Generation of RP Variant Clones for Yeast 2 Hybrid Studies

A THESIS

Submitted by

Akhilesh Kesavan

BS14B003

for the partial fulfillment of the requirements for the award of the degree of

DUAL DEGREE (B.TECH & M.TECH)

in

BIOLOGICAL SCIENCES

DEPARTMENT OF BIOTECHNOLOGY

Bhupat And Jyoti Mehta School Of Biosciences

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

June 2020

THESIS CERTIFICATE

This is to certify that the thesis titled **Generation of RP Variant Clones for Yeast 2 Hybrid Studies** submitted by **Akhilesh Kesavan (BS14B003)**, to the Indian Institute of Technology, Madras, for the award of the degree of **Dual Degree (B.TECH & M.TECH) in Biological Sciences**, is a bonafide record of the research work done by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Himanshu Sinha
Research guide
Associate Professor
Dept. of Biotechnology
IIT Madras, 600 036

Acknowledgments

I thank my guide Prof. Himanshu Sinha, Department of Biotechnology, IIT Madras for providing guidance, assistance, and the opportunity to work on this project. I would also like to thank the Initiative for Biological Systems Engineering (IBSE) for providing me a platform to present my work as well as exposure to the scientific community during my time working on the project.

I thank the members of our lab, the Systems Genetics Lab; Purnima for the motivation and drive to actually do my work; Viswa for his immensely vast global knowledge and vocabulary; Balakumaran for teaching me that the time of the day should not affect your work; Ashley for being a friend I could discuss things I normally can't with others; Ramya and Aditya for being the entertaining people that they always are; Sai Guha for the wondrous example of automation; Nikhita, Sakshi, Nayana, Indu and Srijith for providing a positive research environment.

I also thank Anupama Yadav, Dana-Farber Cancer Institute, for providing us the pDONR 223 plasmids.

Table of Contents

Generation of RP Variant Clones for Y2H Studies

1. Introduction

1.1. What are ribosomes?

Protein synthesis is a vital part of a cell's housekeeping function and takes up a major portion of the cell's GTP pool[1,2]. Being such an energy-intensive process, the protein synthesis is carried out by specific macromolecular machinery called ribosomes and assisted by the transfer RNA (tRNA). The ribosome is an ancient protein making machine made up of ribosomal proteins (RP) and ribosomal RNA (rRNA). These ribosomal proteins are traditionally considered to be highly conserved across species and have very low evolution rates

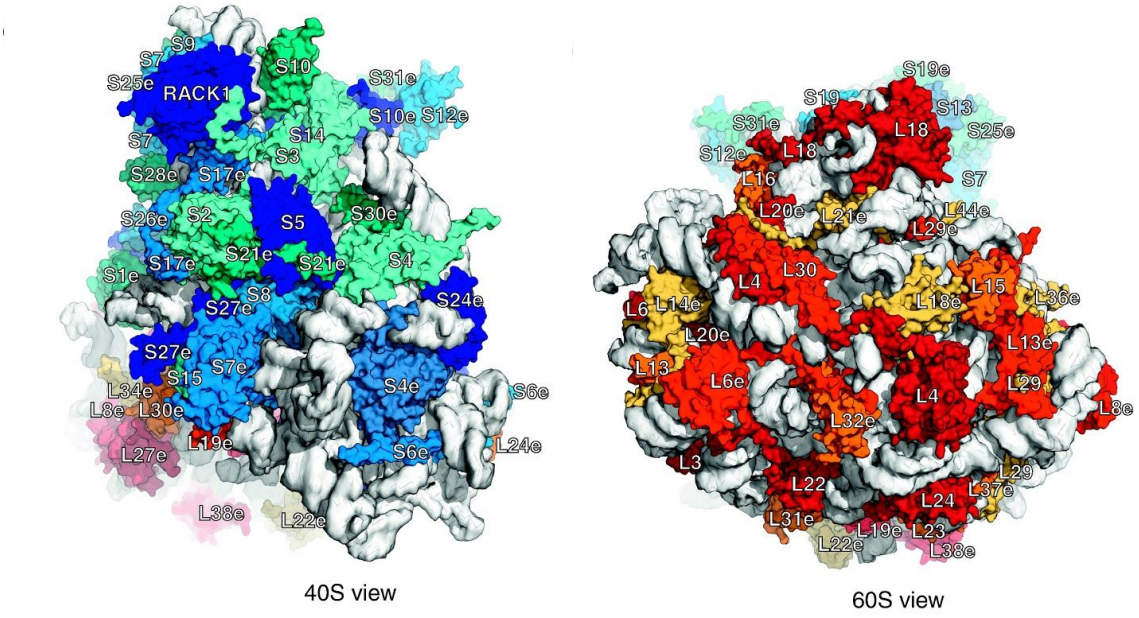


Figure 1: Crystal structures of the two subunits[3]

The ribosome is made up of 2 subunits: the 60S large subunit and the 40S small subunit in eukaryotes or the 50S Large subunit and 30S small subunit in prokaryotes.

In eukaryotes, the large subunit consists of 47 ribosomal proteins (denoted by RPL in yeast and L in humans), 28S rRNA, 5.8S rRNA, and 5S rRNA while the small subunit consists of 32 ribosomal proteins (denoted by RPS in yeast and S in humans), and 18s rRNA.

In prokaryotes, the large subunit consists of 33 ribosomal proteins, 23S rRNA, and 5S rRNA while the small subunit consists of 20 ribosomal proteins and 16s rRNA.

The 18S rRNA in eukaryotes and its homolog in prokaryotes, the 16S rRNA is flanked by highly conserved sequences making them very useful for phylogenic studies

Within a eukaryotic cell, ribosomes are found as free-floating subunits in the cytoplasm or attached to the endoplasmic reticulum. The translation of mRNA into proteins happens here in the cytoplasm. It consists of four steps:

Initiation
When a 5’ capped mRNA is exported out of the nucleus for translation, the small subunit is bound to the mRNA by some initiation factors and scans for the start codon AUG.

Elongation
Once the start codon is identified, the large subunit is recruited and the elongation of the protein begins. There are 3 sites in the ribosomal complex important of elongation: A-site (aminoacyl site) the first site where the incoming tRNA and aminoacid arrive, the P-site (peptidyl site) where the elongating protein is presently attached to the tRNA of the last amino acid and the E-site (exit site) where the tRNA detaches from the complex.

Termination
Once a stop codon is reached, the protein detaches from the P-site, and the translation is terminated.

Recycling
After termination, the ribosome complex breaks up into its subunits and the process is repeated

