



# Max Mass Reclamation

The MinGenome algorithm contains the following objective function:

$$\max \sum_{k \in K} y_k d'_k - \sum_{k \in K} x_k d_k$$

where the variables are

$$x_k = \begin{cases} 1, & \text{if gene or promoter } k \text{ is the first gene or promoter within the deleted segment} \\ 0, & \text{otherwise} \end{cases}$$

and

$$y_k = \begin{cases} 1, & \text{if gene or promoter } k \text{ is immediately after the end of the deleted segment} \\ 0, & \text{otherwise} \end{cases}$$

where

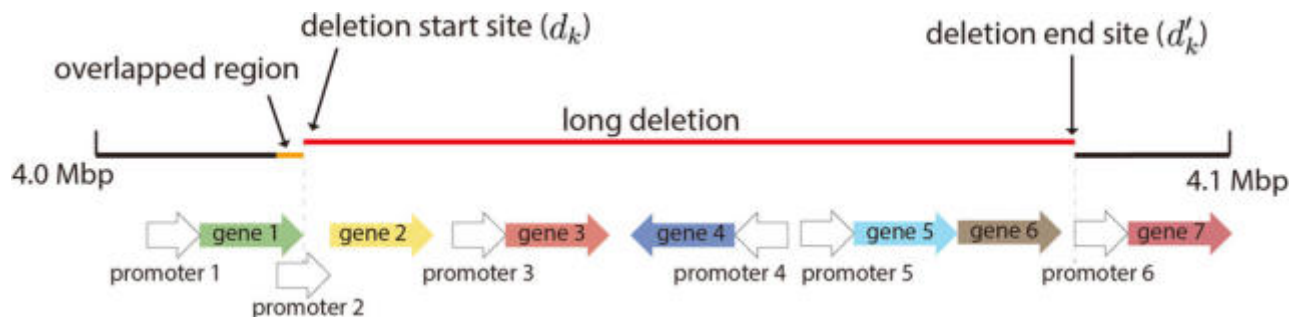
$$\sum_{k \in K} x_k = 1$$

and

$$\sum_{k \in K} y_k = 1$$

and the parameters are

- $d_k$ : Position of the first nucleotide of the deleted sequence starting from the origin of replication when gene or promoter  $k$  is selected to be deleted in the beginning of the stretch. Note that  $d_k$  is not always the start site of a gene or promoter. It is the first nucleotide of the nonoverlapped region between the gene/promoter  $k$  and gene/promoter  $k - 1$
- $d'_k$ : Position of the first nucleotide of the gene or promoter  $k$  immediately after the deleted sequence



MaxMassReclamation is the MinGenome algorithm with the objective function above replaced with the following objective function:

$$\max \sum_{k \in K} y_k \sum_{i=0}^{k-1} m_i - \sum_{k \in K} x_k \sum_{i=0}^{k-1} m_i$$

where  $m_i$  is the measured protein mass of gene  $i$  and  $\sum_{i=1}^k m_i$  is the cumulative protein mass of all genes from the origin of replication to the  $k$ th gene. By subtracting the cumulative protein mass of the start gene ( $x_k$ ) from the cumulative protein mass of the end gene ( $y_k$ ), we obtain the cumulative protein mass of the interval for the optimal solution

We maximize the mass that is knocked out based on absolute quantitative proteomics [Schmidt et al 2015](http://www.nature.com/articles/nbt.3418) (<http://www.nature.com/articles/nbt.3418>) in (fg/cell)

We then run `Max mass reclamation` iteratively without replacement.

Instead of maximizing the length of the knockout, we maximize the amount of protein mass that is reclaimed by

```
In [28]: import random
import numpy as np
genes = range (1,50)
essen_genes= [1,2,23,48]
mass = np.random.randint(1, 10, 50)
cum_mass = np.cumsum(mass)
cum_mass[:4]
mass[:4]
```

```
Out[28]: array([ 9, 12, 20, 26])
```

```
Out[28]: array([9, 3, 8, 6])
```

```
In [26]: x_k =22
y_k =47
cum_mass[y_k-1] - cum_mass[x_k-1], np.sum(mass[x_k:y_k])
```

```
Out[26]: (108, 108)
```

```
In [ ]: Hammr model for the allignment
whole genome allignment
Validation and error message.
```

```
In [32]: import random
import numpy as np
genes = range (1,50)
essen_genes= [1,2,23,48]
mass = np.random.randint(1, 10, 50)
cum_mass = np.insert(np.cumsum(mass), 0, 0)
cum_mass[:5], mass[:4]
```

```
Out[32]: (array([ 0,  8, 14, 21, 28]), array([8, 6, 7, 7]))
```

```
In [33]: x_k =22
y_k =47
cum_mass[y_k] - cum_mass[x_k], np.sum(mass[x_k:y_k])
```

```
Out[33]: (130, 130)
```

```
In [1]: class PDF(object):
        def __init__(self, pdf, size=(1200,600)):
            self.pdf = pdf
            self.size = size

        def _repr_html_(self):
            return '<iframe src={0} width={1[0]} height={1[1]}></iframe>'.format(self.pdf, self.size)

        def _repr_latex_(self):
            return r'\includegraphics[width=1.0\textwidth]{{{0}}}'.format(self.pdf)
```

## Adding violacein pathway to E. coli

```

In [2]: import cobra, pandas as pd, os, sys
        from cobra.core import Reaction, Metabolite
        ecolidir = os.path.join('data', 'Ecoli')
        khk = cobra.io.read_sbml_model(os.path.join(ecolidir, 'iJ01366.xml'))
        khk_induced = khk.copy()
        vioA = cobra.core.Reaction('vioA',
                                    name='L-tryptophan oxidase',
                                    subsystem='Violacein biosynthesis',
                                    lower_bound=0,
                                    upper_bound = 100)
        vioB = cobra.core.Reaction('vioB',
                                    name='N-[2-(carboxylatoamino)-1,2-bis(1H-indol-3-yl)ethyl]carbamate synthase',
                                    subsystem='Violacein biosynthesis',
                                    lower_bound=0,
                                    upper_bound = 100)
        vioE = cobra.core.Reaction('vioE',
                                    name='protodeoxyviolaceinate synthase',
                                    subsystem='Violacein biosynthesis',
                                    lower_bound=0,
                                    upper_bound = 100)
        EX_violacein_e = cobra.core.Reaction('EX_violacein_e',
                                              name='protodeoxyviolaceinate export',
                                              subsystem='Violacein biosynthesis',
                                              lower_bound=0,
                                              upper_bound = 100)
        CPD_11890_c = cobra.core.Metabolite('CPD_11890_c',
                                              name='2-imino-3-(indol-3-yl)propanoate',
                                              formula='C11H9N2O2',
                                              charge=-1,
                                              compartment='c'
                                              )
        CPD_11890_c.annotation = {'biocyc': 'META:CPD-11890'}
        CPD_19471_c = cobra.core.Metabolite('CPD_19471_c',
                                              name='N-[2-(carboxylatoamino)-1,2-bis(1H-indol-3-yl)ethyl]carbamate',
                                              formula='C22H16N4O4',
                                              charge=-2,
                                              compartment='c'
                                              )
        CPD_19471_c.annotation = {'biocyc': 'META:CPD-19471'}
        CPD_14320_c = cobra.core.Metabolite('CPD_14320_c',
                                              name='protodeoxyviolaceinate',
                                              formula='C21H14N3O2',
                                              charge=-1,
                                              compartment='c'
                                              )
        CPD_14320_c.annotation = {'biocyc': 'META:CPD-14320'}
        khk_induced.add_metabolites([CPD_11890_c, CPD_19471_c, CPD_14320_c])
        khk_induced.add_reactions([vioA, vioB, vioE, EX_violacein_e])
        vioA.build_reaction_from_string('trp_L_c + o2_c --> h2o2_c + CPD_11890_c + h_c')
        vioB.build_reaction_from_string('2 CPD_11890_c + h2o2_c --> CPD_19471_c + 2 h2o_c')

