ELUCIDATION OF R GENE MEDIATED YELLOW RUST DISEASE RESISTANCE MECHANISM IN WHEAT BY DUAL BAIT YEAST TWO-HYBRID ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

FİGEN YILDIRIM

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

AUGUST 2005

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

Assoc. Prof. Dr. Dilek Sanin
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Gülay Özcengiz
Co-Supervisor

Prof. Dr. Mahinur S. Akkaya
Supervisor

Examining Committee Members

Prof. Dr. Ufuk Bakır (METU, CHE)

Prof. Dr. Mahinur S. Akkaya (METU, CHEM)

Prof. Dr. Zümrüt B. Ögel (METU, FDE)

Asst. Prof.Dr. M.Cengiz Yakıcıer (Bilkent Unv., MBG)

Asst. Prof. Dr. Şenay Vural Korkut (Yıldız Tech. Unv., BIOL)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Figen YILDIRIM

Signature :

ABSTRACT

ELUCIDATION OF R GENE MEDIATED YELLOW RUST DISEASE RESISTANCE MECHANISM IN WHEAT BY DUAL BAIT YEAST TWO-HYBRID ANALYSIS

Yıldırım, Figen

Ph. D., Department of Biotechnology

Supervisor : Prof. Dr.Mahinur S. Akkaya

Co-Supervisor: Prof. Dr. Gülay Özcengiz

August 2005, 272 pages

Yellow rust, caused by *Puccinia striiformis* Westend. f. sp. *tritici* Eriksson is one of the most severe leaf diseases of wheat. Aim of this study is to illuminate the downstream signaling pathways upon incompetible infection of rust pathogen in wheat, thus to understand the genes involved in resistance mechanism. The strategy used is the dual bait yeast two-hybrid analysis which is the most powerful method for *in vivo* detection of protein-protein interactions. The bait proteins used are; the domains of *Yr10* yellow rust resistance gene, *Rad6* gene which is considered to have a critical role in *R* gene mediated signaling pathway, and WR5 gene fragment which is an unknown

protein having homology to the WD40 repeat containing protein with apoptosis related activity.

Screening of a yeast prey library with these baits revealed proteins having mostly apoptosis related functions (SRP72, POR1, CSE1), translation initiation control in response to stress conditions (Gcn2p, Eap1p), phosphorylation (SKY1) and dephosphorylation activities (GAC1), cell cycle control (FAR1), oxidative stress control (OXR1), protein degradation control (TOM1), protein folding control (CPR7) and ion homeostasis in the cell (POR1, GAC1). The significance of the study can be summarized as i) being the first yeast two hybrid analysis of a wheat *R* gene, ii) being able to detect interacting partners with anticipated functions, iii) most importantly, initiating further detailed analysis of the key interactors.

Keywords: Yeast two-hybrid analysis, Yr10, Rad6, disease resistance, wheat yellow rust

ÖZ

BUĞDAYDA R GENİ İLE KONTROL EDİLEN SARI PAS HASTALIĞI DİRENÇLİLİK MEKANİZMASININ MAYA İKİLİ HİBRİD YÖNTEMİ İLE AYDINLATILMASI

Yıldırım, Figen
Doktora, Biyoteknoloji Bölümü Bölümü
Tez Yöneticisi : Prof. Dr. Mahinur S. Akkaya
Ortak Tez Yöneticisi: Prof. Dr. Gülay Özcengiz

Ağustos 2005, 272 sayfa

Puccinia striiformis Westend. f. sp. *tritici* Eriksson tarafından oluşturulan sarı pas, bugdayın en ölümcül yaprak hastalıklarından biridir. Bu çalışmanın amacı buğdayda pas enfeksiyonuna karşı oluşan dirençlilik alt sinyal mekanizmasının aydınlatılması, böylece dirençlilik mekanizmasında rol alan genlerin anlaşılmasıdır. Çalışmada kullanılan, maya ikili-hibrid analizi hücre içinde protein-protein etkileşimlerini saptamak için en kuvvetli yöntemlerden biridir. Çalışmada yem olarak kullanılan proteinler; i) *Yr10* sarı pas dirençlilik geninin parçaları, ii) *R* geni ile kontrol edilen sinyal mekanizmasında kritik rolü olduğu düşünülen *Rad6* geni, iii)

programlanmış hücre ölümü ile ilişkili aktivitesi bulunan WD40 tekrarı içeren proteinler ile homoloji gösteren ve henüz fonksiyonu bilinmeyen WR5 gen parçasıdır.

Bu yem proteinleri kullanılarak maya av kütüphanesinin taranması sonucunda; programlı hücre ölümü ile ilişkili fonksiyonları bulunan proteinler (SRP72, POR1, CSE1), stres koşullarında translasyon başlangıcını kontrol eden proteinler (Gcn2p, Eap1p), fosforilasyon (SKY1) ve defosforilasyon (GAC1) aktivitesi gösteren proteinler, hücre döngüsü kontrolü ile ilişkili protein (FAR1), oksidatif stres kontrolü ile ilişkili protein (OXR1), protein parçalanmasını kontrol eden protein (TOM1), protein katlanması ile ilgili protein (CPR7) ve hücrede iyon dengesi kontrolü ile ilişkili proteinler (POR1, GAC1) saptanmıştır. Yapılan çalışmanın önemi i) buğday R geni ile yapılan ilk maya ikili hybrid analizi olması, ii) tahmin edilen fonksiyonlarda etkileşimlerin saptanması, iii) en önemlisi anahtar etkileşimlerin detaylı olarak incelenmesi için yeni çalışmalara olanak sağlaması, olarak özetlenebilir.

Anahtar Kelimeler: Maya ikili hybrid analizi, Yr10, Rad6, hastalık direçliliği, buğdayda sarı pas.

To My Father

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my supervisor *Prof. Dr. Mahinur S. Akkaya* for allowing me to take this extremely exciting research for my Ph.D. degree under her admirable supervision; for her kindly friendship, all kinds of guidance, support and training throughout the research.

I gratefully acknowledge the kind supports and helps of *Dr. Ilya Serebriiskii* and *Dr. Erica Golemis* of Fox Chase Cancer Center, USA; for inviting me to their lab and supplying an environment which considerably extended my experience in the yeast two-hybrid analysis work.

The financial supports of the following institutions are acknowledged:

BAP (Middle East Technical University Research Fund), BAP-2002-07-02-00-30.

DPT (Turkish State Organization), DPT-2004K120750

I offer sincere thanks to *Prof. Dr. Gülay Özcengiz* for being my co-supervisor.

I owe thanks to *my parents* for their care, encouragement and support throughout my life.

I am indebted deeply to *Altuğ Bodur* who always believed in and supported me.

I am grateful to my best friends *Mehtap S. Toklu* and *Gözde Ataç* for their supports and helps throughout my life.

I owe thanks to all my friends and colleagues in the lab for their contribution and creation of an enjoyable environment.

I would like to thank *Elif Uz*, *Osman Bozkurt*, *Özge Gürsoy*, and *Dr. Aslıhan Günel* for their helps and supports during my studies.

I owe thanks to all those gave me support and encouragement during my study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

A_{600nm}	: Absorbance at 600nm
AD	: Activation Domain
ARC	: Apaf-1, R protein and CED4 homology
Avr	: Avirulance
bp	: base pair
CC	: Coiled Coil
CM	: Complete Minimal
CSN	: COP9 signalosome
DBD	: DNA Binding Domain
dH2O	: distilled water
DTT	: Dithiothreitol
Eds1	: Enhanced disease susceptibility1
Gal-Raff	: Galactose-Raffinose
Glu	: Glucose
His	: Histidine
HR	: Hypersensitive Response
HSP90	: Heat Shock Protein 90
kDa	: kilo Dalton
LB	: Luria Broth
Leu	: Leucine
LLR	: Leucine Rich Repeat
Lys	: Lysine
mg	: mili gram
min	: minute

mM	: Milimolar
NBS	: Nucleotide Binding Site
NCBI	: National Center for Biotechnology Information
NDR1	: Nonrace-specific Disease Resistance1
NEB	: New England Biolabs
ng	: Nanogram
nm	: nanometer
NPR1/NIM1	: Nonexpressor of PR genes1
ORF	: Open Reading Frame
PAD4	: Phytoalexin Deficient4
PCR	: Polymerase Chain Reaction
pmol	: Pico mole
PR	: Pathogenesis Related
R	: Resistance
Rar1	: Required for mla-dependent resistance
ROI	: Reactive Oxygen Intermediates
Rpm	: Rotation per minute
SGT1	: Supressor of G2 allele of SKP1
SKP1	: S-phase Kinase associated Protein
TA	: Transcriptional Activator
TIR	: Toll and Interleukin-1 Receptor
TM	: Transmembrane
Trp	: Tryptophane
U	: Unit
Ura	: Uracil
v/v	: volume/volume
v/w	: volume/ weight
X-Gal	: 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

X-Gluc	: 5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid
YPD	: Yeast extract/Peptone/Dextrose
YR	: Yellow Rust

CHAPTER 1

INTRODUCTION

1.1 Wheat

"Wheat is the most valuable single crop in the modern world" (Diamond, 1997).

Turkey has a unique importance in terms of wheat. Anatolia is considered to be the center of origin and diversity of wheat (Heun *et al*., 1997 and Gökgöl, 1935) possessing potential sources of disease resistance genes. Wheat production is vitally important in Turkey, and is the number one country in bread wheat consumption per capita. Turkey is among the highest wheat producers in the world and one of the countries with the largest land (%42 of the arable land) dedicated for wheat cultivation (FAO statistical databases, 2004).

Plants are attacked by many disease causing organisms like; bacteria, fungi, viruses, insects and nematodes. There are more than 40 diseases and insects that can cause economic losses to the wheat crop under various circumstances. However, one of the greatest danger is the rust epidemics that can be formed as a result of the airborne pathogens; such as *Puccinia graminis tritici*, *Puccinia recondita*, and *Puccinia striiformis*, that cause stem (black) rust, leaf (brown) rust, and stripe (yellow) rust disease of wheat, respectively. Epidemic of wheat rust can occur on a continental scale because of the extensive spreading of uredinospores (Roelfs, 1989).

1.1 Yellow rust disease of wheat

Stripe rust, caused by *Puccinia striiformis* Westend. f. sp. *tritici* Eriksson, is one of the most serious leaf diseases of wheat in the world. In an infected field, the reduction in yield can reach 90% (when environmental conditions are suitable) due to the shrunken grains and reduced spikes caused by the disease (Roelfs *et al*., 1992). The yellow rust disease may develop when free moisture (rain or dew) occurs; minimum, optimum and maximum temperatures for stripe rust infection are 0°, 11° and 23°C, respectively (Hogg *et al*., 1969) (Table 1.1). Stripe rust can attack wheat, barley and many other related grasses. The disease is found in all highland and/or temperate areas where cereals are grown (Stubbs, 1988). *Puccinia striiformis* does not have a sexual stage and does not require an alternate host to complete its life cycle.

Table 1.1 Stripe rust development conditions in nature (Roelfs *et al*., 1992)

Stripe rust	Temperature (°C)			Light	Free water
Stage	Minimum	Optimum	Maximum		
Germination	0	9-13	23	Low	Essential
Germling	-	10-15	-	Low	Essential
Appressorium	-	-	(not formed)	None	Essential
Penetration	2	9-13	23	Low	Essential
Growth	3	12-15	20	High	None
Sporulation	5	12-15	20	High	None

Several resistance genes have been defined, but most of them only protect wheat against specific races of the rust fungus. When wheat varieties carrying such resistance genes are grown over large areas, rapid evolution to virulence occurs in

pathogen populations. Although the development in breeding and research by the utilization of these genes is clear, specific resistances are often overcome by rapid adaptation of the pathogen (Bayles and Stigwood, 1994, 1995, 1996). As a result, few naturally occurring resistance genes remain useful to breeders.

Yellow rust is the most environmentally sensitive of the rusts (Robbelen and Sharp, 1978), a feature that makes evaluating host plant resistance difficult. Moreover, resistance genes also appear to be sensitive to the environment. Recessive and dominant monogenic resistance, and resistance controlled by minor genes have been reported (Stubbs, 1985).

1.1.2 Yield loss in wheat due to yellow rust disease in Turkey

Yellow rust disease is damaging mainly in winter and spring when the whether conditions are suitable for widespread infection. The coastal, Middle Anatolia and the passage regions have mild or semi-severe winters and cold rainy springs. These conditions are optimum for the infection and the development of yellow rust. Thus, most of the time disease is present, but when the conditions are optimum the enormous epidemics occur. In the last 25 years, the yellow rust has been a reason of many epidemics in Turkey and the neighboring countries (Braun and Saari, 1992). The most serious epidemics were observed in the years of 1975, 1976, 1977, and 1984. In 1991, where Gerek-79 cultivar grew more than 1.5 ha, the yield loss was determined to be 26.5 %. In 1995, the high level of yellow rust epidemic on 'Seri-82' resulted in half a billion ton of yield loss. In years 1996 and 1997 due to losses in the west and Mediterranean region the planting of the high quality 'Seri-82' was drastically diminished. More recently, in 1998, the infection was very high in the fields sown to

'Gerek-79' with 50 % yield loss. Yield loss over the years on the average is about 200 million US dollars per year at the farmers' level over 20 years.

Chemical controls have been used effectively in some areas against rust diseases, but the basic control mechanism of the cereal rusts can be through the use of durable resistant cultivars. There are many disadvantages of fungicide application such as; cost of chemicals, limited shelf-life, known or unknown environmental hazards, occurrence of pathogens developing resistance to applied chemicals and inadequate control on susceptible cultivars. Growing resistant cultivars reduces the cost and there is no requirement of action by the farmer after cultivar's selection, and no environmental impact. Therefore, developing and planting resistant cultivars are the most economical method of controlling the disease.

1.2 Disease Resistance

1.2.1'Gene-for-gene' disease resistance

The 'gene-for-gene' disease resistance was first identified by the studies of flax and it's interaction with flax rust (Flor, 1971). The 'gene-for-gene' interaction states that for every dominant resistance (*R*) gene in the plant, there is a matching dominant avirulence (*Avr)* gene in the pathogen (Figure 1.1). Like animals, plants have the ability to distinguish potential pathogens. Plants have evolved resistance (*R*) genes that specifically recognize corresponding pathogen avirulance (*avr*) genes to start plant defense mechanism known as 'gene-for-gene resistance' (Dangl and Jones, 2001). There is little information about the molecular mechanisms controlling R–Avr recognition or how R proteins activate disease responses.

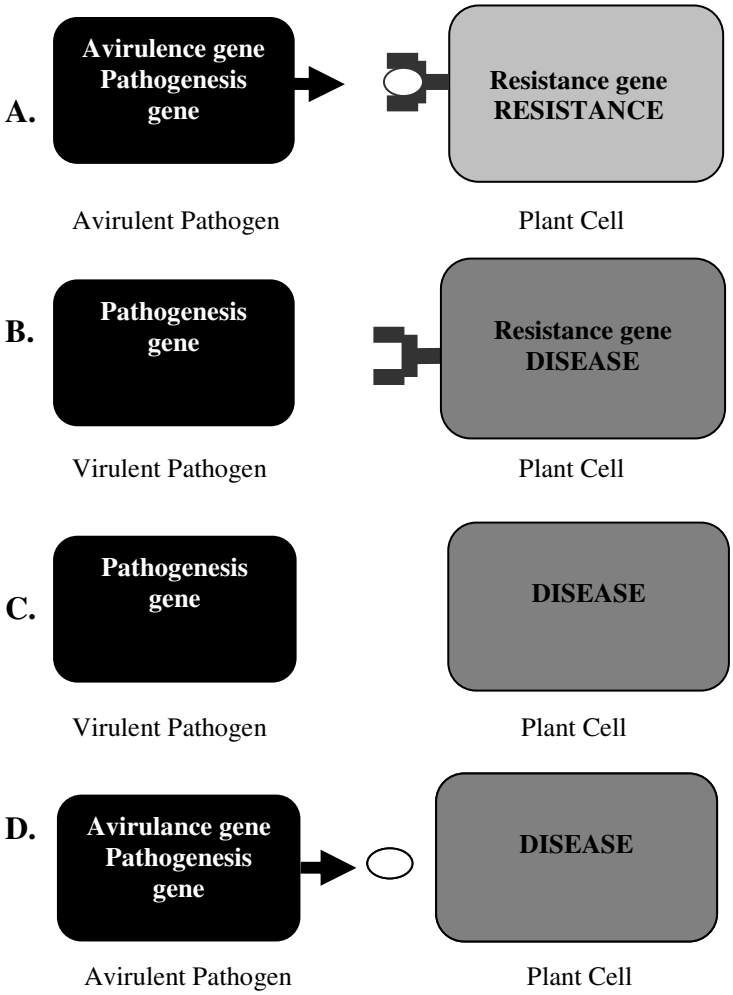


Figure 1.1 A schematic diagram explaining the 'gene-for-gene' interactions between the host and the pathogen. **A:** the pathogen possesses the *avr* (avirulence) gene and the host (plant) possesses the corresponding *R* (resistance) *gene* resulting in disease resistance (incompatible) **B:** host carries the *R* gene but the pathogen lacks the *avr* gene that the *R* gene can recognize so interaction results with disease **C:** both the pathogen and the host lack the matching *Avr* and *R* genes respectively which lead to disease. **D:** the pathogen produces the Avr signal but there is not a matching R product to receive the signal produced by the Avr so interaction results in disease (compatible)

1.2.2 Interactions of R and Avr proteins

How *R* genes recognize pathogen Avr products is still an unsolved question. Direct interaction of R protein and Avr protein is shown in the rice Pi-ta and AVR-Pita (*Magnaporthe grisea*, IB-49 and IC-17) (Jia *et al*., 2000), tomato Pto with AvrPto (*Pseudomonas syringe,* pathovar *tomato*) or AvrPtoB (*Pseudomonas syringae* B) (Kim, 2002; Scofield, 1996 and Tang, 1996) and *Arabidopsis thaliana* RPS2 with AvrRpt2 (*Pseudomonas syringae* expressing AvrRpt2) and fascinatingly with the non-associated effector AvrB (*Pseudomonas syringae* B) (Leister and Katagiri, 2000). However, experimental evidence of direct physical interactions between R proteins and Avr proteins is limited. This lack of direct interaction raised 'the guard hypothesis' (Bogdanove, 2002) which states that there can be a third protein (or more) that is responsible for the recognition of Avr protein by R gene product.

For the interaction between the R proteins and the corresponding Avr proteins a number of models have been proposed (Figure 1.2) (reviewed by Martin *et al*., 2003):

Receptor-ligand model: This model is the 'gene-for-gene' interaction model. The R protein directly interacts with the Avr effector protein and activates the downstream signaling pathways for resistance.

Bridge model: The pathogen effector binds to R protein and a third protein joining one to the other. Examples of the third proteins required for resistance are Prf (tomato, Pto mediated resistance to *Pseudomonas syringae*), PBS1 (*Arabidopsis thaliana*, RPS5 mediated resistance to *Pseudomonas syringae*), TIP (*Arabidopsis thaliana*, HRT-mediated resistance to *Turnip crinkle virus*), RIN4 (*Arabidopsis thaliana*, RPM1 mediated disease resistance to *Pseudomonas syringae*) and Rcr3 (tomato, Cf2 mediated disease resistance to *Cladosporium fulvum* (Rathjen *et al*., 1999;

Swiderski and Innes, 2001; Ren *et al*., 2000; Mackey *et al*., 2002; and Kruger *et al*., 2002).

Matchmaker: The pathogen effector induces a direct interaction between the R protein and a third protein by causing a conformational change in one or the other, or both. By this conformational change resistance mechanism is induced.

Affinity enhancement: The interaction of the pathogen effector with the R protein, a third protein, or both, stabilizes an existing, weak interaction between the two plant proteins, such that the abundance of the complex increases and drives downstream signaling to activate the induced defense response.

Depression: The pathogen effector depresses the defense response by disturbing an interaction of the R protein and a third protein that negatively regulates activity of the R protein.

Dual recognition: The independent interactions between the pathogen effector and the *R* gene and the effector and a third protein are both required for resistance function is a possibility, strengthened by the dual requirement of Pto and Prf and the interaction of AvrPto and Pto (tomato, *Pseudomonas syringae*) (Rathjen *et al*., 1999).

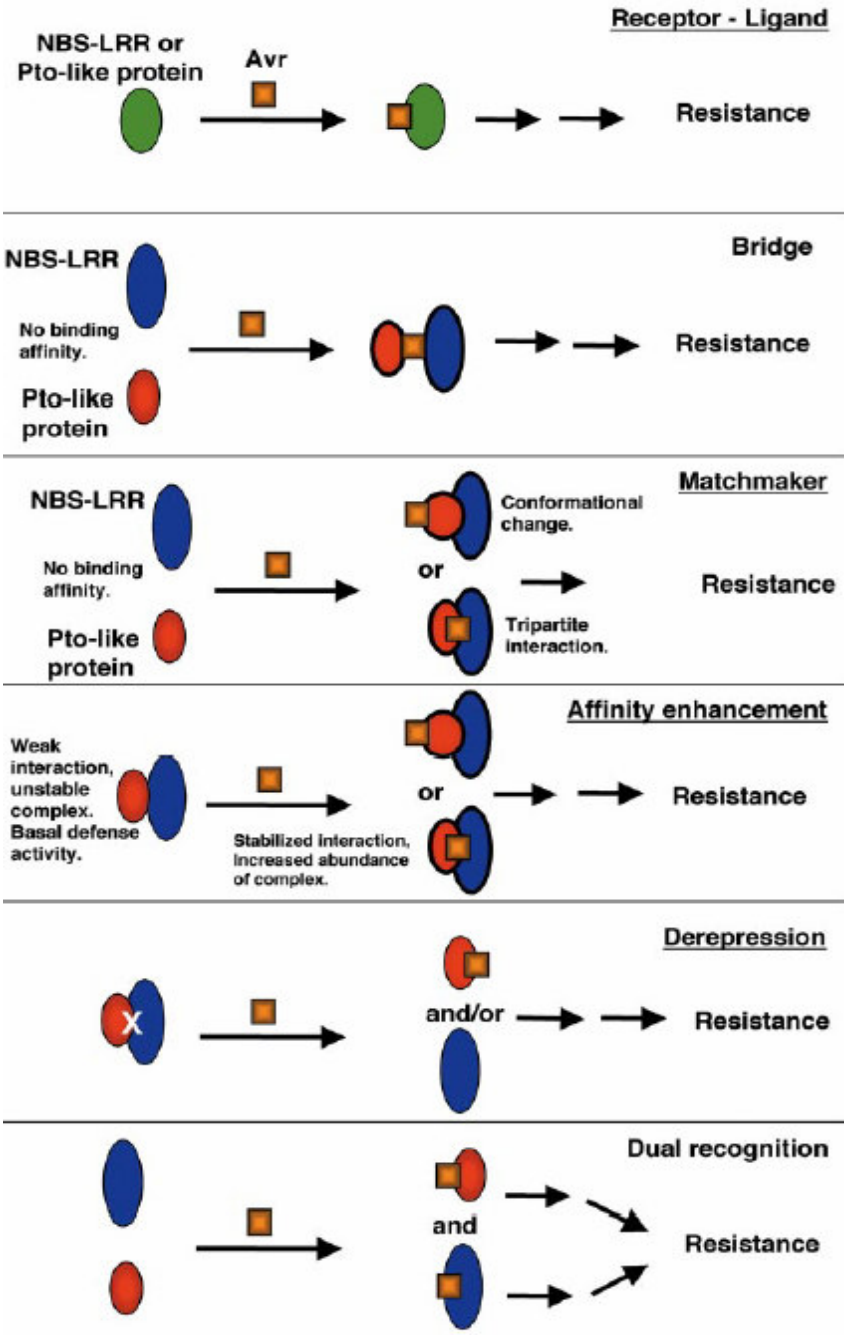


Figure 1.2 Models of R-Avr interactions (Martin *et al*., 2003)

1.2.3 Hypersensitive response (HR)

R–Avr recognition activates the defense signaling pathways frequently promote localized programmed plant cell death (the hypersensitive response [HR]), an oxidative rupture producing reactive oxygen intermediates (ROI), (Cohn *et al*., 2001) and the accumulation of salicylic acid (SA), a phenolic molecule necessary for the induction of systemic acquired resistance (SAR) (Feys and Parker, 2000).

The HR is characterized by rapid, local cell death at the site of infection (Hammond-Kossack and Jones, 1996). The cell death response is for the limitation of pathogen growth by limiting its access to further nutrient sources and the propagation. HR starts by plant's recognition of specific signal molecules produced by the pathogen. These elicitors are produced directly or indirectly by *avr* genes and *R* genes are thought to encode receptors for those elicitors. After recognition of avirulent signal, R gene products activate a series of host genes that leads to HR and inhibit pathogen growth.

During HR a broad range of physiological changes (production of reactive oxygen species [ROS], transient ion-flux resulting in intracellular pH changes, cell wall strengthening near the infection site, release of secondary signal molecules such as nitric oxide, and the synthesis of antimicrobial compounds including phytoalexins and pathogenesis-related proteins [PR]) occur, resulting in inhibition of the further attack of the pathogen (Cohn *et al*., 2001).

1.2.4 Systemic Acquired Resistance (SAR)

HR and other necrotic reactions are thought to start the development of a consequent response, referred to as systemic acquired resistance (SAR) which acts non-

specifically against a broad spectrum of pathogens (Dempsey, 1999). SAR is related with the activation of genes encoding pathogenesis-related (PR) proteins, some with antimicrobial activity (Van Loon and Van Strien, 1999), and with the ability to induce cellular defense responses more rapidly and to a greater degree than in non-induced plants (Mur *et al*., 1996). SAR reduces the harshness of disease caused by all classes of pathogens. Transgenic *Arabidopsis* plants carrying the *Pseudomonas fluorescens nah*G gene that encodes the salicylate hydroxylase (inactivates SA), are more susceptible to both pathogens that normally induce a resistance response and to pathogens that normally cause disease (Dempsey, 1999).

1.2.5 Classification of *R* genes based on structurally conserved domains

Five classes of disease resistance genes have been defined (Figure 1.3) according to the structural characteristics of their predicted protein products: (reviewed by Martin *et al*., 2003)

Class 1 is the serine/ threonine kinase catalytic region and a mrystylation motif at N terminus. The only examples of this class are the *Pto* from tomato (Martin *et al*., 1993) and *Rpg1* from barley (Brueggeman *et al*., 2002).

Class 2 is the most common class. These R proteins have a region of carboxy-terminal leucine-rich-repeats (LRRs), a central nucleotide binding (NBS) domain, and an N-terminal putative leucine-zipper (LZ) or other coiled-coil (CC) sequence.

The LLR domain contains 9-41 repeats (each is about 25 amino acids long) with a consensus amino acid sequence of xx(L)x(L)xxxx (Cooley *et al.*, 2000). It is

believed to be responsible for the recognition specificity (Ellis *et al*, 1999; Dodds *et al*., 2001).

The NBS domain contains three motifs: kinase-1a (p-loop), a kinase-2, and a putative kinase-3a (Traut 1994; Tameling *et al*., 2002). The consensus sequence of GxxGxGK(T/S)T, LxxxDDVW and GxxxxTxR for the *R* gene motifs p-loop, kinase-2, and the putative kinase-3a, respectively, is different from that present in other NBS-encoding proteins (Hammond-Kosack and Jones, 1997; Meyers *et al.*, 1999). Many R proteins' NBS domain show homology to nearly 320 amino acid region of the APAF-1 and CED-4 proteins which are involved in regulating programmed cell death in animals (Van der Biezen and Jones, 1998). The alignment contains five other short motifs of undefined function and was designated the NB-ARC (nucleotide binding in APAF-1, R gene products, and CED-4) domain. Other motifs present in the NBS domain of NBS/LRR-type *R* genes are GLPL, RNBS-D and MHD (Meyers *et al.*, 1999). The sequences spread between these motifs and domains can be very different even among homologues of an *R* gene (Michelmore and Meyers, 1998; Pan *et al.*, 2000). Recently, a new unspecified motif (VMVCVS) is found between the kinase-1a and kinase-2 domains within the Resistance Gene Analogs of wheat (Bozkurt *et al*., unpublished data).

Like leucine zipper, the coiled coil (CC) structure is a repeated heptad sequence with interspersed hydrophobic amino acid residues. The role(s) of the CC domain in resistance remains to be unsolved, but general dependence on the downstream signaling components of CC-containing R proteins in *Arabidopsis is* distinct from those required for TIR-NBS-LRR proteins. This suggests that CC domain may be involved in signaling rather than recognition (Aarts *et al.*, 1998; Van der Biezen *et al.*, 2002, Warren *et al.*, 1999).

Class 3 is similar to class 2 but has a "Toll-interleukin-1 receptor" (TIR) domain similar to intracellular regions of human and *Drosophila* Toll-like receptors instead of CC motif. The highly variable LRR part of plant R proteins contribute to specificity in pathogen recognition, whereas the more conserved TIR-NBS or CC-NBS regions are responsible for transducing the pathogen-derived signal (Dangl and Jones, 2001). These first three classes do not have transmembrane (TM) domains and localized in the cell.

Class 4 lacks NBS site and have TM domain and an extracellular LLR region. Cf proteins from tomato are examples of this class (Dixon, 1996).

Class 5 has extracellular LLR, TM domain and a cytoplasmic serine/threonine kinase region. The example of this class is the *Xa21* gene of rice (Song, 1995).

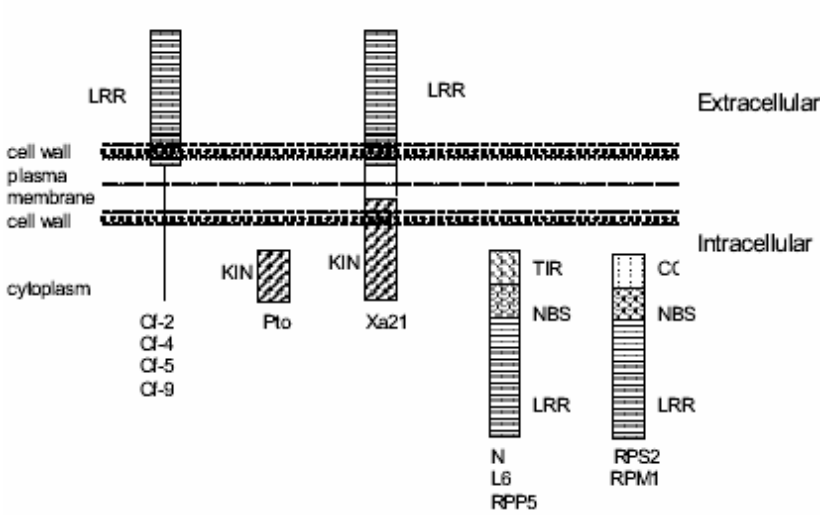


Figure 1.3 Representation of the location and structure of the five main classes of plant disease resistance proteins. Conserved structural domains: LRR-leucine-rich repeats; NBS-nucleotide-binding site; KIN-ser/thr protein kinase; CC-coiled-coil domain; TIR-Toll and interleukin receptor domain (Lehmann, 2002)

1.2.6 *R* genes in cereals

Although several disease-resistant genes have been isolated in dicots, only a small number of have been reported in monocots. There are increasing number of disease resistance genes isolated from wheat (*Lr21*, *Lr10*, *Pm3*), rice (*Xa21*, *Xa1*, *Xa26*, *Pi-b*, *Pi-ta*) and maize (*Rp1-D*, *Rp3*, *Hm1*) (Ayliffe and Lagudah, 2004). These three crops account for 85% of cereal production (FAO statistical databases, 2004).

The majority of the isolated cereal *R* genes encode proteins having a NBS domain and LRR regions. A major distinction between monocotyledonous and dicotyledonous species is that, only CC-NBS-LRR genes have been identified in

monocots whereas both CC and TIR types are found in dicots. There are some *R* genes in monocots (*Xa21* and *Xa26*) that have extracellular LLR, a transmembrane domain and a serine/threonine protein kinase domain and barley *Rpg1* gene only encodes a protein kinase.

1.2.7 *R* genes in wheat

Common wheat is an allohexaploid ($2n = 6x = 42$) containing the three genomes A, B and D. It has an extremely large genome (16 million kb per haploid cell) (Bennett and Smith, 1976) with more than 80% repetitive DNA sequences. These properties of the wheat genome make cloning of agronomically important genes by genetic methods very difficult.

Up to date there are only three *R* genes isolated and published from wheat. 50 *Lr* genes, all are known only from their phenotype and-or map positions, but only two of them have been functionally characterized. These are *Lr21* (Huang *et al*., 2003) and *Lr10* (Feuillet *et al*., 2003) genes. The third *R* gene of wheat is the Pm3 (Yahiaoui *et al*., 2004) which is responsible for the resistance to powdery mildew disease.

Yr10 gene is the only gene which is responsible for the yellow rust disease resistance in wheat. Its sequence is available in the GenBank, but only the abstract is published to date (Laroche *et al*., 2002). The stripe-rust-resistant locus *Yr10* is located on chromosome 1B in 'Moro' and originates from the Turkish line PI 178383. Two full-length NBS– LRR (nucleotide binding site–leucine-rich repeat) genomic sequences 4B and 4E have been isolated. It is found that the sequence 4B appears to be expressed while 4E is not. The expressed NBS–LRR sequence had linkage in segregating populations totaling 874 lines. The functional *Yr10* cDNA sequence of 2.5 kb

(kilobases) was isolated by RT– PCR. The genomic clone 4B was introduced in the susceptible wheat cv. Fielder, *via* microprojectile bombardment, and resulted in resistant phenotype (Laroche *et al*., 2002).

1.2.8 Common *R* gene signaling pathway components

Some signaling components are required for resistance conferred by multiple *R* genes, and are considered to function downstream of the initial pathogen recognition event. Important resistance components include the plant-specific proteins like EDS1 (Enhanced disease susceptibility1), PAD4 (phytoalexin deficient4) NDR1 (Nonrace-specific Disease Resistance1), NPR1/NIM1 (Nonexpressor of PR genes1), Rar1 (Required for mla-dependent resistance), SGT1 (Suppressor of G2 allele of SKP1).

The *Arabidopsis EDS1* and *PAD4* genes are specifically required for resistance to *P. syringae* and *Peronospora parasitica* conditioned by the same spectrum of TIR-NBS-LRR *R* genes (Feys *et al*., 2001). *EDS1* and *PAD4* genes encode lipase-like proteins (Falk *et al*., 1999; Jirage *et al*., 1999) and participate in the same signaling pathway and have different signal transduction functions (Aarts, 1998; Feys *et al*., 2001). Although *EDS1* is required for TIR-NBS-LRR class of *R* genes, *Arabidopsis NDR1* gene encodes a putative membrane-bound protein is essential for the function of most, but not all CC-NBS-LRR class R genes (Aarts, 1998; Century *et al*., 1995). A subset of TIR-NBS-LRR genes, including *RPP5*, *RPP1*and *RPP4*, are affected by mutations in the SAR-associated gene *NPR1/NIM1* (Rairdan and Delaney, 2002; Van der Biezen *et al*., 2002).

The requirement for the *Rar1* gene in resistance to powdery mildew by several unlinked CC-NBS-LRR genes is another example of a converging point in disease

resistance (Schulze-Lefert and Vogel, 2000). *Nicotiana benthamiana Rar1* (*NbRar1*) is an essential component of the *N* gene–mediated resistance response to *Tobacco mosaic virus* (TMV) (Liu et al., 2002). As *N* gene belongs to the TIR-NBS-LRR class of *R* genes (Whitham et al., 1994) *Rar1* may represent an example of a signaling component involved in pathways shared by both the CC-NBS-LRR and the TIR-NBS-LRR resistance gene classes.

Studies in barley suggest that, *Rar1* functions downstream of pathogen recognition and upstream of H_2O_2 accumulation and host cell death (Shirasu *et al*., 1999). The *Rar1* gene encodes a protein with a novel zinc finger motif called CHORD (Cys- and His-rich domain) and lacks a CS (CHORD and SGT1) domain that is found in metazoan CHORD proteins (Figure 2). The CS domain shares homology with yeast SGT1 which is a subunits of a multiprotein ubiquitin ligase (E3) complex (the SCF [Skp/Cullin/F-box] complex) (Kitagawa *et al*., 1999; Matsuzawa and Reed, 2001). This complex mediates the degradation of multiple proteins involved in diverse signaling pathways through an ubiquitin-proteasome protein degradation pathway (Deshaies, 1999).

It is found that only a subset of R proteins require Rar1, whereas SGT1 is essential for disease resistance and HR triggered by a wide range of R proteins including non-leucine-rich repeats (LRR) type R proteins (Peart *et al*., 2002)

Arabidopsis SGT1 was identified in mutational analysis for loss of *RPP7*- and *RPP5*-mediated resistance (Austin *et al.*, 2002; Tor *et al.*, 2002), and in a yeast two-hybrid screen SGT1 interacts with the Rar1 (Azevedo *et al*., 2002). Deletion analysis in the yeast two-hybrid system showed that SGT1 and Rar1 proteins interact, respectively, *via* their CS and CHORD-II domains (Figure 1.4). Rar1-SGT1 physical interaction (Figure1.3) and their interactions with ubiquitination-related proteins suggests that, the

two proteins are signaling components that may be involved in protein degradation processes (Austin *et al*.,2002; Azevedo *et al*., 2002; Muskett *et al*., 2002; Shirasu *et al*., 1999; Tor *et al*., 2002, Tornero *et al*,. 2002) (Figure 1.5).

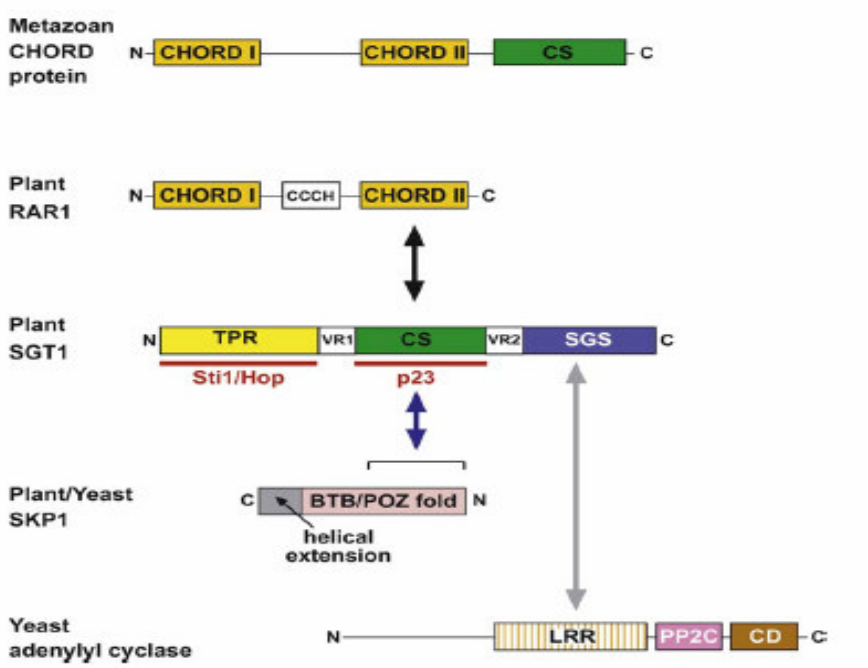


Figure 1.4 SGT1 conserved domains and interactions (Muskett and Parker, 2003). Interactions are shown with double-ended arrows. **Domains of SGT1 protein:** TPR (tetratricopeptide) (homology to the co-chaperones Sti1/Hop), VR1 (variable region 1), CS (CHORD and SGT1 motif) (homology to p23), VR2 (variable region 2), SGS (SGT1-specific motif). **Domains of Rar1:** CHORD I (cysteine- and histidine-rich domain), CCCH, and CHORDII. **Domains of SKP1:** BTB/POZ fold (bric a brac, tramtrack, broad complex/poxvirus and zinc finger), and helical extension. **Domains of Adenylyl cyclase:** LRR (leucine rich repeat), PP2C (protein phosphatase 2C domain), CD (catalytic domain).

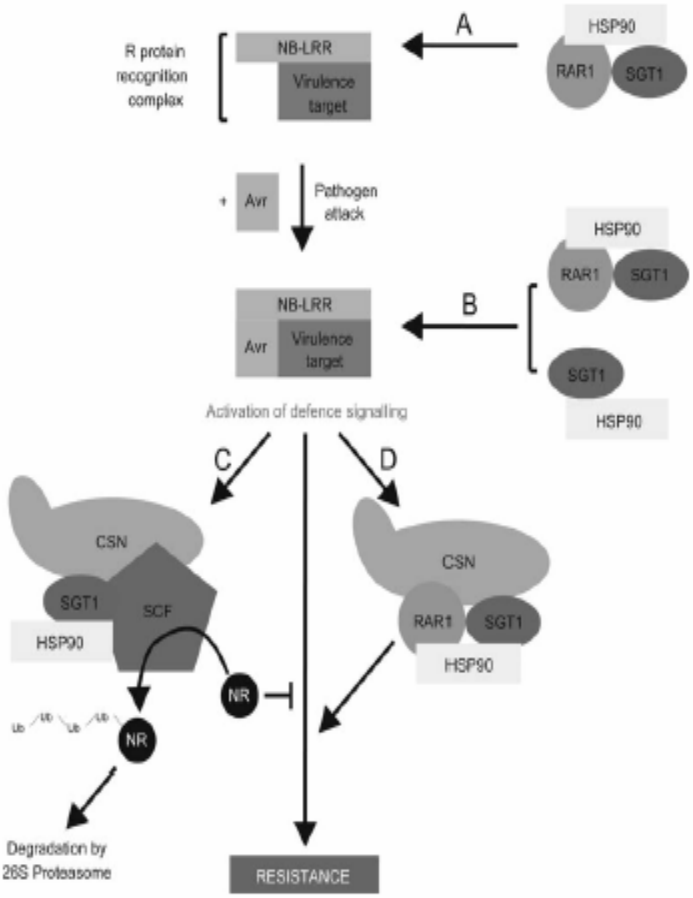


Figure 1.5 Possible roles for SGT1 in R gene product mediated resistance (Muskett and Parker, 2003). **A:** R protein complex formation for the recognition of virulence target (Possible co-operation with Rar1 and HSP90). **B:** After activation of the R protein by the Avr effector, might be responsible for regulation of the correct folding or intermolecular associations of R protein recognition complex following (Possible co-operation with HSP90 +/– Rar1). **C:** Removal of the negative regulators (NR) of the resistance response via SCF-mediated ubiquitination (Rar1 independent). **D:** For the removal of negative regulators of activation of the positive regulators, a downstream function with Rar1 and CSN (COP9 signalosome), possibly part of a non-SCF-type E3 ligase complex.

Cytosolic heat shock protein 90 (HSP90), a molecular chaperone, is found to be another Rar1 interacting protein by yeast two-hybrid screening (Takahashi *et al*., 2003). It is shown that Rar1 interacts with the N-terminal half of HSP90 that contains the ATPase domain. HSP90 also found to be specifically interacting with SGT1 that contains a tetratricopeptide repeat (TPR) motif and a domain with similarity to the co-chaperone p23 (Takahashi *et al*., 2003).

HSP90 functions in protein complexes with a large set of co-chaperones (Picard, 2002). Thus, it is possible that the HSP90 chaperone complex together with Rar1 and SGT1 is responsible for modulating the activity and/or stability of substrate proteins that are essential for disease resistance signaling (Takahashi *et al*., 2003).

1.2.9 Ubiquitination and plant disease resistance

Ubiquitin is a small, 76 amino acid protein and is present in all cells. Covalent attachment of the ubiquitin C terminus to the substrate lysine residues, a process known as ubiquitylation, targets the substrate for a range of possible fates, the best known of which is degradation by the 26S proteasome, but which also include endocytosis, targeting to lysosomes, and modification of protein function (Hershko and Ciechanover, 1998).

Activities of ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2) and ubiquitin-ligating enzyme (E3) initiates polyubiqutination, usually targets a substrate for degradation by proteasome. The genome encodes a few E1, but diversity of ubiquitin system arises from 20 different E2 and even greater number of E3 enzymes (Weissman, 2001). As a matter of fact one of the important component of R protein mediated resistance, SGT1, is involved in such activity through SCF-mediated

ubiquitination. However, no ubiquitin ligase targets associated with disease resistance have yet been identified in plants (Muskett and Parker, 2003).

SCF complexes (Figure1.6) are a large class of RING-type ubiquitin E3 ligases that facilitate the attachment of ubiquitin to specific protein substrates, targeting them for subsequent degradation by the 26S proteasome (Deshaies, 1999). The RING-finger protein RBX1 (REGULATOR OF CULLINS1 [ROC1]) is a core component of the SCF complex, and interacts directly with CULLIN1. SKP1 links CUL1 to a F-box protein, which acts as an adaptor molecule that recruits specific substrates to the SCF complex for ubiquitination (Craig and Tyers, 1999). Structural analysis of the RBX1–CUL1–SKP1–F-box complex revealed that CUL1 serves as a scaffold that orientates the other subunits to allow transfer of the ubiquitin moiety from E2 to the target protein (Zheng *et al*., 2002)

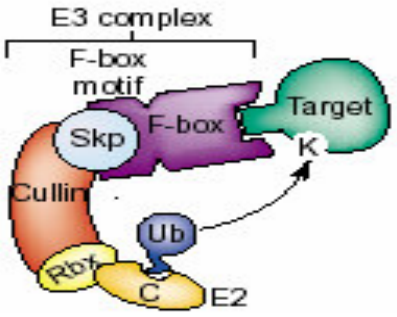


Figure1.6 The SCF complex. The pathway of ubiqiutin transfer from the E2 (ubiquitin-conjugating enzyme) to accessible lysines (K) in the target is indicated by the arrows (Vierstra, 2003).

Ubiquitination plays an important role in the regulation of biological processes such as cell cycle regulation, signal transduction, apoptosis, and DNA repair (Ciechanover, 1994; Hiyama *et al*., 1999; Hauser *et al*., 2000; Hershko and Ciechanover, 1998). Thus, it is also very important for the *R* gene mediated disease resistance.

The role of ubiquitination and protein degradation in disease resistance is still a question. Possible roles might be the removal of negative regulators of plant defense responses, the modulation of R protein levels (Martin *et al*., 2003), the modification of signal proteins and the regulation of cellular processes such as transcription, protein trafficking, membrane transport or activation of protein kinases (Pickart, 2001).

1.2.10 *Rad6* gene

Rad6 is an ubiquitin-conjugating enzyme (E2) (Sung *et al*., 1990). It is shown that yeast *rad6* mutants have defects in DNA repair, proteolysis of N-end rule protein substrates, cellular proliferation, cell cycle progression, and an inability to sporulate (Watkins *et al*., 1993). Rad6 can ubiquitylate histone H2B in yeast (Robzyk *et al*., 2000), and plays a role in the post-replication repair (PRR) pathway (Prakash *et al*., 1993; Xiao *et al*., 2000). Recently, Rad6 cDNA and recombinant protein was isolated from rice, *Oryza sativa* L. cv. Nipponbare (Yamamoto *et al*., 2004). In this report, it is shown that there is interaction between the SGT1 and Rad6. There is an increased expression of Rad6 due to H_2O_2 treatment. Thus, it is believed that Rad6 may have a function to stop the cell growth. Although this report is not related with disease resistance in plants, all these findings suggest that Rad6 protein might have a role in disease resistance. The support comes from our studies, in which *Rad6* gene has

differentially expressed in resistant wheat to yellow rust disease (Bozkurt *et al*., unpublished data).

In tobacco *Rad6* homolog, named as "tobacco ubiquitin conjugating enzyme (NtUBC2)", is found to be interacting with ERF3 (Ethylene Responsive Factor3) (Koyama *et al*., 2003). ERF3 is a member of the ERF-domain transcription factors and has a transcriptional repressor activity. AtERF3 genes were found to be induced by ethylene, jasmonic acid, and an incompatible pathogen (Fujimoto et. al., 2000). It is suggested that NtUBC2 might have a involvement in the regulation of repression activity of ERF3 (Koyama *et al*., 2003). These findings strongly suggest that *Rad6* gene might have regulatory role in plant disease resistance pathway.

1.3 Yeast two-hybrid system (Interaction trap)

Protein-protein interactions are responsible for all kinds of processes and signaling in a cell. Proteins rarely function in separately. In a given cell proteins are connected through large interaction networks, where non-covalent interactions are continuously forming and dissociating. Interactions changes during development and in response to external stimuli. To understand the cellular mechanism all the interactions between proteins need to be defined.

There are genetic, biochemical, microbiological and biophysical techniques to understand the protein-protein interactions. Co-immuno precipitation, yeast and bacterial two-hybrid selection systems, imaging protein-protein interactions by FRET microscopy, Tandem affinity purification and mass spectroscopy of the purified complexes, *In silico* (computed) predictions are some methods used for the protein-protein interactions.

To study protein-protein interactions, the yeast two-hybrid system (Fields and Song, 1989) is one of the most effective tool. The method has now been extensively used both to identify novel protein-protein interactions and to further analyze known interactions. Many classes of proteins can be used as baits in this system. There are two main requirements in order to perform the yeast two-hybrid analysis: i) bait protein must be effectively transported to nucleus (not a membrane protein), ii) do not activate transcription (reviewed by Golemis, 2002).

The first application of two-hybrid methods to probe protein function was to examine the interactions between proteins isolated by two hybrid methods and relatively small numbers of test proteins (Durfee *et al*., 1993; Gyuris *et al*., 1993; Harper *et al*., 1993; Zervos *et al*., 1993), but their use quickly spread to the analysis of many other proteins.

1.3.1 Basic mechanism of the yeast two-hybrid technique

The complex process of transcription initiation requires the ordered assembly of numerous interacting transcription factors with RNA polymerase and ancillary proteins, into a protein machine that initiates transcription (Guarente, 1996; Tjian and Maniatis, 1994). The yeast two-hybrid technique uses two protein domains that have specific functions: a DNA-binding domain (DBD) that is able to bind to DNA, and an activation domain (AD), that is capable of activating transcription of the DNA. In order for DNA to be transcribed, it requires a protein called a transcriptional activator (TA). This protein binds to the promoter region and it is then able to activate transcription *via* its activation domain (Figure 1.7). For the activity of a TA both DBD and AD are required and if either of these domains is absent, then transcription of the gene will fail.

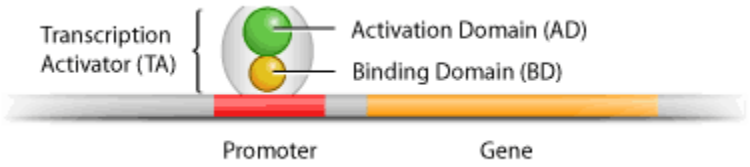


Figure 1.7 Control of transcription (Long and Sobhanifar, 2005)

By using the information on the control of transcription, the yeast two-hybrid system was developed (Fields and Song, 1989). The method is based on the two chimeric proteins, a fusion of a protein of interest "**x**" to a DNA binding domain (DBD-x) and a fusion of a cDNA library "**X**" to a transcriptional activation domain (AD-X) (Figure 1.8). There most frequently used transcription controls are bacterial repressor protein LexA, bacteriophage lambda cI and the amino-terminal end of the yeast transcription factor Gal4 systems. They all have the same principal mechanism.

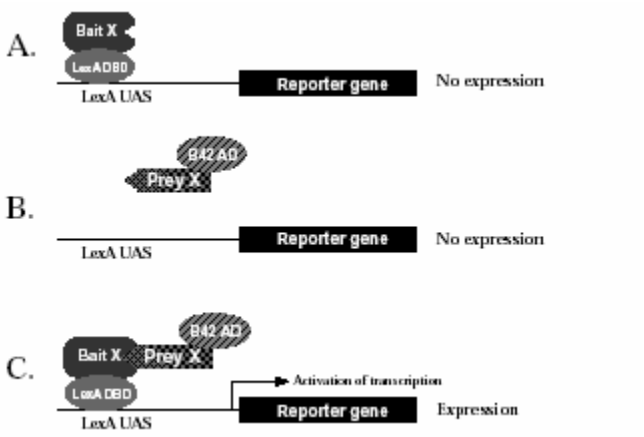


Figure 1.8 Transcriptional control and its usage in yeast two-hybrid system. **A:** There is only bait protein (**x**) which is fused to LexA DBD so there is no activation of the transcription and expression of the reporter gene. **B:** There is only library (prey) protein (**X**) fused to B42 AD so there is no activation and expression of the reporter gene. **C:** The bait and the prey interaction results in the formation of transcription activator. There is expression of the reporter gene (Dual Bait Hybrid Hunter Manual, Catalog no: K5200-01, Invitrogen).

1.3.2 LexA based yeast two-hybrid system

The DBD used is a bacterial repressor protein LexA and the AD is a protein encoded by random *E. coli* sequences, the B42 (acid blob). First reporter contains *lexA* operator sites (DNA-binding sites) upstream of the *lacZ* gene (encodes for β-galactosidase) and a second reporter contains *lexA* operators upstream of *LEU2* (a required gene in the yeast leucine biosynthetic pathway).

Yeast having the non-transcriptional activator bait (DBD-**x**) do not turn on either of these reporters, so in the absence of LacZ or Leu2 proteins cells are white on medium containing X-Gal and are not able to grow on medium lacking leucine.

An AD-cDNA library (AD-**y**) is introduced into yeast having non-transcriptional activator bait (either by mating or direct transformation) and plated on selective media. When library-encoded protein interacts with the bait, AD is brought to the *lexA* sites flanking the *lacZ* and *LEU2* reporters, resulting in colony growing on medium lacking leucine, and blue color on medium containing X-Gal (Figure1.7).

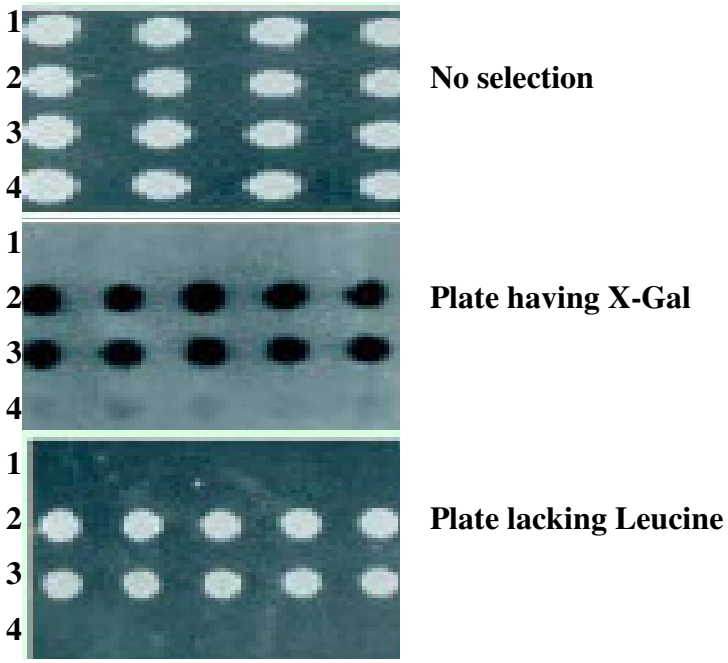


Figure 1.9 Selection of interactions. Colony number1 has no interacting partners as it does not for blue color on X- Gal containing medium and is not able to grow on medium lacking leucine. Colony numbers 2 and 3 have interacting bait and library proteins so it forms blue color on X-Gal containing medium and can grow in the absence of leucine. Colony number 4 shows a very weak interaction as it shows slightly blue color on X-Gal containing medium.

1.3.3 Two-hybrid Dual Bait System

This method is developed by Serebriiskii *et. al*. in 1999. This system incorporates controls for false positive or nonspecific interactions in a single step and allows the immediate assay of a protein interaction with two related or unrelated

partners in a single cell, which is also useful for a variety of high throughput and genome-oriented studies.

In dual bait system there two bait constructs. The first bait protein consists of a LexA DBD/Bait **x** fusion while the prey protein consists of a B42 AD/Prey **X** fusion. Prey X can be replaced with a cDNA library to screen for unknown interactors. An interaction between bait **x** and prey **X** brings the B42 AD together with the LexA DBD to reconstitute transcriptional activation and allow expression of two reporter genes, the *LEU2* auxotrophic marker and *lacZ*, *via* binding to the LexA operator sites (Figure1.10). Positive interactions can be detected by selecting for leucine prototrophy and β-galactosidase activity. This is the standard two-hybrid method.

A further level of complexity is introduced to the system with the addition of a second bait protein consisting of a fusion between the bacteriophage lambda cI (cI repressor protein) DBD and bait Y. Screening a cDNA library will then allow identification of prey Y proteins that interact with the bait Y. In this case, an interaction between bait Y and prey Y brings the B42 AD together with the cI DBD (Figure1.8) to reconstitute transcriptional activation and allow expression of two reporter genes, the *LYS2* auxotrophic marker and *gusA, via* binding to the cI operator sites. Positive interactions can be detected by selecting for lysine prototrophy and β-glucuronidase activity.

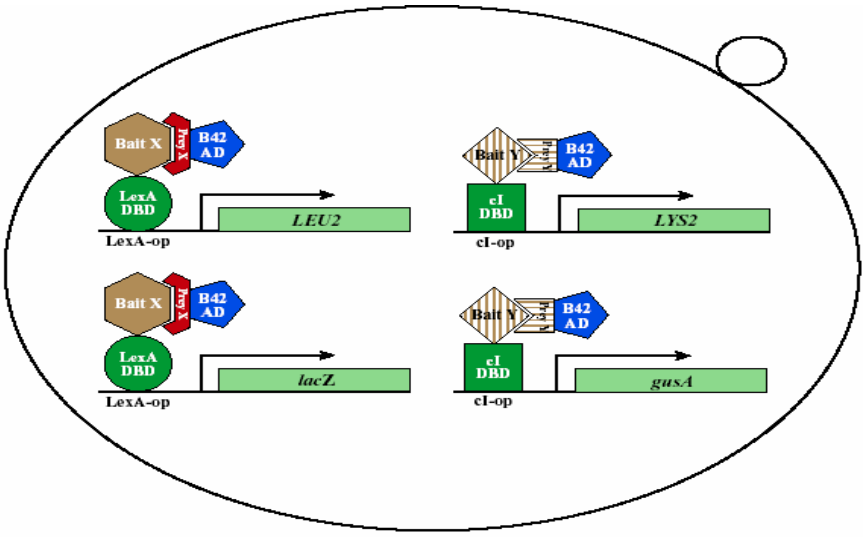


Figure 1.10 The dual bait yeast twp-hybrid system (Dual Bait Hybrid Hunter Manual, Catalog no: K5200-01, Invitrogen).

1.4 Aim of the study

Aim of our study is to understand the *Yr10* gene mediated yellow rust disease resistance mechanism in wheat. *Yr10* is the only gene that is cloned to date and its sequence is available in the GenBank. Although, there are many R genes in wheat only three of them are molecularly characterized. It is difficult to clone genes in wheat by map based cloning methods or other strategies, since the wheat geneome is very large and complex. Having a hexaploid genome with more than 80% repetitive DNA sequences makes the cloning even more difficult. To this date there is no report in literature that targets to understand the downstream signaling in wheat disease resistance. These types of studies are mostly conducted in model plant system as

Arabidopsis. Wheat is a monocot species therefore it may have different or additional signaling systems in the resistance. This study group believes that molecular level characterizations of disease resistance must also be conducted in the economically important plants if we would like to make direct use of these fundamental research results for practical applications in improving wheat.

In order to understand the signaling cascades, it is important to find protein-protein interactions. To enlighten the yellow rust disease resistance pathway in wheat, we try to find novel protein-protein interactions that may have a role in disease signaling. The method used in this study for the detection of protein-protein interactions is the dual bait yeast two-hybrid analysis. This is a powerful technique that is commonly used for the discovery of the novel protein interactions. As baits we have used the *Yr10* gene fragments, the *Rad6* gene (that is differentially expressed during yellow rust disease resistance in Avocet-YR10 differential line), and WR5 (which aligns with a WD40 repeat containing protein). The yeast library is screened with these baits, to find probable conserved protein-protein interactions.

Although the signaling pathway during the resistance response is very complex, still there is involvement of much conserved protein such as SGT1, Rar1, and HSP90. After the recognition of the Avr protein by the R protein, the signaling for resistance starts in the plant cell. This recognition triggers the hypersensitive cell death, which is a kind of programmed cell death. The similarity of the R proteins' NB-ARC domain with APAF-1 and CED-4 proteins (which are involved in regulating programmed cell death in animals) supports the idea that there might be some conserved proteins involved in controlled cell death downstream signaling both in plants and animals. To find these conserved proteins we used the yeast model system, as its genome is widely known.

Our findings are very important as this is the first study in literature which aims to understand the *Yr10* mediated signal transduction events. The obtained results might be involved in many *R* gene mediated disease responses, as we target to find the conserved pathways.

CHAPTER 2

MATERIALS AND METHODS

2.1 Plant materials

Avocet-Yellow Rust differential line YR10 developed by Wellings Colin (Australia) was used as plant material for the amplification of *Yr10* gene fragments.

Pastor (Dr. Singh), Avocet (Dr. Wellings), Gerek79 (Dr. Yıldırım) and Bezostaja (Dr. Yıldırım) plants were used for the amplification of wheat *SGT1* and *Rar1* gene fragments.

2.2 Germination of plants

Single seeds were wrapped by filter papers. They were wetted by tap water and incubated at 4°C for 3 days in order to vernalize. After 3 days period, they were incubated at room temperature in a dark environment for one week. Finally, seedlings were transferred to soil and incubated at growth chamber for 14 days. Growth conditions were 18°C at light period for 12 h and 12°C at dark period for 12 h. Samples were collected from leaf tissues of plants and transferred to – 80°C until use.

2.3 DNA isolation of plant material

For the PCR amplifications of *Rar1* and *SGT1* gene fragments, DNA samples from Pastor, Avocet, Gerek79 and Bezostaja plants were isolated. DNA isolation was performed using 'Qiagen DNeasy Plant Mini Kit' as suggested by the manufacturer's instructions. In the DNeasy plant DNA isolation procedure, plant material is first mechanically disrupted and then lysed by addition of lysis buffer and incubation at 65^{o}C. RNase in the lysis buffer digests the RNA in the sample. After lysis, proteins and polysaccharides are salt precipitated. Cell debris and precipitates are removed in a single step by a brief spin through QIAshredder, a unique filtration and homogenization unit. The cleared lysate is transferred to new tube and binding buffer and ethanol were added to promote binding of the DNA to the DNeasy membrane. The sample is then applied to a DNeasy spin column and spun briefly in a microfuge. DNA binds to the membrane while two wash steps efficiently remove contaminants such as proteins and polysaccharides. Pure DNA was eluted in a small volume of PCR quality water.

Protocol for isolation of DNA from plant tissue with the 'DNeasy Plant Mini Kit':

Wet plant tissue (100 mg) was ground under liquid nitrogen to a fine powder using a pestle in a microfuge tube and allowed the liquid nitrogen to evaporate without allowing the samples to be thawn. To the ground plant tissue 400 µL of buffer AP1 (lysis buffer) and 4 µL of RNase A stock solution (100 mg/mL) was added immediately and shaked vigorously. The mixture was incubated for 10 minutes at 65^{o}C and mixed 2-3 times by inverting the tubes. Buffer AP2 (130 µL) was added to the lysate, mixed and incubated for 5 minutes on ice for the precipitation of detergent, proteins, and polysaccharides. The lysate was applied to the QIAshredder spin column sitting in a 2 mL collection tube and it was centrifuge for 2 minutes at 15,000 rpm. The flow-through fraction from previous step was transferred to a new tube without

disturbing the cell-debris pellet and on to the cleared lysate 0.5 volume of Buffer AP3 and 1 volume of ethanol (96-100%) was added and mixed by pipetting. The mixture (650 µL) was applied to DNeasy mini spin column and centrifuged for 1 minute at 10,000 rpm and the flow-through was discarded. This part was performed once more with the remaining mixture. DNeasy column was placed in a new 2 mL collection tube and 500 µL Buffer AW was added onto the column and centrifuged for 1 minute at 10,000 rpm. The flow-trough was discarded and again 500 µL Buffer AW was added to column and centrifuged for 2 minutes at 15,000 rpm to dry the column membrane. The column was transferred to a microfuge tube and 100 µL of preheated (65^{o}C) Buffer AE was directly pipetted on to the column membrane and incubated for 5 minutes at room temperature and then centrifuged for 1 minute at 10,000 rpm for elution of bound DNA. This step was repeated once more in order to obtain the entire bound DNA.

2.4 Total RNA isolation from plant leaf tissue

From Pastor, Avocet, Gerek79, Bezostaja and Avocet-YR10 plants RNA was isolated. RNA isolations were done using the TRIzol reagent (Invitrogen) according to the manufacturer's protocol.

Homogenization: Tissue samples were powdered using a mortar and liquid nitrogen. Powdered tissue was homogenized in 1 mL of TRIzol reagent per 50-100 mg of tissue in a 2 mL sterile tube.

Phase separation: Homogenized samples were incubated for 5 minutes at room temperature to permit the complete dissociation of nucleoprotein complexes. Then 0.2 mL of chloroform per 1 mL of TRIzol reagent was added. Sample tubes were capped securely. Tubes were shaked vigorously by hand for 15 seconds and incubated at room

temperature for 2-3 min. The samples were centrifuged at 15,000 rpm for 15 min. at 4^{o}C. Following centrifugation, the mixture separates into a lower red, phenol-chloroform phase, an interphase, and a colorless upper aqueous phase. RNA remains exclusively in the aqueous phase.

RNA precipitation: The upper phase was transferred to a fresh tube. RNA was precipitated from the aqueous phase by mixing with isopropyl alcohol. 0.5 mL of isopropyl alcohol was used per 1 mL of TRIzol reagent used for the initial homogenization. Samples were incubated at room temperature for 10 min. and centrifuged at 15,000 rpm for 10 min. at 4^{o}C. The RNA precipitate observed was a gel like pellet on the side and bottom of the tube.

RNA wash: Supernatant was removed and the RNA pellet was washed once with 75% ethanol, adding at least 1 mL of 75% ethanol per 1 mL of TRIzol Reagent used for the initial homogenization. Samples were mixed by vortexing and centrifuged at 10,000 rpm for 5 min. at 4^{o}C.

Re-dissolving the RNA: At the end of the procedure, the RNA pellet was briefly dried for 10-15 min. RNA was dissolved in DEPC treated sterile water and incubated for 10 min. at 55^{o}C-60^{o}C.

2.5 Determination of concentration of DNA and RNA samples

RNA and DNA samples were diluted by 1200 folds to 0.5 mL in double distilled water and their absorbance values were measured at 230 nm, 260 nm, and 280 nm in Schimadzu UV-1601 spectrophotometer. UV light is absorbed by DNA and

RNA at 260 nm, however it is also required to know the absorbance at 230 nm and absorbance at 260 nm values in order to evaluate the contamination level of phenolic compounds and proteins, respectively.

Concentration determination of DNA and RNA samples was made according to the equation given below:

$1A_{260nm}$ = 50µg/mL DNA (or 40 µg/mL RNA)

Concentration of RNA (or DNA) (µg/µL) = A_{260} value x dilution factor x 50 µg/mL DNA (or x 40 µg/mL RNA) /A_{260nm}

After determination of concentrations of DNA and RNA samples with UV spectrophotometer, the intactness of RNA was checked on 1% formaldehyde-agarose RNA gel and the intactness of DNA was checked on 1% agarose gel.

2.6 Synthesis of the first strand cDNA

Synthesis of first strand was achieved by using Superscript III reverse transcriptase (Gibco BRL) according to the manufacturer's procedure. Following components were combined in a 200 µL sterile PCR tube: 50 pmol oligodT (Gibco BRL), total RNA (1-5 µg), 0.625 mM dNTP (DNA Amp.) mix, Sterile distilled water up to 12 µL. Mixture was incubated at 65^{o}C for 5 min. and quickly chilled on ice, spinned briefly and the following components were added: 1X First Strand Buffer

(Gibco BRL), 0.005 M DTT (Gibco BRL), 30U RNase inhibitor (Ambion). The content of the tube was spinned briefly and incubated at 42°C for 2 min. Finally 200U of SuperScript III reverse transcriptase enzyme was added (Gibco BRL). The reaction was carried out at 42°C for 2 hours and stopped by incubating at 70°C for 15 min.

2.7 PCR amplifications

In a 200 µL sterile PCR tube, following components were mixed; 1X PCR Buffer (75 mM Tris-HCl with pH 8.8 at 25°C, 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween20), 0.25 mM dNTP mix (DNAmp), 1.5 mM $MgCl_2$ (DNAmp), 2U of *Taq* DNA polymerase, 15 pmol forward primer, 15 pmol reverse primer and sterile distilled water up to 20 µL volume in a 200 µL PCR tube. PCR cycling conditions were 94 °C for 2 min. as initial denaturation, 35 cycles of three steps as; denaturation at 94°C for 1 min., annealing at 55°C for 1 min. and extension at 72°C for 1 min. and 1 cycle of 5 min. extension at 72°C. Sequences of primers that were used in PCR amplifications and the gene fragments obtained in these amplifications are listed in Table 2.1 and Table 2.2, respectively.

SGT1-F and SGT1-R primers were synthesized from conserved regions of barley (Accession number: AF439974) and rice (Accession number: AF192467) *SGT1* genes. Rar1-F and Rar1-R primers were synthesized from conserved regions of barley (Accession number: AF192261) and tobacco *Rar1* (Accession number: AF480488) genes. Conserved regions were detected using ClustalX software (Jeanmougin *et al*., 1998). Primer Detective 1.01 version (Lowe *et al*., 1990) was used for the primer design. YR10 primers were synthesized according to the correct open reading frames using the *Yr10* gene sequence in literature (Accession number: AF149114).

Table 2.1 PCR primers for cloning

Primers	**5' to 3' sequence**
YR10-F-ECO	AT GGG GAA TTC ATG GAG GTC GTG ACC GG
YR10-R-XHO	CG TTT CTC GAG TCA GCG TGG AGT TAC CTT CA
YR10-1-F-ECO	GAA TTC GTG ACC GGG GCG ATG AGC A
YR10-1-R-XHO	CTC GAG AAG AGT TGT CTT CCC TAA GC
YR10-2-F-ECO	GAA TTC GGT AGT AGA GTA ATC GCA A
YR10-2-R-XHO	CTC GAG GTC GTG TAC ACG GAC AGA GCT
WR5-F-ECO	GAA TTC GGG CGA AGG GGA AGC CGT
WR5-R-XHO	CTC GAG CGC AAC CAC TAG AGT AAC AA
Rad6-F-ECO	GAA TTC ATG TCG ACT CCT TCA AGG AA
Rad6-R-XHO	CTC GAG GAG TTA GTC TGC CGT CCA GCT C
SGT1-F	ACA CTG AGG CTG TAG C
SGT1-R	CTT WGA AAA CAG ACG AGG CTG G
Rar1-F	GRA AGC AYA CAA CWG ARA AAC C
Rar1-R	TCA CAC AGC ATC AGC ATT GTG

Table2.2 Cloned gene products

Name	Size (bp)	Place of the domains on the gene	Domains		Primers used for amplification
YR10-1506	1506	10..75	putative leucine zipper domain	CC	
		97..183	putative leucine zipper domain	CC	YR10-F-Eco
		601..630	P-loop domain	NBS	YR10-2-R-Xho
		796..833,1989	domain 2	NBS	
		829..833,1989..2013	kinase domain 2a	NBS	
		2068..2106	kinase domain 3a	NBS	
		2434..2460	domain 3	ARC	
		2521..2529	domain Q/EGF	ARC	
		2656..2661	domain HD	ARC	
YR10-1563	1563	2068..2106	kinase domain 3a		YR10-2-F-Eco
		2434..2460	domain 3	ARC	YR10-R-Xho
		2521..2529	domain Q/EGF	ARC	
		2656..2661	domain HD	ARC	
		2661..3630	LRR		
YR10-1	621	10..75	putative leucine zipper domain	CC	YR10-1-F-Eco
		97..183	putative leucine zipper domain	CC	YR10-1-R-Xho
		601..630	P-loop domain		
YR10-2	594	2068..2106	kinase domain 3a		
		2434..2460	domain 3	ARC	YR10-F-Eco
		2521..2529	domain Q/EGF	ARC	YR10-2-R-Xho
		2656..2661	domain HD	ARC	
WR5	267		WD40 domain		WR5-F-Eco
					WR5-R-Xho
Rad6	459		UBCc domain		Rad6-F-Eco
					Rad6-R-Xho
SGT1	523		TPR, CS domains		SGT1-R
					SGT1-F
Rar1	452		CHORD Domain		Rar1-F
					Rar1-R

2.8 Cloning of the amplified gene fragments for sequencing

2.8.1 Ligation of re-amplified fragments to pGEM-T Easy vector

Re-dissolved DNA that was sliced from agarose gel was ligated to pGEM-T Easy vector (Figure 2.1) (Promega). Final volume of the reaction was 10 µL. Following components were combined; 6 µL redissolved DNA, 5 ng pGEM-T Easy vector (Promega), 1X Ligase Buffer (Promega) and 2U of T_4 DNA Ligase enzyme (Promega). Mixture was incubated at 4°C overnight (approximately 18 hours).

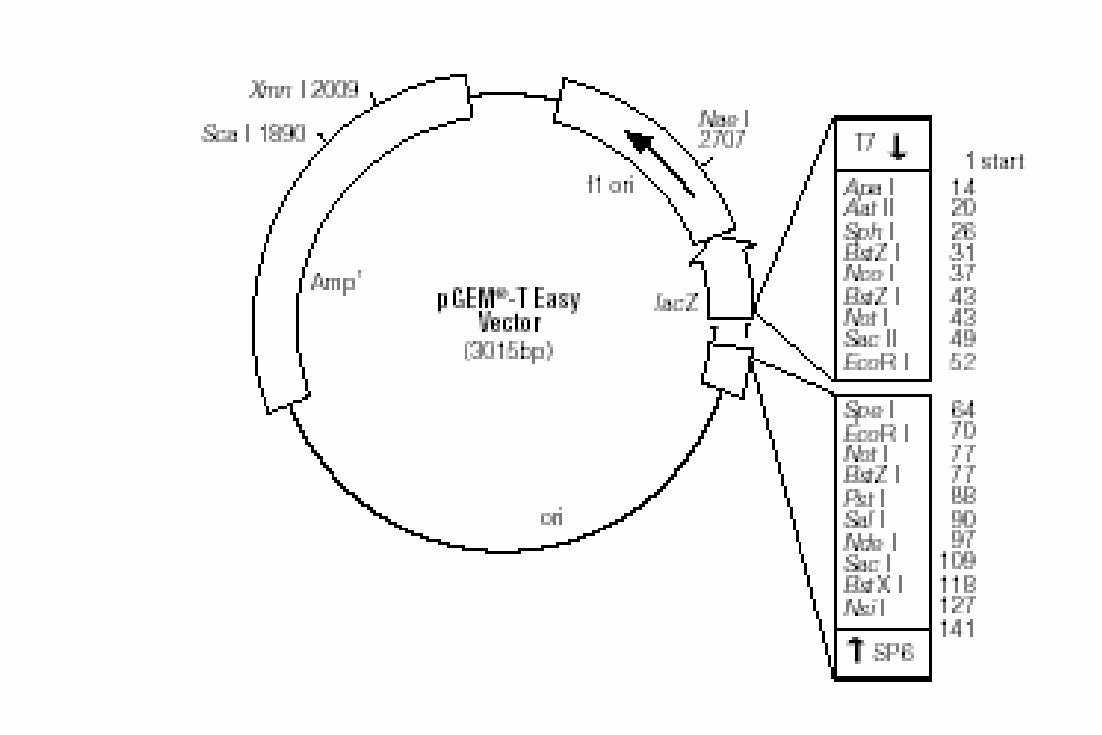


Figure 2.1 pGEM-T Easy vector circle map and sequence reference points.

2.8.2 Preparation of *E. coli* competent cells

Competent cells were prepared according to the following procedure (Fredrick *et al*. 1994). A single colony of *E. coli* DH5-α cells was inoculated into 2 mL LB medium. Cells were incubated at 37°C with moderate shaking (250rpm) overnight. Overnight grown culture (1mL) was inoculated into 100 mL of LB medium in a sterile 2-liter flask and grown at 37°C, shaking (250rpm), to an A_{590nm} of 0.375. Culture was aliquated into two 50 mL pre-chilled tubes and leaved on ice for 10 min. Cells were centrifuged for 7 min. at 3000 rpm. Supernatant was poured off and each pellet was re-suspended in 5 mL ice-cold $CaCl_2$ (4mM) solution. Cells were centrifuged for 5 min. at 2500 rpm and supernatant was discarded. Each pellet was re-suspended in 2 mL of ice-cold $CaCl_2$ (4mM) solution. Again cells were centrifuged for 5 min. at 2500 rpm and supernatant was discarded. Each pellet was re-suspended in 1.6 mL of ice-cold $CaCl_2$ (4mM) solution. Cells were dispensed into pre-chilled 500 µL PCR tube, leaved on ice for 24 hours and transferred to - 80°C after use.

2.8.3 Transformation of *E. coli* competent cells with the ligation products

E. coli DH5-α competent cells were transformed with ligation products according to the procedure of manufacturer of pGEM-T Easy vector. In a sterile 2 mL tube following components were combined; 5 µL ligation product, 0.02M β-Mercaptoethanol and 30 µL *E. coli* DH5-α competent cells. Mixture was placed on ice for 30 min., and then heat shocked at 42°C for 45 sec. SOC Medium was added up to 100 µL.

Mixture was incubated at 37°C for 45 minutes and quickly chilled on ice for 2 minutes. Mixture was spread on plates as 50 µL/plate (plates contain LB Agar). Plates

were incubated at 37°C for overnight. After the incubation, white colonies were selected among grown colonies and these colonies transferred to 100 µL LB medium containing sterile 2 mL tubes. Selected colonies were PCR amplified.

2.8.4 PCR amplification of DNA from recombinants

1X PCR buffer (DNAmp), 0.2 mM dNTP mix (DNAmp), 0.8 pmol M13 forward and reverse primers (TIB Molecular Biology), 1.2 mM $MgCl_2$ (DNAmp), 1U *Taq* Polymerase enzyme, 1 µL DNA from colonies and sterile PCR water up to 25 µL final volume combined in a sterile PCR tube. PCR cycling conditions were; initial denaturation at 94°C for 2 minutes followed by 30 cycles, 94°C for 30 sec., 55°C for 30 sec. and 72°C for 1 min.

PCR products were run on 1% agarose gel to detect positive and false positive clones (colonies that do not carry the expected sized DNA fragments). Positive clones were selected and used in plasmid isolation step.

2.8.5 Plasmid isolation from colonies

Plasmids were isolated using QIAGEN QIAprep Spin Miniprep Kit according to the manufacturer's protocol.