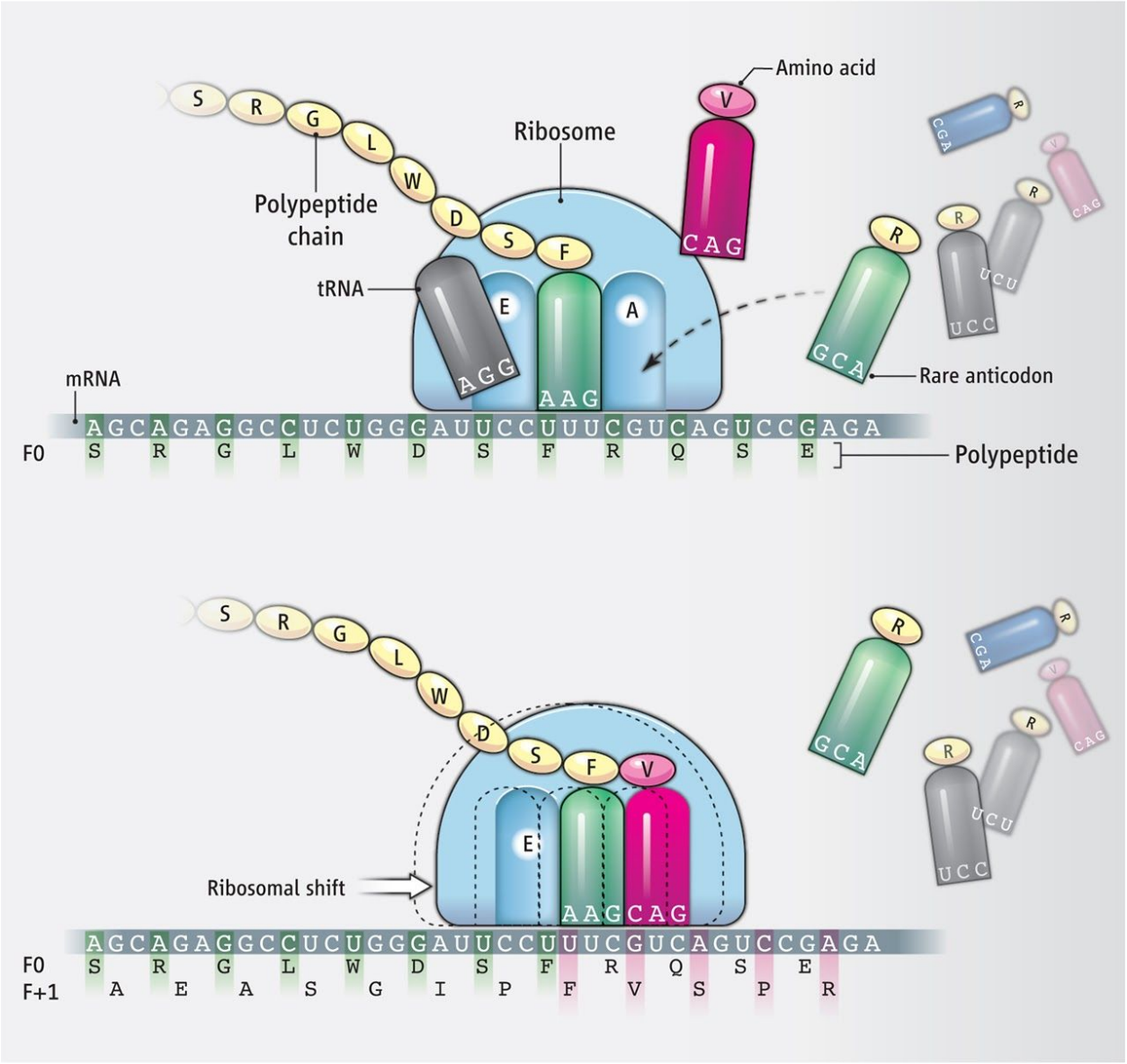


Figure 2: Ribosomal shift during translation[4]

According to the central dogma of life, the sole purpose of the ribosome is the translation of mRNA transcripts into proteins. It is not considered to have any regulation by itself on what transcripts are being translated either. However, emerging studies suggest a greater role for the ribosomal proteins. Using *Saccharomyces cerevisiae* as a model we can investigate these extraribosomal functions of the ribosomal proteins.

1.2. Why *Saccharomyces cerevisiae*?

Saccharomyces cerevisiae is a species of yeast that is commonly used in brewing and baking. It is a single-cell eukaryote with a doubling time of 90 minutes. Naturally, there are 38 strains of *S.cerevisiae* found across the world all of which have been

sequenced and are publicly available[5]. It's also easy to genetically manipulate and overall is a great model organism to study eukaryotic genes.

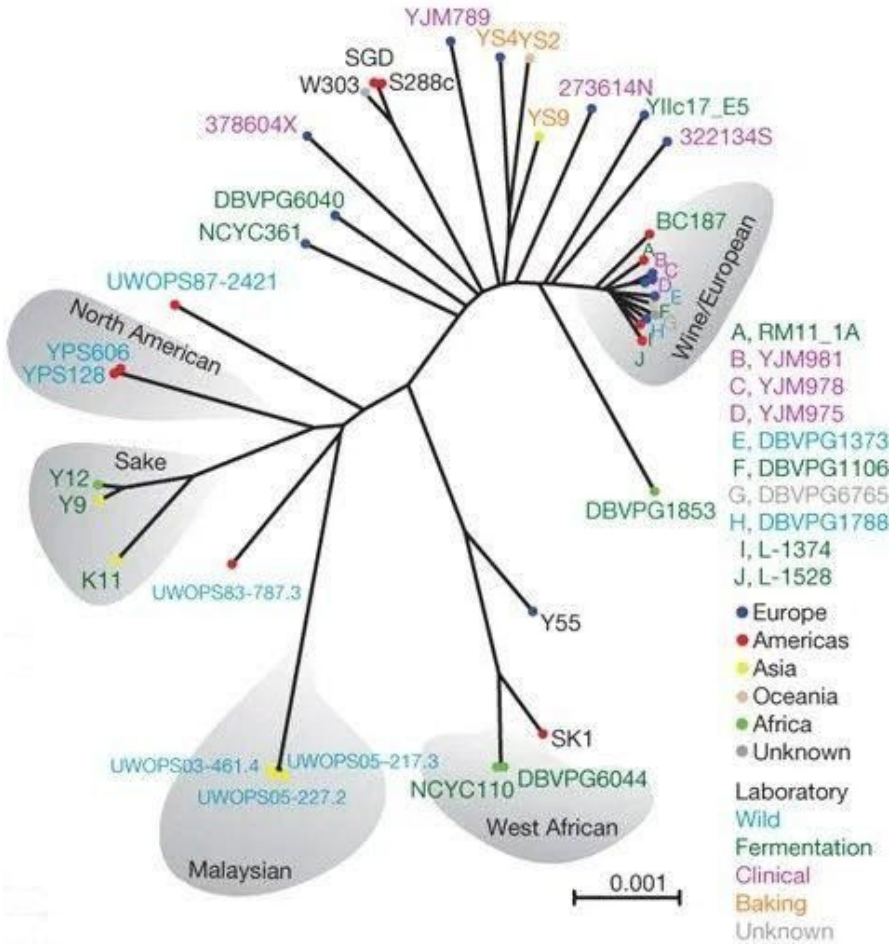


Figure 3: Neighbour-joining tree of all S. cerevisiae strains and their places of origin[5]

Being a eukaryote, *S. cerevisiae* has many homologous proteins with other eukaryotes. They share similar sequences and indicate common ancestry. Investigation of these proteins can lead to insights on their homologs and their function in higher eukaryotes like humans.

Unlike most eukaryotes, due to a whole-genome duplication event, most ribosomal proteins in *S. cerevisiae* have paralogs, nearly identical sequences present in different chromosomes[6]. As a consequence, there are a total of 132 ribosomal proteins in *S.cerevisiae*; 76 RPLs and 56 RPSs

List of all RPLs				List of all RPSs		
RPL01A	RPL15A	RPL28	RPL43B	RPS00A	RPS15	RPS28B
RPL01B	RPL15B	RPL29		RPS00B	RPS16A	RPS29A
RPL02A	RPL16A	RPL30		RPS01A	RPS16B	RPS29B
RPL02B	RPL16B	RPL31A		RPS01B	RPS17A	RPS30A
RPL03	RPL17A	RPL31B		RPS02	RPS17B	RPS30B
RPL04A	RPL17B	RPL32		RPS03	RPS18A	RPS31
RPL04B	RPL18A	RPL33A		RPS04A	RPS18B	
RPL05	RPL18B	RPL33B		RPS04B	RPS19A	
RPL06A	RPL19A	RPL34A		RPS05	RPS19B	
RPL06B	RPL19B	RPL34B		RPS06A	RPS20	
RPL07A	RPL20A	RPL35A		RPS06B	RPS21A	
RPL07B	RPL20B	RPL35B		RPS07A	RPS21B	
RPL08A	RPL21A	RPL36A		RPS07B	RPS22A	
RPL08B	RPL21B	RPL36B		RPS08A	RPS22B	
RPL09A	RPL22A	RPL37A		RPS08B	RPS23A	
RPL09B	RPL22B	RPL37B		RPS09A	RPS23B	
RPL10	RPL23A	RPL38		RPS09B	RPS24A	
RPL11A	RPL23B	RPL39		RPS10A	RPS24B	
RPL11B	RPL24A	RPL40A		RPS10B	RPS25A	
RPL12A	RPL24B	RPL40B		RPS11A	RPS25B	
RPL12B	RPL25	RPL41A		RPS11B	RPS26A	
RPL13A	RPL26A	RPL41B		RPS12	RPS26B	
RPL13B	RPL26B	RPL42A		RPS13	RPS27A	
RPL14A	RPL27A	RPL42B		RPS14A	RPS27B	
RPL14B	RPL27B	RPL43A		RPS14B	RPS28A	

Table 1: List of all ribosomal proteins in *S. cerevisiae*

1.3. The motivation of the study

In a reanalysis of data from a previous study[7], it was observed that when strains of yeast whose ribosomal proteins were deleted were grown in different conditions, they show a huge phenotypic variance in growth (unit: replicative lifespan) (figure 4)[8].