```

```

vioE.build_reaction_from_string('CPD_19471_c + 2 h_c --> co2_c + nh4_c +
CPD_14320_c')
EX_violacein_e.build_reaction_from_string('CPD_14320_c -->')
cobra.io.save_json_model(khk_induced,
                          os.path.join(ecolidir,
                                         'khk_induced.json'),
                          pretty=True,
                          sort=True)

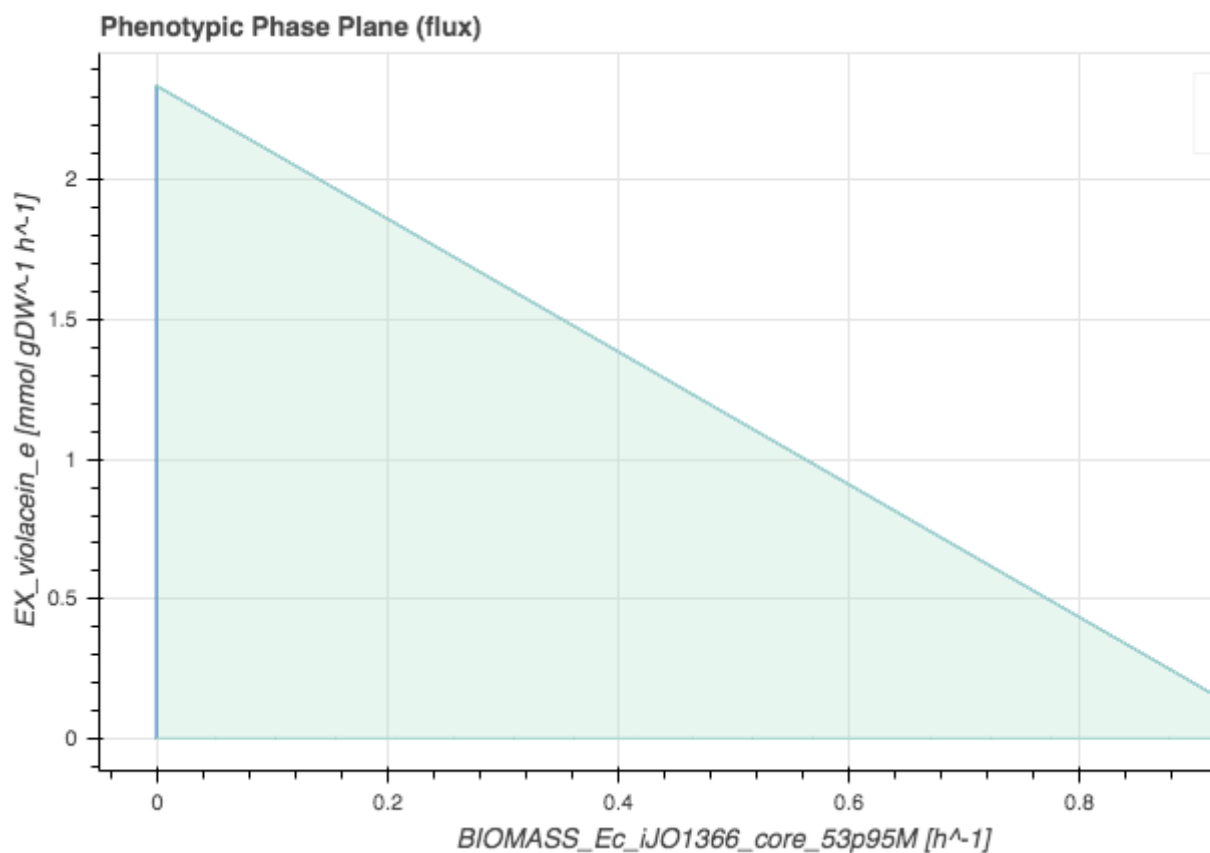
```

```

In [3]: %matplotlib inline
import cameo
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"
from cameo import phenotypic_phase_plane

p = phenotypic_phase_plane(khk_induced, variables=["BIOMASS_Ec_iJO1366_c
ore_53p95M"], objective='EX_violacein_e')
p.plot()

```



```
In [28]: from IPython.display import HTML
with khk_induced:
    khk_induced.objective = 'vioE'
    violacein_obj = khk_induced.optimize().fluxes
    tol = 1e-3
    display(HTML(violacein_obj[(violacein_obj > tol) | (violacein_obj <
-tol)].to_frame('fluxes').sort_values('fluxes', ascending=False).to_html
()))
```

	<b>fluxes</b>
<b>EX_h2o_e</b>	52.982482
<b>ATPS4rpp</b>	20.646837
<b>CYTBO3_4pp</b>	14.737226
<b>NADH16pp</b>	14.128954
<b>EX_co2_e</b>	10.877372
<b>GLCtex_copy1</b>	10.000000
<b>GAPD</b>	9.964964
<b>ENO</b>	9.964964
<b>O2tpp</b>	9.707786
<b>O2tex</b>	9.707786
<b>GLCt2pp</b>	9.391727
<b>EX_h_e</b>	9.356691
<b>XYLI2</b>	7.964964
<b>HEX7</b>	7.964964
<b>NH4tex</b>	7.017518
<b>NH4tpp</b>	7.017518
<b>TRPS3</b>	4.678345
<b>vioA</b>	4.678345
<b>SHKK</b>	4.678345
<b>DDPA</b>	4.678345
<b>DHQTi</b>	4.678345
<b>PSCVT</b>	4.678345
<b>PRAIi</b>	4.678345
<b>IGPS</b>	4.678345
<b>GLNS</b>	4.678345
<b>SHK3Dr</b>	4.678345
<b>PRPPS</b>	4.678345
<b>CHORS</b>	4.678345
<b>ANPRT</b>	4.678345
<b>ANS</b>	4.678345
<b>DHQS</b>	4.678345
<b>ADK1</b>	4.678345