Bacterial culture of 2 mL was harvested by centrifugation for 5 min. at 15,000 rpm. Supernatant was poured off; tube was inverted and blotted to on a paper towel to remove excess media. Buffer P1 (250 µL) was added to tube and cell pellet was completely re-suspended by vortexing. Buffer P2 (250 µL) was added, and then the

tube mixed by gently inverting four times, incubated at RT for approximately 5 minutes. N3 solution (350 µL) was added and mixed by inverting 4 times. Sample tubes were centrifuged at 15,000 rpm for 10 minutes at RT. A spin column, provided by manufacturer, inserted into collection tube. Cleared lysate, the upper phase was transferred to the spin column, centrifuged at 15,000 rpm for 1 minute at room temperature.

After the centrifugation, lower phase in the collection tube was discarded and collection tube reinserted. 750 µL of Column Washing solution Buffer PE was added to spin column, centrifuged at 15,000 rpm for 1 minute at RT, flow through was discarded and the collection tube reinserted again and to remove residual wash buffer completely additional 1 min. centrifugation was involved. The contents of spin column was transferred to a new sterile 1.5 mL tube, the plasmid DNA was eluted by adding 50 µL of Buffer EB (elution buffer) and centrifugation at 15,000 rpm for 1 min. Spin column assembly was removed and harvest stored at –20°C.

2.8.6 Visualization of isolated plasmids

In order to determine DNA integration to the plasmid and to be informed about the size of the insert, plasmid DNA was digested with *Eco*RI restriction enzyme. Purified plasmid (5 µL), 1X NE Buffer (New England Biolabs), 1 U of *Eco*RI enzyme (New England Biolabs) and PCR ddH_2O were combined in a PCR tube with a final volume of 10 µL. Mixture was incubated at 37°C overnight (approximately 18 hours). After the incubation samples were run on 1% agarose gel.

2.8.7 DNA sequencing reactions

Plasmids were purified as described in Section 2.8.5. Inserts were custom sequenced, and were read using T7 and SP6 primer. Sequencing reactions of the 600 ng purified pGEM-T Easy (Promega) recombinant colonies were performed and the sequences were read on ABI prism-310 Genetic Analyzer in Ohio State University, sequencing facility.

2.8.8 DNA sequence and homology analysis

NCBI databanks were searched for finding homolog sequences available in literature (Zhang and Madden, 1997). Conserved domains were found by searching the NCBI databanks (Marchler-Bauer and Bryant, 2004). Sequence analysis was performed using Editseq 4.0 (DNA Star Inc.) expert sequence analysis software. The sequence alignments were obtained using ClustalX 1.81 (Jeanmougin *et al*., 1998) software.

2.9 Basic flow chart of the dual bait yeast two-hybrid system

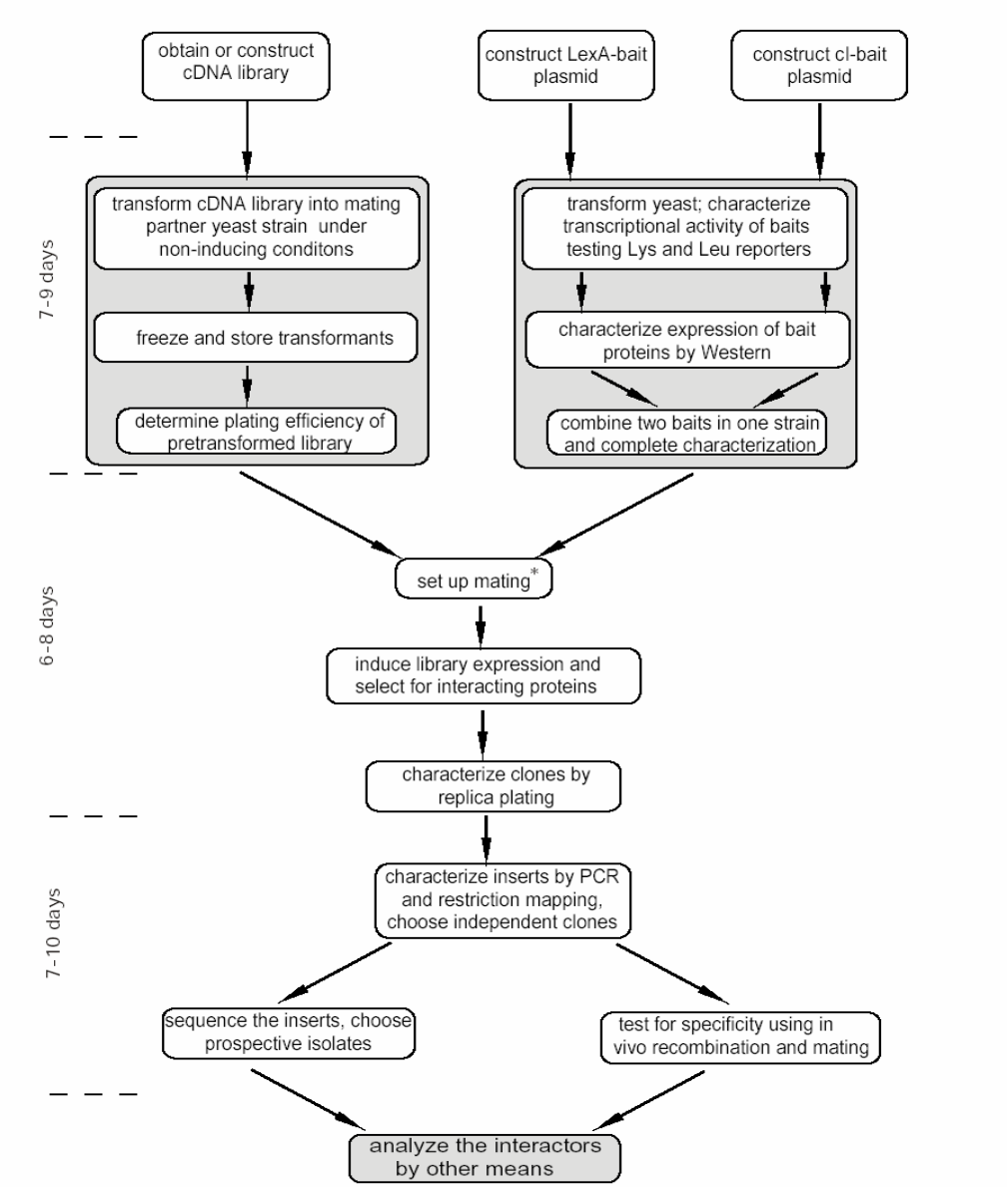


Figure 2.2 Basic flow chart of the dual bait yeast-two hybrid system (Golemis, 2002)

2.10 Constructing the bait vectors for yeast two-hybrid analysis

For the construction of the bait vector there were two approaches; one requires the band purification from the agarose gel and the other does not. During band purification the sticky ends might be disrupted. Thus, the method that does not require any purification step can be used if the first approach does not work.

2.10.1 Constructing the bait vectors with band purification from agarose gel

2.10.1.1 Restriction digestion of isolated plasmids

Isolated plasmids were cut in order to be cloned into bait vectors pMW103 and pGLS23 (Figure 2.3). Bait plasmids pMW103 and pGLS23, pGEM-T plasmids having the genes of interest (YR10-1506, YR10-1563, YR10-1, YR10-2, WR5 and Rad6.) were cut using *Eco*RI and *Xho*I enzymes. 5 µL purified plasmid, 1X NE Buffer (New England Biolabs), 5 U of *Eco*RI enzyme (New England Biolabs), 5 U of *Xho*I enzyme (New England Biolabs) and PCR ddH_2O were combined in a PCR tube with a final volume of 10 µL. Mixture was incubated at 37 °C for 3 hours. After the incubation samples were run on 1% agarose gel and band were excised with a clean, sharp scalpel to be purified and cloned into pMW103 and pGLS23 vectors.

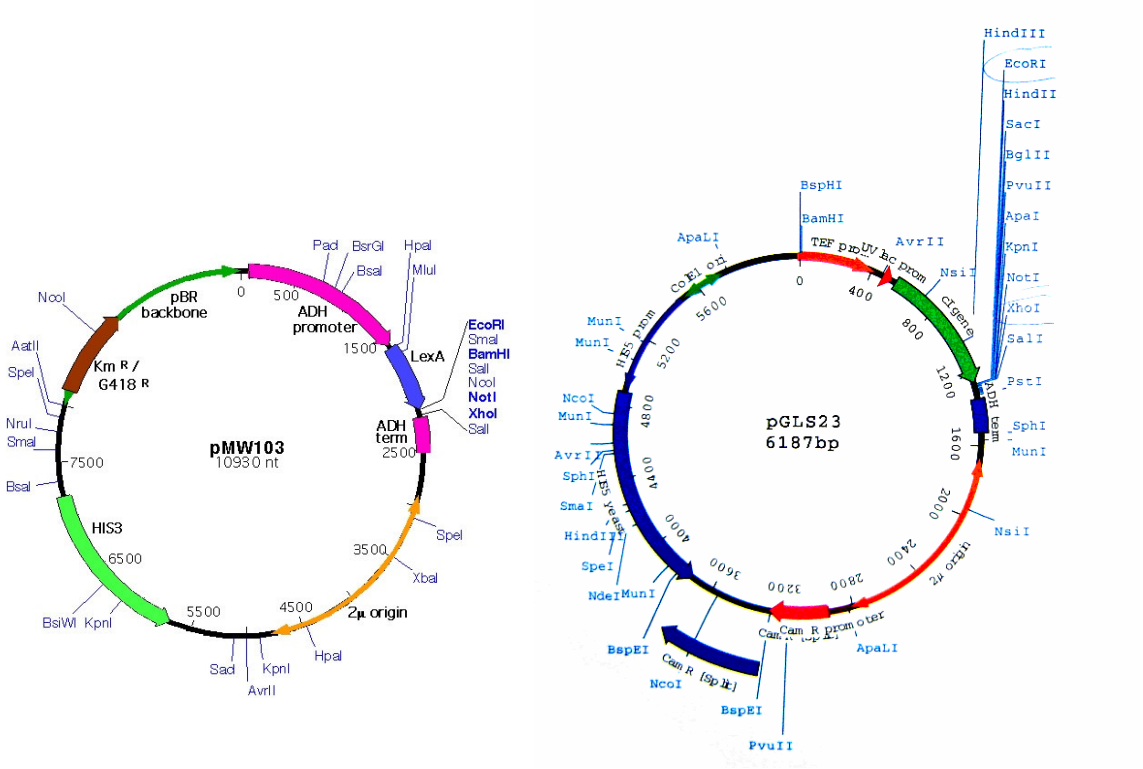


Figure 2.3 Bait vectors circle maps and sequence reference points. **LexA-DBD fusion vector (pMW103):** The strong ADH promoter was used to express bait proteins as fusions to the DNA-binding protein LexA. The plasmid contains *HIS3* selectable marker and the 2µ origin of replication to allow propagation in yeast, and *kanamycin*^{R} gene and the pBR origin of replication to allow propagation in *E. coli*. **cI- DBD fusion vector (pGLS23):** The plasmid contains *HIS5* selectable marker and the 2µ origin of replication to allow propagation in yeast, and *chloramphenicol*^{R} gene and the colE1 origin of replication to allow propagation in *E. coli*.

2.10.1.2 Purifications of DNA fragments from agarose gels

The bands that were excised from agorose gels were purified using QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's protocol.

The gel slice was weighted and 3 volume of Buffer QG was added to 1 volume of gel (100mg~100μL). Tube was incubated at 50°C for 10 min. (until the gel dissolves completely). 1 volume of isopropanol was added and mixed. QIAquick spin column was placed in a provided 2 mL collection tube. To bind DNA, the samples were applied to the QIAquick column, and centrifuged for 1 min. The flow-through was discarded and QIAquick column was placed back in the same collection tube. 0.5 mL of Buffer QG was added to QIAquick column and centrifuged for 1 min. To wash, 0.75 mL of Buffer PE was added to QIAquick column and centrifuged for 1 min. The flow-through was discarded and the QIAquick column was centrifuged for an additional 1 min. at 13000 rpm. QIAquick column was placed into a clean 1.5 mL microcentrifuge tube and to elute DNA, 50 μL of Buffer EB was added to center of the QIAquick membrane and centrifuged for 1 min.

2.10.1.3 Ligation reactions

Purified bands and RE digestion products were run on 1% agarose gel to understand the concentrations after purification. Purified DNA bands of cut vector or RE digestion product of pMW103 and the purified insert or RE digestion products of the pGEM-T easy vectors were combined with ratio of 1 to 4 in with a final volume of 10 μL in a PCR tube. Following components were added: 1X Ligase Buffer (MBI, Fermentas) and 6U T_4 DNA Ligase enzyme (MBI, Fermentas). Mixture was incubated at 22°C for 1 hour.

2.10.2 Constructing the bait vectors with direct cloning approach

2.10.2.1 Restriction digestion of isolated plasmids

Isolated plasmids were cut in order to clone into bait vector pMW103. pGEM-T Easy plasmids having the genes of interest (YR10-1506, YR10-1563, YR10-1, YR10-2, WR5 and Rad6) were cut using *Eco*RI and *Xho*I enzymes. 5 µL purified plasmid, 1X NE Buffer (New England Biolabs), 5U of *Eco*RI enzyme (New England Biolabs), 5U of *Xho*I enzyme (New England Biolabs) and PCR ddH_2O were combined in a PCR tube with a final volume of 10 µL. Mixture was incubated at 37°C for 3 hours.

Bait plasmid pMW103 was cut with *Eco*RI and *Sal*I enzymes, 5 µL purified plasmid, 1X NE Buffer (New England Biolabs), 5U of *Eco*RI enzyme (New England Biolabs), 5U of *Sal*I enzyme (New England Biolabs) and PCR ddH_2O were combined in a PCR tube with a final volume of 10 µL. Mixture was incubated at 37°C for 3 hours.

2.10.2.2 Ligation of plasmid and inserts

Digestion products were run on 1% agarose gel for concentration determination and RE digestion products were directly used for the ligation.

The vector and the inserts were combined with ratio of 1 to 4 in a final volume of 10 µL and following components were added: 1X Ligase Buffer (MBI, Fermentas) and 6U T_4 DNA Ligase enzyme (MBI, Fermentas). The mixture was incubated at 22°C for 1 hour.

For direct ligation of the RE digestion products, after ligation there was one more double restriction digestion with *Apa*I (which only cuts pGEM-T Easy vector but not pMW103 or inserts) and *Xho*I (to get rid of self ligated pMW103 vector). This RE digested ligation product was transformed to bacteria for selection. 5 µL ligation product, 1X NE Buffer (New England Biolabs), 5U of *Xho*RI enzyme (New England Biolabs), 5U of *Apa*I enzyme (New England Biolabs) and PCR ddH_2O were combined in a PCR tube with a final volume of 10 µL. Mixture was incubated at 37°C for 3 hours.

Ligation products (either obtained from with or without band isolation methods) were transformed to *E. coli* competent cells as described in section 2.8.3. LB Plates were incubated at 37 °C overnight. After the incubation, colonies were selected among grown colonies on Kanamycin (35 mg/mL) for pMW103 plasmids or Chloramphenicol (25 mg/mL) for pGLS23 plasmids. Plasmids were purified by using QIAGEN QIAprep Spin Miniprep Kit according to the manufacturer's protocol.

2.11 Restriction digestion of the bait vectors having the inserts

After the plasmids having the bait constructs were isolated from *E. coli* in order to check the insert size several types of digestions were performed depending on the insert size and ligation types (with or without band isolation).

The inserts (YR10-1, YR10-2, WR5 and Rad6) in the pMW103 vector were confirmed by *Eco*RI and *Pst*I double digestion. The insert YR10-1506 was confirmed by digestion with *Xba*I and YR10-1563 was confirmed by restriction digestion with *Sac*I enzyme.

The inserts were confirmed by either double restriction digestion of pGLS23-bait constructs (WR5 and Rad6) with *Bam*HI and *Xho*I or with the PCR amplification of the isolated plasmids (pGLS23/YR10-1, pGLS23/YR10-2, pGLS23/WR5 and pGLS23/Rad6) using the primers designed for the vector pGLS23.

2.12 Transformation into yeast cells

There are two protocols used to transform the yeast cells. The first one is used for the transformation of the bait vectors. This technique has less efficiency with respect to the second one but it is sufficient the bait vector transformation as a few transformed yeast cells would be enough. The second technique is much more efficient than the first one so mainly used for the library transformation of the yeast cells.

2.12.1 Technique 1 (for bait vectors transformation)

Yeast cells were inoculated using a loop and grown overnight at 30°C in 5 mL YPD. A_{600nm} of grown cells were measured and diluted with YPD to A_{600nm}: 0.5. The cells were grown until they reach to A_{600nm}: 2 (3-4 hours). The cells were collected by centrifugation at 3000 rpm. The medium was poured out and the pellet was dissolved in 1.5 mL sterile dH_2O. The cells were centrifuged at 3000 rpm pour the water and re-dissolved in 1.5 mL of water. There was three times washing step with water. For each transformation 5 mL of cell with A_{600nm}: 2 was used. So accordingly, the cells were separated into eppendorfs, centrifuged at 3000rpm and the water was poured out. Into this eppendorf, 240 µL 50% PEG3500 (w/v), 36 µL 1.0 M LiAc, 50 µL of 2.0 mg/mL single-stranded carrier DNA (Salmon Sperm DNA, Applichem) 0.1-10 µg plasmid DNA plus water up to 34 µL was added. The cells were heat shocked at 42°C for 40

min. Sterile dH_2O was added to fill the eppendorfs, centrifuged at 14000rpm for 1 min. and the supernatant was discarded. The pellet was re-dissolved in 100-200 µL sterile dH_2O. The transformed cells were plated on selective media according to the plasmids they posses.

2.12.2 Technique 2 (for library vector transformation)

Yeast cells were inoculated using a loop and grown overnight at 30°C in 5 mL YPD. On the next day, the cells were diluted A_{600nm}: 0.15. The cells were grown to A_{600nm}: 0.5- 0.7 (5-6 hours) and harvested in a 50 mL Falcon tube by spinning down at 2500 rpm for 5 min. The supernatant was decanted and cells were washed with sterile dH_2O. The cells were washed for two more times. For each transformation 5 mL of A_{600nm}: 0.7 cells were used. 5 mL of these cells were re-dissolved in 50 µL TE/Lithium Acetate solution (filter sterilized [10mM Tris-HCl pH.8.0, 1 mM EDTA, 0.1 M Lithium Acetate]). 50 µL of cells dissolved in TE/Lithium Acetate solution, 5 µL of Plasmid DNA (2 µg), 1 µL of 2.0 mg/mL single-stranded carrier DNA (Salmon Sperm DNA, Applichem) and 300 µL of TE/Lithium Acetate/PEG4000 (filter sterilized [10mM Tris-HCl pH.8.0, 1 mM EDTA, 0.1 M Lithium Acetate, 40% PEG4000]) were combined in a tube. The tube was incubated at 30°C for 30 minutes- 6hours. 40 µL of DMSO was added into the tube and cells were heat shocked at 42°C for 10 min.

Bait vector constructs are; pMW103/YR10-1506, pMW103/YR10-1563, pMW103/YR10-1, pMW103/YR10-2, pMW103/WR5, pMW103/Rad6, pGLS23/YR10-1, pGLS23/YR10-2, pGLS23/WR5, and pGLS23/Rad6.

For dual bait yeast two-hybrid studies pMW103/YR10-1506 and pGLS23/YR10-1; pMW103/YR10-1563 and pGLS23/YR10-2; pMW103/YR10-1 and

pGLS23/YR10-2; pMW103/YR10-2 and pGLS23/YR10-1; pMW103/WR5 and pGLS23/Rad6; were co-transformed with pDR8 (pLacGus) (Figure 2.4) reporter plasmid into SKY191 (MATα; *trp1*, *his3*, *ura3*, *lexAop-LEU2, cIop-Lys2*) yeast cells. For checking the self transcriptional activation of the bait plasmid, pEG202-Krit was used. pEG202-Krit is a weak self activator, meaning that although there is no interacting protein it slightly activates the reporter genes.

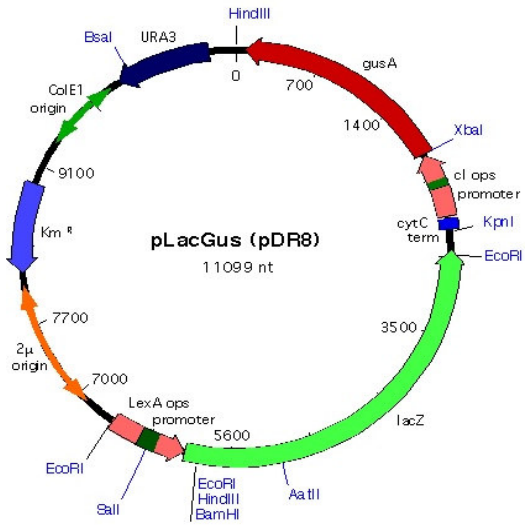


Figure 2.4 Reporter plasmid pDR8 (Invitrogen) vector circle map and sequence reference points. The plasmid has *kanamycin*^{R} gene for bacterial selection, *URA3* gene for the selection in yeast, 8 *Lex-A* and 3 *cI* operator sequences for activation tests. (*LexA*-responsive *lacZ* reporter the same as that in pMW112, while *cI*-responsive *gusA* reporter has lower background level that in pGKS8)

2.13 Characterization of the bait expression using Western Blotting

2.13.1 Preparation of SDS-PAGE gels

The glass plates were assembled properly and sealed using agar. The Resolving gel was prepared in needed concentration as in Table 2.3. The acrylamide solution was poured into the gap between the glass plates and sufficient space the stacking gel (the length of teeth of the comb plus 1cm) was left. Using a pasteur pipette, carefully the acrylamide solution was overlaid with 0.1% SDS (for gels containing ≤ 8% acrylamide) or isobutanol (for gels containing ≥ 10% acrylamide). After the polymerization was complete (30 minutes) the overlay was poured off and washed several times with deionized water to remove any unpolymerized acrylamide. Top of the gel was drained and stacking gel was prepared as in Table 2.4. The stacking gel was poured on to the polymerized resolving gel and a clean comb was inserted into the stacking gel solution.

Table 2.3 Solutions for preparing resolving gels for Tris–glycine SDS-Polyacrylamide gel electrophoresis.

Solution components	**6% gel (10mL)**	**8%gel (10mL)**	**10%gel (10mL)**	**12%gel (10mL)**	**15%gel (10mL)**
H_2O(mL)	5.3	4.6	4.0	3.3	2.3
30%Acrylamide mix (mL)	2.0	2.7	3.3	4.0	5.0
1.5 M Tris (pH 8.8) (mL)	2.5	2.5	2.5	2.5	2.5
10% SDS(mL)	0.1	0.1	0.1	0.1	0.1
10% APS(mL)	0.1	0.1	0.1	0.1	0.1
TEMED(mL)	0.008	0.006	0.004	0.004	0.004

Table 2.4 Solutions for 5% stacking gels for Tris-glycine SDS-Polyacrylamide gel electrophoresis.

Solution components For 5% Stacking Gels	**(2mL)**	**(5mL)**
H_2O(mL)	1.4	3.4
30%Acrylamide mix (mL)	0.33	0.83
1.0 M Tris(pH 6.8) (mL)	0.25	0.63
10% SDS(mL)	0.02	0.05
10% APS(mL)	0.02	0.05
TEMED(mL)	0.002	0.005

2.13.2 Sample preparation for immuno-blotting

Yeast cells having the proper plasmids were grown overnight at 30°C in the liquid minimal medium. The next day A_{600nm} values were measured and cells were diluted with minimal medium to A_{600nm}: 0.15 and grown until the absorbance reached to 0.5-0.7. Cells were collected by centrifugation at 15000rpm for 1 min., resuspended in 50 µL 2X SDS gel loading buffer (100 mM Tris-HCl pH 6.8, 200 mM dithiothreitol, 4% SDS, 0.2% bromophenol blue, 20% glycerol) and boiled for 5 min. Samples were put back to ice and spinned for 5 min. at 4°C , 13000 rpm.

2.13.3 SDS-polyacrylamide gel electrophoresis

1X Tris-glycine electrophoresis buffer (25 mM Tris, 250 mM glycine pH 8.3, 0.1% SDS) was added to the reservoirs. After preparation, 15 µL of samples were loaded into the wells. 10 µL of protein marker (BenchMark Pre-Stained Protein Ladder,

Invitrogen) is loaded to gel in order to understand the size of proteins. The gel was run at 100-115V for 1-1.5 hours.

2.13.4 Membrane blotting

Semi-wet transfer apparatus, Biometra Fastbol B33, was used to transfer the proteins to the PVDF membrane (Duralon-UV Membranes, Stratagene). 6 pieces of Whatman paper (3MM chromatography paper, Whatman) were soaked in transfer buffer (39mM, glycine, 48 mM Tris base, 0.037% SDS, 20% methanol). One piece of PVDF membrane was cut same size of the gel and soaked in methanol. On to the anode (+) of the transfer apparatus 3 pieces of pre-wet Whatman papers, pre-wet membrane, SDS-polyacrylamide gel and 3 pieces of pre-wet Whatman papers were put, in this order. Cathode of the apparatus was settled and samples were transferred at 0.65mA/sq. cm of the gel for a period of 1- 1.5 hours. When the transfer was finished, membrane was immersed in blocking buffer (PBS-Tween20 solution [1X PBS, 0.5% Tween20 (v/v)], 5% (w/v) nonfat dried milk), on a rocker for 2 hours at room temperature or over night at 4°C.

2.13.5 Immuno-blotting

The blocking buffer poured of and membrane was incubated with appropriate quantity of the primary antibody (LexA antibody, Invitrogen, 1/5000 dilution; cI antibody, Invitrogen, 1/5000 dilution; and HA antibody, Bethyl Laboratories Inc, 1/5000 dilution) diluted in blocking buffer for 1 hour. The membrane was washed with PBS-Tween20 solution 3 times for 15 min. After washing the membrane, it was incubated with appropriate amount of the secondary antibody (Immunopure Goat Anti-

Rabbit IgG (H+L) Horse Radish Peroxidase Conjugated Antibody, Pierce, 1/10000 dilution) diluted in blocking buffer for 1 hour and the membrane was washed with PBS-Tween20 solution 3 times for 15 min. The liquid was dumped and membrane was soaked in Super Signal West Pico Stable Peroxide solution 0.75 mL/Super Signal West Pico Luminol Enhancer Solution 0.75 mL mix (Pierce) for 5 min.

2.13.6 Detection of the immuno-blots

The membrane was inserted into a plastic bag, and fixed to X-Ray cassette. In the dark room the membrane was exposed to the films for 1 min., 5 min. and 30 min. The films were developed in METU Health Center. The sizes of the proteins were detected according to the protein ladder (BenchMark Pre-Stained Protein Ladder, Invitrogen).

2.14 Library transformation and characterization

The mating partner yeast strain (SKY473; MAT a *his3, leu2, trp1, ura3, lexAop-LEU2 cIop-LYS2*) was inoculated to 20 mL in liquid YPD medium overnight at 30°C on a shaker. It is important to inoculate a fresh single colony and maintain sterile conditions throughout all subsequent procedures. Overnight grown culture was diluted to an A_{600nm}: 0.15 and incubated at 30°C on a shaker until the culture has reached absorbance of 0.50- 0.70. The culture was transferred to sterile 50 mL Falcon tubes, and spinned for 5 min at 2500 rpm at room temperature. The cells were gently re-suspended in 5 mL of sterile water (for each Falcon), and all slurries were combined in one of the Falcon tubes. The sterile water was added to the top of the tube and mixed. Cells were re-spinned for 5 min at 2500 rpm for 5 min, water was poured of and cells

were re-suspended in 1.5 mL TE buffer/lithium acetate (filter sterilized [10mM Tris-HCl pH.8.0, 1mM EDTA, 0.1 M Lithium Acetate]). 30 micrograms of *S. cerevisiae* S288C library DNA which is obtained from Dr. Ilya Serebriiskii, (in pJG4-5 plasmid, Figure 3.5) and 1.5 mg of freshly denatured carrier DNA was mixed in an eppendorf tube, and added on to the yeast cells. This mixture was mixed gently by inverting the tube a number of times and aliquated into 30 microfuge tubes (~60 µL each). To each tube 300 µL TE/ Lithium Acetate/ PEG4000 (filter sterilized [10mM Tris-HCl pH.8.0, 1mM EDTA, 0.1M Lithium Acetate, 40% PEG4000]) was added and placed at 30°C for 30-60 minutes. After incubation, ~40 µL of DMSO was added to each tube, mixed by inversion and incubated at 42°C for 10 min. The contents of each tube were pipetted onto a 24x24 cm Glu/CM –Trp dropout plates and spreaded evenly using sterile glass beads, plates were inverted (glass beads were left in the plates as they will be needed to harvest the library transformants) and incubated at 30°C until colonies appear (usually 2-3 days).

After the colonies appear, 23x23 mm square (1% of plate bottom surface) was drawn and colonies were counted to calculate the transformation efficiency. A good transformation should yield approximately 20.000 to 40.000 colonies per large plate, and represent ~10^5 transformants per microgram of library DNA.

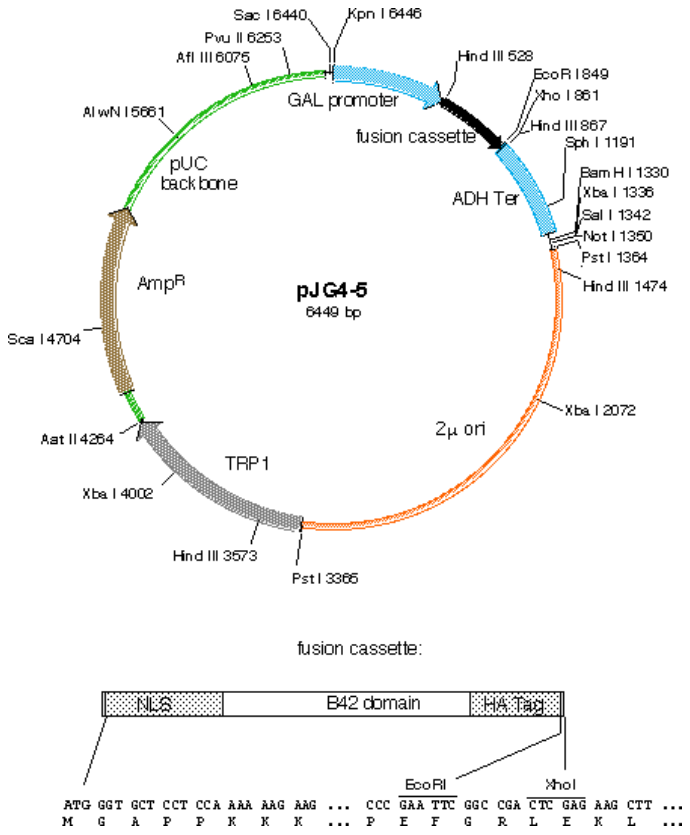


Figure 2.5 The library construction plasmid, pJG4-5 vector circle map and sequence reference points. It codes for B42 Activation Domain, HA (hemagglutinin) epitope tag, SV40 Nuclear localization signal (NLS). The plasmid contains TRP1 selectable marker and the 2µ origin of replication to allow propagation in yeast, and *ampicilin*^{R} gene and the pBR origin of replication to allow propagation in *E. coli*. Expression of the library proteins were under control of the GAL promoter.

2.15 Collecting primary transformant cells

All the plates were inspected if there was a visible contamination of mold or other contaminants. If there is contamination the contaminants must be excised using a sterile razor blade prior to beginning harvest of library transformants. 10 mL of sterile water was poured on to the plates and shaked with the help of glass beads until all colonies were re-suspended (1-2 min). Using a sterile pipette, all 5 mL of yeast slurries were collected into a falcon tube. Fill the falcon tube with sterile water and suspend the cells. Cells were spinned down at 2500g for 5 min at room temperature and supernatant was poured of. Cells were washed by adding sterile water, re-suspending and centrifuging for 5min at 2500rpm and discarding the supernatant. The washing step was repeated for a second time. After the second wash, cumulative pellet was about 25 mL derived from 1.5x 10^6 transformants. Cells were resuspended in 1 volume glycerol solution (65% v/v sterile glycerol, 0.1 M Mg SO_4, 25mM Tris-HCl pH=8.0). Cells were distributed to eppendorf tubes and stored at -70ºC. If the library is directly used for screening viability of the culture is 100%. If it is used after freezing the viability of cells can be calculated by replating them on Glu/CM –Trp dropout plates (all medium contents were represented in Appendix J) and determined from the number of colony forming units per mL.

2.16 Selection of interacting clones

2.16.1 Mating the bait strain and the pretransformed library

To 30 mL Glu/Complete minimal (CM) –Ura-His liquid culture, bait strain from the Glu/CM –Ura-His master plate was inoculated. The cells were grown at 30ºC to mid- to late log phase (A_{600nm}: 1.0- 2.0). Cells were collected by centrifugation at

2500 rpm for 5 min at room temperature. Cell pellet was re-suspended in 1 mL of sterile water and transferred to a sterile 1.5 mL microfuge tube. The A_{600nm} values of 1:100 diluted aliquot was measured to ensure that the A_{600nm} of the undiluted suspension 30-50. This corresponds to ~$1x10^9$cfu/mL. An aliquot of the pre-transformed library strain was thawed to room temperature and 200 µL of the bait stain with approximately 10^8 cells of the pre-transformed library strain. The mixture was plated on a single 100 mm diameter YPD and incubated for 12-15 hours at 30ºC. 1-2 mL sterile water was added on YPD plate; mated cells were suspended and collected by the help of glass beads. Collected cells were centrifuged at 2500 rpm for 5 min. and re-suspended in 1 volume of glycerol solution. The cells were divided into 200µL aliquots and frozen at -80ºC. Before titering, the mated cells were thawed; serial dilutions were made and plated on Glu/CM- Trp-His-Ura plates. After 2-3 days, cells were counted to determine the plating efficiency.

2.16.2 Screening for interacting proteins

One aliquot of the mated yeast cells was thawed and diluted with Gal-Raff/CM -Ura-His-Trp to an A_{600nm}: 0.15 and grown at 30ºC to absorbance value 0.5-0.7 in order to induce the GAL1 promoter on the library. 10^6 cells were spreaded on 10 (100mm plates) and 10^7 cells were spreaded 10 more plates for each bait (20 Gal-Raff/CM -Ura-His-Trp-Leu plates for pMW103 bait vectors and 20 Gal-Raff/CM -Ura-His-Trp-Lys plates for pGLS23 bait vectors). The plates were incubated at 30°C for five days. Common appearance of colonies is usually at 2-4 days after plating. Positive clones were selected and replica plated. Each single colony was picked up with a sterile toothpick and put in into a microtiter plate which contains 50-75 µL of sterile water per well. By using the fogger/replicator the cells were printed on Glu/CM -Ura-His-Trp plate and grown at 30ºC until patches form, to have a master plate.

2.16.3 First confirmation of positive interactors

The first step is testing for the galactose dependent transcriptional activation of lexAop-LEU2, cIop-LYS2, lexAop-lacZ and cIop- gusA reporters.

In dual bait system two bait constructs (LexA fusion plasmid, having the *HIS3* yeast selection marker and cI fusion plasmid having the *HIS5* yeast selection marker) and a reporter plasmid (having yeast selectable marker of *URA3*, *lacZ* under control of *LexA* operator sequence and *GusA* under control of *cI* operator sequences) were co-transformed to mating type α yeast cells. The library plasmid having the yeast selection marker *TRP1* was transformed to a type yeast cell. When these cells mate library and bait plasmids were joined in one cell. Because of the selectable markers, the transformed cells do not need the uracil, histidine and tryptophan.

The fogger was inverted on the laboratory bench, then master plate was placed upside-down on the spokes, making sure that a proper alignment of the spokes and the colonies was made. The fogger was inserted into a microtiter plate which contains 50-75μL of sterile water per well; and fogger was sit for 5-10 min. When all yeast cells were suspended, they were printed on the following plates: Glu/CM -Ura-His-Trp (3 plates: 1 for master plate, 1 for *lacZ* activation tests, 1 for *GusA* activation tests), Gal-Raff/CM -Ura-His-Trp (2 plates: 1 for *lacZ* activation tests, 1 for *GusA* activation tests), Glu/CM -Ura-His- Trp- Leu (1 plate for *LEU2* activation tests), Gal-Raff/CM -Ura-His-Trp-Leu (1 plate for *LEU2* activation tests), Glu/CM -Ura-His-Trp-Lys (1 plate for *LYS2* activation tests), and Gal-Raff/CM -Ura-His-Trp-Lys (1 plate for *LYS2* activation tests).

The plates were incubated for 3-4 days at 30ºC. After 1 day of incubation, all of the CM-Ura-His-Trp plates were taken out from the incubator and one was stored as

fresh master plate and other plates were checked for the *GusA* and *LacZ* transcriptional activation. CM -Ura-His-Trp-Leu and CM -Ura-His-Trp-Lys plates were left in the incubator and growth of cells were monitored between 48-72 hours.

2.16.4 X-Gal and X-Gluc overlays

Each plate was overlaid with chloroform gently and colonies were left under chloroform for 5 min. Left chloroform was dumped; plates were washed with chloroform and air dried for another 5 min. in the hood at 37ºC. 1% low melting agarose in 100mM $KHPO_4$, pH 7 was prepared. For each plate ~7 mL of agorose was cooled to ~60ºC, X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside, Sigma Aldrich) and X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, Sigma Aldrich) was added (to different agars according to the baits to be checked) to have a final concentration of 0.25mg/mL. Plates were overlaid with agar having the substrate (either X-Gal or X-Gluc). Cells were incubated at 30ºC and monitored for the development of the blue color. Very strong interactors will produce blue color in 20 minutes but it is good to check the color development 1-3 hours.

Positive clones were selected according to blue color formation on Gal-Raff/CM -Ura-His-Trp but not on Glu/CM -Ura-His-Trp plates (X-Gal test for pMW103-baits, X-Gluc test for pGLS23-Baits); and better growth on Gal-Raff/CM -Ura-His-Trp-Leu but not on Glu/CM -Ura-His-Trp-Leu plates (for the pMW103-baits) or Gal-Raff/CM -Ura-His-Trp-Lys but not on Glu/CM -Ura-His-Trp-Lys plates (for the pGLS23-baits).

2.16.5 Isolation of library plasmid inserts and second confirmation of positive interactions

There were two basic approaches:

i) Making PCR on colonies, sequencing the PCR fragments, transforming the bait, reporter and the interactor (homologous recombination of the library vector and PCR product) back to yeast and checking the interaction once more.

ii) Isolation of interactors plasmids from yeast, transforming this plasmid to *E. coli* in order to have high number of plasmid. Isolation of the plasmid from *E. coli* and sequencing this plasmid. Transforming the interactor, reporter and the bait plasmids to yeast cell and checking interaction once more. This method was preferred more than the 1st method because separated plasmid isolations from interacting colonies will be needed for further studies.

2.16.5.1 Isolation of plasmid from yeast cells

Yeast interactor plasmids were isolated from cells by using the Yeast Plasmid Isolation Kit (USB) according to the manufacturer's procedure. A colony was picked up and transferred to 1.5 mL of drop out medium (to maintain selection for the plasmid) in a 17x100 mm sterile, polystyrene round bottom tube. Cells were grown overnight or longer at 30°C to stationary phase. Culture was transferred to a 1.5 mL microcentrifuge tube and centrifuged at 7,500 rpm for 2–3 minutes to pellet the cells. Supernatant was removed and cells were washed with 1 mL dH2O (re-suspended and

centrifuged at 7.500 rpm for 2–3 minutes, supernatant was removed). Cells were re-suspended in 100 µL Spheroplast Buffer, 4 µl Enzyme Solution was added and incubated at 37°C for 1/2 to 2 hours. Mixture was centrifuged at 7.500 rpm for 4 minutes and supernatant was removed very carefully. Lysis Solution (150 µL) was added to the spheroplast pellet but was not re-suspended. Using the tip pellet was grasped and transferred to a fresh 0.5 mL microcentrifuge tube. The remaining lysis buffer was also transferred and pellet was re-suspended. The mixture was incubated at 95°C for 10 minutes and chilled on ice. Precipitation Solution (50 µL) was added and mixed gently by inverting the tube 3–5 times. A white fluffy precipitated and incubated on ice for 10–20 minutes. Tubes were centrifuged at 7.500 rpm for 4 minutes and supernatant was transferred to a fresh 0.5 mL microcentrifuge tube. Pellet was discarded and the supernatant was centrifuged at 14.000 rpm for 2 minutes to remove residual precipitate. To the supernatant, an equal volume of chilled 100% isopropanol was added, mixed by inversion and incubated at -80°C for at least 1 hour. Tubes were centrifuged at 14.000 rpm for 20 minutes at 4°C and supernatant was removed. The pellet was washed once with an equal volume of 70% ethanol, centrifuged at 14,000 rpm for 5 minutes, dried and re-suspended in 13 µL of dH2O.

2.16.5.2 PCR amplification of the interactors

In a 200 µL sterile PCR tube, following components were mixed ; 1 µL of isolated plasmid, 1X PCR Buffer (75 mM Tris-HCl with pH 8.8 at 25^{o}C, 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween20), 0.25 mM dNTP mix (DNA Amp), 1.5 mM $MgCl_2$ (DNA Amp), 2U *Taq* DNA polymerase, 15 pmol pJG4-5 forward primer (5'-CTG AGT GGA GAT GCC TCC-3'), 15 pmol pJG4-5 reverse primer (5'- CTG GCA AGG TAG ACA AGC CG-3') and sterile distilled water up to 30 µL volume were mixed in a 200 µL PCR tube. PCR cycling conditions were 94°C for 2 min. as initial denaturation,

35 cycles of three steps as denaturation at 94°C for 45 sec., annealing at 56°C for 45 sec. and extension at 72°C for 1 min. and 1 cycle of 5 min. extension at 72°C.

2.16.5.3 Second confirmation of the inserts

The transformed and characterized bait vectors (Section 2.12 and 2.13) were used for the preparation of competent cells as described in Section 2.11.2. All the purified plasmids of the interactors were transformed into these cells. Colonies appear in 2-3 days and these transformants were replica plated as described in Section 2.16.3. The activation tests were re-performed (Sections 2.16.3 and 2.16.4) on these cells.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 DNA isolations

DNA isolations from Pastor, Avocet, Gerek-79 and Bezostaja were performed in order to amplify the *Rar1* and *SGT1* gene fragments from wheat. The ethidium bromide stained agarose gels of Pastor, Avocet, Gerek-79 and Bezostaja are presented in Figure 3.1. The appearances of DNA samples on the agarose gels are satisfactory. We have observed the expected correlation with the intensity of the bands and the concentration values. Based on both evaluations and composition with molecular weight marker bands, the samples were diluted to 30 ng/µL and stock DNA samples and diluents were stored at –80^{o}C.

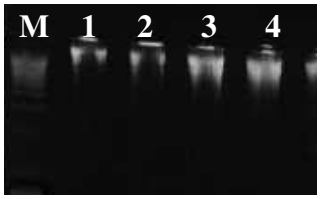


Figure 3.1 Concentration determinations of the DNA samples on 1% agarose gel. **M:** Molecular weight marker Lamda/ *Pst*I (MBI, Fermentas), **Lane 1:** Pastor-DNA (200 ng/µL), **Lane 2:** Avocet-DNA (200 ng/µL), **Lane 3:** Gerek79-DNA (400 ng/µL), **Lane 4:** Bezostaja-DNA (400 ng/µL).

3.2 Total RNA isolation from plant leaf tissue for cDNA synthesis

Total RNA isolations were performed from the Pastor, Avocet, Gerek79, Bezostaja (Figure 3.2) and Avocet-YR10 differential line plants (Figure 3.3) for the cDNA synthesis. Isolations were done using the TRIzol reagent. After determination of the concentration by spectrophotometer, the concentrations and the intactness of the RNA were verified by running RNA samples on 1% formaldehyde-agarose RNA gel. The appearances of the RNA samples on the gels and correlation with the intensity of the bands and the concentration values were satisfactory.

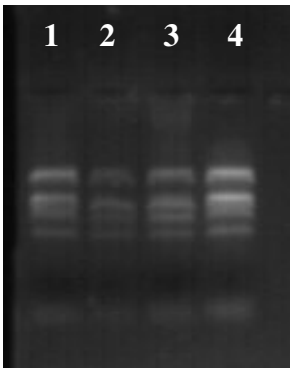


Figure 3.2 Appearances of total RNA isolates from Pastor, Avocet, Bezostaja and Gerek79 plants on 1% formaldehyde-agarose RNA gel. **Lane1:** Pastor total RNA (0.7µg/µl), **Lane 2:** Avocet total RNA (0.5µg/µl), **Lane 3:** Gerek79 total RNA (0.7µg/µl), **Lane 4:** Bezostaja total RNA (1µg/µl).

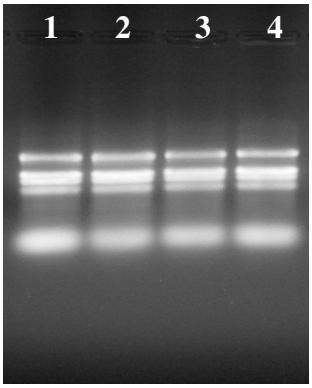


Figure 3.3 Appearances of total RNA isolates from Avocet-YR10 differential line on 1% formaldehyde-agarose RNA gel. All lanes are Avocet-YR10 total RNA with concentration of 1µg/µl.

3.3 PCR amplifications

3.3.1 PCR amplification reactions of *SGT1* gene fragments in wheat

In order to amplify the *SGT1* gene fragment in wheat, both DNA and cDNA of plant samples were used. Since this gene is not cloned in wheat before, primers were designed according to the conserved regions of barley (Accession number: AF439974) and rice (Accession number: AF192467) *SGT1* genes. The expected size of the gene fragment was 523 bp and targeted region for PCR amplification does not have any intron sequences. Thus, we expected to see same banding pattern on DNA and cDNA samples. The ethidium bromide stained 1% agarose gels of the amplifications are presented in Figure 3.4.

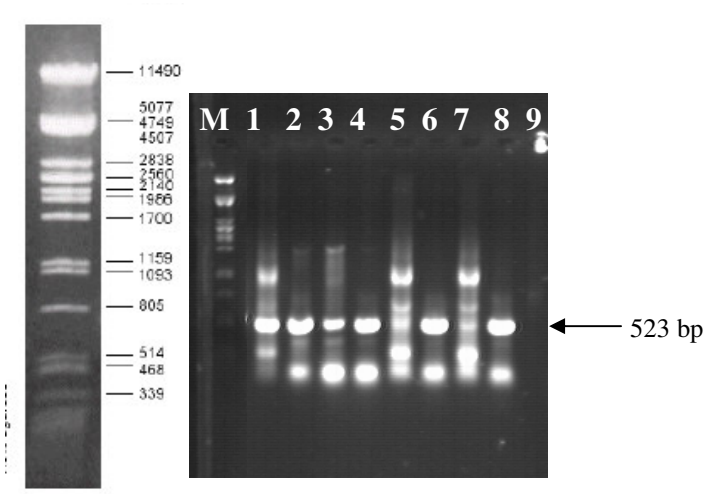


Figure 3.4 Appearances of PCR amplifications of SGT1 gene fragment (523bp) in wheat, on 1% agarose gel. The DNA template used for amplifications; **Lane 1:** Pastor-DNA, **Lane 2:** Pastor-cDNA, **Lane 3:** Avocet-DNA, **Lane 4:** Avocet-DNA, **Lane 5:** Bezostaja-DNA, **Lane 6:** Bezostaja-cDNA, **Lane 7:** Gerek79-DNA, **Lane 8**: Gerek79-cDNA. **Lane 9:** Negative control and **M:** Molecular weight marker Lamda/*Pst*I (MBI, Fermentas).

3.3.2 PCR amplification reactions of *Rar1* gene fragments in wheat

For the amplification of the *Rar1* gene fragment in wheat, we have used the degenerate primers synthesized from the sequences of barley (Accession number: AF192261) and tobacco (Accession number: AF480488) *Rar1*. The region that we amplified has intron sequence of ~100bp. We have observed the difference between DNA and cDNA amplification as presented in Figure 3.5.

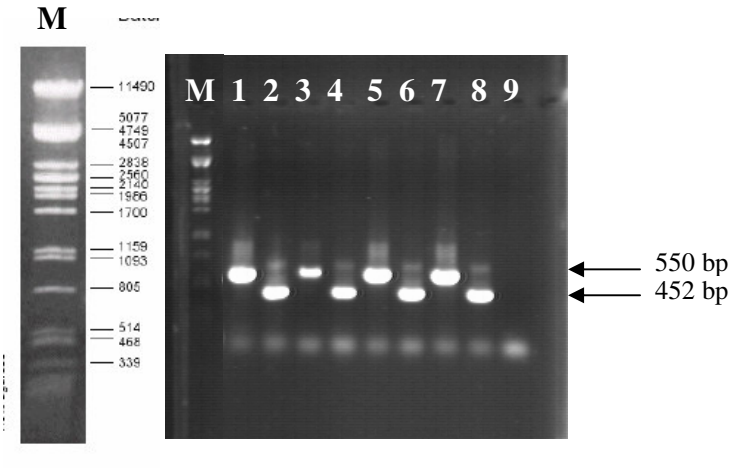


Figure 3.5 Appearances of PCR amplifications of *Rar1* gene fragments (550 bp with intron sequence, 452bp without intron sequence) in wheat, on 1% agarose gel. The DNA template for the amplifications; **Lane 1:** Pastor-DNA, **Lane 2:** Pastor-cDNA, **Lane 3:** Avocet-DNA, **Lane 4:** Avocet-DNA, **Lane 5:** Bezostaja-DNA, **Lane 6:** Bezostaja-cDNA, **Lane 7:** Gerek79-DNA, **Lane 8:** Gerek79-cDNA. **Lane 9:** Negative control and **M:** Molecular weight marker Lamda/*Pst*I (MBI, Fermentas).

For the sequencing of these gene fragments, the bands were cut from agarose gel and cloned into pGEM-T-Easy vector. For the sequencing of *SGT1* gene fragment, amplification on Avocet cDNA was cloned and for the sequencing of *Rar1* gene fragment Pastor, Gerek79 and Bezostaja cDNA amplifications were used. Sequencing results and alignments of genes are presented in Appendix A. According to the results Avocet-*SGT1* gene fragment showed 95% homology with the barley *SGT1* gene. Pastor, Gerek79 and Bezostaja *Rar1* gene fragments show 94%, 93% and 93% homology to barley *Rar1* gene, respectively.

As *SGT1* and *Rar1* genes are common and very important elements of *R* gene mediated resistance downstream signaling. They are highly conserved in plant species. There is no wheat sequences of these genes in literature, thus this is the first sequence information of these genes in wheat.

3.3.3 PCR amplification reactions of *Yr10* gene fragments (YR10-1506, YR10 1563, YR10-1, YR10-2), WR5 gene fragment and *Rad6* gene

For the amplifications of gene fragments of *Yr10* gene several primers were designed according to the sequence information of *Yr10* gene available (Accession number: AF149114). The distributions of the conserved *R* gene domains on the *Yr10* gene region are; 10-600 bp the CC region; 601-833, 1989-2106 bp the NB region; 2106-2661 bp NB spanning region (ARC Domain); and 2661-3630 bp the LLR region.

In yeast two-hybrid system, if the gene fragment is big (more than 1kb) sometimes it is difficult to detect interacting proteins. For, this reason we have used 4 different bait gene fragments (YR10-1506, YR10-1563, YR10-1 and YR10-2) for detecting the Yr10 interacting proteins.

YR10-1506 gene fragment starts from the 10^{th} base and ends with 2661^{st} base. This gene fragment encodes for the putative leucine zipper domain (10-75bp), putative leucine zipper domain (97-183bp), P-loop domain (601-630bp), domain 2 (796-833,1989bp), kinase domain 2a (829-833, 1989-2013bp), kinase domain 3a (2068-2106bp), domain 3 (2434-2460bp), domain Q/EGF (2521-2529bp), and domain HD (2656-2661bp). The YR10-1563 is from 2068^{th} base to the end of the gene. The fragment has the kinase domain 3a (2068-2106bp), domain 3 (2434-2460bp), domain

Q/EGF (2521-2529bp), domain HD (2656-2661bp) in common with YR10-1506 fragment plus the LLR domain of the *Yr10* gene.

The ethidium bromide stained 1% agarose gels of the amplifications in various annealing temperatures are presented in Figure 3.6. The annealing temperature of 55°C worked best for both of the primer sets.

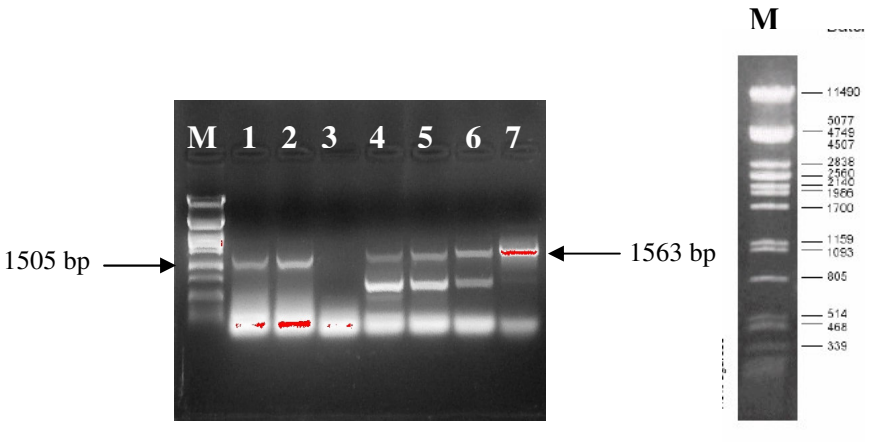


Figure 3.6 Appearances of PCR amplifications of the YR10-1505 and YR10-1563 on 1% agarose gel. **M:** Molecular weight marker Lamda/*Pst*I (MBI, Fermentas), **Lane 1:** YR10-1505 (annealing at 53°C), **Lane 2:** YR10-1505 (annealing at 55°C), **Lane 3:** YR10-1505 (annealing at 60°C), **Lane 4:** YR10-1563 (annealing at 47.7°C), **Lane 5:** YR10-1563 (annealing at 50.4°C), **Lane 6:** YR10-1563 (annealing at 53°C), **Lane 7:** YR10-1563 (annealing at 55°C).

The WR5 (267 bp) gene fragment was amplified from the plasmid cloned before (Figure 3.7) This gene fragment was found in another study in our laboratory (Osman Bozkurt, PhD thesis studies). The band was present in resistant bulk but absent in susceptible bulk. This sequence aligns with a protein having WD40 repeat, which is an important domain having function in apoptosis. Thus, in order to understand the possible role(s) of this unknown sequence, it is important to find some interacting proteins. The restriction sites *Eco*RI and *Xho*I were added by the help of PCR primers and become suitable for cloning into bait vectors for yeast two-hybrid analysis.

YR10-1 (621 bp) gene fragment was amplified from the cDNA of Avocet-YR10 plant (Figure 3.7). The annealing temperature for the amplifications was 55°C. The fragment is from 10^{th} bp to 630^{th} bp of the *Yr10* gene. YR10-1 fragment encodes for the putative leucine zipper domain (10-75bp), putative leucine zipper domain (97-183bp), and P-loop domain (601-630bp).

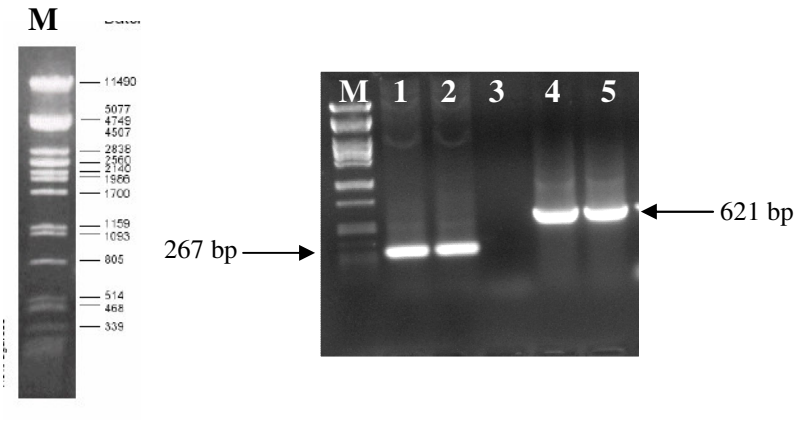


Figure 3.7 Appearances of PCR amplifications of WR5 (267 bp) and YR10-1 (621 bp) gene fragments on 1% agarose gel. **M:** Molecular weight marker Lamda/*Pst*I (MBI, Fermentas), **Lanes 1, 2:** WR5, **Lane 3:** Negative control, **Lanes 4, 5:** YR10-1.

YR10-2 (594 bp) gene fragment was amplified from the cDNA of Avocet-YR10 plant (Figure 3.8). The amplifications were performed at annealing temperature of 55°C. The gene fragment is from 2068th bp to 2661st bp of the *Yr10* gene. This gene fragment encodes for the kinase domain 3a (2068-2106 bp), domain 3 (2434-2460 bp), domain Q/EGF (2521-2529 bp), and domain HD (2656-2661 bp).

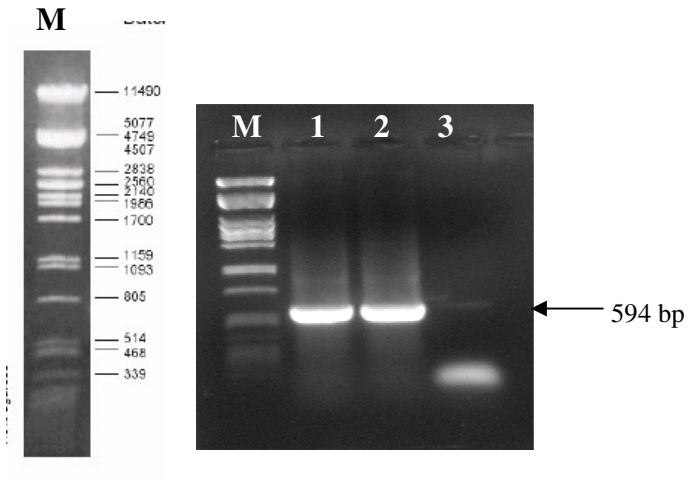


Figure 3.8 Appearances of PCR amplifications of the YR10-2 gene fragment (594bp) on 1% agarose gel. **M:** Molecular weight marker Lamda/*Pst*I (MBI, Fermentas), **Lanes 1, 2:** Yr10-2, **Lane 3:** Negative control.

Rad6 gene was found in our laboratory (Osman Bozkurt, PhD thesis study). The gene found to be induced in Avocet-YR10 plant after infection with the virulent yellow rust. *Rad6* is an ubiquitin conjugating enzyme (E2). To understand the molecular function of the *Rad6* gene in disease resistance pathway, by means of identifying interacting proteins of it, this gene was also included for yeast two-hybrid analysis. The gene was amplified from the plasmid having *Rad6* insert (Figure3.9). The amplification were performed at annealing temperature of 55°C and with the primers having *Eco*RI and *Xho*I recognition, so that it can be cloned into the bait vectors.

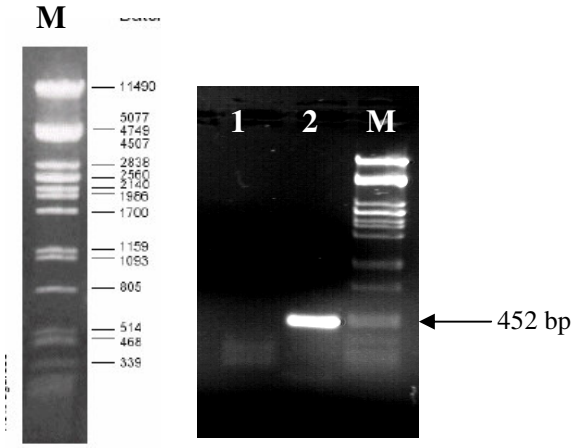


Figure 3.9 Appearance of PCR amplification of the *Rad6* gene (452 bp) on 1% agarose gel. **M:** Molecular weight marker Lamda/*Pst*I (MBI, Fermentas), **Lane 1:** Negative control, **Lane 2:** *Rad6* gene.

The PCR amplified YR10-1505, YR10-1562, YR10-1, YR10-2, WR5 gene fragments and *Rad6* gene were sliced form the 1% agarose gels and cloned into the pGEM-T Easy (Promega) vector as described in section 2.8.

3.4 Restriction digestions of the pGEM-T Easy vector having gene of interests for verification of cloning

In order to control sequence and the restriction sites inserted (*Eco*RI and *Xho*I), PCR amplifications were cloned into pGEM-T Easy vector. To confirm the insert sizes plasmids were cut with EcoRI (Figure 3.10).

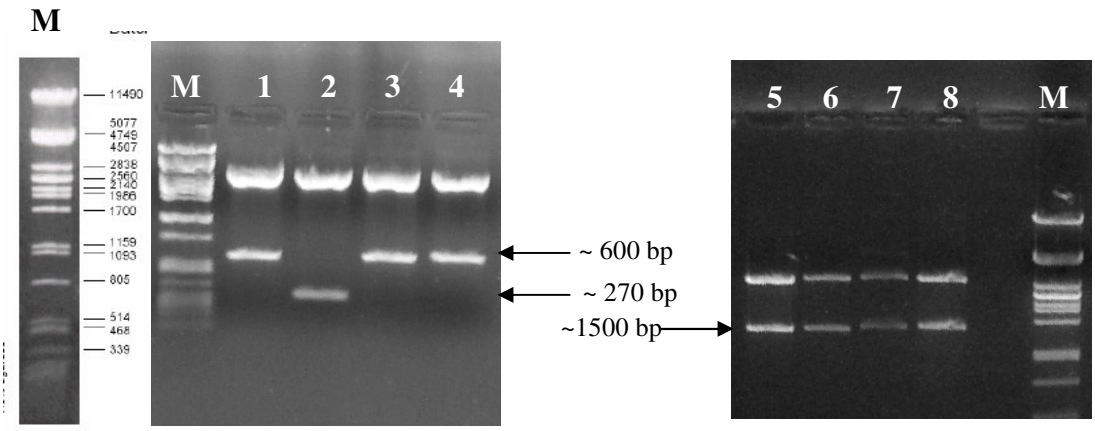


Figure 3.10 Appearances of pGEM-T Easy recombinants digested with *Eco*RI on 1% agarose gel. **M:** Molecular weight marker Lamda/*Pst*I (MBI,Fermentas), **Lane 1, 3:** pGEM-T Easy/YR10-1, **Lane 2:** pGEM-T Easy/WR5, **Lane 4:** pGEM-T Easy/YR10-2, **Lanes 5, 6:** pGEM-T Easy/YR10-1506, **Lanes 7, 8:** pGEM-T Easy/YR10-1563.

The sequences of these fragments results are presented in Appendix B. All the plasmids had the restriction sites for cloning and the coding sequences were on the correct open reading frame (ORF). In expression studies this is a critical point to be verified in order to express functional proteins. After verification of the correct ORF and restriction sites, the pGEM-T Easy recombinants were double digested with *Eco*RI and *Xho*I in order to be cloned into the bait vectors.

3.5 Cloning of the baits into pMW103 from pGEM-T Easy recombinant

In order to clone the inserts to the bait vector pMW103, pGEM-T Easy vectors having the inserts were double digested with *Eco*RI and *Xho*I (Figure 3.11).

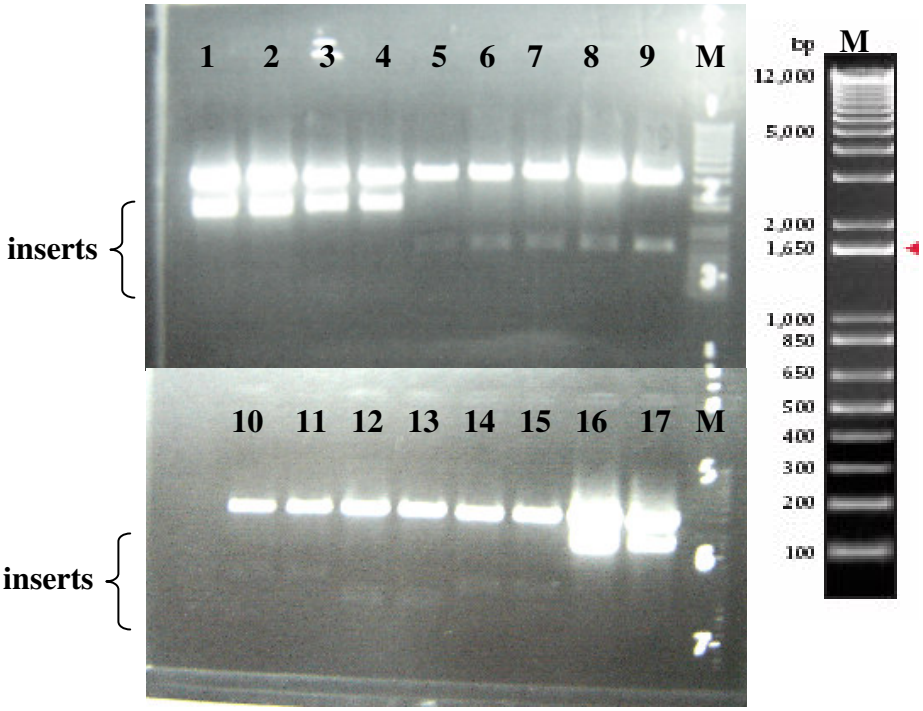


Figure 3.11 Appearances of pGEM-T Easy recombinants double digested with *Eco*RI and *Xho*I, on 1% agarose gel. **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen), **Lanes 1, 2, 16:** pGEM-T Easy/YR10-1506, **Lanes 3, 4, 17:** pGEM-T Easy/YR10-1563, **Lanes 5, 6, 10, 11:** pGEM-T Easy/YR10-1, **Lanes 7, 8, 9:** pGEM-T Easy/YR10-2, **Lanes 12, 13:** pGEM-T Easy/WR5, **Lanes 14, 15:** pGEM-T Easy/Rad6.

All of the inserts were sliced from the agarose and purified. Some of the purified samples were loaded to agarose gel for concentration determination. (Figure 3.12) In cloning experiments, insert: vector ratio is very important and must be 4:1. After determination of the insert concentration ligation reaction was performed.

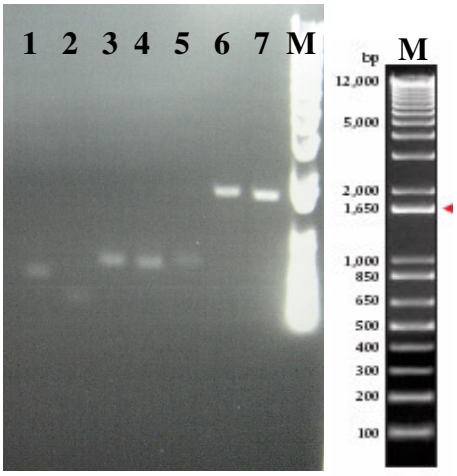


Figure 3.12 Appearances of the purified inserts on 1% agarose gel. Inserts; **Lane 1:** Rad6, **Lane 2:** WR5, **Lane 3, 4:** YR10-2, **Lane 5:** YR10-1, **Lane 6:** YR10-1563, **Lane 7:** YR10-1506, **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen).

However the ligation reaction did not work with the band isolated inserts. The reason of this might be the deteriation of the restriction sites of the vector and/or the inserts during purification steps.

An alternative strategy was used for cloning the inserts. The pMW103 with an insert of X gene fragment was double digested with *Eco*RI and *Sal*I (Figure 3.13) *Sal*I is present in the multiple cloning site of the pMW103 and produces the same sticky ends with the *Xho*I. pGEM-T Easy vector having the inserts were double digested with *Eco*RI and *Xho*I (Figure 3.13) and their concentrations were measured from the agarose gel. This digestion produces two bands, the insert and the pGEM-T Easy vector. This

product was directly ligated to *Eco*RI and *Sal*I double digested pMW103 in ratio of 4:1. There were no purification steps for either of the digestions. Thus, the risk of restriction site disruption was minimized. The possible combinations for this ligation were: (1) pMW103 having the inserts that we want to clone, (2) pMW103 having insert of X gene fragment, (3) pMW103 having the pGEM-T Easy vector as insert and (4) pGEM-T Easy vector having our inserts. As pGEM-T easy vector do not have a kanamycin resistance gene the cells having number 4 type of vector will not be able to grow on LB kanamycin plates. After ligation (before transformation) the ligation product was double digested with *Apa*I and *Xho*I. *Apa*I digests the pGEM-T Easy vector but not the pMW103 or inserts that we want to clone into. By this digestion number 3 type of vector will not be transformed at all. *Xho*I digests the pMW103 having the X gene fragment as this has the original multiple cloning site but it will not digest the plasmids having our inserts as the *Xho*I site was lost. Thus, number 2 type of vector will not be transformed, since it is cut. After this double digestion, the only remaining type of vector was the one that we want so transformed to *E. coli* and the transformants were selected on LB kanamycin plates.

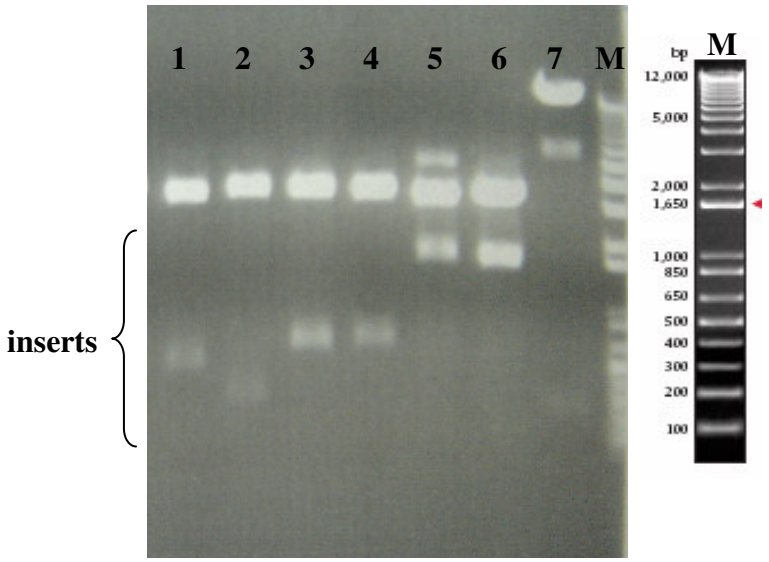


Figure 3.13 Appearances of the double restriction digestion of pGEM-T Easy recombinants with *Eco*RI plus *Xho*I, and pMW103 with *Sal*I plus *Eco*RI on 1% agarose gel. **Lane 1:** pGEM-T Easy/Rad6 (100 ng/ µL), **Lane 2:** pGEM-T Easy/WR5 (100 ng/ µL), **Lane 3:** pGEM-T Easy/YR10-2 (100 ng/ µL), **Lane 4:** pGEM-T Easy/YR10-1 (100 ng/ µL, **Lane 5:** pGEM-T Easy/YR10-1563 (100 ng/ µL), **Lane 6:** pGEM-T Easy/YR10-1506 (150 ng/ µL), **Lane 7:** pGEM-T Easy/pMW103 (having an insert X) (100 ng/ µL), **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen).

After ligation and transformation of the baits into the pMW103 vector, in order to verify the presence of the inserts in pMW103, recombinant vectors were double digested with *Eco*RI and *Pst*I (as *Eco*RI, XhoI double digestion would not work any more). The expected banding pattern of the fragments was: 5496 bp, 2500 bp, 1500 bp, 1300 bp and our insert (Figure 3.14). By this digestion, the inserts having sizes

different than 5496 bp, 2500 bp, 1500 bp and1300 bp were detected. The YR10-1506 and YR10-1563 could not be detected as the bands of plasmid and inserts (~ 1500bp) will overlap.

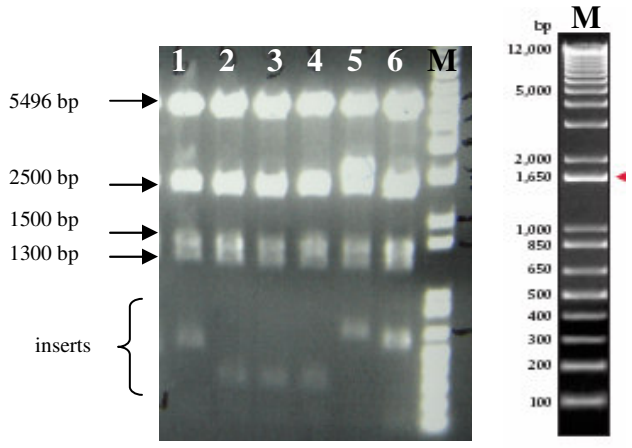


Figure 3.14 Appearances of double restriction digestion of the pMW103 recombinants with both *Eco*RI and *Pst*I, on 1% agarose gel. **Lane 1:** pMW103/Rad6, **Lanes 2, 3, 4:** pMW103/WR5, **Lane 5:** pMW103/YR10-1, **Lane 6:** pMW103/YR10-2, **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen).

In order to verify the insert YR10-1506, pMW103 recombinant was digested with *Xba*I. Expected banding pattern gives two bands with sizes 10300bp and 1800 bp (Figure 3.15). For the detection of YR10-1563, pMW103 recombinant was digested with *Sac*I. Expected banding pattern gives bands with 8500 bp, 4000 bp, and 50 bp (Figure 3.15).

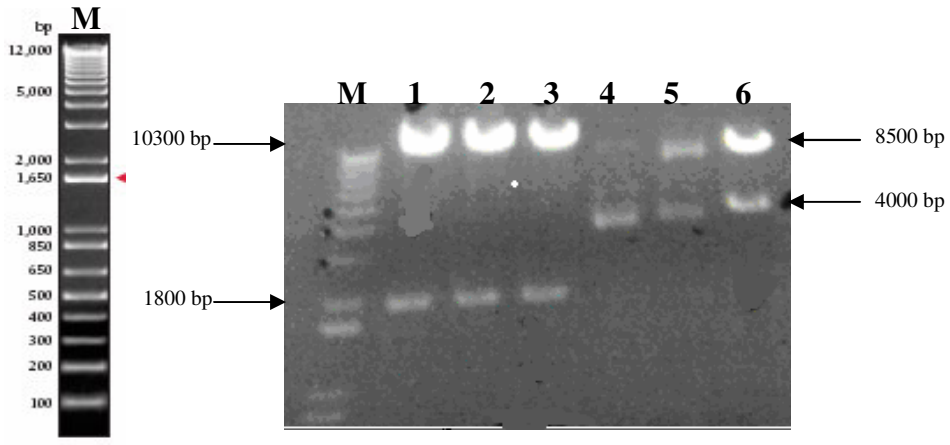


Figure 3.15 Appearances of restriction digested pMW103 recombinants on 1% agarose gel. **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen), **Lanes 1, 2, 3:** pMW103/YR10-1506 (*Xba*I digestion), **Lanes 4, 5, 6:** pMW103/YR10-1563 (*Sac*I digestion)

3.6 Cloning of the inserts into pGLS23 from pGEM-T Easy recombinants

In order to clone the baits into pGLS23, the other bait vector of the dual system, ligation after band purification method was used. This method worked in these samples so there was no need to perform the direct ligation method. Restriction digested and purified fragments (Figure 3.12) were ligated into *Eco*RI and *Xho*I double digested plasmids (Figure 3.16).

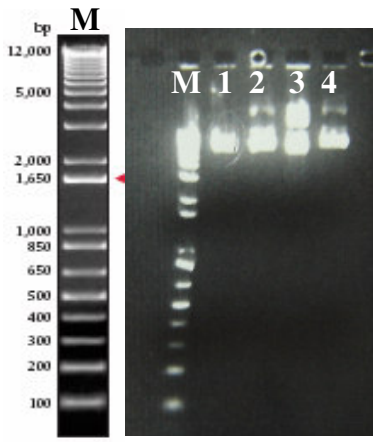


Figure 3.16 Appearances of the restriction digested pGLS23 vectors on 1% agarose gel. **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen), **Lane 1:** pGLS23 cut with *Eco*RI+*Xho*I, **Lane 2:** pGLS23 cut with *Eco*RI only, **Lane 3:** uncut pGLS23, **Lane 4:** pGLS23 cut with *Eco*RI+*Xho*I.

Cloning of the inserts was confirmed by either double restriction digestion of pGLS23-bait constructs with *Bam*HI and *Xho*I (Figure 3.17) or with the PCR amplification of the isolated plasmids (Figure 3.18).

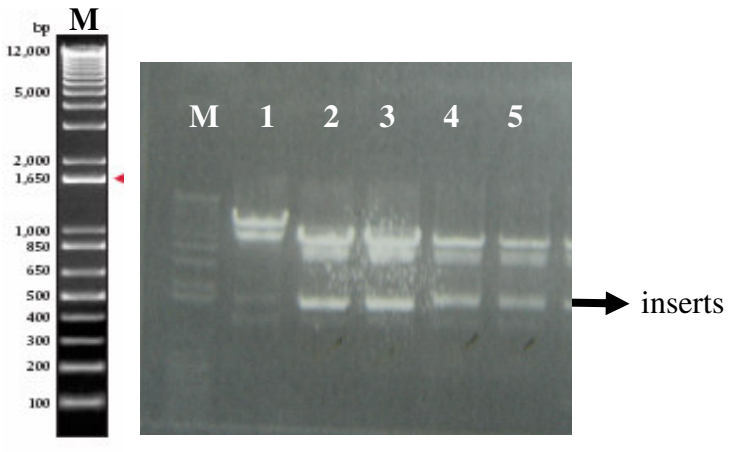


Figure 3.17 Appearances of double restriction digested of pGLS23 recombinants with *Bam*HI and *Xho*I, on 1% agarose gel. **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen), **Lane 1:** Uncut pGLS23, **Lanes 2, 3:** pGLS23/WR5, **Lanes 4, 5:** pGLS23/Rad6.

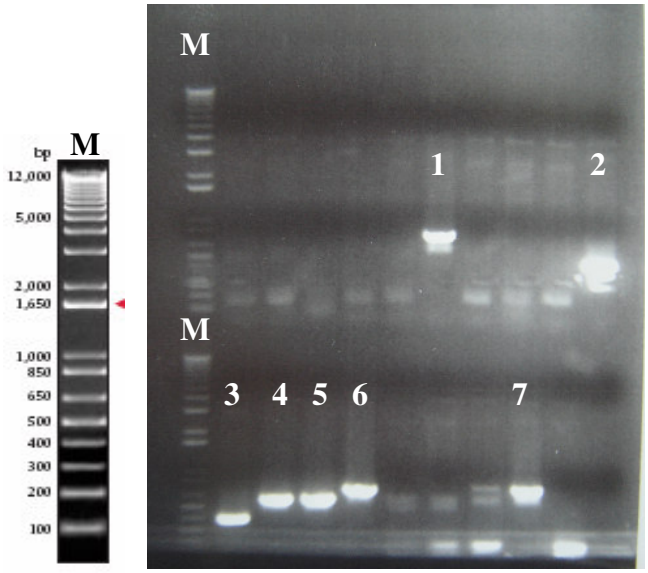


Figure 3.18 Appearances of the PCR amplifications of pGLS23 recombinant plasmids. **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen), **Lane 1:** pGLS23/YR10-1, **Lanes 2, 3:** pGLS23/WR5, **Lanes 4, 5:** pGLS23/Rad6, **Lane 6, 7:** pGLS23/YR10-2.