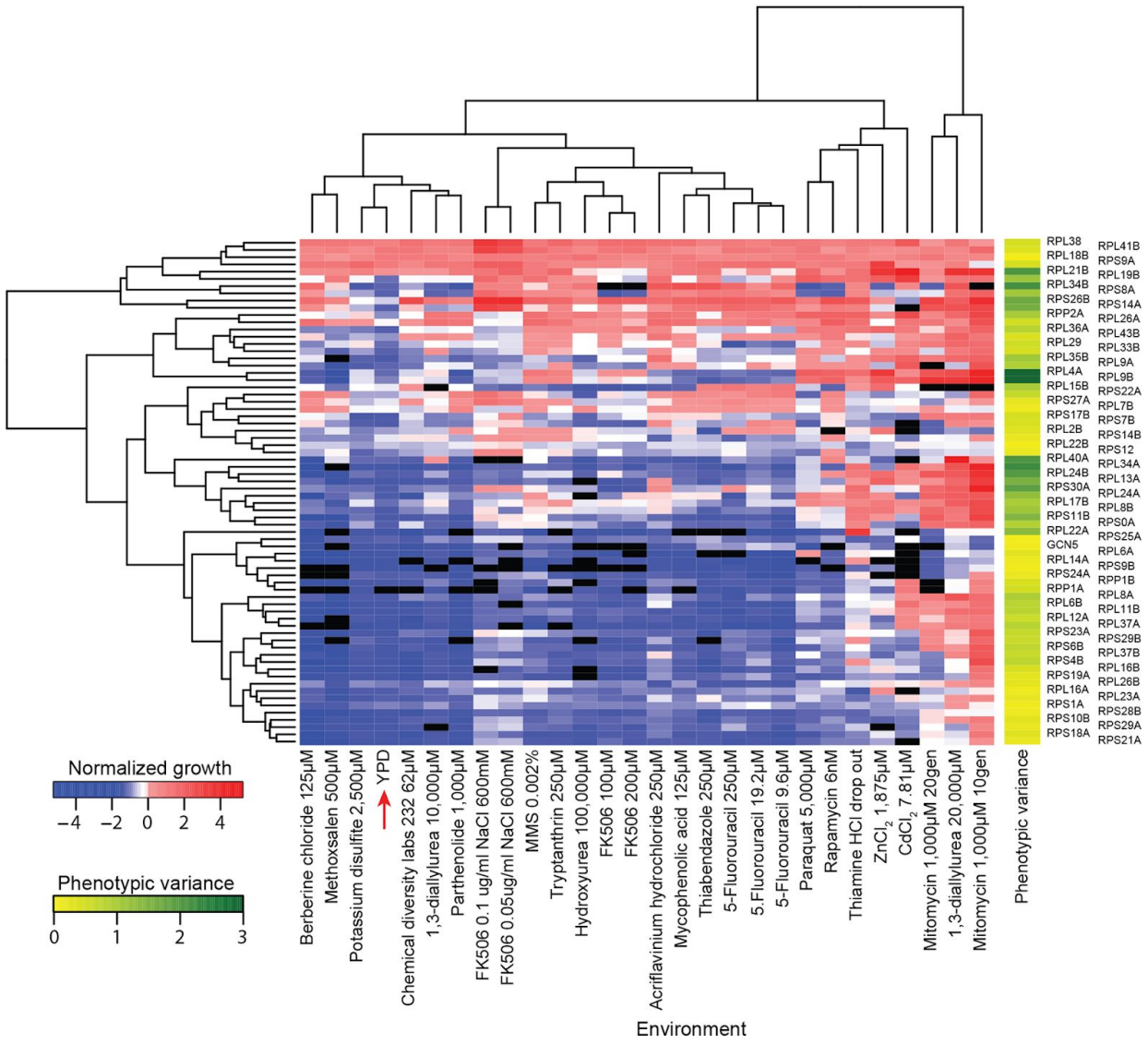


Figure 4: Phenotypic variability of yeast RP genes: Heat map showing hierarchical clustering of normalized growth of yeast strains for 68 single deletions of RP genes in 26 environments. The red arrow indicates YPD (rich medium)[8]

The X-axis of the heat map (figure 4) records the 26 different environments and the Y-axis shows the ribosomal protein deleted. The heat map shows the normalized growth of the deletion strain in a certain medium. As we can see (in figure 4) some deletion strains grow better across all environments, some show worse growth and most of them vary depending on the growth condition.

Recent studies also provide evidence of some extra ribosomal activity of the ribosomal proteins like cell development[9–13], ribosome assembly, [14–16], selection of mRNA for translation[17], cell signaling[18] as well as possible involvement in diseases like Diamond-Blackfan anemia[18,19] to name a few. All this evidence poses an intriguing question: How do the ribosomal proteins exhibit their extra ribosomal functions?

Although most of the studies are performed on deletion strains, in nature most variation occurs through coding and non-coding variants of the protein. While non-coding variants have a more regulatory effect[20], coding variants can alter the final protein giving us functional variants. This leads us to question: Do the ribosomal proteins have functional coding variants? Can the function difference be studied by identifying the interacting partners of the ribosomal proteins?

1.4. Objectives

- Search the existing *Saccharomyces* Genome Database (SGD) and identifying in naturally existing non-synonymous variant alleles of RP genes across all 38 strains in the *Saccharomyces* Genome Resequencing Project database (SGRP)
- Isolate and amplify the genes of interest
- Clone the amplified genes of interest into pDONR223 vectors to generate a library of pENTR vectors
- Conduct yeast 2 hybrid studies using the RP genes as baits and the 5000 yeast proteins as preys and generate the list of allele-specific interacting partners for all RPs

2. Identification of Naturally Occurring Ribosomal Protein Variants

2.1. Introduction

There are 38 identified strains of *Saccharomyces cerevisiae*, from the *Saccharomyces* Genome Resequencing Project database[5]. These strains, over the years, have grown in different environments and accumulated different mutations. A small portion of these mutations is present in the ribosomal proteins. Since these mutations appear naturally in different strains, there is a reasonable possibility that these variable alleles confer some benefit to the stain with respect to its native environment over other strains. While both synonymous and non-synonymous variants are known to have an effect of the proteome, synonymous variants don't change the final ribosomal protein, thereby making its effect indirect, limited to the synthesis of the ribosomal proteins. Non-synonymous variants, however, provide a more viable opportunity for the ribosomal protein to affect the proteome actively by potentially altering key interaction of the ribosomal protein with other proteins in the cell.

2.2. Database

The *Saccharomyces* Genome Database (SGD)[21] is a community resource for studying the budding yeast *Saccharomyces cerevisiae.* The SGD project provides a piece of encyclopedic information on the yeast genome and its genes, proteins, and encoded features. Along with these, peer-reviewed literature of experimental results of functions and interactions is also available. All this data is provided through Locus Summary pages, a powerful query engine, and a rich genome browser. The SGD resource provides the gold standard for functional description of budding yeast and a platform from which to research related genes and pathways in higher organisms.

The *Saccharomyces* Genome Resequencing Project (SGRP)[11] is a collaborative project to explore the genetic and phenotypic diversity of *Saccharomyces.* The goal of the project was to advance our understanding of genetic variation and evolution by analyzing sequences from multiple strains of the two *Saccharomyces* species, *S.cerevisiae* and *S.paradoxus*.The project sequenced haploids of 37 *cerevisiae* strains and 27 *paradoxus* strains. The sequence data has been aligned to the respective reference genome sequences. We use the blast server supplemented by this data to get the RP gene sequences across the varied strains.

Translation of the genomic sequences is done using the online resource, EMBOSS Transeq[22]. For alignment of the various sequences, we used the online resource, Clustal Omega by EMBL[23]

2.3. Procedure

1. The sequence of the gene of interest is obtained from SGD. In case of the presence of intronic regions, the coding sequence is taken i.e. only exons (figure 5)
2. The sequence is BLASTed in the SGRP blast server to obtain the gene sequence in all the S.cerevisiae strains(figure 6). If the gene is present in the non-sense strand of the DNA, the reverse complement is taken
3. The sequences are translated using Transeq. In case the sequence doesn't start from the beginning of the gene, the translation frame is adjusted accordingly
4. The translated sequences are aligned using the online multiple sequence alignment tool, Clustal Omega and the screened for variant amino acids
5. The variant amino acid along with the corresponding variant nucleotide, their positions wrt the start codon, and strains exhibiting the variant are listed