	<b>fluxes</b>
<b>TPI</b>	4.643309
<b>FBA</b>	3.964964
<b>PFK</b>	3.964964
<b>ATPM</b>	3.150000
<b>MDH</b>	2.947445
<b>EX_violacein_e</b>	2.339173
<b>vioE</b>	2.339173
<b>vioB</b>	2.339173
<b>G6PDH2r</b>	2.035036
<b>PGL</b>	2.035036
<b>GND</b>	2.035036
<b>HEX1</b>	1.426764
<b>FBA3</b>	0.678345
<b>TKT1</b>	0.678345
<b>PFK_3</b>	0.678345
<b>ICDHyr</b>	0.608273
<b>AKGDH</b>	0.608273
<b>CS</b>	0.608273
<b>SUCDi</b>	0.608273
<b>GLCptspp</b>	0.608273
<b>PDH</b>	0.608273
<b>ACONTa</b>	0.608273
<b>FUM</b>	0.608273
<b>ACONTb</b>	0.608273
<b>SUCOAS</b>	-0.608273
<b>MOX</b>	-2.339173
<b>RPE</b>	-3.321655
<b>TKT2</b>	-4.000000
<b>PPKr</b>	-4.678345
<b>TRPAS2</b>	-4.678345
<b>RPI</b>	-5.356691
<b>EX_nh4_e</b>	-7.017518

	fluxes
Htex	-9.356691
EX_o2_e	-9.707786
PGM	-9.964964
PGK	-9.964964
EX_glc__D_e	-10.000000
CO2tex	-10.877372
CO2tp	-10.877372
H2Otex	-52.982482
H2Otp	-52.982482

In [ ]:


In [20]:  $10 \times 6 - 10.877 \times 1 - 2.33 \times 21$

Out[20]: 0.192999999999999784

# Genome view of predicted KOs

```
In [5]: PDF('Visualizations/ProteinMassReclaimed.pdf')
```

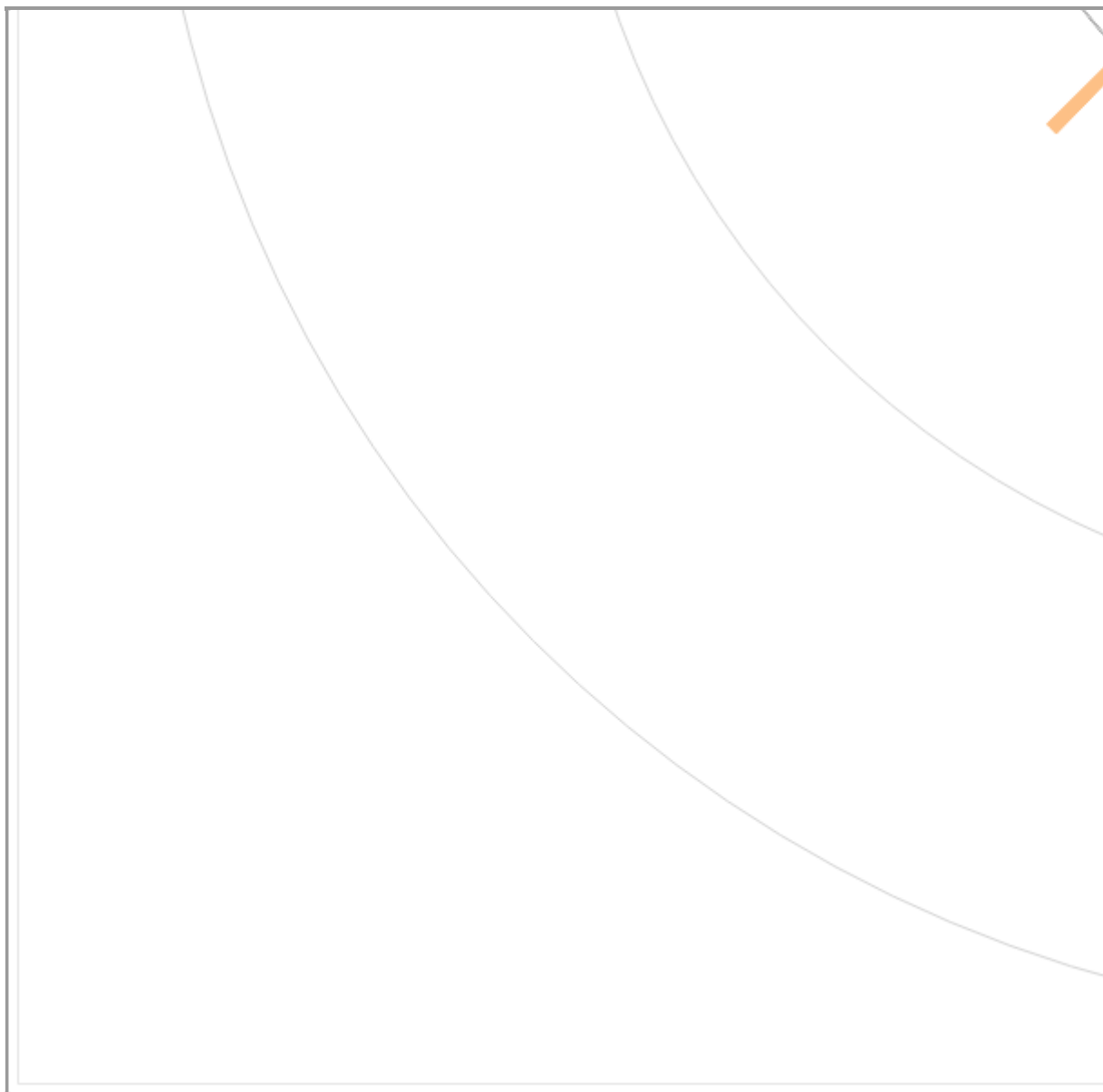
```
Out[5]:
```



## Metabolic map of predicted KOs

```
In [6]: PDF('Visualizations/reclaimed.png')
```

```
Out[6]:
```



## E. coli genes and promoters

```
In [35]: import pandas as pd
df = pd.read_excel('data/Ecoli/genes_and_promoters.xlsx', sheet_name='al
l_clear_v2')
df

start, end = 190, 9191
genes_of_interval = df[(df['start'] >= start) & \
                        (df['end'] <= end) & \
                        (df['class'] == 'gene')]
genes_of_interval['gene_or_promoter'].tolist()
```