3.7 Characterization of the baits

For dual bait yeast two-hybrid studies pMW103/YR10-1506 and pGLS23/YR10-1; pMW103/YR10-1563 and pGLS23/YR10-2; pMW103/YR10-1 and pGLS23/YR10-2; pMW103/YR10-2 and pGLS23/YR10-1; pMW103/WR5 and pGLS23/Rad6; pMW103-Rad6 were co-transformed with pDR8 (pLacGus) reporter plasmid into SKY191 (MATα; *trp1*, *his3*, *ura3*, *lexAop-LEU2, cIop-Lys2*) yeast cells.

In choosing the method of bait construction, not only the assays require for the bait to enter the nucleus, but also the bait should be transcriptional non-activator. The baits that were used in this study do not have sequences that confer the attachment to membrane. Their transcriptional activation was also tested on their growth on leucine lacking medium (*LEU2* transcription activation test) and X-Gal overlay (*LacZ* transcription activation) for pMW103 constructs. For pGLS23 constructs, their growth on lysine lacking medium (*LYS2* transcription activation) and X-Gluc overlay (*GusA* transcription activation) were tested (Figure 3.19). For checking the self transcriptional activation of the bait plasmid, pEG202-Krit was used. pEG202-Krit is a weak self activator, meaning that although there is no interacting protein it slightly activates the reporter genes.

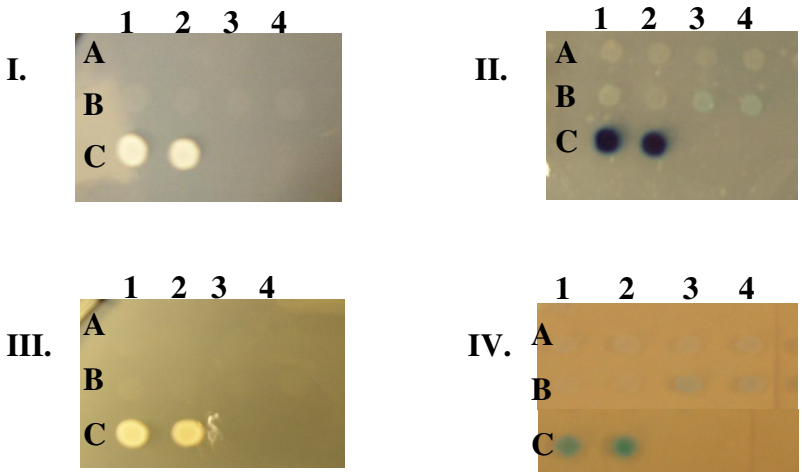


Figure 3.19 Test of transcriptional activation **I:** Test on Glu/CM –Leu medium, **II:** X-Gal overlay, **III:** Test on Glu/CM –Lys medium, **IV:** X-Gluc overlay. (Yeast cells having vector constructs; **A1:** pMW103/YR1506 and pGLS23/YR10-1, **A2:** pMW103/YR1563 and pGLS23/YR10-2, **A3:** pMW103/YR10-1 and pGLS23/YR10-2, **A4:** pMW103/YR10-2 and pGLS23/YR10-1, **B1, B2, B3 and B4:** pMW103/WR5 and pGLS23/Rad6, **C1 and C2:** positive controls of self transcription activation)

Another critical step is characterizing the bait proteins if they are detectably expressed or not. To determine the expression levels western blot analysis was performed (Section 2.13). The samples for western blot analysis were prepared as in Section 2.13.2 and were run on 12 % polyacrylamide gels. The polyacrylamide gels were membrane blotted. After immuno-blotting (the primary antibodies used were, against Lex-A and cI proteins), the membranes are exposed to films for detection. The results of the exposures are presented in Figure 3.20 and Figure 3.21. The expressions

of the bait proteins were found to be satisfactory enough to perform yeast two-hybrid analysis.

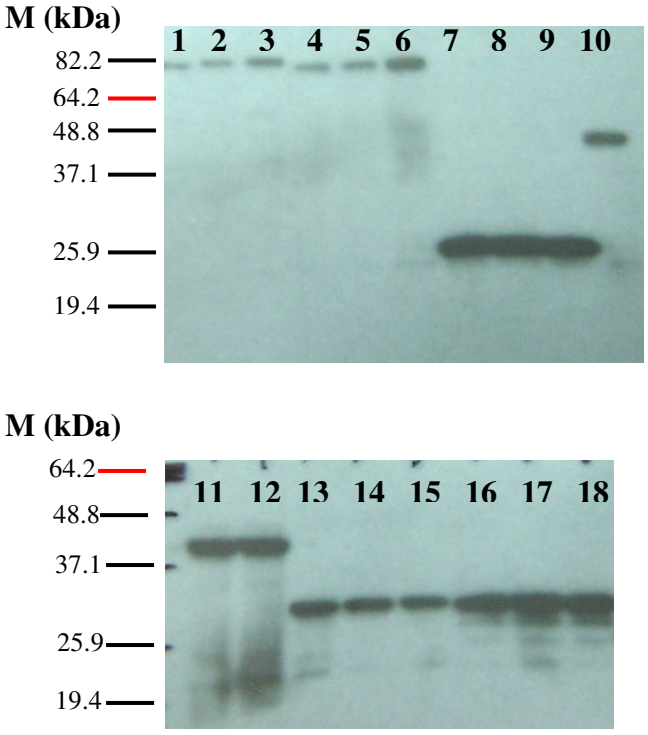


Figure 3.20 Detection of the LexA fused baits on the X-Ray film. **M:** Marker (BenchMark pre-stained protein ladder, Invitrogen), **Lanes 1, 2, 3:** LexA fused YR10-1506 protein, **Lanes 4, 5, 6:** LexAfused YR10-1563 protein, **Lanes 7, 8, 9:** LexA fused YR10-1 protein, **Lanes 10, 11, 12:** LexA fused YR10-2 protein, **Lanes 13, 14, 15:** LexA fused WR5 protein, **Lanes 16, 17, 18:** LexA fused Rad6 protein.

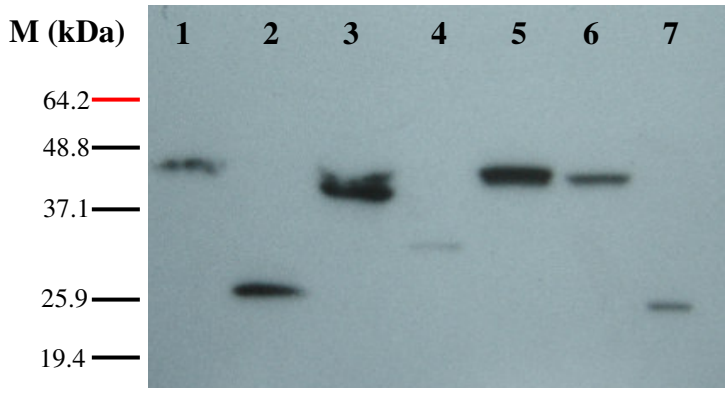


Figure 3.21 Detection of the cI fused baits on the X-Ray film. **M:** Marker BenchMark pre-stained protein ladder, Invitrogen) **Lanes 1, 6:** cI fused YR10-2 protein, **Lanes 2, 7:** cI fused YR10-1 protein, **Lane 3:** cI fused Rad6 protein, **Lane 4:** cI fused WR5 protein.

3.8 Transformation and characterization of the library

Isolated yeast prey library was loaded onto the 1% agarose gel (Figure 3.22) to verify the intactness of the library before transforming into yeast cells (SKY473; MAT a *his3, leu2, trp1, ura3, lexAop-LEU2 cIop-LYS2*).

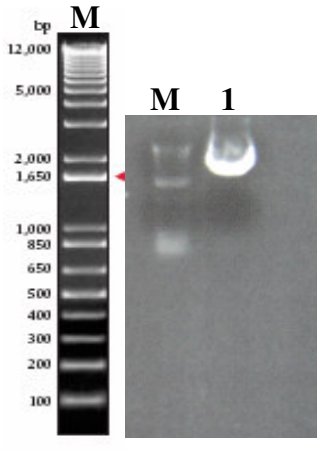


Figure 3.22 Appearance of the yeast library DNA on 1% agarose gel. **M:** Molecular weight marker (Ready-Load 1kb plus DNA ladder, Invitrogen), **Lane 1:** Yeast library DNA.

For the library transformation, mating is a satisfactory method. The main advantage of this approach is that if the same library to be screened with multiple baits, only a single large-scale transformation is required.

After the library was transformed into yeast cells as in Section 2.14, transformed cells were grown on medium lacking tryptophan. 23x23 mm square (1% of plate bottom surface) was drawn and colonies were counted to calculate the transformation efficiency (Figure 3.23). The average number on these squares was found to be 375, which makes about 37500 colonies per plate. A good transformation should yield approximately 20.000 to 40.000 colonies per large plate, and correspond

to ~10^5 transformants per microgram of library DNA. Thus, the transformation of yeast library was satisfactory enough to continue yeast two-hybrid analysis.

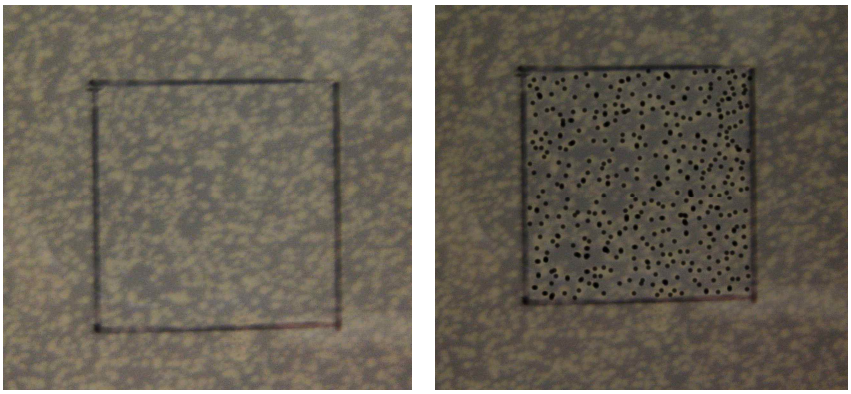


Figure 3.23 Appearances of the library yeast prey library transformed SKY473 cells on Glu/CM –Trp plates. 23X23 mm square drawn on plates and the colonies on these squares were counted to calculate the transformation efficiency.

3.9 Selection of the interactors

After characterization of the bait strain and the efficiency of the library transformation, two yeast cells were mated on YPD medium and grown overnight. During this time, individual cells of the bait strain were fused with individual cells of the library strain to form diploid cells. After mating cells, A_{600nm} values of the cells were obtained and serial dilutions of the cells were plated on Glu/CM -Trp-His-Ura plates. After 2-3 days, cells were counted to determine the mating efficiency as only mated cells can be grown on this medium. Typical mating efficiency is about 1 to 10%

range. After the mating, to select for the interactors, mated cells were diluted with Gal-Raff/CM-Ura- His- Trp to A_{600nm}: 0.15 and grown at 30ºC until the absorbance value reach to 0.5- 0.7 in order to induce the GAL1 promoter on the library. $110x10^6$ colonies were screened for each of the baits (in dual bait system there are two baits). Gal-Raff/ CM- Ura-His-Trp-Leu plates (20 plates, $110x10^6$ cells for each bait) were used for pMW103 bait vectors and Gal-Raff/ CM- Ura-His-Trp-Lys plates (20 plates, $110x10^6$ cells for each bait) for pGLS23 bait vectors. The cells were incubated for 5 days and the growing cells on selective media were collected.

- From the yeast cell having LexA-YR10-1506 and cI-YR10-1 baits, 96 cells were selected from Glu/CM Ura-His-Trp-Lys plates and 6 cells were selected from Glu/CM Ura-His-Trp-Leu plates.

- From the yeast colonies having LexA-YR10-1563 and cI-YR10-2 baits, 18 colonies were selected from Glu/CM Ura-His-Trp-Lys plates and 90 colonies were selected from Glu/CM Ura-His-Trp-Leu plates.

- From the yeast colonies having LexA-YR10-1 and cI-YR10-2 baits, 96 colonies were selected from Glu/CM Ura-His-Trp-Lys plates and 96 colonies were selected from Glu/CM Ura-His-Trp-Leu plates.

- From the yeast colonies having LexA-YR10-2 and cI-YR10-1 baits, 96 colonies were selected from Glu/CM Ura-His-Trp-Lys plates and 57 colonies were selected from Glu/CM Ura-His-Trp-Leu plates.

- From the yeast colonies having LexA-WR5 and cI-Rad6 baits, 96 colonies were selected from Glu/CM Ura-His-Trp-Lys plates and 96 colonies were selected from Glu/CM Ura-His-Trp-Leu plates.

All of the colonies were replica plated on:

- Glu/ CM-Ura-His-Trp (3 plates: 1 for master plate, 1 for *LacZ*, 1 for *GusA* transcriptional activation)
- Gal-Raff/CM -Ura-His-Trp (2 plates: 1 for *LacZ*, 1 for *GusA* transcriptional activation)
- Glu/CM -Ura-His-Trp- Leu (1 plate for *LEU2* transcriptional activation)
- Gal-Raff/CM -Ura-His-Trp-Leu (1 plate for *LEU2* transcriptional activation)
- Glu/CM -Ura-His-Trp-Lys (1 plate for *LYS2* transcriptional activation)
- Gal-Raff/CM -Ura-His-Trp-Lys (1 plate for *LYS2* transcriptional activation)

The plates were incubated for 3-4 days at 30ºC. After 1 day of incubation, all of the -Ura-His-Trp plates were taken out from the incubator and one was stored as fresh master plate and other plates were checked for the *GusA* and *LacZ* transcription activation. -Ura-His- Trp- Leu and Ura-His- Trp- Lys plates were left in the incubator to monitor the growth of the cells for more days (48-72 hours).

The results of growth differences in –Leu and –Lys plates, X-Gal and X-Gluc overlays are presented in Appendix I.

3.10 Nomenclature of the selected colonies

The first number presents the baits that the cell poses. Number **1** stands for the baits LexA/YR10-1506 and cI/YR10-1, number **2** stands for the LexA/YR10-1563 and cI/YR10-2, number **3** stands for the baits LexA/YR10-1 and cI/YR10-2, number **4**

stands for the baits LexA/YR10-1 and cI/YR10-2, and number **5** stands for the baits LexA/WR5 and cI-Rad6.

The letter and the number stand for the type and number of the plate, respectively. L2 means; cells were first collected from Glu/CM Ura-His-Trp-Leu (L) plates and the plate number is 2. K1 means the cells were collected from Glu/CM Ura-His-Trp- Lys (K) and the plate number is 1.

1K2 totally means, cell has the baits number 1 (LexA/YR10-1506 and cI/YR10-1) and collected from Glu/CM Ura-His-Trp-Lys plates, and this is the 2nd plate of this type.

The letter and number after the "-" sign, present the cells' position on plate.

1**K2-**G6: **1** stands for the baits LexA/YR10-1506 and cI/YR10-1, **K2** stands for the first positive clones selected were from Glu/CM Ura-His-Trp-Lys (K) plate number 2, G6 is the position of the clone on plates (Figure 3.24).

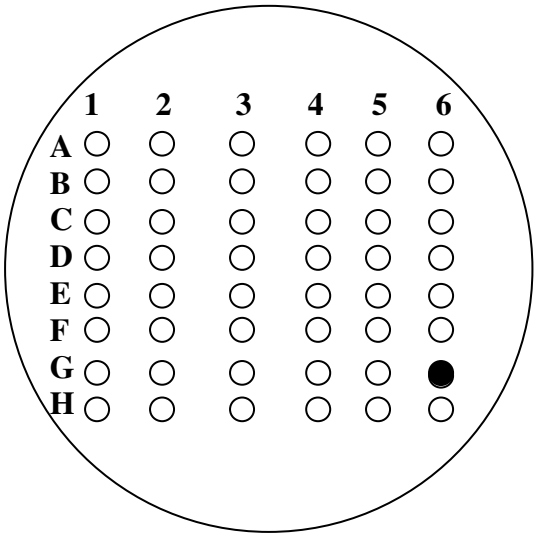


Figure 3.24 Position of the example colony on the plate.

For only 2KL2 plates positions from A1-A6 were selected from the Glu/CM Ura-His-Trp-Leu, the rest of the cells (B1-H6) were initially selected from Glu/CM Ura-His-Trp-Lys plates.

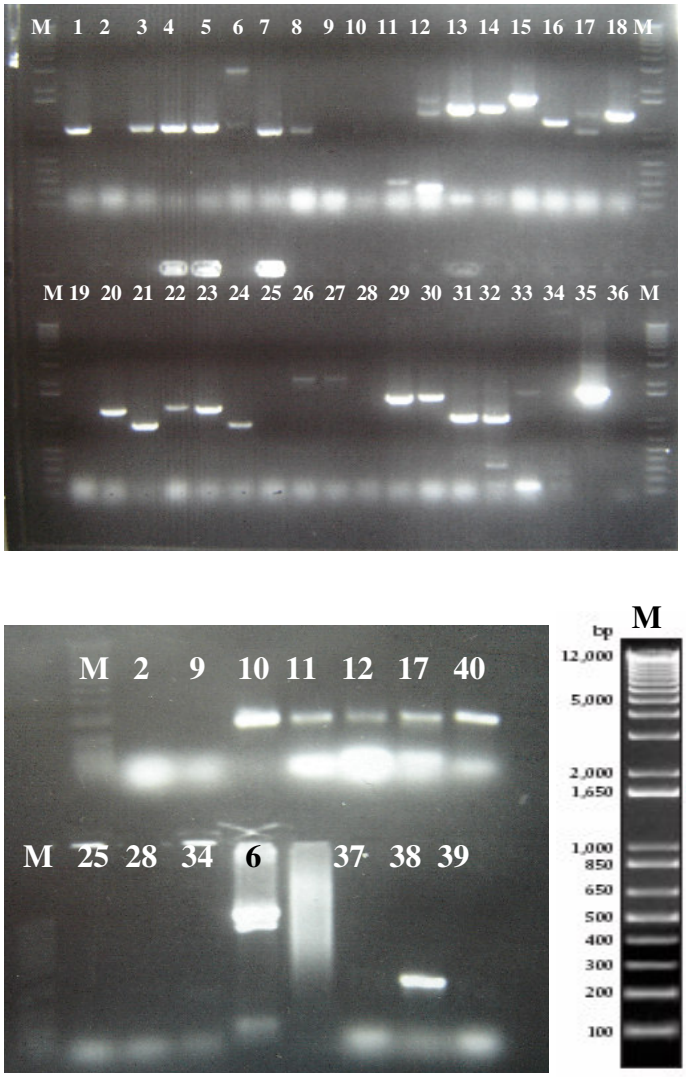
All colonies were named according to this style so that none of them can mix up. Some examples of colony nomenclature are presented in Table 3.1.

Table 3.1 Examples of the colony nomenclature

Baits	For nomenclature	Selective medium	For nomenclature	Plate number	Colony position	Nomenclature
LexA/YR10-1506 cI/YR10-1	1	Glu/CM -Ura-His-Trp-**Leu**	L	1	H1	1L1-H1
LexA/YR10-1506 cI/YR10-1	1	Glu/CM -Ura-His-Trp-**Lys**	K	2	G6	1K2- G6
LexA/YR10-1563 cI/YR10-2	2	Glu/CM -Ura-His-Trp-**Leu**	L	2	A6	2L2-A6
LexA/YR10-1563 cI/YR10-2	2	Glu/CM -Ura-His-Trp-**Lys**	K	1	C5	2K1-C5
LexA/YR10-1 cI/YR10-2	3	Glu/CM -Ura-His-Trp-**Leu**	L	1	D3	3L1-D3
LexA/YR10-1 cI/YR10-2	3	Glu/CM -Ura-His-Trp-**Lys**	K	1	E2	3K1-E2
LexA/YR10-2 cI/YR10-1	4	Glu/CM -Ura-His-Trp-**Leu**	L	2	F4	4L2-F4
LexA/YR10-2 cI/YR10-1	4	Glu/CM -Ura-His-Trp-**Lys**	K	2	B2	4K2-B2
LexA/WR5 cI-Rad6	5	Glu/CM -Ura-His-Trp-**Leu**	L	2	C3	5L2-C3
LexA/WR5 cI-Rad6	5	Glu/CM -Ura-His-Trp-**Lys**	K	1	G6	5K1-G6

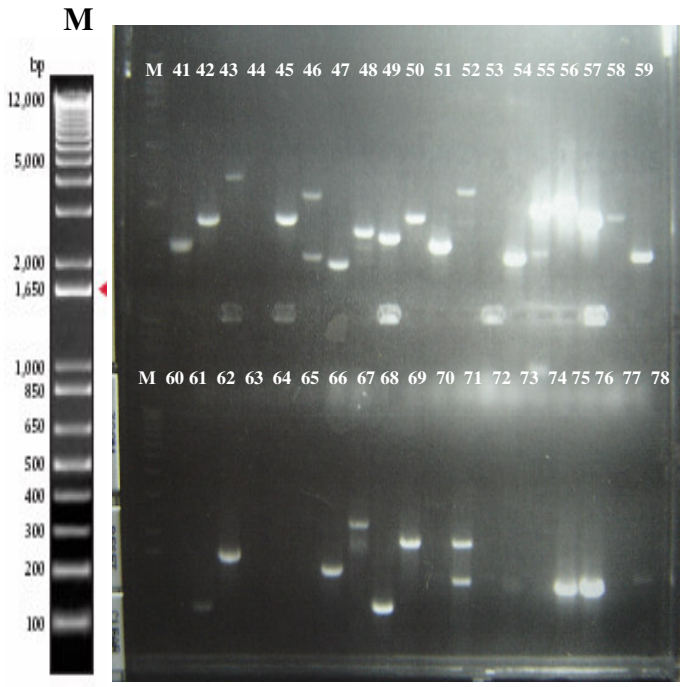
3.11 PCR amplification reactions on selected colonies

According to the transcription activation levels, the most promising interactions were selected and PCR was performed on 78 interactions using the primers of the library plasmid pJG4-5 (Figure 3.25). The PCR products were visualized on 1% agarose gel.



Lane	Colony name
1	1L1-A1
2	1L1-A3
3	2L1-A1
4	2L1-A4
5	2L1-A6
6	1K1-F3
7	1K1-F5
8	3L1-A3
9	3L1-B6
10	3L1-E3
11	3L1-G1
12	4K1-A6
13	4K1-B2
14	4K1-C1
15	4K1-D2
16	4K1-F5
17	4K1-G1
18	4K1-G2
19	4K1-G3
20	2K1-A1
21	2K1-A6
22	5L1-A6
23	5L1-C1
24	5L1-D3
25	5L1-E1
26	5L1-E5
27	5L1-G2
28	5L1-G3
29	5L1-A1
30	5L1-B1
31	5K1-A1
32	5K1-C2
33	5K1D3
34	pJG4-5/BRAF (PC)
35	pJG4-5/Krit2A(PC)
36	NC
37	1K2-D4
38	1K2D5
39	1K2E3
40	1K2F1

Figure 3.25 Appearances of the PCR amplifications of the colonies with positive signals on 1% agarose gel. NC: negative control, PC: Positive control.



Lane	Colony name	Lane	Colony name
41	1K2-F5	**60**	5L2-H1
42	4L2-B4	**61**	5L2-A1
43	4L2-D1	**62**	5L2-F1
44	4L2-E1	**63**	5L2-D4
45	4L2-F5	**64**	5L2-C2
46	4L2-G1	**65**	5L2-E2
47	2KL2-B1	**66**	5L2-G6
48	2KL2-D6	**67**	5L2-C3
49	2KL2-F3	**68**	5L2-G4
50	2KL2-G6	**69**	5K2-D2
51	2KL2-H1	**70**	5K2-D6
52	2KL2-H3	**71**	5K2-E3
53	3L2-A5	**72**	5K2-F6
54	3L2-B4	**73**	5K2-G1
55	3L2-F1	**74**	5K2-H1
56	3L2-G5	**75**	5K2-G3
57	4L2-B4	**76**	5K2-G4
58	5L2-F5	**77**	5K2-G6
59	5L2-G5	**78**	NC

Figure 3.25 (continues) Appearances of the PCR amplifications of the colonies with positive signals on 1% agarose gel. NC: Negative control.

From the PCR amplifications the inserts having different sizes were mainly selected for sequencing (to minimize the risk of sequencing the same fragment). Total of 39 bands were sequenced. The sequenced colonies were; 1K1-F3, 1K1-F5, 4K1-B2, 4K1-C1, 1L1-A1, 4K1-D2, 2L1-A1, 4K1-G1 (upper band), 2L1-A4, 4K1-G2, 2L1-A6, 2K1-A1, 2K1-A6, 5L1-A6, 5L1-C1, 5L1-A1, 5L1-B1, 5K1-A1, 1K2-D5, 1K2-F1, 1K2-F5, 2KL2-H3, 4L2-D1, 3L2-B4, 4L2-F5, 3L2-F1, 2KL2-B1, 3L2-G5, 2KL2-D6, 4L2-B4, 2KL2-F3, 2KL2-G6, 5L2-G5, 2KL2-H1, 5K2-D2, 5K2-E3/1 (upper band), 5K2-E3/2 (lower band), 5K2G3, and 5K2G4.

Sequencing results are presented in Appendix C. Obtained sequences were searched in the data base for alignments and the alignments found are presented in Appendix D.

3.12 Appearances of the sequenced interacting colonies on appropriate growth and analysis conditions

3.12.1 YR10-1505 and YR10-1563 interactors

There were very few colonies grown on 1L1 (Figure 3.30) and 2L1 (Figure 3.31) plates. The reason for this might be that these gene fragments were big in size. According to the way that the proteins may be folded, interactions might be limited in the cell. Although the cells that were sequenced for interaction did not show better growth on Gal-Raff/CM- Ura-His-Trp-Leu plates, three of the colonies (1L1-A1, 2L1-A1,and 2L1-A4) were sequenced and all three was the same sequence “Extracellular matrix protein 9” (Appendix D). As the cells showed no growth difference (Figure 3.26 and Figure 3.27), it is expected that these might be some kind of contamination.

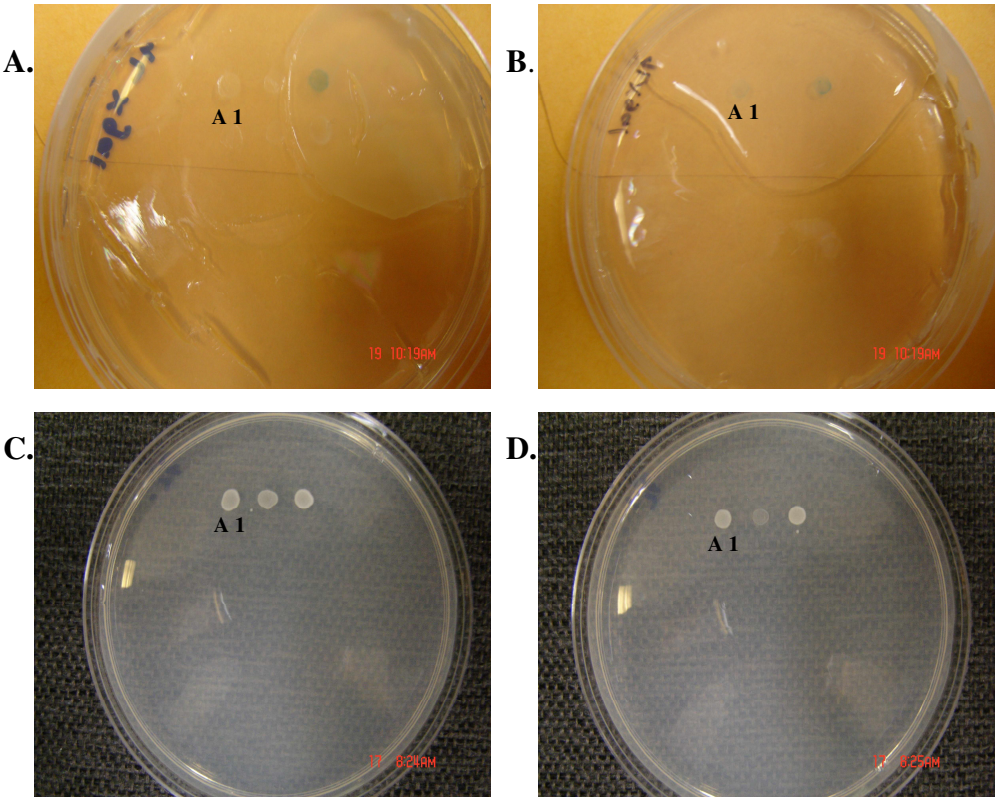


Figure 3.26 *LEU2* and *LacZ* transcription activation tests of 1L1 plate. **A:** X-Gal overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gal overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM -Ura-His-Trp-Leu medium **D:** Growth of cells on Gal-Raff/CM- Ura-His-Trp-Leu medium

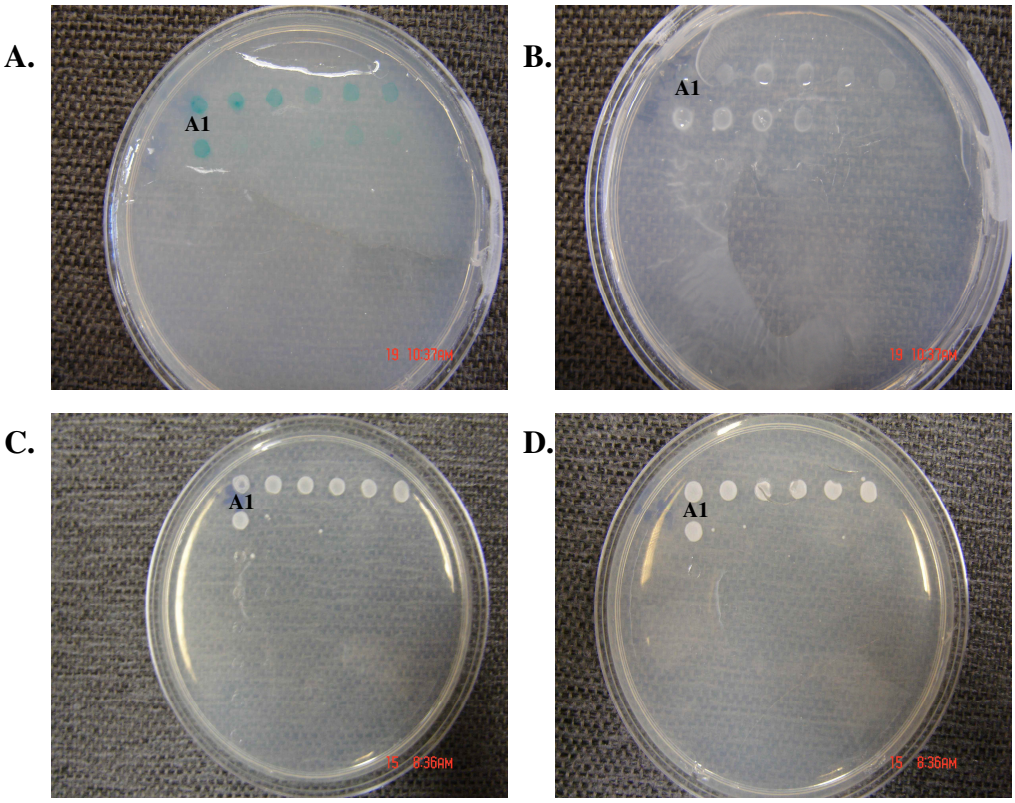


Figure 3.27 *LEU2* and *LacZ* transcription activation tests of 2L1 plate. **A:** X-Gal overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gal overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM -Ura-His-Trp-Leu medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Leu medium

3.11.2 YR10-1 interactors

Total of 13 interacting proteins with YR10-2 were selected for sequencing (Appendix E).

The promising interactors of the YR10-1 on the 1K1 plate were F3 and F5 positions which were yeast genes *SRP72* and *LAS17*, respectively. Although many cells showed back ground activity of β-glucuronidase on glucose plates (Figure 3.28 A), only the true interactors showed activity on the Gal-Raff medium (Figure 3.28 B) The cells having *GusA* transcription activation on Gal-Raff plates, were also better grown on Gal-Raff/CM- Ura-His-Trp-Lys plates (Figure 3.28 D) than Glu/CM- Ura-His-Trp-Lys plates (the control of *LYS2* transcription activation) (Figure 3.28 C).

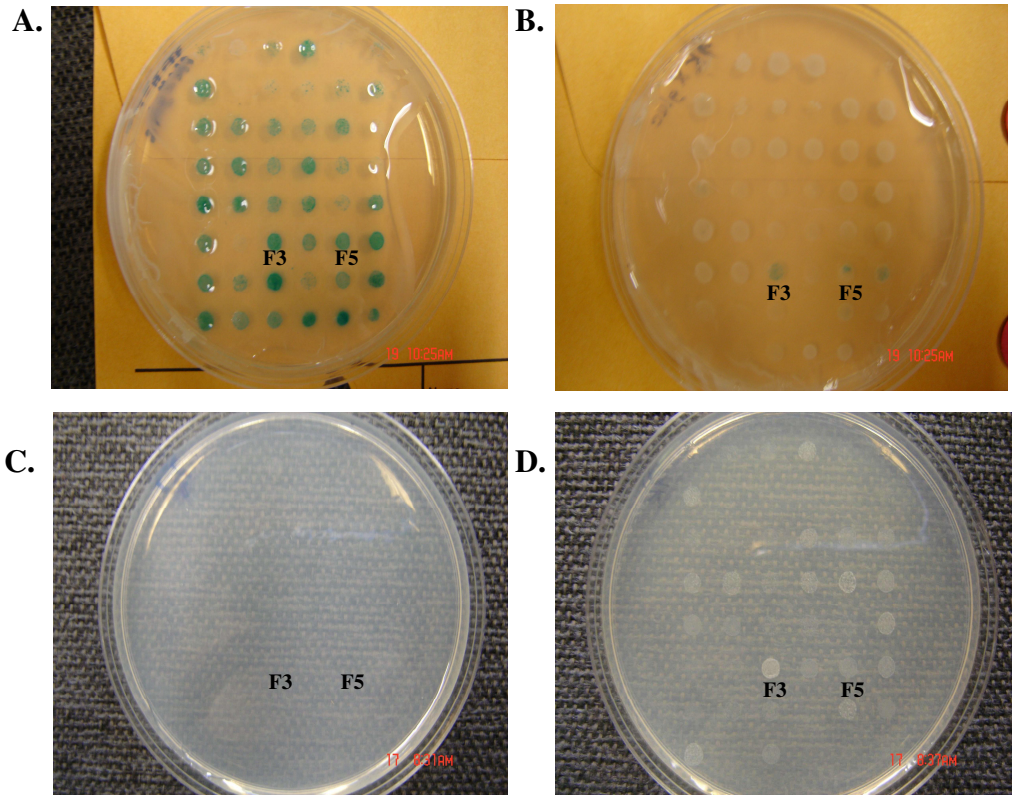


Figure 3.28 *LYS2* and *GusA* transcription activation tests of 1K1 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

The sequenced interactors of the YR10-1 on the 4K1 plate were positions B2, C1, D2, G1 and G2. The library plasmids in these colonies code for yeast open reading frames GAC1, TOM1, AZF1, TyB and YPL207w, respectively. There was background activity of β-glucuronidase on glucose plate (Figure 3.29 A) but this background

activity was lost on Gal-Raff plate (Figure 3.29 B). Selected colonies for sequencing also showed better *LYS2* transcription activation on Gal-Raff/CM- Ura-His-Trp-Lys (Figure 3.29 D) plate than on Glu/CM- Ura-His-Trp-Lys plate (Figure 3.29 C).

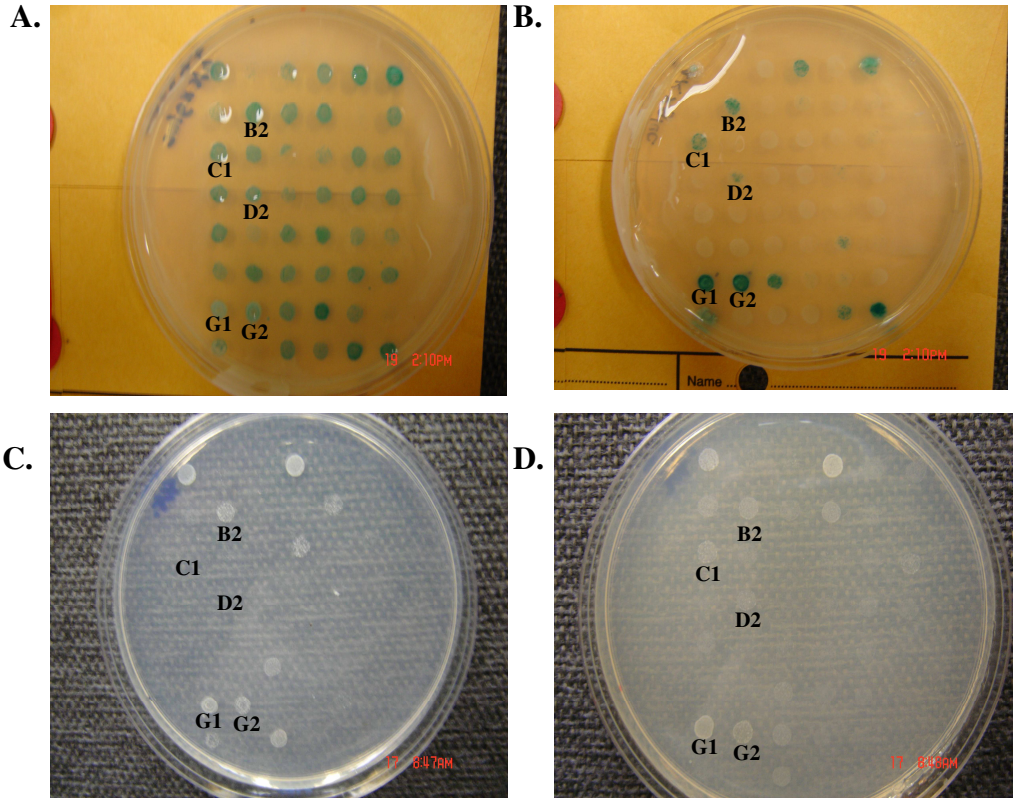


Figure 3.29 *LYS2* and *GusA* transcription activation tests of 4K1 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

The sequenced interactors of the YR10-1 on the 4L2 plate were positions B4, F5, and D1. The library plasmids in these colonies code for yeast open reading frames of CPR7, PYC2, and J protein type 2, respectively. There was again background activity of β-glucuronidase on glucose plates (Figure 3.30 A), but it was lost on Gal-Raff plate (Figure 3.30 B). The sequenced interactors showed better transcription activation of *LYS2* on Gal-Raff/CM- Ura-His-Trp-Lys plate (Figure 3.30 C) than Glu/CM- Ura-His-Trp-Lys plate (Figure 3.30 D).

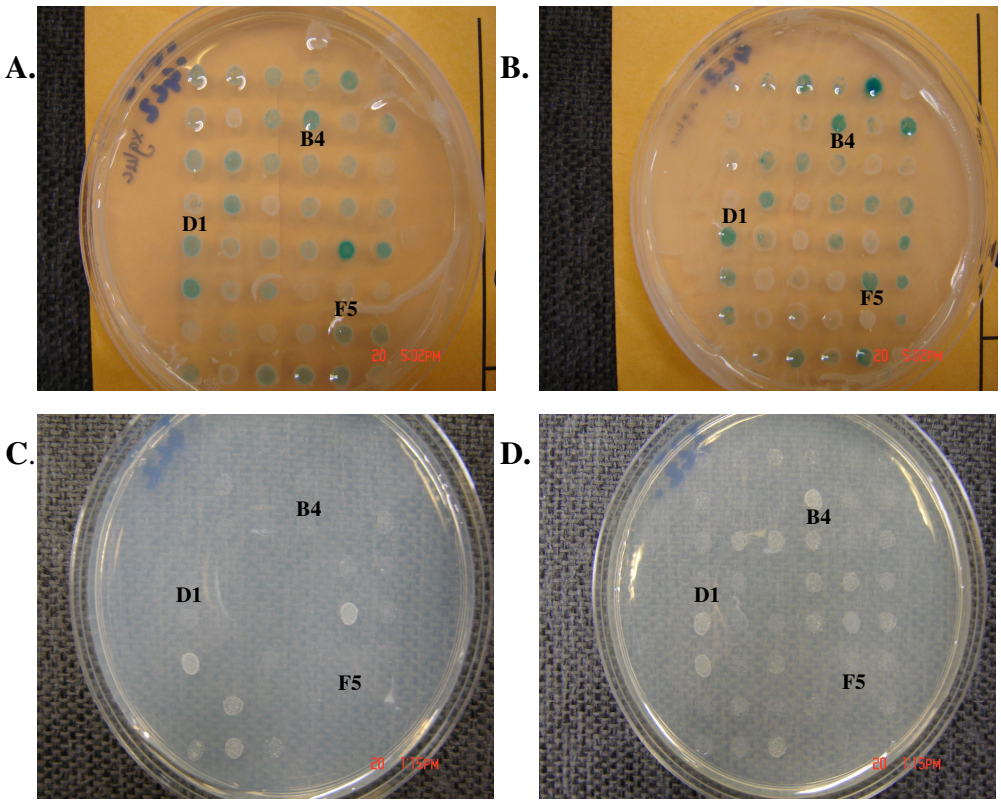


Figure 3.30 *LYS2* and *GusA* transcription activation tests of 4L2 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM -Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

The sequenced interactors of the YR10-1 on the 1K2 plate are the positions D5, F1 and F5. The library plasmids in these colonies code for yeast open reading frames of YKR 106w, Yll066cp and PYC2, respectively. The background activity of β-glucuronidase on glucose plates (Figure 3.31 A), was lost on Gal-Raff plate (Figure 3.31 B). The colonies having sequenced interactors were better grown on Gal-

Raff/CM- Ura-His-Trp-Lys plate (Figure 3.31 C) than Glu/CM- Ura-His-Trp-Lys plate (Figure 3.31 D).

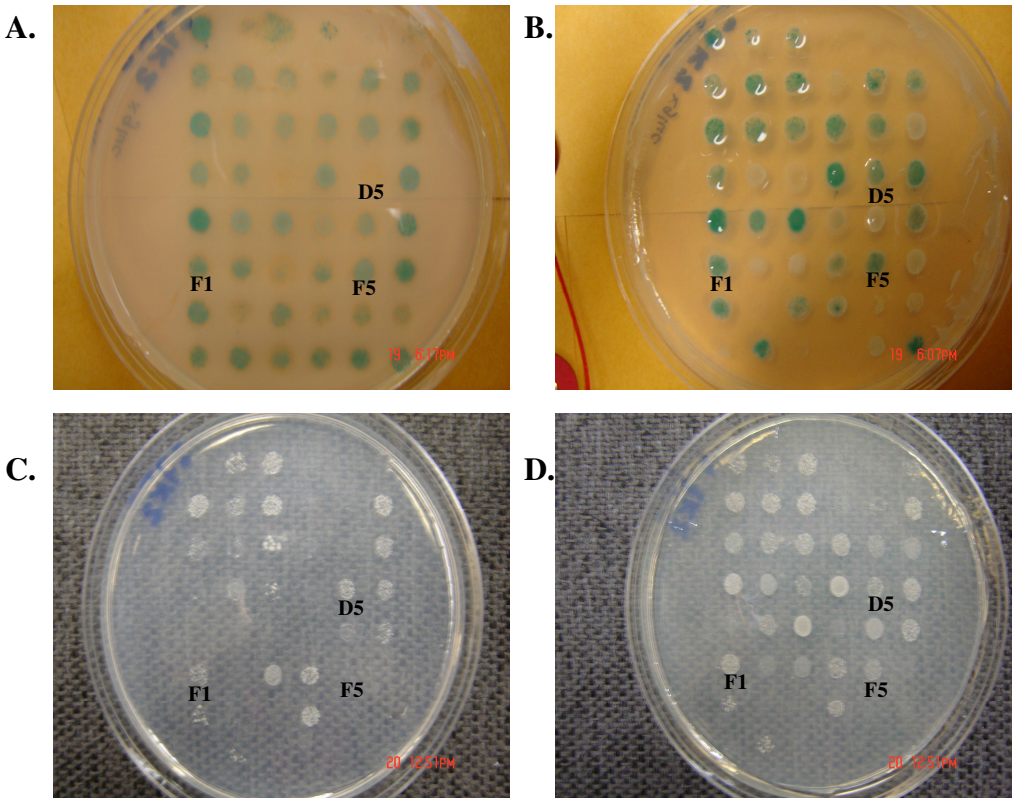


Figure 3.31 *LYS2* and *GusA* transcription activation tests of 1K2 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

3.12.3 YR10-2 interactors

Total of 9 YR10-2 interactors were sequenced (AppendixF).

On 2K1 plate, selected colonies for sequencing were positions A1 and A6. The library plasmids in these colonies code for yeast open reading frames of VPS35 and TyB, respectively. These two were the only blue colonies as a result of the X-Gluc overlay on Gal-Raff plate (Figure 3.32 B). The background activity of β-glucuronidase on the glucose plate (Figure 3.32 A), was lost on Gal-Raff plate (Figure 3.32 B). The selected colonies were slightly better grown on Gal-Raff/CM- Ura-His-Trp-Lys plate (Figure 3.32 D) than Glu/CM- Ura-His-Trp-Lys plate (Figure 3.32 C).

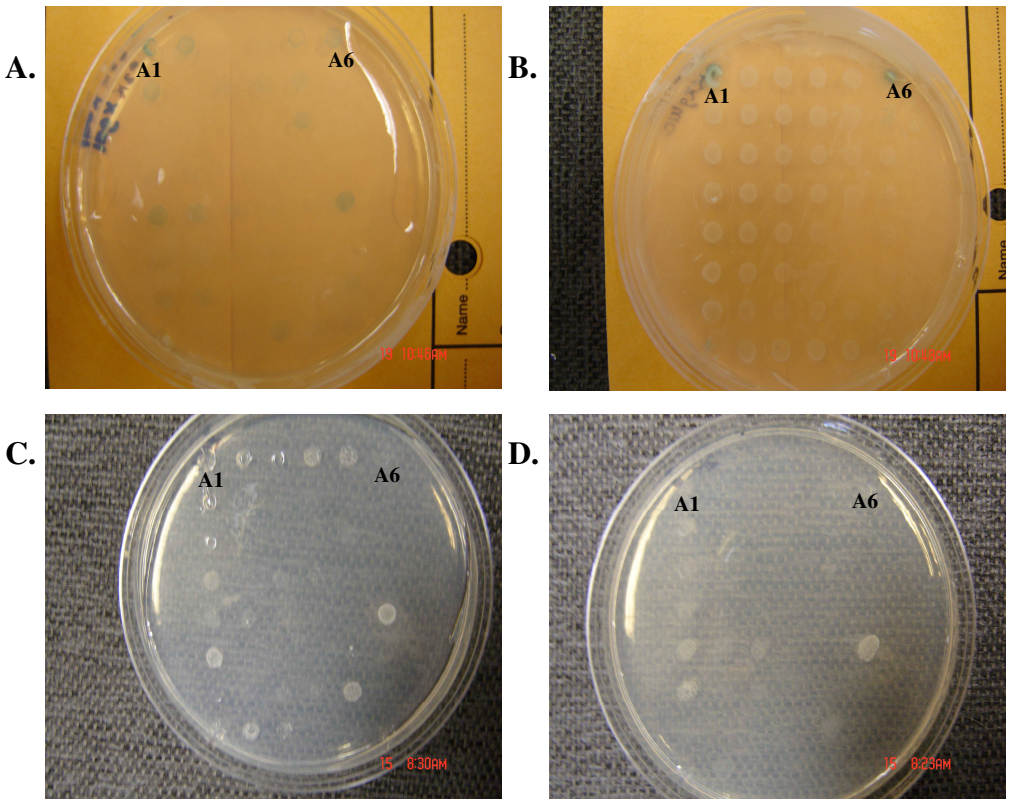


Figure 3.32 *LYS2* and *GusA* transcription activation tests of 2K1 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

The colonies selected for sequencing on 3L2 plate were positions B4, F1 and G5. The library plasmids in these colonies code for yeast open reading frames of YLR162w, GCN2, and NCP1, respectively. The selected cells both showed better transcription activation of *GusA* and *LYS2* on Gal-Raff plates (Figure 3.33 B and D) than glucose plates (Figure 3.33 A and C).

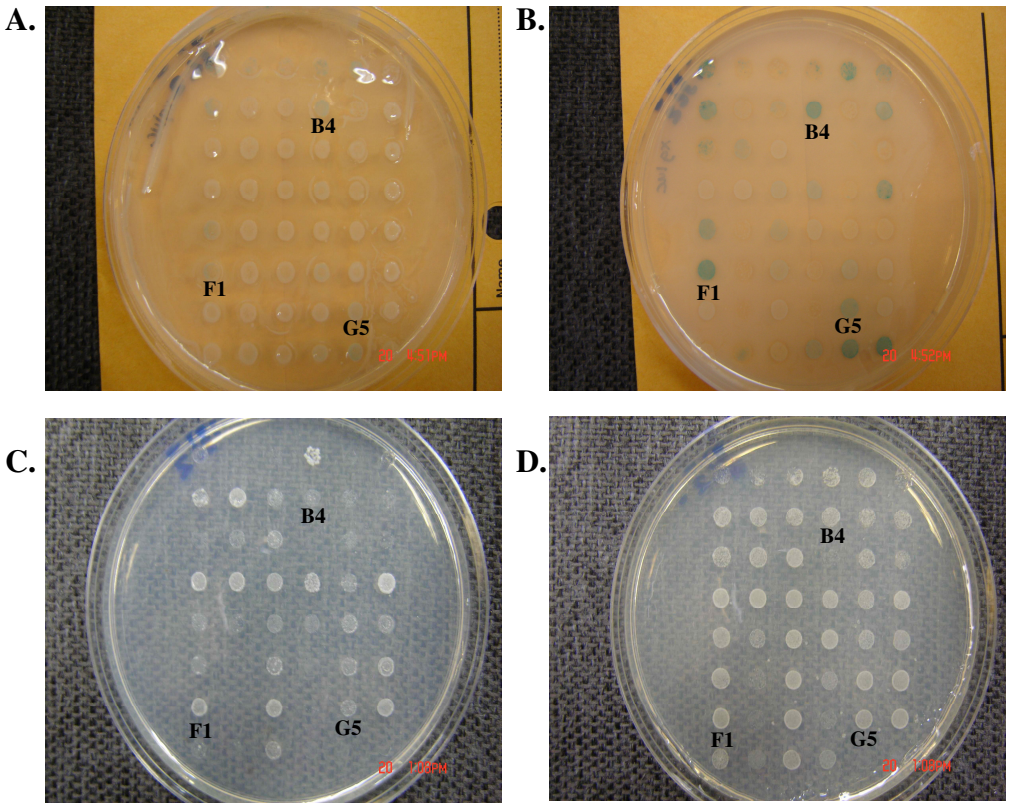


Figure 3.33 *LYS2* and *GusA* transcription activation tests of 3L2 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

The interactors sequenced from 2KL2 plate were positions B1, D6, G6, F3, and H1. The library plasmids in these colonies code for yeast open reading frames of EAP1, PEX8, SKY1, POR1, and CSE1, respectively. Selected colonies showed better *GusA* transcription activation on Gal-Raff plate (Figure 3.34 B) than glucose plate (Figure

3.34 A), and slightly better growth on Gal-Raff/CM- Ura-His-Trp-Lys plate (Figure 3.34 D) than Glu/CM- Ura-His-Trp-Lys plate (Figure 3.34 C).

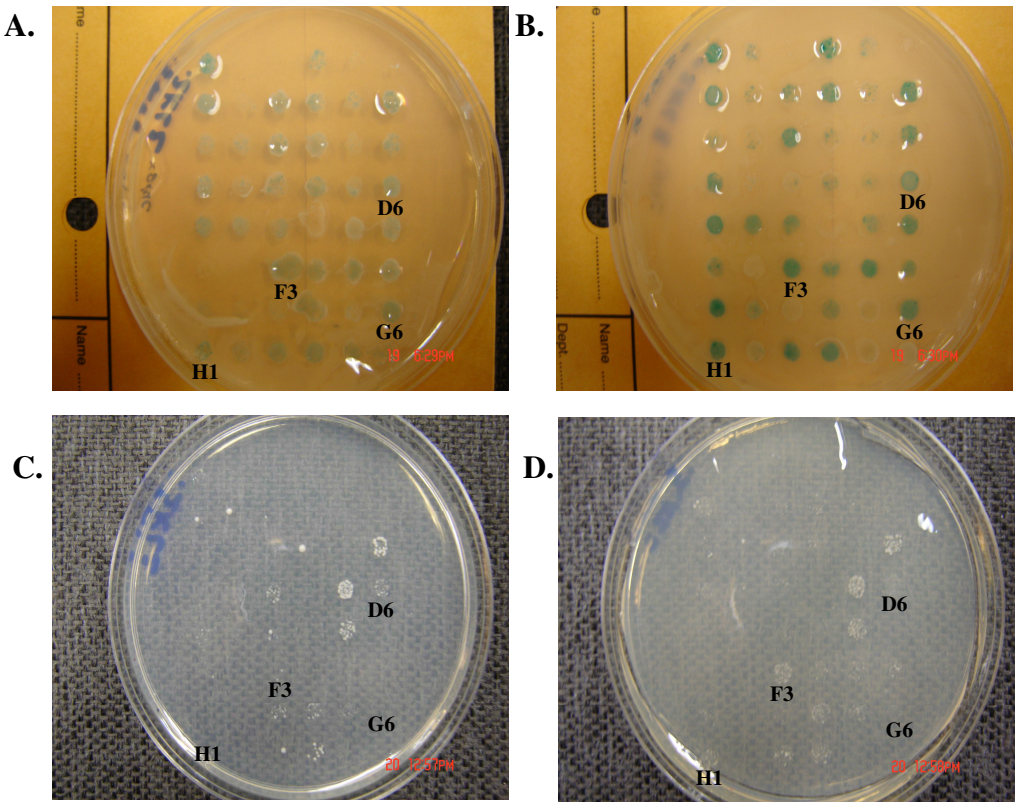


Figure 3.34 *LYS2* and *GusA* transcription activation tests of 2KL2 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

3.12.4 WR5 interactors

There were 3 WR5 interactors sequenced (Appendix G).

The interactors sequenced from 5L1 plate were positions A6 and C1. The library plasmids in these colonies code for yeast open reading frames of OXR1 and Yjl057cp, respectively. Selected colonies showed better *LacZ* transcription activation on Gal-Raff plate (as there was no background activity) (Figure 3.35 B) than glucose plate (Figure 3.35 A), and apparent better growth on Gal-Raff/CM -Ura-His-Trp-Leu plate (Figure 3.35 D) than Glu/CM-Ura-His-Trp-Leu plate (Figure 3.35 C).

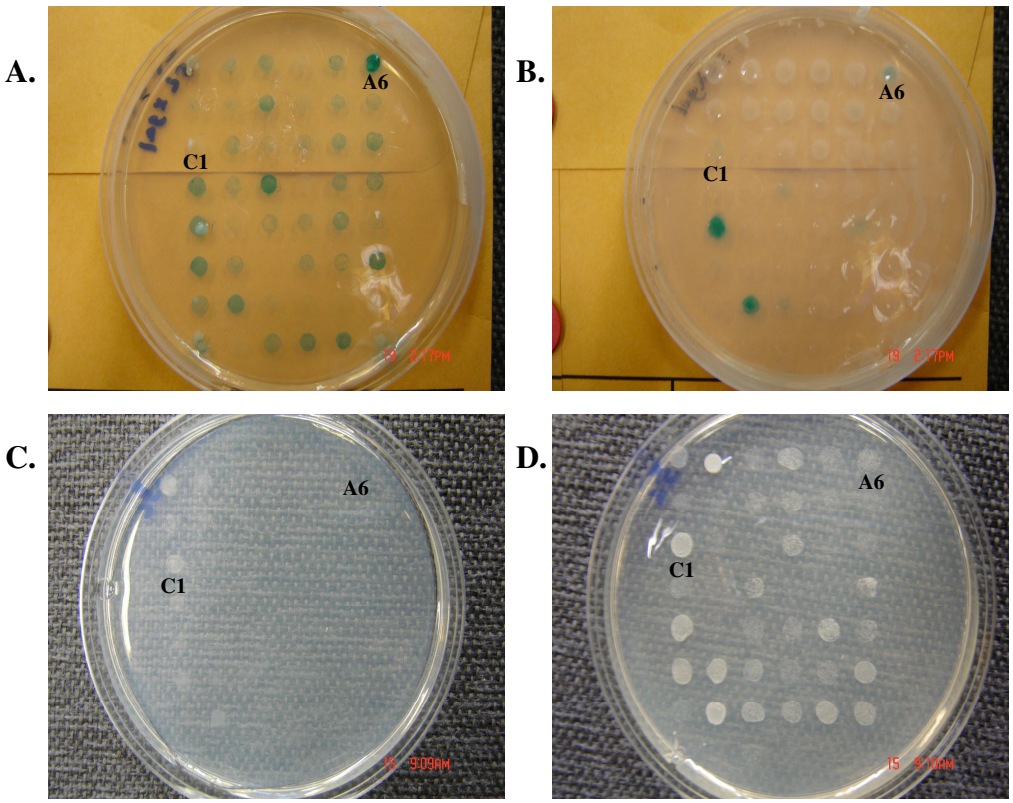


Figure 3.35 *LEU2* and *LacZ* transcription activation tests of 5L1 plate. **A:** X-Gal overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gal overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Leu medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Leu medium

The interactor sequenced from 5L2 plate was the position G5 (GWT1). Selected colonies showed better *LacZ* and *LEU2* transcription activation on Gal-Raff plates (Figure 3.36 B and D.) than glucose plates (Figure 3.36 A and C).

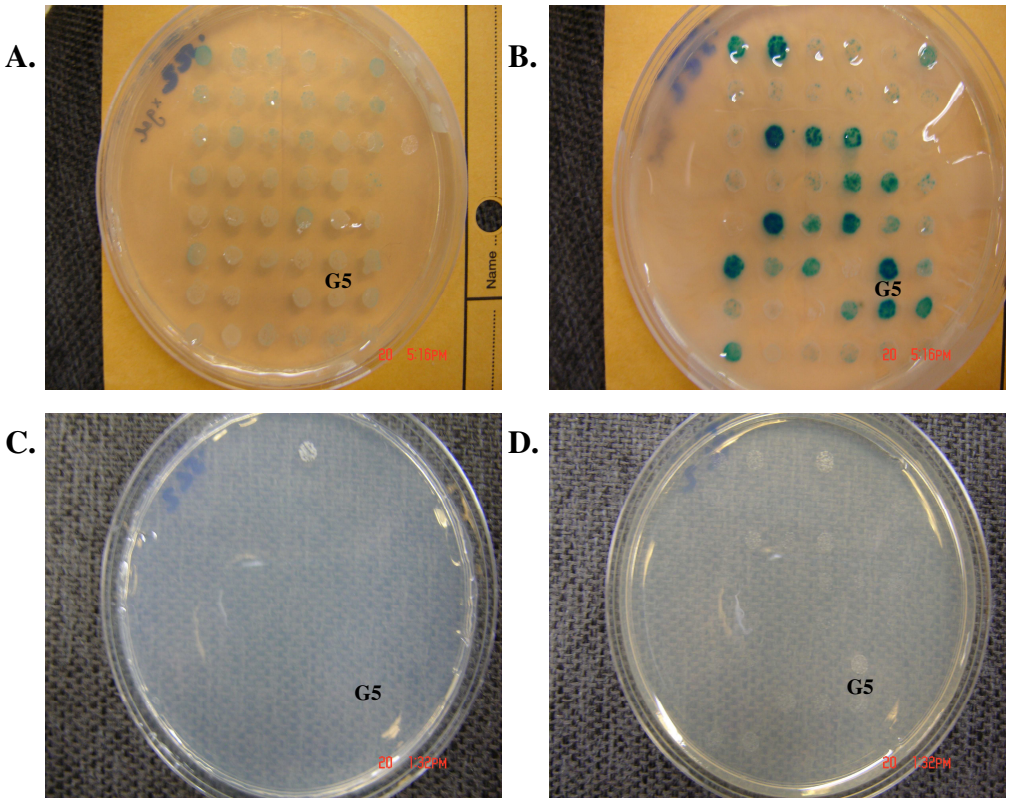


Figure 3.36 *LEU2* and *LacZ* transcription activation tests of 5L2 plate. **A:** X-Gal overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gal overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Leu medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Leu medium

3.12.5 Rad6 interactors

There were total of 8 cells selected for sequencing (Appendix H) as the interactors of Rad6 protein.

The colonies selected for sequencing from 5L1 plate were positions A1, A6 and B1. The library plasmids in these colonies code for yeast open reading frames of MAF1, OXR1, and Yer087wp, respectively. These cells showed better *GusA* transcription activation on Gal-Raff plate (Figure 3.37 B) than glucose plate (Figure 3.37 A). Also these cells apparently better are grown on Gal-Raff/CM- Ura-His-Trp-Lys plates (Figure 3.37 D) than Glu /CM- Ura-His-Trp-Lys plate (Figure 3.37 C).

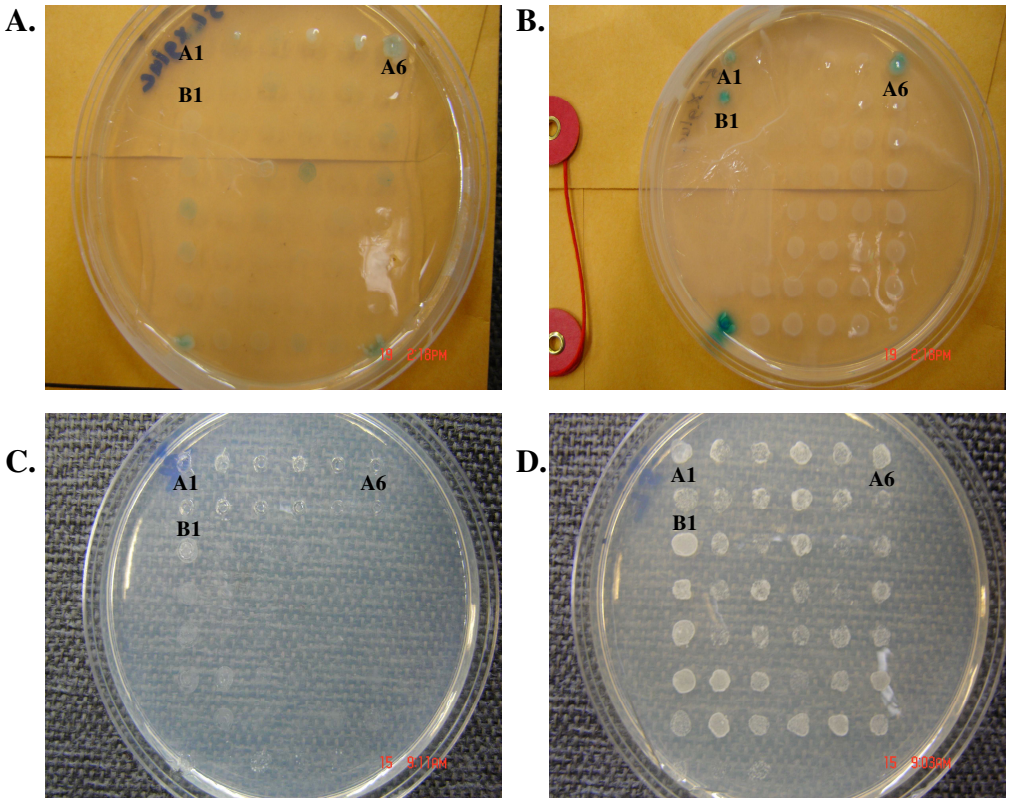


Figure 3.37 *LYS2* and *GusA* transcription activation tests of 5L1 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

From 5K2 plate positions D2, E2, E3, G3 and G4 were the cells selected for sequencing. The PCR amplification of E3 colony with library plasmids' primers gave 2 bands (Figure 3.24) and both of the bands were sequenced. The library plasmids in these colonies code for yeast open reading frames of cosmid 8082, cosmid 8082 & an

unnamed protein, Yll066cp, and PRY3, respectively. These colonies demonstrated better growth on Gal-Raff/CM- Ura-His-Trp-Lys plate (Figure 3.38 D) than Glu /CM-Ura-His-Trp-Lys (Figure 3.38 C). *GusA* transcription activation on Gal-Raff plate (Figure 3.38 B) was also higher than the activation on glucose plate (Figure 3.38 A).

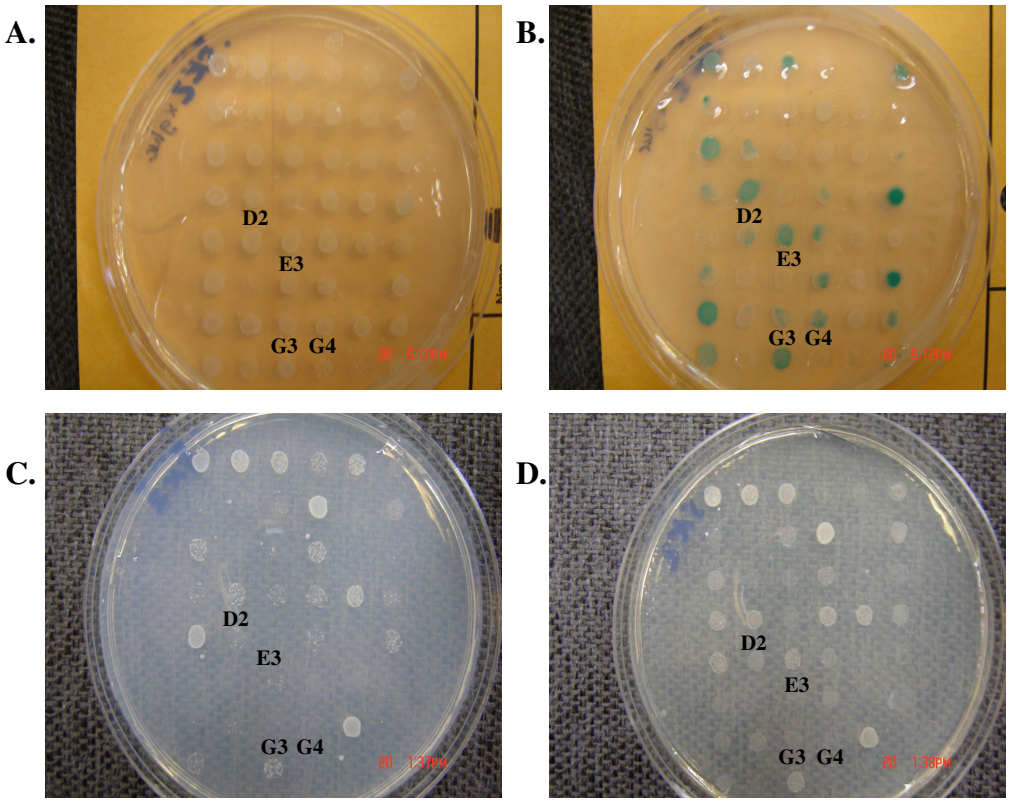


Figure 3.38 *LYS2* and *GusA* transcription activation tests of 5K2 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

From the 5K1 plate only colony on A1 position (FAR1) was selected for sequencing as it was the only one that confers blue color on X-Gluc overlaid Gal-Raff plate (Figure 3.39 B). There was very little transcription activation of *LYS2* on Gal-Raff/CM -Ura-His-Trp-Lys medium (Figure 3.39 D)

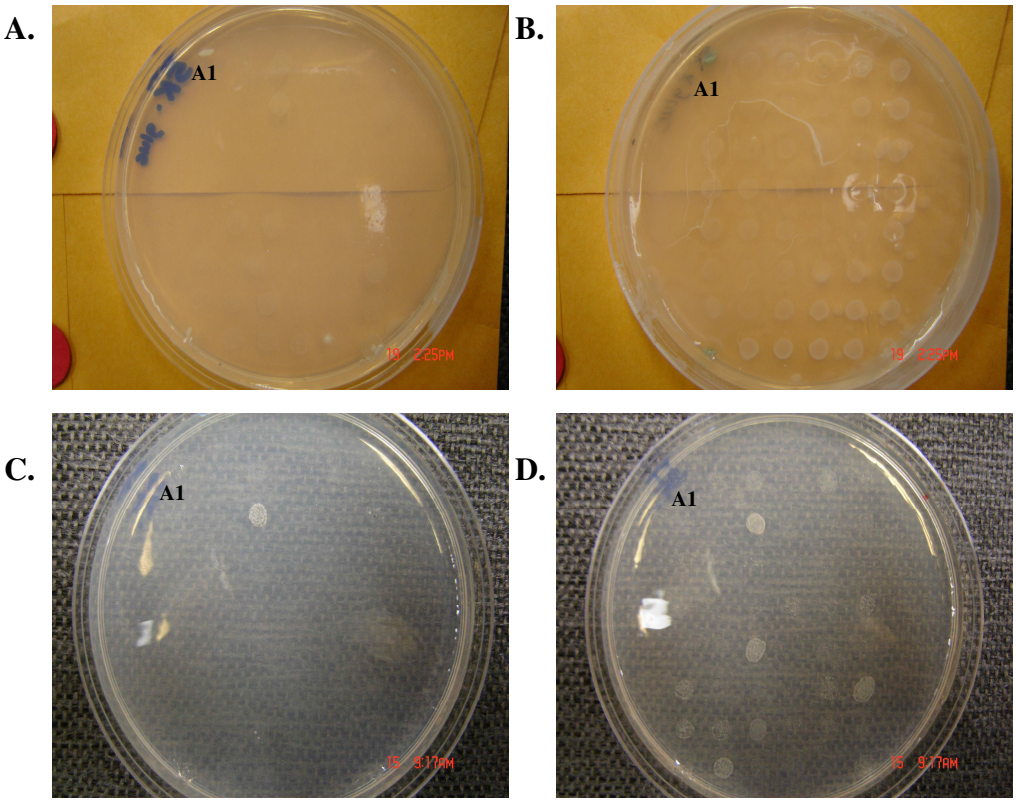


Figure 3.39 *LYS2* and *GusA* transcription activation tests of 5K1 plate. **A:** X-Gluc overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/CM -Ura-His-Trp medium **C:** Growth of cells on Glu/CM-Ura-His-Trp-Lys medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium

3.13 Second confirmation of the positive clones

For the second round of confirmation of the positive clones all interactors (pJG4-5-interactor) in the yeast cells were isolated. The isolated plasmids were transformed into *E. coli* for amplification of the plasmids. The amplified interactor plasmids were isolated from *E. coli* cells and transformed into the yeast cell having corresponding baits and the reporter plasmid. The interactions were tested either with *LacZ* and *LEU2* (pMW103 constructs) or *GusA* and *LYS2* (pGLS23 constructs) transcriptional activation, depending on the plasmid of the bait.

5L1-C1, 5L2-G5 and 5L1-A6 were the colonies that were selected as the interactors of WR5 gene fragment which was cloned into pMW103 bait vector. The second confirmations of these interactions are presented in Figure 3.40. The background activity of β-Galactosidase was detected on the Glu/CM -Ura-His-Trp medium (Figure 3.40 A), but this background was lost in the Gal-Raff/CM -Ura-His-Trp (Figure 3.40 B). Positive clones also showed better growth profile on Gal-Raff/CM -Ura-His-Trp-Leu medium (Figure 3.40 D) than Glu/CM -Ura-His-Trp-Leu (Figure 3.40C). All of the positive interactions were confirmed.

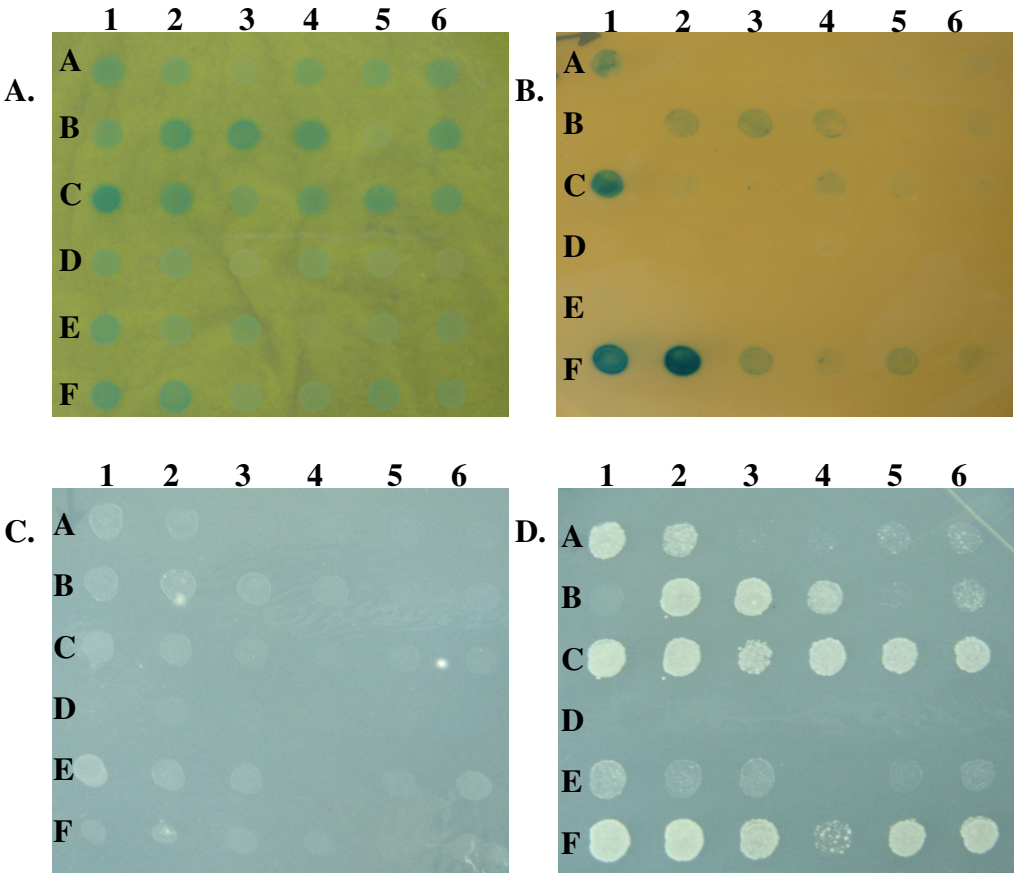


Figure 3.40 Second confirmations of positive clones (LEU2 and *LacZ* transcription activation tests) **A:** X-Gal overlay on Glu/CM -Ura-His-Trp medium **B:** X-Gal overlay on Gal-Raff/CM -Ura-His-Trp **C:** Growth of cells on Glu/CM -Ura-His-Trp-Leu medium **D:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Leu medium **Samples:** Lane A (1-6): 5L1-C1, Lane B (1-6): 5L2-G5, Lane C (1-6): 5L1-A6, Lane D(1-6): 1K1-F3 (negative control), Lane E (1-6): Negative control (pGLS22-Ras, pJG4-5-Raf), Lane F (1-6): Positive control (pMW103-Krit, pJG4-5-Krev)

For the second confirmation of the interaction between pGLS-bait and pJG4-5-prey, *GusA* (Figure 3.41 and Figure 3.42) and *LYS2* (Figure 3.43 and Figure 3.44) transcription activations are tested. There was β-Glucoronidase back ground activity on Glu/CM -Ura-His-Trp medium (Figure 3.41 A and Figure 3.42 A), but disappeared on Gal-Raff/ CM-Ura-His-Trp medium (Figure 3.41 B and Figure 3.42 B). The growth of colonies on Gal-Raff/ CM -Ura-His-Trp-Lys (Figure 3.43 B and Figure 3.44 B) medium were better with respect to the growth of cells on Glu/CM -Ura-His-Trp-Lys medium (Figure 3.43 A and Figure 3.44 A). All of the positive interactions were further confirmed.

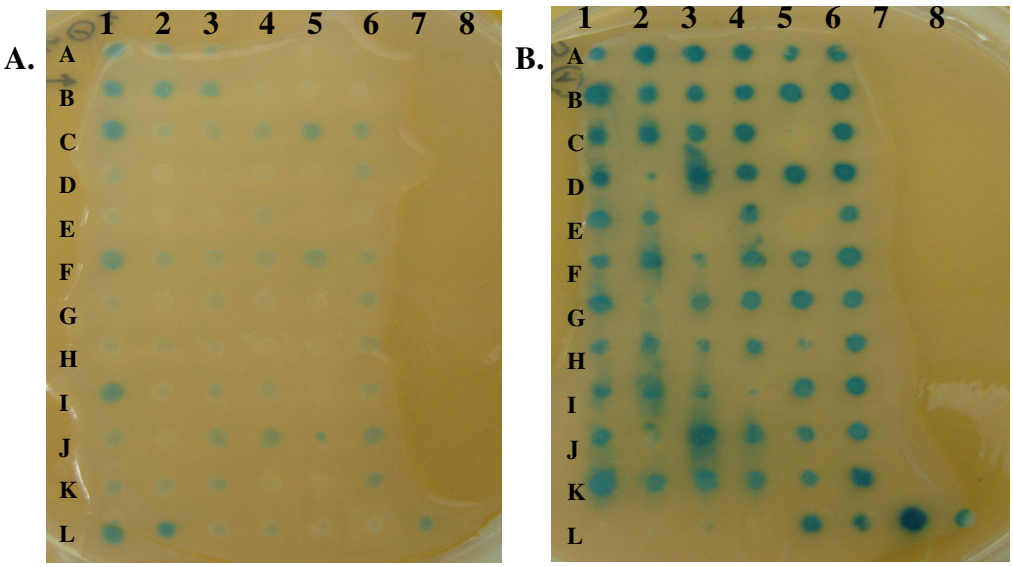


Figure 3.41 Second confirmations of positive clones (*GusA* transcription activation tests) **A:** X-Gluc overlay on Glu/ CM-Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/ CM-Ura-His-Trp medium. **Samples:** Lane A (1-6): 5L1-A6, Lane B (1-6): 5K2-G4, Lane C (1-6): 5L1-A1, Lane D (1-6): 5K1-A1, Lane E (1-6): 2K1-A1, Lane F (1-6): 3L2-F1, Lane G (1-6): 3L2-G5, Lane H (1-6): 2KL2-D6, Lane I (1-6): 2KL2-F3, Lane J (1-6): 2KL2-H1, Lane K (1-6): 2KL2-B1, Lane L (1-4): Negative control (pMW103-Krit, pJG4-5-Krev), Lane L (5-8): Positive control (pGLS22-Ras, pJG4-5-Raf).