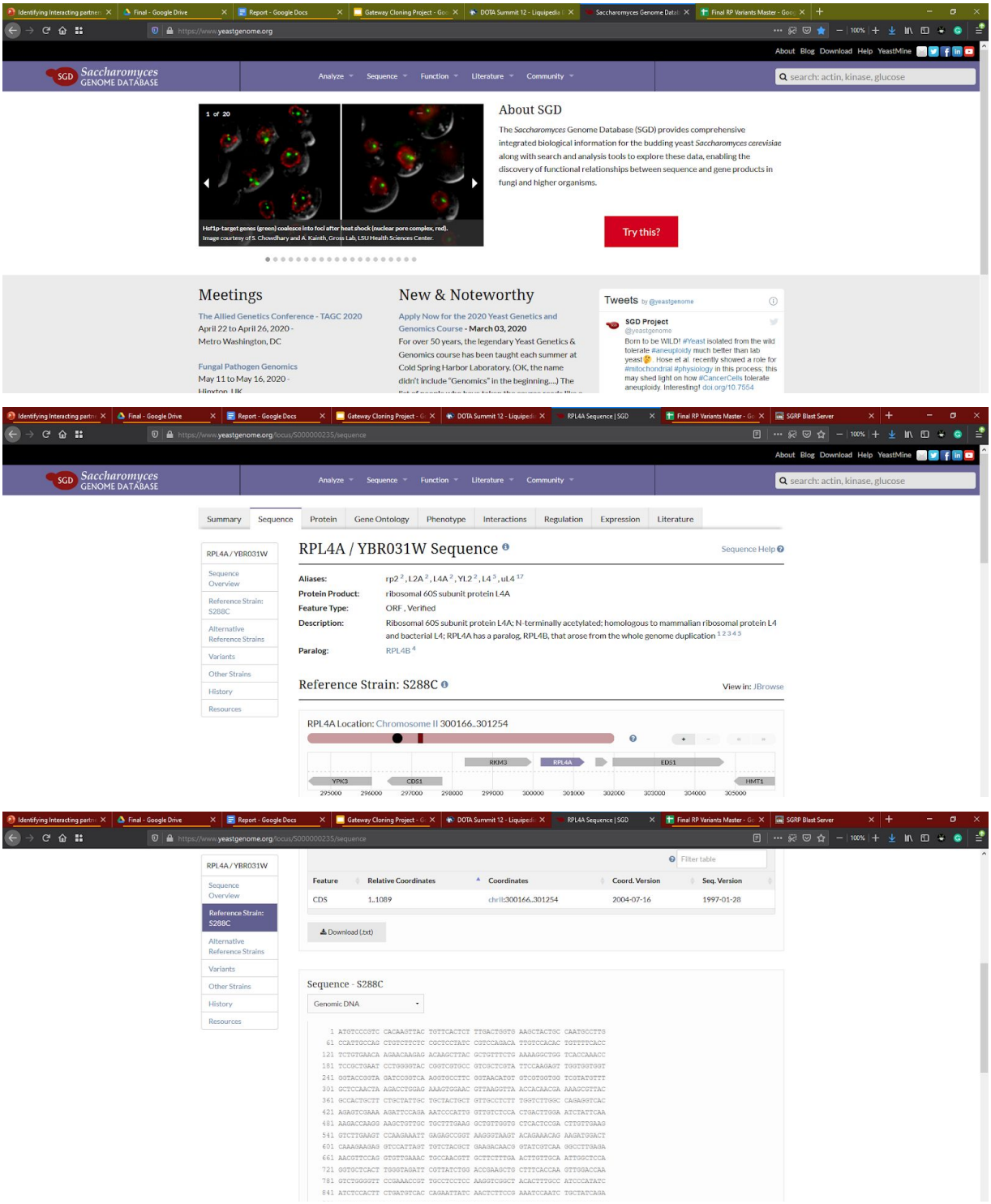
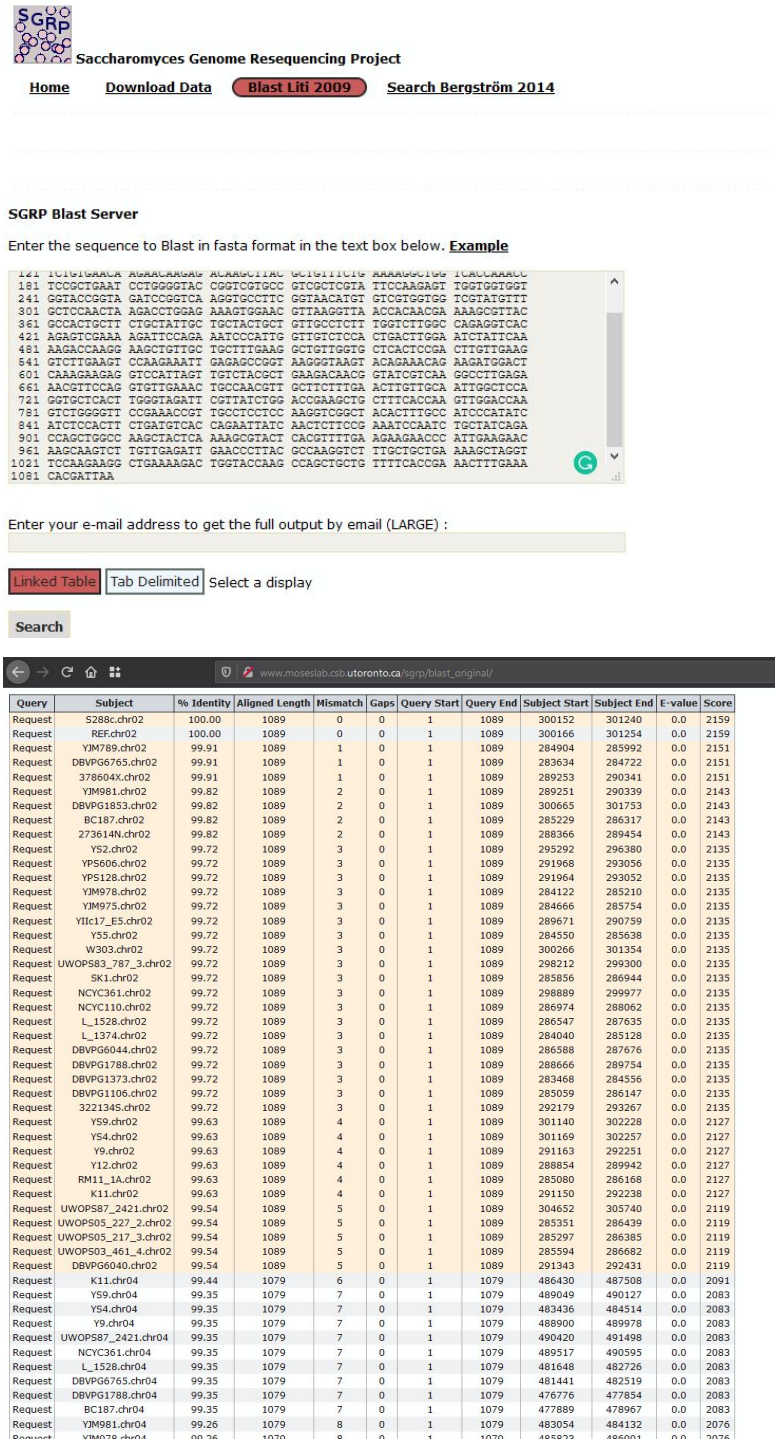


Figure 5: Screenshots from the SGD website for the RPL4A query[21]



Query	Subject	% Identity	Aligned Length	Mismatch	Gaps	Query Start	Query End	Subject Start	Subject End	E-value	Score
Request	S288c.chr02	100.00	1089	0	0	1	1089	300152	301240	0.0	2159
Request	REF.chr02	100.00	1089	0	0	1	1089	300166	301254	0.0	2159
Request	YJM789.chr02	99.91	1089	1	0	1	1089	284904	285992	0.0	2151
Request	DBVPG6765.chr02	99.91	1089	1	0	1	1089	283634	284722	0.0	2151
Request	378604X.chr02	99.91	1089	1	0	1	1089	289253	290341	0.0	2151
Request	YJM981.chr02	99.82	1089	2	0	1	1089	289251	290339	0.0	2143
Request	DBVPG1853.chr02	99.82	1089	2	0	1	1089	300665	301753	0.0	2143
Request	BC187.chr02	99.82	1089	2	0	1	1089	285229	286317	0.0	2143
Request	273614N.chr02	99.82	1089	2	0	1	1089	288366	289454	0.0	2143
Request	YS2.chr02	99.72	1089	3	0	1	1089	295292	296380	0.0	2135
Request	YPS606.chr02	99.72	1089	3	0	1	1089	291968	293056	0.0	2135
Request	YPS128.chr02	99.72	1089	3	0	1	1089	291964	293052	0.0	2135
Request	YJM978.chr02	99.72	1089	3	0	1	1089	284122	285210	0.0	2135
Request	YJM975.chr02	99.72	1089	3	0	1	1089	284666	285754	0.0	2135
Request	YIIc17_E5.chr02	99.72	1089	3	0	1	1089	289671	290759	0.0	2135
Request	Y55.chr02	99.72	1089	3	0	1	1089	284550	285638	0.0	2135
Request	W303.chr02	99.72	1089	3	0	1	1089	300266	301354	0.0	2135
Request	UWOPS83_787_3.chr02	99.72	1089	3	0	1	1089	298212	299300	0.0	2135
Request	SK1.chr02	99.72	1089	3	0	1	1089	285856	286944	0.0	2135
Request	NCYC361.chr02	99.72	1089	3	0	1	1089	298889	299977	0.0	2135
Request	NCYC110.chr02	99.72	1089	3	0	1	1089	286974	288062	0.0	2135
Request	L_1528.chr02	99.72	1089	3	0	1	1089	286547	287635	0.0	2135
Request	L_1374.chr02	99.72	1089	3	0	1	1089	284040	285128	0.0	2135
Request	DBVPG6044.chr02	99.72	1089	3	0	1	1089	286588	287676	0.0	2135
Request	DBVPG1788.chr02	99.72	1089	3	0	1	1089	288666	289754	0.0	2135
Request	DBVPG1373.chr02	99.72	1089	3	0	1	1089	283468	284556	0.0	2135
Request	DBVPG1106.chr02	99.72	1089	3	0	1	1089	285059	286147	0.0	2135
Request	322134S.chr02	99.72	1089	3	0	1	1089	292179	293267	0.0	2135
Request	YS9.chr02	99.63	1089	4	0	1	1089	301140	302228	0.0	2127
Request	YS4.chr02	99.63	1089	4	0	1	1089	301169	302257	0.0	2127
Request	Y9.chr02	99.63	1089	4	0	1	1089	291163	292251	0.0	2127
Request	Y12.chr02	99.63	1089	4	0	1	1089	288854	289942	0.0	2127
Request	RM11_1A.chr02	99.63	1089	4	0	1	1089	285080	286168	0.0	2127
Request	K11.chr02	99.63	1089	4	0	1	1089	291150	292238	0.0	2127
Request	UWOPS87_2421.chr02	99.54	1089	5	0	1	1089	304652	305740	0.0	2119
Request	UWOPS05_227_2.chr02	99.54	1089	5	0	1	1089	285351	286439	0.0	2119
Request	UWOPS05_217_3.chr02	99.54	1089	5	0	1	1089	285297	286385	0.0	2119
Request	UWOPS03_461_4.chr02	99.54	1089	5	0	1	1089	285594	286682	0.0	2119
Request	DBVPG6040.chr02	99.54	1089	5	0	1	1089	291343	292431	0.0	2119
Request	K11.chr04	99.44	1079	6	0	1	1079	486430	487508	0.0	2091
Request	YS9.chr04	99.35	1079	7	0	1	1079	489049	490127	0.0	2083
Request	YS4.chr04	99.35	1079	7	0	1	1079	483436	484514	0.0	2083
Request	Y9.chr04	99.35	1079	7	0	1	1079	488900	489978	0.0	2083
Request	UWOPS87_2421.chr04	99.35	1079	7	0	1	1079	490420	491498	0.0	2083
Request	NCYC361.chr04	99.35	1079	7	0	1	1079	489517	490595	0.0	2083
Request	L_1528.chr04	99.35	1079	7	0	1	1079	481648	482726	0.0	2083
Request	DBVPG6765.chr04	99.35	1079	7	0	1	1079	481441	482519	0.0	2083
Request	DBVPG1788.chr04	99.35	1079	7	0	1	1079	476776	477854	0.0	2083
Request	BC187.chr04	99.35	1079	7	0	1	1079	477889	478967	0.0	2083
Request	YJM981.chr04	99.26	1079	8	0	1	1079	483054	484132	0.0	2076

