

DEVELOPING A CHIP-SEQ PIPELINE TO DETERMINE FUNCTION OF
CHROMATIN REMODELING AND TRANSCRIPTION RELATED COMPLEXES IN
A MODEL ORGANISM

By

Alejandro Saettone Chipana

Master of Science - Ryerson University, Toronto, Ontario, Canada - 2015

Bachiller en Ciencias Biológicas - Universidad Nacional Mayor de San Marcos, Lima-Perú - 2008

A dissertation presented to Ryerson University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the program of Molecular Sciences

Toronto, Ontario, Canada, 2020

© Alejandro Saettone Chipana, 2020

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A DISSERTATION

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

ABSTRACT

Title: DEVELOPING A CHIP-SEQ PIPELINE TO DETERMINE FUNCTION OF CHROMATIN REMODELING AND TRANSCRIPTION RELATED COMPLEXES IN A MODEL ORGANISM

Student: Alejandro Saettone Chipana

Degree: Doctor of Philosophy in the program of Molecular Sciences

University: Ryerson University

Year: 2020

In eukaryotes, chromatin is the complex of DNA and proteins that packages genomic DNA into chromosomes. Chromatin remodeling is the term used to indicate that the accessibility of DNA within chromatin must be regulated for transcription to happen. Cellular health depends on processes such as chromatin remodeling and transcription occurring in a timely and precise manner. The protein complexes that perform chromatin remodeling and transcription are poorly understood. The ciliate protozoa *Tetrahymena thermophila* has a unique nuclear biology with structurally and functionally distinct nuclei in a common cytoplasm. Recently we used a proteomic approach to characterize the protein complement of *Tetrahymena* several chromatin remodeling complexes that have in common a bromodomain containing protein we have named Ibd1, in addition to a divergent version of the Mediator complex that functions in transcriptional co-activation in yeast and humans. In order to understand the function of these protein complexes, I developed ChIP-Seq for use in *Tetrahymena*.

Chromatin Immunoprecipitation followed by Next Generation Sequencing (ChIP-Seq) is a molecular method widely used in yeast and human cells to investigate the function of chromatin-related proteins by identifying their associated DNA sequences on a genomic scale. ChIP-Seq generates large quantities of data that is difficult to process and analyze, particularly for organisms with contig-based sequenced genomes. Whereas human and yeast genomes have advanced genomic annotations such as, promoter and enhancer regions and transcription binding sites, a contig-based genome typically are poorly annotated or have minimal annotation on their associated set of genes. These annotations are associated coordinates primarily predicted by gene finding programs. Poorly annotated genome sequence, such as *Tetrahymena*'s, makes comprehensive analysis of ChIP-Seq data difficult and as such standardized analysis pipelines are lacking.

In this thesis I designed a ChIP-Seq pipeline for contig-based genomes to complement current proteomic approaches and to determine protein function of two proteins associated with chromatin remodeling and transcription in *Tetrahymena*, Ibd1 and Med31. The bromodomain-containing Ibd1 co-purifies in *Tetrahymena* with four protein complexes that function in yeast and humans in chromatin remodeling. Med31 is a conserved member of the Mediator complex in yeast, humans and *Tetrahymena*. ChIP-Seq analysis of Ibd1 suggests a role for the protein and its associated complexes during transcription, suggesting that it coordinates high levels of transcription of highly expressed genes in *Tetrahymena thermophila*. Furthermore, Med31 ChIP-Seq analysis suggested a global role for the Mediator complex in transcription regulation. Reduced levels of Med31 conduces to ectopic expression of developmental genes important for programmed DNA rearrangements and irreversible gene silencing. The ChIP-Seq computational

pipeline presented in this thesis is an efficient and reliable tool to analyze genome-wide raw ChIP-Seq data generated in model organisms with poorly annotated contig-based genome sequence.

TABLE OF CONTENTS

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A DISSERTATION	ii
ABSTRACT.....	iii
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	xiii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS.....	xvii
1 INTRODUCTION	1
1.1 Expression, Functional and Comparative Proteomics.....	1
1.2 Genomics approaches to aid in functional annotation of evolutionarily conserved proteins of unknown function	2
1.2.1 Next Generation Sequencing, advantages and disadvantages	2
1.2.2 Definition of genic and intergenic regions.....	3
1.2.3 Chromatin Immunoprecipitation coupled to NGS Next Generation Sequencing.....	4
1.3 Tetrahymena thermophila: A Useful Proteomic Model for Nuclear Events.....	7
1.3.1 Nuclear dimorphism.....	8
1.3.2 Conjugation.....	9
1.4 Identification of Core Histones and their Variants.....	10
1.5 Telomeres and Telomerase.....	12

1.6	Programmed DNA Rearrangements.....	15
1.7	Epigenetics	17
1.8	Chromatin remodeling.....	20
1.9	Identification of chromatin remodeling complexes in <i>Tetrahymena thermophila</i> (SWI/SNF, SWR, SAGA and COMPASS)	23
1.9.1	Identification of the SWI/SNF complex in <i>Tetrahymena thermophila</i>	23
1.9.2	Interactive Bromodomain containing Protein 1 (Ibd1).....	24
1.9.3	Ibd1 Recognizes Kac	26
1.9.4	Ibd1 Interacts with Multiple Chromatin Related Proteins, SWR, SAGA and COMPASS.....	27
1.10	Mediator and Transcription	32
1.11	Identification of a Transcription facilitating complex in <i>Tetrahymena thermophila</i> (Mediator)	34
1.11.1	Rpb1 lacks canonical CTD heptad repeats in <i>Tetrahymena</i>	34
1.11.2	Mediator subunit Med31 is conserved in <i>Tetrahymena</i>	36
1.11.3	Purification of <i>Tetrahymena</i> Mediator.....	37
1.11.4	<i>Tetrahymena</i> 's putative mediator complex.....	40
1.12	Rationale.....	43
1.13	Thesis Objective	44
1.13.1	Aim 1: Development of a wet lab protocol for ChIP-Seq using <i>Tetrahymena</i> <i>thermophila</i>	44

1.13.2	Aim2: Development of a ChIP-Seq pipeline to determine function of chromatin remodeling and transcription related complexes in <i>Tetrahymena thermophila</i>	44
1.13.3	Aim 3: Determine the function of Ibd1 by Ibd1-FZZ ChIP-Seq	44
1.13.4	Aim 4: Determining the function of Med31 by Med31-FZZ ChIP-Seq	44
2	Materials and Methods.....	45
2.1	Cell strains.....	45
2.2	Macronuclear gene replacement.....	45
2.3	ChIP-qPCR: Validation of RACS results for Ibd1	46
2.4	RT-qPCR: Validation of RACS results for Med31.....	46
2.5	Generation of cell extracts and Western blotting	48
2.6	Indirect immunofluorescence	48
2.7	Bioinformatics and molecular evolutionary analyses	49
2.8	ChIP-Seq and data analysis	50
2.9	Data Deposition.....	51
2.10	Materials Table	51
3	CHAPTER 1: Development of a ChIP-Seq pipeline to determine function of chromatin remodeling and transcription related complexes in a model organism.....	53
3.1	Introduction to RACS (rapid analysis of ChIP-Seq data for contig based genomes)	53
3.2	WetLab: Chromatin Preparation for Chromatin Immunoprecipitation (ChIP) in <i>Tetrahymena thermophila</i>	54

3.3	Library preparation and NGS	55
3.4	Implementation of RACS: rapid analysis of ChIP-Seq data for contig based genomes	56
3.5	Core Pipeline Tools	58
3.6	Determination of the Genic Regions	58
3.7	Determination of the Intergenic Regions	60
3.8	Post-processing: Normalization of Reads accumulation and Enrichment Calculation..	61
3.9	Utilities: Validation and Quality Checks	63
3.9.1	Visualization of Reads accumulation.....	63
4	CHAPTER 2: ChIP-Seq analysis indicates that Ibd1 localizes to highly expressed gene in <i>Tetrahymena</i>	65
4.1	Ibd1 Localizes to Transcriptionally Active Chromatin.....	65
4.2	Ibd1 Localizes to a set of highly expressed genes	67
4.3	Gene Ontology (GO) annotation analysis of Ibd1 bound genes suggests function in fundamental biological processes	68
4.4	Localization of Ibd1 in <i>Tetrahymena</i> 's genome	69
4.5	Validation of Ibd1-FZZ ChIP-Seq	72
5	CHAPTER 3: ChIP-Seq of Med31 indicates that Mediator regulates activation and repression of genes in <i>Tetrahymena</i>	74
5.1	Med31 chromatin occupancy indicates a global role of Mediator in transcription.....	74
5.2	Med31 function may be related to gene activation and repression	75

5.3	Gene Ontology (GO) annotation analysis of Med31 bound genes suggests function in diverse cellular processes.....	77
5.4	Validation of Med31-FZZ ChIP-Seq	78
5.4.1	Med31 negatively regulates the expression of developmental genes during growth	78
5.4.2	MED31 KD impacts Tetrahymena cellular proliferation	85
5.4.3	Med31 regulates the expression of the developmental gene Pdd1	86
5.5	MED31 KD show a range of DAPI-positive structures that could be related to ectopic expression of development specific genes.....	87
5.6	MED31 KD cells shows γ H2A.X foci in the MAC.....	91
6	DISCUSSION	94
6.1	RACS as a new methodology for analysis of ChIP-Seq data originating from contig-based genomes	94
6.2	Model for Ibd1 function	96
6.3	Model for Med31 function	99
7	FUTURE WORK.....	102
7.1	Testing the model for Ibd1	102
7.2	Characterization of BRD containing proteins in Tetrahymena thermophila.....	104
7.3	Characterization of SAGA activity in Tetrahymena thermophila.....	105
7.4	Characterization of DNA motif associated with Ibd1	106

7.5	Genome wide analysis of Mediator-associated transcribed genes to validate the Med31 model.....	106
7.6	Tetrahymena's Mediator proximity biotinylation mapping	107
7.7	Potential common link between Ibd1 and Med31.....	108
7.8	Characterization of PHD domain containing proteins in Tetrahymena	108
8	CONCLUSIONS.....	110
	APPENDICES	112
	A.1: Chromatin Preparation for Chromatin Immunoprecipitation (ChIP) in Tetrahymena thermophila	112
	A.2: Ibd1ChIP-Seq post-processing files.....	118
	RACS processed data for the 354 genic regions occupied by that Ibd1: Refer to Digital Spreadsheet (TET_Ibd1_MAC_Genome_Genic) for complete results.....	118
	RACS processed data for the 842 intergenic regions occupied by that Ibd1: Refer to Digital Spreadsheet (TET_Ibd1_MAC_Genome_Intergenic) for complete results.	138
	A.3: Med31ChIP-Seq post-processing files.....	150
	RACS processed data for the first 1000 (out of 13715) genic regions occupied by that Med31: Refer to Digital Spreadsheet (TET_Med31_MAC_Genome_Genic) for complete results.	150
	RACS processed data for the first 1000 (out of 5019) intergenic regions occupied by that Med31: Refer to Digital Spreadsheet (TET_Med31_MAC_Genome_Intergenic) for complete results.	206

A.4: Common link between Ibd1 and Med31	220
Common potential genes controlled by Ibd1 and Med31	220
Common potential intergenic regions occupied by Ibd1 and Med31	221
BIBLIOGRAPHY	226

LIST OF TABLES

Table 1: AP-MS data for Ibd1-FZZ, Ada2-FZZ and Swc4-FZZ purified from vegetative cells.	30
Table 2: AP-MS data for MED31 purified from vegetative cells.	39
Table 3: Definitions.....	60
Table 4: ChIP-qPCR validates ChIP-Seq results.	73
Table 5: HTA3 ChIP-qPCR data.	80
Table 6: Med31 ChIP-qPCR data.	81
Table 7: Pdd1 ChIP-qPCR data.	82

LIST OF FIGURES

Figure 1: Chromatin immunoprecipitation followed by NGS summary.....	6
Figure 2: Conjugation (Mating) and Nuclear Development in <i>Tetrahymena thermophila</i> . .	10
Figure 3: Extensive Programmed DNA Rearrangements in <i>Tetrahymena</i>	16
Figure 4. One example of a mechanism of chromatin remodeling.....	23
Figure 5: <i>Tetrahymena</i> TTHERM_00729230 (tlbd1) BRD shows homology to <i>Tetrahymena</i> (tGcn5) and <i>Saccharomyces</i> (yGcn5p, yBDF1, yBDF2 and ySWI2/SNF2) BRDs.	25
Figure 6: Domain architecture of the identified BRD-containing proteins.....	26
Figure 7: Network view of Ibd1 protein-protein interactions.	28
Figure 8. One example of a mechanism for transcription with Mediator complex.	34
Figure 9: CTD diversity in representative protist lineages.	35
Figure 10: Protein phylogeny of Med31 in Protista under LG+G+I model of evolution.	36
Figure 11: Comparative analysis of Med31 domain sequence identity in different species.	37
Figure 12: Med31-FZZ expression and affinity purification.	38
Figure 13: <i>Tetrahymena</i> 's putative mediator complex.	41
Figure 14: Med31-FZZ immunofluorescence.....	43
Figure 15: RACS Core Pipeline overview.	57
Figure 16: Ibd1 Distribution.....	66
Figure 17: <i>Tetrahymena</i> 's expression distribution and Ibd1 localization.....	68
Figure 18: Ibd1_GO.....	69
Figure 19: Ibd1 is localized in promoters, ORF and terminators.	71
Figure 20: ChIP-qPCR validates ChIP-Seq results.....	72

Figure 21: MED31-FZZ ChIP-Seq analysis in growing <i>Tetrahymena</i>	75
Figure 22: Med31 Expression profile.....	77
Figure 23: GO analysis of genes enriched by Med31-FZZ.	78
Figure 24: Med31 regulates the expression of developmental genes in <i>Tetrahymena</i>	79
Figure 25: RT-qPCR analysis to examine the expression levels of Med31.....	82
Figure 26: RT-qPCR analysis of <i>PDD1</i> , <i>PDD2</i> and <i>PDD3</i> expression using RNA extracted either from wild-type cells or from <i>MED31</i> KD cells.	84
Figure 27: Genome browser screenshots for Med31 ChIP-Seq peaks over <i>PDD1</i> , <i>PDD2</i> and <i>PDD3</i> loci.	85
Figure 28: <i>MED31</i> KD impacts <i>Tetrahymena</i> cellular proliferation.....	86
Figure 29: Western blot analysis of Pdd1 expression.	87
Figure 30: Immunofluorescence analysis using either <i>MED31</i> KD or wild-type growing <i>Tetrahymena</i> cells.	88
Figure 31: Quantification of cells with Pdd1 staining in <i>MED31</i> KD and wild-type controls.	89
Figure 32: Quantification of cells with DAPI positive structures upon <i>MED31</i> KD.....	90
Figure 33: Immunofluorescence analysis of <i>MED31</i> KD <i>Tetrahymena</i> cells shows γ H2A.X foci in the MAC.	92
Figure 34: <i>MED31</i> KD results in abnormal γ H2A.X foci in the MAC.	93
Figure 35. Ibd1 localizes to many intergenic regions but accumulates to genic regions.	97
Figure 36: Model for Ibd1 Function recruiting multiple chromatin remodeling complexes.	99
Figure 37. Med31 localizes and accumulates to genic regions.	100

Figure 38: After Sonication, DNA fragments between 300 and 600 base pairs.	113
Figure 39: Recommended WB layout for ChIP experiment.....	115

LIST OF ABBREVIATIONS

Kac	Acetyllysine
ADP	adenosine diphosphate
AP	affinity purification
ATP	adenosine triphosphate
BRD	bromodomain
C-	carboxy
ChIP-Seq	Chromatin immunoprecipitation coupled to Next Generation Sequencing
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
DCL1	dicer-like protein 1
ddH ₂ O	double distilled water
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
FCS	Flowcell summary
FZZ	3x FLAG-ZZ epitope tag
g	grams
GEO	Gene Expression Omnibus
GO	Gene Oncology
H	hours
H2A	histone 2A
H2A.Z	histone 2A variant z

H2B	histone 2B
H3	histone 3
H4	histone 4
HAT	histone acetyltransferase
HMT	histone methyltransferase
HPC	High Performance Computing
HPLC	high pressure liquid chromatography
Ibd1	Interactive Bromodomain Protein 1
IBD1-FZZ	refers to the FZZ strains
Ibd1-FZZ	refers to the tagged protein
IES	internal eliminated sequence
IF	indirect immunofluorescence
IgG	immunoglobulin G
IGR	Intergenic Region
INI1	integrase interactor 1
IP	Immunoprecipitated
IP	immunoprecipitation
k	kilo
KAT	lysine acetyltransferase
KD	knock-down gene
LC	liquid chromatography
Log	logarithmic
LTQ	Linear Trap Quadrupole

M	molar
MAC	macronucleus
MED31-FZZ	refers to the FZZ strains
Med31-FZZ	refers to the tagged protein
MIC	micronucleus
MS	mass spectrometry
MS-AP	mass spectrometry- affinity purification
mt	mating type
N-	amino
n	haploid number
N	Normalized
ncRNA	non-coding RNA
NGS	Next Generation Sequencing
ORF	Open Region Frame
PCR	polymerase chain reaction
PF	Passing Filtering
pH	potential hydrogen
pm-r	paramomycin resistant
PTM	posttranslational modification
RACS	Rapid Analysis of ChIP-Seq data
rDNA	ribosomal deoxyribonucleic acid
RNA	ribonucleic acid
RP-HPLC	reversed-phase high-pressure liquid chromatography

rpm	rotations per minute
SAGA	multi-subunit histone modifying complex
SAINT	significance analysis of interactome
scnRNA	scan RNA
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNF5-FZZ	refers to the FZZ strains
Snf5-FZZ	refers to the tagged protein
SPP	sequestering protease peptone
SWI/SNF	SWItching/Sucrose Non-Fermenting
SWR	histone variant exchange complex
TCA	trichloroacetic acid
TET	<i>Tetrahymena thermophila</i>
TEV	tobacco etch virus
TF	transcription factor
TGD	wiki <i>Tetrahymena</i> genome database wiki
U	unit
UTR	untranslated region
UV	ultraviolet
Veg	vegetative growth
w/v	weight/ volume
WB	Western Blot
WCE	whole cell extract
WT	wild type

1 INTRODUCTION

1.1 Expression, Functional and Comparative Proteomics

The term proteome describes the complete set of proteins in a cell, organelle or organism (James, 1997). Proteomics refers to the study of proteins encompassing the identification, quantification, and characterization of their composition, concentration, structure and activity. Several sub-fields have emerged including expression, functional, and comparative proteomics (Köcher & Superti-Furga, 2007; Rost, 2002). Expression proteomics refers to the identification and differential analysis under different stimuli/developmental stages/disease stage of all proteins present in a particular cell/organelle/organism (Köcher & Superti-Furga, 2007). More recently, these types of proteomic analyses have been extended to include the cataloguing of post-translational modifications (PTMs) on proteins. The importance of assemblies of proteins into functional units (protein complexes) to the cell in directing cellular function is well established (Alberts, 1998). Functional proteomics as an experimental approach relies on “guilt by association” (Köcher & Superti-Furga, 2007). If a protein of unknown function is shown to physically interact with a protein of known function, it can be inferred as a first approximation that the proteins have similar function and specific hypotheses can be derived about the function of the first protein. Affinity purification (AP) exploits the high specificity of antibody-antigen interactions or other affinity reagents by allowing selective purification via an antibody-conjugated chromatographic medium or other affinity matrix of its cognate epitope tag that has been attached to the ‘bait’ protein. Co-purifying proteins that physically interact with the ‘bait’ protein, the so-called set of ‘prey’ proteins, may then be identified via various means, most

commonly mass spectrometry (MS). Affinity purification and mass spectrometry (AP-MS) when combined with high throughput methods for the systematic discovery of protein-protein interaction (PPI), and combined with bioinformatic analysis (Ashraf et al., 2019; Saettone et al., 2018; Shannon et al., 2003; Warde-Farley et al., 2010) has led to the development of comprehensive interaction maps for proteins from many organisms including the prokaryote *Escherichia coli* (Butland et al., 2005) in addition to eukaryotes such as the yeast *Saccharomyces cerevisiae* (Krogan et al., 2006), and human cells (Huttlin et al., 2017). Comparative studies of protein-protein interaction networks among the limited number of eukaryotic species for which proteomic data exists show that many circuits embedded within the protein networks are conserved over evolution (Sharan et al., 2005). Comparative proteomics has the potential to permit functional annotation of evolutionarily conserved proteins of unknown function (Wan et al., 2015).

1.2 Genomics approaches to aid in functional annotation of evolutionarily conserved proteins of unknown function

1.2.1 Next Generation Sequencing, advantages and disadvantages

The advantage of Next Generation Sequencing (NGS) methodologies for researchers is that high-throughput sequencing allows millions of DNA molecules to be read at the same time (Johnson, Mortazavi, Myers, & Wold, 2007; Loman et al., 2012; Schuster, 2008). NGS can help us to understand mechanisms related to transcription (Ng et al., 2009; Z. Wang, Gerstein, & Snyder, 2009). However, the output of NGS is therefore substantial and can be overwhelming for

analyses (Mardis, 2008; Park, 2009). These analyses are facilitated in model organisms that feature well-annotated genomes such as humans and yeast where the genomic sequences are presented in full chromosomal form, and the DNA sequences of which can be found as individual files. These genomes have available annotation files that depict the chromosome-specific DNA base pair coordinates of cis-acting DNA sequences including, open reading frames (ORFs), untranslated regions, transcription start sites, and promoter sequences as well as information about the genes themselves taken from the scientific literature making the interpretation of ChIP-Seq data of transcription proteins more accessible. The difficulties during NGS analyses can be compounded if the genome under study is presented as contig-based (contiguous) sequence assemblies, as is the case in the model ciliate *Tetrahymena thermophila*. A contig-based genome sequence is structured and presented as a basic assembly of consensus regions based on overlapping DNA sequences obtained from DNA sequencing. Contig-based genome sequences are usually available as a conglomerate of individual contigs in a large file. These genome sequences frequently provide files with minimal annotations of predicted genes usually reflecting the lack of available information in the literature.

1.2.2 Definition of genic and intergenic regions

A genic region begins at the start codon ATG and ends at the stop codon TGA. An intergenic region starts after the end of a genic region and ends before the beginning of the next contiguous gene.

1.2.3 Chromatin Immunoprecipitation coupled to NGS Next Generation Sequencing

Chromatin immunoprecipitation coupled to NGS (ChIP-Seq) assess global DNA-binding sites to understand the mechanisms underlying transcription. The main steps of ChIP-Seq (**Figure 1**) are crosslinking of proteins and DNA followed by DNA shearing to obtain smaller fragments of proteins and DNA. Using immunoprecipitation, the Protein of interest (POI), which is normally a protein that has an epitope tag, plus the attached DNA can be separated from the whole cells extract. The precipitated POI is unbound from the DNA and the DNA can be sequenced using NGS (**Figure 1**). ChIP-Seq is used in gene expression studies to make predictions about the function(s) a protein in transcription based on its position within a gene (Johnson et al., 2007; Schmidt et al., 2010). For example, if the Protein of Interest (POI) accumulates within genes rather than intergenic regions, it could be inferred that it might have a direct role in transcription regulation. An enrichment of the ChIP peaks near the 5'UTRs would suggest that the POI likely functions in transcription initiation. On the other hand, accumulation of ChIP peaks at the 3' ends would suggest a role in transcription termination while proteins involved in elongation are typically found throughout the coding region. Note this is only a first approximation since gene expression can also be coordinated by elements that are not in close proximity to the specific gene (Cremer & Cremer, 2001).

To determine POI position(s) within a genome from raw ChIP-Seq data, the files containing gene coordinates are needed. It is important to note that less developed genomes such as that of *Tetrahymena thermophila* provide files containing the predicted coordinates for gene positions as minimum annotation. Current ChIP-Seq applications such as MACS2 (Y. Zhang et al., 2008) do

not directly address whether the accumulation of the POI is in a specific area such as genic or intergenic region. To obtain a genome file that can be used by a software like MACS2 many other computational steps are required. After the initial alignment, the data is typically analyzed by a peak calling software, such as MACS2, which provide with peaks coordinates. The user then needs to further process the peaks obtained with third-party software such as (Quinlan & Hall, 2010) to assess the local enrichment within genic and/or intergenic regions. The advantages of ChIP-Seq include high resolution (single nucleotide) and genome wide coverage. The challenges include experimental design and data analysis and cost. ChIP-Seq produces substantial amounts of data thus computational analyses are critical to discover protein function (Park, 2009).

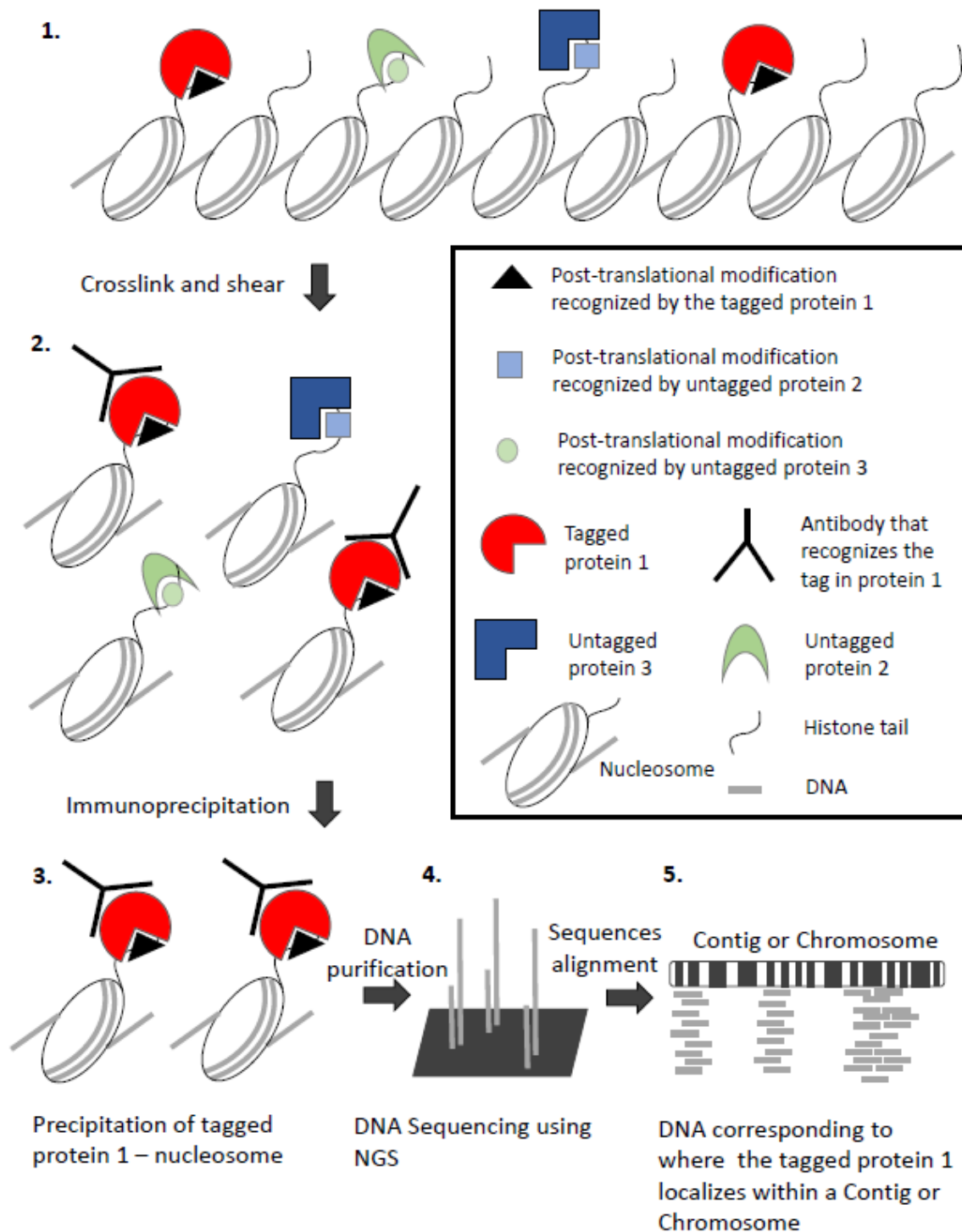


Figure 1: Chromatin immunoprecipitation followed by NGS summary.

1) Native state of chromatin. 2) Specific antibodies recognize the tagged proteins. 3) Isolation of tagged protein plus its interacting chromatin. 4) After DNA purification and library preparation NGS is performed. 5) The output data from NGS is aligned to *Tetrahymena thermophila*'s genome assembly (Saettone, Ponce, Nabeel-Shah, & Fillingham, 2019).

1.3 *Tetrahymena thermophila*: A Useful Proteomic Model for Nuclear Events

Most model species exploited for proteomic analysis to date belong to either the prokaryotes (Hu et al., 2009), fungi (particularly yeast) (Krogan et al., 2006), and animals such as *Drosophila* (Guruharsha et al., 2011), worm (Simonis et al., 2009), and human (Huttlin et al., 2017), members of the Opisthokont lineage of eukaryotes. Large-scale PPI data gaps currently exist for a number of other eukaryotic lineages including protists, and green plants. The protist *Tetrahymena thermophila* is a ciliate, a member of the Alveolate group of eukaryotes, along with dinoflagellates and the parasitic apicomplexans that include *Plasmodium* species that cause malaria.

Tetrahymena grows axenically with a generation time of approximately 2.5 h in inexpensive rich media (Ruehle, Orias, & Pearson, 2016) and large-scale culture is routine. The macronucleus (MAC) and micronucleus (MIC) genomes are sequenced (Eisen et al., 2006; Hamilton et al., 2016) and molecular genetic methods developed for *Tetrahymena* include homologous recombination mediated exact gene replacement in the MAC or MIC allowing gene knockout and epitope tagging (Douglas L Chalker, 2012), in addition to inducible RNAi gene knock down (Akematsu et al., 2018, 2017; Ruehle et al., 2016). The *Tetrahymena* Genome Database (TGD and its associated Wiki, www.ciliate.org), contains the MAC and MIC genome sequences as well as predicted gene models (Stover et al., 2006; Stover, Punia, Bowen, Dolins, & Clark, 2012). The *Tetrahymena* Functional Genomics Database (TetraFGD, <http://tfgd.ihb.ac.cn>) contains

functional genomic data sets relating to gene expression in addition to gene networks (Xiong et al., 2012, 2013).

1.3.1 Nuclear dimorphism

A typical eukaryotic nucleus is composed of regions of transcriptionally inert heterochromatin as well as euchromatic areas which are considered competent for transcription. *Tetrahymena* is a unicellular eukaryote with a variety of unique biological features that have made it a powerful model system for many aspects of molecular and cellular biology (Ruehle et al., 2016), particularly with respect to nuclear biology where it segregates germ-line specific transcriptionally silent, and somatic transcriptionally active chromatin into two structurally and functionally distinct nuclei contained within its single cell (Orias, Cervantes, & Hamilton, 2011). The micronucleus (MIC) is diploid, divides by mitosis and is not transcribed during growth. In addition, the MIC undergoes meiosis during the sexual phase of the life cycle, conjugation, and is analogous to a germ-line nucleus. The macronucleus (MAC) is polyploid, divides amitotically, is transcriptionally active and is not inherited sexually, an analog of a somatic nucleus. The MAC is rich in acetylated chromatin which was critical in the discovery of the first lysine acetyl transferase (KAT) p55/Gcn5 (J. E. Brownell & Allis, 1995; James E Brownell et al., 1996). The different chromatin structures of the MAC and MIC have their origins in the sexual phase (conjugation) of the life cycle (Duane W. Martindale, Allis, & Bruns, 1982).

1.3.2 Conjugation

During the conjugation (**Figure 2**), genome-wide transcription of non-coding RNAs (ncRNAs) occurs that directs large-scale RNAi-dependent assembly of distinct chromatin domains in the new MAC as a prelude to programmed DNA rearrangements including site-specific DNA deletion, analogous to irreversible gene silencing (Douglas L Chalker & Yao, 2011) that are similar to the enigmatic developmentally regulated chromatin diminution that occurs in presomatic cells of the parasitic nematode *Ascaris* (Streit, Wang, Kang, & Davis, 2016).

After pairing, the MIC in each of the two cells undergoes meiosis, generating four haploid meiotic products, only one of which is retained. This gametic nucleus divides mitotically, and one of the two resulting identical haploid nuclei is reciprocally exchanged and fuses with that of its partner to form a genetically identical diploid zygotic nucleus in each cell. The zygotic nucleus divides twice, resulting in four identical products at which point two begin to develop into new MACs (NM). MAC development in the NM of each exconjugant involves extensive programmed DNA rearrangements/irreversible genome silencing that are directly linked to ncRNA-based changes in chromatin structure. These DNA rearrangements include site-specific chromosome fragmentation as well as the deletion of MIC-limited sequences called Internal Eliminated Sequences (IESs), that together result in the loss of ~15% of the germ-line genome (M. C. Yao, Choi, Yokoyama, Austerberry, & Yao, 1984). IES deletion begins with the bidirectional transcription of RNAs from the meiotic MIC (Douglas L. Chalker & Yao, 2001; Schoeberl, Kurth, Noto, & Mochizuki, 2012). Meiosis is the only stage of the *Tetrahymena* life cycle where the MIC is transcribed (Douglas L. Chalker & Yao, 2001; Mochizuki & Gorovsky,

2004a). This meiotic MIC-specific transcription is catalyzed by RNAPII (Mochizuki & Gorovsky, 2004a). A global MIC-specific nuclear run-on analysis showed that meiotic MIC-specific transcription is biased towards IES DNA, implying that initiation/start-site selection of the MIC-specific transcription is regulated and not simply a result of global or random transcription (Gao & Liu, 2012; Schoeberl et al., 2012). The underlying molecular mechanisms underlying any transcription in *Tetrahymena* remain poorly understood.

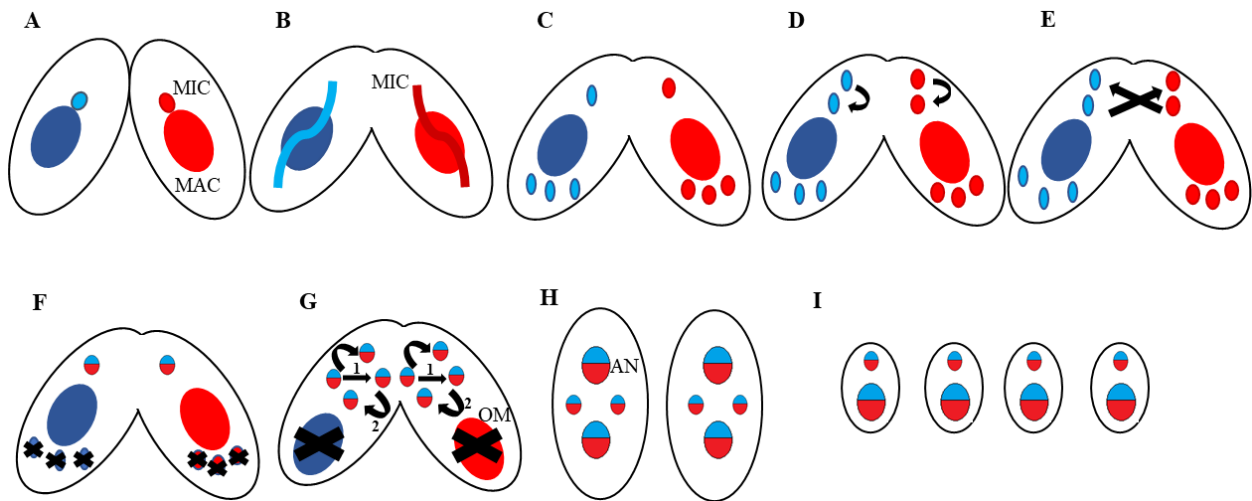


Figure 2: Conjugation (Mating) and Nuclear Development in *Tetrahymena thermophila*.

(A) Cell adhesion occurs at 1.5 h. (B) Meiosis (Prophase I) occurs at 3 h where the micronucleus (MIC) crescent is formed. (C) At 4.5 h, four gametic nuclei are formed. (D) Duplication of one of the posterior gametic nuclei occurs. (E) At 6 h, nuclear exchange followed by fusion of the gametic nuclei occurs, followed by, (F) Formation of the zygotic nucleus and elimination of the other three gametic nuclei. (G) At 8 h two rounds of zygotic duplication occurs, old macronucleus (MAC) (OM) degradation and new developing MAC (Anlagen (AN)) begins. (H) 2 Anlagen 2 MIC stage. (I) At 10 h separation of exconjugants occurs and at this point is when chromosome fragmentation, Internal Eliminated Sequences (IES) deletion and DNA amplification starts (~45°C). At the end of conjugation there are four daughter cells. (Saettone, Nabeel-Shah, et al., 2019)

1.4 Identification of Core Histones and their Variants

The biochemical identification and characterization of *Tetrahymena* core and variant histones are early examples of the utility of expression proteomics. After developing a differential centrifugation-based method to separate MACs and MICs (Gorovsky, 1970b), Gorovsky compared acid-soluble fractions enriched in positively charged DNA-binding chromatin proteins from the two nuclei (Gorovsky, 1970a). Using one dimensional (1D) SDS-PAGE, they observed differences between the two nuclei, providing the first indications that their chromatin composition differed. Allis et al. (Allis, Glover, & Gorovsky, 1979), extended the study by using two dimensional (2D) SDS-PAGE to look at differences among the chromatin proteins and found that the MIC contain two types of histone H3 (H3_S and H3_F), one with a slightly faster (H3_F) gel mobility that was subsequently found to be a consequence of a regulated proteolytic event that removes the first six amino acids from the H3 N-terminus (Allis, Bowen, Abraham, Glover, & Gorovsky, 1980). Other acid-soluble MAC proteins identified in this type of analysis include histone variants hv1 (H2A.Z) and hv2 (H3.3) (Allis, Glover, Bowen, & Gorovsky, 1980), a finding that at the time provided important information concerning the function of these conserved eukaryotic proteins in transcription. When MAC and MIC acid-soluble nuclear proteins were further separated into perchloric acid (PCA) soluble and insoluble fractions, 1D SDS-PAGE expression proteomic analysis led to the discovery of two high mobility group (HMG) proteins (Schulman, Cook, Richman, & Allis, 1987), distinct MAC and MIC specific histone H1 proteins (Allis, Allen, Wiggins, Chicoine, & Richman, 1984; Schulman et al., 1987), and a MAC-specific HP1 (Huang, Wiley, Lending, & Allis, 1998). The MIC specific H1 is unique in that it is processed into different peptides (α , β , γ , and δ) (Allis et al., 1984). To date the function of regulated processing events (MIC-specific H1 and H3_F), as well as the identity of the respective processing protease(s), is unknown.

1.5 *Telomeres and Telomerase*

Proteomic analysis in *Tetrahymena* has yielded information concerning the nature of telomerase. Telomeres are repetitive DNA sequences located at the ends of linear chromosomes, which are required for stability and replication (Blackburn, 1984). It was reasoned that the large number of telomeres in the *Tetrahymena* MAC should require an amount of telomerase that could be detected in cell extracts. Using a novel primer extension based assay, a telomere terminal transferase activity was identified in nuclear extracts from conjugating cells (Greider & Blackburn, 1985). Gel-filtration and ion exchange chromatography were used to partially purify the activity and demonstrate the existence of the co-purifying RNA component (Greider & Blackburn, 1987). A similar approach led to the identification of regulatory proteins p80 and p95 that co-purified with the activity (Collins, Kobayashi, & Greider, 1995). The catalytic subunit proved recalcitrant to identification until similar analysis was performed using a hypotrichous ciliate with millions of gene-sized MAC DNA molecules. The then newly developed nanoelectrospray tandem MS was used to analyze the fraction containing telomerase activity resulting in the identification of p123, the catalytic subunit of telomerase that contained reverse transcriptase motifs (Nakamura, 1997) whose orthologs were soon identified using comparative sequence analysis to design degenerate primers which led to the molecular cloning of the cDNA and genomic locus of TERT in *Tetrahymena* (Collins & Gandhi, 1998). With TERT identified, a rapid and gentle AP strategy based on the newly developed tandem affinity purification (TAP) method of Rigaut et al. (Rigaut et al., 1999) was utilized (Witkin & Collins, 2004) to affinity purify the complex in order to assess the full protein complement. TAP-tagged TERT was

integrated into the *Tetrahymena* genome into the non-essential β -tubulin 1 (BTU1) locus, replacing the coding region in the strain Cu522, a strain engineered so that complete replacement of all copies of BTU1 results in resistance to taxol. Genes inserted into the locus are over-expressed under the control of the strong *BTU1* promoter (Gaertig, Gao, Tishgarten, Clark, & Dickerson, 1999). To demonstrate functionality of the C-terminal tagged TERT, the gene encoding WT *TERT*, which was previously demonstrated to be essential, was deleted from TERT-TAP (*BTU1*) and shown that TERT-TAP could completely replace function of the essential WT TERT (Witkin & Collins, 2004). SDS-PAGE analysis of the tandem affinity purified TERT-TAP revealed four major polypeptides (p20, p45, p65 and p75) in addition to p123/TERT. Molecular analysis of p45 and p65 demonstrated essential roles in the maintenance of telomere length (Prathapam, Witkin, O'Connor, & Collins, 2005; Witkin & Collins, 2004) showing the validity of the approach. Initial attempts to obtain peptide sequences from gel slices containing p75 and p20 were unsuccessful, therefore the authors repeated the AP as a one-step purification to increase yield for sequencing by Edman degradation and were able to identify the corresponding cDNAs (Witkin, Prathapam, & Collins, 2007). Reciprocal purification of over-expressed ZZ-p75 (ZZ epitope tag in tandem protein A domains followed by TEV cleavage site) in a Cu522 background identified p123, p65, p45 and p20, and its genetic analysis revealed a role in telomere homeostasis (Witkin et al., 2007). For these and subsequent AP, MS was used to identify co-purifying proteins guided by gene predictions from the *Tetrahymena* macronuclear genome sequencing project (Eisen et al., 2006) and annotated in the *Tetrahymena* Genome Database (TGD; www.ciliate.org) (Stover et al., 2006). Subsequent AP identified a ubiquitin ligase as the major interacting partner of p20, consistent with its molecular identity as the Skp1 component of SCF (Skp1/Cullin/F-box protein) ubiquitin ligase suggesting that the telomerase-

p20 interaction was a result of over-expression of telomerase subunits (Min & Collins, 2009). A C-terminal FZZ (3xFLAG—TEV cleavage—FZZ) epitope tag was added to p123/TERT at the endogenous locus so that the tagged protein's expression was dependent on the gene's endogenous promoter. The set of p123 co-purifying proteins included p45, p65 and p75 identified previously, in addition to three new proteins, p19, p50 and p82/Teb1 that is paralogous to the Rpa1 large subunit of the general single-stranded DNA binding heterotrimer replication protein A (RPA) (Min & Collins, 2009). Finally, the existence of two additional proteins in a functional telomerase complex was predicted by cryo-EM studies of telomerase before their subsequent identification as Teb2 and Teb3 when TERT-FZZ was affinity purified and analyzed by the more sensitive liquid chromatography–tandem mass spectrometry (LC-MS/MS) (Jiang et al., 2015). AP of Teb2 and Teb3 both co-purified low levels of telomerase activity and subsequent MS identified several telomerase holoenzyme proteins confirming that Teb1, Teb2, and Teb3 forms a module within telomerase (Upton, Chan, Feigon, & Collins, 2017). Unlike Teb1, subsequent AP-MS of epitope tagged Teb2 and Teb3 revealed them to not be telomerase-specific, with Rpa1 and each other as their major binding proteins leading to the suggestion that Rpa1, Teb2, and Teb3 together compose *Tetrahymena* RPA (Upton et al., 2017).

The identification of proteins that bind the *Tetrahymena* telomere began with Pot1a, identified in 2007 by its similarity with that of the human orthologue (Jacob, Lescasse, Linger, & Price, 2007). MS was used to analyze proteins that co-purified with TAP-Pot1, expressed from its endogenous locus. MS analysis of a one-step AP from nuclear extracts found two co-purifying proteins Tpt1 (TPP1/Tpz1 in *Tetrahymena*) and Pat1 (Pot1-associated *Tetrahymena*) (Linger, Morin, & Price, 2011). To confirm these interactions, cell lines were engineered expressing either TAP-tagged Tpt1 (TAP-Tpt1) or Pat1 (TAP-Pat1), recombined into the respective

endogenous gene locus but expressed under the heterologous cadmium-regulated metallothionein (MTT1) promoter which was done in order to insert the selectable marker upstream of each gene due to N-terminal placement of the tag. Conventional pull-downs were performed using Ni-agarose to purify 6x-HIS tagged proteins from whole cell extracts and an antibody that recognizes Pot1 was used in Western blotting of SDS-PAGE to demonstrate the interaction. Conditional knockouts of either genes affected telomere length demonstrating their importance to telomerase activity (Linger et al., 2011). Another study used MS to identify proteins that co-purify with TAP-Tpt1 and TAP-Pat1 (Premkumar et al., 2014). This identified peptides from Tpt1, Pat1, Pot1a, and a new protein named Pat2 (Pot1a-associated *Tetrahymena* 2), suggesting that the proteins present at the 3' overhangs of *Tetrahymena* telomeres exist in a four-protein complex consisting of Pot1a, Tpt1, Pat1, and Pat2. The detailed mechanism by which this complex regulates telomere length remains to be determined.

1.6 Programmed DNA Rearrangements

During MAC development (**Figure 2**), extensive programmed genomic rearrangements occur including fragmentation of chromosomes and concomitant addition of telomeres, DNA amplification, and site specific deletion of ~10–15% of the original genome, comprising ~6000 individual deletion events (M.-C. Yao, Chao, & Cheng, 2014) (**Figure 3**). Proteomics has generated novel insights into the mechanism of these programmed DNA re-arrangements. Using an expression proteomic approach, Madireddi et al. (Madireddi, Davis, & Allis, 1994) briefly pulsed conjugating cells with a radio-labeled amino acid before separating enriched fractions of MIC, old MACs, and new developing MAC (Anlagen). Fluorography of proteins separated on

1D SDS-PAGE identified four abundant proteins (p32/Pdd3, p43/Pdd2, p65/Pdd1 and p85) specifically in the Anlagen, but not in the MIC or the parental MAC (Madireddi et al., 1994). Subsequent molecular cloning and functional analysis of the genes encoding Pdd1, Pdd2 and Pdd3 linked them to the heterochromatin formation and programmed DNA deletion during MAC development (Madireddi et al., 1996; Nikiforov, Gorovsky, & Allis, 2000; Smothers, Mizzen, Tubbert, Cook, & Allis, 1997). Although not conclusively addressed in the literature, it is probable that the fourth protein, p85, is Twi1, a PIWI/PAZ protein that was found independently to link the RNAi pathway to heterochromatin formation and programmed DNA deletion (Mochizuki, Fine, Fujisawa, & Gorovsky, 2002; Taverna, Coyne, & Allis, 2002).

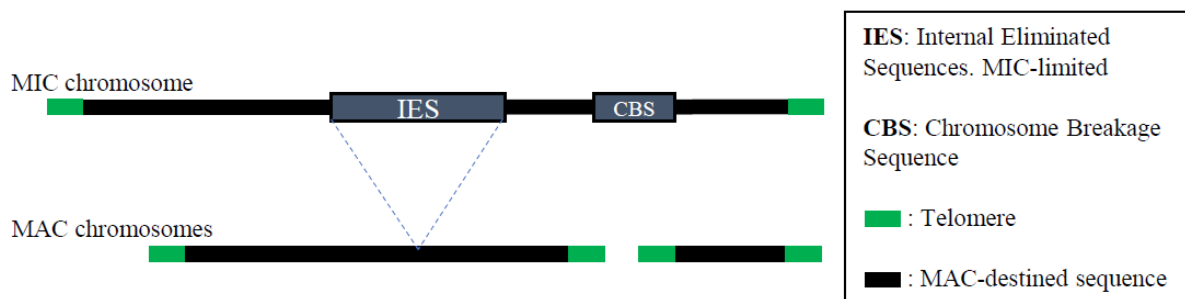


Figure 3: Extensive Programmed DNA Rearrangements in *Tetrahymena*.

These events occur during MAC development where ~10–15% of MIC genome is deleted. See text for additional details (Saettone, Nabeel-Shah, et al., 2019).

Twi1 was initially identified and its gene cloned based upon its similarity to members of the PPD gene family defined by conserved PAZ and Piwi domains (Mochizuki et al., 2002). Somatic knockout analysis of *TWI1* functionally implicated it in the production of conjugation specific small ~28 nt RNAs named scan RNA (scn RNA) and programmed DNA deletion (Mochizuki et al., 2002). The scan RNA model was developed to account for the phenotype of the *TWI1* knockout (Mochizuki et al., 2002) linking scan RNAs, Twi1 and heterochromatin and excision of

IESs, while accounting for their previously demonstrated epigenetic regulation (D L Chalker & Yao, 1996). A functional proteomic approach was adopted to further understand how Twi1p and scnRNA interact (Mochizuki et al., 2002). AP-MS of Flag-HA-Twi1p expressed from its endogenous locus during conjugation, identified *EMA1*, a DExH box RNA helicase (Aronica et al., 2008). Functional analysis of *EMA1* showed it to be required for selective down-regulation of scnRNAs homologous to MAC-destined sequences which are then unavailable to template deletion when transported to the anlagen. Three additional proteins were found to co-purify with Flag-HA-Twi1p. Wag1 and CnjB interact with Twi1 and each other (Bednenko et al., 2009). Interestingly, *CNJB* had previously been identified by sequencing of a conjugation-specific cDNA library (D W Martindale & Bruns, 1983). Molecular analysis of *CNJB* and *WAG1* demonstrated that expression of both was required for proper excision of selected IESs (Bednenko et al., 2009). The final Flag-HA-Twi1 interacting protein Giw1 was found by performing AP-MS using gentle cell lysis prior to AP favoring the isolation of cytoplasmic specific proteins (Noto et al., 2010) and its molecular analysis demonstrated a function in Twi1 nuclear localization (Noto et al., 2010). AP-MS of Pdd1, Pdd2 or Pdd3 has not been reported but will likely result in the identification of new proteins involved in IES deletion/programmed DNA rearrangements/irreversible genome silencing.

1.7 Epigenetics

Proteomic studies have revealed insight into the identity of histone PTMs. In a 1985 study, conjugating cells were pulse-labeled with [³H] lysine for 30 min from 5.0 to 5.5 h of conjugation, and MIC were purified, acid-extracted, and histones analyzed on long acid-urea gels by

fluorography (Allis, Chicoine, Richman, & Schulman, 1985). It was observed that newly synthesized histone H3 and histone H4 are deposited as mono or di-acetylated species. In several subsequent studies, distinct histone H4-acetyl or H3-acetyl species were separated, excised and eluted from an acid urea gel before being micro-sequenced to demonstrate that lysines 4 and 11 of histone H4 (H4K4/H4K11, analogous to H4K5/H4K12 in yeast/humans) (Chicoine, Schulman, Richman, Cook, & Allis, 1986), and H3K9 and H3K14 (Sobel, Cook, Perry, Annunziato, & Allis, 1995) are acetylated in *Tetrahymena* associated with new histone synthesis and deposition. A similar approach demonstrated that H4K7, analogous to H4K8 in yeast/humans, is associated with transcription in the MAC (Chicoine et al., 1986). These *Tetrahymena* proteomic studies represent some of the first evidence that histone acetylation occurred in a dynamic and non-random manner. A similar approach was used to demonstrate that MAC H1 (Hho1) possessed five distinct phosphorylation sites in growing but not starved *Tetrahymena* (Mizzen et al., 1999). Identification of the sites permitted their functional analysis through mutation revealing a role for H1 phosphorylation in the regulation of transcription (Dou, Mizzen, Abrams, Allis, & Gorovsky, 1999; Dou, Song, Liu, & Gorovsky, 2005). A proteomic study provided early important information concerning histone methylation in *Tetrahymena* and its link to transcription. Histone H3 from ³H-AdoMet-labeled MAC was micro-sequenced to demonstrate that lysine 4 is the major site of methylation in MAC in growing cells (B D Strahl, Ohba, Cook, & Allis, 1999), which was the first correlation between H3K4me and active transcription.

The development of MS was a powerful tool leading to the identification of new histone PTMs. One study examined PTM profiles on histone H3 from several model organisms, including *Tetrahymena*, yeast, mouse, and human (Garcia et al., 2007). Whole cell (MAC and MIC)

histone H3 was purified by reversed phase chromatography and analyzed by MS. The study found that *Tetrahymena* and budding yeast both possessed a preponderance of H3K4me1, -me2, and -me3 when compared to human cells (Garcia et al., 2007). *Tetrahymena* was also found to possess H3K56ac (Garcia et al., 2007), a modification that was first reported in *S. cerevisiae* in a MS analysis of its histone PTMs (Masumoto, Hawke, Kobayashi, & Verreault, 2005; Recht et al., 2006). In contrast to budding yeast, *Tetrahymena* was found to have very little H3K79me but was found to contain H3K27me (Garcia et al., 2007). H3K27me3 was subsequently found to be required for programmed DNA deletion that occurs during MAC development in *Tetrahymena* conjugation (Y. Liu et al., 2007).

Individual and distinct histone PTMs have well established roles in certain biological processes, notably regulation of transcription (Brian D Strahl & Allis, 2000). A MS approach relying on electron capture dissociation (ECD), a high energy fragmentation technique, was used in combination with a second method termed proton transfer charge reduction (PTR) to identify combinations of histone PTMs on the same MAC H3 molecule associated with transcription (Taverna et al., 2007). More specifically, the first 50 amino acids of H3 (H3₁₋₅₀) were sequenced MS using ETD/PTR to reveal that H3K27me2 is more correlated with hypo-acetylated H3 and can co-occur with H3K4me1, whereas H3K4me3 was found to be correlated with hyper-acetylated H3 (Taverna et al., 2007). A related study aimed to correlate histone PTM patterns across 15 different conditions from different physiological states and to mutations of key histone modifying enzymes (C. Zhang et al., 2014). Five specific chromatin states were identified; each associated with a set of specific histone PTMs that likely map to distinct functions. To identify MIC specific histone PTMs, MIC were isolated and total histone extracted and separated by reverse-phase HPLC (Papazyan et al., 2014). H3_S and H3_F were analyzed using pH gradient

hydrophilic interaction liquid chromatography (HILIC) directly coupled to high-resolution MS using electron transfer dissociation (ETD) (Papazyan et al., 2014). Both displayed hypo-acetylation, enrichment for H3K27me without H3K4me. Known PTMs found included H3K27me₃, and the mitotic PTM histone H3 serine 10 phosphorylation (H3S10ph), were only detected on H3_F. Additionally, methylation of H3K23, a novel PTM was identified. H3K23me₃ in particular was found to be exclusive to H3_F and occurred in combination with H3K27me. Functional analysis of the PTM revealed Ezl3 to be the histone methyltransferase (HMT) in addition to having a functional role in meiosis (Papazyan et al., 2014).

1.8 Chromatin remodeling

Eukaryotic cells possess multiple levels of regulation of mRNA transcription by RNA polymerase II. Many co-activators of transcription exert their function through chromatin modifying activities. In budding yeast, the SAGA histone acetyl transferase complex co-activates transcription by acetylating specific lysine residues in the N-terminus of histone H3 within the nucleosome, which can then serve as a platform to recruit the SWI/SNF complex via the bromodomain (BRD) present in SNF2/Brg1 (Hassan et al., 2002). The BRD specifically binds acetyl lysine (Kac) within proteins such as histones (Jain & Barton, 2017). When recruited to a genomic region, the SWI/SNF complex co-activates transcription in part by hydrolyzing ATP via the Snf2 subunit and remodeling nucleosomes to make promoter sequences available to be bound by general transcription factors (TFs) such as TFIID. Some other histone modifying complexes that function in promoting transcription include the NuA4 histone acetyl transferase that acetylates nucleosomal H4 (Ginsburg et al., 2014), and the Set1 and Set2 histone methyl

transferases that methylate nucleosomal H3K4 and H3K36 (B. D. Strahl et al., 2002) respectively.

Chromatin remodeling (CR) is typically performed by large multi-protein complexes and is required for essential DNA transactions such as transcription, replication, and repair to occur. Mechanisms of CR involve the ATP-dependent histone sliding (e.g., SWI/SNF, and INO80) (Clapier, Iwasa, Cairns, & Peterson, 2017) or selective insertion of histone variants (e.g., SWR complex (Krogan et al., 2003) inserts H2A.Z) in addition to histone PTMs that include acetylation, methylation, and phosphorylation (Berger, 2007; Biterge & Schneider, 2014; Gardner, Allis, & Strahl, 2011). Acetylation occurs on lysine residues and is catalyzed by KATs (Agalioti, Chen, & Thanos, 2002) and removed, or ‘erased’, by lysine deacetylases (KDACs) (Muller, Filippakopoulos, & Knapp, 2011). KATs and KDACs that acetylate/deacetylate histone substrates were initially termed histone acetyltransferases/deacetylase (HATs/HDACs) since histones were their first substrates identified (Richters & Koehler, 2017). Histone acetylation occurs either at the nucleosomal level (SAGA (Baker & Grant, 2007) and NuA4 (Allard et al., 1999) complexes) or on histones prior to their deposition into chromatin (Hat1 (Parthun, 2007), Rtt109 (Berndsen & Denu, 2008)). Bromodomains (BRD) (Jain & Barton, 2017) recognize, or ‘read’ lysine acetylation in order to condense chromatin, regulate transcription or organize protein complexes. Additional protein domains that function in transcription complexes by recognizing some of the diverse histone PTMs include the methyl lysine-recognizing PHD and chromodomains (Steunou et al., 2016; Taverna et al., 2006). ATP-dependent CR complexes that function in transcription include the SWR complex that exchanges core H2A in the nucleosome for the transcription-friendly histone H2A variant Htz1 (X. Liu, Li, & Gorovsky, 1996;

Papamichos-chronakis, Watanabe, Rando, & Craig, 2011) and the INO80 complex, one function of which is to catalyze the reverse reaction (Poli, Gasser, & Papamichos-Chronakis, 2017).

Proteomic methods using *Tetrahymena* have made significant contributions to understanding fundamental CR mechanisms. The first molecular cloning of the gene encoding a KAT took advantage of *Tetrahymena*'s polyploid MAC specialized for transcription, a source of abundant KAT activity (Vavra, Allis, & Gorovsky, 1982). An acetyltransferase activity in-gel assay was developed (Hutchcroft, Anostario, Harrison, & Geahlen, 1991) to directly identify a catalytically active HAT polypeptide of 55kD from MAC extracts by virtue of its ability to transfer [³H]-acetate from [³H]-acetyl CoA to core histone H3 incorporated within a polyacrylamide gel (J. E. Brownell & Allis, 1995). The protein was purified via conventional chromatography using an in-gel HAT assay and proteolytic analysis resulted in the identification of several internal peptide sequences. Molecular cloning of the gene and comparison of its derived amino acid sequence with available databases revealed it to be highly similar to the yeast protein Gcn5, a transcriptional co-activator implicated in mediating transcriptional activation by acidic activators such as yeast Gcn4p (Georgakopoulos & Thireos, 1992; Guarente, 1995). In budding yeast, Gcn5 was found to be the catalyst of the SAGA histone acetyl transferase complex, co-activating transcription by acetylating specific lysine residues in the N-terminus of histone H3 within the nucleosome, which can then serve as a platform to recruit the SWI/SNF complex via the BRD present in SNF2/Brg1 (Reisman, Glaros, & Thompson, 2009). When recruited to a genomic region, the SWI/SNF complex co-activates transcription by remodeling nucleosomes to make promoter sequences available to be bound by general transcription factors (GTFs) such as TFIID (**Figure 4**).

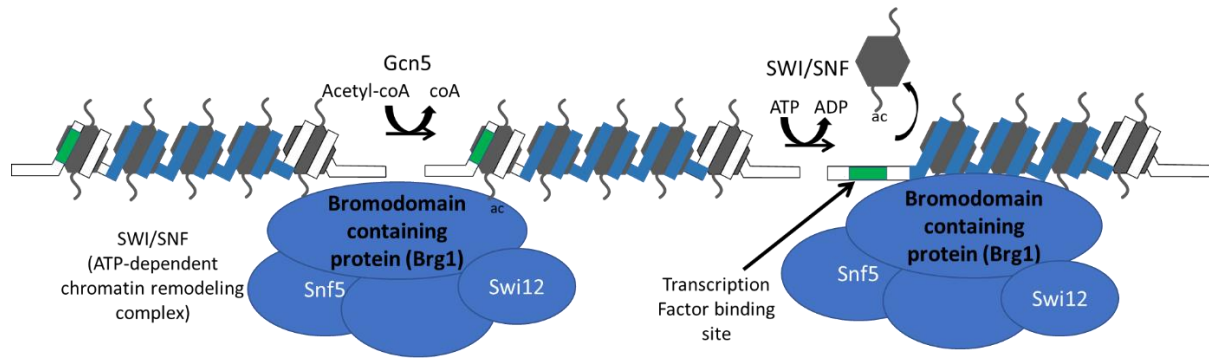


Figure 4. One example of a mechanism of chromatin remodeling

The ATPase subunit (Brg1) of most eukaryotic SWI/SNF complex contains a bromodomain (BRD) that recognizes the PTM Acetyllysine. There are likely many other mechanisms that are unknown and others that are poorly understood. Knowledge gaps exist concerning the role of post-translation modifications and other ATP-dependent chromatin remodeling activities in other organisms including parasitic protozoa.

1.9 Identification of chromatin remodeling complexes in *Tetrahymena thermophila* (SWI/SNF, SWR, SAGA and COMPASS)

1.9.1 Identification of the SWI/SNF complex in *Tetrahymena thermophila*

A SNF2-related gene in *Tetrahymena thermophila* was characterized by (J. Fillingham et al., 2006b). Despite high primary sequence similarity of *Tetrahymena*'s Brg1 to the budding yeast Snf2 and human Brg1 through most of the protein, Brg1 does not possess a recognizable BRD, and its C-terminal region, unlike the entire protein, is dispensable for growth and development (J. Fillingham et al., 2006b) raising the possibility that SWI/SNF functions independently of histone acetylation.

In order to better understand function of SWI/SNF, AP-MS was performed on a conserved component of SWI/SNF in yeast and human cells, Snf5 (Saettone et al., 2018). MS data resulting

from AP of Snf5-FZZ expressed from its native locus was curated using SAINTexpress (Teo et al., 2014). The data revealed Snf5 to co-purify with Brg1, and *Tetrahymena* orthologues of fundamental SWI/SNF proteins Swi1, Swi3, and Snf12 in addition to four other proteins without clear orthologues, Saf1-4, in other described SWI/SNF complexes, along with two novel proteins with clear chromatin-interacting domains Saf5 and Ibd1 (Saettone et al., 2018). Saf5 possesses two tandem plant homeodomains (PHD domain), one function of which is to mediate specific interactions with methylated lysine on histone proteins (Shi et al., 2006). Ibd1 possesses a single C-terminal BRD suggesting that the *Tetrahymena* SWI/SNF complex functions downstream of histone acetylation (at least in growing cells) despite the lack of a C-terminal BRD on Brg1/Snf2.

Of note, three of the novel proteins that co-purify with *Tetrahymena* SWI/SNF (Saf2, Saf3 and Saf4) do not possess conserved domains outside of glutamine-rich regions. Yeast and mammalian Swi1 possess an AT-rich Interactive (ARID) and also a Q-rich domain (Mao, Cao, Nie, Liu, & Chen, 2006). Swi1 possesses an ARID but not a Q-rich domain. The data suggest that in SWI/SNF, the Q-rich proteins Saf2, Saf3, and Saf4 act in conjunction with Swi1. The Q-rich domain in animal Sp1 functions as an activation domain for transcription factors through recruitment of general transcription factor(s) (Emili, Greenblatt, & Ingles, 1994). The data suggest that the function of Saf2-4 is to function in co-activation by recruiting general transcription factors and/or RNA polymerase to promoter regions of highly expressed genes in growing *Tetrahymena*.

1.9.2 Interactive Bromodomain containing Protein 1 (*Ibd1*)

Ibd1 is a bromodomain containing protein that was identified to be a component of SWI/SNF during vegetative growth but not during conjugation (Saettone et al., 2018). The BRD of Ibd1 shows including aminoacid level homology to the bromodomains of *Tetrahymena* and *Saccharomyces* bromodomain containing proteins (**Figure 5**).

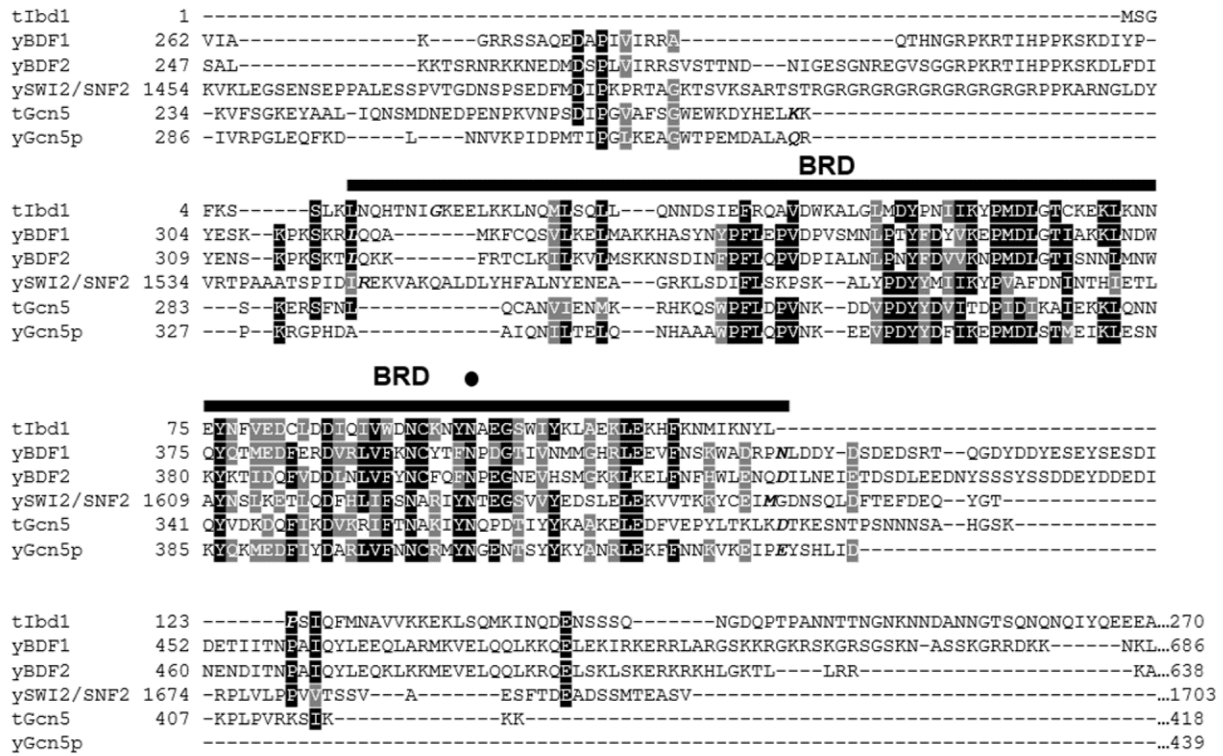


Figure 5: *Tetrahymena* TTHERM_00729230 (tIbd1) BRD shows homology to *Tetrahymena* (tGcn5) and *Saccharomyces* (yGcn5p, yBDF1, yBDF2 and ySWI2/SNF2) BRDs.

Clustal Omega Multiple Sequence Alignment from EMBL-EBI, BoxShade from ExPASy and SMART from EMBL-HEIDELBERG were used for this figure. The start and end of the bromodomains are as follows: tIbd1 BRD:17-123; yBDF1 BRD1:124-139;BRD2:314-423; YBDF2 BRD1:135-241;BRD2:319-428; ySWI2/SNF2 BRD:1547-1657; tGcn5 BRD:281-389; yGcn5p BRD:325-433 (SMART). Absolutely conserved residues are represented by a black shade and conserved residues are represented by a grey shade (BoxShade). The last number of last line represents the last residue for each given sequence. The black circle represents the highly conserved Asparagine (N) that makes contact with acetyl-lysine.

Recombinant Ibd1 recognizes several Kac marks on histones that are correlated with active transcription in *Tetrahymena* including acetylated H3K9 and H3K14, acetylated H2AK9 and

H2AK13 and tri-acetylated H4K5, H4K8 and H4K12 (Saettone et al., 2018). AP-MS analysis of Ibd1 revealed it to interact with the protein complexes SWI/SNF, SWR, SAGA, as well as with a novel putative H3K4-specific histone methyltransferase (**Figure 7**). There are 14 additional BRD-containing proteins encoded in the *Tetrahymena* genome (**Figure 6**) (Saettone et al., 2018).

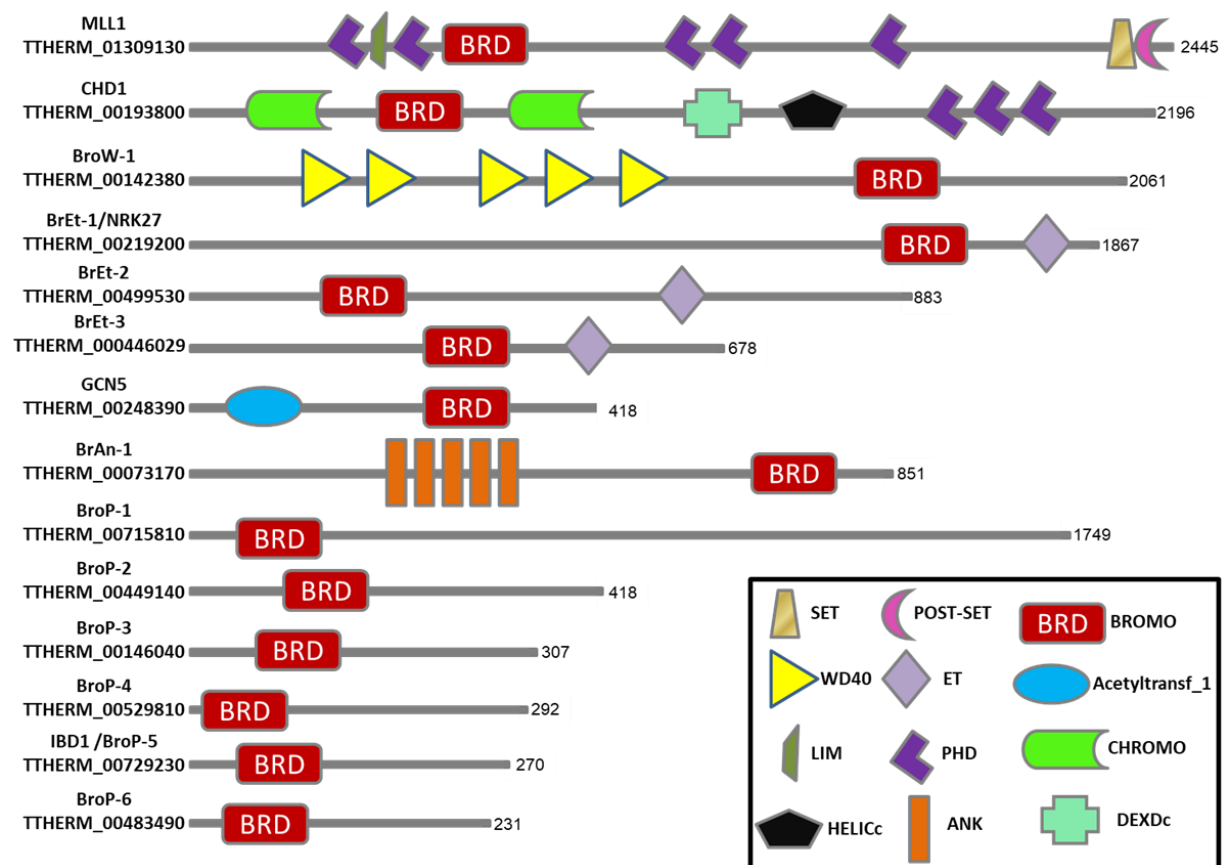


Figure 6: Domain architecture of the identified BRD-containing proteins. Domains were predicted using the SMART web tool and Pfam domain analysis (Saettone et al., 2018).

1.9.3 Ibd1 Recognizes Kac

Our finding of a distinct BRD-containing protein in SWI/SNF is consistent with the fact that a BRD in the catalytic subunit (Snf2/Brg1) has important functions in eukaryotic SWI/SNF complexes. The primary sequence of the BRD of Ibd1 was aligned to those of Gcn5, yGcn5p, yBDF1, yBDF2 and ySWI2/SNF2 which are functional BRD-containing proteins (**Figure 5**) (Saettone et al., 2018). The alignment showed a number of conserved amino acids in the BRD including the highly conserved asparagine (N) that makes contact with Kac (Filippakopoulos & Knapp, 2014; Filippakopoulos et al., 2012) suggesting that Ibd1 as other BRD-containing proteins is likely to bind this mark similar to that observed for Snf2/Sth1 in yeast (Hassan et al., 2007) and Brahma/Brg1 in humans (Shen et al., 2007). Recombinant 6xHIS-Ibd1 was expressed, purified, and incubated with a commercially available peptide array that includes a large number of possible histone post-translational modifications, including many histone acetylation sites. Recombinant 6xHIS-Ibd1 displayed strong specificity for acetylated H3K9 and H3K14, acetylated H2AK9 and H2AK13 and tri-acetylated H4K5, H4K8 and H4K12 (Saettone et al., 2018), which are all acetylation patterns associated with the transcriptionally active MAC in *Tetrahymena thermophila* (Chicoine, Richman, Cook, Gorovsky, & Allis, 1987; Taverna et al., 2007). When incubated on the same peptide array, control recombinant histone methyltransferase 6xHIS-G9a recognized mono- and di-methylated H3K9 (Saettone et al., 2018), as previously demonstrated (Sampath et al., 2007).

1.9.4 *Ibd1 Interacts with Multiple Chromatin Related Proteins, SWR, SAGA and COMPASS*

A stable line expressing Ibd1-FZZ from its MAC locus was generated. The *IBD1-FZZ* tagging construct (Saettone et al., 2018) was used to transform growing *Tetrahymena thermophila* strains

using biolistic transformation. After selection and phenotypic assortment (Merriam & Bruns, 1988), Western blotting demonstrated expression of Ibd1-FZZ in whole cell extracts of transformed strains. Similar to Snf5-FZZ, Ibd1-FZZ also co-purifies with Brg1 as assessed by Western blotting of affinity purified material. Gel-free LC-MS/MS based analysis on affinity purified proteins identified 28 high confidence Ibd1-FZZ co-purifying proteins (Saettone et al., 2018). Comparison of the interaction partners recovered from the purification of Snf5-FZZ, Saf5-FZZ and Ibd1-FZZ interacting proteins (**Figure 7, Table 1**), showed 11 common proteins that co-purify with Ibd1, Saf5 and Snf5 including Swi1, Swi3, Snf5, Snf12 and Brg1, Ibd1, and Saf1-5 that together form a *Tetrahymena thermophila* SWI/SNF complex.

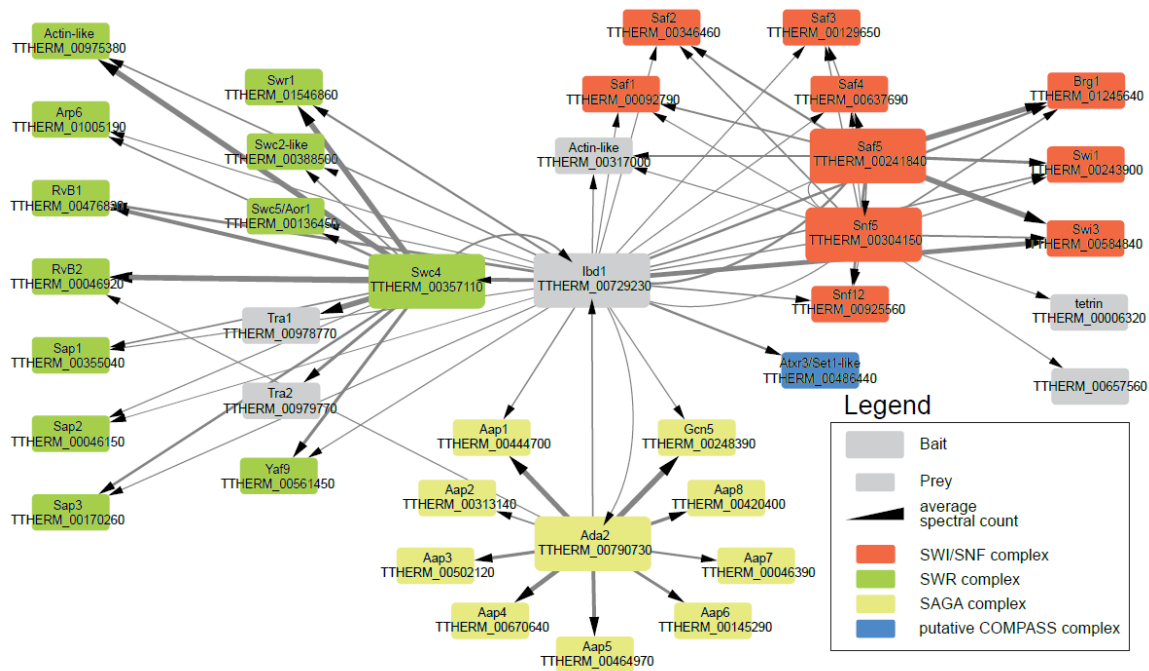


Figure 7: Network view of Ibd1 protein-protein interactions.