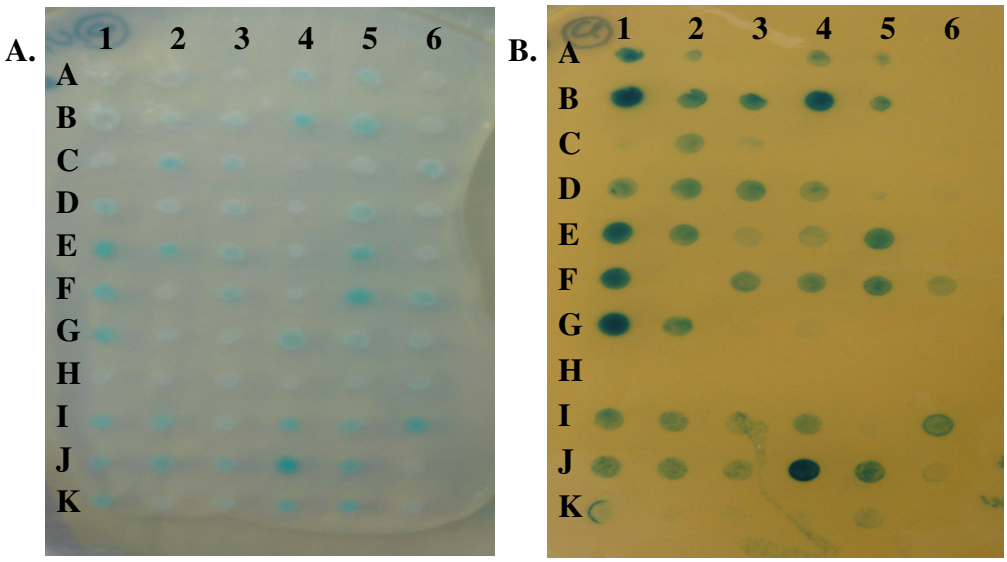


Figure 3.42 Second confirmations of positive clones (*GusA* transcription activation tests) **A:** X-Gluc overlay on Glu/ CM-Ura-His-Trp medium **B:** X-Gluc overlay on Gal-Raff/ CM-Ura-His-Trp medium. **Samples:** Lane A (1-6): 1K1-F3, Lane B (1-6): 1K1-F5, Lane C (1-6): 4K1-B2, Lane D (1-6): 4K1-C1, Lane E (1-6): 4K1-D1, Lane F (1-6): 4K1-G2, Lane G (1-6): 4L2-F5, Lane H (1-6): 5L1-C1 (Negative control), Lane I (1-6): 4L2-B4, Lane J (1-6): Positive control (pGLS22-Ras, pJG4-5-Raf), Lane K (1-6): Negative control (pMW103-Krit, pJG4-5-Krev).

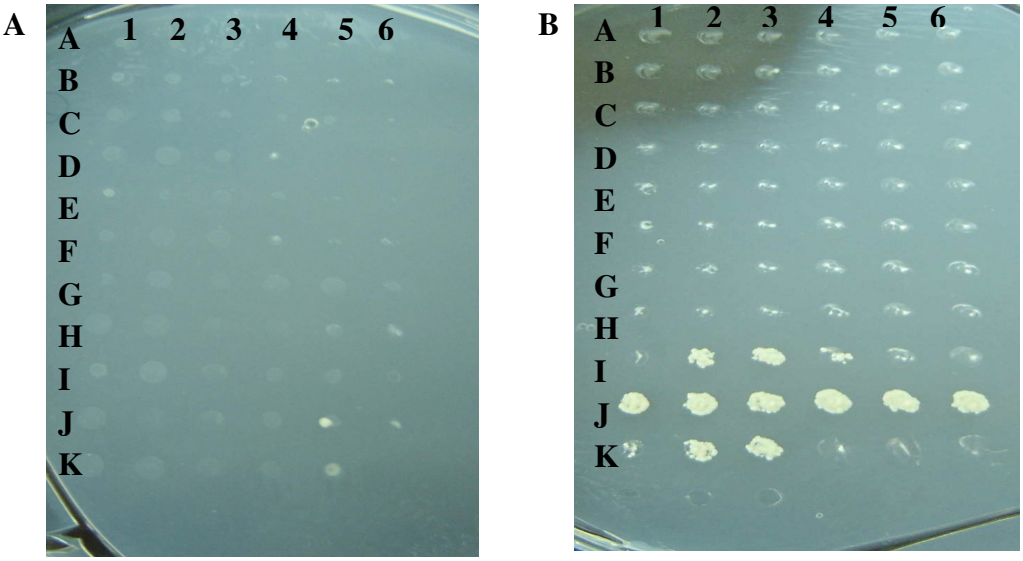


Figure 3.43 Second confirmations of positive clones (*LYS2* transcription activation tests) **A:** Growth of cells on Glu/CM -Ura-His-Trp-Lys medium **B:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium. **Samples:** Lane A (1-6): 1K1-F3, Lane B (1-6): 1K1-F5, Lane C (1-6): 4K1-B2, Lane D (1-6): 4K1-C1, Lane E (1-6): 4K1-D1, Lane F (1-6): 4K1-G2, Lane G (1-6): 4L2-F5, Lane H (1-6): 5L1-C1 (Negative control), Lane I (1-6): 4L2-B4, Lane J (1-6): Positive control (pGLS22-Ras, pJG4-5-Raf). Lane K (1-3): Positive control (pGLS22-Ras, pJG4-5-Raf), Lane K (4-6): Negative control (pMW103-Krit, pJG4-5-Krev).

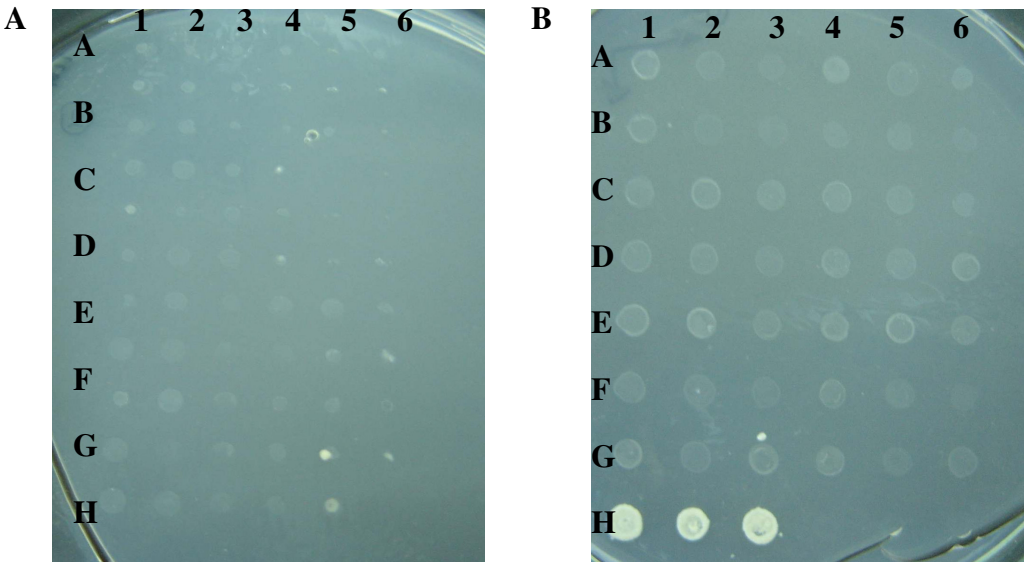


Figure 3.44 Second confirmations of positive clones (*LYS2* transcription activation tests) **A:** Growth of cells on Glu/CM -Ura-His-Trp-Lys medium **B:** Growth of cells on Gal-Raff/CM -Ura-His-Trp-Lys medium. **Samples:** Lane A (1-6): 2K1-A1, Lane B (1-6): 3L2-F1, Lane C (1-3): 3L2-G5, Lane C (4-6): 5L1-A6, Lane D (1-3) : 2KL2-D6, Lane D (4-5): 5K2-G4, Lane E (1-3): 2KL2-F3, Lane E (4-6): 5L1-A1, Lane F (1-3): 2KL2-H1, Lane F (4-6): 5K1-A1, Lane G: 2KL2-B1 Lane H (1-3): Positive control (pGLS22-Ras, pJG4-5-Raf) Lane H (4-6): Negative control (pMW103-Krit, pJG4-5-Krev).

3.14 An overview of the results

SGT1 and *Rar1* gene fragments are cloned in wheat for the first time. The sequence alignments showed that these two genes are highly conserved (more than 90% sequence homology with the barley *SGT1* and *Rar1* genes). These two genes are

very important in *R* gene mediated disease responses. The studies in literature state that *Rar1* is a required gene in powdery mildew resistance by several CC-NBS-LLR genes (Schulze-Lefert and Vogel, 2000) and *N* gene–mediated resistance response to *Tobacco mosaic virus* (TMV) in tobacco plants (Liu et al., 2002). *Arabidopsis SGT1* was identified in mutational analysis for loss of *RPP7*- and *RPP5*-mediated resistance (Austin *et al.*, 2002; Tor *et al.*, 2002), and in a yeast two-hybrid screen SGT1 is found to be interacting with the Rar1 (Azevedo *et al*., 2002). Both of these proteins are important signaling components in diverse *R* gene mediated disease responses.

The sequence information of *Rar1* and *SGT1* genes in wheat were used for further applications, such as quantitative PCR for the analysis of the mRNA level changes occurrence upon avirulent and virulent infections of Avocet differential lines (Osman Bozkurt, PhD thesis study).

The dual bait yeast two-hybrid analysis was successfully performed in order to investigate the possible conserved signaling cascade in *Yr10* mediated disease response. *Yr10*, being the only gene cloned against the yellow rust disease of wheat, is very important as the initial regulator of the disease response. The downstream reactions are all initiated with recognition of the pathogen Avr protein by the plant R protein. There are very promising interactions found that may inform us by means of R protein mediated signaling. The baits used, and the preys investigated in this study are summarized in Table 3.2.

Table 3.2 An overview of the interactions determined in the study.

Bait	Prey	Colony nb.	Function of prey in yeast
YR10-1	SRP72	1K1-F3	targeting nascent secretory proteins to the endoplasmic reticulum (ER) membrane
YR10-1	LAS17	1K1-F5	Actin assembly factor
YR10-1	Gac1	4K1-B2	Protein phosphatase 1 regulatory subunit
YR10-1	TOM1	4K1-C1	E3 ubiquitin protein ligase
YR10-1	AZF1	4K1-D2	Asparagine-rich zinc finger protein
YR10-1	YPL207w	4K1-G2	unnamed protein product
YR10-1	ARD1	4L2-F5	subunit of the major N alpha-acetyltransferase
YR10-1	CPR7	4L2-B4	cyclophilin-like protein CPR7
YR10-1	TyB	4K1-G1	yeast transposon Ty1-H3
YR10-1	Yll066c	1K2-F1	Y'helicase protein 1
YR10-1	PYC2	1K2-F5	Pyruvate carboxylase 2
YR10-1	J protein type 2	4L2-D1	Associated with hsp70 heat-shock system
YR10-2	Vps35	2K1-A1	Vacuolar protein sorting-associated protein 35
YR10-2	Gcn2	3L2-F1	Serine/threonine-protein kinase
YR10-2	NCP1	3L2-G5	NADPH cytochrome P450 reductase
YR10-2	PEX8	2KL2-D6	Intraperoxisomal organizer of the peroxisomal import machinery
YR10-2	POR1	2KL2-F3	Mitochondrial porin (voltage-dependent anion channel)
YR10-2	CSE1	2KL2-H1	Importin alpha re-exporter
YR10-2	Eap1	2KL2-B1	eIF4E binding protein
YR10-2	TyB	2K1-A6	yeast transposon Ty1-H3
YR10-2	YLR162w	3L2-B4	Protein of unknown function
YR10-2	SKY1	2KL2-G6	SR protein kinase (SRPK)
WR5	YJL057C	5L1-C1	Probable serine/threonine-protein kinase
WR5	Gwt1	5L2-G5	GWT1 involved in GPI anchor biosynthesis; required for inositol acylation in yeast
WR5	OXR1	5L1-A6	Protein of unknown function required for normal levels of resistance to oxidative damage
Rad6	OXR1	5L1-A6	Protein of unknown function required for normal levels of resistance to oxidative damage
Rad6	PRY3	5K2-G4	Protein of unknown function, has similarity to Pry1p and Pry2p and to the plant PR-1 class of pathogen related proteins
Rad6	MAF1	5L1-A1	Negative regulator of RNA polymerase III
Rad6	YER087w	5L1-B1	Probable prolyl-tRNA synthetase
Rad6	FAR1	5K1-A1	Cyclin-dependent kinase inhibitor
Rad6	cosmid 8082	5K2-D2 5K2-E3/1	No function determined
Rad6	unnamed protein product	5K2-E3/2	Hypothetical 14.1 kDa protein in SEC21-MRPL10 intergenic region
Rad6	Yll066c	5K2-G3	subtelomerically-encoded DNA helicase

3.15 Final discussions

The interactions found in this study were literature searched in order to explain the possible roles in R protein mediated disease resistance mechanisms. There are very promising results presenting interactors that possibly have significant roles during the resistance signaling cascade.

3.15.1 Post translational and translation initiation controls

The first important protein is the **SRP72** (colony number: 1K1-F3) which interacts with the YR10-1 gene fragment. Signal Recognition Particle (SRP) is a highly conserved cytoplasmic complex. It has several components that work as binding to the newly synthesized proteins having signal peptide, arrest of the elongation during translation, and binding to the SRP receptor so that elongation arrest ends and the transport of the secretory protein into the ER becomes possible (Walter and Blobel, 1980, 1982). SRP72 is responsible for the binding to the receptor and it targets nascent secretory proteins into the endoplasmic reticulum (ER) membrane (Brown *et al.,* 1994; Grosshans, *et al.*, 2001). The human SRP72 is found to be cleaved during apoptosis (Utz *et al*., 1998). This is a post-translational modification during the programmed cell death. It is shown that SRP72 cleavage is prevented by chemical and peptide caspase inhibitors, and by over expression of *bcl-2*, which is an inhibitor of apoptotic cell death. Cleavage of SRP72 produces a 66-kDa amino-terminal fragment and a 6-kDa carboxyl-terminal fragment, selectively phosphorylated on serine residues. This phosphorylation and caspase cleavage believed to regulate targeting of secretory proteins into the ER lumen during apoptosis (Utz *et al*., 1998). In this study authors suggest that cleavage of SRP72 may result in inhibition of is binding to SRP, so that elongation is arrested and secretory proteins can not enter into the ER. This model is not valid in *in vitro* studies

as secretion is not completely blocked after cleavage. In our case, this proposed model might be important and possible if the pathogen needs to secrete some proteins outside the cell for its spread. It is also declared that, yeast SRP72 has serine residues on the 620^{th} and 627^{th} amino acids, which might be possible phosphorylation sites (Utz *et al*., 1998). In our experiment, it can be predicted that there might be induction of a protease activity (caspase is a cysteine protease) by the action of the p-loop (like Apaf-1 and CED4 activating caspase 9). By this model the post translational modification of the SRP72 during apoptosis and the indirect interaction with YR10-1 can be explained (Figure 3.45).

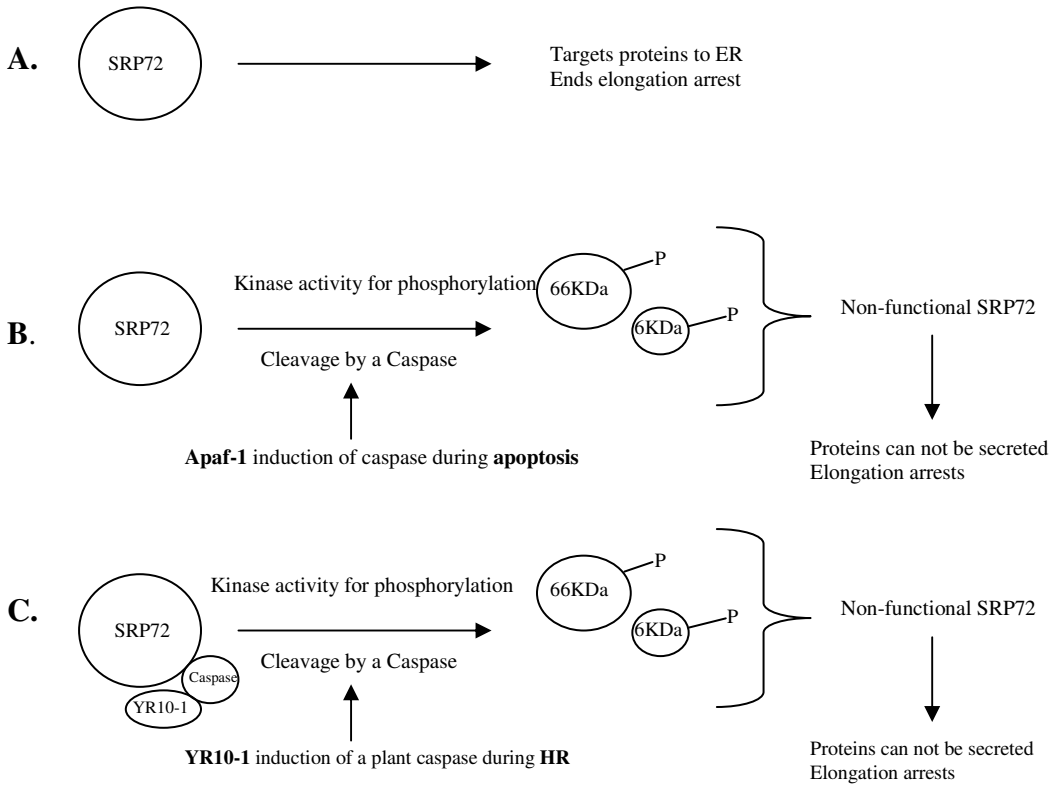


Figure 3.45 SRP72 functions and probable function in YR10 protein mediated HR. **A:** Functional SRP72 protein. **B:** Non-functional SRP72 during apoptosis. **C:** Possible mechanism in YR10-1 mediated HR.

A part from post-translational controls, translation mechanism is regulated at the initiation in response to stress (Dever, 2002).

Lipid and protein transport is a highly regulated route and it is required to sustain the integrity of a variety of intracellular organelles in eukaryotes. Recently, it is

found that mutations in the secretory pathway, lead to quick and precise attenuation of the translation (Deloche *et al*., 2004).

In higher eukaryotes translation initiation control mechanism is achieved by several kinases like mitogen-activated protein kinases (MAPK), which inhibits translation by phosphorylation of eIF4E (mRNA cap-binding protein) (Sonenberg and Gingras, 1998); phosphorylation of the eIF2 initiation factor by double stranded RNA-regulated protein kinase (de Haro *et al*., 1996; Clemens 1997). Many different eIF-2a kinases present in mammals, but the yeast has only one eIF-2a kinase, which is the **Gcn2p** (colony number: 3L2-F1, interactor of YR10-2) (Hinnebusch, 1996; Wek *et al*., 1992). Phosphorylation of eIF2 down regulates the protein synthesis as it blocks the conversion eIF-2-GDP to eIF-2-GTP at translation initiation. Only GTP bound form can bind to initiator $tRNA_i^{Met}$ (reviewed by Hershey, 1991). Studies on eIF4E binding proteins (4EBPs) show that this binding to eIF4E, blocks the ribosome binding to mRNA (Poulin *et al*., 1998). Functional homolog of EBPs in yeast is the **Eap1p** (colony number: 2KL2-B1, interactor of YR10-2) and it has a negative role in translation initiation (Cosentino *et al*., 2000). Translation initiation, in membrane stressed wild-type cells (by chemical treatment with Chlorpromazine) and in mutant yeast cells (affected in diverse stages of the secretory (*sec*) and endocytic (*end*) pathways) show rapid translational attenuation which is mediated by Gcn2p and Eap1p (Deloche *et al*., 2004).

It is interesting that both of these proteins are interacting with YR10-2. These interactions state that, there might be a translational control mechanism related to Yr10 mediated disease resistance response. During disease response there might be vesicular transport defects (which is also evident in SRP72), and these might induce the translational negative control (Figure 3.46).

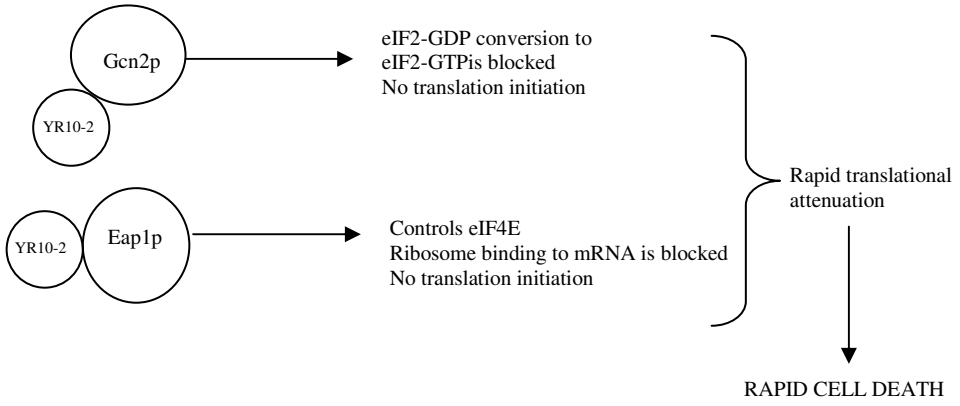


Figure 3.46 Possible roles of Gcn2p and Eap1p in Yr10 protein mediated resistance response. Gcn2p and Eap1p may control the translation initiation in response to the Yr10 mediated signaling. The rapid attenuation of protein synthesis may result in the rapid plant cell death.

Another key finding is that Gcn2p interacts with the Gcn4p through the histidyl-tRNA synthetase (HisRS) domain and increases Gcn4p synthesis during amino acids starvation (Dever *et al*., 1992). Increased levels of Gcn4p stimulate the expression of enzymes required for the synthesis of amino acids (Hinnebusch, 1996). Another significant finding is that, translational control of Gcn4p involves a reduction in eIF-2B function, a mechanism used in mammalian cells to regulate total protein synthesis in response to stress (Cigan *et al*., 1993). This result means that if there is translational control of the Gcn4p, there will be less control on eIF2 which increases the translation of several proteins, or just the opposite.

These findings raise two other possibilities related to the disease resistance mechanism.

One possibility is, during the disease resistance pathway there might be a high level of protein synthesis is needed and Gcn4p homolog in plant needs to be activated to fulfill this demand. In this possibility Yr10 somehow activates the Gcn2p (Figure 3.47 A).

The other possibility is that Yr10 may inactivate the Gcn2p so that protein Gcn4p synthesis will not be activated. This might be happening as the cell kills rapidly itself during the HR. To support this idea, there is a significant finding that ubiqutination of Gcn4p is achieved by Rad6 and Cdc34 in yeast (Kornitzer *et al*., 1994). The Rad6 protein interacts with the SGT1 and it is expected to have a role during the disease resistance. The possible role of Rad6 might be the degradation of Gcn4p, and regulation of the protein synthesis. Another important protein having role in plant disease resistance mechanism is the HSP90 (Takahashi *et al*., 2003). Interestingly, Gcn2p requires HSP90 for proper regulation (Donze and Picard, 1999). Gcn4p and Gcn2p interacting with these important R gene mediated resistance genes raises the idea that, protein synthesis and degradations might be controlled by Gcn4 and Gcn2 proteins during resistance response (Figure 3.47 B).

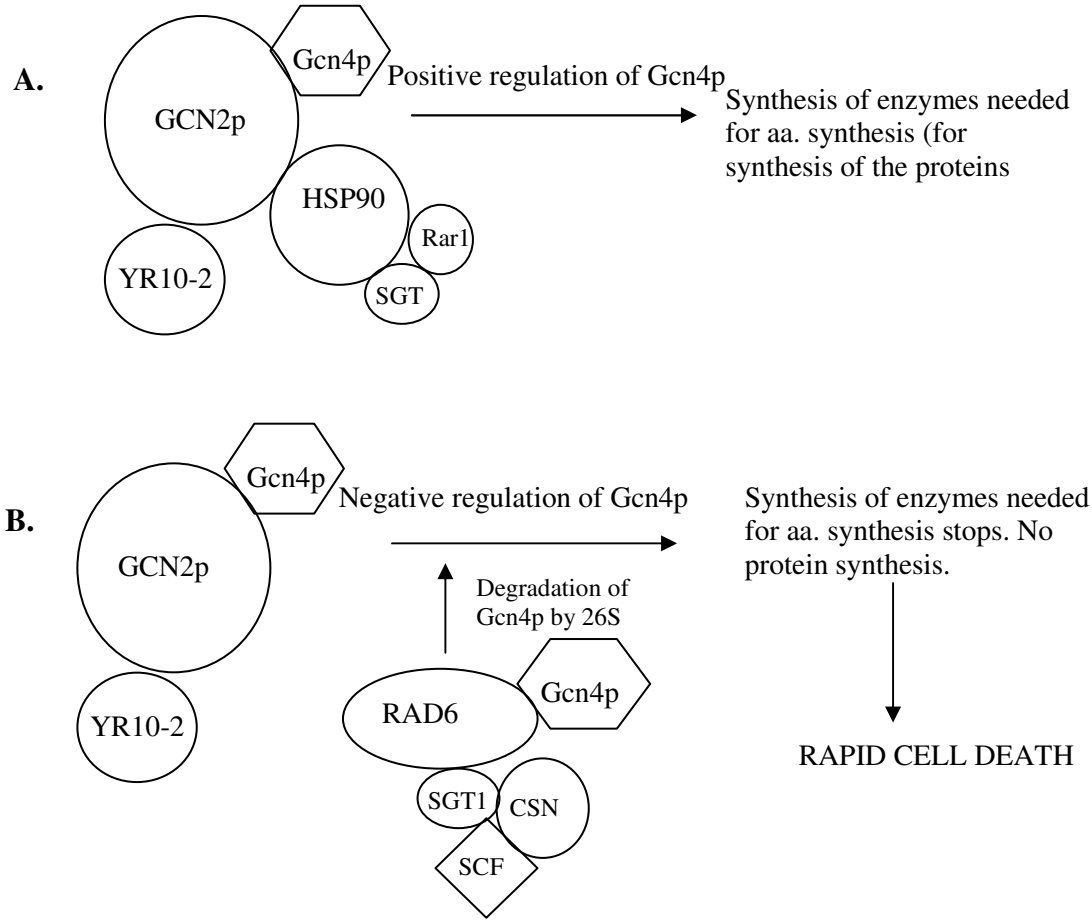


Figure 3.47 Possible roles of YR10-2 and Gcn2p interaction on the Gcn4p expression and degradation. **A:** YR10-2 may control the activity of Gcn2p plant homolog in the way that Gcn4p expression is positively regulated. During this regulation there might be involvements of HSP90, SGT1 and Rar1 proteins (as it is known that Gcn2p needs HSP90 for its activity on Gcn4p). **B:** YR10-2 may control the activity of Gcn2p plant homolog (blocking) in the way that Gcn4p expression is negatively regulated. There might be involvement of Rad6, SGT1, SCF and CSN (COP9 signalosome) (as it is known that Gcn4p degradation is controlled by Rad6 in yeast, and also Rad6 is an interacting partner of SGT1).

To sum up; a SRP72 plant homolog, a Gcn2p plant homolog, Gcn4p plant homolog, an Eap1p plant homolog, eIF2, eIF4E, Rad6, Hsp90 and SGT1 might work collectively in the plant disease resistance pathway, for the control of protein translation and post-translational modifications.

3.15.2 Kinase and phosphatase activities

Serine/threonine-protein kinase **SKY1** (colony number: 2KL2-G6), which also interacts with YR10-2, is involved in regulating proteins involved mRNA metabolism and cation homeostasis (Erez and Kahana, 2002) in yeast. Both of these properties are also important in *R* gene mediated disease response. A serine/threonine-protein kinase may have diverse function in every kind of pathway as signaling components require the involvements of kinases.

The clone 4K1- B2 sequence interacting with the bait YR10-1 encodes for the yeast Protein phosphatase 1 (PP1) regulatory subunit **GAC1** (Wu *et al*., 2001). *GAC1* encodes a protein phosphatase that forms a haloenzyme with PP1, and controls the dephosphorylation of glycogen synthase (Ramaswamy *et al*., 1998) GAC1 is found to be induced by heat shock and osmotic stress (Parrou *et al*., 1997). PP1 is a serine/threonine phosphatase. Phosphorylation and dephosphorylation events are important in all kinds of signal transduction mechanisms. PP1 has a role in several pathways including cell cycle regulation, glycogen accumulation, lipid metabolism, Ca^{+2} transport, and regulation of translational initiation (reviewed by Williams-Hart *et al*., 2002). It also stated that PP1 regulates ion homeostasis (Williams-Hart *et al*., 2002). All these data suggest that Yr10 interaction with GAC1 might be important for the regulation of PP1 plant homolog.

In a study, it is found that yeast differentially express **GAC1** and **PRY3** (colony number: 5K2-G4, Rad6 interactor) during isooctane tolerance response (Miura *et al*., 2000). This finding is important as these two genes work in the same tolerance pathway. PRY3 has no known function in yeast but it shows similarity to plant PR1 (pathogen related proteins) genes. GAC1 is found to be induced, while PRY3 is found to be depressed during isooctane tolerance response. This finding is also related with our study as we had found the PRY3 as an interactor of Rad6. This means that stability of PRY3 plant homolog might be controlled during the Yr10 mediated disease response.

3.15.3 Protein degradation controls

4K1-C1 sequence (interactor of the bait, YR10-1) codes for the **TOM1** protein which is an E3 protein ligase. Ubiquitination plays an important role in the regulation of biological processes such as cell cycle regulation, signal transduction, apoptosis, and DNA repair (Ciechanover, 1994; Hiyama *et al*., 1999; Hauser *et al*., 2000; Hershko and Ciechanover, 1998). Thus, it is also very important for the *R* gene mediated disease resistance in plants. Recently, it is stated that there is antagonistic control of resistance protein stability. In this study it is predicted that Rar1 controls preactivation of R protein accumulation, and SGT1 antagonizes Rar1 to negatively regulate the R protein accumulation before infection (Holt *et al*., 2005). By the interaction of TOM1 and Yr10 gene we can speculate that there might be a control on Yr10 stability *via* TOM1 plant homolog in wheat.

3.15.4 Apoptosis related functions

A very important finding in our study is the interaction of **POR1** (colony number: 2KL2-F3) with the YR10-2 (NB-ARC domain, Appendix B). ARC domain (Apaf-1, R protein and CED4 homology) shows homology to the Apaf-1 and CED4 proteins which are involved in cell-death signaling. It is expected to have some conserved relations between the animal cell apoptosis and plant HR (Van der Biezen and Jones, 1998). POR1 is a mitochondrial porin (voltage-dependent anion channel), and required for the maintenance of mitochondrial osmotic stability and mitochondrial membrane permeability (Lee *et al*., 1998; Sanches *et al*., 2001). Mitochondrial membrane potential and the release of cytochrome *c* are important features of apoptosis (Figure 3.48). Yr10-POR1 interaction raise the idea that this *R* gene can have a role in cell death by controlling the mitochondria regulated apoptosis.

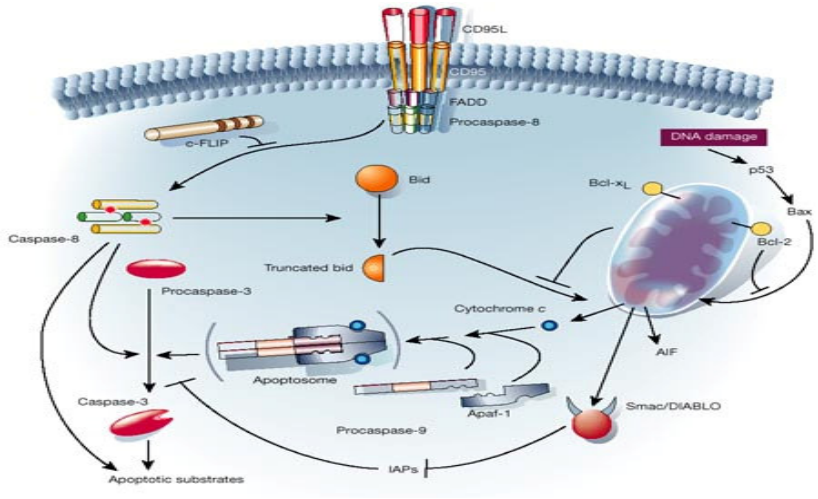


Figure 3.48 Apoptotic pathways in mammalian cells (Hengartner, 2000)

CSE1 (colony number: 2KL2-H1) is an importin alpha re-exporter which is related with the chromosome segregation in yeast (Xiao *et al*., 1993). The human homolog of this protein is the cellular apoptosis–susceptibility (CAS) protein (Brinkmann *et al*., 1995) is associated with both apoptosis and cell proliferation. It is stated that CAS participates in apoptosis when correct chromosome segregation cannot be achieved (Brinkmann, 1998). CSE1 homolog is also found in plants (putative cellular apoptosis susceptibility protein in *Arabidopsis thaliana*, Appendix F). CSE1 interaction with the YR10-1 brings the idea that Yr10 protein might be controlling the cell cycle and apoptosis in plant cell, in response to pathogen attack.

SRP72 is the other protein belonging to this group as it is post translationally modified during apoptosis (discussed 3.15.1).

3.15.5 Protein folding activities

4L2-B4 codes for **CPR7** (cyclophilin-like protein) in yeast which is a co-chaperone of HSP90, a protein chaperone that functions in several signal transduction pathways (including R protein mediated plant disease resistance response) (Takahashi *et al*., 2003). Also, CPR7 mutants show a slow-growth phenotype (Duina *et al*., 1996). The interaction of CPR7 with YR10-1 raises the idea that, Yr10 may mediate the activity of HSP90 through its co-chaperone CPR7 (Figure 3.49 B). Recently, in a study it is stated that Rar1 and HSP90 may determine whether NB-LLR proteins are functional in disease resistance or marked for degradation (Holt *et al*, 2005). Cytosolic HSP90 binds to proteins and responsible for the proper folding, which is regulated by ATP binding and hydrolysis. This function is controlled by the co-chaperones. Also, the specificity of HSP90 arises from these co-chaperones (reviewed by Holt *et al*,

2005). In turn, CPR7 plant homolog might be acting as co-chaperone in the YR10 maintenance (Figure 3.49 A).

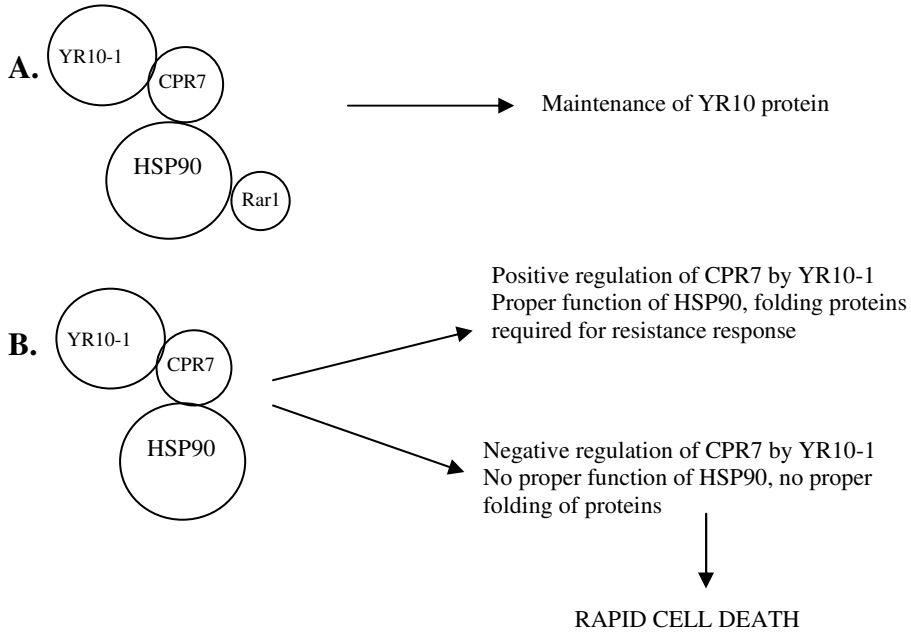


Figure 3.49 Possible functions of CPR7-YR10-1 interaction. **A:** HSP90-Rar1 complex may interact with the Yr10 protein *via* CPR7 plant homolog, for the maintenance of Yr10 protein during resistance response. **B:** Yr10-CPR7 plant homolog interaction may positively or negatively control the activity of HSP90 during the R protein mediated disease responses.

The involvement of cyclophilins in plant disease response is discussed in a study, which shows that a pepper cyclophilin gene is differentially expressed (induced)

during the pathogen infection (Kong *et al*., 2001). They discussed that cyclophilins may have the role of accelerating the folding and maturation of the newly synthesized proteins induced during stress.

3.15.6 Cell cycle controls

Factor arrest protein 1 (**Far1**) (colony number, 5K1-A1; Rad6 interactor) has two functions, it is a cyclin-dependent kinase inhibitor that mediates cell cycle arrest in response to pheromone and also forms a complex with Cdc24p, Ste4p, and Ste18p establish cell polarity during yeast mating (Shimada *et al*., 2000). Pheromones activate a mitogen-activated protein (MAP) kinase signal transduction pathway altering gene transcriptions, cellular morphology and as a result cell cycle is arrested in G_1 (Gustin *et al*., 1998). If pheromones are not available Far1 is phosphorylated at serine 87 by Cdc28. After phosphorylation it becomes a target of SCF^{Cdc4} and degraded as a result of ubiqutinylation (O'Shea and Herskowitz, 2000). Other role of Far1 is the inhibition of cyclin-dependent kinase (Cdc28) (Figure 3.50 B). Far1 is phosphorylated by Fus3 this increases it's binding to the Cdc28–Cln complex (Peter *et al*., 1993; Gartner *et al*., 1998). Far1 is shown to interact with the F-box protein Cdc4. It is also known that SCF is responsible for its degradation. According to our results, Far1 interacts with the Rad6 ubiqutin conjugating enzyme (Figure 3.50 A). This result is important as it shows that Rad6 is involved in the SCF (Section 1.2.9). It is known that SGT1, which is involved in plant disease resistance pathways, functions in SCF-mediated ubiquitination. Our findings generate the idea that Rad6 may also work together with SCF complex in *R* gene mediated disease resistance signaling. Since, SCF^{Cdc4} complex is required to degrade Far1, Cdc6 and Gcn4 (Deshaies, 1999), this finding also supports our finding that Gcn4 also found to be interacting with Rad6 (section 3.15.1) (Kornitzer *et al*., 1994).

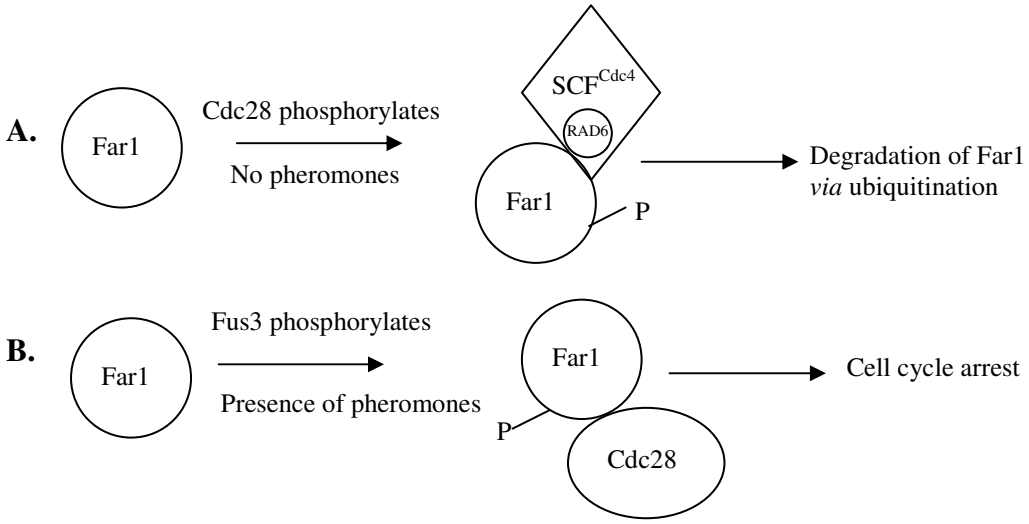


Figure 3.50 Far1 protein level controls in yeast cell. **A:** In the absence of pheromones it is degraded by SCF complex via ubiquitination. The SCF complex involves an E3 (ubiquitin ligase) complex (F-Box [Cdc34 in this case], a cullin, Rbx, Skp), and an E2 (ubiquitin conjugating enzyme). In this study, E2 of this complex is found as the RAD6 (as it interacts with the Far1). **B:** In the presence of pheromones phosphorylated Far1 binds to Cdc28 and arrests cell cycle.

3.15.7 Oxidative stress responses

OXR1 (Colony number: 5L1-A6; interactor of both WR5 and Rad6) has no known function in yeast but it is required for normal levels of resistance to oxidative damage, and mutants show sensitivity to hydrogen peroxide (Volkert *et al*., 2000). This protein is much conserved in eukaryotes. Recently human OXR1 is identified. In the study, it is found that OXR1 is induced during heat and oxidative response and targeted to mitochondria which are the major source of ROS within the cell (Elliott and Volkert,

2004). In our study, OXR1 interacts with Rad6 that generates the idea that OXR1 wheat homolog might be degraded during hypersensitive response in plant disease resistance mechanism. By this mechanism the cell would be more susceptible to the reactive oxygen species and facilitates the cell death. As we do not know the function of theWR5 gene fragment, we can only speculate that OXR1 might be controlled by a WD40 repeat containing protein.

3.15.8 Other interactors with important roles

Among the interactors of YR10-1 another promising one is the, 4L2-F5 which codes for the Arrest-defective protein 1 (**ARD1**) in yeast. The protein, N-terminally acetylates many proteins, which influences multiple processes such as the cell cycle, heat-shock resistance, mating, sporulation, and telomeric silencing (Park and Szostak, 1992; Mullen *et al.*, 1989; Polevoda *et al.*, 1999; Whiteway *et al.*, 1987; Aparicio *et al.*, 1991). It has the conserved domains Acetyltransf_1 and RimI. The plant homolog of this gene sequence is silencing group B protein of *Zea mays*. This finding states that there might be silencing of some proteins though acetylation mechanism.

Pyruvate carboxylase 2 **(PYC2)** (clone number: 1K2-F5) is another significant interactor of the YR10-1. This protein is important in our point of view because it is found that PYC2 expression increased 3.6 folds to provide cells with an adaptation function in response to oxidative stress (Chauhan *et al*., 2003). Yr10 protein somehow may be responsible for the inhibition of this increase, so that cell will not be able to adapt oxidative stress and die as expected in the HR.

1K1-F5 codes for **LAS17** which is an actin assembly factor. In addition to cell cycle-dependent actin regulation, yeast change their physiological state by reorganizing

the actin cytoskeleton in response to various stimuli, including mating pheromones, heat stress, salt stress, changes in nutrient composition, and so on (Naqvi *et al*., 1998). LAS17 activates the Arp2/3 protein complex. The Arp2/3 complex is essential for several actin-mediated cellular and developmental functions in eukaryotes, including plants (Kandasamy *et al*., 2004). Thus, LAS17 might be related with disease resistance as actin assembly is very important for many events in the cell.

4K1-G2 codes for the Hypothetical UPF0026 protein **YPL207w**. This protein has the conserved domains of Radical_SAM, COG0731 (Fe-S oxidoreductases), COG2100 (predicted Fe-S oxidoreductase). The sequence gives hit to the plant proteins, flavodoxin family protein / radical SAM domain-containing protein from *Arabidopsis thaliana* and putative oxidoreductase from *Oryza sativa* (Appendix E). Radical-SAM superfamily of proteins uses iron-sulfur clusters and S-adenosylmethionine to initiate H atom abstraction reactions (Sofia *et al*., 2001). This protein interacts with the YR10-1 and this interaction states that there might be a control of Yr10 on the radical formation.

MAF1 (colony number: 5L1-A1, interactor of Rad6) is a negative regulator of RNA polymerase III and component of several signaling pathways that repress polymerase III transcription in response to changes in cellular environment (Pluta et al., 2001; Upadhya *et al*.,2002). This Rad6 may be involved in transcriptional control during alterations in the cellular environment.

Glycosylphosphatidylinositol (GPI) anchoring is a conserved post-translational modification to attach cell surface proteins to plasma membrane in eukaryotes. **GWT1** (GPI-anchored wall transfer protein 1) (colony number: 5L2-G5, WR5 interactor) is involved in GPI anchor biosynthesis; it is required for inositol acylation in yeast. Glycosylphosphatidylinositol (GPI)-anchored cell wall mannoproteins are required for

the adhesion of pathogenic fungi, such as *Candida albicans*, to human epithelium. In a study they stated that small molecular inhibitors of the cell surface presentation of GPI-anchored mannoproteins would be candidate drugs to block the establishment of fungal infections (Tsukahara et al., 2003). According to our results GWT1 might be inhibited by the action of WR5 and fungal (*Puccinia striiformis*) adhesion to the plant cell might be blocked.

CHAPTER 4

CONCLUSIONS

Stripe (yellow) rust, caused by *Puccinia striiformis* is one of the most severe leaf diseases of wheat. The disease results in great losses during epidemics. Although there are chemical control agents, this approach is not preferable due to high cost and environmental hazard potentials. Developing and planting the resistant cultivars are the most convenient approach. To develop resistance cultivars, the *R* gene mediated disease resistance must be fully understood. R protein recognition of the pathogen elicitor is the very first reaction that starts a resistance cascade. Understanding the downstream signaling of this recognition is as important as the discovering the *R* genes themselves. Although the *R* genes carry similar conserved domains, the *R* gene mediated recognition of the pathogen by plant is very specific to the pathogen. The mode of action varies based on the plant and the pathogen, even the strain of the pathogen.

One strategy for achieving durable resistance might be by using downstream signaling components that are common for many disease resistance pathways. For this reason, the aim of this study is to find some possible conserved proteins having important roles in the *R* gene mediated yellow rust disease resistance in wheat.

The dual bait yeast two-hybrid system is used in order to screen the yeast library by the baits unique to this study, which are designed for understanding the signal transduction pathway in disease resistance. During the *R* gene mediated disease

response, hyper sensitive cell death is the first thing that happens in the plant cell following recognition of an avirulent signal by the R gene product. This hypersensitive cell death leads to the inhibition of the pathogen spreading. There might be a conserved pathway for the plant cell to develop sudden hypersensitive response. The reason of screening a yeast library is that, the yeast is the simplest eukaryote. If the hypersensitive cell death is that much conserved, it is possible to find a model pathway in yeast.

Yr10 gene (Accession number: AF149114) is the only one which is cloned as a wheat stripe rust resistance gene. According to the sequence information four baits were designed in the correct open reading frame, and the yeast library was screened for the interacting proteins with the *Yr10* gene fragments. WR5 sequence, which was found in our laboratory during differential display studies previously (Bozkurt and Akkaya, unpublished data), is used as a bait to understand its role in disease resistance mechanism. The *Rad6* gene, which is highly conserved from yeast to mammalian genome (Yamamoto *et al*., 2004), was also found in our laboratory as an induced gene during resistance to stripe rust disease in Avocet-Yr10 plant (Bozkurt, *et al*., unpublished data). To understand the function of the *Rad6* gene in resistance pathway, it is important to find its potential interacting partners.

No interacting proteins with the YR10-1506 and YR10-1563 gene fragments were found. This might be due to large sizes of the fragment sizes of the insert. This type of observation with large fragments is common in yeast two-hybrid analysis. Additionally, the proteins might be not folded properly in yeast (Serebriiskii, personal communication). There were a few colonies grown on the plates after screening, 3 of these colonies were sequenced and all gave hit to the same protein, the Extracellular Matrix Protein 9 (ECM9). It was thought to be a kind of contamination because it did not activate the expression of the reporter genes (*LEU2* and *LacZ*).

YR10-1 gene fragment has the conserved domains of two putative leucine zipper domains and p-loop domain. The two leucine zipper domains are involved in the coiled coil (CC) domain of the Yr10 protein. The studies on TIR region of Arabidopsis R genes revealed that TIR region is involved in signaling rather than recognition of the Avr protein (Aarts *et al*., 1998; Van der Biezen *et al*., 2002; Warren *et al*., 1999). As the TIR region is replaced by CC structure in some of the NB-LLR type R proteins, this suggests that CC domain may be involved in signaling rather than in recognition (Martin *et al*., 2003). P-loop domain, another domain in YR10-1 gene fragment, is highly conserved in R proteins. The domain is also called as the kinase1a and functions as the AP (apoptotic)-ATPase. Analyses of the animal apoptotic proteins APAF/Ced4 and plant pathogen-resistance ATPases led defined them as AP-ATPase family (Van der Biezen and Jones, 1998). The animal members of the AP-ATPase family, like nematode CED4 and mammalian Apaf-1, are involved in cell-death signaling by activating caspases (Chaudhary *et al*., 1998).

YR10-2 gene fragment has the kinase domain 3a, domain 3, domain Q/EGF, and domain HD as the conserved domains. Kinase 3a is the part of NB conserved domain and other domains are collectively called as the ARC (Apaf-1, R protein and CED4 homology) domain (Appendix B).

Total 13 and 10 positive clones were sequenced as the interactors of YR10-1 and YR10-2 gene fragments, respectively. Among these interactors, some of them are identified as extremely promising ones that may have role in *R* gene mediated disease resistance in plants. The most important interacting proteins with *Yr10* gene fragments are:

- **SRP72** (YR10-1 interactor), is post translationally modified during apoptosis in human (Utz *et al*., 1998), is thought to be having a role in R protein mediated disease response. SRP72 plant homolog might be modified in response to yellow rust infection by the Yr10 protein, so that it possibly regulates the programmed cell death of the wheat cell.

- **Gcn2p** and **Eap1p** (YR10-2 interactors) proteins are related with translation initiation controls. Their interaction with the Yr10 protein raises the idea that Yr10 protein might regulate the translation initiation during the R protein mediated disease resistance response.

- **SKY1** (YR10-2 interactor) and **GAC1** (YR10-1 interactor) proteins have kinase and phosphatase regulator activities, respectively. Plant homologs of these proteins might control the signaling events during R protein mediated resistance.

- **TOM1** (YR10-1 interactor) is a yeast E3 ligase. The interaction of TOM1 with Yr10 protein raise the suggestion that plant homolog of TOM1 might be controlling the Yr10 protein level in the plant cell.

- **POR1** (YR10-2 interactor) is a mitochondrial porin. Yr10-POR1 interaction raise the idea that this *R* gene can have a role in cell death by controlling the mitochondria regulated apoptosis.

- **CSE1** (YR10-2 interactor) which has the plant homolog putative cellular apoptosis susceptibility protein in *Arabidopsis thaliana*, might be controlling the apoptosis in plant cell via Yr10 protein.

- **CPR7** (YR10-1 interactor) is a co chaperone of the HSP90 in yeast. The CPR7 plant homolog might be CPR7 plant homolog might be acting as co-chaperone in the Yr10 maintenance.

- **ARD1** (YR10-1 interactor) shows homology to the silencing group B protein of *Zea mays*. This finding and its interaction with Yr10 protein states that there might be silencing of some proteins though acetylation mechanism during resistance responces mediated by Yr10 protein.

- **PYC2** (YR10-1 interactor) protein level increases in response to oxidative stress (Chauhan *et al*., 2003). Yr10 protein somehow may be responsible for the inhibition of this increase in plant cell, so that the cell will not be able to adapt oxidative stress and die as expected in the HR.

- **LAS17** (YR10-1 interactor) is an actin assembly factor. Its plant homolog might be having an important role in pathogen resistance response.

Total 8 and 3 positive clones were sequenced as the interactors of Rad6 protein and WR5 gene fragment, respectively. Among these interactors the most important interactions are:

- **FAR1** (Rad6 interactor) mediates cell cycle and its interaction with Rad6 protein raises the idea that FAR1 might be degraded via Rad6 involved ubiqutination.

- **OXR1** (both Rad6 and WR5 interactor) it is found in a study that human OXR1 is induced during heat and oxidative response and targeted to mitochondria which are the major source of ROS within the cell (Elliott and Volkert, 2004).

This raises the idea that plant OXR1 might be degraded in during the plant disease resistance responses.

- **MAF1**-Rad6 interaction increases the thought that Rad6 may be involved in transcriptional control during alterations in the cellular environment.

- **GWT1** (WR5 interactor) is a GPI-anchored wall transfer protein. Recently in a study it is found that these proteins are needed for adhesion of the pathogenic fungi to human epithelium (Tsukahara et al., 2003). According to our results GWT1 might be inhibited by the action of WR5 and fungal (*Puccinia striiformis*) adhesion to the plant cell might be blocked.

The importance of this study arises from its being the first yeast two hybrid analysis of a wheat *R* gene in literature up to date. The study was able to detect interacting partners with anticipated functions. The interactions found in this study not only enlighten the possible mechanisms in Yr10 protein mediated disease resistance in wheat, but also initiates further detailed analysis of the key interactors.

REFERENCES

Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daniels, M.J., Parker, J.E. **(1998)** "Different requirements for EDS1 and NDR1 by disease resistance genes define at least two *R* gene mediated signaling pathways in *Arabidopsis*" *Proc. Natl. Acad. Sci. USA* **95**: 10306-10311

Aparicio, O.M., Billington, B.L., and Gottschling, D.E. **(1991)** "Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*" *Cell* **66**:1279-1287

Austin, M.J., Muskett, P., Kahn, K., Feys, B.J., Jones, J.D., and Parker, J.E. **(2002)** "Regulatory role of SGT1 in early *R* gene-mediated plant defenses" *Science* **295**: 2077-2080

Ayliffe M.A. and Lagudah E.S. **(2004)** "Molecular genetics of disease resistance in cereals" *Annals of Botany* **94**: 765-773

Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K., and Schulze-Lefert, P. **(2002)** "The RAR1 interactor SGT1, an essential component of *R* gene-triggered diseaseresistance" *Science* **295**: 2073-2076

Bayles, R.A. and Stigwood, P.L. **(1995)** "Yellow rust of wheat." *Annu. Report, UK Cereal Pathogen Virulence Survey*, pp 13–17

Bayles, R.A. and Stigwood, P.L. **(1996)** "Yellow rust of wheat." *Annu. Report, UK Cereal Pathogen Virulence Survey*, pp 13–17

Bayles, R.A.and Stigwood, P.L. **(1994)** "Yellow rust of wheat." *Annu. Report, UK Cereal Pathogen Virulence Survey*, pp 16–21

Bodanove, A.J. **(2002)** "Protein-protein interactions in pathogen recognition by plants" *Plant Mol. Biol.* **50**: 981-989

Braun, H.J. and Saari, E.E. **(1992)** "An assessment of the potential of *Puccinia striiformis* f.sp. *tritici* to cause yield losses in wheat on the Anatolian Plateau of Turkey." *Proc. of the 8th. Europ. and Med. Cereal Rusts and Mildew Conf.* pp 121-124

Brinkmann, U., Brinkmann, E., Gallo, M., and Pastan, I. **(1995)** "Cloning and characterization of a cellular apoptosis susceptibility gene, the human homologue to the yeast chromosome segregation gene CSE1" *Proc. Natl. Acad. Sci. USA* **92**: 10427-10431

Brinkmann*, U.* **(1998)** "*CAS,* the human homologue of the yeast chromosome-segregation gene *CSE1,* in proliferation, apoptosis, and cancer" *Am. J. Hum. Genet.* **62**: 509-513

Brown,J.D., Hann, B.C., Medzihradszky, K.F., Niwa, M., Burlingame, A.L., Walter, P.**(1994)** "Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression" *EMBO J*. **18**: 4390-400

Brueggeman, R., Rostoks, N., Kudrna, N., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., Kleinhofs, A. **(2002)** "The barley stem rust resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases" *Proc. Natl. Acad. Sci. USA* **95**: 9328-9333

Century, K.S., Holub, E.B., and Staskawicz, B.J. **(1995)** "*NDR1*, a locus of *Arabidopsis thaliana* that is required for disease resistance to both a bacterial and a fungal pathogen" *Proc. Natl. Acad. Sci. USA* **92**: 6597-6601

Chaudhary, D., O'Rourke, K., Chinnaiyan, A.M. and Dixit, V.M. **(1998)** "The death inhibitory molecules CED-9 and CED-4L use a common mechanism to inhibit the CED-3 death protease" *J. Biol. Chem.* **273**: 17708-17712

Chauhan, N., Inglis, D., Roman, E., Pla, J., Li, D., Calera, J.A., and Calderone, R. **(2003)** *"Candida albicans* response regulator gene *SSK1* regulates a subset of genes whose functions are associated with cell wall biosynthesis and adaptation to oxidative stress" *Eukaryotic Cell* **2**: 1018-1024

Ciechanover, A. **(1994)** "The ubiquitin–proteasome proteolytic pathway" *Cell* **79**: 13-21

Cigan, M.A., Bushman, J.L., Boal, T.R., and Hinnebusch, A.G. **(1993)** "A protein complex of translational regulators of GCN4 mRNA is the guanine nucleotide-exchange factor for translation initiation factor 2 in yeast" *Proc. Natl. Acad. Sci. USA* **90**: 5350-5354

Clemens, M.J. **(1997)** "PKR-A protein kinase regulated by double-stranded RNA" *Int. J. Biochem. Cell Biol.* **29**: 945-949

Cohn, J., Sessa, G. and Martin, G.B. **(2001)** "Innate immunity in plants" *Curr. Opin Immunol* **13**: 55-62

Cooley, M.B., Pathirana, S., Wu, H.J., Kachroo, P. and Klessig,D.F. **(2000)** "Members of the *Arabidopsis* HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens" *Plant Cell* **12**: 663-676

Cosentino, G.P., Schmelzle, T., Haghighat, A., Helliwell, S.B., Hall, M.N., and Sonenberg, N. **(2000)** "Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in *Saccharomyces cerevisiae*" *Mol. Cell. Biol.* **20**: 4604-4613

Craig, K.L., and Tyers, M. **(1999)** "The F-box: a new motif for ubiquitindependent proteolysis in cell cycle regulation and signal transduction" *Prog. Biophys. Mol. Biol.* **72**: 299-328

Dangl, J.L. and Jones, J.D.G. **(2001)** "Plant pathogens and integrated defense responses to infection" *Nature* **411**: 826-833

de Haro, C., Mendez, R., and Santoyo, J. **(1996)** "The eIF-2alpha kinases and the control of protein synthesis" *FASEB J.* **10**:1378-1387

Deloche, O., de la Cruz, J., Kressler, D., Doere, M., and Linder, P. **(2004)** "Membrane transport defect leads to a rapid attenuation of translation initiation in *Saccharomyces cerevisiae*" *Mol. Cell* **13**: 357-366

Dempsey, D.A., Shah, J., and Klessig, D.F. **(1999)** "Salicylic acid and disease resistance in plants" *Crit. Rev. Plant Sci.* **18**: 547-575

Deshaies, R.J. **(1999)** "SCF and Cullin/Ring H2–based ubiquitin ligases" *Annu. Rev. Cell Dev. Biol.* **15**: 435-467

Dever, T.E., Feng, L., Wek, R.C., Cigan, A.M., Donahue, T.F., and Hinnebusch, A.G. **(1992)** "Phosphorylation of initiation factor 2 alpha by protein kinase *GCN2* mediates gene-specific translational control of GCN4 in yeast" *Cell* **66**: 585-596.

Dever, T.E. **(2002)** "Gene-specific regulation by general translation factors" *Cell* **108**: 545-556.

Diamond, J. **(1997)** "Location, location, location: The first farmers" *Science* **278**: 1243-1244

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K., Jones, J.D.G. **(1996)** "The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine rich repeat proteins" *Cell* **84**: 451-459

Dodds, P., Lawrence, G., and Ellis, J. **(2001)** "Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax" *Plant Cell* **13**:163-178

Donze, O., and Picard, D. **(1999)** "Hsp90 binds and regulates the ligand-inducible α subunit of eukaryotic translation initiation factor kinase Gcn2" *Mol. Cell Biol.* **19**: 8422-8432

Duina, A.A., Marsh, J.A., and Gaber, R.F. **(1996)** "Identification of two CyP-40-like cyclophilins in *Saccharomyces cerevisiae*, one of which is required for normal growth" *Yeast* **12**: 943-952

Durfee, T., Becherer, K., Chen, P.-L., Yeh, S.-H., Yang, Y., Kilburn, A. E., Lee, W.-H., and Elledge, S.J. **(1993) "**The retinoblastoma protein associates with the protein phophatase type 1 catalytic subunit" *Genes and Dev*. **7**: 555-569

Elliott, N.A., and Volkert, M.R. **(2004)** "Stress induction and mitochondrial localization of Oxr1 proteins in yeast and humans" *Mol. Cellu. Biol.* **24**:3180–3187

Ellis, J.G., Lawrence, G.J., Luck, J.E., and Dodds, P.N. **(1999)** "Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity" *Plant Cell* **11**: 495–506

Erez, O., and Kahana, C. **(2002)** "Deletions of SKY1 or PTK2 in the *Saccharomyces cerevisiae* trk1Δ trk2Δ mutant cells exert dual effect on ion homeostasis" Biochem. Biophy. Res. Commu. **295**: 1142-1149

Falk, A., Feys, B.J., Frost, L.N., Jones, J.D., Daniels, M.J., Parker, J.E. **(1999)** "EDS1, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases" *Proc. Natl. Acad. Sci. USA* **96**: 3292-3297

FAOSTAT data, **(2004)**
http://faostat.fao.org/faostat/servlet/XteServlet3?Areas=862&Items=%3E1717&Elements=51&Years=2004&Format=Table&Xaxis=Years&Yaxis=Countries&Aggregate=&Calculate=&Domain=SUA&ItemTypes=Production.Crops.Primary&language=EN, last accessed August 2005

FAOSTAT data, **(2004)**
http://faostat.fao.org/faostat/servlet/XteServlet3?Areas=223&Items=%3E1717&Elements=31&Years=2004&Format=Table&Xaxis=Years&Yaxis=Countries&Aggregate=&Calculate=&Domain=SUA&ItemTypes=Production.Crops.Primary&language=EN, last accessed August 2005

Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat A., and Keller, B. **(2003)** "Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome" *Proc. Natl. Acad. Sci. USA* **100**: 15253-15258

Feys, B.J. and Parker, J.E. **(2000)** "Interplay of signaling pathways in plant disease resistance" *Trends Genet*. **16**: 449-455

Feys, B.J., Moisan, L.J., Newman, M.A., and Parker, J.E. **(2001)** "Direct interaction between the *Arabidopsis* disease resistance signaling proteins, EDS1 and PAD4" *EMBO J.* **20**: 5400-5411

Fields, S., and Song, O-K. **(1989)** "A novel genetic system to detect protein-protein interactions" *Nature* **340**: 245-246

Flor, H. **(1971)** "Current status of the gene-for-gene concept" *Annu. Rev Phytopathol*, **9**: 275-296

Fredrick, K.L., and Helmann, J.D. **(1994)** "Dual chemotaxis signaling pathways in *Bacillus subtilis*: a α-dependent gene encodes a novel protein with both CheW and CheY homologous domains" *J. Bacteriol.* **176**: 2727-2735

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., Ohme-Takagi, M. **(2000)** "Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box mediated gene expression" *Plant Cell* **12**: 393-404

Gartner, A., Jovanovic, A., Jeoung, D.I., Bourlat, S., Cross, F.R. and Ammerer, G. **(1998)** "Pheromone-dependent G_1 cell cycle arrest requires Far1 phosphorylation, but

may not involve inhibition of Cdc28-Cln2 kinase, *in vivo*" *Mol. Cell. Biol.* **18**: 3681–3691

Gökgöl, M. **(1935)** *Türkiye Buğdayları*. Tarım Bakanlığı İstanbul Yeşilköy Tohum Islah ve Yetiştirme Enstitüsü Yayını **7** Devlet Matbaası, İstanbul

Golemis, E.A. **(2002)** "A yeast two-hybrid system based on transcriptional activation" in Golemis E.A. (ed) *Protein-Protein Interactions, A molecular cloning manual*. Cold Spring Harbor Laboratory Press, New York, pp96

Grosshans, H., Deinert, K., Hurt, E., and Simos, G. **(2001)** "Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export" *J. Cell Biol.* **4**: 745-62

Guarente, L. **(1996)** "Transcriptional coactivators in yeast and beyond" *Trends in Biochem*. **20**: 517-521
Gustin, M.C., Albertyn, J., Alexander, M. and Davenport, K. **(1998)** "MAP kinase pathways in the yeast *Saccharomyces cerevisiae*" *Microbiol. Mol. Biol. Rev.* **62**: 1264-1300

Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. **(1993)** "Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2" *Cell* **75**: 791-803

Hammond-Kosack, K. and Jones, J. **(1997)** "Plant disease resistance genes" *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**: 575-608

Hauser, S., Adelmant, G., Sarraf, P., Wright, H.M., Mueller, E., and Spiegelman, B.M. **(2000)** "Degradation of the peroxisome proliferatoractivated receptor gamma is linked to ligand-dependent activation" *J. Biol. Chem.* **275**: 18527-18533

Hengartner, M.O. **(2000)** "The biochemistry of apoptosis" *Nature* **407**: 770-776

Hershey, J.W.B. **(1991)** "Translational control in mammalian cells" *Annu Rev Biochem*. **60**: 717-755.

Hershko, A., and Ciechanover, A. **(1998)** "The ubiquitin system" *Annu Rev. Biochem.* **67**: 425-479

Heun M., Schafer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B. and Salamini, F. **(1997)** "Site of einkorn wheat domestication identified by DNA fingerprinting" *Science* **278**: 1312-1314

Hinnebusch, A.G. **(1996)** "Translational control of GCN4: gene-specific regulation by phosphorylation of eIF-2". *In* J. W. B. Hershey, M. B. Mathews, and N. Sonenberg (ed.), *Translational control*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. pp 199-244

Hiyama, H., Yokoi, M., Masutani, C., Sugasawa, K., Maekawa, T., Tanaka, K., Hoeijmakers, J., and Hanaoka, F. **(1999)** "Interaction of hHR23 with S5a. The ubiquitin-like domain of hHR23 mediates interaction with S5a subunit of 26 S proteasome" *J. Biol. Chem.* **274**: 28019-28025

Hogg, W.H., Hounam, C.E., Mallik, A.K. and Zadoks, J.C. **(1969)** "Meteorological factors affecting the epidemiology of wheat rusts" *WMO Tech. Note No. 99* pp 143

Holt, B.F., Belkhadir, Y., and Dangl, J.L. **(2005)** "Antagonistic control of disease resistance protein stability in the plant immune system" *Science* **309**: 929-932

Huang, L., Brooks, S.A., Li, W.L., Fellers, J.P., Trick, H.N., and Gill, B.S. **(2003)** "Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat" *Genetics* **164**: 655-664

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J. **(1998)** "Multiple sequence alignment with Clustal X" *Trends Biochem. Sci.* **23**: 403–405

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P., Valent, B. **(2000)** "Direct interaction of resistance gene and avirulence gene products confers rice blast resistance" *EMBO J.* **19**: 4004-4014

Jirage, D., Tootle, T.L., Reuber, T.L., Frost, L.N., Feys, B.J., *et al*. **(1999)** "*Arabidopsis thaliana PAD4* encodes a lipase-like gene that is important for salicylic acid signaling" *Proc. Natl. Acad. Sci. USA* **96**: 13583-13588

Kandasamy, M.K., Deal, R.B., McKinney, E.C., and Meagher, R.B. **(2004)** "Plant actin-related proteins" *TRENDS Plant Sci.* **9**: 196-202

Kim, Y.J., Lin, N.C., and Martin G.B. **(2002)** "Two distinct *Pseudomonas* effector proteins interact with the Pto kinase and activate plant immunity" *Cell* **109**: 589-598

Kitagawa, K., Skowyra, D., Elledge, S.J., Harper, J.W., and Hieter, P. **(1999)** "SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex" *Mol. Cell* **4**: 21-33

Kong, H.Y., Lee, S.C., and Hwang, B.K. **(2001)** "Expression of pepper cyclophilin gene is differentially regulated during the pathogen infection and abiotic stress conditions" *Physiol. Mol. Plant Path*. **59**: 189-199

Kornitzer, D., Raboy, B., Kulka, R.G., and Finkl, G.R. **(1994)** "Regulated degradation of the transcription factor Gcn4" *EMBO J.* **13**: 6021-6030

Koyama, T., Okada, T., Kitajima, S., Ohme-Takagi, M., Shinshi, H., and Sato, F. **(2003)** "Isolation of tobacco ubiquitin-conjugating enzyme cDNA in a yeast two-hybrid system with tobacco ERF3 as bait and its characterization of specific interaction" *J. Experimental Botany* **54**: 1175-1181

Kruger, J., Thomas, C.M., Golstein, C., Dion, M.S., Smoker, M., *et al.* **(2002)** "A tomato cysteine protease required for Cf-2- dependent disease resistance and suppression of autonecrosis. *Science* **296**: 744–747

Lee, A.C., Xu, X., Blachly-Dyson, E., Forte, M., and Colombini, M. **(1998)** "The role of yeast VDAC genes on the permeability of the mitochondrial outer membrane. *J. Membr. Biol.* **161**: 173-181

Laroche, A., Eudes, F., Frick, M.M., Huel, R., Nykiforuk, C.L., Conner, R.L., Kuzyk, A., Acharya, S. and Jordan, M. **(2002)** "A wheat resistance gene against stripe rust" *Can. J. Plant Pathol.* **24**: 504-507

Lehmann, P. **(2002)** "Structure and evolution of plant disease resistance genes" *J. Appl. Genet.* **43**(4): 403-414

Leister, R.T., and Katagiri, F. **(2000)** "A resistance gene product of the nucleotide binding site-leucine rich repeats class can form a complex with bacterial avirulence proteins *in vivo*" *Plant J.* **22**: 345-54

Liu, Y., Schiff, M., Marathe, R., and Dinesh-Kumar, S.P. **(2002)** "Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus" *Plant J.* **30**: 415-429

Long, J. and Sobhanifar, S. **(2005)** "The Yeast Two-Hybrid Assay an Exercise in Experimental Eloquence"
http://www.bioteach.ubc.ca/MolecularBiology/AYeastTwoHybridAssay/, last accessed on August 2005

Lowe T., Sharefkin, J., Yang, S.Q., and Dieffenbach, C.W. **(1990)** "A computer program for selection of oligonucleotide primers for polymerase chain reactions" Nucleic Acids Res. **18**: 1757–1761

Mackey, D., Holt, B.F., Wiig, A., Dangl, J.L. **(2002)** "RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*" *Cell* **108**: 743-754

Marchler-Bauer, A., and Bryant, S.H. **(2004)** "CD-Search: protein domain annotations on the fly" *Nucleic Acids Res.* **32**: W327-331

Martin, G.B., Bogdanove, A.J., and Sessa, G. **(2003)** "Understanding the functions of plant disease resistance proteins" *Annu. Rev. Plant Biol.* **54**: 23-61

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., *et al.* **(1993)**. "Map-based cloning of a protein kinase gene conferring disease resistance in tomato" *Science* **262**: 1432-1436

Matsuzawa, S.-I., and Reed, J.C. **(2001)** "Siah-1, SIP, and Ebi collaborate in a novel pathway for β-catenin degradation linked to p53 responses" *Mol. Cell* **7**: 915-926

Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral, B.W. and Young, N.D. **(1999)** "Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily" *Plant J.* **20**: 317-332

Miura, S., Zou, W., Ueda, M., aand Tanaka, A. **(2000)** "Screening of genes involved in isooctane tolerance in *Saccharomyces cerevisiae* by using mRNA differential display" *Appl. Env. Microbiol.* **66**: 4883-4889

Mullen, J.R., Kayne, P.S., Moerschell, R.P., Tsunasawa, S., Gribskov, M., Colavito-Shepanski, M., Grunstein, M., Sherman, F., and Sternglanz, R. **(1989)** "Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast" *EMBO J* **8**: 2067-2075

Mur, L.A., Naylor, G., Warner, S.A.J., Sugars, J.M., White, R.F., and Draper, J. **(1996)** "Salicylic acid potentiates defense gene expression in tissue exhibiting acquired resistance to pathogen attack" *Plant J.* **9**: 559-571

Muskett, P., and Parker, J. **(2003)** "Role of *SGT1* in the regulation of plant *R* gene signalling" *Microbes and Infection* **5**: 969-976

Muskett, P.R., Kahn, K., Austin, M.J., Moisan, L.J., Sadanandom, A., *et al*. **(2002)** "*Arabidopsis* RAR1 exerts rate-limiting control of *R* gene-mediated defenses against multiple pathogens" *Plant Cell* **14**: 979-992

Naqvi, S.N., Zahn, R., Mitchell, D.A., Stevenson, B.J., Munn, A.L. **(1998)** "The WASP homologue Las17p functions with the WIP homologue End5p/verprolin and is essential for endocytosis in yeast" *Curr Biol* **8**: 959-962

Newcomb, L.L., Hall, D.D., and Heideman, W. **(2002)** "AZF1 is a glucose-dependent positive regulator of *CLN3* transcription in *Saccharomyces cerevisiae*" *Mol. Cellu. Biol.* **22:** 1607-1614

O'Shea, E.K. and Herskowitz, I. **(2000)** "The ins and outs of cell-polarity decisions" *Nature Cell Biol.* **2**: E39–E41

Pan, Q., Wendel, J. and Fluhr, R. **(2000)** "Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes" *J. Mol. Evol.* **50**: 203-213

Park, E.C. and Szostak, J.W. **(1992)** "ARD1 and NAT1 proteins form a complex that has N-terminal acetyltransferase activity" *EMBO J.* **11**: 2087-2093

Parrou, J. L., Teste, M. A., and Francois, J. **(1997)** "Effects of various types of stress on the metabolism of reserve carbohydrates in Saccharomyces cerevisiae: genetic evidence for a stress-induced recycling of glycogen and trehalose" *Microbiology* **143**: 1891-1900

Peart, J.R., Lui, R., Sadanandom, A., Malcuit, I., Moffett, P., Brice, D.C., Schauser, L., Jaggard, D.A. W., Xiao, S., Coleman, M.J., Dow, *M.,* Jones J.D.G., Shirasu, K.. Baulcombe, D.C. **(2002)** "Ubiquitin ligase-associated protein SGT1 is required for host and nonhost disease resistance in plants" *Proc. Acad. Nat. Sci USA* **99**: 10865-10869

Peter, M., Gartner, A., Horecka, J., Ammerer, G. and Herskowitz, I. **(1993)** "FAR1 links the signal transduction pathway to the cell cycle machinery in yeast" *Cell* **73**: 747-760

Picard, D. **(2002)** "Heat-shock protein 90, a chaperone for folding and regulation" *Cell. Mol. Life Sci.* **59:** 1640–1648

Pickart, C.M. **(2001)** "Ubiquitin enters the new millennium" *Mol. Cell* **8**: 499-504

Polevoda, B., Norbeck, J., Takakura, H., Blomberg, A., and Sherman, F. **(1999)** "Identification and specificities of N-terminal acetyltransferases from *Saccharomyces cerevisiae*" *EMBO J.* **18**: 6155-6168

Pluta, K., Lefebvre, O., Martin, N.C., Smagowicz, W.J., Stanford, D.R., Ellis, S.R., Hopper A.S. and Boguta M.A. **(2001)** "Maf1p, a Negative Effector of RNA Polymerase III in *Saccharomyces cerevisiae*" Mol. Cellu. Biol. **21**: 5031-5040

Poulin, F., Gingras, A.C., Olsen, H., Chevalier, S., and Sonenberg, N. **(1998)** "4E–BP3, a new member of the eukaryotic initiation factor 4E-binding protein family" *J. Biol. Chem.* **273**: 14002-14007

Prakash, S., Sung, P., and Prakash, L. **(1993)** "DNA repair genes and proteinsm of *Saccharomyces cerevisiae*" *Annu. Rev. Genet.* **27**: 33-70

Rairdan, G.J. and Delaney, T.P. **(2002)** "Role of salicylic acid and NIM1/NPR1 in race-specific resistance in *Arabidopsis*" *Genetics* **161**: 803-811

Ramaswamy, N.T., Li, L., Khalil, M. and Cannon, J.F. **(1998)** "Regulation of yeast glycogen metabolism and sporulation by Glc7p protein phosphatase" *Genetics* **149**: 57-72

Rathjen, J.P., Chang, J.H., Staskawicz, B.J., and Michelmore, R.W. **(1999)** "Constitutively active Pto induces a Prf-dependent hypersensitive response in the absence of AvrPto" *EMBO J.* **18**:3232-3240

Ren, T., Qu, F., and Morris, T.J. **(2000)** "*HRT* gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus" *Plant Cell* **12**: 1917–1926

Robbelen, G. and Sharp, E.L., **(1978)** "Mode of inheritance, interaction and application of genes conditioning resistance to yellow rust." *Adv. Plant Breeding* **9**: 88

Robzyk, K., Recht, J., and Osley, M.A. **(2000)** "Rad6-dependent ubiquitination of histone H2B in yeast" *Science* **287**: 501-504

Roelfs, A.P. **(1989)** "Epidemiology of the cereal rusts in North America." *Can J Plant Patho.l* **11**: 86-90

Roelfs, A.P., Huerta-Espino, J. and Marshall, D. **(1992)** "Barley stripe rust in Texas" *Plant Dis.* **76**: 538

Sanchez, N.S., Pearce, D.A., Cardillo, T.S., Uribe, S., and Sherman, F. **(2001)** "Requirements of Cyc2p and the porin, Por1p, for ionic stability and mitochondrial integrity in Saccharomyces cerevisiae" *Arc. Biochem. Biophys.* **392**: 326-332

Schulze-Lefert, P., and Vogel, J. **(2000)** "Closing the ranks to attack by powdery mildew" *Trends Plant Sci.* **5**: 343-348

Scofield, S.R., Tobias, C.M., Rathjen, J.P., Chang, J.H., Lavelle, D.T., *et al.* **(1996)** "Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato" *Science* **274**: 2063-2065

Serebriiskii, I., Khazak, V., and Golemis, E.A. **(1999)** "A Two-hybrid dual bait system to discriminate specificity of protein interactions" *J. Biol. Chem.* **274**: 17080-17087

Shimada, Y., Gulli, M.P., and Peter, M. **(2000)** "Nuclear sequestration of the exchange factor Cdc24 by Far1 regulates cell polarity during yeast mating" Nature *Cell Biol.* **2**: 117-124

Shirasu, K., Lahaye, T., Tan, M.W., Zhou, F., Azevedo, C., and Schulze-Lefert, P. **(1999)** "A novel class of eukaryotic zinc-binding proteins is required for disease resistance signaling in barley and development in *C. Elegans*" *Cell* **99**: 355-366