Figure 6: Screenshot of the BLAST search of the RPL4A sequence on the SGRP server[24]

2.4. Results

Of the 132 RP genes present in S.cerevisiae, 53 show at least one non-synonymous variant in at least one strain. A total of 81 variants were observed

Gene	Variant	Gene	Variant	Gene	Variant	Gene	Variant
RPL02A	S53G	RPL09B	Q121K	RPL18A	V117A	RPL36B	I81T
RPL04A	S168A	RPL10	I148V	RPL18A	S172F	RPL37B	A33T
RPL04A	A256T	RPL10	R183K	RPL18B	T51A	RPS00B	G66A
RPL04B	V19A	RPL12A	V49A	RPL19A	T12A	RPS01B	V168I
RPL04B	A356T	RPL13A	L150P	RPL21B	V111A	RPS04B	I76V
RPL05	E238D	RPL13A	G161D	RPL21B	K146N	RPS06B	I16F
RPL05	V292A	RPL14A	A67P	RPL22B	V46A	RPS07B	Q122H
RPL06B	N117K	RPL14A	G89A	RPL22B	Q55E	RPS07B	N123D
RPL07A	T50A	RPL15A	D153E	RPL22B	S57T	RPS08A	T45S
RPL07B	T171A	RPL15B	A18V	RPL26A	Q26E	RPS09B	R40K
RPL08A	L58V	RPL15B	H34N	RPL27A	G26V	RPS11B	I64V
RPL08A	H95N	RPL15B	A35V	RPL27A	S37P	RPS14A	T40A
RPL08A	T252N	RPL15B	Y178H	RPL27A	L75V	RPS14B	D76G
RPL08B	T25P	RPL16A	V6F	RPL28	T66A	RPS14B	V77I
RPL08B	N122K	RPL16A	A152V	RPL29	Q58K	RPS15	A7V
RPL08B	V153I	RPL16A	R163S	RPL33A	S38P	RPS17B	S104N
RPL08B	V180I	RPL16B	V62A	RPL34B	V92A	RPS17B	L110V
RPL08B	I252T	RPL16B	T174A	RPL35A	S112P	RPS25A	R90K
RPL09A	D61G	RPL17A	S17A	RPL35B	Q14K	RPS26A	T63A
RPL09A	G153D	RPL17A	N53D	RPL36A	G19S	RPS30A	I50V
						RPS30B	I50V

Table 2: List of all Ribosomal Protein Variants

2.5. Discussion

The existence of the non-synonymous variants in what is considered to be a highly conserved and essential macromolecule further supports our hypothesis that RPs are capable of differentially interacting with other proteins in the cell. These interactions can be involved in extra ribosomal functions, i.e, the ability of the ribosomal proteins to work outside of translation like modulating the translation of differential transcription pools

3. Generation of RP allele clones using Gateway cloning

3.1. Introduction

In order to study the effect of the variant RPs on the cell, multiple approaches can be used. One possible way is to generate allele-specific populations for each variant and compare the growth in different conditions. While this approach is viable for in-depth studies of an RP, it is not advisable for a large scale screen of the effect. To this effect, a different approach is considered.

In this approach, the variant alleles are expressed in a system and screened for all its interacting partners in the entire yeast genome. This process is more scalable to the number of variants observed and provides an understanding of the effect of the variant at a whole-cell level

Keeping this in mind, strains are chosen such that for a given RP gene there exists only one non-synonymous SNP difference between them. The genes are then extracted and cloned thereby generating two copies of the gene that, when translated differ by only one amino acid.

3.2. What is gateway cloning

Gateway cloning is a novel cloning technique developed by Invitrogen, where the gene of interest is inserted into the plasmid using recombination rather than the traditional restriction enzyme based ligation. This is achieved by flanking the gene of interest with specific sequences called att-B sites that can recombine with a complimentary site on the plasmid, called att-P site in a reaction termed BP reaction. The technique was developed inspired by the way lambda phages insert and extract their genomic DNA from the host during the course of an infection. Its genome is inserted into the host using an enzyme, integrase, and an integration host factor. The reverse reaction called LR reaction is catalyzed by excisionase along with the integrase and IHF.

In our BP reaction, only the pENTR vector can be taken up by DH5α cells and survive on a spectinomycin plate. This is because the PCR product and the by-product lack the resistance marker while the ccdB toxin in pDONR kills the DH5α cells that pick it up

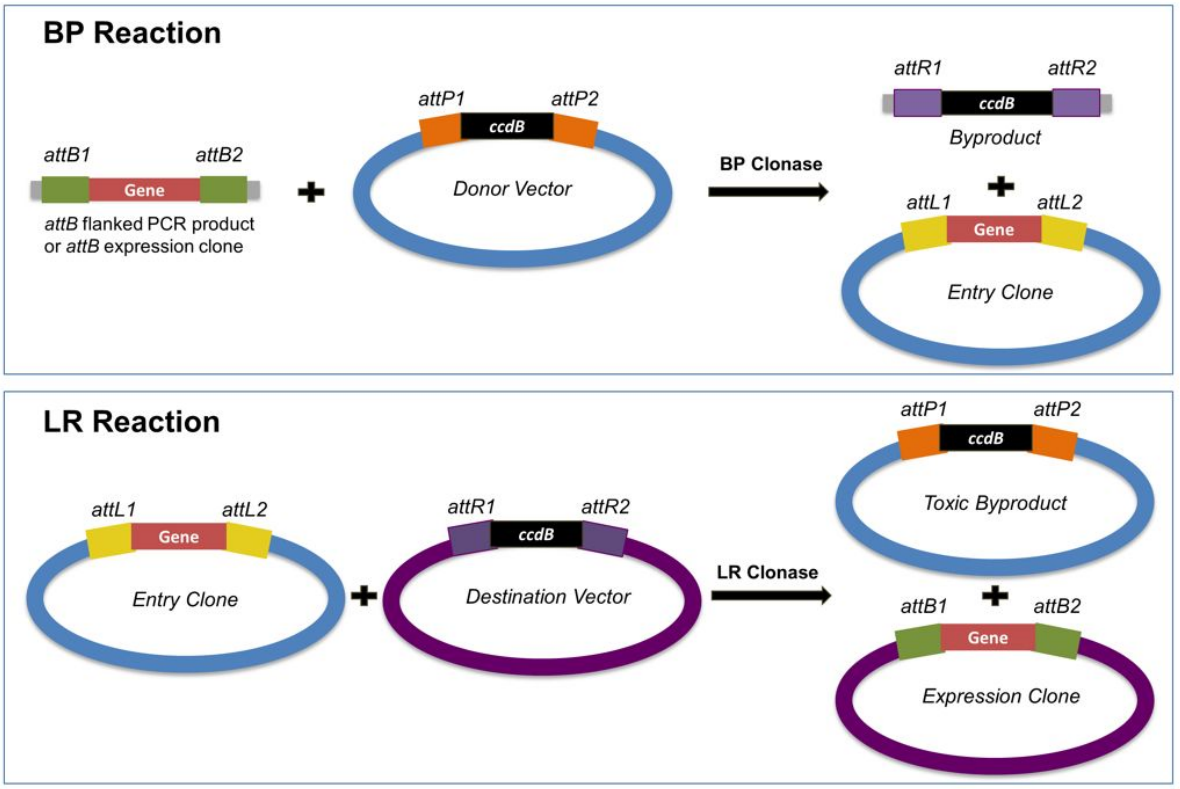


Figure 7: Gateway Cloning [25]

