Crystal structure of conserved domains 1 and 2 of the human DEAD-box helicase DDX3X in complex with the mononucleotide AMP. *J Mol Biol* 372:150–159.

Hu K, Scheer FAJL, Buijs RM, Shea SA (2008) The endogenous circadian pacemaker imparts a scale-invariant pattern of heart rate fluctuations across time scales spanning minutes to 24 hours. *J Biol Rhythms* 23:265–273

Ishiura M, Kutsuna S, Aoki S, Iwasaki H, Andersson CR, Tanabe A, Golden SS, Johnson CH, and Kondo T (1998) Expression of a gene cluster kaiABC as a circadian feedback process in *Cyanobacteria*. *Science* 281:1519–1523.

Iwasaki H, Kondo T (2000) The current state and problems of circadian clock studies in *Cyanobacteria. Plant Cell Physiol* 41 (9): 1013-1020.

Iwasaki H, Nishiwaki T, Kitayama Y, Nakajima M, and Kondo T (2002) KaiA-stimulated KaiC phosphorylation in circadian timing loops in *Cyanobacteria*. *Proc Natl Acad Sci U S A* 99:15788–15793.

Jacobs AM, Nicol SM, Hislop RG, Jaffray EG, Hay RT, Fuller-Pace FV (2007) SUMO modification of the DEAD box protein p68 modulates its transcriptional activity and promotes its interaction with HDAC1. *Oncogene* 24:543–553.

Jacobs JS, Anderson AR, Parker RP (1998) The 3 to 5degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3 to 5 exonucleases of the exosome complex. *EMBO J* 17:1497–1506.

Janknecht R (2010) Multi-talented DEAD-box proteins and potential tumor promoters: p68 RNA helicase (DDX5) and its paralog, p72 RNA helicase (DDX17). *Am J Transl Res* 2:223–234.

Jarmoskaite I, Russell R (2011) DEAD-box proteins as RNA helicases and chaperones, *Wiley Interdiscip Rev RNA* 2:135–152.

Jin Y, Allan S, Baber L, Bhattarai EK, Lamb TM, Versaw WK (2007) Rapid genetic mapping in *Neurosporacrassa*. *Fungal Genet and Biol* 44:455-465.

Johnson CH, Mori T, Xu Y (2008) A *Cyanobacterial* circadian clockwork. *Current Biology* 18: R816–R825.

Kadener S, Stoleru D, McDonald M, Nawathean P, Rosbash M (2007) Clockwork Orange is a transcriptional repressor and a new *Drosophila* circadian pacemaker component. *Genes Dev* 21:1675–86.

Kevei E, Gyula P, Hall A, Kozma-Bognar L, Kim WY, et al (2006) Forward genetic analysis of the circadian clock separates the multiple functions of ZEITLUPE. *Plant Physiol* 140:933–45.

Kim EY, Ko HW, Yu W, Hardin PE, Edery I (2007) A DOUBLETIME kinase binding domain on the *Drosophila* PERIOD protein is essential for its hyperphosphorylation, transcriptional repression, and circadian clock function. *Mol Cell Biol* 27(13):5014–28.

Kloss B, Price JL, Saez L, Blau J, Rothenfluh A,Young MW (1998) The *Drosophila* clock gene double-time encodes a protein closely related to human casein kinase Ie. *Cell* 94:97–107.

Kojima S, Matsumoto K, Hirose M, Shimada M, Nagano M, Shigeyoshi Y, Hoshino S, Ui-Tei K, Saigo K, Green CB, Sakaki Y, Tei H (2007) LARK activates posttranscriptional expression of an essential mammalian clock protein, PERIOD1. *Proc Natl Acad Sci U S A* 104(6):1859-64.

Konopka R, Benzer S (1971) Clock mutants of *Drosophila melanogaster*. *Proc Nat Acad Sci U S A* 68 (9):2112–2116.

Lakin-Thomas PL (1992) Phase resetting of the *Neurospora crassa* circadian oscillator: effects of inositol depletion on sensitivity to light. *J Biol Rhythms* 7(3):227-39.

Lakin-Thomas PL (1996) Effects of choline depletion on the circadian rhythm in *Neurospora crassa*. *Biological Rhythm Research* 27:12-30.

Lakin-Thomas PL (1998) Choline depletion, frq mutations, and temperature compensation of the circadian rhythm in *Neurospora crassa*. *J Biol Rhythms* 13:268.

Lakin-Thomas PL, Brody S (1985) Circadian rhythms in *Neurospora crassa*: interactions between clock mutations. *Genetics* 109(1):49-66.

Lakin-Thomas PL, Brody S (2000) Circadian rhythms in *Neurospora crassa*: lipid deficiencies restore robust rhythmicity to null frequency and white-collar mutants. *Proc Natl Acad Sci U S A* 97: 256–261.

Lakin-Thomas PL (2006) Transcriptional feedback oscillators: maybe, maybe not. *J Biol Rhythms* 21(2):83-92.

Levi F, Schibler U (2007) Circadian rhythms: mechanisms and therapeutic implications. *Annu Rev Pharmacol Toxicol* 47:593–628.

Li S, Lakin-Thomas P (2010) Effects of prd circadian clock mutations on FRQ-Less rhythms in *Neurospora*. *J Biol Rhythms* 25(2):71-80.

Li S, Motavaze K, Kafes E, Suntharalingam S, Lakin-Thomas P (2011) A new mutation affecting FRQ-Less rhythms in the circadian system of *Neurospora crassa*. *PLoS Genet* 7(6):e1002151.

Liang S, Hitomi M, Hu YH, Liu Y, Tartakoff AM (1996). A DEAD-box-family protein is required for nucleocytoplasmic transport of yeast mRNA. *Mol Cell Biol* 16: 5139–5146.

