



Representing Transcriptional Heterogeneity and Inter-Strain Variation at the *Saccharomyces* Genome Database



Patrick C. Ng, Edith D. Wong, Stacia R. Engel, Robert S. Nash, Marek S. Skrzypek, Suzi Aleksander, Kevin A. MacPherson, Travis Sheppard, Shuai Weng, Kalpana Karra, J. Michael Cherry and the SGD Project

Department of Genetics, School of Medicine, Stanford University, Palo Alto, CA

One of missions of the *Saccharomyces* Genome Database (SGD; <https://www.yeastgenome.org>) is to catalog genome information about *Saccharomyces cerevisiae* that includes variations in DNA sequence and RNA expression. To present a high resolution view of the transcriptional dynamics in *S. cerevisiae*, SGD utilized data generated by collaborators (Pelechano et al., 2013) through Transcript IsoForm Sequencing (TIF-seq) technology. This technique captures the heterogeneity of RNA transcription in *S. cerevisiae*, which can vary widely across experimental perturbations, within different environments, and between cells. The YeastMine data warehouse (<https://yeastmine.yeastgenome.org>) provides downloadable tables mapping these isoforms to the locations of transcripts identified in a curated group of other genomic datasets. Transcript abundance and outer boundaries can be visualized in the JBrowse genome navigator (<https://browse.yeastgenome.org>), and annotation files of all transcripts that fully encompass annotated genes are available from SGD through Amazon Web Services (AWS). The AWS cloud server also hosts variant genotyping files for 25 *S. cerevisiae* strains (Song et al., 2015) assembled by the Automated Genome Analysis Pipeline (AGAPE), and these single nucleotide polymorphisms and small insertions/deletions are also shown in JBrowse. SGD's Variant Viewer (<https://www.yeastgenome.org/variant-viewer/>) presents the molecular consequence of variants within annotated genes and infers evolutionary relatedness across SGD's 12 reference strains. With these resources in place, we seek to incorporate more transcriptome and strain variant data in concert with improving technical capabilities and a growing collection of datasets. This work is supported by a grant from the NHGRI (U41 HG001315).