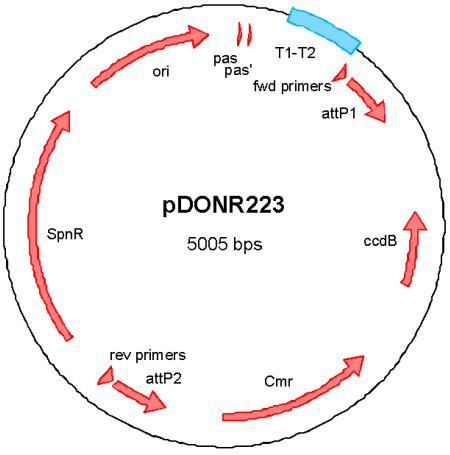


Figure 8: The plasmid map of pDONR223

3.3. Maintainance of pDONR 223

The pDONR 223 plasmid contains the ccdB gene which is a gyrase inhibitor. It acts as a toxin and kills any cell that doesn't have a gyrase mutation like *E.coli* strain DH5α. We use the **XL 10 gold** strain of *E. coli* to maintain and propagate the vector. We use the chloramphenicol resistance as the selective marker for the maintenance

3.4. Methodology

3.4.1. Extraction of Genomic DNA[26]

1. Select the strain of *S. cerevisiae* such that it contains only one variant in the gene of interest
2. Grow 10 mL of the strain culture in YPD overnight
3. Ensure the OD at 600 is about 0.4 after a 1 in 10 dilutions. Spin the culture at 1200*g for 5 min. Discard the supernatant and resuspend in 500µL water
4. Transfer to a microcentrifuge tube and spin for 5 sec. Discard the supernatant and disrupt the pellet by vortexing
5. Resuspend cells in 200µL of 25:24:1 (v/v/v) phenol/ chloroform/ isoamyl alcohol buffered with phenol. Add 300µg of glass beads. Vortex at max speed for 3 minutes
6. Add 200µL TE and vortex briefly. Centrifuge for 5 min and transfer the aqueous layer to a clean tube
7. Repeat steps 5 and 6 one more time
8. Add 1mL 100% ethanol and mix gently by inversion
9. Centrifuge for 3 min at max speed. Discard supernatant and re-suspend in 400µL TE
10. Add 30µL of 1 mg/mL RNase A, mix and incubate for 5 min at 37℃
11. Add 10µL 5M ammonium acetate and 1mL 100% ethanol. Centrifuge for 3 min at max speed
12. Discard the supernatant and air dry the pellet. Resuspend DNA in 100µL TE

Result

The genomic DNA of the required strains were extracted in the following concentrations

Strain	Concentration (µg/mL)	A260/A280	A260/230
YS2	266.5	1.73	1.52
YS4	316.5	1.74	1.58
YS9	324	1.76	1.62
S288C	296.75	1.8	1.9
UWOPS05_227_2	211.5	1.7	1.45
K11	256.25	1.73	1.57
SK1	200	1.7	2.6
322134S	194	1.8	2.7
273614N	333	1.92	2.98
DBVPG1788	484.75	1.8	2.3
YJM975	700.25	1.77	2.06
273614N	590	1.75	2.11
RM11_1A	562.75	1.7	2
Y9	596	1.75	2.17
DBVPG1853	583.25	1.79	2.03
YJM789	235.25	1.77	1.8
L_1528	328.25	1.77	1.9
UWOPS83_787_3	664	1.77	1.88

Table 3: Concentrations of the extracted genomic DNA

3.4.2. Primer Design

1. The first and last 30 nucleotides of the ribosomal protein gene are screened such that there is at least a difference of 3 bases with the paralog. The resulting sequence is taken as our nucleotide specific sequence. The reverse complement is taken for the reverse primer
2. The attB sequences are added to the start of the primers
    a. Forward attB:GGGGACAACTTTGTACAAAAAAGTTGGCACC
    b. Reverse attB:GGGGACAACTTTGTACAAGAAAGTTGGCAA

Result

The primers for the amplification of the gene of interest are designed and given for manufacturing

3.4.3. PCR Amplification of Gene of Interest

1. Set up the reaction as follows (if making master mix add primers and template separately):

Component	Conc.	Units	Final Conc.	Units	20	µL rxn	X +1	rxn Master mix
5x Q5 Buffer	5	X	1	X	4	µL	4*(X+1)	µL
dNTPs(2mM)	2	mM	0.2	mM	2	µL	2*(X+1)	µL
5x GC Enhancer	5	X	1	X	4	µL	4*(X+1)	µL
Q5	2	U/µL	0.02	U/µL	0.2	µL	0.2*(X+1)	µL
FP	5	µM	0.5	µM	2	µL		
RP	5	µM	0.5	µM	2	µL		
Template(50ng/µL)	50	ng/µL	1	ng/µL	0.4	µL		
H2O					5.4	µL	5.4*(X+1)	µL
Final					20	µL	15.6	µL/rxn

Table 4: PCR preparation

2. Place the PCR tubes in the thermocycler
3. Run the protocol as follows (use NEBTm Calculator[27] preferably):

Name	Temperature (°C)	Time	Cycles
Initial Denaturation	98	30 sec	
Denaturation 1	98	10 sec	
Annealing 1	Ta of nucleotide specific sequence	30 sec	
Extension 1	72	30 sec/kb	10
Denaturation 2	98	10 sec	
Annealing 2	Ta of the whole primer	30 sec	
Extension 2	72	30 sec/kb	35
Final Extension	72	2 min	
Hold	4		