The edge thickness represents the averaged spectral counts for the prey. Bait proteins are shown in larger nodes which are colored according to predicted complexes as indicated (Saettone et al., 2018).

The other 17 high-confidence Ibd1 interacting proteins (**Figure 7, Table 1**) could be divided in 3 groups, based on similarity to predicted *S. cerevisiae* orthologs: 1] the SAGA histone acetyl transferase co-activator complex containing Gcn5, Ada2 and a Phd-containing protein, designated Aap1 (**Ada2-Associated Protein 1**), 2] the SWR ATP-dependent chromatin remodeling complex that in yeast and human cells deposits histone variant Htz1/H2A.Z onto chromatin (Swr1, Yaf9, Rvb1, Rvb2, Swc2 and Swc4, Swc5 (C-terminal BCNT domain), two actin-like, and three predicted **Swc4-Associated Proteins** (Sap1-3), one of which possess an AT-hook (Sap1), the other two (Sap2 and Sap3) contain no recognizable domains; and 3] a putative H3K4 methyl transferase (Atrx3/Set1-like). Sap3 shares similarity only on a small portion of the protein with hypothetical proteins in *P. tetraurelia* and *Pseudocohnilembus persalinus*. Sap4 shares similarity throughout the entire protein with a hypothetical protein in *P. persalinus*. The Ibd1 protein therefore appears to be a component of several chromatin remodeling complexes (SWI/SNF, SAGA, SWR) and one containing an Atrx3/Set1-like HMT (COMPASS).

To further delineate the Ibd1 protein interaction network, separate stable lines expressing Ada2-FZZ and Swc4-FZZ from their respective MAC loci following an identical strategy as outlined above were generated. SAINTexpress analysis of AP-MS data from growing cells showed that Ada2 co-purifies with Ibd1 in addition to the Ibd1-interacting Aap1 and Gcn5. Additionally, Ada2 co-purified with three PHD domain-containing proteins (Aap2, Aap3 and Aap4; **Figure 7, Table 1**) and four *Tetrahymena thermophila*-specific hypothetical proteins (Aap5, Aap6, Aap7 and Aap8; **Figure 7, Table 1**) that did not find to co-purify with Ibd1-FZZ. It is suggested that the Ada2 interacting proteins together represent a *Tetrahymena* SAGA complex (**Figure 7, Table 1**).

SAINTexpress analysis of Swc4-FZZ AP-MS revealed it to co-purify a subset of Ibd1 interacting proteins that were predicted to be SWR complex proteins (**Figure 7, Table 1**). Swc4-FZZ further interacts with *Tetrahymena thermophila* orthologs of the Tra1 and Tra2 PI3 kinases (**Figure 7, Table 1**), neither of which co-purified with Ibd1. In yeast, Swc4 co-purifies with Tra1 via the NuA4 histone acetyltransferase complex of which Swc4 is a component, in addition to SWR-C. Swc4-FZZ was not observed to co-purify with any protein that would indicate it to be a member of a *Tetrahymena thermophila* NuA4 complex. The set of proteins that are hypothesized to constitute SWR are listed in **Table 1**. Although the *Tetrahymena thermophila* genome encodes a predicted orthologue of Swc6/Vps71 (TTHERM_01298590), it was not found to co-purify with Swc4 or Ibd1 in growing cells (Saettone et al., 2018).

Table 1: AP-MS data for Ibd1-FZZ, Ada2-FZZ and Swc4-FZZ purified from vegetative cells.

Curated SAINTexpress data from 2 biological replicates of Ibd1-FZZ, Ada2-FZZ and Swc4-FZZ AP-MS samples (Saettone et al., 2018).

TTHERM	Gene Name	Spectral Count Sum Ibd1 (BAIT)	Spectral Count Sum Swc4 (BAIT)	Spectral Count Sum Ada2 (BAIT)	Possible complex	Notes
TTHERM_00729230	<i>IBD1</i>	827	99	58	All listed below	Bromodomain-containing protein
TTHERM_00486440	<i>Atr3/SET1-</i> like	125	---	---	COMPASS	---
TTHERM_00584840	<i>SWI3</i>	411	---	---	SWI/SNF	---
TTHERM_01245640	<i>BRG1</i>	202	---	---	SWI/SNF	SNF2 catalytic subunit
TTHERM_00925560	<i>SNF12</i>	81	---	---	SWI/SNF	---
TTHERM_00243900	<i>SWI1</i>	92	---	---	SWI/SNF	---
TTHERM_00304150	<i>SNF5</i>	64	---	---	SWI/SNF	---
TTHERM_00092790	<i>SAF1</i>	47	---	---	SWI/SNF	Transmembrane protein, putative
TTHERM_00346460	<i>SAF2</i>	39	---	---	SWI/SNF	Hypothetical protein – 13% Glutamine
TTHERM_00637690	<i>SAF4</i>	36	---	---	SWI/SNF	Hypothetical protein – 31% Glutamine
TTHERM_00129650	<i>SAF3</i>	24	---	---	SWI/SNF	Hypothetical protein – 26% Glutamine
TTHERM_00241840	<i>SAF5</i>	29	---	---	SWI/SNF	PHD finger containing protein
TTHERM_00317000	Actin-like	63	---	---	Undefined	---
TTHERM_00046920	<i>RVB2</i>	174	893	---	SWR	---
TTHERM_00476820	<i>RVB1</i>	121	333	---	SWR	---
TTHERM_01546860	<i>SWR1</i>	113	561	---	SWR	---
TTHERM_00975380	Actin-like	101	542	---	SWR	---
TTHERM_01005190	<i>ARP6</i>	18	82	---	SWR	---
TTHERM_00170260	<i>Sap3</i>	23	144	---	SWR	Hypothetical protein

TTHERM_00357110	SWC4	33	419	---	SWR	---
TTHERM_00136450	SWC5/AOR1	32	121	---	SWR	Bucentaur or craniofacial development containing protein
TTHERM_00355040	Sap1	20	78	---	SWR	AT-hook containing protein
TTHERM_00388500	SWC2-like	19	75	---	SWR	---
TTHERM_00561450	YAF9	15	200	---	SWR	---
TTHERM_00046150	Sap2	9	32	---	SWR	Hypothetical protein
TTHERM_00978770	TRA1	---	767	---	Undefined	---
TTHERM_00979770	TRA2	---	179	---	Undefined	---
TTHERM_00444700	Aap1	51	---	356	SAGA	PHD finger containing protein
TTHERM_00248390	GCN5	32	---	605	SAGA	---
TTHERM_00790730	ADA2	28	---	429	SAGA	---
TTHERM_00145290	Aap6	---	---	152	SAGA	---
TTHERM_00313140	Aap2	---	---	69	SAGA	PHD finger containing protein
TTHERM_00670640	Aap4	---	---	314	SAGA	PHD finger containing protein
TTHERM_00502120	Aap3	---	---	184	SAGA	PHD finger containing protein
TTHERM_00420400	Aap8	---	---	194	SAGA	---
TTHERM_00046390	Aap7	---	---	93	SAGA	---
TTHERM_00464970	Aap5	---	---	247	SAGA	---

The function of the SWR complex in fission (H. Kim et al., 2009) and budding (Krogan et al., 2003) yeasts is the deposition of the histone H2A variant Pht1/Htz1 (H2A.Z in humans, and Hv1 in *Tetrahymena*). Deposition of Htz1 in budding yeast is linked to NuA4-dependent histone acetylation via the BRD-containing Bdf1 subunit of SWR (Raisner et al., 2005). In yeast, Bdf1 is also a component of TFIID linking histone acetylation to pre-initiation complex assembly (Matangkasombut, Buratowski, Swilling, & Buratowski, 2000). In *Tetrahymena*, Ibd1 did not co-purify with any proteins similar to components of the general transcription apparatus. Like Ibd1, Hv1 is localized to transcriptionally active MAC in growing cells (Allis et al., 1986). Unlike Ibd1, Hv1 localizes also to the crescent MIC corresponding to meiotic prophase (Stargell et al., 1993), a time period in *Tetrahymena* where large genome-wide transcription of the MIC by RNAPII occurs (reviewed in (Schoeberl & Mochizuki, 2011)).

In budding yeast, SWR is functionally linked to the NuA4 histone acetyl transferase complex via shared subunits Swc4 and Yaf9. In *Tetrahymena*, Swc4 did not co-purify with a histone acetyl transferase subunit and may not be a member of a NuA4-type complex. In fact, a strict NuA4

type complex in *Tetrahymena* is unlikely to exist, despite the presence of 3 genes encoding MYST family histone acetyl transferases. A previous study did identify a H2A/H4 nucleosomal HAT similar to the activity of NuA4 but also showed by glycerol gradient analysis that the activity purifies at ~80kD (Ohba et al., 1999). Consistent with this observation, the MAC does not appear to encode a gene that is a clear ortholog of the conserved NuA4 subunit such as Epl1/EPC so it is unclear if there exists a “piccolo” NuA4 (Chittuluru et al., 2012). Swc4 did co-purify with orthologs of Tra1 and Tra2 kinases that did not purify with Ibd1 (**Table 1**). In *S. cerevisiae* Tra1 co-purifies with NuA4 (Allard et al., 1999) and SAGA (Grant, Schieltz, Pray-Grant, Yates III, & Workman, 1998) that contribute to their co-activator function (Knutson & Hahn, 2011). It will be interesting to determine whether SAGA fulfills the function of SAGA and NuA4 in budding yeast or whether there exists a divergent version of NuA4 in *Tetrahymena*.

Ibd1 co-purifies with Gcn5 and Ada2 in addition to the PHD domain-containing A2A1. Ada2 co-purifies with these proteins in addition to seven others including three additional PHD domain-containing proteins A2A2-4. Thus, Ada2 co-purifies with four distinct PHD domain-containing proteins. Further work will be necessary to determine whether the set of Ada2 interacting proteins represent a single assemblage or if Ibd1, Ada2 and Gcn5 represent a ‘core’ to the *Tetrahymena* SAGA complex that can have different specificity depending on which PHD protein it is interacting with at a particular time.

1.10 Mediator and Transcription

RNAPII-dependent transcription in eukaryotic cells begins when a transcription factor (TF) binds a specific enhancer DNA and, in conjunction with coactivators, specifies the recruitment of the GTFs and ultimately RNAPII to the promoter of the gene. Mediator is a multiprotein complex with 26 core subunits in humans (21 in *Saccharomyces cerevisiae*) and 29 subunits as the “Cdk8-Mediator” complex (25 subunits in *S. cerevisiae*) (Jeronimo & Robert, 2017; Julie Soutourina, 2017). Mediator has been shown to be important for mammalian development (Yin & Wang, 2014) and is often mutated or deregulated in many developmental disorders, as well as cancers (Clark, Oldenbroek, & Boyer, 2015; Schiano et al., 2014). Mediator serves as a bridge between gene-specific activators and RNAPII to initiate transcription (**Figure 8**) (J. Soutourina, Wydau, Ambroise, Boschiero, & Werner, 2011; Thompson & Young, 1995). Because many of its subunits were initially uncovered in genetic screens for suppressors of a growth defect associated with truncation of the Rpb1 CTD (Hengartner et al., 1995; Liao et al., 1995; Nonet & Young, 1989), Mediator has been linked to and extensively studied in eukaryotes that have an RNAPII with a CTD carrying conserved heptad repeats (consensus Tyr1–Ser2–Pro3–Thr4–Ser5–Pro6–Ser7) (Y. Yang, Li, & Qu, 2016)

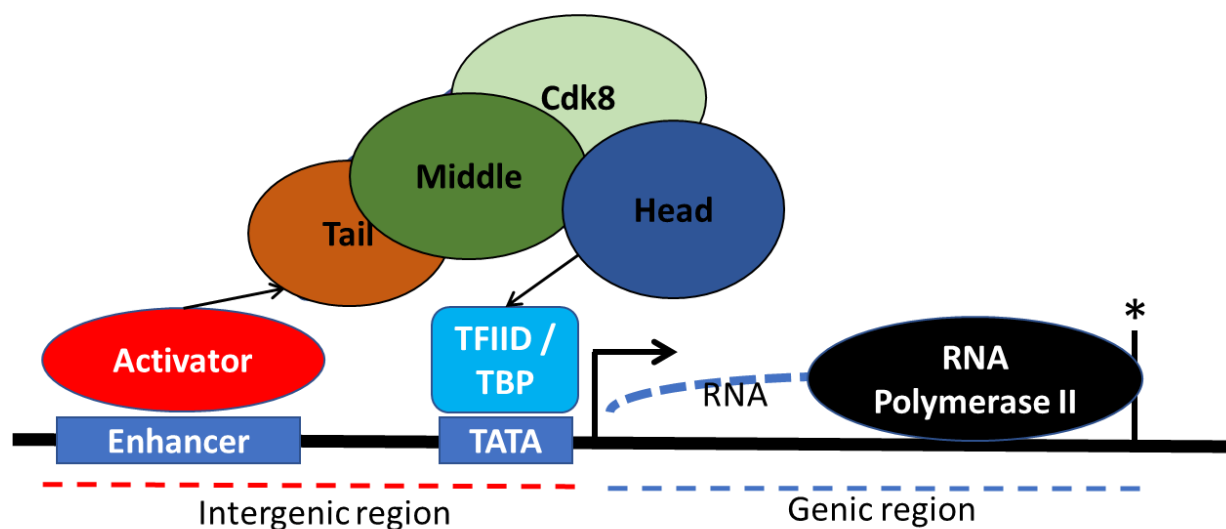


Figure 8. One example of a mechanism for transcription with Mediator complex.

The transcription factor binding sites (regulatory regions) within DNA are exposed. Activator will bind its corresponding regulatory region. Mediator will interact with the activator and recruit transcription factors to recruit RNAPII. The gene's DNA sequence will be copied to make an RNA molecule. Multi-protein complex is a transcriptional co-activator that mediates communication between enhancer, transcription factors and RNAPII. Critical transcription facilitator that is well studied in yeast and humans (knowledge gap in protozoa).

1.11 Identification of a Transcription facilitating complex in *Tetrahymena*

***thermophila* (Mediator)**

1.11.1 *Rpb1* lacks canonical CTD heptad repeats in *Tetrahymena*

Mediator function remains mostly uncharacterized in evolutionarily deep branching eukaryotes (Lee et al., 2010), such as ciliates, which lack canonical CTD heptad repeats.

Tetrahymena thermophila, a unicellular eukaryotic model system, provides an excellent system to study the functional conservation or diversification of chromatin-related and gene expression regulatory complexes (Ashraf et al., 2019). The *Tetrahymena* genome encodes a Rpb1 (human *POLR2A*; yeast *RPO21*) that lacks canonical heptad repeats at its CTD, a feature that is shared among all ciliates (**Figure 9**). These observations are consistent with a previous evolutionary study suggesting loss of CTD heptad repeats among deeply branching eukaryotes (C. Yang & Stiller, 2014). Furthermore, our multiple sequence alignment analysis revealed that the *Tetrahymena* Rpb1 CTD contains 24% fewer serines, 41% fewer threonines, 27% fewer tyrosines, and half the number of prolines compared to yeast, suggesting a relative depletion of phosphorylatable residues. High throughput phosphoproteomic analysis has however identified

phosphorylated serine and tyrosine residues present within the *Tetrahymena* CTD (Tian et al., 2014).

Mediator is conserved across eukaryotes and is thought to have played important role(s) during eukaryotic diversification (Bourbon, 2008; Karijolic & Hampsey, 2012). Mediator subunits, however, have diverged significantly in terms of their identity and/or sequence similarity (Bourbon, 2008), suggesting lineage-specific functional evolution. Despite the lack of canonical CTD heptad repeats, Mediator subunit Med31 was identified in several protist genomes (Linder & Gustafsson, 2004), which prompted us to identify Mediator and analyze its function in the divergent eukaryote *Tetrahymena thermophila*.

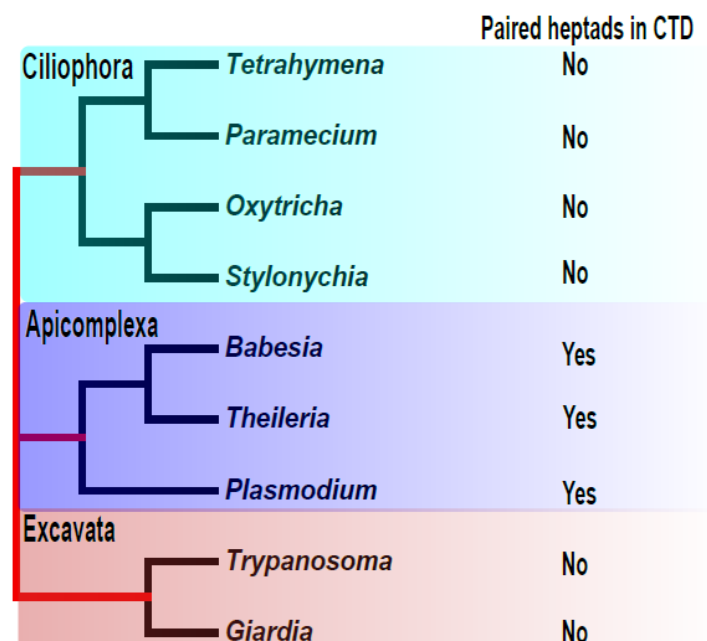


Figure 9: CTD diversity in representative protist lineages.

The tree represents the species relationship based on their taxonomic groupings (Adl et al., 2012). Different taxonomic groups are highlighted in different colors. The presence or absence of canonical CTD heptad repeats is indicated in the right margins based on previous findings (C. Yang & Stiller, 2014) (Garg et al., 2019).

1.11.2 Mediator subunit Med31 is conserved in Tetrahymena

Med31 (Soh1) is the most conserved Mediator subunit (Bourbon, 2008; Linder & Gustafsson, 2004) and has been shown to function as a protein-bridge to connect head and middle modules of the Mediator (Guglielmi et al., 2004). Gene encoding only *MED31* in *Tetrahymena* by sequence similarity using budding yeast and human Mediator subunits as queries were identified. Multiple sequence alignment analysis revealed a high degree of sequence similarity between *Tetrahymena* Med31 and its human/yeast counterparts (Garg et al., 2019), suggesting functional conservation between the orthologs. A phylogenetic analysis using a panel of deeply branching protist lineages established the monophyly of Med31 (**Figure 10**).

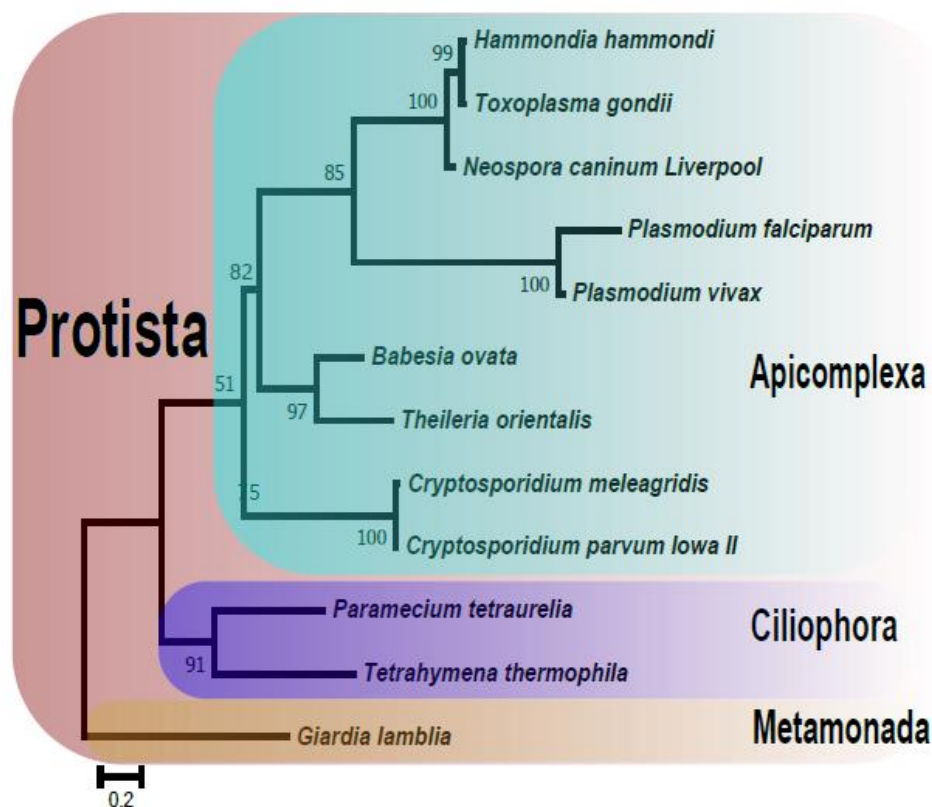


Figure 10: Protein phylogeny of Med31 in Protista under LG+G+I model of evolution. Different taxonomic groups are highlighted in colors. Tree topology represents the ML estimations based on 1000 bootstrap replicas (confidence value only reported when at least $\geq$

50%) and is rooted with *Giardia lamblia*. The scale bar indicates the number of substitutions per site (Garg et al., 2019).

The sequence identities among the *Tetrahymena*, human, *S. cerevisiae* and *S. pombe* Med31 proteins (**Figure 11**) were calculated. *Tetrahymena* Med31 shares 42%, 35% and 25% sequence identities in its core region with human, *S. pombe* and *S. cerevisiae* proteins, respectively. In contrast, *S. cerevisiae* Med31 has 37% and 32% sequence identity with human and *S. pombe*, respectively. *S. pombe* Med31 shares 48% identity to its human counterpart.

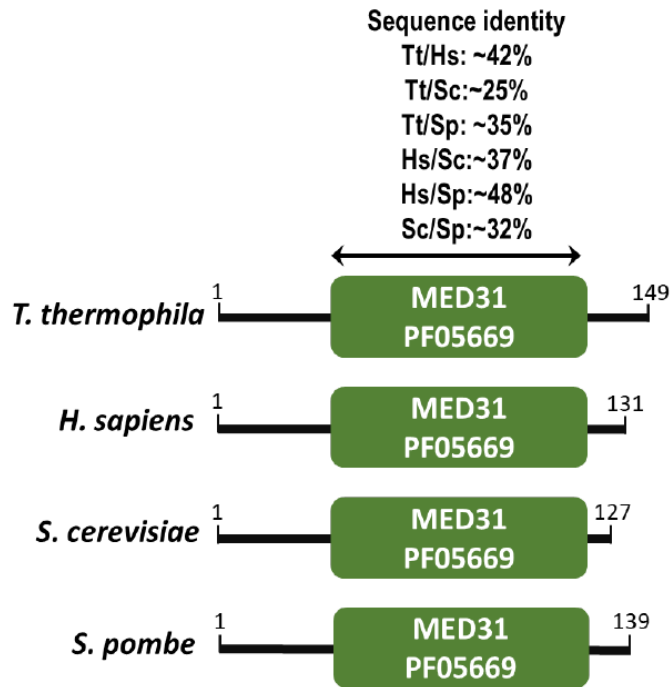


Figure 11: Comparative analysis of Med31 domain sequence identity in different species. The sequence identities identified for the core domain region (pfam domain: PF05669) are indicated above the diagram (Garg et al., 2019).

1.11.3 Purification of *Tetrahymena* Mediator

In order to functionally characterize the Mediator in *Tetrahymena*, a cell line expressing *MED31* with a C-terminal FZZ epitope tag from its native macronuclear (MAC) chromosomal locus was

generated. Affinity purification (AP) of Med31-FZZ and untagged control lines coupled to LC-MS/MS (MS) analysis was performed to identify the *Tetrahymena* Mediator complex. The resulting MS data were curated using SAINTexpress which utilizes semi-quantitative spectral counts for assigning a confidence value to individual protein-protein interactions (PPIs) (Teo et al., 2014). The SAINTexpress analysis of seven biological replicates of Med31-FZZ AP-MS data filtered against numerous control purifications identified 27 high-confidence co-purifying proteins ($\leq 1\%$ false discovery rate (FDR)) (**Figure 12, Table 2**).

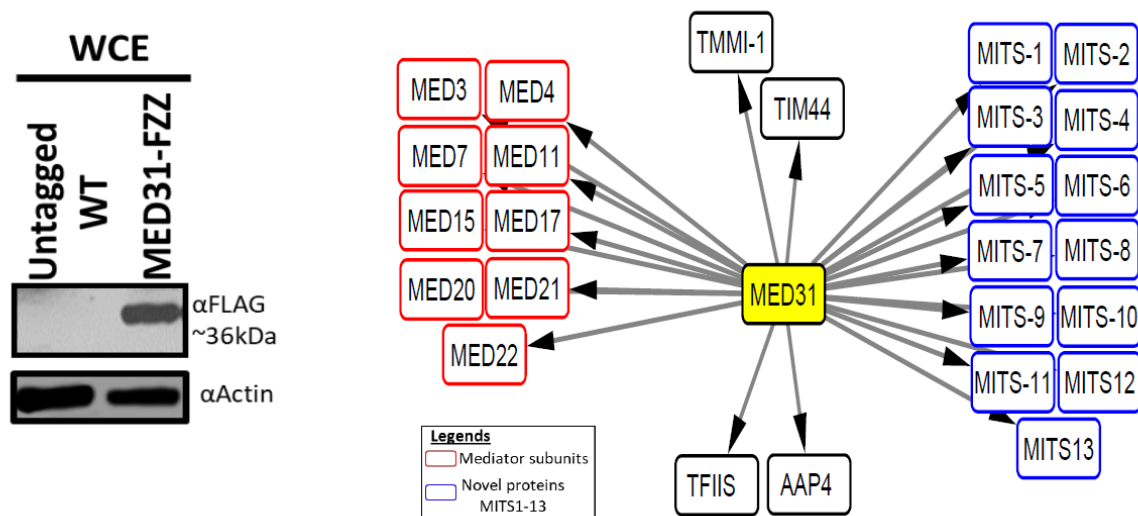


Figure 12: Med31-FZZ expression and affinity purification.

Left - Expression analysis of Med31-FZZ in comparison to the untagged controls by Western blotting using whole cell extracts (WCEs). The top panel was probed with anti-FLAG to detect Med31-FZZ and the bottom panel presents anti-Actin as loading control. **Right** - Med31-FZZ protein-protein interactions network (FDR ≤ 1%). Bait is shown in yellow. Prey node borders are colored according to their putative protein complexes. Network legend is given in the box (Garg et al., 2019).

Med31-FZZ co-purifying proteins include several *Tetrahymena* orthologs of human and yeast Mediator subunits Med 3, 4, 7, 11, 15, 17, and 20-22 that were identified using signature sequence motifs (Bourbon, 2008). Our analysis also identified 13 hypothetical proteins without

any recognizable domains. These proteins do not share similarity to any known protein and appear to be *Tetrahymena*-specific. These proteins were named as **Med31-Interacting *Tetrahymena*-specific proteins (MITS 1-13) (Figure 12, Table 2)**. Among Med31-FZZ co-purifying proteins, the transcription factor IIS (TFIIS) was also identified as a high confidence interacting protein. Med31 and TFIIS in yeast have been reported to genetically interact with each other (Malagon, Tong, Shafer, & Strathern, 2004). Yeast TFIIS is a component of the RNAPII preinitiation complex, and its association with gene promoters has been shown to depend on the Mediator and SAGA complexes (B. Kim et al., 2007). Additional Med31-FZZ co-purifying proteins include a ciliate-specific transmembrane protein (TMMI-1), Tim44, and the PHD containing protein Aap4, that co-purifies with the Ada2 (Saettone et al., 2018).

Table 2: AP-MS data for MED31 purified from vegetative cells.
Curated SAINTexpress data from 7 biological replicates (Garg et al., 2019).

TTHERM	Names	Spectral Count Sum Med31 (BAIT)	Notes
TTHERM_00052180	MITS-1	864	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00780600	MED17	727	Conserved Mediator domain was identified
TTHERM_00752180	MITS-3	534	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00052189	MED15	459	Conserved Mediator domain was identified
TTHERM_00295380	MED21	435	Conserved Mediator domain was identified
TTHERM_00657539	MITS-5	426	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00147570	MED7	417	Conserved Mediator domain was identified
TTHERM_01014520	MITS-2	397	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00490630	MED3	391	Conserved Mediator domain was identified
TTHERM_00670380	MED22	346	Conserved Mediator domain was identified
TTHERM_00989470	MITS13	341	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00922930	MED20	330	Conserved Mediator domain was identified
TTHERM_00316620	MITS-4	290	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00691210	MED4	282	Conserved Mediator domain was identified
TTHERM_00355460	MED31	218	Med31
TTHERM_01002760	MITS-11	206	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00419920	MITS-6	184	MED31-interacting <i>Tetrahymena</i> specific protein (MITS)

TTHERM_00918460	MED11	165	Conserved Mediator domain was identified
TTHERM_00028490	MITS-9	162	MED31-interacing <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00829330	MITS-10	152	MED31-interacing <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00334350	MITS-7	147	MED31-interacing <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00444720	MITS12	108	MED31-interacing <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00467799	MITS-8	100	MED31-interacing <i>Tetrahymena</i> specific protein (MITS)
TTHERM_00732830	TMMI-1	94	Transmembrane domain protein co-purifies with MED31
TTHERM_00670640	AAP4	68	Ibd1 Associated protein (Saettone, 2018), Aap4, SAGA, PHD containing
TTHERM_00691200	TF SII	15	Transcription Factor elongation
TTHERM_00052630	TIM44	6	Transmembrane domain protein co-purifies with MED31

1.11.4 *Tetrahymena*'s putative mediator complex.

The expression profiles of Med31-FZZ co-purifying proteins during growth and development was analyzed using publicly available micro-array data (Xiong et al., 2011) and despite the lack of sequence similarity, found a striking similarity in the temporal expression patterns of divergent MITS proteins and conserved Mediator subunits (**Figure 13, left**), suggesting a functional link among the Med31 interacting partners. The conserved subunits that were identified in *Tetrahymena* may represent the minimal primitive or core Mediator complex composition (**Figure 13, right**), consistent with previous evolutionary predictions (Bourbon, 2008). Considering their interactions and similarities in expression profiles, it suggests that these novel, divergent MITS proteins represent lineage-specific divergent Mediator subunits, consistent with the diverged nature of the *Tetrahymena* Rpb1 CTD. In addition, Med31-FZZ did not co-purify with the CDK8 module which predicted function is to repress transcription (Samuelson et al., 2003).

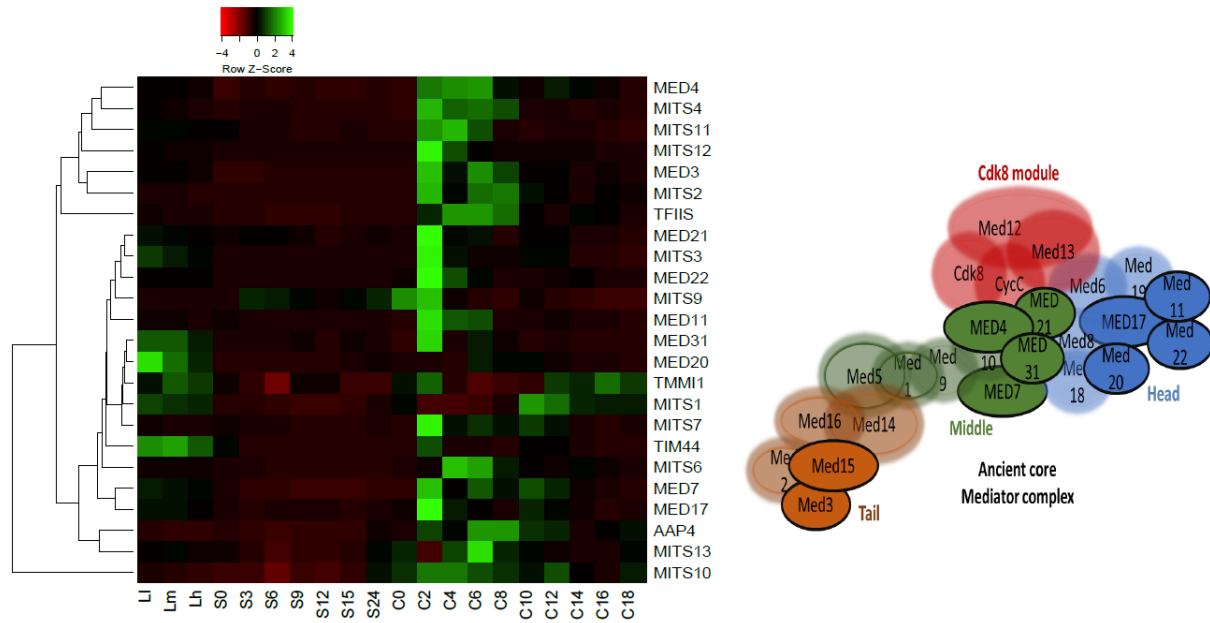


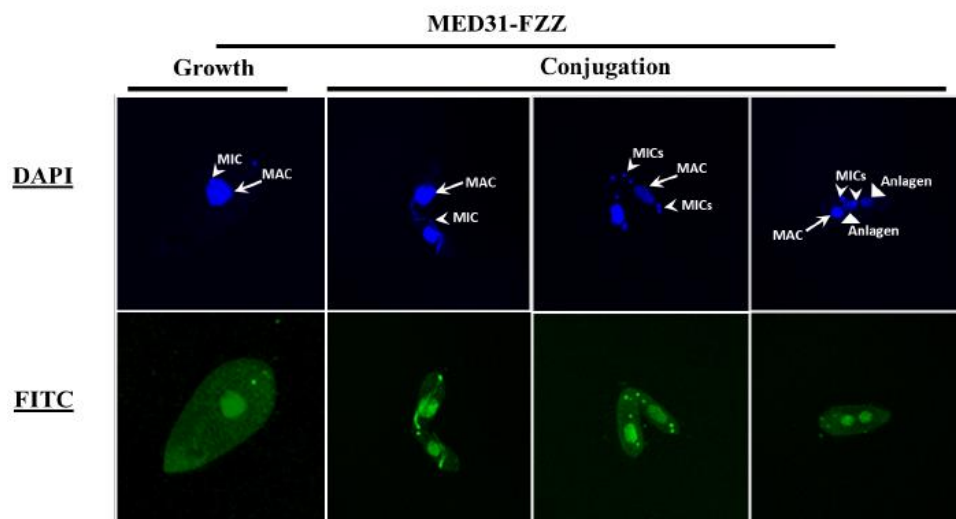
Figure 13: *Tetrahymena*'s putative mediator complex.

Left - The heatmap represents the hierarchical clustering of microarray expression values. Z-scores were calculated across the rows for each protein to examine its differential expression across growth, starvation and developmental stages. L1-LH represent vegetative growth, S0-24 represent starvation for 24 hours and C stands for conjugation (0-18 hours post mixing the cells of different mating types). **Right**- Proposed ancient core Mediator based on conserved subunits identified in *Tetrahymena*. Tail, middle and head modules are color coded as described (Julie Soutourina, 2017). The subunits identified in *Tetrahymena* are solid filled whereas remaining subunits of yeast Mediator (as reviewed in (Julie Soutourina, 2017)) are depicted with colored borders. (Garg et al., 2019).

Tetrahymena features a physical separation of chromatin states in the form of two distinct nuclei, a transcriptionally active Macronucleus (MAC) and a silent germline Micronucleus (MIC), present within the same cell. The polyploid MAC essentially controls all gene expression during vegetative growth whereas the diploid MIC ensures stable genetic inheritance during sexual development (conjugation) (Orias et al., 2011). Both the MAC and MIC originate during conjugation from the same zygotic nucleus (Duane W. Martindale et al., 1982). During conjugation, genome-wide RNAPII-transcription of non-coding RNAs (ncRNAs) occurs from the MICs (Douglas L Chalker & Yao, 2011; Mochizuki & Gorovsky, 2004a). These ncRNAs

direct large-scale RNAi-dependent assembly of distinct chromatin domains in the new MAC as a prelude to programmed DNA rearrangements and irreversible genome silencing (Douglas L Chalker & Yao, 2011). The transcription regulation of these ncRNAs remains poorly understood.

Indirect immunofluorescence (IF) studies using Med31-FZZ were used to examine its localization pattern to gain insights into possible development related roles of *Tetrahymena* Mediator. During *Tetrahymena* growth, Med31-FZZ exclusively localized to the MAC, consistent with a role in transcription. Med31-FZZ localized to both the MAC and MIC during early conjugation. Specifically, in addition to the MAC, Med31-FZZ is also found in the elongated MIC (crescent stage) undergoing meiosis, a stage where the MIC is transcriptionally active. Med31-FZZ staining persisted in the MIC until the period of new MAC (anlagen) development, at which time the signal is only observed within the anlagen (**Figure 14**).



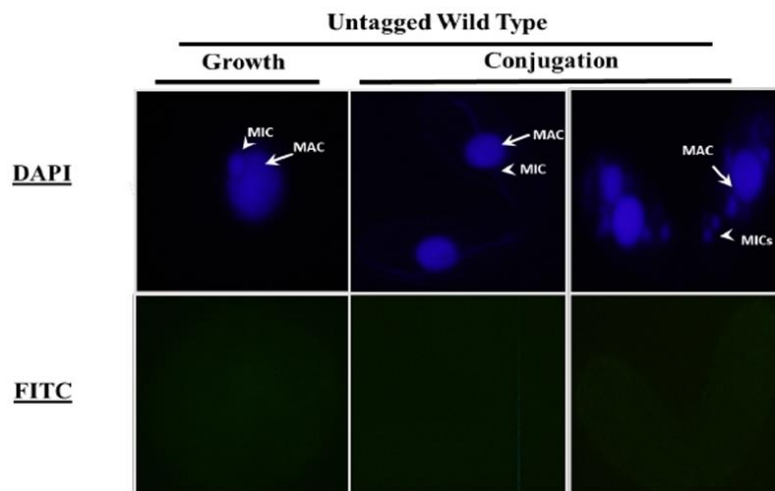


Figure 14: Med31-FZZ immunofluorescence.

Top - Med31-FZZ localizes to the MAC during vegetative growth and to both MAC and MIC during development stages (2, 4 and 6.5 hours post mixing). **Bottom** - Untagged wild type *Tetrahymena* mating cells were used as control. DAPI was used to stain the nuclei (Garg et al., 2019).

1.12 Rationale

To further expand our knowledge in nuclear events, such as chromatin remodeling and transcription, we need to determine the function of these proteins. The selected genomics approach to understand their function is ChIP-Seq. ChIP-Seq is a widely-used molecular method to investigate the function of chromatin-related proteins and can add to their functional understanding since by identifying their associated DNA sequences on a genomic scale. ChIP-Seq permits inference of a protein of interest (POI) function based on its chromatin occupancy, for instance, if the POI is found within expressed genes it can be inferred that the function of this

protein may be related to gene activation. On the other hand, if the POI is found in genes that are not active it can be inferred that its function is related to gene repression. In either case downstream experiments can then be designed to test ChIP-Seq generated hypotheses.

1.13 Thesis Objective

The objective of this thesis is to develop a Chip-Seq pipeline in order to determine function of putative chromatin remodeling protein Ibd1, as well as the Med31-containing Mediator complex in the model organism *Tetrahymena thermophila*. The results obtained will test and improve the models derived from the current proteomics pipeline.

1.13.1 Aim 1: Development of a wet lab protocol for ChIP-Seq using Tetrahymena thermophila

1.13.2 Aim2: Development of a ChIP-Seq pipeline to determine function of chromatin remodeling and transcription related complexes in Tetrahymena thermophila

1.13.3 Aim 3: Determine the function of Ibd1 by Ibd1-FZZ ChIP-Seq

1.13.4 Aim 4: Determining the function of Med31 by Med31-FZZ ChIP-Seq

2 Materials and Methods

2.1 Cell strains

Tetrahymena strains CU428 [Mpr/Mpr (VII, mp-s)] and B2086 [Mpr+/Mpr+ (II, mp-s)] of inbreeding line B were acquired from the *Tetrahymena* Stock Center, Cornell University, Ithaca N.Y. (<http://tetrahymena.vet.cornell.edu/>). Cells cultured in 1 × SPP were maintained axenically at 30 °C as described (J. S. Fillingham, Bruno, & Pearlman, 2001). The Ibd1-FZZ and Med31-FZZ cells were the same used in (Garg et al., 2019; Saettone et al., 2018)

2.2 Macronuclear gene replacement

MED31 KD vectors for *MED31* were generated by amplifying two separate ~ 1 kb DNA fragments up and downstream of the ORF (TTHERM_00355460). The PCR products were digested with KpnI/XhoI and NotI/SacI and cloned into the gene targeting vector p4T2-1, which contained the *Neo2* drug resistance gene provided by Dr. Kathleen Collins (University of California, Berkeley, CA). The 5µg of plasmid DNA was linearized and coated onto one micrometer gold particles (60 mg/ml; Bio-Rad), which was introduced into *T. thermophila* MAC using biolistic transformation with a PDS-1000/He Biolistic particle delivery system (Bio-Rad). The transformants were selected using paromomycin (60µg/ml). To achieve MAC homozygosity cells were grown in increasing concentrations of paromomycin to a final concentration of 1 mg/ml.

2.3 ChIP-qPCR: Validation of RACS results for *Ibd1*

Four ChIP biological repetitions for the *Ibd1*-FZZ and three ChIP repetitions for the untagged cell lines were quantified (Nanodrop, Thermo Scientific) and diluted to reach the smallest DNA concentration found in a sample (1 to 3.1ng/μL of DNA). Master mixes with a final volume of 20μL were prepared (Sybr Green Supermix, Cat. #1708880, Bio-Rad) to amplify: the top 3 genes that presented the highest fold enrichment from *Ibd1*-FZZ ChIP-Seq and are highly expressed and a gene that is not expressing during vegetative growth (*PDD1*) (Saettone et al., 2018) using qPCR (CFX 96-well Real-Time System, Bio-Rad) with the following parameters: Initial denaturation at 98°C for 3 minutes; 40 cycles of amplification at 95°C for 15 seconds and 60°C for 60 seconds followed by acquisition in the SYBR/FAM channel; and melting curve from 65°C to 95°C increasing 0.5°C/cycle and acquisition every 0.5 seconds in the SYBR/FAM channel. Each targeted gene was considered as an individual experiment each with its own standard curve. The standard curve for each target has 3 points representing 100%, 10%, and 1% of the corresponding input sample. The largest point of the curve was undiluted input sample and was followed by serial dilutions. Raw Cq values for input DNA and IP DNA were analyzed using the BioRad Prime PCR program, which normalizes these data to the generated standard curve that it was represented as % with respect to the INPUT. Ultimately, these normalized ChIP data are expressed as fold enrichment, by dividing normalized IP over normalized Input. The standard error of the mean (SEM) was calculated for each duplicate (**Table 4**).

2.4 RT-qPCR: Validation of RACS results for *Med31*

RNA was extracted from untagged and Med31-KD growing cells using TRIzol (Life Technologies) as by protocol provided by Dr. Kathleen Collins (University of California, Berkeley, CA) followed by Deoxyribonuclease I (RNase-free, Thermo) treatment. cDNA and negative controls were generated using iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad). The negative controls have the RNA plus the reagents provided in the kit to generate cDNA minus the reverse transcriptase. Master mixes with a final volume of 20µL were prepared using Sybr Green Supermix (Bio-Rad) to amplify the negative controls and cDNA of the reference gene HTA3 and the target genes Med31, Pdd1, Pdd2 and Pdd3 (Garg et al., 2019). The instrument used was a CFX 96-well Real-Time System (Bio-Rad) with the following parameters: Initial denaturation at 98°C for 3 minutes; 40 cycles of amplification at 95°C for 15 seconds and 60°C for 60 seconds followed by acquisition in the SYBR/FAM channel; and melting curve from 65°C to 95°C increasing 0.5°C/cycle and acquisition every 0.5 seconds in the SYBR/FAM channel. The amplified cDNA and negatives controls from untagged and Med31-KD were three replicates of HTA3 and Pdd1 and two replicates of Med31-KD. The prepared negative controls were three replicates of HTA3, Pdd1 and Med31-KD. It is important noting that *Tetrahymena*'s genomic DNA was used for generation of the standard curves (**Table 5**, **Table 6**, **Table 7**, green panels). One was made for each target and reference having 3 points representing 100% (undiluted DNA), 10% (1:10), and 1% (1:100) of the untagged DNA samples. Each targeted and reference gene were considered as an individual experiment each with its own standard curve. The assigned and Cq values can be found in **Table 5**, **Table 6** and **Table 7**. Raw Cq values were analyzed using the BioRad Prime PCR program, which normalizes these data to the generated standard curve that was represented as %. Ultimately, the data was analyzed using relative quantification (Ratio=Reference/Target) being the values for HTA3 the reference and Pdd1 and

Med31 the targets respectively for the independent experiments. T-test were calculated for each experiment. In the supplemental file rt-qPCR the SQ (starting quantity) columns contain the values of each target sequence. These values were divided by the SQ Average of the cDNA and negative control of HTA3 and deposited in the Ratio (T:SQ/R:SQAverage) column for the Pdd1 and Med31 cDNA and negatives controls respectively. T-test were calculated to assess statistical differences between the untagged and Med31-KD strains (**Table 6, Table 7**).

2.5 Generation of cell extracts and Western blotting

10% trichloroacetic acid was used to prepare whole-cell extracts (WCE) by incubation on ice for 30 minutes. The WCEs were re-suspended in 10µl of SDS loading buffer. WCEs were separated by electrophoresis through 10% SDS-PAGE. The proteins were transferred to nitrocellulose and probed with indicated antibodies after blocking in 5% skim milk. Antibodies and dilutions used were anti-Pdd1 (1:1000; Abcam), anti-Brg1 (1:5000; (J. Fillingham et al., 2006a)).

2.6 Indirect immunofluorescence

Cells were grown and fixed during vegetative growth, 24 hours starvation, and 2, 4, 6 and 7.5 hours post mixing after starvation to perform indirect immunofluorescence as previously described (Garg et al., 2013). Briefly, we washed the cells with 10mM Tris-HCl, pH 7.7, fixed with 4% paraformaldehyde and membrane-permeabilized in cold acetone for 20 min. Anti-Pdd1 (Abcam) and Anti-H2A.X were used at dilutions 1:5000, 1:500 and 1:1000 respectively and incubation was at 4°C overnight in 1×PBST. Cells were washed three times in 1×PBS. Incubation in secondary

antibody fluorescein isothiocyanate-conjugated (FITC - Pierce), Alexa Fluor 488 or Cy3 (Jackson) goat anti-mouse or anti-rabbit accordingly was for 1 h at room temperature. For nuclear counterstaining 4,6-diamidino-2-phenylindole dihydrochloride (DAPI) was used. Immunofluorescence analysis was carried out using Olympus, DP70 equipped with a fluorescent microscope (Reichert-Jung, POLYVER) at 100x magnification and no oil was used. Final image preparation was carried out using ImageJ (1.50i) software (Schneider, Rasband, & Eliceiri, 2012).

2.7 Bioinformatics and molecular evolutionary analyses

Amino acid sequences for yeast and human Med31 were acquired from the UniprotKB and were used as a query against the *Tetrahymena* genome (<http://ciliate.org/index.php/home/welcome>). To identify orthologs across the eukaryotes, human Med31 sequence was searched against NCBI non-redundant database using PSI-BLAST with default parameters. Sequences retrieved were analyzed using SMART (<http://smart.embl-heidelberg.de/>) (Letunic & Bork, 2018). To reconstruct a protein phylogeny, we used amino acid sequences of the identified Med31 orthologs. Multiple sequence alignments were built using MUSCLE with default parameters. All protein phylogenetic analyses were carried out using the maximum likelihood (ML) method under LG+G+1 model using MEGA 7 (Kumar, Stecher, & Tamura, 2016) with 1000 replicas for each tree. Med31 structural prediction was performed using the I-TASSER server (J. Yang et al., 2015). The secondary structure predictions were carried out using PSIPRED (Buchan & Jones, 2019). Furthermore, *Tetrahymena* phosphoproteomic data for Rpb1 analysis was acquired from http://tfgd.ihb.ac.cn/search/phos?locus=TTHERM_00538940&Submit=SEARCH.

2.8 ChIP-Seq and data analysis

ChIP seq data generated in this paper can be found online at Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) GSE126344. NGS and reads files (Private data, not to be shared or distributed without permission) produced in this study were deposited at <https://www.ncbi.nlm.nih.gov/geo/> with unique identifier GSE126344. Use the following token: otovcyaannwtjch to access until publication of the results. Direct link: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE126344>

The ChIP-Seq experiments were performed as described in (Saettone et al., 2018). The ChIP-seq data was normalized to inputs and was analyzed essentially as described in (Saettone et al., 2018), with the following modifications. Instead of using arbitrary cut-offs to determine enriched regions, untagged controls were used to account for unspecific binding to the M2 agarose matrix. The obtained untagged strains values of each genic and intergenic regions were subtracted from the correspondent tagged strains values (See Supplementary file 2A, B for details). The values with less than 1.5 X enrichment were filtered. The accumulation of reads on the genic and intergenic regions were calculated by adding the total obtained values of reads and expressed as percentage. The ChIP-seq data for the untagged controls has been deposited in GEO repository and will be reported elsewhere for the ChIP analysis software development (manuscript in preparation). The metagene analysis was performed using ChIP-Seq reads normalized over the inputs and by ‘Reads Per Kilobase of transcript per Million mapped reads (RPKM)’ values. The plots were generated used Deeptools (Ramírez, Dündar, Diehl, Grüning, & Manke, 2014).

2.9 Data Deposition

ChIP-seq data generated in this paper can be found online at Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) GSE125576, GSE103318 and GSE126344.

2.10 Materials Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Anti-Pdd1 Polyclonal Antibody	Abcam	Abcam Cat# ab5338, RRID:AB_304827
Mouse Anti-H2A.X	BioLegend	BioLegend Cat#613401; RRID:AB_315794
Rabbit Anti-BRG1	Pearlman Lab [Supplemental reference#8] – York University, ON, Canada	N/A
HRP conjugated Anti-IgG	Millipore Sigma	AP160P, Lot# 2901304
Anti- GFP	Millipore Sigma	Cat#11814460001
Cy TM 3 AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch	RRID: AB_2307443
Alexa Fluor® 488 AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch	RRID: AB_2341099
Alexa Fluor® 488 AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch	RRID: AB_2313584
Bacterial and Virus Strains		
Subcloning Efficiency DH5Alpha Competent Cells, 2mL	Invitrogen	Catalog number: 18265017
Critical Commercial Assays		
iScript TM Reverse Transcription Supermix for RT-qPCR	Bio-Rad	1708840
SsoAdvanced TM Universal SYBR® Green Supermix	Bio-Rad	1725271
TruSeq ChIP Library Preparation Kit	Illumina	IP-202-1012
Experimental Models: Organisms/Strains		
<i>T.thermophila</i> : B2086 [<i>Mpr</i> ⁺ / <i>Mpr</i> ⁺ (mp-s, mt II)]	Tetrahymena stock center of Cornell University, Ithaca NY	N/A
<i>T.thermophila</i> : CU428 [<i>Mpr</i> / <i>Mpr</i> (mp-s, mt VII)]	Tetrahymena stock center of Cornell University, Ithaca NY	N/A
<i>T.thermophila</i> : Med31-FZZ: B2086 [<i>Mpr</i> ⁺ / <i>Mpr</i> ⁺ (mp-s, mt II)]	(Garg et al., 2019)	N/A
<i>T.thermophila</i> : Ibd1-FZZ: B2086 [<i>Mpr</i> ⁺ / <i>Mpr</i> ⁺ (mp-s, mt II)]	(Saettone et al., 2018)	

<i>T.thermophila</i> : Med31-KD: CU428 [<i>Mpr/Mpr</i> (mp-s, mt VII)]	(Garg et al., 2019)	N/A
<i>T.thermophila</i> : Med31-GFP: B2086 [<i>Mpr+/Mpr+</i> (mp-s, mt II)]	(Garg et al., 2019)	N/A
Oligonucleotides		
qPCR: Pdd1 Forward: CGCTTCTAAAATGTCTCTTAGCC	This study	N/A
qPCR: Pdd1 Reverse: TTGAGCTAACTCTAAGCAGCAA	This study	N/A
qPCR: HTA3 Forward: GCTGGTTTATAATTCCCAGTCG	This study	N/A
qPCR: HTA3 Reverse: AGCAGCAGCATAAACAGCAG	This study	N/A
qPCR: Med31 Forward: CAAATTCGTTAAGTATCCACTAGGC	This study	N/A
qPCR: Med31 Reverse: TTCATTTTATTGCTAATTCAATTCC	This study	N/A
qPCR: Pdd2 Forward: AGAAATGCAAAGAGGCCAAA	This study	N/A
qPCR: Pdd2 Reverse: ATGTTTGTGTCAGCATACTTTTCCA	This study	N/A
qPCR: Pdd3 Forward: TTCAATTTAAAAGGAGCAAACCTGA	This study	N/A
qPCR: Pdd3 Reverse: TCTTAAGGTAATTTACTCCCACTTTCA	This study	N/A
3 MED31-KD_UP_F 5`- GGTACCTTCTTTTGATTTGTTTATTAATTTTTT	This study	N/A
MED31-KD_UP_R 5`- CTCGAGTAGCTTTGAATTTGCATGC	This study	N/A
4 MED31-KD_DOWN_F 5`- GGATCCTCTAGCTTGTAAGAAAGATATTAAGAAAG	This study	N/A
5 MED31-KD_DOWN_F 5`- CCGCGGTTTTTGATAGATCTAATTGAAATTTG	This study	N/A
Recombinant DNA		
p4T2-1 Neo2	Jacek Gaertig(University of Georgia,Athens,GA)	
Software and Algorithms		
RACS	(Saettone, Ponce, et al., 2019)	N/A
CFX Manager™ Software	Bio-Rad	1845000
Integrative Genomics Viewer	Broad Institute	https://software.broadinstitute.org/software/igv/download

3 CHAPTER 1: Development of a ChIP-Seq pipeline to determine function of chromatin remodeling and transcription related complexes in a model organism

3.1 Introduction to RACS (rapid analysis of ChIP-Seq data for contig based genomes)

RACS is a one-stop methodology to analyze ChIP-Seq data to find the set of genome coordinates for a given Protein of Interest (POI). This methodology utilizes open-source tools such as BWA (H. Li & Durbin, 2009), SAMtools (H. Li et al., 2009), Linux shell and R scripts (R Core, 2013). RACS was developed to answer whether the POI localized to a given set of coordinates (genes) or to the remaining regions in the genome that were not given by the user (intergenic). RACS was designed in a user-friendly manner to accommodate researchers with basic knowledge in Linux shell and R. RACS provides accessible downstream analyses of ChIP-Seq data obtained from Illumina instruments. RACS follows a unique approach to tackle this problem, is widely applicable and useful enough to analyze ChIP-Seq related data from a variety of different organisms generated by NGS.

After processing ChIP-Seq data, RACS will output two tables, the first containing all found reads accumulation in the genic region corresponding to the annotated genes and the second containing the accumulation of reads in the intergenic regions. An intergenic region will be calculated as a region that starts at the end of a given gene coordinate and ends at the beginning of the next contiguous given gene coordinate. RACS will calculate the beginning of a contig as the

beginning of an intergenic region (as long as there is not an annotated gene at the beginning of the contig) which ends at the coordinate of the first encountered gene, and it will do the same at the end of each contig. These intergenic or adjacent regions are newly generated each time to account for modifications or improvements in the files containing gene annotations. The obtained results from both tables are normalized to the number of clusters that passed Illumina's "Chastity filter" also called clusters PF. These numbers represent the reads obtained per sample. The normalized values are further filtered by using the data obtained from the mock samples.

3.2 WetLab: Chromatin Preparation for Chromatin Immunoprecipitation (ChIP) in *Tetrahymena thermophila*

90 mL of tagged cells in exponential growth were collected and cross-linked with 2.5 mL of 37% formaldehyde for 30 min at room temperature and neutralized with 13mL of 2.5M glycine. The chromatin was washed with 10mM Tris pH7.4 followed with cold lysis buffer (50mM Tris pH 8.0, 5mM EDTA, 1% SDS in H₂O plus protease inhibitor (cOmplete™ Protease Inhibitor Cocktail) and 100mM PMSF). The chromatin was fragmented to 300 to 600 base pairs on ice by sonication (Branson Sonifier). The lysis buffer containing the sonicated chromatin was diluted 10 times using dilution buffer (50mM Tris pH 8.0, 100mM NaCl, 5mM EDTA, 2% Triton X and 0.2% deoxycholate. Chromatin immunoprecipitation was performed using 25uL of packed bead volume M2 agarose beads (Sigma) rotating at 4°C for 4hrs. The beads were wash one time with each of the following buffers 1X FA, 1X FA plus 500mM NaCl, LiCl buffer and 1X TE. DNA elution started in 2X Proteinase K buffer (20mM Tris pH8, 10mM EDTA and 1% SDS in H₂O) at 65°C for 20 minutes and diluted to 1X Proteinase K by H₂O addition. RNase (Fermentas) was

added and incubated for 1hr at 37°C. Proteinase K (Fermentas) was added and incubated at 42°C for 2 hrs and 65°C for 8hrs. The DNA was isolated (Qiagen PCR purification kit) and H₂O was used as elution (Garg et al., 2019; Saettone et al., 2018). The full protocol can be found in the Appendix 1.

3.3 Library preparation and NGS

Two biological replicate analyses of INPUT and IP samples were sequenced utilizing high-throughput sequencing for untagged, Ibd1-FZZ and Med31-FZZ. The NGS data can be found in the following Gene Expression Omnibus (GEO: <https://www.ncbi.nlm.nih.gov/geo/>) accessions respectively: GSE125576, GSE103318 and GSE126344.

Library preparation was performed using Illumina's TruSeq ChIP Library Preparation Kit (cat# IP-202-1012). Samples were sequenced on the Illumina HiSeq2500 platform as follows: single-end 58bp read length using v4 chemistry. The samples were demultiplexed using bcl2fastq2 Conversion Software v2.17 (https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html) allowing for 1 barcode mismatch. Passed Filter reads were used in downstream sample analysis. NGS was performed by the Donnelly Sequencing Centre, CSPro at University of Toronto.

3.4 Implementation of RACS: rapid analysis of ChIP-Seq data for contig based genomes

The RACS pipeline utility was described and demonstrated using ChIP-Seq data sets generated in the model organism *Tetrahymena thermophila*. The *Tetrahymena thermophila* ChIP-Seq data set originates from our recent studies (Garg et al., 2019; Saettone et al., 2018). The RACS pipeline is an open source set of shell and R scripts, which allows the user to compute reads differentiating between genic and intergenic regions automatically. RACS also includes auxiliary spreadsheets that are used to perform the post-processing and normalization manually, as well as, to check against negative controls. For normalization using the Cluster Passing Filtering (PF) values given by the NGS instrument. RACS generates files to validate results by visualizing the reads accumulation and run comparisons with other software tools, such as IGV and MACS2 respectively. The RACS repository (<https://bitbucket.org/mjponce/racs/src/master/>) includes the core or main scripts placed in the core directory. The comparison and auxiliary tools are placed in the tools directory. Additionally, a datasets directory containing scripts that allow the user to download the data used in these analyses is included. Details about the pipeline implementation and how to use it are included in the README file available within the RACS repositories. A generic top-down overview of the pipeline implementation for the data analysis, is shown in **Figure 15**.

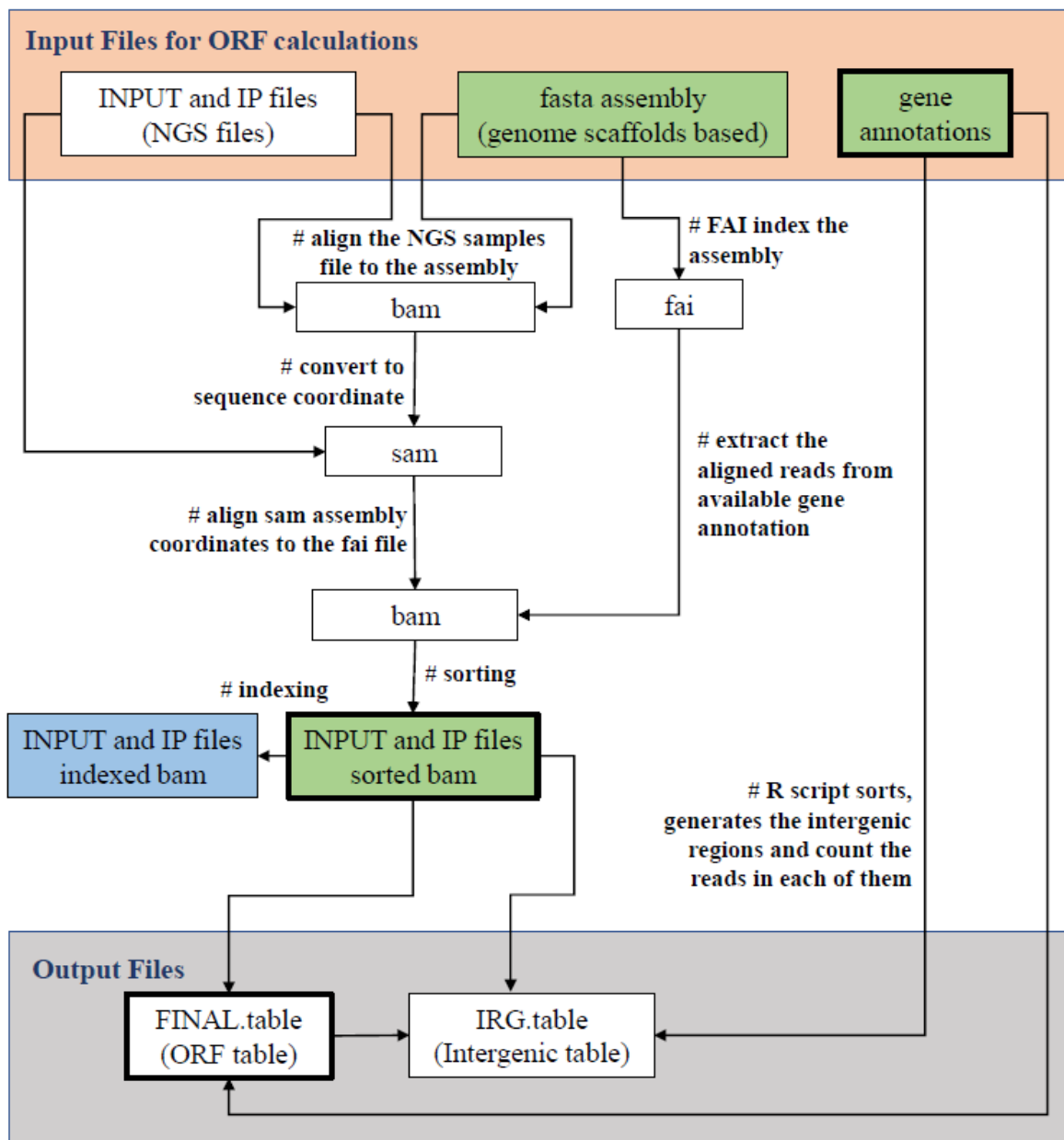


Figure 15: RACS Core Pipeline overview.

This flowchart represents the logic steps implemented in the core pipeline. Boxes represent files and file types as indicated in the text. Files with thick boxes represent the Input Files for intergenic calculations. Files in green are to be uploaded to IGV. File in blue is needed for IGV but it does not have to be uploaded to IGV. This file has to be kept in the same folder directory than the sorted bam (Saettone, Ponce, et al., 2019).

The RACS pipeline will run in any standard workstation with a Linux-type operating system. In addition, the following open source tools are needed by the RACS core scripts: i] Burrows-Wheeler Alignment (BWA) version 0.7.13 (H. Li & Durbin, 2009); ii] Sequence Alignment/Map (SAMtools) version 1.3.1 (H. Li et al., 2009) and iii] the R statistical language (R Core, 2013). The pipeline requires as input the *.fastq* files (obtained from NGS) from the ChIP-Seq experiments and the specific genome assembly files and a file containing the gene annotations (e.g. *.gff3* files containing genic regions) corresponding to the organism. For *T.thermophila* these files are: *T_thermophila_June2014.assembly.fasta* and *T_thermophila_June2014.gff3*. Both files can be found at <http://ciliate.org/index.php/home/downloads> (Stover et al., 2006).

3.5 Core Pipeline Tools

The core scripts do not require any additional packages other than the ones mentioned above; however, the comparison tools, depending on what format the data to compare with is given, might use some additional R packages, such as a spreadsheet reader package. The results of the genic and intergenic regions (IGR) are generated in two *.csv* files. These are standard text ASCII files, which can be read with any typical spreadsheet software or R.

3.6 Determination of the Genic Regions

To count the amount of reads in each genic region the core pipeline script was implemented using Linux shell commands combined with the usage of BWA and SAMtools. The input files

are the genome of reference (T_thermophila_June2014.assembly.fasta), the gene annotation file T_thermophila_June2014.gff3 and the INPUT and IP files obtained from NGS. The INPUT files contain the information obtained from NGS prior to the immunoprecipitation; thus, this file contains the initial reference amount of DNA reads. The IP file contains the data after the immunoprecipitation; thus, this file contains the DNA that were enriched by the POI. After the INPUT and IP sequences are aligned with the genome and sorted, the script uses a loop to count the reads in each genic region and deposits the obtained data in a file named FINAL.table. INPUTfile-IPfile; where INPUTfile and IPfile are the INPUT and IP files respectively. **Figure 15** depicts a flowchart representing the required steps to obtain the final table containing the number of reads found in each of the genic regions, in relation to *Tetrahymena thermophila* scaffold database and the breakdown of each these steps.

The RACS pipeline was implemented to specifically target data from the *Tetrahymena thermophila* organism in particular utilizing a specific .gff3 file. This file contains the General Feature Format of gene annotations. However, due to the modular fashion in which RACS was implemented, it is possible for users targeting different organisms and even different markers, to modify RACS to do so. At the level of the IGR, if the reference file follows the usual .gff3 structure, nothing has to be modified in the code.

At the level of the ORF, the user will need to specify a few parameters that will be used when the targets depart from the ones used by default in the pipeline. The terms and filters allow the user to target either genes or mRNA or any other item within the reference file, making the pipeline available to use with other organisms. In order to achieve this, the user should provide a

'definition' file, indicating the targets for the pipeline for which to filter for the reference file. The subdirectory in the repository "core/defns", contains some files exemplifying the implementation of different cases and organisms.

In particular, the variables filter1, filter2, as well as, delim1, delim2, delim3; should be adjusted correspondingly to the organism of interest and the way the data is organized within the reference file. Table 3 shows an example of how the variables should be adjusted for *Tetrahymena thermophila*.

Table 3: Definitions

This table shows the definitions that can be modified based on the .gff3 tables for different organisms. "defn" file for *Tetrahymena thermophila*; see: RACS/core/defns/TT_gene.id in the RACS repository (Saettone, Ponce, et al., 2019).

filter1=gene
filter2="Name=TTHERM_"
.....
delim1="TTHERM"
delim2=";Note"
delim3="Note="

3.7 Determination of the Intergenic Regions

The intergenic regions were not available neither determined by the standard packages. For this reason, an R script was developed to determine these regions. In this pipeline these sequences are calculated during each run to account for further genome actualizations. The inputs for this script

are the files generated by the genic regions pipeline discussed in the previous section (i.e. FINAL.table, INPUTfile-IPfile, the INPUT and IP .bam files --which are generated as intermediate files of the Genic Region pipeline) plus the gene annotation file (e.g. T_thermophila_June2014.gff3). First, the script determines the intergenic regions by calculating the beginning and end of each annotated gene within each available scaffold and subtracts these values. The algorithm only reports intergenic regions that are equal or greater to zero. Second, the script uses the newly generated intergenic regions to count the number of reads in each of them. Finally, the data is deposited in the intergenic table for each of the intergenic regions **Figure 15**.

3.8 Post-processing: Normalization of Reads accumulation and Enrichment

Calculation

To account for differences in the amount of clusters PF (reads) presented among samples, each of the obtained INPUT and IP values were normalized by dividing them by the corresponding clusters PF value of the Flowcell summary (obtained from the NGS run) or from the Total Sequences (obtained from the fastQC file). These calculations can be calculated employing the following spreadsheets that can be found in the datasets of Appendices 3 and 4 and subdirectory within the RACS repository (<https://bitbucket.org/mjponce/racs/src/master/datasets/>):

- for the genic regions TET_Ibd1_MAC_Genome_Genic.xlsx, and
- for the intergenic regions TET_Ibd1_MAC_Genome_Intergenic.xlsx;

Notice that there are several spreadsheets provided in this subdirectory, each of them will be used for different organisms/cases and can be used as templates for other datasets.

These spreadsheets contain the reads found in the untagged (or mock purification/negative control) samples are in the *Untagged* tab (data from GEO GSE125576, compiled in Appendices 3 and 4). The user can also add the Flowcell summary (FCS) details in the *Add_FCS_for_(SAMPLE_ID)* tab. The user can manually introduce the read values for the samples being analyzed in the *Add_(SAMPLE_ID)_ChiP_Seq* tab. In this tab the user can divide the number found by RACS by the corresponding cluster PF number found in the previous tab. This data can be deposited in the *Normalized_INPUT* or *_IP* (FCS) columns. After the required reads normalization, the accumulation can be obtained as the number of IP reads divided by the number of INPUT reads (IP/INPUT). This can be deposited in the *Enrichment_(N_IP(FCS) N_INPUT(FCS))* column of the same tabs. The obtained values are filtered (Filter 1) by the user by subtracting the corresponding number found in the *Untagged* tab and deposit the values in the *Enrichment_Minus_AVERAGE_untagged* column. If there are more than two samples the values can be averaged and values that are less than 1.5 can be filtered (Filter 2) and deposited in the *Enrichment_Average_Sample* tab. For the genic region table, in this tab there is a column containing the Expression profile obtained from the *RNA_Seq* tab. To ease the analysis, it is recommended to copy the filtered cells to the *Results* tab. The distribution of the protein of interest can be calculated in this tab. For the Intergenic table there is a *ORF_vs_IGR* (Intergenic) tab where the number of regions and reads can be calculated. The number of regions is represented by the number of genic and intergenic regions that passed the 2 filters. The number of reads found in the genic and intergenic regions can be calculated by adding all the available values from the *Normalized IP* (FCS) columns and deposit them in the *ORF_vs_IGR* tab of the Intergenic table.

During the post-processing steps, it is important to note that some regions presented in the processed table may have very few reads after subtracting the values obtained from untagged samples and they may seem as real interactors when they are not. For instance, a sample that has 2 reads in the INPUT and 10 reads in the IP will return an enrichment of 5 and it may pass the filter of 1.5 enrichment but they may not be significantly enriched.

3.9 Utilities: Validation and Quality Checks

To account for biological and experimental variability in the wet lab, 2 independent ChIP-Seq samples are typically performed using for each distinct protein/strain and average their Enrichment. To validate the findings, it is important to determine the genic and intergenic regions of the untagged (negative control) INPUT and IP samples. After this determination, the obtained average enrichment from untagged to the obtained tagged average of the samples is subtracted. Then values that had an enrichment greater than or equal to 1.5 are filtered in the final enrichment column. These are the enriched regions and represent genomic regions to which the POI binds.

3.9.1 Visualization of Reads accumulation

The browser IGV (Robinson et al., 2011a) can be used to visually inspect and validate the obtained reads based on their ranked enrichment. The files needed are illustrated in **Figure 15** and the README file included in the RACS' repository. MACS2, a main-stream application to

call peaks, can also be used as specified in (Y. Zhang et al., 2008). MACS2 uses the same intermediate files (.bam) obtained from the RACS pipeline, hence it can be a good reference to be considered for comparison purposes (**Figure 16**, A and B).

4 CHAPTER 2: ChIP-Seq analysis indicates that Ibd1 localizes to highly expressed gene in *Tetrahymena*

The objective of this chapter is to test and improve the model derived from the proteomics pipeline presented in the introduction. The RACS pipeline was utilized to further understand the function of Ibd1.

4.1 *Ibd1 Localizes to Transcriptionally Active Chromatin*

Ibd1 co-purifies with multiple protein complexes involved in gene expression regulation and *in vitro* recognizes histone marks associated with an active chromatin state (Saettone et al., 2018). These observations suggest an intimate role of Ibd1 in transcription regulation. To examine this possibility in more detail, I employed chromatin immunoprecipitation followed by next generation sequencing (ChIP-Seq). Specifically, I asked whether Ibd1 localizes to specific regions of the genome that correlate with transcriptionally active chromatin.

Data for two biological replicates that include DNA from input chromatin as well as Ibd1-FZZ precipitate from two independent experiments were analyzed. Our ChIP-Seq (GEO accession GSE103318) data set utilizing the available genome annotations (Stover et al., 2006) was composed of all annotated genic or open reading frames (ORF) and intergenic regions. The two generated lists displayed greater than or equal to 1.5-fold enrichment of Ibd1 and were ranked in descending order (Appendix 2). From these lists it was observed that Ibd1 occupies 354 ORF and 842 intergenic regions with an enrichment (IP/INPUT) greater than or equal to 1.5-fold. However, the majority of reads are found in genic regions. (**Figure 16**, C and D, Appendix 2).

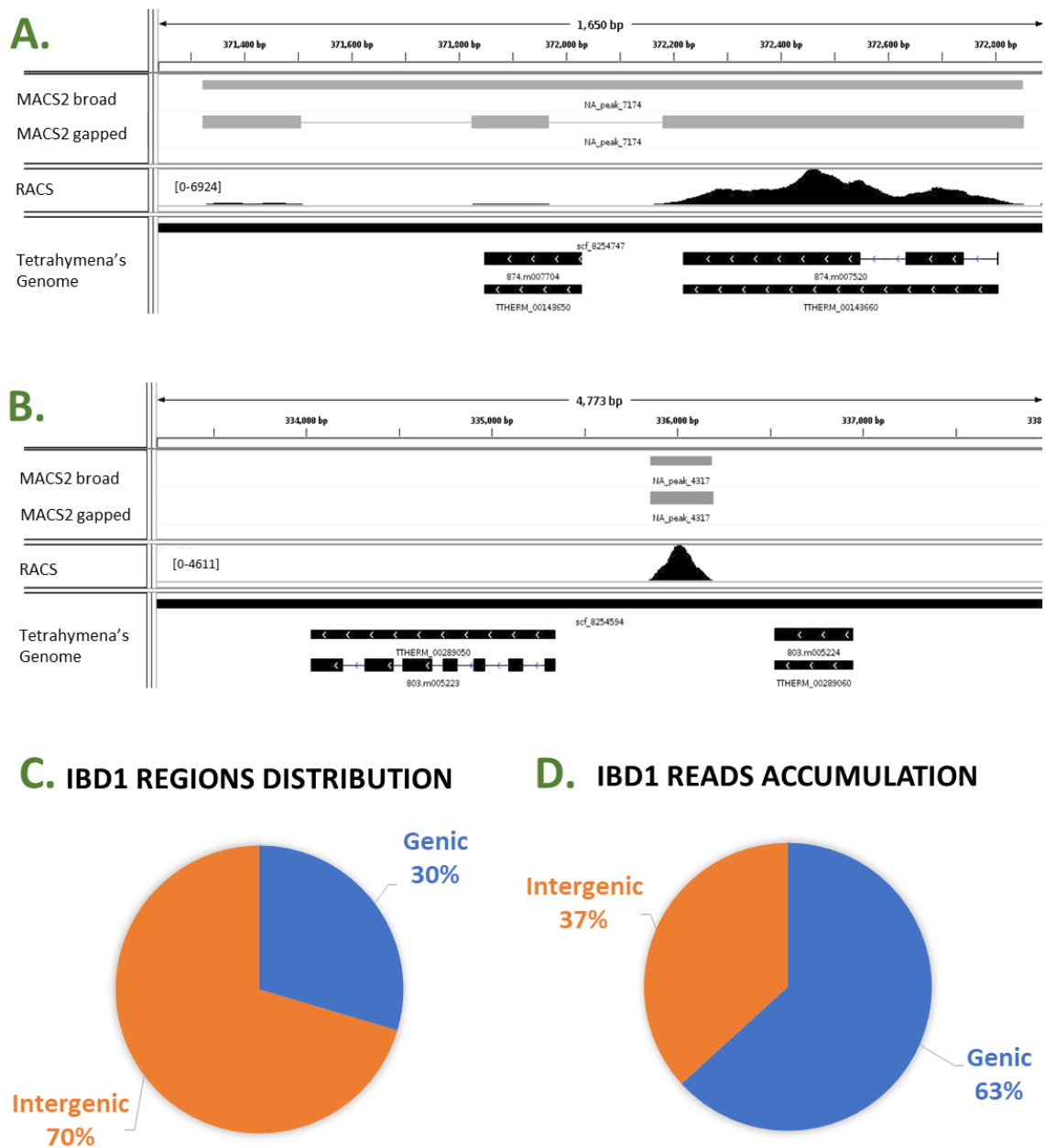


Figure 16: Ibd1 Distribution.

