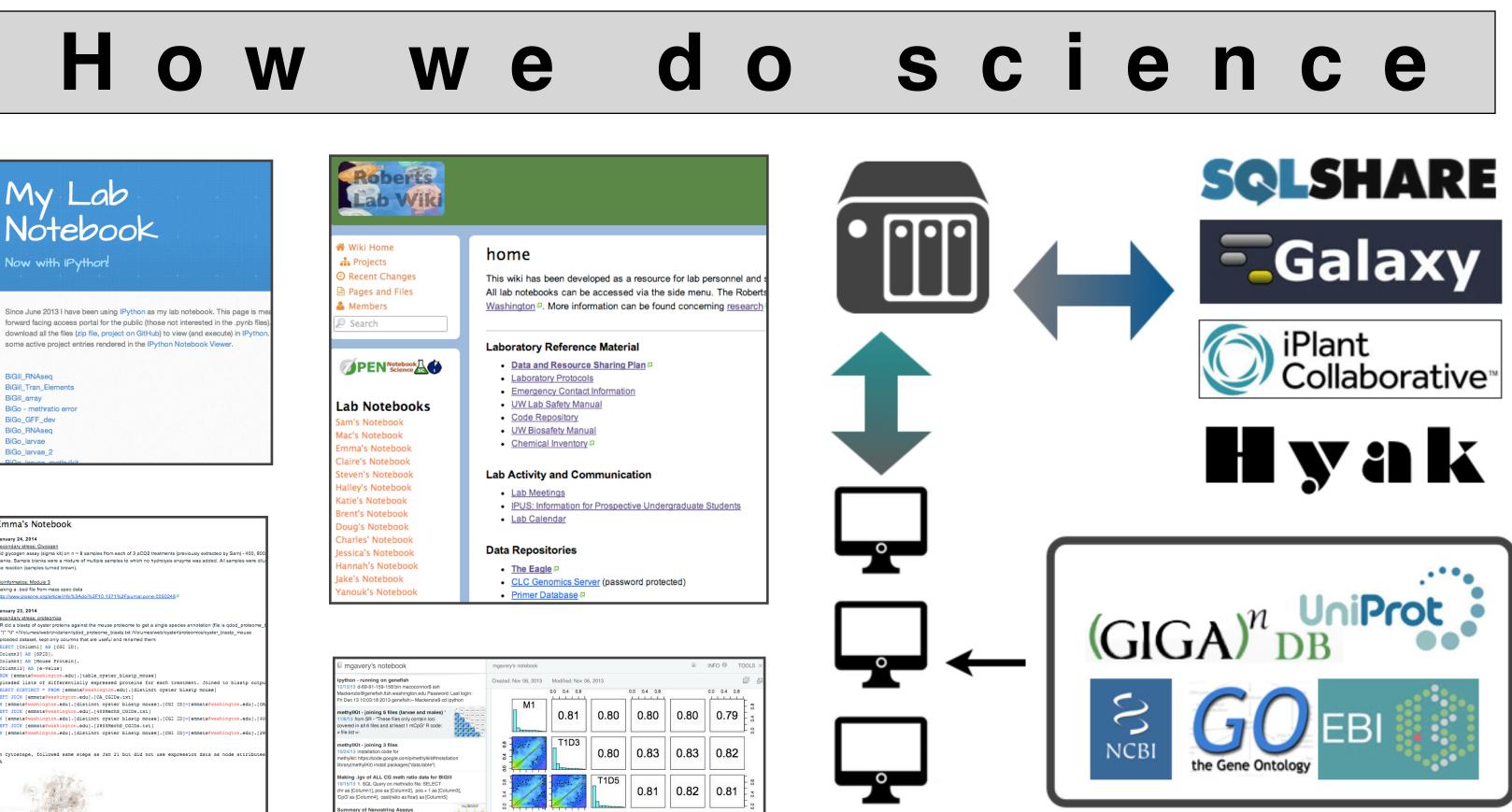
# **Exploring the Biology of Oysters** a few million base pairs at a time