Lim C, Chung BY, Pitman JL et al (2007) Clockwork orange encodes a transcriptional repressor important for circadian-clock amplitude in *Drosophila*. *Curr Biol* 17:1082–9.

Lin Y, Han M, Shimada B et al (2002) Influence of the period dependent circadian clock on diurnal, circadian, and aperiodic gene expression in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 99(14):9562–7.

Linden H, Macino G (1997) White-collar-2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. *EMBO J* 16:98-109.

Linder P, Jankowsky E (2011) From unwinding to clamping — the DEAD box RNA helicase family, *Nat Rev Mol Cell Biol* 12:505–516.

Liu Y, Bell-Pedersen D (2006) Circadian Rhythms in *Neurospora crassa* and other filamentous fungi. *Eukaryotic cell* 5:1184-1193.

Lombardi L, Schneider K, Tsukamoto M, Brody S (2007) Circadian Rhythms in *Neurosporacrassa*: Clock mutant effects in the absence of a frq-based oscillator. *Genetics* 175:1175–1183.

Loros JJ and Dunlap JC (2001) Genetic and molecular analysis of circadian rhythms in *Neurospora*. *Annu Rev Physiol* 63:757-794.

Loros JJ, Feldman JF (1986) Loss of temperature compensation of circadian period length in the frq9 mutant of *Neurospora crassa*. *J Biol Rhythms* 1:187-98.

Lowrey PL, Takahashi JS (2004) Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annu Rev Genomics Hum Genet* 5:407–41.

Ma WK, Cloutier SC, Tran EJ (2013) The DEAD-box protein Dbp2 functions with the RNA binding protein Yra1 to promote mRNP assembly. *J Mol Biol* (in press).

Mallam AL, Del Campo M, Gilman B, Sidote DJ, Lambowitz AM (2012) Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p, *Nature* 490:121–125.

McClung CR (2006) Plant circadian rhythms. *Plant Cell* 18:792-803.

McDonald MJ, Rosbash M (2001) Microarray analysis and organization of circadian gene expression in *Drosophila. Cell* 107:567–78.

McWatters HG, Bastow RM, Hall A, Millar AJ (2000). The ELF3 zeitnehmer regulates light signalling to the circadian clock. *Nature* 408:716–20.

McWatters HG, Kolmos E, Hall A, Doyle MR, Amasino RM, et al (2007) ELF4 is required for oscillatory properties of the circadian clock. *Plant Physiol* 144:391–401.

Mehra A, Shi M, Baker CL, Colot HV, Loros JJ, Dunlap JC (2009) A role for casein kinase 2 in the mechanism underlying circadian temperature compensation. *Cell* 137, 749–760.

Merrow M, Brunner M, Roenneberg T (1999) Assignment of circadian function for the *Neurospora* clock gene frequency. *Nature* 399(6736):584-6.

Merrow M, Franchi L, Dragovic Z, Gorl M, Johnson J, Brunner M, Macino G, and Roenneberg T (2001) Circadian regulation of the light input pathway in *Neurospora crassa*. *EMBO J* 20:307-315.

Michael TP, McClung CR (2003) Enhancer trapping reveals widespread circadian clock transcriptional control in Arabidopsis. *Plant Physiol* 132:629–39.

Millar AJ, Carre IA, Strayer CA, Chua NH, Kay SA (1995) Circadian clock mutants in *Arabidopsis* identified by luciferase imaging. *Science* 267:1161–63.

Mohawk JA, Green CB, Takahashi JS (2012) Central and peripheral circadian clocks in mammals. *Annu Rev Neurosci* 35:445-62.

Morgan LW, Feldman JF, Bell-Pedersen D (2001) Genetic interactions between clock mutations in Neurospora crassa: can they help us to understand complexity? *Philos Trans R Soc Lond B Biol Sci* 356(1415): 1717–1724.

Nakahira Y, Katayama M, Miyashita H, Kutsuna S, Iwasaki H, Oyama T, and Kondo T (2004) Global gene repression by KaiC as a master process of prokaryotic circadian system. *Proc Natl Acad Sci U S A* 101:881-885.

Nakajima M, Imai K, Ito H, Nishiwaki T, Murayama Y, Iwasaki H, Oyama T, and Kondo T (2005) Reconstitution of circadian oscillation of *Cyanobacterial* KaiC phosphorylation in vitro. *Science* 308:414–415.

Nawathean P, Rosbash M (2004) The doubletime and CKII kinases collaborate to potentiate *Drosophila* PER transcriptional repressor activity. *Mol Cell* 13:213–23.

Newby LM, Jackson FR (1993) A new biological rhythm mutant of *Drosophila melanogaster* that identifies a gene with an essential embryonic function. *Genetics* 135(4):1077-90.

Padmanabhan K, Robles MS, Westerling T, Weitz CJ (2012) Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex. *Science* 337, 599.

Panda S, Hogenesch JB, Kay SA (2002) Circadian rhythms from flies to human. *Nature* 417(6886):329-35.

Park DH, Somers DE, Kim YS, Choy YH, Lim HK, et al (1999) Control of circadian rhythms and photoperiodic flowering by the Arabidopsis GIGANTEA gene. *Science* 285:1579–82.

Parvatiyar K, Zhang Z, Teles RM, Ouyang S, Jiang Y, Iyer SS, Zaver SA, Schenk M, Zeng S,Zhong W, Liu ZJ, Modlin RL, Liu YJ, Cheng G (2012) The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. *Nat Immunol* 13:1155–1161.

Penev PD, Kolker DE, Zee PC and Turek FW (1998) Chronic circadian desynchronization decreases the survival of animals with cardiomyopathic heart disease. *Am J Physiol* 275:H2334-7.