Table 5: PCR Protocol

4. Load 1µL of the PCR product using a 6x loading buffer in a 1% gel. Use appropriate ladder

Result

The amplified PCR product can be visualized on the gel

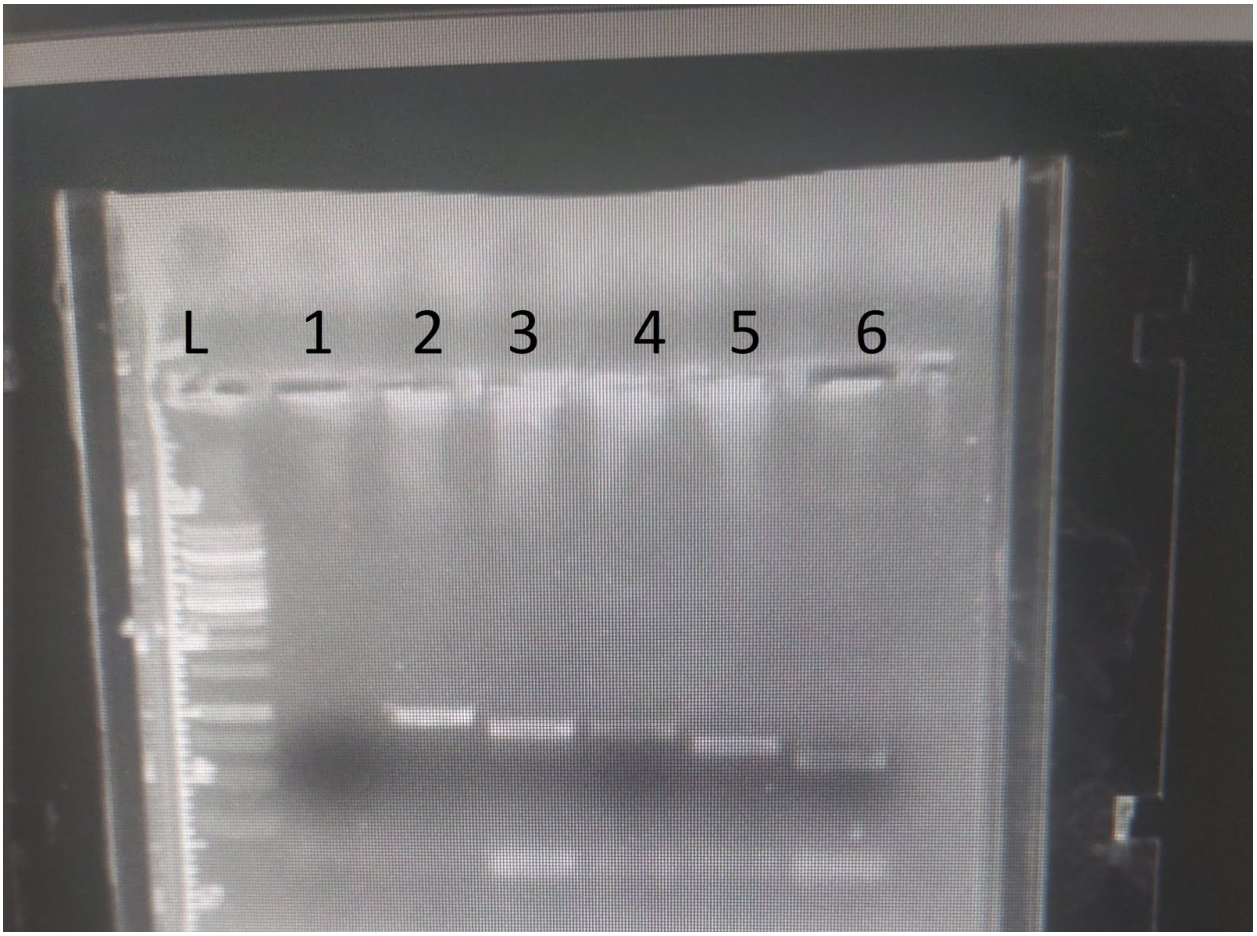


Figure 9:Gel electrophoresis image of PCR amplified products;
Legend L – Ladder; 1 – RPL13A P150L D161G; 2 – RPL06B K117N; 3 – RPS14B G76D;
4 – RPS14B I77V; 5 – RPS14A A40T; 6 – RPS08A S45T

3.4.4. BP Reaction and Transformation

1. Dilute the PCR product 4-fold with TE
2. Add ½ volume of 30% PEG 8000/ 30mM MgCl2 and vortex
3. Centrifuge for 15 min at full speed in a microcentrifuge
4. Discard the supernatant and suspend clear pellet in TE
5. Thaw Clonase™ II enzyme mix on ice for 2 minutes. Vortex briefly twice
6. Incubate the purified PCR product with the pDONR vector and the BP Clonase™ II enzyme mix at 25°C overnight (BP Reaction) as follows

Reagent	Volume (µL)
PCR product	0.6
pDONR	0.3
Clonase™ II Buffer	0.6
TE	1.5
Total	3

Table 6: BP reaction setup

7. Add 0.3µL of Proteinase K to each reaction and incubate at 37°C for 10 minutes
8. Thaw CaCl2 washed DH5α competent cells on ice for half an hour
9. Add 1µL BP reaction mixture to the competent cells and incubate on ice for 30 min
10. Heat-shock the cells by incubating at 42°C for 30 seconds
11. Place on ice for 2 min and add 250µL of S.O.C. medium and incubate at 37°C for 1 hour
12. Spin down the cells, discard part of the supernatant and resuspend in 100µL media
13. Plate the cells on YPD plated contain Spectinomycin
14. Observe the growth of transformants and restreak the culture
15. Make a 1mL overnight culture and prepare glycerol stocks by adding 250 mL of 100% glycerol to 750 mL culture
16. Store at -80°C

3.5. Results

The required variants are cloned into DH5α cells

Gene	Strain	Variant 1	Variant 2
RPL06B	DBVPG1853	K117N	
RPL07B	322134S	A171T	
RPL09A	DBVPG1853	G61D	
RPL09A	RM11_1A	D153G	
RPL09B	SK1	K121Q	
RPL13A	Y9	P150L	D161G
RPL14A	YJM789	P67A	
RPL15B	DBVPG1853	V18A	H178Y
RPL15B	SK1	V35A	
RPL15B	YJM975	N34H	
RPL21B	YS9	N146K	
RPL35B	YS9	K14Q	
RPS01B	YJM975	I168V	
RPS06B	K11	F16I	
RPS08A	UWOPS05_227_2	S45T	
RPS11B	UWOPS05_227_2	V64I	
RPS14A	DBVPG1788	A40T	
RPS14B	L_1528	G76D	
RPS14B	SK1	I77V	
RPS17B	L_1528	V110L	
RPS17B	Y9	N104S	

Table 7: List of Ribosomal Proteins Cloned into DH5α

3.6. Discussion

Since the variants of a gene of interest differ by one SNP, we can study the effect of the variant allele individually. This lets us establish the effect each variant has in further studies

4. Future Work

4.1. Introduction

Following the creation of the variant clones, we aim to establish the interacting partners of each of the genes of interest and study the effect the variants have on them. Since we aim to screen a large number of variants across the entire proteome of S.cerevisae, we opt for the Yeast 2 Hybrid (Y2H) system for the interaction studies.

In a Yeast 2 Hybrid system, the host cell is modified by replacing a native transcription activator, like Gal4 is replaced with two synthetic hybrid proteins[28] [29](figure 10); one containing the N-terminal DNA binding domain of Gal 4 fused to one of our cloned variants (called bait), and the other containing the C-terminal transcription activating domain fused to one of the 5000 proteins present in the yeast genome (called prey). The Gal 4 transcription factor enables the transcription of HIS3 which is required for the cell to grow in media lacking histidine. If the bait and prey don not interact, the transcription of HIS3 does not occur and the cell won't grow. If the bait and prey interact, the transcription factor is reconstituted and the cell grows on media lacking histidine.

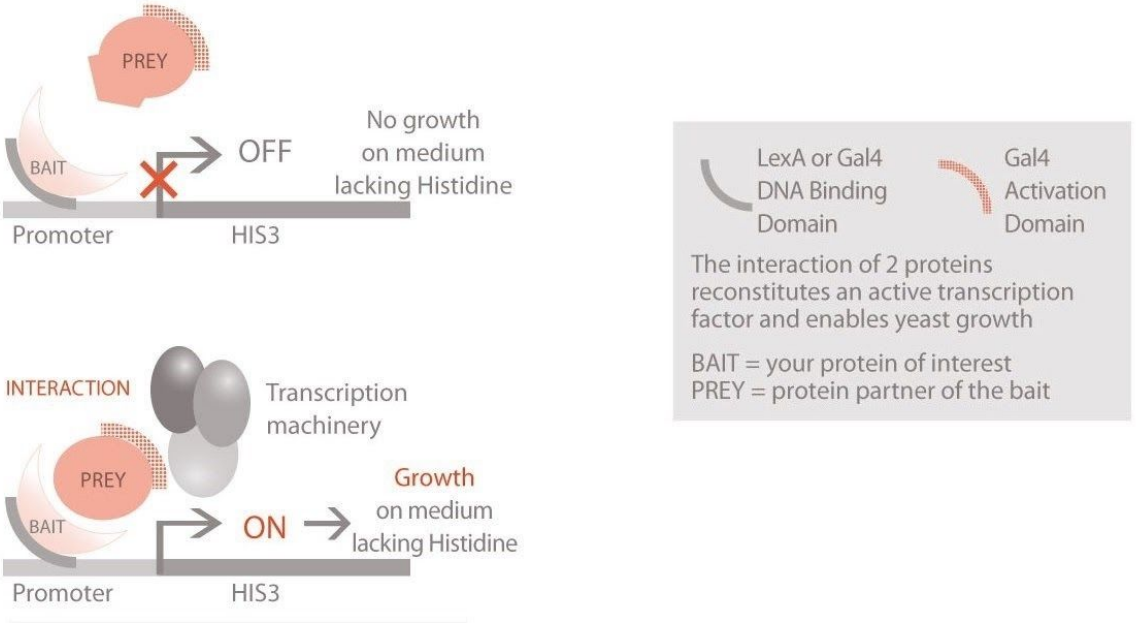


Figure 10: Yeast 2 Hybrid[30]

4.2. Why Yeast 2 Hybrid

While there are several techniques available to study protein interactions, it is incredibly hard to scale them to a large number of samples. The Yeast 2 Hybrid

system enables us to perform a large scale study using a library of the yeast genome simultaneously. Being an in-vivo system it is better suited to replicate the conditions within the cell making the model closer to the actual yeast cell. The Yeast 2 hybrid system also takes care of all post-transcriptional and translational modifications that the protein would normally undergo. Additionally, we do not need to make proteins and antibodies of high purity which is laborious as well as expensive.