**Steven Roberts** College of the Environment School of Aquatic and Fishery Sciences



ocuse ' sure fi Meeting



0.88



Data

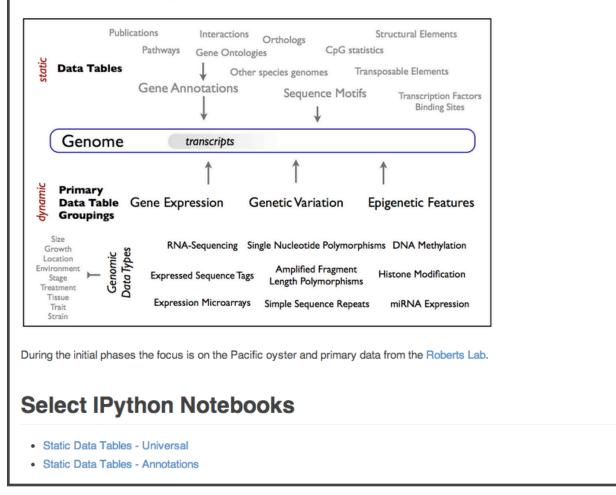
Emma's Notebook	
January 24, 2014	
Secondary stress: Glycogen	
Did glycogen assay (sigma kit) on n = 8	3 samples from each of 3 pCO2 treatments (previously extracted by Sam) - 400, 8
blanks. Sample blanks were a mixture	of multiple samples to which no hydrolysis enzyme was added. All samples were d
the reaction (samples turned brown).	
Bioinformatics: Module 3	
making a .bed file from mass spec data	a
http://www.plosone.org/article/info%3A	doi%2F10.1371%2Fjournal.pone.0050246#
January 23, 2014	
Secondary stress: proteomics	
	nst the mouse proteome to get a single species annotation (file is qdod_proteome
	d_proteome_blastp.txt /Volumes/web/oyster/proteomics/oyster_blastp_mouse
Uploaded dataset, kept only columns t	hat are useful and renamed them:
SELECT [Column1] AS [CGI ID],	
[Column3] AS [SPID], [Column4] AS [Mouse Protein],	
[Column13] AS [e-value]	
FROM [emmats@washington.edu].[	table oveter blastn mousel
	ly expressed proteins for each treatment. Joined to blastp out
	Washington.edul.[distinct ovster blastp mouse]
LEFT JOIN [emmats@washington.e	
ON [emmats@washington.edu].[di	istinct oyster blastp mouse].[CGI ID]=[emmats@washington.edu].[(
LEFT JOIN [emmats@washington.e	du].[400MechS_CGIDs.txt]
ON [emmats@washington.edu].[di	istinct oyster blastp mouse].[CGI ID]=[emmats@washington.edu].[-
LEFT JOIN [emmats@washington.e	du].[2800Mech5_CGIDs.txt]
ON [emmats@washington.edu].[di	istinct oyster blastp mouse].[CGI ID]=[emmats@washington.edu].[2
In Cytoscape, followed same st	eps as Jan 21 but did not use expression data as node attribute
OA	
	x
	4
	1
T YANG H	
A States	
14 C C C	
Olim Co .	

Halley's Notebook Katie's Notebook Brent's Notebook Doug's Notebook Charles' Notebook	Lab Meetings     IPUS: Information for Prospec     Lab Calendar		graduate :	<u>Students</u>
Jessica's Notebook Hannah's Notebook Jake's Notebook Yanouk's Notebook	Data Repositories <u>The Eagle</u> <u>CLC Genomics Server</u> (passw     Primer Database 7	word protect	ted)	
Imgavery's notebook	mgavery's notebook Created: Nov 06, 2013 Modified: Nov 06, 2013			INFO () TOOLS
115/13 c4-9-159-158 bin macconnor5 sh Mackenzie@genefish.fish.washington.edu Password: Last login: Fri De t3 10:03:18 2013 genefish- MackenzieS od loython methylKi - joining 6 files (larvae and males) 115/13 from SR - 'These files only contain loci covered in all 6 files and at least 1 mCpG'R code:	M1 0.81 0.80	0.0 0.4 0.8	0.80	0.0 0.4 0.8
methylkit - joining 3 files 10/24/13 installation code for methylkit https://code.google.com/p/methylkit/#Installation library(methylkit) install.packages("data.table")		0.83	0.83	0.82
Making .lgv of ALL CG meth ratio data for BiGili 10/15/13 1. SQL Query on methratio file: SELECT chr as [Column1], pos as [Column2], pos + 1 as [Column3], "CpG" as [Column4], cast(ratio as float) as [Column5]		0.81	0.82	0.81
Summary of Nanostring Assays 10/7/13 A total of 24 assays were performing using		T3D5		<sup>°</sup>