Sofia H.J., Chen, G., Hetzler, B.G., Reyes-Spindola, J.F., and Miller, N.E. **(2001)** "Radical SAM, a novel superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: Functional characterization using new analysis and information visualization methods" *Nucleic Acids Res.* **29**: 1097-1106

Sonenberg, N., and Gingras, A. **(1998)** "The mRNA 59 cap-binding protein eIF4E and control of cell growth" *Curr. Opin. Cell Biol.* **10**: 268-275

Song, W.Y., Wang, G-L., Chen L-L., Kim H-S., Pi L-Y., et al**. (1995)** "A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*" *Science* **270**: 1804-1806

Stubbs, R.W. **(1985)** "Stripe rust." in Roelfs A.P., Bushnell W.R. (eds) *The cereal rusts*. Academic Press, London, pp 61-101

Stubbs, R.W. **(1988)** "Pathogenicity analysis of yellow (stripe) rust of wheat and its significance in a global context." In Simmonds NW, Rajaram S (eds) *Breeding strategies for resistance to the rusts of wheat.* CIMMYT, Mexico, pp 23-38

Sung, P., Prakash, S., and Prakash, L. **(1990)** "Mutation of cysteine-88 in the *Saccharomyces cerevisiae* RAD6 protein abolishes its ubiquitin conjugating activity and its various biological functions" *Proc. Natl. Acad. Sci. USA* **87**: 2695-2699

Swiderski, M.R., and Innes, R.W. **(2001)** "The *Arabidopsis PBS1* resistance gene encodes a member of a novel protein kinase subfamily" *Plant J.* **26**: 101-112

Takahashi, A., Casais, C., Ichimura, K., and Shirasu K. **(2003)** "HSP90 interacts with RAR1 and SGT1 and is essential for RPS2-mediated disease resistance in Arabidopsis" *Proc. Natl. Acad. Sci. USA* **100**: 11777-11782

Tameling, W.I., Elzinga, S.D., Darmin, P.S., Vossen, J.H., Takken,F.L., Haring, M.A. and Cornelissen, B.J. **(2002)** "The tomato *R* gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity" *Plant Cell* **14**: 2929-2939

Tang, X., Frederick, R.D., Zhou, J., Halterman, D.A., Jia, Y., Martin, G.B. **(1996)** "Initiation of plant disease resistance by physical interaction of AvrPto and Pto kinase" *Science* **274**: 2060-2063

Tjian, R., and Maniatis, T. **(1994)** "Transcriptional activation: a complex puzzle with few easy pieces" *Cell* **77**: 5-8

Tor, M., Gordon, P., Cuzick, A., Eulgem, T., Sinapidou, E., *et al*. **(2002)** "*Arabidopsis* SGT1b is required for defense signaling conferred by several downy mildew resistance genes" *Plant Cell* **14**: 993-1003

Tornero, P., Merritt, P., Sadanandom, A., Shirasu, K., Innes, R.W., Dangl, J.L. **(2002) "**RAR1 and NDR1 contribute quantitatively to disease resistance in *Arabidopsis*, and their relative contributions are dependent on the *R* gene assayed" *Plant Cell* **14**: 1005-1015

Traut, T.W. **(1994)** "The functions and consensus motifs of nine types of peptide segments that form different types of nucleotidebinding sites" *Eur. J. Biochem*. **222**: 9-19

Tsukahara, K., Hata, K., Nakamoto, K., Sagane, K., Watanabe, N., Kuromitsu, J., Kai, J., Tsuchiya, M., Ohba, F., Jigami, Y., Yoshimatsu, K., and Nagasu, T. **(2003)** "Medicinal genetics approach towards identifying the molecular target of a novel inhibitor of fungal cell wall assembly" *Mol. Micro*. **48**: 1029-1042

Upadhya, R., Lee, J.H., and Willis, I.M. **(2002)** "Maf1 is an essential mediator of diverse signals that repress RNA Polymerase III transcription" *Molecular Cell* **10**: 1489-1494

Utz, P.J., Hottelet, M., Le, T.M., Kim, S.J., Geiger, M.E., van Venrooij, W.J., and Anderson, P. **(1998)** "The 72-kDa component of signal recognition particle is cleaved during apoptosis" *J. Biol. Chem*. **273**: 35362-35370

Van der Biezen, E.A., and Jones, J.D.G. **(1998)** "The NB-ARC domain: a novel signaling motif shared by plant resistance gene products and regulators of cell death in animals" *Curr.Biol.* **8**: R226-R227

Van der Biezen, E.A., Freddie, C.T., Kahn, K., Parker, J.E., and Jones, J.D. **(2002)** "*Arabidopsis RPP4* is a member of the *RPP5* multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signaling components" *Plant J.* **29**: 439-451

Van Loon, L.C., and Van Strien, E.A. **(1999)** "The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins" *Physiol. Mol. Plant Pathol.* **55**: 85-97

Vierstra, R.D. **(2003)** "The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins" *TRENDS in Plant Sci.* **8**: 135-142

Volkert, M. R., Elliott, N. A., and Housman, D. E. **(2000)** "Functional genomics reveals a family of eukaryotic oxidation protection genes" *Proc. Natl. Acad. Sci. USA* **97**:14530-14535

Walter, P., and Blobel, G. **(1980)** "Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum" *Proc. Natl. Acad. Sci. U. S. A.* **77**: 7112-7116

Walter, P., and Blobel, G. **(1982)** "Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum" *Nature* **299**: 691-698

Warren R.F., Merritt, P.M., Holub, E., and Innes, R.W. **(1999)** "Identification of three putative signal transduction genes involved in *R* gene-specified disease resistance in *Arabidopsis*" *Genetics.* **152**: 401-412

Watkins, J.F., Sung, P., Prakash, S., and Prakash, L. **(1993)** "The extremely conserved amino terminus of RAD6 ubiquitin-conjugating enzyme is essential for amino-end rule-dependent protein degradation" *Genes Dev.* **7**: 250-261

Weissman, A.M. **(2001)** "Themes and variations on ubiqutination" *Nat Rev Mol Cell Biol.* **2**: 169-178

Wek, R.C., Cannon, J.F., Dever, T.E., and Hinnebusch, A.G. **(1992)** "Truncated protein phosphatase GLC7 restores translational activation of *GCN4* expression in yeast mutants defective for the eIF-2a kinase GCN2" *Mol. Cell. Biol.* **12**: 5700-5710

Whiteway, M., Freedman, R., Van Arsdell., S, Szostak., J.W., and Thorner, J. **(1987)** "The yeast ARD1 gene product is required for repression of cryptic mating-type information at the HML locus" *Mol. Cell Biol.* **7**: 3713-3722

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. **(1994)** "The product of the tobacco mosaic virus resistance gene *N*: Similarity to toll and the interleukin-1 receptor" *Cell* **78**: 1101-1115

Williams-Hart, T., Wu, X., and Tatchell, K. **(2002)** "Protein Phosphatase type 1 regulates ion homeostasis in *Saccharomyces cerevisiae*" *Genetics* **160**: 1423-1437

Wu, X., Hart, H., Cheng, C., Roach, P.J., and Tatchell, K. **(2001)** "Characterization of Gac1p, a regulatory subunit of protein phosphatase type I involved in glycogen accumulation in *Saccharomyces cerevisiae*" *Mol. Genet. Genomics* **265**: 622-635

Xiao, W., Chow, B.L., Broomfield, S., and Hanna, M. **(2000)** "The *Saccharomyces cerevisiae* RAD6 group is composed of an error-prone and two error-free postreplication repair pathways" *Genetics* **155**:1633-1641

Xiao, Z., McGrew, J.T., Schroeder, A.J., Fitzgerald-Hayes, M. **(1993)** "CSE1 and CSE2, two new genes required for accurate mitotic chromosome segregation in *Saccharomyces cerevisiae*" Mol. Cell Biol. **13**: 4691-4702

Yahiaoui, N., Srichumpa, P., Dudler, R., and Keller, B. **(2004)** "Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene Pm3b form hexaploid wheat" *Plant J.* **37**: 528-538

Yamamoto, T., Mori, Y., Ishibashi, T., Uchiyama, Y., Sakaguchi, N., Furukawa, T., Hashimoto, J., Kimura, S., and Sakaguchi, K. **(2004)** "Characterization of Rad6 from a higher plant, rice (*Oryza sativa* L.) and its interaction with Sgt1, a subunit of the SCF ubituitin ligase complex " *Biochem. Biophys. Research Commu*. **314**: 434-439

Zervos, A.S., Gyuris, J., and Brent, R. **(1993)** "Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites" *Cell* **72**: 223-232

Zhang, J. and Madden, T.L. **(1997)** "PowerBLAST: a newnetwork BLAST application for interactive or automated sequence analysis and annotation" *Genome Res.* **7**: 649-656

Zheng, N., Schulman, B.A., Song, L.Z., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M. *et al*. **(2002)** "Structure of the Cul1-Rbx1-Skp1-F box (Skp2) SCF ubiquitin ligase complex" *Nature* **416**: 703-709

APPENDIX A

DNA SEQUENCING AND ALIGNMENTS OF RAR1 AND SGT1

Sequencing result of *Rar*1 gene fragment from Bezostaja

```
Tcacacagcatcagcattgtgccacccctttgtgcacggaggtatctccataaattcgtcaaactccttgacgtggatatcgcaaca
cttccaccctctattcctgtcatggaaaaccgcaggacctggatggtattcacatgcagcatcatggttatccttctccttgtaggt
tttaccacatcctttattcttacaaaccctaggctcatttatattaacaacttttttcctgggctgtggaactgagcattcttcgac
aggttcgtatttgtaccatttacagcaactggtttttgtgccttcggctgtgatccatggtcggagcaaaagaaaccctgacggcac
ctggagcaggcctcggtttccacaccctgcttagaagactggattggagctaactttggtggggttgcctttgagttaagagaaaca
gcttttgtgactggnttctctgttgtatgcttcca
```

Alignment of *Rar*1 gene fragment from Bezostaja with barley *Rar1*

gi|6581045|gb|AF192261.1|AF192261 Hordeum vulgare Rar1 (Rar1) mRNA, complete cds

```
Score =  682 bits (344)
Identities = 441/473 (93%), Gaps = 3/473 (0%)
Strand = Plus / Minus


Query: 1   tcacacagcatcagcattgtgccacccctttgtgcacggaggtatctccataaattcgtc 60
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct: 789 tcacacagcatcagcattgtgccacccctttgtgcatggaggtatctccataaattcgtc 730


Query: 61  aaactccttgacgtggatatcgcaacacttccaccctctattcctgtcatggaaaaccgc 120
           |||||||||||| |||| ||| |||||||||||||||||| |||||||||||||| ||
Sbjct: 729 aaactccttgacatggacatcacaacacttccaccctctatttctgtcatggaaaactgc 670


Query: 121 aggacctggatggtattcacatgcagcatcatggttatccttctccttgtaggttttacc 180
           ||||||||||||||| || ||||||||||||||||||||||||||||||| ||||| ||
Sbjct: 669 aggacctggatggtaatcgcatgcagcatcatggttatccttctccttgtacgttttgcc 610


Query: 181 acatcctttattcttacaaaccctaggctcatttatattaacaacttttttcctgggctg 240
           |||||||||||||||||||||||||| |||||||||| |||||| |||||| ||||||
Sbjct: 609 acatcctttattcttacaaaccctaggttcatttatatcaacaacctttttcttgggctc 550
```

```
Query: 241 tggaactgagcattcttcgacaggttccgtatttgtaccatttacagcaactggtttttg 300
           |||||||||| |||  | ||||| ||||||||||||||||||||||||||||||||||||
Sbjct: 549 tggaactgaggatttctggacagattccgtatttgtaccatttacagcaactggtttttg 490

Query: 301 tgccttcggctgtgatccatggtcggagcaaaagaaaccctgacggcacctggagcaggc 360
           |||||| |||||||||||||||| || ||||||||||||||||||||| ||||||||
Sbjct: 489 tgccttgggctgtgatccatggtcagaacaaaagaaaccctgacggcaccttgagcaggc 430

Query: 361 ctcggtttccacaccctgcttagaagactggattggagctaactttggtggggttgcctt 420
           ||| |||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Sbjct: 429 ctcagtttccacaccctgcttagaagactggactggagctgactttggtggggttgcctt 370

Query: 421 t---gagttaagagaaacagcttttgtgactggnttctctgttgtatgcttcc 470
           |   |||||||||||||||||||||||||||| ||||| ||||||||||||
Sbjct: 369 tgaggagttaagagaaacagcttttgtgactggtttctcagttgtatgcttcc 317
```

Sequencing result of *Rar*1 gene fragment from Gerek79

```
tcacacagcatcagcattgtgccagccctttgtgcatggaggtatctccataaattcgtcaaactccttgacgtggacatcgcaaca
cttccaccctctattcctgtcatggaaaaccgcaggacctggatggtattcacatgcagcatcatggttatccttctctttgtaggt
tttaccacatcctttattcttacaaaccctaggctcatttatattaacaacttttttcttgggctgtggaactgagcatttttcgac
aggttccgtatttgtaccatttacagcaactggtttttgtgccttgggctgtgatccatggtcggagcaaaagaaaccctgacggca
cctggagcaggcctcggtttccacaccctgcttagaagactggattggagctaactttggtggggttgcctttgagttaagagaaac
agcttttgtggctggtttctctgttgtatgcttcca
```

Alignment of *Rar*1 gene fragment from Gerek79 with barley *Rar1*

gi|6581045|gb|AF192261.1|AF192261 Hordeum vulgare Rar1 (Rar1) mRNA, complete cds

```
Score =  704 bits (355)
Identities = 444/473 (93%), Gaps = 3/473 (0%)


Query: 1   tcacacagcatcagcattgtgccagccctttgtgcatggaggtatctccataaattcgtc 60
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct: 789 tcacacagcatcagcattgtgccacccctttgtgcatggaggtatctccataaattcgtc 730

Query: 61  aaactccttgacgtggacatcgcaacacttccaccctctattcctgtcatggaaaaccgc 120
           |||||||||||| |||||||| |||||||||||||||||| |||||||||||||| ||
Sbjct: 729 aaactccttgacatggacatcacaacacttccaccctctatttctgtcatggaaaactgc 670

Query: 121 aggacctggatggtattcacatgcagcatcatggttatccttctctttgtaggttttacc 180
           ||||||||||||||| || |||||||||||||||||||||||| ||||| ||||| ||
Sbjct: 669 aggacctggatggtaatcgcatgcagcatcatggttatccttctccttgtacgttttgcc 610

Query: 181 acatcctttattcttacaaaccctaggctcatttatattaacaacttttttcttgggctg 240
           |||||||||||||||||||||||||| |||||||||| |||||| ||||||||||||
Sbjct: 609 acatcctttattcttacaaaccctaggttcatttatatcaacaacctttttcttgggctc 550

Query: 241 tggaactgagcatttttcgacaggttccgtatttgtaccatttacagcaactggtttttg 300
           |||||||||| |||| | ||||| ||||||||||||||||||||||||||||||||||||
Sbjct: 549 tggaactgaggatttctggacagattccgtatttgtaccatttacagcaactggtttttg 490
```

```
Query: 301 tgccttgggctgtgatccatggtcggagcaaaagaaaccctgacggcacctggagcaggc 360
           ||||||||||||||||||||||| || ||||||||||||||||||||| ||||||||
Sbjct: 489 tgccttgggctgtgatccatggtcagaacaaaagaaaccctgacggcaccttgagcaggc 430

Query: 361 ctcggtttccacaccctgcttagaagactggattggagctaactttggtggggttgcctt 420
           ||| |||||||||||||||||||||||||| ||||||| ||||||||||||||||||
Sbjct: 429 ctcagtttccacaccctgcttagaagactggactggagctgactttggtggggttgcctt 370

Query: 421 t---gagttaagagaaacagcttttgtggctggtttctctgttgtatgcttcc 470
           |   |||||||||||||||||||| |||||||||| ||||||||||||
Sbjct: 369 tgaggagttaagagaaacagcttttgtgactggtttctcagttgtatgcttcc 317
```

Sequencing result of *Rar1* gene fragment from Pastor

```
tcacacagcatcagcattgtgccacccctttgtgcatggaggtatctccataaattcgtcaaactccttgacatggacatcgcaaca
cttccaccctctattcctgtcatggaaaaccgcagggcctggatggtattcgcatgcagcatcatggttatccttctccttgtaggt
tttaccacatcctttattcttacaaaccctaggctcgtttatattaacaacctttttcttgggctgtggaaccgagcatttctcgac
aggttccgtatttgtaccatttacagcaactggtttttgtgccttgggctgtgatccatggtcggagcaaaagaaaccctgacggca
cctggagcaggcctcggtttccacaccctgcttagaagactggactggagctaactttggtggggttgcctttgagttaagagaaac
agcttttgtgatctggnttctctgttgtgtgcttcca
```

Alignment of *Rar1* gene fragment from Gerek79 with barley *Rar1*

gi|6581045|gb|AF192261.1|AF192261 Hordeum vulgare Rar1 (Rar1) mRNA, complete cds

```
Score =  716 bits (361)
Identities = 447/474 (94%), Gaps = 4/474 (0%)
Strand = Plus / Minus


Query: 1   tcacacagcatcagcattgtgccacccctttgtgcatggaggtatctccataaattcgtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 789 tcacacagcatcagcattgtgccacccctttgtgcatggaggtatctccataaattcgtc 730

Query: 61  aaactccttgacatggacatcgcaacacttccaccctctattcctgtcatggaaaaccgc 120
           |||||||||||||||||||| |||||||||||||||||| |||||||||||||| ||
Sbjct: 729 aaactccttgacatggacatcacaacacttccaccctctatttctgtcatggaaaactgc 670

Query: 121 agggcctggatggtattcgcatgcagcatcatggttatccttctccttgtaggttttacc 180
           ||| |||||||||||| ||||||||||||||||||||||||||||||||||| ||||| ||
Sbjct: 669 aggacctggatggtaatcgcatgcagcatcatggttatccttctccttgtacgttttgcc 610

Query: 181 acatcctttattcttacaaaccctaggctcgtttatattaacaacctttttcttgggctg 240
           |||||||||||||||||||||||||| || ||||||| ||||||||||||||||||||
Sbjct: 609 acatcctttattcttacaaaccctaggttcatttatatcaacaacctttttcttgggctc 550

Query: 241 tggaaccgagcatttctcgacaggttccgtatttgtaccatttacagcaactggtttttg 300
           |||||| ||| |||||| ||||| ||||||||||||||||||||||||||||||||||
Sbjct: 549 tggaactgaggatttctggacagattccgtatttgtaccatttacagcaactggtttttg 490
```

```
Query: 301 tgccttgggctgtgatccatggtcggagcaaaagaaaccctgacggcacctggagcaggc 360
           ||||||||||||||||||||||| || |||||||||||||||||||||| ||||||||
Sbjct: 489 tgccttgggctgtgatccatggtcagaacaaaagaaaccctgacggcaccttgagcaggc 430

Query: 361 ctcggtttccacaccctgcttagaagactggactggagctaactttggtggggttgcctt 420
           ||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Sbjct: 429 ctcagtttccacaccctgcttagaagactggactggagctgactttggtggggttgcctt 370

Query: 421 t---gagttaagagaaacagcttttgtgatctggnttctctgttgtgtgcttcc 471
           |   |||||||||||||||||||||| |||| ||||| ||||| |||||||
Sbjct: 369 tgaggagttaagagaaacagcttttgtga-ctggtttctcagttgtatgcttcc 317
```

Sequencing result of *SGT1* gene fragment from Avocet

```
acactgaggctgtagctgatgccaacaaagcaattgaacttgatccttcgatgcataaagcataccttcggaagggctctgcttgca
tcaagctggaggaataccaaactgcaaaggctgctcttgaagtgggttcttcttatgcatctggtgactcgaggtttactcgtctga
tgaaggagtgtgatgatcgtattgctgaggaggctagccaggtgccagtaaagaatgccgctgcggctgttgcttcagctacatctt
cgggggcatcttccggggctacaactgtggctactgaagctgaggaccaggatggtgcaaatatggagaatgcacagccaacgatgg
aagtgccaagcaagcccaaatacaggcatgactactacaatactcctacagaagtggtactgactatatttgctaagggtgttccag
ctgacagcgtggttgttgactttggtgaacagattntgangngtctcaattgaacttcccggtgaggaaccataccntttaangggg
nctcgtctgttttcaaaga
```

Alignment of *SGT1* from gene fragment Gerek79 with barley *SGT1*

gi|17017305|gb|AF439974.1|AF439974 Hordeum vulgare SGT1 (SGT1) mRNA, complete cds

```
Score =  436 bits (220)
Identities = 256/268 (95%)
Strand = Plus / Plus

Query: 1   acactgaggctgtagctgatgccaacaaagcaattgaacttgatccttcgatgcataaag 60
           ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||
Sbjct: 175 acactgaggctgtagctgatgccaacaaagcaattgaacttgatccctcgatgcacaaag 234

Query: 61  cataccttcggaagggctctgcttgcatcaagctggaggaataccaaactgcaaaggctg 120
           | |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct: 235 cgtaccttcggaagggttctgcttgcatcaagctggaggaataccaaactgcaaaggctg 294

Query: 121 ctcttgaagtgggttcttcttatgcatctggtgactcgaggtttactcgtctgatgaagg 180
           |||||||||| |||||||| |||||||||||||||| |||||||||||||| |||||||
Sbjct: 295 ctcttgaagttggttcttcctatgcatctggtgactcaaggtttactcgtcttatgaagg 354

Query: 181 agtgtgatgatcgtattgctgaggaggctagccaggtgccagtaaagaatgccgctgcgg 240
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct: 355 agtgtgatgatcgtattgctgaggaggctagccaggcgccagtaaagaatgccgctgcgg 414

Query: 241 ctgttgcttcagctacatcttcgggggc 268
           |||||||| | |||||||||| |||||
Sbjct: 415 ctgttgctcctgctacatcttccggggc 442
```

APPENDIX B

DNA SEQUENCING AND ALIGNMENTS OF THE BAITS

Sequencing result of YR10-1506 gene fragment

atggaggtcgtgaccggggcgatgagcacgctcctgcccttgcttggcgacctgcttaaggaggagtacaacctgcagaagagcacca
agggtgagatcaagttcctcaaagcagagctggagagcatggaggctgccctcatcaagatctcggaggcacccttggatcagccacc
taacattcaggtcaagctctgggcgagggacgtcaaggacctgtcctatgagatcgaagatggcatcgacaaattccgggtgcacctt
gagtgccggcaacaaaagaagccacacagctttatgggtttcatccacaaaagcatggacatgctgacaaagggcaagatccgacaca
aaataggcatcgatatcaaagacatcaagagccgcatcaaggaggtcagtgacaggcgtgaaaggtacaaggttgatagcgttgcgcc
caagcccaccggcacaagtactgatacacttcgccagttagccttgttcaaaaaggcggaagagcttattggcaccaaagaaaagagc
cttgacatagtcaagatgctgacggaaggagatgaggtcttcaagaaacatcttaagatggtctctattgttggctttggaggcttag
ggaagacaactcttgctaacgtggtatatgagaagcttcgcggggactttgattgtgcagcttttgtctctgtgtctcttaatcctga
catgaagaagcttttcaagtgtttgctccatcagcttgacaaAggcgagtacaagaacatcatggacgagtcagcgtggagcgaaaca
caactcataagtgagataagagatttccttcgagacaagaggtacttcattctcattgatgacatatgggataaatctgtgtggaata
atattagatgtgctctgattgagaatgaatgtggtagtagagtaatcgcaacaactcgcattctagatgttgccaaagaagttggtgg
tgtttatcagcttaaacctctttctactagtgactcaAgacaattattctatcaaagaatatttggaattggcgacaagcgtcctcct
attcagttggctgaagtaagtgagaaaattttagggaaatgtggcggagtaccattagctatcattacactggctagtatgttggctg
gtaaaaaggaacatgaaaatacatatacttattggtacaaggtgtaccaatctatgggttctgggctagaaaataatcctggcctgat
ggacatgaggaggatactacatgtcagttactatgacctacctccaaatctgaagacttgtttactgtatctcagtttgtatccagag
gattataatattgaaaccaaagagttgatatggaaatggataggcgaaggattcattcatgaagagcaagggaagagcttgtatgaag
taggagaggattacattgctgagctcattaGcaaaagcttggtccaaccaatgtatatcaatattgctaataaggcgagctctgtccg
tgtacacgac

Figure B.1 YR10-1506 structure and conserved domains

```
gnl|CDD|24492 pfam00931, NB-ARC, NB-ARC domain.

CD-Length = 305 residues,  99.3% aligned
Score =  192 bits (488), Expect = 1e-49

Query:  155  DTLRQLALFKKAEELIGTKEKSLDIVKMLTEGDEVFKKHLKMVSIVGFGGLGKTTLANVV  214
Sbjct:  3    NTVRSSGLVPDESTVVGREDMVEAVIEKLLEMSE----NLGVVGIVGMGGVGKTTLAKQI  58

Query:  215  YEK--LRGDFDCAAFVSVSLNPDMKKLFKCLLHQLDK-GEYKNIMDESAWSETQLISEIR  271
Sbjct:  59   YNDFSVGGHFDSVAWVVVSKTYTEFDLQKTILQELGSEDESWDHKNEG-----ELAVKLK  113

Query:  272  DFLRDKRYFILIDDIWDKSVWNNIRCALIENECGSRVIATTRILDVAKEVGG---VYQLK  328
Sbjct:  114  ELLLRKRFLLVLDDVWEKEDWDDIGVPFPDGENGSRVIVTTRSESVAGRMGGTSKPHEVE  173

Query:  329  PLSTSDSRQLFYQRIFGIGDKRPPIQLAEVSEKILGKCGGVPLAIITLASMLAGKKEHEN  388
Sbjct:  174  SLEPEESWELFSNKVFPKNLPSEHPELEEVAKEIVEKCKGLPLALKVLGGLLASKSTVPE  233


Query:  389  TYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDLPPNLKTCLLYLSLYPEDYNIETKE  448
Sbjct:  234  ----WKHVLEQLNNELAGRDNLNEVLSILSLSYDNLPMHLKRCFLYLALFPEDYDIPVKQ  289

Query:  449  LIWKWIGEGFIHEEQG  464
Sbjct:  290  LISLWIAEGFVEAENE  305
```

Amino acid alignment of YR10-1506 with YR10 protein

```
gi|11990500|gb|AAG42168.1|   stripe rust resistance protein Yr10 [Triticum aestivum]

Identities = 516/520 (99%), Positives = 517/520 (99%)

Query: 1   MEVVTGAMSTLLPLLGDLLKEEYNLQKSTKGEIKFLKAELESMEAALIKISEAPLDQPPN 60
           MEVVTGAMSTLLPLLGDLLKEEYNLQKSTKGEIKFLKAELESMEAALIKISEAPLDQPPN
Sbjct: 1   MEVVTGAMSTLLPLLGDLLKEEYNLQKSTKGEIKFLKAELESMEAALIKISEAPLDQPPN 60

Query: 61  IQVKLWARDVKDLSYEIEDGIDKFRVHLECRQQKKPHSFMGFIHKSMDMLTKGKIRHKIG 120
           IQVKLWARDVKDLSYEIEDGIDKFRVHLECRQQKKPHSFMGFIHKSMDMLTKGKIRHKIG
Sbjct: 61  IQVKLWARDVKDLSYEIEDGIDKFRVHLECRQQKKPHSFMGFIHKSMDMLTKGKIRHKIG 120

Query: 121 IDIKDIKSRIKEVSDRRERYKVDSVAPKPTGTSTDTLRQLALFKKAEELIGTKEKSLDIV 180
           IDIKDIKSRIKEVSDRRERYKVDSVAPKPTGTSTDTLRQLALFKKAEELIGTKEKSLDIV
Sbjct: 121 IDIKDIKSRIKEVSDRRERYKVDSVAPKPTGTSTDTLRQLALFKKAEELIGTKEKSLDIV 180

Query: 181 KMLTEGDEVFKKHLKMVSIVGFGGLGKTTLANVVYEKLRGDFDCAAFVSVSLNPDMKKLF 240
           KMLTEGDEVFKKHLKMVSIVGFGGLGKTTLANVVYEKLRGDFDCAAFVSVSLNPDMKKLF
Sbjct: 181 KMLTEGDEVFKKHLKMVSIVGFGGLGKTTLANVVYEKLRGDFDCAAFVSVSLNPDMKKLF 240

Query: 241 KCLLHQLDKGEYKNIMDESAWSETQLISEIRDFLRDKRYFILIDDIWDKSVWNNIRCALI 300
           KCLLHQLDKGEYKNIMDESAWSETQLISEIRDFLRDKRYFILIDDIWDKSVWNNIRCALI
Sbjct: 241 KCLLHQLDKGEYKNIMDESAWSETQLISEIRDFLRDKRYFILIDDIWDKSVWNNIRCALI 300

Query: 301 ENECGSRVIATTRILDVAKEVGGVYQLKPLSTSDSRQLFYQRIFGIGDKRPPIQLAEVSE 360
           ENECGSRVIATTRILDVAKEVGGVYQLKPLSTSDSRQLFYQRIFGIGDKRPPIQLAEVSE
Sbjct: 301 ENECGSRVIATTRILDVAKEVGGVYQLKPLSTSDSRQLFYQRIFGIGDKRPPIQLAEVSE 360

Query: 361 KILGKCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVS 420
           KILGKCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVS
Sbjct: 361 KILGKCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVS 420
```

```
Query: 421 YYDLPPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELIS 480
           YYDLPPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELI+
Sbjct: 421 YYDLPPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELIN 480

Query: 481 KSLVQPMYINIANKASSVRVHD 502
           KSLVQPMYINIANKASSVRVHD
Sbjct: 481 KSLVQPMYINIANKASSVRVHD 502
```

Sequencing result of YR10-1562 gene fragment

```
ggtagtagagtaatcgcaacaactcgcattctagatgttgccaaagaagttggtggtgtttatcagcttaaacctctttctactagt
gactcaggacaattattctatcaaagaatatttggaattggcgacaagcgtcctcctattcagttggctgaagtaagtgagaaaatt
ttagggaaatgtggcggagtaccattagctatcattacactggctagtatgttggctggtaaaaaggaacatgaaaatacatatact
tattggtacaaggtgtaccaatctatgggttctgggctagaaaataatcctggcctgatggacatgaggaggatactacatgtcagt
tactatgacctacctccaaatctgaagacttgtttactgtatctcagtttgtatccagaggattataatattgaaaccaaagagttg
atatggaaatggataggcgaaggattcattcatgaagagcaagggaagagcttgtatgaagtaggagaggattacattgctgagctc
attaacaaaagcttggtccaaccaatgtatatcaatattgctaataaggcgagctctgtccgtgtacacgacatggtgcttgacctt
atcCcttccttgtcaaatgaggagaactttctcgcaacattgggtggtcagcagaccaggtcactgccaAgtaagatccgtcgactg
tctctccaaagcagtGatgaagaggatgtccagccaatgccaaccatgagcagcttgtcccatgtgaggtcacttactgtgttcagt
aaagacctcagtttgctgtcggcactttcaggttttcttgtcctgcgtgcattggatttaagtggttgtgaggaagtgggtaatcat
cacatgaaggacatttgcaatttatttcacttgaggtatctgagtctagaagggacatctatcCctgagatcccgaaagagataagc
aatctacggcttctgcaattgctagtcataaggtccacaaaaatgaaaaaatttccatcaacctttgttcagctaggacaactggtg
ttcattgacatgggtaatagggaagtctctagattgcttctaaaatcaatgtccaccttgccctccctctcttcacttgcaatcgga
ataggagaactgagagaggaagacctccaaatacttgggagcatgccgtctctgcatgacctctctattgatgtaggttattgggaa
agaggcagagataaaaggctagtcattgacagtggttctcccttccggtctctgacaaggttcagtataaagggttgcggcttcatt
gatttcatgtttgcacaaggaaccttgcaaaagctccagatcctggagttaagtatttttggtaaggcaataaaagacagatttggt
gattttcaatttgggctggagaatctctcatcacttgagcatgtctatgtggacgctcgtggtcgtggtataatcccaagtcaagaa
gctgagctgagcggtgcacttgagaaagagcttgatattaatcccaacaagcccacactgacggtgaaggtaactccacgctga
```

Figure B.2 YR10-1563 structure and conserved domains

```
gnl|CDD|24492 pfam00931, NB-ARC, NB-ARC domain.

CD-Length = 305 residues, only  52.1% aligned
Score =  122 bits (307), Expect = 1e-28

Query:  1    GSRVIATTRILDVAKEVGG---VYQLKPLSTSDSGQLFYQRIFGIGDKRPPIQLAEVSEK  57
Sbjct:  147  GSRVIVTTRSESVAGRMGGTSKPHEVESLEPEESWELFSNKVFPKNLPSEHPELEEVAKE  206
```

```
Query:  58   ILGKCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSY  117
Sbjct:  207  IVEKCKGLPLALKVLGGLLASKSTVPE----WKHVLEQLNNELAGRDNLNEVLSILSLSY  262


Query:  118  YDLPPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQG  160
Sbjct:  263  DNLPMHLKRCFLYLALFPEDYDIPVKQLISLWIAEGFVEAENE  305


gnl|CDD|14020 COG4886, COG4886, Leucine-rich repeat (LRR) protein
[Function unknown].

CD-Length = 394 residues, only  67.3% aligned
Score = 44.6 bits (104), Expect = 3e-05

Query:  177  KSLVQPMYINIANKASSVRVHDMVLDLIPSLSNEENFLATLGGQQTRSLPSKIRRLSLQS  236
Sbjct:  44   SVAVNRLALNLSSNTLLLLPSSLSRLLSLDLLSP-SGISSLDGSENLLNLLPLPSLDLNL  102

Query:  237  SDEEDVQPMPTMSSLSHVRSLTVFSKDLSLLLSALSGFLV--LRALDLSGCEEVGNHHMKD  294
Sbjct:  103  N--RLRSNISELLELTNLTSLDLDNNNITDIPPLIGLLKSNLKELDLSDNKI--ESLPSP  158

Query:  295  ICNLFHLRYLSLEGTSIPEIPKEISNLRLLQLLVIRSTKMKKFPSTFVQLGQLVFIDMGN  354
Sbjct:  159  LRNLPNLKNLDLSFNDLSDLPKLLSNLSNLNNLDLSGNKISDLPPEIELLSALEELDLSN  218

Query:  355  REVSRLLLKSMSTLPSLSSLAIGIGELREEDLQILGSMPSLHDLSIDVGYWERGRDKRLV  414
Sbjct:  219  NSIIELLS-SLSNLKNLSGLELSNNKL-EDLPESIGNLSNLETLDLS-------NNQISS  269

Query:  415  IDSGSPFRSLTRFSIKGCGFIDFMFAQGTLQKLQILELS  453
Sbjct: 270  ISSLGSLTNLRELDLSGNSLSNALPLIALLLLLLELLLN 308
```

Amino acid alignment of YR10-1563 with YR10 protein

```
gi|11990500|gb|AAG42168.1|   stripe rust resistance protein Yr10 [Triticum aestivum]

Identities = 516/520 (99%), Positives = 517/520 (99%)

Query: 1   GSRVIATTRILDVAKEVGGVYQLKPLSTSDSGQLFYQRIFGIGDKRPPIQLAEVSEKILG 60
           GSRVIATTRILDVAKEVGGVYQLKPLSTSDSGQLFYQRIFGIGDKRPPIQLAEVSEKILG
Sbjct: 305 GSRVIATTRILDVAKEVGGVYQLKPLSTSDSGQLFYQRIFGIGDKRPPIQLAEVSEKILG 364

Query: 61  KCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDL 120
           KCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDL
Sbjct: 365 KCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDL 424

Query: 121 PPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELINKSLV 180
           PPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELINKSLV
Sbjct: 425 PPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELINKSLV 484

Query: 181 QPMYINIANKASSVRVHDMVLDLIPSLSNEENFLATLGGQQTRSLPSKIRRLSLQSSDEE 240
           QPMYINIANKASSVRVHDMVLDLI SLSNEENFLATLGGQQTRSLP KIRRLSLQSS+EE
Sbjct: 485 QPMYINIANKASSVRVHDMVLDLITSLSNEENFLATLGGQQTRSLPRKIRRLSLQSSNEE 544

Query: 241 DVQPMPTMSSLSHVRSLTVFSKDLSLLSALSGFLVLRALDLSGCEEVGNHHMKDICNLFH 300
           DVQPMPTMSSLSHVRSLTVFSKDLSLLSALSGFLVLRALDLSGCEEVGNHHMKDICNLFH
Sbjct: 545 DVQPMPTMSSLSHVRSLTVFSKDLSLLSALSGFLVLRALDLSGCEEVGNHHMKDICNLFH 604

Query: 301 LRYLSLEGTSIPEIPKEISNLRLLQLLVIRSTKMKKFPSTFVQLGQLVFIDMGNREVSRL 360
           LRYLSLEGTSI EIPKEISNLRLLQLLVIRSTKMKKFPSTFVQLGQLVFIDMGNREVSRL
Sbjct: 605 LRYLSLEGTSITEIPKEISNLRLLQLLVIRSTKMKKFPSTFVQLGQLVFIDMGNREVSRL 664
```

```
Query: 361 LLKSMSTLPSLSSLAIGIGELREEDLQILGSMPSLHDLSIDVGYWERGRDKRLVIDSGSP 420
           LLKSMSTLPSLSSLAIGIGELREEDLQILGSMPSLHDLSIDVGYWERGRDKRLVIDSGSP
Sbjct: 665 LLKSMSTLPSLSSLAIGIGELREEDLQILGSMPSLHDLSIDVGYWERGRDKRLVIDSGSP 724

Query: 421 FRSLTRFSIKGCGFIDFMFAQGTLQKLQILELSIFGKAIKDRFGDFQFGLENLSSLEHVY 480
           FRSLTRFSIKGCGFIDFMFAQGTLQKLQILELSIFGKAIKDRFGDFQFGLENLSSLEHVY
Sbjct: 725 FRSLTRFSIKGCGFIDFMFAQGTLQKLQILELSIFGKAIKDRFGDFQFGLENLSSLEHVY 784


Query: 481 VDARGRGIIPSQEAELSGALEKELDINPNKPTLTVKVTPR 520
           VDARGRGIIPSQEAELSGALEKELDINPNKPTLTVKVTPR
Sbjct: 785 VDARGRGIIPSQEAELSGALEKELDINPNKPTLTVKVTPR 824
```

Sequencing result of YR10-1

```
Gtgaccggggcgatgagcacgctcctgcccttgcttggcgacctgcttaaggaggagtacaacctgcagaagagcaccaagggtgaga
tcaagttcctcaaagcagagctggagagcatggaggctgccctcatcaagatctcggaggcacccttggatcagccacctaacattca
ggtcaagctctgggcgagggacgtcaaggacctgtcctatgagatcgaagatggcatcgacaaattccgggtgcaccttgagtgccgg
caacaaaagaagccacacagctttatgggtttcatccacaaaagcatggacatgctgacaaagggcaagatccgacacaaaataggca
tcgatatcaaagacatcaagagccgcatcaaggaggtcagtgacaggcgtgaaaggtacaaggttgatagcgttgcgcccaagcccac
cggcacaagtactgatacacttcgccagttagccttgttcaaaaaggcggaagagcttattggcaccaaagaaaagagccttgacata
gtcaagatgctgacggaaggagatgaggtcttcaagaaacatcttaagatggtctctattgttggctttggaggcttagggaagacaa
ctctt
```

Amino acid alignment of YR10-1 with YR10 protein

```
gi|11990500|gb|AAG42168.1|   stripe rust resistance protein Yr10 [Triticum aestivum]

Identities = 207/207 (100%), Positives = 207/207 (100%)

Query: 1   VTGAMSTLLPLLGDLLKEEYNLQKSTKGEIKFLKAELESMEAALIKISEAPLDQPPNIQV 60
           VTGAMSTLLPLLGDLLKEEYNLQKSTKGEIKFLKAELESMEAALIKISEAPLDQPPNIQV
Sbjct: 4   VTGAMSTLLPLLGDLLKEEYNLQKSTKGEIKFLKAELESMEAALIKISEAPLDQPPNIQV 63

Query: 61  KLWARDVKDLSYEIEDGIDKFRVHLECRQQKKPHSFMGFIHKSMDMLTKGKIRHKIGIDI 120
           KLWARDVKDLSYEIEDGIDKFRVHLECRQQKKPHSFMGFIHKSMDMLTKGKIRHKIGIDI
Sbjct: 64  KLWARDVKDLSYEIEDGIDKFRVHLECRQQKKPHSFMGFIHKSMDMLTKGKIRHKIGIDI 123

Query: 121 KDIKSRIKEVSDRRERYKVDSVAPKPTGTSTDTLRQLALFKKAEELIGTKEKSLDIVKML 180
           KDIKSRIKEVSDRRERYKVDSVAPKPTGTSTDTLRQLALFKKAEELIGTKEKSLDIVKML
Sbjct: 124 KDIKSRIKEVSDRRERYKVDSVAPKPTGTSTDTLRQLALFKKAEELIGTKEKSLDIVKML 183

Query: 181 TEGDEVFKKHLKMVSIVGFGGLGKTTL 207
           TEGDEVFKKHLKMVSIVGFGGLGKTTL
Sbjct: 184 TEGDEVFKKHLKMVSIVGFGGLGKTTL 210
```

Sequencing result of YR10-2 gene fragment

```
Ggtagtagagtaatcgcaacaactcgcattctagatgttgccaaagaagttggtggtgtttatcagcttaaacctctttctactagtg
actcaagacaattattctatcaaagaatatttggaattggcgacaagcgtcctcctattcagttggctgaagtaagtgagaaaatttt
agggaaatgtggcggagtaccattagctatcattacactggctagtatgttggctggtaaaaaggaacatgaaaatacatatacttat
tggtacaaggtgtaccaatctatgggttctgggctagaaaataatcctggcctgatggacatgaggaggatactacatgtcagttact
atgacctacctccaaatctgaagacttgtttactgtatctcagtttgtatccagaggattataatattgaaaccaaagagttgatatg
gaaatggataggcgaaggattcattcatgaagagcaagggaagagcttgtatgaagtaggaggggattacattgctgagctcattaac
aaaagcttggtccaaccaatgtatatcaatattgctaataaggcgagctctgtccgtgtacacgac
```

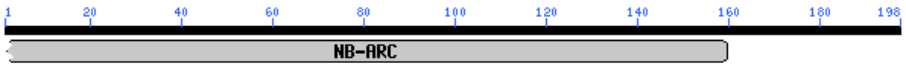


Figure B.3 YR10-2 structure and conserved domains

```
gnl|CDD|24492 pfam00931, NB-ARC, NB-ARC domain.

CD-Length = 305 residues, only  52.1% aligned
Score =  124 bits (312), Expect = 9e-30

Query:  1    GSRVIATTRILDVAKEVGG---VYQLKPLSTSDSRQLFYQRIFGIGDKRPPIQLAEVSEK  57
Sbjct:  147  GSRVIVTTRSESVAGRMGGTSKPHEVESLEPEESWELFSNKVFPKNLPSEHPELEEVAKE  206

Query:  58   ILGKCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSY  117
Sbjct:  207  IVEKCKGLPLALKVLGGLLASKSTVPE----WKHVLEQLNNELAGRDNLNEVLSILSLSY  262

Query:  118  YDLPPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQG  160
Sbjct:  263  DNLPMHLKRCFLYLALFPEDYDIPVKQLISLWIAEGFVEAENE  305

gi|11990500|gb|AAG42168.1|   stripe rust resistance protein Yr10 [Triticum aestivum]


Score =  411 bits (1056), Expect = e-114
Identities = 196/198 (98%), Positives = 196/198 (98%)

Query: 1   GSRVIATTRILDVAKEVGGVYQLKPLSTSDSRQLFYQRIFGIGDKRPPIQLAEVSEKILG 60
           GSRVIATTRILDVAKEVGGVYQLKPLSTSDS QLFYQRIFGIGDKRPPIQLAEVSEKILG
Sbjct: 305 GSRVIATTRILDVAKEVGGVYQLKPLSTSDSGQLFYQRIFGIGDKRPPIQLAEVSEKILG 364

Query: 61  KCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDL 120
           KCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDL
Sbjct: 365 KCGGVPLAIITLASMLAGKKEHENTYTYWYKVYQSMGSGLENNPGLMDMRRILHVSYYDL 424

Query: 121 PPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGGDYIAELINKSLV 180
           PPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVG DYIAELINKSLV
Sbjct: 425 PPNLKTCLLYLSLYPEDYNIETKELIWKWIGEGFIHEEQGKSLYEVGEDYIAELINKSLV 484
```

```
Query: 181 QPMYINIANKASSVRVHD 198
           QPMYINIANKASSVRVHD
Sbjct: 485 QPMYINIANKASSVRVHD 502
```

Sequencing result of *RAD6* gene

```
atgtcgactccttcaaggaagaggctgatgagggacttcaagcggctgatgcaggaccctcctgcgggcataagcggggcgccgcag
gacaacaacataatgctgtggaatgctgtgatttttggccctgacgatagcccgtgggatggaggcacgtttaagctgactctccag
tttaatgaagaatatcctaataagccaccaacagttcggtttatttctcggatgtttcaccctaacatttatgctgatggaagcata
tgcttagatattctacagaatcagtggagcccaatatatgatgtagctgctatacttacatctatccagtcgctgctgtgtgatcct
aacccaaattcgcctgctaactcagaagctgcccgcatgttcagtgagaacaagcgagagtacaaccgcaaagtgcgggagattgtt
gagcagagctggacggcagactaa
```

Amino acid alignment of RAD6 with *Nicotiana tabacum* and *Capsicum chinense* ubiquitin-conjugating enzymes

```
gi|58578278|emb|CAI48075.1|  ubiquitin-conjugating enzyme [Capsicum chinense]
gi|13516451|dbj|BAB40310.1| ubiquitin-conjugating enzyme (E2) [Nicotiana tabacum]
Identities = 146/152 (96%), Positives = 150/152 (98%)

Query: 1   MSTPSRKRLMRDFKRLMQDPPAGISGAPQDNNIMLWNAVIFGPDDSPWDGGTFKLTLQFN 60
           MSTP+RKRLMRDFKRL QDPPAGISGAP DNNIMLWNAVIFGPDD+PWDGGTFKLTLQF+
Sbjct: 1   MSTPARKRLMRDFKRLQQDPPAGISGAPYDNNIMLWNAVIFGPDDTPWDGGTFKLTLQFS 60

Query: 61  EEYPNKPPTVRFISRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS 120
           E+YPNKPPTVRFISRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS
Sbjct: 61  EDYPNKPPTVRFISRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS 120

Query: 121 PANSEAARMFSENKREYNRKVREIVEQSWTAD 152
           PANSEAARMFSENKREYNRKVREIVEQSWTAD
Sbjct: 121 PANSEAARMFSENKREYNRKVREIVEQSWTAD 152
```

Amino acid alignment of RAD6 with *Avicennia marina* ubiquitin-conjugating protein

```
gi|8118527|gb|AAF73016.1|   ubiquitin conjugating protein [Avicennia marina]

Identities = 144/152 (94%), Positives = 150/152 (98%)

Query: 1   MSTPSRKRLMRDFKRLMQDPPAGISGAPQDNNIMLWNAVIFGPDDSPWDGGTFKLTLQFN 60
           MSTP+RKRLMRDFKRL QDPPAGISGAPQDNNIMLWNAVIFGPDD+PWDGGTFKLTLQF
Sbjct: 1   MSTPARKRLMRDFKRLQQDPPAGISGAPQDNNIMLWNAVIFGPDDTPWDGGTFKLTLQFT 60
```

```
Query: 61  EEYPNKPPTVRFISRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS 120
           E+YPNKPPTVRF+SRMFHPNIYADGSICLDILQNQWSPIYD+AAILTSIQSLLCDPNPNS
Sbjct: 61  EDYPNKPPTVRFVSRMFHPNIYADGSICLDILQNQWSPIYDIAAILTSIQSLLCDPNPNS 120

Query: 121 PANSEAARMFSENKREYNRKVREIVEQSWTAD 152
           PANSEAARMFSENKREYNR+VREIVEQSWTAD
Sbjct: 121 PANSEAARMFSENKREYNRRVREIVEQSWTAD 152
```

Amino acid alignment of RAD6 with *Oryza sativa* RAD6 protein

```
gi|50919097|ref|XP_469945.1|  ubiquitin carrier protein [Oryza sativa (japonica
cultivar-group)]
gi|18844990|dbj|BAB85469.1| Rad6 [Oryza sativa (japonica cultivar-group)]
gi|28269456|gb|AAO37999.1| ubiquitin carrier protein [Oryza sativa (japonica cultivar-
group)]
Identities = 143/152 (94%), Positives = 148/152 (97%)

Query: 1   MSTPSRKRLMRDFKRLMQDPPAGISGAPQDNNIMLWNAVIFGPDDSPWDGGTFKLTLQFN 60
           MSTP+RKRLMRDFKRL QDPPAGISGAP DNNIMLWNAVIFGPDD+PWDGGTFKLTLQF
Sbjct: 1   MSTPARKRLMRDFKRLQQDPPAGISGAPHDNNIMLWNAVIFGPDDTPWDGGTFKLTLQFT 60

Query: 61  EEYPNKPPTVRFISRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS 120
           E+YPNKPP VRF+SRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS
Sbjct: 61  EDYPNKPPVVRFVSRMFHPNIYADGSICLDILQNQWSPIYDVAAILTSIQSLLCDPNPNS 120

Query: 121 PANSEAARMFSENKREYNRKVREIVEQSWTAD 152
           PANSEAAR+FSENKREYNRKVREIVEQSWTAD
Sbjct: 121 PANSEAARLFSENKREYNRKVREIVEQSWTAD 152
```

Sequencing result of WR5 gene fragment

```
Gggcgaaggggaagccgtgagataaacagggagtttagcgggatcaaagtgccccgcagggatcgccaatttatatacagcaggttc
aagccatgccgtgtttgccgtgatgaggcctcgctgttgacttgcatgacttttattggaggtgcatcaagggtggcggctgggaac
catagtggtgaattaagaatatttgacagcaaccttgcaaatctcattgagacgcacacttgtcaccagaatcttgttactctagtg
gttgcg
```

Amino acid alignment of WR5 with *Oryza sativa* OSJNBb0020J19.2

```
gi|50929885|ref|XP_474470.1|  OSJNBb0020J19.2 [Oryza sativa (japonica cultivar-group)]
gi|32487488|emb|CAE05773.1|  OSJNBb0020J19.2 [Oryza sativa (japonica cultivar-group)]
Length = 1878

Identities = 68/87 (78%), Positives = 78/87 (89%)
```

```
Query: 1    GRRGSREINREFSGIKVPRRDRQFIYSRFKPCRVCRDEASLLTCMTFIGGASRVAAGNHS 60
             R GSREI R+FSGI++PRRDRQFIYSRFK CRVCRDE+SLLTCMTF+G ASRVAAGNH+
Sbjct: 1468 ARMGSREIRRQFSGIQIPRRDRQFIYSRFKLCRVCRDESSLLTCMTFLGDASRVAAGNHT 1527

Query: 61   GELRIFDSNLANLIETHTCHQNLVTLV 87
             GELR+FD N AN++ET TCHQ LVT+V
Sbjct: 1528 GELRVFDCNTANILETQTCHQQLVTIV 1554
```

Amino acid alignment of WR5 with *Arabidopsis thaliana* transducin family protein / WD-40 repeat family protein

```
gi|7270018|emb|CAB79834.1|  putative protein [Arabidopsis thaliana]
gi|15235843|ref|NP_194845.1|  transducin family protein / WD-40 repeat family protein
[Arabidopsis  thaliana]
gi|7486499|pir||T10670  hypothetical protein F6E21.80 - Arabidopsis thaliana
Length = 1846


Identities = 40/87 (45%), Positives = 63/87 (72%)

Query: 1    GRRGSREINREFSGIKVPRRDRQFIYSRFKPCRVCRDEASLLTCMTFIGGASRVAAGNHS 60
             GR G+RE+   +SG+   RRDRQF++SRFK  R  RDE +L TC+  +GG + +A G+H+
Sbjct: 1391 GRLGTRELQSFYSGVHGNRRDRQFVFSRFKSWRSFRDETALFTCIALLGGTNHIAVGSHA 1450

Query: 61   GELRIFDSNLANLIETHTCHQNLVTLV 87
             GE++IF+++  +++E+ + HQ  VTLV
Sbjct: 1451 GEIKIFEASSGSMLESVSGHQAPVTLV 1477
```

APPENDIX C

DNA SEQUENCING RESULTS OF THE INTERACTORS

1K1-F3

ggatgaatcttcctacgatttattattcaatgaatcattcatcatggcatccgtaggtaaatatgacaaggctatcgaactattagaaaaagcgcttcaagga
gccacgaatgaaggctatcaaaacgatattaatactattaaattgcaattatcatttgttttgcaaatggtggggaaaactgcacaaagtaaagaaatcttaa
agggcttattgcaggaactgaaagccgattctccattctctttaatttgtcaaaataacctaaatgcgtttgttgatttctctaaatacaatacaaacttcaattta
ctgttaagggaattaaacgtagaaaaactaaacacttttaatttacaaaccttcacccatgaacaatggtcaaacattcaacgcaatgtacttttcctgcgttt
gttcaacaatgtcaaaattcactctcaagagtccctttttatcaagaacg

1K1-F5

cacttttaaagagagttcaaaagcgtgaaagatatgctaacaggaaaactttgttgaacaaaaatgctgtagcattaaccaagaaagtaagagaagaac
aaaaatctcaagtggtgcacggcccaagaggggagtcattgattgacaatcaaaggaaaagatataattatgaagatgtggacacaattccaactacaa
agcataaggctcctccccctcctccgccaacggccgaaacatttgattcagaccaaacaagttcattttccgatatcaattcgacaacagcatccgcaccg
actaccccagcccctgctcttcctcctgcatctcctgaagtaagaaaagaagaaacgcatccaaagcatagtttaccgcctttaccaaatcagtttgcgcc
attaccagaccctccacaacataactctccacctcaaaataacgcgccttcgcaaccccaaagcaatccatttccattcccaattcctgaaattccctcgac
acagtctgcaacaaacccatttccatttccggtacctcagcagcagtttaatcaagc

4K1-B2

acttcatttttgtaaacaagttggaactggaacttactatagaacaagagcttaaaccatggtaatacaaactgctactacgttatcgccagctaaggcgag
accttctttccctcataacgatttgattaaaagtatgtcagacagccttataagtcgccccactcaccctccgattaggaaattaaagtcatccttgaaaatatc
acaccctgaacctatttcaagatcaaagtcagaaattttctgcacatcacctgagaaaaatgtacgttttgctattgaattaacaactgttaagagatttgataa
aaacgcggagccaagttctatctcaaatgaaaactcacccacattatcccctgttgatagcaataccgctgctgatgacgtccagcttttcaacaatgagg
actgctggttcaatgattcttcgttggtcacaaacttgctgaaaaatgaaaagaagtttcgttatatgaattccttaaacaacatgttcaagttggacctttatga
ttcagaagatgaggatgatattgacgagcacatcaatagccagg

4K1-C1

acatcacgcatacttattttctactctagggatgagttgtatgctaacaacatagctagatcgggaatattatctagattgttgaaagtaattggaagtttccaa
aaactagacaaaataaacttttttagagtcatctttccttttattaaccaggagatgttttgaaacgactgaaaatgttgatgctttaatacgtgcagaaattaata
gaagctttacagcaagaccgttaggtggtggggatgatgctgtacgtgagttaaccactattctagaagaaaaggcgcatgtggtcatgcgtagtccctct
caatttattgacgttctttgcgaaacagctcggttccatgaatttgatgatcaaggtgcgttggttgattactcattgaaaagatttcttggtgaaaaagataag
aatacgcaagcctcaagtacagaaaaatcagatatctacgagagaacggggatcatgcatttattactgtcacaacttatggctgcttctgagaaagactg
gttgtc

1L1-A1

gctcaattcttttattgtaaggtaaattatcgtagaaccacctagcagtattccaacagtagtaattagagaaatgttgggatccagaagaatgaaatacaaa
taaaatatctggaagaaccagtgtattagcatcaagagataaaacgaatagttttcggtataatatccaaagagatgacgacttgttgatacgattgtttgat
gaacaaagaagccgttgaatcagtcttacttctttgaccaataaatccggtattactaaaacactgttgtcctggaagtatctcttgagtaaatcttcgtgtaca
ttatataccgttttgttttcaggtgttgtaagtagaaatccaacagtcatgtaataaatattctcattagaaatgccgcctaatttacatcaatatgtagaattgtta
gtagccgtaacaatgaatccaaaatgcgtattgtaaaatagcaatagttttttgtattttaccgttgagacatactttgatcagagaaagttttgttaaaatatgt
atgtgcttcttggaaaagagataaaaatgttgttttgaagataattagttctatgga

4K1-D2

gcctgattttgaatttttcagtatgcagcagtcacagcaaccgcagttccagcctagcagtagaagagaatcaaactctatgagacctccactgttaatacc
tgcagcaactactaaaagccagtccaatggcaccaataatagtgggaatatgaacacaaatgcagattatgaatcatttttttaatactggtacaaacaaca
gtaattctaatcagaacccgtacttttttgagttcaagaaataattctttgaagtttaatcctgaagatttcgattttcaattcaaaaggcggaattcttttgttagag
gtactttggaccatagcagccaaaatgcgtttatacctgaatcaagattaaactcactatccgttaacaacaaagctaatggtgatcctgtcgcggataatgt
taccaataacatgaaaggaaaaagtaatgaagttgacaatgatgatggcaatgacagtagcaataacaataacaacaacaacaacaacaacaacaatg
aaaacaacaatgacaacaataatgataataacgacaatagtattaattccgccaccagtac

2L1-A1

Gctcaattcttttattgtaaggtaaattatcgtagaaccacctagcagtattccaacagtagtaattagagaaatgttgggatccagaagaatgaaatacaaa
taaaatatctggaagaaccagtgtattagcatcaagagataaaacgaatagttttcggtataatatccaaagagatgacgacttgttgatacgattgtttgat
gaacaaagaagccgttgaatcagtcttacttctttgaccaataaatccggtattactaaaacactgttgtcctggaagtatctcttgagtaaatcttcgtgtaca
ttatataccgttttgttttcaggtgttgtaagtagaaatccaacagtcatgtaataaatattctcattagaaatgccgcctaatttacatcaatatgtagaattgtta
gtagccgtaacaatgaatccaaaatgcgtattgtaaaatagcaatagttttttgtattttaccgttgagacatactttgatcagagaaagttttgttaaaatatgt
atgtgcttcttggaaaagagataaaaatgttgttttgaagataattagttctatgga

4K1-G1

actttgtttcaatccataaagtgatttcttcaaacgtatcaacttatcattcattcctaaatgtggtggaggtcttatgtataattcttctttgatgtctgcatacaaat
atgccgaagatatgtctaattgtgtaatatagtagttattgtctaatgcaagtgacagggatgtcattaatgcatagtgatgtacggtattggattgcatgcctg
agtcgtaagtgtcaggatgctgaatatcacctcttgcaacaaatctagctttatgagtaccgtcacgtttcttgttgaagataaacattgagtttattactctttta
gggtctatttctttttctgtcataatattcgtcagtgtcccaagttttcatcttcaacagttgattgacttctttgtggtatgcctcgatatatttttctttttctttaatatct
ttattataggtgattgcctcatcgtatcgtaaggttgtccgtattgg

2L1-A4

gctcaattcttttattgtaaggtaaattatcgtagaaccacctagcagtattccaacagtagtaattagagaaatgttgggatccagaagaatgaaatacaaa
taaaatatctggaagaaccagtgtattagcatcaagagataaaacgaatagttttcggtataatatccaaagagatgacgacttgttgatacgattgtttgat
gaacaaagaagccgttgaatcagtcttacttctttgaccaataaatccggtattactaaaacactgttgtcctggaagtatctcttgagtaaatcttcgtgtaca
ttatataccgttttgttttcaggtgttgtaagtagaaatccaacagtcatgtaataaatattctcattagaaatgccgcctaatttacatcaatatgtagaattgtta
gtagccgtaacaatgaatccaaaatgcgtattgtaaaatagcaatagttttttgtattttaccgttgagacatactttgatcagagaaagttttgttaaaatatgt
atgtgcttcttggaaaagagataaaaatgttgttttgaagataattagttctatgga

4K1-G2

gatgaaatcaaacaaatggtggcgaaagacagtccaacgtacaagaacttgacaaagcaaggttacaaagttattggttcccattcaggcgttaagatct
gccgatggactaagaatgaactacgtggtaaagggtcttgttataaaaaatcgctcttcaatatcgcgtccagtagatgcatggaattgactccttctttggc
atgttcatccaaatgcgttttctgttggaggcatggtacaaatcccgtgtcaaaaaactggagatgggaagtagatgaaccagaatacattctagaaaatg
ctctgaagggacactattccatgatcaaacagatgagaggtgtgcccggtgttattgcagaaagatttgcgaaagcgtttgaagttcgtcattgtgccttgt
ctcttgtcggcgagcctattctttatcctcatatcaataaatttatccaattattacaccaaaagggcataactagtttccttgtgtgcaatgcccaacatccaga
ggccttgagaaatattgttaaagtaacacaattatatgtttctattgacgcacccaccaagacagagctgaaaaaagtggatagacc

2L1-A6

gctcaattcttttattgtaaggtaaattatcgtagaaccacctagcagtattccaacagtagtaattagagaaatgttgggatccagaagaatgaaatacaaa
taaaatatctggaagaaccagtgtattagcatcaagagataaaacgaatagttttcggtataatatccaaagagatgacgacttgttgatacgattgtttgat
gaacaaagaagccgttgaatcagtcttacttctttgaccaataaatccggtattactaaaacactgttgtcctggaagtatctcttgagtaaatcttcgtgtaca
ttatataccgttttgttttcaggtgttgtaagtagaaatccaacagtcatgtaataaatattctcattagaaatgccgcctaatttacatcaatatgtagaattgtta
gtagccgtaacaatgaatccaaaatgcgtattgtaaaatagcaatagttttttgtattttaccgttgagacatactttgatcagagaaagttttgttaaaatatgt
atgtgcttcttggaaaagagataaaaatgttgttttgaagataattagttctatgga

2K1-A1

tacaggaaaagttggcacatttgattcattggataatgaacactacttctcggaaacaaacaatgaaaaacaaaattcaatttagcttagaagcccaattgg
aaatattacttcttattaaatcttcgttcattaaaggtggcatcaatgttaaatacacttttccagcaataatcacaaatttttggaaactgatgaggaaatgccgt
atgatacaagagtacctttttgaaaaaaagacccgataacaagacgttactttcccattattccaatcttttaaagcaaatgtttaaatttgtttctcgttgtatcaa
tgatatctttaattcttgcaacaactcatgcacagatctgattctgaaactgaatttacaatgtgccattctagctgagcaattgcaattaaacgaaatttcatat
gatttttttctcacaggccttcacaatatttgaagagtctctaagtgattcaaagactcagttacaggctttaatatatattgctcagtctttacaaaagacaagat
cactctacaaagaagcttattatgattctttgattgtcagatgcacactccatggatccaaattattaaagaaacaagaccaatgtcgtgctgtttatttatgctc
ccacctctggtgggcaacggaaatttcaaatattggtgaggaagaaggtatcacagacaacttctacagagatggtaaacgggtgttagaatgtctacaa
aggtctctccgtgtggcggattctataatggataacgaacagagttgtgagctaatggtcgaaatattaaatagatgcctctactattttattcacggggatg
agtccgaaactcatatatcgataaaatacatcaatggtctgatcgaactaatcaaaacaaacttaaaatcgctaaa

2K1-A6

ttccggctgcaatcaggtgaattcgtttcttcgatctcggaggttctaaactacgcatattcttagtattccatgtgtctcgtgataccttaatttcagtttcattat
cttctaatgatcttttcttactgttgatagtagtataggcattagagtcaccaataccacccaaactggaattagtttgatgagaatttatcggtgggagttcttta
aatgggtcagggaattcggtaggagattctggaggtagatcagggagtgggagatcagcgatgatagattcctcggtattctgttcagaaacagtagttg
gcgttttgataggaacaatattgtgcgatgaattattttccggtggagaagcatcgattgaaggtgaacggtgtataatccttttctcagtctcttggtcacttat
ctgcggaacagttttgttgttggtaccacccgtactggatattggtacgtttgtatgattagtctcattttcactgtacgagtctgagtgtctgaaatctttagattt
actggcgtgcgacgactcatgt

5L1-A6

gctggataaagtacctgatgttaatatatccgaaaaagaagaatctgagcaagaaggtaaagaaggtaaggaggaaggggataaagaggagagatgg
aggttttcaggttatccttacactggagttaacgaatttgctatatactgtacatctgaatttctttcaatgggtgcaggtgatggtcattatggtcttttatgcgat
gacggcctactccatggtgtttcaaacccttgtcaaacatatggcaatgaagttttgagtaaggagggcaaaaaattttctatagtagctttggaagtatggc
gtgtaggatagtgtcatctatttaaatatcatactatgtatttatgaatgagaacctcatgattgtaaacgagaaacactccttttaaatagcctaaaccgtatcct
cagtaagaaaaactagaaaatacggtcctaatacttgcatgtgttttatgagtgacgaaataaataggaaaagaaaaatattttaatacaaactatcatttatt
gaaatgctctttttcattttcttcattaaatcattgtatgcaggctg

5L1-C1

aaggacattgtatatatagatcctatgagtgacaactaggtaaccgaagaaacactattagatatatagtttgagggatagcataagtagaacttagaaatg
agtttagtaccatatgaagaaggttcgctgattttagatgatcctaattctaaatcagtggtggttgttaatccaacctcaggaactttatcatttttccaacagg
acaacagtaataatgactctgaactaaatgaagaccagactgcatctttatcagcactggactttccttcagggattcatcaatacaaaagccctattgcttc
ctatgtgtgcccgcaatgtggtacggaaattaaccctgatattataaacagaagacaacttcatcgtcgagcctctgctggagtagaatcagaaagcagta

ggctcagtattccagagaacactgtgccacttggatttgaatttgccaattcaagcttttcaagaagatattttcagtcattggaaagaaatcatcggcactac
g

5L1-A1

Ggccaagcaatttaagtttgggtccgtttggtcccataaacgaaccgtcaagccgcaaaatatttgcttatctgattgctatcctcaacgcttcttatcctgac
catgatttttcatcggttgagccaacggattttgtcaaaacatcattgaaaacttttatttccaaatttgaaaacaccttatattctcttggtagacaaccagagg
aatgggtctgggaggtaattaattctcacatgactctttctgattgcgtccttttttcaatattcaccttcaaactctttttttggaagatgagcctggctatctttgga
atcttataggttttctttacaacaggaaaaggaaaagagtggcttacctttacttgatttgctcgcgtctaaattcgagtacaggcgaagtggaagatgcctt
ggcaaaaaaacctcagggaaagcttataatagatgatggctcaaatgaatacgaaggagaatacgatttcacttatgatgagaatgtaatagatgataaat
cagatcaagaagaatccctacagtagttttaaactttaatttgtga

5L1-B1

gatacttacaataagatattcaaagatctgaaaatccccttttgttagtgcgtgggcagatagcggtgacattggcggtgagtttagcaaagaatttcacctg
attcatgaatccggtgaggataccttgatgagttgcaaacattgcggcgatatatcaacttggatatgtcccagtcatatcctgaaaaagatggtcaatattc
aggtgatgtcgattgtaagtatgctttgactaaggatcactcaacattgatatgtttctactatccaaaggatagacaattgaattggaatttagcattaaatgc
tatggataaagacatcgatttgacgttgaggaacaaaccaaatgatcacgtcttacaagtctacgaaaaagataacgaagatattatgttcagtaaaatatt
aagagttatggactgtagattgaattcaaaatccaactttcctgattttccactaaagaaatacctgaaaaataattttggtcaaatcagtgacgtttctatagtg
gatgcacaggaaaacgaaatctgtggcaaatgtgaggaaggaagg

5K1-A1

gctatcttgatagcccatttttaaattcaccattcgttaataaaatggcaactacagacccttttgatttaagtgatgatgaaaaactggactgtgatgatgaaa
tagacgaaagtgctgctgaagtgtggtttagtaaaacgggtggtgaacatgttatggtcagtgtgaagtttcaagaaatgcggaccagtgatgatttaggg
gtcttgcaagatgtgaatcatgtggaccacgaagaactagaggaaagggaaaaggaatggaaaaagaaaattgatcaatatatcgaaacaaatgtgga
taaggatagcgaatttggcagtttgatcttgtttgacaaactaatgtattctgatgatggcgaacaatgggtagacaataaccttgtaatattattctccaaattt
ctagttttgtttgattttgaagaaatgaagatcctgggtaagattcctagagaccagttttaccaggtaatcaagttcaatgaggacgtactactttgtagtttg
aagagcactaatatacctgaaatatactt

1K2-D5

tgatacacacaaataatgccatctcttccagtaaaagataagtaagataaactggctatgttatgctgtctgtttatatgttgatacaaggaattaagatctatct
tatatgccttagcgtagcgcaagaattgaacatttctatataaaataagtataagacacgtagggttactacttctctgtgtgcatatacctgtgttacggtatt
gcatgcgagatggggctagacgagaccgtagggacatatagcaccttagaaataaccttgtatgaaatctttcattatttcctgtttagacacttgcgtcaa
ggtatttttttttatcgtggtgttgaatgtggtaacccaatagcataatatgagtaatgctttagtattgtttcagagctgttttagataaagaaaacatatagtagc

aaacctctaatccggtagtacttaagaaactatagtttctatgtacaaaacggtaactatgtaattcttacatttacataacatgtagaagaatccaataaacttactacaatatgacatataagctagatcgtaattcattacgtcaacataccgtaaaattatttggcgtttcgttaactcctacaaaaacaacaaacctatcaccgtatcgtcttctttagata

1K2-F1

gatttggcaaacgttgaagttctcgctgcagataacacgagggtaccgctgtacatgctgatggtagcggttcacaaagagctggatagcgatgatgttccagacggtagatttgatataatattactatgtagagattcgagcagagaagttggagagtgaaggaaattgttgttacgaaagtcagtgattatgtattgtgtagtatagtatattgtaagaaatttttttttctagggaatatgcgttttgatgtagtagtatttcactgttttgatttagtgtttgttgcacggcagtagcgagagacaagtgggaaagagtaggataaaaagacaatctataaaaagtaaacataaaataaaggtagtaagtagcttttggttgaacatccgggtaagagacaacaggggcttggaggagacgtacatgagggctatttagggctatttagggctatgtagaagtgctgtagggctaaagaacagggtttcattttcatttttttttttttaattcggtcagaaagccgggtaaggtatgacagcgagagtagaggtagatgtgagagagtgtgtgggtgtggtgtgt

1K2-F5

atgtgtacccatttgaactatttttttcaacaacttctaggtagtagtctaagttgtatttcttacctggctgaagcatgtcaccagagtaacaaacagtagcttcgacaacaccaccggccttcttgacagcattcacaccaacttttaattgttctaaatcattcaaggcatcaaaaactctaaatatatcaacaccattatccttggcttgcttgacaaaatggtcaatagcattgtcaggtaatgaagagtaagccacaccgttggcaccacgtaataacatttggaatggaatattaggcaccagagatcttaattttctcagacgttcccatggatcctcatgcaagaatctcattgcaacgtcgaatgtagcaccaccccaacattctaaagcgaaagcacctgcaagggcatgtgcggttgttggagcgattgtagccaaatcgtgggttctgactcttgttgcaagtagagattgatgagcgtctctccaggtggtgtccatcagtagagtaccattgaactgtctgacttgcttggcaaat

2KL2-H3

ggcgcgatttcagattttgaaattgaaatcctgttttccttacaagaacaaattctgagaagaaagactcaactcactgcttacaatatactgcttagtgaactggaaatattattatcgttcgctcaagtgtctgctgaaaggaattatgcagagcctcaattggtggaagatgagtgcatattggaaattattaacggaagacatgctttgtatgaaacattccttgataattatatccccaatagcacaatgattgatggcgggctattttctgaattaagttggtgtgaacaaaataaaggaagaatcattgtagtcactggcgctaatgcatctggaaagtctgtatatcttacacagaatggtttaattgtgtacttggcacaaattggttgttttgttccagcagagagagcgagaattggaatagcggataaaatattaactanaatcaggactcangaaactgtttataagactcaaagttccttttgctagattctcaaca

4L2-D1

atcccaaacatgaatttggatcttcagagtaagatttatgtgtgtacgcatttttacccaatatttcccatatttcttggaactgagagttcgtatccattttaccggcatcagtcattcttacttttgaggctccattggtttcatggaacaaagaatcaatatcaggacttccatgtttcctttccttttttggatgaataatagtggtttcgt

tggtagctttttaaattgaaggacttgaatatcggaaccttactcttagagcttgaggtcatttttttgcctacaccgaaaccgtagggctgttgctcataaggt
ttattttgcctatgataagcttctgattgaccaagtgtgggactagccccttgtgattcctttgccttggtcttgaaaatatgacagtttttcttcggctggtaagt
gtgtaaacctttgattttcaaatctcgatcatatcgaagtttttgatcttcatcagttaaaattgaatgtgcatgcaccacagctttgaataattcctcagacttatc
agattttgttttatctgggtgaagtaatctggcc

3L2-B4

ataggaaatgagcttttctcaattctctaaacttatacaagcactcatgtttgccgctctgatggtgcggaaaaaactgctccatgaagcaaactgtccggg
caaatcctttcacgctcgggaagctttgtgaaagcccttctctttcaacccatctttgcaacgaaaaaaaaaaaaaaaaataaaaaataaaaagaccaaata
gtaaatagtaacttacatacattagtaaatggtacactcttacacactatcatcctcatcgtatattataatagatatatacaatacatgtttttacccggatcata
gaattcttaagacaaataaaatttatagagacttgttcagtctacttctctctaaactaggccccggctcctgccagtacccacttagaaagaaataaaaaac
aaatcagacaacaaaggcttaatctcagcagatcgtaacaacaaggctactctactgcttacaataccccgttgtacatctaagtcgtatacaaatgatttat
ccccacgcaaaatgacattgcaattcgccagcaagcacccaaggcctttccgccaagtgcaccgttgc

4L2-F5

gccagaggcttcgtttgttgctactaccaccactttggactgtgaagatagtgatgaacaagatgagaacgacaaattggaattgaccttagatgggacca
atgacggcaggacaatcaagttggatccaacatacttggctccgggcgaaaaattagttggatacgttctcgtgaagatgaacgacgaccccgaccag
caaaacgagccaccaaatggtcatattacttccttaagtgtcatgagaacctatagaaggatggggattgccgaaaacttaatgagacaagctttatttgcg
ttgagagaagttcatcaagcggaatatgtttccttgcatgtgagacaatctaatagagcagctttacacttgtacagggacacattggcgtttgaggtgttga
gtatcgagaaaagctactatcaggatggtgaagatgcatacgccatgaagaaggttttaaaattagaggagttgcagataagcaacttcacccatcgccg
tttgaaagagaatgaggaaaaactagaagacgatctagaaagcgatctactagaggatatc

3L2-F1

tttcacgaatgttgtcgcaagtaggctttgcaaggtccttcaaggaagtaaagaatgaactaaaggcgcaactgaacatatcttctacggcattgaatgattt
ggagttatttgattttagactggacttcgaagcagccaaaaaacgcctgtataaattgatgattgatagtccgcatctaaaaaaaattgaggactctttgtccc
atatatcaaaggttctcagttacctaaaaacccttagaagttgcaagaaatgttgtgatatctccttgagtaactacaatagcgctttttacaaaggaggtatca
tgtttcatgcagtttatgacgatggatcctcacgtaatatgatagctgctggagggaggtatgacactttgatatccttttttgccagaccatcaggaaaaaa
gagcagcaatactcgtaaggctgtaggtttcaacttagcgtgggaaacaatattcggtatagcccaaaactatttcaaactcgcttctggaaataggataaa
gaagagaaataggtttttgaaagataca

2KL2-B1

cccaatgccaccattaccacaggggttcccaatcccaccaaatggtatgttaccggtaacaggacaacagccacagccaccttatccaaatatgatgcta
caaggaaatttccccccctaatttccagcaaggttttggtagtaactcccctatgccaattccatctattattaatgccaatggcaaaaatgttacaaaccagtt
accaccaggattaaactctaaaaagaatataaaatgagatgcgtaacgagcgagtacttgacaggaaacgacaaaaacacagtgaacatactatctaaa
agcatgttcttttccttttcatttttcatctttattatctagagtttgtatataacatgtatcctatattatcttatcctatcatcatcacttaacttgttataattctcccata
tacttatgttttcgcctttgatatttactagtcgaaggaacttttttcgcagttagtaacgtcacgctcggaactaaacattaatgaatgagatgcatttttataca
acacttttacaggctatatacaactaagtgattttcaatacatt

3L2-G5

gctgatgatggtgcaggaactacagacgaagattacatggcctggaaggactccatcctggaggttttgaaagacgaactgcatttggacgaacagga
agccaagttcacctctcaattccagtacactgtgttgaacgaaatcactgactccatgtcgcttggtgaaccctctgctcactatttgccctcgcatcagttga
accgcaacgcagacggcatccaattgggtcccttcgatttgtctcaaccgtatattgcacccatcgtgaaatctcgcgaactgttctcttccaatgaccgta
attgcatccactctgaatttgacttgtccggctctaacatcaagtactccactggtgaccatcttgctgtttggccttccaacccattggaaaaggtcgaaca
gttcttatccatattcaacctggaccctgaaaccatttttgacttgaagcccctggatcccaccgtcaaagtgcccttcccaacgccaactactattggcgct
gctattaaacactatttggaaattacaggacctgtctcca

2KL2-D6

gctttacttttaatcagtcaaaccgaagcgccctttgcatggagttgacgtatcatggaatttggtcacagagaaaactattggacctttggaagagtatgt
acacggcattgtgcagccgatggaaatttttcgactgattcagtcttatcaaccaacttgaatcatttggccagttgcttgaccagttcaatcacgaggtcaa
atgaggcaactttggtaaattcagtacgcaaactagaaaggatatgtcgatacttgtccgatactgtggcttcgttgaaagaacaacaactagatttcaagtt
ccaaaatgtatttatcctgataatcctcgcgttaaaagaattatccgccatgaatatgacaatactgcctaaccataaagacacttttattctatgatctgtttaa
gccttttccacgtccacgtattgactcaaaaaatcgggacagtcggttttcccagttacgattatgtttacgataatttagtaacgtatttcattgtcatggatga
cttgtctaaaatcactacagttttggaactaatgaaaagaaacaacacc

4L2-B4

Ccaatttaggttctccaaatacgaacaattcccagtttttcataactacatacgccgcacctcatttgaacggcaagcattctatttttgggcaggtggttcat
ggaaaatcagtggttcgcactatagaaaactgcagggtagattctgatggagtccctgaatcagacgttagaatcagcgactgtggagtgtgggaaaaa
actatgggtgtcccgctttacaacgcttctaatgaccaaattggcggtgacgtatacgaagaatatccagacgatgatacacatttcggtgacgatgatttt
ggtaaggctcttgaggctgctaatataatcaaagaatctgggacgctcttatttaagaaaaaggattactctaatgccttttttcaaatacagaaaatctttaaat
tatatcaatgaatacatgcctgaacccgacgtagacaaagaaagaaacatccaatttattaacctaaagatgaaaatttacctaaatttatctttggtgctcttt
aatctggaaagatacgatgatgctatcatgtatgctacgtatttattggaga

2KL2-F3

gcccacgaaggtattgttggtggcgcagagtttggttacgatatcagcgccggttccatttctcgttatgccatggctttaagttatttcgccaaagactactc
cttgggcgctacattgaacaacgagcaaataactaccgttgacttcttccaaaacgtcaacgccttttacaggtcggtgctaaggctacaatgaactgca
aactacctaactccaatgtcaacatcgaattcgccactagatatttgcctgatgcatcttcccaagttaaggctaaggtgtccgattccggtattgtcactttg
gcttacaagcaattgttaagacctggcgtcactctgggtgtcggttcctctttcgatgctttgaagttgtctgaacctgttcacaagctaggttggtctttgtcct
tcgacgcttgaacg

2KL2-G6

ttccggccccaatggcgtccatgtggttntggtatttgaagtcctcggcgagaacttactggctttgataaagaagtatgaacataggggcattccattgatt
tatgtcaagcaaatatcgaagcagttactgcttggtcttgattacatgcataggcggtgtggtattatacacacagatattaaaccagaaaacgtgttgatgg
agattggagatgttgagggtatcgttcaaatggtcgaggcactggataaacagaagagagaggcaaaaaggttgcagaggcacgtttcgcggtcttctg
atattaccgcgaacgatagcagtgatgagaaatgggcagaatgccaaacaagtatgccctgcggctcaagctcaaattccaaatctcgttctatagagaa
ggacttgtcaaaaagatgcttnngaagacctagacgtcatacaattatcacagggtcgcaaccattaccttccccaattagttcatcaaatttctttgaaatg
agggcgcacttttgtggtagtagtcannaatagcttttcctccgtatctggggaataggaacattccat

5L2-G5

Gcttcttaaaaaatgcatttaatgccttaaaatcaggaggaactctattgttcctaggattgctgaggttgtttttgtaaaaaatttggaatatcaagaacatgt
cacagaatatggggttcattggaatttttttatcaccctatcattgttgccacttgtattgacctttattgatcccgtcacaagaatggttccacgctgctcaattg
caatattcatttcatgcatttatgaatggctacttttaaaggacgatcgcactttaaacttttaattttggctgatagaaattgtttcttcagtgctaatagagaag
gcatcttctcatttctaggttattgctcgatttttctttggggccaaaacacgggattttacttgttgggaaataaaccaactttaaacaatctttataagccttcta
cgcaagacgtagttgcagcatcaaagaagtcttcgacttgggactattggacttcagtaaccccattaagtggcctctgtatatggagtacaatttttcttgtt
atcagccagttggtttttcaataccatccttatag

2KL2-H1

gcctacacttttgagtgatttagcttccagattgagtaatgatgatatggtgacgaataaaggtgtccttacagtggcacattctatttttaaaagatggagac
ctttatttagatcagatgaacttttttttggagattaaattggttcttgacgtgtttactgctccatttttgaacttattgaaaacggtcgatgaacagataacagcg
aatgaaaataacaaggcatcgctaaatattttatttgatgtattgctagtattaattaaactatactacgattttaattgtcaagatataccagagtttttgaggat
aacattcaagtgggtatgggtatcttccataagtatttgtcatattctaatcctttattggaagaccctgacgaaactgaacatgcgtctgtcctaataaaagta
aagtcctctatccaggagctggttcaattgtacacaacaagatatgaagatgtctttggacctatgatcaatgaattcatacaaataacttggaatcttctgac
ctcaatttcaaaccaacctaaatacgacatcttagtatccaagtccttgtcatttttt

5K2-D2

tttatcagattaattcacggaatgttacttatcttatatattatataaaatatgaatcatactaagtggtggaagcgcggaatctcggatctaaactaattgttcag
gcatttatacttttgggtagttcagctagggaaggacgggttttgtctcatgttgttcgttttgttataaggttgtttcatatgtgttttatgaacgtttaggatgac
gtattgtcatactgacgtatctcattttgagatacaacagaatgagtggtattaaagaaataaggaataccataaggttgatcaaaattaaatattggaagga
aggtaactagctttttgactttattattgttggaactaataatcattaatattcgacactttaaaaatttttttataagattaatatcgttgcaattaatacaccgtttgtt
gctatgctaacaatgtgaatacgcacaccgtatatttatgttcatttcgtatatgtttttgcatgggtgtgccttttgcagaagacagaaatgttgttctggcaag
ggagacaatagttcagttgaggtaaggccctatta

5K2-E3/1

tttatcagattaattcacggaatgttacttatcttatatattatataaaatatgaatcatactaagtggtggaagcgcggaatctcggatctaaactaattgttcag
gcatttatacttttgggtagttcagctagggaaggacgggttttgtctcatgttgttcgttttgttataaggttgtttcatatgtgttttatgaacgtttaggatgac
gtattgtcatactgacgtatctcattttgagatacaacagaatgagtggtattaaagaaataaggaataccataaggttgatcaaaattaaatattggaagga
aggtaactagctttttgactttattattgttggaactaataatcattaatattcgacactttaaaaatttttttataagattaatatcgttgcaattaatacaccgtttgtt
gctatgctaacaatgtgaatacgcacaccgtatatttatgttcatttcgtatatgtttttgcatgggtgtgccttttgcagaagacagaaatgttgttctggcaag
ggagacaatagttcagttgaggtaaggccctatta

5K2-E3/2

tgtattacccccatatctttggtacttaagtaaaataagtgaacacaatgtaataaacagaatggattagggctttatgacttcccagtacaaaataaatgtcg
atataatacttgaaaatgcaagcaatgccatcaaaaaatttgagcggtacgagaatcgaacccgtgtccccaccttggaagggtgggatgataaccacta
cactaaccgcccatgtgttggatggtgtaactgtaatgacgaggtaataagtgatggcgactatatcaccaccgaaacacagcgattacttaaaagggta
atcactactgcgattctaattgttggaataaaccaaacctcttccgataatgttggaacgagagtaattaatagtgacatgtgttgctatggtaacaatctaat
gcttacatcgtatattaatgtacacctcgtatacgtttaagtgtgattgcacctattgcagaaggaatgttaaacgagaagctcag

5K2-G3

gggccagcatgttggatgctgtggctccaggacagacctgtctgctgacacagtggaactgatagaaagaatggacagattggctgaaaaacaggcg
acagcttccatgtcgatcgttgcgttaccgtctagcttccaggagagcaatagcagtgacaggtgcagaaagtattgcagcagtgatgaggacagcaac
acgtgcattcatggtagtgctaatgccagtaccaatgcgactaccaactccagcactaatgctactaccactgccagcaccaacgtcaggactagtgcta
ctaccactgccagcatcaacgtcaggactagtgcgactaccactgaaagtaccaactccagcactaatgctactaccactgccagcaccaacgtcagg
actagtgctactaccactgccagcatcaacgtcaggactagtgcgactaccactgaaagtaccaactccaacactagtgctactaccaccgaaagtacc
gactccaacactagtgctactaccactgaaagtaccaactccagcactaatgctactaccactgc

5K2-G4

Ccttcactagaacaatctcgacatcttgtagcactttgaacggcgcctcaacccaaaccagtgagctaaccacatcgcctatgaaaaccaacacggtgg
ttccagcttcttctttcccttcaactacaaccacttgtctagaaaatgatgacactgccttttctagtatctacactgaagtcaacgccgcaactatcattaaccc
cggagaaacatcttctctcgctagcgatttcgccacatctgaaaagccaaacgagcccacttctgtcaaatccacctcaaacgaaggcacctcttccaca
acaacaacctaccaacagactgttgctacactgtatgccaagccctccagcacaagcctaggtgcaagaacaactactggtagcaacggtcgttcaact
accagccaacaagacgggtctgccatgcatcagccaacttcctcgatctacactcaactaaaagaaggcacatcaaccaccgcaaaactttctgcatac
gaaggtgctgcaacacctctttccattttccagtgcaatagtctagctggaacgat

APPENDIX D

DNA AND PROTEIN SEQUENCE ALIGNMENTS OF YR10-1506 AND YR10-1562 INTERACTORS

1L-A1, 2L1-A1 and 2L1-A4 have the same sequences which is the *S. cerevisiae* centromeric region CEN11. The region codes for yeast "Extracellular matrix protein 9" (ECM9).