4.3. Possible outcomes and hypotheses

The Yeast 2 Hybrid studies will generate a network of genes that our protein of interest interacts with (figure 11). The variant can potentially alter this network by removing and/ or adding new genes. This difference in the set of genes should be caused entirely by the variant allele. This gives us an idea of what pathways our ribosomal protein is potentially involved in and what sort of extra ribosomal function it possesses

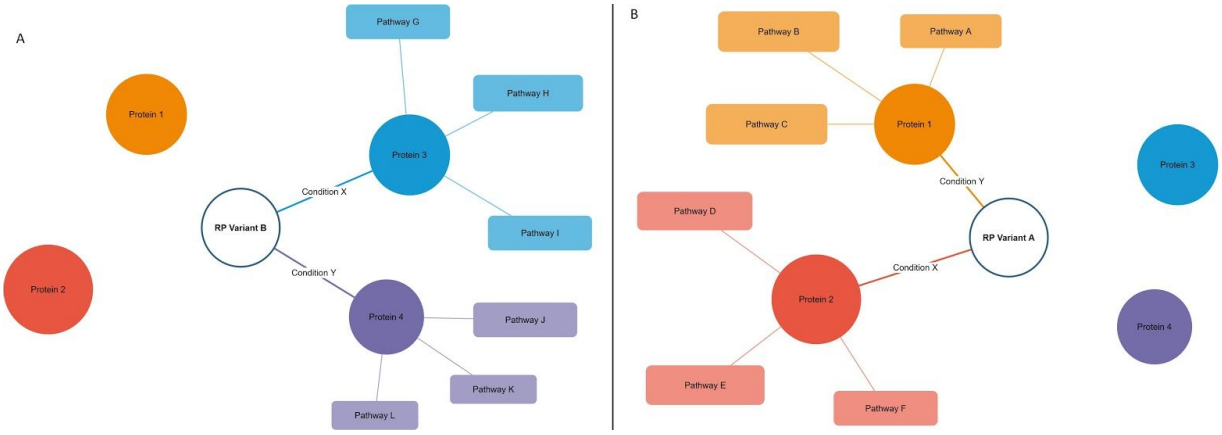


Figure 11:Possible interactions, (A) Possible interactions of variant B, (B) Possible interactions of variant A

5. Highlights of the work

The key features of this work are mentioned below

1. This work builds on upcoming hypotheses that support extra ribosomal functions for the ribosomal proteins

2. Though thought to be highly conserved, variant alleles have been identified in the ribosomal proteins
3. A novel high-efficiency approach to cloning is employed in the form of gateway cloning
4. Set up protocols for BP reactions

6. Reagents

6.1. 1M Tris Buffer (10 mL) (M.W. 121.13 g/mol)

- Weigh 1.2114 g of Tris base
- Dissolve in water and adjunct pH to 8 using HCl
- Makeup to 10 mL

6.2. TE Buffer (50 mL)

- Mix 0.5 mL of 1M Tris HCl and 0.1 mL of 0.5 M EDTA
- Makeup to 50 mL volume

6.3. LB Agar Plate (100 mL)

- Dissolve 2.5g LB and 1.5g Agar in 100 mL water

6.4. SOC Base (1 L)

- Tryptone - 20 g
- Yeast extract - 5g

6.5. SOC Media Additions (44.5 mL)(44.5 µL per mL of SOC Base)

- 5M NaCl - 2 mL
- 1M KCl - 2.5 mL
- 1M MgCl2 - 10 mL
- 1M MgSO4 - 10 mL
- 1M Glucose - 20 mL

6.6. Spectinomycin (Stock 100 mg/ mL) (Working 100 µg/ mL)

- Weigh 100 mg and dissolve in 1 mL water
- Filter sterilize
- Use 100 µL per 100 ml of media

6.7. Chloramphenicol (Stock 35mg / mL) (Working 35 µg/ mL)

- Weigh 35 mg in a dark tube and dissolve in 1 mL 100% Ethanol
- Filter sterilize
- Use 100 µL per 100 mL of media

7. References

1. Kafri M, Metzl-Raz E, Jona G, Barkai N. The Cost of Protein Production. Cell Reports. 2016. pp. 22–31. doi:10.1016/j.celrep.2015.12.015

2. Thomson E, Ferreira-Cerca S, Hurt E. Eukaryotic ribosome biogenesis at a glance. Journal of Cell Science. 2013. pp. 4815–4821. doi:10.1242/jcs.111948

3. Jenner L, Melnikov S, Garreau de Loubresse N, Ben-Shem A, Iskakova M, Urzhumtsev A, et al. Crystal structure of the 80S yeast ribosome. Curr Opin Struct Biol. 2012;22: 759–767.

4. Yewdell JW, Ince WL. Frameshifting to PA-X Influenza. Science. 2012. pp. 164–165. doi:10.1126/science.1225539

5. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, et al. Population genomics of domestic and wild yeasts. Nature. 2009;458: 337–341.

6. Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature. 1997;387: 708–713.

7. Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, et al. The chemical genomic portrait of yeast: uncovering a phenotype for all genes. Science. 2008;320: 362–365.

8. Yadav A, Radhakrishnan A, Panda A, Singh A, Sinha H, Bhanot G. The Modular Adaptive Ribosome. doi:10.1101/068361

9. Strittmatter AW, Fischer C, Kleinschmidt M, Braus GH. FLO11 mediated filamentous growth of the yeast Saccharomyces cerevisiae depends on the expression of the ribosomal RPS26 genes. Mol Genet Genomics. 2006;276: 113–125.

10. Mao-De L, Jing X. Ribosomal Proteins and Colorectal Cancer. Current Genomics. 2007. pp. 43–49. doi:10.2174/138920207780076938

11. Danilova N, Sakamoto KM, Lin S. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. Blood. 2008. pp. 5228–5237. doi:10.1182/blood-2008-01-132290

12. Byrne ME. A role for the ribosome in development. Trends Plant Sci. 2009;14: 512–519.

13. Szakonyi D, Byrne ME. Ribosomal protein L27a is required for growth and patterning in Arabidopsis thaliana. The Plant Journal. 2011. pp. 269–281.

doi:10.1111/j.1365-313x.2010.04422.x

14. Fernández-Pevida A, Rodríguez-Galán O, Díaz-Quintana A, Kressler D, de la Cruz J. Yeast Ribosomal Protein L40 Assembles Late into Precursor 60 S Ribosomes and Is Required for Their Cytoplasmic Maturation. Journal of Biological Chemistry. 2012. pp. 38390–38407. doi:10.1074/jbc.m112.400564

15. Martín-Marcos P, Hinnebusch AG, Tamame M. Ribosomal Protein L33 Is Required for Ribosome Biogenesis, Subunit Joining, and Repression of GCN4 Translation. Molecular and Cellular Biology. 2007. pp. 5968–5985. doi:10.1128/mcb.00019-07

16. Nissan TA. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. The EMBO Journal. 2002. pp. 5539–5547. doi:10.1093/emboj/cdf547

17. Bauer JW, Brandl C, Haubenreisser O, Wimmer B, Weber M, Karl T, et al. Specialized Yeast Ribosomes: A Customized Tool for Selective mRNA Translation. PLoS ONE. 2013. p. e67609. doi:10.1371/journal.pone.0067609

18. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, et al. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol. 1999;17: 676–682.

19. Raiser DM, Narla A, Ebert BL. The emerging importance of ribosomal dysfunction in the pathogenesis of hematologic disorders. Leukemia & Lymphoma. 2014. pp. 491–500. doi:10.3109/10428194.2013.812786

20. She R, Jarosz DF. Mapping Causal Variants with Single-Nucleotide Resolution Reveals Biochemical Drivers of Phenotypic Change. Cell. 2018;172: 478–490.e15.

21. Saccharomyces Genome Database | SGD. [cited 7 Jun 2020]. Available: https://www.yeastgenome.org/

22. EMBL-EBI. EMBOSS Transeq < Sequence Translation Sites < EMBL-EBI. [cited 7 Jun 2020]. Available: https://www.ebi.ac.uk/Tools/st/emboss_transeq/

23. EMBL-EBI. Clustal Omega < Multiple Sequence Alignment < EMBL-EBI. [cited 7 Jun 2020]. Available: https://www.ebi.ac.uk/Tools/msa/clustalo/

24. SGRP Blast Server. [cited 7 Jun 2020]. Available: http://www.moseslab.csb.utoronto.ca/sgrp/blast_original/

25. Soriano M. Plasmids 101: Gateway Cloning. [cited 7 Jun 2020]. Available: https://blog.addgene.org/plasmids-101-gateway-cloning

26. Hoffman CS, Winston F. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. Gene. 1987;57: 267–272.

27. NEB Tm Calculator. [cited 11 Jun 2020]. Available: http://tmcalculator.neb.com/#!/main

28. Fields S, Song O-K. A novel genetic system to detect protein–protein interactions. Nature. 1989. pp. 245–246. doi:10.1038/340245a0

29. yeast two-hybrid (Y2H) principle. [cited 11 Jun 2020]. Available: https://www.hybrigenics-services.com/contents/resources/yeast-two-hybrid-principle

30. yeast two-hybrid (Y2H) principle. [cited 11 Jun 2020]. Available: https://www.hybrigenics-services.com/contents/resources/yeast-two-hybrid-principle