Perkins DD, Radford AS, Saches MS (2001) The *Neurospora* Compendium: Chromosomal Loci. *Academic Press California*.

Pregueiro AM, Liu Q, Baker CL, Dunlap JC, Loros JJ (2006) The *Neurospora*checkpoint kinase 2: a regulatory link between the circadian and cell cycles. *Science* 313: 644–649.

Price JL, Dembinska ME, Young MW, Rosbash M (1995) Suppression of PERIOD protein abundance and circadian cycling by the *Drosophila* clock mutation timeless. *EMBO J* 14: 4044–9.

Putnam AA, Jankowsky E (2013) DEAD-box helicases as integrators of RNA, nucleotide and protein binding. *Biochim.t Biophys. Acta* 1829:884–893.

Reddy AB, Wong GK, O'Neill J, Maywood ES, Hastings MH (2005) Circadian clocks: neural and peripheral pacemakers that impact upon the cell division cycle. *Mutat Res* 574, 76–91.

Richier B, Michard-Vanhee C, Lamouroux A, Papin C, Rouyer F (2008) The clockwork orange *Drosophila* protein functions as both an activator and a repressor of clock gene expression. *J Biol Rhythms* 23: 103–16.

Roenneberg T, Morse D (1993) Two circadian oscillators in one cell. *Nature* 362:362–64.

Rothenfluh A, Abodeely M, Young MW (2000) Short-period mutations of per affect a double-time-dependent step in the *Drosophila* circadian clock. *Curr Biol* 10: 1399–402.

Sai J, Johnson CH (1999) Different circadian oscillators control Ca2+ fluxes and Lhcb gene expression. *Proc Natl Acad Sci U S A* 96:11659–63.

Schernhammer ES, Kroenke, CH, Laden F, Hankinson SE (2006) Night work and risk of breast cancer. *Epidemiology* 17:108–111.

Schernhammer ES, Laden F, Speizer FE, Willett WC, Hunter DJ, Kawachi I, Colditz GA (2001) Rotating night shifts and risk of breast cancer in women participating in the nurses' health study. *J Natl Cancer Inst* 93:1563–1568.

Shi M, Collett M, Loros JJ, Dunlap JC (2010) FRQ-interacting RNA helicase mediates negative and positive feedback in the *Neurospora* circadian clock. *Genetics* 184:351–361.

Sundram V, Ng FS, Roberts MR, Millán C, Ewer J, Jackson FR (2012) Cellular requirements for LARK in the *Drosophila* circadian system. *J Biol Rhythms* 27: 183.

Takahashi JS, Hong HK, Ko CH, McDearmon EL (2008) The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat Rev Genet* 9:764–75.

Torchet C, Bousquet-Antonelli C, Milligan L, Thompson E, Kufel J, and Tollervey D (2002) Processing of 3-extended read-through transcripts by the exosome can generate functional mRNAs. *Mol Cell* 9:1285–1296.

Vitaterna MH, King DP, Chang AM, Kornhauser JM, Lowrey PL, McDonald JC, Dove WF, Pinto LH, Turek FW, Takahashi JS (1994) Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* 264(5159):719-725.

Welsh DK, Takahashi JS, Kay SA (2010) Suprachiasmatic nucleus: cell autonomy and network properties. *Annu Rev Physiol* 72:551–77.

Woelfle MA, Ouyang Y, Phanvijhitsiri K, Johnson CH (2004) The adaptive value of circadian clocks; an experimental assessment in *Cyanobacteria*. *Curr Biol* 14:1481–1486.

Woodson SA (2010) Compact intermediates in RNA folding. *Annu Rev Biophys* 39, 61–77.

Xu Y, Mori T, and Johnson CH (2000) Circadian clock-protein expression in *Cyanobacteria*: rhythms and phase setting. *EMBO J* 19:3349–3357.

Xu Y, Mori T, and Johnson CH (2003) *Cyanobacterial* circadian clockwork: roles of KaiA, KaiB and the kaiBC promoter in regulating KaiC. *EMBO J* 22:2117-2126.

Yang YH, Cheng P, Liu Y (2002) Regulation of the *Neurospora*circadian clock by casein kinase II. *Genes Dev* 16, 994–1006.

Yang Z, Sehgal A (2001) Role of molecular oscillations in generating behavioral rhythms in *Drosophila*. *Neuron* 29: 453–67.

Yi C, Mu L, de la Longrais IA, Sochirca O, Arisio R, Yu H, Hoffman AE, Zhu Y, Katsaro D (2010) The circadian gene NPAS2 is a novel prognostic biomarker for breast cancer. *Breast Cancer Res Treat* 120(3):663–669.

Yoo SH, Yamazaki S, Lowrey PL, Shimomura K, Ko CH, et al (2004) PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc Natl Acad Sci U S A* 101:5339–46.

Young MW, Kay SA (2001) Time zones: a comparative genetics of circadian clocks. *Nat Rev Genet* 2:702–15.

Yu W, Zheng H, Houl JH, Dauwalder B, Hardin PE (2006) PER-dependent rhythms in CLK phosphorylation and E-box binding regulate circadian transcription. *Genes Dev* 20: 723–33.

Zemora G, Waldsich C (2010) RNA folding in living cells. *RNA Biol* 7:634–641.