A. and B. Visualization of genic and intergenic regions using IGV. The top track shows MACS2 broad and gapped peaks. The middle track shows RACS visual representation of reads accumulation. Note that RACS shows graphical reads behaviour and accumulation preferences. The bottom track shows *Tetrahymena thermophila*'s genes. On the other hand, MACS2 found two weak peaks that can be interpreted as background by our pipeline. The range inside the brackets represents the highest number of reads for that specific track. **C.** Ibd1 localizes to more intergenic than genic regions (354 ORF and 842 intergenic). **D.** However, the majority of reads are found in genic regions. These results take into consideration the information provided by the mock samples (Saettone, Ponce, et al., 2019).

4.2 *Ibd1* Localizes to a set of highly expressed genes

The 354 ORFs found were assessed for their transcriptional state. Previously published RNA-Seq data was utilized to designate and rank *Tetrahymena* genes in four groups based on the number of found RNA counts (expression level) during vegetative growth (GEO accession GSM692081, (Xiong et al., 2012)). High expression are genes that have 1000 or more RNA counts. Moderate expression are genes that have between 100 and 999 RNA counts. Low to No-Expression genes are genes that have less than 100 counts of RNA. The final group represents genes that are predicted to exist but they do not have available expression data during grow. Based on this data it was found that during grow, 9% and 29% of genes in *Tetrahymena* are highly and moderately expressed respectively (**Figure 17, A**, Appendix 2). On the other hand, it was found that 51% (181 ORF) and 6% (21 ORF) of genes occupied by Ibd1 are highly and moderately expressed respectively. These observations are consistent with our histone peptide-array data and further strengthen the idea that Ibd1 primarily occupies active chromatin regions. Interestingly, 16% (55 ORF) and 27% (97 ORF) of genes occupied by Ibd1 present low, no expression or non-available expression data during grow (**Figure 17, B**, Appendix 2).

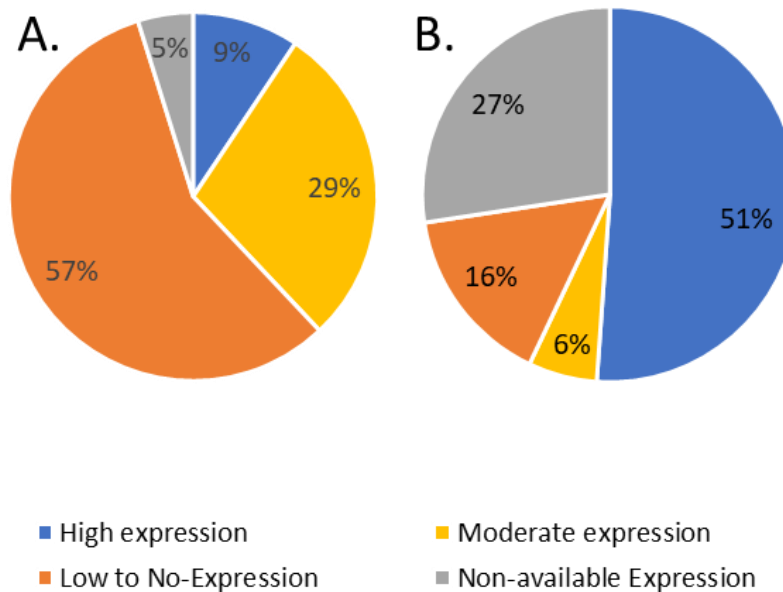


Figure 17: *Tetrahymena*'s expression distribution and Ibd1 localization.

A. 9% of *Tetrahymena*'s genes are highly expressed. **B.** 181 ORF (51%) that are occupied by Ibd1 are highly expressed. Ibd1 also localizes to 97 (27%) coding regions that do not present available data for the RNA-Seq data (GEO accession GSM692081, (Xiong et al., 2012)) (Saettone, Ponce, et al., 2019).

4.3 Gene Ontology (GO) annotation analysis of Ibd1 bound genes suggests function in fundamental biological processes

To examine whether these 354 genes are enriched for any particular functional categories, they were grouped using STRING (Snel, Lehmann, Bork, & Huynen, 2000) based on their predicted Gene Ontology (GO) terms (Gene Ontology, 2000). It was found that these genes are grouped to terms related to housekeeping genes which are genes that are highly expressed in log phase, such as cellular process, translation, metabolic processes and gene expression (**Figure 18**, Appendix 2). These housekeeping genes are generally highly expressed during growth consistent with our findings that Ibd1 primarily occupies transcriptionally active chromatin.

BIOLOGICAL PROCESSES (GO) CONTROLLED BY IBD1

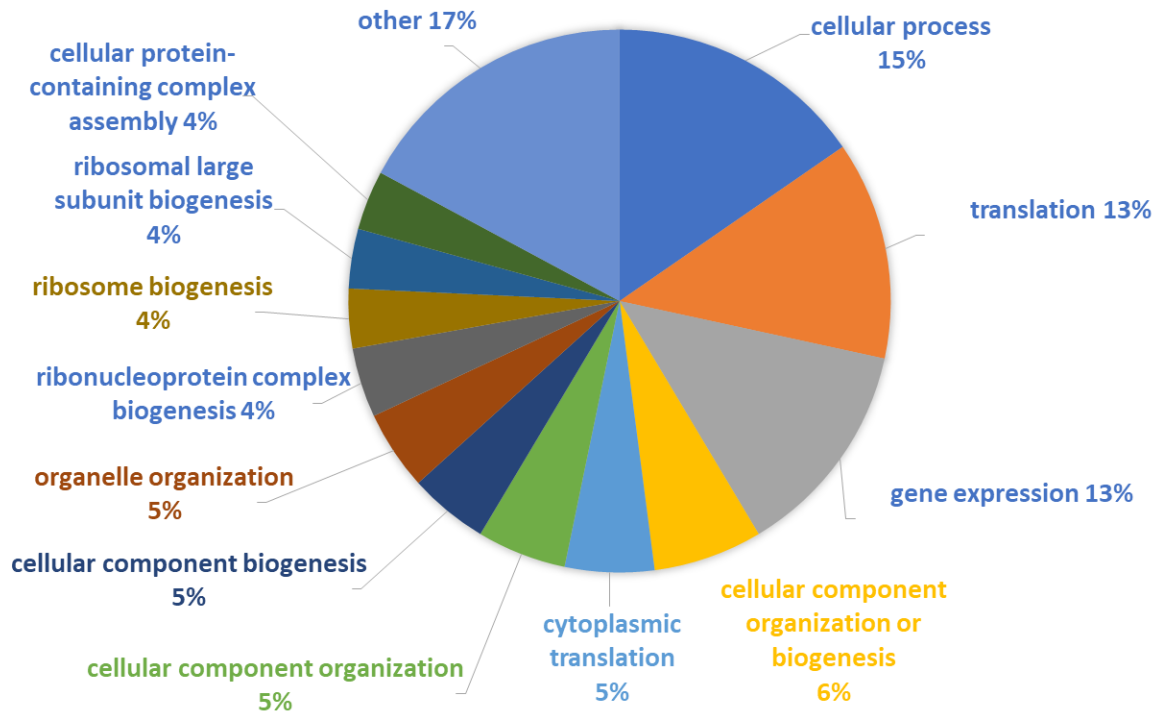


Figure 18: Ibd1_GO.

Gene Ontology (GO) analysis of genes controlled by Ibd1. GO predicted that the majority of Ibd1 bounded genes are related to housekeeping functions (Saettone, Ponce, et al., 2019).

4.4 Localization of Ibd1 in Tetrahymena's genome

The ChIP-Seq data was examined for both ORFs and intergenic regions that showed greater than or equal to 4-fold enrichment to determine how Ibd1 is situated in the genome relative to ORF and intergenic regions. Using this fold-enrichment cut-off, 39 genic and 48 intergenic regions were obtained. The genic regions were first investigated to assess the Ibd1 peak distribution. (Figure 19) shows a representative example of Ibd1 ORF-specific localization where peaks are

primarily enriched within the gene-body (**Figure 19**, A, Appendix 2 for the full list). Next, to classify 48 intergenic regions, the ChIP-Seq read accumulation were manually inspected using the genome browser (Robinson et al., 2011b) and categorized them into five groups based on their localization (**Figure 19**, B-F, Appendix 2 for the full list). The promoter group showed intergenic localization that was proximal to the 5' region single predicted genes (e.g. **Figure 19**, B). The Ibd1 terminator group showed intergenic localization proximal to the 3' region of single predicted genes (e.g. **Figure 19**, C). The third intergenic group showed Ibd1 localization to regions where there is an overlap between the promoter of one predicted gene and the terminator of another (e.g. **Figure 19**, D). The fourth group showed localization of Ibd1 to single 5' promoter regions potentially controlling expression of two predicted genes (**Figure 19**, E). The fifth group showed localization of Ibd1 to single terminator 3' regions of two distinct predicted genes (**Figure 19**, F). It was found that among the 39 ORF showing $\geq 4X$ Ibd1 enrichment, some genes additionally showed enrichment through the promoter and the terminator region (**Figure 19**, G and H). Collectively these data suggest that Ibd1 appears to bind near the promoters and within gene bodies, consistent with a role in transcription regulation through its potential role in organizing multiple protein complexes (Saettone et al., 2018).

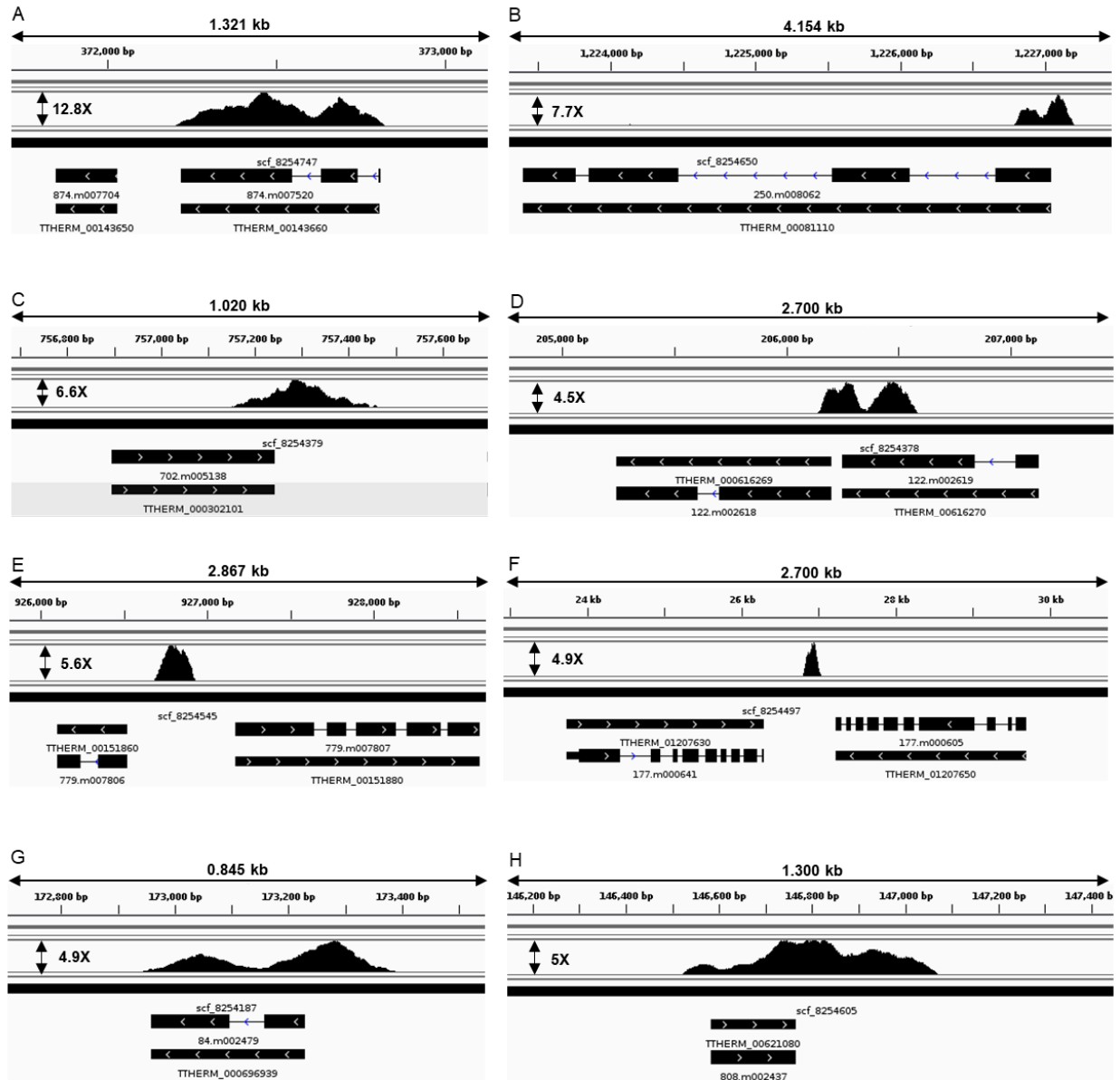


Figure 19: Ibd1 is localized in promoters, ORF and terminators.

In regions with more than or equal to 4-fold enrichment (IP/NPUT), Ibd1 localizes to 8 specific type-regions, including: **A.** ORF, **B.** promoters, **C.** terminators, **D.** localization in 2 regions where there is overlap between the promoter of one predicted gene and the terminator of another, **E.** regions showed localization to a single 5' promoter region potentially controlling expression of two predicted genes, **F.** localization to single terminators 3' regions of two distinct predicted genes. Combining these data for genes that present enrichment in the ORF and intergenic region, it was found that there is mutual enrichment in: **G.** regions that occupy from the promoter to the ORF, and **H.** regions that present enrichment from the ORF to the terminator region, (see Additional file 7 and 8 for Raw Data). The fold enrichments are presented beside each peak (Saettone et al., 2018).

4.5 Validation of Ibd1-FZZ ChIP-Seq

To validate our ChIP-Seq analysis of Ibd1 enriched chromatin, primers were designed for the three genes that showed the highest Ibd1-FZZ fold enrichment (Appendix 2) as well as a fourth, *PDD1* which is exclusively developmentally expressed (Smothers et al., 1997) and did not show enrichment for Ibd1-FZZ during growth (see GEO accession GSE103318). Our ChIP-qPCR analysis of the four genes confirmed specific enrichment of Ibd1-FZZ in *HTA3*, *RPS22*, and *HHF1* but not *PDD1* relative to chromatin made from untagged cells (**Figure 20, Table 4**). It was concluded that Ibd1 occupies transcriptionally active chromatin and might have a role in regulating the expression of a subset of genes involved in basal cellular housekeeping functions.

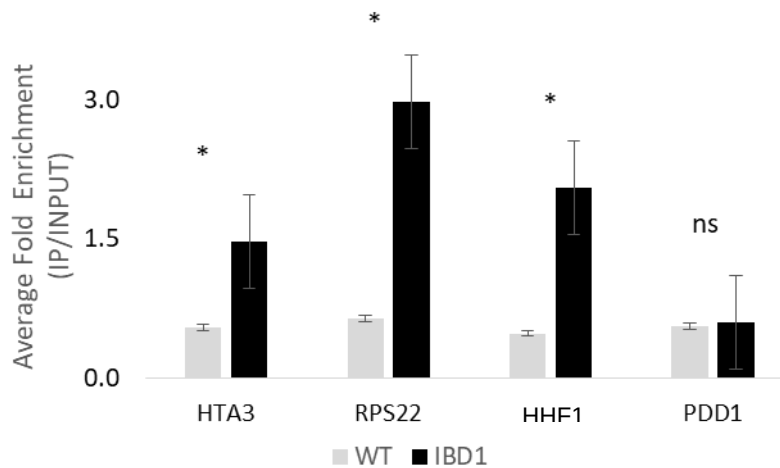


Figure 20: ChIP-qPCR validates ChIP-Seq results.

Anti-FLAG ChIP was performed in the 3 replicas of untagged and 4 replicas of Ibd1-FZZ during vegetative growth. ChIP DNA was amplified using primers to amplify *HTA3*, *RPS22*, *HHF1* and *PDD1* by real-time PCR using SYBR green. The significant p-values from the t-test are represented by a * (p-value <0.05). These significant p-values are 0.043 for *HTA3*, 0.041 for *RPS22* and 0.015 for *HHF1* this confirmed enrichment of Ibd1 in these genes. Our negative control, *Pdd1* shows a no significant p-value (ns) meaning no enrichment at this gene. The error bars represent the standard error of the mean for each sample. Primers used in (Saettone et al., 2018).

Table 4: ChIP-qPCR validates ChIP-Seq results.

Ibd1 is enriched in *HTA3*, *RPS22*, and *HHF1* but not *PDD1*. Each panel on the left represents the values obtained for each gene based on arbitrary ratios. On the right t-Test was used to compare the ratio values obtained for untagged (WT, Variable 1) and tagged (Ibd1, Variable 2) samples. Primers used in (Saettone et al., 2018).

HTA3								t-Test: Two-Sample Assuming Unequal Variances		
	WT1	WT2	WT3	IBD1_1	IBD1_2	IBD1_3	IBD1_4	Variable 1	Variable 2	
INPUT	4.27	2.67	2.85	4.13	3.24	3.28	3.50	Mean	0.544859021	1.469954763
IP	0.06	2.17	2.30	2.73	3.69	6.54	7.31	Variance	0.21060211	0.473280988
Fold Enrichment	0.01	0.81	0.81	0.66	1.14	1.99	2.09	Observations	3	4
Average Enrichment	0.54			1.47				Hypothesized Mean Difference	0	
STD	0.46			0.69				df	5	
SEM	0.32			0.49				t Stat	-2.130624279	
								P(T<=t) one-tail	0.04316793	
								t Critical one-tail	2.015048373	

RPS22								t-Test: Two-Sample Assuming Unequal Variances		
	WT1	WT2	WT3	IBD1_1	IBD1_2	IBD1_3	IBD1_4	Variable 1	Variable 2	
INPUT	4.67	2.71	2.97	4.11	2.64	3.036	3.225	Mean	0.646336249	2.976994645
IP	1.78	2.24	2.17	15.32	13.60	4.507	4.975	Variance	0.055202734	3.193128254
Fold Enrichment	0.38	0.83	0.73	3.73	5.15	1.48	1.54	Observations	3	4
Average Enrichment	0.65			2.98				Hypothesized Mean Difference	0	
STD	0.23			1.79				df	3	
SEM	0.17			1.26				t Stat	-2.579003151	
								P(T<=t) one-tail	0.040925287	
								t Critical one-tail	2.353363435	

HHF1								t-Test: Two-Sample Assuming Unequal Variances		
	WT1	WT2	WT3	IBD1_1	IBD1_2	IBD1_3	IBD1_4	Variable 1	Variable 2	
INPUT	4.21	2.70	3.04	3.16	3.05	3.68	3.90	Mean	0.48420265	2.052887298
IP	0.25	2.02	1.96	6.55	9.93	5.33	5.61	Variance	0.137853614	0.727543478
Fold Enrichment	0.06	0.75	0.64	2.07	3.25	1.45	1.44	Observations	3	4
Average Enrichment	0.48			2.05				Hypothesized Mean Difference	0	
STD	0.37			0.85				df	4	
SEM	0.26			0.60				t Stat	-3.286422905	
								P(T<=t) one-tail	0.015158588	
								t Critical one-tail	2.131846786	

PDD1								t-Test: Two-Sample Assuming Unequal Variances		
	WT1	WT2	WT3	IBD1_1	IBD1_2	IBD1_3	IBD1_4	Variable 1	Variable 2	
INPUT	1.43	3.13	3.40	15.12	0.88	2.78	3.62	Mean	0.563527113	0.601325044
IP	0.02	2.73	2.73	3.79	0.05	3.17	3.48	Variance	0.229419028	0.281011823
Fold Enrichment	0.01	0.87	0.80	0.25	0.05	1.14	0.96	Observations	3	4
Average Enrichment	0.56			0.60				Hypothesized Mean Difference	0	
STD	0.48			0.53				df	5	
SEM	0.34			0.37				t Stat	-0.098676684	
								P(T<=t) one-tail	0.462614448	
								t Critical one-tail	2.015048373	

5 CHAPTER 3: ChIP-Seq of Med31 indicates that Mediator regulates activation and repression of genes in *Tetrahymena*

The objective of this chapter is to test and improve the model derived from the proteomics pipeline presented in the introduction. The RACS pipeline was utilized to further understand the function of Med31.

5.1 *Med31 chromatin occupancy indicates a global role of Mediator in transcription*

To investigate the global role of *Tetrahymena* Mediator complex in transcription regulation, ChIP-Seq experiments in biological replicates using Med31-FZZ in vegetative cells were performed. ChIP-Seq data revealed that Med31 predominantly localizes to predicted genic regions (~70% of Med31-bound regions are genic, and ~90% of Med31 reads are in genic regions), covering roughly half of *Tetrahymena* predicted open reading frames (13715 /26996 ORFs) (**Figure 21**, Appendix 3). In contrast, ChIP-Seq analysis performed in control untagged *Tetrahymena* cells did not show such enrichments (Appendix 3).

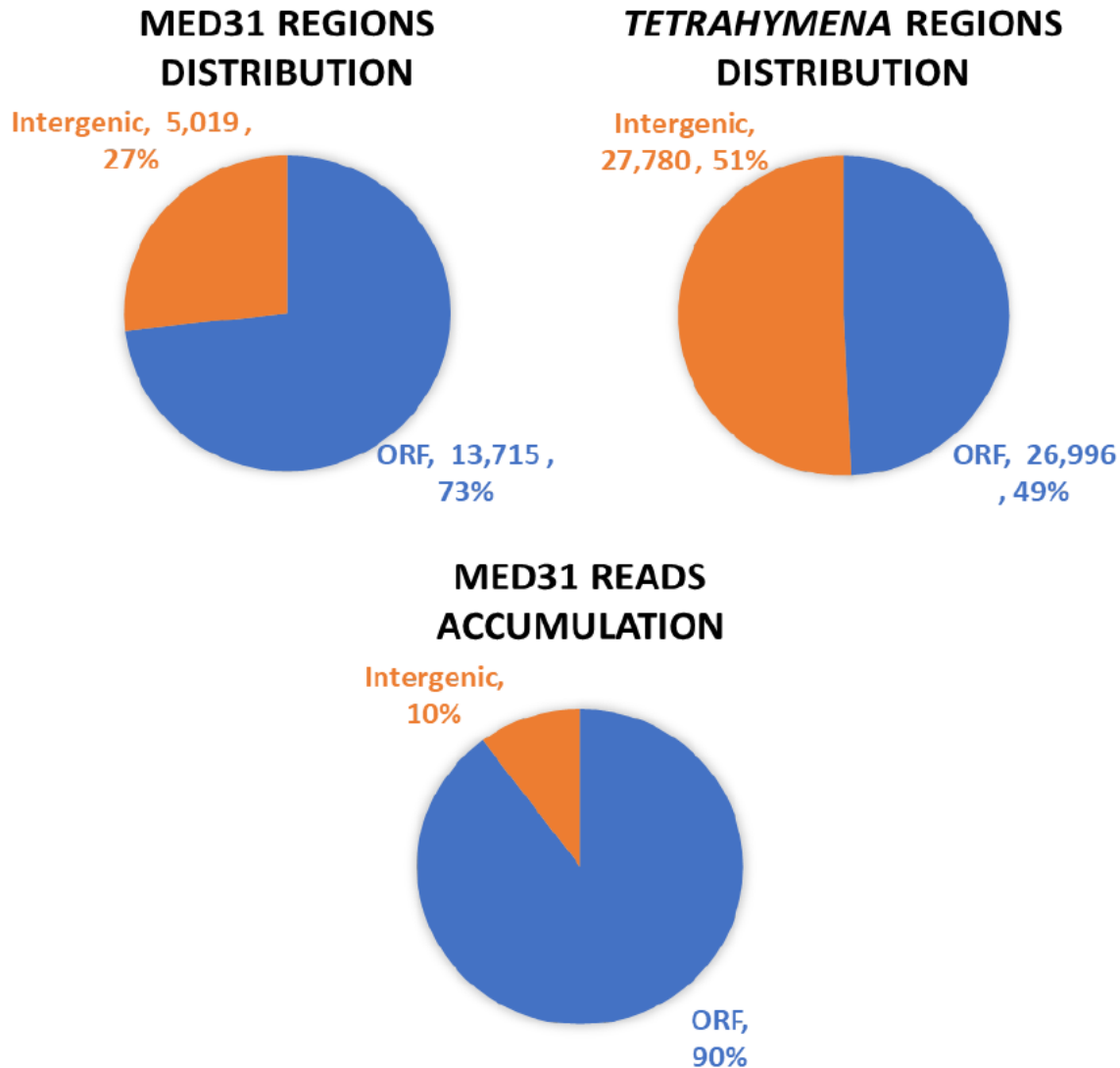


Figure 21: MED31-FZZ ChIP-Seq analysis in growing *Tetrahymena*.

A-Top left: Distribution of MED31 in terms of ORF (genic) and intergenic bound regions. **Top right:** Distribution of *Tetrahymena* genomic regions. MED31 localizes to 13,715 open reading frames (ORFs) (left) out of the 26,996 (right) annotated genes (Eisen et al., 2006). **Bottom:** MED31 reads predominantly are found within ORFs (Garg et al., 2019).

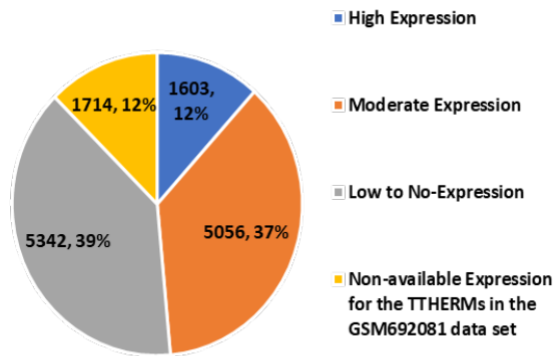
5.2 *Med31* function may be related to gene activation and repression

RNA-Seq data that have been used to rank *Tetrahymena* genes based on their expression profile during vegetative growth (Xiong et al., 2012) (GEO accession GSM692081) indicated that the

majority of the Med31-FZZ bound genes (76% of the total bound genes) are either moderately or very weakly expressed during *Tetrahymena* growth, i.e., 37% and 39% of the Med31-FZZ bound genes are classified as moderately and weakly expressed respectively (**Figure 21**).

It was observed that ~49% of *Tetrahymena* highly and moderately expressed genes are bound by Med31, suggesting a positive role for Mediator in the transcriptional activation of these genes. In contrast, 39% of *Tetrahymena* low to no-expression genes are found within the Med31 targets (**Figure 22**, top left panel). The Mediator-binding profile was examined across the target ORFs using metagene analysis. Our analysis indicated that Med31 binds throughout the predicted ORFs and the binding density is significantly higher for highly expressed genes in comparison to the other expression groups (**Figure 22**, lower panels). The metagene analysis was performed using ChIP-Seq reads normalized over the inputs and by ‘Reads Per Kilobase of transcript per Million mapped reads (RPKM)’ values. The plots were generated using Deeptools (Ramírez et al., 2014).

Med31 ORF Expression Profile



Tetrahymena ORF Expression Profile

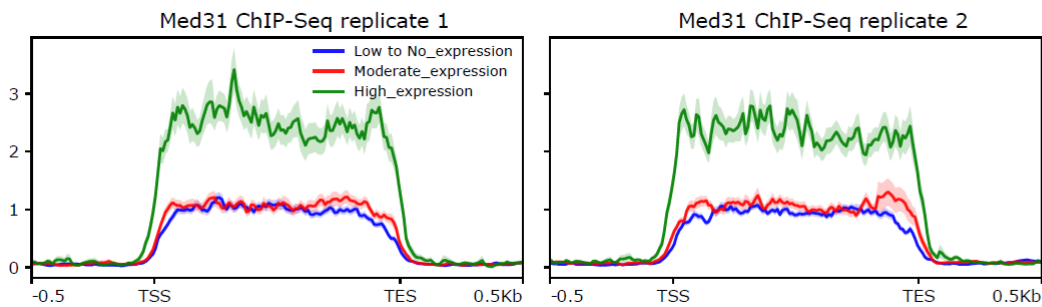
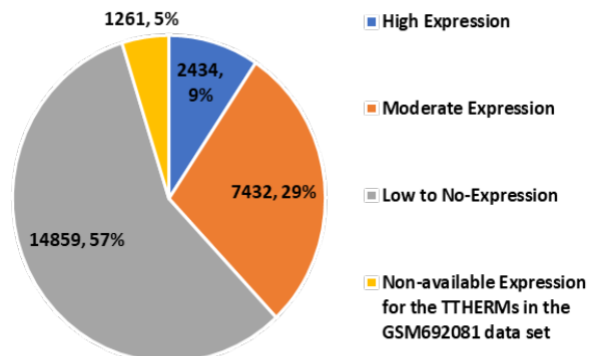


Figure 22: Med31 Expression profile.

Top left -During vegetative growth MED31 targets ORFs that are either moderately, low, or not expressed. **Top right:** *Tetrahymena* ORFs are grouped based on their expression levels as previously reported (Xiong et al., 2012). **Bottom:** Metagene analysis of Med31 ChIP-Seq two replicates. The plot was calculated by taking ChIP over the input ratio. The shaded area represents the standard error of the mean (Garg et al., 2019).

5.3 Gene Ontology (GO) annotation analysis of Med31 bound genes suggests

function in diverse cellular processes

The wide-spread binding of Med31 across genic regions suggests a general role for *Tetrahymena* Mediator complex in transcription regulation. Consistent with this, gene ontology (GO) analysis indicated that Med31-bound genes have diverse cellular functions (**Figure 23**, Appendix 3).

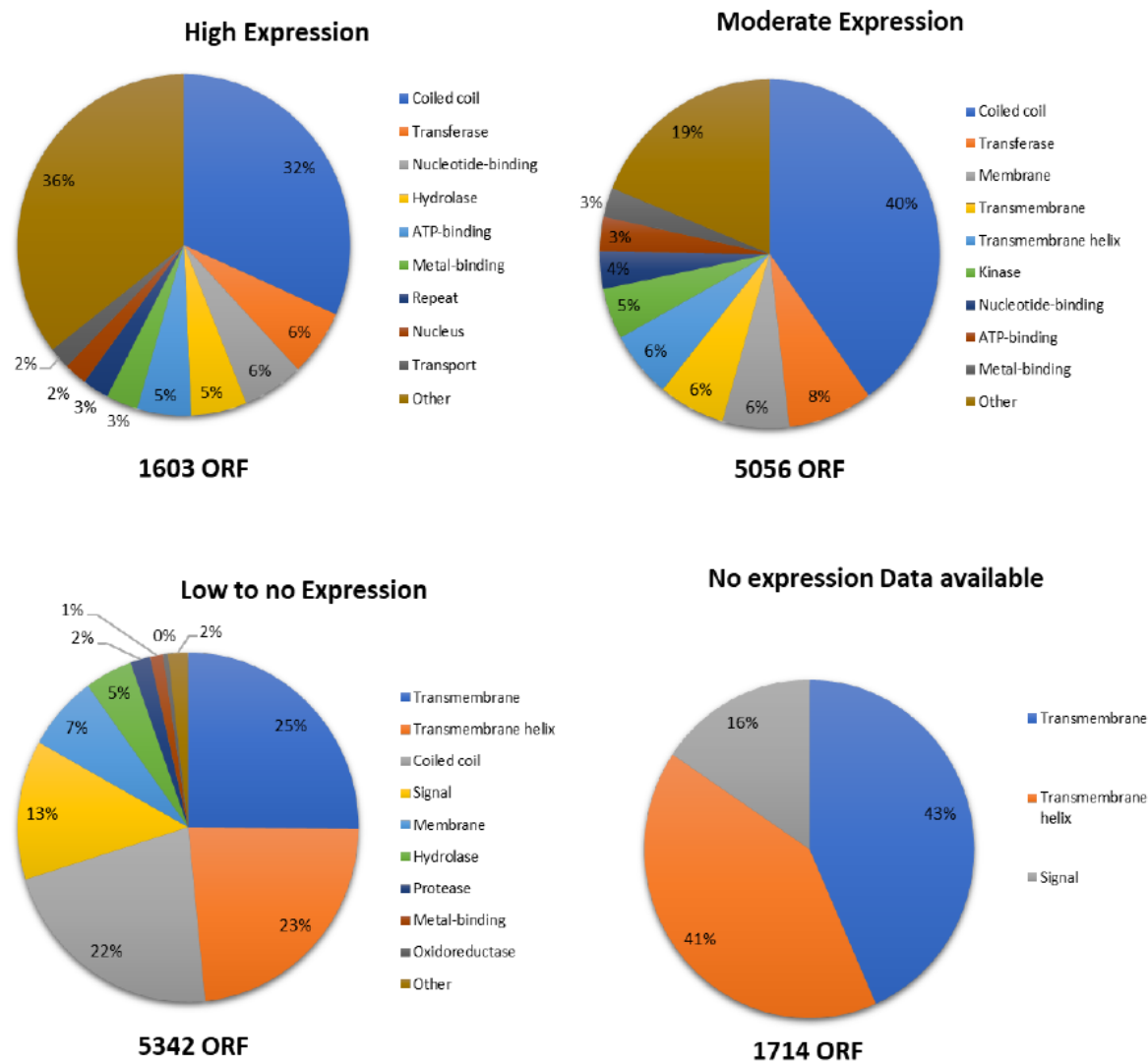


Figure 23: GO analysis of genes enriched by Med31-FZZ.

GO enrichment analysis of Med31 bound genes based on their expression levels during growth. The analysis was carried out using UniprotKB keywords (Garg et al., 2019).

5.4 Validation of Med31-FZZ ChIP-Seq

5.4.1 Med31 negatively regulates the expression of developmental genes during growth

The observation that a large subset of Med31 bound genes are weakly expressed during growth suggested that *Tetrahymena* Mediator complex might negatively regulate the expression of at

least some of its target genes. Developmentally expressed genes such as *PDD1* remain transcriptionally silent during *Tetrahymena* vegetative growth via a poorly understood mechanism. Among Med31 ChIP-Seq identified targets were many developmentally regulated genes including those with known functions in programmed DNA rearrangements and irreversible gene silencing such as *PDD1*, *PDD2* and *TWI1* (Kataoka & Mochizuki, 2011; Smothers et al., 1997) (**Figure 24**). This observation suggested that Mediator might have a role in the regulated expression of developmental genes.

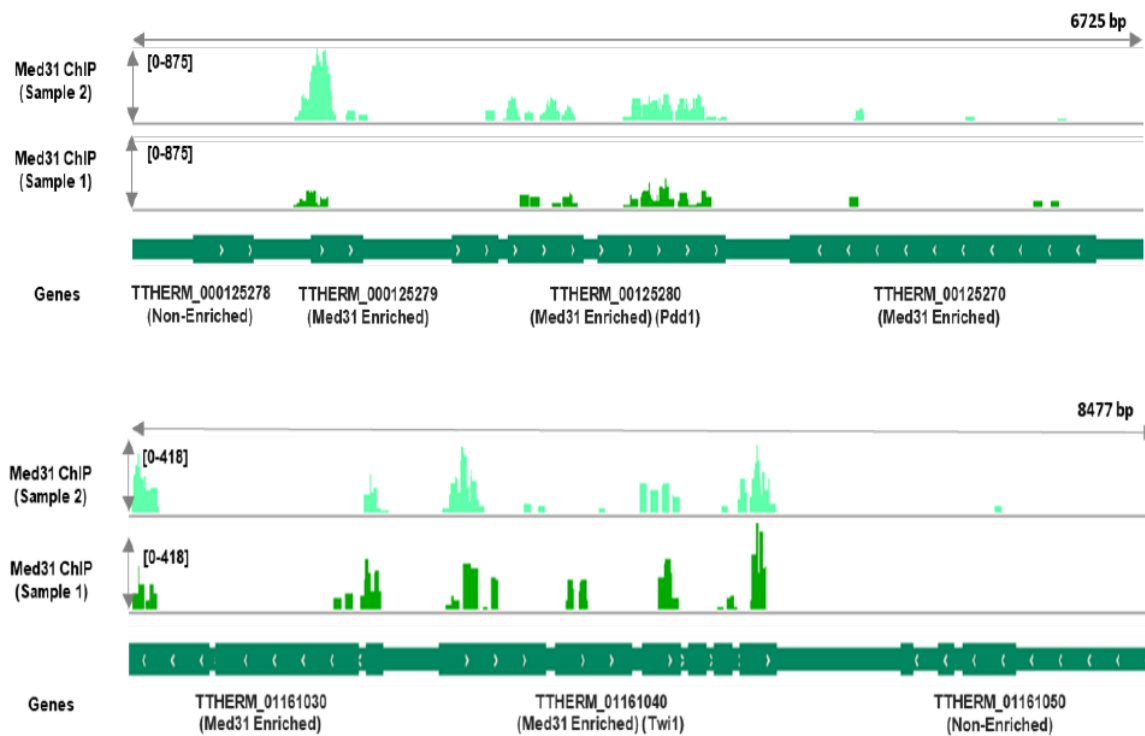


Figure 24: Med31 regulates the expression of developmental genes in *Tetrahymena*.

Genome browser screenshots of Med31 ChIP-Seq peaks on representative genes. Peaks for two independent ChIP replicates are shown on Med31 targets (*PDD1* and *TWI1*) as well as on non-target genes (Garg et al., 2019).

To further examine this possibility, the *MED31* locus was deleted via homologous recombination mediated gene replacement. Southern blot analysis revealed that most of the *MED31* wild-type

gene copies were replaced with our knockout construct (Garg et al., 2019). RT-qPCR experiments demonstrated a significantly reduced, but not completely abolished, expression of *MED31* (Table 5, Table 6, Figure 25) suggesting that *MED31* might be essential for growth, as it is in most human cell lines (Meyers et al., 2017).

To quantify the abundance of targeted RNA, a relative quantification approach was employed. The objective of this experiment was to compare the levels of RNA expression for untagged vs the Med31-KD strains and to relative quantify the target RNA to verify expression levels in comparison to a reference gene based on the formula Ratio= Target/Reference. The targeted RNA used in this study were Med31, Pdd1, Pdd2 and Pdd3 and the reference gene was the histone HTA3.

Table 5: HTA3 ChIP-qPCR data.

Unkn: Represents the qPCR using reverse transcriptase. NegCtrl: Represents qPCR not using reverse transcriptase. Green panel represents the standard curve made with dilutions using genomic DNA as template. Primers used in (Saettone et al., 2018) (Garg et al., 2019).

Reference: HTA3						
<i>Fluor</i>	<i>Target</i>	<i>Content</i>	<i>Sample</i>	<i>Cq</i>	<i>SQ (starting quantity)</i>	<i>SQ Average</i>
SYBR	HTA3	Unkn	WT	21.23	137.70	144.50
SYBR	HTA3	Unkn	WT	21.12	148.20	
SYBR	HTA3	Unkn	WT	21.13	147.60	
SYBR	HTA3	Unkn	Med31KO	21.16	145.10	107.29
SYBR	HTA3	Unkn	Med31KO	21.97	83.40	
SYBR	HTA3	Unkn	Med31KO	21.80	93.37	
SYBR	HTA3	Neg Ctrl	Med31KO	36.60	0.00	0.00
SYBR	HTA3	Neg Ctrl	Med31KO	36.15	0.01	
SYBR	HTA3	Neg Ctrl	Med31KO	36.10	0.01	
SYBR	HTA3	Neg Ctrl	WT	39.54	0.00	0.00
SYBR	HTA3	Neg Ctrl	WT	39.94	0.00	
SYBR	HTA3	Neg Ctrl	WT	N/A	0.00	
SYBR	HTA3	Std		21.76	100.00	
SYBR	HTA3	Std		24.89	10.00	
SYBR	HTA3	Std		28.55	1.00	
SYBR	HTA3	Std		21.76	100.00	
SYBR	HTA3	Std		25.02	10.00	
SYBR	HTA3	Std		28.45	1.00	

Table 6: Med31 ChIP-qPCR data.

Unkn: Represents the qPCR using reverse transcriptase. NegCtrl: Represents qPCR not using reverse transcriptase. Green panel represents the standard curve made with dilutions using genomic DNA as template. Bottom right panel has t-Test was used to compare the ratio values obtained for Variable 1: Med31/HTA3 (Table 5) in WT cells; and Variable 2: Med31/HTA3 in Med31KD cells. Primers used in (Saettone et al., 2018). Data in red was not used (Garg et al., 2019).

Target 1: Med31

Fluor	Target	Content	Sample	Cq	SQ (starting quantity)	(T:SQ/R:SQ) Ave	Ratio Average	Standard deviation
SYBR	Med-2	Unkn	WT	30.93	0.18	1.23E-03		
SYBR	Med-2	Unkn	WT	31.49	0.11	7.54E-04	7.81E-04	4.40E-04
SYBR	Med-2	Unkn	WT	32.34	0.05	3.56E-04		
SYBR	Med-2	Unkn	Med31KO	35.40	0.00	3.24E-05		
SYBR	Med-2	Unkn	Med31KO	35.25	0.00	3.72E-05	3.48E-05	3.43E-06
SYBR	Med-2	Unkn	Med31KO	30.62	0.23	2.18E-03		
SYBR	Med-2	Neg Ctrl	WT	N/A	0.00	0.00E+00		
SYBR	Med-2	Neg Ctrl	WT	N/A	0.00	0.00E+00	0.00E+00	
SYBR	Med-2	Neg Ctrl	WT	N/A	0.00	0.00E+00		
SYBR	Med-2	Neg Ctrl	Med31KO	N/A	0.00	0.00E+00		
SYBR	Med-2	Neg Ctrl	Med31KO	N/A	0.00	0.00E+00	0.00E+00	
SYBR	Med-2	Neg Ctrl	Med31KO	N/A	0.00	0.00E+00		
SYBR	Med-2	Std		24.37	100.00	t-Test: Two-Sample Assuming Unequal Variances		
SYBR	Med-2	Std		25.75	10.00		Variable 1	Variable 2
SYBR	Med-2	Std		29.26	1.00	Mean	0.000781361	3.47749E-05
SYBR	Med-2	Std		23.82	100.00	Variance	1.93297E-07	1.17451E-11
SYBR	Med-2	Std		25.49	10.00	Observation	3	2
SYBR	Med-2	Std		29.40	1.00	Hypothesize	0	
						df	2	
						t Stat	2.941092591	
						P(T<=t) one-t	0.049387355	*
						t Critical one	2.91998558	
						P(T<=t) two-t	0.098774709	
						t Critical two	4.30265273	

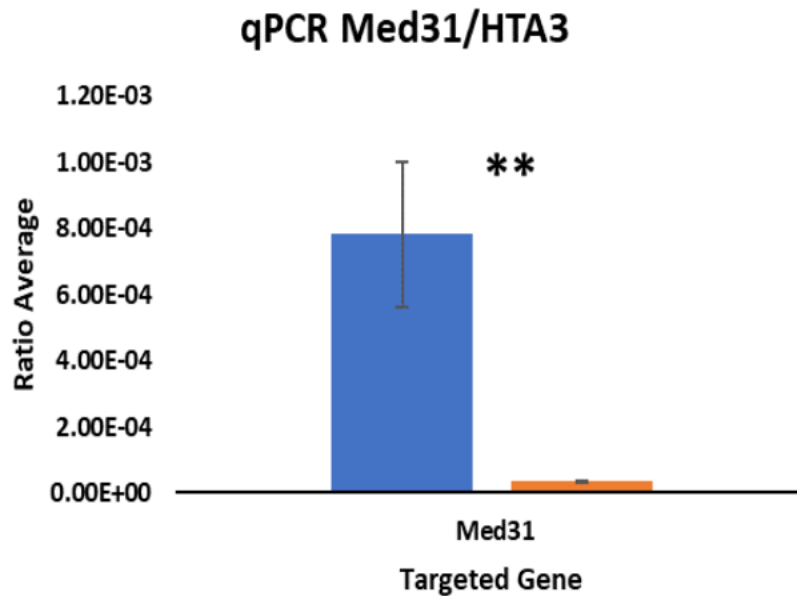


Figure 25: RT-qPCR analysis to examine the expression levels of Med31.

These results are upon homologous recombination mediated gene replacement experiments (Garg et al., 2019). RNA was isolated either from the control wild type or from the cells treated with vectors for *MED31* gene deletion (*MED31* KD). The RT-qPCR data were normalized using *HTA3*. ** indicates p-value of <0.05 (Garg et al., 2019).

Table 7: Pdd1 ChIP-qPCR data.

Unkn: Represents the qPCR using reverse transcriptase. NegCtrl: Represents qPCR not using reverse transcriptase. Green panel represents the standard curve made with dilutions using genomic DNA as template. Bottom right panel has t-Test was used to compare the ratio values obtained for Variable 1: Pdd1/HTA3 (Table 5) in WT cells; and Variable 2: Pdd1/HTA3 in Med31KD cells. Primers used in (Saettone et al., 2018) (Garg et al., 2019)..

Target 2: Pdd1

Fluor	Target	Content	Sample	Cq	SQ (starting quantity)	Ratio (T:SQ/R Ratio Average	Standard deviation
SYBR	Pdd1	Unkn	WT	32.58	0.05	3.61E-04	
SYBR	Pdd1	Unkn	WT	33.02	0.04	2.65E-04	3.21E-04
SYBR	Pdd1	Unkn	WT	32.68	0.05	3.37E-04	5.02215E-05
SYBR	Pdd1	Unkn	Med31KO	32.49	0.06	5.20E-04	
SYBR	Pdd1	Unkn	Med31KO	32.11	0.07	6.78E-04	5.91E-04
SYBR	Pdd1	Unkn	Med31KO	32.34	0.06	5.76E-04	8.04475E-05
SYBR	Pdd1	Neg Ctrl	WT	32.60	0.00	0.00E+00	
SYBR	Pdd1	Neg Ctrl	WT	32.81	0.00	0.00E+00	0.00E+00
SYBR	Pdd1	Neg Ctrl	WT	33.34	0.00	0.00E+00	
SYBR	Pdd1	Neg Ctrl	Med31KO	32.19	0.00	0.00E+00	
SYBR	Pdd1	Neg Ctrl	Med31KO	33.12	0.00	0.00E+00	0.00E+00
SYBR	Pdd1	Neg Ctrl	Med31KO	32.74	0.00	0.00E+00	
SYBR	Pdd1	Std		21.99	100.00	t-Test: Two-Sample Assuming Unequal Variances	
SYBR	Pdd1	Std		25.10	10.00	Variable 1	Variable 2
SYBR	Pdd1	Std		28.19	1.00	Mean	0.000320946
SYBR	Pdd1	Std		21.73	100.00	Variance	2.5222E-09
SYBR	Pdd1	Std		25.17	10.00	Observation	3
SYBR	Pdd1	Std		28.59	1.00	Hypothesize	0
						df	3
						t Stat	-4.939217429
						P(T<=t) one-t	0.007958349 **
						t Critical one	2.353363435
						P(T<=t) two-t	0.015916699
						t Critical two	3.182446305

Similar results were obtained by RT-qPCR experiments using the RNA isolated from either the control or *MED31* KD cells (**Table 7, Figure 26**). The expression of *PDD1*, *PDD2* and *PDD3* was examined during vegetative growth by RT-qPCR experiments using RNA isolated from either the control or *MED31* KD cells (**Figure 26**). Both *PDD1* and *PDD2* exhibited an upregulation in their expression levels during vegetative growth whereas *PDD3* did not.

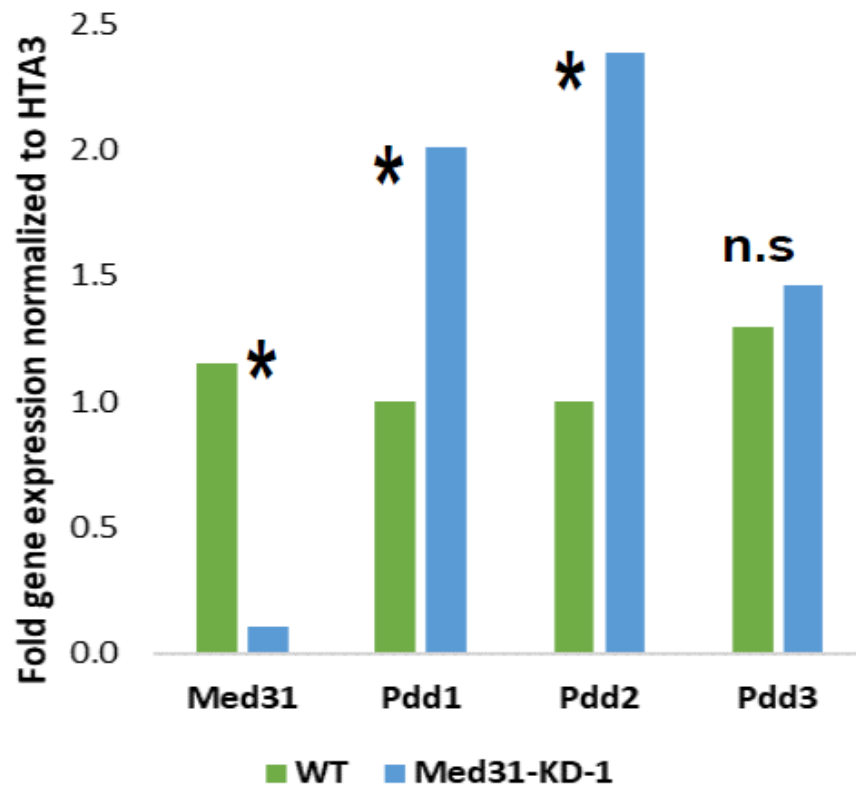


Figure 26: RT-qPCR analysis of *PDD1*, *PDD2* and *PDD3* expression using RNA extracted either from wild-type cells or from *MED31* KD cells.

The data were normalized to *HTA3*. * indicated P-value<0.05 (Garg et al., 2019).

This result is consistent with the observation that *PDD1* and *PDD2* harbor Med31 ChIP peaks while *PDD3* does not (**Figure 27**).



Figure 27: Genome browser screenshots for Med31 ChIP-Seq peaks over *PDD1*, *PDD2* and *PDD3* loci.

PDD1 and *PDD2* have Med31 enriched ChIP-Seq peaks whereas *PDD3* does not show any Med31 binding (Garg et al., 2019).

5.4.2 *MED31* KD impacts *Tetrahymena* cellular proliferation

Since expression of *MED31* was greatly reduced (>90%) in these knockdown (KD) cells (**Figure 25**), it was decided to characterize them further. Consistent with the idea that Mediator functions as a general transcriptional regulator, *MED31* KD in *Tetrahymena* results in a slower growth phenotype during vegetative growth (**Figure 28**).

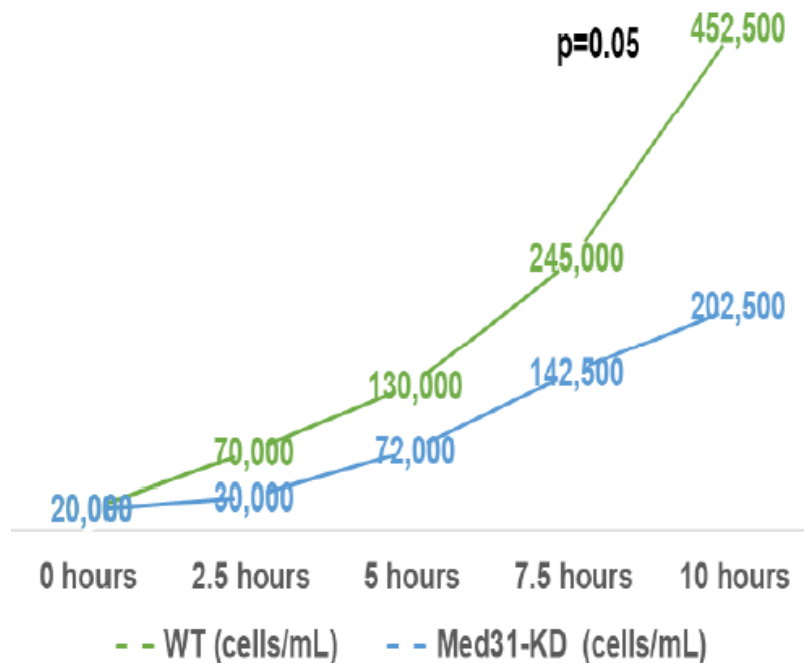


Figure 28: *MED31* KD impacts *Tetrahymena* cellular proliferation.

Equal number of log phase *MED31* KD and wild-type cells were inoculated and were counted at the indicated time points. *MED31* KD cells grew slower than their wild-type counterparts (Garg et al., 2019).

5.4.3 *Med31* regulates the expression of the developmental gene *Pdd1*

The expression of *Pdd1* was further examined in growing cells by Western blotting using whole cell extracts prepared either from *MED31* KD or control cells. A measurable amount of *Pdd1* was expressed in the *MED31* KD cells in contrast to the control cells where no expression was observed (**Figure 29**).

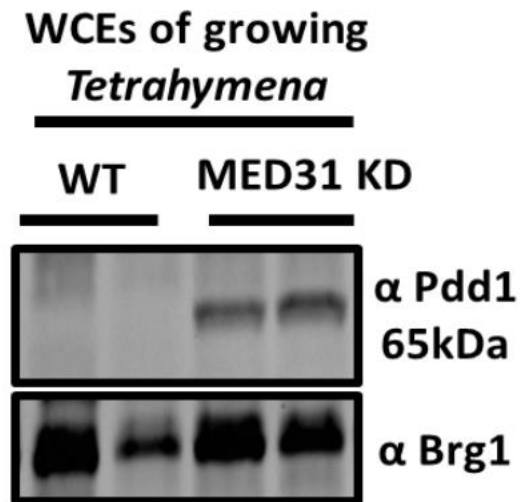


Figure 29: Western blot analysis of Pdd1 expression.

WCEs prepared from duplicate samples of either wild-type or *MED31* KD *Tetrahymena* cells during vegetative growth. The blot was probed with anti-Pdd1 antibody. Anti-Brg1 was used as loading control (Garg et al., 2019).

5.5 MED31 KD show a range of DAPI-positive structures that could be related to ectopic expression of development specific genes

Since *PDD1*, *PDD2* and *TWI1* have essential roles during development in programmed DNA rearrangements and irreversible gene silencing (Madireddi et al., 1996; Mochizuki et al., 2002), the localization patterns of Pdd1 in *MED31* KD cells were examined. An anti-Pdd1 antibody was used to monitor the localization of Pdd1 and observed that it localized to both the MAC and MIC in ~50% of growing cells. The signal was observed only upon *MED31* knockdown, consistent with our Western blotting and RT-qPCR experiments (**Figure 30, Figure 31**).

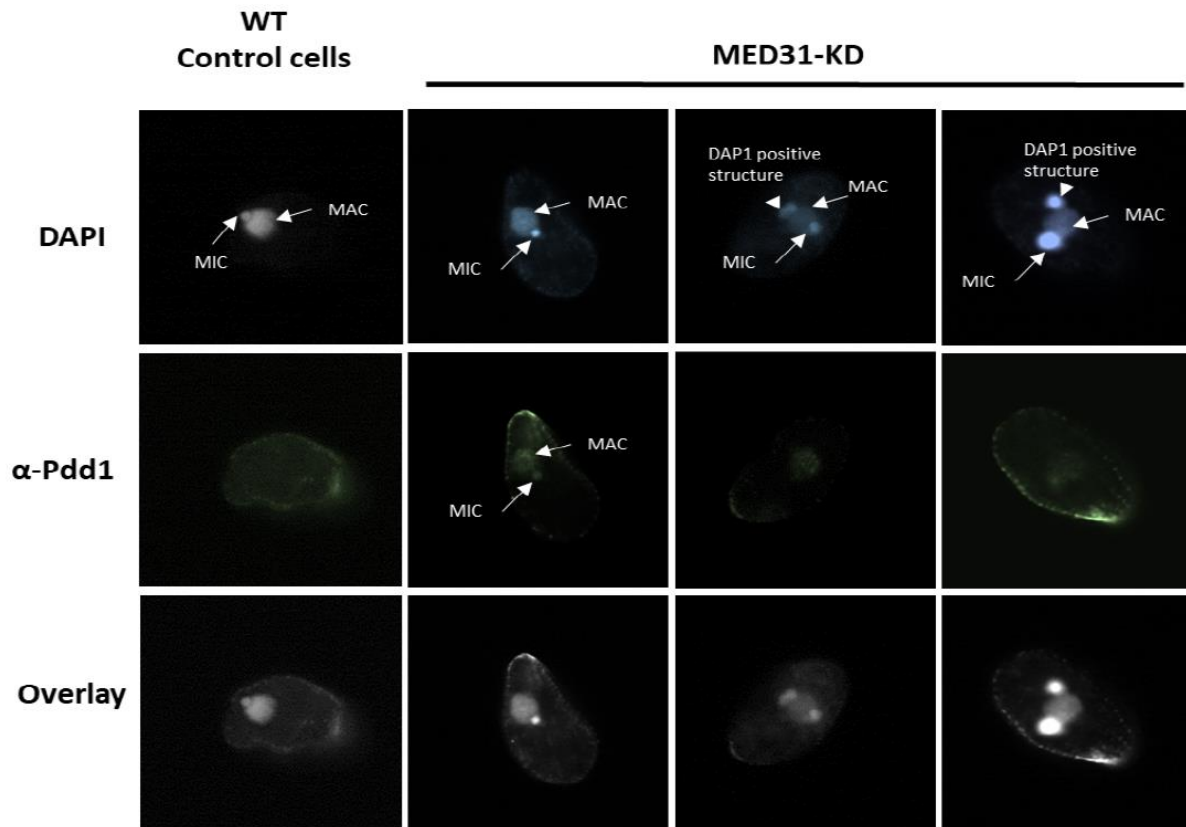


Figure 30: Immunofluorescence analysis using either *MED31* KD or wild-type growing *Tetrahymena* cells.

DAPI was used to stain nuclear material. Anti-Pdd1 antibody was used to examine the expression and localization of Pdd1. The position of MAC, MIC and additional DAPI staining structures is indicated on the images (Garg et al., 2019).

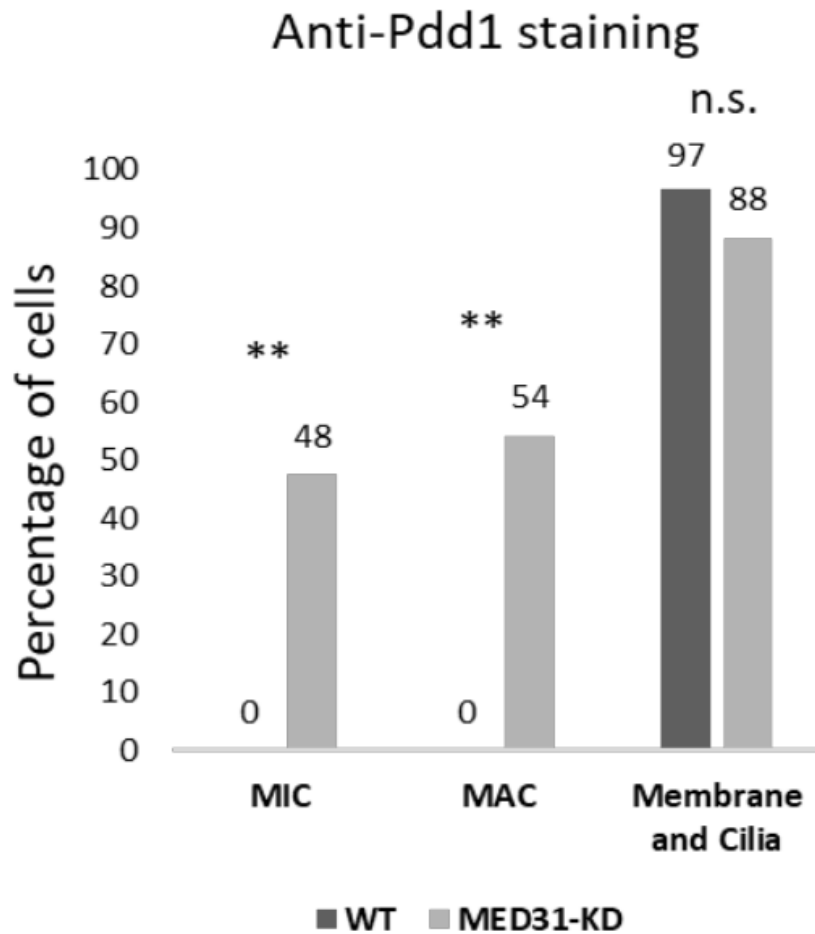


Figure 31: Quantification of cells with Pdd1 staining in *MED31* KD and wild-type controls. Cells were counted (100 cells for each condition) and ** represent $p < 0.01$ for staining in the MIC and MAC. The non-specific staining was observed in the outer membrane and cilia in both the *MED31* KD and wild-type cells (Garg et al., 2019).

DAPI staining also revealed the presence of DAPI positive structures in roughly half of the examined *MED31* KD cells (**Figure 30** and **Figure 32**). The origin and exact nature of these structures remains unclear. However, these structures are reminiscent of previously reported extranuclear chromatin bodies (extrusion bodies or EBs) and differentially staining nuclear material associated with nuclear integrity and chromatin condensation upon loss of histone deacetylase Thd1p in *Tetrahymena* (Parker, Maxson, Mooney, & Wiley, 2007; Wiley, Myers,

Parker, Braun, & Yao, 2005). These data suggest that the altered expression of one or more genes in *MED31* KD cells might result in the formation of these structures. Further experiments will be required to examine the nature and mechanism of the formation of these structures.

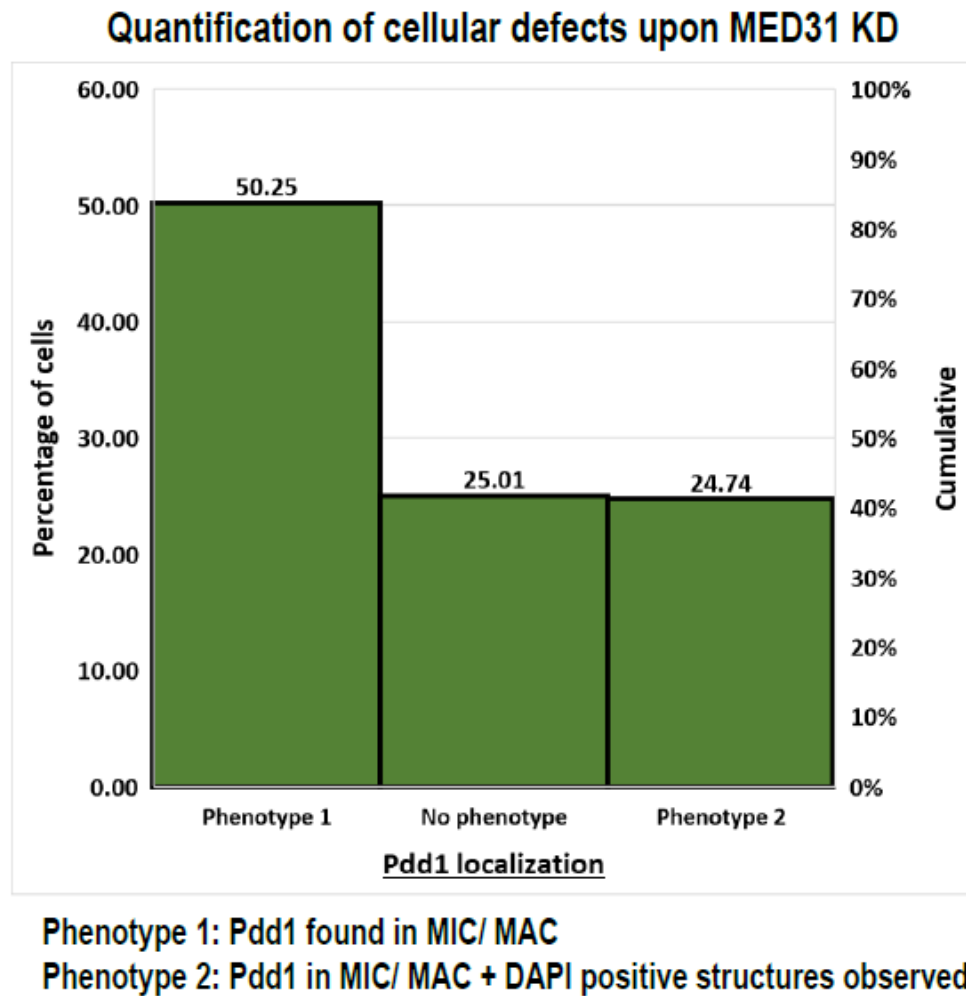


Figure 32: Quantification of cells with DAPI positive structures upon *MED31* KD. Cells were counted (100 cells) and percentage was calculated using all the cells under the field (Garg et al., 2019).

5.6 *MED31* KD cells shows γ H2A.X foci in the MAC

Furthermore, it was observed that 26% of the *MED31* KD cells exhibited MAC-specific γ H2A.X staining. This suggests that the ectopic Pdd1 may be recognizing H3K27me3 in MAC to direct elimination of DNA in a similar manner to IES elimination as it is performed during MAC development. RACS shows that there are other proteins Pdd2, Lia1, Lia2, Lia4, Lia5 and Tpb6 involved in IES elimination that may be collaborating with Pdd1 to direct DNA elimination as a result γ H2A.X is found in MAC (**Figure 33** and **Figure 34**). In contrast WT cells only showed weak MIC staining for γ H2A.X, consistent with previous observations (Gao et al., 2013) (**Figure 34**). The data suggest that the altered expression of one or more genes in *MED31* KD cells might result in the formation of these DAPI positive structures.

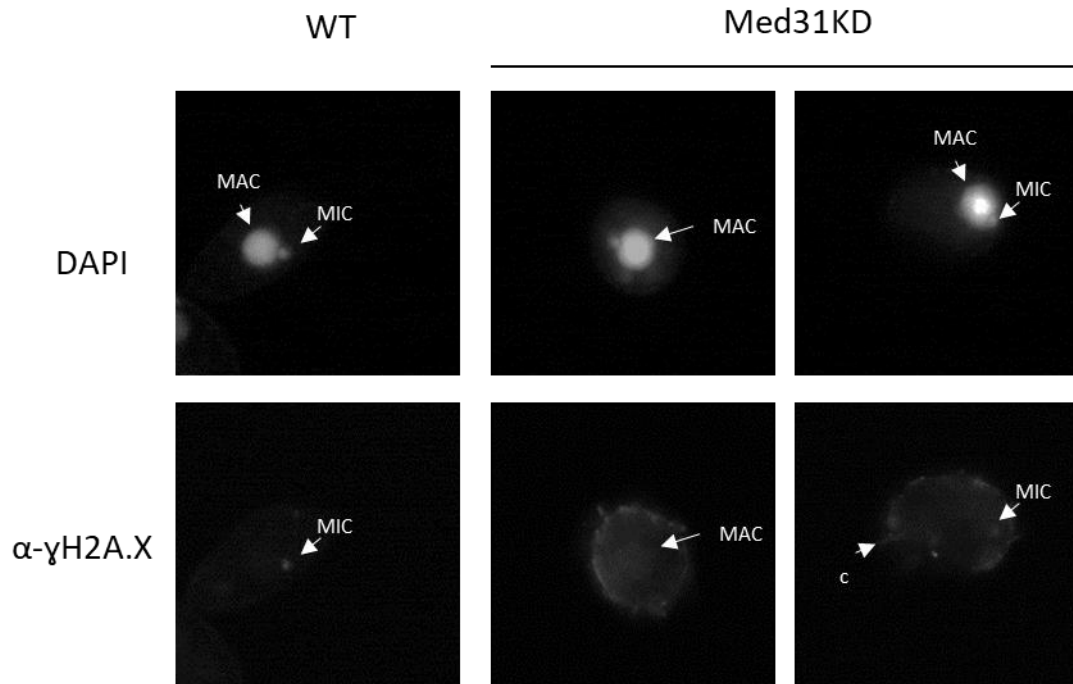


Figure 33: Immunofluorescence analysis of MED31 KD Tetrahymena cells shows γ H2A.X foci in the MAC.

DAPI was used to stain nuclear material. Anti- γ H2A.X antibody was used to examine the expression and localization of γ H2A.X. The position of MAC, MIC and additional DAPI staining structures is indicated on the images (Garg et al., 2019).

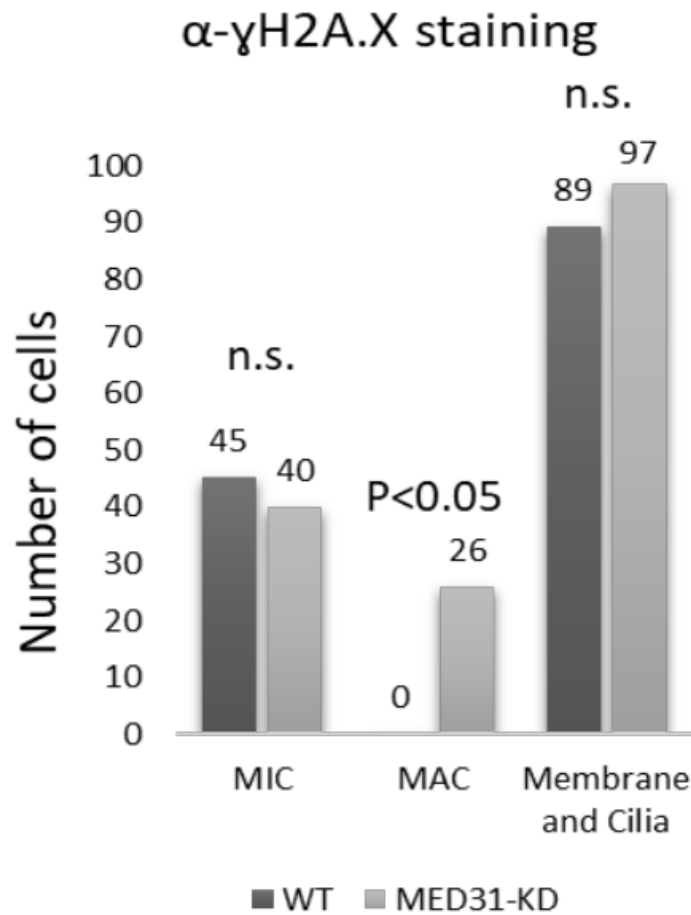


Figure 34: *MED31* KD results in abnormal γ H2A.X foci in the MAC.

Quantification of cells with α - γ H2A.X staining in *MED31* KD and wild-type controls. Signal was observed in MIC of both the *MED31* KD and wild-type cells whereas the MAC was only stained in *MED31* KD cells. The non-specific staining in the outer membrane and cilia was observed in both the *MED31* KD and wild-type cells (Garg et al., 2019).

6 DISCUSSION

6.1 *RACS as a new methodology for analysis of ChIP-Seq data originating from contig-based genomes*

RACS is a pipeline that utilizes open source tools for the rapid analysis of ChIP-Seq data for a POI from an organism with a contig-based genome sequence. RACS utilizes the predicted gene coordinates to group and rank the reads accumulation for genic and intergenic regions. The objective is to be able to infer hypotheses concerning POI function based on its chromatin occupancy. This pipeline has been applied to *Tetrahymena thermophila*'s ChIP-Seq data, but its application can be extended to ChIP-Seq datasets generated in any other organisms (Saettone, Ponce, et al., 2019). Initially, when peaks were called for Ibd1 ChIP-Seq data using MACS2, the output did not indicate whether the protein localizes to a genic or an intergenic region. MACS2 provides the coordinates where the peaks were found. MACS2 calls all peaks regardless of their position in a genic or intergenic region, which makes interpretation difficult when combined with the minimal annotation of the *Tetrahymena* genome. RACS uses ChIP-Seq to account for enrichment of DNA reads found after the immunoprecipitation (IP) versus DNA reads found in the whole cell extracts (INPUT). RACS segregates the ranked enrichment of genic and intergenic regions in two tables (Appendices 2 and 3). This can help to quickly assign biological function to a POI based on the predicted function of a set of genes or regulatory (intergenic) regions it occupies. It was noted that other programs, such as BEDTools, can be used to perform this task in combination with MACS2. Without the need of any additional “external” software, RACS calls peaks and segregates them into two tables based on the given set of coordinates (genes) or

the remaining regions in the genome that were not provided by the user (intergenic). Thus, our pipeline is appropriate to address biological questions regarding protein function based on genome position.

MACS2 and RACS are complementary to each other, but are *not* dependent on each other for analysis. For example, MACS2 can be used to establish or to generate a set of coordinates for a specific transcription protein binding to the genome. Thus, it can be inferred that the POI is attaching to specific areas in the genome to control transcription and these regions could be annotated as binding sites for the specific transcription protein. Subsequently, if ChIP-Seq is performed on a different protein that has also been shown to physically interact with the transcription protein previously mentioned, the user could use the coordinates given by MACS2 to generate a .gff3 file to input alongside the genome file to the RACS pipeline. This will allow us to rapidly determine the degree of overlap and potential co-localization in some or all binding sites. In that, MACS and RACS can synergize to provide a powerful tool for the analysis of less developed genome sequences.

Even though there are many computational tools available for processing ChIP-Seq data, RACS is particularly suitable for the analysis of contig-based genome sequence with associated minimal annotation. Other tools, such as, MACS2 and metagene using deepTools analysis (Ramírez et al., 2014) complement RACS. Recently several ChIP-Seq studies (Garg et al., 2019; Kataoka & Mochizuki, 2015; Saettone et al., 2018; Y. Wang, Chen, Sheng, Liu, & Gao, 2017), have emerged for *Tetrahymena thermophila*. However, there is a lack of standardized computational

methods for this model organism, hence it becomes difficult to reliably reach the same conclusions when replicating the findings. Our tool is the first application designed for use with *Tetrahymena thermophila* to provide a community resource for genome-wide ChIP-Seq studies. This allows for RACS to potentially contribute to the standardization of ChIP-Seq analyses in ciliates. The refinement of RACS is intended to continue based on community needs. For example, recently, single-molecule sequencing based on nanopores has emerged as a promising technology with a potential to revolutionize the field of genomics (Loose, 2017). Nanopore sequencing provides the advantages of 1) long reads, enabling the de novo transcriptome analysis (Byrne et al., 2017), 2) point-of-care, making real-time analysis possible (Lu, Giordano, & Ning, 2016), and 3) PCR free, allowing the direct identification of epigenetics (Simpson et al., 2017). Considering its promising outcomes, studies using model organisms with divergent genomes, such ciliates and parasitic organisms including *Trypanosoma*, will be particularly benefited from the nanopore sequencing technology (Díaz-Viraqué et al., 2019). Currently, a major challenge is to develop sophisticated and high-performance computational tools to interpret and analyze the nanopore sequencing data (Han, Li, Gao, & Wang, 2018; Y. Li et al., 2018; Teng et al., 2018). In the future, the goal is to improve and implement the RACS pipeline for the analysis of nanopore sequencing data.

6.2 Model for Ibd1 function

Bromodomains (BRD) are PTM readers that specifically binds acetyl lysines (Kac) within proteins, such as histones, to regulate transcription (Jain & Barton, 2017). Consistent with a function in promoting transcription, the BRD containing protein Ibd1-FZZ localizes to several

regulatory (intergenic) regions and can also be found in coding regions (**Figure 16, C**).