DNA sequence alignment

gi|3517|emb|X65124.1|SCCEN11D S.cerevisiae centromeric region **CEN11**

```
Score = 1154 bits (582),  Expect = 0.0
Identities = 582/582 (100%), Gaps = 0/582 (0%)

Query  1      GCTCAATTCTTTTATTGTAAGGTAAATTATCGTAGAACCACCTAGCAGTATTCCAACAGT  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14961  GCTCAATTCTTTTATTGTAAGGTAAATTATCGTAGAACCACCTAGCAGTATTCCAACAGT  14902

Query  61     AGTAATTAGAGAAATGTTGGGATCCAGAAGAATGAAATACAAATAAAATATCTGGAAGAA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14901  AGTAATTAGAGAAATGTTGGGATCCAGAAGAATGAAATACAAATAAAATATCTGGAAGAA  14842

Query  121    CCAGTGTATTAGCATCAAGAGATAAAACGAATAGTTTTCGGTATAATATCCAAAGAGATG  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14841  CCAGTGTATTAGCATCAAGAGATAAAACGAATAGTTTTCGGTATAATATCCAAAGAGATG  14782

Query  181    ACGACTTGTTGATACGATTGTTTGATGAACAAAGAAGCCGTTGAATCAGTCTTACTTCTT  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14781  ACGACTTGTTGATACGATTGTTTGATGAACAAAGAAGCCGTTGAATCAGTCTTACTTCTT  14722

Query  241    TGACCAATAAATCCGGTATTACTAAAACACTGTTGTCCTGGAAGTATCTCTTGAGTAAAT  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14721  TGACCAATAAATCCGGTATTACTAAAACACTGTTGTCCTGGAAGTATCTCTTGAGTAAAT  14662

Query  301    CTTCGTGTACATTATATACCGTTTTGTTTTCAGGTGTTGTAAGTAGAAATCCAACAGTCA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14661  CTTCGTGTACATTATATACCGTTTTGTTTTCAGGTGTTGTAAGTAGAAATCCAACAGTCA  14602
```

```
Query  361    TGTAATAAATATTCTCATTAGAAATGCCGCCTAATTTACATCAATATGTAGAATTGTTAG  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14601  TGTAATAAATATTCTCATTAGAAATGCCGCCTAATTTACATCAATATGTAGAATTGTTAG  14542

Query  421    TAGCCGTAACAATGAATCCAAAATGCGTATTGTAAAATAGCAATAGTTTTTTGTATTTTA  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14541  TAGCCGTAACAATGAATCCAAAATGCGTATTGTAAAATAGCAATAGTTTTTTGTATTTTA  14482
Query  481    CCGTTGAGACATACTTTGATCAGAGAAAGTTTTGTTAAAATATGTATGTGCTTCTTGGAA  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  14481  CCGTTGAGACATACTTTGATCAGAGAAAGTTTTGTTAAAATATGTATGTGCTTCTTGGAA  14422

Query  541    AAGAGATAAAAATGTTGTTTTGAAGATAATTAGTTCTATGGA  582
              ||||||||||||||||||||||||||||||||||||||||||
Sbjct  14421  AAGAGATAAAAATGTTGTTTTGAAGATAATTAGTTCTATGGA  14380
```

DNA sequence alignment

gi|60416363|sp|Q02202|ECM9_YEAST Extracellular matrix protein 9

Identities = 129/129 (100%), Positives = 129/129 (100%), Gaps = 0/129 (0%)

```
Query  388  GISNENIYYMTVGFLLTTPENKTVYNVHEDLLKRYFQDNSVLVIPDLLVKEVRLIQRLLC  209
            GISNENIYYMTVGFLLTTPENKTVYNVHEDLLKRYFQDNSVLVIPDLLVKEVRLIQRLLC
Sbjct  77   GISNENIYYMTVGFLLTTPENKTVYNVHEDLLKRYFQDNSVLVIPDLLVKEVRLIQRLLC  136

Query  208  SSNNRINKSSSLWILYRKLFVLSLDANTLVLPDILFVFHSSGSQHFSNYYCWNTARWFYD  29
            SSNNRINKSSSLWILYRKLFVLSLDANTLVLPDILFVFHSSGSQHFSNYYCWNTARWFYD
Sbjct  137  SSNNRINKSSSLWILYRKLFVLSLDANTLVLPDILFVFHSSGSQHFSNYYCWNTARWFYD  196

Query  28   NLPYNKRIE  2
            NLPYNKRIE
Sbjct  197  NLPYNKRIE  205
```

Identities = 30/30 (100%), Positives = 30/30 (100%), Gaps = 0/30 (0%)

```
Query  582  SIELIIFKTTFLSLFQEAHTYFNKTFSDQS  493
            SIELIIFKTTFLSLFQEAHTYFNKTFSDQS
Sbjct  47   SIELIIFKTTFLSLFQEAHTYFNKTFSDQS  7
```

APPENDIX E

PROTEIN SEQUENCE ALIGNMENTS AND CONSERVED DOMAINS OF YR10-1 INTERACTORS

1K1-F3 protein sequence alignment

gi|37362700|ref|NP_015114.2| Core component of the signal recognition particle (SRP) ribonucleoprotein (RNP) complex that functions in targeting nascent secretory proteins to the endoplasmic reticulum (ER) membrane [Saccharomyces cerevisiae]

Identities = 155/155 (100%), Positives = 155/155 (100%)

```
Query: 1   DESSYDLLFNESFIMASVGKYDKAIELLEKALQGATNEGYQNDINTIKLQLSFVLQMVGK 60
           DESSYDLLFNESFIMASVGKYDKAIELLEKALQGATNEGYQNDINTIKLQLSFVLQMVGK
Sbjct: 184 DESSYDLLFNESFIMASVGKYDKAIELLEKALQGATNEGYQNDINTIKLQLSFVLQMVGK 243

Query: 61  TAQSKEILKGLLQELKADSPFSLICQNNLNAFVDFSKYNTNFNLLLRELNVEKLNTFNLQ 120
           TAQSKEILKGLLQELKADSPFSLICQNNLNAFVDFSKYNTNFNLLLRELNVEKLNTFNLQ
Sbjct: 244 TAQSKEILKGLLQELKADSPFSLICQNNLNAFVDFSKYNTNFNLLLRELNVEKLNTFNLQ 303

Query: 121 TFTHEQWSNIQRNVLFLRLFNNVKIHSQESLLSRT 155
           TFTHEQWSNIQRNVLFLRLFNNVKIHSQESLLSRT
Sbjct: 304 TFTHEQWSNIQRNVLFLRLFNNVKIHSQESLLSRT 338
```

gi|1370436|emb|CAA97925.1| **SRP72** [Saccharomyces cerevisiae]
gi|730825|sp|P38688|SRP72_YEAST Signal recognition particle 72 kDa protein homolog (SRP72)
gi|531018|gb|AAA53400.1| signal recognition particle protein

Identities = 155/155 (100%), Positives = 155/155 (100%)

```
Query: 1   DESSYDLLFNESFIMASVGKYDKAIELLEKALQGATNEGYQNDINTIKLQLSFVLQMVGK 60
           DESSYDLLFNESFIMASVGKYDKAIELLEKALQGATNEGYQNDINTIKLQLSFVLQMVGK
Sbjct: 190 DESSYDLLFNESFIMASVGKYDKAIELLEKALQGATNEGYQNDINTIKLQLSFVLQMVGK 249

Query: 61  TAQSKEILKGLLQELKADSPFSLICQNNLNAFVDFSKYNTNFNLLLRELNVEKLNTFNLQ 120
           TAQSKEILKGLLQELKADSPFSLICQNNLNAFVDFSKYNTNFNLLLRELNVEKLNTFNLQ
Sbjct: 250 TAQSKEILKGLLQELKADSPFSLICQNNLNAFVDFSKYNTNFNLLLRELNVEKLNTFNLQ 309

Query: 121 TFTHEQWSNIQRNVLFLRLFNNVKIHSQESLLSRT 155
           TFTHEQWSNIQRNVLFLRLFNNVKIHSQESLLSRT
Sbjct: 310 TFTHEQWSNIQRNVLFLRLFNNVKIHSQESLLSRT 344
```

1K1-F5 protein sequence alignment

gi|1420437|emb|CAA99390.1| **LAS17** [Saccharomyces cerevisiae]
gi|2498506|sp|Q12446|LAS17_YEAST Proline-rich protein **LAS17**
gi|6324755|ref|NP_014824.1| Actin assembly factor, activates the Arp2/3 protein complex that nucleates branched actin filaments; localizes with the Arp2/3 complex to actin patches; homolog of the human Wiskott-Aldrich syndrome protein (WASP) [Saccharomyces cerevisiae]

gi|1101757|dbj|BAA11386.1| prolin rich protein [Saccharomyces cerevisiae]

Identities = 133/133 (100%), Positives = 133/133 (100%)

```
Query: 1   HFLKRVQKRERYANRKTLLNKNAVALTKKVREEQKSQVVHGPRGESLIDNQRKRYNYEDV 60
           HFLKRVQKRERYANRKTLLNKNAVALTKKVREEQKSQVVHGPRGESLIDNQRKRYNYEDV
Sbjct: 115 HFLKRVQKRERYANRKTLLNKNAVALTKKVREEQKSQVVHGPRGESLIDNQRKRYNYEDV 174

Query: 61  DTIPTTKHKAPPPPPPTAETFDSDQTSSFSDINSTTASAPTTPAPALPPASPEVRKEETH 120
           DTIPTTKHKAPPPPPPTAETFDSDQTSSFSDINSTTASAPTTPAPALPPASPEVRKEETH
Sbjct: 175 DTIPTTKHKAPPPPPPTAETFDSDQTSSFSDINSTTASAPTTPAPALPPASPEVRKEETH 234

Query: 121 PKHSLPPLPNQFA 133
           PKHSLPPLPNQFA
Sbjct: 235 PKHSLPPLPNQFA 247
```

4K1-B2 protein sequence alignment

gi|1420431|emb|CAA99387.1| **GAC1** [Saccharomyces cerevisiae]
gi|2507224|sp|P28006|GAC1_YEAST Protein phosphatase 1 regulatory subunit **GAC1**
gi|6324752|ref|NP_014821.1| **Gac1p** [Saccharomyces cerevisiae]

Identities = 168/168 (100%), Positives = 168/168 (100%)

```
Query: 1   MVIQTATTLSPAKARPSFPHNDLIKSMSDSLISRPTHPPIRKLKSSLKISHPEPISRSKS 60
           MVIQTATTLSPAKARPSFPHNDLIKSMSDSLISRPTHPPIRKLKSSLKISHPEPISRSKS
Sbjct: 1   MVIQTATTLSPAKARPSFPHNDLIKSMSDSLISRPTHPPIRKLKSSLKISHPEPISRSKS 60

Query: 61  EIFCTSPEKNVRFAIELTTVKRFDKNAEPSSISNENSPTLSPVDSNTAADDVQLFNNEDC 120
           EIFCTSPEKNVRFAIELTTVKRFDKNAEPSSISNENSPTLSPVDSNTAADDVQLFNNEDC
Sbjct: 61  EIFCTSPEKNVRFAIELTTVKRFDKNAEPSSISNENSPTLSPVDSNTAADDVQLFNNEDC 120

Query: 121 WFNDSSLVTNLLKNEKKFRYMNSLNNMFKLDLYDSEDEDDIDEHINSQ 168
           WFNDSSLVTNLLKNEKKFRYMNSLNNMFKLDLYDSEDEDDIDEHINSQ
Sbjct: 121 WFNDSSLVTNLLKNEKKFRYMNSLNNMFKLDLYDSEDEDDIDEHINSQ 168
```

gi|3714|emb|CAA45371.1| **GAC1** [Saccharomyces cerevisiae]

Identities = 168/168 (100%), Positives = 168/168 (100%)

```
Query: 1   MVIQTATTLSPAKARPSFPHNDLIKSMSDSLISRPTHPPIRKLKSSLKISHPEPISRSKS 60
           MVIQTATTLSPAKARPSFPHNDLIKSMSDSLISRPTHPPIRKLKSSLKISHPEPISRSKS
Sbjct: 1   MVIQTATTLSPAKARPSFPHNDLIKSMSDSLISRPTHPPIRKLKSSLKISHPEPISRSKS 60

Query: 61  EIFCTSPEKNVRFAIELTTVKRFDKNAEPSSISNENSPTLSPVDSNTAADDVQLFNNEDC 120
           EIFCTSPEKNVRFAIELTTVKRFDKNAEPSSISNENSPTLSPVDSNTAADDVQLFNNEDC
Sbjct: 61  EIFCTSPEKNVRFAIELTTVKRFDKNAEPSSISNENSPTLSPVDSNTAADDVQLFNNEDC 120
```

```
Query: 121 WFNDSSLVTNLLKNEKKFRYMNSLNNMFKLDLYDSEDEDDIDEHINSQ 168
           WFNDSSLVTNLLKNEKKFRYMNSLNNMFKLDLYDSEDEDDIDEHINSQ
Sbjct: 121 WFNDSSLVTNLLKNEKKFRYMNSLNNMFKLDLYDSEDEDDIDEHINSQ 168
```

4K1-C1 protein sequence alignment

gi|927738|gb|AAB64910.1| **Tom1p**; CAI: 0.16 [Saccharomyces cerevisiae]
gi|50401412|sp|Q03280|TOM1_YEAST **E3 ubiquitin protein ligase TOM1** (Temperature dependent-organization in mitotic nucleus protein 1) (Suppressor of snRNA protein 2)
gi|6320665|ref|NP_010745.1| Temperature dependent Organization in Mitotic nucleus; hect-domain-containing protein, containing kinase motifs; similar to Rsp5 [Saccharomyces cerevisiae]

```
Identities = 174/174 (100%), Positives = 174/174 (100%)

Query: 1    TSRILIFYSRDELYANNIARSGILSRLLKVIGSFQKLDKINFLESSFLLLTRRCFETTEN 60
           TSRILIFYSRDELYANNIARSGILSRLLKVIGSFQKLDKINFLESSFLLLTRRCFETTEN
Sbjct: 1456 TSRILIFYSRDELYANNIARSGILSRLLKVIGSFQKLDKINFLESSFLLLTRRCFETTEN 1515

Query: 61   VDALIRAEINRSFTARPLGGGDDAVRELTTILEEKAHVVMRSPSQFIDVLCETARFHEFD 120
           VDALIRAEINRSFTARPLGGGDDAVRELTTILEEKAHVVMRSPSQFIDVLCETARFHEFD
Sbjct: 1516 VDALIRAEINRSFTARPLGGGDDAVRELTTILEEKAHVVMRSPSQFIDVLCETARFHEFD 1575

Query: 121  DQGALVDYSLKRFLGEKDKNTQASSTEKSDIYERTGIMHLLLSQLMAASEKDWL 174
           DQGALVDYSLKRFLGEKDKNTQASSTEKSDIYERTGIMHLLLSQLMAASEKDWL
Sbjct: 1576 DQGALVDYSLKRFLGEKDKNTQASSTEKSDIYERTGIMHLLLSQLMAASEKDWL 1629
```

Figure E.1 TOM1 structure and conserved domains

gnl|CDD|14812 cd00078, HECTc, HECT domain; C-terminal catalytic domain of a subclass of Ubiquitin-protein ligase (E3). It binds specific ubiquitin-conjugating enzymes (E2), accepts ubiquitin from E2, transfers ubiquitin to substrate lysine side chains, and transfers additional ubiquitin molecules to the end of growing ubiquitin chains.

```
CD-Length = 352 residues,  99.4% aligned
Score =  391 bits (1006), Expect = 7e-109
```

```
Query: 2913  ITVRREQVFLDSYRALFFKTNDEIKNSKLEITFKGESGVDAGGVTREWYQVLSRQMFNPD  2972
Sbjct:  3     ITVRRDRILEDALRQLSKVSSSDLK-KVLEVEFVGEEGIDAGGVTREFFTLVSKELFNPS  61

Query: 2973  YALFLPVPSDKTTFHPNRTSGINPEHLSFFKFIGMIIGKAIRDQCFLDCHFSREVYKNIL  3032
Sbjct:  62    YGLFRYTPDDSGLLYPNPSSFADEDHLKLFRFLGRLLGKALYEGRLLDLPFSRAFYKKLL  121

Query:  3033  GRPVSLKDMESLDPDYYKSLVWILENDI-TDIIEETFSVETDD-YGEHKVINLIEGGKDI  309
Sbjct:  122   GKPLSLEDLEELDPELYKSLKELLDNDGDEDDLELTFTIELDSSFGGAVTVELKPGGRDI  181

Query: 3091  IVTEANKQDYVKKVVEYKLQTSVKEQMDNFLVGFYALISKDLITIFDEQELELLISGLPD  3150
Sbjct:  182  PVTNENKEEYVDLYVDYRLNKGIEEQVEAFRDGFSEVIPEELLSLFTPEELELLICGSED  241


Query: 3151  IDVDDWKNNTTYVN-YTATCKEVSYFWRAVRSFDAEERAKLLQFVTGTSKVPLNGFKELS  3209
Sbjct:  242  IDLEDLKKNTEYKGGYSSDSPTIQWFWEVLESFTNEERKKFLQFVTGSSRLPVGGFADLN  301

Query:  3210  GVNGVCKFSIHRDFGSSERLPSSHTCFNQLNLPPYESYETLRGSLLLAINEGHEGFG  3266
Sbjct:  302   P-----KFTIRRVGSPDDRLPTAHTCFNLLKLPPYSSKEILREKLLYAINEG-AGFG  352


gnl|CDD|24898 pfam06025, DUF913, Domain of Unknown Function (DUF913).

CD-Length = 372 residues,  98.7% aligned
Score =  363 bits (934), Expect = 1e-100

Query:  445   LASATHLLETFIDNSETTTEFIENDGFTMLITSVANEIDFTL----AHPETWQPPKYSVV  500
Sbjct:  6     VTRAVRILDLFTDYNIAFALFRENGGLDMLINRVVHEVDECLKSSDSKLGTGPPEQLSYV  65

Query:  501   YYSISFRELAYIRSLLKLVLKLLSTDSGDR-IRNLIDSPILVSLKKILENKLVFGLTLIT  559
Sbjct:  66    DYSISYQRAALLKSLLKFIKKLIQDPGSFDGLRNLMDSSLPESLCHILRNAEYYGPSLFS  125

Query:  560   YTLDVVQKVINSEPTIYPVLVEAGLIPYVIDNFP-KLIGPSAELLSLLPDVVSAICLNPE  618
Sbjct:  126   LATTVVSDFIHNEPTCLSSLQEAGLTSAFLDALLRKDIPPSAEVITSLPNVFSAICLNNR  185

Query:  619   GLKQVKEKGLINNLFDFLLDADHARILTGG-DRSTEYGTDIDELARHYPDLKANIVEALC  677
Sbjct:  186   GLKQFKSKNPINQLFKIFLSPDHLVALRGRRDTATNLGSALDELMRHQPSLKTDIVKAII  245

Query:  678   NVIRKMPSTFRNEREFLFTSPK----------DQKYFFHRKNEEILTDKEEHEPAYW  724
Sbjct:  246   KILNEVLILGRNPEATCSKSAKLPADAASVSPATVDQGAAEERDSEDIDVDEEEASSSVS  305

Query:  725   ELLDKGTMLDTFTSVLFGMSLGNGSFSQVPQHLEARDFLAIIFMENPPYEYFTSVAISNV  784
Sbjct:  306   EEQQPATA----------TSPGNIELPLIDYILNVARFLETILSNSPPYEHCREF-----  350

Query:  785   TEVLQYLDEKYEDYAFMDVMKVLNDQLENLNDFLNSPNDRSFF  827
Sbjct:  351   ----------------------VSKGGLEPLLDLLSLPNLPLDF  372


gnl|CDD|24885 pfam06012, DUF908, Domain of Unknown Function (DUF908).

CD-Length = 290 residues, 100.0% aligned
Score =  269 bits (690), Expect = 3e-72

Query:  90    EAHDVDYCLKMLKFTRRLLLNTENRFVYSSGDVLMYLLNCPNFTIKLAVMRILAILGERF  149
Sbjct:  1     EEIDKDLLLLAILRFTRLLLLENCFNRSIYNSIEHLTLLLNSPNMDVVAAVLNLLYVFGKRS  60

Query:  150   VIAREKIVAHNIFGDHNLRKKTLKLALSLSSSVMDEDGEHFSLVDLYFDKKKVPQKWRKL  209
Sbjct:  61    NISERKINLPSKKQDLLLPKLFSLAQR------WGGKGYGFGLVDCYFSEPKLEQLWKNL  114

Query:  210   RFTHYTSNDFKKSSQQKNNINETQTSIKKVTMTTQELCEHSLQQIFDKGMAL--LPAESW  267
Sbjct:  115   DFTYYDSDGQKINS----EINETDESFKVLHIPTVKKISKSLNELFNKLMEDYNIPEQYK  170

Query:  268   FDFSIKASVAKAFSDDSGENIDLRNIIIETKLNAIAFVNTIFSPPQVSSKLFELDPYAFN  327
Sbjct:  171   FDLLHRLRLAVAFSDPKR-----RQQAIRIRLLAISNLVYSSADQDVSSKFLYSDPGFLE  225
```

```
Query:  328   SLTDLISLSETKIPKELRTDALFTLECISLKHVWCSDIIRNLGGNISHGLLFQILRYIAK  387
Sbjct:  226   ELVELLK-YEDKVPKEIRAICLRTLEAISHKHRKIPSVLAALGVNVHHGLLSVLVRYIID  284

Query:  388   TLREAT  393
Sbjct:  285   ELKSPT  290
```

gnl|CDD|14151 COG5021, HUL4, Ubiquitin-protein ligase [Posttranslational modification, protein turnover, chaperones].

```
CD-Length = 872 residues, 100.0% aligned
Score =  753 bits (1944), Expect = 0.0

Query: 2385  DIREGILDTEAEEQRMFGRIGSSADVIRA*DDDVSNNDEEVEN*GLDHGNSNDRNNADPEKK  2444
Sbjct:  1     DLRVGGLLLEDLSCRLFSELFRSNSSVRAEFDLSKDESGVRNSLDYGAAGNKNMSLSDEK  60

Query: 2445  KPARIYFAPLIDRAGIASLMKSVFISKPYIQREIYHELFYRLCSSKQNRNDLMNTFLFIL  2504
Sbjct:  61    GLVRSSIAALDGLQNRDCLRSLDPLSVLSVDGLQTSETSFRSSALNPYVNEFLCENDVRL  120

Query: 2505  SEGIIDQHSLEKVYNIISSRAMGHAKTTTVRQLPSDCTPLTVANQTIEILQSLIDADSRL  2564
Sbjct:  121   SSSITIQVSDESKQNVIEDVFSGLENLGSVDVLS------TEATKGIDFLEILITRDFLF  174

Query: 2565  KYFLIAEHDNLIVNKANNKSRKEALPDKKLRWPLWHLFSLLDRKLITDESVLMDLLTRIL  2624
Sbjct:  175   SSCSLNSDFLKIISGSSVKSRKLAVSNVEKSEPDNVLFGLPYRSALTMRTNFLKLDTGNL  234

Query: 2625  QVCTKTLAVLSTSSNGKENLSKKFHLPSFDEDDLMKILSIIMLDSCTTRVFQQTLNIIYN  2684
Sbjct:  235   SGEVQALLARYISIKLVIK--KLYLGPGPDASSRISTLIIRLSNTNLNRRLSYILSHSSF  292

Query: 2685  LSKLQGCMSIFTKHLVSLAISIMSKLKSALDGLSREVGTITTGMEINSELLQKFTLPSSD  2744
Sbjct:  293   EDSLLRLNSLFSTRADSFGRTYYLDHDRILTQYSRPLLEETLGESTSFLVVNNDD--SSS  350

Query: 2745  QAKLLKILTTVDFLYTHKRKEEERNVKDLQSLYDKMNGGPVWSSLSECLSQFEKSQAINT  2804
Sbjct:  351   IKDLPHQVGSNPFLEAHPEFSELLKNQSRGTTRDFRNKPTGWSSSIEDLGQFLFSDFLTS  410

Query: 2805  SATILLPLIESLMVVCRRSDLSQNRNTAV-KYEDAKLLDFSKTRVENLFFPFTDAHKKLL  2863
Sbjct:  411   SSTYEDLRREQLGRESDESFYVASNVQQQRASREGPLLSGWKTRLNNLYRFYFVEHRKKT  470

Query: 2864  NQMIRSNPKLMSGPFALLVKNPKVLDFDNKRYFFNAKLK-SDNQERPKLPITVRREQVFL  2922
Sbjct:  471   LTKNDSR----LGSFISLNKLDIRRIKEDKRRKLFYSLKQKAKIFDPYLHIKVRRDRVFE  526

Query: 2923  DSYRALFFKTNDEIKNSKLEITFKGESGVDAGGVTREWYQVLSRQMFNPDYALFLPVPSD  2982
Sbjct:  527   DSYREIMDESGDDLK-KTLEIEFVGEEGIDAGGLTREWLFLLSKEMFNPDYGLFEYITED  585

Query: 2983  KTTFHPNRTSGINPEHLSFFKFIGMIIGKAIRDQCFLDCHFSREVYKNILGRPVSLKDME  3042
Sbjct:  586   LYTLPINPLSSINPEHLSYFKFLGRVIGKAIYDSRILDVQFSKAFYKKLLGKPVSLVDLE  645

Query: 3043  SLDPDYYKSLVWILENDITD-IIEETFSVETDDYGEHKVINLIEGGKDIIVTEANKQDYV  3101
Sbjct:  646   SLDPELYRSLVWLLNNDIDETILDLTFTVEDDSFGESRTVELIPNGRNISVTNENKKEYV  705

Query: 3102  KKVVEYKLQTSVKEQMDNFLVGFYALISKDLITIFDEQELELLISGLPD-IDVDDWKNNT  3160
Sbjct:  706   KKVVDYKLNKRVEKQFSAFKSGFSEIIPPDLLQIFDESELELLIGGIPEDIDIDDWKSNT  765

Query: 3161  TYVNYTATCKEVSYFWRAVRSFDAEERAKLLQFVTGTSKVPLNGFKELSGVNGVCKFSIH  3220
Sbjct:  766   AYHGYTEDSPIIVWFWEIISEFDFEERAKLLQFVTGTSRIPINGFKDLQGSDGVRKFTIE  825

Query:  3221  RDFGSSERLPSSHTCFNQLNLPPYESYETLRGSLLLAINEGHEGFGLA  3268
Sbjct:  826   KGGTDDDRLPSAHTCFNRLKLPEYSSKEKLRSKLLTAINEGA-GFGLL  872
```

4K1-D2 protein sequence alignment

```
gi|1420303|emb|CAA99311.1|  AZF1 [Saccharomyces cerevisiae]
gi|525274|emb|CAA81212.1|  zinc finger protein of 101170 Da [Saccharomyces cerevisiae]
gi|1164958|emb|CAA64033.1|  YOR3244w [Saccharomyces cerevisiae]
gi|1050816|emb|CAA62111.1|  ORF O3244 [Saccharomyces cerevisiae]
gi|1168610|sp|P41696|AZF1_YEAST   Asparagine-rich zinc finger protein AZF1
gi|6324687|ref|NP_014756.1|  Azf1p [Saccharomyces cerevisiae]


Identities = 121/121 (100%), Positives = 121/121 (100%)

Query: 25  NSMRPPLLIPAATTKSQSNGTNNSGNMNTNADYESFFNTGTNNSNSNQNPYFLSSRNNSL 84
           NSMRPPLLIPAATTKSQSNGTNNSGNMNTNADYESFFNTGTNNSNSNQNPYFLSSRNNSL
Sbjct: 228 NSMRPPLLIPAATTKSQSNGTNNSGNMNTNADYESFFNTGTNNSNSNQNPYFLSSRNNSL 287

Query: 85  KFNPEDFDFQFKRRNSFVRGTLDHSSQNAFIPESRLNSLSVNNKANGDPVADNVTNNMKG 144
           KFNPEDFDFQFKRRNSFVRGTLDHSSQNAFIPESRLNSLSVNNKANGDPVADNVTNNMKG
Sbjct: 288 KFNPEDFDFQFKRRNSFVRGTLDHSSQNAFIPESRLNSLSVNNKANGDPVADNVTNNMKG 347

Query: 145 K 145
           K
Sbjct: 348 K 348
```

4K1-G2 protein sequence alignment

```
gi|1370430|emb|CAA97922.1|  unnamed protein product [Saccharomyces cerevisiae]
gi|2501612|sp|Q08960|YP07_YEAST Hypothetical UPF0026 protein YPL207w
gi|6325049|ref|NP_015117.1|  Hypothetical ORF; Ypl207wp [Saccharomyces cerevisiae]


Identities = 199/199 (100%), Positives = 199/199 (100%)


Query: 1    DEIKQMVAKDSPTYKNLTKQGYKVIGSHSGVKICRWTKNELRGKGSCYKKSLFNIASSRC 180
           DEIKQMVAKDSPTYKNLTKQGYKVIGSHSGVKICRWTKNELRGKGSCYKKSLFNIASSRC
Sbjct: 411  DEIKQMVAKDSPTYKNLTKQGYKVIGSHSGVKICRWTKNELRGKGSCYKKSLFNIASSRC 470

Query: 181  MELTPSLACSSKCVFCWRHGTNPVSKNWRWEVDEPEYILENALKGHYSMIKQMRGVPGVI 360
           MELTPSLACSSKCVFCWRHGTNPVSKNWRWEVDEPEYILENALKGHYSMIKQMRGVPGVI
Sbjct: 471  MELTPSLACSSKCVFCWRHGTNPVSKNWRWEVDEPEYILENALKGHYSMIKQMRGVPGVI 530

Query: 361  AERFAKAFEVRHCALSLVGEPILYPHINKFIQLLHQKGITSFLVCNAQHPEALRNIVKVT 540
           AERFAKAFEVRHCALSLVGEPILYPHINKFIQLLHQKGITSFLVCNAQHPEALRNIVKVT
Sbjct: 531  AERFAKAFEVRHCALSLVGEPILYPHINKFIQLLHQKGITSFLVCNAQHPEALRNIVKVT 590

Query: 541  QLYVSIDAPTKTELKKVDR 597
           QLYVSIDAPTKTELKKVDR
Sbjct: 591  QLYVSIDAPTKTELKKVDR 609
```

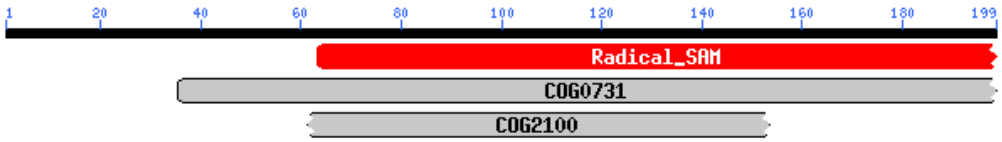


Figure E.2 YPL207w structure and conserved domains

```
gnl|CDD|26181 pfam04055, Radical_SAM, Radical SAM superfamily.

CD-Length = 167 residues, only  73.7% aligned
Score = 36.5 bits (83), Expect = 0.002

Query:  63   LTPSLACSSKCVFCWRHGTNPVSKNWRWEVDEPEYILENALKGHYSMIKQMRGVPGVIAE  122
Sbjct:  1    IIITPGCNLRCTYCAFPSIGRARGKGREL--SLEEILE------------------EAK  39

Query:  123  RFAKAFEVRHCALSLVGEPILYPHINKFIQLLHQK----GITSFLVCNAQH--PEALRNI  176
Sbjct:  40   ELKARLGVEVVILT-GGEPLLLPDLVELLERLLKLEEDEGIRITLETNGTLLDEELLEEL  98

Query:  177  VK--VTQLYVSIDAPTKTELKKVDR  199
Sbjct:  99   KEAGLDRVSIGLQSGDDEVLKLINR  123


gnl|CDD|10600  COG0731,  COG0731,  Fe-S  oxidoreductases  [Energy  production  and
conversion].

CD-Length = 296 residues, only  50.7% aligned
Score =  153 bits (387), Expect = 2e-38

Query:  35   RWTKNELRGKGSCYKKSLFNIASSRCMELTPSLA-CSSKCVFCWRHGTNPVSKNWRWEVD  93
Sbjct:  1    HWTKKALPIVFGPVKSRRYGISLG--IQMTPSKKWCSYNCVYCWR-GRTKKGTPER----  53

Query:  94   EPEYILENALKGHYSMIKQMRGVPGVIAERFAKAFEVRHCALSLVGEPILYPHINKFIQL  153
Sbjct:  54   -PEFIVEESILEELKLLLGYKG---------DEATEPDHVTISLSGEPTLYPNLGELIEE  103

Query:  154  LHQKG-ITSFLVCNAQHPEALRNIVKVTQLYVSIDAPTKTELKKVDR  199
Sbjct:  104  IKKRGKKTTFLVTNGSLPDVLEELKLPDQLYVSLDAPDEKTFRRINR  150


gnl|CDD|11808  COG2100,  COG2100,  Predicted  Fe-S  oxidoreductase  [General  function
prediction only].

CD-Length = 414 residues, only  18.4% aligned
Score = 37.6 bits (87), Expect = 0.001

Query:  61   MELTPSLACSSKCVFCWRHGTNPVSK-NWRWEVDEPEYILENALKGHYSMIKQMRGVPGV  119
Sbjct:  109  IQVRPSTGCNLNCIFC-SVDEGPYSRTRKLDYVVDPEYLLEWFEK---------------  152

Query:  120  IAERFAKAFEVRHCALSLVGEPILYPHINKFIQLL  154
Sbjct:  153  VARFKGKGLEAH---LDGQGEPLLYPHLVDLVQAL  184
```

Plant homologs

gi|21281261|gb|AAM45053.1| unknown protein [Arabidopsis thaliana]
gi|19423975|gb|AAL87267.1| unknown protein [Arabidopsis thaliana]
gi|22330642|ref|NP_177656.2| flavodoxin family protein / radical SAM domain-containing protein [Arabidopsis thaliana]

Identities = 120/200 (60%), Positives = 151/200 (75%), Gaps = 4/200 (2%)

```
Query: 1   DEIKQMVAKDSPTYK-NLTKQGYKVIGSHSGVKICRWTKNELRGKGSCYKKSLFNIASSR 59
           D  K+MV   +P  + +LTKQGYK+IGSHSGVKICRWTK++LRG+G CYK S + I S R
Sbjct: 266 DGKKEMV---TPVIRASLTKQGYKIIGSHSGVKICRWTKSQLRGRGGCYKHSFYGIESHR 322

Query: 60  CMELTPSLACSSKCVFCWRHGTNPVSKNWRWEVDEPEYILENALKGHYSMIKQMRGVPGV 119
           CME TPSLAC++KCVFCWRH TNPV K+W+W++DEP  I++ AL  H +MIKQM+GVPGV
Sbjct: 323 CMETTPSLACANKCVFCWRHHTNPVGKSWQWKMDEPSVIVKGALDLHKNMIKQMKGVPGV 382

Query: 120 IAERFAKAFEVRHCALSLVGEPILYPHINKFIQLLHQKGITSFLVCNAQHPEALRNIVKV 179
             E+  +    RHCALSLVGEPI+YP IN  +  LH + I++FLV NAQ PE +  +  +
Sbjct: 383 TPEKLQEGLNPRHCALSLVGEPIMYPEINALVDELHGRRISTFLVTNAQFPEKILMMKPI 442

Query: 180 TQLYVSIDAPTKTELKKVDR 199
           TQLYVS+DA TK  LK +DR
Sbjct: 443 TQLYVSVDAATKESLKAIDR 462
```

gi|34896894|ref|NP_909791.1| putative oxidoreductase [Oryza sativa (japonica cultivar-group)]
gi|24899457|gb|AAN65027.1| putative oxidoreductase [Oryza sativa (japonica cultivar-group)]
Length = 653

Identities = 114/184 (61%), Positives = 143/184 (77%)

```
Query: 16  NLTKQGYKVIGSHSGVKICRWTKNELRGKGSCYKKSLFNIASSRCMELTPSLACSSKCVF 75
           +L KQGYK+IGSHSGVKICRWTK++LRG+G CYK S + I S RCME TPSLAC++KCVF
Sbjct: 287 SLEKQGYKIIGSHSGVKICRWTKSQLRGRGGCYKHSFYGIESHRCMEATPSLACANKCVF 346

Query: 76  CWRHGTNPVSKNWRWEVDEPEYILENALKGHYSMIKQMRGVPGVIAERFAKAFEVRHCAL 135
           CWRH TNPV K+W+W++D+P  I+  A+  H  M+KQM+GVPGV  ER A+    RHCAL
Sbjct: 347 CWRHHTNPVGKSWKWKMDDPLDIVNAAIDQHTKMVKQMKGVPGVKPERLAEGLSPRHCAL 406

Query: 136 SLVGEPILYPHINKFIQLLHQKGITSFLVCNAQHPEALRNIVKVTQLYVSIDAPTKTELK 195
           SLVGEPI+YP IN  I  LH++ I++FLV NAQ P+ ++ +  +TQLYVS+DA TK  LK
Sbjct: 407 SLVGEPIMYPEINVLIDELHRRHISTFLVTNAQFPDKIKTLKPITQLYVSVDAATKESLK 466

Query: 196 KVDR 199
            VDR
Sbjct: 467 AVDR 470
```

4L2-F5 protein sequence alignment

gi|6321801|ref|NP_011877.1| Subunit of the N-terminal acetyltransferase NatA (Nat1p, Ard1p, Nat5p); N-terminally acetylates many proteins, which influences multiple

processes such as the cell cycle, heat-shock resistance, mating, sporulation, and telomeric silencing; Ard1p [Saccharomyces cerevisiae]
gi|1070532|pir||TWBYA1 protein N-acetyltransferase (EC 2.3.1.-) chain ARD1 - yeast Saccharomyces cerevisiae)
gi|45269535|gb|AAS56148.1| YHR013C [Saccharomyces cerevisiae]
gi|500702|gb|AAB68937.1| Ard1p: subunit of the major N alpha-acetyltransferase [Saccharomyces cerevisiae]
gi|728879|sp|P07347|ARD1_YEAST N-terminal acetyltransferase complex **ARD1** subunit **(Arrest-defective protein 1)**

Identities = 168/168 (100%), Positives = 168/168 (100%)

```
Query: 1   PEASFVATTTTLDCEDSDEQDENDKLELTLDGTNDGRTIKLDPTYLAPGEKLVGYVLVKM 60
           PEASFVATTTTLDCEDSDEQDENDKLELTLDGTNDGRTIKLDPTYLAPGEKLVGYVLVKM
Sbjct: 41  PEASFVATTTTLDCEDSDEQDENDKLELTLDGTNDGRTIKLDPTYLAPGEKLVGYVLVKM 100

Query: 61  NDDPDQQNEPPNGHITSLSVMRTYRRMGIAENLMRQALFALREVHQAEYVSLHVRQSNRA 120
           NDDPDQQNEPPNGHITSLSVMRTYRRMGIAENLMRQALFALREVHQAEYVSLHVRQSNRA
Sbjct: 101 NDDPDQQNEPPNGHITSLSVMRTYRRMGIAENLMRQALFALREVHQAEYVSLHVRQSNRA 160

Query: 121 ALHLYRDTLAFEVLSIEKSYYQDGEDAYAMKKVLKLEELQISNFTHRR 168
           ALHLYRDTLAFEVLSIEKSYYQDGEDAYAMKKVLKLEELQISNFTHRR
Sbjct: 161 ALHLYRDTLAFEVLSIEKSYYQDGEDAYAMKKVLKLEELQISNFTHRR 208
```

Figure E.3 ARD1 structure and conserved domains

gnl|CDD|25558 pfam00583, Acetyltransf_1, Acetyltransferase (GNAT) family. This family contains proteins with N-acetyltransferase functions.

```
CD-Length = 82 residues,  87.8% aligned
Score = 59.1 bits (143), Expect = 4e-10

Query:  89   GEKLVGYVLVKMNDDPDQQNEPPNGHITSLSVMRTYRRMGIAENLMRQALFALREVHQAE  148
Sbjct:  6    DGELVGFASLRPIDE-----EGNVAEIEGLAVDPEYRGKGIGTALLEALLEYARE-LGLK  59

Query:  149  YVSLHVRQSNRAALHLYR  166
Sbjct:  60   RIELEVLEDNEAAIALYE  77
```

gnl|CDD|10329 COG0456, RimI, Acetyltransferases [General function prediction only].

```
CD-Length = 177 residues,  94.9% aligned
Score = 86.2 bits (212), Expect = 3e-18
```

```
Query:  1    MPINIRRATINDIICMQNANLHNLPEN----YMMKYYMYHILSWPEASFVATTTTLDCED  56
Sbjct:  10   DKVTIREAINKDLLDVALAALEARTFDIRLPWSREYFEKDLTQAPELLLVAETGGLDGL-  68

Query:  57   SDEQDENDKLELTLDGTNDGRTIKLDPTYLAPGEKLVGYVLVKMNDDPDQQNEPPNGHIT  116
Sbjct:  69   ------------------------------LDGKVVGFLLVRVVDGRP--SADHEGHIY  95

Query:  117  SLSVMRTYRRMGIAENLMRQALFALREVHQAEYVSLHVRQSNRAALHLYRDTLAFEVLSI  176
Sbjct:  96   NLAVDPEYRGRGIGRALLDEALERLRERGLADKIVLEVRESNEAAIGLYRK-LGFEVVKI  154

Query:  177  EKSYYQDG-EDAYAMKKVLKLEE  198
Sbjct:  155  RKNYYADGNGDALLMLKMLNGLA  177
```

Plant homologs

gi|14550116|gb|AAK67148.1| silencing group B protein [Zea mays]

```
Identities = 58/107 (54%), Positives = 80/107 (74%), Gaps = 3/107 (2%)

Query: 49  GEKLVGYVLVKMNDDPDQQNEPPNGHITSLSVMRTYRRMGIAENLMRQALFALREVHQAE 108
           G ++VGYVL KM +DP   +EP +GHITSL+V+R++R++G+A  LM  A  A+ +V  AE
Sbjct: 50  GGRIVGYVLAKMEEDP---SEPCHGHITSLAVLRSHRKLGLATKLMSAAQAAMDQVFGAE 106

Query: 109 YVSLHVRQSNRAALHLYRDTLAFEVLSIEKSYYQDGEDAYAMKKVLK 155
           YVSLHVR+SNRAA +LY  TL +++  IE  YY DGEDA+ M+K L+
Sbjct: 107 YVSLHVRRSNRAAFNLYTSTLGYQIHDIEAKYYADGEDAFDMRKPLR 153
```

4L2-B4 protein sequence alignment

gi|1015677|emb|CAA89559.1| **CPR7** [Saccharomyces cerevisiae]
gi|1129166|emb|CAA60725.1| J1585 [Saccharomyces cerevisiae] gi|51013657|gb|AAT93122.1| YJR032W [Saccharomyces cerevisiae]
gi|6322492|ref|NP_012566.1| Peptidyl-prolyl cis-trans isomerase (cyclophilin), catalyzes the cis-trans isomerization of peptide bonds N-terminal to proline residues; binds to Hsp82p and contributes to chaperone activity; Cpr7p [Saccharomyces cerevisiae]
gi|1077877|pir||S57050 cyclophilin-like protein CPR7 - yeast (Saccharomyces cerevisiae)
gi|1176336|sp|P47103|CYP7_YEAST Peptidyl-prolyl cis-trans isomerase CYP7 (PPIase) (Rotamase)

```
Identities = 188/188 (100%), Positives = 188/188 (100%)

Query: 1   NLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRIS 60
           NLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRIS
Sbjct: 133 NLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRIS 192

Query: 61  DCGVWEKTMGVPLYNASNDQIGGDVYEEYPDDDTHFGDDDFGKALEAANIIKESGTLLFK 120
           DCGVWEKTMGVPLYNASNDQIGGDVYEEYPDDDTHFGDDDFGKALEAANIIKESGTLLFK
Sbjct: 193 DCGVWEKTMGVPLYNASNDQIGGDVYEEYPDDDTHFGDDDFGKALEAANIIKESGTLLFK 252

Query: 121 KKDYSNAFFKYRKSLNYINEYMPEPDVDKERNIQFINLKMKIYLNLSLVLFNLERYDDAI 180
           KKDYSNAFFKYRKSLNYINEYMPEPDVDKERNIQFINLKMKIYLNLSLVLFNLERYDDAI
Sbjct: 253 KKDYSNAFFKYRKSLNYINEYMPEPDVDKERNIQFINLKMKIYLNLSLVLFNLERYDDAI 312
```

```
Query: 181 MYATYLLE 188
           MYATYLLE
Sbjct: 313 MYATYLLE 320
```

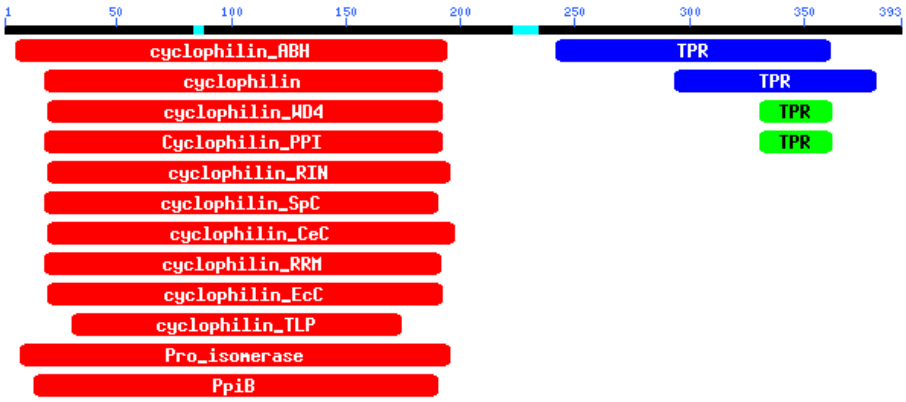


Figure E.4 CPR7 structure and conserved domains

gnl|CDD|28268 cd01926, cyclophilin_ABH_like, cyclophilin_ABH_like: Cyclophilin A, B and H-like cyclophilin-type peptidylprolyl cis- trans isomerase (PPIase) domain. This family represents the archetypal cystolic cyclophilin similar to human cyclophilins A, B and H. PPIase is an enzyme which accelerates protein folding by catalyzing the cis-trans isomerization of the peptide bonds preceding proline residues. These enzymes have been implicated in protein folding processes which depend on catalytic /chaperone-like activities. As cyclophilins, Human hCyP-A, human cyclophilin-B (hCyP-19), S. cerevisiae Cpr1 and C. elegans Cyp-3, are inhibited by the immunosuppressive drug cyclopsporin A (CsA). CsA binds to the PPIase active site. Cyp-3. S. cerevisiae Cpr1 interacts with the Rpd3 - Sin3 complex and in addition is a component of the Set3 complex. S. cerevisiae Cpr1 has also been shown to have a role in Zpr1p nuclear transport. Human cyclophilin H associates with the [U4/U6.U5] tri-snRNP particles of the splicesome.

```
CD-Length = 164 residues, 100.0% aligned
Score =  173 bits (439), Expect = 4e-44

Query:  5    PLVYLDISIDKKPIGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHR  64
Sbjct:  1    PKVFFDITIGGEPAGRIVMELFADVVPKTAENFRALCTGEKG---KGGKPFGYKGSTFHR  57

Query:  65   VVKNFMIQAGDIVFGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENL-GEF  123
Sbjct:  58   VIPDFMIQGGDFTRGN---------GTGGKSIYGEK--------------FPDENFKLKH  94

Query:  124  VEPFTLGMANLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDG  183
Sbjct:  95   TGPGLLSMANAG-PNTNGSQFFITTVKTPWLDGKHVVFGKVVEGMDVVKKIENVGSGN-G  152
```

```
Query:  184  VPESDVRISDCG  195
Sbjct:  153  KPKKKVVIADCG  164
```

gnl|CDD|28261 cd00317, cyclophilin, cyclophilin: cyclophilin-type peptidylprolyl cis-trans isomerases. This family contains eukaryotic, bacterial and archeal proteins which exhibit a peptidylprolyl cis- trans isomerases activity (PPIase, Rotamase) and in addition bind the immunosuppressive drug cyclosporin (CsA). Immunosuppression in vertebrates is believed to be the result of the cyclophilin A-cyclosporin protein drug complex binding to and inhibiting the protein-phosphatase calcineurin. PPIase is an enzyme which accelerates protein folding by catalyzing the cis-trans isomerization of the peptide bonds preceding proline residues. Cyclophilins are a diverse family in terms of function and have been implicated in protein folding processes which depend on catalytic /chaperone-like activities. This group contains human cyclophilin 40, a co-chaperone of the hsp90 chaperone system; human cyclophilin A, a chaperone in the HIV-1 infectious process and; human cyclophilin H, a component of the U4/U6 snRNP, whose isomerization or chaperoning activities may play a role in RNA splicing.

```
CD-Length = 146 residues,  96.6% aligned
Score =  146 bits (371), Expect = 3e-36

Query:  18   IGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIV  77
Sbjct:  6    KGRIVIELYGDEAPKTVENFLSLARGG-----------FYDGTTFHRVIPGFMIQGGDPT  54

Query:  78   FGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDEN--LGEFVEPFTLGMANLG  135
Sbjct:  55   GT-----------GGGGSGPGYK--------------FPDENFPLKYHHRRGTLSMANAG  89

Query:  136  SPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRISD  193
Sbjct:  90   -PNTNGSQFFITTAPTPHLDGKHTVFGKVVEGMDVVDKIERGDTDENGRPIKPVTISD  146
```

gnl|CDD|28269 cd01927, cyclophilin_WD40, cyclophilin_WD40: cyclophilin-type peptidylprolyl cis- trans isomerases (cyclophilins) having a WD40 domain. This group consists of several hypothetical and putative eukaryotic and bacterial proteins which have a cyclophilin domain and a WD40 domain. Function of the protein is not known.

```
CD-Length = 148 residues,  95.3% aligned
Score =  123 bits (309), Expect = 5e-29

Query:  19   GRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIVF  78
Sbjct:  7    GDIHIRLFPEEAPKTVENFTTHARNGY-----------YNNTIFHRVIKGFMIQTGD-PT  54

Query:  79   GTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDE---NLGEFVEPFTLGMANLG  135
Sbjct:  55   GD---------GTGGESIWG--------------KEFEDEFSPSL-KHDRPYTLSMANAG  90

Query:  136  SPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRISD  193
Sbjct:  91   -PNTNGSQFFITTVATPWLDNKHTVFGRVVKGMDVVQRIENVKTDKNDRPYEDIKIIN  147
```

gnl|CDD|28270 cd01928, Cyclophilin_PPIL3_like, Cyclophilin_PPIL3_like. Proteins similar to Human cyclophilin-like peptidylprolyl cis- trans isomerase (PPIL3). Members of this family lack a key residue important for cyclosporin binding: the tryptophan residue corresponding to W121 in human hCyP-18a; most members have a histidine at this position. The exact function of the protein is not known.

```
CD-Length = 153 residues,  92.8% aligned
Score =  118 bits (297), Expect = 1e-27

Query:  18   IGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIV  77
Sbjct:  9    LGDIKIELFCDDCPKACENFLALCASGY-----------YNGCIFHRNIKGFMVQTGD--  55

Query:  78   FGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGE--FVEPFTLGMANLG  135
Sbjct:  56   --------PTGTGKGGESIWGKK--------------FEDEFRETLKHDSRGVVSMANNG  93
```

```
Query:  136  SPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRISD  193
Sbjct:  94   -PNTNGSQFFITYAKQPHLDGKYTVFGKVIDGFETLDTLEKLPVDKKYRPLEEIRIKD  150
```

gnl|CDD|28265 cd01923, cyclophilin_RING, cyclophilin_RING: cyclophilin-type peptidylprolyl cis- trans isomerases (cyclophilins) having a modified RING finger domain. This group includes the nuclear proteins, Human hCyP-60 and Caenorhabditis elegans MOG-6 which, compared to the archetypal cyclophilin Human cyclophilin A exhibit reduced peptidylprolyl cis- trans isomerase activity and lack a residue important for cyclophilin binding. Human hCyP-60 has been shown to physically interact with the proteinase inhibitor peptide eglin c and; C. elegans MOG-6 to physically interact with MEP-1, a nuclear zinc finger protein. MOG-6 has been shown to function in germline sex determination.

```
CD-Length = 159 residues,  90.6% aligned
Score =  109 bits (275), Expect = 4e-25

Query:  19   GRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIVF  78
Sbjct:  9    GDLNLELHCDKAPKACENFIKLCK---------KGY--YDGTIFHRSIRNFMIQGGDPT-  56

Query:  79   GTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGEFVEPFT------LGMA  132
Sbjct:  57   GT---------GRGGESIWGK--------------PFKDE----FKPNLSHDGRGVLSMA  89

Query:  133  NLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRIS  192
Sbjct:  90   NSG-PNTNGSQFFITYRSCKHLDGKHTVFGRVVGGLETLEAMENVPDPGTDRPKEEIKIE  148

Query:  193  DCGV  196
Sbjct:  149  DTSV  152
```

gnl|CDD|28264 cd01922, cyclophilin_SpCYP2_like, cyclophilin_SpCYP2_like: cyclophilin 2-like peptidylprolyl cis- trans isomerase (PPIase) domain similar to Schizosaccharomyces pombe cyp-2. These proteins bind their respective SNW chromatin binding protein in autologous systems, in a CsA independent manner indicating interaction with a surface outside the PPIase active site. SNW proteins play a basic and broad range role in signaling.

```
CD-Length = 146 residues,  95.2% aligned
Score =  105 bits (262), Expect = 1e-23

Query:  18   IGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIV  77
Sbjct:  6    MGEITLELYWNHAPKTCKNFYELAK---------RGY--YNGTIFHRLIKDFMIQGGDPT  54

Query:  78   FGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGE--FVEPFTLGMANLG  135
Sbjct:  55   -GT---------GRGGASIYGKK--------------FEDEIHPELKHTGAGILSMANAG  90

Query:  136  SPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRI  191
Sbjct:  91   -PNTNGSQFFITLAPTPWLDGKHTIFGRVSKGMKVIENMVEVQTQTDR-PIDEVKI  144
```

gnl|CDD|28267 cd01925, cyclophilin_CeCYP16-like, cyclophilin_CeCYP16-like: cyclophilin-type peptidylprolyl cis- trans isomerase) (PPIase) domain similar to Caenorhabditis elegans cyclophilin 16. C. elegans CeCYP-16, compared to the archetypal cyclophilin Human cyclophilin A has, a reduced peptidylprolyl cis- trans isomerase activity, is cyclosporin insensitive and shows an altered substrate preference favoring, hydrophobic, acidic or amide amino acids. Most members of this subfamily have a glutamate residue in the active site at the position equivalent to a tryptophan (W121 in Human cyclophilin A), which has been shown to be important for cyclophilin binding.

```
CD-Length = 171 residues,  86.0% aligned
Score = 89.5 bits (222), Expect = 6e-19

Query:  19   GRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIVF  78
Sbjct:  15   GDIDIELWSKEAPKACRNFIQLCL---------EGY--YDNTIFHRVVPGFIIQGGD---  60
```

```
Query:  79   GTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENLG--EFVEPFTLGMANLGS  136
Sbjct:  61   -------PTGTGTGGESIYG--------------EPFKDEFHSRLRFNRRGLVGMANAGD  99

Query:  137  PNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSV--VRTIENCRVDSDGVPESDVRISDC  194
Sbjct:  100  D-SNGSQFFFTLDKADELNNKHTLFGKVT-GDTIYNLLKLAEVETDKDERPVYPPKITSV  157

Query:  195  GVWE  198
Sbjct:  158  EVLE  161
```

gnl|CDD|28263 cd01921, cyclophilin_RRM, cyclophilin_RRM: cyclophilin-type peptidylprolyl cis- trans isomerase domain occuring with a C-terminal RNA recognition motif domain (RRM). This subfamily of the cyclophilin domain family contains a number of eukaryotic cyclophilins having the RRM domain including the nuclear proteins: human hCyP-57, Arabidopsis thaliana AtCYP59, Caenorhabditis elegans CeCyP-44 and Paramecium tetrurelia Kin241. The Kin241 protein has been shown to have a role in cell morphogenesis.

```
CD-Length = 166 residues,  89.8% aligned
Score = 85.3 bits (211), Expect = 1e-17

Query:  18   IGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIV  77
Sbjct:  6    LGDLVIDLFTDECPLACLNFLKLC---------KLKY--YNFCLFYNVQKDFIAQTGDPT  54

Query:  78   FGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGEFV--EPFTLGMANLG  135
Sbjct:  55   GT----------GAGGESIYSQ---LYGRQARF----FEPEILPLLKHSKKGTVSMVNAG  97

Query:  136  SPNTNNSQFFIT-TYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPESDVRIS  192
Sbjct:  98   -DNLNGSQFYITLGENLDYLDGKHTVFGQVVEGFDVLEKINDAIVDDDGRPLKDIRIK  154
```

gnl|CDD|28262 cd01920, cyclophilin_EcCYP_like, cyclophilin_EcCYP_like: cyclophilin-type A-like peptidylprolyl cis- trans isomerase (PPIase) domain similar to the cytosolic E. coli cyclophilin A and Streptomyces antibioticus SanCyp18. Compared to the archetypal cyclophilin Human cyclophilin A, these have reduced affinity for cyclosporin A. E. coli cyclophilin A has a similar peptidylprolyl cis- trans isomerase activity to the human cyclophilin A. Most members of this subfamily contain a phenylalanine residue at the position equivalent to Human cyclophilin W121, where a tyrptophan has been shown to be important for cyclophilin binding.

```
CD-Length = 155 residues,  95.5% aligned
Score = 62.6 bits (152), Expect = 7e-11

Query:  19   GRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIVF  78
Sbjct:  7    GDIVVELYDDKAPITVENFLAYVRKG-----------FYDNTIFHRVISGFVIQGG----  51

Query:  79   GTQKDSSSSSVGKGGCSIYADKEEVKTDD----ESFCYGNFEDENLGEFVEPFTLGMANL  134
Sbjct:  52   ----------------GFTPDLAQKETLKPIKNEA---GNGLSNTRG------TIAMART  86

Query:  135  GSPNTNNSQFFITTYAAPHLN-----GKHSIFGQVVHGKSVVRTIENCRVDSDG----VP  185
Sbjct:  87   NAPDSATSQFFINLKDNASLDYQNEQWGYTVFGEVTEGMDVVDKIAGVETYSFGSYQDVP  146

Query:  186  ESDVRISD  193
Sbjct:  147  VQDVIIES  154
```

gnl|CDD|5390 cd00189, TPR, Tetratricopeptide repeat domain; typically contains 34 amino acids [WLF]-X(2)-[LIM]-[GAS]-X(2)-[YLF]-X(8)-[ASE]-X(3)-[FYL]-X(2)-[ASL]-X(4)-[PKE] is the consensus sequence; found in a variety of organisms including bacteria, cyanobacteria, yeast, fungi, plants, and humans in various subcellular locations; involved in a variety of functions including protein-protein interactions, but common features in the interaction partners have not been defined; involved in chaperone, cell-cycle, transciption, and protein transport complexes; the number of TPR motifs varies among proteins (1,3-11,13 15,16,19); 5-6 tandem repeats generate a right-handed helical structure with an amphipathic channel that is thought to accomodate an alpha-

helix of a target protein; it has been proposed that TPR proteins preferably interact with WD-40 repeat proteins, but in many instances several TPR-proteins seem to aggregate to multi-protein complexes; examples of TPR-proteins include, Cdc16p, Cdc23p and Cdc27p components of the cyclosome/APC, the Pex5p/Pas10p receptor for peroxisomal targeting signals, the Tom70p co-receptor for mitochondrial targeting signals, Ser/Thr phosphatase 5C and the p110 subunit of O-GlcNAc transferase; three copies of the repeat are present here.

CD-Length = 100 residues, 100.0% aligned
Score = 49.7 bits (118), Expect = 6e-07

```
Query:  241  NIIKESGTLLFKKKDYSNAFFKYRKSLNYINEYMPEPDVDKERNIQFINLKMKIYLNLSL  300
Sbjct:  1    EALLNLGNLYYKLGDYDEALEYYEKALELDPDNA------------------DAYYNLAA  42

Query:  301  VLFNLERYDDAIMYATYLLEMDNVPNRDQAKAYYRRGNSYLKKKRLDEALQDYIFCKEKN  360
Sbjct:  43   AYYKLGKYEEALEDYEKALEL----DPDNAKAYYNLGLAYYKLGKYEEALEAYEKALELD  98

Query:  361  PD  362
Sbjct:  99   PN  100
```

gnl|CDD|5390 cd00189, TPR, Tetratricopeptide repeat domain; typically contains 34 amino acids [WLF]-X(2)-[LIM]-[GAS]-X(2)-[YLF]-X(8)-[ASE]-X(3)-[FYL]-X(2)-[ASL]-X(4)-[PKE] is the consensus sequence; found in a variety of organisms including bacteria, cyanobacteria, yeast, fungi, plants, and humans in various subcellular locations; involved in a variety of functions including protein-protein interactions, but common features in the interaction partners have not been defined; involved in chaperone, cell-cycle, transciption, and protein transport complexes; the number of TPR motifs varies among proteins (1,3-11,13 15,16,19); 5-6 tandem repeats generate a right-handed helical structure with an amphipathic channel that is thought to accomodate an alpha-helix of a target protein; it has been proposed that TPR proteins preferably interact with WD-40 repeat proteins, but in many instances several TPR-proteins seem to aggregate to multi-protein complexes; examples of TPR-proteins include, Cdc16p, Cdc23p and Cdc27p components of the cyclosome/APC, the Pex5p/Pas10p receptor for peroxisomal targeting signals, the Tom70p co-receptor for mitochondrial targeting signals, Ser/Thr phosphatase 5C and the p110 subunit of O-GlcNAc transferase; three copies of the repeat are present here.

CD-Length = 100 residues, 86.0% aligned
Score = 42.0 bits (98), Expect = 1e-04

```
Query:  293  KIYLNLSLVLFNLERYDDAIMYATYLLEMDNVPNRDQAKAYYRRGNSYLKKKRLDEALQD  352
Sbjct:  1    EALLNLGNLYYKLGDYDEALEYYEKALELDP----DNADAYYNLAAAYYKLGKYEEALED  56

Query:  353  YIFCKEKNPDDEVIEQRIEYVNRLIEENKE  382
Sbjct:  57   YEKALELDPDNAKAYYNLGLAYYKLGKYEE  86
```

gnl|CDD|28266 cd01924, cyclophilin_TLP40_like, cyclophilin_TLP40_like: cyclophilin-type peptidylprolyl cis- trans isomerases (cyclophilins) similar ot the Spinach thylakoid lumen protein TLP40. Compared to the archetypal cyclophilin Human cyclophilin A, these proteins have similar peptidylprolyl cis- trans isomerase activity and reduced affinity for cyclosporin A. Spinach TLP40 has been shown to have a dual function as a folding catalyst and regulator of dephosphorylation.