Appendix

Appendix-1

Media

Liquid minimal medium (for microtiter wells or shaker flasks):

2ml 50X Vogels

2 g D-glucose

100 ml water

Minimal agar medium (for stock tubes and babies, 3 ml per tube or 1 ml per baby):

2 ml 50X Vogels

2 g D-glucose

2 g agar

100 ml water

MA medium (maltose-argenine for race tubes, 6 ml per race tubes):

2 ml 50X Vogels

0.5 ml L-arginine, 20 mg/ml

0.5 g maltose

2 g agar

100 ml water

Growth supplements:

Choline for stock tubes: 2 mg/ml, add 1.5 ml per 100 ml = 215 µM

Histidine: 25 mg/ml, use 2 ml per 100 ml = 500 µg/ml

Hygromycin: 50 mg/ml stock, use 0.4 ml per 100 ml = 200 µg/ml

50X Vogels stock medium:

Add successively to 250 ml flask, stirring continuously:

188 ml distilled water, stir and heat

32.8 g Na_3-citrate.2H2O, let it dissolve

62.5 g KH_2PO_4 anhydrous, let dissolve

25 g NH_4NO_3 anhydrous, let dissolve

Turn off heat and add water to 250 ml.

2.5 g $MgSO_4.7H_2O$, let dissolve

1.25 g $CaCl_2.2H_2O$ dissolved in 5 ml water

1.25 ml trace elements solution

0.625 ml biotin solution

Allow to cool, add about 2 ml choloroform, cap tightly. Store at room temperature.

Trace Elements:

Add successively to 9.5 ml water, with stirring:

0.5 g citric acid.H_2O

0.5 g $ZnSO_4.7H_2O$

0.1 g Fe $(NH_4)_2(SO_4)_2.6H_2O$

0.025 g CuSO4.5H_2O

0.066 g $MnSO_4.4H_2O$

0.005 g H_3BO_3 anhydrous

0.0037 g $(NH_4)_6MO_7O_{24}.4H_2O$

10X TAE Buffer:

48.4 g Tris

3.72 g EDTA disodium salt, dihydrate

11.4 ml Acetic Acid

~ 600 ml dH_2O

Measure out approximately 600 ml water in a beaker. Add one at a time of Tris, EDTA and Acetic Acid, stirring to dissolve completely after each. Adjust pH to 8.0. Pour into a graduated cylinder and add more water to equal 1 L total.

DNA extraction Buffer:

1.21 g Tris (final concentration 100mM)

1.86 g EDTA (final concentration 50 mM)

1 ml 10% SDS

Dissolve Tris and EDTA with stirring in 80 ml dd H_2O. Adjust PH to 8.0. Add water for a total volume of 100 ml. To make DNA extraction buffer wit 1% SDS, measure 1 ml 10% SDS with 9 ml DNA extraction buffer.

TE Buffer:

0.121 g Tris (final concentration 10mM)

0.0372 g EDTA (final concentration 1 mM)

Dissolve Tris and EDTA with stirring in 80 ml dd H_2O. Adjust pH to 8.0. Add water for a final volume of 100 ml.

LB Broth:

Make 1 L, autoclave and use as required

1% NaCl = 10 g

0.5% Yeast Extract = 5 g

1% Tryptone/peptone = 10 g

Dissolve in 1L water and autoclave.

LB agar plates:

1% NaCl = 10 g

0.5% Yeast Extract = 5 g

1% Tryptone/peptone = 10 g

1.5% agar = 15 g

Dissolve frist three and add agar and autoclave. Cool it to 60^{o}C in oven then add any antibiotics (if required). Add 1 ml ampicillin (1 ml/L) for 1000 ml LB agar.

1X TSS:

Make 50 ml LB broth at 2X concentration and autoclave:

2% NaCl

1% Yeast Extract

2% Tryptone/peptone

Add the following as power or from stock

10 g Polyethylene Glycol as power

5 ml DMSO

0.5 ml of 1M stock to get a final concentration of 50 mM $MgCl_2$

Make up the volume to final 100 ml; all this is in the sterile bottle

Store at 4^{o}C overnight before using the next day.

Transformation of Neurospora:

All the solution need to be sterile (autoclaved):

1 M sorbitol

9.1 g sorbitol, dissolved in about 30 ml water, then top up volume to 50 ml

Recovery medium

0.2 ml 50x Vogels 1.82 g sorbitol 9 ml water

Choline (150 µl of 2 mg/ml) or any other supplement as needed (No histidine)

Autoclave, allow to cool, add 1ml FIGS

Plating media

91 g sorbitol, dissolved in about 350 ml water, then add water to 450 ml total

10 ml 50x Vogels

Choline (7.5 ml of 2 mg/ml) or other supplement as needed (No histidine)

14 g agar

Autoclave, cool in oven to 50-60^{o}C, add 50 ml FIGS

FIGS

10 g sorbose

0.25 g fructose

0.25 g glucose

50 ml water

Autoclave, cool to 50-60^{o}C

100 ml flask medium for growing conidia:

2 g glucose

2g agar

2 ml 50x Vogels

2 ml histidine 25 mg/ml

Plus any other supplement needed, such as choline (1.5 ml of 2 mg/ml)

Dissolve in 100 ml water and autoclave.

Microconida plating media:

X SC (synthetic crossing medium)

0.5% sucrose

2% agar

215 µM choline (if required)

For 100 ml (5 plates):