Out[35]:

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
0	PM00249	148	189	1	promoter	[b0001, b0002, b0003, b0004]	148
1	b0001	190	255	1	gene	NaN	190
2	b0002	337	2799	1	gene	NaN	337
3	b0003	2801	3733	1	gene	NaN	2801
4	b0004	3734	5020	1	gene	NaN	3734
5	b0005	5234	5530	1	gene	NaN	5234
6	b0006	5683	6459	-1	gene	NaN	5683
7	b0007	6529	7959	-1	gene	NaN	6529
8	PM0-9956	8191	8237	1	promoter	[b0008]	8191
9	b0008	8238	9191	1	gene	NaN	8238
10	b0009	9306	9893	1	gene	NaN	9306
11	b0010	9928	10494	-1	gene	NaN	9928
12	b0011	10643	11356	-1	gene	NaN	10643
13	b0013	11382	11786	-1	gene	NaN	11382
14	PM00316	12048	12162	1	promoter	[b0014, b0015]	12048
15	b0014	12163	14079	1	gene	NaN	12163
16	b0015	14168	15298	1	gene	NaN	14168
17	b0016	15445	16557	1	gene	NaN	15445
18	b0018	16751	16960	-1	gene	NaN	16751
19	b4412	16751	16903	-1	gene	NaN	16960
20	b4413	16952	17006	1	gene	NaN	16960
21	b0019	17489	18655	1	gene	NaN	17489
22	b0020	18715	19620	1	gene	NaN	18715
23	b0021	19811	20314	-1	gene	NaN	19811
24	b0022	20233	20508	-1	gene	NaN	20314
25	b0023	20815	21078	-1	gene	NaN	20815
26	PM0-8222	21079	21210	-1	promoter	[b0023]	21079
27	b0024	21181	21399	1	gene	NaN	21210

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>28</b>	PM0-8866	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399
<b>29</b>	b0025	21407	22348	1	gene	NaN	21407
...	...	...	...	...	...	...	...
<b>6051</b>	b4383	4619603	4620826	1	gene	NaN	4619603
<b>6052</b>	b4384	4620883	4621602	1	gene	NaN	4620883
<b>6053</b>	b4385	4621769	4623100	1	gene	NaN	4621769
<b>6054</b>	b4386	4623101	4624117	-1	gene	NaN	4623101
<b>6055</b>	b4387	4624145	4624789	-1	gene	NaN	4624145
<b>6056</b>	PM00508	4624790	4624799	-1	promoter	[b4386, b4387]	4624790
<b>6057</b>	PM00509	4624856	4624894	1	promoter	[b4389, b4388]	4624856
<b>6058</b>	b4388	4624895	4625863	1	gene	NaN	4624895
<b>6059</b>	b4389	4625912	4627294	1	gene	NaN	4625912
<b>6060</b>	b4390	4627315	4628547	1	gene	NaN	4627315
<b>6061</b>	b4391	4628855	4630522	-1	gene	NaN	4628855
<b>6062</b>	PM0-46112	4630700	4630732	1	promoter	[b4392]	4630700
<b>6063</b>	b4392	4630733	4632670	1	gene	NaN	4630733
<b>6064</b>	PM00095	4632704	4632759	1	promoter	[b4393]	4632704
<b>6065</b>	b4393	4632760	4633086	1	gene	NaN	4632760
<b>6066</b>	b4394	4633233	4633745	-1	gene	NaN	4633233
<b>6067</b>	b4395	4633797	4634444	1	gene	NaN	4633797
<b>6068</b>	b4396	4634441	4635310	-1	gene	NaN	4634444
<b>6069</b>	PM0-9728	4635243	4635520	1	promoter	[b4399, b4398, b4397]	4635310
<b>6070</b>	PM0-8184	4635311	4635353	-1	promoter	[b4396]	4635520
<b>6071</b>	PM00603	4635477	4635520	1	promoter	[b4397, b4398, b4399, b4400]	4635520
<b>6072</b>	b4397	4635521	4635994	1	gene	NaN	4635521

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>6073</b>	b4398	4636007	4636696	1	gene	NaN	4636007
<b>6074</b>	b4399	4636696	4638120	1	gene	NaN	4636696
<b>6075</b>	PM0-10215	4638160	4638177	1	promoter	[b4400]	4638160
<b>6076</b>	b4400	4638178	4639530	1	gene	NaN	4638178
<b>6077</b>	b4401	4639590	4640306	-1	gene	NaN	4639590
<b>6078</b>	PM205	4640307	4640801	-1	promoter	[b4401]	4640307
<b>6079</b>	b4402	4640402	4640542	1	gene	NaN	4640801
<b>6080</b>	b4403	4640942	4641628	1	gene	NaN	4640942

6081 rows × 8 columns

```
Out[35]: ['b0001', 'b0002', 'b0003', 'b0004', 'b0005', 'b0006', 'b0007', 'b0008']
```



In [2]: df

Out[ 2 ]:

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
0	PM00249	148	189	1	promoter	[b0001, b0002, b0003, b0004]	148
1	b0001	190	255	1	gene	NaN	190
2	b0002	337	2799	1	gene	NaN	337
3	b0003	2801	3733	1	gene	NaN	2801
4	b0004	3734	5020	1	gene	NaN	3734
5	b0005	5234	5530	1	gene	NaN	5234
6	b0006	5683	6459	-1	gene	NaN	5683
7	b0007	6529	7959	-1	gene	NaN	6529
8	PM0-9956	8191	8237	1	promoter	[b0008]	8191
9	b0008	8238	9191	1	gene	NaN	8238
10	b0009	9306	9893	1	gene	NaN	9306
11	b0010	9928	10494	-1	gene	NaN	9928
12	b0011	10643	11356	-1	gene	NaN	10643
13	b0013	11382	11786	-1	gene	NaN	11382
14	PM00316	12048	12162	1	promoter	[b0014, b0015]	12048
15	b0014	12163	14079	1	gene	NaN	12163
16	b0015	14168	15298	1	gene	NaN	14168
17	b0016	15445	16557	1	gene	NaN	15445
18	b0018	16751	16960	-1	gene	NaN	16751
19	b4412	16751	16903	-1	gene	NaN	16960
20	b4413	16952	17006	1	gene	NaN	16960
21	b0019	17489	18655	1	gene	NaN	17489
22	b0020	18715	19620	1	gene	NaN	18715
23	b0021	19811	20314	-1	gene	NaN	19811
24	b0022	20233	20508	-1	gene	NaN	20314
25	b0023	20815	21078	-1	gene	NaN	20815
26	PM0-8222	21079	21210	-1	promoter	[b0023]	21079
27	b0024	21181	21399	1	gene	NaN	21210