Particularly, Ibd1-FZZ accumulates within coding regions of multiple genes (**Figure 16, D, Figure 35**).

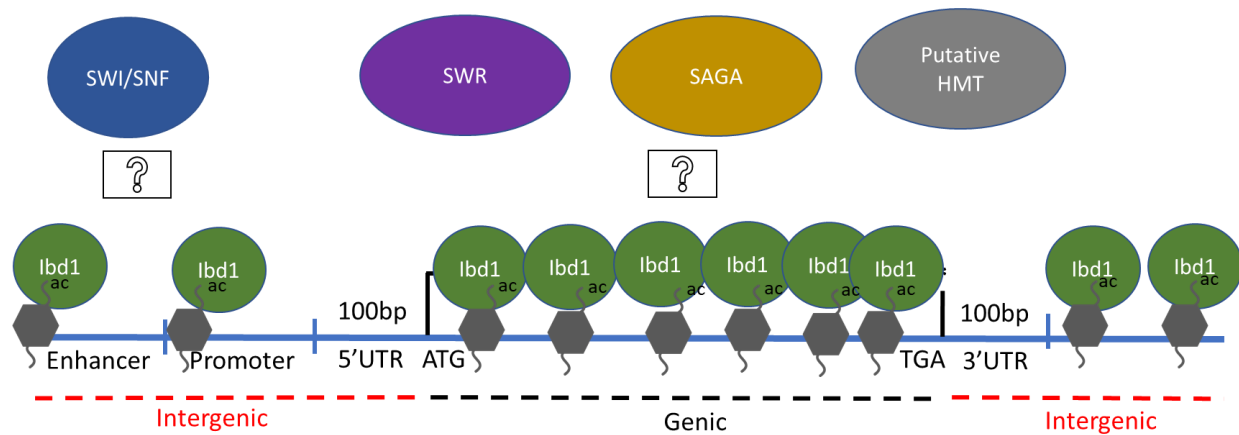


Figure 35. Ibd1 localizes to many intergenic regions but accumulates to genic regions.

The abundance distribution of Ibd1 is related to its localization within genic or intergenic regions. Ibd1 recruits a specific chromatin remodeling complex to a specific region to promote transcription. It is hypothesized that due to the high abundance of Ibd1 in genic regions, there is more remodeling machinery recruited there. This could be due to a higher abundance of Acetyllysines within the genic region.

Ibd1 specifically localizes to the coding regions of multiple highly transcribed genes during vegetative growth (**Figure 17**) (Saettone et al., 2018). Gene Ontology (GO) analysis of genes controlled by Ibd1 predicted that the majority of Ibd1-FZZ bounded genes are related to housekeeping functions (**Figure 18**), including strong occupancy of the coding regions of genes encoding core histones HHT2/H3 and HHF1/H4 (Appendix 2).

The *Tetrahymena* SWI/SNF complex is defined by the set of proteins that co-purify with Ibd1, Snf5, and Saf5 (**Figure 7**) (Saettone et al., 2018). SWI/SNF includes orthologues of canonical SWI/SNF proteins Swi1, Swi3, Snf5, Snf12, and Snf2/Brg1, the PHD domain-containing Saf5,

the bromodomain containing protein Ibd1, as well as several ciliate and species- specific novel proteins, Saf1-4 (**Table 1**). The findings that Ibd1 localizes to the highly acetylated MAC and is a member of the SWI/SNF complex, and that the recombinant 6xHIS-Ibd1 interacts with Kac of human histone proteins *in vitro* (Saettone et al., 2018) are consistent with Ibd1's role in gene expression. In addition to being a member of SWI/SNF, Ibd1 is also a distinct component of the SWR and SAGA complexes and interacts with an uncharacterized H3K4-specific histone methyl transferase that is similar to human Atrx3 and yeast Set1; possibly a member of a putative COMPASS complex (**Table 1, Figure 7**).

A model for Ibd1 function is that it generally targets highly transcribed genes containing one or more specific histone Kac marks associated with transcription. Subsequently, Ibd1 recruits multiple chromatin related complexes that are all linked to transcription in other experimental systems, including SWI/SNF, SAGA, SWR, and COMPASS. These complexes are recruited to the region to further acetylate nearby chromatin (SAGA), to remodel nucleosomes (SWI/SNF), to deposit Hv1 (SWR), and to di- or tri-methylate histone H3K4 (COMPASS - Atrx3/Set1-like histone methyl transferase) (**Figure 36**).

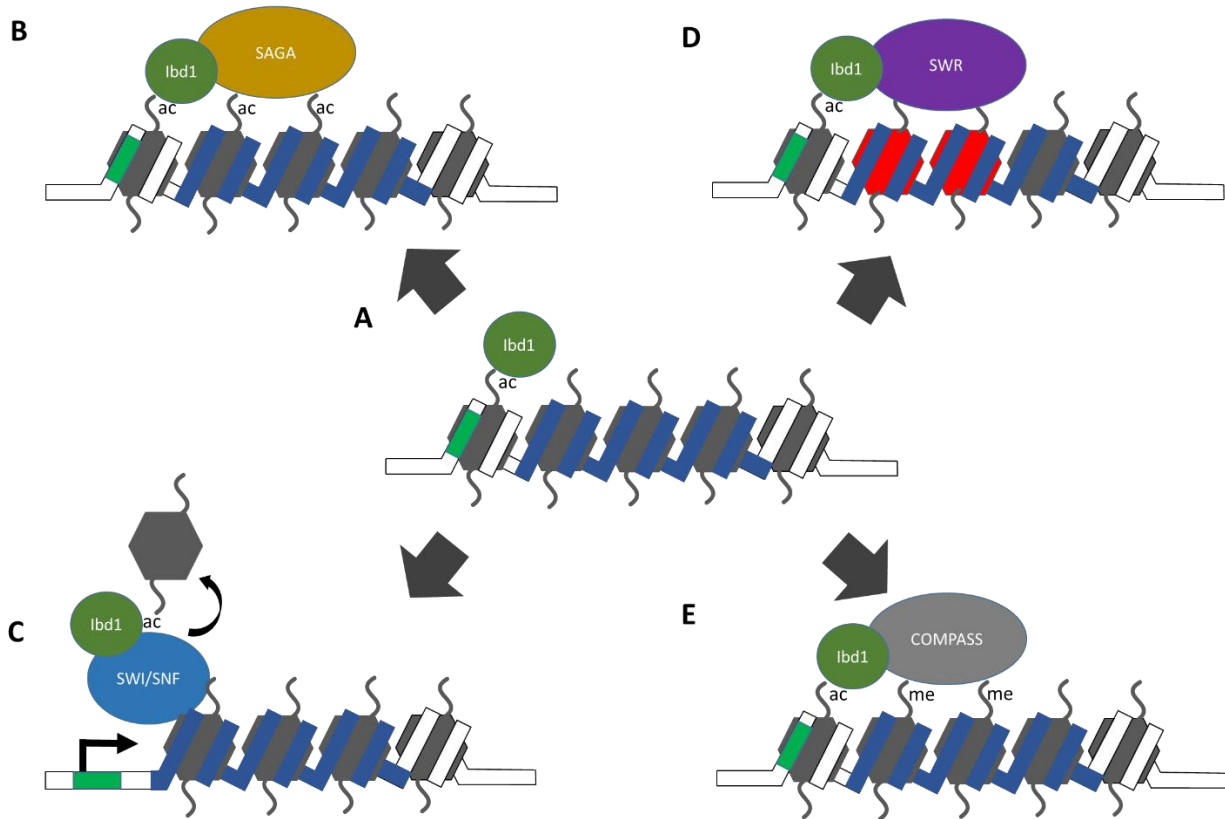


Figure 36: Model for Ibd1 Function recruiting multiple chromatin remodeling complexes. A model for Ibd1 function is that it generally targets highly transcribed genes **A.** containing one or more specific histone Kac marks associated with transcription and recruits multiple chromatin related complexes to the region to either **B.** further acetylate nearby chromatin via SAGA, **C.** to eject nucleosomes via SWI/SNF, to expose transcription factor binding sites (green box), **D.** to deposit Hv1 (red hexagons) via SWR, and **E.** to di- or tri-methylate histone H3K4 via COMPASS (Atrx3/Set1-like histone methyl transferase).

6.3 Model for Med31 function

The canonical function of Mediator is to co-activate transcription for genes transcribed by RNAPII (Tsai et al., 2017). Consistent with a function in controlling transcription,

Tetrahymena's Med31-FZZ localized particularly to coding regions rather than regulatory (intergenic) regions of multiple genes showing a global role in transcription regulation (**Figure 21, Figure 37**).

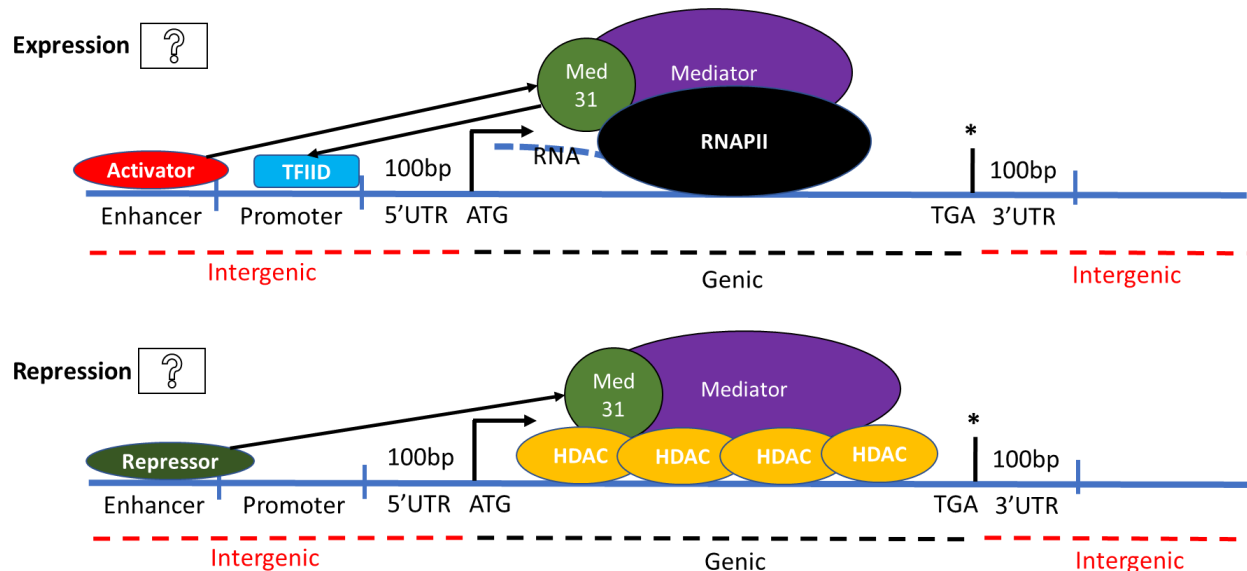


Figure 37. Med31 localizes and accumulates to genic regions.

It is hypothesized that the abundance distribution of Med31 is related to the mechanism of action of Mediator mostly within the coding region. Top panel shows Mediator as a co-activator working with an activator and TF. Bottom panel shows Mediator as a co-repressor working with a repressor and HDAC.

Med31-FZZ particularly localizes to moderate and non-expressed genes during vegetative growth (**Figure 22**). Gene Ontology (GO) analysis of genes controlled by Med31 predicted that the majority of Med31-FZZ bounded genes are of diverse functions, including housekeeping and structural functions (**Figure 23**). Supporting the role of Mediator in *Tetrahymena* programmed DNA rearrangements is the observation that Med31 targets many critical developmental genes including *PDD1*, *PDD2* and *TWI1* to control their transcription (**Figure 24** and **Figure 27**). As an example, reduced *MED31* levels results in ectopic expression of *PDD1* during growth (**Figure 26** and **Figure 29**) and aberrant nuclear structures (**Figure 33**). It remains unclear

however, whether the altered expression of *PDD1* and *PDD2* (**Figure 26**) is directly caused by the loss of a functional Mediator itself or whether it is a consequence of expression changes in other transcription regulators upon *MED31* KD.

The purification and characterization of Mediator complex was reported, via its highly conserved Med31 subunit as bait in AP-MS experiments (Garg et al., 2019). The MS data shows that Med31-FZZ co-purifies with several *Tetrahymena* orthologues of human and yeast Mediator subunits (Bourbon, 2008; Garg et al., 2019), the 13 *Tetrahymena*-specific hypothetical proteins (MITS 1-13) that do not share similarity to any known protein and with Aap4 and TFIIS (**Figure 12, Table 2**). Aap4 is a PHD containing protein that co-purifies with the *Tetrahymena*'s SAGA complex (**Table 1**) and TFIIS is a transcription factor involved in elongation (Garg et al., 2019). In addition, the *Tetrahymena*'s Mediator complex lacks the suppression module CDK8 which could be an evolutionary consequence of Rpb1 missing the CTD heptad repeats (**Figure 13**).

The observation that Med31 localizes to MAC and co-purifies with TFIIS (**Figure 12, Table 2, Figure 14**) suggests a role for Mediator in ncRNA production in *Tetrahymena* during vegetative growth (Garg et al., 2019). Recent evidence indicates that Mediator functions in ncRNAs expression in yeast, plants and mammals (Y. J. Kim et al., 2011; Krebs et al., 2010; Thorsen, Hansen, Venturi, Holmberg, & Thon, 2012).

Med31 localizes to the crescent-shaped meiotic prophase MIC during conjugation (**Figure 14**, 2 hours post mixing) (Garg et al., 2019). This is a stage when the MIC undergoes bursts of ncRNA transcription required for programmed DNA rearrangements (Mochizuki et al., 2002; Mochizuki

& Gorovsky, 2004a, 2004b). Furthermore, in *Paramecium tetraurelia*, a closely related ciliate to *Tetrahymena*, TFIIS has been implicated in scan-RNA production during development (Maliszewska-Olejniczak et al., 2015). This suggests that the interaction between Med31 and TFIIS present during vegetative growth could be prolonged until conjugation. In addition, the localization of Med31-FZZ to the developing MAC during conjugation (**Figure 14**, 6.5 hours post mixing), suggests that Mediator plays a role in programmed DNA rearrangements in *Tetrahymena*, a stage when the developing MAC undergoes irreversible genome silencing (Y. Liu et al., 2007; Mochizuki et al., 2002; Taverna et al., 2002).

A model for Med31 function is that Mediator in *Tetrahymena* might function in transcriptional repression as well as activation, consistent with its known roles in yeast and humans (Poss, Ebmeier, & Taatjes, 2013). *Tetrahymena*'s Mediator is important to maintain low or silenced rates of transcription in genes such as those encoding the proteins involved in programmed DNA degradation, Pdd1, Pdd2 and Twi1. Further work will reveal the extent to which Mediator functions in the co-activation of gene expression as well as deciphering the full extent of Mediator's possible regulatory potential in *Tetrahymena* development.

7 FUTURE WORK

7.1 Testing the model for Ibd1

The first part of the proposed model for Ibd1 states that acetylation of histones leads to Ibd1-dependent recruitment of multiple CR activities that remodel chromatin to facilitate high levels

of transcription (**Figure 36**) (Saettone et al., 2018). However, the chromatin-binding specificity of Ibd1 is unknown. Preliminary data using a human array, containing histone PTM, shows that the full length of recombinant 6xHIS-Ibd1 recognizes acetylated H3K9 and H3K14, acetylated H2AK9 and H2AK13 and tri-acetylated H4K5, H4K8 and H4K12 (Saettone et al., 2018). However, it is unknown if the BRD of Ibd1 is binding the PTMs found in these peptides. It is important to note that Ibd1 has a BRD and a region of unknown function (Grey region in **Figure 36**). Thus, it is possible that this region of unknown function within Ibd1 recognizes peptide motifs in the peptides containing these PTMs. A similar case was shown in (Wai et al., 2018) where the ET domain of human BRD3 recognizes proteins via a short peptide motif. To test which acetyllysines are being recognized by Ibd1, a MS approach utilizing recombinant bromodomains can be used. The MAC core histones will be isolated as specified in (Taverna et al., 2007). Briefly, cells are disrupted, and differential centrifugation will be used to separate the MAC and MIC. Total histones will be purified using H₂SO₄ extraction of isolated nuclei. Quality of histones will be assessed using SDS-PAGE. Next, recombinant 6xHis-BRD proteins will be expressed and purified following the protocol developed in (Saettone et al., 2018). Recombinant proteins will be incubated with acid-extracted MAC total histones as specified in (Gradolatto et al., 2009). The recombinant proteins will be affinity purified using Ni-agarose in condition that favor co-purifications of bound histones. These purifications will be resolved on 4%-20% Tris-Glycine SDS-PAGE and visualized by colloidal Coomassie staining. The bands corresponding to histones will be treated in-gel with d6-acetic anhydride to convert free lysines to D3-acetyl lysines (“heavy” form) and then digested with trypsin. The amount of acetylation catalyzed by a HAT enzyme on a specific lysine is calculated using the ratio of “light” (enzymatic) and “heavy” (chemical) acetylation signal detected at that lysine position. Resulting

tryptic histone peptides will be subjected to quantitative tandem MS analysis to assess predominant PTMs as described (Cieniewicz et al., 2014). An aliquot of input histones will be similarly treated and assessed and compared with the immunoprecipitation results to verify enrichment of the identified histone PTM. To confirm the in vivo relevance of any identified PTM recognized by BRD proteins, peptide pull down experiments will be performed as in (Taverna et al., 2006). We will incubate the *Tetrahymena* whole cell extracts and the biotinylated peptide representing the specific PTM identified above. Streptavidin agarose will then be used to purify the peptide and any additional proteins will be identified using MS. The set of co-purifying proteins will be compared with our AP-MS results. For example, it is predicted that the peptide representing the histone PTM bound by Ibd1 should co-purify with a set of Ibd1-interacting proteins (SWI/SNF, SWR-C, SAGA and COMPASS).

The second part of the hypothesis states that Ibd1 is particularly important in maintaining high rates of transcription on highly expressed genes. This includes genes encoding core histones or ribosomal proteins by recruitment of SWI/SNF, SAGA, SWR, and the HMT to the ORFs and regulatory regions identified in our ChIP seq analysis. ChIP-Seq of the Ibd1-containing complex members Saf5-FZZ, Swc4-FZZ, Ada2-FZZ and Atrx3/Set1-like-FZZ will be required to test the validity of this hypothesis. If the model presented in **Figure 36** is correct, the individual complexes should be found in highly expressed genes. Note that it is possible that these proteins could only localize to subset of specific genes.

7.2 Characterization of BRD containing proteins in *Tetrahymena thermophila*

Utilizing an approach that combines ChIP- Seq and AP- MS analyses, Ibd1 was correlated with transcription. Ibd1 discovery prompted the identification of the full repertoire of the 14 *Tetrahymena* BRD proteins (**Figure 6**) (Saettone et al., 2018). To functionally characterize the other 13 BRD proteins in *Tetrahymena*, the developed ChIP-Seq and MS approach for Ibd1 can be applied. The characterization will start by generating stable *Tetrahymena* strains expressing the BRD-FZZ protein, as it was performed for Ibd1 (Saettone et al., 2018). RACS will be an important tool to understand what genes are being controlled by these BRD proteins.

7.3 Characterization of SAGA activity in *Tetrahymena thermophila*

Another future research can involve a Mass Spectrometry strategy to target the HAT1 and SAGA HAT complexes. During this research, the *Tetrahymena* version of SAGA containing the p55/Gcn5 histone acetyl-transferase (HAT) and Ada2 was identified (Saettone et al., 2018). A Mass spectrometry strategy similar to the one developed in (Cieniewicz et al., 2014) will be used in order to identify the site of action of *Tetrahymena* SAGA on histone H3. In (Saettone et al., 2018) *Tetrahymena* SAGA was co-purified via separate Ibd1 and Ada2 AP-MS (**Table 1**). The AP will be performed, and functional complexes will be recovered using native elution from M2 agarose i.e. TEV protease or 3xFLAG peptide. HAT assays will be performed in vitro using recombinant *Tetrahymena* histone H3 with co-factor acetyl CoA (Cieniewicz et al., 2014). Products from the HAT assay will be resolved and visualized on acid-urea gels that separate by both charge and molecular weight (Cieniewicz et al., 2014). The ablation of charge that occurs when a single lysine is acetylated (from positive to neutral) results in a specific decrease of migration rate on acid-urea gels. The distinct acetyl forms of H3 will be excised to be

treated with d6-acetic anhydride and be analyzed by mass spectrometry. Identified sites will be validated using Western blotting of the products of the HAT assay blotted with commercially available antibodies. This data will be used in conjunction of ChIP-Seq experiments to functionally characterize the role of *Tetrahymena* SAGA in gene expression.

7.4 Characterization of DNA motif associated with Ibd1

It will also be interesting to determine whether the regulatory regions enriched by Ibd1 contain conserved DNA sequences. This may indicate whether specific DNA-binding transcription factors recruit Ibd1-containing protein complexes to regulatory regions. To test this, the intergenic regions could be assessed using motif finding software.

7.5 Genome wide analysis of Mediator-associated transcribed genes to validate the Med31 model

ChIP-Seq of Med31 in growing cells shows it to localize to genes that are expressed during growth as well as development-specific genes, like PDD1 (Garg et al., 2019). In addition, KD of MED31 resulted in ectopic expression of PDD1, suggesting that one of Mediator functions is to repress expression of PDD1 during growth (Garg et al., 2019). Based on these findings, it was hypothesized that *Tetrahymena*'s Mediator is involved in suppression and activation of genes throughout *Tetrahymena*'s life-cycle. To test this hypothesis, a genome wide analysis of transcription by comparing mRNA profiles (RNA-Seq) from WT and MED31 KD will be performed. This data will be analyzed to define the full set of genes that are activated and

repressed by Mediator. RACS capabilities can be used to analyze the NGS data and the genes will be ranked by fold increase and decrease relative to the WT. The objective is to determine the mechanism by identifying all Mediator regulated genes during growth. These experiments will show the ectopic presence of mRNA from development-specific genes that are not normally transcribed during growth, such as PDD1. The RNA-Seq data will be compared with the data obtained with Med31-FZZ ChIP-Seq (Garg et al., 2019) to identify a list of genes for further analysis. With these two data sets, a comprehensive list of genes that are being directly affected by Mediator will be observed. It is expected to observe a strong increase/decrease in expression of specific genes in the MED31 KD when compared to the WT as well as Med31-FZZ binding for the same genes. These genes will be confirmed for expression levels by qRT-PCR analysis in the WT and the MED31 KD strains.

7.6 Tetrahymena's Mediator proximity biotinylation mapping

It was hypothesized that Mediator will be proximal to conserved and unidentified non-Mediator proteins including general transcription factors (TBP, TFIIB) or enhancer binding transcription factors to promote transcription. It is also hypothesized that Mediator will be proximal to proteins with repressive functions as it represses developmental genes during growth (Garg et al., 2019). To find these proteins, a proximity-dependent biotinylation mapping method known as BioID can be used (Gingras, Abe, & Raught, 2019). BioID provides information regarding a specific protein's physical or proximal interactions being complementary to other AP-MS approaches. BioID is particularly useful for transient interactions or to capture proteins that are difficult to solubilize (Lambert, Tucholska, Go, Knight, & Gingras, 2015). For this, a codon

optimized BirA* sequence (Urbanska et al., 2018) will be added to the Med31 C-terminus to obtain a Med31- BirA* cell line. A GFP-BirA* strain will be engineered to be used as a negative control cell line. The transformation process will use the same *Tetrahymena* approach as in (Garg et al., 2019; Saettone et al., 2018). The BirA* C-terminus of Med31 C-terminus will add biotin to the proximal proteins. The Med31-BirA* and GFP-BirA* cell lines will be grown with biotin and in duplicate. Harsh lysis condition will be used, allowing for the solubilization of most cellular proteins. It is important noting that these conditions do not interfere with streptavidin-biotin binding (Urbanska et al., 2018). Mass spectrometry and data curation by SAINTexpress (Teo et al., 2014) will be performed. SAINTexpress will find proteins that are significantly enriched in Med31-BirA* samples compared to controls. It is expected that the Med31-BirA* will add biotin to the proteins that interact with Med31-FZZ (Garg et al., 2019), as well as to transcription factors and suppression proteins.

7.7 Potential common link between Ibd1 and Med31

An interesting finding is that the ChIP-Seq analysis also revealed a set of 20 genes and 132 regulatory (intergenic) regions that are occupied by Ibd1 and Med31 (Appendix 4). Even though there is no evidence of Ibd1 interacting with Med31, it is possible that Ibd1 recruits SAGA containing Aap4 to highly expressed genes. Moreover, Aap4 recruits Mediator to promote transcription.

7.8 Characterization of PHD domain containing proteins in *Tetrahymena*

An intriguing finding is that the transcription related complexes, SWI/SNF, SAGA and Mediator, have at least one PHD domain containing protein (Garg et al., 2019; Saettone et al., 2018). One function attributed to PHD domains is recognizing methylated lysines in proteins such as histones; like how the PHD domain of human ING2 recognizes H3K4me3 to suppress gene expression via its deacetylase activity (Shi et al., 2006). On the other hand, the PHD domain of human PHD3 recognizes H3K4me3 to promote transcription by recruiting the MLL complex (Chang et al., 2010). The *Tetrahymena* transcriptionally active MAC contains hyper-acetylated histone H3 that is also di- or tri-methylated on H3K4 (Taverna et al., 2007).

In addition to the BRD-containing Ibd1, *Tetrahymena* SWI/SNF also contains a PHD domain-containing protein, Saf5. The SWI/SNF complex in *Tetrahymena* contains two proteins that potentially recognize PTM on histones; Saf5 likely recognizes methyl lysine (and possibly acetyl-lysine (Lange et al., 2008)), and Ibd1 that recognizes Kac.

Med31-FZZ co-purifies with, Aap4, one of the four PHD domain containing proteins (Aap1-4) found in the SAGA complex (Garg et al., 2019; Saettone et al., 2018). There are two hypothesized functions of *Tetrahymena*'s Mediator complex including a PHD domain containing protein (Aap4). Firstly, Aap4 recognizes PTMs to promote transcription, as human PHD3 (Chang et al., 2010) and similar to the predicted function of Saf5 in the SWI/SNF complex (Saettone et al., 2018). Alternatively, Aap4 could be acting as a substitute of the missing suppression module CDK8. In that Aap4 could be recognizing specific PTMs to block Mediator's function and suppressing gene expression of developmental genes during growth.

The function of these PHD containing proteins is unclear. It is hypothesized that their function is to recognize specific PTMs in histones to increase or repress transcription rates. To test this hypothesis the functional characterization of these PHD containing proteins in *Tetrahymena* using the developed ChIP-Seq and MS approach for Ibd1 and Med31 will be applied. This characterization will start by generating stable *Tetrahymena* strains expressing the PHD-FZZ protein as it was performed for Ibd1 and Med31 (Garg et al., 2019; Saettone et al., 2018). RACS will be an important tool to understand what genes are being controlled by these PHD proteins. To understand these PHD proteins and their chromatin-binding specificity, a similar strategy as suggested for BRD proteins can be employed.

8 CONCLUSIONS

In multi-cellular eukaryotes, the precise function of how chromatin remodeling complexes work is poorly understood. Alteration or loss of factors involved in these complexes through mutation has been shown to be associated with cancer. A protist model was utilized, the Alveolate *Tetrahymena thermophila*, which segregates transcriptionally active and silent chromatin into two distinct nuclei; the macronucleus (MAC) and micronucleus (MIC) respectively, contained in the same cell. Proteomic and genomic analyses has advanced knowledge of nuclear structure/function in many systems. Studies in *Tetrahymena* have been detailed and comprehensive and, due mainly to the unique nuclear dimorphism of ciliates, have in many cases informed studies in other eukaryotes. Proteomic and genomic studies will continue to yield new information concerning nuclear functions in *Tetrahymena*, particularly in several relatively unexplored areas like global transcription of ncRNAs.

A ChIP-Seq pipeline (RACS) was developed to demonstrate that Ibd1 and Med31 are enriched in ORFs of highly and low expressed genes, respectively. Through the development of RACS, a computational pipeline to analyze ChIP-Seq data for genomes that are contig-based and/or have poor annotations, the knowledge of chromatin remodeling complexes and transcription co-activators (mediator) in protists was advanced. RACS is complementary to other tools, such as MACS2, as it can help to discriminate complex regions improving the overall analysis. RACS offers an alternative tool with a different approach focused on a simple, modular and open approach. RACS offers a versatile, agile and modular pipeline that cover many of the steps needed in the process of analyzing ChIP-Seq data.

RACS allowed for the development of two functional models for two transcription related proteins, Ibd1 and Med31. The first model was presented where a single protein, Ibd1, localizes to highly expressed genes coordinating the action of multiple chromatin remodeling complexes to achieve high levels of transcription. The second model presents a Mediator containing Med31 that regulates gene expression, and disruption of Med31 expression leads to ectopic expression of developmental genes such as PDD1. This research contributes to the current understanding of transcription in ciliates, and more broadly the function and diversity of chromatin remodeling complexes in eukaryotes. This research builds knowledge about transcription function in a model system known for its abundant histone acetylation. Understanding how transcription related proteins function in the *Tetrahymena* model will be of general interest to the chromatin research community.

APPENDICES

A.1: Chromatin Preparation for Chromatin Immunoprecipitation (ChIP) in Tetrahymena thermophila

Day 1

1. Inoculate 500uL of bench culture in 10mL of SPP overnight.

Day 2

1. Inoculate 3 mL of the overnight culture in 50mL of SPP in 1L flask
2. Grow overnight to mid-log phase (2×10^5 to 5×10^5 cells/mL)

Day 3 - Harvesting Cells

1. Turn big centrifuge ON and set to 4°C
2. In a 250mL flask, add 2.5mL of formaldehyde (supplied at 37%) using a disposable pipette (inside the fume-hood) and using a 100mL graduated cylinder add 90mL of *Tetrahymena* culture to the flask.
3. Rotate on a slow rotator for 30mins
4. Add 13mL of 2.5M glycine solution (Make fresh) and incubate rotating for 5 mins
5. Divide 100mL in to 5x 20mL in 50mL Falcon tubes and spin down 5min, 4000rpm, 4 degrees
6. Collect and discard supernatant as formaldehyde waste in fume hood. Dab on a paper towel to get rid of the rest of the formaldehyde (solutions go in ice after this point).
7. Re-suspend in 10mL ice-cold 10mM Tris pH 7.4 and spin as above (x2)
8. Wash 1x with lysis buffer (w/o protease inhibitors)
9. Freeze pellet on dry ice and store -80 freezer

Day 4 - Preparation of Chromatin

1. Turn micro-centrifuge ON and set to 4°C
2. Prepare the lysis buffer with protease inhibitors
3. Thaw cell pellet on ice. Transfer pellet in 750uL of lysis buffer in an Eppendorf tube. **Let it stand for 10mins**
4. To obtain DNA fragments between 300 and 600 base pairs (**Figure 38**), sonicate TIME=0:01:15 hr/m/s; MAX.TEMP=TEMP OFF; AMPLITUDE=30%; PULSE ON=0.3 SEC; PULSE OFF= 6.0 SEC; PULSE TEMP=TEMP OFF; PROBE TEMP=TEMP OFF (Fisher Scientific Sonic Dismembrator Model 500)
5. Leave samples as cold as possible on wet ice for 1min-2min between bursts (to prevent the heat build-up in the sonication sample)
6. Spin the samples 30min at max speed in 4 degrees
7. Transfer the supernatant to a new micro-centrifuge tube on ice (pour it, only pipette if the pellet is moving)

8. Save 2 aliquots (**A** and **B**) of 50uL of chromatin each and save the pellets and store at -80 (one aliquot for input WB for next step A and the other for input DNA for purification for next step B).

Day 5

A soluble sample – SDS-PAGE analysis of soluble chromatin

1. Add **50uL** of fresh 2XSDS loading dye buffer to the sample A and **750uL** for the correspondent pellet.
2. Boil for 30mins (To ensure complete de-crosslinking of proteins) and run 10uL on SDS-PAGE to observe the protein of interest.
3. Load 10uL per well. **Save the rest -80 °C for use in the last step as INPUT for WB. Save the pellet and add 650uL of 2XSDS loading dye buffer for use in the last step for WB.**

B soluble sample – Purification of Input DNA

1. Add 200uL of 2x proteinase K buffer and 250uL of sterile ddH₂O to input DNA sample
2. Add 1uL of RNase (Thermo Scientific 10mg/mL) 1hr 37°C
3. Add 10uL of Proteinase K (Thermo Scientific : 20mg/mL) 42°C for 2hrs and 65°C for 8hrs
4. Perform Qiagen purification of DNA
 - Add 2.5mL of PB buffer in 15mL falcon tube. Load the same column 4 times
 - The rest of the protocol remains the same; however, **use 30uL of ddH₂O for elution** instead EB.
5. Run 5uL of the 30uL on 1% agarose gel with a 100 base pairs size standard (**Figure 38**) and **save the remaining 25uL in -20 °C for desired ChIP application.**

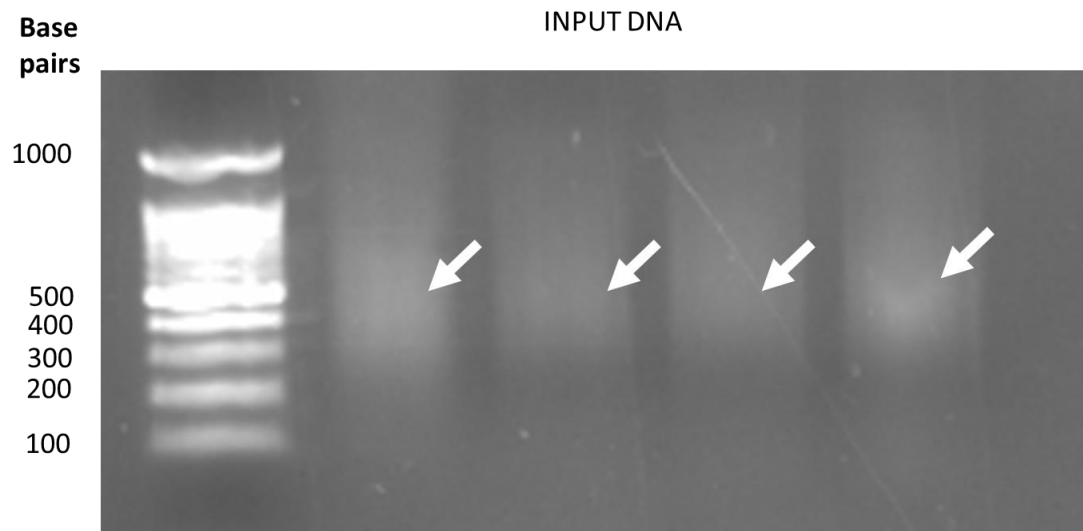


Figure 38: After Sonication, DNA fragments between 300 and 600 base pairs.

Day 6

1. Turn the 4 °C micro-centrifuge on.
2. Thaw Chromatin
3. Measure the remaining volume with a pipette (Should be around 500uL)
4. Dilute this remaining volume 1:10 in dilution buffer in a 15mL falcon tube (around 5.5mL)
5. Rotate for 30 minutes at 4 °C.
6. Preparation of M2 Agarose beads
 - Packed bead volume (PBV) needed for ChIP = 25uL (50uL from the bottle)
 1. Pipette the total amount of PBV required for samples + for 1 extra sample in to an E. tube using cut blue/yellow tip
 2. Spin 2000rpm 4°C 2min, aspirate the supernatant and add 750uL of ice cold 1x dilution buffer
 3. Mix by inversion
 4. Repeat steps 2 and 3, 3 times
 5. After the last aspiration, re-suspend equal volume of ice-cold dilution buffer
7. Aliquot 25uL of PBV or 50uL 1:1 slurry in to a fresh 15mL falcon tube.
8. Ensure that proper amount of beads is aliquoted in each falcon tube by spinning in the centrifuge
9. Pour the diluted soluble chromatin into the tube w/beads
10. Incubate 4°C w/ rotation 4hrs – o/n

Day 7

1. Turn the heat block on and set to 65°C
2. Turn the micro-centrifuge and big centrifuge on and set to 4 °C.
3. 1st spin – 4000rpm 4°C 5min in big centrifuge [15mL Falcon tubes]. **Save 100uL of the supernatant and add 100uL of 2XSDS loading dye buffer for use in the last step for WB as unbound**
4. After the first wash, use a cut blue tip + 1X FA buffer to transfer the beads to a fresh E. tube

All washes performed at 4 degrees using ice-cold wash buffers – washes 5 min with rotation, spins 2000rpm, 4 degrees, 2 min

- i. Wash 1X with 1x FA
 - ii. Wash 1X with FA 500mM NaCl
 - iii. Wash 1X with LiCl buffer
 - iv. Wash 1X TE
5. Elute protein-DNA complexes using 250uL of 2x ProK buffer by heating 65 degrees 20min (flick beads to mix every 5 min)
 6. Spin 5000rpm RT 1 min and transfer supernatant to a 0.5mL PCR tube. **Take 25uL and add 25uL of fresh SDS loading dye. Keep it at -80°C until running. Load the 30uL of sample for WB.**
 7. Wash beads with 225uL sterile ddH₂O and transfer to the same 0.5mL PCR tube. **Add 250uL of fresh 2XSDS loading dye buffer to the beads and keep them in -80°C until running the WB**

8. Boil all WB samples for 30minutes and load the 30uL of each sample in the SDS gel as follows (Figure 39).

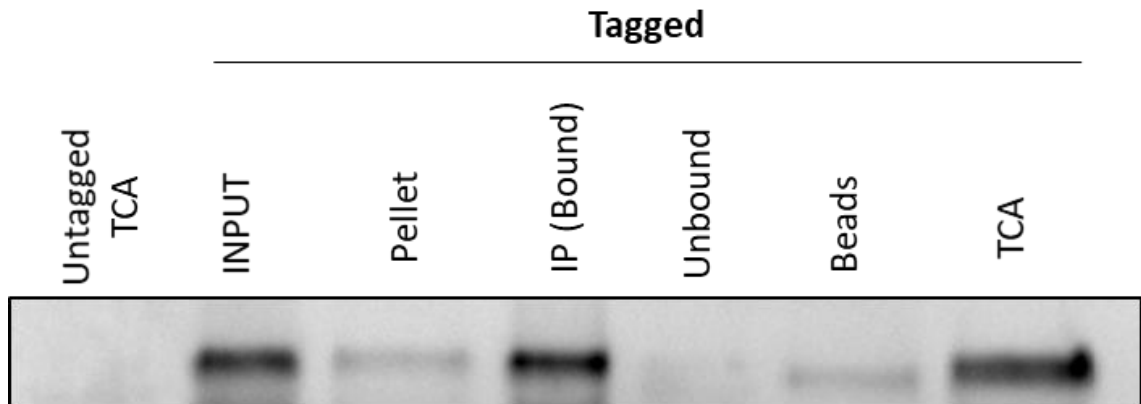


Figure 39: Recommended WB layout for ChIP experiment.

9. Treat the sample in the 0.5mL PCR tube with 1uL of RNase (Fermentas) for 1hr at 37°C
10. Add 10uL of ProK (Fermentas)
11. Incubate 42°C for 2 hrs and 65°C for 8hrs in PCR machine

Day 8

1. Use Qiagen PCR purification kit to clean up DNA solution
 - Add 2.5mL of PB buffer in 15mL falcon tube. Load the same column 4 times
 - The rest of the protocol remains the same; however, use 30uL of ddH₂O for elution instead EB.
2. Save the sample in -20°C for desired ChIP application.

REAGENTS

FLAG Beads - *Beads are generally suspended in ethanol.*

Buffers

2.5M Glycine (200mL)

Glycine - 45.04g

Tris (1M pH8) - 4mL

ddH₂O to 200mL

heat with a stir bar to mix. Make fresh every time

Lysis Buffer

1M Tris pH 8.0 - 2.5mL (50mM)

0.5M EDTA - 0.5mL (5mM)

10% SDS - 5mL (1%)

Complete up to 50mL using sterile ddH₂O

(+ Protease inhibitors - 1 tablet+ 250uL of 100mM PMSF per 50mL of lysis buffer)

2X Proteinase K Buffer

1M Tris pH8 -1mL

0.5M EDTA - 1mL

10% SDS - 5mL

ddH₂O - 43mL

5M NaCl

146.1g in 500mL of ddH₂O

Dilution Buffer

1M Tris pH 8 - 2.5mL

5M NaCl - 1.5mL

0.5M EDTA - 500uL

100% Triton X - 500uL

10% NaDOX - 500uL

ddH₂O to 50mL

10% Deoxycholate

Deoxycholate - 2g

Sterile ddH₂O to 20mL in a 50mL falcon

Vortex, store RT for up to a few weeks

1M HEPES

59.5g HEPES

ddH₂O- 200mL

Adjust pH to 7.5 using 10N KOH

Adjust to 250mL with ddH₂O and *filter sterilize*

2X FA Buffer

1M HEPES pH7.5 - 50mL
5M NaCl - 30mL
0.5M EDTA pH 8.0 - 2mL
Triton X-100 (100%) - 10mL
NaDeoxycholate (10%) - 10mL
ddH₂O to 500mL, *filter sterilize and store at 4degrees*

1X FA

2X FA buffer 25mL
10% SDS - 0.5mL
ddH₂O to 50mL (24.5 mL)

1X FA (500mM NaCl)

2X FA buffer 25mL
10% SDS - 0.5mL
5M NaCl- 3.5mL
ddH₂O - 21mL

LiCl Buffer

1M Tris pH 8 - 0.5mL
4M LiCl - 3.125mL
0.5M EDTA - 0.1mL
10% NP40 - 2.5mL
10% Nadeoxycholate 2.5mL
ddH₂O – 41.3mL

10N KOH - 10M (the solution will headup really quickly, therefore use a lower mixing scale) -
56.11g in 100mL to ddH₂O

4M LiCl

LiCl 8.48g
up to 50mL ddH₂O

1X TE (50mL)

1M Tris-HCl pH 8 – 500uL
0.5M EDTA pH 8 – 100uL
up to 50mL ddH₂O

2X SDS loading dye

10% SDS - 1150uL
2M Tris pH 6.8 - 37.5uL
BME – 62.5uL
Pinch of bromophenol blue

A.2: *Ibd1*ChIP-Seq post-processing files

RACS processed data for the 354 genic regions occupied by that Ibd1: Refer to Digital Spreadsheet (TET_Ibd1_MAC_Genome_Genic) for complete results.

#	Gene region	Gene_ID	Note	Enrichment Average	Expression Profile
1	scf_8254786:36 8577-368782	TTHERM_00554 480	Note="hypothetical_protein"	14.68	Low to No- Expression
2	scf_8254747:37 2219-372806	TTHERM_00143 660	Note="HTA3_histone_H2A"	8.80	High expression
3	scf_8254431:10 42526-1042867	TTHERM_00056 9259	Note="hypothetical_protein"	8.24	#N/A
4	scf_8254611:80 381-80857	TTHERM_00476 640	Note="60S_ribosomal_protein_L24"	7.09	High expression
5	scf_8254776:27 9565-279945	TTHERM_00161 060	Note="40S_ribosomal_protein_S12"	6.55	High expression
6	scf_8254708:88 146-88793	TTHERM_01053 000	Note="RPL10_60S_ribosomal_protein_L10"	6.50	High expression
7	scf_8254582:10 59131-1059598	TTHERM_00131 110	Note="RPS18_40S_ribosomal_protein_S18_putative"	6.18	High expression
8	scf_8254359:14 6904-148054	TTHERM_00837 950	Note="RPL37_60S_acidic_ribosomal_protein"	6.16	High expression
9	scf_8254369:35 664-36176	TTHERM_00992 760	Note="RPS20_40S_ribosomal_protein_S20"	6.12	High expression
10	scf_8254757:13 9703-141282	TTHERM_01005 320	Note="60S_acidic_ribosomal_protein"	5.98	High expression
11	scf_8254579:16 2130-162603	TTHERM_00666 540	Note="RPL21_60S_ribosomal_protein_L21"	5.72	High expression
12	scf_8254489:64 193-64372	TTHERM_00106 8151	Note="hypothetical_protein"	5.72	#N/A
13	scf_8253815:48 7792-488103	TTHERM_00498 190	Note="HHF1_predicted_protein"	5.66	High expression
14	scf_8254397:90 244-90711	TTHERM_00823 670	Note="RPS19_40S_ribosomal_protein_S19"	5.65	High expression

15	scf_8254582:75 8139-758591	TTHERM_00134 940	Note="RPL25_60S_ribosomal_protein_L23a"	5.57	High expression
16	scf_8254593:38 0888-381577	TTHERM_00463 490	Note="RPS26_hypothetical_protein"	5.56	High expression
17	scf_8254767:71 36-7738	TTHERM_01386 050	Note="RPS5_40S_ribosomal_protein_S5"	5.52	High expression
18	scf_8254697:46 0806-461132	TTHERM_00008 8069	Note="hypothetical_protein"	5.50	#N/A
19	scf_8254764:20 9225-209593	TTHERM_00633 360	Note="HTB1_histone_H2B.1"	5.50	High expression
20	scf_8254814:16 1952-162386	TTHERM_00418 120	Note="transmembrane_protein_putative"	5.50	High expression
21	scf_8254718:96 55-9873	TTHERM_00805 850	Note="hypothetical_protein"	5.50	High expression
22	scf_8254002:13 886-14374	TTHERM_01128 560	Note="RPS10_40S_ribosomal_protein_S10_putative"	5.40	High expression
23	scf_8254610:43 5938-436459	TTHERM_00339 620	Note="RPL40_ubiquitin/ribosomal_fusion_protein_putative"	5.34	High expression
24	scf_8254459:33 139-33663	TTHERM_00052 9532	Note="hypothetical_protein"	4.94	#N/A
25	scf_8254582:67 7901-679496	TTHERM_00136 120	Note="RPL6_60S_ribosomal_protein_L6"	4.88	High expression
26	scf_8254359:13 8222-139472	TTHERM_00083 7939	Note="hypothetical_protein"	4.85	#N/A
27	scf_8254393:27 790-28959	TTHERM_01289 110	Note="RPL15_60S_ribosomal_protein_L15"	4.79	High expression
28	scf_8254600:41 1013-411381	TTHERM_00283 180	Note="HTB2_histone_H2B.1"	4.64	High expression
29	scf_8254731:22 8078-229028	TTHERM_00594 350	Note="SerL_immobilization_antigen"	4.62	High expression
30	scf_8254181:10 2105-102719	TTHERM_00929 590	Note="Gar1/Naf1_RNA-binding_region_protein"	4.58	High expression
31	scf_8254527:50 6843-508503	TTHERM_00487 140	Note="60S_ribosomal_protein_L28"	4.52	High expression
32	scf_8254665:69 5910-696542	TTHERM_00227 770	Note="eukaryotic_translation_initiation_factor_5A"	4.42	High expression

33	scf_8254460:40 4021-405352	TTHERM_00348 510	Note="BTU1_tubulin"	4.33	High expression
34	scf_8254392:78 587-80473	TTHERM_01151 600	Note="40S_ribosomal_protein_S8-2"	4.28	High expression
35	scf_8254365:29 9646-300215	TTHERM_00686 150	Note="RPS31_ubiquitin-40S_ribosomal_protein_S27a"	4.26	High expression
36	scf_8254582:90 8503-909508	TTHERM_00133 510	Note="iron-binding_zinc_finger_CDGSH_type_protein"	4.24	High expression
37	scf_8254594:15 4942-155842	TTHERM_00285 620	Note="NAC_domain_protein"	4.14	High expression
38	scf_8254377:15 6310-156984	TTHERM_00716 040	Note="transmembrane_protein_putative"	4.05	High expression
39	scf_8254187:17 2960-173229	TTHERM_00069 6939	Note="hypothetical_protein"	4.01	#N/A
40	scf_8254584:31 824-32951	TTHERM_00628 610	Note="CTH34_papain_family_cysteine_protease"	3.98	High expression
41	scf_8254417:14 8320-150331	TTHERM_00732 640	Note="RPL16_60S_ribosomal_protein_L13/L16"	3.98	High expression
42	scf_8254697:86 0090-860774	TTHERM_00085 180	Note="RPL32_60S_ribosomal_protein_L32"	3.97	High expression
43	scf_8254565:23 8239-240195	TTHERM_00471 090	Note="RPS7_40S_ribosomal_protein_S7"	3.95	High expression
44	scf_8254665:27 458-28036	TTHERM_00222 240	Note="transmembrane_protein_putative"	3.94	Low to No- Expression
45	scf_8254638:32 6145-326456	TTHERM_00189 170	Note="HHF2_histone_H4"	3.94	High expression
46	scf_8254002:50 651-51127	TTHERM_01129 660	Note="RPL36_60S_ribosomal_protein_L36"	3.92	High expression
47	scf_8254577:32 8591-329022	TTHERM_00660 180	Note="HMGB_high_mobility_group_(HMG)_box_protein"	3.91	High expression
48	scf_8254748:10 5327-106993	TTHERM_00706 300	Note="RPS16_ribosomal_protein_S9"	3.91	High expression
49	scf_8254610:51 4250-516295	TTHERM_00033 9809	Note="60S_ribosomal_protein_L17"	3.90	#N/A
50	scf_8254605:35 6575-357531	TTHERM_00621 600	Note="MCF11_2-oxoglutarate/malate_carrier_protein"	3.88	High expression

51	scf_8254515:18 629-19270	TTHERM_00721 220	Note="hypothetical_protein"	3.81	High expression
52	scf_8254776:29 3996-294451	TTHERM_00016 0999	Note="nucleoside_diphosphate_kinase"	3.80	#N/A
53	scf_8254431:94 4986-945396	TTHERM_00057 0549	Note="histone_H3.4"	3.78	#N/A
54	scf_8254649:37 1205-372167	TTHERM_00268 060	Note="CTH110_papain_family_cysteine_protease"	3.70	High expression
55	scf_8254650:22 4903-226271	TTHERM_00075 670	Note="RPL43_60S_ribosomal_protein_L37a"	3.66	High expression
56	scf_8254731:23 1693-232139	TTHERM_00594 360	Note="immobilization_antigen"	3.65	High expression
57	scf_8254359:15 133-16464	TTHERM_00836 580	Note="BTU2_tubulin"	3.61	High expression
58	scf_8254095:52 286-53860	TTHERM_01113 100	Note="RACK1_receptor_of_activated_kinase_C_1A_component_of_40S_small_ribosomal_subunit"	3.60	High expression
59	scf_8253815:38 9185-391042	TTHERM_00497 880	Note="RPL3_60S_ribosomal_protein_L3_putative"	3.60	High expression
60	scf_8254610:55 552-55907	TTHERM_00334 350	Note="hypothetical_protein"	3.58	Low to No- Expression
61	scf_8254365:32 1683-323225	TTHERM_00686 220	Note="RPS25_S25_ribosomal_protein"	3.56	High expression
62	scf_8254397:17 3159-173703	TTHERM_00082 3969	Note="MFS_transporter"	3.56	#N/A
63	scf_8254763:38 8416-389993	TTHERM_00561 680	Note="RPL35_ribosomal_protein_L35"	3.55	High expression
64	scf_8254591:13 603-14313	TTHERM_00120 9760	Note="transmembrane_protein_putative"	3.51	#N/A
65	scf_8254686:54 6179-546571	TTHERM_00454 080	Note="RPS22_predicted_protein"	3.50	High expression
66	scf_8254649:86 121-87702	TTHERM_00266 390	Note="RPS15_40S_ribosomal_protein_S15"	3.50	High expression
67	scf_8254667:14 64557-1466483	TTHERM_00037 090	Note="RPS29_40S_ribosomal_protein_S29"	3.48	High expression
68	scf_8254610:22 4519-225949	TTHERM_00335 830	Note="hypothetical_protein"	3.46	High expression

69	scf_8254671:56 3426-564478	TTHERM_00427 580	Note="transmembrane_protein_putative"	3.45	Low to No- Expression
70	scf_8254691:28 8589-290072	TTHERM_00032 1589	Note="MFS_transporter"	3.44	#N/A
71	scf_8254665:46 8755-469111	TTHERM_00227 270	Note="RPL22_60S_ribosomal_protein_L22_putative"	3.44	High expression
72	scf_8254428:20 9665-210579	TTHERM_00535 740	Note="phosphate_carrier_protein"	3.43	High expression
73	scf_8254406:78 0108-782830	TTHERM_00248 370	Note="hypothetical_protein"	3.41	Low to No- Expression
74	scf_8254486:18 4479-185223	TTHERM_00055 8148	Note="MFS_transporter"	3.39	#N/A
75	scf_8254691:29 9356-300025	TTHERM_00032 1629	Note="MFS_transporter"	3.36	#N/A
76	scf_8254373:39 9913-400296	TTHERM_00240 460	Note="hypothetical_protein"	3.34	Low to No- Expression
77	scf_8254449:88 490-90227	TTHERM_00773 330	Note="RPS13_40S_ribosomal_protein_S13"	3.33	High expression
78	scf_8254638:32 6796-327206	TTHERM_00189 180	Note="HHT2_predicted_protein"	3.29	High expression
79	scf_8254564:37 7535-379842	TTHERM_00218 790	Note="vacuolar_H+-translocating_inorganic_pyrophosphatase"	3.27	High expression
80	scf_8254814:36 829-37653	TTHERM_00041 8549	Note="U1_zinc_finger_protein"	3.26	#N/A
81	scf_8254751:15 2028-152564	TTHERM_00943 010	Note="RPL31_60S_ribosomal_protein_L31_putative"	3.22	High expression
82	scf_8254607:42 8287-429190	TTHERM_00043 7558	Note="MFS_transporter"	3.22	#N/A
83	scf_8254371:26 8309-269806	TTHERM_00467 660	Note="RPS24_40S_ribosomal_protein_S24"	3.22	High expression
84	scf_8254548:14 4450-144755	TTHERM_00031 6229	Note="transmembrane_protein_putative"	3.21	#N/A
85	scf_8254648:44 142-44732	TTHERM_00877 050	Note="PPI1_peptidyl-prolyl_cis-trans_isomerase_cyclophilin-type_protein"	3.18	High expression
86	scf_8254688:23 526-25022	TTHERM_00297 140	Note="RPL23_60S_ribosomal_protein_L23"	3.18	High expression

87	scf_8254435:19 0829-192136	TTHERM_00655 820	Note="translation_elongation_factor_EF-1alpha"	3.18	High expression
88	scf_8254563:12 8273-128602	TTHERM_00060 0018	Note="transmembrane_protein_putative"	3.18	#N/A
89	scf_8254748:38 68-5738	TTHERM_00705 050	Note="transcription_factor_BTF3a"	3.16	High expression
90	scf_8254659:16 15975-1617509	TTHERM_01142 730	Note="RPL9_60S_ribosomal_protein_L9"	3.15	High expression
91	scf_8254638:62 0048-620952	TTHERM_00019 0979	Note="MFS_transporter"	3.13	#N/A
92	scf_8254718:84 094-85174	TTHERM_00808 000	Note="hypothetical_protein"	3.12	Low to No- Expression
93	scf_8254486:18 5386-185871	TTHERM_00055 8149	Note="MFS_transporter"	3.11	#N/A
94	scf_8254545:32 1047-322226	TTHERM_00149 260	Note="hypothetical_protein"	3.06	High expression
95	scf_8254716:73 803-75077	TTHERM_00095 540	Note="CBF5_tRNA_pseudouridine_synthase"	3.05	High expression
96	scf_8254548:23 2430-232831	TTHERM_00316 500	Note="HTA2_histone_H2B.1"	3.05	High expression
97	scf_8254579:16 5867-166595	TTHERM_00666 560	Note="PRE6_proteasome_subunit_alpha_type_protein"	3.03	High expression
98	scf_8254649:52 9884-530502	TTHERM_00270 350	Note="HVA22/TB2/DP1_family_protein"	3.02	High expression
99	scf_8254284:18 7805-189438	TTHERM_00765 300	Note="RPS14_40S_ribosomal_protein_S14"	3.02	High expression
100	scf_8254590:23 1028-231333	TTHERM_00251 150	Note="transmembrane_protein_putative"	3.01	Moderate expression
101	scf_8254667:15 77844-1579027	TTHERM_00037 300	Note="MFS_transporter"	3.00	Low to No- Expression
102	scf_8254814:10 1163-103322	TTHERM_00041 8279	Note="hypothetical_protein"	3.00	#N/A
103	scf_8254815:24 3646-244225	TTHERM_00198 207	Note="hypothetical_protein"	2.98	Low to No- Expression
104	scf_8254649:43 0285-430707	TTHERM_00026 8189	Note="hypothetical_protein"	2.97	#N/A

105	scf_8253811:53 282-54343	TTHERM_00052 5180	Note="MFS_transporter"	2.96	#N/A
106	scf_8254420:59 369-59588	TTHERM_00069 4439	Note="hypothetical_protein"	2.95	#N/A
107	scf_8254644:16 978-17888	TTHERM_00064 8561	Note="MFS_transporter"	2.95	#N/A
108	scf_8254584:92 75-9448	TTHERM_00628 700	Note="hypothetical_protein"	2.94	Low to No- Expression
109	scf_8254470:66 5725-667098	TTHERM_00433 940	Note="LBP/BPI/CETP_family_carboxy-terminal_domain_protein"	2.92	Moderate expression
110	scf_8254582:29 4970-295440	TTHERM_00138 140	Note="hypothetical_protein"	2.90	Low to No- Expression
111	scf_8254622:64 851-65462	TTHERM_01194 830	Note="BBC23_antioxidant_AhpC/TSA_family_protein"	2.90	High expression
112	scf_8254691:22 671-23382	TTHERM_00031 9933	Note="hypothetical_protein"	2.88	#N/A
113	scf_8254652:49 7798-500130	TTHERM_00382 330	Note="hypothetical_protein"	2.88	High expression
114	scf_8254401:51 283-52619	TTHERM_00108 0449	Note="MFS_transporter"	2.87	#N/A
115	scf_8254449:91 451-93997	TTHERM_00773 340	Note="RPL8_60S_ribosomal_L7Ae-like_protein"	2.87	High expression
116	scf_8254587:51 698-53673	TTHERM_00384 800	Note="Nopp52_RNA_recognition_motif_1_in_heterogeneous_nuclear_ribonucleoprotein_hnRNP_protein"	2.87	High expression
117	scf_8254667:11 32636-1133043	TTHERM_00003 5237	Note="transmembrane_protein_putative"	2.86	#N/A
118	scf_8254479:10 7561-107728	TTHERM_00062 2779	Note="hypothetical_protein"	2.82	#N/A
119	scf_8254652:38 8818-389404	TTHERM_00379 060	Note="transmembrane_protein_putative"	2.82	Moderate expression
120	scf_8254010:66 2427-663043	TTHERM_00332 160	Note="thioredoxin_domain_protein"	2.81	High expression
121	scf_8254332:12 626-12985	TTHERM_01299 660	Note="SCP-2_sterol_transfer_family_protein"	2.81	High expression
122	scf_8254402:39 287-39679	TTHERM_00065 7292	Note="hypothetical_protein"	2.80	#N/A

123	scf_8254373:74 4709-746133	TTHERM_00242 320	Note="MFS_transporter"	2.79	Low to No- Expression
124	scf_8254385:99 741-100852	TTHERM_00532 050	Note="MFS_transporter"	2.79	Low to No- Expression
125	scf_8254688:60 2205-602543	TTHERM_00029 4739	Note="hypothetical_protein"	2.78	#N/A
126	scf_8254653:11 8093-119118	TTHERM_00641 150	Note="CTH40_papain_family_cysteine_protease"	2.78	High expression
127	scf_8254527:35 0241-350693	TTHERM_00486 690	Note="hypothetical_protein"	2.77	High expression
128	scf_8254607:42 9426-429670	TTHERM_00043 7559	Note="hypothetical_protein"	2.77	#N/A
129	scf_8254636:43 3266-433619	TTHERM_00491 100	Note="cell_surface_immobilization_antigen_SerH6_putative"	2.77	Low to No- Expression
130	scf_8254752:11 51717-1153069	TTHERM_00121 000	Note="transmembrane_protein_putative"	2.75	High expression
131	scf_8254597:12 5917-127159	TTHERM_00034 3608	Note="MFS_transporter"	2.74	#N/A
132	scf_8253957:60 8-1155	TTHERM_00106 9309	Note="transmembrane_protein_putative"	2.74	#N/A
133	scf_8254047:16 5496-165912	TTHERM_00790 790	Note="HTA1_histone_H2A"	2.71	High expression
134	scf_8254403:43 6388-438603	TTHERM_00475 140	Note="PDI1_protein_disulfide-isomerase"	2.70	High expression
135	scf_8254504:27 58-4658	TTHERM_01109 770	Note="RPS11_40S_ribosomal_protein_S11_putative"	2.69	High expression
136	scf_8254752:10 16165-1017932	TTHERM_00117 590	Note="eukaryotic_porin_protein"	2.67	High expression
137	scf_8254034:86 656-87325	TTHERM_01124 030	Note="hypothetical_protein"	2.67	Low to No- Expression
138	scf_8254406:80 6753-807000	TTHERM_00024 8449	Note="transmembrane_protein_putative"	2.66	#N/A
139	scf_8254403:37 6315-377994	TTHERM_00474 970	Note="60S_acidic_ribosomal_protein"	2.65	High expression
140	scf_8254464:93 016-95144	TTHERM_00578 640	Note="RPL13_60S_ribosomal_protein_L13"	2.64	High expression

141	scf_8254547:14 350-14847	TTHERM_00141 4123	Note="hypothetical_protein"	2.63	#N/A
142	scf_8253957:67 734-67901	TTHERM_00107 1451	Note="hypothetical_protein"	2.63	#N/A
143	scf_8254380:13 011-14054	TTHERM_01125 120	Note="hypothetical_protein"	2.63	High expression
144	scf_8254811:64 7067-647407	TTHERM_00062 680	Note="hypothetical_protein"	2.62	Low to No- Expression
145	scf_8254691:42 1749-422476	TTHERM_00322 920	Note="hypothetical_protein"	2.61	Moderate expression
146	scf_8254600:14 6507-147199	TTHERM_00277 560	Note="MFS_transporter"	2.60	Low to No- Expression
147	scf_8254731:23 4580-235530	TTHERM_00594 370	Note="immobilization_antigen"	2.60	High expression
148	scf_8254674:21 6874-218083	TTHERM_00074 8948	Note="MFS_transporter"	2.60	#N/A
149	scf_8254686:24 4565-246344	TTHERM_00449 230	Note="MFS_transporter"	2.60	Low to No- Expression
150	scf_8254674:26 3009-263448	TTHERM_00074 9021	Note="hypothetical_protein"	2.59	#N/A
151	scf_8254431:90 2340-902645	TTHERM_00057 1689	Note="hypothetical_protein"	2.58	#N/A
152	scf_8254491:76 219-77542	TTHERM_00786 930	Note="NAP1_nucleosome_assembly_protein"	2.58	High expression
153	scf_8254373:42 9125-429508	TTHERM_00024 1579	Note="hypothetical_protein"	2.58	#N/A
154	scf_8254436:28 232-30840	TTHERM_00938 820	Note="EEF2_elongation_factor_2_putative"	2.57	High expression
155	scf_8254652:50 1260-503152	TTHERM_00382 350	Note="RPL26_60S_ribosomal_protein_L26"	2.57	High expression
156	scf_8254496:17 3589-174703	TTHERM_00859 300	Note="MFS_transporter"	2.56	Low to No- Expression
157	scf_8254666:39 6120-397352	TTHERM_00402 120	Note="glutathione_S-transferase"	2.56	High expression
158	scf_8254594:19 707-21259	TTHERM_00285 290	Note="ubiquitin-conjugating_enzyme_E2"	2.56	High expression

159	scf_8254672:45 362-47270	TTHERM_00829 380	Note="RPL11_60S_ribosomal_protein_L11"	2.55	High expression
160	scf_8254397:99 618-100654	TTHERM_00823 720	Note="HHO1_hypothetical_protein"	2.55	High expression
161	scf_8254736:24 0292-240681	TTHERM_00073 6519	Note="transmembrane_protein_putative"	2.53	#N/A
162	scf_8254608:58 226-58534	TTHERM_01009 880	Note="transmembrane_protein_putative"	2.53	Low to No- Expression
163	scf_8254437:11 9283-119681	TTHERM_00044 4469	Note="hypothetical_protein"	2.52	#N/A
164	scf_8254034:13 9220-139519	TTHERM_00852 780	Note="transmembrane_protein_putative"	2.52	Moderate expression
165	scf_8254564:66 598-67395	TTHERM_00216 040	Note="high_mobility_group_(HMG)_box_protein"	2.51	High expression
166	scf_8254431:10 17142-1019541	TTHERM_00569 340	Note="poly(ADP-ribose)_polymerase_and_DNA-ligase_Zn- finger_region_protein"	2.51	Low to No- Expression
167	scf_8254496:14 6216-147720	TTHERM_00859 260	Note="EF-1_guanine_nucleotide_exchange_domain_protein"	2.50	High expression
168	scf_8254674:24 1789-243671	TTHERM_00748 980	Note="MFS_transporter"	2.50	Low to No- Expression
169	scf_8254659:14 03468-1404397	TTHERM_00052 310	Note="AAC_ADG/ATP_transporter_on_adenylate_translocase"	2.50	High expression
170	scf_8254457:11 1904-114756	TTHERM_01014 750	Note="chaperone_protein_DnaK"	2.50	High expression
171	scf_8254716:77 2649-774607	TTHERM_00105 110	Note="HSP70_heat_shock_70_kDa_protein"	2.50	High expression
172	scf_8254487:37 7317-380618	TTHERM_00502 380	Note="peptidase_M16_inactive_domain_protein"	2.50	High expression
173	scf_8254776:23 5829-237134	TTHERM_00161 200	Note="hypothetical_protein"	2.50	High expression
174	scf_8254582:19 4918-196060	TTHERM_00013 8409	Note="hypothetical_protein"	2.49	#N/A
175	scf_8254674:24 1260-241594	TTHERM_00074 8979	Note="MFS_transporter"	2.48	#N/A
176	scf_8254719:85 1998-852369	TTHERM_00782 060	Note="TCT1B_dynein_light_chain_tctex-type_1_protein"	2.47	Moderate expression

177	scf_8254688:59 211-59505	TTHERM_00029 7069	Note="hypothetical_protein"	2.47	#N/A
178	scf_8254716:21 7658-217843	TTHERM_00099 850	Note="hypothetical_protein"	2.46	Low to No- Expression
179	scf_8254528:43 438-45952	TTHERM_00636 970	Note="RPP0_60S_acidic_ribosomal_protein_P0"	2.45	High expression
180	scf_8254600:44 8021-448440	TTHERM_00028 3279	Note="hypothetical_protein"	2.45	#N/A
181	scf_8254492:52 533-54573	TTHERM_01150 360	Note="MFS_transporter"	2.44	Low to No- Expression
182	scf_8254654:50 6509-507187	TTHERM_00051 9879	Note="transmembrane_protein_putative"	2.44	#N/A
183	scf_8254815:39 4508-395392	TTHERM_00020 1617	Note="MFS_transporter"	2.44	#N/A
184	scf_8254435:11 0723-111551	TTHERM_00655 520	Note="superoxide_dismutase_[Cu-Zn]_protein"	2.41	High expression
185	scf_8254757:83 21-10988	TTHERM_01004 930	Note="metallopeptidase_family_M24_containing_protein"	2.40	High expression
186	scf_8254497:60 34-6329	TTHERM_01206 400	Note="hypothetical_protein"	2.40	Low to No- Expression
187	scf_8254777:98 347-98646	TTHERM_00374 950	Note="hypothetical_protein"	2.40	High expression
188	scf_8254582:31 2280-313956	TTHERM_00138 080	Note="FIB1_rRNA_2-prime-O-methyltransferase_fibrillarin"	2.40	High expression
189	scf_8254811:43 7910-440189	TTHERM_00068 140	Note="LSC1_succinyl-CoA_synthetase_alpha_subunit"	2.39	High expression
190	scf_8254486:38 3371-384720	TTHERM_00558 620	Note="ATU1_tubulin/FtsZ_family_GTPase_domain_protein"	2.39	High expression
191	scf_8254479:43 4190-435096	TTHERM_00624 860	Note="thioredoxin_protein"	2.38	High expression
192	scf_8254379:19 7365-197668	TTHERM_00029 9602	Note="hypothetical_protein"	2.37	#N/A
193	scf_8254610:12 54-3549	TTHERM_00333 210	Note="RPL4_60S_ribosomal_protein_L4"	2.37	High expression
194	scf_8254752:11 09953-1110668	TTHERM_00012 0879	Note="hypothetical_protein"	2.36	#N/A

195	scf_8254444:29 3932-296957	TTHERM_00585 260	Note="ATP2_ATP_synthase_F1_beta_subunit"	2.36	High expression
196	scf_8254403:11 6322-117179	TTHERM_00047 3239	Note="kinase_domain_protein"	2.35	#N/A
197	scf_8254545:92 6102-926522	TTHERM_00151 860	Note="complex_1_protein_LYR_family_protein"	2.34	Low to No- Expression
198	scf_8254776:25 4862-256439	TTHERM_00161 130	Note="CTH90_papain_family_cysteine_protease"	2.34	High expression
199	scf_8254007:30 87-5898	TTHERM_01146 030	Note="CAT1_catalase_heme-binding_enzyme"	2.33	High expression
200	scf_8254609:18 3617-184881	TTHERM_00442 200	Note="hypothetical_protein"	2.33	High expression
201	scf_8254813:40 3378-404292	TTHERM_00363 210	Note="ADP_ATP_transporter_on_adenylate_translocase"	2.33	High expression
202	scf_8253864:11 89-3444	TTHERM_01339 580	Note="hyaluronan/mRNA-binding_family_protein"	2.32	High expression
203	scf_8254515:44 94-6734	TTHERM_00721 150	Note="60S_ribosomal_protein_L7"	2.31	High expression
204	scf_8254665:36 9888-370449	TTHERM_00022 7039	Note="hypothetical_protein"	2.30	#N/A
205	scf_8254496:11 2486-113145	TTHERM_00085 8140	Note="transmembrane_protein_putative"	2.29	#N/A
206	scf_8254666:39 4310-394915	TTHERM_00040 2119	Note="transmembrane_protein_putative"	2.29	#N/A
207	scf_8254431:83 9502-843687	TTHERM_00571 860	Note="ATP1_ATP_synthase_F1_alpha_subunit"	2.26	High expression
208	scf_8254417:39 803-39955	TTHERM_00073 0368	Note="hypothetical_protein"	2.25	#N/A
209	scf_8254047:20 4760-205097	TTHERM_00079 0889	Note="hypothetical_protein"	2.25	#N/A
210	scf_8254545:47 3367-474856	TTHERM_00149 630	Note="RPL34_60S_ribosomal_protein_L34"	2.25	High expression
211	scf_8254431:98 0734-981210	TTHERM_00569 490	Note="EF-hand_protein"	2.25	Moderate expression
212	scf_8254814:74 549-75230	TTHERM_00418 400	Note="hypothetical_protein"	2.25	Low to No- Expression

213	scf_8253817:40 9-822	TTHERM_00091 8449	Note="transmembrane_protein_putative"	2.25	#N/A
214	scf_8254815:83 3659-833964	TTHERM_00021 4698	Note="hypothetical_protein"	2.24	#N/A
215	scf_8254708:93 259-99017	TTHERM_01053 030	Note="CCT_motif_protein"	2.24	High expression
216	scf_8254665:37 1741-372070	TTHERM_00227 040	Note="FKB1_peptidyl-prolyl_cis-trans_isomerase"	2.24	High expression
217	scf_8254763:11 4398-116128	TTHERM_00055 9961	Note="hypothetical_protein"	2.23	#N/A
218	scf_8254760:60 759-61250	TTHERM_01018 350	Note="hypothetical_protein"	2.22	Low to No- Expression
219	scf_8254667:76 3932-764210	TTHERM_00003 0389	Note="interferon-induced_6-16_family_protein"	2.21	#N/A
220	scf_8253891:18 5783-186745	TTHERM_00683 010	Note="PLA1_papain_family_cysteine_protease"	2.20	High expression
221	scf_8254373:71 3497-713934	TTHERM_00242 240	Note="hypothetical_protein"	2.20	Moderate expression
222	scf_8254470:79 5333-795659	TTHERM_00433 520	Note="MTT2_copper-induced_metallothionein"	2.20	Moderate expression
223	scf_8254736:24 0096-240269	TTHERM_00736 510	Note="hypothetical_protein"	2.19	Low to No- Expression
224	scf_8254650:62 4645-625187	TTHERM_00077 590	Note="hypothetical_protein"	2.18	Moderate expression
225	scf_8254691:37 215-37490	TTHERM_00319 970	Note="acyl-CoA-binding_protein"	2.18	High expression
226	scf_8254385:44 9407-451780	TTHERM_00534 010	Note="zinc_finger_C-x8-C-x5-C-x3-H_type_protein"	2.17	High expression
227	scf_8254813:27 3382-274812	TTHERM_00361 900	Note="immobilization_antigen"	2.17	Low to No- Expression
228	scf_8254527:36 441-36629	TTHERM_00485 830	Note="hypothetical_protein"	2.17	Low to No- Expression
229	scf_8254515:23 9160-239522	TTHERM_00072 1789	Note="transmembrane_protein_putative"	2.16	#N/A
230	scf_8254776:48 7631-489846	TTHERM_00158 520	Note="HSP82_predicted_protein"	2.16	High expression

231	scf_8254752:10 27278-1028177	TTHERM_00118 610	Note="ATP_synthase_F1_gamma_subunit"	2.16	High expression
232	scf_8254748:10 8669-109723	TTHERM_00706 330	Note="RRM45_RNA_recognition_motif_protein"	2.16	High expression
233	scf_8254666:14 9420-149743	TTHERM_00399 450	Note="Tim10/DDP_family_zinc_finger_protein"	2.15	Moderate expression
234	scf_8254362:12 9804-130818	TTHERM_00723 290	Note="Sm_protein_B_domain_protein"	2.15	High expression
235	scf_8254813:45 3523-454175	TTHERM_00364 320	Note="hypothetical_protein"	2.15	Low to No- Expression
236	scf_8254428:27 0802-272768	TTHERM_00535 929	Note="DNA_double-strand_break_repair_rad50_ATPase_putative"	2.14	High expression
237	scf_8254815:38 6045-387219	TTHERM_00020 0570	Note="hypothetical_protein"	2.12	#N/A
238	scf_8254548:54 9134-551137	TTHERM_00317 440	Note="MFE1_acyl_carrier_reductase"	2.12	High expression
239	scf_8254187:68 508-69062	TTHERM_00069 7299	Note="hypothetical_protein"	2.10	#N/A
240	scf_8254751:13 7302-137658	TTHERM_00094 2949	Note="transmembrane_protein_putative"	2.09	#N/A
241	scf_8254006:11 0997-112298	TTHERM_00945 210	Note="transmembrane_protein_putative"	2.09	High expression
242	scf_8254659:69 5183-697628	TTHERM_00047 480	Note="RPS1_40S_ribosomal_protein_S3a"	2.08	High expression
243	scf_8254028:17 742-18459	TTHERM_00067 0107	Note="poxvirus_A32_protein"	2.07	#N/A
244	scf_8254373:13 0883-131388	TTHERM_00023 7681	Note="hypothetical_protein"	2.07	#N/A
245	scf_8254382:91 777-92208	TTHERM_00094 7479	Note="transmembrane_protein_putative"	2.06	#N/A
246	scf_8254610:23 425-23754	TTHERM_00033 4255	Note="hypothetical_protein"	2.06	#N/A
247	scf_8254470:63 4674-636088	TTHERM_00434 040	Note="RPS12_40S_ribosomal_protein_S23"	2.06	High expression
248	scf_8254697:10 92559-1093599	TTHERM_00083 480	Note="CTH39_papain_family_cysteine_protease"	2.06	High expression