CD-Length = 176 residues, 83.0% aligned
Score = 46.9 bits (111), Expect = 4e-06

```
Query:  30   APKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQAGDIVFGTQKDSSSSSV  89
Sbjct:  18   APVTAGNFVDLVERGF-----------YDGMEFHRVEGGFVVQTGD-PQGKNPGFPDPET  65

Query:  90   GKGGCSIYADKEEVKTDDESFCYGNFEDENLGEFVEPF--------TLGMA-NLGSPNTN  140
Sbjct:  66   GK----SRTIPLEIKPEGQKQPVYGKTLEEAGRYDEQPVLPFNAFGAIAMARTEFDPNSA  121
```

```
Query:  141  NSQFFITTYA-------APHLNGKHSIFGQVVHGKSVVRTIE  175
Sbjct:  122  SSQFFFLLKDNELTPSRNNVLDGRYAVFGYVTDGLDILRELK  163
```

gnl|CDD|24194 smart00028, TPR, Tetratricopeptide repeats; Repeats present in 4 or more copies in proteins. Contain a minimum of 34 amino acids each and self-associate via a "knobs and holes" mechanism. .

```
CD-Length = 34 residues, 100.0% aligned
Score = 38.2 bits (90), Expect = 0.002

Query:  330  AKAYYRRGNSYLKKKRLDEALQDYIFCKEKNPDD  363
Sbjct:  1    AEALYNLGNAYLKLGDYDEALEYYEKALELDPNN  34
```

gnl|CDD|25421 pfam00160, Pro_isomerase, Cyclophilin type peptidyl-prolyl cis-trans isomerase.

```
CD-Length = 161 residues,  98.8% aligned
Score =  175 bits (445), Expect = 8e-45

Query:  7    VYLDISIDKKPIGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVV  66
Sbjct:  2    VYFDIEIGGKPAGRIVFELFDDVVPKTVENFRALCTGEKG--------FGYKGSTFHRVI  53

Query:  67   KNFMIQAGDIVFGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENL-GEFVE  125
Sbjct:  54   PGFMIQGGDFTRGN---------GTGGKSIYGEK--------------FKDENFNLKHDR  90

Query:  126  PFTLGMANLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVP  185
Sbjct:  91   PGTLSMANAG-PNTNGSQFFITTVPTPWLDGKHVVFGKVVEGMDVVKKIESVGTDSGGKP  149

Query:  186  ESDVRISDCGV  196
Sbjct:  150  SKPVKIADCGE  160
```

gnl|CDD|16751 pfam00515, TPR, TPR Domain.

```
CD-Length = 34 residues, 100.0% aligned
Score = 39.6 bits (93), Expect = 6e-04

Query:  330  AKAYYRRGNSYLKKKRLDEALQDYIFCKEKNPDD  363
Sbjct:  1    AEAYYNLGNAYLKLGKYDEAIEYYEKALELNPNN  34
```

gnl|CDD|10522 COG0652, PpiB, Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family [Posttranslational modification, protein turnover, chaperones].

```
CD-Length = 158 residues,  94.3% aligned
Score =  150 bits (379), Expect = 4e-37

Query:  13   IDKKPIGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVKNFMIQ  72
Sbjct:  3    ILETNKGDITIELYPDKAPKTVANFLQLV----------KEGFYDGTIFHRVIPGFMIQ  51

Query:  73   AGDIVFGTQKDSSSSSVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGEFVEPFTLGMA  132
Sbjct:  52   GGDPTGGD---------GTGGPGP-------PFKDENFALNGDRHK-------RGTLSMA  88

Query:  133  NLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDG----VPESD  188
Sbjct:  89   RAGDPNSNGSQFFITVVDNPFLDGKYTVFGQVVEGMDVVDKIKNGDTDDSGYVQDVPADP  148

Query:  189  VRI  191
Sbjct:  149  VKI  151
```

Plant homologs

```
gi|37788308|gb|AAP44535.1|  cyclophilin-like protein [Triticum aestivum]

Identities = 84/194 (43%), Positives = 103/194 (53%), Gaps = 38/194 (19%)

Query: 7   VYLDISIDK---KPIGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFH 63
           VY DISI     K +GRIV  L+ +  P+T ENF  LCAG        ++   YKG+ FH
Sbjct: 80  VYFDISIGNPVGKNVGRIVIGLYGDDVPQTVENFRALCAG--------EKGFGYKGSSFH 131

Query: 64  RVVKNFMIQAGDIVFGTQKDXXXXXVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGEF 123
           RV+K+FMIQ GD   G          G GG SIY                F+DEN  +
Sbjct: 132 RVIKDFMIQGGDFDKGN---------GTGGKSIYGR--------------TFKDENF-QL 167

Query: 124 VE--PFTLGMANLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDS 181
           V   P  L MAN G PNTN SQFFI T   P L+G+H +FGQV+ G  +VRTIE+   D
Sbjct: 168 VHTGPGVLSMANAG-PNTNGSQFFICTMKTPWLDGRHVVFGQVLEGMDIVRTIESSETDR 226

Query: 182 DGVPESDVRISDCG 195
              P+  V IS+CG
Sbjct: 227 GDRPKKKVVISECG 240


gi|13442983|gb|AAK02067.1|  cyclophilin-40 [Arabidopsis thaliana]
gi|20197802|gb|AAD41985.2|   expressed protein [Arabidopsis thaliana]
gi|18397936|ref|NP_565381.1|  peptidyl-prolyl  cis-trans  isomerase  /  cyclophilin-40
(CYP40) /rotamase [Arabidopsis thaliana]

Identities = 128/385 (33%), Positives = 186/385 (48%), Gaps = 31/385 (8%)

Query: 8   YLDISIDKKPIGRIVCKLFREKAPKTTENFYKLCAGDVKSPLKDQQYLSYKGNGFHRVVK 67
           ++DISI  +  GRIV +L+ +  PKT ENF  LC G+          L YKGN FHRV+K
Sbjct: 7   FMDISIGGELEGRIVIELYDDVVPKTAENFRLLCTGEKGLGPNTGVPLHYKGNRFHRVIK 66

Query: 68  NFMIQAGDIVFGTQKDXXXXXVGKGGCSIYADKEEVKTDDESFCYGNFEDENLGEFVEPF 127
            FMIQ GDI   +  D      G GG SIY     +K DDE+F     + E  G
Sbjct: 67  GFMIQGGDI---SAND------GTGGESIYG----LKFDDENF---ELKHERKG------ 104

Query: 128 TLGMANLGSPNTNNSQFFITTYAAPHLNGKHSIFGQVVHGKSVVRTIENCRVDSDGVPES 187
            L MAN G PNTN SQFFITT    HL+GKH +FG+V  G  VVR+IE+  ++    P
Sbjct: 105 MLSMANSG-PNTNGSQFFITTTRTSHLDGKHVVFGRVTKGMGVVRSIEHVSIEEQSCPSQ 163

Query: 188 DVRISDCGVWEKTMGVPLYNASNDQIGGDVYEEYPXXXXXXXXXXXXKALEAANIIKESG 247
           DV I DCG   +     + +   D   GDVY ++P              +E  + +K  G
Sbjct: 164 DVVIHDCGEIPEGADDGICDFFKD---GDVYPDWPIDLNESPAELSW-WMETVDFVKAHG 219

Query: 248 TLLFKKKDYSNAFFKYRKSLNYINEYMPEPDVDKERNIQFINLKMKIYLNLSLVLFNLER 307
              FKK+DY  A  KYRK+L Y++    +  +D+E +      K +I+ N +
Sbjct: 220 NEHFKKQDYKMALRKYRKALRYLDICWEKEGIDEETSTALRKTKSQIFTNSAACKLKFGD 279

Query: 308 YDDAIMYATYLLEMDNVPNRDQAKAYYRRGNSYLKKKRLDEALQDYIFCKEKNPDDEVIE 367
              A++   + +  ++    +  KA +R+G +Y+    +D A +      +  P+D  I+
Sbjct: 280 AKGALLDTEFAMRDED----NNVKALFRQGQAYMALNNVDAAAESLEKALQFEPNDAGIK 335

Query: 368 QRIEYVNRLIEENKEKTRKNISKFF 392
           +    V + I     + +K   K F
Sbjct: 336 KEYAAVMKKIAFRDNEEKKQYRKMF 360
```

4K1-G1 protein sequence alignment

gi|663250|emb|CAA88158.1| **TyB** protein - yeast transposon Ty1-H3 [Saccharomyces cerevisiae]

Identities = 161/161 (100%), Positives = 161/161 (100%), Gaps = 0/161 (0%)

```
Query  483  PIRTTLRYDEAITYNKDIKEKEKYIEAYHKEVNQLLKMKTWDTDEYYDRKEIDPKRVINS  304
            PIRTTLRYDEAITYNKDIKEKEKYIEAYHKEVNQLLKMKTWDTDEYYDRKEIDPKRVINS
Sbjct  627  PIRTTLRYDEAITYNKDIKEKEKYIEAYHKEVNQLLKMKTWDTDEYYDRKEIDPKRVINS  686

Query  303  MFIFNKKRDGTHKARFVARGDIQHPDTYDSGMQSNTVHHYALMTSLSLALDNNYYITQLD  124
            MFIFNKKRDGTHKARFVARGDIQHPDTYDSGMQSNTVHHYALMTSLSLALDNNYYITQLD
Sbjct  687  MFIFNKKRDGTHKARFVARGDIQHPDTYDSGMQSNTVHHYALMTSLSLALDNNYYITQLD  746

Query  123  ISSAYLYADIKEELYIRPPPHLGMNDKLIRLKKSLYGLKQS  1
            ISSAYLYADIKEELYIRPPPHLGMNDKLIRLKKSLYGLKQS
Sbjct  747  ISSAYLYADIKEELYIRPPPHLGMNDKLIRLKKSLYGLKQS  787
```

1K2-D5 protein sequence alignment

gi|3876417|emb|CAB03002.1| Hypothetical protein F26E4.9 [Caenorhabditis elegans]
gi|17506881|ref|NP_492601.1| cytochrome C Oxidase, subunit Vb, LiPid Depleted **LPD-4** (lpd-4) [Caenorhabditis elegans]
gi|7499931|pir||T21416 hypothetical protein F26E4.9 - Caenorhabditis elegans

Identities = 23/77 (29%), Positives = 39/77 (50%), Gaps = 5/77 (6%)

```
Query  392  SSETILKHYSYYAIGLPHSTPR*KKYLDASV*TGNNERFHTRLFLRCYMSLRSR--LAPS  219
            ++E   + Y YY   L H+T R KK L A +    ++R+  +++ R   S + +  L PS
Sbjct  28   ATEASPEDYGYYPDPLEHATGREKKMLLARL--AGDDRYEPKVYYRAEASTKQKPNLVPS  85

Query  218  RMQYRNTG-ICTQRSSN  171
               +R  G +C Q S +
Sbjct  86   HYDFRIIGCMCEQDSGH  102
```

1K2-F1 protein sequence alignment

gi|6322962|ref|NP_013034.1| hypothetical protein; **Yll066cp** [Saccharomyces cerevisiae]
gi|1360282|emb|CAA97520.1| unnamed protein product [Saccharomyces cerevisiae]
gi|642317|emb|CAA87990.1| ORF L0519 [Saccharomyces cerevisiae]

Identities = 53/53 (100%), Positives = 53/53 (100%), Gaps = 0/53 (0%)

```
Query  1     DLANVEVLAADNTRVPLYMLMVAVHKELDSDDVPDGRFDIILLCRDSSREVGE  159
            DLANVEVLAADNTRVPLYMLMVAVHKELDSDDVPDGRFDIILLCRDSSREVGE
Sbjct  1153  DLANVEVLAADNTRVPLYMLMVAVHKELDSDDVPDGRFDIILLCRDSSREVGE  1205
```

gi|6323500|ref|NP_013572.1| Helicase encoded by the Y' element of subtelomeric regions, highly expressed in the mutants lacking the telomerase component TLC1; potentially phosphorylated by Cdc28p; Yrf1-5p [Saccharomyces cerevisiae]
gi|6320754|ref|NP_010834.1| Helicase encoded by the Y' element of subtelomeric regions, highly expressed in the mutants lacking the telomerase component TLC1; potentially phosphorylated by Cdc28p; Yrf1-1p [Saccharomyces cerevisiae]
gi|1420848|emb|CAA99727.1| unnamed protein product [Saccharomyces cerevisiae]
gi|1431574|emb|CAA98916.1| unnamed protein product [Saccharomyces cerevisiae]
gi|1360590|emb|CAA97727.1| unnamed protein product [Saccharomyces cerevisiae]
gi|17380338|sp|P24088|YRF11_YEAST Y'helicase protein 1 copies 1/5/8

Identities = 53/53 (100%), Positives = 53/53 (100%), Gaps = 0/53 (0%)

```
Query  1     DLANVEVLAADNTRVPLYMLMVAVHKELDSDDVPDGRFDIILLCRDSSREVGE  159
             DLANVEVLAADNTRVPLYMLMVAVHKELDSDDVPDGRFDIILLCRDSSREVGE
Sbjct  1744  DLANVEVLAADNTRVPLYMLMVAVHKELDSDDVPDGRFDIILLCRDSSREVGE  1796
```

1K2-F5 protein sequence alignment

gi|6319695|ref|NP_009777.1| Pyruvate carboxylase isoform, cytoplasmic enzyme that converts pyruvate to oxaloacetate; highly similar to isoform Pyc1p but differentially regulated; mutations in the human homolog are associated with lactic acidosis; Pyc2p [Saccharomyces cerevisiae]
gi|536608|emb|CAA85182.1| **PYC2** [Saccharomyces cerevisiae]
gi|585765|sp|P32327|PYC2_YEAST Pyruvate carboxylase 2 (Pyruvic carboxylase 2) (PCB 2)
gi|1041735|gb|AAC49147.1| pyruvate carboxylase 2

Identities = 183/183 (100%), Positives = 183/183 (100%), Gaps = 0/183 (0%)

```
Query  549  FAKQVRQFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATF  370
            FAKQVRQFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATF
Sbjct  548  FAKQVRQFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATF  607

Query  369  DVAMRFLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVD  190
            DVAMRFLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVD
Sbjct  608  DVAMRFLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVD  667

Query  189  IFRVFDALNDLEQLKVGVNAVKKAGGVVEATVCYSGDMLQPGKKYNLDYYLEVVEKIVQM  10
            IFRVFDALNDLEQLKVGVNAVKKAGGVVEATVCYSGDMLQPGKKYNLDYYLEVVEKIVQM
Sbjct  668  IFRVFDALNDLEQLKVGVNAVKKAGGVVEATVCYSGDMLQPGKKYNLDYYLEVVEKIVQM  727

Query  9    GTH  1
            GTH
Sbjct  728  GTH  730
```

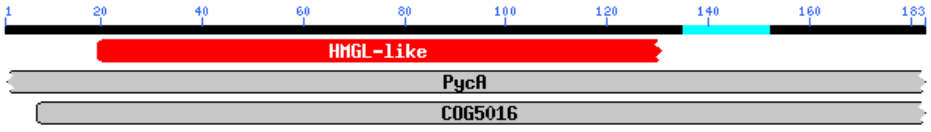


Figure E.5 PYC2 structure and conserved domains

gnl|CDD|25586 pfam00682, HMGL-like, HMGL-like. This family contains a diverse set of enzymes. These include various aldolases and a region of pyruvate carboxylase. .

CD-Length = 260 residues, only 35.8% aligned
Score = 57.2 bits (138), Expect = 1e-09

```
Query:  19   RDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATFDVAMRFLHEDPWERLRKL  78
Sbjct:  1    RDGEQALGV-----------AFSLEEKLAIARALDEAGVDEIEVGFPFMSPTDFESVRAI  49

Query:  79   RSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRVFDALNDL  131
Sbjct:  50   AELGPK---------AKILALCRPRDDDIDAAVEAAVGAGVDRVHVFIATSDL  93
```

gnl|CDD|10765 COG1038, PycA, Pyruvate carboxylase [Energy production and conversion].

CD-Length = 1149 residues, only 16.0% aligned
Score = 307 bits (787), Expect = 6e-85

```
Query:  1     FAKQVRQFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATF  60
Sbjct:  526   FARWVREQKAVLLTDTTFRDAHQSLLATRVRTHDLARIAPATARALPQLFSLEMWGGATF  585

Query:  61    DVAMRFLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVD  120
Sbjct:  586   DVAMRFLKEDPWERLERLRKAVPNVLFQMLLRGANGVGYKNYPDNVIREFVKQAAKSGID  645

Query:  121   IFRVFDALNDLEQLKVGVNAVKKAGGVVEATVCYSGDMLQPGK-KYNLDYYLEVVEKIVQ  179
Sbjct:  646   VFRIFDSLNWVEQMRVAIDAVREAGKVAEATICYTGDILDPGRKKYTLDYYVKLAKELEK  705

Query:  180   MGTH  183
Sbjct:  706   AGAH  709
```

gnl|CDD|14146 COG5016, COG5016, Pyruvate/oxaloacetate carboxyltransferase [Energy production and conversion].

CD-Length = 472 residues, only 36.0% aligned
Score = 180 bits (457), Expect = 1e-46

```
Query:  7    QFNGTLLMDTTWRDAHQSLLATRVRTHDLATIAPTTAHAL--AGAFALECWGGATFDVAM  64
Sbjct:  2    AMKKIKITDTVLRDGHQSLLATRMRTEDMLPIA----EALDKVGYWSLEVWGGATFDACI  57

Query:  65   RFLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSLPDNAIDHFVKQAKDNGVDIFRV  124
Sbjct:  58   RFLNEDPWERLRELKKAVPNTKLQMLLRGQNLVGYRHYADDVVEKFVEKAAENGIDVFRI  117


Query:  125  FDALNDLEQLKVGVNAVKKAGGVVEATVCYSGDMLQPGKKYNLDYYLEVVEKIVQMGTH  183
Sbjct:  118  FDALNDVRNLKTAIKAAKKHGAHVQGTISYT---TSP--VHTLEYYVELAKELLEMGVD  171
```

4L2-D1 protein sequence alignment

gi|6322299|ref|NP_012373.1| Protein of unknown function, contains a **J-domain**, which is a region with homology to the E. coli DnaJ protein; Jjj2p [Saccharomyces cerevisiae]
gi|1008348|emb|CAA89457.1| unnamed protein product [Saccharomyces cerevisiae]
gi|1353029|sp|P46997|JJJ2_YEAST J protein type 2

```
Identities = 189/189 (100%), Positives = 189/189 (100%), Gaps = 0/189 (0%)

Query  568  ARLLHPDKTKSDKSEELFKAVVHAHSILTDEDQKLRYDRDLKIKGLHTYQPKKNCHIFKT  389
            ARLLHPDKTKSDKSEELFKAVVHAHSILTDEDQKLRYDRDLKIKGLHTYQPKKNCHIFKT
Sbjct  37   ARLLHPDKTKSDKSEELFKAVVHAHSILTDEDQKLRYDRDLKIKGLHTYQPKKNCHIFKT  96

Query  388  KAKESQGASPTLGQSEAYHRQNKPYEQQPYGFGVGKKMTSSSKSKVPIFKSFNLKSYQRN  209
            KAKESQGASPTLGQSEAYHRQNKPYEQQPYGFGVGKKMTSSSKSKVPIFKSFNLKSYQRN
Sbjct  97   KAKESQGASPTLGQSEAYHRQNKPYEQQPYGFGVGKKMTSSSKSKVPIFKSFNLKSYQRN  156

Query  208  HYYSSKKERKHGSPDIDSLFHETNGASKVRMTDAGKMDTNSQFQEIWEILGKNAYTHKSY  29
            HYYSSKKERKHGSPDIDSLFHETNGASKVRMTDAGKMDTNSQFQEIWEILGKNAYTHKSY
Sbjct  157  HYYSSKKERKHGSPDIDSLFHETNGASKVRMTDAGKMDTNSQFQEIWEILGKNAYTHKSY  216

Query  28   SEDPNSCLG  2
            SEDPNSCLG
Sbjct  217  SEDPNSCLG  225
```

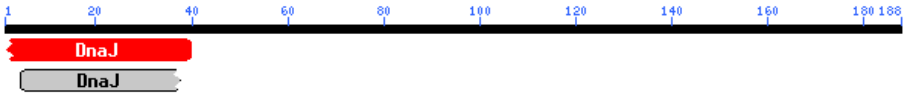


Figure E.6 J protein type 2 structure and conserved domains

gnl|CDD|24391 pfam00226, DnaJ, DnaJ domain. DnaJ domains (J-domains) are associated with hsp70 heat-shock system and it is thought that this domain mediates the interaction. DnaJ-domain is therefore part of a chaperone (protein folding) system. The T-antigens are confirmed as DnaJ containing domains from literature.

```
CD-Length = 66 residues, only  62.1% aligned
Score = 37.5 bits (87), Expect = 0.001

Query:  1    RLLHPDKTKSDK-SEELFKAVVHAHSILTDEDQKLRYDRDL  40
Sbjct:  26   LKYHPDKNPGDPEAEEKFKEINEAYEVLSDPEKRAIYDQYG  66
```

gnl|CDD|10357 COG0484, DnaJ, DnaJ-class molecular chaperone with C-terminal Zn finger domain [Posttranslational modification, protein turnover, chaperones].

```
CD-Length = 371 residues, only   9.7% aligned
Score = 38.3 bits (89), Expect = 5e-04

Query:  4    HPDKTKSDK-SEELFKAVVHAHSILTDEDQKLRYDR  38
Sbjct:  33   HPDRNPGDKEAEEKFKEINEAYEVLSDPEKRAAYDQ  68
```

APPENDIX F

PROTEIN SEQUENCE ALIGNMENTS AND CONSERVED DOMAINS OF YR10-2 INTERACTORS

2K1-A1 protein sequence alignment

gi|253435|gb|AAB22844.1| **Vps35p** [Saccharomyces cerevisiae]

Identities = 303/303 (100%), Positives = 303/303 (100%)

```
Query: 1   QEKLAHLIHWIMNTTSRKQTMKNKIQFSLEAQLEILLLIKSSFIKGGINVKYTFPAIITN 60
           QEKLAHLIHWIMNTTSRKQTMKNKIQFSLEAQLEILLLIKSSFIKGGINVKYTFPAIITN
Sbjct: 524 QEKLAHLIHWIMNTTSRKQTMKNKIQFSLEAQLEILLLIKSSFIKGGINVKYTFPAIITN 583

Query: 61  FWKLMRKCRMIQEYLLKKRPDNKTLLSHYSNLLKQMFKFVSRCINDIFNSCNNSCTDLIL 120
           FWKLMRKCRMIQEYLLKKRPDNKTLLSHYSNLLKQMFKFVSRCINDIFNSCNNSCTDLIL
Sbjct: 584 FWKLMRKCRMIQEYLLKKRPDNKTLLSHYSNLLKQMFKFVSRCINDIFNSCNNSCTDLIL 643

Query: 121 KLNLQCAILAEQLQLNEISYDFFSQAFTIFEESLSDSKTQLQALIYIAQSLQKTRSLYKE 180
           KLNLQCAILAEQLQLNEISYDFFSQAFTIFEESLSDSKTQLQALIYIAQSLQKTRSLYKE
Sbjct: 644 KLNLQCAILAEQLQLNEISYDFFSQAFTIFEESLSDSKTQLQALIYIAQSLQKTRSLYKE 703

Query: 181 AYYDSLIVRCTLHGSKLLKKQDQCRAVYLCSHLWWATEISNIGEEEGITDNFYRDGKRVL 240
           AYYDSLIVRCTLHGSKLLKKQDQCRAVYLCSHLWWATEISNIGEEEGITDNFYRDGKRVL
Sbjct: 704 AYYDSLIVRCTLHGSKLLKKQDQCRAVYLCSHLWWATEISNIGEEEGITDNFYRDGKRVL 763

Query: 241 ECLQRSLRVADSIMDNEQSCELMVEILNRCLYYFIHGDESETHISIKYINGLIELIKTNL 300
           ECLQRSLRVADSIMDNEQSCELMVEILNRCLYYFIHGDESETHISIKYINGLIELIKTNL
Sbjct: 764 ECLQRSLRVADSIMDNEQSCELMVEILNRCLYYFIHGDESETHISIKYINGLIELIKTNL 823

Query: 301 KSL 303
           KSL
Sbjct: 824 KSL 826
```

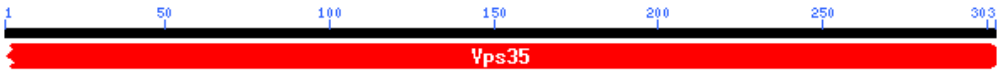


Figure F.1 Vps35 structure and conserved domains

gnl|CDD|26119 pfam03635, Vps35, Vacuolar protein sorting-associated protein 35. Vacuolar protein sorting-associated protein (Vps) 35 is one of around 50 proteins involved in protein trafficking. In particular, Vps35 assembles into a retromer complex with at least four other proteins Vps5, Vps17, Vps26 and Vps29. Vps35

contains a central region of weaker sequence similarity, thought to indicate the presence of at least three domains.

CD-Length = 729 residues, only 36.9% aligned
Score = 348 bits (893), Expect = 6e-97

```
Query:  1    QEKLAHLIHWIMNTTSRKQTMKNKIQFSLEAQLEILLLIKSSFIKGGIN-VKYTFPAIIT  59
Sbjct:  455  QNLVARLLHLLRNDDP-------------EVQAEILNSARKTFIKGGARRVPYTLPPLVF  501

Query:  60   NFWKLMRKCRMIQEYLLKKRPDNKTLLSHYSNLLKQMFKFVSRCINDIFNSCNNSCTDLI  119
Sbjct:  502  AALRLARQYRRL---------DNKRDGDKADAKAKQMFKFVSKTINVLLK--NVEAPELA  550

Query:  120  LKLNLQCAILAEQLQLNEISYDFFSQAFTIFEESLSDSKTQLQALIYIAQSLQKTRSLYK  179
Sbjct:  551  LKLYLQCAITADQCELETVAYEFFSQAFSLYEESVSDSKAQVQALVLIIGTLQKMRS-FP  609

Query:  180  EAYYDSLIVRCTLHGSKLLKKQDQCRAVYLCSHLWWATEISNIGEEEGITDNFYRDGKRV  239
Sbjct:  610  EENYDTLATKCTLYASKLLKKPDQCRAVYLCAHLFWAGQLSNINE-------PLKDGKRV  662

Query:  240  LECLQRSLRVADSIMDNEQSCELMVEILNRCLYYFIHGDESETHISIKYINGLIELIKTN  299
Sbjct:  663  LECLRRSLKIADQCMDPLQSLQLFVELLNRYLYFY---EKSCYLVSVKHINELIELIKEN  719

Query:  300  LKSL  303
Sbjct:  720  LKSL  723
```

3L2-F1 protein sequence alignment

gi|121064|sp|P15442|GCN2_YEAST Serine/threonine-protein kinase **GCN2**
gi|171578|gb|AAA34636.1| GCN2

Identities = 181/181 (100%), Positives = 181/181 (100%)

```
Query: 3    SRMLSQVGFARSFKEVKNELKAQLNISSTALNDLELFDFRLDFEAAKKRLYKLMIDSPHL 182
            SRMLSQVGFARSFKEVKNELKAQLNISSTALNDLELFDFRLDFEAAKKRLYKLMIDSPHL
Sbjct: 1134 SRMLSQVGFARSFKEVKNELKAQLNISSTALNDLELFDFRLDFEAAKKRLYKLMIDSPHL 1193
```

```
Query: 183  KKIEDSLSHISKVLSYLKPLEVARNVVISPLSNYNSAFYKGGIMFHAVYDDGSSRNMIAA 362
            KKIEDSLSHISKVLSYLKPLEVARNVVISPLSNYNSAFYKGGIMFHAVYDDGSSRNMIAA
Sbjct: 1194 KKIEDSLSHISKVLSYLKPLEVARNVVISPLSNYNSAFYKGGIMFHAVYDDGSSRNMIAA 1253

Query: 363  GGRYDTLISFFARPSGKKSSNTRKAVGFNLAWETIFGIAQNYFKLASGNRIKKRNRFLKD 542
            GGRYDTLISFFARPSGKKSSNTRKAVGFNLAWETIFGIAQNYFKLASGNRIKKRNRFLKD
Sbjct: 1254 GGRYDTLISFFARPSGKKSSNTRKAVGFNLAWETIFGIAQNYFKLASGNRIKKRNRFLKD 1313

Query: 543  T 545
            T
Sbjct: 1314 T 1314
```

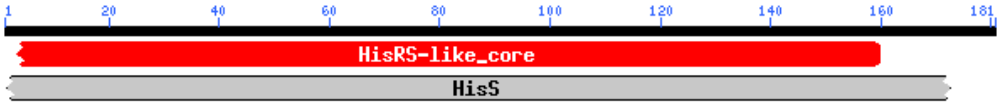


Figure F.2 GCN2 structure and conserved domains

gnl|CDD|27671 cd00773, HisRS-like_core, Class II Histidinyl-tRNA synthetase (HisRS)-like catalytic core domain. HisRS is a homodimer. It is responsible for the attachment of histidine to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs. This domain is also found at the C-terminus of eukaryotic GCN2 protein kinase and at the N-terminus of the ATP phosphoribosyltransferase accessory subunit, HisZ. HisZ along with HisG catalyze the first reaction in histidine biosynthesis. HisZ is found only in a subset of bacteria and differs from HisRS in lacking a C-terminal anti-codon binding domain..

```
CD-Length = 309 residues, only  45.3% aligned
Score = 91.1 bits (226), Expect = 7e-20

Query:  3    MLSQVGFARSFKEVKNELKAQLNISSTALNDLELFDFRLDFEAAKKRLYKLMIDSPHLKK  62
Sbjct:  170  KLDKKGLELVEDELARLGK----------NPLRILDLLLSSKGAGEEELLLAAPLILDKL  219

Query:  63   IEDSLSHISKVLSYLKPLEVARNVVISPLSNYNSAFYKGGIMFHAVYDDGSSRNMIAAGG  122
Sbjct:  220  AEEALAHLEKLLDYLEALGVDIKYSI-DLSLVRGLDYYTGIVFEAVADGLGAQGSIAGGG  278

Query:  123  RYDTLISFFARPSGKKSSNTRKAVGFNLAWETIFGIAQ  160
Sbjct:  279  RYDGLLEEF-------GGEDVPAVGFAIGLERLLLALE  309
```

gnl|CDD|9999 COG0124, HisS, Histidyl-tRNA synthetase [Translation, ribosomal structure and biogenesis].

```
CD-Length = 429 residues, only  35.7% aligned
Score = 83.4 bits (206), Expect = 1e-17
```

```
Query:  1    SRMLSQVGFARSFKEVKNELKAQLNISSTALNDLELFDFRLDFEAAKKRLYKLMIDSPHL  60
Sbjct:  185  LRYLDKLDKI-GKLELDEDSKRRLKT-----NPLRVLDSKKDSDQE------LLKNAPEL  232

Query:  61   KKIED--SLSHISKVLSYLKPLEVARNVVISPLSNYNSAFYKGGIMFHAVYDDGSSRNMI  118
Sbjct:  233  LDYLDEESLEHLEELLALLDALGI--SYEIDPSLVRGLDYYTG-TVFEAVTDGLGAQGSV  289

Query:  119  AAGGRYDTLISFFARPSGKKSSNTRKAVGFNLAWETIFGIAQNYFKLASGNRIKK  173
Sbjct:  290  CGGGRYDGLVEEFGGKP-------TPAVGFAIGVERLILALEEEGKEDPVETRVD  337
```

3L2-G5 protein sequence alignment

gi|218453|dbj|BAA02936.1| NADPH-cytochrome P450 reductase precursor [Saccharomyces cerevisiae]
gi|361356|prf||1408205A NADPH cytochrome P450 reductase

```
Identities = 183/183 (100%), Positives = 183/183 (100%)


Query: 1   ADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAH 60
           ADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAH
Sbjct: 185 ADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAH 244

Query: 61  YLPSHQLNRNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTG 120
           YLPSHQLNRNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTG
Sbjct: 245 YLPSHQLNRNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTG 304

Query: 121 DHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITG 180
           DHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITG
Sbjct: 305 DHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITG 364

Query: 181 PVS 183
           PVS
Sbjct: 365 PVS 367
```

gi|51013633|gb|AAT93110.1| YHR042W [Saccharomyces cerevisiae]
gi|6321832|ref|NP_011908.1| NADP-cytochrome P450 reductase; involved in ergosterol biosynthesis; associated and coordinately regulated with Erg11p [Saccharomyces cerevisiae]
gi|630034|pir||S46735 NADPH-ferrihemoprotein reductase (EC 1.6.2.4) - yeast (Saccharomyces cerevisiae)
gi|488169|gb|AAB68904.1| Ncp1p: NADP-cytochrome P450 reductase [Saccharomyces cerevisiae]
gi|730126|sp|P16603|NCPR_YEAST NADPH--cytochrome P450 reductase (CPR) (P450R)

```
Identities = 183/183 (100%), Positives = 183/183 (100%)

Query: 1   ADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAH 60
           ADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAH
Sbjct: 185 ADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAH 244

Query: 61  YLPSHQLNRNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTG 120
           YLPSHQLNRNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTG
Sbjct: 245 YLPSHQLNRNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTG 304

Query: 121 DHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITG 180
           DHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITG
Sbjct: 305 DHLAVWPSNPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITG 364
```

```
Query: 181 PVS 183
           PVS
Sbjct: 365 PVS 367
```

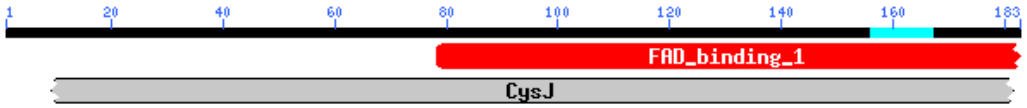


Figure F.3 NCP1 structure and conserved domains

gnl|CDD|4294 pfam00667, FAD_binding_1, FAD binding domain. This domain is found in sulfite reductase, NADPH cytochrome P450 reductase, Nitric oxide synthase and methionine synthase reductase.

```
CD-Length = 217 residues, only  49.8% aligned
Score =  120 bits (303), Expect = 8e-29

Query:  78   PFDLSQPYIAPIVKSRELFS-SNDRNCIHSEFDLSGSNIKYSTGDHLAVWPSNPLEKVEQ  136
Sbjct:  1    PFDAKNPFLASVVVNRELTKPGSDRSCIHLELDISGSGLRYQTGDHLGVYPPNNEELVEE  60

Query:  137  FLSIFNLDPETIFDLKPLDPT-VKVPFPTPTTIGAAIKHYLEITGPVS  183
Sbjct:  61   LLELLGLDPDTVVSLEALDEDWVKPPFLPPTTYRQALTYYLDITGPPT  108
```

gnl|CDD|10243 COG0369, CysJ, Sulfite reductase, alpha subunit (flavoprotein) [Inorganic ion transport and metabolism].

```
CD-Length = 587 residues, only  22.7% aligned
Score = 86.2 bits (213), Expect = 2e-18

Query:  9    DEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAHYLPSHQLN  68
Sbjct:  178  EAAAAPWRDDVLELLKSKFPGQEAAPAQVATS--------PQSESPYSKPAPSVAILLEN  229

Query:  69   RNADGIQLGPFDLSQPYIAPIVKSRELFSSNDRNCIHSEFDLSGSNIKYSTGDHLAVWPS  128
Sbjct:  230  RKLTGR----------------------DSDKDVRHIELDLPDSGLRYEPGDALGVWPE  266

Query:  129  NPLEKVEQFLSIFNLDPETIFDLKPLDPTVKVPFPTPTTIGAAIKHYLEITGPV  182
Sbjct:  267  NDPELVDEFLELLGLDPEEPVTVDG----------ETLPLVEALKSHFEFTSAP  310
```

Plant homologs

gi|10442765|gb|AAG17471.1| NADPH-cytochrome P450 reductase [Triticum aestivum]

```
Identities = 43/144 (29%), Positives = 74/144 (51%), Gaps = 7/144 (4%)
```

```
Query: 9   DEDYMAWKDSILEVLKDELHLDEQEAKFTSQFQYTVLNEITDSMSLGEPSAHYLPSHQLN 68
           ++D+ AWK+ +   L D+L  DE +    +   YT        + + +    +
Sbjct: 185 EDDFTAWKELVWPEL-DQLLRDEDDTTGATT-PYTAAIPEYRVVFIDKSDLEFEDKSWTL 242

Query: 69  RNADGIQLGPFDLSQPYIAPIVKSRELFS-SNDRNCIHSEFDLSGSNIKYSTGDHLAVWP 127
            N +G+     D   P  A +   +EL   ++DR+CIH EFD+SG+ + Y TGDH+ V+
Sbjct: 243 ANGNGV----IDAQHPCRANVAVRKELHKPASDRSCIHLEFDISGTGLVYETGDHVGVYS 298

Query: 128 SNPLEKVEQFLSIFNLDPETIFDL 151
            N +E VEQ  ++ +L P+T+F +
Sbjct: 299 ENSIETVEQAETLLDLSPDTVFSI 322
```

2KL2-D6 protein sequence alignment

gi|1323107|emb|CAA97079.1| unnamed protein product [Saccharomyces cerevisiae]
gi|1723681|sp|P53248|PEX8_YEAST **Peroxisomal biogenesis factor 8** (Peroxin-8) (Peroxisomal protein
PAS6)
gi|6321514|ref|NP_011591.1| Intraperoxisomal organizer of the peroxisomal import machinery,tightly associated with the lumenal face of the peroxisomal membrane, essential for peroxisome biogenesis, binds PTS1-signal receptor Pex5p [Saccharomyces cerevisiae]

Identities = 191/191 (100%), Positives = 191/191 (100%)

```
Query: 1   ALLFNQSNRSALLHGVDVSWNLVTEKLLDLLEEYVHGIVQPMEIFSTDSVLSTNLNHLAS 180
           ALLFNQSNRSALLHGVDVSWNLVTEKLLDLLEEYVHGIVQPMEIFSTDSVLSTNLNHLAS
Sbjct: 180 ALLFNQSNRSALLHGVDVSWNLVTEKLLDLLEEYVHGIVQPMEIFSTDSVLSTNLNHLAS 239

Query: 181 CLTSSITRSNEATLVNSVRKLERICRYLSDTVASLKEQQLDFKFQNVFILIILALKELSA 360
           CLTSSITRSNEATLVNSVRKLERICRYLSDTVASLKEQQLDFKFQNVFILIILALKELSA
Sbjct: 240 CLTSSITRSNEATLVNSVRKLERICRYLSDTVASLKEQQLDFKFQNVFILIILALKELSA 299

Query: 361 MNMTILPNHKDTFYSMICLSLFHVHVLTQKIGTVGFPSYDYVYDNLVTYFIVMDDLSKIT 540
           MNMTILPNHKDTFYSMICLSLFHVHVLTQKIGTVGFPSYDYVYDNLVTYFIVMDDLSKIT
Sbjct: 300 MNMTILPNHKDTFYSMICLSLFHVHVLTQKIGTVGFPSYDYVYDNLVTYFIVMDDLSKIT 359

Query: 541 TVLELMKRNNT 573
           TVLELMKRNNT
Sbjct: 360 TVLELMKRNNT 370
```

2KL2-F3 protein sequence alignment

gi|1301923|emb|CAA95926.1| **POR1** [Saccharomyces cerevisiae]
gi|6324273|ref|NP_014343.1| Mitochondrial porin (voltage-dependent anion channel), outer membrane protein required for the maintenance of mitochondrial osmotic stability and mitochondrial membrane permeability [Saccharomyces cerevisiae]
gi|994830|gb|AAA99656.1| outer mitochondrial membrane protein porin

Identities = 141/141 (100%), Positives = 141/141 (100%)

```
Query: 1   AHEGIVGGAEFGYDISAGSISRYAMALSYFAKDYSLGATLNNEQITTVDFFQNVNAFLQV 180
           AHEGIVGGAEFGYDISAGSISRYAMALSYFAKDYSLGATLNNEQITTVDFFQNVNAFLQV
Sbjct: 143 AHEGIVGGAEFGYDISAGSISRYAMALSYFAKDYSLGATLNNEQITTVDFFQNVNAFLQV 202

Query: 181 GAKATMNCKLPNSNVNIEFATRYLPDASSQVKAKVSDSGIVTLAYKQLLRPGVTLGVGSS 360
           GAKATMNCKLPNSNVNIEFATRYLPDASSQVKAKVSDSGIVTLAYKQLLRPGVTLGVGSS
Sbjct: 203 GAKATMNCKLPNSNVNIEFATRYLPDASSQVKAKVSDSGIVTLAYKQLLRPGVTLGVGSS 262

Query: 361 FDALKLSEPVHKLGWSLSFDA 423
           FDALKLSEPVHKLGWSLSFDA
Sbjct: 263 FDALKLSEPVHKLGWSLSFDA 283
```

Figure F.4 POR1 structure and conserved domains

```
gnl|CDD|16959 pfam01459, Porin_3, Eukaryotic porin.

CD-Length = 273 residues, only  49.8% aligned
Score =  126 bits (317), Expect = 1e-30

Query:  1    AHEGIVG------GAEFGYDISAGSISRYAMALSYFAKDYSLGATLNNEQITTVDFFQNV  54
Sbjct:  138  GNPGFLGTPRLALGAEASYDTAKGKPGRYNAALSYATADYWLASATNNAGGAHASYYHKV  197

Query:  55   NAFLQVGAKATMNCKLPNSNVNIEFATRYLPDASSQVKAKVSDSGIVTLAYKQLLRPGVT  114
Sbjct:  198  NDQLQTGVEAELNH--GTNETTFTFGTKYDLDPDTTVRAKVNSNGKVSALLEKRLRPGVT  255

Query:  115  LGVGSSFDALKLSEPVHKLG  134
Sbjct:  256  LTLSAEVDHKKNDA--HKFG  273
```

2KL2-H1 protein sequence alignment

```
gi|1322903|emb|CAA96957.1|  CSE1 [Saccharomyces cerevisiae]
gi|1706161|sp|P33307|CSE1_YEAST   Importin alpha re-exporter (Chromosome segregation
protein CSE1)
gi|6321198|ref|NP_011276.1|  Cse1p [Saccharomyces cerevisiae]
gi|58177143|pdb|1WA5|C     Chain C, Crystal Structure Of The Exportin Cse1p Complexed
With Its Cargo (Kap60p) And Rangtp


Identities = 194/194 (100%), Positives = 194/194 (100%)
```

```
Query: 1   PTLLSDLASRLSNDDMVTNKGVLTVAHSIFKRWRPLFRSDELFLEIKLVLDVFTAPFLNL 60
           PTLLSDLASRLSNDDMVTNKGVLTVAHSIFKRWRPLFRSDELFLEIKLVLDVFTAPFLNL
Sbjct: 122 PTLLSDLASRLSNDDMVTNKGVLTVAHSIFKRWRPLFRSDELFLEIKLVLDVFTAPFLNL 181

Query: 61  LKTVDEQITANENNKASLNILFDVLLVLIKLYYDFNCQDIPEFFEDNIQVGMGIFHKYLS 120
           LKTVDEQITANENNKASLNILFDVLLVLIKLYYDFNCQDIPEFFEDNIQVGMGIFHKYLS
Sbjct: 182 LKTVDEQITANENNKASLNILFDVLLVLIKLYYDFNCQDIPEFFEDNIQVGMGIFHKYLS 241


Query: 121 YSNPLLEDPDETEHASVLIKVKSSIQELVQLYTTRYEDVFGPMINEFIQITWNLLTSISN 180
           YSNPLLEDPDETEHASVLIKVKSSIQELVQLYTTRYEDVFGPMINEFIQITWNLLTSISN
Sbjct: 242 YSNPLLEDPDETEHASVLIKVKSSIQELVQLYTTRYEDVFGPMINEFIQITWNLLTSISN 301

Query: 181 QPKYDILVSKSLSF 194
           QPKYDILVSKSLSF
Sbjct: 302 QPKYDILVSKSLSF 315
```

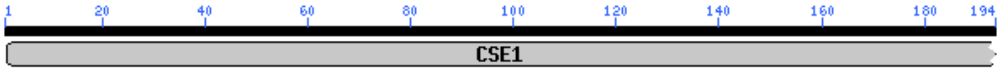


Figure F.5 CSE1 structure and conserved domains

```
gnl|CDD|14741 COG5657, CSE1, CAS/CSE protein involved in chromosome segregation [Cell
division and chromosome partitioning].

CD-Length = 947 residues, only  20.0% aligned
Score =  214 bits (547), Expect = 5e-57

Query:  1    PTLLSDLASRLSNDDMVTNKGVLTVAHSIFKRWRPLFRSDELFLEIKLVLDVFTAPFLNL  60
Sbjct:  122  PTLVPDLLSLLSEKDMVTNENSLRVLHHIFKRLRRLFRSDALFLEIAPVLLSILCPFLFS  181

Query:  61   LKTVDEQITANENNKASLNILFDVLLVLIKLYYDFNCQDIPEFFEDNIQVGMGIFHKYLS  120
Sbjct:  182  SAYFWSMS---ENLDESLLSLFQVCLKLIRRYYDLGFQDIPEFFEDNLDKFMEHFCKLLS  238

Query:  121  YSNPLLEDPDETEHASVLIKVKSSIQELVQLYTTRYEDVFGPMINEFIQITWNLLTSISN  180
Sbjct:  239  YSNPVLQK-DCLED-CVYFKLKGSICEIFNLYTTKYPEVITYLIYDFVEIVWNLLTTITR  296

Query:  181  QPKYDILVSKSLSF  194
Sbjct:  297  PYIRDYLVSKSLTV  310
```

Plant homologs

gi|20197825|gb|AAM15266.1| putative **cellular apoptosis susceptibility protein** [Arabidopsis thaliana]

gi|4415933|gb|AAD20163.1| putative cellular apoptosis susceptibility protein [Arabidopsis thaliana]
gi|15226001|ref|NP_182175.1| cellular apoptosis susceptibility protein, putative / importin-alpha re-exporter, putative [Arabidopsis thaliana]
gi|25349694|pir||H84903 hypothetical protein At2g46520 [imported] - Arabidopsis thaliana
gi|18077710|emb|CAC83300.1| cellular apoptosis susceptibility protein homologue [Arabidopsis thaliana]
gi|20138095|sp|Q9ZPY7|CSE1_ARATH Importin-alpha re-exporter (Cellular apoptosis susceptibility protein homolog)

Identities = 67/201 (33%), Positives = 108/201 (53%), Gaps = 9/201 (4%)

```
Query: 1   PTLLSDLASRLSN----DDMVTNKGVLTVAHSIFKRWRPLFRSDELFLEIKLVLDVFTAP 56
           P LL +L + L N     D V+  G+L  A SIFK++   +R+D LF+++K  LD F AP
Sbjct: 130 PALLPELIANLQNAALAGDYVSVNGILGTASSIFKKFSYEYRTDALFVDLKYCLDNFAAP 189

Query: 57  FLNLLKTVDEQITANENNKAS---LNILFDVLLVLIKLYYDFNCQDIPEFFEDNIQVGMG 113
              +       I +  ++  S   L  LF+   +   ++Y  N QD+PEFFED+++  MG
Sbjct: 190 LTEIFLKTSSLIDSAASSGGSPPILKPLFESQRLCCTIFYSLNFQDLPEFFEDHMKEWMG 249

Query: 114 IFHKYLSYSNPLLEDPDETEHASVLIKVKSSIQELVQLYTTRYEDVFGPMINEFIQITWN 173
            F KYLS + P LE  +  E  +++  ++++I E +  Y  + E+ F   +NEF  + W
Sbjct: 250 EFKKYLSSNYPALESTE--EGLTLVDDLRAAICENINHYIEKNEEEFQGFLNEFASVVWT 307

Query: 174 LLTSISNQPKYDILVSKSLSF 194
           LL  +S  P  D L + ++ F
Sbjct: 308 LLRDVSKSPSRDQLATTAIKF 328
```

2KL2-B1 protein sequence alignment

gi|486363|emb|CAA82049.1| unnamed protein product [Saccharomyces cerevisiae]
gi|549748|sp|P36041|YKU4_YEAST Hypothetical 69.8 kDa protein in LOS1-TOR2 intergenic region
gi|6322645|ref|NP_012718.1| **Eap1p** [Saccharomyces cerevisiae]
Length = 632

Identities = 79/79 (100%), Positives = 79/79 (100%)

```
Query: 2   PMPPLPQGFPIPPNGMLPVTGQQPQPPYPNMMLQGNFPPNFQQGFGSNSPMPIPSIINAN 181
           PMPPLPQGFPIPPNGMLPVTGQQPQPPYPNMMLQGNFPPNFQQGFGSNSPMPIPSIINAN
Sbjct: 554 PMPPLPQGFPIPPNGMLPVTGQQPQPPYPNMMLQGNFPPNFQQGFGSNSPMPIPSIINAN 613

Query: 182 GKNVTNQLPPGLNSKKNIK 238
           GKNVTNQLPPGLNSKKNIK
Sbjct: 614 GKNVTNQLPPGLNSKKNIK 632
```

2K1-A6 protein sequence alignment

gi|663250|emb|CAA88158.1| **TyB protein** - yeast transposon Ty1-H3 [Saccharomyces cerevisiae]

Identities = 182/182 (100%), Positives = 182/182 (100%), Gaps = 0/182 (0%)

```
Query  550  HESSHASKSKDFRHSDSYSENETNHTNVPISSTGGTNNKTVPQISDQETEKRIIHRSPSI  371
            HESSHASKSKDFRHSDSYSENETNHTNVPISSTGGTNNKTVPQISDQETEKRIIHRSPSI
Sbjct  437  HESSHASKSKDFRHSDSYSENETNHTNVPISSTGGTNNKTVPQISDQETEKRIIHRSPSI  496

Query  370  DASPPENNSSHNIVPIKTPTTVSEQNTEESIIAXXXXXXXXPESPTEFPDPFKELPPINS  191
            DASPPENNSSHNIVPIKTPTTVSEQNTEESIIADLPLPDLPPESPTEFPDPFKELPPINS
Sbjct  497  DASPPENNSSHNIVPIKTPTTVSEQNTEESIIADLPLPDLPPESPTEFPDPFKELPPINS  556

Query  190  HQTNSSLGGIGDSNAYTTINSKKRSLEDNETEIKVSRDTWNTKNMRSLEPPRSKKRIHLI  11
            HQTNSSLGGIGDSNAYTTINSKKRSLEDNETEIKVSRDTWNTKNMRSLEPPRSKKRIHLI
Sbjct  557  HQTNSSLGGIGDSNAYTTINSKKRSLEDNETEIKVSRDTWNTKNMRSLEPPRSKKRIHLI  616

Query  10   AA  5
            AA
Sbjct  617  AA  618
```

3L2-B4 protein sequence alignment

gi|6323191|ref|NP_013263.1| Protein of unknown function; overexpression confers resistance to the antimicrobial peptide MiAMP1; Ylr162wp [Saccharomyces cerevisiae]
gi|1234844|gb|AAB67486.1| L9632.2 gene product
gi|2132698|pir||S68478 probable membrane protein YLR162w - yeast (Saccharomyces cerevisiae)

Identities = 53/58 (91%), Positives = 53/58 (91%), Gaps = 0/58 (0%)

```
Query  411  NKSDNKGLISADRNNKATLLLTIPRCTSKSYTNDLSPRKMTLQFASKHPRPFRQVHRC  584
            NKSDNKGLISADRNNKATLLLTIPRCTSKSYTNDLSP KMTL  A KHPRPFRQ HRC
Sbjct  61   NKSDNKGLISADRNNKATLLLTIPRCTSKSYTNDLSPLKMTLLSAGKHPRPFRQEHRC  118
```

2KL2-G6 protein sequence alignment

gi|6323872|ref|NP_013943.1| **SR protein kinase** (SRPK) involved in regulating proteins involved mRNA metabolism and cation homeostasis; similar to human SRPK1; Sky1p [Saccharomyces cerevisiae]
gi|854468|emb|CAA89931.1| unknown [Saccharomyces cerevisiae]
gi|2499619|sp|Q03656|SKY1_YEAST Serine/threonine-protein kinase SKY1 (SRPK)

Identities = 173/176 (98%), Positives = 173/176 (98%), Gaps = 0/176 (0%)

```
Query  3    GPNGVHVVXVFEVLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDI  182
            GPNGVHVV VFEVLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDI
Sbjct  236  GPNGVHVVMVFEVLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDI  295

Query  183  KPENVLMEIGDVEGIVQMVEALDKQKREAKRLQRHVSRSSDITANDSSDEKWAECQTSMP  362
            KPENVLMEIGDVEGIVQMVEALDKQKREAKRLQRHVSRSSDITANDSSDEKWAECQTSMP
Sbjct  296  KPENVLMEIGDVEGIVQMVEALDKQKREAKRLQRHVSRSSDITANDSSDEKWAECQTSMP  355

Query  363  CGSSSNSKSRSIEKDLSKRCXXRPRRHTIITGSQPLPSPISSSNFFEMRAHFCGSS  530
            CGSSSNSKSRSIEKDLSKRC  RPRRHTIITGSQPLPSPISSSNFFEMRAHFCGSS
Sbjct  356  CGSSSNSKSRSIEKDLSKRCFRRPRRHTIITGSQPLPSPISSSNFFEMRAHFCGSS  411
```

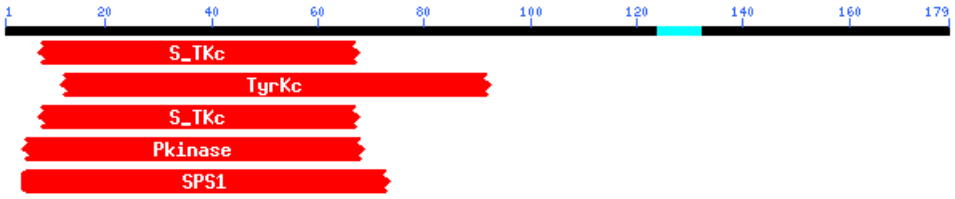


Figure F.6 SRPK structure and conserved domains

gnl|CDD|17776 cd00180, S_TKc, Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail.

```
CD-Length = 256 residues, only  23.4% aligned
Score = 59.8 bits (145), Expect = 2e-10

Query:  7    HVVXVFE-VLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDIKPEN  65
Sbjct:  73   KLYLVMEYMSGGDLFDLLKK--RGRLSEDEARFYFRQILSGLEYLHSK-GIIHRDLKPEN  129

Query:  66   VLM  68
Sbjct:  130  ILL  132
```

gnl|CDD|5392 cd00192, TyrKc, Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-specific kinase subfamily. Enzymes with TyrKc domains belong to an extensive family of proteins which share a conserved catalytic core common to both serine/threonine and tyrosine protein kinases. Enzymatic activity of tyrosine protein kinases is controlled by phosphorylation of specific tyrosine residues in the activation segment of the catalytic domain or a C-terminal tyrosine (tail) residue with reversible conformational changes.

```
CD-Length = 269 residues, only  32.7% aligned
Score = 35.2 bits (81), Expect = 0.005
```

```
Query:  11   VFE-VLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDIKPENVL--  67
Sbjct:  87   VMEYMEGGDLLDYLRKNRPDSLSLSDLLSFALQIAKGMEYLESK-NFVHRDLAARNCLVG  145

Query:  68   ----MEIGDVEGIVQMVEALDKQKREAKRL  93
Sbjct:  146  ENKVVKIADF-GLARDLYDDDYYRKGGGKL  174
```

gnl|CDD|25296 smart00220, S_TKc, Serine/Threonine protein kinases, catalytic domain; Phosphotransferases. Serine or threonine-specific kinase subfamily.

```
CD-Length = 256 residues, only  23.4% aligned
Score = 58.3 bits (141), Expect = 5e-10

Query:  7    HVVXVFEVL-GENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDIKPEN  65
Sbjct:  72   KLYLVMEYCEGGDLFDLLKK--RGRLSEDEARFYFRQILSALEYLHSN-GIIHRDLKPEN  128

Query:  66   VLM  68
Sbjct:  129  ILL  131
```

gnl|CDD|25380 pfam00069, Pkinase, Protein kinase domain.

```
CD-Length = 258 residues, only  24.8% aligned
Score = 52.6 bits (126), Expect = 3e-08

Query:  4    NGVHVVXVFE-VLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTDIK  62
Sbjct:  69   DGDHLYLVMEYMEGGDLFDYLRK--KGPLSEKEAKKIALQILRGLEYLHSR-GIVHRDLK  125

Query:  63   PENVLME  69
Sbjct:  126  PENILLD  132
```

gnl|CDD|10386 COG0515, SPS1, Serine/threonine protein kinase [General function prediction only / Signal transduction mechanisms / Transcription / DNA replication, recombination, and repair].

```
CD-Length = 384 residues, only  18.5% aligned
Score = 41.3 bits (94), Expect = 6e-05

Query:  4    NGVHVVXVFEVLGENLLALIKKYEHRG-IPLIYVKQISKQLLLGLDYMHRRCGIIHTDIK  62
Sbjct:  70   EGSLYLVMEYVDGGSLEDLLKKIGRKGPLSESEALFILAQILSALEYLHSK-GIIHRDIK  128

Query:  63   PENVLMEIGDVE  74
Sbjct:  129  PENILLDRDGRV  140
```

Plant homologs

gi|9843649|emb|CAC03678.1| SRPK4 [Arabidopsis thaliana]

```
Identities = 41/68 (60%), Positives = 54/68 (79%), Gaps = 0/68 (0%)

Query  1    SGPNGVHVVXVFEVLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTD  60
            SGPNG HV  VFE LG+NLL LIK  ++RG+P+  VK+I   +L+GLDY+H++  IIHTD
Sbjct  114  SGPNGQHVCMVFEYLGDNLLTLIKYSDYRGLPIPMVKEICYHMLVGLDYLHKQLSIIHTD  173

Query  61   IKPENVLM  68
            +KPENVL+
Sbjct  174  LKPENVLL  181
```

gi|18409750|ref|NP_566977.1| protein kinase family protein [Arabidopsis thaliana]

Identities = 41/68 (60%), Positives = 54/68 (79%), Gaps = 0/68 (0%)

```
Query  1    SGPNGVHVVXVFEVLGENLLALIKKYEHRGIPLIYVKQISKQLLLGLDYMHRRCGIIHTD  60
            SGPNG HV  VFE LG+NLL LIK  ++RG+P+  VK+I   +L+GLDY+H++  IIHTD
Sbjct  114  SGPNGQHVCMVFEYLGDNLLTLIKYSDYRGLPIPMVKEICYHMLVGLDYLHKQLSIIHTD  173

Query  61   IKPENVLM  68
            +KPENVL+
Sbjct  174  LKPENVLL  181
```

APPENDIX G

PROTEIN SEQUENCE ALIGNMENTS AND CONSERVED DOMAINS OF WR5 INTERACTORS

5L1-C1 protein sequence alignment

gi|1008199|emb|CAA89348.1| unnamed protein product [Saccharomyces cerevisiae]
gi|6322404|ref|NP_012478.1| **probable serine/threonine kinase; Yjl057cp** [Saccharomyces cerevisiae]
gi|1346384|sp|P47042|KJF7_YEAST Probable serine/threonine-protein kinase YJL057C

Identities = 139/139 (100%), Positives = 139/139 (100%)

```
Query: 336 MSLVPYEEGSLILDDPNSKSVVVVNPTSGTLSFFQQDNSNNDSELNEDQTASLSALDFPS 515
           MSLVPYEEGSLILDDPNSKSVVVVNPTSGTLSFFQQDNSNNDSELNEDQTASLSALDFPS
Sbjct: 1   MSLVPYEEGSLILDDPNSKSVVVVNPTSGTLSFFQQDNSNNDSELNEDQTASLSALDFPS 60

Query: 516 GIHQYKSPIASYVCPQCGTEINPDIINRRQLHRRASAGVESESSRLSIPENTVPLGFEFA 695
           GIHQYKSPIASYVCPQCGTEINPDIINRRQLHRRASAGVESESSRLSIPENTVPLGFEFA
Sbjct: 61  GIHQYKSPIASYVCPQCGTEINPDIINRRQLHRRASAGVESESSRLSIPENTVPLGFEFA 120

Query: 696 NSSFSRRYFQSLERNHRHY 752
           NSSFSRRYFQSLERNHRHY
Sbjct: 121 NSSFSRRYFQSLERNHRHY 139
```

Figure G.1 YJL057C structure and conserved domains

gnl|CDD|17776 cd00180, S_TKc, Serine/Threonine protein kinases, catalytic domain. Phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail.

CD-Length = 256 residues, 97.3% aligned
Score = 123 bits (310), Expect = 7e-29

```
Query:  173  FKILSLLGNGARGSVYKVVHTIGNTELGVFALKKIPIGNDMEWFNKCI-REVKALSSLTH  231
Sbjct:  2    YELLEKLGKGAFGKVYLARDKKTGEL---VAIKIIKKKKEKKKQVERILREIKILKRLNH  58

Query:  232  KSANLITYNHVWLEMDSSVGFVRSIDGSQSDSQEEVPCIFILQQYCSGGNLEDCILRKVF  291
Sbjct:  59   P--NIVKLYDVFEDED----------------------KLYLVMEYMSGGDLFDLLKKR--  93

Query:  292  NRFSDTESPEE*RKKKFRTRKK*NHGKSGEVGLSTEQLVSIIRDIARGLHELHSIGLIHRDL  351
Sbjct:  94   ----------------------------GRLSEDEARFYFRQILSGLEYLHSKGIIHRDL  125

Query:  352  KPSNCLLLTPFKSDNENDNVYDREHNSDEFFPSIVIGDLGES-QLEGESRLGTGCTGTLE  410
Sbjct:  126  KPENILL-------DEDGHVK--------------IADFGLAKQLDSGGRKLTTFVGTPW  164

Query:  411  FTAPDLIIQGRPV*SSSTLPSRSS*HTYNEYTFASDMYSLGMICYFIVFGELPFEPQLDIVD  470
Sbjct:  165  YMAPE-VLLGKG----------------YGKKVDIWSLGVILYELLTGKPPF-PGDNEED  206

Query:  471  LKVRIKNFRFDTEGMIEKHQAMKLKPIDRRIFHLMDALLQPNNDARPTAKTV  522
Sbjct:  207  LLEKILKGKGTPDELPPN--------ISPEAKDLIKKLLVKDPEKRPTAEEL  250
```

gnl|CDD|5392 cd00192, TyrKc, Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-specific kinase subfamily. Enzymes with TyrKc domains belong to an extensive family of proteins which share a conserved catalytic core common to both serine/threonine and tyrosine protein kinases. Enzymatic activity of tyrosine protein kinases is controlled by phosphorylation of specific tyrosine residues in the activation segment of the catalytic domain or a C-terminal tyrosine (tail) residue with reversible conformational changes.

CD-Length = 269 residues, 98.1% aligned
Score = 81.4 bits (201), Expect = 3e-16

```
Query:  170  RKFFKILSLLGNGARGSVYK--VVHTIGNTELGVFALKKIPIGNDMEWFNKCIREVKALS  227
Sbjct:  6    REDLTLGKKLGEGAFGEVYKGTLKGKGDNKSIDV-AVKTLKEDASEEQIKEFLREAKIMR  64

Query:  228  SLTHKSANLITYNHVWLEMDSSVGFVRSIDGSQSDSQEEVPCIFILQQYCSGGNLEDCIL  287
Sbjct:  65   KLKHP--NIVRLLGVCTEEEP----------------------LMIVMEYMEGGDLLDY-L  100

Query:  288  RKVFNRFSDTESPEE*RKKKFRTRKK*NHGKSGEVGLSTEQLVSIIRDIARGLHELHSIGLI  347
Sbjct:  101  R--------------------------KNRPDSLSLSDLLSFALQIAKGMEYLESKNFV  133

Query:  348  HRDLKPSNCLLLTPFKSDNENDNVYDREHNSDEFFPSIVIGDLGESQL--EGESRLGTGC  405
Sbjct:  134  HRDLAARNCLV-------GENKVVK--------------IADFGLARDLYDDDYYRKGGG  172

Query:  406  TGTLEFTAPDLIIQGRPV*SSSTLPSRSS*HTYNEYTFASDMYSLGMICYFIV-FGELPFEP  464
Sbjct:  173  KLPIRWMAPESLKDGK-----------------FTSKSDVWSFGVLLWEIFTLGEQPY-P  214

Query:  465  QLDIVDLKVRIKNFRFDTEGMIEKHQAMKLKPIDRRIFHLMDALLQPNNDARPTAKTVEE  524
Sbjct:  215  GMSNQEVLEYLKK----------GYRLPQPENCPDEIYDLMLQCWAADPEDRPTFSELVE  264

Query:  525  TLDEM  529
Sbjct:  265  RLEAL  269
```

gnl|CDD|25296 smart00220, S_TKc, Serine/Threonine protein kinases, catalytic domain; Phosphotransferases. Serine or threonine-specific kinase subfamily.

```
CD-Length = 256 residues,  97.7% aligned
Score =  124 bits (313), Expect = 3e-29


Query:  173  FKILSLLGNGARGSVYKVVHTIGNTELGVFALKKIPIGNDMEWFNKCI-REVKALSSLTH  231
Sbjct:  1    YELLEVLGKGAFGKVYLARDKKTGK---LVAIKVIKKEKLKKKKRERILREIKILKKLDH  57

Query:  232  KSANLITYNHVWLEMDSSVGFVRSIDGSQSDSQEEVPCIFILQQYCSGGNLEDCILRKVF  291
Sbjct:  58   P--NIVKLYDVFEDKD----------------------KLYLVMEYCEGGDLFDLLKKR--  92

Query:  292  NRFSDTESPEE*RKKKFRTRKK*NHGKSGEVGLSTEQLVSIIRDIARGLHELHSIGLIHRDL  351
Sbjct:  93   ----------------------------GRLSEDEARFYFRQILSALEYLHSNGIIHRDL  124

Query:  352  KPSNCLLLTPFKSDNENDNVYDREHNSDEFFPSIVIGDLGES-QLEGESRLGTGCTGTLE  410
Sbjct:  125  KPENILL-------DSDGHVK--------------LADFGLAKQLDSGGTLLTTFVGTPE  163

Query:  411  FTAPDLIIQGRPV*SSSTLPSRSS*HTYNEYTFASDMYSLGMICYFIVFGELPFEPQLDIVD  470
Sbjct:  164  YMAPEVL-----------------LGKGYGKAVDIWSLGVILYELLTGKPPFPGDDQLDA  206

Query:  471  LKVRIKNFRFDTEGMIEKHQAMKLKPIDRRIFHLMDALLQPNNDARPTAKTV  522
Sbjct:  207  LFKKIGKPPPPFP--------PPEWKISPEAKDLIRKLLVKDPEKRLTAEEA  250


gnl|CDD|24220  smart00219,  TyrKc,  Tyrosine  kinase,  catalytic  domain;
Phosphotransferases. Tyrosine-specific kinase subfamily.

CD-Length = 255 residues,  99.6% aligned
Score = 76.4 bits (188), Expect = 1e-14

Query:  174  KILSLLGNGARGSVYKVVHTIGNTELGVFALKKIPIGNDMEWFNKCIREVKALSSLTHKS  233
Sbjct:  2    TLGKKLGEGAFGEVYKGTLKGKSGKEVEVAVKTLKEDASEQI-EEFLREAKIMRKLDHP-  59

Query:  234  ANLITYNHVWLEMDSSVGFVRSIDGSQSDSQEEVPCIFILQQYCSGGNLEDCILRKVFNR  293
Sbjct:  60   -NIVKL-----------LGVC---------TEEEPLMI--VMEYMEGGDLLDYL-------  90

Query:  294  FSDTESPEE*RKKKFRTRKK*NHGKSGEVGLSTEQLVSIIRDIARGLHELHSIGLIHRDLKP  353
Sbjct:  91   ----------------------RKNRPKLSLSDLLSFALQIARGMEYLESKNFVHRDLAA  128

Query:  354  SNCLLLTPFKSDNENDNVYDREHNSDEFFPSIVIGDLGESQLEGESRL--GTGCTGTLEF  411
Sbjct:  129  RNCLV-------GENKTVK--------------IADFGLSRDLYSDDYYKVKGGKLPIRW  167

Query:  412  TAPDLIIQGRPV*SSSTLPSRSS*HTYNEYTFASDMYSLGMICYFIV-FGELPFEPQLDIVD  470
Sbjct:  168  MAPESL-----------------KYGKFTSKSDVWSFGVLLWEIFTLGESPY-PGMSNEE  209

Query:  471  LKVRIKNFRFDTEGMIEKHQAMKLKPIDRRIFHLMDALLQPNNDARPTAKTVEETL  526
Sbjct:  210  VLEYLKK----------GYRLPQPPNCPDEIYDLMLQCWAEDPEDRPSFSELVERL  255


gnl|CDD|25380 pfam00069, Pkinase, Protein kinase domain.

CD-Length = 258 residues,  96.9% aligned
Score =  127 bits (321), Expect = 3e-30

Query:  173  FKILSLLGNGARGSVYKVVHTIGNTELGVFALKKIPIGNDMEWFNK-CIREVKALSSLTH  231
Sbjct:  1    YELGRKLGSGSFGKVYKAKHKDTGKI---VAVKILKKRSEKSKKDEQALREIQILRRLSH  57

Query:  232  KSANLITYNHVWLEMDSSVGFVRSIDGSQSDSQEEVPCIFILQQYCSGGNLEDCILRKVF  291
Sbjct:  58   P--NIVRLLGVFEDGDH----------------------LYLVMEYMEGGDLFDYLRKK--  92

Query:  292  NRFSDTESPEE*RKKKFRTRKK*NHGKSGEVGLSTEQLVSIIRDIARGLHELHSIGLIHRDL  351
Sbjct:  93   ----------------------------GPLSEKEAKKIALQILRGLEYLHSRGIVHRDL  124

Query:  352  KPSNCLLLTPFKSDNENDNVYDREHNSDEFFPSIVIGDLGES-QLEGESRLGTGCTGTLE  410
Sbjct:  125  KPENILL-------DENGQVK--------------IADFGLAKKLESSYSKLTTFVGTPR  163
```

```
Query:  411  FTAPDLIIQGRPVSSSTLPSRSSHTYNEYTFASDMYSLGMICYFIVFGELPFEPQLDIVD  470
Sbjct:  164  YMAPEVLLGGRG----------------YSSKVDVWSLGVILYELLTGKPPF-PGEGLID  206

Query:  471  LKVRIKNFRFDTEGMIEKHQAMKLKPIDRRIFHLMDALLQPNNDARPTAK  520
Sbjct:  207  QLQLIER------ILRPPLPFDCPKSDSEELKDLIKKCLNKDPEKRPTAK  250
```

gnl|CDD|10386 COG0515, SPS1, Serine/threonine protein kinase [General function prediction only / Signal transduction mechanisms / Transcription / DNA replication, recombination, and repair].

```
CD-Length = 384 residues,  94.3% aligned
Score = 69.0 bits (166), Expect = 2e-12

Query:  173  FKILSLLGNGARGSVYKVVHTIGNTELGVFALKKIPIG--NDMEWFNKCIREVKALSSLT  230
Sbjct:  2    YRILRKLGEGSFGEVYLARDRK------LVALKVLAKKLESKSKEVERFLREIQILASLN  55

Query:  231  HKSANLITYNHVWLEMDSSVGFVRSIDGSQSDSQEEVPCIFILQQYCSGGNLEDCILRKV  290
Sbjct:  56   HPP-----------------NIVKLYDFFQDEGS-----LYLVMEYVDGGSLEDLLKKIG  93

Query:  291  FNRFSDTESPEERKKKFRTRKKNHGKSGEVGLSTEQLVSIIRDIARGLHELHSIGLIHRD  350
Sbjct:  94   RKGP--------------------------LSESEALFILAQILSALEYLHSKGIIHRD  126

Query:  351  LKPSNCLLLTPFKSDNENDNVYDREHNSDEFFPSIVIGDLGESQLEGESRLGTGCTGTLE  410
Sbjct:  127  IKPENILL----DRDGRVVKLID--------FGLAKLLPDPGSTSSIPALPSTSV-GTPG  173

Query:  411  FTAPDLIIQGRPVSSSTLPSRSSHTYNEYTFASDMYSLGMICYFIVFGELPFEPQLDIVD  470
Sbjct:  174  YMAPEVLLGLSL--------------AYASSSSDIWSLGITLYELLTGLPPFEGEKNSSA  219

Query:  471  LKVRIKNFRFDTEGMIEKHQAMKLKP-IDRRIFHLMDALLQPNNDARPTAKTVEETLDEM  529
Sbjct:  220  TSQTLKIILELPTPSLASPLSPSNPELISKAASDLLKKLLAKDPKNRLSSS--SDLSHDL  277

Query:  530  LINSKPGKKFWKENVDSTLNFSTISEVNENTNSFTDDYIDGDNVTLSLPAPEENLSTVST  589
Sbjct:  278  LAHLKLKESDLSDLLKPDDSAPLRLSLPPSLEALISSLNSLAISGSDLKLDDSNFSKELA  337

Query:  590  QNLQKYSALNRTIQVCYKLVSMILTI  615
Sbjct:  338  PNGVSSSPHNSSSLLLSTASSKRSSL  363
```

5L2-G5

gi|37362667|ref|NP_012444.2| Protein involved in the inositol acylation of glucosaminyl phosphatidylinositol (GlcN-PI) to form glucosaminyl(acyl)phosphatidylinositol (GlcN(acyl)PI), an intermediate in the biosynthesis of glycosylphosphatidylinositol (GPI) anchors; Gwt1p [Saccharomyces cerevisiae]

```
Identities = 188/188 (100%), Positives = 188/188 (100%)

Query: 1   FLKNAFNALKSGGTLLFLGLLRLFFVKNLEYQEHVTEYGVHWNFFITLSLLPLVLTFIDP 60
           FLKNAFNALKSGGTLLFLGLLRLFFVKNLEYQEHVTEYGVHWNFFITLSLLPLVLTFIDP
Sbjct: 195 FLKNAFNALKSGGTLLFLGLLRLFFVKNLEYQEHVTEYGVHWNFFITLSLLPLVLTFIDP 254

Query: 61  VTRMVPRCSIAIFISCIYEWLLLKDDRTLNFLILADRNCFFSANREGIFSFLGYCSIFLW 120
           VTRMVPRCSIAIFISCIYEWLLLKDDRTLNFLILADRNCFFSANREGIFSFLGYCSIFLW
Sbjct: 255 VTRMVPRCSIAIFISCIYEWLLLKDDRTLNFLILADRNCFFSANREGIFSFLGYCSIFLW 314

Query: 121 GQNTGFYLLGNKPTLNNLYKPSTQDVVAASKKSSTWDYWTSVTPLSGLCIWSTIFLVISQ 180
           GQNTGFYLLGNKPTLNNLYKPSTQDVVAASKKSSTWDYWTSVTPLSGLCIWSTIFLVISQ
Sbjct: 315 GQNTGFYLLGNKPTLNNLYKPSTQDVVAASKKSSTWDYWTSVTPLSGLCIWSTIFLVISQ 374
```

```
Query: 181 LVFQYHPY 188
           LVFQYHPY
Sbjct: 375 LVFQYHPY 382
```

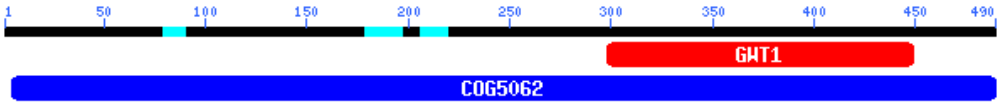


Figure G.2 GWT1 structure and conserved domains

gnl|CDD|26536 pfam06423, GWT1, GWT1. Glycosylphosphatidylinositol (GPI) is a conserved post-translational modification to anchor cell surface proteins to plasma membrane in eukaryotes. GWT1 is involved in GPI anchor biosynthesis; it is required for inositol acylation in yeast.

```
CD-Length = 149 residues, 100.0% aligned
Score =  144 bits (365), Expect = 2e-35

Query:  297  ANREGIFSFLGYCSIFLWGQNTGFYLLGNKPTLNNLYKPSTQDVVAASKKSSTWDYWTSV  356
Sbjct:  1    ANREGLVSLLGYVAIFLAGIYTGRYMLAKDPLR---YKSWIRKVRLLFALSLCLFLSFVS  57

Query:  357  TPLSGLCIWSTIFLVISQLVFQYHPYSVSRRFANLPYTLWVITYNLLFLTGYCLTDKIFG  416
Sbjct:  58   NFLSEGVSRVTARLGNVIWVFAILSLMLFASGIYL--TFLIDTYLFPSSWKLILSPEIFK  115

Query:  417  NSSEYYKVAECLE-SINSNGLFLFLLANVSTGLVN  450
Sbjct:  116  KKS-GSKVPECLEDSLNRNGLFFFLLSNVLTGLVN  149
```

gnl|CDD|14191 COG5062, COG5062, Uncharacterized membrane protein [Function unknown].

```
CD-Length = 429 residues, 100.0% aligned
Score =  403 bits (1036), Expect = 3e-113

Query:  4    LKQRKEDFVTGLNGGSITEINAVTSIALVTYISWNLLKNSNLMPPGISSVQYIIDFALNW  63
Sbjct:  1    LKPCKEGFVTNLTL-SEIELFAVTSITHLSIVVYHLM-----LPKS-----TLLEFMF--  47

Query:  64   VALLLSITIYASEPYLLNTLILLPCLLAFIYGKFTSSSKPSNP-IYNKK-KMITQRFQ--  119
Sbjct:  48   -CLVPLYSIIVFLKYSGLTYALFLCLLGFLFLFWRRLRKPVNPGNPSSKPKQSTDAPQDD  106

Query:  120  LEKKPY-ITAYRGGMLILTAIAILAVDFPIFPRRFAKVETWGTSLMDLGVGSFVFSNGIV  178
Sbjct:  107  EPYTSMAITRYRFLIIGCTVIAILAVDFPFFPRRLGKSETWGTSLMDIGVGSFVYNSGIV  166

Query:  179  SSRALLKNLSLKSKPSFLKNAFNALKSGGTLLFLGLLRLFFVKNLEYQEHVTEYGVHWNF  238
Sbjct:  167  STRAKSKR---------------KLKNALILLFLGFLRYFSVKLLNYQVHVREYGVHWNF  211

Query:  239  FITLSLLPLVLTFIDPVTRMVPRCSIAIFISCIYEWLLLKDDRTLNFLILADRNCFFSAN  298
Sbjct:  212  FFTLGLLNLASLFIRTRA----NFLLGFFICLTHELLLKFFG-LEKFIYSAARSSILTSN  266
```

```
Query:  299  REGIFSFLGYCSIFLWGQNTGFYLLGNKPTLNNLYKPSTQDVVAASKKSSTWDYWTSVTP  358
Sbjct:  267  REGITSLLPYISIFLMGADTGKVVFKKKPT-----------------RKKAWKIIILYN-  308

Query:  359  LSGLCIWSTIFLVISQLVFQYHPYSVSRRFANLPYTLWVITYNLLFLTGYCLTDKIFGNS  418
Sbjct:  309  AFFLCV---------YLVFNFYSTS-SRRLANLPFVMWIMLLHTFHLTVYELFDR---TS  355

Query:  419  SEYYKVAECLESINSNGLFLFLLANVSTGLVNMSMVTIDS--SPLKSFLVLLAYCSFIAV  476
Sbjct:  356  KIYNLVMHRFESKNLNFLLVFSNANVLTGKVNSSSKTMSSCSSHAKCLTYLVAVFSIPCA  415

Query:  477  ISVFLYRKRIFIKL  490
Sbjct:  416  INSKLAGKLIVSGL  429
```

Plant homologs

gi|15236699|ref|NP_193525.1| **zinc finger** (C3HC4-type RING finger) family protein /pentatricopeptide (PPR) repeat-containing protein [Arabidopsis thaliana]
Length = 1208

```
Identities = 74/209 (35%), Positives = 109/209 (52%), Gaps = 25/209 (11%)

Query: 126 ITAYRGGMLILTAIAILAVDFPIFPRRFAKVETWGTSLMDLGVGSFVFSNGIVXXXXXXX 185
           +++YR  ++++T + ILAVDF IFPRR+AK ET+GTSLMDLGVGSFV +N +V
Sbjct: 664 VSSYRVALMLITCLCILAVDFTIFPRRYAKTETYGTSLMDLGVGSFVLANAVV------- 716

Query: 186 XXXXXXXXXXXXNAFNALKSXXXXXXXXXXXXXXVKNLEYQEHVTEYGVHWNFFITLSLL 245
                       N    +K+                 ++YQ HVTEYGVHWNFF TL+ +
Sbjct: 717 --SRQARDVSSGNWITGIKATAPLLLLGFIRLVTTSGVDYQVHVTEYGVHWNFFFTLAAI 774



Query: 246 PLVLTFIDPVTRMVPRCSIAIFISCIYEWLLLKDDRTLNFLILADR--NCFFSANREGIF 303
            ++ +F++   +       A+ ++    WLL      LN  +L+D       S N+EG++
Sbjct: 775 SILTSFVNIPAKYCGLLGFAV-LAGYQTWLL----SGLNTYLLSDERGTDIISKNKEGVY 829

Query: 304 SFLGYCSIFLWGQNTGFYLLGNKPTLNNL 332
           S LG   I  W       LLG+ P+  +L
Sbjct: 830 SILG--KILTW-------LLGHVPSWRSL 849
```

gi|11994148|dbj|BAB01169.1| **non-race specific disease resistance protein-like** [Arabidopsis thaliana]
gi|22331216|ref|NP_683581.1| non-race specific disease resistance protein, putative [Arabidopsis thaliana]
Length = 222

```
Identities = 18/63 (28%), Positives = 32/63 (50%), Gaps = 11/63 (17%)

Query: 42 WNFFITL----------SLLPLVLTFIDPVTRMVPRCSIAIF-ISCIYEWLLLKDDRTLN 90
          W +F+T+          S L  +  ++  +   +PRCSI  F I  + + L+  D+ TLN
Sbjct: 7  WEWFVTIVGSLMTLLYVSFLLALCLWLSTLVHHIPRCSIHYFYIPALNKSLISSDNTTLN 66

Query: 91 FLI 93
          F++
Sbjct: 67 FMV 69
```

5L1-A6 protein sequence alignment (also an interactor of RAD6)

gi|6325060|ref|NP_015128.1| Protein of unknown function required for normal levels of resistance to oxidative damage, null mutants are sensitive to hydrogen peroxide; member of a conserved family of proteins found in eukaryotes but not in prokaryotes; **Oxr1p** [Saccharomyces cerevisiae]
gi|2132227|pir||S65215 hypothetical protein YPL196w - yeast (Saccharomyces cerevisiae)

Identities = 105/105 (100%), Positives = 105/105 (100%)

```
Query: 2   LDKVPDVNISEKEESEQEGKEGKEEGDKEERWRFSGYPYTGVNEFAIYCTSEFLSMGAGD 181
           LDKVPDVNISEKEESEQEGKEGKEEGDKEERWRFSGYPYTGVNEFAIYCTSEFLSMGAGD
Sbjct: 169 LDKVPDVNISEKEESEQEGKEGKEEGDKEERWRFSGYPYTGVNEFAIYCTSEFLSMGAGD 228


Query: 182 GHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVWRVG 316
           GHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVWRVG
Sbjct: 229 GHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVWRVG 273
```

Figure G.3 OXR1 structure and conserved domains

gnl|CDD|3951 smart00584, TLDc, domain in TBC and LysM domain containing proteins;

CD-Length = 166 residues, 100.0% aligned
Score = 158 bits (402), Expect = 5e-40

```
Query:  73   NKLLTPEMCDEIRTLMPTRIQLYTEWNLLYSLEQHGSSLHSLYSNVAPDSKEFRRVGYVL  132
Sbjct:  1    SSILSVEILALINSLLPTRTRGYP-WTLLYSSSQHGYSLNTLYRKVEGYGP-----PTLL  54

Query:  133  VIKDRKNGIFGAYSNEAFHPNEHRQYTGNGECFLWKLDKVPDVNIS*EKEESEQEGKEGKE*  192
Sbjct:  55   IIKDTDGEVFGAYASQAWRVSPH--FYGTGESFLFQLN----------------------  90

Query:  193  *EGDKEE*RWRFSGYPYTGVN--EFAIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTY  250
Sbjct:  91   -------PKFVVYDATGKNKNYYYINGTPDSLPIGGGGGGFGLWIDEDLNHGSSSSCKTF  143

Query:  251  GNEVLSKEGKKFSIVALEVWRVG  273
Sbjct:  144  GNPPLLTKDEDFLILDIEVWGFG  166
```