#### qdod: Querying Disparate Oyster Datasets

This respository provides access to genomic data and workflows (IPython notebooks) that are being integrated as part of effort to ncrease effeciency of biological discovery. The wiki associated with this repository will serve as the primary means for documentation. Currently the documentation is focused on 1) describing current datasets and 2) providing workflow tutorials.

In brief, data in the form of delimited text files is aggregated into SQLShare where they can be easily queried. Below is schematic representation of the different types of datasets.



Crass	sostrea gigas (v	/9, id22701): ι	unmasked											
JBro	owse File	View Help											œ⊃ Share	filter by tet Solution 9 tracks shown
0	100,000	200,000	300,000	400,000	500,000	600, <mark>000</mark>	700,000	800,000	900,000	1,000,000	1,100,000	1,200,000	1,300,000	
						$\Theta \oplus \Theta$	scaffold1179	scaffold11	79:36100 Go	<b>A</b>				Sequence

P branch: master - LabDocs / D	ataSharePlan.md	6
--------------------------------	-----------------	---

2 contributors 📑 👬 file | 26 lines (13 sloc) | 2.218 kb Open Edit Raw Blame History Delet

**Data and Resource Sharing Plan** 

Data Sharing

**Routine Activity** 

including Evernote, Wikispaces, IPython, and GitHub. Noteboooks will be made centrally accessible at

#### Product Publicatio

Ill published data will be submitted to journals indexed by the major biomedical indexing sites including PubMed a on) will be made available on the lab website (http://faculty.washington.edu/sr320/?page\_id=246). For all studies, the raw data, as well as the relevant metadata and computational algorithms used to generate statistical and aphical summaries which were used to generate the final research data will be provided. This data will be provided wit permissive (CC0 or CC-BY) use license. Data will be deposited as supplementary data to the relevant publications where ble and where permitted, deposited to web-accessible data repositories such as Figshare (http://figsha similar services. Finally data will be posted on, or linked to, from the principal investigator's website. If further data is btained relevant to these studies, which can be combined with the published data sets, these evolving data sets will be Ided to the available online repositories, separately from the published data sets

opular Datasets erring Hepatic Transcr highly discussed viewed figshare. highly discussed highly viewed highly di highly viewed highly discussed discussed

	Figshare Profiles	
highly discussed highly viewed highly discussed figshare.	Steven Roberts on figshare	9982
e Cod highly viewed highly discussed discussed figshare.	E. Timmins-Schi on figshare	1460
viewed saved discussed discussed figshare.	Claire Ellis on figshare	381
s disease resistance in the eastern oyster	Mackenzie Gavery on figshare	1243
shii RE22 assembly	elene dorfmeier on figshare	2455
discussed figshare.	Caroline Storer on figshare	701

Products

Data | Presentations | Peer-Reviewed Publications | PrePrints | Theses | Grants | Media