249	scf_8254470:64 3636-646946	TTHERM_00434 000	Note="transmembrane_protein_putative"	2.06	High expression
250	scf_8253915:18 5394-188074	TTHERM_00688 720	Note="dual_specificity_protein_phosphatase_CDC14A"	2.05	Moderate expression
251	scf_8254622:65 807-66882	TTHERM_00119 4835	Note="hypothetical_protein"	2.04	#N/A
252	scf_8254460:35 7740-358453	TTHERM_00348 650	Note="dynein_light_chain_roadblock-type_2_protein"	2.04	Moderate expression
253	scf_8254688:41 3884-414840	TTHERM_00295 170	Note="thioredoxin-dependent_peroxide_reductase"	2.04	High expression
254	scf_8254617:88 0635-880937	TTHERM_00019 6499	Note="hypothetical_protein"	2.03	#N/A
255	scf_8254034:95 080-96210	TTHERM_00851 650	Note="hypothetical_protein"	2.03	Low to No- Expression
256	scf_8253880:68 855-70296	TTHERM_01055 600	Note="GRL8_Ndc1_protein"	2.03	High expression
257	scf_8254638:49 8241-498621	TTHERM_00189 610	Note="hypothetical_protein"	2.03	Low to No- Expression
258	scf_8254549:17 7508-181672	TTHERM_00777 260	Note="hypothetical_protein"	2.02	Moderate expression
259	scf_8254475:15 7456-158160	TTHERM_00420 130	Note="hypothetical_protein"	2.02	High expression
260	scf_8254440:19 0276-191376	TTHERM_00825 590	Note="NAD/FAD-dependent_oxidoreductase"	2.02	Moderate expression
261	scf_8254433:98 77-10167	TTHERM_01085 460	Note="hypothetical_protein"	2.01	Moderate expression
262	scf_8254551:40 3130-403646	TTHERM_00031 0009	Note="hypothetical_protein"	2.00	#N/A
263	scf_8254645:33 6048-336344	TTHERM_00129 180	Note="hypothetical_protein"	1.99	Low to No- Expression
264	scf_8254627:59 067-59426	TTHERM_00086 0479	Note="hypothetical_protein"	1.98	#N/A
265	scf_8254803:14 00645-1403842	TTHERM_00013 300	Note="hypothetical_protein"	1.96	Low to No- Expression
266	scf_8254545:51 6976-519501	TTHERM_00149 798	Note="RPS6_40S_ribosomal_protein_S6"	1.96	High expression

267	scf_8254649:80 1117-805468	TTHERM_00276 030	Note="hypothetical_protein"	1.95	Low to No- Expression
268	scf_8254776:10 8980-109805	TTHERM_00161 620	Note="ankyrin_repeat_protein"	1.95	Low to No- Expression
269	scf_8254659:54 3004-543376	TTHERM_00004 7049	Note="hypothetical_protein"	1.93	#N/A
270	scf_8254594:52 5540-526293	TTHERM_00028 9579	Note="hypothetical_protein"	1.93	#N/A
271	scf_8254617:56 3949-564460	TTHERM_00194 690	Note="ubiquinol-cytochrome_C_reductase_hinge_protein"	1.92	High expression
272	scf_8254371:94 025-94718	TTHERM_00046 6140	Note="hypothetical_protein"	1.91	#N/A
273	scf_8254527:53 487-54517	TTHERM_00048 5879	Note="transmembrane_protein_putative"	1.90	#N/A
274	scf_8254449:17 6875-177219	TTHERM_00773 580	Note="cytochrome_C-like_protein"	1.90	High expression
275	scf_8254468:40 569-40874	TTHERM_00104 4599	Note="transmembrane_protein_putative"	1.89	#N/A
276	scf_8253811:16 0789-161582	TTHERM_00526 450	Note="RPS30_40S_ribosomal_protein_S30"	1.88	High expression
277	scf_8254752:34 186-35208	TTHERM_00057 7279	Note="transmembrane_protein_putative"	1.88	#N/A
278	scf_8254610:33 6933-338845	TTHERM_00033 8239	Note="60S_ribosomal_protein_L18"	1.87	#N/A
279	scf_8254664:32 5448-327500	TTHERM_00522 630	Note="RPS0_40S_ribosomal_protein_S0"	1.87	High expression
280	scf_8254716:11 70535-1170945	TTHERM_00107 160	Note="transmembrane_protein_putative"	1.86	Low to No- Expression
281	scf_8254652:13 5577-136364	TTHERM_00378 520	Note="dual_specificity_phosphatase_domain_protein"	1.86	Low to No- Expression
282	scf_8254757:10 1441-101705	TTHERM_00100 5219	Note="hypothetical_protein"	1.85	#N/A
283	scf_8254428:42 741-45053	TTHERM_00535 260	Note="succinyl-CoA_ligase_[GDP-forming]_subunit_beta"	1.84	High expression
284	scf_8254697:15 51179-1551688	TTHERM_00035 5369	Note="hypothetical_protein"	1.84	#N/A

285	scf_8254731:66 114-66674	TTHERM_00592 860	Note="hypothetical_protein"	1.82	Low to No- Expression
286	scf_8254479:10 8562-110691	TTHERM_00622 780	Note="hypothetical_protein"	1.81	Low to No- Expression
287	scf_8254600:78 2719-783703	TTHERM_00284 150	Note="hypothetical_protein"	1.81	Low to No- Expression
288	scf_8254716:84 0233-840778	TTHERM_00105 330	Note="RPS28_40S_ribosomal_protein_S28e_putative"	1.81	High expression
289	scf_8254597:63 7247-638985	TTHERM_00346 900	Note="serine_carboxypeptidase_family_protein"	1.80	High expression
290	scf_8254799:10 9525-111115	TTHERM_00812 780	Note="RPL38_60S_ribosomal_protein_L38"	1.79	High expression
291	scf_8254010:51 3899-514291	TTHERM_00032 9909	Note="cell_surface_immobilization_antigen"	1.79	#N/A
292	scf_8254645:33 8348-339077	TTHERM_00129 090	Note="transmembrane_protein_putative"	1.79	Low to No- Expression
293	scf_8254645:60 6036-607752	TTHERM_00128 320	Note="CTH5_eukaryotic_aspartyl_protease"	1.79	High expression
294	scf_8254716:89 8388-899522	TTHERM_00105 430	Note="PRF1_profilin"	1.79	High expression
295	scf_8254543:13 7227-139219	TTHERM_00899 460	Note="peroxisomal_3-ketoacyl-CoA_thiolase_B"	1.77	High expression
296	scf_8254395:31 9614-320211	TTHERM_00456 980	Note="transmembrane_protein_putative"	1.76	Low to No- Expression
297	scf_8254652:29 0736-292184	TTHERM_00378 890	Note="GRL5_granule_lattice_protein"	1.75	High expression
298	scf_8254617:82 3374-826083	TTHERM_00196 370	Note="HSP60_chaperonin_CPN60-1"	1.75	High expression
299	scf_8254401:66 793-67153	TTHERM_01080 490	Note="hypothetical_protein"	1.75	High expression
300	scf_8254811:88 0913-882047	TTHERM_00005 9119	Note="transmembrane_protein_putative"	1.75	#N/A
301	scf_8254645:63 4345-636845	TTHERM_00127 260	Note="transmembrane_protein_putative"	1.74	High expression
302	scf_8254436:60 93-7046	TTHERM_00938 750	Note="DnaJ_domain_protein"	1.74	High expression

303	scf_8254582:55 6425-558373	TTHERM_00136 440	Note="transcription_factor_apfi_protein_putative"	1.74	High expression
304	scf_8254658:14 7089-148071	TTHERM_00673 300	Note="hypothetical_protein"	1.74	Low to No- Expression
305	scf_8254437:12 0945-122135	TTHERM_00444 470	Note="S-adenosylmethionine_synthetase_protein"	1.74	High expression
306	scf_8254667:70 2946-704700	TTHERM_00030 190	Note="MDH1_cytoplasmic_and_cytosolic_malate_dehydrogenase_family_protein"	1.74	High expression
307	scf_8254798:40 6778-409877	TTHERM_00171 850	Note="SSA6_heat_shock-binding_protein_70_ER_luminal_protein"	1.73	High expression
308	scf_8254776:67 7754-678311	TTHERM_00158 040	Note="Ran-binding_protein_2_Ran-binding_domain_protein"	1.73	High expression
309	scf_8254803:10 24054-1025385	TTHERM_00011 400	Note="Myb-like_DNA-binding_domain_protein"	1.73	Moderate expression
310	scf_8254377:24 5390-247749	TTHERM_00716 240	Note="RPL2_60S_ribosomal_protein_L2/L8"	1.72	High expression
311	scf_8254598:20 4395-204706	TTHERM_00075 3491	Note="hypothetical_protein"	1.71	#N/A
312	scf_8254580:22 8364-230775	TTHERM_00590 090	Note="CMB1_hypothetical_protein"	1.71	High expression
313	scf_8254620:65 37-7795	TTHERM_00146 9326	Note="hypothetical_protein"	1.69	#N/A
314	scf_8254379:64 9461-651958	TTHERM_00301 770	Note="hypothetical_protein"	1.67	Moderate expression
315	scf_8254564:52 4406-530182	TTHERM_00219 210	Note="Serine/Threonine_kinase_domain_protein"	1.67	Moderate expression
316	scf_8254431:90 0013-900453	TTHERM_00571 700	Note="transmembrane_protein_putative"	1.66	Low to No- Expression
317	scf_8254719:68 1350-681684	TTHERM_00078 4569	Note="hypothetical_protein"	1.66	#N/A
318	scf_8254814:19 1715-192564	TTHERM_00041 8058	Note="transmembrane_protein_putative"	1.65	#N/A
319	scf_8254475:11 6817-118391	TTHERM_00420 000	Note="KHD4_predicted_protein"	1.64	High expression
320	scf_8254783:13 6651-137295	TTHERM_00084 7039	Note="transmembrane_protein_putative"	1.64	#N/A

321	scf_8254803:20 42794-2044236	TTHERM_00023 980	Note="RAN1_GTP-binding_nuclear_protein_Ran-1"	1.64	High expression
322	scf_8254464:32 2523-322903	TTHERM_00579 270	Note="SSS1_protein_transporter_SEC61_subunit_gamma"	1.63	High expression
323	scf_8254470:10 40235-1041347	TTHERM_00429 890	Note="nucleoplasmin_protein"	1.63	High expression
324	scf_8254437:47 6586-477375	TTHERM_00044 6389	Note="hypothetical_protein"	1.63	#N/A
325	scf_8254776:14 8766-148930	TTHERM_00161 450	Note="hypothetical_protein"	1.62	Low to No- Expression
326	scf_8254671:43 8673-441360	TTHERM_00426 260	Note="hypothetical_protein"	1.62	High expression
327	scf_8254719:28 2366-283381	TTHERM_00393 180	Note="hypothetical_protein"	1.61	Low to No- Expression
328	scf_8254803:10 98124-1098522	TTHERM_00001 1599	Note="hypothetical_protein"	1.60	#N/A
329	scf_8254676:49 842-51511	TTHERM_01190 440	Note="HOPP2_TBP1_interacting_protein"	1.59	Moderate expression
330	scf_8254590:42 5273-425698	TTHERM_00253 430	Note="hypothetical_protein"	1.58	High expression
331	scf_8253992:10 1052-101909	TTHERM_00100 8728	Note="transmembrane_protein_putative"	1.58	#N/A
332	scf_8254803:19 86009-1987557	TTHERM_00022 810	Note="phosphoenolpyruvate_phosphomutase"	1.58	High expression
333	scf_8254597:28 4020-286760	TTHERM_00344 030	Note="NAD-dependent_isocitrate_dehydrogenase"	1.57	High expression
334	scf_8254377:25 4150-254782	TTHERM_00716 270	Note="scavenger_mRNA_decapping_enzyme_carboxy-term-binding_protein"	1.57	High expression
335	scf_8253978:33 46-4360	TTHERM_01165 210	Note="hypothetical_protein"	1.57	Low to No- Expression
336	scf_8254049:77 774-78686	TTHERM_01145 000	Note="thioredoxin_protein"	1.57	High expression
337	scf_8253908:26 686-27231	TTHERM_00090 1634	Note="kinase_domain_protein_putative"	1.56	#N/A
338	scf_8254679:63 679-63952	TTHERM_01227 890	Note="hypothetical_protein"	1.56	Low to No- Expression

339	scf_8255706:80 2-1112	TTHERM_02653 301	Note="Cyp-like_protein_putative"	1.56	High expression
340	scf_8254813:26 9984-271730	TTHERM_00361 890	Note="hypothetical_protein"	1.56	Low to No- Expression
341	scf_8254803:17 30885-1732357	TTHERM_00016 060	Note="short_chain_dehydrogenase_reductase_family_protein_putative"	1.56	High expression
342	scf_8254646:13 2670-133178	TTHERM_00819 520	Note="hypothetical_protein"	1.56	Low to No- Expression
343	scf_8254667:75 8579-759627	TTHERM_00003 0359	Note="hypothetical_protein"	1.55	#N/A
344	scf_8253974:25 816-26808	TTHERM_01276 400	Note="CTH52_papain_family_cysteine_protease"	1.55	Moderate expression
345	scf_8254564:78 6946-787239	TTHERM_00022 0969	Note="transmembrane_protein_putative"	1.55	#N/A
346	scf_8254717:13 3694-134247	TTHERM_00098 6269	Note="transmembrane_protein_putative"	1.54	#N/A
347	scf_8254667:49 2706-494389	TTHERM_00028 740	Note="RPL12_ribosomal_protein_L11_amine-terminal_domain_protein"	1.54	High expression
348	scf_8253880:11 0804-111513	TTHERM_00105 5719	Note="transmembrane_protein_putative"	1.53	#N/A
349	scf_8254777:15 1904-152323	TTHERM_00373 770	Note="transmembrane_protein_putative"	1.53	High expression
350	scf_8254712:18 3571-184077	TTHERM_00803 670	Note="transmembrane_protein_putative"	1.52	Low to No- Expression
351	scf_8254617:21 7566-219580	TTHERM_00193 740	Note="RPS2_40S_ribosomal_protein_S2_putative"	1.52	High expression
352	scf_8254776:38 3303-384808	TTHERM_00160 770	Note="FTT49_tyrosine_3-monooxygenase/tryptophan_5- monooxygenase_activation_protein_epsilon_protein"	1.52	High expression
353	scf_8254440:33 958-34979	TTHERM_00825 140	Note="hypothetical_protein"	1.52	Low to No- Expression
354	scf_8254736:17 8649-181091	TTHERM_00735 340	Note="vacuolar_H+-translocating_inorganic_pyrophosphatase"	1.51	High expression

RACS processed data for the 842 intergenic regions occupied by that Ibd1: Refer to Digital Spreadsheet (TET_Ibd1_MAC_Genome_Intergenic) for complete results.

#	Scaffold	Enrichment Average	#	Scaffold	Enrichment Average
1	scf_8254688:663530-663965	18.17	421	scf_8254803:34848-36701	2.10
2	scf_8254736:240270-240291	16.00	422	scf_8254716:1102950-1104136	2.10
3	scf_8254410:100350-100779	15.19	423	scf_8254379:716153-718285	2.10
4	scf_8254297:1-2450	12.23	424	scf_8254392:44132-47841	2.10
5	scf_8254823:448673-449008	11.76	425	scf_8254187:221236-221879	2.09
6	scf_8254688:186225-186315	10.36	426	scf_8254588:209638-210396	2.09
7	scf_8254697:28876-29485	10.30	427	scf_8254670:96972-97470	2.09
8	scf_8254719:258830-258854	9.40	428	scf_8254674:263449-264934	2.09
9	scf_8254666:234688-235147	8.16	429	scf_8254666:266819-268379	2.09
10	scf_8254582:322414-322732	7.68	430	scf_8254379:1-1213	2.08
11	scf_8254688:140258-140635	6.97	431	scf_8254638:324415-326144	2.08
12	scf_8254437:133715-133753	6.89	432	scf_8253894:89068-93871	2.08
13	scf_8254564:283445-283717	6.79	433	scf_8254697:1272010-1273785	2.07
14	scf_8254659:1646126-1646764	5.71	434	scf_8254649:684647-686341	2.07
15	scf_8254564:621208-621725	5.69	435	scf_8254469:35823-36522	2.07
16	scf_8254594:335348-336525	5.68	436	scf_8254587:348268-349436	2.06
17	scf_8254650:1227049-1230315	5.68	437	scf_8254752:290803-291558	2.06
18	scf_8254582:837903-838644	5.64	438	scf_8254563:360239-360545	2.06
19	scf_8254600:746329-746908	5.53	439	scf_8254549:28866-30263	2.06
20	scf_8254401:115330-115639	5.39	440	scf_8254665:305443-306107	2.06
21	scf_8254397:90712-90809	5.34	441	scf_8253899:60903-61682	2.05
22	scf_8254095:53861-54722	5.30	442	scf_8254646:36957-37561	2.05
23	scf_8254359:139473-140428	5.28	443	scf_8254574:70375-70861	2.05
24	scf_8254659:917911-918509	5.18	444	scf_8254776:861295-863977	2.05
25	scf_8254691:569722-569741	5.05	445	scf_8253992:98388-99243	2.05
26	scf_8254667:2094720-2094798	5.04	446	scf_8254649:693242-694266	2.05
27	scf_8253891:162391-163693	5.04	447	scf_8254751:52271-53059	2.05
28	scf_8254590:565657-566989	4.94	448	scf_8254468:28563-31765	2.05
29	scf_8254667:675192-675592	4.93	449	scf_8254095:50451-52285	2.04
30	scf_8254497:1482-2129	4.79	450	scf_8254653:340274-342787	2.04
31	scf_8254600:159363-159688	4.74	451	scf_8254580:393089-393507	2.04

32	scf_8254630:102675-102975	4.71	452	scf_8254636:367365-368021	2.04
33	scf_8253815:391043-391976	4.68	453	scf_8254365:298371-299645	2.04
34	scf_8254373:386895-387151	4.65	454	scf_8254547:1-1489	2.03
35	scf_8254731:24114-26115	4.55	455	scf_8254670:241502-242262	2.03
36	scf_8254752:34141-34185	4.53	456	scf_8254617:410253-410497	2.03
37	scf_8254493:1-491	4.52	457	scf_8254645:135414-136056	2.03
38	scf_8254564:431690-432322	4.49	458	scf_8254600:435187-436434	2.03
39	scf_8254378:231384-233735	4.43	459	scf_8254373:428451-429124	2.02
40	scf_8254555:181545-181756	4.29	460	scf_8254788:523864-524274	2.02
41	scf_8254450:5582-6165	4.29	461	scf_8254688:215346-216881	2.02
42	scf_8254652:109766-110387	4.29	462	scf_8254650:33836-35671	2.01
43	scf_8254591:14314-14582	4.27	463	scf_8254609:71806-72269	2.00
44	scf_8254449:28967-29386	4.12	464	scf_8254747:1013750-1014492	2.00
45	scf_8254406:387412-387673	4.10	465	scf_8254592:516297-516841	2.00
46	scf_8254650:226272-227703	4.04	466	scf_8254486:321767-322528	2.00
47	scf_8254391:156473-156612	4.03	467	scf_8254788:779666-783365	2.00
48	scf_8254515:238677-239159	4.02	468	scf_8254429:17002-17906	1.99
49	scf_8254385:337896-338250	4.00	469	scf_8253894:212787-214095	1.99
50	scf_8254444:232671-233328	3.99	470	scf_8254787:12811-13631	1.99
51	scf_8254373:746134-746442	3.97	471	scf_8254548:398922-400591	1.99
52	scf_8254811:1273468-1273591	3.96	472	scf_8254486:340116-342027	1.99
53	scf_8254747:803895-804124	3.93	473	scf_8254716:155794-157915	1.99
54	scf_8254558:12513-12987	3.89	474	scf_8254697:1752087-1753648	1.99
55	scf_8253815:515570-515884	3.88	475	scf_8254475:146808-147074	1.98
56	scf_8254181:52724-53079	3.87	476	scf_8254002:51128-52970	1.98
57	scf_8254545:359158-359535	3.83	477	scf_8254611:23081-24994	1.98
58	scf_8254359:170776-171721	3.83	478	scf_8254370:220964-221526	1.98
59	scf_8254385:382964-383402	3.81	479	scf_8254607:299721-300143	1.98
60	scf_8254459:201981-202253	3.79	480	scf_8254370:6848-7113	1.98
61	scf_8254431:912334-913205	3.77	481	scf_8254549:183533-184613	1.98
62	scf_8254565:343422-343878	3.73	482	scf_8254187:33538-34349	1.98
63	scf_8254437:442385-443780	3.73	483	scf_8254814:204562-208630	1.98
64	scf_8254667:379483-380749	3.73	484	scf_8254489:70101-71634	1.97
65	scf_8254476:192413-193798	3.70	485	scf_8254476:210214-210810	1.97
66	scf_8254406:90838-91575	3.67	486	scf_8254397:60704-61453	1.97
67	scf_8254362:128806-129803	3.66	487	scf_8254592:117430-117850	1.97

68	scf_8254665:188656-188995	3.65	488	scf_8254617:145638-146862	1.97
69	scf_8254590:39894-40397	3.59	489	scf_8254747:820985-821887	1.97
70	scf_8254688:601867-602204	3.57	490	scf_8254428:154290-154700	1.96
71	scf_8254006:112299-113123	3.56	491	scf_8254610:372195-373960	1.96
72	scf_8254373:161362-161599	3.53	492	scf_8254483:214034-217129	1.96
73	scf_8254788:381572-381713	3.52	493	scf_8254598:111873-113760	1.96
74	scf_8254572:43902-44139	3.52	494	scf_8253811:232038-232423	1.96
75	scf_8254403:236870-239184	3.51	495	scf_8254644:100233-100457	1.96
76	scf_8254607:429191-429425	3.50	496	scf_8254584:1-620	1.96
77	scf_8254649:428675-430284	3.46	497	scf_8254728:70321-71567	1.95
78	scf_8254610:653935-654649	3.46	498	scf_8254691:258605-258825	1.95
79	scf_8254671:564479-564843	3.45	499	scf_8254403:408852-409228	1.95
80	scf_8254600:409839-411012	3.45	500	scf_8254667:1215139-1216158	1.95
81	scf_8254719:859579-860620	3.45	501	scf_8254395:362153-363164	1.95
82	scf_8254503:102570-102835	3.44	502	scf_8254373:429509-429691	1.94
83	scf_8254718:16302-16873	3.43	503	scf_8254470:324251-326796	1.94
84	scf_8254365:165549-166450	3.43	504	scf_8254486:75807-76812	1.94
85	scf_8254763:227565-228339	3.42	505	scf_8254716:418459-419104	1.94
86	scf_8254712:140910-141291	3.42	506	scf_8254763:243032-243926	1.94
87	scf_8253815:486894-487791	3.39	507	scf_8254645:109758-110669	1.94
88	scf_8254420:290155-291123	3.39	508	scf_8254811:1027924-1029147	1.94
89	scf_8254611:80858-82367	3.39	509	scf_8254367:165454-169235	1.93
90	scf_8254486:183205-184478	3.38	510	scf_8254688:1-322	1.93
91	scf_8254437:108324-108797	3.37	511	scf_8254688:524617-525060	1.92
92	scf_8253815:195474-196793	3.36	512	scf_8254385:1-950	1.92
93	scf_8254718:222816-223482	3.36	513	scf_8254638:183928-184600	1.92
94	scf_8254479:433339-434189	3.35	514	scf_8254010:526477-527413	1.91
95	scf_8253957:67082-67733	3.35	515	scf_8254028:197955-200532	1.91
96	scf_8253984:196361-199651	3.33	516	scf_8254645:339078-339614	1.91
97	scf_8254449:87659-88489	3.33	517	scf_8254051:51641-53638	1.90
98	scf_8254622:66883-68258	3.33	518	scf_8254650:320693-321318	1.90
99	scf_8254420:125344-128868	3.32	519	scf_8254406:457607-458144	1.90
100	scf_8254397:172333-173158	3.32	520	scf_8254644:1-227	1.90
101	scf_8254610:131074-131345	3.31	521	scf_8254403:305216-305870	1.90
102	scf_8254777:604572-605401	3.28	522	scf_8255662:632-1020	1.90
103	scf_8254459:150815-151193	3.27	523	scf_8254417:39956-40585	1.90

104	scf_8254007:12403-13870	3.25	524	scf_8254763:169165-170751	1.90
105	scf_8254187:173230-174179	3.25	525	scf_8254444:94016-94959	1.90
106	scf_8254034:103687-104124	3.24	526	scf_8254594:153295-154941	1.90
107	scf_8254406:782831-783216	3.24	527	scf_8254187:69063-69712	1.90
108	scf_8254788:659215-659919	3.23	528	scf_8254682:78833-80341	1.90
109	scf_8254653:81790-82173	3.22	529	scf_8254460:33937-34199	1.90
110	scf_8254610:575380-575872	3.22	530	scf_8254551:190620-192572	1.89
111	scf_8254811:1249785-1250868	3.22	531	scf_8254752:1598297-1599409	1.89
112	scf_8254803:2205448-2206307	3.22	532	scf_8253823:230381-231612	1.89
113	scf_8254718:8558-9654	3.21	533	scf_8254449:13377-13714	1.89
114	scf_8254487:380619-381486	3.21	534	scf_8254811:195486-196276	1.89
115	scf_8254043:198868-199390	3.20	535	scf_8254707:164319-173643	1.89
116	scf_8254571:204290-204785	3.20	536	scf_8254815:224802-225355	1.89
117	scf_8254564:90814-91823	3.19	537	scf_8254410:137635-138568	1.89
118	scf_8254803:1800594-1800887	3.18	538	scf_8253984:60507-60913	1.89
119	scf_8254428:272769-273366	3.17	539	scf_8254691:313499-315567	1.89
120	scf_8254752:1109078-1109952	3.17	540	scf_8254788:203397-204875	1.89
121	scf_8254437:179593-179981	3.16	541	scf_8254788:800088-800956	1.88
122	scf_8254468:69809-70975	3.16	542	scf_8254475:74762-76491	1.88
123	scf_8254582:909509-911004	3.14	543	scf_8254650:995759-996930	1.88
124	scf_8254377:199363-201550	3.14	544	scf_8254389:5481-6044	1.88
125	scf_8254592:309006-309549	3.12	545	scf_8254564:730718-731022	1.88
126	scf_8254638:219360-219762	3.12	546	scf_8254653:119119-120514	1.87
127	scf_8254181:102720-103847	3.11	547	scf_8254590:21912-22276	1.87
128	scf_8254367:170736-171089	3.10	548	scf_8254697:1189931-1190636	1.87
129	scf_8254459:310132-312343	3.10	549	scf_8254095:9007-9653	1.87
130	scf_8253815:376838-377553	3.10	550	scf_8254716:160527-161051	1.87
131	scf_8254496:173280-173588	3.08	551	scf_8254284:246322-247167	1.87
132	scf_8254659:596024-596235	3.08	552	scf_8254486:384721-386943	1.86
133	scf_8254512:8060-9199	3.08	553	scf_8254393:28960-30307	1.86
134	scf_8253930:135182-135750	3.07	554	scf_8254449:2420-3008	1.86
135	scf_8254763:389994-390522	3.07	555	scf_8254610:29058-30044	1.86
136	scf_8254466:49835-50102	3.07	556	scf_8254548:654611-655175	1.86
137	scf_8254663:63238-63724	3.06	557	scf_8254814:68426-72471	1.86
138	scf_8254719:392569-392974	3.05	558	scf_8254627:14809-15169	1.86
139	scf_8254382:111608-113024	3.04	559	scf_8254811:112920-113334	1.86

140	scf_8254582:758592-759687	3.04	560	scf_8254597:146930-147603	1.86
141	scf_8254481:1-1578	3.03	561	scf_8254716:75078-76491	1.85
142	scf_8254650:985346-985830	3.02	562	scf_8254747:462345-463026	1.85
143	scf_8254492:44661-45727	3.01	563	scf_8254594:738268-740553	1.85
144	scf_8254444:1-477	3.01	564	scf_8254617:34000-34500	1.85
145	scf_8254373:456785-459437	3.00	565	scf_8254610:369917-370662	1.85
146	scf_8254752:190712-193134	3.00	566	scf_8253899:49983-51603	1.84
147	scf_8253964:79728-80675	3.00	567	scf_8254803:927004-929291	1.84
148	scf_8254716:840779-844382	3.00	568	scf_8254564:37277-37689	1.84
149	scf_8254007:1-442	2.98	569	scf_8254393:21627-22449	1.84
150	scf_8254584:90955-91300	2.95	570	scf_8254551:526588-527675	1.84
151	scf_8254470:109329-110484	2.95	571	scf_8254811:401285-402561	1.84
152	scf_8254487:201625-202141	2.95	572	scf_8254697:1790541-1790992	1.83
153	scf_8254814:419229-422183	2.95	573	scf_8254590:752669-753452	1.83
154	scf_8254396:89662-90735	2.94	574	scf_8254403:358205-358757	1.83
155	scf_8254431:843688-845288	2.94	575	scf_8254593:181355-182713	1.83
156	scf_8254811:835962-836609	2.92	576	scf_8254549:99494-100484	1.83
157	scf_8254470:350634-351375	2.91	577	scf_8254594:761551-762266	1.83
158	scf_8254575:87685-87844	2.91	578	scf_8254716:1094247-1094827	1.83
159	scf_8254010:1-670	2.91	579	scf_8254580:47053-48930	1.83
160	scf_8254379:251626-252624	2.90	580	scf_8254652:1-1093	1.83
161	scf_8254379:458702-459424	2.89	581	scf_8254047:95053-95670	1.82
162	scf_8254515:239523-239568	2.89	582	scf_8254719:657683-658240	1.82
163	scf_8254752:356209-358939	2.88	583	scf_8253904:100522-101334	1.82
164	scf_8254645:379966-380480	2.85	584	scf_8254431:1016661-1017141	1.82
165	scf_8254811:1295310-1295683	2.85	585	scf_8254051:1-598	1.82
166	scf_8254776:675877-677753	2.85	586	scf_8254788:85165-86720	1.82
167	scf_8254028:97178-97863	2.85	587	scf_8254609:12839-13672	1.82
168	scf_8254378:386477-386758	2.84	588	scf_8254406:333287-333977	1.81
169	scf_8253934:30437-31027	2.83	589	scf_8254440:106326-106914	1.81
170	scf_8254752:194929-195696	2.83	590	scf_8254431:809802-810178	1.81
171	scf_8254719:626079-626648	2.83	591	scf_8254812:41376-44931	1.81
172	scf_8254401:50787-51282	2.82	592	scf_8253815:290932-290967	1.81
173	scf_8254814:103323-104356	2.82	593	scf_8254395:120919-122412	1.81
174	scf_8254379:712897-713740	2.81	594	scf_8254592:250700-252069	1.81
175	scf_8254645:632242-634344	2.78	595	scf_8254811:225140-226398	1.81

176	scf_8253815:254305-254858	2.78	596	scf_8254708:88794-90962	1.80
177	scf_8254437:541588-542079	2.78	597	scf_8254034:96211-97049	1.80
178	scf_8254609:155308-156216	2.77	598	scf_8254688:2029-2682	1.80
179	scf_8254607:1940-2360	2.77	599	scf_8254450:13893-14286	1.80
180	scf_8254385:100853-102274	2.77	600	scf_8254437:119682-120944	1.80
181	scf_8254403:296870-298376	2.76	601	scf_8254803:1647187-1648982	1.79
182	scf_8254435:179869-180962	2.75	602	scf_8254433:4688-5867	1.79
183	scf_8254665:370450-371740	2.75	603	scf_8254611:498097-498972	1.79
184	scf_8254707:132719-136970	2.75	604	scf_8254551:496625-497599	1.79
185	scf_8253887:15793-16679	2.73	605	scf_8254607:429671-430031	1.79
186	scf_8254811:1044684-1045277	2.73	606	scf_8254719:621619-623047	1.79
187	scf_8254617:637400-637706	2.73	607	scf_8254798:830547-830754	1.79
188	scf_8254602:53304-54323	2.73	608	scf_8254679:13975-15316	1.79
189	scf_8254665:467378-468754	2.71	609	scf_8254748:290712-291460	1.78
190	scf_8254396:1-690	2.70	610	scf_8254464:22441-23258	1.78
191	scf_8254010:414440-415861	2.70	611	scf_8254786:192750-194022	1.78
192	scf_8254666:75196-76130	2.69	612	scf_8254392:35493-36994	1.78
193	scf_8254379:644746-645409	2.69	613	scf_8254435:188524-190828	1.78
194	scf_8254607:557657-558682	2.69	614	scf_8254392:38266-39753	1.77
195	scf_8254747:131531-131935	2.69	615	scf_8254667:1158054-1159995	1.77
196	scf_8254582:1056500-1059130	2.68	616	scf_8254470:337751-339418	1.77
197	scf_8254460:358454-358858	2.68	617	scf_8254752:1104250-1106305	1.77
198	scf_8254371:352110-352595	2.68	618	scf_8254649:798652-801116	1.77
199	scf_8254470:65712-66066	2.68	619	scf_8254645:664670-665718	1.77
200	scf_8253930:97862-98791	2.67	620	scf_8254814:424299-425006	1.77
201	scf_8254489:96554-98836	2.67	621	scf_8254431:685646-686695	1.76
202	scf_8254649:528510-529883	2.67	622	scf_8254686:450537-450856	1.76
203	scf_8254548:251873-253052	2.66	623	scf_8254543:180295-183375	1.76
204	scf_8254401:106993-107097	2.65	624	scf_8254563:381483-381965	1.76
205	scf_8254528:111090-111410	2.64	625	scf_8254565:48548-49393	1.76
206	scf_8254747:520059-520499	2.64	626	scf_8254436:125643-126185	1.76
207	scf_8254665:373387-373847	2.63	627	scf_8254373:131389-133040	1.76
208	scf_8254418:32424-33170	2.63	628	scf_8254527:263488-263740	1.76
209	scf_8254373:488568-489025	2.63	629	scf_8254645:311015-312837	1.75
210	scf_8254815:243089-243645	2.63	630	scf_8254361:31479-32061	1.75
211	scf_8254504:83672-83923	2.63	631	scf_8254428:1-1120	1.75

212	scf_8255740:510-1327	2.62	632	scf_8254617:917540-918217	1.75
213	scf_8254711:40706-41340	2.62	633	scf_8254777:97069-98346	1.75
214	scf_8254747:575844-576605	2.62	634	scf_8254788:384251-384926	1.75
215	scf_8254545:997844-998697	2.61	635	scf_8254716:380611-381023	1.75
216	scf_8254645:468858-469665	2.61	636	scf_8254719:279909-282365	1.75
217	scf_8254610:336172-336932	2.61	637	scf_8254415:50344-51137	1.74
218	scf_8254610:329401-330350	2.61	638	scf_8254365:18843-19436	1.74
219	scf_8254752:891461-892443	2.61	639	scf_8254656:60916-61806	1.74
220	scf_8254373:809780-811798	2.60	640	scf_8254617:507687-508657	1.74
221	scf_8254691:300026-300877	2.60	641	scf_8254769:151106-151777	1.73
222	scf_8254691:100259-103502	2.60	642	scf_8254587:559428-560330	1.73
223	scf_8253891:109694-110564	2.60	643	scf_8254497:3895-6033	1.73
224	scf_8253915:188075-191456	2.59	644	scf_8254667:2043791-2045379	1.73
225	scf_8254551:118312-118674	2.59	645	scf_8254777:106673-108603	1.73
226	scf_8254670:253979-254290	2.58	646	scf_8253803:4624-5649	1.73
227	scf_8254594:155843-156917	2.58	647	scf_8254823:200343-200603	1.73
228	scf_8254395:103661-103880	2.58	648	scf_8254803:1560632-1561210	1.73
229	scf_8254371:411143-411873	2.58	649	scf_8254435:111552-113070	1.73
230	scf_8254674:241595-241788	2.57	650	scf_8254545:203442-203841	1.72
231	scf_8254712:67045-69567	2.56	651	scf_8254677:70864-72028	1.72
232	scf_8254799:92073-92481	2.56	652	scf_8254607:244728-245542	1.72
233	scf_8254486:192291-193487	2.55	653	scf_8254010:640460-641140	1.72
234	scf_8254650:617587-618176	2.55	654	scf_8254747:65255-66917	1.72
235	scf_8254665:28037-28609	2.55	655	scf_8254752:730791-731549	1.72
236	scf_8254752:739437-740988	2.55	656	scf_8254422:219517-220379	1.72
237	scf_8254486:185224-185385	2.54	657	scf_8254582:994040-995444	1.71
238	scf_8254422:6770-7527	2.54	658	scf_8254648:189357-189817	1.71
239	scf_8254392:41578-42453	2.54	659	scf_8254697:207936-208909	1.71
240	scf_8254431:1024940-1025761	2.53	660	scf_8254379:195848-197364	1.71
241	scf_8254379:514636-514969	2.53	661	scf_8254610:309148-311622	1.71
242	scf_8254590:386171-386677	2.53	662	scf_8254752:1015033-1016164	1.71
243	scf_8254814:75231-76124	2.53	663	scf_8254597:82119-82430	1.71
244	scf_8254564:379843-382368	2.52	664	scf_8254811:530665-531778	1.71
245	scf_8254446:348797-351798	2.52	665	scf_8254697:1359492-1360447	1.70
246	scf_8254697:1275239-1275689	2.52	666	scf_8254448:69908-73128	1.70
247	scf_8254719:137740-139795	2.52	667	scf_8254028:342644-343190	1.70

248	scf_8254718:107061-107580	2.52	668	scf_8254610:3550-5967	1.70
249	scf_8253915:237745-238009	2.52	669	scf_8254813:404293-405921	1.69
250	scf_8254181:50470-51320	2.52	670	scf_8253887:187036-187551	1.69
251	scf_8254446:43514-43923	2.50	671	scf_8254454:6744-8174	1.69
252	scf_8254665:474001-474868	2.49	672	scf_8254671:377847-380221	1.69
253	scf_8254437:411027-413344	2.49	673	scf_8254807:451542-452421	1.69
254	scf_8253904:243628-244348	2.49	674	scf_8254476:124587-125702	1.69
255	scf_8254686:500967-501843	2.49	675	scf_8254403:435748-436387	1.69
256	scf_8254645:245904-246582	2.48	676	scf_8254649:340560-342749	1.69
257	scf_8254565:115649-116205	2.48	677	scf_8254716:133974-136261	1.69
258	scf_8254483:196894-198043	2.48	678	scf_8254609:487101-487653	1.69
259	scf_8254582:266194-266860	2.47	679	scf_8254437:118903-119282	1.69
260	scf_8254815:244226-244465	2.47	680	scf_8254403:438604-439136	1.68
261	scf_8254479:106986-107560	2.47	681	scf_8254644:272388-272712	1.68
262	scf_8254373:34850-35000	2.46	682	scf_8254406:36035-37045	1.68
263	scf_8254448:84650-85642	2.46	683	scf_8254803:1926389-1926878	1.68
264	scf_8254716:1182377-1183626	2.46	684	scf_8254659:1253886-1254164	1.68
265	scf_8254650:756204-756490	2.44	685	scf_8254777:585057-585639	1.68
266	scf_8254763:308562-310787	2.44	686	scf_8254600:115449-115929	1.68
267	scf_8254460:590595-593210	2.44	687	scf_8254403:495492-496762	1.68
268	scf_8254077:35457-36204	2.44	688	scf_8254284:137894-138513	1.68
269	scf_8254652:134791-135576	2.44	689	scf_8254677:48898-49205	1.68
270	scf_8254449:1-453	2.44	690	scf_8254610:295417-296299	1.67
271	scf_8254645:961441-961910	2.44	691	scf_8254460:405353-407661	1.67
272	scf_8254397:173704-174617	2.43	692	scf_8254719:283382-284482	1.67
273	scf_8254598:122492-124234	2.42	693	scf_8254712:218046-219781	1.67
274	scf_8254548:189353-189870	2.42	694	scf_8254776:681910-682447	1.67
275	scf_8254677:44468-45113	2.42	695	scf_8254446:385727-388592	1.66
276	scf_8253894:54841-55450	2.42	696	scf_8254653:115715-118092	1.66
277	scf_8254440:143670-144481	2.42	697	scf_8254798:283661-285100	1.66
278	scf_8254459:130948-132016	2.41	698	scf_8254446:190532-191374	1.66
279	scf_8254170:437-1104	2.41	699	scf_8254383:9193-10789	1.66
280	scf_8254483:85437-88885	2.41	700	scf_8254373:262656-264838	1.66
281	scf_8254034:87326-89920	2.41	701	scf_8254552:147318-148079	1.66
282	scf_8254391:100961-101728	2.41	702	scf_8254584:153362-153870	1.66
283	scf_8254691:36337-37214	2.41	703	scf_8254373:450820-451446	1.65

284	scf_8254551:612872-618206	2.40	704	scf_8254697:406641-407516	1.65
285	scf_8254028:50730-51142	2.40	705	scf_8254593:461230-462173	1.65
286	scf_8254513:1372-3140	2.40	706	scf_8254652:503153-504794	1.65
287	scf_8254476:187476-188135	2.40	707	scf_8254475:429845-430254	1.65
288	scf_8254638:168286-170970	2.40	708	scf_8254464:95145-96072	1.65
289	scf_8254371:267484-268308	2.40	709	scf_8254590:773634-775253	1.65
290	scf_8254564:366801-367441	2.40	710	scf_8254431:722985-723341	1.64
291	scf_8254823:227391-228784	2.39	711	scf_8254489:52224-53998	1.64
292	scf_8254470:518571-518796	2.39	712	scf_8254459:108812-109882	1.64
293	scf_8254638:978327-979535	2.39	713	scf_8254597:148054-153269	1.64
294	scf_8254816:41766-42605	2.39	714	scf_8254179:50193-51096	1.63
295	scf_8254667:1576952-1577843	2.38	715	scf_8254747:525748-527467	1.63
296	scf_8254187:67829-68507	2.38	716	scf_8254551:265595-269573	1.63
297	scf_8254470:294045-294440	2.38	717	scf_8254527:130747-132790	1.63
298	scf_8254486:77781-78641	2.38	718	scf_8254638:955851-957052	1.63
299	scf_8254385:67122-69300	2.37	719	scf_8254545:776231-780647	1.63
300	scf_8254763:47739-48218	2.37	720	scf_8254454:91885-92792	1.63
301	scf_8254771:25925-26120	2.37	721	scf_8254542:55538-56413	1.63
302	scf_8254688:414841-415306	2.37	722	scf_8253815:120447-120894	1.63
303	scf_8254475:567207-567846	2.37	723	scf_8254659:500769-501875	1.63
304	scf_8254786:1-936	2.37	724	scf_8254545:860297-860722	1.63
305	scf_8254475:551089-551422	2.36	725	scf_8254199:4165-4702	1.63
306	scf_8253815:423945-424508	2.36	726	scf_8254551:580838-581891	1.63
307	scf_8254565:270953-271896	2.36	727	scf_8254776:702267-703532	1.62
308	scf_8254803:1426989-1427307	2.36	728	scf_8254752:1289031-1289785	1.62
309	scf_8254598:204707-206755	2.36	729	scf_8254659:912070-913538	1.62
310	scf_8254803:133667-135341	2.35	730	scf_8254181:185263-185740	1.62
311	scf_8254528:364559-364884	2.35	731	scf_8254617:79717-80561	1.62
312	scf_8254459:1-658	2.35	732	scf_8254663:140946-142367	1.62
313	scf_8254763:315258-315875	2.35	733	scf_8253930:23975-24119	1.62
314	scf_8254002:11872-13885	2.34	734	scf_8254707:46940-48170	1.62
315	scf_8254644:174086-175430	2.34	735	scf_8254770:4742-5037	1.62
316	scf_8254497:6330-6949	2.34	736	scf_8254719:854358-854860	1.61
317	scf_8254611:363707-364324	2.33	737	scf_8254752:686712-687715	1.61
318	scf_8254811:1100735-1101470	2.33	738	scf_8254543:47147-47962	1.61
319	scf_8254763:229984-231099	2.33	739	scf_8254476:233066-233475	1.61

320	scf_8254716:955713-956144	2.33	740	scf_8254771:24927-25462	1.61
321	scf_8254444:249224-249940	2.33	741	scf_8254431:966599-966844	1.61
322	scf_8254496:110693-112485	2.33	742	scf_8254587:100436-101752	1.61
323	scf_8254764:207767-209224	2.32	743	scf_8254470:505474-507480	1.61
324	scf_8254028:96272-96376	2.32	744	scf_8254373:648710-654805	1.61
325	scf_8254718:130301-131805	2.32	745	scf_8254649:95880-96071	1.61
326	scf_8254607:65059-66093	2.31	746	scf_8254667:784964-785790	1.61
327	scf_8254747:853696-853869	2.31	747	scf_8254592:114104-114940	1.60
328	scf_8254688:707460-708574	2.31	748	scf_8254716:150950-153398	1.60
329	scf_8254564:644771-645773	2.31	749	scf_8254807:123028-126113	1.60
330	scf_8254674:218084-218407	2.31	750	scf_8254406:803092-804848	1.60
331	scf_8254688:634668-635098	2.30	751	scf_8254552:200552-200977	1.60
332	scf_8254095:38074-38715	2.29	752	scf_8254666:397353-398434	1.60
333	scf_8254645:392724-393531	2.29	753	scf_8254459:303132-304040	1.60
334	scf_8254666:80748-83713	2.29	754	scf_8254580:65922-66795	1.60
335	scf_8254777:15309-16335	2.28	755	scf_8254199:101624-102451	1.60
336	scf_8254385:227177-228092	2.28	756	scf_8254564:748146-748638	1.60
337	scf_8253856:16509-16722	2.28	757	scf_8254551:561786-562469	1.60
338	scf_8254369:95964-97658	2.28	758	scf_8254545:815795-817097	1.59
339	scf_8254470:493541-494287	2.28	759	scf_8254362:210737-211432	1.59
340	scf_8254379:516128-516888	2.28	760	scf_8254674:103489-109112	1.59
341	scf_8254545:472611-473366	2.28	761	scf_8254798:229864-233920	1.59
342	scf_8254563:78038-78401	2.27	762	scf_8254464:364534-365013	1.58
343	scf_8254608:71007-71691	2.27	763	scf_8254605:228563-230427	1.58
344	scf_8254489:64373-66864	2.27	764	scf_8254650:868463-869130	1.58
345	scf_8254470:412970-415300	2.26	765	scf_8254479:360149-360602	1.58
346	scf_8254688:531361-531612	2.26	766	scf_8254417:243650-244627	1.58
347	scf_8254650:578930-580379	2.26	767	scf_8254431:585177-585519	1.58
348	scf_8254582:1059599-1061247	2.25	768	scf_8254571:25472-27855	1.57
349	scf_8254371:316980-317926	2.25	769	scf_8254652:379089-379600	1.57
350	scf_8254543:171508-171836	2.25	770	scf_8254679:47905-48628	1.57
351	scf_8254365:291162-291808	2.25	771	scf_8254449:231939-232399	1.57
352	scf_8254652:263270-265224	2.24	772	scf_8254747:236310-236603	1.57
353	scf_8254362:22429-22863	2.24	773	scf_8254609:336487-337018	1.57
354	scf_8254607:427671-428286	2.24	774	scf_8254412:31464-32088	1.57
355	scf_8254786:81166-82593	2.24	775	scf_8254674:139490-143708	1.57

356	scf_8254433:23727-24417	2.24	776	scf_8254788:90913-94032	1.57
357	scf_8254495:389596-390558	2.23	777	scf_8254776:829920-830873	1.57
358	scf_8254719:615204-616432	2.23	778	scf_8254915:1-219	1.57
359	scf_8254786:400531-401552	2.22	779	scf_8254391:286587-287335	1.56
360	scf_8254575:92857-93425	2.22	780	scf_8254683:12217-15976	1.56
361	scf_8254597:124715-125916	2.22	781	scf_8254717:34167-34707	1.56
362	scf_8254579:162604-165866	2.22	782	scf_8254645:319085-321387	1.56
363	scf_8254815:882594-882878	2.22	783	scf_8254617:826084-827203	1.56
364	scf_8254449:165710-166772	2.22	784	scf_8254674:97732-101310	1.56
365	scf_8254594:756288-757235	2.22	785	scf_8254373:695343-696399	1.56
366	scf_8254470:93361-95870	2.22	786	scf_8254185:113978-114561	1.56
367	scf_8254592:401964-402769	2.22	787	scf_8254659:1093455-1094138	1.56
368	scf_8254607:470558-471733	2.21	788	scf_8254691:298902-299355	1.55
369	scf_8254593:623338-627379	2.21	789	scf_8254565:259186-260727	1.55
370	scf_8254095:78037-78869	2.19	790	scf_8254688:25023-26394	1.55
371	scf_8254431:185678-186203	2.19	791	scf_8254594:419302-420467	1.55
372	scf_8254631:107977-108191	2.19	792	scf_8254688:629103-629543	1.55
373	scf_8254824:116641-117230	2.19	793	scf_8254788:44352-46270	1.55
374	scf_8254417:174347-175049	2.19	794	scf_8254711:16923-19086	1.55
375	scf_8254428:210580-211529	2.19	795	scf_8254593:559771-561166	1.55
376	scf_8254682:31419-32262	2.19	796	scf_8254479:110692-111429	1.54
377	scf_8254378:318812-319277	2.18	797	scf_8254548:19739-20476	1.54
378	scf_8254697:1371582-1372077	2.18	798	scf_8254598:211013-216815	1.54
379	scf_8254391:37370-38449	2.18	799	scf_8254814:84744-85199	1.54
380	scf_8254588:190863-191911	2.18	800	scf_8254645:305652-307734	1.54
381	scf_8254667:1679866-1680721	2.18	801	scf_8254757:10989-13030	1.54
382	scf_8254697:1519429-1519815	2.18	802	scf_8253970:9835-12715	1.54
383	scf_8254551:194474-197602	2.17	803	scf_8254010:186053-186745	1.54
384	scf_8254437:93736-93911	2.17	804	scf_8254659:73601-74008	1.54
385	scf_8254284:33129-34047	2.17	805	scf_8254613:2935-4199	1.53
386	scf_8254803:2041078-2042793	2.17	806	scf_8254437:490510-491290	1.53
387	scf_8253891:327546-327598	2.17	807	scf_8254579:154497-155568	1.53
388	scf_8254437:153330-153618	2.16	808	scf_8254548:548588-549133	1.53
389	scf_8254460:302852-303556	2.16	809	scf_8254593:115150-116828	1.53
390	scf_8254176:29916-30540	2.16	810	scf_8254803:425148-429620	1.53
391	scf_8254551:713696-714275	2.16	811	scf_8254563:393951-394042	1.53

392	scf_8254697:368679-369021	2.16	812	scf_8254577:329023-329770	1.53
393	scf_8254503:97759-97976	2.15	813	scf_8254406:75949-77087	1.53
394	scf_8255716:1-256	2.15	814	scf_8254803:1398951-1400644	1.52
395	scf_8254597:498182-498548	2.15	815	scf_8254645:336345-336714	1.52
396	scf_8254587:435464-436554	2.15	816	scf_8254764:80733-81764	1.52
397	scf_8254187:131411-132416	2.15	817	scf_8254697:1040071-1042614	1.52
398	scf_8254788:811889-813515	2.15	818	scf_8254776:304842-305452	1.52
399	scf_8254406:237288-238266	2.14	819	scf_8254043:157412-158436	1.52
400	scf_8254444:123843-125690	2.14	820	scf_8253815:488104-489261	1.52
401	scf_8254600:147200-148231	2.14	821	scf_8254716:856054-857181	1.52
402	scf_8254664:491952-492999	2.14	822	scf_8254719:214124-217179	1.52
403	scf_8254496:41044-41504	2.14	823	scf_8254528:30582-31236	1.52
404	scf_8254788:564168-565059	2.13	824	scf_8254470:341036-342066	1.51
405	scf_8254583:60234-62892	2.13	825	scf_8254611:1-583	1.51
406	scf_8254470:247462-248430	2.13	826	scf_8254085:80990-82241	1.51
407	scf_8254610:436460-438794	2.13	827	scf_8254811:545405-546658	1.51
408	scf_8254486:203004-203338	2.13	828	scf_8254752:663804-664544	1.51
409	scf_8254600:39404-40023	2.12	829	scf_8254505:110479-110750	1.51
410	scf_8254731:183458-184879	2.12	830	scf_8253886:32878-33490	1.51
411	scf_8254636:445793-446281	2.12	831	scf_8253887:72895-74836	1.51
412	scf_8254104:205988-206899	2.12	832	scf_8254104:141575-141905	1.51
413	scf_8254385:113832-116707	2.11	833	scf_8254688:131930-132985	1.51
414	scf_8254362:11946-12909	2.11	834	scf_8254588:208721-209086	1.51
415	scf_8253891:329980-331213	2.11	835	scf_8254549:97886-98461	1.51
416	scf_8253855:32955-35332	2.11	836	scf_8253823:112888-113796	1.51
417	scf_8254047:105175-105628	2.11	837	scf_8254446:483154-483764	1.50
418	scf_8254580:2626-2905	2.11	838	scf_8254362:1-982	1.50
419	scf_8254811:327840-329254	2.11	839	scf_8254104:286722-287030	1.50
420	scf_8254610:392332-393659	2.10	840	scf_8254470:809453-811134	1.50
421	scf_8254803:34848-36701	2.10	841	scf_8254428:328765-329525	1.50
422	scf_8254716:1102950-1104136	2.10	842	scf_8253843:93660-94275	1.50

A.3: Med31ChIP-Seq post-processing files

RACS processed data for the first 1000 (out of 13715) genic regions occupied by that Med31: Refer to Digital Spreadsheet (TET_Med31_MAC_Genome_Genic) for complete results.

#	Gene region	Gene_ID	Note	Enrichment Average	Expression Profile
1	scf_8254728:16008 1-160623	TTHERM_008354 00	Note="hypothetical_protein"	733.76587	Low Expression
2	scf_8254470:17938 5-179675	TTHERM_006303 50	Note="hypothetical_protein"	469.99651	No-Expression
3	scf_8254731:12825 5-128638	TTHERM_000593 068	Note="hypothetical_protein"	413.31855	#N/A
4	scf_8254644:32484 8-325123	TTHERM_000649 439	Note="cytoplasmic_dynein_light_chain_2"	348.29514	#N/A
5	scf_8254528:17250 4-172983	TTHERM_006373 30	Note="hypothetical_protein"	310.36963	Low Expression
6	scf_8254716:76780 9-768598	TTHERM_001050 90	Note="phosducin_(Phd)- like_thioredoxin_(TRX)_domain_protein"	304.73125	High expression
7	scf_8254437:23922 2-240061	TTHERM_004447 30	Note="hypothetical_protein"	291.57709	Low Expression
8	scf_8254558:12251 -12512	TTHERM_001359 453	Note="hypothetical_protein"	285.19112	#N/A
9	scf_8254646:93783 -94172	TTHERM_008183 50	Note="hypothetical_protein"	271.74232	No-Expression
10	scf_8254697:19530 5-195850	TTHERM_000092 857	Note="emopamil-binding_protein"	261.55927	#N/A
11	scf_8254382:12021 6-120404	TTHERM_000947 569	Note="hypothetical_protein"	245.14707	#N/A
12	scf_8254617:61307 4-614048	TTHERM_001947 90	Note="transmembrane_protein_putative"	240.1089	Moderate expression
13	scf_8254545:31273 -31693	TTHERM_000147 532	Note="hypothetical_protein"	232.05626	#N/A
14	scf_8254428:13384 1-134266	TTHERM_005355 50	Note="ribosomal_protein_S6"	219.60123	Moderate expression
15	scf_8254638:90356 3-904117	TTHERM_001918 80	Note="hypothetical_protein"	216.14464	Moderate expression

16	scf_8254594:27873 5-279085	TTHERM_000287 919	Note="hypothetical_protein"	213.07702	#N/A
17	scf_8254371:16838 3-169267	TTHERM_000467 319	Note="endo_alpha-1,4_polygalactosaminidase_precursor"	208.5198	#N/A
18	scf_8254431:11153 8-111795	TTHERM_000312 259	Note="hypothetical_protein"	201.19587	#N/A
19	scf_8254751:23130 -24261	TTHERM_009426 60	Note="transmembrane_protein_putative"	192.00948	Moderate expression
20	scf_8254584:33995 1-340409	TTHERM_006270 50	Note="kazal-type_proteinase_inhibitor_1"	188.99711	No-Expression
21	scf_8254582:61316 1-613898	TTHERM_001363 00	Note="2,3-diketo-5-methylthio-1- phosphopentane_phosphatase_family_protein"	182.79535	Moderate expression
22	scf_8254654:6278- 6555	TTHERM_000503 751	Note="transmembrane_protein_putative"	182.39488	#N/A
23	scf_8254688:37479 2-375115	TTHERM_000295 249	Note="transmembrane_protein_putative"	181.61756	#N/A
24	scf_8254644:22486 5-225155	TTHERM_006491 30	Note="transmembrane_protein_putative"	180.98322	Low Expression
25	scf_8254798:69362 8-694038	TTHERM_000185 487	Note="transmembrane_protein_putative"	180.84239	#N/A
26	scf_8254748:17304 -17783	TTHERM_007050 90	Note="hypothetical_protein"	177.52463	No-Expression
27	scf_8254747:39459 -39686	TTHERM_000140 792	Note="dolichol_phosphate- mannose_biosynthesis_regulatory_protein"	174.16291	#N/A
28	scf_8254464:17886 -18413	TTHERM_000578 441	Note="transmembrane_protein_putative"	173.96288	#N/A
29	scf_8253805:13299 -13508	TTHERM_014315 60	Note="hypothetical_protein"	167.91763	No-Expression
30	scf_8254751:61071 -61571	TTHERM_000942 773	Note="Sex-determining_region_protein"	167.83489	#N/A
31	scf_8254650:61688 2-617586	TTHERM_000775 50	Note="GST45_glutathione_S-transferase_amine- terminal_domain_protein"	165.34929	High expression
32	scf_8254402:60565 -61149	TTHERM_006573 60	Note="trafficking_protein_particle_complex_subunit_6b"	163.16049	Moderate expression
33	scf_8254638:29175 -29558	TTHERM_001883 80	Note="hypothetical_protein"	160.80125	Low Expression

34	scf_8254529:22627 -22941	TTHERM_010541 60	Note="hypothetical_protein"	157.09103	Low Expression
35	scf_8254600:19050 3-191318	TTHERM_002796 60	Note="hypothetical_protein"	154.14284	No-Expression
36	scf_8254708:40499 -41045	TTHERM_001051 841	Note="transmembrane_protein_putative"	151.13873	#N/A
37	scf_8254485:50947 -51426	TTHERM_009894 00	Note="LITAF-like_zinc_ribbon_domain_protein"	150.43058	Low Expression
38	scf_8254401:62384 -63049	TTHERM_001080 487	Note="hypothetical_protein"	149.1556	#N/A
39	scf_8254692:10540 -11007	TTHERM_014999 40	Note="hypothetical_protein"	146.48686	Low Expression
40	scf_8254587:16637 2-166581	TTHERM_003870 60	Note="hypothetical_protein"	146.30019	High expression
41	scf_8254788:48589 7-486517	TTHERM_002633 40	Note="Serine/Threonine_kinase_putative"	144.64156	No-Expression
42	scf_8254811:44201 2-442925	TTHERM_000681 20	Note="hypothetical_protein"	144.22489	High expression
43	scf_8254584:33071 3-331144	TTHERM_006270 70	Note="kazal-type_proteinase_inhibitor_1"	142.98767	Low Expression
44	scf_8254627:13521 -14808	TTHERM_000860 415	Note="hypothetical_protein"	139.48028	#N/A
45	scf_8254811:85021 3-850743	TTHERM_000059 189	Note="hypothetical_protein"	137.63672	#N/A
46	scf_8254671:22011 -22427	TTHERM_000423 264	Note="hypothetical_protein"	135.88615	#N/A
47	scf_8254803:17923 91-1793065	TTHERM_000016 298	Note="transmembrane_protein_putative"	135.65921	#N/A
48	scf_8254636:33388 2-335024	TTHERM_004908 50	Note="UBA/TS-N_domain_protein"	133.52791	Low Expression
49	scf_8254406:21760 6-219217	TTHERM_002452 50	Note="MFS_transporter"	132.7497	High expression
50	scf_8254564:87046 3-870948	TTHERM_002211 70	Note="hypothetical_protein"	132.31108	High expression
51	scf_8254811:86636 5-867792	TTHERM_000591 50	Note="methyltransferase_domain_protein"	131.81191	High expression

52	scf_8254555:49618 4-496657	TTHERM_004118 70	Note="ribose_5-phosphate_isomerase_B"	131.30813	Moderate expression
53	scf_8254379:20956 1-209920	TTHERM_000299 729	Note="transmembrane_protein_putative"	128.45112	#N/A
54	scf_8254798:33851 1-338768	TTHERM_001716 50	Note="hypothetical_protein"	127.3424	Low Expression
55	scf_8254658:20720 5-207717	TTHERM_006755 40	Note="hypothetical_protein"	125.62064	No-Expression
56	scf_8254763:21312 9-213479	TTHERM_000561 209	Note="hypothetical_protein"	122.66158	#N/A
57	scf_8254607:22848 3-228809	TTHERM_004391 10	Note="hypothetical_protein"	122.59693	Low Expression
58	scf_8254611:6865- 7956	TTHERM_004764 30	Note="RimK-like_ATP-grasp_domain_protein"	120.52618	Low Expression
59	scf_8254617:72105 4-722016	TTHERM_000196 029	Note="radial_spoke_head-like_protein"	120.45171	#N/A
60	scf_8254571:19532 5-197265	TTHERM_007981 20	Note="oxalate/formate_antipporter_family_transporter"	120.1477	Moderate expression
61	scf_8254371:26284 3-264161	TTHERM_004676 40	Note="ankyrin_repeat_protein"	118.80588	Moderate expression
62	scf_8254431:11252 06-1126934	TTHERM_005680 00	Note="oxysterol-binding_protein"	118.68755	High expression
63	scf_8254359:93826 -94038	TTHERM_000836 782	Note="hypothetical_protein"	117.36015	#N/A
64	scf_8254457:10278 0-103271	TTHERM_001014 719	Note="hypothetical_protein"	116.39557	#N/A
65	scf_8254823:26847 -27176	TTHERM_000549 619	Note="hypothetical_protein"	113.82207	#N/A
66	scf_8254645:92028 6-920837	TTHERM_001254 80	Note="Mpv17/PMP22_family_protein"	113.34767	Moderate expression
67	scf_8254719:75799 1-759183	TTHERM_000784 329	Note="FAD/FMN-binding_family_oxidoreductase"	113.158	#N/A
68	scf_8253904:17567 4-176339	TTHERM_007261 80	Note="Mpv17/PMP22_family_protein"	112.48961	Moderate expression
69	scf_8254617:17208 8-172522	TTHERM_001936 20	Note="transmembrane_protein_putative"	111.77822	Low Expression

70	scf_8254496:85391 -85810	TTHERM_008580 70	Note="hypothetical_protein"	111.63207	Moderate expression
71	scf_8254667:14154 20-1415581	TTHERM_000036 939	Note="hypothetical_protein"	110.3998	#N/A
72	scf_8254361:69867 -70946	TTHERM_011936 30	Note="ornithine_decarboxylase_antizyme_protein"	109.88012	High expression
73	scf_8254638:19185 -19592	TTHERM_000188 352	Note="hypothetical_protein"	109.04462	#N/A
74	scf_8254047:20336 3-204532	TTHERM_007908 80	Note="linear_amide_C- N_hydrolase_choloylglycine_hydrolase_family_protein"	108.82398	No-Expression
75	scf_8254033:6753- 7352	TTHERM_001362 507	Note="hypothetical_protein"	107.99342	#N/A
76	scf_8254436:11841 4-119034	TTHERM_009390 50	Note="START_domain_protein"	107.80756	Moderate expression
77	scf_8254515:12111 4-121647	TTHERM_007214 80	Note="hypothetical_protein"	106.7874	Low Expression
78	scf_8254779:18725 -19069	TTHERM_010262 00	Note="hypothetical_protein"	105.73315	Low Expression
79	scf_8254577:26059 -27341	TTHERM_006639 90	Note="3- beta_hydroxysteroid_dehydrogenase/isomerase_family_protein "	105.5417	Low Expression
80	scf_8253815:42812 1-428702	TTHERM_004979 90	Note="eukaryotic_translation_initiation_factor_1A"	104.64922	High expression
81	scf_8254667:18614 44-1861776	TTHERM_000390 20	Note="Pre-mRNA-splicing_factor_ini1"	104.00113	Moderate expression
82	scf_8254010:10128 -11307	TTHERM_003255 90	Note="enoyl-CoA_hydratase/delta3,5-delta2,4-dienoyl- CoA_isomerase"	103.91633	High expression
83	scf_8254807:44622 4-448609	TTHERM_005404 10	Note="tetratricopeptide_repeat_protein"	103.77246	Moderate expression
84	scf_8254667:70708 6-707565	TTHERM_000302 10	Note="LC7B_dynein_light_chain_roadblock-type_2_protein"	102.98926	Moderate expression
85	scf_8254010:79469 -79819	TTHERM_003268 10	Note="transmembrane_protein_putative"	102.82052	Low Expression
86	scf_8254594:32666 5-326976	TTHERM_000289 038	Note="transmembrane_protein_putative"	102.65041	#N/A
87	scf_8254410:30643 -31726	TTHERM_001015 921	Note="oxidoreductase_(short-chain_dehydrogenase_family)"	102.17089	#N/A

88	scf_8254719:41883 5-419959	TTHERM_003927 20	Note="SPFH_domain/band_7_family_protein"	100.34725	Moderate expression
89	scf_8254811:44035 3-440667	TTHERM_000681 30	Note="transmembrane_protein_putative"	98.65349	Low Expression
90	scf_8254579:59155 -59450	TTHERM_000666 223	Note="H/ACA_ribonucleoprotein_complex_protein_putative"	98.203129	#N/A
91	scf_8254600:38049 5-382132	TTHERM_002831 00	Note="hypothetical_protein"	95.961556	Moderate expression
92	scf_8254658:32968 3-330048	TTHERM_006758 70	Note="hypothetical_protein"	95.159764	High expression
93	scf_8254555:89004 -89486	TTHERM_004088 40	Note="OsmC_family_protein"	94.072835	High expression
94	scf_8254475:55030 5-551088	TTHERM_000421 119	Note="NUDIX_hydrolase"	93.86598	#N/A
95	scf_8254748:28652 -29107	TTHERM_007051 20	Note="hypothetical_protein"	93.650834	Low Expression
96	scf_8254666:28710 6-288047	TTHERM_004018 50	Note="integral_membrane_protein_DUF6_containing_protein"	93.632817	Moderate expression
97	scf_8254527:15395 -15667	TTHERM_004857 90	Note="hypothetical_protein"	93.626425	High expression
98	scf_8254667:49929 -51104	TTHERM_006541 50	Note="nucleotide-sugar_transporter"	93.209428	High expression
99	scf_8254028:22503 8-225751	TTHERM_006706 80	Note="hypothetical_protein"	92.803567	No-Expression
100	scf_8254665:28551 9-286329	TTHERM_002258 40	Note="C2_domain_protein"	91.813762	No-Expression
101	scf_8254638:82910 7-830093	TTHERM_001916 20	Note="deoxyribose-phosphate_aldolase"	91.505049	High expression
102	scf_8254638:18057 2-180889	TTHERM_000188 839	Note="hypothetical_protein"	91.24858	#N/A
103	scf_8254515:15458 4-154928	TTHERM_007215 60	Note="transmembrane_protein_putative"	90.972764	Low Expression
104	scf_8253974:54625 -54975	TTHERM_012764 50	Note="transmembrane_protein_putative"	90.938611	No-Expression
105	scf_8254505:7249- 8740	TTHERM_010274 90	Note="Di-glucose- binding_within_endoplasmic_reticulum_protein"	90.889904	High expression

106	scf_8253915:6800-7129	TTHERM_006882 80	Note="hypothetical_protein"	90.214423	Moderate expression
107	scf_8254716:3286-3774	TTHERM_000095 358	Note="hypothetical_protein"	90.162902	#N/A
108	scf_8254379:71026-6-710776	TTHERM_003019 40	Note="hypothetical_protein"	89.558826	No-Expression
109	scf_8254548:47853-4-479634	TTHERM_003172 40	Note="hypothetical_protein"	89.319683	Low Expression
110	scf_8254605:37323-0-373577	TTHERM_006216 20	Note="transmembrane_protein_putative"	89.000307	No-Expression
111	scf_8254544:7728-8417	TTHERM_001098 988	Note="transmembrane_protein_putative"	88.198376	#N/A
112	scf_8254371:30530-7-306167	TTHERM_004677 70	Note="poly_polymerase_and_DNA-ligase_Zn-finger_region_family_protein_putative"	87.644558	Moderate expression
113	scf_8254460:58757-59310	TTHERM_003533 79	Note="ML_domain_protein"	87.367407	Moderate expression
114	scf_8254544:67913-68695	TTHERM_010991 70	Note="outer_membrane_protein"	87.049559	Moderate expression
115	scf_8254379:21740-8-218488	TTHERM_002997 60	Note="alpha/beta_fold_hydrolase"	86.415779	Moderate expression
116	scf_8254638:26966-7-270203	TTHERM_000189 048	Note="hypothetical_protein"	86.2978	#N/A
117	scf_8254431:46166-46678	TTHERM_000312 029	Note="hypothetical_protein"	85.899762	#N/A
118	scf_8254006:56842-57389	TTHERM_000944 156	Note="hypothetical_protein"	85.887213	#N/A
119	scf_8254752:10061-16-1006460	TTHERM_001175 60	Note="hypothetical_protein"	85.871341	Low Expression
120	scf_8254403:15133-5-152255	TTHERM_004733 00	Note="hypothetical_protein"	85.640588	High expression
121	scf_8254617:91821-8-920911	TTHERM_001966 20	Note="MIF4G_domain_protein"	85.634803	High expression
122	scf_8254600:36783-7-368643	TTHERM_002830 50	Note="hypothetical_protein"	85.615941	Low Expression
123	scf_8254783:63938-64564	TTHERM_000845 819	Note="transmembrane_protein_putative"	85.316004	#N/A

124	scf_8254760:85582 -86739	TTHERM_010184 50	Note="THD12_SIR2_family_transcriptional_regulator"	85.074388	Low Expression
125	scf_8254377:17741 1-178340	TTHERM_007160 80	Note="Mps_one_binder_kinase_activator-like_protein"	85.070895	Moderate expression
126	scf_8254788:16067 0-160876	TTHERM_002596 30	Note="hypothetical_protein"	84.09006	Low Expression
127	scf_8254466:13070 6-132892	TTHERM_009519 10	Note="chitinase-like_protein_cluster_protein"	83.972331	Moderate expression
128	scf_8254470:42673 2-427010	TTHERM_001178 679	Note="hypothetical_protein"	83.747691	#N/A
129	scf_8254382:12101 9-122929	TTHERM_009475 70	Note="hypothetical_protein"	83.539671	Moderate expression
130	scf_8254813:51762 3-521978	TTHERM_003654 50	Note="Na,H_K_antipporter_P- type_ATPase_alpha_subunit_family_protein"	83.46543	High expression
131	scf_8254459:19469 8-195600	TTHERM_000530 017	Note="transmembrane_protein_putative"	83.212892	#N/A
132	scf_8254807:15302 7-153521	TTHERM_000538 849	Note="hypothetical_protein"	82.726055	#N/A
133	scf_8254803:92016 4-920478	TTHERM_000110 30	Note="transmembrane_protein_putative"	82.652898	Moderate expression
134	scf_8254592:50510 6-510622	TTHERM_007176 40	Note="ABC_transporter_family_protein"	82.241875	Moderate expression
135	scf_8254551:45952 9-461678	TTHERM_003101 30	Note="hypothetical_protein"	81.264668	Moderate expression
136	scf_8254548:69498 1-695928	TTHERM_003188 30	Note="choline-phosphate_cytidyltransferase"	81.070798	Moderate expression
137	scf_8254617:76433 6-764848	TTHERM_000196 169	Note="transmembrane_protein_putative"	80.822402	#N/A
138	scf_8254610:19521 1-195627	TTHERM_003357 30	Note="hypothetical_protein"	80.093556	Low Expression
139	scf_8254582:60334 0-603903	TTHERM_001363 30	Note="methionine-R-sulfoxide_reductase"	79.727925	Moderate expression
140	scf_8254667:18049 05-1805774	TTHERM_000388 30	Note="SNF7_family_protein"	79.53368	Moderate expression
141	scf_8254803:56494 0-565630	TTHERM_000062 50	Note="ABC_transporter_family_protein_putative"	79.307153	Low Expression

142	scf_8254609:43443 2-436555	TTHERM_000442 899	Note="kinase_domain_protein"	79.125311	#N/A
143	scf_8254650:72646 0-726858	TTHERM_000777 90	Note="prefoldin"	78.682996	Moderate expression
144	scf_8254752:12944 81-1297028	TTHERM_001223 40	Note="hypothetical_protein"	78.602051	Moderate expression
145	scf_8254431:39189 7-393368	TTHERM_003129 90	Note="DnaJ_domain_protein"	78.389196	No-Expression
146	scf_8254435:29415 6-294609	TTHERM_006560 40	Note="import_inner_membrane_translocase_subunit_TIM14"	77.643593	Low Expression
147	scf_8254803:55622 1-556733	TTHERM_000062 20	Note="HVA22/TB2/DP1_family_protein"	77.371476	High expression
148	scf_8254378:20625 0-207125	TTHERM_006162 70	Note="UbiE/COQ5_family_methyltransferase"	77.178128	High expression
149	scf_8254654:11371 -12217	TTHERM_005037 70	Note="transmembrane_protein_putative"	77.016184	No-Expression
150	scf_8254666:29207 1-292580	TTHERM_000401 869	Note="hypothetical_protein"	76.258965	#N/A
151	scf_8254593:53118 5-532153	TTHERM_004639 10	Note="oxidoreductase_short_chain_dehydrogenase/reductase_f amily_protein"	75.971656	Moderate expression
152	scf_8254636:48469 8-484889	TTHERM_004912 30	Note="hypothetical_protein"	75.148194	Low Expression
153	scf_8254747:73212 -73817	TTHERM_001409 20	Note="helix_turn-helix_protein"	75.013293	High expression
154	scf_8254814:76125 -78008	TTHERM_004183 80	Note="plasma_platelet-activating_factor_acetylhydrolase"	74.908527	High expression
155	scf_8254051:15335 4-154418	TTHERM_009719 50	Note="transmembrane_protein_putative"	74.798298	Moderate expression
156	scf_8254645:84029 0-840607	TTHERM_001257 50	Note="transmembrane_protein_putative"	74.625747	No-Expression
157	scf_8254545:26541 7-265779	TTHERM_001491 50	Note="hypothetical_protein"	74.617189	Low Expression
158	scf_8254466:13575 0-138029	TTHERM_009519 20	Note="chitinase-like_protein_cluster_protein"	74.070092	High expression
159	scf_8254564:41477 3-415255	TTHERM_002188 99	Note="cytidine_and_deoxycytidylate_deaminase_zinc- binding_region_protein"	73.944696	Low Expression

160	scf_8254617:14530 2-145637	TTHERM_000193 549	Note="transmembrane_protein_putative"	73.724461	#N/A
161	scf_8254747:24940 0-250681	TTHERM_001423 30	Note="RAD51_DNA_repair_RAD51-like_protein"	72.834824	Moderate expression
162	scf_8254465:6665- 6977	TTHERM_001498 869	Note="hypothetical_protein"	72.409706	#N/A
163	scf_8254487:15860 0-159459	TTHERM_005007 40	Note="hypothetical_protein"	72.339089	Moderate expression
164	scf_8254606:99544 -99708	TTHERM_001087 961	Note="hypothetical_protein"	72.067596	#N/A
165	scf_8254582:34048 9-342591	TTHERM_001380 00	Note="DHHC_zinc_finger_protein"	71.304542	Low Expression
166	scf_8254650:12807 -13760	TTHERM_000731 50	Note="hypothetical_protein"	71.115215	Low Expression
167	scf_8254584:18061 6-181578	TTHERM_006283 90	Note="transmembrane_protein_putative"	70.81384	High expression
168	scf_8254609:48018 0-480776	TTHERM_000442 979	Note="transmembrane_protein_putative"	70.682891	#N/A
169	scf_8254302:30932 -31390	TTHERM_001306 871	Note="hypothetical_protein"	70.565801	#N/A
170	scf_8254719:52594 6-526464	TTHERM_000391 439	Note="OsmC_family_protein"	70.440178	#N/A
171	scf_8254665:83651 3-836851	TTHERM_002341 60	Note="transmembrane_protein_putative"	70.384902	No-Expression
172	scf_8254656:61807 -62937	TTHERM_010298 90	Note="hypothetical_protein"	70.375519	No-Expression
173	scf_8254464:60588 -61157	TTHERM_005785 50	Note="transmembrane_protein_putative"	69.701542	High expression
174	scf_8254605:17360 0-174256	TTHERM_000621 168	Note="transmembrane_protein_putative"	69.347221	#N/A
175	scf_8254803:14485 36-1451079	TTHERM_000134 70	Note="ATP-NAD_kinase"	69.10176	Moderate expression
176	scf_8254582:48486 2-486577	TTHERM_001376 50	Note="hypothetical_protein"	69.086683	Moderate expression
177	scf_8254594:72519 0-726869	TTHERM_002921 60	Note="xylulose_kinase"	68.896181	Low Expression