Plant homologs

gi|51535088|dbj|BAD37677.1| **nucleolar protein-like** [Oryza sativa (japonica cultivar-group)]

```
Identities = 64/241 (26%), Positives = 106/241 (43%), Gaps = 49/241 (20%)

Query: 43  FRNKTTSE---RTIVEEGSLPPVRLNGYLPSTKNKLL----TPEM--CDEIRTLMPTRIQ 93
           FR K +      T+  E   P + L G   S  +  L     P M   + +RT++ T +
Sbjct: 139 FRQKPSYSIDGETVQTEYDAPGLELKGSKESASHDKLPAMSEPSMLLSETMRTVLYTSLP 198

Query: 94  LYTE---WNLLYSLEQHGSSLHSLYSNVAPDSKEFRRVGY-VLVIKDRKNGIFGAYSNEA 149
           +  +   W L+YS  +HG SL +LY       +     GY +L++ DRK  +FG
Sbjct: 199 VLVQGRNWMLVYSTWRHGISLSTLYR------RSMLCAGYSLLIVGDRKGAVFGGLVEAP 252

Query: 150 FHPNEHRQYTGNGECFLWKLDKVPDVNISXXXXXXXXXXXXXXXXXXXXRWRFSGYPYTG 209
             P   ++Y G   CF++        NI+                          Y  TG
Sbjct: 253 LQPLIKKKYQGTNNCFVF-------TNIAGRPVI----------------------YRPTG 284

Query: 210 VNEFAIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEV 269
            N +  +C++++L+MG G GH+ L  D  LL+G S+  +T+ N  LS+  ++F +  +E+
Sbjct: 285 ANNYFTFCSTDYLAMGGG-GHFALYLDGDLLNGSSSTSETFNNPCLSRS-REFEVKDVEL 342

Query: 270 W 270
           W
Sbjct: 343 W 343
```

gi|50912883|ref|XP_467849.1| **oxidation protection protein-like** [Oryza sativa (japonica cultivar-group)]
gi|46805939|dbj|BAD17233.1| oxidation protection protein-like [Oryza sativa (japonica cultivar-group)]
gi|46390139|dbj|BAD15574.1| oxidation protection protein-like [Oryza sativa (japonica cultivar-group)]

```
Identities = 50/177 (28%), Positives = 79/177 (44%), Gaps = 43/177 (24%)

Query: 98  WNLLYSLEQHGSSLHSLY--SNVAPDSKEFRRVGY-VLVIKDRKNGIFGAYSNEAFHPNE 154
           W LLYS  +HG SL +LY  S + P        GY +LV+ D++  +FG        P
Sbjct: 202 WVLLYSTWRHGISLSTLYRRSMLCP--------GYSLLVVGDKEGAVFGGLVEAPLQPTS 253

Query: 155 HRQYTGNGECFLW-KLDKVPDVNISXXXXXXXXXXXXXXXXXXXXRWRFSGYPYTGVNEF 213
            ++Y G+  CF++  L   P +                          Y  TG N +
Sbjct: 254 AKKYQGSNSCFVFTNLHSNPSI----------------------------YRPTGANNY 284

Query: 214 AIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVW 270
              C++++L++G G GH+ L  D  LL G S+  +T+ N  LS     F++  +E+W
Sbjct: 285 FTVCSTDYLALGGG-GHFALYLDADLLSGSSSNSETFNNMCLS-HSPDFAVKDVELW 339
```

gi|50936597|ref|XP_477826.1| **oxidation resistance 1-like protein** [Oryza sativa (japonica cultivar-group)]
gi|34394846|dbj|BAC84293.1| oxidation resistance 1-like protein [Oryza sativa (japonica cultivar-group)]
gi|50508536|dbj|BAD30835.1| oxidation resistance 1-like protein [Oryza sativa (japonica cultivar-group)]

```
Identities = 21/49 (42%), Positives = 29/49 (59%), Gaps = 1/49 (2%)

Query  40   TGVNEFAIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLS  88
           TG N +   C ++ L+ G G G + L  D+ LLHG S  CQT+GN  L+
Sbjct  137  TGANRYYYLCLNDALAFGGG-GSFALCLDEDLLHGTSGSCQTFGNSCLA  184
```

gi|26983810|gb|AAN86157.1| unknown protein [Arabidopsis thaliana]
gi|30692556|ref|NP_195697.3| expressed protein [Arabidopsis thaliana]
Length = 394

Identities = 63/227 (27%), Positives = 95/227 (41%), Gaps = 46/227 (20%)

```
Query: 45  NKTTSERTIVEEGSLPPVRLNGYLPSTKNKLLTPEMCDEIRTLMPTRIQLYTEWNLLYSL 104
           N    E T+V +  +P +       S  + LL+ +    + T +P  +Q   +W LLYS
Sbjct: 187 NNVKEEVTVVVQAIIPEI-------SEPSLLLSEQSRRSLYTSLPALVQ-GRKWILLYST 238

Query: 105 EQHGSSLHSLYSNVAPDSKEFRRVGY-VLVIKDRKNGIFGAYSNEAFHPNEHRQYTGNGE 163
            +HG SL +LY       K     G  +LV+ DRK  +FG        P + ++Y G
Sbjct: 239 WRHGISLSTLYR------KSLLWPGLSLLVVGDRKGSVFGGLVEAPLIPTD-KKYQGTNS 291

Query: 164 CFLWKLDKVPDVNISXXXXXXXXXXXXXXXXXXXXRWRFSGYPYTGVNEFAIYCTSEFLS 223
            F++        N S                          Y  TG N F   C+ EFL+
Sbjct: 292 TFVF-------TNKSGQPTI----------------------YRPTGANRFYTLCSKEFLA 323

Query: 224 MGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVW 270
           +G G G + L  D  LL G S   +TYGN  L+ + + F +  +E+W
Sbjct: 324 LGGG-GRFALYLDSELLSGSSAYSETYGNSCLA-DSQDFDVKEVELW 368
```

APPENDIX H

PROTEIN SEQUENCE ALIGNMENTS AND CONSERVED DOMAINS OF RAD6 INTERACTORS

5L1-A6 protein sequence alignment (also an interactor of WR5)

gi|6325060|ref|NP_015128.1| Protein of unknown function required for normal levels of resistance to oxidative damage, null mutants are sensitive to hydrogen peroxide; member of a conserved family of proteins found in eukaryotes but not in prokaryotes; **Oxr1p** [Saccharomyces cerevisiae]
gi|2132227|pir||S65215 hypothetical protein YPL196w - yeast (Saccharomyces cerevisiae)

Identities = 105/105 (100%), Positives = 105/105 (100%)

```
Query: 2   LDKVPDVNISEKEESEQEGKEGKEEGDKEERWRFSGYPYTGVNEFAIYCTSEFLSMGAGD 181
           LDKVPDVNISEKEESEQEGKEGKEEGDKEERWRFSGYPYTGVNEFAIYCTSEFLSMGAGD
Sbjct: 169 LDKVPDVNISEKEESEQEGKEGKEEGDKEERWRFSGYPYTGVNEFAIYCTSEFLSMGAGD 228

Query: 182 GHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVWRVG 316
           GHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVWRVG
Sbjct: 229 GHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVWRVG 273
```

Figure H.1 OXR1 structure and conserved domains

gnl|CDD|3951 smart00584, TLDc, domain in TBC and LysM domain containing proteins;

CD-Length = 166 residues, 100.0% aligned

```
Score =  158 bits (402), Expect = 5e-40

Query:  73   NKLLTPEMCDEIRTLMPTRIQLYTEWNLLYSLEQHGSSLHSLYSNVAPDSKEFRRVGYVL  132
Sbjct:  1    SSILSVEILALINSLLPTRTRGYP-WTLLYSSSQHGYSLNTLYRKVEGYGP-----PTLL  54

Query:  133  VIKDRKNGIFGAYSNEAFHPNEHRQYTGNGECFLWKLDKVPDVNIS*EKEESEQEGKEGKE*  192
Sbjct:  55   IIKDTDGEVFGAYASQAWRVSPH--FYGTGESFLFQLN----------------------  90

Query:  193  *EGDKEE*RWRFSGYPYTGVN--EFAIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTY  250
Sbjct:  91   -------PKFVVYDATGKNKNYYYINGTPDSLPIGGGGGGFGLWIDEDLNHGSSSSCKTF  143

Query:  251  GNEVLSKEGKKFSIVALEVWRVG  273
Sbjct:  144  GNPPLLTKDEDFLILDIEVWGFG  166
```

Plant homologs

gi|51535088|dbj|BAD37677.1| **nucleolar protein-like** [Oryza sativa (japonica cultivar-group)]

```
Identities = 64/241 (26%), Positives = 106/241 (43%), Gaps = 49/241 (20%)

Query: 43  FRNKTTSE---RTIVEEGSLPPVRLNGYLPSTKNKLL----TPEM--CDEIRTLMPTRIQ 93
           FR K +      T+  E   P + L G   S  +  L     P M   + +RT++ T +
Sbjct: 139 FRQKPSYSIDGETVQTEYDAPGLELKGSKESASHDKLPAMSEPSMLLSETMRTVLYTSLP 198

Query: 94  LYTE---WNLLYSLEQHGSSLHSLYSNVAPDSKEFRRVGY-VLVIKDRKNGIFGAYSNEA 149
           +  +   W L+YS  +HG SL +LY       +     GY +L++ DRK  +FG
Sbjct: 199 VLVQGRNWMLVYSTWRHGISLSTLYR------RSMLCAGYSLLIVGDRKGAVFGGLVEAP 252

Query: 150 FHPNEHRQYTGNGECFLWKLDKVPDVNISXXXXXXXXXXXXXXXXXXXXRWRFSGYPYTG 209
             P   ++Y G   CF++        NI+                          Y  TG
Sbjct: 253 LQPLIKKKYQGTNNCFVF-------TNIAGRPVI----------------------YRPTG 284

Query: 210 VNEFAIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEV 269
            N +  +C++++L+MG G GH+ L  D  LL+G S+  +T+ N  LS+  ++F +  +E+
Sbjct: 285 ANNYFTFCSTDYLAMGGG-GHFALYLDGDLLNGSSSTSETFNNPCLSRS-REFEVKDVEL 342

Query: 270 W 270
           W
Sbjct: 343 W 343
```

gi|50912883|ref|XP_467849.1| **oxidation protection protein-like** [Oryza sativa (japonica cultivar-group)]
gi|46805939|dbj|BAD17233.1| oxidation protection protein-like [Oryza sativa (japonica cultivar-group)]
gi|46390139|dbj|BAD15574.1| oxidation protection protein-like [Oryza sativa (japonica cultivar-group)]

```
Identities = 50/177 (28%), Positives = 79/177 (44%), Gaps = 43/177 (24%)

Query: 98  WNLLYSLEQHGSSLHSLY--SNVAPDSKEFRRVGY-VLVIKDRKNGIFGAYSNEAFHPNE 154
           W LLYS  +HG SL +LY  S + P        GY +LV+ D++  +FG        P
Sbjct: 202 WVLLYSTWRHGISLSTLYRRSMLCP--------GYSLLVVGDKEGAVFGGLVEAPLQPTS 253


Query: 155 HRQYTGNGECFLW-KLDKVPDVNISXXXXXXXXXXXXXXXXXXXXRWRFSGYPYTGVNEF 213
```

```
            ++Y G+  CF++  L   P +                              Y  TG N +
Sbjct: 254 AKKYQGSNSCFVFTNLHSNPSI-----------------------------YRPTGANNY 284

Query: 214 AIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVW 270
              C++++L++G G GH+ L  D  LL G S+  +T+ N  LS     F++  +E+W
Sbjct: 285 FTVCSTDYLALGGG-GHFALYLDADLLSGSSSNSETFNNMCLS-HSPDFAVKDVELW 339
```

gi|50936597|ref|XP_477826.1| **oxidation resistance 1-like protein** [Oryza sativa (japonica cultivar-group)]
gi|34394846|dbj|BAC84293.1| oxidation resistance 1-like protein [Oryza sativa (japonica cultivar-group)]
gi|50508536|dbj|BAD30835.1| oxidation resistance 1-like protein [Oryza sativa (japonica cultivar-group)]

```
Identities = 21/49 (42%), Positives = 29/49 (59%), Gaps = 1/49 (2%)

Query  40   TGVNEFAIYCTSEFLSMGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLS  88
            TG N +   C ++ L+ G G G + L  D+ LLHG S  CQT+GN  L+
Sbjct  137  TGANRYYYLCLNDALAFGGG-GSFALCLDEDLLHGTSGSCQTFGNSCLA  184
```

gi|26983810|gb|AAN86157.1| unknown protein [Arabidopsis thaliana]
gi|30692556|ref|NP_195697.3| expressed protein [Arabidopsis thaliana]

```
Length = 394

Identities = 63/227 (27%), Positives = 95/227 (41%), Gaps = 46/227 (20%)

Query: 45  NKTTSERTIVEEGSLPPVRLNGYLPSTKNKLLTPEMCDEIRTLMPTRIQLYTEWNLLYSL 104
            N    E T+V +  +P +       S  + LL+ +    + T +P  +Q   +W LLYS
Sbjct: 187 NNVKEEVTVVVQAIIPEI-------SEPSLLLSEQSRRSLYTSLPALVQ-GRKWILLYST 238

Query: 105 EQHGSSLHSLYSNVAPDSKEFRRVGY-VLVIKDRKNGIFGAYSNEAFHPNEHRQYTGNGE 163
             +HG SL +LY       K     G  +LV+ DRK  +FG        P + ++Y G
Sbjct: 239 WRHGISLSTLYR------KSLLWPGLSLLVVGDRKGSVFGGLVEAPLIPTD-KKYQGTNS 291

Query: 164 CFLWKLDKVPDVNISXXXXXXXXXXXXXXXXXXXXRWRFSGYPYTGVNEFAIYCTSEFLS 223
             F++        N S                          Y  TG N F   C+ EFL+
Sbjct: 292 TFVF-------TNKSGQPTI----------------------YRPTGANRFYTLCSKEFLA 323

Query: 224 MGAGDGHYGLLCDDGLLHGVSNPCQTYGNEVLSKEGKKFSIVALEVW 270
            +G G G + L  D  LL G S   +TYGN  L+ + + F +  +E+W
Sbjct: 324 LGGG-GRFALYLDSELLSGSSAYSETYGNSCLA-DSQDFDVKEVELW 368
```

5K2-G4 protein sequence alignment

gi|6322383|ref|NP_012457.1| Protein of unknown function, has similarity to Pry1p and Pry2p and to the plant PR-1 class of pathogen related proteins; Pry3p [Saccharomyces cerevisiae]
gi|1078240|pir||S56032 probable membrane protein YJL078c - yeast (Saccharomyces cerevisiae)
gi|1352993|sp|P47033|PRY3_YEAST **PRY3** protein (Pathogen related in Sc 3)

```
Identities = 185/185 (100%), Positives = 185/185 (100%)

Query: 1   FTRTISTSCSTLNGASTQTSELTTSPMKTNTVVPASSFPSTTTTCLENDDTAFSSIYTEV 60
```

```
           FTRTISTSCSTLNGASTQTSELTTSPMKTNTVVPASSFPSTTTTCLENDDTAFSSIYTEV
Sbjct: 685 FTRTISTSCSTLNGASTQTSELTTSPMKTNTVVPASSFPSTTTTCLENDDTAFSSIYTEV 744

Query: 61  NAATIINPGETSSLASDFATSEKPNEPTSVKSTSNEGTSSTTTTYQQTVATLYAKPSSTS 120
           NAATIINPGETSSLASDFATSEKPNEPTSVKSTSNEGTSSTTTTYQQTVATLYAKPSSTS
Sbjct: 745 NAATIINPGETSSLASDFATSEKPNEPTSVKSTSNEGTSSTTTTYQQTVATLYAKPSSTS 804

Query: 121 LGARTTTGSNGRSTTSQQDGSAMHQPTSSIYTQLKEGTSTTAKLSAYEGAATPLSIFQCN 180
           LGARTTTGSNGRSTTSQQDGSAMHQPTSSIYTQLKEGTSTTAKLSAYEGAATPLSIFQCN
Sbjct: 805 LGARTTTGSNGRSTTSQQDGSAMHQPTSSIYTQLKEGTSTTAKLSAYEGAATPLSIFQCN 864


Query: 181 SLAGT 185
           SLAGT
Sbjct: 865 SLAGT 869
```

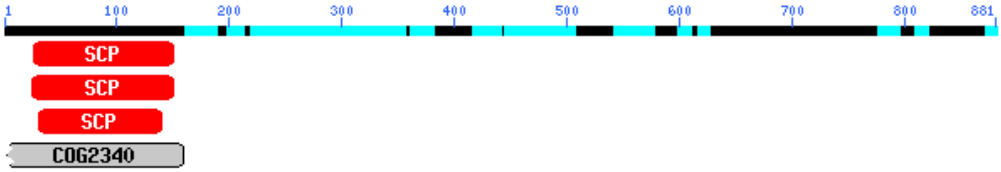


Figure H.2 PRY3 structure and conserved domains

gnl|CDD|5279 cd00168, SCP, SCP / Tpx-1 / Ag5 / PR-1 / Sc7 family of extracellular domains.; Human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein represent functional links between plant defense systems and human immune system. This family has no known function.

```
CD-Length = 152 residues,  98.0% aligned
Score =  137 bits (346), Expect = 5e-33

Query:  25   FESDVLNEHNKFRALHV-----------DTAPLTWSDTLATYAQNYADQYDCSGVLTHSD  73
Sbjct:  1    ERQDILNAHNKLRSRVAKGLEVNKPAASNMLKMTWDDELAASAQNYANQ--CPFGHSNGR  58

Query:  74   GPYGENLALGYTD-------TGAVDAWYGEISKYNYSNPGFSEST--GHFTQVVWKSTAE  124
Sbjct:  59   EGYGENLYWGSSSNPKDSYGSKAVDSWVDEKQDYNYNSNTCTTGKGIGHYTQVVWAKTYK  118

Query:  125  IGCGYKYCGTTWNN--YIVCSYNPPGNYLGE  153
Sbjct:  119  IGCGVVNCPDDNGTKTYVVCNYGPPGNYIGQ  149
```

gnl|CDD|14917 smart00198, SCP, SCP / Tpx-1 / Ag5 / PR-1 / Sc7 family of extracellular domains. Human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein represent functional links between plant defense systems and human immune system. This family has no known function.

```
CD-Length = 145 residues, 100.0% aligned
Score =  133 bits (336), Expect = 8e-32

Query:  24   NFESDVLNEHNKFRALHV--------DTAPLTWSDTLATYAQNYADQYDCSGVLTHSDG-  74
Sbjct:  1    NEQQEILDAHNKLRSQVAKGLYPPASNMLKLTWDCELAASAQNWANQ--C--PFGHSNPR  56

Query:  75   PYGENLALGY-------TDTGAVDAWYGEISKYNYSNPGFS--ESTGHFTQVVWKSTAEI  125
Sbjct:  57   GYGENLAWGSSSADPDTYASAAVKSWYDEFQDYGYTSNTCNPGKGIGHYTQMVWAKTYKV  116

Query:  126  GCGYKYCGTTWNN--YIVCSYNPPGNYLG  152
Sbjct:  117  GCGVSNCPDGTGNKVYVVCNYDPPGNYLG  145


gnl|CDD|25434 pfam00188, SCP, SCP-like extracellular protein. This domain is also found
in prokaryotes.
CD-Length = 131 residues,  99.2% aligned
Score =  113 bits (283), Expect = 1e-25

Query:  30   LNEHNKFRALHVDTAP-------LTWSDTLATYAQNYADQY-----DCSGVLTHSDGPYG  77
Sbjct:  1    LDAHNELRSQVAKGLPPASNMLKLKWDDELAALAQNWANQCPLAHSSDLCRNTLGFYYVG  60

Query:  78   ENLALGYTD-----TGAVDAWYGEISKYNYSNPGFSESTGHFTQVVWKSTAEIGCGYKYC  132
Sbjct:  61   ENLALSSGGNDLSGSDAVKAWFSEKKDYGY-NTNFCPGCGHYTQLVWDKTTKVGCAVARC  119

Query:  133  GTTWNN-YIVC  142
Sbjct:  120  GDGNRGTFVVC  130


gnl|CDD|12015 COG2340, COG2340, Uncharacterized protein with SCP/PR1 domains [Function
unknown].

CD-Length = 207 residues, only  72.9% aligned
Score =  104 bits (259), Expect = 5e-23

Query:  1    MLEFPISVLLGCLVAVKAQTTFPNFESDVLNEHNKFRALHVDTAPLTWSDTLATYAQNYA  60
Sbjct:  56   ATRFPTQATALSSSPVESQSTLAQFEKAVVAETNQERAKH-GLPPLAWNATLAKAARNHA  114

Query:  61   DQYDCSGVLTHSDGPYGENLALGYTDTGAVDAWYGEISKYNYSNPGFSESTGHFTQVVWK  120
Sbjct:  115  RDMAKNGYFSHT-SPTGETPADRLKKYGISGATAGENIAYGSNDP------PEAAVDGWL  167

Query:  121  STAEIGCGYKYCGTTWNNYIVCSYNPPGNYLGEFAEEVEPL  161
Sbjct:  168  NSP--GHRKNLLNPAYTEIGVGVAYDASNRYGVYWTQVFGL  206
```

5L1-A1 protein sequence alignment

```
gi|642810|emb|CAA88065.1|  Maf1p [Saccharomyces cerevisiae]
gi|6320208|ref|NP_010288.1| Negative regulator of RNA polymerase III; component of
several signaling pathways that repress polymerase III transcription in response to
changes in cellular environment; targets the initiation factor TFIIIB;
Maf1p[Saccharomyces cerevisiae]
gi|1170854|sp|P41910|MAF1_YEAST Repressor of RNA polymerase III transcription MAF1
gi|829377|gb|AAB51655.1|  Maf1p

Identities = 181/181 (100%), Positives = 181/181 (100%)

Query: 1   PSNLSLGPFGPINEPSSRKIFAYLIAILNASYPDHDFSSVEPTDFVKTSLKTFISKFENT 60
           PSNLSLGPFGPINEPSSRKIFAYLIAILNASYPDHDFSSVEPTDFVKTSLKTFISKFENT
Sbjct: 215 PSNLSLGPFGPINEPSSRKIFAYLIAILNASYPDHDFSSVEPTDFVKTSLKTFISKFENT 274
Query: 61  LYSLGRQPEEWVWEVINSHMTLSDCVLFQYSPSNSFLEDEPGYLWNLIGFLYNRKRKRVA 120
```

```
            LYSLGRQPEEWVWEVINSHMTLSDCVLFQYSPSNSFLEDEPGYLWNLIGFLYNRKRKRVA
Sbjct: 275 LYSLGRQPEEWVWEVINSHMTLSDCVLFQYSPSNSFLEDEPGYLWNLIGFLYNRKRKRVA 334

Query: 121 YLYLICSRLNSSTGEVEDALAKKPQGKLIIDDGSNEYEGEYDFTYDENVIDDKSDQEESL 180
            YLYLICSRLNSSTGEVEDALAKKPQGKLIIDDGSNEYEGEYDFTYDENVIDDKSDQEESL
Sbjct: 335 YLYLICSRLNSSTGEVEDALAKKPQGKLIIDDGSNEYEGEYDFTYDENVIDDKSDQEESL 394

Query: 181 Q 181
            Q
Sbjct: 395 Q 395
```

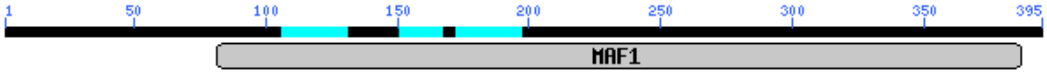


Figure H.3 MAF 1 structure and conserved domains

gnl|CDD|14176 COG5046, MAF1, Protein involved in Mod5 protein sorting [Posttranslational modification, protein turnover, chaperones].

CD-Length = 282 residues, 100.0% aligned
Score = 249 bits (636), Expect = 6e-67

```
Query:  81   SPFYSNRRDSNSFWEQKRRISFSEYNSNNNTNNSNGNSSNNNNYSGPNGSSPATFPKSAK  140
Sbjct:  1    SPFYSNRRDSNSFWEQKRRISFSEYNSNNNTNNSNGNSSNNNNYSGPNGSSPATFPKSAK  60

Query:  141  LNDQNLKELVSNYDSGSMSSSSLDSSSKNDERIRRRSSSSISSFKSGKSSNNNYSSGTAT  200
Sbjct:  61   LNDQNLKELMITIVVYQMEYRSTNSDKKLFKAIENRCQEDLFALSSSKSSEYAFSLTQQS  120

Query:  201  NNVNKRRKSSINERPSNLSLGPFGPINEPSSRKIFAYLIAILNASYPDHDFSSVEPTDFV  260
Sbjct:  121  ----------------------PFGPYLELSSRSKFNYLFAKLNASYPDHDFSSEAPTDFS  159

Query:  261  K-TSLKTFISKFENTLYSLGRQPE--EWVWEVINSHMTLSDCVLFQYSP-SNSFLEDEPG  316
Sbjct:  160  KRRSLGRFVEKPLGTLLSAGRLRFPYDFCWESIGSHRRLSDCEVLNYMPYSISTTHKNSD  219

Query:  317  YLWNLIGFLYNRKRKRVAYLYLICSRLNSSTGEVEDALAKKPQGKLIIDDGSNEYEGEYD  376
Sbjct:  220  DVWEFVGFLENILHKRVSYLGLEIFSYENRMGPFEDCLWYFS---FLFDDKK-----QRD  271

Query:  377  FTYDENVIDDK  387
Sbjct:  272  FVMDNAFMSDS  282
```

5L1-B1 protein sequence alignment

```
gi|6320931|ref|NP_011010.1|  Hypothetical ORF; Yer087wp [Saccharomyces cerevisiae]
gi|603325|gb|AAB64642.1|  Yer087wp [Saccharomyces cerevisiae]
gi|1077681|pir||S50590 hypothetical protein YER087w - yeast (Saccharomyces cerevisiae)
gi|731489|sp|P39965|SYPC_YEAST  Probable prolyl-tRNA synthetase, cytoplasmic (Proline--
tRNA ligase)(ProRS)

Identities = 188/188 (100%), Positives = 188/188 (100%)

Query: 1   DTYNKIFKDLKIPFVSAWADSGDIGGEFSKEFHLIHESGEDTLMSCKHCGDISTLDMSQS 60
           DTYNKIFKDLKIPFVSAWADSGDIGGEFSKEFHLIHESGEDTLMSCKHCGDISTLDMSQS
Sbjct: 185 DTYNKIFKDLKIPFVSAWADSGDIGGEFSKEFHLIHESGEDTLMSCKHCGDISTLDMSQS 244

Query: 61  YPEKDGQYSGDVDCKYALTKDHSTLICFYYPKDRQLNWNLALNAMDKDIDLTLRNKPNDH 120
           YPEKDGQYSGDVDCKYALTKDHSTLICFYYPKDRQLNWNLALNAMDKDIDLTLRNKPNDH
Sbjct: 245 YPEKDGQYSGDVDCKYALTKDHSTLICFYYPKDRQLNWNLALNAMDKDIDLTLRNKPNDH 304

Query: 121 VLQVYEKDNEDIMFSKILRVMDCRLNSKSNFPDFPLKKYLKNNFGQISDVSIVDAQENEI 180
           VLQVYEKDNEDIMFSKILRVMDCRLNSKSNFPDFPLKKYLKNNFGQISDVSIVDAQENEI
Sbjct: 305 VLQVYEKDNEDIMFSKILRVMDCRLNSKSNFPDFPLKKYLKNNFGQISDVSIVDAQENEI 364

Query: 181 CGKCEEGR 188
           CGKCEEGR
Sbjct: 365 CGKCEEGR 372
```

Figure H.4 YER087w structure and conserved domains

gnl|CDD|27677 cd00779, ProRS_core_prok, Prolyl-tRNA synthetase (ProRS) class II core catalytic domain. ProRS is a homodimer. It is responsible for the attachment of proline to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs in the core domain. This subfamily contains the core domain of ProRS from prokaryotes and from the mitochondria of eukaryotes.

CD-Length = 437 residues, 99.5% aligned
Score = 456 bits (1175), Expect = 3e-129

```
Query:  25   KSETTQELLQTVGFVRRSQVGLFQWLPLGLRSLNKVSNAIRNRMDSDGGAIEVSLSAISS  84
Sbjct:  2    AEIISHKLLLRAGFIRQTSSGLYSWLPLGLRVLKKIENIIREEMNKIG-AQEILMPILQP  60

Query:  85   KALWQATDRWNN--SELFKLKDSKGKQYCLTATCEEDITDLMKNYIASYKDMPITIYQMT  142
Sbjct:  61   AELWKESGRWDAYGPELLRLKDRHGKEFLLGPTHEEVITDLVANEIKSYKQLPLNLYQIQ  120

Query:  143  RKYRDEIRPRGGILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKIFKDLKIPFVSAW  202
Sbjct:  121  TKFRDEIRPRFGLMRGREFLMKDAYSFDIDEESLEETYEKMYQAYSRIFKRLGLPFVKVE  180

Query:  203  ADSGDIGGEFSKEFHLIHESGEDTLMSCKHCGDISTLDMSQSYPEKDGQYSGDVDCKYAL  262
Sbjct:  181  ADSGAIGGSLSHEFHVLSPIGEDTIVYCKSCGYAANIEKATSKPPKIDKNEEHVEEKKVT  240

Query:  263  TKDHSTLICFYYPKDRQLNWNLALNAMDKDIDLTLRNKPNDH------------------  304
Sbjct:  241  TPDTSTIEEVEFFNRVPKTKTKKIELYDKDKKLILVLIRGDIELDENKTKLKAADSPLAA  300

Query:  305  -------------VLQVYEKDNEDIMFSKILRVMDCRLNSKSNFPDFPLKKYLKNNFGQI  351
Sbjct:  301  AEEDKILNLTGLGFLFGGPLPNDVPVIADISVVMMKNFVIGANEDDHHYVNINGNRDFQI  360

Query:  352  SDVSIVDAQENEICGKCEEGRLEPLKSIEVGHIFLLGNKYSKPLNVKFVDKENKNETFVH  411
Sbjct:  361  DEVADLRLVKEGDCCPNCSGPLKITKGIEVGHIFQLGTKYSKALGATFLD-ENGKPKPLE  419

Query:  412  MGCYGIGVSRLVGAIAE  428
Sbjct:  420  MGCYGIGVSRLLAAIIE  436
```

gnl|CDD|27670 cd00772, ProRS_core, Prolyl-tRNA synthetase (ProRS) class II core catalytic domain. ProRS is a homodimer. It is responsible for the attachment of proline to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs in the core domain..

CD-Length = 455 residues, 99.3% aligned
Score = 181 bits (459), Expect = 2e-46

```
Query:  26   SETTQELLQTVGFVRRSQV-GLFQWLPLGLRSLNKVSNAIRNRMDSDGGAIEVSLSAISS  84
Sbjct:  3    SEKSLEHIGKAELADQGPGRGIINFLPLAKAILDKIENVLDKMF-KEHGAQNALFPFFIL  61

Query:  85   KALWQAT---DRWNNSELFKLKDSKGKQ----YCLTATCEEDITDLMKNYIASYKDMPIT  137
Sbjct:  62   ASFLEKEAEHDEGFSKELAVFKDAGDEELEEDFALRPTLEENIGEIAAKFIKSWKDLPQH  121

Query:  138  IYQMTRKYRDEIRPRGGILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKIFKDL-KI  196
Sbjct:  122  LNQIGNKFRDEIRPRFGFLRAREFIMKDGHSAHADAEEADEEFLNMLSAYAEIARDLAAI  181

Query:  197  PFVSAWADSGD--IGGEFSKEFHLIHESGEDTLMSCKHCGDISTLDMSQSYPEKDGQYSG  254
Sbjct:  182  DFIEGEADEGAKFAGASKSREFEALMECGEDTIVVCQNCDYAANIEIAKRSKRTEPLMSP  241

Query:  255  DVDCKY-----------------------------ALTKDHSTLICFYYPKDRQLNWNLA  285
Sbjct:  242  SALAKFPTPNTTSAPSVAEFFKTEPYFVLKALVNKVIHKDKETLACFFVRGDDNLEETKA  301

Query:  286  LNAMD--KDIDLTLRNKPNDHVLQV-----------YEKDNEDIMFSKILRVMDCRLNSK  332
Sbjct:  302  LNTLNLLGANALELREANEEDLNKAGLIAGFIGPYGLKKHVCYIIFDEDLKEGDC-LIVG  360
```

```
Query:  333  SNFPDFPLKKYLKNNFGQISDVSIVDAQENEICGKCEEGRLEPLKSIEVGHIFLLGNKYS  392
Sbjct:  361  ANEKDFHAVGVDLKGFENLVYADIVQVKESDCCPNCQ-GADGKAKQAETGHIFGEGFARA  419

Query:  393  KPLNVKFVDKENKNETFVHMGCYGIGVSRLVGAIAE  428
Sbjct:  420  FDLKAKFLDKDGK-EKFFEMGCWGIGISRFIGAIIE  454
```

gnl|CDD|27663 cd00670, Gly_His_Pro_Ser_Thr_tRS_core, Gly_His_Pro_Ser_Thr_tRNA synthetase class II core domain. This domain is the core catalytic domain of tRNA synthetases of the subgroup containing glycyl, histidyl, prolyl, seryl and threonyl tRNA synthetases. It is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. These enzymes belong to class II aminoacyl-tRNA synthetases (aaRS) based upon their structure and the presence of three characteristic sequence motifs in the core domain. This domain is also found at the C-terminus of eukaryotic GCN2 protein kinase and at the N-terminus of the ATP phosphoribosyltransferase accessory subunit, HisZ and the accessory subunit of mitochondrial polymerase gamma (Pol gamma b) . Most class II tRNA synthetases are dimers, with this subgroup consisting of mostly homodimers. These enzymes attach a specific amino acid to the 3' OH group of ribose of the appropriate tRNA..

```
CD-Length = 235 residues, only  66.4% aligned
Score = 76.9 bits (189), Expect = 7e-15

Query:  58   NKVSNAIRNRMDSDGGAIEVSLSAISSKALWQATDRWNN--SELFKLKDSKGKQ----YC  111
Sbjct:  6    RALERFLDDRMAEYGYQ-EILFPFLAPTVLFFKGGHLDGYRKEMYTFEDKGRELRDTDLV  64

Query:  112  LTATCEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRGGILRGREFLMKDAYSFAS  171
Sbjct:  65   LRPAACEPIYQIFSGEILSYRALPLRLDQIGPCFRHEPSGRRGLMRVREFRQVEYVVFGE  124

Query:  172  NEEDAFASFQKLDDTYNKIFKDLKIPFVSAWADSGDIG  209
Sbjct:  125  -PEEAEEERREWLELAEEIARELGLPVRVVVADDPFFG  161
```

gnl|CDD|27666 cd00768, class_II_aaRS-like_core, Class II tRNA amino-acyl synthetase-like catalytic core domain. Class II amino acyl-tRNA synthetases (aaRS) share a common fold and generally attach an amino acid to the 3' OH of ribose of the appropriate tRNA. PheRS is an exception in that it attaches the amino acid at the 2'-OH group, like class I aaRSs. These enzymes are usually homodimers. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. The substrate specificity of this reaction is further determined by additional domains. Intererestingly, this domain is also found is asparagine synthase A (AsnA), in the accessory subunit of mitochondrial polymerase gamma and in the bacterial ATP phosphoribosyltransferase regulatory subunit HisZ..

```
CD-Length = 211 residues, only  78.7% aligned
Score = 64.7 bits (157), Expect = 3e-11

Query:  56   SLNKVSNAIRNRMDSDGGAIEVSLSAISSKALWQATDRWNNSELFKLKDSKGKQYCLTAT  115
Sbjct:  1    IRSKIEQKLRRFMAELG-FQEVETPIVEREPLLEKAGHEPKDLLP-VGAENEEDLYLRPT  58

Query:  116  CEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRGGILRGREFLMKDAYSFASNEED  175
Sbjct:  59   LEPGLVRLFVSHI---RKLPLRLAEIGPAFRNEGGRRG-LRRVREFTQLEGEVFGEDGEE  114

Query:  176  AFASFQKLDDTYNKIFKDL-----KIPFVSAWADSGDIGGEFSKEFHLIHESG  223
Sbjct:  115  ASEFEELIELTE-ELLRALGIKLDIVFVEKTPGEFSPGGAGPGFEIEVDHPEG  166
```

gnl|CDD|27676 cd00778, ProRS_core_arch_euk, Prolyl-tRNA synthetase (ProRS) class II core catalytic domain. ProRS is a homodimer. It is responsible for the attachment of proline to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs in the core domain. This subfamily contains the core domain of ProRS from archaea, the cytoplasm of eukaryotes and some bacteria.

```
CD-Length = 261 residues, only  69.7% aligned
Score = 64.4 bits (157), Expect = 3e-11

Query:  26   SETTQELLQTVGFVRRSQV-GLFQWLPLGLRSLNKVSNAIRNRMDSDGGAIEVSLSAISS  84
Sbjct:  3    SEWYTEVITKAELIDYGPVKGCMVFRPYGYAIWENIQKILDKEIKETGHE-NVYFPLLIP  61

Query:  85   KALWQATD---RWNNSELFKLKDSKGKQ----YCLTATCEEDITDLMKNYIASYKDMPIT  137
Sbjct:  62   ESELEKEKEHIEGFAPEVAWVTHGGLEELEEPLALRPTSETAIYPMFSKWIRSYRDLPLK  121

Query:  138  IYQMTRKYRDEIRPRGGILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKIFKD-LKI  196
Sbjct:  122  INQWVNVFRWETKTTRPFLRTREFLWQEGHTAHATEEEAEEEVLQILDLYKEFYEDLLAI  181

Query:  197  PFV  199
Sbjct:  182  PVV  184
```

gnl|CDD|27676 cd00778, ProRS_core_arch_euk, Prolyl-tRNA synthetase (ProRS) class II core catalytic domain. ProRS is a homodimer. It is responsible for the attachment of proline to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs in the core domain. This subfamily contains the core domain of ProRS from archaea, the cytoplasm of eukaryotes and some bacteria.

```
CD-Length = 261 residues, only  18.4% aligned
Score = 39.0 bits (91), Expect = 0.002


Query:  377  KSIEVGHIFLLGNKYSKPLNVKFVDKENKNETFVHMGCYGIGVSRLVGAI  426
Sbjct:  211  RALQSGTSHNLGQNFSKAFDIKYQDKDGQKE-YVHQTSWGIS-TRLIGAI  258
```

gnl|CDD|27669 cd00771, ThrRS_core, Threonyl-tRNA synthetase (ThrRS) class II core catalytic domain. ThrRS is a homodimer. It is responsible for the attachment of threonine to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs in the core domain.

```
CD-Length = 298 residues, only  52.7% aligned
Score = 57.5 bits (139), Expect = 4e-09

Query:  45   GLFQWLPLGLRSLNKVSNAIRNRMDSDGGAIEVSLSAISSKALWQATDRWNN-SELFKLK  103
Sbjct:  21   GLPFWLPKGAIIRNELEDFLRELQRKRG-YQEVETPIIYNKELWETSGHWDHYRENMFPF  79

Query:  104  DSKGKQYCLT-ATCEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRG---GILRGR  159
Sbjct:  80   EEEDEEYGLKPMNCPGHCL-IFKSKPRSYRDLPLRLAEFGTVHRYE--QSGALHGLTRVR  136

Query:  160  EFLMKDAYSFASNE---EDAFASFQKLDDTYNKI-FKDLKI  196
Sbjct:  137  GFTQDDAHIFCTPDQIKEEIKGVLDLIKEVYSDFGFFDYKV  177
```

gnl|CDD|27668 cd00770, SerRS_core, Seryl-tRNA synthetase (SerRS) class II core catalytic domain. SerRS is responsible for the attachment of serine to the 3' OH group of ribose of the appropriate tRNA. This domain It is primarily responsible for ATP-dependent formation of the enzyme bound aminoacyl-adenylate. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs in the core domain. SerRS synthetase is a homodimer.

```
CD-Length = 297 residues, only  54.5% aligned
Score = 47.8 bits (114), Expect = 3e-06

Query:  95   NNSELFKLKDSKGKQYCLTATCEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRG-  153
Sbjct:  94   FDEQLYKVEGED---LYLIATAEVPLAALHRDEILEEEELPLKYAGYSPCFRKEAGSAGR  150
```

```
Query:  154  ---GILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKIFKDLKIPFVSAWADSGDIGG  210
Sbjct:  151  DTRGLFRVHQFEKVEQFVFTKPEE-SWEELEELISNAEEILQELGLPYRVVNICTGDLGF  209

Query:  211  EFSKEFHL------IHESGEDTLMSCKHCGDISTLDMSQSY-PEKDGQ  251
Sbjct:  210  AAAKKYDIEAWMPGQGKYRE--ISSCSNCTDFQARRLNIRYRDKKDGK  255
```

gnl|CDD|27672 cd00774, GlyRS-like_core, Glycyl-tRNA synthetase (GlyRS)-like class II core catalytic domain. GlyRS functions as a homodimer in eukaryotes, archaea and some bacteria and as a heterotetramer in the remainder of prokaryotes. It is responsible for the attachment of glycine to the 3' OH group of ribose of the appropriate tRNA. This domain is primarily responsible for ATP binding and hydrolysis. This alignment contains only sequences from the GlyRS form which homodimerizes. The heterotetramer glyQ is in a different family of class II aaRS. Class II assignment is based upon its structure and the presence of three characteristic sequence motifs. This domain is also found at the N-terminus of the accessory subunit of mitochondrial polymerase gamma (Pol gamma b). Pol gamma b stimulates processive DNA synthesis and is functional as a homodimer, which can associate with the catalytic subunit Pol gamma alpha to form a heterotrimer. Despite significant both structural and sequence similarity with GlyRS, Pol gamma b lacks conservation of several class II functional residues.

```
CD-Length = 327 residues, only  19.9% aligned
Score = 39.8 bits (93), Expect = 0.001

Query:  134  MPITIYQMTRKYRDEIRPRGGILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKIFKD  193
Sbjct:  179  LPFGVAQIGKSFRNEISPRNGLFRVREFTQAEIEFFVDPEKS----HPWFDY-----WAD  229

Query:  194  LKIPFVSAWADSGD  207
Sbjct:  230  QRLKWLPKFAQSPE  243
```

gnl|CDD|25560 pfam00587, tRNA-synt_2b, tRNA synthetase class II core domain (G, H, P, S and T). Other tRNA synthetase sub-families are too dissimilar to be included. This domain is the core catalytic domain of tRNA synthetases and includes glycyl, histidyl, prolyl, seryl and threonyl tRNA synthetases.

```
CD-Length = 157 residues, 100.0% aligned
Score =  132 bits (333), Expect = 1e-31

Query:  45   GLFQWLPLGLRSLNKVSNAIRNRMDSDGGAIEVSLSAISSKALWQATDRWNN---SELFK  101
Sbjct:  1    GFYDLLPLGARLRRALENFIRDLFKRYG-YQEVDTPILEPEELWKKSGHLDKFVEKELYR  59

Query:  102  LKDSKGKQYCLTATCEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRGGILRGREF  161
Sbjct:  60   FDDEGREELYLRPTAEVGIARLFKNEILSYRDLPLKLYQIGPVFRYEASPRRGLGRVREF  119

Query:  162  L-MKDAYSFASNEEDAFASFQKLDDTYNKIFKDLKIPF  198
Sbjct:  120  TQVEDEHIFHPDEEQSEEADEELLLLAEEILQALGLPY  157
```

gnl|CDD|10316 COG0442, ProS, Prolyl-tRNA synthetase [Translation, ribosomal structure and biogenesis].

```
CD-Length = 500 residues,  97.0% aligned
Score =  392 bits (1008), Expect = 8e-110

Query:  9    RSCHIFHPKSLSNNTLKSETTQELLQTVGFVRRSQVGLFQWLPLGLRSLNKVSNAIRNRM  68
Sbjct:  2    RMSKLFLPTLKEKPADASEWSHQLLLRAGMIRKPVKGLYVWLPLGLRVLEKIENIIREEM  61

Query:  69   DSDGGAIEVSLSAISSKALWQATDRWN--NSELFKLKDSKGKQYCLTATCEEDITDLMKN  126
Sbjct:  62   DKIG-AQEVLFPTLIPAELWKESGRWEGFGPELFRVKDRGDRPLALRPTSEEVITDMFRK  120

Query:  127  YIASYKDMPITIYQMTRKYRDEIRPRGGILRGREFLMKDAYSFASNEEDAFASFQKLDDT  186
Sbjct:  121  WIRSYKDLPLKLYQIQSKFRDEKRPRFGLLRGREFLMKDAYSFHADEEDAEETYEKMLDA  180
```

```
Query:  187  YNKIFKDLKIPFVSAWADSGDIGGEFSKEFHLIH-ESGEDTLMSCKHCGDISTLDMSQSY  245
Sbjct:  181  YSRIFLRLPLIFGPVPADEGFIGGSYSHEFEALMPDGGEDQIATSHHYGA----NFEKAF  236

Query:  246  PEKDGQYSGDVDCKYALTKDHSTLICFYYPKDRQLNWNLALNAMDKDIDLTLRNKPNDHV  305
Sbjct:  237  IDIKFEDEEEGELEYVHTT------------------SYGISTRIIGAAILIHGDNEGLV  278

Query:  306  LQVYEKDNEDIMFSKILRVMDCRLNSKSNFPDFPLKKYLKNNFGQISDVSIVDAQENEIC  365
Sbjct:  279  LP------PIVADIQVVIVPIFIKGANE---HYKVVNYGRDVAEPLEKLGIR-VEGDDRS  328

Query:  366  GKCEEGRLEPLKSIEVGHIFLLGNKYSKPLNVKFVDKENKNETFVHMGCYGIGVSRLVGA  425
Sbjct:  329  PDGPGFKLNIWEGIEVGHIFELGTKYSEAMNATVLD-RDGKEQPKTMGCYGIGVSRLVAA  387

Query:  426  IAELGRDSNGFRWPAIMAPYKVSICTGPNNPENSQRLQDVKSELLNDPTMQNLQNDILDQ  485
Sbjct:  388  LLEQIHDENGIIWPKAIAPFDVHIV---PVNTKDFKQAEAAEKLYVELPWCGTVEVLLDD  444

Query:  486  FNEKLGIGARIKLSHAMGIPLCVIVGSKSWPNVEIEVRGIRWGEK  530
Sbjct:  445  RDER--PGVKFADADLIGIPLRIVVG-KRLAEGEVEVKCRKCGEK  486
```

gnl|CDD|10315 COG0441, ThrS, Threonyl-tRNA synthetase [Translation, ribosomal structure and biogenesis].

```
CD-Length = 589 residues, only  27.8% aligned
Score = 53.7 bits (129), Expect = 6e-08

Query:  38   FVRRSQVGLFQWLPLGLRSLNKVSNAIRNRMDSDGGAIEVSLSAISSKALWQATDRWNN-  96
Sbjct:  204  FSPEEGPGLPFWHPKGATIRNLLEDYVRTKLRSYG-YQEVKTPVLADLELWELSGHWDNY  262

Query:  97   SELFKLKDSKGKQYCL-TATCEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRG--  153
Sbjct:  263  KEDMFLTESDDREYALKPMNCPGHIL-IFKSGLRSYRELPLRLAEFGYVYRYE--KSGAL  319

Query:  154  -GILRGREFLMKDAYSFASNE---EDAFASFQKLDDTYNKI-FKDLKI  196
Sbjct:  320  HGLMRVRGFTQDDAHIFCTPDQIKDEFKGILELILEVYKDFGFTDYEV  367
```

gnl|CDD|10315 COG0441, ThrS, Threonyl-tRNA synthetase [Translation, ribosomal structure and biogenesis].

```
CD-Length = 589 residues, only  23.9% aligned
Score = 37.9 bits (88), Expect = 0.003

Query:  371  GRLEPLKSIEVGHIFLLGNKYSKPLNVKFVDKENKNETFV--HMGCYGIGVSRLVGAIAE  428
Sbjct:  423  GREWQLGTIQL--DFNLPER----FDLEYVDEDGEKKRPVIIHRAILG-SIERFIGILLE  475

Query:  429  LGRDSNGFRWPAIMAPYKVSICTGPNNPENSQRLQDVKSELLNDPTMQNLQNDILDQFNE  488
Sbjct:  476  ----HYAGALPTWLAPVQVRVI--PVADEHLDYAKEVAEKLRK----AGIRVDI-DDRNE  524

Query:  489  KLGIGARIKLSHAMGIPLCVIVGSKSWPNVEIEVRGIRWGEK  530
Sbjct:  525  KLG--KKIREAGTQKIPYVIVVGDKEVETGTVVVR-RRGGKQ  563
```

gnl|CDD|10047 COG0172, SerS, Seryl-tRNA synthetase [Translation, ribosomal structure and biogenesis].

```
CD-Length = 429 residues, only  33.1% aligned
Score = 45.2 bits (107), Expect = 2e-05

Query:  98   ELFKLKDSKGKQYCLTATCEEDITDLMKNYIASYKDMPITIYQMTRKYRDEIRPRG----  153
Sbjct:  219  DLYKVEDPD---LYLIPTAEVPLTNLHRDEILDEEDLPIKYTAYSPCFRSEAGSAGKDTR  275

Query:  154  GILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKIFKDLKIPFVSAWADSGDIGGEFS  213
Sbjct:  276  GLIRVHQFDKVELVVITK-PEESEEELEEMLGNAEEVLQELELPYRVVNLCTGDLGFSAA  334

Query:  214  K----EFHLIHESGEDTLMSCKHCGD  235
Sbjct:  335  KKYDLEVWLPGQNKYREISSCSNCTD  360
```

gnl|CDD|10297 COG0423, GRS1, Glycyl-tRNA synthetase (class II) [Translation, ribosomal structure and biogenesis].

CD-Length = 558 residues, only 10.2% aligned
Score = 44.1 bits (104), Expect = 4e-05

```
Query:  134  MPITIYQMTRKYRDEIRPRGGILRGREFLMKDAYSFASNEEDAFASFQKLDDTYNKI  190
Sbjct:  186  LPFGIAQIGKSFRNEISPRNGLFRTREFEQAEIEFFVDPEEKEHPKFNEVKDEKLPL  242
```

5K1-A1 protein sequence alignment

gi|1015579|emb|CAA89452.1| **FAR1** [Saccharomyces cerevisiae]
gi|6322304|ref|NP_012378.1| Cyclin-dependent kinase inhibitor that mediates cell cycle arrest in response to pheromone; also forms a complex with Cdc24p, Ste4p, and Ste18p that may specify the direction of polarized growth during mating; potential Cdc28p substrate; Far1p [Saccharomyces cerevisiae]
gi|1169610|sp|P21268|FAR1_YEAST Cyclin-dependent kinase inhibitor FAR1 (CKI FAR1) (Factor arrest protein)

Identities = 180/180 (100%), Positives = 180/180 (100%)

```
Query: 1   YLDSPFLNSPFVNKMATTDPFDLSDDEKLDCDDEIDESAAEVWFSKTGGEHVMVSVKFQE 60
           YLDSPFLNSPFVNKMATTDPFDLSDDEKLDCDDEIDESAAEVWFSKTGGEHVMVSVKFQE
Sbjct: 338 YLDSPFLNSPFVNKMATTDPFDLSDDEKLDCDDEIDESAAEVWFSKTGGEHVMVSVKFQE 397

Query: 61  MRTSDDLGVLQDVNHVDHEELEEREKEWKKKIDQYIETNVDKDSEFGSLILFDKLMYSDD 120
           MRTSDDLGVLQDVNHVDHEELEEREKEWKKKIDQYIETNVDKDSEFGSLILFDKLMYSDD
Sbjct: 398 MRTSDDLGVLQDVNHVDHEELEEREKEWKKKIDQYIETNVDKDSEFGSLILFDKLMYSDD 457

Query: 121 GEQWVDNNLVILFSKFLVLFDFEEMKILGKIPRDQFYQVIKFNEDVLLCSLKSTNIPEIY 180
           GEQWVDNNLVILFSKFLVLFDFEEMKILGKIPRDQFYQVIKFNEDVLLCSLKSTNIPEIY
Sbjct: 458 GEQWVDNNLVILFSKFLVLFDFEEMKILGKIPRDQFYQVIKFNEDVLLCSLKSTNIPEIY 517
```

5K2-D2 and 5K2-E3/1 protein sequence alignment (cosmid 8082 in yeast)

gi|23481569|gb|EAA17807.1| hypothetical protein [Plasmodium yoelii yoelii]

Identities = 36/138 (26%), Positives = 53/138 (38%), Gaps = 15/138 (10%)

```
Query  565  GPYLN*TIV--SLARTTFLSSAKGTPMQKHIRNEHKYTVCVFTLLA*QQTVY*LQRY*SY  392
            GP LN +++  SL        +       H   ++KY    F  +      Y ++
Sbjct  434  GPLLNSSLLNSSLFDRPLFDKSLVAQFDSHKNGQNKYLENEFLFICLNNKTYKIRIEKLI  493

Query  391  KKFL-KCRILMIISSNNNKVKKLVTFLPIFNFDQPYG------------IPYFFNTTHSV  251
              F  K  IL + S +N    K  T L IF   Q Y             I YF  T+H+
Sbjct  494  PNFFHKKNILNLDSEHNIYTNKTETDLLIFFKSQNYKKNTILKKYNHSIINYFNKTSHNY  553

Query  250  VSQNEIRQYDNTSS*TFI  197
            V +N+  Q+DN    TF+
Sbjct  554  VDENKNEQFDNDECNTFL  571
```

5K2-E3/2

gi|1302363|emb|CAA96200.1| unnamed protein product [Saccharomyces cerevisiae]
gi|1730648|sp|P53831|YN25_YEAST Hypothetical 14.1 kDa protein in SEC21-MRPL10 intergenic region

Identities = 117/117 (100%), Positives = 117/117 (100%), Gaps = 0/117 (0%)

```
Query  77   MTSQYKINVDIILENASNAIKKFERYENRTRVPTLEGWDDNHYTNRPCVGWCNCNDEVIS  256
            MTSQYKINVDIILENASNAIKKFERYENRTRVPTLEGWDDNHYTNRPCVGWCNCNDEVIS
Sbjct  1    MTSQYKINVDIILENASNAIKKFERYENRTRVPTLEGWDDNHYTNRPCVGWCNCNDEVIS  60

Query  257  DGDYITTETQRLLKRVITTAILIVGINQTSSDNVGTRVINSDMCCYGNNLMLTSYIN  427
            DGDYITTETQRLLKRVITTAILIVGINQTSSDNVGTRVINSDMCCYGNNLMLTSYIN
Sbjct  61   DGDYITTETQRLLKRVITTAILIVGINQTSSDNVGTRVINSDMCCYGNNLMLTSYIN  117
```

5K2-G3

gi|6322962|ref|NP_013034.1| hypothetical protein; Yll066cp [Saccharomyces cerevisiae]
gi|1360282|emb|CAA97520.1| unnamed protein product [Saccharomyces cerevisiae]
gi|642317|emb|CAA87990.1| ORF L0519 [Saccharomyces cerevisiae]

Identities = 143/143 (100%), Positives = 143/143 (100%), Gaps = 0/143 (0%)

```
Query  2    GQHVGCCGSRTDLSADTVELIERMDRLAEKQATASMSIVALPSSFQESNSSDRCRKYCSS  181
            GQHVGCCGSRTDLSADTVELIERMDRLAEKQATASMSIVALPSSFQESNSSDRCRKYCSS
Sbjct  592  GQHVGCCGSRTDLSADTVELIERMDRLAEKQATASMSIVALPSSFQESNSSDRCRKYCSS  651

Query  182  DEDSNTCIHGSANASTNATTNSSTNATTTASTNVRTSATTTASINVRTSATTTESTNSST  361
            DEDSNTCIHGSANASTNATTNSSTNATTTASTNVRTSATTTASINVRTSATTTESTNSST
Sbjct  652  DEDSNTCIHGSANASTNATTNSSTNATTTASTNVRTSATTTASINVRTSATTTESTNSST  711

Query  362  NATTTASTNVRTSATTTASINVR  430
            NATTTASTNVRTSATTTASINVR
Sbjct  712  NATTTASTNVRTSATTTASINVR  734
```

gi|6325462|ref|NP_015530.1| subtelomerically-encoded DNA helicase; Ypr204wp [Saccharomyces cerevisiae]
gi|1370587|emb|CAA97896.1| unnamed protein product [Saccharomyces cerevisiae]
gi|2133041|pir||S65341 probable membrane protein YPR204w - yeast (Saccharomyces cerevisiae) Length=1032

Identities = 143/143 (100%), Positives = 143/143 (100%), Gaps = 0/143 (0%)

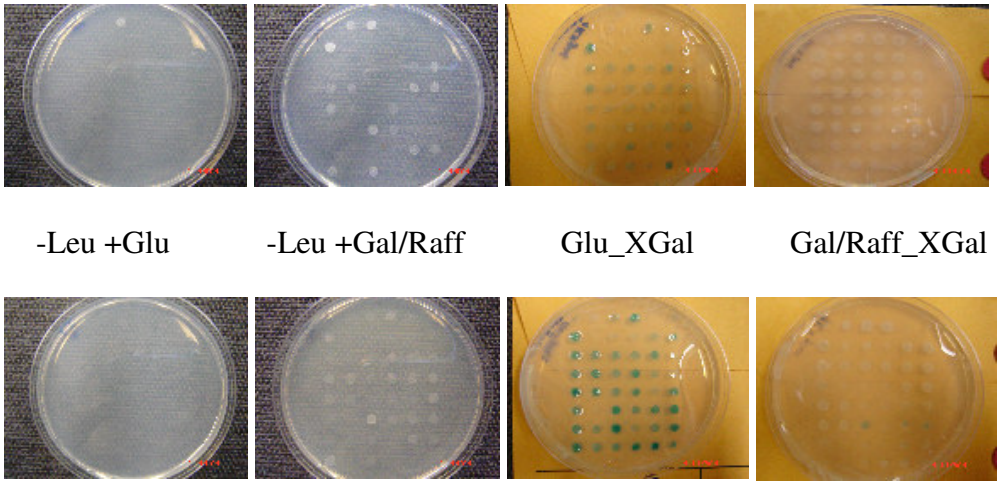
```
Query  2    GQHVGCCGSRTDLSADTVELIERMDRLAEKQATASMSIVALPSSFQESNSSDRCRKYCSS  181
            GQHVGCCGSRTDLSADTVELIERMDRLAEKQATASMSIVALPSSFQESNSSDRCRKYCSS
Sbjct  384  GQHVGCCGSRTDLSADTVELIERMDRLAEKQATASMSIVALPSSFQESNSSDRCRKYCSS  443

Query  182  DEDSNTCIHGANASTNATTNSSTNATTTASTNVRTSATTTASINVRTSATTTESTNSST  361
            DEDSNTCIHGSANASTNATTNSSTNATTTASTNVRTSATTTASINVRTSATTTESTNSST
Sbjct  444  DEDSNTCIHGSANASTNATTNSSTNATTTASTNVRTSATTTASINVRTSATTTESTNSST  503
```

```
Query  362  NATTTASTNVRTSATTTASINVR  430
            NATTTASTNVRTSATTTASINVR
Sbjct  504  NATTTASTNVRTSATTTASINVR  526
```

APPENDIX I

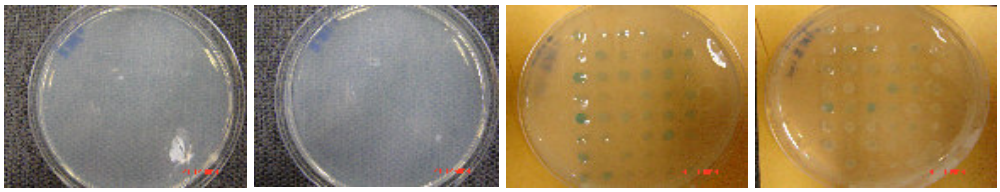
PLATES



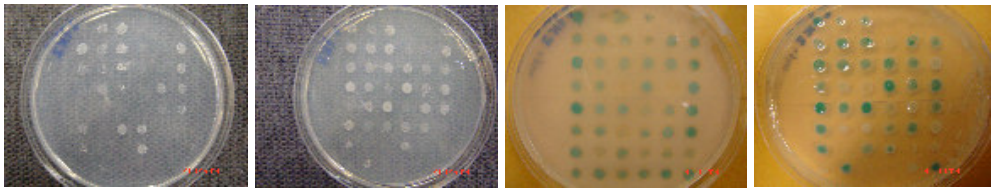
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.1 1K1 plate results

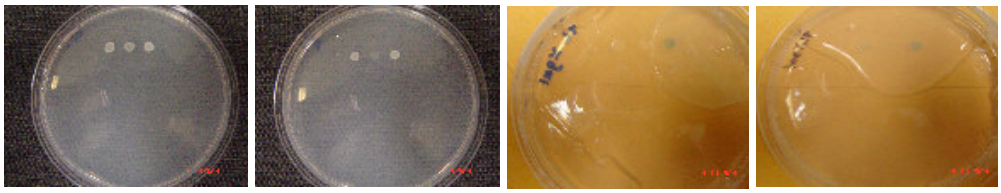


-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

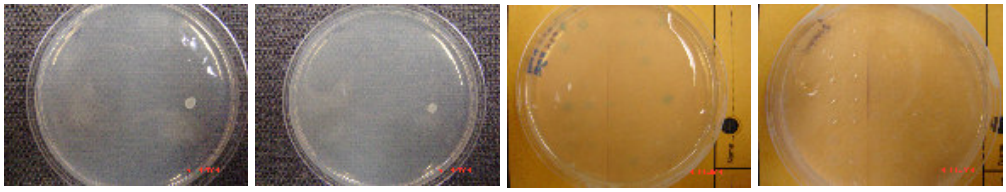
Figure I.2 1K2 plate results



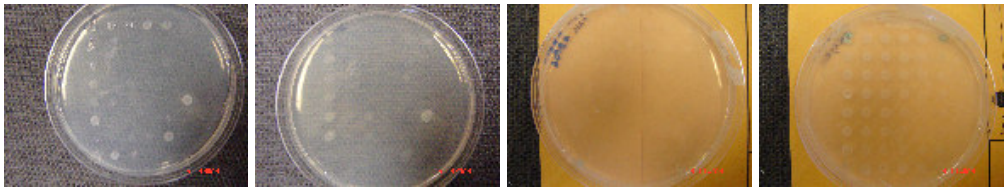
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.3 1L1 plate results

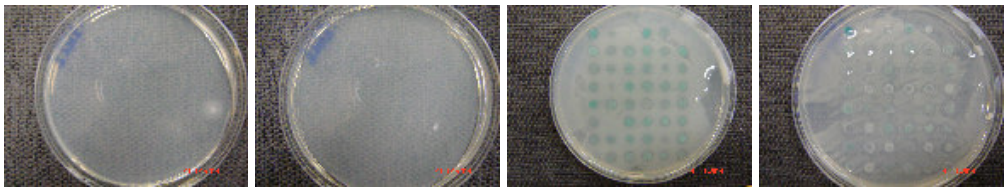


-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

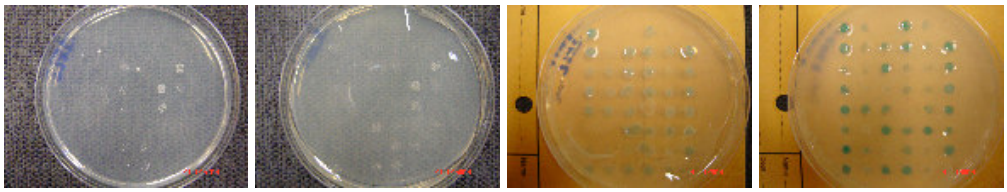


-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.4 2K1 plate results

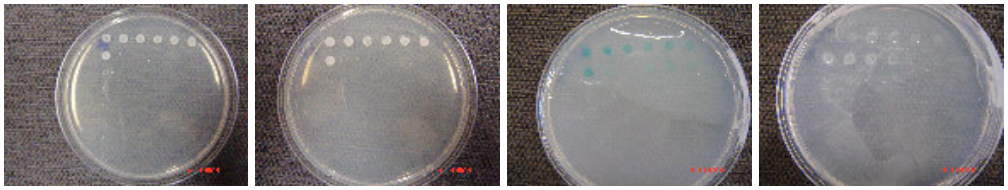


-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

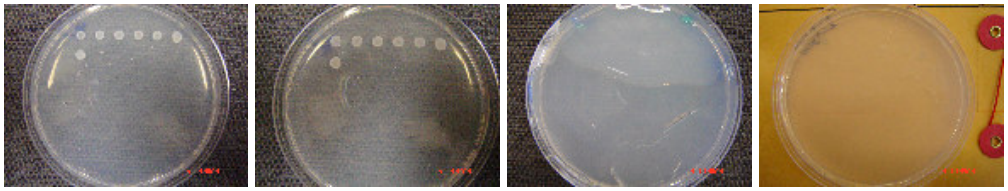


-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.5 2KL2 plate results



-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

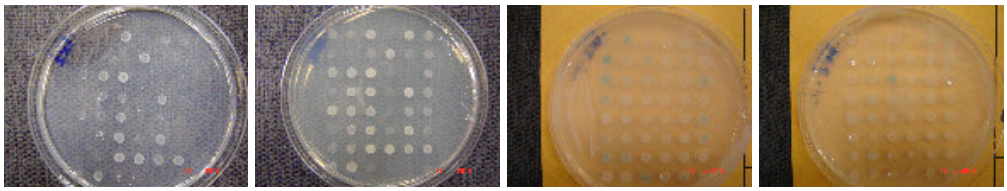
Figure I.6 2L1 plate results

-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

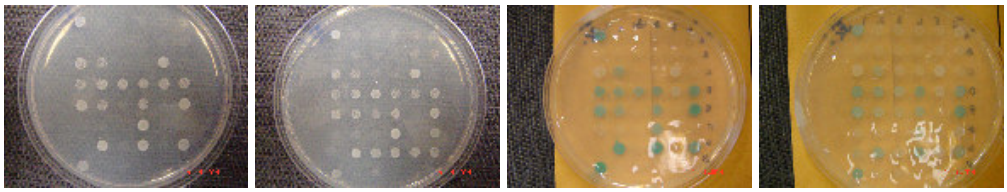
Figure I.7 3K1 plate results

-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

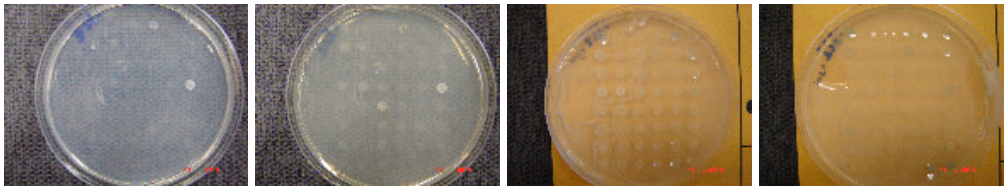
Figure I.8 3K2 plate results



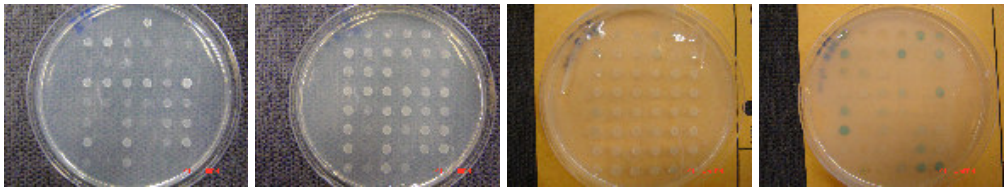
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.9 3L1 plate results



-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



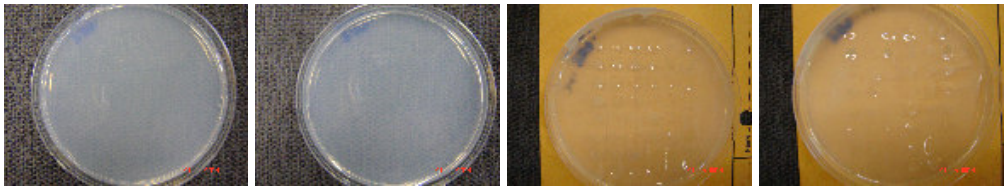
-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.10 3L2 plate results

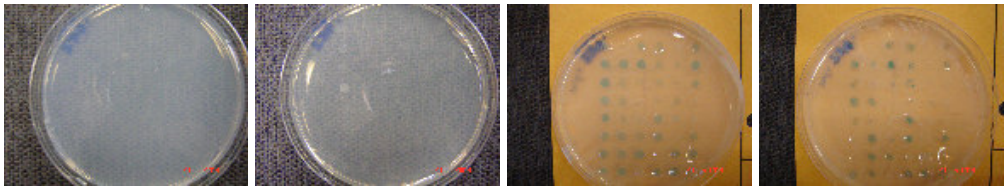
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.11 4K1 plate results

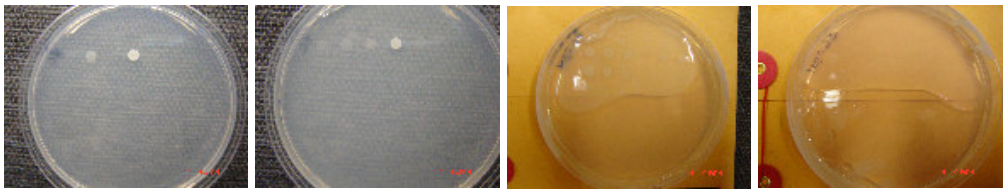


-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

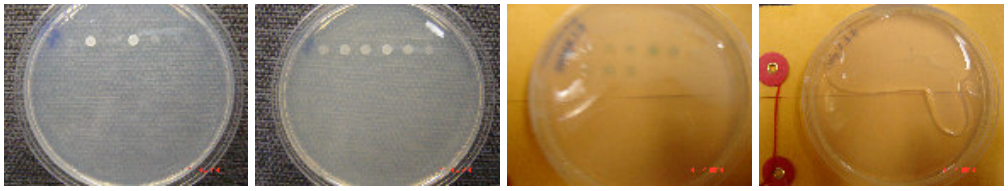


-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.12 4K2 plate results



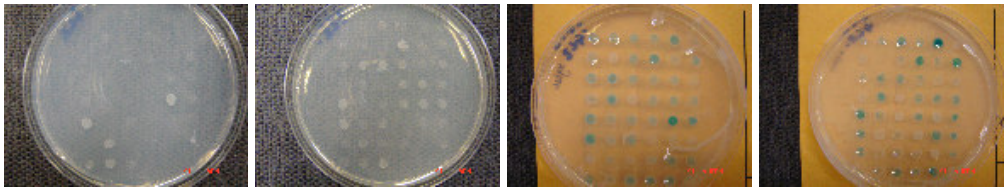
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.13 4L1 plate results

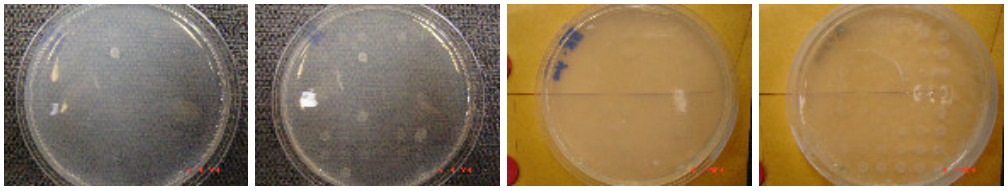
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.14 4L2 plate results

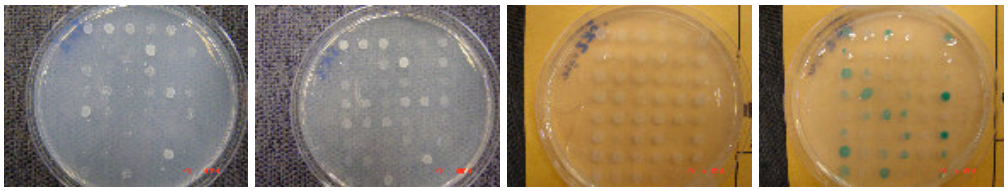
-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

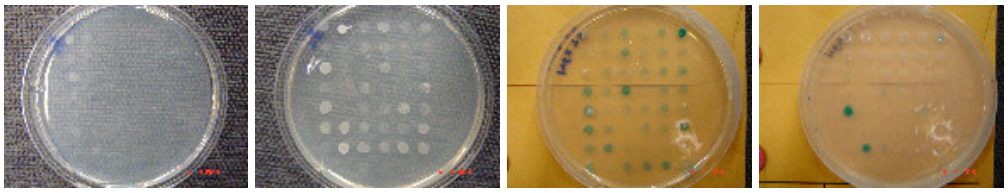
Figure I.15 5K1 plate results

-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

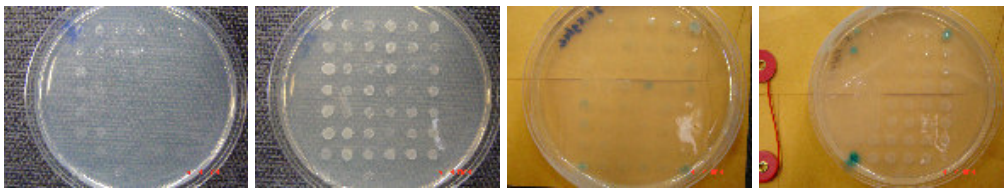


-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.16 5K2 plate results

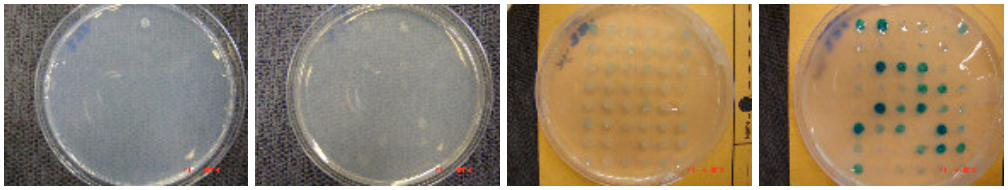


-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal

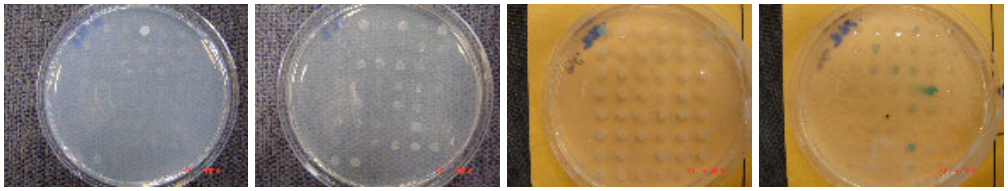


-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.17 5L1 plate results



-Leu +Glu -Leu +Gal/Raff Glu_XGal Gal/Raff_XGal



-Lys+Glu -Lys+ Gal/Raff Glu_XGluc Gal/Raff_XGluc

Figure I.18 5L2 plate results

APPENDIX J

GROWTH MEDIA FOR YEAST

YPD Medium

20 g of peptone (Q.BIOgene), 10g of yeast extract (Q.BIOgene), 20g of glucose (AppliChem) and 20g of agar (LAB M) (if for plates) are dissolved in 1L of distilled water and autoclaved.

Yeast complete dropout powder

All amino acids are obtained from Sigma Aldrich. 1.2g Uracil (not present in –Ura mediums), 1.2g Arginine, 1.2g Histidine (not present in –His mediums), 1.2g Methionine, 1.8g Tyrosine, 1.8g Isoleucine, 1.8g Lysine (not present in –Lys mediums), 2.4 g Tryptophane (not present in –Trp mediums), 2.5g Adenine, 3.0g Phenylalenine, 3.6g Leucine (not present in –Leu mediums), 6.0g Glutamic acid, 6.0g Aspartic acid, 9.0g Valine, 12.0 g Threonine and 22.0g Serine is mixed to make the yeast dropout powder.

Yeast dropout medium

1.5 g dropout powder, 6.7 g yeast nitrogen base without aminoacids (Q.BIOgene), 20g glucose (AppliChem) (for Glu/CM mediums) or 20g galactose (AppliChem) and 10g raffinose (Q.BIOgene) (for Gal-Raff CM mediums) and 20g agar (LAB M) (for plates) are dissolved in 1L distilled water and autoclaved.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Yıldırım, Figen
Nationality: Turkish (TC)
Date and Place of Birth: 30 April 1977 , Elazığ
Marital Status: Single
Phone: +90 312 210 31 97
Fax: +90 312 210 12 80
email: figeny@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Biotechnology	2001
BS	METU Biology	1999
High School	TED Ankara College	1995

WORK EXPERIENCE

Year	Place	Enrollment
1999-Present	METU Graduate School of Natural & Applied Sciences	Research Assistant
1998 June	Sevgi Hospital, Virology and Microbiology Department	Intern student

FOREIGN LANGUAGES

Advanced English

PUBLICATIONS

1. Yildirim, F**.** and Akkaya, M.S., ‘’DNA fingerprinting and genetic characterization of Anatolian *Triticum* sp. using AFLP markers’’ Manuscript accepted by *Genetic Resources and Crop Evolution*, (2005)

2. Yildirim, A., Karadag, Y., Sakin, M.A., Gokmen, S., Kandemir, N., Akkaya, M.S., Yildirim, F. "Transfer of the stripe rust resistance gene *Yr26*, into wheat by using microsatellite markers" *Cereal Research Communications* , 32, 25-30 (2004)

3. Yildirim, A. Sakin, M.A., Karadag, Y., Gokmen, S., Kandemir, N., Akkaya, M.S., Yildirim, F. "Genetic Marker mediated transfer of an alien gene*, Pm21*, into wheat conferring resistance to powdery mildew" *Biotechnol. & Biotechnol. Eq.*, 18,15-19 (2004)

POSTER PRESENTATIONS

1. Mahinur S. Akkaya, Erdogan E. Hakki, Figen Yildirim, Elif Uz , “A novel isolation method OF SSRs: NO CLONING, NO SCREENING” Plant & Animal Genome IX Conference, San Diego, USA 13-17, January, 2001. Page 304

2. Akkaya MS, Hakki EE, Bilgic H, Buyukunal EB, Yildirim F, Uz E, Bozkurt OT, Barbaros Y, “Using molecular markers for the genetic relationship analysis of Turkish wheat cultivars, A novel method for isolation of microstallites and investigations on resistance genes.” 12. National Biotechnology Congress, 17-21 September, 2001, Ayvalık, Balıkesir,TURKEY. Page 103

3. Yildirim F, Akkaya MS, “DNA fingerprinting and genetic characterization of Anatolian *Triticum* sp. using AFLP markers.” 12. National Biotechnology Congress, 17-21 Eylül, 2001, Ayvalık, Balıkesir, TURKEY. PA40

4. Mahinur S. Akkaya, Figen Yildirim, Osman Bozkurt, Mine Turktas, Yasemin Aktas, Mehmet Somel, Turgay Unver, Adnan El-Asbhi “Strategies for the determination of genes involved in disease resistance and zinc uptake in plants” 13. National Biotechnology Congress, 25-29 August, 2003, Canakkale, TURKEY.

5. Mahinur S. Akkaya, Xianming Chen, Osman Tolga Bozkurt, Figen Yildirim, Turgay Uver, Mehmet Somel “Isolation of RGAs and disease related gene fragments from wheat stripe rust resistant differential lines” 11th Internatioal Cereal Rusts and Powdery Mildews Conference, 22-27 August, 2004, Norwich, UK, Poster A2.1.

6. Mahinur S.Akkaya, Osman Bozkurt, Turgay Unver, Figen Yildirim "Determination of differentially expressed wheat genes upon infection with avirulent races of *Puccinia striiformis*" 3rd Plant Genomics European Meetings, 22-25 September, 2004, Lyon, FRANCE, P 173.

PRACTICAL COURSE:

"Expression of heterelogous genes in bacteria and yeast" 28 October- 8 November 2002, ICGEB; New Delhi, INDIA

VISITING SCIENTIST

" Yeast two-hybrid system" 2 August- 3 November 2004, Golemis Lab., Fox Chase Cancer Center, Philadelphia, USA

PROGRAMMING SKILLS

MS Office (Word, Excel, Power Point), Popgen32 vs., SPSS for windows, Ntsys, Codehop, Primer detective, Clustal, Blast.