		CpG	124	125	0		
	methratio	CpG	128	129	0	+	
C10009	methratio	CpG	136	137	0	+	
C10011	methratio	CpG	51	52	0	+	
C10011	methratio	CpG	62	63	0	+	
	methratio	CpG			0	+	
C10011	methratio	CpG	108	109	0	+	
: Iwe /Ve	lumes/web/cnic	darian/Bi	Go_methra	atio_boop	p.gff		
764281	17 68785353 378	8876013 /	Volumes/1	web/cnid	arian/Bi	.Go_meth:	ratio_
]:from pa	undas import *						
	data from data						
BiGOboo	op = read_csv(						ian/Bi
		# what c					
	na_values	в=["", "	"]) # wh	at value	s should	i be con	sidere
1 BiGObor	onf'score'l.his	at (bine=5)	n				
#Axis 1 plt.axi	<pre>pp['score'].hig limits are chan ls([0, 1, 0, 40 0, 400000]</pre>	nged usin		is([xmin	, xmax,	ymin, yı	max])
#Axis i plt.axi	limits are chan s([0, 1, 0, 40	nged usin		is([xmin	, xmax,	ymin, yn	max])
#Axis 1 plt.axi	limits are chan s([0, 1, 0, 40	nged usin		is([xmin	, xmax,	ymin, yı	max])
#Axis 1 plt.axi ]:[0, 1, 400000	limits are chan s([0, 1, 0, 40	nged usin		is([xmin	, xmax,	ymin, yn	max]) .
#Axis J plt.axi ]:[0, 1, 400000 350000	limits are chan s([0, 1, 0, 40	nged usin		is([xmin	, xmax,	ymin, yı	nax]) .
#Axis 1 plt.axi 5]:[0, 1, 400000 350000 300000	limits are chan s([0, 1, 0, 40	nged usin		is([xmin	, xmax,	ymin, yı	max]) .
#Axis i plt.axi ]:[0, 1, 40000 350000 250000	limits are chan s([0, 1, 0, 40	nged usin		is([xmin	, xmax,	ymin, yı	max]) .
#Axis ] plt.axi ]:[0, 1, 400000 350000 250000 200000	limits are chan s([0, 1, 0, 40	nged usin		is([xmin.	, xmax,	ymin, yn	max]) .
#Axis 1 plt.axi :[0, 1, 400000 350000 300000 250000 150000	limits are chan s([0, 1, 0, 40	nged usin		is([xmin,	, xmax,	ymin, yn	nax])

BiGo\_methratio Last Checkpoint: Sep 13 09:5

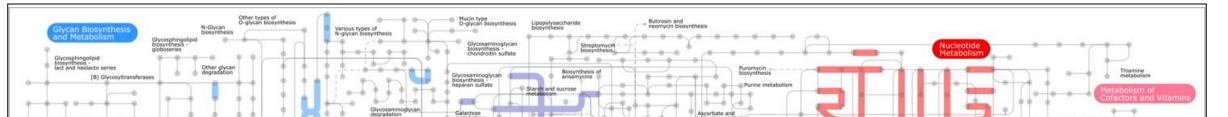
View Insert Cell Kernel Help

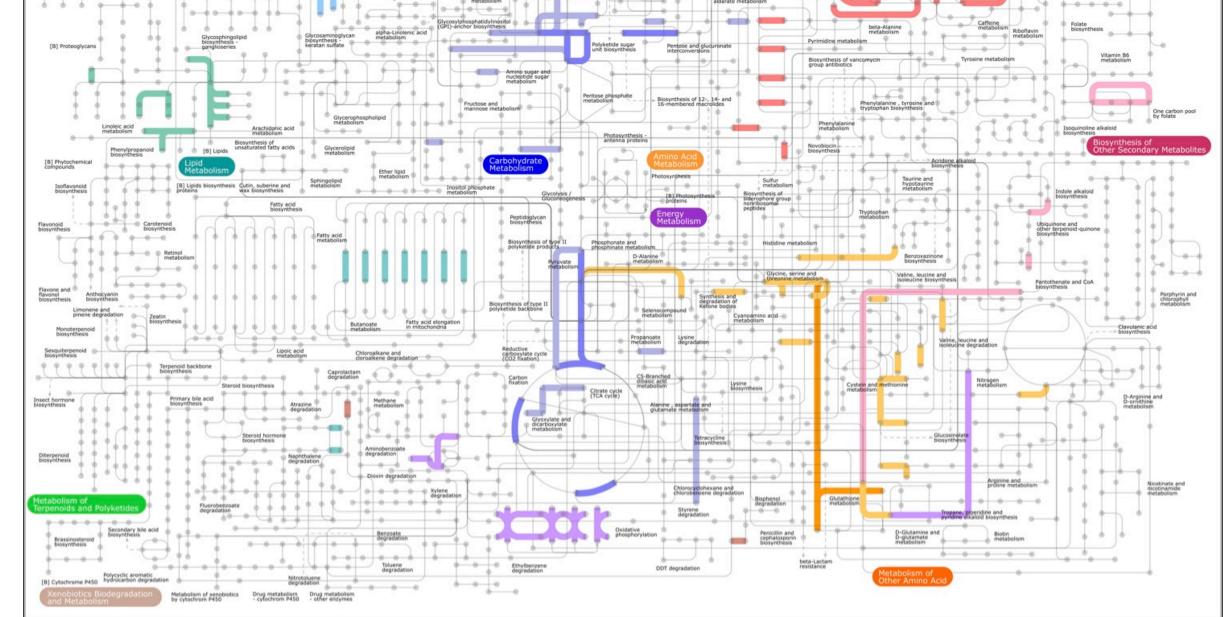
B 3< 2 1 ★ ↓ 0 0 ► ■ Code \$ Cell Toolbar: None \$</p>