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>28</b>	PM0-8866	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399
<b>29</b>	b0025	21407	22348	1	gene	NaN	21407
...	...	...	...	...	...	...	...
<b>6051</b>	b4383	4619603	4620826	1	gene	NaN	4619603
<b>6052</b>	b4384	4620883	4621602	1	gene	NaN	4620883
<b>6053</b>	b4385	4621769	4623100	1	gene	NaN	4621769
<b>6054</b>	b4386	4623101	4624117	-1	gene	NaN	4623101
<b>6055</b>	b4387	4624145	4624789	-1	gene	NaN	4624145
<b>6056</b>	PM00508	4624790	4624799	-1	promoter	[b4386, b4387]	4624790
<b>6057</b>	PM00509	4624856	4624894	1	promoter	[b4389, b4388]	4624856
<b>6058</b>	b4388	4624895	4625863	1	gene	NaN	4624895
<b>6059</b>	b4389	4625912	4627294	1	gene	NaN	4625912
<b>6060</b>	b4390	4627315	4628547	1	gene	NaN	4627315
<b>6061</b>	b4391	4628855	4630522	-1	gene	NaN	4628855
<b>6062</b>	PM0-46112	4630700	4630732	1	promoter	[b4392]	4630700
<b>6063</b>	b4392	4630733	4632670	1	gene	NaN	4630733
<b>6064</b>	PM00095	4632704	4632759	1	promoter	[b4393]	4632704
<b>6065</b>	b4393	4632760	4633086	1	gene	NaN	4632760
<b>6066</b>	b4394	4633233	4633745	-1	gene	NaN	4633233
<b>6067</b>	b4395	4633797	4634444	1	gene	NaN	4633797
<b>6068</b>	b4396	4634441	4635310	-1	gene	NaN	4634444
<b>6069</b>	PM0-9728	4635243	4635520	1	promoter	[b4399, b4398, b4397]	4635310
<b>6070</b>	PM0-8184	4635311	4635353	-1	promoter	[b4396]	4635520
<b>6071</b>	PM00603	4635477	4635520	1	promoter	[b4397, b4398, b4399, b4400]	4635520
<b>6072</b>	b4397	4635521	4635994	1	gene	NaN	4635521

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_selec
<b>6073</b>	b4398	4636007	4636696	1	gene	NaN	4636007
<b>6074</b>	b4399	4636696	4638120	1	gene	NaN	4636696
<b>6075</b>	PM0-10215	4638160	4638177	1	promoter	[b4400]	4638160
<b>6076</b>	b4400	4638178	4639530	1	gene	NaN	4638178
<b>6077</b>	b4401	4639590	4640306	-1	gene	NaN	4639590
<b>6078</b>	PM205	4640307	4640801	-1	promoter	[b4401]	4640307
<b>6079</b>	b4402	4640402	4640542	1	gene	NaN	4640801
<b>6080</b>	b4403	4640942	4641628	1	gene	NaN	4640942

6081 rows × 8 columns

## Protein measurements in femtograms from Schmidt et al Table S13

```
In [40]: proteins_fg = pd.read_excel('data/Ecoli/Schmidt/nbt.3418-S2.xlsx', sheet
      _name='Table S13', skiprows=3)
      proteins_fg
```

Out[40]:

	Uniprot Accession	Gene	Description	Cellular protein location (according to www.uniprot.org)	A14.07032	A14.07032
0	P04825	pepN	Aminopeptidase N OS=Escherichia coli (strain K...	Cell inner membrane	4.633803e-01	4.113094e-01
1	P0C0V0	degP	Protease do OS=Escherichia coli (strain K12) G...	Cell inner membrane	2.128826e-01	2.076923e-01
2	P0AAI3	ftsH	ATP-dependent zinc metalloprotease FtsH OS=Esc...	Cell inner membrane	4.640302e-01	4.143873e-01
3	P0ABC7	hflK	Modulator of FtsH protease HflK OS=Escherichia...	Cell inner membrane	1.772662e-01	1.711126e-01
4	P08506	dacC	D-alanyl-D-alanine carboxypeptidase dacC OS=Es...	Cell inner membrane	8.301700e-02	5.860804e-02
5	P0ABC3	hflC	Modulator of FtsH protease HflC OS=Escherichia...	Cell inner membrane	1.300773e-01	1.328812e-01
6	P0AEB2	dacA	D-alanyl-D-alanine carboxypeptidase dacA OS=Es...	Cell inner membrane	1.552737e-01	1.336722e-01
7	P0AG14	sohB	Probable protease sohB OS=Escherichia coli (st...	Cell inner membrane	2.783188e-02	2.469125e-02
8	P23865	prc	Tail-specific protease OS=Escherichia coli (st...	Cell inner membrane	7.892680e-02	7.982559e-02
9	P23894	htpX	Protease HtpX OS=Escherichia coli (strain K12)...	Cell inner membrane	2.854932e-02	3.011286e-02
10	P00803	lepB	Signal peptidase I OS=Escherichia coli (strain...	Cell inner membrane	1.444528e-02	1.591279e-02

	Uniprot Accession	Gene	Description	Cellular protein location (according to www.uniprot.org)	A14.07032	A14.07032
11	P0AEE3	degS	Protease degS OS=Escherichia coli (strain K12)...	Cell inner membrane	2.546474e-03	2.717599e-03
12	P08395	sppA	Protease 4 OS=Escherichia coli (strain K12) GN...	Cell inner membrane	9.592818e-03	7.771863e-03
13	P0AEH1	rseP	Regulator of sigma E protease OS=Escherichia C...	Cell inner membrane	1.117086e-03	9.208512e-04
14	P0ABB4	atpD	ATP synthase subunit beta OS=Escherichia coli ...	Cell inner membrane	3.271261e+00	3.541148e+00
15	P0ABB0	atpA	ATP synthase subunit alpha OS=Escherichia coli...	Cell inner membrane	2.085016e+00	2.419878e+00
16	P0ABA0	atpF	ATP synthase subunit b OS=Escherichia coli (st...	Cell inner membrane	3.491107e-01	3.987848e-01
17	P0A6E6	atpC	ATP synthase epsilon chain OS=Escherichia coli...	Cell inner membrane	3.374377e-01	4.792207e-01
18	P0ABA4	atpH	ATP synthase subunit delta OS=Escherichia coli...	Cell inner membrane	3.809520e-01	3.950800e-01
19	P0AE06	acrA	Acriflavine resistance protein A OS=Escherichi...	Cell inner membrane	2.532360e-01	3.165534e-01
20	P0ABA6	atpG	ATP synthase gamma chain OS=Escherichia coli (...)	Cell inner membrane	1.925043e-01	2.094286e-01