Select Tracks

My Tracks

Currently Active

Recently Used

Category

20 DNA replication, recombination and repair

3 RNA catabolism

16 RNA structure

9 Reference sequence

1 carbon utilization

84 chromatin organization

102 histone modification

14 mRNA processing

1 mitotic cell cycle

17 nutrient utilization and stress

78 transcription

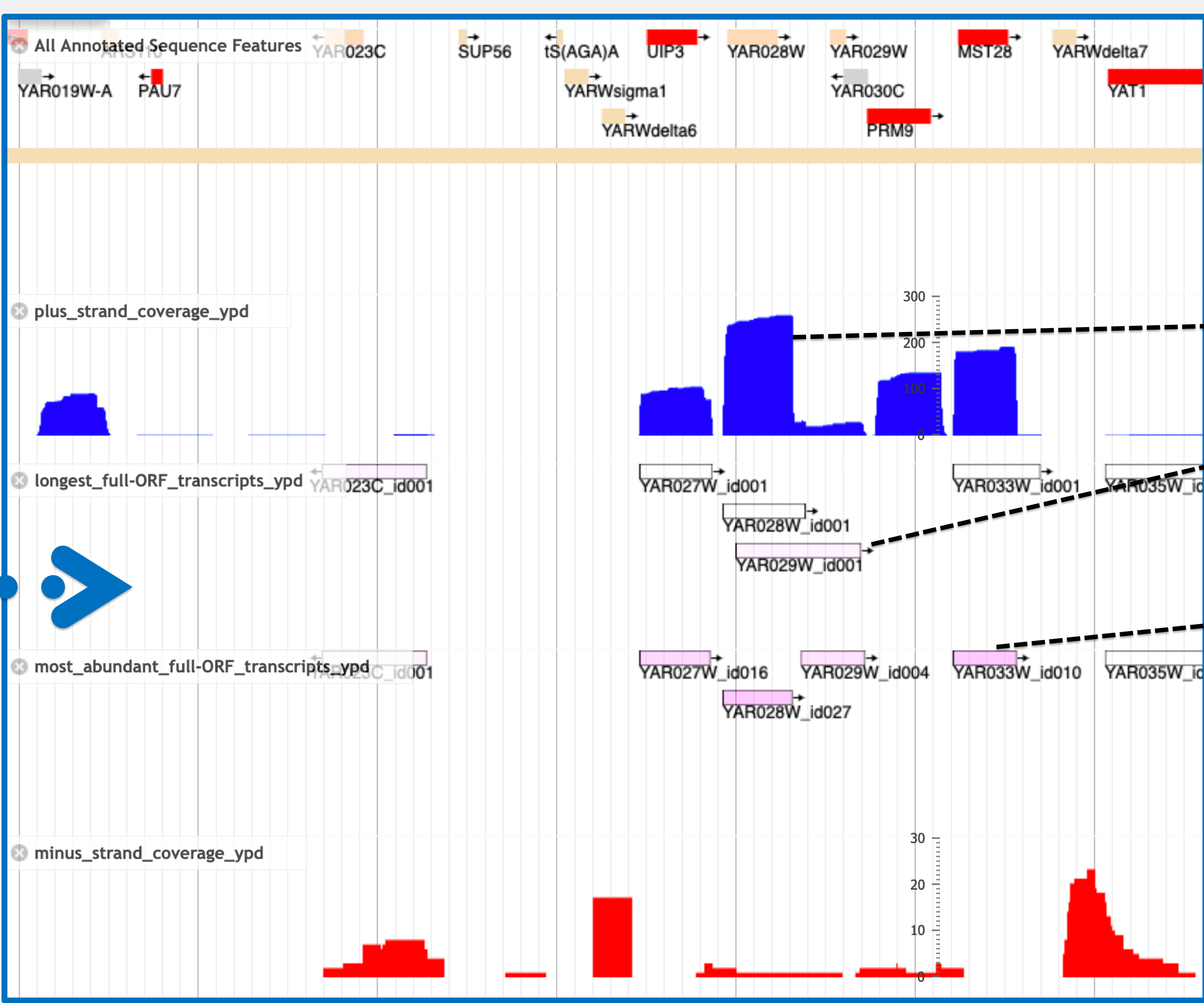
32 transcriptional regulation

14 transcriptome

2 translation regulation

8 transposons

50 variants



Investigate transcripts and variants relative to your gene or strain of interest in **JBrowse**, **YeastMine**, and **Variant Viewer**

Transcript coverage per nucleotide position depicted on plus (blue) and minus (red) strands

Longest and most abundant isoforms pinpointed and shaded according to raw transcript count

primary_transcript

Primary Data

Name

YAR033W_id010

Type

primary_transcript

Position

chrI:188043..188926 (+ strand)

Length

884 bp

Attributes

Id

YAR033W_id010

Seq_id

chrI

Source

SGD

Ypd

22

Region sequence

>chrI chrI:188043..188926 (+ strand)
class=primary transcript length=884
ATGCAGTGGCACAGAAACACCGAGTAATAGTATAGTCAAGCAATACAAGGTATATAT
TGGCATGCGACAGCCCTCCAGAAAGTACCGAGTCAAGTATAGTACACTTAACGAACCT
AGTGCACATTAAATGAGAAAATGTGGCTCTTCCCTAAGACATATTCGCTTCTACT
GAGTTATTGGATCTATGAATCGCTCGCTATACACAGCATGATGTTTGTCCCTGGTAA
TAGGGGTTTGTGTTTATTAATATATTTTAAATGACACAGAGCTTGTGTTTCAAT
TCTGCAATATTGCTTTTACTTCTCTGTAGCTTTGTTAATAATATTAGTGATGTTAA
TCCAAAGCTACTGCTGCTGCAATTTTAGGACCGAGCTTTAGTGGATGTCATCACAC
GTAAACCGCGGTGAAGGGAAGAAATGAGGATCATCACATACATGAACCAATAT
TTGTTAATCATGGCAATGGCATACTCGTATTACTTTTACAGCATGAGGATTGCTA

Transcript isoform dialog box displays raw count in glucose (ypd) or galactose (gal) media, and predicted sequence

a functional protein

RPL4A
Ribosomal 60S subunit protein L4A; N...

validate cytidyltran...ase (CDP-diglyceride synthetase); an enzyme that...

RKM3
Ribosomal lysine methyltransferase; specific for monomethylation of R...

enase/reduct...

dependent protein kinase (PKA) in a TORC1-depe...