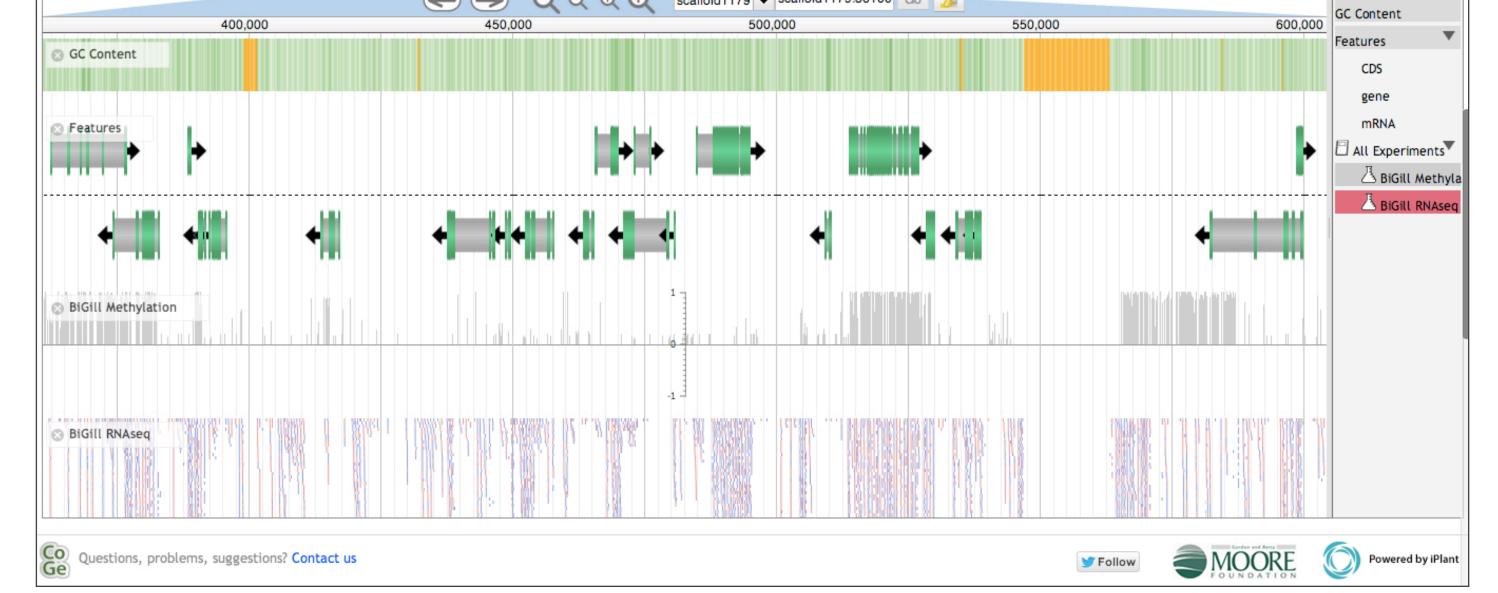
C10009 methratio CpG 70 71 0

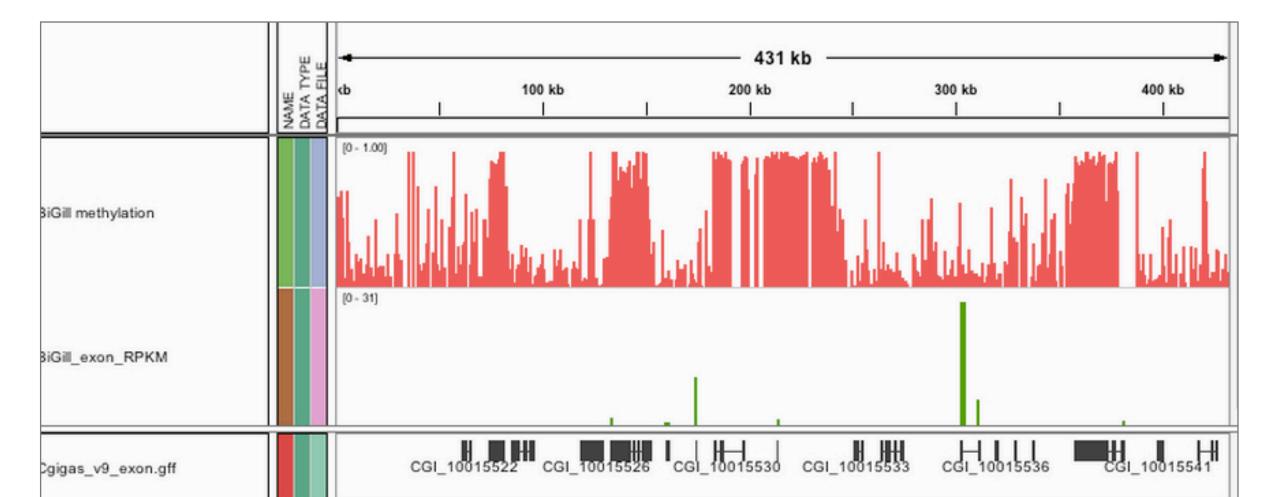
che625's notebook	che625's notebook 🔒 INFO 🖲 TOOLS 🗙						
stacked bar plots code Yesterday gene<- read.csvl"/Users/claireolson/Desktop/promoterandgene/stackedba	Created: Dec 30, 2013 Modified: Dec 30, 2013						
vauxav; (rosensulainebisinibeskupipininiterialingerietsaackebba Iolc_R.csv; header=T) head(gene) par(mar=c(5.1,4.1,4.1,8.5), romoter Region Analysis days ago SELECT * FROM (che625@washington.edu).	legend("topright", lty=1, lwd=2, col=c("red","blue"), legend=c("mRNA", "Promoter"))						
Methylated_Promoter.txt] FULL OUTER JOIN [che625@washington.edu].	Promoter vs Gene Body Methylation						
Notes- bioinformatics 4 days ago Get tables into CLC and use either DEseq, DEseq2 or EdgeR	9 - mRNA - Promoter						
BioInformatics 2014- Module 2 Last week Goals: -Creating visualization from BLAST results -Attach Gene Ontology information to the too blast this obtained for Module 1.	0 - 29						