	Uniprot Accession	Gene	Description	Cellular protein location (according to www.uniprot.org)	A14.07032	A14.07032
21	P0AB98	atpB	ATP synthase subunit a OS=Escherichia coli (st...	Cell inner membrane	1.124002e-01	8.509506e-02
22	P0AAF6	artP	Arginine transport ATP-binding protein ArtP OS...	Cell inner membrane	4.511575e-02	3.379879e-02
23	P31224	acrB	Acriflavine resistance protein B OS=Escherichi...	Cell inner membrane	1.211155e-01	9.604174e-02
24	P68699	atpE	ATP synthase subunit c OS=Escherichia coli (st...	Cell inner membrane	7.510506e-03	7.719950e-03
25	P32705	actP	Cation/acetate symporter ActP OS=Escherichia c...	Cell inner membrane	2.921024e-02	5.817646e-03
26	P0ABB8	mgtA	Magnesium- transporting ATPase, P-type 1 OS=Esc...	Cell inner membrane	2.552800e-02	1.317152e-02
27	P0AEC3	arcB	Aerobic respiration control sensor protein Arc...	Cell inner membrane	2.063283e-02	1.494044e-02
28	P31466	adeP	Probable adenine permease PurP OS=Escherichia ...	Cell inner membrane	4.120856e-03	3.361424e-03
29	P15993	aroP	Aromatic amino acid transport protein AroP OS=...	Cell inner membrane	2.101542e-03	2.063010e-03
...	...	...	...	...	...	...
1144	P0AE85	cpxP	Periplasmic protein CpxP OS=Escherichia coli (...)	Periplasm	5.819712e-03	2.761288e-03



	Uniprot Accession	Gene	Description	Cellular protein location (according to www.uniprot.org)	A14.07032	A14.07032
1145	P78067	ynjE	Thiosulfate sulfurtransferase YnjE OS=Escheric...	Periplasm	9.458087e-04	1.097532e-03
1146	P75954	ycfS	Probable L,D- transpeptidase YcfS OS=Escherichi...	Periplasm	1.577965e-03	8.271217e-04
1147	P28249	asmA	Protein AsmA OS=Escherichia coli (strain K12) ...	Periplasm	4.712922e-03	4.015742e-03
1148	P21338	rna	Ribonuclease I OS=Escherichia coli (strain K12...	Periplasm	4.128799e-03	4.835157e-03
1149	P39176	erfK	Probable L,D- transpeptidase ErfK/SrfK OS=Esche...	Periplasm	4.268847e-03	3.406840e-03
1150	P76128	ddpA	Probable D,D- dipeptide-binding periplasmic pro...	Periplasm	3.095786e-03	1.683892e-03
1151	P31550	thiB	Thiamine-binding periplasmic protein OS=Escher...	Periplasm	4.194574e-03	5.067506e-03
1152	P00811	ampC	Beta-lactamase OS=Escherichia coli (strain K12...	Periplasm	2.107232e-03	2.823426e-03
1153	Q47622	sapA	Peptide transport periplasmic protein sapA OS=...	Periplasm	4.274385e-03	5.000386e-03
1154	P07102	appA	Periplasmic AppA protein OS=Escherichia coli (...)	Periplasm	1.678385e-03	1.120738e-03

	Uniprot Accession	Gene	Description	Cellular protein location (according to www.uniprot.org)	A14.07032	A14.07032
1155	P76341	hiuH	5-hydroxyisourate hydrolase OS=Escherichia col...	Periplasm	1.992851e-03	1.392178e-03
1156	P26365	amiB	N- acetylmuramoyl-L- alanine amidase AmiB OS=Esc...	Periplasm	8.730257e-04	9.634147e-04
1157	P0AA57	yobA	Protein YobA OS=Escherichia coli (strain K12) ...	Periplasm	1.668325e-03	3.333927e-03
1158	P24228	dacB	D-alanyl-D-alanine carboxypeptidase dacB OS=Es...	Periplasm	1.010670e-03	8.245424e-04
1159	P76342	yedY	Sulfoxide reductase catalytic subunit yedY OS=...	Periplasm	6.456927e-04	2.099756e-04
1160	P63883	amiC	N- acetylmuramoyl-L- alanine amidase AmiC OS=Esc...	Periplasm	2.291386e-04	6.179431e-05
1161	Q47537	tauA	Taurine-binding periplasmic protein OS=Escheri...	Periplasm	9.619339e-05	2.116520e-04
1162	P0C0T5	mepA	Penicillin- insensitive murein endopeptidase OS...	Periplasm	3.677398e-04	4.714331e-04
1163	P00805	ansB	L-asparaginase 2 OS=Escherichia coli (strain K...	Periplasm	3.765238e-03	3.937850e-03
1164	P33590	nikA	Nickel-binding periplasmic protein OS=Escheric...	Periplasm	4.456321e-05	4.388987e-05

	Uniprot Accession	Gene	Description	Cellular protein location (according to www.uniprot.org)	A14.07032	A14.07032
1165	P0AAA9	zraP	Zinc resistance-associated protein OS=Escheric...	Periplasm	2.629344e-05	1.332028e-05
1166	P0AAX3	ybiJ	Uncharacterized protein YbiJ OS=Escherichia co...	Periplasm	7.906730e-05	2.362067e-05
1167	P36560	asr	Acid shock protein OS=Escherichia coli (strain...	Periplasm	2.230201e-06	1.392737e-04
1168	P0ABK9	nrfA	Cytochrome c-552 OS=Escherichia coli (strain K...	Periplasm	8.239733e-06	1.232533e-05
1169	P0AFI5	pbpG	D-alanyl-D-alanine endopeptidase OS=Escherichi...	Periplasm	7.438443e-07	7.231266e-07
1170	P31697	fimC	Chaperone protein fimC OS=Escherichia coli (st...	Periplasm	3.801773e-04	3.714249e-04
1171	P33219	yebF	Protein yebF OS=Escherichia coli (strain K12) ...	Secreted	4.241761e-02	1.825890e-02
1172	P37773	mpl	UDP-N- acetylmuramate:L- alanyl-gamma-D- glutamyl...	Secreted	5.095097e-02	5.834843e-02
1173	P04949	fliC	Flagellin OS=Escherichia coli (strain K12) GN=...	Secreted	5.996419e-04	4.275764e-04

1174 rows × 73 columns

## Switch to Uniprot

```
In [41]: BW25113_fg = proteins_fg[['Uniprot Accession',  
                                   'Description', 'Gene',  
                                   'Cellular protein location (according to www.u  
niproto.org)',  
                                   'A14.07036', 'A14.07037', 'A14.07038' ]].set_i  
ndex('Uniprot Accession')  
BW25113_fg
```