178	scf_8254593:51423 7-515828	TTHERM_004638 50	Note="fatty_acid_hydroxylase_putative"	68.874914	High expression
179	scf_8254716:48985 6-491028	TTHERM_001014 60	Note="peptidase_M76_family_protein"	68.833679	Moderate expression
180	scf_8254107:978- 1355	TTHERM_018088 70	Note="transmembrane_protein_putative"	68.536648	No-Expression
181	scf_8254815:64286 1-643535	TTHERM_002052 00	Note="GST49_glutathione_S-transferase_amine- terminal_domain_protein"	68.422	Moderate expression
182	scf_8254659:67987 4-680500	TTHERM_000474 20	Note="oxidoreductase_short_chain_dehydrogenase/reductase_f amily_protein"	68.02879	Moderate expression
183	scf_8254667:12944 49-1295779	TTHERM_000356 30	Note="brix_domain_protein"	67.84226	Moderate expression
184	scf_8254370:22876 2-229898	TTHERM_007607 70	Note="inhibitor_of_apoptosis-promoting_Bax1_protein"	67.543079	Moderate expression
185	scf_8254776:14507 9-145465	TTHERM_001614 80	Note="CHY_zinc_finger_protein"	67.500968	Moderate expression
186	scf_8254491:14964 2-149893	TTHERM_007871 60	Note="hypothetical_protein"	67.457713	No-Expression
187	scf_8254811:64706 7-647407	TTHERM_000626 80	Note="hypothetical_protein"	67.375186	No-Expression
188	scf_8253915:23987 0-240382	TTHERM_006898 50	Note="CEN3_caltractin"	67.174906	High expression
189	scf_8254803:55060 6-551567	TTHERM_000062 00	Note="rhomboid_protease"	67.036893	Low Expression
190	scf_8254638:90417 9-904597	TTHERM_000191 889	Note="hypothetical_protein"	66.707796	#N/A
191	scf_8254663:49443 -49997	TTHERM_008543 00	Note="Myb-like_DNA-binding_domain_protein"	66.668761	Moderate expression
192	scf_8254551:45068 3-453854	TTHERM_003101 10	Note="hypothetical_protein"	66.637815	Moderate expression
193	scf_8254486:37990 2-380864	TTHERM_005586 00	Note="thioredoxin_domain_protein"	66.569771	Moderate expression
194	scf_8254813:21213 9-212633	TTHERM_003617 80	Note="hypothetical_protein"	66.270693	Low Expression
195	scf_8254686:30477 5-305302	TTHERM_004494 90	Note="hypothetical_protein"	65.778287	Low Expression

196	scf_8254570:55350 -55824	TTHERM_000935 563	Note="hypothetical_protein"	65.628685	#N/A
197	scf_8254752:11110 28-1111608	TTHERM_001208 80	Note="protozoan/cyanobacterial_globin_family_protein"	65.486562	High expression
198	scf_8254666:72920 -73354	TTHERM_003991 90	Note="transmembrane_protein_putative"	65.021722	Moderate expression
199	scf_8254776:58130 2-582158	TTHERM_001582 80	Note="DOMON_domain_protein"	64.129673	Moderate expression
200	scf_8254728:98495 -98917	TTHERM_008352 10	Note="transmembrane_protein_putative"	64.082783	No-Expression
201	scf_8254527:33452 1-335285	TTHERM_004866 10	Note="hypothetical_protein"	64.081771	Moderate expression
202	scf_8254593:44164 6-444267	TTHERM_004636 50	Note="glycosyl_hydrolase_family_31_protein"	64.07086	Low Expression
203	scf_8254476:14123 5-142245	TTHERM_000728 889	Note="patatin_family_phospholipase"	63.697901	#N/A
204	scf_8254470:14026 7-141774	TTHERM_006305 00	Note="CAM1_calmodulin_1"	63.682636	High expression
205	scf_8254667:19933 87-1993821	TTHERM_000403 30	Note="hypothetical_protein"	63.423756	Low Expression
206	scf_8254752:80176 1-802087	TTHERM_000113 149	Note="hypothetical_protein"	63.14108	#N/A
207	scf_8254402:12838 3-129127	TTHERM_006575 30	Note="iron- sulfur_cluster_assembly_scaffold_protein_putative"	63.095425	High expression
208	scf_8254752:13692 04-1369782	TTHERM_001225 00	Note="c11orf73-like_protein_putative"	62.997247	Low Expression
209	scf_8254644:11183 6-112045	TTHERM_000648 879	Note="ion_channel_inhibitory_toxin_protein"	62.97699	#N/A
210	scf_8254584:14699 6-147541	TTHERM_006284 90	Note="hypothetical_protein"	62.296248	Low Expression
211	scf_8254638:95408 4-954731	TTHERM_000192 088	Note="C3HC4_type_(RING_finger)_zinc_finger_protein"	62.251294	#N/A
212	scf_8254776:11331 7-113640	TTHERM_000161 579	Note="hypothetical_protein"	62.027048	#N/A
213	scf_8254380:35727 -35915	TTHERM_001125 183	Note="cytochrome_C_oxidase_copper_chaperone"	61.955796	#N/A

214	scf_8254590:8815-10180	TTHERM_00249570	Note="serine_carboxypeptidase_family_protein"	61.94755	No-Expression
215	scf_8254731:244355-244663	TTHERM_00594400	Note="hypothetical_protein"	61.943297	No-Expression
216	scf_8253823:229470-230380	TTHERM_00678350	Note="CDP-diacylglycerol-inositol_3-phosphatidyltransferase"	61.668167	Moderate expression
217	scf_8254844:121-516	TTHERM_02186720	Note="ABC_transporter_family_protein"	61.575264	No-Expression
218	scf_8254545:908155-908526	TTHERM_000151799	Note="transmembrane_protein_putative"	61.477127	#N/A
219	scf_8254359:98887-99090	TTHERM_00837800	Note="transmembrane_protein_putative"	61.175042	Low Expression
220	scf_8253994:37607-38029	TTHERM_01311260	Note="hypothetical_protein"	61.126762	No-Expression
221	scf_8254665:385071-385793	TTHERM_000227089	Note="hypothetical_protein"	61.090982	#N/A
222	scf_8254564:409964-411931	TTHERM_00218890	Note="Na-dependent_neurotransmitter_transporter"	60.777367	Low Expression
223	scf_8254085:28469-30313	TTHERM_01039810	Note="protein_kinase"	60.517176	Moderate expression
224	scf_8254595:153667-154704	TTHERM_000999150	Note="PLD-like_domain_protein"	60.459942	#N/A
225	scf_8254645:548229-548990	TTHERM_00128470	Note="zinc_finger_C2H2_type_family_protein"	60.443079	Low Expression
226	scf_8254607:302518-303165	TTHERM_00438890	Note="dual_specificity_phosphatase_domain_protein"	60.202393	Moderate expression
227	scf_8253844:85-956	TTHERM_02302910	Note="hypothetical_protein"	60.157926	No-Expression
228	scf_8254666:25619-26986	TTHERM_000398101	Note="hypothetical_protein"	60.147442	#N/A
229	scf_8254431:680894-681559	TTHERM_00313710	Note="HRT1_anaphase-promoting_complex_subunit_11_RING-H2_finger_protein"	60.096458	High expression
230	scf_8254496:59581-59979	TTHERM_00857960	Note="LITAF-like_zinc_ribbon_domain_protein"	59.933487	No-Expression
231	scf_8254010:476100-476570	TTHERM_00329800	Note="guanylate_kinase"	59.453269	High expression

232	scf_8254611:67585 -67926	TTHERM_000476 562	Note="hypothetical_protein"	59.225496	#N/A
233	scf_8254665:79416 0-794501	TTHERM_002330 20	Note="transmembrane_protein_putative"	59.204178	No-Expression
234	scf_8254470:73812 5-738481	TTHERM_004337 40	Note="hypothetical_protein"	58.69085	No-Expression
235	scf_8254104:21906 7-220579	TTHERM_007134 00	Note="transmembrane_protein_putative"	58.675424	Moderate expression
236	scf_8254610:37588 4-376264	TTHERM_003383 50	Note="hypothetical_protein"	58.606151	Low Expression
237	scf_8254380:31510 -32256	TTHERM_011251 70	Note="hypothetical_protein"	58.52305	Low Expression
238	scf_8254377:20788 5-208217	TTHERM_007161 70	Note="hypothetical_protein"	58.52177	Low Expression
239	scf_8254592:13831 3-139416	TTHERM_005817 70	Note="Nop53_(60S_ribosomal_biogenesis)_protein"	58.494513	Moderate expression
240	scf_8254582:78400 1-784732	TTHERM_000134 817	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	58.349793	#N/A
241	scf_8254379:90976 -92552	TTHERM_002983 60	Note="hypothetical_protein"	58.26291	High expression
242	scf_8254370:16336 3-164871	TTHERM_007606 20	Note="LETM1-like_protein"	58.19122	High expression
243	scf_8254645:55058 -56572	TTHERM_001298 90	Note="cytochrome_P450_family_monooxygenase"	58.179601	High expression
244	scf_8254479:16343 0-164221	TTHERM_006229 60	Note="RING_finger_and_CHY_zinc_finger_protein"	58.170919	Moderate expression
245	scf_8254163:124- 1161	TTHERM_021887 20	Note="PLD-like_domain_protein"	57.987381	Low Expression
246	scf_8254563:23009 9-231062	TTHERM_006003 59	Note="transmembrane_protein_putative"	57.955263	#N/A
247	scf_8254667:11769 46-1178748	TTHERM_000353 60	Note="ABC-2_family_transporter_protein"	57.851079	Moderate expression
248	scf_8254630:82354 -82764	TTHERM_007588 10	Note="LITAF-like_zinc_ribbon_domain_protein"	57.755714	No-Expression
249	scf_8254486:42539 1-425852	TTHERM_005586 90	Note="transmembrane_protein_putative"	57.701559	No-Expression

250	scf_8254043:54868 -57634	TTHERM_007809 10	Note="Serine/Threonine_kinase_domain_protein"	57.607452	Low Expression
251	scf_8254092:99647 -100164	TTHERM_011050 30	Note="UbiE/COQ5_methyltransferase_putative"	57.480378	Moderate expression
252	scf_8254545:54547 9-546157	TTHERM_001498 59	Note="thioredoxin"	57.472226	Moderate expression
253	scf_8254652:25723 9-258152	TTHERM_003787 90	Note="scavenger_mRNA_decapping_enzyme_carboxy-term- binding_protein"	56.970365	Moderate expression
254	scf_8254479:27010 7-271144	TTHERM_006243 10	Note="DTD1_D-tyrosyl-tRNA(Tyr)_deacylase"	56.912157	Moderate expression
255	scf_8254486:28737 4-288372	TTHERM_005583 60	Note="kinase_domain_protein"	56.802463	Moderate expression
256	scf_8254487:12225 8-122607	TTHERM_000499 609	Note="hypothetical_protein"	56.720862	#N/A
257	scf_8253856:19685 -21718	TTHERM_012847 30	Note="oxalate/formate_antipporter_family_transporter"	56.667373	Moderate expression
258	scf_8253887:70479 -70895	TTHERM_006469 40	Note="hypothetical_protein"	56.628926	No-Expression
259	scf_8254644:24055 0-240867	TTHERM_006491 50	Note="hypothetical_protein"	56.396121	High expression
260	scf_8253856:16723 -17969	TTHERM_012847 20	Note="hypothetical_protein"	56.388015	No-Expression
261	scf_8254449:21319 2-215316	TTHERM_007736 90	Note="adenylosuccinate_synthetase_protein"	56.35932	High expression
262	scf_8254590:15675 7-159916	TTHERM_002509 70	Note="glutamate-cysteine_ligase_catalytic_subunit"	56.229841	Moderate expression
263	scf_8254607:42514 0-425379	TTHERM_004375 80	Note="hypothetical_protein"	56.163528	Low Expression
264	scf_8254367:26147 2-262767	TTHERM_007565 30	Note="transmembrane_protein_putative"	56.075786	High expression
265	scf_8254811:22441 7-225139	TTHERM_000696 40	Note="hypothetical_protein"	56.0427	No-Expression
266	scf_8254752:72929 5-730790	TTHERM_001129 40	Note="TGL7_Ab-hydrolase_associated_lipase_region_protein"	55.927537	Low Expression
267	scf_8254470:10137 52-1014030	TTHERM_004299 80	Note="hypothetical_protein"	55.908796	Low Expression

268	scf_8254617:89376 7-894096	TTHERM_000196 549	Note="hypothetical_protein"	55.674043	#N/A
269	scf_8254365:19296 6-197174	TTHERM_006859 00	Note="multidrug_resistance_protein- like_transporter_family_ABC_domain_protein"	55.623401	High expression
270	scf_8254697:17302 46-1730641	TTHERM_000355 919	Note="hypothetical_protein"	55.557252	#N/A
271	scf_8254495:84885 -85562	TTHERM_005640 80	Note="Erv1/Alr_family_protein"	55.535488	Moderate expression
272	scf_8254590:62370 -62789	TTHERM_002497 40	Note="hypothetical_protein"	55.378896	Low Expression
273	scf_8254645:99935 8-1001125	TTHERM_001252 80	Note="PDD1_chromodomain_protein"	55.365943	No-Expression
274	scf_8254371:45646 6-457422	TTHERM_004691 90	Note="carrier_protein"	55.109822	Moderate expression
275	scf_8254659:70877 6-709363	TTHERM_000475 30	Note="YRS1_predicted_protein"	54.936874	Low Expression
276	scf_8254644:25493 8-256362	TTHERM_006492 20	Note="histidine_acid_phosphatase_family_protein"	54.806102	High expression
277	scf_8254610:51955 9-521195	TTHERM_003398 20	Note="methylene- tetrahydrofolate_dehydrogenase/cyclohydrolase_NADP- binding_domain_protein"	54.630087	Moderate expression
278	scf_8254489:20575 -20982	TTHERM_010680 60	Note="LITAF-like_zinc_ribbon_domain_protein"	54.525426	No-Expression
279	scf_8254659:15858 8-160152	TTHERM_000461 80	Note="tyrosine/nicotianamine_aminotransferase"	54.491493	High expression
280	scf_8254468:4899- 5484	TTHERM_001044 376	Note="transmembrane_protein_putative"	54.478769	#N/A
281	scf_8254415:57768 -59822	TTHERM_012053 70	Note="phosphoribulokinase/uridine_kinase"	54.476476	High expression
282	scf_8254464:36501 4-366018	TTHERM_000580 366	Note="ABC_transporter_family_protein"	54.428434	#N/A
283	scf_8254638:61666 0-617067	TTHERM_001909 70	Note="hypothetical_protein"	54.193287	Low Expression
284	scf_8254813:27626 0-277627	TTHERM_003619 10	Note="serine_carboxypeptidase_family_protein"	54.137291	Low Expression
285	scf_8254645:71010 6-710849	TTHERM_001270 40	Note="hypothetical_protein"	54.091636	Moderate expression

286	scf_8254610:41349 2-414118	TTHERM_003385 20	Note="hypothetical_protein"	53.632595	Low Expression
287	scf_8254716:82563 3-827261	TTHERM_001052 80	Note="casein_kinase"	53.598874	Moderate expression
288	scf_8254659:15180 03-1519120	TTHERM_000526 20	Note="OLE2_fatty_acid_desaturase"	53.557339	Moderate expression
289	scf_8254679:27395 -28353	TTHERM_012277 60	Note="hypothetical_protein"	53.23571	High expression
290	scf_8254377:20500 1-205285	TTHERM_007161 60	Note="hypothetical_protein"	53.063818	No-Expression
291	scf_8254747:52746 8-528519	TTHERM_001450 80	Note="zinc-binding_dehydrogenase_family_oxidoreductase"	53.040364	High expression
292	scf_8254776:78413 0-784380	TTHERM_001567 20	Note="hypothetical_protein"	52.860199	No-Expression
293	scf_8254373:25646 1-257470	TTHERM_002390 80	Note="SNF7_family_protein"	52.802276	Moderate expression
294	scf_8254644:34376 4-345325	TTHERM_006494 70	Note="AN1-like_zinc_finger_protein"	52.71362	High expression
295	scf_8254241:99831 -103453	TTHERM_010924 80	Note="hypothetical_protein"	52.635897	Moderate expression
296	scf_8254638:40079 -40249	TTHERM_001884 20	Note="hypothetical_protein"	52.564361	Low Expression
297	scf_8254010:63589 2-636512	TTHERM_003320 90	Note="GST24_glutathione_S-transferase_amine- terminal_domain_protein"	52.477189	High expression
298	scf_8254385:43953 5-440248	TTHERM_000533 999	Note="hypothetical_protein"	52.224192	#N/A
299	scf_8254803:72740 1-727757	TTHERM_000086 10	Note="transmembrane_protein_putative"	52.197542	Low Expression
300	scf_8254645:64171 9-642584	TTHERM_001272 50	Note="hypothetical_protein"	51.894677	High expression
301	scf_8254010:29533 -30573	TTHERM_003256 30	Note="integral_membrane_protein_DUF6_containing_protein"	51.882295	Moderate expression
302	scf_8254752:14460 57-1447221	TTHERM_001236 80	Note="senescence-associated_protein"	51.864042	Moderate expression
303	scf_8254638:89567 2-896782	TTHERM_001918 60	Note="diphthine-ammonia_ligase"	51.837526	Moderate expression

304	scf_8254397:47169-47636	TTHERM_008235 20	Note="TKI1_kazal-type_serine_protease_inhibitor_domain_protein"	51.738751	Moderate expression
305	scf_8254406:224260-226276	TTHERM_002452 70	Note="MFS_transporter"	51.718182	High expression
306	scf_8254645:642707-642895	TTHERM_001272 40	Note="hypothetical_protein"	51.594331	No-Expression
307	scf_8254650:278105-279160	TTHERM_000757 90	Note="hypothetical_protein"	51.46259	Low Expression
308	scf_8254650:755415-756203	TTHERM_000788 80	Note="hypothetical_protein"	51.33836	Low Expression
309	scf_8254139:2076-3944	TTHERM_015071 10	Note="Sec1_family_protein"	51.306486	Low Expression
310	scf_8255775:585-924	TTHERM_002653 414	Note="zinc_finger_LSD1_subclass_family_protein_putative"	51.234984	#N/A
311	scf_8254422:201366-201932	TTHERM_006916 00	Note="hypothetical_protein"	51.131663	Low Expression
312	scf_8254475:430255-432310	TTHERM_004208 30	Note="small-conductance_calcium-activated_potassium_channel_protein"	51.074379	Moderate expression
313	scf_8254548:632501-633622	TTHERM_003186 50	Note="hypothetical_protein"	51.054759	Moderate expression
314	scf_8254454:54563-55678	TTHERM_011061 20	Note="NTG1_base_excision_DNA_repair_protein_HhH-GPD_family_protein"	50.862283	Low Expression
315	scf_8254747:524574-525747	TTHERM_001450 70	Note="zinc-binding_dehydrogenase_family_oxidoreductase"	50.749548	Low Expression
316	scf_8254393:16649-16855	TTHERM_012890 60	Note="hypothetical_protein"	50.702621	High expression
317	scf_8254515:223777-225114	TTHERM_007217 40	Note="hypothetical_protein"	50.64394	Moderate expression
318	scf_8254043:70430-73198	TTHERM_007808 50	Note="FEN1_hypothetical_protein"	50.622867	High expression
319	scf_8254593:197269-197730	TTHERM_004630 00	Note="UPF0678_fatty_acid-binding_protein-like_protein_putative"	50.310037	Moderate expression
320	scf_8254043:26640-27460	TTHERM_007809 90	Note="S54_family_peptidase"	50.200864	Low Expression
321	scf_8254811:1310034-1311086	TTHERM_000560 10	Note="hypothetical_protein"	50.080826	High expression

322	scf_8254440:19688 5-198211	TTHERM_008256 10	Note="endonuclease/exonuclease/phosphatase_family_protein"	50.04844	Moderate expression
323	scf_8254548:23852 4-239132	TTHERM_000316 519	Note="hypothetical_protein"	50.03259	#N/A
324	scf_8254637:82273 -83538	TTHERM_008269 40	Note="melanocyte_proliferating_protein_putative"	49.824511	High expression
325	scf_8254422:1662- 2732	TTHERM_006910 90	Note="integral_membrane_protein_DUF6_containing_protein"	49.649001	High expression
326	scf_8254590:13529 8-136077	TTHERM_002509 29	Note="transmembrane_protein_putative"	49.642017	No-Expression
327	scf_8254610:52800 3-529043	TTHERM_003398 50	Note="delta-6_fatty_acid_desaturase"	49.568577	High expression
328	scf_8254446:51483 6-515869	TTHERM_004949 00	Note="transmembrane_protein_putative"	49.462281	Moderate expression
329	scf_8254811:58237 -59640	TTHERM_000709 90	Note="ALG7_UDP-N-acetylglucosamine-dolichyl- phosphate_N-acetylglucosaminophosphotransferase_putative"	49.420206	Moderate expression
330	scf_8254555:54436 9-545127	TTHERM_004120 00	Note="hypothetical_protein"	49.327318	Low Expression
331	scf_8254470:11975 23-1197858	TTHERM_013385 30	Note="transmembrane_protein_putative"	49.228554	No-Expression
332	scf_8254548:68795 6-688961	TTHERM_003188 00	Note="MAP_kinase_organizer_1"	49.165013	Moderate expression
333	scf_8254428:41283 1-414788	TTHERM_005373 00	Note="ribose-phosphate_pyrophosphokinase"	49.099095	High expression
334	scf_8254362:16776 9-170011	TTHERM_007234 10	Note="Serine/Threonine_kinase_domain_protein"	49.044275	Moderate expression
335	scf_8254548:97978 -98844	TTHERM_003161 60	Note="ribosomal_protein_L27"	49.008969	High expression
336	scf_8254711:11411 8-115101	TTHERM_010344 40	Note="transmembrane_protein_putative"	48.99634	Low Expression
337	scf_8254045:1936- 2349	TTHERM_017877 90	Note="hypothetical_protein"	48.919596	Low Expression
338	scf_8254667:17092 61-1710196	TTHERM_000376 40	Note="hydrogenosomal_ATP/ADP_carrier_protein"	48.886389	High expression
339	scf_8254431:42775 8-429412	TTHERM_003130 80	Note="acyl-CoA_dehydrogenase"	48.829074	Moderate expression

340	scf_8254654:44012 1-441992	TTHERM_005186 90	Note="hypothetical_protein"	48.800355	Moderate expression
341	scf_8253908:12043 7-120922	TTHERM_009039 00	Note="hypothetical_protein"	48.723427	No-Expression
342	scf_8254757:75435 -76358	TTHERM_010051 40	Note="hypothetical_protein"	48.523845	Moderate expression
343	scf_8254359:19561 1-195931	TTHERM_000840 099	Note="transmembrane_protein_putative"	48.400827	#N/A
344	scf_8254665:84039 8-841118	TTHERM_002341 70	Note="transmembrane_protein_putative"	48.243858	No-Expression
345	scf_8254497:34610 -36142	TTHERM_012076 60	Note="CDC2_cyclin-dependent_kinase- like_Serine/Threonine_kinase_family_protein"	48.172086	High expression
346	scf_8254361:61954 -63876	TTHERM_011935 90	Note="transmembrane_protein_putative"	48.170565	Moderate expression
347	scf_8254610:53924 5-540312	TTHERM_003398 90	Note="XYPPX_repeat_protein"	48.153638	Moderate expression
348	scf_8254644:18130 3-182743	TTHERM_006490 10	Note="ubiquitin-associated_domain_protein_putative"	47.957373	High expression
349	scf_8254650:11561 01-1156871	TTHERM_000079 949	Note="hypothetical_protein"	47.901301	#N/A
350	scf_8253891:23010 5-230518	TTHERM_000683 149	Note="transmembrane_protein_putative"	47.879281	#N/A
351	scf_8254659:51984 7-522072	TTHERM_000469 80	Note="membrane_bound_O-acyltransferase_MBOAT_protein"	47.837882	Moderate expression
352	scf_8254748:27198 9-272291	TTHERM_000709 669	Note="hypothetical_protein"	47.818233	#N/A
353	scf_8254479:33964 -35382	TTHERM_012650 30	Note="kinase_domain_protein"	47.74463	Moderate expression
354	scf_8254565:35958 0-360496	TTHERM_004715 40	Note="hypothetical_protein"	47.694146	Moderate expression
355	scf_8254043:96989 -97243	TTHERM_000780 769	Note="hypothetical_protein"	47.692899	#N/A
356	scf_8254476:87153 -87467	TTHERM_000727 749	Note="mitotic-spindle_organizing_protein"	47.621902	#N/A
357	scf_8254431:70731 2-707575	TTHERM_003138 20	Note="hypothetical_protein"	47.479717	Moderate expression

358	scf_8254370:21170 1-213226	TTHERM_000760 739	Note="kinase_domain_protein_putative"	47.464287	#N/A
359	scf_8253815:49380 5-495150	TTHERM_000498 219	Note="hypothetical_protein"	47.437731	#N/A
360	scf_8254487:22779 4-229309	TTHERM_005009 60	Note="Serine/Threonine_kinase_domain_protein"	47.407541	Moderate expression
361	scf_8254645:10877 4-109757	TTHERM_001297 80	Note="lipase_family_protein"	47.329875	Low Expression
362	scf_8254659:49749 0-498676	TTHERM_000469 30	Note="hypothetical_protein"	47.286797	Low Expression
363	scf_8254811:41553 -42797	TTHERM_000710 50	Note="hypothetical_protein"	47.269112	Moderate expression
364	scf_8254776:31848 2-321221	TTHERM_001609 30	Note="calmodulin-domain_kinase"	47.236133	Moderate expression
365	scf_8254587:10175 3-102341	TTHERM_003849 10	Note="CEN1_basal_body_centrin_protein"	47.156022	High expression
366	scf_8254370:13923 9-140147	TTHERM_007605 50	Note="zinc-binding_alcohol_dehydrogenase_family_protein"	47.116231	Low Expression
367	scf_8254475:10486 3-105420	TTHERM_004199 60	Note="hypothetical_protein"	47.114716	High expression
368	scf_8254645:73797 4-738393	TTHERM_000126 959	Note="transmembrane_protein_putative"	47.094966	#N/A
369	scf_8254378:11148 4-112316	TTHERM_006159 60	Note="dienelactone_hydrolase_family_protein"	47.017062	Low Expression
370	scf_8254659:91851 0-920989	TTHERM_000490 70	Note="CIT2_peroxisomal_(Per)_like_protein"	46.995703	High expression
371	scf_8254686:35478 5-357808	TTHERM_004496 30	Note="DRH35_type_III_restriction_enzyme_res_subunit_famil y_protein"	46.952377	High expression
372	scf_8254034:7808- 8161	TTHERM_001123 846	Note="hypothetical_protein"	46.69076	#N/A
373	scf_8254552:10718 6-108646	TTHERM_007942 40	Note="hypothetical_protein"	46.660532	No-Expression
374	scf_8254776:81696 8-818551	TTHERM_001556 40	Note="WD_domain_G-beta_repeat_protein"	46.438288	Low Expression
375	scf_8254512:37460 -38819	TTHERM_001228 989	Note="aspartyl/asparaginyl_beta-hydroxylase_family_protein"	46.252924	#N/A

376	scf_8254104:16100 9-161455	TTHERM_007132 20	Note="hypothetical_protein"	46.240749	No-Expression
377	scf_8254803:18944 83-1896320	TTHERM_000166 00	Note="methionine_aminopeptidase"	46.231522	High expression
378	scf_8254582:10363 94-1039341	TTHERM_001311 80	Note="alix_V-shaped_domain_binding_to_HIV_protein"	46.188814	High expression
379	scf_8254645:99850 3-998847	TTHERM_000125 279	Note="hypothetical_protein"	46.170539	#N/A
380	scf_8254587:42950 0-431456	TTHERM_000389 699	Note="transmembrane_protein_putative"	46.064492	#N/A
381	scf_8254610:41261 2-412989	TTHERM_003385 10	Note="cytochrome_b5-like_heme/steroid-binding_domain_protein"	46.005382	High expression
382	scf_8254638:9808- 10740	TTHERM_001883 20	Note="transmembrane_protein_putative"	45.926152	Low Expression
383	scf_8254667:83658 9-838481	TTHERM_000305 90	Note="poly(ADP-ribose)_glycohydrolase"	45.923346	High expression
384	scf_8254799:18554 6-186190	TTHERM_008130 00	Note="chitin_recognition_protein_putative"	45.912863	No-Expression
385	scf_8254803:10885 44-1089686	TTHERM_000115 40	Note="enoyl-(acyl_carrier)_reductase"	45.887603	High expression
386	scf_8254402:26781 8-268777	TTHERM_000658 889	Note="hypothetical_protein"	45.797306	#N/A
387	scf_8254551:59919 3-599579	TTHERM_003105 60	Note="transmembrane_protein_putative"	45.625048	High expression
388	scf_8254798:60066 4-601051	TTHERM_001851 50	Note="hypothetical_protein"	45.474122	Low Expression
389	scf_8254565:34387 9-344643	TTHERM_004714 90	Note="small_guanosine_triphosphatase_family_Ras_family_protein"	45.469578	Moderate expression
390	scf_8254686:24749 2-248547	TTHERM_004492 40	Note="divergent_CRAL/TRIO_domain_protein"	45.437991	Low Expression
391	scf_8254104:53381 -54016	TTHERM_007118 60	Note="hypothetical_protein"	45.414951	Moderate expression
392	scf_8254645:56883 5-569137	TTHERM_000128 379	Note="hypothetical_protein"	45.340652	#N/A
393	scf_8254647:35096 -35797	TTHERM_001323 739	Note="hypothetical_protein"	45.266073	#N/A

394	scf_8254437:29291 1-295391	TTHERM_000444 859	Note="electron_transfer_flavoprotein_ubiquinone_oxidoreductase"	45.241975	#N/A
395	scf_8254444:51249 -52046	TTHERM_000584 621	Note="transmembrane_protein_putative"	45.102329	#N/A
396	scf_8254582:91368 4-914142	TTHERM_001334 90	Note="transmembrane_protein_putative"	45.068901	Low Expression
397	scf_8254761:97614 -98351	TTHERM_010506 00	Note="OsmC_family_protein"	45.067299	Moderate expression
398	scf_8254752:15290 44-1530963	TTHERM_001238 60	Note="hypothetical_protein"	45.010261	Low Expression
399	scf_8254028:31216 6-312582	TTHERM_000670 869	Note="transmembrane_protein_putative"	44.995126	#N/A
400	scf_8254555:55784 7-560450	TTHERM_004120 50	Note="MFS_transporter"	44.822623	Moderate expression
401	scf_8254362:63300 -65516	TTHERM_007230 90	Note="TGL1_Ab-hydrolase_associated_lipase_region_protein"	44.77316	Moderate expression
402	scf_8254824:4105- 4440	TTHERM_001012 064	Note="hypothetical_protein"	44.770036	#N/A
403	scf_8254719:36666 9-367567	TTHERM_003929 00	Note="hypothetical_protein"	44.677754	Moderate expression
404	scf_8254659:74701 8-748876	TTHERM_000476 00	Note="WD40_domain_protein"	44.600326	Moderate expression
405	scf_8254611:94174 -94578	TTHERM_000476 689	Note="transmembrane_protein_putative"	44.533164	#N/A
406	scf_8254798:60593 0-606695	TTHERM_001851 70	Note="hypothetical_protein"	44.516051	No-Expression
407	scf_8254085:66512 -67180	TTHERM_010419 20	Note="RAB58_small_guanosine_triphosphatase_family_Ras_family_protein"	44.48586	Moderate expression
408	scf_8254811:22260 1-223260	TTHERM_000696 50	Note="transmembrane_protein_putative"	44.479495	No-Expression
409	scf_8254667:57244 6-574050	TTHERM_000028 908	Note="hypothetical_protein"	44.470695	#N/A
410	scf_8254803:21478 44-2149177	TTHERM_000243 00	Note="hypothetical_protein"	44.458154	Low Expression
411	scf_8254528:26322 5-264402	TTHERM_006375 90	Note="transmembrane_protein_putative"	44.432472	High expression

412	scf_8254764:20720 0-207766	TTHERM_006333 50	Note="ciliary_basal_body-associated_B9_protein"	44.262381	Moderate expression
413	scf_8254362:71905 -72624	TTHERM_007231 20	Note="von_willebrand_factor_type_A_domain_protein"	44.067359	Low Expression
414	scf_8254798:52968 2-530674	TTHERM_001810 00	Note="transmembrane_protein_putative"	44.063238	No-Expression
415	scf_8254570:37087 -37494	TTHERM_009354 80	Note="transmembrane_protein_putative"	44.018454	No-Expression
416	scf_8254545:70637 2-708857	TTHERM_001513 10	Note="hypothetical_protein"	43.993021	High expression
417	scf_8254496:47271 -48023	TTHERM_008579 20	Note="LITAF-like_zinc_ribbon_domain_protein"	43.962718	Moderate expression
418	scf_8254545:34559 -35718	TTHERM_001475 60	Note="PRE1_proteasome_subunit_beta_type_protein"	43.941781	High expression
419	scf_8254104:22625 9-226645	TTHERM_007134 30	Note="hypothetical_protein"	43.938663	Low Expression
420	scf_8254385:37635 4-376812	TTHERM_005328 00	Note="cytochrome_C_domain_protein"	43.829036	No-Expression
421	scf_8254579:12257 6-123665	TTHERM_006664 30	Note="inosine-uridine_preferring_nucleoside_hydrolase"	43.801979	Moderate expression
422	scf_8254776:78464 5-785145	TTHERM_000156 699	Note="hypothetical_protein"	43.792351	#N/A
423	scf_8254638:16137 9-161558	TTHERM_001887 50	Note="hypothetical_protein"	43.701179	Low Expression
424	scf_8254812:2821- 2976	TTHERM_001170 539	Note="hypothetical_protein"	43.617768	#N/A
425	scf_8254590:81313 1-813888	TTHERM_002571 80	Note="kinase_binding_protein"	43.586067	Moderate expression
426	scf_8254760:82749 -83126	TTHERM_010184 30	Note="hypothetical_protein"	43.560662	No-Expression
427	scf_8253886:33491 -34604	TTHERM_001001 223	Note="PLP-dependent_L-allo-threonine_aldolase"	43.495911	#N/A
428	scf_8253915:23491 4-235669	TTHERM_006898 20	Note="ER17_exonuclease"	43.422419	Low Expression
429	scf_8254428:18310 7-184684	TTHERM_005356 80	Note="fatty_acid_desaturase"	43.412616	High expression

430	scf_8254600:53721 4-538706	TTHERM_002835 30	Note="PDD2_programmed_DNA_degradation_2_protein"	43.337419	No-Expression
431	scf_8254446:17146 1-171847	TTHERM_004927 70	Note="hypothetical_protein"	43.318455	Moderate expression
432	scf_8254777:11802 6-118892	TTHERM_003748 70	Note="carbonic_anhydrase"	43.31544	Moderate expression
433	scf_8254650:99067 1-991096	TTHERM_000079 488	Note="hypothetical_protein"	43.294946	#N/A
434	scf_8253894:33406 -35187	TTHERM_008412 60	Note="hypothetical_protein"	43.282592	Moderate expression
435	scf_8254593:65174 8-652368	TTHERM_011610 00	Note="ETC_complex_I_subunit_motif_protein"	43.255611	High expression
436	scf_8254577:28640 8-287337	TTHERM_006603 20	Note="hypothetical_protein"	43.23848	Moderate expression
437	scf_8254716:80357 0-804696	TTHERM_001052 30	Note="CAAX_protease_self-immunity_protein"	43.202603	Moderate expression
438	scf_8254034:83217 -84098	TTHERM_011240 10	Note="14-3-3_family_protein_14-3-3_beta/zeta"	43.19642	Moderate expression
439	scf_8253815:12835 8-129957	TTHERM_004971 20	Note="cytochrome_C_oxidase_assembly_protein_COX15_putative"	43.129285	High expression
440	scf_8254822:20191 -21060	TTHERM_008665 00	Note="glycosyl_hydrolase_family_16_laminarinase"	42.945951	No-Expression
441	scf_8254659:10287 71-1029959	TTHERM_000494 50	Note="RPN11_Mov34/MPN/PAD-1_family:_proteasomal_regulatory_protein_rpn11_and_signalosome_complex_protein"	42.855038	High expression
442	scf_8254652:14141 3-144043	TTHERM_003785 40	Note="PX_domain_protein"	42.821977	Low Expression
443	scf_8254555:49457 3-495282	TTHERM_004118 60	Note="hypothetical_protein"	42.814305	Moderate expression
444	scf_8254051:50723 -51640	TTHERM_009716 90	Note="diacylglycerol_acyltransferase"	42.735366	High expression
445	scf_8254659:35065 7-352715	TTHERM_000465 50	Note="hypothetical_protein"	42.653169	High expression
446	scf_8254622:8326- 9103	TTHERM_011946 50	Note="histone-like_transcription_factor_(CBF/NFY)_and_archaeal_histone"	42.640241	Moderate expression
447	scf_8254446:71773 -74563	TTHERM_004924 60	Note="TKU80_Ku70/Ku80_beta-barrel_domain_protein"	42.590179	Moderate expression

448	scf_8254194:15345 -17610	TTHERM_010195 90	Note="acyl-CoA_dehydrogenase"	42.560662	High expression
449	scf_8254811:10983 91-1100734	TTHERM_000585 50	Note="molecular_chaperone_DnaK"	42.497638	High expression
450	scf_8254754:16233 1-162675	TTHERM_000607 169	Note="transmembrane_protein_putative"	42.405067	#N/A
451	scf_8254431:38368 2-385389	TTHERM_003129 50	Note="pfkB_family_carbohydrate_kinase"	42.386724	Moderate expression
452	scf_8254565:52596 1-526260	TTHERM_004720 09	Note="transmembrane_protein_putative"	42.339766	High expression
453	scf_8254707:16204 3-164318	TTHERM_008959 70	Note="adenosine/AMP_deaminase"	42.296894	High expression
454	scf_8253811:11945 3-120340	TTHERM_005263 30	Note="aldo/keto_reductase_family_oxidoreductase"	42.272719	No-Expression
455	scf_8254459:37492 0-376074	TTHERM_005305 40	Note="peroxisomal_targeting_signal_2_receptor"	42.224758	Moderate expression
456	scf_8254654:25195 -25683	TTHERM_000503 802	Note="transmembrane_protein_putative"	42.188433	#N/A
457	scf_8254803:18768 67-1877256	TTHERM_000016 579	Note="transmembrane_protein_putative"	42.148556	#N/A
458	scf_8254422:63592 -64113	TTHERM_000691 228	Note="hypothetical_protein"	42.122963	#N/A
459	scf_8254664:29503 7-295760	TTHERM_005225 40	Note="nucleoside_triphosphate_pyrophosphohydrolase"	42.100555	Moderate expression
460	scf_8254139:22- 1049	TTHERM_015071 00	Note="TRM6_tRNA_(guanine(37)-N(1))-methyltransferase"	42.045667	No-Expression
461	scf_8254783:10176 0-102857	TTHERM_008469 40	Note="transmembrane_protein_putative"	42.003377	Low Expression
462	scf_8254379:24204 2-242971	TTHERM_000299 839	Note="hypothetical_protein"	41.977059	#N/A
463	scf_8254428:50848 -51828	TTHERM_005353 20	Note="transmembrane_protein_putative"	41.949194	High expression
464	scf_8254636:14376 6-144833	TTHERM_004894 60	Note="cell_surface_immobilization_antigen"	41.909107	No-Expression
465	scf_8254776:36146 0-362662	TTHERM_001608 20	Note="FAD/FMN-binding_family_oxidoreductase"	41.875492	Low Expression

466	scf_8254798:97626 6-978868	TTHERM_001872 70	Note="AMP-binding_enzyme_family_protein"	41.848479	High expression
467	scf_8254697:18195 02-1819837	TTHERM_003571 20	Note="transmembrane_protein_putative"	41.8363	No-Expression
468	scf_8254813:26335 7-265096	TTHERM_003618 60	Note="serine_carboxypeptidase_family_protein"	41.832826	Low Expression
469	scf_8254667:78166 9-782352	TTHERM_000304 40	Note="PRE3_proteasome_subunit_beta_type_protein"	41.704668	High expression
470	scf_8254597:24802 3-248256	TTHERM_000343 969	Note="hypothetical_protein"	41.614057	#N/A
471	scf_8254650:93648 1-937219	TTHERM_000793 20	Note="Fop_carboxy-terminal_duplication_domain_protein"	41.601514	High expression
472	scf_8254549:55297 -56307	TTHERM_007759 40	Note="transmembrane_protein_putative"	41.505102	Moderate expression
473	scf_8254598:23577 0-236147	TTHERM_007536 30	Note="transmembrane_protein_putative"	41.498998	No-Expression
474	scf_8254688:34802 4-349937	TTHERM_002953 40	Note="RRM59_polyadenylate-binding_protein"	41.445336	High expression
475	scf_8254719:84571 0-846702	TTHERM_007820 90	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	41.371139	Low Expression
476	scf_8254605:20458 0-205585	TTHERM_006212 40	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	41.246677	Moderate expression
477	scf_8253930:14598 1-146595	TTHERM_009415 80	Note="PfpI_family_intracellular_protease"	41.244695	High expression
478	scf_8254610:37190 1-372194	TTHERM_003383 30	Note="hypothetical_protein"	41.244415	Moderate expression
479	scf_8254671:9093- 9290	TTHERM_004222 40	Note="transmembrane_protein_putative"	41.238869	High expression
480	scf_8254565:33394 7-335644	TTHERM_004714 70	Note="EF_hand_protein"	41.23791	High expression
481	scf_8254582:68491 9-685843	TTHERM_001361 00	Note="D-alanyl-D-alanine_dipeptidase"	41.225524	Low Expression
482	scf_8254803:10154 45-1015789	TTHERM_000113 70	Note="transmembrane_protein_putative"	41.209962	High expression
483	scf_8254495:78407 -81903	TTHERM_005640 60	Note="basic_region_leucine_zipper_protein"	41.186815	High expression

484	scf_8254564:66209 1-663636	TTHERM_002206 00	Note="hypothetical_protein"	41.130068	High expression
485	scf_8254551:59641 3-596775	TTHERM_003105 40	Note="transmembrane_protein_putative"	41.122195	High expression
486	scf_8253880:33936 -37559	TTHERM_010555 20	Note="ATPase_histidine_kinase- _DNA_gyrase_B"	40.959596	Moderate expression
487	scf_8254187:50209 -50913	TTHERM_006974 20	Note="50S_ribosomal_protein_L24"	40.753852	High expression
488	scf_8254588:39817 -40395	TTHERM_008493 20	Note="ADP-ribosylation_factor(Arf)/Arf- like_(Arl)_small_GTPase_family_protein"	40.75315	Moderate expression
489	scf_8254645:53741 0-537733	TTHERM_000128 499	Note="hypothetical_protein"	40.552748	#N/A
490	scf_8254645:88154 8-882507	TTHERM_001255 90	Note="carrier_protein"	40.434735	High expression
491	scf_8254549:13938 0-139743	TTHERM_000777 133	Note="hypothetical_protein"	40.425673	#N/A
492	scf_8254781:8043- 9248	TTHERM_012533 60	Note="hypothetical_protein"	40.423459	Moderate expression
493	scf_8254410:13274 1-133439	TTHERM_010162 30	Note="dual_specificity_phosphatase_domain_protein"	40.404091	Low Expression
494	scf_8253815:81286 -82227	TTHERM_004970 00	Note="protein_kinase"	40.261243	Low Expression
495	scf_8254644:17022 7-171399	TTHERM_006489 90	Note="hypothetical_protein"	40.192581	Low Expression
496	scf_8254584:15603 4-159307	TTHERM_006284 40	Note="NRK34_plant_dual- specificity_MAP_kinase_kinase_family_domain_protein"	40.1096	Moderate expression
497	scf_8254650:57769 8-578929	TTHERM_000774 50	Note="glycoprotein"	39.920589	High expression
498	scf_8254594:42046 8-420983	TTHERM_002893 00	Note="swarming_motility_protein_ybiA"	39.894988	Low Expression
499	scf_8253815:18939 6-189908	TTHERM_004973 10	Note="transmembrane_protein_putative"	39.857776	High expression
500	scf_8254464:22781 6-228169	TTHERM_005790 59	Note="transmembrane_protein_putative"	39.827886	#N/A
501	scf_8254284:17983 7-181957	TTHERM_007652 80	Note="branched-chain-amino-acid_aminotransferase"	39.74948	High expression

502	scf_8254590:24533 5-246597	TTHERM_002511 90	Note="metal_cation_transporter_ZIP_family_protein"	39.727261	Moderate expression
503	scf_8253994:12686 -14170	TTHERM_013101 80	Note="hypothetical_protein"	39.671823	Moderate expression
504	scf_8254777:27834 5-282540	TTHERM_003724 60	Note="FLP10_phospholipid-translocating_P- type_ATPase_flippase_family_protein"	39.615258	High expression
505	scf_8254823:36901 5-369692	TTHERM_005436 10	Note="hypothetical_protein"	39.612394	Moderate expression
506	scf_8254582:18275 3-183463	TTHERM_000138 428	Note="transmembrane_protein_putative"	39.57159	#N/A
507	scf_8254752:84914 9-851975	TTHERM_001132 60	Note="MFS_transporter"	39.433323	Moderate expression
508	scf_8253886:31867 -32877	TTHERM_001001 221	Note="hypothetical_protein"	39.425736	#N/A
509	scf_8254528:25645 3-257325	TTHERM_006375 50	Note="divergent_CRAL/TRIO_domain_protein"	39.418468	Low Expression
510	scf_8254659:10197 78-1020088	TTHERM_000049 417	Note="hypothetical_protein"	39.351113	#N/A
511	scf_8253803:5650- 7518	TTHERM_016249 00	Note="Sec1_family_protein"	39.346241	High expression
512	scf_8253886:12652 8-128954	TTHERM_010015 20	Note="MHCK/EF2_kinase_domain_protein"	39.342313	Moderate expression
513	scf_8253930:67489 -68936	TTHERM_009414 10	Note="WD_domain_G-beta_repeat_protein"	39.280164	Moderate expression
514	scf_8254555:52905 -53563	TTHERM_000408 751	Note="hypothetical_protein"	39.18884	#N/A
515	scf_8254588:3211- 4509	TTHERM_008492 00	Note="mevalonate_diphosphate_decarboxylase"	39.179974	High expression
516	scf_8254378:26212 0-263352	TTHERM_006164 40	Note="transmembrane_protein_putative"	39.124714	No-Expression
517	scf_8254152:14530 -15036	TTHERM_015122 30	Note="hypothetical_protein"	39.057061	Low Expression
518	scf_8254811:12045 43-1206963	TTHERM_000582 60	Note="aldo/keto_reductase_family_oxidoreductase"	38.960108	High expression
519	scf_8254595:13232 6-133249	TTHERM_009990 90	Note="SLP1_scramblase_family_protein"	38.956014	Moderate expression

520	scf_8254475:45232 5-455864	TTHERM_004208 80	Note="TPA7_Na,H/K_antiporter_P- type_ATPase_alpha_subunit_family_protein"	38.924249	Low Expression
521	scf_8254382:88913 -89581	TTHERM_009474 60	Note="ADP-ribosylation_factor-like_protein"	38.905708	Low Expression
522	scf_8254476:56570 -57698	TTHERM_007276 60	Note="methyltransferase-like_protein"	38.882587	Moderate expression
523	scf_8254803:12133 42-1213743	TTHERM_000118 70	Note="hypothetical_protein"	38.848338	Low Expression
524	scf_8254476:14279 3-146447	TTHERM_007288 90	Note="patatin-like_phospholipase_family_protein"	38.841898	Moderate expression
525	scf_8254550:92397 -92789	TTHERM_000923 053	Note="hypothetical_protein"	38.698066	#N/A
526	scf_8254181:13411 6-136495	TTHERM_009295 20	Note="UDP-sugar_pyrophosphorylase"	38.661774	High expression
527	scf_8254550:75128 -76665	TTHERM_009230 20	Note="pyridoxal-phosphate- dependent_enzyme_family_protein"	38.610532	Moderate expression
528	scf_8254638:46507 1-466168	TTHERM_001895 10	Note="zinc-binding_alcohol_dehydrogenase_family_protein"	38.603594	No-Expression
529	scf_8253894:88519 -89067	TTHERM_008423 90	Note="transmembrane_protein_putative"	38.599621	Low Expression
530	scf_8254590:75345 3-755955	TTHERM_002569 80	Note="zinc_finger_lsd1_subclass_family_protein"	38.598908	High expression
531	scf_8254072:7364- 9779	TTHERM_013457 70	Note="hypothetical_protein"	38.585195	High expression
532	scf_8254564:46383 5-464802	TTHERM_002190 50	Note="MXR1_peptide_methionine_sulfoxide_reductase_msrA "	38.481027	Moderate expression
533	scf_8254429:11060 -12815	TTHERM_010430 80	Note="CYC4_amine-terminal_domain_cyclin"	38.480454	Moderate expression
534	scf_8254659:40383 8-405754	TTHERM_000466 80	Note="RIO1_Serine/Threonine_kinase_RIO1"	38.426855	Moderate expression
535	scf_8254552:49256 -50140	TTHERM_007940 50	Note="hypothetical_protein"	38.39057	Moderate expression
536	scf_8254505:12087 6-121676	TTHERM_010277 30	Note="hypothetical_protein"	38.344862	Low Expression
537	scf_8254487:20214 2-203881	TTHERM_005008 90	Note="WRS1_tryptophanyl-tRNA_synthetase_protein"	38.286561	High expression

538	scf_8254716:35839 6-358816	TTHERM_001011 40	Note="RRM1_splicing_factor"	38.282954	Low Expression
539	scf_8254667:47170 6-473286	TTHERM_000287 00	Note="serine_carboxypeptidase_S28_family_protein"	38.036269	High expression
540	scf_8254587:52809 2-529171	TTHERM_003900 10	Note="hypothetical_protein"	37.999831	Moderate expression
541	scf_8254194:11552 9-116436	TTHERM_010208 40	Note="jacalin-like_lectin_domain_protein"	37.989583	Low Expression
542	scf_8254495:21653 7-216977	TTHERM_000564 499	Note="hypothetical_protein"	37.856746	#N/A
543	scf_8254671:30770 1-309293	TTHERM_004259 50	Note="UTP18_WD_domain_G-beta_repeat_protein"	37.823141	Moderate expression
544	scf_8254545:35834 2-359157	TTHERM_001493 70	Note="hypothetical_protein"	37.816249	No-Expression
545	scf_8254728:16267 2-163756	TTHERM_008354 10	Note="hypothetical_protein"	37.801626	Low Expression
546	scf_8254665:82492 8-825266	TTHERM_002331 20	Note="transmembrane_protein_putative"	37.675433	No-Expression
547	scf_8254798:68911 7-689809	TTHERM_001854 30	Note="transmembrane_protein"	37.663589	Moderate expression
548	scf_8254653:27848 4-282505	TTHERM_006435 50	Note="Na,H/K_antiporter_P- type_ATPase_alpha_subunit_family_protein"	37.604427	High expression
549	scf_8254607:19905 5-201057	TTHERM_004391 50	Note="hypothetical_protein"	37.595057	#N/A
550	scf_8254607:78483 -79412	TTHERM_004405 30	Note="casein_kinase"	37.55384	Moderate expression
551	scf_8254671:42998 5-431033	TTHERM_004262 40	Note="hypothetical_protein"	37.513225	High expression
552	scf_8254610:43119 9-434932	TTHERM_003396 10	Note="RPN1_26S_proteasome_regulatory_subunit"	37.491032	High expression
553	scf_8254719:34964 2-350001	TTHERM_000392 949	Note="hypothetical_protein"	37.477862	#N/A
554	scf_8253807:5869- 7486	TTHERM_015132 60	Note="BBC49_FAD-dependent_pyridine_nucleotide- disulfide_oxidoreductase"	37.304883	High expression
555	scf_8254594:26766 4-268254	TTHERM_002868 91	Note="hypothetical_protein"	37.231032	Low Expression

556	scf_8254605:35915 4-361789	TTHERM_000621 609	Note="histone_deacetylase_family_protein"	37.203998	#N/A
557	scf_8254629:48834 -49703	TTHERM_001081 693	Note="transmembrane_protein_putative"	37.148954	#N/A
558	scf_8254697:24747 3-248133	TTHERM_000091 689	Note="transmembrane_protein_putative"	37.09868	#N/A
559	scf_8254788:49127 2-491691	TTHERM_002633 60	Note="immobilization_antigen_LA_putative"	36.854057	Low Expression
560	scf_8254811:12983 06-1300895	TTHERM_000560 50	Note="E3_ubiquitin-protein_ligase_HECTD1"	36.852231	Moderate expression
561	scf_8254716:72585 3-726677	TTHERM_001049 60	Note="transmembrane_protein_putative"	36.800759	Moderate expression
562	scf_8254788:41328 7-414799	TTHERM_002632 10	Note="TMS_membrane_protein/tumor_differentially_protein"	36.779917	Moderate expression
563	scf_8254786:30997 6-311275	TTHERM_005543 20	Note="tyrosine/nicotianamine_aminotransferase"	36.564653	Moderate expression
564	scf_8253886:14902 8-149792	TTHERM_010015 70	Note="adenylate_kinase"	36.527431	High expression
565	scf_8254527:33711 4-338591	TTHERM_004866 30	Note="CTH105_papain_family_cysteine_protease"	36.511722	Moderate expression
566	scf_8254814:40147 6-403019	TTHERM_004156 00	Note="peptidase_C13_family_protein"	36.49992	No-Expression
567	scf_8254365:14189 9-142615	TTHERM_006847 10	Note="hypothetical_protein"	36.488531	High expression
568	scf_8254691:44119 9-442584	TTHERM_003229 90	Note="hypothetical_protein"	36.487546	Moderate expression
569	scf_8254788:16185 2-162058	TTHERM_002596 40	Note="hypothetical_protein"	36.478454	Low Expression
570	scf_8254667:14156 34-1416962	TTHERM_000369 40	Note="hypothetical_protein"	36.474829	Moderate expression
571	scf_8254665:72021 8-722068	TTHERM_002278 60	Note="patatin_family_phospholipase"	36.407533	Low Expression
572	scf_8254010:52053 -52472	TTHERM_000326 682	Note="hypothetical_protein"	36.395819	#N/A
573	scf_8254377:45780 -46561	TTHERM_007157 20	Note="3-methyladenine_DNA_glycosylase"	36.354448	Low Expression

574	scf_8254607:40214 2-403094	TTHERM_004376 40	Note="membrane-bound_O-acyltransferase_family_MBOAT_protein"	36.340003	Moderate expression
575	scf_8254371:50320 2-503774	TTHERM_004693 20	Note="hypothetical_protein"	36.333467	Low Expression
576	scf_8254385:46605 4-470218	TTHERM_005340 80	Note="glycogen/starch/alpha-glucan_phosphorylase"	36.330994	High expression
577	scf_8253891:28563 0-287492	TTHERM_006833 10	Note="copine_protein"	36.328787	Moderate expression
578	scf_8253887:23253 3-235055	TTHERM_006473 10	Note="Serine/Threonine_kinase_domain_protein"	36.305904	High expression
579	scf_8254697:73451 7-735525	TTHERM_000085 439	Note="hypothetical_protein"	36.305232	#N/A
580	scf_8253974:51458 -53976	TTHERM_012764 40	Note="3-prime-5-prime_exonuclease"	36.261301	Moderate expression
581	scf_8254028:33066 9-330890	TTHERM_000670 919	Note="transmembrane_protein_putative"	36.24605	#N/A
582	scf_8254370:94946 -95833	TTHERM_007604 30	Note="ADP/ATP_transporter_on_adenylate_translocase"	36.164666	Low Expression
583	scf_8254224:43135 -43882	TTHERM_012434 80	Note="glycosyl_hydrolase_family_25_protein"	36.148562	Moderate expression
584	scf_8254777:58232 5-585056	TTHERM_003707 10	Note="peroxisomal_acyl-CoA_oxidase"	36.140941	High expression
585	scf_8254666:26838 0-268940	TTHERM_000400 809	Note="hypothetical_protein"	36.099458	#N/A
586	scf_8254527:31541 2-317010	TTHERM_000486 589	Note="hypothetical_protein"	36.0851	#N/A
587	scf_8254823:14204 7-142232	TTHERM_000548 329	Note="transmembrane_protein_putative"	36.075115	#N/A
588	scf_8254365:27888 3-279861	TTHERM_006860 60	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	36.051861	Moderate expression
589	scf_8254464:20134 7-201996	TTHERM_000578 999	Note="ubiquitin-conjugating_enzyme"	36.006512	#N/A
590	scf_8254565:65389 -65679	TTHERM_000470 563	Note="EF-hand_protein"	36.000617	#N/A
591	scf_8254437:47758 6-478164	TTHERM_004463 90	Note="ADP-ribosylation_factor(Arf)/Arf-like_(Arl)_small_GTPase_family_protein"	35.997386	High expression

592	scf_8254551:63995 7-640802	TTHERM_003107 40	Note="thioesterase_family_protein"	35.9763	Moderate expression
593	scf_8254475:45898 9-462543	TTHERM_004209 00	Note="Na,H/K_antiporter_P- type_ATPase_alpha_subunit_family_protein"	35.964851	Moderate expression
594	scf_8254667:11723 30-1174243	TTHERM_000353 50	Note="ABC-2_family_transporter_protein"	35.851274	Low Expression
595	scf_8254788:25916 5-259500	TTHERM_002618 50	Note="hypothetical_protein"	35.840378	Low Expression
596	scf_8253815:29927 1-300618	TTHERM_000497 689	Note="hypothetical_protein"	35.809342	#N/A
597	scf_8254396:56623 -58782	TTHERM_011086 20	Note="APM1A_clathrin_adapter_protein_AP-1_mu_subunit"	35.787829	High expression
598	scf_8254708:19920 -20987	TTHERM_010518 00	Note="tetraspanin_family_protein"	35.702848	Moderate expression
599	scf_8254665:83168 9-832006	TTHERM_002341 40	Note="hypothetical_protein"	35.701867	Low Expression
600	scf_8254667:19808 89-1982027	TTHERM_000392 80	Note="ankyrin_repeat_protein"	35.6862	Low Expression
601	scf_8254788:78679 3-791676	TTHERM_002651 00	Note="ATPase_P- type_(transporting)_HAD_superfamily_protein"	35.639225	High expression
602	scf_8254665:43798 5-438569	TTHERM_002271 90	Note="hypothetical_protein"	35.633541	No-Expression
603	scf_8254460:33189 8-333950	TTHERM_003487 20	Note="kinase_domain_protein"	35.62463	Moderate expression
604	scf_8253823:16246 7-164784	TTHERM_006782 10	Note="transmembrane_protein_putative"	35.591548	Low Expression
605	scf_8254395:47585 4-477302	TTHERM_004593 40	Note="hypothetical_protein"	35.551327	Moderate expression
606	scf_8254783:88291 -90399	TTHERM_008459 00	Note="tetratricopeptide_repeat_protein"	35.542011	Moderate expression
607	scf_8254610:33197 2-333015	TTHERM_003382 00	Note="aldo/keto_reductase_family_oxidoreductase"	35.522709	High expression
608	scf_8254814:39554 3-396622	TTHERM_004156 40	Note="alpha/beta_fold_hydrolase"	35.442296	Moderate expression
609	scf_8254719:24474 -25057	TTHERM_000396 969	Note="hypothetical_protein"	35.413692	#N/A

610	scf_8254754:20127 5-201628	TTHERM_000607 289	Note="hypothetical_protein"	35.412931	#N/A
611	scf_8254803:23023 -23376	TTHERM_000000 60	Note="zinc_finger_C2H2_type_family_protein"	35.395004	No-Expression
612	scf_8254010:12324 1-124725	TTHERM_003269 09	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	35.367828	Low Expression
613	scf_8254176:13069 -13634	TTHERM_012805 60	Note="hypothetical_protein"	35.345911	No-Expression
614	scf_8254470:69724 0-698248	TTHERM_004338 40	Note="EF-hand_protein"	35.326735	Moderate expression
615	scf_8254487:47861 0-480377	TTHERM_005026 10	Note="THD6_class_I_histone_deacetylase_family_protein"	35.23458	Moderate expression
616	scf_8254486:33367 8-334553	TTHERM_005584 80	Note="endoplasmic_reticulum_membrane_protein"	35.184124	Low Expression
617	scf_8254395:17245 9-173262	TTHERM_000455 598	Note="Mpv17/pmp22_family_protein"	35.181766	#N/A
618	scf_8254075:20447 -20947	TTHERM_001262 811	Note="hypothetical_protein"	35.179967	#N/A
619	scf_8254367:21031 2-212551	TTHERM_007563 70	Note="hemolysin-III-like_protein"	35.174163	Moderate expression
620	scf_8254751:14396 4-145522	TTHERM_009429 70	Note="protein_phosphatase_2C_containing_protein"	35.159167	High expression
621	scf_8254565:49111 8-493192	TTHERM_004719 20	Note="ubiquitin_carboxyl-terminal_hydrolase"	35.155445	High expression
622	scf_8254688:56221 8-563808	TTHERM_002948 60	Note="amine-terminal_domain_cyclin"	35.149312	Moderate expression
623	scf_8254665:81593 7-816275	TTHERM_002331 00	Note="transmembrane_protein_putative"	35.106267	No-Expression
624	scf_8254590:18815 1-189328	TTHERM_002510 90	Note="peroxisomal_biogenesis_factor_11"	35.04721	Moderate expression
625	scf_8254658:19209 -19889	TTHERM_006719 70	Note="RAB51_small_guanosine_triphosphatase_family_Ras_f amily_protein"	35.040874	Low Expression
626	scf_8254824:17023 -19545	TTHERM_010121 00	Note="MFS_transporter"	34.938287	Low Expression
627	scf_8254650:9598- 9990	TTHERM_000731 40	Note="hypothetical_protein"	34.93266	No-Expression

628	scf_8254664:36697 7-368320	TTHERM_005227 60	Note="ribose-phosphate_pyrophosphokinase"	34.919646	Moderate expression
629	scf_8254485:90843 -91501	TTHERM_009895 00	Note="hypothetical_protein"	34.915982	Low Expression
630	scf_8254028:46941 -47954	TTHERM_006702 09	Note="peptidyl-tRNA_hydrolase"	34.849645	High expression
631	scf_8254571:21289 8-214886	TTHERM_007981 50	Note="oxalate/formate_antiporter_family_transporter"	34.844169	Moderate expression
632	scf_8254803:83664 2-837016	TTHERM_000009 877	Note="transmembrane_protein_putative"	34.807872	#N/A
633	scf_8254667:17969 82-1797758	TTHERM_000388 00	Note="histidine_phosphatase_family_(branch_protein_1)"	34.765272	Moderate expression
634	scf_8254587:19806 3-203027	TTHERM_003881 60	Note="MFS_transporter"	34.738411	Low Expression
635	scf_8254667:11840 56-1185858	TTHERM_000353 90	Note="lipid_transporter_family_ABC_domain_protein"	34.729683	Moderate expression
636	scf_8254736:40950 -41183	TTHERM_007339 90	Note="hypothetical_protein"	34.717143	No-Expression
637	scf_8254582:70453 9-705844	TTHERM_001360 40	Note="C21orf59_putative"	34.70316	Moderate expression
638	scf_8254719:42874 4-431251	TTHERM_000392 689	Note="cyclic_nucleotide-binding_domain_protein"	34.695715	#N/A
639	scf_8253887:34426 6-347289	TTHERM_006475 00	Note="hydantoinase/oxoprolinase_amine- terminal_region_protein"	34.662464	High expression
640	scf_8254181:49984 -50469	TTHERM_009308 00	Note="zinc_finger_LSD1_subclass_family_protein"	34.660005	Low Expression
641	scf_8254654:33144 5-332044	TTHERM_000516 450	Note="glutathione_S-transferase_amine- terminal_domain_protein"	34.607732	#N/A
642	scf_8254446:22100 2-221373	TTHERM_004928 70	Note="transmembrane_protein_putative"	34.566618	Low Expression
643	scf_8254688:22327 3-224148	TTHERM_002956 70	Note="embryo-abundant_protein_putative"	34.54565	Moderate expression
644	scf_8254815:73707 1-737756	TTHERM_000209 449	Note="hypothetical_protein"	34.530495	#N/A
645	scf_8254671:17192 -17389	TTHERM_004232 60	Note="transmembrane_protein_putative"	34.419115	Low Expression

646	scf_8254764:18957 9-189980	TTHERM_006333 39	Note="hypothetical_protein"	34.391881	#N/A
647	scf_8254610:12815 7-131073	TTHERM_003345 90	Note="SacI-like_domain_protein"	34.387766	High expression
648	scf_8254666:12514 9-126570	TTHERM_003994 00	Note="hypothetical_protein"	34.374288	Moderate expression
649	scf_8254718:21574 6-218706	TTHERM_008093 80	Note="NRK31_plant_dual- specificity_MAP_kinase_kinase_family_domain_protein"	34.364694	Moderate expression
650	scf_8254515:18858 9-189975	TTHERM_007216 40	Note="2-oxoglutarate/malate_carrier_protein"	34.349552	Low Expression
651	scf_8254653:10406 8-104889	TTHERM_006411 19	Note="RIB43A_protein"	34.348618	High expression
652	scf_8254564:23561 8-238159	TTHERM_002184 10	Note="FRS2_phenylalanyl-tRNA_synthetase_beta_chain"	34.341689	High expression
653	scf_8254667:42946 0-432607	TTHERM_000285 80	Note="SPT5_transcription_elongation_factor_spt5"	34.315717	High expression
654	scf_8254548:13870 5-139166	TTHERM_003162 10	Note="mapeg_family_protein"	34.30323	Low Expression
655	scf_8254605:38095 0-381513	TTHERM_006216 50	Note="transmembrane_protein_putative"	34.288082	No-Expression
656	scf_8254665:30424 1-305442	TTHERM_002258 90	Note="porphyromonas-type_peptidyl-arginine_deiminase"	34.25762	Low Expression
657	scf_8254028:49275 -50729	TTHERM_006702 10	Note="PTH3_WD_domain_G-beta_repeat_protein"	34.242773	Low Expression
658	scf_8254010:51389 9-514291	TTHERM_000329 909	Note="cell_surface_immobilization_antigen"	33.999084	#N/A
659	scf_8254600:18936 2-190080	TTHERM_002796 50	Note="hypothetical_protein"	33.997696	No-Expression
660	scf_8254637:10928 4-109853	TTHERM_000827 039	Note="transmembrane_protein_putative"	33.980501	#N/A
661	scf_8254555:14469 9-145250	TTHERM_000408 989	Note="transmembrane_protein_putative"	33.942522	#N/A
662	scf_8254446:1215- 1911	TTHERM_004923 00	Note="hypothetical_protein"	33.922874	Low Expression
663	scf_8254440:15031 6-153478	TTHERM_008254 60	Note="RAD4_DNA_repair_protein_Rad4"	33.904352	Moderate expression

664	scf_8254491:24799 4-248368	TTHERM_000787 423	Note="hypothetical_protein"	33.883589	#N/A
665	scf_8254563:31689 9-317396	TTHERM_006006 40	Note="hypothetical_protein"	33.88238	No-Expression
666	scf_8254716:10722 60-1074105	TTHERM_001069 10	Note="C3HC4_type_(RING_finger)_zinc_finger_protein"	33.880162	Low Expression
667	scf_8254617:37184 1-372302	TTHERM_000194 179	Note="transmembrane_protein_putative"	33.853201	#N/A
668	scf_8254397:12557 9-127959	TTHERM_008238 10	Note="CTH37_papain_family_cysteine_protease"	33.826652	High expression
669	scf_8254104:10363 9-106771	TTHERM_007119 60	Note="hypothetical_protein"	33.78232	Moderate expression
670	scf_8254528:20586 -22020	TTHERM_006369 00	Note="delta-aminolevulinic_acid_dehydratase"	33.666653	High expression
671	scf_8254549:18461 4-185961	TTHERM_007772 80	Note="transmembrane_protein_putative"	33.65668	Moderate expression
672	scf_8254593:57701 3-577231	TTHERM_001215 028	Note="transmembrane_protein_putative"	33.610574	#N/A
673	scf_8254431:82552 2-827531	TTHERM_005719 00	Note="alpha/beta_fold_hydrolase"	33.579268	Moderate expression
674	scf_8254395:44759 5-448025	TTHERM_004592 60	Note="thioredoxin"	33.562915	High expression
675	scf_8254470:76964 8-771429	TTHERM_004336 10	Note="PCI-domain_protein"	33.522359	High expression
676	scf_8254671:35474 4-355856	TTHERM_004260 80	Note="Von_willebrand_factor_type_A_(vWA)_domain_was_o riginally_protein"	33.521494	Low Expression
677	scf_8254659:48094 2-481340	TTHERM_000468 80	Note="hypothetical_protein"	33.500046	Low Expression
678	scf_8254485:14768 9-150442	TTHERM_009916 20	Note="DRH5_DEAD/DEAH-box_helicase"	33.47663	High expression
679	scf_8254611:28139 1-281831	TTHERM_000479 100	Note="hypothetical_protein"	33.46854	#N/A
680	scf_8254667:14497 1-145138	TTHERM_006539 90	Note="hypothetical_protein"	33.466811	Low Expression
681	scf_8254556:13021 5-130589	TTHERM_009587 00	Note="transmembrane_protein_putative"	33.418021	Low Expression

682	scf_8254468:12095 7-121889	TTHERM_010447 70	Note="STOP_protein"	33.397346	Moderate expression
683	scf_8254798:78756 2-788842	TTHERM_001857 10	Note="transmembrane_protein_putative"	33.371889	Low Expression
684	scf_8254682:77981 -78832	TTHERM_009271 50	Note="integral_membrane_protein"	33.340535	No-Expression
685	scf_8254010:81295 -82541	TTHERM_003268 20	Note="alpha/beta_hydrolase_family_protein"	33.277212	Moderate expression
686	scf_8254752:61474 -62526	TTHERM_005771 90	Note="hypothetical_protein"	33.262344	Low Expression
687	scf_8254587:28804 9-290070	TTHERM_003884 10	Note="PA26_p53-induced_protein_(sestrin)_protein"	33.241284	Moderate expression
688	scf_8254664:46978 6-470547	TTHERM_000522 989	Note="hypothetical_protein"	33.218169	#N/A
689	scf_8254803:10357 48-1039021	TTHERM_000114 40	Note="peptidase_family_M49_protein"	33.197979	High expression
690	scf_8254788:74083 0-741552	TTHERM_002648 60	Note="transmembrane_protein_putative"	33.181492	Moderate expression
691	scf_8254663:16059 -17140	TTHERM_008542 30	Note="phospholipase_carboxylesterase_family_protein"	33.138741	Moderate expression
692	scf_8254638:54508 3-546782	TTHERM_001907 50	Note="Prp19-binding_domain_splicing_factor"	33.125736	Moderate expression
693	scf_8254814:23930 0-239755	TTHERM_004179 50	Note="transmembrane_protein_putative"	33.119231	Low Expression
694	scf_8254747:57988 7-581802	TTHERM_001452 00	Note="sugar_porter_(SP)_family_MFS_transporter"	33.103577	Low Expression
695	scf_8254397:12450 6-125222	TTHERM_008238 00	Note="ubiquitin_carboxyl-terminal_hydrolase"	33.089495	Moderate expression
696	scf_8254373:6404- 6970	TTHERM_002372 80	Note="NUDIX_hydrolase"	33.057294	Moderate expression
697	scf_8254460:33851 8-339484	TTHERM_003487 00	Note="EF_hand_protein"	33.050128	Low Expression
698	scf_8254811:42399 9-426624	TTHERM_000691 80	Note="plasma-membrane_choline_transporter"	33.024137	High expression
699	scf_8254752:67772 0-679469	TTHERM_001128 20	Note="hypothetical_protein"	33.020291	High expression

700	scf_8254564:39324 2-395147	TTHERM_000218 829	Note="saccharopine_dehydrogenase"	32.981013	#N/A
701	scf_8254617:31223 2-313452	TTHERM_001939 69	Note="hypothetical_protein"	32.948977	No-Expression
702	scf_8254716:33044 3-331999	TTHERM_001000 90	Note="hypothetical_protein"	32.90118	Moderate expression
703	scf_8254373:11977 2-122588	TTHERM_002376 40	Note="cytochrome_d1_heme_domain_protein"	32.857781	Moderate expression
704	scf_8254545:76502 0-766495	TTHERM_001514 70	Note="acyl-CoA_dehydrogenase"	32.842357	High expression
705	scf_8254728:14661 6-147147	TTHERM_008353 80	Note="hypothetical_protein"	32.839583	No-Expression
706	scf_8254431:77850 5-779337	TTHERM_005720 20	Note="GST58_glutathione_S-transferase_amine- terminal_domain_protein"	32.815427	Moderate expression
707	scf_8254665:82786 5-828203	TTHERM_002331 30	Note="transmembrane_protein_putative"	32.811588	No-Expression
708	scf_8254811:93124 4-931825	TTHERM_000589 50	Note="hypothetical_protein"	32.801412	Low Expression
709	scf_8253887:11159 8-112086	TTHERM_006470 60	Note="hypothetical_protein"	32.77608	Low Expression
710	scf_8254528:49250 -49486	TTHERM_006370 00	Note="hypothetical_protein"	32.770642	Moderate expression
711	scf_8254788:46271 -47527	TTHERM_002593 50	Note="monothiol_glutaredoxin-S11_protein"	32.756905	High expression
712	scf_8254814:66944 4-670289	TTHERM_008846 70	Note="hypothetical_protein"	32.756472	No-Expression
713	scf_8254617:56163 0-563189	TTHERM_001946 80	Note="cytosolic_iron- sulfur_protein_assembly_protein_CIAO1-B_putative"	32.730292	Moderate expression
714	scf_8254617:80562 -82402	TTHERM_001934 10	Note="serine_hydroxymethyltransferase"	32.686699	High expression
715	scf_8254611:48796 7-489182	TTHERM_004836 40	Note="TCDK3_cyclin-dependent_kinase- like_Serine/Threonine_kinase_family_protein"	32.683592	Low Expression
716	scf_8253811:44474 5-445764	TTHERM_005273 40	Note="transmembrane_protein_putative"	32.6662	Low Expression
717	scf_8254555:45048 6-452330	TTHERM_004117 40	Note="oxidoreductase_NAD-binding_domain_protein"	32.643624	High expression

718	scf_8254486:12534 0-125681	TTHERM_000557 969	Note="hypothetical_protein"	32.583989	#N/A
719	scf_8254711:19087 -19617	TTHERM_010311 20	Note="HVA22/TB2/DP1_family_protein"	32.562689	High expression
720	scf_8254659:38038 8-382684	TTHERM_000466 20	Note="hypothetical_protein"	32.539872	Moderate expression
721	scf_8253900:11678 5-117237	TTHERM_009776 70	Note="transmembrane_protein_putative"	32.530028	Low Expression
722	scf_8254010:51545 0-515944	TTHERM_003299 10	Note="cell_surface_immobilization_antigen"	32.525054	Low Expression
723	scf_8254486:47054 -49161	TTHERM_005577 60	Note="NAD- dependent_epimerase/dehydratase_family_protein"	32.504664	High expression
724	scf_8254373:81179 9-814226	TTHERM_002425 00	Note="nucleolar_GTP-binding_protein"	32.46451	High expression
725	scf_8254728:27150 -27815	TTHERM_008349 70	Note="transmembrane_protein_putative"	32.44235	High expression
726	scf_8254659:12131 39-1213468	TTHERM_000518 80	Note="cytochrome_b5-like_heme/steroid- binding_domain_protein"	32.414952	Low Expression
727	scf_8254590:12623 2-127398	TTHERM_002509 10	Note="NLI_interacting_factor-like_phosphatase"	32.406165	Moderate expression
728	scf_8254460:26993 2-270507	TTHERM_003488 70	Note="optic_atrophy_3_protein"	32.396539	Low Expression
729	scf_8254563:78402 -79778	TTHERM_005999 10	Note="FAD-dependent_oxidoreductase"	32.355351	Low Expression
730	scf_8253992:16614 -18980	TTHERM_010086 30	Note="THC1_squalene-tetrahymanol_cyclase"	32.297501	High expression
731	scf_8254006:14494 4-145951	TTHERM_009453 00	Note="CTH69_papain_family_cysteine_protease"	32.295587	No-Expression
732	scf_8254693:2324- 3973	TTHERM_013984 70	Note="cytochrome_P450_family_monooxygenase"	32.250452	Moderate expression
733	scf_8254815:83840 2-839752	TTHERM_002147 00	Note="transmembrane_protein_putative"	32.241762	Moderate expression
734	scf_8254757:96446 -100160	TTHERM_010052 10	Note="V-ATPase_A_subunit_9-2_isotype_of_the_V0_sector"	32.215035	High expression
735	scf_8254582:62466 8-624991	TTHERM_000136 237	Note="hypothetical_protein"	32.187499	#N/A