5 ml of 2X SC

0.5 g sucrose

2 g agar

1.5 ml of 2 mg/ml stock

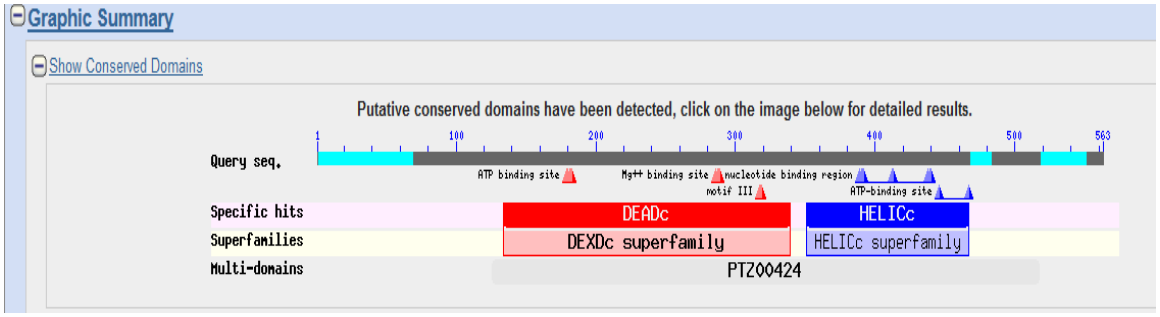
100 ml water

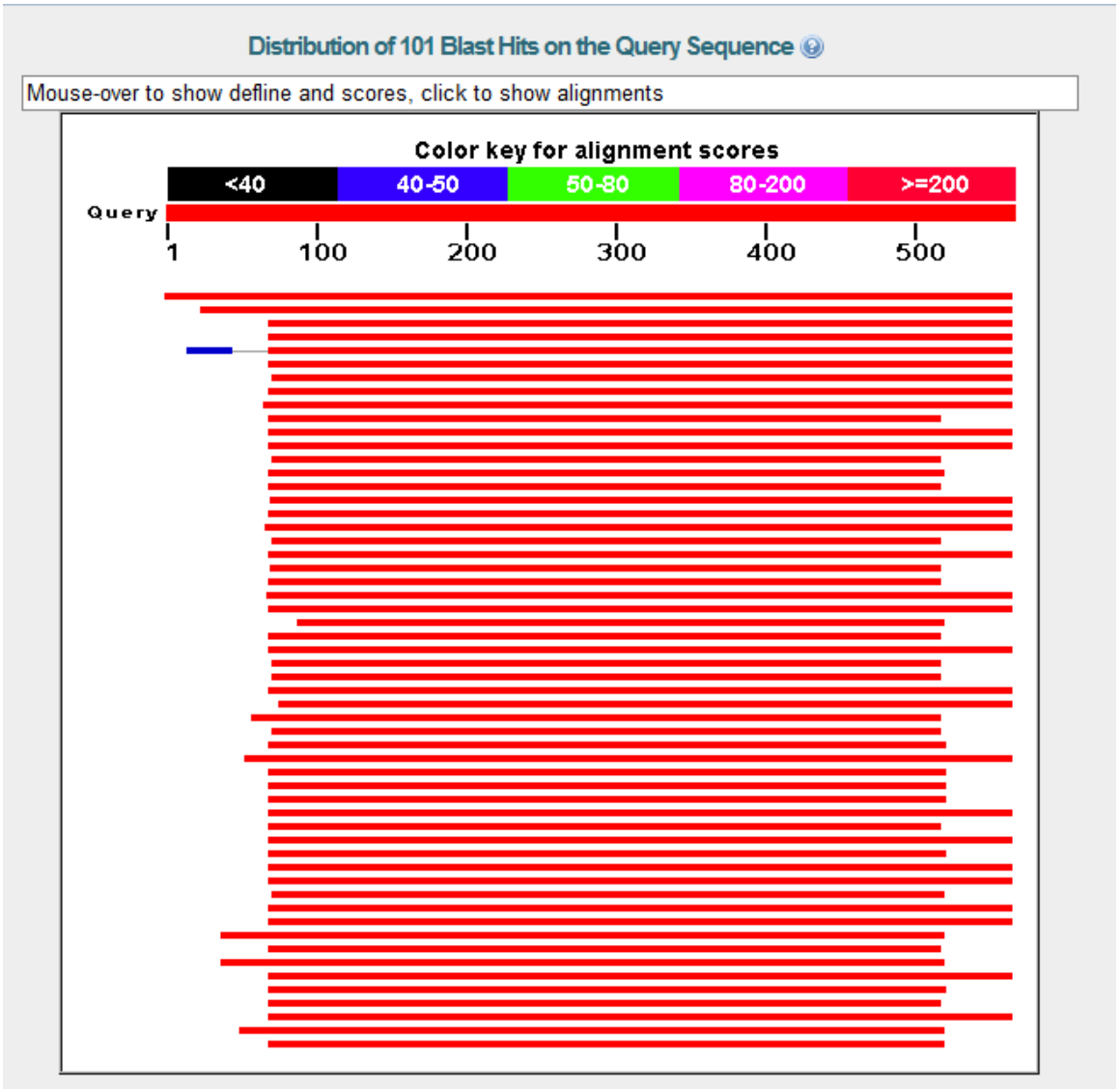
After autoclaving add 0.1 M stock filter- sterilized Na-iodoacetate.

Appendix-2

Results of BLAST search using PRD-1 protein sequence

(http://www.broadinstitute.org/annotation/genome/neurospora/GeneDetails.html?sp=S7000007580564890)





Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
RecName: Full=ATP-dependent RNA helicase dbp-2 >gb|EAA33724.3| ATP-dependent RNA helicase dbp-2 [Neurospora crassa OR74A]	1138	1138	99%	0.0	100%	Q7SBC6.2
ATP-dependent RNA helicase dbp-2 [Neurospora tetrasperma FGSC 2508] >gb|EGZ72342.1| ATP-dependent RNA helicase dbp-2 [Neurospora tetraspern	959	959	95%	0.0	98%	EGO57401.1
ATP-dependent RNA helicase dbp-2 [Neurospora crassa OR74A]	946	946	87%	0.0	100%	XP_962960.2
unnamed protein product [Sordaria macrospora k-hell]	932	932	87%	0.0	97%	CCC11459.1
hypothetical protein SMAC_02116 [Sordaria macrospora k-hell]	929	971	92%	0.0	97%	XP_003350404.1
hypothetical protein THITE_2120154 [Thielavia terrestris NRRL 8126] >gb|AEO69563.1| hypothetical protein THITE_2120154 [Thielavia terrestris NRRL 8:	889	889	87%	0.0	90%	XP_003655899.1
hypothetical protein CHGG_05545 [Chaetomium globosum CBS 148.51] >sp|Q2H720.1|DBP2_CHAGB RecName: Full=ATP-dependent RNA helicase DE	878	878	87%	0.0	88%	XP_001221640.1
hypothetical protein MYCTH_2303449 [Myceliophthora thermophila ATCC 42464] >gb|AEO57374.1| hypothetical protein MYCTH_2303449 [Myceliophthora	875	875	87%	0.0	90%	XP_003662619.1
putative atp-dependent rna helicase dbp-2 protein [Togninia minima UCRPA7]	870	870	88%	0.0	89%	EOO00620.1
ATP-dependent RNA helicase DBP2 [Magnaporthe oryzae 70-15] >sp|A4QSS5.1|DBP2_MAGO7 RecName: Full=ATP-dependent RNA helicase DBP2 >gb|	867	867	79%	0.0	92%	XP_003712846.1
RNA helicase [Grosmannia clavigera kw1407]	861	861	87%	0.0	87%	EFX02419.1
ATP-dependent RNA helicase dbp-2 [Sporothrix schenckii ATCC 58251]	861	861	87%	0.0	87%	ERT01248.1
hypothetical protein [Podospora anserina S mat+] >emb|CAP69180.1| unnamed protein product [Podospora anserina S mat+]	853	853	78%	0.0	91%	XP_001908507.1
ATP-dependent RNA helicase dbp-2 [Metarhizium anisopliae ARSEF 23]	854	854	79%	0.0	91%	EFZ02905.1
ATP-dependent RNA helicase DBP2 [Gaeumannomyces graminis var. tritici R3-111a-1]	854	854	79%	0.0	90%	EJT81100.1
RNA helicase [Colletotrichum orbiculare MAFF 240422]	852	852	87%	0.0	88%	ENH86476.1
ATP-dependent RNA helicase dbp-2 [Metarhizium acridum CQMa 102]	851	851	87%	0.0	87%	EFY92383.1
DEAD/DEAH box helicase [Beauveria bassiana ARSEF 2860]	848	848	88%	0.0	86%	EJP67186.1