Out[41]:

	Description	Gene	Cellular protein location (according to www.uniprot.org)	A14.07036	A14.07037	/
Uniprot Accession						
<b>P04825</b>	Aminopeptidase N OS=Escherichia coli (strain K...	pepN	Cell inner membrane	2.136611e-01	2.127355e-01	2.10
<b>P0C0V0</b>	Protease do OS=Escherichia coli (strain K12) G...	degP	Cell inner membrane	2.587472e-01	2.753042e-01	2.73
<b>P0AAI3</b>	ATP-dependent zinc metalloprotease FtsH OS=Esc...	ftsH	Cell inner membrane	2.440695e-01	2.623177e-01	2.72
<b>P0ABC7</b>	Modulator of FtsH protease HflK OS=Escherichia...	hflK	Cell inner membrane	9.544338e-02	1.122174e-01	1.01
<b>P08506</b>	D-alanyl-D-alanine carboxypeptidase dacC OS=Es...	dacC	Cell inner membrane	6.989639e-02	7.144231e-02	7.43
<b>P0ABC3</b>	Modulator of FtsH protease HflC OS=Escherichia...	hflC	Cell inner membrane	7.555463e-02	8.089087e-02	8.05
<b>P0AEB2</b>	D-alanyl-D-alanine carboxypeptidase dacA OS=Es...	dacA	Cell inner membrane	1.008689e-01	9.944921e-02	9.93
<b>P0AG14</b>	Probable protease sohB OS=Escherichia coli (st...	sohB	Cell inner membrane	2.659054e-02	2.699793e-02	2.71
<b>P23865</b>	Tail-specific protease OS=Escherichia coli (st...	prc	Cell inner membrane	6.429820e-02	6.502070e-02	6.32
<b>P23894</b>	Protease HtpX OS=Escherichia coli (strain K12)...	htpX	Cell inner membrane	2.133862e-02	2.654830e-02	2.38

	Description	Gene	Cellular protein location (according to www.uniprot.org)	A14.07036	A14.07037	I
Uniprot Accession						
<b>P00803</b>	Signal peptidase I OS=Escherichia coli (strain...	lepB	Cell inner membrane	8.768313e-03	8.758116e-03	8.66
<b>P0AEE3</b>	Protease degS OS=Escherichia coli (strain K12)...	degS	Cell inner membrane	4.593809e-03	4.156953e-03	5.50
<b>P08395</b>	Protease 4 OS=Escherichia coli (strain K12) GN...	sppA	Cell inner membrane	4.546397e-03	7.645568e-03	5.86
<b>P0AEH1</b>	Regulator of sigma E protease OS=Escherichia c...	rseP	Cell inner membrane	5.413127e-04	4.640121e-04	5.66
<b>P0ABB4</b>	ATP synthase subunit beta OS=Escherichia coli ...	atpD	Cell inner membrane	1.981909e+00	2.016620e+00	2.02
<b>P0ABB0</b>	ATP synthase subunit alpha OS=Escherichia coli...	atpA	Cell inner membrane	1.395489e+00	1.333149e+00	1.22
<b>P0ABA0</b>	ATP synthase subunit b OS=Escherichia coli (st...	atpF	Cell inner membrane	2.572286e-01	2.507203e-01	2.57
<b>P0A6E6</b>	ATP synthase epsilon chain OS=Escherichia coli...	atpC	Cell inner membrane	2.560332e-01	2.525617e-01	2.30
<b>P0ABA4</b>	ATP synthase subunit delta OS=Escherichia coli...	atpH	Cell inner membrane	2.473332e-01	2.468104e-01	2.34
<b>P0AE06</b>	Acriflavine resistance protein A OS=Escherichi...	acrA	Cell inner membrane	1.696292e-01	1.660571e-01	1.73

	Description	Gene	Cellular protein location (according to www.uniprot.org)	A14.07036	A14.07037	I
Uniprot Accession						
<b>P0ABA6</b>	ATP synthase gamma chain OS=Escherichia coli (...)	atpG	Cell inner membrane	1.095532e-01	1.257258e-01	1.20
<b>P0AB98</b>	ATP synthase subunit a OS=Escherichia coli (st...	atpB	Cell inner membrane	6.345721e-02	5.355321e-02	5.80
<b>P0AAF6</b>	Arginine transport ATP-binding protein ArtP OS...	artP	Cell inner membrane	2.767890e-02	2.636776e-02	2.60
<b>P31224</b>	Acriflavine resistance protein B OS=Escherichi...	acrB	Cell inner membrane	5.319037e-02	5.819845e-02	5.25
<b>P68699</b>	ATP synthase subunit c OS=Escherichia coli (st...	atpE	Cell inner membrane	3.946795e-03	3.909397e-03	3.00
<b>P32705</b>	Cation/acetate symporter ActP OS=Escherichia C...	actP	Cell inner membrane	4.160314e-03	2.220756e-03	3.97
<b>P0ABB8</b>	Magnesium- transporting ATPase, P-type 1 OS=Esc...	mgtA	Cell inner membrane	1.158409e-02	1.073559e-02	1.07
<b>P0AEC3</b>	Aerobic respiration control sensor protein Arc...	arcB	Cell inner membrane	7.114456e-03	7.426018e-03	7.22
<b>P31466</b>	Probable adenine permease PurP OS=Escherichia ...	adeP	Cell inner membrane	6.308420e-03	7.098006e-03	7.65
<b>P15993</b>	Aromatic amino acid transport protein AroP OS=...	aroP	Cell inner membrane	1.750197e-02	1.647115e-02	1.56
...	...	...	...	...	...	...