736	scf_8254187:23161 3-232689	TTHERM_006957 50	Note="hypothetical_protein"	32.140572	High expression
737	scf_8254185:47710 -48501	TTHERM_010791 70	Note="3-methyladenine_DNA_glycosylase"	32.085733	Moderate expression
738	scf_8254587:21593 2-216477	TTHERM_003882 20	Note="transmembrane_protein_putative"	32.073777	Low Expression
739	scf_8254444:30596 2-306774	TTHERM_005852 80	Note="transmembrane_protein_putative"	32.064597	Moderate expression
740	scf_8254551:64422 2-645160	TTHERM_003107 60	Note="integral_membrane_protein_DUF6_containing_protein"	32.050121	Moderate expression
741	scf_8254659:62706 -64602	TTHERM_000439 10	Note="transaldolase"	31.906989	High expression
742	scf_8254528:24097 0-242376	TTHERM_006374 98	Note="hypothetical_protein"	31.90169	Moderate expression
743	scf_8254460:42416 1-424625	TTHERM_003484 30	Note="hypothetical_protein"	31.884265	Low Expression
744	scf_8254593:59273 -60100	TTHERM_004606 10	Note="aromatic_ring- opening_dioxygenase_catalytic_subunit_LigB"	31.879848	Low Expression
745	scf_8254406:46036 2-461423	TTHERM_002457 80	Note="glycoside_hydrolase_family_5_protein"	31.878548	Low Expression
746	scf_8254437:43551 3-435986	TTHERM_000446 269	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	31.869459	#N/A
747	scf_8253815:39557 9-400390	TTHERM_004979 00	Note="FLP1_phospholipid-translocating_P- type_ATPase_flippase_family_protein"	31.837652	High expression
748	scf_8254645:59861 1-605231	TTHERM_000128 319	Note="amylo-alpha-1,6-glucosidase"	31.815516	#N/A
749	scf_8254545:81399 2-815794	TTHERM_001516 10	Note="Serine/Threonine_kinase_domain_protein"	31.792998	Moderate expression
750	scf_8254803:20236 -20835	TTHERM_000000 045	Note="transmembrane_protein_putative"	31.784928	#N/A
751	scf_8254811:26598 0-267065	TTHERM_000695 40	Note="FAD/FMN-binding_family_oxidoreductase"	31.750535	Low Expression
752	scf_8254371:48355 1-484087	TTHERM_000469 289	Note="hypothetical_protein"	31.734781	#N/A
753	scf_8254679:33548 -33994	TTHERM_012277 80	Note="GNAT_family_acetyltransferase"	31.707956	Low Expression

754	scf_8254437:44378 1-445441	TTHERM_004462 90	Note="parkin_co-regulated_protein"	31.707671	High expression
755	scf_8253843:1224- 2371	TTHERM_010354 90	Note="cyclin-dependent_kinase- like_Serine/Threonine_kinase_family_protein"	31.698691	Moderate expression
756	scf_8254422:25558 6-256344	TTHERM_000691 779	Note="glutaredoxin-like_protein"	31.685587	#N/A
757	scf_8254697:12801 2-128422	TTHERM_000939 60	Note="hypothetical_protein"	31.678677	Low Expression
758	scf_8254609:83403 -83717	TTHERM_000441 911	Note="hypothetical_protein"	31.649972	#N/A
759	scf_8254402:15778 0-159100	TTHERM_006576 10	Note="N-acyl ethanolamine- hydrolyzing_acid_amidase_putative"	31.626031	Low Expression
760	scf_8254642:84- 744	TTHERM_001379 938	Note="transmembrane_protein_putative"	31.621603	#N/A
761	scf_8254648:45869 -46768	TTHERM_008770 60	Note="DnaJ_(Hsp40)_family_B_protein"	31.61592	Moderate expression
762	scf_8254811:50970 2-510523	TTHERM_000669 20	Note="hypothetical_protein"	31.564613	Low Expression
763	scf_8254417:13562 1-136943	TTHERM_007325 90	Note="EIF6_eukaryotic_translation_initiation_factor_6"	31.520273	High expression
764	scf_8254688:60287 4-606147	TTHERM_002947 40	Note="vps16_amine-terminal_region_protein"	31.478156	High expression
765	scf_8254736:57173 -57526	TTHERM_007340 30	Note="hypothetical_protein"	31.437925	No-Expression
766	scf_8253992:31936 -32943	TTHERM_010086 60	Note="CTH85_papain_family_cysteine_protease"	31.390165	Low Expression
767	scf_8254691:11848 5-119564	TTHERM_003202 00	Note="lipase_family_protein"	31.389611	No-Expression
768	scf_8254545:58582 3-587770	TTHERM_001499 80	Note="YRS2_tyrosyl-tRNA_synthetase_protein"	31.389395	High expression
769	scf_8254379:75174 4-753967	TTHERM_003021 00	Note="class_I/II_aminotransferase"	31.374925	High expression
770	scf_8254650:98896 6-989400	TTHERM_000794 60	Note="hypothetical_protein"	31.358466	Low Expression
771	scf_8254600:17946 3-181799	TTHERM_002786 30	Note="hypothetical_protein"	31.336033	Low Expression

772	scf_8254496:96179 -97709	TTHERM_008581 10	Note="zinc_finger_protein_putative"	31.3178	No-Expression
773	scf_8254010:18674 6-187558	TTHERM_003270 00	Note="transmembrane_protein_putative"	31.311928	Moderate expression
774	scf_8254367:26322 2-263521	TTHERM_000756 530	Note="transmembrane_protein_putative"	31.265394	#N/A
775	scf_8254550:12942 4-133083	TTHERM_009231 70	Note="TPA9_cation-transporting_ATPase_putative"	31.260048	High expression
776	scf_8254630:14552 6-153440	TTHERM_007589 80	Note="dopey_amine-terminal_domain_protein"	31.249031	High expression
777	scf_8254415:33128 -35236	TTHERM_012053 20	Note="hypothetical_protein"	31.236527	Moderate expression
778	scf_8254594:19163 0-192448	TTHERM_002857 00	Note="RING_finger_protein"	31.229574	Low Expression
779	scf_8254380:24419 -26957	TTHERM_011251 40	Note="serine/threonine_protein_phosphatase_superfamily_prot ein_putative"	31.151017	High expression
780	scf_8254798:75593 3-756286	TTHERM_001856 20	Note="hypothetical_protein"	31.132345	No-Expression
781	scf_8254558:17315 -18508	TTHERM_013594 80	Note="CTH63_papain_family_cysteine_protease"	31.129789	No-Expression
782	scf_8254584:34277 3-343348	TTHERM_006270 30	Note="kazal-type_proteinase_inhibitor_1"	31.03909	No-Expression
783	scf_8254448:18104 0-182098	TTHERM_009258 50	Note="CTH49_papain_family_cysteine_protease"	31.036564	No-Expression
784	scf_8253855:68166 -69407	TTHERM_011339 50	Note="fatty_acid_desaturase_family_protein"	30.964308	No-Expression
785	scf_8254377:14958 4-152022	TTHERM_007160 30	Note="cyclin-dependent_kinase- like_Serine/Threonine_kinase_family_protein"	30.953268	Moderate expression
786	scf_8254446:28602 9-288284	TTHERM_004940 90	Note="hypothetical_protein"	30.927209	High expression
787	scf_8254811:74919 8-751860	TTHERM_000595 30	Note="hypothetical_protein"	30.914733	Low Expression
788	scf_8254667:46158 8-463215	TTHERM_000286 70	Note="deoxyribonuclease_II"	30.906468	Moderate expression
789	scf_8254389:41986 -43532	TTHERM_009941 70	Note="transmembrane_protein_putative"	30.893058	No-Expression

790	scf_8254803:15532 62-1554050	TTHERM_000136 70	Note="transmembrane_protein_putative"	30.874008	No-Expression
791	scf_8253957:76905 -78246	TTHERM_010714 80	Note="hypothetical_protein"	30.849442	Moderate expression
792	scf_8254470:20096 1-201280	TTHERM_000630 179	Note="hypothetical_protein"	30.843928	#N/A
793	scf_8254579:16806 9-171435	TTHERM_006665 70	Note="alpha_amylase_domain_protein"	30.823553	High expression
794	scf_8254803:13407 92-1342845	TTHERM_000131 70	Note="WD40_domain_protein"	30.806422	Moderate expression
795	scf_8254556:11654 0-116941	TTHERM_009576 70	Note="transmembrane_protein_putative"	30.777834	No-Expression
796	scf_8254418:30707 -32423	TTHERM_011003 80	Note="hypothetical_protein"	30.767691	High expression
797	scf_8254587:47004 0-471230	TTHERM_003898 60	Note="XIAP-associated_factor_1"	30.727561	Moderate expression
798	scf_8254072:20532 -22407	TTHERM_013458 10	Note="PPN1_serine/threonine_protein_phosphatase_2A"	30.695756	High expression
799	scf_8254491:15715 7-158437	TTHERM_007871 90	Note="transmembrane_protein_putative"	30.683715	Moderate expression
800	scf_8254373:67808 -71568	TTHERM_002374 90	Note="RAD54_SNF2_family_amine-terminal_protein"	30.68192	Moderate expression
801	scf_8254391:36955 -37369	TTHERM_000703 379	Note="hypothetical_protein"	30.648235	#N/A
802	scf_8254776:47383 0-474459	TTHERM_001605 40	Note="hypothetical_protein"	30.646801	No-Expression
803	scf_8254747:34456 0-345702	TTHERM_001435 70	Note="TGL10_ab- hydrolase_associated_lipase_region_family_protein"	30.562065	No-Expression
804	scf_8254448:18822 3-189455	TTHERM_009258 90	Note="squalene/phytoene_synthase"	30.51921	High expression
805	scf_8254798:13475 7-136004	TTHERM_001691 10	Note="GST36_glutathione_S-transferase_amine- terminal_domain_protein"	30.49894	Low Expression
806	scf_8254565:24223 1-244212	TTHERM_004711 20	Note="Serine/Threonine_kinase_domain_protein"	30.491264	Moderate expression
807	scf_8254594:19707 -21259	TTHERM_002852 90	Note="ubiquitin-conjugating_enzyme_E2"	30.488621	High expression

808	scf_8254564:86749 9-868955	TTHERM_002211 50	Note="NTP6_equilibrative_nucleoside_transporter_family_protein"	30.476654	Moderate expression
809	scf_8254659:76201 3-763125	TTHERM_000476 40	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	30.395158	Moderate expression
810	scf_8254373:27312 7-273588	TTHERM_002391 49	Note="Bola-like_protein"	30.366162	Low Expression
811	scf_8254593:13996 -15505	TTHERM_004605 00	Note="hypothetical_protein"	30.323805	Moderate expression
812	scf_8254650:10067 80-1009122	TTHERM_000795 40	Note="succinyl-CoA_ligase_[GDP-forming]_subunit_beta"	30.297988	High expression
813	scf_8254610:53157 4-532035	TTHERM_000339 869	Note="hypothetical_protein"	30.287323	#N/A
814	scf_8254650:56400 3-565535	TTHERM_000774 29	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	30.286216	Low Expression
815	scf_8254428:92718 -94812	TTHERM_005354 40	Note="hypothetical_protein"	30.224899	Moderate expression
816	scf_8254428:37393 4-375046	TTHERM_005371 90	Note="LAG1_longevity_assurance_protein"	30.224296	Moderate expression
817	scf_8254187:19510 -20986	TTHERM_006975 20	Note="dual_specificity_protein_phosphatase"	30.203296	Moderate expression
818	scf_8254284:80699 -82457	TTHERM_000763 039	Note="MutS_domain_V_family_protein"	30.162003	#N/A
819	scf_8254187:21506 3-218244	TTHERM_006958 20	Note="WD_domain_G-beta_repeat_protein"	30.145813	Moderate expression
820	scf_8254676:7791- 9227	TTHERM_011883 30	Note="transmembrane_amino_acid_transporter_protein"	30.116648	Moderate expression
821	scf_8254549:30264 -31983	TTHERM_007748 60	Note="hypothetical_protein"	30.093409	No-Expression
822	scf_8253817:65850 -66965	TTHERM_009195 80	Note="oxysterol-binding_protein"	30.092188	Moderate expression
823	scf_8254747:12350 6-125207	TTHERM_001410 60	Note="platelet-activating_factor_acetylhydrolase"	30.089034	Moderate expression
824	scf_8254653:15186 7-153665	TTHERM_006412 40	Note="GPM1_2,3-bisphosphoglycerate-dependent_phosphoglycerate_mutase"	30.075943	High expression
825	scf_8254671:29217 3-296114	TTHERM_004259 00	Note="serine/threonine_phosphatase_2C"	30.042796	Moderate expression

826	scf_8254548:59222 8-594273	TTHERM_003185 40	Note="MFS_transporter"	30.032863	No-Expression
827	scf_8254580:64474 -65921	TTHERM_005867 40	Note="translation_elongation_factor_Tu_protein"	30.020146	Moderate expression
828	scf_8253969:1606- 1946	TTHERM_002653 432	Note="transmembrane_protein_putative"	29.959753	#N/A
829	scf_8254479:29876 0-300605	TTHERM_006244 30	Note="type_II_methionine_aminopeptidase"	29.956107	High expression
830	scf_8254644:12668 1-127541	TTHERM_006489 20	Note="hypothetical_protein"	29.939248	Moderate expression
831	scf_8254716:11211 74-1124964	TTHERM_001070 40	Note="tetratricopeptide_repeat_protein"	29.90198	High expression
832	scf_8254611:47496 0-476078	TTHERM_004836 00	Note="hypothetical_protein"	29.889617	Moderate expression
833	scf_8254650:10776 11-1079317	TTHERM_000797 10	Note="MFS_transporter"	29.887751	Moderate expression
834	scf_8254775:6092- 7708	TTHERM_012172 00	Note="poly(A)+_RNA_export_protein"	29.884895	Moderate expression
835	scf_8254034:10751 9-108490	TTHERM_008527 10	Note="hypothetical_protein"	29.88248	Low Expression
836	scf_8254043:77945 -79481	TTHERM_007808 30	Note="ubiquinone_biosynthesis_hydroxylase_UbiH/UbiF/VisC /COQ6_family_protein"	29.879614	High expression
837	scf_8254814:56527 0-565957	TTHERM_004142 40	Note="transmembrane_protein_putative"	29.867443	Low Expression
838	scf_8254776:68892 3-689600	TTHERM_001580 00	Note="transmembrane_protein_putative"	29.833339	High expression
839	scf_8254665:70477 2-705497	TTHERM_002278 10	Note="hypothetical_protein"	29.818065	Low Expression
840	scf_8254587:21411 0-214575	TTHERM_000388 219	Note="transmembrane_protein_putative"	29.783143	#N/A
841	scf_8254654:32872 -34931	TTHERM_005038 39	Note="2-hydroxyphytanoyl-CoA_lyase"	29.738637	Low Expression
842	scf_8254609:21646 -22829	TTHERM_000441 733	Note="transmembrane_protein_putative"	29.738564	#N/A
843	scf_8254597:1224- 1769	TTHERM_003412 50	Note="emopamil-binding_protein"	29.735441	Moderate expression

844	scf_8254664:39067 3-392047	TTHERM_005228 10	Note="hypothetical_protein"	29.711528	High expression
845	scf_8254402:35697 4-357303	TTHERM_000659 131	Note="hypothetical_protein"	29.704572	#N/A
846	scf_8254548:42342 1-423729	TTHERM_000317 059	Note="hypothetical_protein"	29.682385	#N/A
847	scf_8254395:12763 7-129101	TTHERM_004554 50	Note="transmembrane_protein_putative"	29.667007	Low Expression
848	scf_8254431:69025 1-691702	TTHERM_003137 50	Note="PX_domain_protein"	29.641981	Moderate expression
849	scf_8254731:17642 8-176763	TTHERM_000594 179	Note="hypothetical_protein"	29.588674	#N/A
850	scf_8254607:48245 9-484000	TTHERM_004374 20	Note="kinase_domain_protein"	29.577327	Moderate expression
851	scf_8254402:33366 9-335834	TTHERM_006590 80	Note="ubiquitin-conjugating_enzyme"	29.571157	Moderate expression
852	scf_8254565:78782 -79405	TTHERM_004706 20	Note="isochorismatase_hydrolase"	29.559191	Moderate expression
853	scf_8254659:91573 4-917910	TTHERM_000490 60	Note="tetratricopeptide_repeat_protein"	29.530181	High expression
854	scf_8254371:49892 6-500257	TTHERM_004693 10	Note="hypothetical_protein"	29.504233	Low Expression
855	scf_8254671:51310 4-513943	TTHERM_004274 80	Note="endonuclease/exonuclease/phosphatase_family_protein"	29.484486	No-Expression
856	scf_8254667:72440 3-726570	TTHERM_000302 60	Note="divergent_AAA_domain_protein"	29.4475	High expression
857	scf_8254466:29471 -30031	TTHERM_009516 90	Note="hypothetical_protein"	29.398702	No-Expression
858	scf_8254551:11476 1-116179	TTHERM_003043 40	Note="tetratricopeptide_repeat_protein"	29.384716	Moderate expression
859	scf_8254528:20496 6-206153	TTHERM_006374 10	Note="cytochrome_C_oxidase_assembly_protein_COX11"	29.35164	High expression
860	scf_8254486:12342 9-123785	TTHERM_005579 50	Note="hypothetical_protein"	29.326626	Low Expression
861	scf_8254037:13140 -13660	TTHERM_013676 80	Note="hypothetical_protein"	29.321135	No-Expression

862	scf_8254707:48171-49379	TTHERM_00895700	Note="hypothetical_protein"	29.307675	Moderate expression
863	scf_8254807:374153-374677	TTHERM_00540230	Note="hypothetical_protein"	29.296236	Low Expression
864	scf_8254815:603152-604219	TTHERM_00204090	Note="hypothetical_protein"	29.261442	Low Expression
865	scf_8254638:393367-394746	TTHERM_00189320	Note="pyridoxal-phosphate-dependent_protein"	29.239729	Moderate expression
866	scf_8254691:71027-71578	TTHERM_00320040	Note="hypothetical_protein"	29.225164	Moderate expression
867	scf_8254811:489920-492474	TTHERM_00066990	Note="START-2_domain_protein"	29.217471	Moderate expression
868	scf_8254587:122573-124246	TTHERM_00384950	Note="C2H2_type_zinc-finger_protein"	29.215715	Moderate expression
869	scf_8254659:639037-639573	TTHERM_00047350	Note="hypothetical_protein"	29.209701	Low Expression
870	scf_8254590:4183-5559	TTHERM_00249550	Note="serine_carboxypeptidase_family_protein"	29.197277	No-Expression
871	scf_8254437:414911-416938	TTHERM_00446190	Note="RRM25_RNA_recognition_motif_protein"	29.18746	High expression
872	scf_8254597:172851-174212	TTHERM_00343710	Note="hypothetical_protein"	29.111465	Low Expression
873	scf_8254034:39797-41606	TTHERM_01123920	Note="transmembrane_protein_putative"	29.10859	Moderate expression
874	scf_8254075:35013-35771	TTHERM_01262860	Note="hypothetical_protein"	29.096354	Low Expression
875	scf_8254652:51515-52975	TTHERM_00377260	Note="integral_membrane_protein_DUF6_containing_protein"	29.084092	Moderate expression
876	scf_8254551:270327-270998	TTHERM_000307769	Note="transmembrane_protein_putative"	29.04733	#N/A
877	scf_8254431:289492-293310	TTHERM_00312760	Note="ion_channel_protein"	29.035141	Low Expression
878	scf_8254527:352240-352876	TTHERM_00486700	Note="hypothetical_protein"	29.030337	Moderate expression
879	scf_8254544:26004-29161	TTHERM_01099060	Note="WD_domain_G-beta_repeat_protein"	29.029824	High expression

880	scf_8254650:71988 1-720986	TTHERM_000777 70	Note="hypothetical_protein"	29.021343	Low Expression
881	scf_8254564:86104 3-864938	TTHERM_002211 40	Note="ARS2_alanine-tRNA_ligase/alanyl-tRNA_synthetase_protein"	29.015628	High expression
882	scf_8254583:88801 -89301	TTHERM_011228 30	Note="hypothetical_protein"	28.98421	No-Expression
883	scf_8254816:46215 -46553	TTHERM_001237 399	Note="hypothetical_protein"	28.981643	#N/A
884	scf_8254437:36337 1-366852	TTHERM_000446 029	Note="bromodomain_domain_protein"	28.969717	#N/A
885	scf_8253915:65683 -66832	TTHERM_006884 60	Note="Rho_GDP-dissociation_inhibitor_putative"	28.967734	Moderate expression
886	scf_8254788:35459 1-355934	TTHERM_002630 30	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	28.959971	Moderate expression
887	scf_8253855:82049 -82477	TTHERM_011340 00	Note="kazal-type_proteinase_inhibitor_1"	28.930175	Low Expression
888	scf_8254542:53707 -54353	TTHERM_001250 151	Note="transmembrane_protein_putative"	28.888021	#N/A
889	scf_8254799:17991 5-180601	TTHERM_008129 80	Note="hypothetical_protein"	28.876858	Low Expression
890	scf_8253887:13558 7-137060	TTHERM_006471 40	Note="hypothetical_protein"	28.874644	Moderate expression
891	scf_8254181:92282 -93022	TTHERM_009306 40	Note="hypothetical_protein"	28.860011	Moderate expression
892	scf_8254788:81606 4-819081	TTHERM_002651 80	Note="NRK16_Serine/Threonine_kinase_domain_protein"	28.855561	Moderate expression
893	scf_8254663:17694 4-178453	TTHERM_008566 70	Note="senescence-associated_protein"	28.854868	Low Expression
894	scf_8254431:19362 3-196554	TTHERM_003124 90	Note="hypothetical_protein"	28.845858	Moderate expression
895	scf_8254697:21900 9-220436	TTHERM_000928 00	Note="hypothetical_protein"	28.839851	Low Expression
896	scf_8254630:5975- 8390	TTHERM_007575 60	Note="MFS_transporter"	28.836462	Low Expression
897	scf_8254594:10444 6-105959	TTHERM_002855 10	Note="hypothetical_protein"	28.814387	Moderate expression

898	scf_8254402:30228 3-303227	TTHERM_006589 80	Note="ankyrin_repeat_protein"	28.809224	Low Expression
899	scf_8254436:10429 4-104797	TTHERM_000939 009	Note="hypothetical_protein"	28.740777	#N/A
900	scf_8254682:97991 -98818	TTHERM_009272 40	Note="short_chain_dehydrogenase"	28.731924	Moderate expression
901	scf_8254440:64980 -66696	TTHERM_008252 40	Note="MORN_repeat_protein"	28.711619	Moderate expression
902	scf_8254671:12455 0-125002	TTHERM_004244 60	Note="ATP-binding_ABC_transporter"	28.663357	No-Expression
903	scf_8254464:36632 5-366891	TTHERM_000580 367	Note="L-isaspartate_O-methyltransferase"	28.649031	#N/A
904	scf_8254440:18166 3-182463	TTHERM_008255 60	Note="zinc_finger_C2H2_type_family_protein"	28.63094	Moderate expression
905	scf_8254600:48093 6-482044	TTHERM_002833 40	Note="hypothetical_protein"	28.606521	Moderate expression
906	scf_8254814:73283 9-736742	TTHERM_008858 70	Note="calpain_family_cysteine_protease"	28.589957	High expression
907	scf_8254512:35517 -36902	TTHERM_012289 80	Note="hypothetical_protein"	28.586267	Low Expression
908	scf_8254406:20000 8-202796	TTHERM_002452 10	Note="MFS_transporter"	28.553466	High expression
909	scf_8254814:49576 6-496578	TTHERM_004144 10	Note="YhhN_family_protein"	28.507015	Moderate expression
910	scf_8254548:33033 8-330664	TTHERM_000316 723	Note="hypothetical_protein"	28.500426	#N/A
911	scf_8254594:72069 1-723022	TTHERM_002921 40	Note="SEC141_CRAL/TRIO_amine- terminal_domain_protein"	28.480574	High expression
912	scf_8254464:18873 -19869	TTHERM_005784 50	Note="hypothetical_protein"	28.480436	No-Expression
913	scf_8254362:983- 4051	TTHERM_007229 20	Note="calcium-translocating_P-type_ATPase_PMCA- type_protein"	28.458499	High expression
914	scf_8254365:28949 4-291161	TTHERM_006861 00	Note="hydroxymethylglutaryl-CoA_lyase"	28.457744	Moderate expression
915	scf_8254598:21996 3-220349	TTHERM_007535 80	Note="transmembrane_protein_putative"	28.431239	Low Expression

916	scf_8254527:640-2697	TTHERM_004857 40	Note="hypothetical_protein"	28.410101	Low Expression
917	scf_8254475:73524-74761	TTHERM_004198 50	Note="50S_ribosomal_protein_L27"	28.406694	High expression
918	scf_8254811:1403095-1404483	TTHERM_000537 40	Note="hypothetical_protein"	28.376094	Moderate expression
919	scf_8254564:241558-246379	TTHERM_002184 30	Note="transporter_family_ABC_domain_protein"	28.366559	#N/A
920	scf_8254638:928478-932151	TTHERM_001920 00	Note="CYC5_amine-terminal_domain_cyclin"	28.352609	Moderate expression
921	scf_8253815:383864-386925	TTHERM_004978 60	Note="RRM33_RNA-binding_motif_protein"	28.338653	Moderate expression
922	scf_8254077:10764-12689	TTHERM_011263 80	Note="Serine/Threonine_kinase_domain_protein"	28.335958	Moderate expression
923	scf_8254373:598472-600178	TTHERM_002420 20	Note="hypothetical_protein"	28.321616	Moderate expression
924	scf_8254686:238254-239327	TTHERM_004492 00	Note="class_I_glutamine_amidotransferase"	28.289377	Moderate expression
925	scf_8254548:712530-714558	TTHERM_003189 00	Note="MFS_transporter"	28.248205	Moderate expression
926	scf_8254814:404419-405898	TTHERM_004155 90	Note="peptidase_C13_family_protein"	28.224357	Low Expression
927	scf_8254181:179241-180059	TTHERM_000929 442	Note="transmembrane_protein_putative"	28.178351	#N/A
928	scf_8254582:200025-202220	TTHERM_001384 00	Note="amidohydrolase"	28.173815	High expression
929	scf_8254526:9196-9639	TTHERM_014856 80	Note="hypothetical_protein"	28.12571	No-Expression
930	scf_8254814:366663-368546	TTHERM_004156 90	Note="ser/thr_protein_phosphatase_family_protein"	28.101613	No-Expression
931	scf_8254811:826944-828327	TTHERM_000592 80	Note="zinc_finger_C3HC4_type_(RING_finger)_protein"	28.088797	Moderate expression
932	scf_8254470:1013275-1013730	TTHERM_000429 979	Note="transmembrane_protein_putative"	28.055683	#N/A
933	scf_8254460:184362-187229	TTHERM_003490 90	Note="bZIP_transcription_factor"	28.016404	Moderate expression

934	scf_8254418:85748-88305	TTHERM_011005 40	Note="MATE_efflux_family_protein"	27.969273	Low Expression
935	scf_8254638:543647-544552	TTHERM_001907 40	Note="transmembrane_protein_putative"	27.967696	Moderate expression
936	scf_8254194:6068-7078	TTHERM_010195 70	Note="transmembrane_protein"	27.940541	Moderate expression
937	scf_8254803:2139371-2140680	TTHERM_000242 60	Note="FAD/FMN-binding_family_oxidoreductase"	27.916051	Moderate expression
938	scf_8254373:613522-613777	TTHERM_002420 60	Note="hypothetical_protein"	27.898366	Low Expression
939	scf_8254666:83714-84691	TTHERM_003992 60	Note="scramblase"	27.888354	Moderate expression
940	scf_8254479:442147-442485	TTHERM_006248 90	Note="cytochrome_b5-like_heme/steroid-binding_domain_protein"	27.857271	High expression
941	scf_8254593:177981-179468	TTHERM_004629 40	Note="oxidoreductase_short_chain_dehydrogenase/reductase_family_protein"	27.856139	Moderate expression
942	scf_8254752:358940-359464	TTHERM_005753 40	Note="hypothetical_protein"	27.840907	Low Expression
943	scf_8254664:307106-308427	TTHERM_005225 80	Note="SPS1_selenide_water_dikinase"	27.83609	High expression
944	scf_8254798:360560-363334	TTHERM_001717 20	Note="transmembrane_protein_putative"	27.835125	High expression
945	scf_8254603:63226-63948	TTHERM_010469 40	Note="glyoxalase/bleomycin_resistance_protein/dioxygenase_putative"	27.822727	Moderate expression
946	scf_8254716:1145296-1146192	TTHERM_001070 90	Note="oxidoreductase_short_chain_dehydrogenase/reductase_family_protein"	27.81229	Moderate expression
947	scf_8254395:472935-473953	TTHERM_004593 20	Note="coatomer_subunit_epsilon"	27.795426	Moderate expression
948	scf_8254653:316744-317358	TTHERM_006446 50	Note="START_domain_protein"	27.792784	Moderate expression
949	scf_8254659:83770-84804	TTHERM_000449 70	Note="hypothetical_protein"	27.734066	Moderate expression
950	scf_8254659:337156-338078	TTHERM_000465 20	Note="60_kDa_inner_membrane_protein"	27.723898	Moderate expression
951	scf_8254611:463424-464473	TTHERM_004835 50	Note="hypothetical_protein"	27.711617	Low Expression

952	scf_8254777:31555 9-316074	TTHERM_003723 20	Note="hypothetical_protein"	27.703754	Low Expression
953	scf_8254479:43234 9-433338	TTHERM_006248 50	Note="START_domain_protein"	27.701132	Moderate expression
954	scf_8254776:23377 5-235124	TTHERM_001612 10	Note="2_iron_2_sulfur_cluster-binding_protein"	27.689698	High expression
955	scf_8254565:26922 2-269872	TTHERM_004712 30	Note="TRAF-type_zinc_finger_protein"	27.687754	Low Expression
956	scf_8254373:26698 4-268278	TTHERM_002391 30	Note="MSC1_UAA_transporter_family_protein"	27.628062	Moderate expression
957	scf_8254582:84226 4-843241	TTHERM_001336 80	Note="short-chain_dehydrogenase/reductase_family_protein"	27.591996	Low Expression
958	scf_8254548:60222 7-603474	TTHERM_003185 80	Note="arginase"	27.578267	Moderate expression
959	scf_8254551:49305 9-494972	TTHERM_003102 40	Note="MOZ/SAS_family_protein"	27.572383	Moderate expression
960	scf_8254485:21715 -22374	TTHERM_009893 40	Note="hypothetical_protein"	27.55265	Low Expression
961	scf_8254617:12371 -14706	TTHERM_001932 00	Note="Serine/Threonine_kinase_domain_protein"	27.536571	High expression
962	scf_8254652:20179 1-204049	TTHERM_003786 80	Note="transmembrane_protein_putative"	27.504441	Low Expression
963	scf_8254583:64154 -65659	TTHERM_011227 30	Note="von_willebrand_factor_type_A_domain_protein"	27.50365	Moderate expression
964	scf_8254460:1694- 3898	TTHERM_003535 60	Note="hypothetical_protein"	27.49405	High expression
965	scf_8254716:71249 4-712892	TTHERM_000104 938	Note="hypothetical_protein"	27.459095	#N/A
966	scf_8254728:85260 -85685	TTHERM_008351 50	Note="transmembrane_protein_putative"	27.454433	Low Expression
967	scf_8254199:31318 -35003	TTHERM_010492 70	Note="tesmin_TSO1-like_CXC_domain_protein"	27.45097	High expression
968	scf_8254627:92754 -93491	TTHERM_000860 540	Note="hypothetical_protein"	27.449285	#N/A
969	scf_8254646:13675 2-137662	TTHERM_008195 40	Note="membrane_protein_putative"	27.403135	Moderate expression

970	scf_8254610:30074 0-301281	TTHERM_003380 80	Note="hypothetical_protein"	27.398742	Moderate expression
971	scf_8254582:56971 2-571335	TTHERM_001364 10	Note="tesmin_TSO1-like_CXC_domain_protein"	27.393608	Moderate expression
972	scf_8254092:76091 -79533	TTHERM_011049 90	Note="patatin-like_phospholipase_family_protein"	27.39017	Moderate expression
973	scf_8254594:23191 4-232333	TTHERM_000286 801	Note="transmembrane_protein_putative"	27.357201	#N/A
974	scf_8254527:17901 0-181339	TTHERM_004861 70	Note="hypothetical_protein"	27.344244	Low Expression
975	scf_8254428:28922 8-290877	TTHERM_005359 80	Note="GMC_oxidoreductase_family_protein"	27.335416	Moderate expression
976	scf_8254359:36852 -39919	TTHERM_008366 20	Note="tetratricopeptide_repeat_protein"	27.328577	Moderate expression
977	scf_8254814:12558 3-126686	TTHERM_004182 30	Note="oxidoreductase_short_chain_dehydrogenase/reductase_f amily_protein"	27.328065	Moderate expression
978	scf_8254803:58371 1-585033	TTHERM_000062 90	Note="MFS_transporter"	27.30418	No-Expression
979	scf_8254718:10282 4-103295	TTHERM_008080 70	Note="hypothetical_protein"	27.289426	Low Expression
980	scf_8254658:21877 6-220002	TTHERM_006755 80	Note="hypothetical_protein"	27.284989	Moderate expression
981	scf_8254362:10094 8-101877	TTHERM_000723 219	Note="hypothetical_protein"	27.275478	#N/A
982	scf_8254582:11970 9-121668	TTHERM_001385 30	Note="fatty_acid_desaturase"	27.267414	High expression
983	scf_8254650:86762 -90960	TTHERM_000743 20	Note="glycogen/starch/alpha-glucan_phosphorylase"	27.258863	High expression
984	scf_8254555:49769 0-498361	TTHERM_004118 80	Note="transmembrane_protein_putative"	27.24008	Moderate expression
985	scf_8254370:73326 -74528	TTHERM_007603 70	Note="NUDIX_hydrolase"	27.202838	Low Expression
986	scf_8254393:46250 -47785	TTHERM_012891 70	Note="transmembrane_protein_putative"	27.202279	Moderate expression
987	scf_8254814:72472 -74378	TTHERM_004184 30	Note="ABC_transporter_family_protein"	27.155981	Low Expression

988	scf_8253887:16328 3-164805	TTHERM_006471 70	Note="MORN_motif_protein"	27.15329	Moderate expression
989	scf_8254636:34664 0-347308	TTHERM_004908 80	Note="GST15_glutathione_S-transferase_amine- terminal_domain_protein"	27.152365	No-Expression
990	scf_8253811:16950 8-169947	TTHERM_000526 509	Note="hypothetical_protein"	27.136023	#N/A
991	scf_8254406:52614 4-528477	TTHERM_002459 00	Note="transmembrane_protein_putative"	27.118995	No-Expression
992	scf_8254396:5708- 6332	TTHERM_011085 00	Note="ubiquitin-fold_modifier-conjugating_enzyme"	27.101297	Moderate expression
993	scf_8254815:25358 6-254023	TTHERM_001982 50	Note="cytochrome_P450_family_monooxygenase"	27.08745	No-Expression
994	scf_8254716:62923 5-630891	TTHERM_001027 30	Note="2-aminoethylphosphonate-pyruvate_transaminase"	27.054111	High expression
995	scf_8254545:7141- 8420	TTHERM_001474 80	Note="required_for_amino_acid_permease_transport_from_the _to_the_cell_surface_protein"	27.052539	Moderate expression
996	scf_8254761:99575 -100081	TTHERM_001050 619	Note="hypothetical_protein"	27.044324	#N/A
997	scf_8254422:20648 9-207499	TTHERM_006916 30	Note="CTH68_papain_family_cysteine_protease"	27.044236	No-Expression
998	scf_8254752:86779 4-869556	TTHERM_001133 10	Note="RNA_polymerase_II-binding_domain_protein"	27.037869	Moderate expression
999	scf_8254397:11259 2-113401	TTHERM_008237 80	Note="taurine_catabolism_dioxygenase_TauD/TfdA_family_p rotein"	27.034239	Low Expression
100 0	scf_8254650:89955 0-901826	TTHERM_000792 10	Note="Von_willebrand_factor_type_A_(vWA)_domain_was_o riginally_protein"	27.026291	Moderate expression

RACS processed data for the first 1000 (out of 5019) intergenic regions occupied by that Med31: Refer to Digital Spreadsheet (TET_Med31_MAC_Genome_Intergenic) for complete results.

#	Scaffold	Enrichment Average	#	Scaffold	Enrichment Average
1	scf_8254638:904118-904178	937.196378	501	scf_8254377:125677-126866	22.280408
2	scf_8254667:1710197-1711256	326.9548664	502	scf_8254362:255722-256235	22.278691
3	scf_8254611:6704-6864	311.4020157	503	scf_8254551:533715-535292	22.275956
4	scf_8253815:475701-476118	301.2935634	504	scf_8254820:215652-216336	22.262122
5	scf_8254617:720525-721053	279.3181483	505	scf_8253880:102277-103331	22.20474
6	scf_8254652:481006-482770	269.9978848	506	scf_8254459:45355-46951	22.175948
7	scf_8254444:304927-305961	264.3336286	507	scf_8254459:100693-101105	22.143111
8	scf_8254551:619719-620260	224.0045748	508	scf_8254034:44811-45581	22.091437
9	scf_8255782:1071-1340	221.6011768	509	scf_8254545:545036-545478	22.076011
10	scf_8253890:9921-11277	203.6499609	510	scf_8254464:369551-369741	21.980125
11	scf_8254487:51339-51738	200.7341055	511	scf_8254605:284947-287283	21.950689
12	scf_8254460:333951-333998	177.6278927	512	scf_8254767:25893-27186	21.949306
13	scf_8254798:584967-585045	165.8209782	513	scf_8254813:223193-223712	21.927423
14	scf_8254412:1-1515	161.5281734	514	scf_8254551:251526-252256	21.925279
15	scf_8253915:52256-53327	159.1727458	515	scf_8254555:524891-525862	21.864617
16	scf_8254594:551766-551834	152.2902677	516	scf_8254479:234027-235219	21.836506
17	scf_8254247:18285-19608	152.03343	517	scf_8254564:678414-678783	21.824715
18	scf_8254047:66158-66671	149.0100122	518	scf_8253904:28689-30997	21.783563
19	scf_8254697:1317019-1317421	148.6318294	519	scf_8254396:68726-69992	21.706717
20	scf_8254359:36320-36851	147.2392765	520	scf_8254716:835698-836597	21.679755
21	scf_8254033:5540-6752	147.1063569	521	scf_8254664:488394-488953	21.672775
22	scf_8254486:126525-127681	145.0292881	522	scf_8254359:92146-93825	21.661572
23	scf_8254747:248172-249399	140.7628525	523	scf_8254652:214284-217201	21.641824
24	scf_8254600:188670-189361	139.0160814	524	scf_8254587:318541-319392	21.595579
25	scf_8254587:590129-590758	137.0404857	525	scf_8254479:148623-150122	21.574518
26	scf_8254428:118067-118717	131.4587591	526	scf_8254391:1-2731	21.570471
27	scf_8254626:5509-6155	129.1664164	527	scf_8254594:65183-65682	21.54211
28	scf_8254457:102357-102779	128.6898483	528	scf_8254645:161754-162346	21.535674
29	scf_8254719:555178-557204	126.7532946	529	scf_8253957:107416-107999	21.48306
30	scf_8254752:825170-826536	125.9772415	530	scf_8253823:35783-36908	21.474666
31	scf_8254437:267649-268580	123.201926	531	scf_8253811:29565-30309	21.435951

32	scf_8254659:970190-970717	120.2286911	532	scf_8254686:74105-76990	21.399699
33	scf_8254579:108516-109488	115.9961687	533	scf_8254395:146337-147215	21.383384
34	scf_8254284:219647-220237	115.6997975	534	scf_8254763:249689-250426	21.359992
35	scf_8254006:106785-107692	113.7891505	535	scf_8254435:219630-221726	21.3408
36	scf_8254010:451610-452458	112.5670635	536	scf_8254666:283041-287105	21.323858
37	scf_8253887:58727-61374	110.4523284	537	scf_8254649:256127-256534	21.305962
38	scf_8254584:239374-239648	110.2262922	538	scf_8254716:1021878-1022819	21.301019
39	scf_8254666:180396-181319	107.2361655	539	scf_8254378:274804-275656	21.266687
40	scf_8254650:685026-685709	104.0975186	540	scf_8254597:97275-98259	21.265639
41	scf_8254565:524896-525960	99.13819872	541	scf_8254464:273645-275779	21.220088
42	scf_8254373:294335-295433	98.86593811	542	scf_8254716:39829-40903	21.219912
43	scf_8254483:1-769	98.49140193	543	scf_8254617:917540-918217	21.213583
44	scf_8254456:10918-11200	97.9974053	544	scf_8254483:57460-58459	21.167668
45	scf_8254385:375346-376353	97.48789428	545	scf_8254617:817098-818014	21.127988
46	scf_8254814:366026-366662	96.67867686	546	scf_8254515:11938-12438	21.107606
47	scf_8254577:300746-301643	95.38133452	547	scf_8254654:180129-180924	21.105623
48	scf_8254697:1519429-1519815	94.25587225	548	scf_8254367:251408-252052	21.061914
49	scf_8254176:46500-54803	93.35102227	549	scf_8254437:416939-418783	21.053398
50	scf_8254104:248807-250039	91.64105788	550	scf_8254487:14861-15736	21.024622
51	scf_8254622:39499-41973	90.78961579	551	scf_8253900:147275-148676	20.99116
52	scf_8253805:12912-13298	88.06072788	552	scf_8254784:3705-5905	20.93895
53	scf_8254600:389631-390613	87.98582027	553	scf_8254691:200981-202227	20.928397
54	scf_8254449:212667-213191	87.4573757	554	scf_8253894:105365-105688	20.874531
55	scf_8254107:1356-2640	86.78159715	555	scf_8254002:78437-80233	20.787103
56	scf_8254564:663637-665318	86.75583222	556	scf_8254691:439246-441198	20.758231
57	scf_8254803:666303-666963	86.08648506	557	scf_8254476:30474-30884	20.746251
58	scf_8254747:879396-881051	85.06142926	558	scf_8254716:232487-232824	20.740932
59	scf_8254593:153416-153721	84.92671349	559	scf_8253957:105180-105780	20.697492
60	scf_8254406:157254-158539	84.58804196	560	scf_8254649:765982-766896	20.663123
61	scf_8254393:47786-48731	83.47561069	561	scf_8254582:887723-888943	20.646352
62	scf_8254605:174257-175511	82.8361844	562	scf_8254429:124867-125541	20.601962
63	scf_8254602:87885-88246	82.6047368	563	scf_8254379:218489-219292	20.586887
64	scf_8254582:822448-823329	81.49067989	564	scf_8254600:277433-278181	20.565805
65	scf_8254697:1082737-1083517	81.41038579	565	scf_8253900:140100-140834	20.555185
66	scf_8254697:98708-99936	80.13064367	566	scf_8254650:1155592-1156100	20.531839
67	scf_8254582:705845-706762	79.95525972	567	scf_8254382:119340-120215	20.515428

68	scf_8253811:146491-147257	78.09484697	568	scf_8254545:255539-258882	20.477607
69	scf_8254373:236079-236716	77.04497832	569	scf_8253811:41046-41575	20.396703
70	scf_8254043:246998-247402	76.68922311	570	scf_8254707:49380-50110	20.365976
71	scf_8254776:805031-806408	76.00745548	571	scf_8254665:469516-471748	20.34945
72	scf_8254106:104695-104949	75.33105773	572	scf_8254460:254649-255530	20.326369
73	scf_8254365:306855-308068	74.13700181	573	scf_8254469:83317-83950	20.321433
74	scf_8254579:154497-155568	73.84891828	574	scf_8254638:180408-180571	20.321139
75	scf_8254617:490255-491036	72.06608938	575	scf_8254010:244196-245004	20.198814
76	scf_8254587:249466-250500	71.41983465	576	scf_8254653:55114-55801	20.170572
77	scf_8254446:413900-414503	71.12794906	577	scf_8254584:2800-3226	20.13293
78	scf_8254707:111896-112114	71.10842128	578	scf_8254034:2488-3483	20.049675
79	scf_8254760:83127-83443	71.06541403	579	scf_8254667:432608-433467	20.036223
80	scf_8254659:498677-499856	69.99057438	580	scf_8254688:602544-602873	20.034882
81	scf_8254708:7376-8603	69.76548344	581	scf_8254437:238677-239221	20.034776
82	scf_8253811:487531-487913	69.72388049	582	scf_8254617:82403-83466	20.006867
83	scf_8254491:248369-248869	68.6686125	583	scf_8254811:1062994-1063582	19.871853
84	scf_8254468:32066-32196	67.59888302	584	scf_8254472:60535-61955	19.849425
85	scf_8254650:659416-659484	67.17013135	585	scf_8253891:275752-276309	19.819951
86	scf_8254049:15834-16105	66.82609192	586	scf_8254395:18574-20746	19.757146
87	scf_8254460:142591-142991	66.25661737	587	scf_8254597:561489-565102	19.742787
88	scf_8254605:181048-181310	66.0783014	588	scf_8254397:84403-86565	19.742434
89	scf_8254764:344828-345047	65.96490796	589	scf_8254589:119664-120337	19.712004
90	scf_8254377:205286-207884	65.03044503	590	scf_8254659:705232-708775	19.705495
91	scf_8254667:1414296-1415419	64.34122803	591	scf_8254431:22770-23123	19.698989
92	scf_8254665:174584-174861	64.33620863	592	scf_8254528:255433-256452	19.666667
93	scf_8254397:107284-108031	64.04176842	593	scf_8254588:5947-6158	19.659103
94	scf_8254711:37525-37768	63.59140442	594	scf_8254811:1178973-1179500	19.649533
95	scf_8254667:954608-957599	63.49674219	595	scf_8254666:37648-38507	19.639837
96	scf_8254666:377502-378996	63.47256198	596	scf_8254403:330549-331169	19.572622
97	scf_8254650:230218-231111	63.26674543	597	scf_8254688:473434-474254	19.497758
98	scf_8254779:37139-37494	62.60204743	598	scf_8254688:347547-348023	19.484231
99	scf_8255047:1-293	62.59358879	599	scf_8254644:256363-258643	19.469406
100	scf_8254428:416206-417476	62.34546178	600	scf_8254731:284138-285821	19.458172
101	scf_8254716:371578-372459	62.13124881	601	scf_8253921:1-667	19.452181
102	scf_8254607:201058-203331	61.87049681	602	scf_8254179:36869-37944	19.377621
103	scf_8254575:147777-148972	61.56529999	603	scf_8254820:37646-38562	19.362701

104	scf_8254803:2109380-2110204	61.51576164	604	scf_8254594:675495-676284	19.354695
105	scf_8254511:14421-15343	61.22127231	605	scf_8253891:109694-110564	19.353796
106	scf_8254600:434056-434955	60.77295151	606	scf_8254650:137002-138002	19.342503
107	scf_8254667:209737-209919	60.61180843	607	scf_8254752:1179338-1180678	19.340326
108	scf_8254563:23900-24871	60.55614339	608	scf_8254600:559378-560386	19.322597
109	scf_8254401:61074-62383	60.07466777	609	scf_8254528:264403-264843	19.308151
110	scf_8254396:2665-3075	59.8571062	610	scf_8254605:204050-204579	19.283105
111	scf_8254665:59153-61024	59.3689697	611	scf_8254571:77756-80106	19.249993
112	scf_8254803:1791174-1792390	59.22562401	612	scf_8254803:1413419-1413929	19.234206
113	scf_8254650:320053-320338	59.16428935	613	scf_8254652:315327-316472	19.153128
114	scf_8254479:265405-265829	57.99081845	614	scf_8254823:141131-142046	19.10951
115	scf_8254406:504426-504724	57.72118781	615	scf_8254691:114644-115542	19.087197
116	scf_8253815:298879-299270	57.65249639	616	scf_8254028:174030-174732	19.081819
117	scf_8253899:45727-46389	57.50074794	617	scf_8254697:246373-247472	19.078224
118	scf_8254664:470548-471050	57.45321426	618	scf_8254459:395474-397620	19.03958
119	scf_8254475:181391-181854	57.40970727	619	scf_8253891:190015-191343	19.037671
120	scf_8254748:43787-45604	57.01568872	620	scf_8254592:452942-453480	18.864301
121	scf_8254659:1080198-1081476	56.70316322	621	scf_8254731:160679-160984	18.83347
122	scf_8253815:149743-150824	56.63899252	622	scf_8254823:187396-189272	18.802076
123	scf_8254548:521420-523544	56.60665442	623	scf_8254528:98853-99762	18.79777
124	scf_8254565:104436-105961	56.34846883	624	scf_8254527:291754-292593	18.794082
125	scf_8254580:253524-254894	56.31886376	625	scf_8254813:438045-438115	18.77907
126	scf_8254470:85961-87064	55.80659869	626	scf_8254611:482563-486245	18.754866
127	scf_8254545:76682-77285	55.66517349	627	scf_8253957:67902-69486	18.754288
128	scf_8254815:577796-578575	55.48225736	628	scf_8253880:1-1890	18.722957
129	scf_8254803:2039975-2040618	55.22194266	629	scf_8254548:422737-423420	18.704139
130	scf_8254403:469759-472193	54.70272384	630	scf_8254579:139202-140332	18.652705
131	scf_8254389:1639-2899	54.43874062	631	scf_8254638:180890-181341	18.651716
132	scf_8254654:70413-71866	54.09215099	632	scf_8254786:256664-257180	18.568171
133	scf_8254594:9455-10551	54.08083584	633	scf_8254581:22211-22453	18.464242
134	scf_8254719:484634-485565	54.00743028	634	scf_8254373:494990-496202	18.448501
135	scf_8254377:208218-209309	53.84630527	635	scf_8254564:391918-393241	18.413795
136	scf_8254582:603061-603339	53.59549322	636	scf_8254803:589563-590628	18.373358
137	scf_8254798:682597-683345	53.49605089	637	scf_8254600:287351-288891	18.372933
138	scf_8253811:472270-472870	53.49468309	638	scf_8254777:550485-550929	18.350962
139	scf_8254444:368886-369337	53.38710488	639	scf_8254807:99927-101127	18.317727

140	scf_8253891:112817-113633	53.30693935	640	scf_8254551:132434-134278	18.314811
141	scf_8254659:1422982-1425827	53.30156347	641	scf_8254654:265787-267357	18.254633
142	scf_8254555:304696-305777	53.06088854	642	scf_8254391:92606-94387	18.25001
143	scf_8254691:258605-258825	52.89964089	643	scf_8254565:344644-345086	18.239269
144	scf_8254803:1866391-1867806	51.9384519	644	scf_8254486:147183-148447	18.213801
145	scf_8254814:348655-348882	51.72299152	645	scf_8253992:4027-5962	18.193402
146	scf_8254428:250038-251230	51.67541712	646	scf_8254752:367031-367421	18.1919
147	scf_8254446:50119-50681	51.63742938	647	scf_8254636:416971-419672	18.173547
148	scf_8254752:320157-320731	51.12436267	648	scf_8254667:856057-856863	18.171579
149	scf_8254505:107828-109878	50.75715193	649	scf_8254747:237066-238021	18.163561
150	scf_8254823:27177-28355	50.74086396	650	scf_8254435:247094-251603	18.143274
151	scf_8254823:309603-310771	50.19917252	651	scf_8254814:405899-406285	18.140098
152	scf_8254719:726464-727844	49.99922957	652	scf_8254788:358207-360890	18.103887
153	scf_8254751:150970-152027	49.8375729	653	scf_8254446:449466-450198	18.0892
154	scf_8254589:67261-68462	49.47121529	654	scf_8254551:270999-272208	18.070271
155	scf_8254563:217018-218336	49.35075601	655	scf_8254579:312835-314151	18.059432
156	scf_8254823:183756-185716	48.93644264	656	scf_8254788:39889-43751	18.006555
157	scf_8254593:652369-653080	48.83283079	657	scf_8254545:916578-917671	17.9979
158	scf_8254527:333985-334520	48.77010303	658	scf_8254495:243537-244389	17.984327
159	scf_8254788:355935-357022	48.6682791	659	scf_8254565:99818-100573	17.934658
160	scf_8254593:177791-177980	48.61534944	660	scf_8254658:287372-288154	17.911422
161	scf_8254671:505973-507505	48.38880197	661	scf_8253994:36206-36815	17.831391
162	scf_8254495:57999-58403	48.20672275	662	scf_8254672:172600-173677	17.817119
163	scf_8254814:657213-658820	48.1049012	663	scf_8254638:716021-719164	17.795432
164	scf_8254697:892320-892988	47.19632821	664	scf_8254747:413983-414280	17.780246
165	scf_8254617:894097-894410	46.70526359	665	scf_8254648:122680-126948	17.774845
166	scf_8254543:121158-122598	46.66569345	666	scf_8253915:66833-67958	17.733854
167	scf_8254515:200161-200939	46.58838293	667	scf_8253908:84448-85134	17.725024
168	scf_8254811:311879-312615	46.56952132	668	scf_8254077:3042-4318	17.664611
169	scf_8254587:446473-447263	46.3343243	669	scf_8254653:60458-61049	17.663972
170	scf_8253887:266805-267004	46.0961956	670	scf_8254369:6763-7876	17.659819
171	scf_8254403:42780-45296	45.78587323	671	scf_8254659:942876-944095	17.643673
172	scf_8254732:27875-29301	45.75473146	672	scf_8254667:2041903-2042633	17.598877
173	scf_8254373:257471-258639	45.50532806	673	scf_8254813:456166-457874	17.56983
174	scf_8253815:495151-495741	45.4587646	674	scf_8254659:1018549-1019777	17.552291
175	scf_8254051:154419-154750	45.45794918	675	scf_8254527:265549-266120	17.532117

176	scf_8254379:709781-710265	45.28075745	676	scf_8253843:21196-22635	17.53077
177	scf_8254367:181764-182663	45.22062073	677	scf_8254652:356979-358551	17.500923
178	scf_8254644:240868-241495	44.86316831	678	scf_8254449:249881-250803	17.498192
179	scf_8254431:597158-598147	44.33300347	679	scf_8254763:425357-425484	17.496369
180	scf_8254410:31727-31999	43.79987713	680	scf_8254104:185817-186569	17.486468
181	scf_8254600:286757-286828	43.69765858	681	scf_8254495:240925-241543	17.424404
182	scf_8254459:30684-31331	43.36399167	682	scf_8254776:353259-353777	17.353228
183	scf_8254431:483449-487528	43.29789099	683	scf_8254638:771641-772597	17.338152
184	scf_8253815:490285-491998	43.22229341	684	scf_8254572:170101-170921	17.335061
185	scf_8254431:102180-103059	43.21178301	685	scf_8254811:56166-58236	17.329793
186	scf_8254106:949-2169	43.20685547	686	scf_8254491:96244-97029	17.319344
187	scf_8254475:307790-309089	43.20603931	687	scf_8254594:204120-205697	17.313248
188	scf_8254634:100892-103380	43.17208067	688	scf_8254600:701419-702381	17.312912
189	scf_8254085:67181-70217	42.89867623	689	scf_8254803:1084031-1084249	17.3082
190	scf_8254431:45155-46165	42.40701097	690	scf_8254548:30012-32018	17.285226
191	scf_8254417:265444-265741	42.35910439	691	scf_8254638:291460-292282	17.285052
192	scf_8254644:341456-343763	42.0822323	692	scf_8254788:16209-19186	17.258962
193	scf_8254823:348843-349365	41.8313853	693	scf_8254697:1707410-1708551	17.251286
194	scf_8254610:594832-595394	41.76752209	694	scf_8254028:109578-110125	17.247819
195	scf_8254665:624130-625089	41.76195279	695	scf_8254665:676826-679317	17.247025
196	scf_8254645:436149-437397	41.40640415	696	scf_8254752:1177780-1178515	17.218797
197	scf_8254360:58359-60413	40.92196466	697	scf_8254592:686625-686855	17.202484
198	scf_8254821:9525-9969	40.89557003	698	scf_8254010:183316-184861	17.199418
199	scf_8254527:52900-53486	40.87009816	699	scf_8254378:134927-136927	17.16208
200	scf_8253904:58603-61232	40.31103501	700	scf_8254397:35072-36231	17.15654
201	scf_8254798:886377-886775	40.25857235	701	scf_8254645:872014-873410	17.124866
202	scf_8254486:379359-379901	40.22575499	702	scf_8254482:3878-5023	17.112737
203	scf_8254563:32067-32516	39.96686799	703	scf_8254611:94579-98072	17.08078
204	scf_8254617:654890-655685	39.95632389	704	scf_8254667:1990190-1991607	17.075727
205	scf_8254777:410116-410888	39.55148985	705	scf_8254622:12807-13670	17.07401
206	scf_8254587:108915-109325	39.43498497	706	scf_8254803:107637-107909	17.073779
207	scf_8254716:188865-190050	39.33380924	707	scf_8254658:335467-337612	17.057264
208	scf_8254617:145638-146862	39.27597821	708	scf_8254545:86763-87600	17.045888
209	scf_8254370:248104-248819	39.21355224	709	scf_8254582:940927-943636	17.032728
210	scf_8254495:77303-78406	38.92730022	710	scf_8254781:7063-8042	16.993639
211	scf_8254667:1887049-1888691	38.84790109	711	scf_8254776:832887-834152	16.993189

212	scf_8254658:70301-71329	38.74002052	712	scf_8254607:198571-199054	16.974206
213	scf_8254594:666460-668344	38.6128085	713	scf_8254650:757505-758225	16.944365
214	scf_8254803:1799578-1800123	38.39588645	714	scf_8254104:178389-179550	16.935149
215	scf_8253815:261459-261885	38.20173375	715	scf_8254697:192306-193245	16.91613
216	scf_8254527:235156-235702	38.0684989	716	scf_8254542:54354-54820	16.867568
217	scf_8254617:350404-350726	38.05454695	717	scf_8254665:294848-295804	16.845951
218	scf_8254582:486578-492328	37.99198802	718	scf_8254472:1-1559	16.844823
219	scf_8254775:18227-18626	37.91728624	719	scf_8254444:31401-34108	16.835157
220	scf_8254367:263522-264037	37.88370533	720	scf_8253823:161917-162466	16.8273
221	scf_8254748:29108-29223	37.87856309	721	scf_8254671:358240-358931	16.810742
222	scf_8254527:14340-15394	37.8712535	722	scf_8254460:330938-331897	16.788805
223	scf_8254807:463658-463717	37.53935168	723	scf_8254707:127583-127801	16.786739
224	scf_8254648:4366-5185	37.53739278	724	scf_8254564:659074-659859	16.772511
225	scf_8254545:31694-32487	37.48229551	725	scf_8254807:429060-430014	16.73363
226	scf_8254378:149384-152484	37.45738146	726	scf_8254581:22820-24068	16.684285
227	scf_8254649:476794-477003	37.42421561	727	scf_8254460:339485-339785	16.642825
228	scf_8254576:184696-187620	37.26877253	728	scf_8254378:376735-377919	16.62699
229	scf_8254479:329802-329951	37.24324127	729	scf_8254422:193407-194051	16.621461
230	scf_8254689:16425-19542	37.15161737	730	scf_8254594:635765-636407	16.57514
231	scf_8254686:120224-121710	37.10258017	731	scf_8254475:127001-127574	16.562845
232	scf_8254187:224519-225115	37.07275337	732	scf_8254457:135547-137920	16.55449
233	scf_8254580:364277-365762	37.06188061	733	scf_8254803:41459-41651	16.543917
234	scf_8253984:39356-40414	36.98051966	734	scf_8254572:81796-82917	16.53252
235	scf_8254577:162807-163890	36.93318254	735	scf_8254075:20948-21952	16.512366
236	scf_8254803:1423510-1423555	36.8846642	736	scf_8254719:14365-14861	16.506619
237	scf_8254544:86586-87327	36.87347783	737	scf_8253823:48300-49107	16.501363
238	scf_8254645:954605-955433	36.74713337	738	scf_8254187:239012-239357	16.494662
239	scf_8254609:208620-209523	36.74511608	739	scf_8254371:528579-529076	16.489721
240	scf_8254707:20603-21659	36.54431	740	scf_8254028:45969-46940	16.481096
241	scf_8254551:599580-600832	36.48444175	741	scf_8254697:1342717-1343843	16.444457
242	scf_8254185:84327-85665	35.78935066	742	scf_8254716:1108152-1109234	16.427433
243	scf_8254803:1339721-1340791	35.75363019	743	scf_8254638:410062-411257	16.417144
244	scf_8254545:29850-31272	35.72054768	744	scf_8254555:341083-342973	16.370221
245	scf_8254697:248134-249225	35.7182582	745	scf_8254527:365367-367612	16.358345
246	scf_8254747:590382-591218	35.59409195	746	scf_8254555:53564-57435	16.338761
247	scf_8254779:94399-95613	35.50084526	747	scf_8254798:845807-846227	16.32373

248	scf_8254402:273048-274415	35.28897736	748	scf_8254665:399342-400415	16.293996
249	scf_8254486:28807-29367	35.24477898	749	scf_8254437:445442-446578	16.286595
250	scf_8254716:69396-70096	35.12049206	750	scf_8253815:411476-412347	16.275135
251	scf_8254446:1-1214	35.04074363	751	scf_8254776:112397-113316	16.269031
252	scf_8254403:349566-350717	34.97616272	752	scf_8254652:263270-265224	16.263625
253	scf_8254600:463255-464187	34.92587413	753	scf_8254790:7388-9030	16.221798
254	scf_8254654:212462-213961	34.86391823	754	scf_8254515:206071-206739	16.218727
255	scf_8254788:126711-127695	34.83965258	755	scf_8254650:576819-577697	16.216368
256	scf_8254179:34265-34709	34.80714447	756	scf_8254371:269807-271493	16.206066
257	scf_8254617:511100-512649	34.59884474	757	scf_8254600:814316-814862	16.181656
258	scf_8254085:44718-45765	34.49124265	758	scf_8253915:115849-116623	16.170826
259	scf_8254761:96929-97613	34.347393	759	scf_8254422:256345-256944	16.153349
260	scf_8254803:1641323-1643568	34.25150981	760	scf_8254638:171391-173199	16.149123
261	scf_8254515:103821-104483	34.09159996	761	scf_8254811:209508-210282	16.132158
262	scf_8254563:316191-316898	34.01323469	762	scf_8254260:31296-32052	16.096718
263	scf_8254813:460798-461642	33.99149789	763	scf_8254754:382896-383661	16.089803
264	scf_8254650:1156872-1160462	33.94789819	764	scf_8254495:353010-353821	16.062111
265	scf_8254607:196028-196567	33.85549154	765	scf_8254764:222393-222998	16.022181
266	scf_8254594:336952-340028	33.79554999	766	scf_8254582:593295-594515	16.015201
267	scf_8254814:561270-561885	33.66812783	767	scf_8254564:125045-125826	15.962823
268	scf_8254476:96118-97290	33.46532557	768	scf_8254697:95781-96574	15.958715
269	scf_8254787:18246-18694	33.35897664	769	scf_8254666:73355-73879	15.918789
270	scf_8254365:230011-231037	33.29384347	770	scf_8253890:8403-9151	15.896803
271	scf_8254505:130062-131307	33.20306421	771	scf_8254549:8830-8989	15.884288
272	scf_8254491:149894-150823	33.09837862	772	scf_8254476:256440-256792	15.847852
273	scf_8254666:282229-282533	32.88876217	773	scf_8253811:341329-342305	15.844142
274	scf_8254422:181757-182622	32.8690731	774	scf_8254649:236210-237282	15.839286
275	scf_8254194:52748-53553	32.85235944	775	scf_8254564:492966-494166	15.832129
276	scf_8253811:445765-446914	32.78726085	776	scf_8254051:90901-91337	15.823641
277	scf_8253894:44551-45462	32.66738411	777	scf_8254798:352116-354575	15.802758
278	scf_8254803:1802556-1803110	32.47836994	778	scf_8254406:223966-224259	15.794499
279	scf_8254580:49562-54745	32.46679096	779	scf_8254611:112229-112728	15.779499
280	scf_8254470:427011-427960	32.42220709	780	scf_8254718:65140-66645	15.767207
281	scf_8254587:233466-238049	32.36254998	781	scf_8254176:9246-11359	15.743782
282	scf_8254731:31452-32695	32.28757908	782	scf_8254777:24099-24848	15.725433
283	scf_8254580:249725-250035	32.28532998	783	scf_8254728:32296-34076	15.70451

284	scf_8254609:547869-551538	32.20625401	784	scf_8254464:255320-256495	15.69587
285	scf_8253815:267284-267929	32.18620967	785	scf_8254483:104003-104358	15.687302
286	scf_8254600:190081-190502	32.18543547	786	scf_8254697:1168042-1168836	15.671533
287	scf_8254631:1-1764	32.16935306	787	scf_8254776:138910-139677	15.670467
288	scf_8254712:94864-95626	32.15537955	788	scf_8254571:125709-125944	15.646329
289	scf_8254181:115660-116064	31.83545196	789	scf_8254667:306737-308898	15.635994
290	scf_8254636:332696-333881	31.8136114	790	scf_8254659:1034056-1035841	15.605731
291	scf_8254373:476721-477654	31.73149768	791	scf_8254582:976790-977557	15.592788
292	scf_8254417:55221-56734	31.66691805	792	scf_8254638:19593-20197	15.585114
293	scf_8254659:89470-90303	31.63202139	793	scf_8254803:723680-724099	15.566665
294	scf_8254587:427806-429499	31.60184863	794	scf_8254654:205198-206360	15.536365
295	scf_8254697:1448184-1448935	31.42463957	795	scf_8254580:372619-373896	15.4977
296	scf_8254645:938437-939032	31.35156915	796	scf_8254378:140700-141825	15.484862
297	scf_8254545:815795-817097	31.34531599	797	scf_8254592:225222-226207	15.483075
298	scf_8254638:324415-326144	31.28219161	798	scf_8254731:441483-442153	15.463258
299	scf_8254811:1088292-1089439	31.2512685	799	scf_8254582:973789-974519	15.449341
300	scf_8254811:433216-433793	31.19902949	800	scf_8254528:215597-216523	15.444719
301	scf_8254650:90961-91886	31.17726212	801	scf_8254666:150621-151978	15.391021
302	scf_8254752:730791-731549	31.1558497	802	scf_8254617:372303-373930	15.385628
303	scf_8254665:486185-486900	31.13282691	803	scf_8254752:1285202-1286600	15.373027
304	scf_8254607:262139-262879	31.09313907	804	scf_8254688:563809-565763	15.333295
305	scf_8254379:604637-605106	30.99261907	805	scf_8254369:109220-109670	15.302306
306	scf_8254564:269121-269842	30.93468161	806	scf_8254431:398453-401506	15.296658
307	scf_8254476:210214-210810	30.93277674	807	scf_8254010:476571-477836	15.292562
308	scf_8254565:358327-359579	30.86231684	808	scf_8254362:236585-238759	15.28639
309	scf_8254638:695994-696770	30.85581619	809	scf_8254444:74227-74827	15.277627
310	scf_8254379:302243-302506	30.85048521	810	scf_8254607:432511-433401	15.273779
311	scf_8254617:575050-575512	30.7799176	811	scf_8254379:427025-429553	15.258352
312	scf_8254460:611162-611799	30.73101012	812	scf_8254485:113443-114231	15.23572
313	scf_8254667:1249666-1251242	30.55189213	813	scf_8254185:78611-79867	15.220086
314	scf_8253994:19521-21283	30.52837829	814	scf_8254431:912334-913205	15.21949
315	scf_8254814:345452-347184	30.29522902	815	scf_8254622:20896-22061	15.215634
316	scf_8254584:224064-227878	30.15859357	816	scf_8254470:591944-594339	15.213333
317	scf_8254563:34375-35656	30.0936466	817	scf_8254814:436950-440278	15.208484
318	scf_8254605:93612-96505	29.95334208	818	scf_8254811:1318970-1320051	15.199656
319	scf_8254716:75078-76491	29.86055966	819	scf_8254527:414481-414950	15.190531

320	scf_8254379:503513-503648	29.72984448	820	scf_8254563:256827-257283	15.184785
321	scf_8254799:176318-177038	29.72661368	821	scf_8254371:339799-341015	15.177582
322	scf_8254658:216526-218775	29.71956587	822	scf_8254798:243736-244663	15.173725
323	scf_8254664:145294-146162	29.60586186	823	scf_8254697:883606-884438	15.144205
324	scf_8254464:200611-201346	29.60368628	824	scf_8254609:347359-348682	15.14107
325	scf_8254028:101234-102814	29.5806734	825	scf_8254428:10870-11921	15.128142
326	scf_8254788:664844-666016	29.45232135	826	scf_8254470:833873-835388	15.118283
327	scf_8253811:83098-84014	29.41317866	827	scf_8254671:286025-287850	15.107058
328	scf_8254370:128954-129651	29.28481968	828	scf_8254646:127038-128033	15.105341
329	scf_8254592:139417-140458	29.14251116	829	scf_8255478:1-579	15.094443
330	scf_8254437:340969-341793	29.1258708	830	scf_8254104:124572-127184	15.081911
331	scf_8254644:174086-175430	29.11173564	831	scf_8254571:214887-216162	15.081907
332	scf_8254582:434450-434688	28.93112084	832	scf_8254428:405985-407770	15.059055
333	scf_8254736:249946-251240	28.92799197	833	scf_8254645:566953-568834	15.054233
334	scf_8254373:571998-573803	28.90622826	834	scf_8254652:344011-344698	15.053698
335	scf_8254650:1026524-1027548	28.90364296	835	scf_8254665:451823-453852	15.018241
336	scf_8254600:350177-351005	28.86088406	836	scf_8254667:805798-806616	15.013752
337	scf_8253855:75970-77868	28.80933324	837	scf_8254686:177944-178974	15.010728
338	scf_8254422:219517-220379	28.80734069	838	scf_8254582:833403-833840	15.007057
339	scf_8254475:366168-366739	28.76082521	839	scf_8254610:516296-517217	14.996276
340	scf_8254072:19592-20531	28.76043099	840	scf_8254433:36827-37313	14.991436
341	scf_8254475:355388-356388	28.71450403	841	scf_8253904:270721-271542	14.988714
342	scf_8254395:292381-292846	28.71045132	842	scf_8254677:76702-79054	14.972933
343	scf_8254823:302482-303348	28.65115385	843	scf_8254410:82302-82746	14.968564
344	scf_8254417:7100-7835	28.57508648	844	scf_8254803:2181470-2182757	14.945708
345	scf_8254664:299255-302276	28.5725398	845	scf_8254543:64457-66406	14.925727
346	scf_8253843:82307-84066	28.3360617	846	scf_8254776:360857-361459	14.919147
347	scf_8254565:188922-189367	28.33213211	847	scf_8254747:493147-493606	14.909008
348	scf_8254389:54301-56054	28.3020728	848	scf_8254572:176511-178950	14.892497
349	scf_8254609:332059-332983	28.27378585	849	scf_8254764:159553-160287	14.885576
350	scf_8254464:280463-281961	28.25985409	850	scf_8254367:234758-235288	14.878379
351	scf_8254610:330932-331971	28.20638869	851	scf_8254564:476907-478105	14.877957
352	scf_8254665:161809-162870	28.1913759	852	scf_8253891:204896-205466	14.869337
353	scf_8254659:739833-740435	28.05190119	853	scf_8254688:372369-374791	14.865072
354	scf_8254564:736387-737122	27.95151324	854	scf_8254617:526834-527912	14.813589
355	scf_8254776:388952-389847	27.88039406	855	scf_8254437:27009-27605	14.799612

356	scf_8254769:44700-45294	27.85155011	856	scf_8254377:1-1774	14.792241
357	scf_8254752:194929-195696	27.84670504	857	scf_8254776:428031-428977	14.791915
358	scf_8254428:51829-52738	27.84651411	858	scf_8254747:1007928-1008504	14.778371
359	scf_8254645:157906-157950	27.81759294	859	scf_8254813:195991-199252	14.69796
360	scf_8254580:73204-74101	27.80246886	860	scf_8254697:894591-895086	14.685011
361	scf_8254659:1015039-1016618	27.75763872	861	scf_8254747:615711-617753	14.682225
362	scf_8254362:270278-270911	27.70334167	862	scf_8254652:253873-254495	14.666564
363	scf_8253887:318759-320560	27.67088019	863	scf_8254691:178960-179683	14.666031
364	scf_8254752:24273-25146	27.5797557	864	scf_8254751:11230-12252	14.659183
365	scf_8254367:262768-263221	27.49067917	865	scf_8254630:4718-5974	14.658475
366	scf_8254670:212976-213306	27.46595789	866	scf_8254665:292958-293959	14.653979
367	scf_8254194:15049-15344	27.43436912	867	scf_8254431:18641-19273	14.640401
368	scf_8254691:5404-7041	27.40250851	868	scf_8254590:794683-795710	14.62877
369	scf_8253994:37260-37606	27.3259935	869	scf_8254697:858996-860089	14.599795
370	scf_8254798:688016-689116	27.20603232	870	scf_8254609:515326-517293	14.594521
371	scf_8254653:64093-65490	27.17215644	871	scf_8254548:714559-715636	14.573486
372	scf_8254582:782871-784000	27.14703705	872	scf_8254592:159660-160016	14.569156
373	scf_8254373:383862-384526	27.13550427	873	scf_8254659:112084-113757	14.552797
374	scf_8254638:177423-178493	27.08494207	874	scf_8254716:42080-42234	14.550331
375	scf_8254636:60530-61475	27.03515358	875	scf_8254590:826763-826903	14.529075
376	scf_8254719:570406-571227	26.90393437	876	scf_8254593:529671-531184	14.518395
377	scf_8254486:57576-59130	26.81929072	877	scf_8254403:508765-509021	14.507783
378	scf_8254712:173655-174341	26.80829376	878	scf_8254788:767657-769501	14.494294
379	scf_8254584:446002-446723	26.75736035	879	scf_8254686:508093-509038	14.47179
380	scf_8254582:1023201-1024399	26.73326718	880	scf_8254391:289526-290432	14.451898
381	scf_8254649:348967-349811	26.66193376	881	scf_8254637:178360-178832	14.447765
382	scf_8253891:44626-45235	26.58012031	882	scf_8254667:262815-264761	14.433362
383	scf_8254385:185616-187216	26.55408682	883	scf_8254593:357769-359230	14.424798
384	scf_8254395:254384-255006	26.54108042	884	scf_8254646:77054-77400	14.422325
385	scf_8254548:126742-127412	26.52506343	885	scf_8254515:251826-252537	14.404852
386	scf_8254691:626974-627611	26.50088237	886	scf_8254672:181476-182092	14.356097
387	scf_8254587:441863-442326	26.47775173	887	scf_8254716:819334-820524	14.344173
388	scf_8254548:330665-331005	26.42295339	888	scf_8254736:101971-102885	14.317579
389	scf_8254751:58391-60593	26.22943591	889	scf_8254563:276770-278188	14.315097
390	scf_8254431:1087930-1089354	26.21728601	890	scf_8254653:115715-118092	14.311599
391	scf_8254475:339857-341283	26.14639666	891	scf_8254688:729630-730834	14.295085