DEAD/DEAH box helicase [Colletotrichum graminicola M1.001]	847	847	78%	0.0	91%	EFQ25261.1
predicted protein [Trichoderma reesei QM6a]	846	846	87%	0.0	86%	EGR50376.1
probable RNA helicase dbp2 (DEAD box protein) [Claviceps purpurea 20.1]	843	843	79%	0.0	89%	CCE27540.1
phosphomethylpyrimidine kinase [Magnaporthe oryzae Y34] >gb|ELQ64783.1| phosphomethylpyrimidine kinase [Magnaporthe oryzae P131]	860	860	79%	0.0	88%	ELQ37019.1
hypothetical protein TRIVIDRAFT_83250 [Trichoderma virens Gv29-8]	839	839	87%	0.0	85%	EHK19911.1
ATP-dependent RNA helicase dbp-2 [Cordyceps militaris CM01]	839	839	87%	0.0	85%	EGX87816.1
hypothetical protein CGLO_06986 [Colletotrichum gloeosporioides Cg-14]	829	829	76%	0.0	92%	EQB53300.1
hypothetical protein TRIATDRAFT_291734 [Trichoderma atroviride IMI 206040]	832	832	79%	0.0	89%	EHK46590.1
RNA helicase (Dbp), putative [Talaromyces marneffei ATCC 18224] >gb|EEA20986.1| RNA helicase (Dbp), putative [Talaromyces marneffei ATCC 18224]	831	831	87%	0.0	85%	XP_002151986.1
ATP-dependent RNA helicase DBP2 [Fusarium oxysporum f. sp. cubense race 4]	831	831	78%	0.0	89%	EMT65942.1
ATP-dependent RNA helicase DBP2 [Fusarium oxysporum f. sp. cubense race 1]	831	831	78%	0.0	89%	ENH75111.1
RNA helicase (Dbp), putative [Talaromyces stipitatus ATCC 10500] >gb|EED20566.1| RNA helicase (Dbp), putative [Talaromyces stipitatus ATCC 10500]	830	830	87%	0.0	85%	XP_002481000.1
predicted protein [Nectria haematococca mpVI 77-13-4] >gb|EEU39106.1| predicted protein [Nectria haematococca mpVI 77-13-4]	828	828	86%	0.0	84%	XP_003044819.1
hypothetical protein FPSE_00152 [Fusarium pseudograminearum CS3096]	828	828	81%	0.0	88%	EKJ79698.1
probable RNA helicase dbp2 (DEAD box protein) [Fusarium fujikuroi IMI 58289]	827	827	78%	0.0	88%	CCT73617.1
ATP-dependent RNA helicase of the DEAD-box protein family [Aspergillus niger ATCC 1015]	826	826	79%	0.0	88%	EHA22725.1
hypothetical protein FG04132.1 [Fusarium graminearum PH-1] >sp|Q4IF76.1|DBP2_GIBZE RecName: Full=ATP-dependent RNA helicase DBP2	825	825	90%	0.0	84%	XP_384308.1
ATP-dependent RNA helicase dbp2 [Aspergillus niger CBS 513.88]	825	825	79%	0.0	88%	XP_001399394.2
RecName: Full=ATP-dependent RNA helicase dbp2	824	824	79%	0.0	87%	Q4X195.2
RecName: Full=ATP-dependent RNA helicase dbp2 >emb|CAK37534.1| unnamed protein product [Aspergillus niger]	824	824	79%	0.0	88%	A2QC74.1
hypothetical protein HCAG_08709 [Ajellomyces capsulatus NAm1] >sp|A6RGE3.1|DBP2_AJECN RecName: Full=ATP-dependent RNA helicase DBP2 >gl	823	823	87%	0.0	82%	XP_001536388.1
P-loop containing nucleoside triphosphate hydrolase [Glarea lozoyensis ATCC 20868]	824	824	79%	0.0	88%	EPE34606.1
ATP-dependent RNA helicase DBP2 [Ajellomyces dermatitidis ER-3] >gb|EGE86011.1| ATP-dependent RNA helicase DBP2 [Ajellomyces dermatitidis ATC	823	823	87%	0.0	82%	EEQ92003.1
ATP-dependent RNA helicase DBP2 [Ajellomyces dermatitidis SLH14081] >gb|EEQ76010.1| ATP-dependent RNA helicase DBP2 [Ajellomyces dermatitidi	821	821	87%	0.0	82%	XP_002620418.1
RNA helicase (Dbp), putative [Neosartorya fischeri NRRL 181] >sp|A1DGZ7.1|DBP2_NEOFI RecName: Full=ATP-dependent RNA helicase dbp2 >gb|EAW	820	820	87%	0.0	84%	XP_001260551.1
ATP-dependent RNA helicase DBP2 [Verticillium dahliae VdLs.17]	820	820	79%	0.0	83%	EGY14525.1
RNA helicase (Dbp), putative [Aspergillus clavatus NRRL 1] >sp|A1C6C4.1|DBP2_ASPCL RecName: Full=ATP-dependent RNA helicase dbp2 >gb|EAW1:	818	818	87%	0.0	84%	XP_001275371.1
ATP-dependent RNA helicase DBP2 [Ajellomyces capsulatus H88]	817	817	87%	0.0	82%	EGC45307.1
ATP-dependent RNA helicase DBP2, putative [Coccidioides posadasii C735 delta SOWgp] >sp|Q1DP69.2|DBP2_COCIM RecName: Full=ATP-dependent	817	817	85%	0.0	83%	XP_003067830.1
ATP-dependent RNA helicase dbp2 [Coniosporium apollinis CBS 100218]	819	819	79%	0.0	88%	EQN64609.1
conserved hypothetical protein [Coccidioides immitis RS]	817	817	85%	0.0	83%	XP_001240731.1
ATP-dependent RNA helicase DBP2 [Ajellomyces capsulatus G186AR]	817	817	87%	0.0	82%	EEH03974.1
ATP-dependent RNA helicase dbp2 [Exophiala dermatitidis NIH/UT8656]	816	816	79%	0.0	86%	EHY60011.1
ATP-dependent RNA helicase dbp2 [Pseudogymnoascus destructans 20631-21]	816	816	79%	0.0	87%	ELR06415.1
ATP-dependent RNA helicase DBP2 [Paracoccidioides sp. 'lutzii' Pb01] >gb|EEH35437.1| ATP-dependent RNA helicase DBP2 [Paracoccidioides sp. 'lutzii	816	816	87%	0.0	82%	XP_002791745.1
RNA helicase [Colletotrichum gloeosporioides Nara gc5]	815	815	82%	0.0	87%	ELA34444.1
conserved hypothetical protein [Sclerotinia sclerotiorum 1980] >sp|A7E449.1|DBP2_SCLS1 RecName: Full=ATP-dependent RNA helicase dbp2 >gb|EDN	817	817	79%	0.0	87%	XP_001597985.1
putative atp-dependent rna helicase dbp2 protein [Neofusicoccum parvum UCRNP2]	815	815	79%	0.0	87%	EOD49085.1
putative ATP-dependent RNA helicase dbp2 [Marssonina brunnea f. sp. 'multigermtubi' MB_m1]	816	816	79%	0.0	86%	EKD14874.1
RecName: Full=ATP-dependent RNA helicase dbp2	813	813	79%	0.0	87%	A6SFW7.2
hypothetical protein AN5931.2 [Aspergillus nidulans FGSC A4] >sp|Q5B0J9.1|DBP2_EMENI RecName: Full=ATP-dependent RNA helicase dbp2 >gb|EAA4	813	813	87%	0.0	82%	XP_663535.1
RNA helicase (Dbp) [Aspergillus fumigatus Af293] >gb|EAL93370.1| RNA helicase (Dbp), putative [Aspergillus fumigatus Af293] >gb|EDP54594.1| RNA he	813	813	79%	0.0	84%	XP_755408.1
hypothetical protein BofuT4_P016160.1 [Botryotinia fuckeliana T4]	813	813	79%	0.0	87%	CCD51310.1
putative atp-dependent rna helicase dbp2 protein [Botryotinia fuckeliana BcDW1]	813	813	79%	0.0	87%	EMR82333.1
p68 RNA helicase [Botryotinia fuckeliana B05.10]	808	808	78%	0.0	87%	XP_001549899.1