to the top plast his obtained for MODULE 1. Prepping for MBD Last week Getting samples ready for MBD tomorrow. Going to run 16ug of my samples through the MBD protocol. Need samples to be at 15ug at a concentration of 25ng/ul Moving forward with	Density						
MBD- round 2 Last week 1/15/14 MBD- heat shock samples 15ug input DNA sheared to 500bp on Covaris 52 V2.5.4 in Armbrust Iab. Processed samples 2M, 2M H.S., 4M, 4M H.S., 6M, and 6M H.S. Triod to							

## pausing occasionally to make maps.





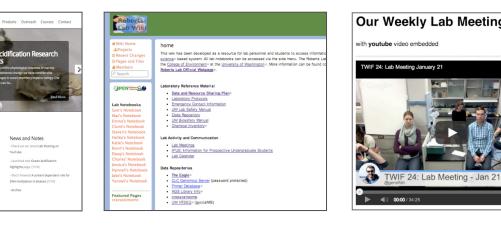


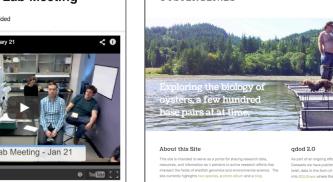


### More information













Poster, Links and lots of fun

e:sr320@uw.edu t:@sr320

Cgigas_v9_gene.gff	CGI_10015522	CGI_10015526	CGI_10015530	CGI_10015533	CGI_10015536	CGI_10015541
--------------------	--------------	--------------	--------------	--------------	--------------	--------------