392	scf_8254731:155909-159877	26.14579558	892	scf_8254764:67064-69044	14.288428
393	scf_8254815:737757-738111	26.12094263	893	scf_8254811:808244-809563	14.274145
394	scf_8254760:46819-47825	26.00941534	894	scf_8254605:159493-160381	14.265369
395	scf_8254697:831923-833877	25.93740023	895	scf_8254747:841473-842261	14.244719
396	scf_8254552:119267-120211	25.8670189	896	scf_8253856:43640-44546	14.234405
397	scf_8254571:191216-192235	25.85442392	897	scf_8254544:24394-26003	14.219907
398	scf_8254752:604165-605189	25.80928211	898	scf_8254470:1182288-1182831	14.201814
399	scf_8254010:475243-476099	25.74788881	899	scf_8254658:164119-165012	14.170731
400	scf_8254775:7709-8128	25.73673225	900	scf_8254043:191707-191952	14.130784
401	scf_8254444:171169-171672	25.71630206	901	scf_8254444:316234-317957	14.106082
402	scf_8254654:324535-325129	25.65513457	902	scf_8254460:139146-140064	14.101229
403	scf_8254666:99976-101508	25.64348701	903	scf_8254605:97079-100489	14.08615
404	scf_8254617:830168-832176	25.64024977	904	scf_8254551:183664-184430	14.086116
405	scf_8254814:469370-469922	25.6263964	905	scf_8254106:81052-81741	14.074238
406	scf_8254496:58905-59580	25.62361963	906	scf_8254592:59736-60626	14.070982
407	scf_8254607:113561-114811	25.62296806	907	scf_8254697:533779-534630	14.061284
408	scf_8254670:141863-142477	25.59860189	908	scf_8254565:156713-156795	14.04712
409	scf_8254644:325124-326211	25.58526343	909	scf_8254187:98221-100400	14.030585
410	scf_8254391:131164-131835	25.52276665	910	scf_8254697:913267-914462	14.023246
411	scf_8254365:170437-171275	25.48590698	911	scf_8254377:83289-84699	14.023048
412	scf_8254682:37761-38216	25.42978355	912	scf_8253964:62933-63398	14.01374
413	scf_8254593:104912-105419	25.39276063	913	scf_8254697:839262-841701	14.007954
414	scf_8254748:64915-68207	25.30740454	914	scf_8254637:141501-142918	13.999957
415	scf_8254653:268862-270530	25.26732255	915	scf_8254491:141503-142037	13.998552
416	scf_8254636:483683-484697	25.25136468	916	scf_8254767:25282-25371	13.997218
417	scf_8254564:624000-624332	25.19028615	917	scf_8254449:234775-235939	13.99607
418	scf_8254645:664670-665718	25.17860271	918	scf_8254588:8922-9541	13.984762
419	scf_8254181:64422-65385	25.07945746	919	scf_8253930:97862-98791	13.984208
420	scf_8254600:611762-612081	24.98307737	920	scf_8254748:96493-97761	13.944621
421	scf_8254564:264125-264516	24.98078402	921	scf_8254649:210505-210873	13.939561
422	scf_8254577:94710-95028	24.93771554	922	scf_8254719:566192-566711	13.921551
423	scf_8254716:31924-33303	24.9145063	923	scf_8253984:169550-170556	13.917778
424	scf_8254585:51131-52380	24.90758131	924	scf_8254263:530-2366	13.901555
425	scf_8254555:589454-590760	24.89708622	925	scf_8254803:673922-674639	13.896666
426	scf_8254752:122679-123462	24.76686207	926	scf_8254403:121061-121336	13.847418
427	scf_8254609:82794-83402	24.75397557	927	scf_8254610:430298-431198	13.84539

428	scf_8254551:302324-303943	24.65893328	928	scf_8254593:579738-580128	13.811841
429	scf_8254487:467498-468915	24.55876992	929	scf_8254814:271149-273009	13.801928
430	scf_8254491:192009-192671	24.54441504	930	scf_8254617:105356-105809	13.798534
431	scf_8254428:155468-156700	24.50769607	931	scf_8254646:134888-136751	13.789245
432	scf_8254667:1651711-1654002	24.42334303	932	scf_8254645:899933-901013	13.780051
433	scf_8254580:305053-307889	24.40117424	933	scf_8254460:344152-345572	13.775492
434	scf_8254610:122205-122829	24.35816912	934	scf_8254716:526698-527321	13.762331
435	scf_8254815:732690-733337	24.35755873	935	scf_8254085:109051-109932	13.750653
436	scf_8254371:421588-422751	24.35648222	936	scf_8254691:585008-585659	13.719269
437	scf_8254646:78619-79598	24.22909973	937	scf_8254754:91874-96125	13.71558
438	scf_8254422:204930-206488	24.18105977	938	scf_8254440:111491-113194	13.714782
439	scf_8254373:272505-273126	24.14675638	939	scf_8254711:44611-46614	13.702681
440	scf_8254659:1060648-1061781	24.11961891	940	scf_8253815:471349-472560	13.696314
441	scf_8254422:61720-63591	24.0734449	941	scf_8254754:162676-164933	13.681678
442	scf_8254548:322395-323230	24.03319562	942	scf_8254440:208121-208699	13.666864
443	scf_8254803:1115570-1116526	23.93994939	943	scf_8254716:859393-860628	13.658644
444	scf_8254495:304407-305170	23.9390421	944	scf_8254459:176049-176553	13.621138
445	scf_8253894:182071-182714	23.86878131	945	scf_8254444:360528-363010	13.612255
446	scf_8254697:787653-788703	23.78281335	946	scf_8254594:442439-445144	13.611692
447	scf_8254697:194326-195304	23.73734362	947	scf_8254716:619207-619618	13.609129
448	scf_8254028:85237-86183	23.7316751	948	scf_8254504:27639-28173	13.595619
449	scf_8254820:280128-281932	23.7303236	949	scf_8254666:449240-450120	13.591298
450	scf_8254486:425853-428979	23.70805501	950	scf_8254751:142711-143963	13.581217
451	scf_8254395:94408-95591	23.68958214	951	scf_8254391:69043-69892	13.580168
452	scf_8254545:34204-34558	23.67407337	952	scf_8254085:119721-122393	13.556162
453	scf_8254716:683716-686314	23.67154132	953	scf_8254659:1580703-1583241	13.510746
454	scf_8254579:201149-202025	23.67032532	954	scf_8254528:40330-43437	13.510387
455	scf_8254697:1157726-1158572	23.64908773	955	scf_8254798:650694-652542	13.505608
456	scf_8254667:887597-888375	23.61354939	956	scf_8254645:28584-30924	13.497697
457	scf_8254747:383332-384473	23.57902032	957	scf_8254194:45011-46146	13.486474
458	scf_8254565:463110-464476	23.47614264	958	scf_8254650:1060202-1061678	13.474137
459	scf_8254565:483571-484363	23.3988754	959	scf_8254593:461230-462173	13.452228
460	scf_8254691:156555-157094	23.38837989	960	scf_8254712:77209-78226	13.446961
461	scf_8254811:1380065-1381368	23.37399103	961	scf_8254650:394880-395481	13.445121
462	scf_8254691:442585-445050	23.34500485	962	scf_8254428:120218-121248	13.440747
463	scf_8254440:191377-192252	23.30578323	963	scf_8254823:244760-245480	13.437289

464	scf_8254749:101936-105691	23.26949621	964	scf_8254600:314145-316160	13.426127
465	scf_8254564:798202-798635	23.24239264	965	scf_8254570:37495-38324	13.419129
466	scf_8254402:19204-19752	23.23969244	966	scf_8254823:293627-295691	13.40776
467	scf_8254579:57308-59154	23.23456162	967	scf_8254072:6614-7363	13.4033
468	scf_8254449:117826-119238	23.22322672	968	scf_8254373:73965-75148	13.385879
469	scf_8254594:456832-458167	23.21704787	969	scf_8254712:90300-90759	13.380055
470	scf_8254362:4052-4925	23.19918115	970	scf_8254798:120452-121383	13.375379
471	scf_8254747:844014-844531	23.19558368	971	scf_8254608:118958-120774	13.37461
472	scf_8253886:140275-140726	23.18442793	972	scf_8254803:1855053-1856514	13.338237
473	scf_8254590:837666-838361	23.11204215	973	scf_8254736:1-2196	13.303641
474	scf_8254397:125223-125578	22.94550367	974	scf_8254542:49385-49631	13.293929
475	scf_8254650:487096-488623	22.89669281	975	scf_8254491:140825-140940	13.283618
476	scf_8254811:65474-66888	22.80785941	976	scf_8254410:121046-122460	13.278866
477	scf_8254365:279862-280312	22.80102247	977	scf_8254686:214262-215240	13.271395
478	scf_8254776:508905-509743	22.79888776	978	scf_8254545:537752-538548	13.258245
479	scf_8254544:1-7727	22.77883424	979	scf_8254548:337487-339934	13.250811
480	scf_8254369:40492-41716	22.73771665	980	scf_8254811:339051-339769	13.249217
481	scf_8254747:799315-800385	22.64984635	981	scf_8254551:595809-596412	13.241666
482	scf_8254645:1056275-1057881	22.60562152	982	scf_8254659:1522221-1523178	13.234665
483	scf_8254707:9925-15393	22.60133423	983	scf_8254798:383608-388059	13.199523
484	scf_8254747:829471-830535	22.59647273	984	scf_8254379:672096-674024	13.19347
485	scf_8254431:111030-111537	22.57007286	985	scf_8254545:669349-670027	13.181964
486	scf_8254075:50216-50532	22.55050362	986	scf_8254752:929408-931830	13.155694
487	scf_8254593:651573-651747	22.54533493	987	scf_8254716:1053376-1054507	13.139365
488	scf_8254010:340825-341897	22.52122693	988	scf_8254590:93058-94313	13.134098
489	scf_8254788:405117-409987	22.51830271	989	scf_8254367:239079-239816	13.123784
490	scf_8254617:641977-642837	22.5179444	990	scf_8254459:468060-469594	13.120518
491	scf_8254649:208612-209897	22.4703833	991	scf_8254594:335348-336525	13.119044
492	scf_8254650:935495-936480	22.4446217	992	scf_8254428:265627-266190	13.11441
493	scf_8254752:1034115-1036770	22.441375	993	scf_8254716:249310-251821	13.110184
494	scf_8254610:280934-282114	22.40443445	994	scf_8254813:235275-239400	13.109632
495	scf_8254360:46787-47645	22.39530178	995	scf_8254637:209368-211541	13.109239
496	scf_8254697:1805372-1805820	22.39018498	996	scf_8254653:46370-48195	13.098211
497	scf_8254814:361507-362343	22.34245186	997	scf_8254592:280098-280970	13.088696
498	scf_8254593:522751-523852	22.31017848	998	scf_8254788:757081-757860	13.086357
499	scf_8254592:203214-204357	22.29958988	999	scf_8254362:235125-236098	13.067203

500	scf_8254611:1434-1915	22.29778208	1000	scf_8254659:1057067-1057985	13.063971
-----	-----------------------	-------------	------	-----------------------------	-----------

A.4: Common link between *Ibd1* and *Med31*

Common potential genes controlled by Ibd1 and Med31

#	Gene region	Gene_ID	Note	Gene Size (bp)	Enrichment Ibd1	Enrichment Med31	Expression Profile
1	scf_8254667:14645 57-1466483	TTHERM_0 0037090	Note="RPS29_40S_ribosomal_protein_S29"	1927	3.48	2.75	High expression
2	scf_8254594:19707- 21259	TTHERM_0 0285290	Note="ubiquitin-conjugating_enzyme_E2"	1553	2.56	30.49	High expression
3	scf_8254776:23582 9-237134	TTHERM_0 0161200	Note="hypothetical_protein"	1306	2.50	3.90	High expression
4	scf_8254776:67775 4-678311	TTHERM_0 0158040	Note="Ran-binding_protein_2_Ran- binding_domain_protein"	558	1.73	3.67	High expression
5	scf_8254431:98073 4-981210	TTHERM_0 0569490	Note="EF-hand_protein"	477	2.25	3.76	Moderate expression
6	scf_8254564:52440 6-530182	TTHERM_0 0219210	Note="Serine/Threonine_kinase_domain_pro tein"	5777	1.67	2.63	Moderate expression
7	scf_8254811:64706 7-647407	TTHERM_0 0062680	Note="hypothetical_protein"	341	2.62	67.38	Low to No- Expression
8	scf_8254716:11705 35-1170945	TTHERM_0 0107160	Note="transmembrane_protein_putative"	411	1.86	2.16	Low to No- Expression
9	scf_8254652:13557 7-136364	TTHERM_0 0378520	Note="dual_specificity_phosphatase_domain _protein"	788	1.86	2.58	Low to No- Expression
10	scf_8254395:31961 4-320211	TTHERM_0 0456980	Note="transmembrane_protein_putative"	598	1.76	1.56	Low to No- Expression
11	scf_8254479:10756 1-107728	TTHERM_0 00622779	Note="hypothetical_protein"	168	2.82	6.31	#N/A
12	scf_8253957:67734- 67901	TTHERM_0 01071451	Note="hypothetical_protein"	168	2.63	7.07	#N/A

13	scf_8254431:90234-0-902645	TTHERM_0 00571689	Note="hypothetical_protein"	306	2.58	2.76	#N/A
14	scf_8254688:59211-59505	TTHERM_0 00297069	Note="hypothetical_protein"	295	2.47	3.59	#N/A
15	scf_8254028:17742-18459	TTHERM_0 00670107	Note="poxvirus_A32_protein"	718	2.07	1.52	#N/A
16	scf_8254468:40569-40874	TTHERM_0 01044599	Note="transmembrane_protein_putative"	306	1.89	3.71	#N/A
17	scf_8254010:51389-9-514291	TTHERM_0 00329909	Note="cell_surface_immobilization_antigen"	393	1.79	34.00	#N/A
18	scf_8254811:88091-3-882047	TTHERM_0 00059119	Note="transmembrane_protein_putative"	1135	1.75	2.07	#N/A
19	scf_8254437:47658-6-477375	TTHERM_0 00446389	Note="hypothetical_protein"	790	1.63	5.84	#N/A
20	scf_8253880:11080-4-111513	TTHERM_0 01055719	Note="transmembrane_protein_putative"	710	1.53	2.00	#N/A

Common potential intergenic regions occupied by Ibd1 and Med31

#	Scaffold	size	Enrichment Ibd1	Enrichment Med31	#	Scaffold	size	Enrichment Ibd1	Enrichment Med31
1	scf_8254594:335348-336525	1178	5.68	2.76	67	scf_8254468:28563-31765	3203	2.05	3.14
2	scf_8254650:1227049-1230315	3267	5.68	13.12	68	scf_8254650:33836-35671	1836	2.01	5.43
3	scf_8254590:565657-566989	1333	4.94	9.42	69	scf_8253894:212787-214095	1309	1.99	2.11
4	scf_8254600:159363-159688	326	4.74	10.15	70	scf_8254716:155794-157915	2122	1.99	7.82
5	scf_8254731:24114-26115	2002	4.55	3.46	71	scf_8254187:33538-34349	812	1.98	2.05
6	scf_8254564:431690-432322	633	4.49	1.54	72	scf_8254476:210214-210810	597	1.97	1.57
7	scf_8254391:156473-156612	140	4.03	1.62	73	scf_8254397:60704-61453	750	1.97	30.93
8	scf_8254385:337896-338250	355	4.00	4.01	74	scf_8254747:820985-821887	903	1.97	39.28

9	scf_8254811:1273468-1273591	124	3.96	4.91	75	scf_8254610:372195-373960	1766	1.96	4.82
10	scf_8254359:170776-171721	946	3.83	5.12	76	scf_8254483:214034-217129	3096	1.96	4.35
11	scf_8254565:343422-343878	457	3.73	15.22	77	scf_8254403:408852-409228	377	1.95	52.90
12	scf_8254476:192413-193798	1386	3.70	7.11	78	scf_8254486:75807-76812	1006	1.94	6.03
13	scf_8254688:601867-602204	338	3.57	1.97	79	scf_8254367:165454-169235	3782	1.93	7.44
14	scf_8254600:409839-411012	1174	3.45	8.51	80	scf_8254403:305216-305870	655	1.90	2.61
15	scf_8253984:196361-199651	3291	3.33	5.23	81	scf_8253823:230381-231612	1232	1.89	4.24
16	scf_8254718:8558-9654	1097	3.21	4.69	82	scf_8254449:13377-13714	338	1.89	2.91
17	scf_8254571:204290-204785	496	3.20	1.97	83	scf_8253984:60507-60913	407	1.89	4.29
18	scf_8254428:272769-273366	598	3.17	1.76	84	scf_8254788:203397-204875	1479	1.89	6.88
19	scf_8254592:309006-309549	544	3.12	10.05	85	scf_8254788:800088-800956	869	1.88	1.53
20	scf_8254367:170736-171089	354	3.10	3.98	86	scf_8254590:21912-22276	365	1.87	4.36
21	scf_8254719:392569-392974	406	3.05	11.36	87	scf_8254449:2420-3008	589	1.86	9.05
22	scf_8254575:87685-87844	160	2.91	3.93	88	scf_8254548:654611-655175	565	1.86	5.54
23	scf_8254379:251626-252624	999	2.90	11.59	89	scf_8254811:112920-113334	415	1.86	5.40
24	scf_8254028:97178-97863	686	2.85	5.21	90	scf_8254747:462345-463026	682	1.85	29.86
25	scf_8254719:626079-626648	570	2.83	27.85	91	scf_8254803:927004-929291	2288	1.84	3.26
26	scf_8254814:103323-104356	1034	2.82	11.93	92	scf_8254551:526588-527675	1088	1.84	3.18

27	scf_8254385:100853-102274	1422	2.77	10.79	93	scf_8254580:47053-48930	1878	1.83	1.98
28	scf_8253887:15793-16679	887	2.73	1.64	94	scf_8254652:1-1093	1093	1.83	6.13
29	scf_8254602:53304-54323	1020	2.73	4.05	95	scf_8254431:1016661-1017141	481	1.82	7.04
30	scf_8254010:414440-415861	1422	2.70	1.92	96	scf_8254551:496625-497599	975	1.79	4.68
31	scf_8254489:96554-98836	2283	2.67	13.98	97	scf_8254470:337751-339418	1668	1.77	7.02
32	scf_8254401:106993-107097	105	2.65	11.64	98	scf_8254814:424299-425006	708	1.77	25.18
33	scf_8253915:188075-191456	3382	2.59	19.35	99	scf_8254686:450537-450856	320	1.76	3.21
34	scf_8254670:253979-254290	312	2.58	9.44	100	scf_8254563:381483-381965	483	1.76	2.41
35	scf_8254486:185224-185385	162	2.54	4.06	101	scf_8254777:97069-98346	1278	1.75	21.21
36	scf_8254422:6770-7527	758	2.54	5.23	102	scf_8254769:151106-151777	672	1.73	2.99
37	scf_8254446:348797-351798	3002	2.52	2.02	103	scf_8254587:559428-560330	903	1.73	4.36
38	scf_8254719:137740-139795	2056	2.52	2.60	104	scf_8254677:70864-72028	1165	1.72	6.18
39	scf_8254373:34850-35000	151	2.46	3.65	105	scf_8254010:640460-641140	681	1.72	1.58
40	scf_8254449:1-453	453	2.44	1.75	106	scf_8254422:219517-220379	863	1.72	31.16
41	scf_8254459:130948-132016	1069	2.41	2.91	107	scf_8254582:994040-995444	1405	1.71	28.81
42	scf_8254667:1576952-1577843	892	2.38	4.57	108	scf_8254648:189357-189817	461	1.71	6.79
43	scf_8254771:25925-26120	196	2.37	5.74	109	scf_8254697:1359492-1360447	956	1.70	6.58
44	scf_8254786:1-936	936	2.37	2.56	110	scf_8254406:36035-37045	1011	1.68	8.57
45	scf_8254497:6330-6949	620	2.34	29.11	111	scf_8254659:1253886-1254164	279	1.68	3.45

46	scf_8254747:853696-853869	174	2.31	12.24	112	scf_8254284:137894-138513	620	1.68	4.94
47	scf_8254688:707460-708574	1115	2.31	4.12	113	scf_8254776:681910-682447	538	1.67	7.91
48	scf_8254095:38074-38715	642	2.29	2.64	114	scf_8254446:385727-388592	2866	1.66	8.04
49	scf_8254563:78038-78401	364	2.27	1.59	115	scf_8254798:283661-285100	1440	1.66	14.31
50	scf_8254608:71007-71691	685	2.27	4.55	116	scf_8254446:190532-191374	843	1.66	5.42
51	scf_8254470:412970-415300	2331	2.26	2.89	117	scf_8254652:503153-504794	1642	1.65	13.45
52	scf_8254688:531361-531612	252	2.26	1.97	118	scf_8254527:130747-132790	2044	1.63	1.75
53	scf_8254362:22429-22863	435	2.24	16.26	119	scf_8254545:776231-780647	4417	1.63	5.22
54	scf_8254470:93361-95870	2510	2.22	2.76	120	scf_8254542:55538-56413	876	1.63	5.44
55	scf_8254378:318812-319277	466	2.18	12.10	121	scf_8254666:397353-398434	1082	1.60	2.85
56	scf_8254697:1371582-1372077	496	2.18	3.46	122	scf_8254199:101624-102451	828	1.60	8.67
57	scf_8254588:190863-191911	1049	2.18	10.49	123	scf_8254362:210737-211432	696	1.59	31.35
58	scf_8254667:1679866-1680721	856	2.18	2.21	124	scf_8254650:868463-869130	668	1.58	6.49
59	scf_8254551:194474-197602	3129	2.17	94.26	125	scf_8254747:236310-236603	294	1.57	4.41
60	scf_8254176:29916-30540	625	2.16	5.78	126	scf_8254717:34167-34707	541	1.56	4.60
61	scf_8254406:237288-238266	979	2.14	3.99	127	scf_8254185:113978-114561	584	1.56	2.69
62	scf_8253855:32955-35332	2378	2.11	2.21	128	scf_8254479:110692-111429	738	1.54	6.95
63	scf_8254610:392332-393659	1328	2.10	2.28	129	scf_8254548:548588-549133	546	1.53	73.85

64	scf_8253894:89068-93871	4804	2.08	31.28	130	scf_8253886:32878-33490	613	1.51	11.65
65	scf_8254549:28866-30263	1398	2.06	1.98	131	scf_8253887:72895-74836	1942	1.51	3.28
66	scf_8254665:305443-306107	665	2.06	2.31	132	scf_8254428:328765-329525	761	1.50	11.34

BIBLIOGRAPHY

- Adl, S. M., Simpson, A. G. B., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., ... Spiegel, F. W. (2012). The revised classification of eukaryotes. *The Journal of Eukaryotic Microbiology*, 59(5), 429–493. <https://doi.org/10.1111/j.1550-7408.2012.00644.x>
- Agalioti, T., Chen, G., & Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. *Cell*, 111(3), 381–392. [https://doi.org/10.1016/S0092-8674\(02\)01077-2](https://doi.org/10.1016/S0092-8674(02)01077-2)
- Akematsu, T., Findlay, A., Fukuda, Y., Pearlman, R. E., Loidl, J., Orias, E., & P Hamilton, E. (2018). Resistance to 6-Methylpurine is Conferred by Defective Adenine Phosphoribosyltransferase in Tetrahymena. *Genes*, 9(4). <https://doi.org/10.3390/genes9040179>
- Akematsu, T., Fukuda, Y., Garg, J., Fillingham, J. S., Pearlman, R. E., & Loidl, J. (2017). Post-meiotic DNA double-strand breaks occur in Tetrahymena, and require Topoisomerase II and Spo11. *ELife*, 6. <https://doi.org/10.7554/eLife.26176>
- Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, 92(3), 291–294. [https://doi.org/10.1016/S0092-8674\(00\)80922-8](https://doi.org/10.1016/S0092-8674(00)80922-8)
- Allard, S., Utley, R. T., Savard, J., Clarke, A., Grant, P., Brandl, C. J., ... Côté, J. (1999). NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO Journal*, 18(18), 5108–5119. <https://doi.org/10.1093/emboj/18.18.5108>
- Allis, C. D., Allen, R. L., Wiggins, J. C., Chicoine, L. G., & Richman, R. (1984). Proteolytic processing of h1-like histones in chromatin: a physiologically and developmentally regulated event in Tetrahymena micronuclei. *The Journal of Cell Biology*, 99(5), 1669–1677. <https://doi.org/10.1083/jcb.99.5.1669>
- Allis, C. D., Bowen, J. K., Abraham, G. N., Glover, C. V. C., & Gorovsky, M. A. (1980). Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in Tetrahymena micronuclei. *Cell*, 20(1), 55–64. [https://doi.org/10.1016/0092-8674\(80\)90234-2](https://doi.org/10.1016/0092-8674(80)90234-2)
- Allis, C. D., Chicoine, L. G., Richman, R., & Schulman, I. G. (1985). Deposition-related histone acetylation in micronuclei of conjugating Tetrahymena. *Proceedings of the National Academy of Sciences of the United States of America*, 82(23), 8048–8052. <https://doi.org/10.1073/pnas.82.23.8048>
- Allis, C. D., Glover, C. V. C., Bowen, J. K., & Gorovsky, M. A. (1980). Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, Tetrahymena thermophila. *Cell*, 20(3), 609–617. [https://doi.org/10.1016/0092-8674\(80\)90307-4](https://doi.org/10.1016/0092-8674(80)90307-4)
- Allis, C. D., Glover, C. V., & Gorovsky, M. A. (1979). Micronuclei of Tetrahymena contain two types of histone H3. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), 4857–4861. <https://doi.org/10.1073/pnas.76.10.4857>

- Allis, C. D., Richman, R., Gorovsky, M. a, Ziegler, Y. S., Touchstone, B., Bradley, W. a, & Cook, R. G. (1986). hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *The Journal of Biological Chemistry*, 261(4), 1941–1948.
- Aronica, L., Bednenko, J., Noto, T., DeSouza, L. V., Siu, K. W. M., Loidl, J., ... Mochizuki, K. (2008). Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in *Tetrahymena*. *Genes and Development*, 22(16), 2228–2241. <https://doi.org/10.1101/gad.481908>
- Ashraf, K., Nabeel-Shah, S., Garg, J., Saettone, A., Derynck, J., Gingras, A.-C., ... Fillingham, J. (2019). Proteomic analysis of histones H2A/H2B and variant Hv1 in *Tetrahymena thermophila* reveals an ancient network of chaperones. *Molecular Biology and Evolution*. <https://doi.org/10.1093/molbev/msz039>
- Baker, S. P., & Grant, P. A. (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene*, 26(37), 5329–5340. <https://doi.org/10.1038/sj.onc.1210603>
- Bednenko, J., Noto, T., DeSouza, L. V., Siu, K. W. M., Pearlman, R. E., Mochizuki, K., & Gorovsky, M. A. (2009). Two GW Repeat Proteins Interact with *Tetrahymena thermophila* Argonaute and Promote Genome Rearrangement. *Molecular and Cellular Biology*, 29(18), 5020–5030. <https://doi.org/10.1128/MCB.00076-09>
- Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, 447(7143), 407–412. <https://doi.org/10.1038/nature05915>
- Berndsen, C. E., & Denu, J. M. (2008). Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Current Opinion in Structural Biology*, 18(6), 682–689. <https://doi.org/10.1016/j.sbi.2008.11.004>
- Biterge, B., & Schneider, R. (2014). Histone variants: Key players of chromatin. *Cell and Tissue Research*, 356, 457–466.
- Blackburn, E. H. (1984). The molecular structure of centromeres and telomeres. *Annual Review of Biochemistry*, 53, 163–194. <https://doi.org/10.1146/annurev.bi.53.070184.001115>
- Bourbon, H.-M. (2008). Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Research*, 36, 3993–4008.
- Brownell, J. E., & Allis, C. D. (1995). An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. *Proceedings of the National Academy of Sciences*, 92(14), 6364–6368. <https://doi.org/10.1073/pnas.92.14.6364>
- Brownell, James E, Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., & Allis, C. D. (1996). *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, 84(6), 843–851. [https://doi.org/10.1016/S0092-8674\(00\)81063-6](https://doi.org/10.1016/S0092-8674(00)81063-6)
- Buchan, D. W. A., & Jones, D. T. (2019). The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkz297>
- Butland, G., Peregrín-Alvarez, J. M., Li, J., Yang, W., Yang, X., Canadien, V., ... Emili, A.

- (2005). Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*, 433(7025), 531–537. <https://doi.org/10.1038/nature03239>
- Byrne, A., Beaudin, A. E., Olsen, H. E., Jain, M., Cole, C., Palmer, T., ... Vollmers, C. (2017). Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nature Communications*, 8, 16027. <https://doi.org/10.1038/ncomms16027>
- Chalker, D L, & Yao, M. C. (1996). Non-Mendelian, heritable blocks to DNA rearrangement are induced by loading the somatic nucleus of *Tetrahymena thermophila* with germ line-limited DNA. *Molecular and Cellular Biology*, 16(7), 3658–3667. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8668182>
- Chalker, Douglas L., & Yao, M. C. (2001). Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*. *Genes & Development*, 15(10), 1287–1298. <https://doi.org/10.1101/gad.884601>
- Chalker, Douglas L. (2012). *Transformation and Strain Engineering of Tetrahymena*. <https://doi.org/10.1016/B978-0-12-385967-9.00011-6>
- Chalker, Douglas L, & Yao, M.-C. (2011). DNA elimination in ciliates: transposon domestication and genome surveillance. *Annual Review of Genetics*, 45, 227–246. <https://doi.org/10.1146/annurev-genet-110410-132432>
- Chang, P.-Y., Hom, R. A., Musselman, C. A., Zhu, L., Kuo, A., Gozani, O., ... Cleary, M. L. (2010). Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. *Journal of Molecular Biology*, 400(2), 137–144. <https://doi.org/10.1016/j.jmb.2010.05.005>
- Chicoine, L. G., Richman, R., Cook, R. G., Gorovsky, M. A., & Allis, C. D. (1987). A single histone acetyltransferase from *Tetrahymena* macronuclei catalyzes deposition-related acetylation of free histones and transcription-related acetylation of nucleosomal histones. *Journal of Cell Biology*, 105(1), 127–135.
- Chicoine, L. G., Schulman, I. G., Richman, R., Cook, R. G., & Allis, C. D. (1986). Nonrandom utilization of acetylation sites in histones isolated from *Tetrahymena*. Evidence for functionally distinct H4 acetylation sites. *The Journal of Biological Chemistry*, 261(3), 1071–1076. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3080415>
- Chittuluru, J. R., Chaban, Y., Monnet-Saksouk, J., Carrozza, M. J., Sapountzi, V., Selleck, W., ... Asturias, F. J. (2012). Structure and nucleosome interaction of the yeast NuA4 and Piccolo-NuA4 histone acetyltransferase complexes. *Nat Struct Mol Biol*, 18(11), 1196–1203. <https://doi.org/10.1038/nsmb.2128.Structure>
- Cieniewicz, A. M., Moreland, L., Ringel, A. E., Mackintosh, S. G., Raman, A., Gilbert, T. M., ... Taverna, S. D. (2014). The bromodomain of Gcn5 regulates site specificity of lysine acetylation on histone H3. *Molecular & Cellular Proteomics : MCP*, 13(11), 2896–2910. <https://doi.org/10.1074/mcp.M114.038174>
- Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature Reviews*.

- Molecular Cell Biology*, 18(7), 407–422. <https://doi.org/10.1038/nrm.2017.26>
- Clark, A. D., Oldenbroek, M., & Boyer, T. G. (2015). Mediator kinase module and human tumorigenesis. *Critical Reviews in Biochemistry and Molecular Biology*, 50(5), 393–426. <https://doi.org/10.3109/10409238.2015.1064854>
- Collins, K., & Gandhi, L. (1998). The reverse transcriptase component of the Tetrahymena telomerase ribonucleoprotein complex. *Proceedings of the National Academy of Sciences of the United States of America*, 95(15), 8485–8490. <https://doi.org/10.1073/pnas.95.15.8485>
- Collins, K., Kobayashi, R., & Greider, C. W. (1995). Purification of tetrahymena telomerase and cloning of genes encoding the two protein components of the enzyme. *Cell*, 81(5), 677–686. [https://doi.org/10.1016/0092-8674\(95\)90529-4](https://doi.org/10.1016/0092-8674(95)90529-4)
- Cremer, T., & Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Reviews. Genetics*, 2(4), 292–301. <https://doi.org/10.1038/35066075>
- Díaz-Viraqué, F., Pita, S., Greif, G., de Souza, R. de C. M., Iraola, G., & Robello, C. (2019). Nanopore Sequencing Significantly Improves Genome Assembly of the Protozoan Parasite *Trypanosoma cruzi*. *Genome Biology and Evolution*, 11(7), 1952–1957. <https://doi.org/10.1093/gbe/evz129>
- Dou, Y., Mizzen, C. A., Abrams, M., Allis, C. D., & Gorovsky, M. A. (1999). Phosphorylation of linker histone H1 regulates gene expression in vivo by mimicking H1 removal. *Molecular Cell*, 4(4), 641–647. [https://doi.org/10.1016/S1097-2765\(00\)80215-4](https://doi.org/10.1016/S1097-2765(00)80215-4)
- Dou, Y., Song, X., Liu, Y., & Gorovsky, M. A. (2005). The H1 phosphorylation state regulates expression of CDC2 and other genes in response to starvation in *Tetrahymena thermophila*. *Molecular and Cellular Biology*, 25(10), 3914–3922. <https://doi.org/10.1128/MCB.25.10.3914-3922.2005>
- Eisen, J. A., Coyne, R. S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J. R., ... Orias, E. (2006). Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biology*, 4(9), e286. <https://doi.org/10.1371/journal.pbio.0040286>
- Emili, A., Greenblatt, J., & Ingles, C. J. (1994). Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. *Molecular and Cellular Biology*, 14(3), 1582–1593. <https://doi.org/PMC358517>
- Filippakopoulos, P., & Knapp, S. (2014). Targeting bromodomains: epigenetic readers of lysine acetylation. *Nature Reviews Drug Discovery*, 13(5), 337–356. <https://doi.org/10.1038/nrd4286>
- Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., ... Knapp, S. (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*, 149(1), 214–231. <https://doi.org/10.1016/j.cell.2012.02.013>
- Fillingham, J., Garg, J., Tsao, N., Vythilingum, N., Nishikawa, T., & Pearlman, R. E. (2006a). Molecular genetic analysis of an SNF2/brhma-related gene in *Tetrahymena thermophila* suggests roles in growth and nuclear development. *Eukaryotic Cell*, 5, 1347–1359. <https://doi.org/10.1128/EC.00149-06>

- Fillingham, J., Garg, J., Tsao, N., Vythilingum, N., Nishikawa, T., & Pearlman, R. E. (2006b). Molecular genetic analysis of an SNF2/brahma-related gene in *Tetrahymena thermophila* suggests roles in growth and nuclear development. *Eukaryotic Cell*, 5(8), 1347–1359. <https://doi.org/10.1128/EC.00149-06>
- Fillingham, J. S., Bruno, D., & Pearlman, R. E. (2001). Cis-acting requirements in flanking DNA for the programmed elimination of mse2.9: a common mechanism for deletion of internal eliminated sequences from the developing macronucleus of *Tetrahymena thermophila*. *Nucleic Acids Res*, 29(2), 488–498.
- Gaertig, J., Gao, Y., Tishgarten, T., Clark, T. G., & Dickerson, H. W. (1999). Surface display of a parasite antigen in the ciliate *Tetrahymena thermophila*. *Nature Biotechnology*, 17(5), 462–465. <https://doi.org/10.1038/8638>
- Gao, S., & Liu, Y. (2012). Intercepting noncoding messages between germline and soma. *Genes and Development*, 26(16), 1774–1779. <https://doi.org/10.1101/gad.199992.112>
- Gao, S., Xiong, J., Zhang, C., Berquist, B. R., Yang, R., Zhao, M., ... Liu, Y. (2013). Impaired replication elongation in *Tetrahymena* mutants deficient in histone H3 Lys 27 monomethylation. *Genes & Development*, 27(15), 1662–1679. <https://doi.org/10.1101/gad.218966.113>
- Garcia, B. A., Hake, S. B., Diaz, R. L., Kauer, M., Morris, S. A., Recht, J., ... Hunt, D. F. (2007). Organismal differences in post-translational modifications in histones H3 and H4. *The Journal of Biological Chemistry*, 282(10), 7641–7655. <https://doi.org/10.1074/jbc.M607900200>
- Gardner, K. E., Allis, C. D., & Strahl, B. D. (2011). Operating on chromatin, a colorful language where context matters. *Journal of Molecular Biology*, 409(1), 36–46. <https://doi.org/10.1016/j.jmb.2011.01.040>
- Garg, J., Lambert, J.-P., Karsou, A., Marquez, S., Nabeel-Shah, S., Bertucci, V., ... Fillingham, J. S. (2013). Conserved Asf1-importin β physical interaction in growth and sexual development in the ciliate *Tetrahymena thermophila*. *Journal of Proteomics*, 94, 311–326. <https://doi.org/10.1016/j.jprot.2013.09.018>
- Garg, J., Saettone, A., Nabeel-Shah, S., Cadorin, M., Ponce, M., Marquez, S., ... Fillingham, J. (2019). The Med31 Conserved Component of the Divergent Mediator Complex in *Tetrahymena thermophila* Participates in Developmental Regulation. *Current Biology : CB*, 29(14), 2371–2379.e6. <https://doi.org/10.1016/j.cub.2019.06.052>
- Gene Ontology, C. (2000). Gene ontology: Tool for the identification of biology. *Natural Genetics*, 25(may), 25–29.
- Georgakopoulos, T., & Thireos, G. (1992). Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *The EMBO Journal*, 11(11), 4145–4152. <https://doi.org/10.1002/j.1460-2075.1992.tb05507.x>
- Gingras, A.-C., Abe, K. T., & Raught, B. (2019). Getting to know the neighborhood: using proximity-dependent biotinylation to characterize protein complexes and map organelles. *Current Opinion in Chemical Biology*, 48, 44–54.

<https://doi.org/10.1016/j.cbpa.2018.10.017>

- Ginsburg, D. S., Anlembom, T. E., Wang, J., Patel, S. R., Li, B., & Hinnebusch, A. G. (2014). NuA4 links methylation of histone H3 lysines 4 and 36 to acetylation of histones H4 and H3. *Journal of Biological Chemistry*, 289(47), 32656–32670. <https://doi.org/10.1074/jbc.M114.585588>
- Gorovsky, M. A. (1970a). Studies on nuclear structure and function in *Tetrahymena pyriformis*. 3. Comparison of the histones of macro- and micronuclei by quantitative polyacrylamide gel electrophoresis. *The Journal of Cell Biology*, 47(3), 631–636. <https://doi.org/10.1083/jcb.47.3.631>
- Gorovsky, M. A. (1970b). Studies on nuclear structure and function in *Tetrahymena pyriformis*. II. Isolation of macro- and micronuclei. *The Journal of Cell Biology*, 47(3), 619–630. <https://doi.org/10.1083/jcb.47.3.619>
- Gradolatto, A., Smart, S. K., Byrum, S., Blair, L. P., Rogers, R. S., Kolar, E. A., ... Tackett, A. J. (2009). A noncanonical bromodomain in the AAA ATPase protein Yta7 directs chromosomal positioning and barrier chromatin activity. *Molecular and Cellular Biology*, 29(17), 4604–4611. <https://doi.org/10.1128/MCB.00160-09>
- Grant, P. A., Schieltz, D., Pray-Grant, M. G., Yates III, J. R., & Workman, J. L. (1998). The ATM-related cofactor Tra1 is a component of the purified SAGA complex. *Molecular Cell*, 2(6), 863–867. [https://doi.org/10.1016/s1097-2765\(00\)80300-7](https://doi.org/10.1016/s1097-2765(00)80300-7)
- Greider, C. W., & Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in *tetrahymena* extracts. *Cell*, 43(2 PART 1), 405–413. [https://doi.org/10.1016/0092-8674\(85\)90170-9](https://doi.org/10.1016/0092-8674(85)90170-9)
- Greider, C. W., & Blackburn, E. H. (1987). The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*, 51(6), 887–898. [https://doi.org/10.1016/0092-8674\(87\)90576-9](https://doi.org/10.1016/0092-8674(87)90576-9)
- Guarente, L. (1995). Transcriptional coactivators in yeast and beyond. *Trends in Biochemical Sciences*, 20(12), 517–521. [https://doi.org/10.1016/S0968-0004\(00\)89120-3](https://doi.org/10.1016/S0968-0004(00)89120-3)
- Guglielmi, B., van Berkum, N. L., Klapholz, B., Bijma, T., Boube, M., Boschiero, C., ... Werner, M. (2004). A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Research*, 32(18), 5379–5391. <https://doi.org/10.1093/nar/gkh878>
- Guruharsha, K. G., Rual, J.-F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N., ... Artavanis-Tsakonas, S. (2011). A protein complex network of *Drosophila melanogaster*. *Cell*, 147(3), 690–703. <https://doi.org/10.1016/j.cell.2011.08.047>
- Hamilton, E. P., Kapusta, A., Huvos, P. E., Bidwell, S. L., Zafar, N., Tang, H., ... Coyne, R. S. (2016). Structure of the germline genome of *Tetrahymena thermophila* and relationship to the massively rearranged somatic genome. *eLife*, 5. <https://doi.org/10.7554/eLife.19090>
- Han, R., Li, Y., Gao, X., & Wang, S. (2018). An accurate and rapid continuous wavelet dynamic time warping algorithm for end-to-end mapping in ultra-long nanopore sequencing. *Bioinformatics (Oxford, England)*, 34(17), i722–i731. <https://doi.org/10.1093/bioinformatics/bty555>

- Hassan, A. H., Awad, S., Al-Natour, Z., Othman, S., Mustafa, F., & Rizvi, T. A. (2007). Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. *Biochemical Journal*, 402(1), 125–133. <https://doi.org/10.1042/BJ20060907>
- Hassan, A. H., Prochasson, P., Neely, K. E., Galasinski, S. C., Chandy, M., Carrozza, M. J., & Workman, J. L. (2002). Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*, 111(3), 369–379. [https://doi.org/10.1016/S0092-8674\(02\)01005-X](https://doi.org/10.1016/S0092-8674(02)01005-X)
- Hengartner, C. J., Thompson, C. M., Zhang, J., Chao, D. M., Liao, S. M., Koleske, A. J., ... Young, R. A. (1995). Association of an activator with an RNA polymerase II holoenzyme. *Genes & Development*, 9(8), 897–910.
- Hu, P., Janga, S. C., Babu, M., Díaz-Mejía, J. J., Butland, G., Yang, W., ... Emili, A. (2009). Global functional atlas of Escherichia coli encompassing previously uncharacterized proteins. *PLoS Biology*, 7(4), e96. <https://doi.org/10.1371/journal.pbio.1000096>
- Huang, H., Wiley, E. A., Lending, C. R., & Allis, C. D. (1998). An HP1-like protein is missing from transcriptionally silent micronuclei of Tetrahymena. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23), 13624–13629. <https://doi.org/10.1073/pnas.95.23.13624>
- Hutchcroft, J. E., Anostario, M., Harrison, M. L., & Geahlen, R. L. (1991). Renaturation and assay of protein kinases after electrophoresis in sodium dodecyl sulfate-polyacrylamide gels. *Methods in Enzymology*, 200, 417–423. [https://doi.org/10.1016/0076-6879\(91\)00157-R](https://doi.org/10.1016/0076-6879(91)00157-R)
- Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Cannon, J. R., Ting, L., Baltier, K., ... Harper, J. W. (2017). Architecture of the human interactome defines protein communities and disease networks. *Nature*, 545(7655), 505–509. <https://doi.org/10.1038/nature22366>
- Jacob, N. K., Lescasse, R., Linger, B. R., & Price, C. M. (2007). Tetrahymena POT1a regulates telomere length and prevents activation of a cell cycle checkpoint. *Molecular and Cellular Biology*, 27(5), 1592–1601. <https://doi.org/10.1128/MCB.01975-06>
- Jain, A. K., & Barton, M. C. (2017). Bromodomain Histone Readers and Cancer. *Journal of Molecular Biology*, 429(13), 2003–2010. <https://doi.org/10.1016/j.jmb.2016.11.020>
- James, P. (1997). Protein identification in the post-genome era: the rapid rise of proteomics. *Quarterly Reviews of Biophysics*, 30(4), 279–331. <https://doi.org/10.1017/S0033583597003399>
- Jeronimo, C., & Robert, F. (2017). The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. *Trends in Cell Biology*, 27(10), 765–783. <https://doi.org/10.1016/J.TCB.2017.07.001>
- Jiang, J., Chan, H., Cash, D. D., Miracco, E. J., Loo, R. R. O., Upton, H. E., ... Feigon, J. (2015). Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science (New York, N.Y.)*, 350(6260), aab4070. <https://doi.org/10.1126/science.aab4070>
- Johnson, D. S., Mortazavi, A., Myers, R. M., & Wold, B. (2007). Genome-Wide Mapping of in

- Vivo Protein-DNA Interactions. *Science*, 316(5830), 1497–1502.
<https://doi.org/10.1126/science.1141319>
- Karijolich, J. J., & Hampsey, M. (2012). The Mediator complex. *Current Biology*, 22(24), R1030–R1031. <https://doi.org/10.1016/J.CUB.2012.11.011>
- Kataoka, K., & Mochizuki, K. (2011). Programmed DNA elimination in *Tetrahymena*: a small RNA-mediated genome surveillance mechanism. *Advances in Experimental Medicine and Biology*, 722, 156–173. https://doi.org/10.1007/978-1-4614-0332-6_10
- Kataoka, K., & Mochizuki, K. (2015). Phosphorylation of an HP1-like Protein Regulates Heterochromatin Body Assembly for DNA Elimination. *Developmental Cell*, 35(6), 775–788. <https://doi.org/10.1016/j.devcel.2015.11.017>
- Kim, B., Nesvizhskii, A. I., Rani, P. G., Hahn, S., Aebersold, R., & Ranish, J. A. (2007). The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41), 16068–16073. <https://doi.org/10.1073/pnas.0704573104>
- Kim, H., Vanoosthuyse, V., Fillingham, J., Roguev, A., Watt, S., Kislinger, T., ... Keogh, M.-C. (2009). An acetylated form of histone H2A.Z regulates chromosome architecture in *Schizosaccharomyces pombe*. *Nature Structural & Molecular Biology*, 16(12), 1286–1293. <https://doi.org/10.1038/nsmb.1688>
- Kim, Y. J., Zheng, B., Yu, Y., Won, S. Y., Mo, B., & Chen, X. (2011). The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. *The EMBO Journal*, 30(5), 814–822. <https://doi.org/10.1038/emboj.2011.3>
- Knutson, B. A., & Hahn, S. (2011). Domains of Tra1 important for activator recruitment and transcription coactivator functions of SAGA and NuA4 complexes. *Molecular and Cellular Biology*, 31(4), 818–831. <https://doi.org/10.1128/MCB.00687-10>
- Köcher, T., & Superti-Furga, G. (2007). Mass spectrometry-based functional proteomics: from molecular machines to protein networks. *Nature Methods*, 4(10), 807–815. <https://doi.org/10.1038/nmeth1093>
- Krebs, A. R., Demmers, J., Karmodiya, K., Chang, N.-C., Chang, A. C., & Tora, L. (2010). ATAC and Mediator coactivators form a stable complex and regulate a set of non-coding RNA genes. *EMBO Reports*, 11(7), 541–547. <https://doi.org/10.1038/embor.2010.75>
- Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., ... Greenblatt, J. F. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*, 440(7084), 637–643. <https://doi.org/10.1038/nature04670>
- Krogan, N. J., Keogh, M.-C., Datta, N., Sawa, C., Ryan, O. W., Ding, H., ... Greenblatt, J. F. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Molecular Cell*, 12, 1565–1576.
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>

- Lambert, J.-P., Tucholska, M., Go, C., Knight, J. D. R., & Gingras, A.-C. (2015). Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *Journal of Proteomics*, 118(12), 81–94. <https://doi.org/10.1016/j.jprot.2014.09.011>
- Lange, M., Kaynak, B., Forster, U. B., Tönjes, M., Fischer, J. J., Grimm, C., ... Sperling, S. (2008). Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes and Development*, 22(17), 2370–2384. <https://doi.org/10.1101/gad.471408>
- Lee, J. H., Cai, G., Panigrahi, A. K., Dunham-Ems, S., Nguyen, T. N., Radolf, J. D., ... Günzl, A. (2010). A TFIIH-associated mediator head is a basal factor of small nuclear spliced leader RNA gene transcription in early-diverged trypanosomes. *Molecular and Cellular Biology*, 30(23), 5502–5513. <https://doi.org/10.1128/MCB.00966-10>
- Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*, 46(D1), D493–D496. <https://doi.org/10.1093/nar/gkx922>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, Y., Han, R., Bi, C., Li, M., Wang, S., & Gao, X. (2018). DeepSimulator: a deep simulator for Nanopore sequencing. *Bioinformatics (Oxford, England)*, 34(17), 2899–2908. <https://doi.org/10.1093/bioinformatics/bty223>
- Liao, S. M., Zhang, J., Jeffery, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., ... Young, R. A. (1995). A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature*, 374(6518), 193–196. <https://doi.org/10.1038/374193a0>
- Linder, T., & Gustafsson, C. M. (2004). The Soh1/MED31 protein is an ancient component of *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* Mediator. *The Journal of Biological Chemistry*, 279(47), 49455–49459. <https://doi.org/10.1074/jbc.M409046200>
- Linger, B. R., Morin, G. B., & Price, C. M. (2011). The Pot1a-associated proteins Tpt1 and Pat1 coordinate telomere protection and length regulation in *Tetrahymena*. *Molecular Biology of the Cell*, 22(21), 4161–4170. <https://doi.org/10.1091/mbc.E11-06-0551>
- Liu, X., Li, B., & GorovskyMA. (1996). Essential and nonessential histone H2A variants in *Tetrahymena thermophila*. *Molecular and Cellular Biology*, 16, 4305–4311.
- Liu, Y., Taverna, S. D., Muratore, T. L., Shabanowitz, J., Hunt, D. F., & Allis, C. D. (2007). RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes & Development*, 21(12), 1530–1545. <https://doi.org/10.1101/gad.1544207>
- Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., & Pallen, M. J. (2012). Erratum: Corrigendum: Performance comparison of benchtop high-throughput

- sequencing platforms. *Nature Biotechnology*, 30(6), 562–562.
<https://doi.org/10.1038/nbt0612-562f>
- Loose, M. W. (2017). The potential impact of nanopore sequencing on human genetics. *Human Molecular Genetics*, 26(R2), R202–R207. <https://doi.org/10.1093/hmg/ddx287>
- Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics*, 14(5), 265–279.
<https://doi.org/10.1016/j.gpb.2016.05.004>
- Madireddi, M. T., Coyne, R. S., Smothers, J. F., Mickey, K. M., Yao, M.-C., & Allis, C. D. (1996). Pdd1p, A Novel Chromodomain-Containing Protein, Links Heterochromatin Assembly and DNA Elimination in Tetrahymena. *Cell*, 87(1), 75–84.
[https://doi.org/10.1016/S0092-8674\(00\)81324-0](https://doi.org/10.1016/S0092-8674(00)81324-0)
- Madireddi, M. T., Davis, M. C., & Allis, C. D. (1994). Identification of a novel polypeptide involved in the formation of DNA-containing vesicles during macronuclear development in Tetrahymena. *Developmental Biology*, 165(2), 418–431.
<https://doi.org/10.1006/dbio.1994.1264>
- Malagon, F., Tong, A. H., Shafer, B. K., & Strathern, J. N. (2004). Genetic interactions of DST1 in *Saccharomyces cerevisiae* suggest a role of TFIIS in the initiation-elongation transition. *Genetics*, 166(3), 1215–1227.
- Maliszewska-Olejniczak, K., Gruchota, J., Gromadka, R., Denby Wilkes, C., Arnaiz, O., Mathy, N., ... Nowak, J. K. (2015). TFIIS-Dependent Non-coding Transcription Regulates Developmental Genome Rearrangements. *PLoS Genetics*, 11(7), e1005383.
<https://doi.org/10.1371/journal.pgen.1005383>
- Mao, X., Cao, F., Nie, X., Liu, H., & Chen, J. (2006). *The Swi / Snf chromatin remodeling complex is essential for hyphal development in Candida albicans*. 580, 2615–2622.
<https://doi.org/10.1016/j.febslet.2006.04.009>
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24(3), 133–141. <https://doi.org/10.1016/j.tig.2007.12.007>
- Martindale, D W, & Bruns, P. J. (1983). Cloning of abundant mRNA species present during conjugation of Tetrahymena thermophila: identification of mRNA species present exclusively during meiosis. *Molecular and Cellular Biology*, 3(10), 1857–1865.
<https://doi.org/10.1128/MCB.3.10.1857>
- Martindale, Duane W., Allis, C. D., & Bruns, P. J. (1982). Conjugation in Tetrahymena thermophila. A temporal analysis of cytological stages. *Experimental Cell Research*, 140(1), 227–236. [https://doi.org/10.1016/0014-4827\(82\)90172-0](https://doi.org/10.1016/0014-4827(82)90172-0)
- Masumoto, H., Hawke, D., Kobayashi, R., & Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature*, 436(7048), 294–298. <https://doi.org/10.1038/nature03714>
- Matangkasombut, O., Buratowski, R. M., Swilling, N. W., & Buratowski, S. (2000). Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes and Development*, 14(8), 951–962. <https://doi.org/10.1101/gad.14.8.951>

- Merriam, E. V, & Bruns, P. J. (1988). Phenotypic assortment in *Tetrahymena thermophila*: assortment kinetics of antibiotic-resistance markers, tsA, death, and the highly amplified rDNA locus. *Genetics*, 120, 389–395.
- Meyers, R. M., Bryan, J. G., McFarland, J. M., Weir, B. A., Sizemore, A. E., Xu, H., ... Tsherniak, A. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature Genetics*, 49(12), 1779–1784. <https://doi.org/10.1038/ng.3984>
- Min, B., & Collins, K. (2009). An RPA-Related Sequence-Specific DNA-Binding Subunit of Telomerase Holoenzyme Is Required for Elongation Processivity and Telomere Maintenance. *Molecular Cell*, 36(4), 609–619. <https://doi.org/10.1016/j.molcel.2009.09.041>
- Mizzen, C. A., Dou, Y., Liu, Y., Cook, R. G., Gorovsky, M. A., & Allis, C. D. (1999). Identification and mutation of phosphorylation sites in a linker histone. Phosphorylation of macronuclear H1 is not essential for viability in tetrahymena. *The Journal of Biological Chemistry*, 274(21), 14533–14536. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10329641>
- Mochizuki, K., Fine, N. A., Fujisawa, T., & Gorovsky, M. a. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell*, 110(6), 689–699. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12297043>
- Mochizuki, K., & Gorovsky, M. a. (2004a). RNA polymerase II localizes in *Tetrahymena thermophila* meiotic micronuclei when micronuclear transcription associated with genome rearrangement occurs. *Eukaryotic Cell*, 3, 1233–1240.
- Mochizuki, K., & Gorovsky, M. a. (2004b). Small RNAs in genome rearrangement in *Tetrahymena*. *Current Opinion in Genetics & Development*, 14(2), 181–187. <https://doi.org/10.1016/j.gde.2004.01.004>
- Muller, S., Filippakopoulos, P., & Knapp, S. (2011). Bromodomains as therapeutic targets. *Expert Reviews in Molecular Medicine*, 13(September), e29. <https://doi.org/10.1017/S1462399411001992>
- Nakamura, T. M. (1997). Telomerase Catalytic Subunit Homologs from Fission Yeast and Human. *Science*, 277(5328), 955–959. <https://doi.org/10.1126/science.277.5328.955>
- Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., ... Shendure, J. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, 461(7261), 272–276. <https://doi.org/10.1038/nature08250>
- Nikiforov, M. A., Gorovsky, M. A., & Allis, C. D. (2000). A novel chromodomain protein, pdd3p, associates with internal eliminated sequences during macronuclear development in *Tetrahymena thermophila*. *Molecular and Cellular Biology*, 20(11), 4128–4134. <https://doi.org/10.1128/MCB.20.11.4128-4134.2000>
- Nonet, M. L., & Young, R. A. (1989). Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics*, 123(4), 715–724.
- Noto, T., Kurth, H. M., Kataoka, K., Aronica, L., DeSouza, L. V., Siu, K. W. M., ... Mochizuki,

- K. (2010). The Tetrahymena argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus. *Cell*, 140(5), 692–703. <https://doi.org/10.1016/j.cell.2010.02.010>
- Ohba, R., Steger, D. J., Brownell, J. E., Mizzen, C. a, Cook, R. G., Côté, J., ... Allis, C. D. (1999). A novel H2A/H4 nucleosomal histone acetyltransferase in Tetrahymena thermophila. *Molecular and Cellular Biology*, 19(3), 2061–2068. <https://doi.org/10.1128/MCB.19.3.2061>
- Orias, E., Cervantes, M. D., & Hamilton, E. P. (2011). Tetrahymena thermophila, a unicellular eukaryote with separate germline and somatic genomes. *Research in Microbiology*, 162(6), 578–586. <https://doi.org/10.1016/j.resmic.2011.05.001>
- Papamichos-chronakis, M., Watanabe, S., Rando, O. J., & Craig, L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin remodeling enzyme is essential for genome integrity. *Cell*, 144(2), 200–213. <https://doi.org/10.1016/j.cell.2010.12.021>.Global
- Papazyan, R., Voronina, E., Chapman, J. R., Luperchio, T. R., Gilbert, T. M., Meier, E., ... Taverna, S. D. (2014). Methylation of histone H3K23 blocks DNA damage in pericentric heterochromatin during meiosis. *ELife*, 3, e02996. <https://doi.org/10.7554/eLife.02996>
- Park, P. J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*, 10(10), 669–680. <https://doi.org/10.1038/nrg2641>
- Parker, K., Maxson, J., Mooney, A., & Wiley, E. A. (2007). Class I histone deacetylase Thd1p promotes global chromatin condensation in Tetrahymena thermophila. *Eukaryotic Cell*, 6(10), 1913–1924. <https://doi.org/10.1128/EC.00217-07>
- Parthun, M. R. (2007). Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene*, 26(37), 5319–5328. <https://doi.org/10.1038/sj.onc.1210602>
- Poli, J., Gasser, S. M., & Papamichos-Chronakis, M. (2017). The INO80 remodeller in transcription, replication and repair. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372(1731), 20160290. <https://doi.org/10.1098/rstb.2016.0290>
- Poss, Z. C., Ebmeier, C. C., & Taatjes, D. J. (2013). The Mediator complex and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology*, 48(6), 575–608. <https://doi.org/10.3109/10409238.2013.840259>
- Prathapam, R., Witkin, K. L., O'Connor, C. M., & Collins, K. (2005). A telomerase holoenzyme protein enhances telomerase RNA assembly with telomerase reverse transcriptase. *Nature Structural & Molecular Biology*, 12(3), 252–257. <https://doi.org/10.1038/nsmb900>
- Premkumar, V. L., Cranert, S., Linger, B. R., Morin, G. B., Minium, S., & Price, C. (2014). The 3' overhangs at Tetrahymena thermophila telomeres are packaged by four proteins, Pot1a, Tpt1, Pat1, and Pat2. *Eukaryotic Cell*, 13(2), 240–245. <https://doi.org/10.1128/EC.00275-13>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>

- R Core, T. (2013). *R: A language and environment for statistical computing*. Retrieved from <http://www.r-project.org/>.
- Raisner, R. M., Hartley, P. D., Meneghini, M. D., Bao, M. Z., Liu, C. L., Schreiber, S. L., ... Madhani, H. D. (2005). Histone Variant H2A.Z Marks the 5' Ends of Both Active and Inactive Genes in Euchromatin. *Cell*, 123(2), 233–248.
- Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., & Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, 42(Web Server issue), W187-91. <https://doi.org/10.1093/nar/gku365>
- Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., ... Allis, C. D. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proceedings of the National Academy of Sciences of the United States of America*, 103(18), 6988–6993. <https://doi.org/10.1073/pnas.0601676103>
- Reisman, D., Glaros, S., & Thompson, E. a. (2009). The SWI/SNF complex and cancer. *Oncogene*, 28(14), 1653–1668. <https://doi.org/10.1038/onc.2009.4>
- Richters, A., & Koehler, A. N. (2017). Epigenetic Modulation Using Small Molecules - Targeting Histone Acetyltransferases in Disease. *Current Medicinal Chemistry*, 24(37), 4121–4150. <https://doi.org/10.2174/0929867324666170223153115>
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., & Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnology*, 17(10), 1030–1032. <https://doi.org/10.1038/13732>
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011a). Integrative genomics viewer. *Nature Biotechnology*, 29(1), 24–26. <https://doi.org/10.1038/nbt.1754>
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011b). Integrative Genomics Viewer. *Nature Biotechnology*, 29, 24–26.
- Rost, B. (2002). Did evolution leap to create the protein universe? *Current Opinion in Structural Biology*, 12(3), 409–416. [https://doi.org/10.1016/S0959-440X\(02\)00337-8](https://doi.org/10.1016/S0959-440X(02)00337-8)
- Ruehle, M. D., Orias, E., & Pearson, C. G. (2016). Tetrahymena as a Unicellular Model Eukaryote: Genetic and Genomic Tools. *Genetics*, 203(2), 649–665. <https://doi.org/10.1534/genetics.114.169748>
- Saettone, A., Garg, J., Lambert, J.-P., Nabeel-Shah, S., Ponce, M., Burtch, A., ... Fillingham, J. (2018). The bromodomain-containing protein Ibd1 links multiple chromatin-related protein complexes to highly expressed genes in Tetrahymena thermophila. *Epigenetics & Chromatin*, 11(1), 10. <https://doi.org/10.1186/s13072-018-0180-6>
- Saettone, A., Nabeel-Shah, S., Garg, J., Lambert, J.-P., Pearlman, R. E., & Fillingham, J. (2019). Functional Proteomics of Nuclear Proteins in Tetrahymena thermophila: A Review. *Genes*, 10(5), 333. <https://doi.org/10.3390/genes10050333>
- Saettone, A., Ponce, M., Nabeel-Shah, S., & Fillingham, J. (2019). RACS: rapid analysis of

- ChIP-Seq data for contig based genomes. *BMC Bioinformatics*, 20(1), 533.
<https://doi.org/10.1186/s12859-019-3100-2>
- Sampath, S. C., Marazzi, I., Yap, K. L., Sampath, S. C., Krutchinsky, A. N., Mecklenbräuker, I., ... Tarakhovsky, A. (2007). Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Molecular Cell*, 27(4), 596–608. <https://doi.org/10.1016/j.molcel.2007.06.026>
- Samuelsen, C. O., Baraznenok, V., Khorosjutina, O., Spahr, H., Kieselbach, T., Holmberg, S., & Gustafsson, C. M. (2003). TRAP230/ARC240 and TRAP240/ARC250 Mediator subunits are functionally conserved through evolution. *Proceedings of the National Academy of Sciences*, 100(11), 6422–6427. <https://doi.org/10.1073/pnas.1030497100>
- Schiano, C., Casamassimi, A., Rienzo, M., de Nigris, F., Sommese, L., & Napoli, C. (2014). Involvement of Mediator complex in malignancy. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1845(1), 66–83. <https://doi.org/10.1016/j.bbcan.2013.12.001>
- Schmidt, D., Wilson, M. D., Ballester, B., Schwalie, P. C., Brown, G. D., Marshall, A., ... Odom, D. T. (2010). Five-Vertebrate ChIP-seq Reveals the Evolutionary Dynamics of Transcription Factor Binding. *Science*, 328(5981), 1036–1040.
<https://doi.org/10.1126/science.1186176>
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7), 671–675.
- Schoeberl, U. E., Kurth, H. M., Noto, T., & Mochizuki, K. (2012). Biased transcription and selective degradation of small RNAs shape the pattern of DNA elimination in Tetrahymena. *Genes & Development*, 26(15), 1729–1742. <https://doi.org/10.1101/gad.196493.112>
- Schoeberl, U. E., & Mochizuki, K. (2011). Keeping the soma free of transposons: Programmed DNA elimination in ciliates. *Journal of Biological Chemistry*, 286(43), 37045–37052.
<https://doi.org/10.1074/jbc.R111.276964>
- Schulman, I. G., Cook, R. G., Richman, R., & Allis, C. D. (1987). Tetrahymena contain two distinct and unusual high mobility group (HMG)-like proteins. *The Journal of Cell Biology*, 104(6), 1485–1494. <https://doi.org/10.1128/MCB.11.1.166>
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods*, 5(1), 16–18. <https://doi.org/10.1038/nmeth1156>
- Shannon, P., Markiel, A., Owen Ozier, 2, Baliga, N. S., Wang, J. T., Ramage, D., ... Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, (13), 2498–2504.
<https://doi.org/10.1101/gr.1239303.metabolite>
- Sharan, R., Suthram, S., Kelley, R. M., Kuhn, T., McCuine, S., Uetz, P., ... Ideker, T. (2005). Conserved patterns of protein interaction in multiple species. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 1974–1979.
<https://doi.org/10.1073/pnas.0409522102>
- Shen, W., Xu, C., Huang, W., Zhang, J., Carlson, J. E., Tu, X., ... Shi, Y. (2007). Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails.

- Biochemistry*, 46(8), 2100–2110. <https://doi.org/10.1021/bi0611208>
- Shi, X., Hong, T., Walter, K., Ewalt, M., Michishita, E., Hung, T., ... Gozani, O. (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature*, 442(6), 96–99.
- Simonis, N., Rual, J.-F., Carvunis, A.-R., Tasan, M., Lemmens, I., Hirozane-Kishikawa, T., ... Vidal, M. (2009). Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network. *Nature Methods*, 6(1), 47–54. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19123269>
- Simpson, J. T., Workman, R. E., Zuzarte, P. C., David, M., Dursi, L. J., & Timp, W. (2017). Detecting DNA cytosine methylation using nanopore sequencing. *Nature Methods*, 14(4), 407–410. <https://doi.org/10.1038/nmeth.4184>
- Smothers, J. F., Mizzen, C. A., Tubbert, M. M., Cook, R. G., & Allis, C. D. (1997). Pdd1p associates with germline-restricted chromatin and a second novel anlagen-enriched protein in developmentally programmed DNA elimination structures. *Development (Cambridge, England)*, 124, 4537–4545.
- Snel, B., Lehmann, G., Bork, P., & Huynen, M. A. (2000). STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Research*, 28(18), 3442–3444. <https://doi.org/10.1093/nar/28.18.3442>
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., & Allis, C. D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proceedings of the National Academy of Sciences of the United States of America*, 92(4), 1237–1241. <https://doi.org/10.1073/pnas.92.4.1237>
- Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C., & Werner, M. (2011). Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo. *Science*, 331(6023), 1451–1454. <https://doi.org/10.1126/science.1200188>
- Soutourina, Julie. (2017). Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology*, 19(4), 262–274. <https://doi.org/10.1038/nrm.2017.115>
- Stargell, L. a., Bowen, J., Dadd, C. a., Dedon, P. C., Davis, M., Cook, R. G., ... Gorovsky, M. a. (1993). Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in *Tetrahymena thermophila*. *Genes and Development*, 7, 2641–2651.
- Steunou, A.-L., Cramet, M., Rossetto, D., Aristizabal, M. J., Lacoste, N., Drouin, S., ... Côté, J. (2016). Combined Action of Histone Reader Modules Regulates NuA4 Local Acetyltransferase Function but Not Its Recruitment on the Genome. *Molecular and Cellular Biology*, 36(22), 2768–2781. <https://doi.org/10.1128/MCB.00112-16>
- Stover, N. A., Krieger, C. J., Binkley, G., Dong, Q., Fisk, D. G., Nash, R., ... Cherry, J. M. (2006). Tetrahymena Genome Database (TGD): a new genomic resource for *Tetrahymena thermophila* research. *Nucleic Acids Research*, 34(Database issue), D500-3. <https://doi.org/10.1093/nar/gkj054>
- Stover, N. A., Punia, R. S., Bowen, M. S., Dolins, S. B., & Clark, T. G. (2012). Tetrahymena

- Genome Database Wiki: a community-maintained model organism database. *Database : The Journal of Biological Databases and Curation*, 2012, bas007. <https://doi.org/10.1093/database/bas007>
- Strahl, B. D., Grant, P. A., Briggs, S. D., Sun, Z.-W., Bone, J. R., Caldwell, J. A., ... Allis, C. D. (2002). Set2 Is a Nucleosomal Histone H3-Selective Methyltransferase That Mediates Transcriptional Repression. *Molecular and Cellular Biology*, 22(5), 1298–1306. <https://doi.org/10.1128/MCB.22.5.1298-1306.2002>
- Strahl, B D, Ohba, R., Cook, R. G., & Allis, C. D. (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proceedings of the National Academy of Sciences of the United States of America*, 96(26), 14967–14972. <https://doi.org/10.1073/pnas.96.26.14967>
- Strahl, Brian D, & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, 403(6765), 41–45. <https://doi.org/10.1038/47412>
- Streit, A., Wang, J., Kang, Y., & Davis, R. E. (2016). Gene silencing and sex determination by programmed DNA elimination in parasitic nematodes. *Current Opinion in Microbiology*, 32, 120–127. <https://doi.org/10.1016/j.mib.2016.05.012>
- Taverna, S. D., Coyne, R. S., & Allis, C. D. (2002). Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. *Cell*, 110(6), 701–711. [https://doi.org/10.1016/S0092-8674\(02\)00941-8](https://doi.org/10.1016/S0092-8674(02)00941-8)
- Taverna, S. D., Ilin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., ... Allis, C. D. (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Molecular Cell*, 24(5), 785–796. <https://doi.org/10.1016/j.molcel.2006.10.026>
- Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., ... Allis, C. D. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proceedings of the National Academy of Sciences of the United States of America*, 104(7), 2086–2091. <https://doi.org/10.1073/pnas.0610993104>
- Teng, H., Cao, M. D., Hall, M. B., Duarte, T., Wang, S., & Coin, L. J. M. (2018). Chiron: translating nanopore raw signal directly into nucleotide sequence using deep learning. *GigaScience*, 7(5). <https://doi.org/10.1093/gigascience/giy037>
- Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A.-C., & Choi, H. (2014). SAINTexpress: Improvements and additional features in Significance Analysis of INTeractome software. *Journal of Proteomics*, 100, 37–43. <https://doi.org/10.1016/j.jprot.2013.10.023>
- Thompson, C. M., & Young, R. A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 92(10), 4587–4590.
- Thorsen, M., Hansen, H., Venturi, M., Holmberg, S., & Thon, G. (2012). Mediator regulates non-coding RNA transcription at fission yeast centromeres. *Epigenetics & Chromatin*, 5(1), 19. <https://doi.org/10.1186/1756-8935-5-19>

- Tian, M., Chen, X., Xiong, Q., Xiong, J., Xiao, C., Ge, F., ... Miao, W. (2014). Phosphoproteomic Analysis of Protein Phosphorylation Networks in *Tetrahymena thermophila*, a Model Single-celled Organism. *Molecular & Cellular Proteomics*, 13(2), 503–519. <https://doi.org/10.1074/mcp.M112.026575>
- Tsai, K.-L., Yu, X., Gopalan, S., Chao, T.-C., Zhang, Y., Florens, L., ... Asturias, F. J. (2017). Mediator structure and rearrangements required for holoenzyme formation. *Nature*, 544(7649), 196–201. <https://doi.org/10.1038/nature21393>
- Upton, H. E., Chan, H., Feigon, J., & Collins, K. (2017). Shared Subunits of Tetrahymena Telomerase Holoenzyme and Replication Protein A Have Different Functions in Different Cellular Complexes. *The Journal of Biological Chemistry*, 292(1), 217–228. <https://doi.org/10.1074/jbc.M116.763664>
- Urbanska, P., Joachimiak, E., Bazan, R., Fu, G., Poprzeczko, M., Fabczak, H., ... Wloga, D. (2018). Ciliary proteins Fap43 and Fap44 interact with each other and are essential for proper cilia and flagella beating. *Cellular and Molecular Life Sciences*, (0123456789). <https://doi.org/10.1007/s00018-018-2819-7>
- Vavra, K. J., Allis, C. D., & Gorovsky, M. A. (1982). Regulation of histone acetylation in *Tetrahymena* macro- and micronuclei. *The Journal of Biological Chemistry*, 257(5), 2591–2598. <https://doi.org/10.1073/pnas.82.23.8048>
- Wai, D. C. C., Szyszk, T. N., Campbell, A. E., Kwong, C., Wilkinson-White, L. E., Silva, A. P. G., ... Mackay, J. P. (2018). The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *The Journal of Biological Chemistry*, 293(19), 7160–7175. <https://doi.org/10.1074/jbc.RA117.000678>
- Wan, C., Borgeson, B., Phanse, S., Tu, F., Drew, K., Clark, G., ... Emili, A. (2015). Panorama of ancient metazoan macromolecular complexes. *Nature*, 525(7569), 339–344. <https://doi.org/10.1038/nature14877>
- Wang, Y., Chen, X., Sheng, Y., Liu, Y., & Gao, S. (2017). N6-adenine DNA methylation is associated with the linker DNA of H2A.Z-containing well-positioned nucleosomes in Pol II-transcribed genes in *Tetrahymena*. *Nucleic Acids Research*, 45(20), 11594–11606. <https://doi.org/10.1093/nar/gkx883>
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1), 57–63. <https://doi.org/10.1038/nrg2484>
- Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., ... Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research*, 38(Web Server issue), W214–20. <https://doi.org/10.1093/nar/gkq537>
- Wiley, E. A., Myers, T., Parker, K., Braun, T., & Yao, M.-C. (2005). Class I histone deacetylase Thd1p affects nuclear integrity in *Tetrahymena thermophila*. *Eukaryotic Cell*, 4(5), 981–990. <https://doi.org/10.1128/EC.4.5.981-990.2005>
- Witkin, K. L., & Collins, K. (2004). Holoenzyme proteins required for the physiological

- assembly and activity of telomerase. *Genes & Development*, 18(10), 1107–1118.
<https://doi.org/10.1101/gad.1201704>
- Witkin, K. L., Prathapam, R., & Collins, K. (2007). Positive and negative regulation of Tetrahymena telomerase holoenzyme. *Molecular and Cellular Biology*, 27(6), 2074–2083.
<https://doi.org/10.1128/MCB.02105-06>
- Xiong, J., Lu, X., Zhou, Z., Chang, Y., Yuan, D., Tian, M., ... Miao, W. (2012). Transcriptome analysis of the model protozoan, tetrahymena thermophila, using deep RNA sequencing. *PLoS ONE*, 7(2), 1–13. <https://doi.org/10.1371/journal.pone.0030630>
- Xiong, J., Lu, Y., Feng, J., Yuan, D., Tian, M., Chang, Y., ... Miao, W. (2013). Tetrahymena functional genomics database (TetraFGD): an integrated resource for Tetrahymena functional genomics. *Database : The Journal of Biological Databases and Curation*, 2013, bat008. <https://doi.org/10.1093/database/bat008>
- Xiong, J., Yuan, D., Fillingham, J. S., Garg, J., Lu, X., Chang, Y., ... Miao, W. (2011). Gene network landscape of the ciliate tetrahymena thermophila. *PLoS ONE*, 6(5).
<https://doi.org/10.1371/journal.pone.0020124>
- Yang, C., & Stiller, J. W. (2014). Evolutionary diversity and taxon-specific modifications of the RNA polymerase II C-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America*, 111(16), 5920–5925.
<https://doi.org/10.1073/pnas.1323616111>
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, 12(1), 7–8.
<https://doi.org/10.1038/nmeth.3213>
- Yang, Y., Li, L., & Qu, L.-J. (2016). Plant Mediator complex and its critical functions in transcription regulation. *Journal of Integrative Plant Biology*, 58(2), 106–118.
<https://doi.org/10.1111/jipb.12377>
- Yao, M.-C., Chao, J.-L., & Cheng, C.-Y. (2014). Programmed Genome Rearrangements in Tetrahymena. *Microbiology Spectrum*, 2(6).
<https://doi.org/10.1128/microbiolspec.MDNA3-0012-2014>
- Yao, M. C., Choi, J., Yokoyama, S., Austerberry, C. F., & Yao, C. H. (1984). DNA elimination in Tetrahymena: a developmental process involving extensive breakage and rejoining of DNA at defined sites. *Cell*, 36(2), 433–440. [https://doi.org/10.1016/0092-8674\(84\)90236-8](https://doi.org/10.1016/0092-8674(84)90236-8)
- Yin, J. -w., & Wang, G. (2014). The Mediator complex: a master coordinator of transcription and cell lineage development. *Development*, 141(5), 977–987.
<https://doi.org/10.1242/dev.098392>
- Zhang, C., Gao, S., Molascon, A. J., Wang, Z., Gorovsky, M. A., Liu, Y., & Andrews, P. C. (2014). Bioinformatic and proteomic analysis of bulk histones reveals PTM crosstalk and chromatin features. *Journal of Proteome Research*, 13(7), 3330–3337.
<https://doi.org/10.1021/pr5001829>
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., ... Liu, X. S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9), R137.

<https://doi.org/10.1186/gb-2008-9-9-r137>