chrI:g.299551..299552insGTG
insertion A -> AGTGG

chrI:g.299555..299557del
indel GGAA -> G.GGAAGAGAGCGAA

chrI:g.300034..300040del
deletion GAAAAA -> G.GA

chrI:g.297970A>T

chrI:g.298141C>T

chrI:g.298721G>A

chrI:g.298966T>C

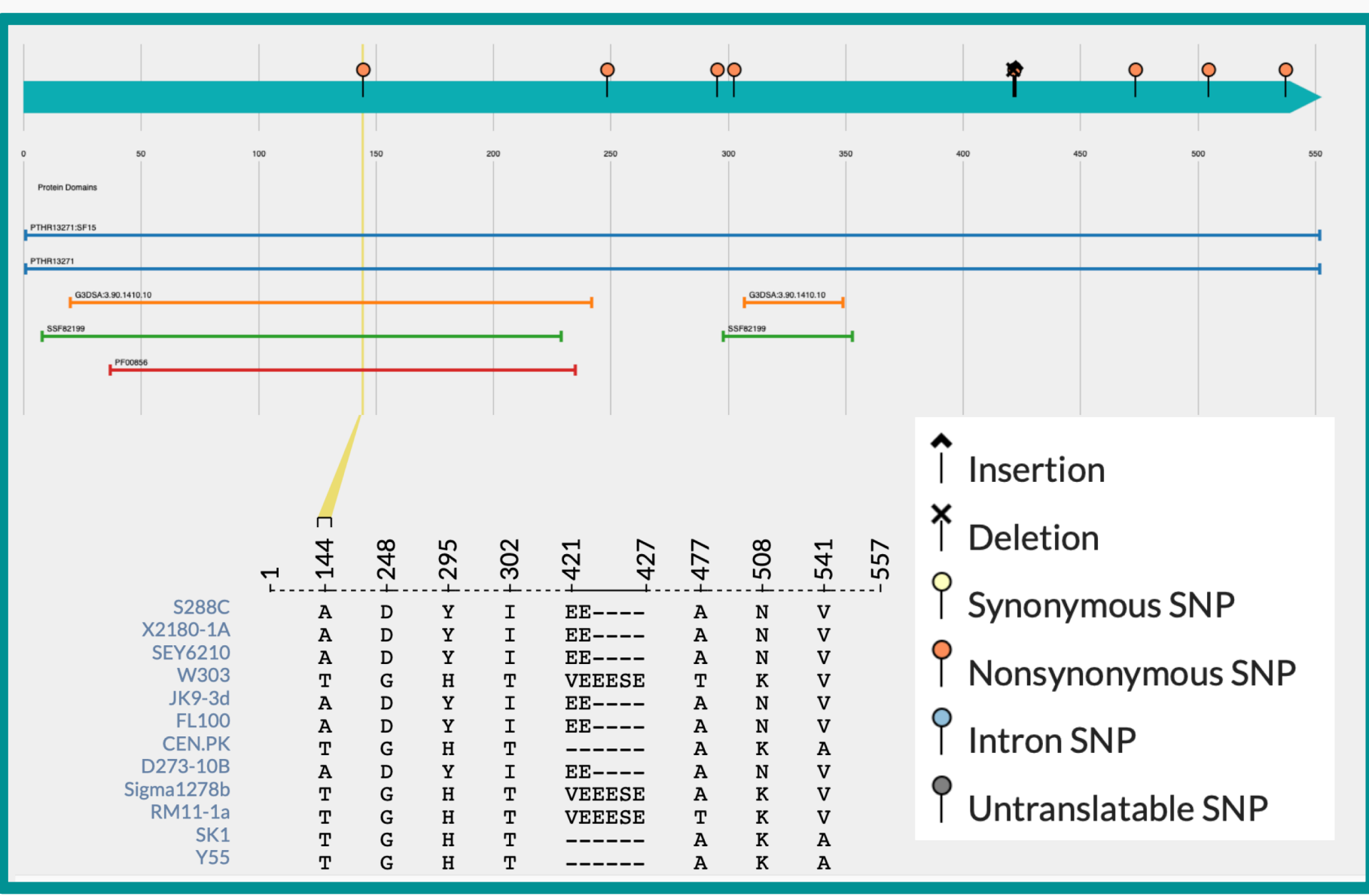
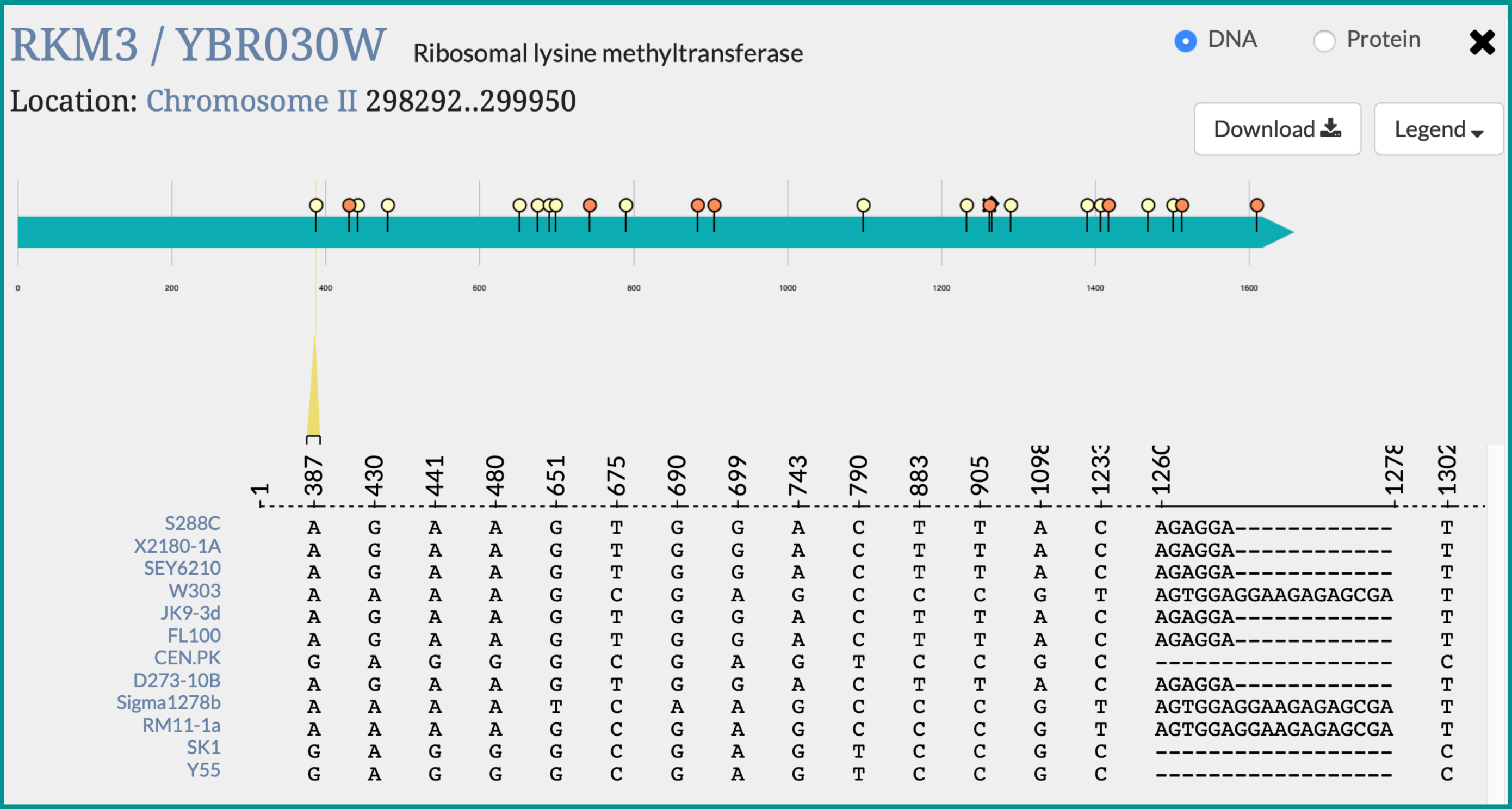
chrI:g.299524C>T