hypothetical protein DOTSEDRAFT_73654 [Dothistroma septosporum NZE10]	810	810	79%	0.0	87%	EME41324.1
putative ATP-dependent RNA helicase dbp2 [Blumeria graminis f. sp. hordei DH14]	811	811	79%	0.0	85%	CCU76732.1
hypothetical protein PDE_03848 [Penicillium oxalicum 114-2]	810	810	79%	0.0	86%	EPS28902.1
ATP-dependent RNA helicase-like protein [Chaetomium thermophilum var. thermophilum DSM 1495]	812	812	87%	0.0	76%	EGS17787.1
Pc22g21030 [Penicillium chrysogenum Wisconsin 54-1255] >emb|CAP99391.1| Pc22g21030 [Penicillium chrysogenum Wisconsin 54-1255]	808	808	79%	0.0	85%	XP_002566000.1
RNA helicase (Dbp), putative [Penicillium digitatum Pd1] >gb|EKV14870.1| RNA helicase (Dbp), putative [Penicillium digitatum PHI26]	808	808	87%	0.0	82%	EKV09591.1
ATP-dependent RNA helicase DBP2 [Trichophyton rubrum CBS 118892] >gb|EGD86827.1| ATP-dependent RNA helicase DBP2 [Trichophyton rubrum CBS	803	803	90%	0.0	79%	XP_003236032.1
ATP-dependent RNA helicase DBP2 [Endocarpon pusillum Z07020]	801	801	90%	0.0	79%	ERF69514.1
hypothetical protein PTT_11286 [Pyrenophora teres f. teres 0-1] >gb|EFQ91808.1| hypothetical protein PTT_11286 [Pyrenophora teres f. teres 0-1]	798	798	79%	0.0	85%	XP_003300130.1
ATP-dependent RNA helicase DBP2 [Arthroderma gypseum CBS 118893] >gb|EFR05021.1| ATP-dependent RNA helicase DBP2 [Arthroderma gypseum C	800	800	79%	0.0	84%	XP_003169856.1
RNA helicase (Dbp), putative [Aspergillus flavus NRRL3357] >ref|XP_001826178.2| ATP-dependent RNA helicase dbp2 [Aspergillus oryzae RIB40] >gb|EE	800	800	88%	0.0	83%	XP_002377846.1
RecName: Full=ATP-dependent RNA helicase dbp2 >dbj|BAE65045.1| unnamed protein product [Aspergillus oryzae RIB40]	800	800	88%	0.0	83%	Q2U070.1
hypothetical protein FOXB_00888 [Fusarium oxysporum Fo5176]	815	815	78%	0.0	83%	EGU88639.1
ATP-dependent RNA helicase DBP2 [Trichophyton equinum CBS 127.97]	796	796	79%	0.0	84%	EGE03831.1
hypothetical protein BAUCODRAFT_95810 [Baudoinia compniacensis UAMH 10762]	790	790	78%	0.0	86%	EMC92651.1
hypothetical protein MYCFIDRAFT_58038 [Pseudocercospora fijiensis CIRAD86]	789	789	79%	0.0	85%	EME88749.1
P-loop containing nucleoside triphosphate hydrolase protein [Sphaerulina musiva SO2202]	788	788	79%	0.0	84%	EMF11116.1
ATP-dependent RNA helicase dbp2 [Pyrenophora tritici-repentis Pt-1C-BFP] >gb|EDU51509.1| ATP-dependent RNA helicase dbp2 [Pyrenophora tritici-rep	789	789	79%	0.0	81%	XP_001938922.1
hypothetical protein MYCGRDRAFT_36977 [Zymoseptoria tritici IPO323] >gb|EGP89753.1| hypothetical protein MYCGRDRAFT_36977 [Zymoseptoria tritici	784	784	79%	0.0	84%	XP_003854777.1
hypothetical protein COCHEDRAFT_1141168 [Bipolaris maydis C5]	784	784	81%	0.0	85%	EMD89203.1
Similar to ATP-dependent RNA helicase dbp2; acc. no. A1DGZ7 [Pyrenema omphalodes CBS 100304]	783	783	79%	0.0	83%	CCX12941.1
putative ATP-dependent RNA helicase dbp2 [Glarea lozoyensis 74030]	778	778	79%	0.0	77%	EHL00001.1

cytosolic regulator pianissimo [Ophiostoma piceae UAMH 11346]	827	827	84%	0.0	86%	EPE10997.1
ATP-dependent RNA helicase dbp-2 [Ophiocordyceps sinensis CO18]	763	763	70%	0.0	92%	EQK99950.1
ATP-dependent RNA helicase DBP2 [Verticillium alfalfae VaMs.102] >gb|EEY20525.1| ATP-dependent RNA helicase DBP2 [Verticillium alfalfae VaMs.102]	764	764	79%	0.0	79%	XP_003003073.1
ATP-dependent RNA helicase p62 [Uncinocarpus reesii 1704] >gb|EEP77362.1| ATP-dependent RNA helicase p62 [Uncinocarpus reesii 1704]	787	787	96%	0.0	75%	XP_002542695.1
hypothetical protein [Tuber melanosporum Mel28] >emb|CAZ80056.1| unnamed protein product [Tuber melanosporum]	755	755	79%	0.0	82%	XP_002835899.1
hypothetical protein H072_9936 [Dactylellina haptotyla CBS 200.50]	753	753	86%	0.0	76%	EPS36514.1
hypothetical protein LEMA_P074610.1 [Leptosphaeria maculans JN3] >emb|CBX99872.1| hypothetical protein LEMA_P074610.1 [Leptosphaeria maculan	776	776	79%	0.0	82%	XP_003843351.1
hypothetical protein SETTUDRAFT_156478 [Setosphaeria turcica Et28A]	766	766	88%	0.0	79%	EOA82807.1
hypothetical protein COCSADRAFT_174289 [Bipolaris sorokiniana ND90Pr]	763	763	79%	0.0	83%	EMD60965.1
hypothetical protein COCC4DRAFT_60914 [Bipolaris maydis ATCC 48331]	761	761	88%	0.0	78%	ENI05079.1
ATP-dependent RNA helicase DBP2 [Paracoccidioides brasiliensis Pb18]	735	735	71%	0.0	87%	EEH45234.1
ATP-dependent RNA helicase Dbp2 [Schizosaccharomyces pombe 972h-] >sp|P24782.2|DBP2_SCHPO RecName: Full=ATP-dependent RNA helicase d	731	731	99%	0.0	67%	NP_596523.1
hypothetical protein AOL_s00043g11 [Arthrobotrys oligospora ATCC 24927]	727	727	84%	0.0	75%	EGX52517.1
p68 protein [Schizosaccharomyces pombe]	728	728	99%	0.0	67%	CAA36873.1