chrI:g.299680T>G

chrI:g.301044C>T

chrI:g.300849C>T

chrI:g.301002T>C



Variant Viewer

Determine nucleotide and peptide sequence alignments of 12 well-studied alternative reference strains for your gene of interest.

SGD

YeastMine

Gene

Transcripts

Retrieve transcript iso-forms for a given gene.

Explore transcript isoforms for your gene of interest that map to other studies, or build your own queries

| Gene Systematic Name | Transcript Chromosome Identifier | Transcript Start Location | Transcript End Location | Transcript Chromosome Strand | Transcripts Five Prime Data Set | Transcripts Three Prime Data Set |
|----------------------|----------------------------------|---------------------------|-------------------------|------------------------------|---------------------------------|----------------------------------|
| YAL003W | chrI | 142153 | 143310 | 1 | Miura_2006 | Miura_2006 |
| YAL003W | chrI | 142153 | 143368 | 1 | Miura_2006 | Nagalakshmi_2008 |
| YAL003W | chrI | 142153 | 143373 | 1 | Miura_2006 | Nagalakshmi_2008 |
| YAL003W | chrI | 142161 | 143199 | 1 | Nagalakshmi_2008 | Miura_2006 |
| YAL003W | chrI | 142161 | 143205 | 1 | Nagalakshmi_2008 | Miura_2006 |
| YAL003W | chrI | 142161 | 143223 | 1 | Nagalakshmi_2008 | Miura_2006 |
| YAL003W | chrI | 142161 | 143368 | 1 | Nagalakshmi_2008 | Nagalakshmi_2008 |
| YAL003W | chrI | 142167 | 143199 | 1 | Zhang_2005 | Miura_2006 |
| YAL003W | chrI | 142167 | 143213 | 1 | Zhang_2005 | Miura_2006 |
| YAL003W | chrI | 142167 | 143223 | 1 | Zhang_2005 | Miura_2006 |
| YAL003W | chrI | 142167 | 143310 | 1 | Zhang_2005 | Miura_2006 |
| YAL003W | chrI | 142167 | 143368 | 1 | Zhang_2005 | Nagalakshmi_2008 |



www.yeastgenome.org
browse.yeastgenome.org
yeastmine.yeastgenome.org



Stanford University