	Description	Gene	Cellular protein location (according to www.uniprot.org)	A14.07036	A14.07037	I
Uniprot Accession						
<b>P0AE85</b>	Periplasmic protein CpxP OS=Escherichia coli (...)	cpxP	Periplasm	1.355103e-02	1.030928e-02	1.01
<b>P78067</b>	Thiosulfate sulfurtransferase YnjE OS=Escheric...	ynjE	Periplasm	1.342539e-03	1.081293e-03	2.62
<b>P75954</b>	Probable L,D-transpeptidase YcfS OS=Escherichi...	ycfS	Periplasm	1.158072e-03	1.849483e-03	1.06
<b>P28249</b>	Protein AsmA OS=Escherichia coli (strain K12) ...	asmA	Periplasm	2.805216e-03	3.166096e-03	2.54
<b>P21338</b>	Ribonuclease I OS=Escherichia coli (strain K12...	rna	Periplasm	2.477588e-03	2.469443e-03	2.74
<b>P39176</b>	Probable L,D-transpeptidase ErfK/SrfK OS=Esche...	erfK	Periplasm	3.641027e-03	4.208494e-03	3.95
<b>P76128</b>	Probable D,D-dipeptide-binding periplasmic pro...	ddpA	Periplasm	1.449479e-03	1.511605e-03	1.60
<b>P31550</b>	Thiamine-binding periplasmic protein OS=Escher...	thiB	Periplasm	2.144131e-03	1.108433e-03	8.97
<b>P00811</b>	Beta-lactamase OS=Escherichia coli (strain K12...	ampC	Periplasm	1.448446e-03	1.321616e-03	1.34
<b>Q47622</b>	Peptide transport periplasmic protein sapA OS=...	sapA	Periplasm	2.012290e-03	1.930414e-03	2.20



	Description	Gene	Cellular protein location (according to www.uniprot.org)	A14.07036	A14.07037	I
Uniprot Accession						
<b>P07102</b>	Periplasmic AppA protein OS=Escherichia coli (...)	appA	Periplasm	7.091263e-04	7.149016e-04	5.68
<b>P76341</b>	5-hydroxyisourate hydrolase OS=Escherichia col...	hiuH	Periplasm	9.835966e-03	8.832063e-03	9.71
<b>P26365</b>	N-acetylmuramoyl-L-alanine amidase AmiB OS=Esc...	amiB	Periplasm	1.218253e-03	1.105550e-03	1.32
<b>P0AA57</b>	Protein YobA OS=Escherichia coli (strain K12) ...	yobA	Periplasm	3.441222e-03	3.996124e-03	4.53
<b>P24228</b>	D-alanyl-D-alanine carboxypeptidase dacB OS=Es...	dacB	Periplasm	9.404004e-04	5.792969e-04	7.18
<b>P76342</b>	Sulfoxide reductase catalytic subunit yedY OS=...	yedY	Periplasm	1.187550e-03	1.230277e-03	1.33
<b>P63883</b>	N-acetylmuramoyl-L-alanine amidase AmiC OS=Esc...	amiC	Periplasm	5.476362e-07	5.506001e-07	1.22
<b>Q47537</b>	Taurine-binding periplasmic protein OS=Escheri...	tauA	Periplasm	2.940778e-04	1.483319e-04	3.57
<b>P0C0T5</b>	Penicillin-insensitive murein endopeptidase OS...	mepA	Periplasm	3.257470e-04	3.826091e-04	3.43
<b>P00805</b>	L-asparaginase 2 OS=Escherichia coli (strain K...	ansB	Periplasm	3.325940e-04	2.879741e-04	2.57

	Description	Gene	Cellular protein location (according to www.uniprot.org)	A14.07036	A14.07037	I
Uniprot Accession						
<b>P33590</b>	Nickel-binding periplasmic protein OS=Escheric...	nikA	Periplasm	4.086822e-04	3.153117e-04	4.66
<b>P0AAA9</b>	Zinc resistance-associated protein OS=Escheric...	zraP	Periplasm	7.818795e-06	7.629708e-05	5.23
<b>P0AAX3</b>	Uncharacterized protein YbiJ OS=Escherichia co...	ybiJ	Periplasm	1.296436e-03	1.068501e-03	7.85
<b>P36560</b>	Acid shock protein OS=Escherichia coli (strain...	asr	Periplasm	1.459242e-06	1.449694e-06	1.46
<b>P0ABK9</b>	Cytochrome c-552 OS=Escherichia coli (strain K...	nrfA	Periplasm	4.388781e-06	5.515976e-07	6.66
<b>P0AFI5</b>	D-alanyl-D-alanine endopeptidase OS=Escherichi...	pbpG	Periplasm	4.747779e-07	4.719466e-07	4.81
<b>P31697</b>	Chaperone protein fimC OS=Escherichia coli (st...	fimC	Periplasm	2.139856e-04	2.193095e-05	4.55
<b>P33219</b>	Protein yebF OS=Escherichia coli (strain K12) ...	yebF	Secreted	3.686415e-02	3.455457e-02	3.77
<b>P37773</b>	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl...	mpl	Secreted	3.146944e-02	2.922845e-02	2.82
<b>P04949</b>	Flagellin OS=Escherichia coli (strain K12) GN=...	fliC	Secreted	5.969345e-04	2.472486e-04	2.59

1174 rows × 6 columns

## Compute cumulative distribution of mass across the genome

```
In [54]: import numpy as np
bw_glucose = ['A14.07036', 'A14.07037', 'A14.07038']
b2u = pd.read_table('data/Ecoli/blatter-to-uniprot.tab', index_col = 'Blattner')
ecoli = df.join(b2u, on='gene_or_promoter').join(BW25113_fg, on='Uniprot')

minAbundance = ecoli[bw_glucose[0]].dropna().min()
display(minAbundance)
ecoli.loc[ecoli['class']== 'promoter', bw_glucose[0]] = minAbundance
ecoli[bw_glucose[0]].fillna(minAbundance,inplace=True)
ecoli['cumulativeMass'] = pd.concat([pd.Series([0]),
                                     np.cumsum(ecoli[bw_glucose[0]]),
                                     ignore_index=True).iloc[:-1]

ecoli.to_csv('data/Ecoli/{0}_cumulative_mass.tab'.format(bw_glucose[0]),
sep='\t', index=False)
ecoli
```

4.591655779049255e-07

Out[54]:

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
0	PM00249	148	189	1	promoter	[b0001, b0002, b0003, b0004]	148
1	b0001	190	255	1	gene	NaN	190
2	b0002	337	2799	1	gene	NaN	337
3	b0003	2801	3733	1	gene	NaN	2801
4	b0004	3734	5020	1	gene	NaN	3734
5	b0005	5234	5530	1	gene	NaN	5234
6	b0006	5683	6459	-1	gene	NaN	5683
7	b0007	6529	7959	-1	gene	NaN	6529
8	PM0-9956	8191	8237	1	promoter	[b0008]	8191
9	b0008	8238	9191	1	gene	NaN	8238
10	b0009	9306	9893	1	gene	NaN	9306
11	b0010	9928	10494	-1	gene	NaN	9928
12	b0011	10643	11356	-1	gene	NaN	10643
13	b0013	11382	11786	-1	gene	NaN	11382
14	PM00316	12048	12162	1	promoter	[b0014, b0015]	12048
15	b0014	12163	14079	1	gene	NaN	12163
16	b0015	14168	15298	1	gene	NaN	14168
17	b0016	15445	16557	1	gene	NaN	15445
18	b0018	16751	16960	-1	gene	NaN	16751
19	b4412	16751	16903	-1	gene	NaN	16960
20	b4413	16952	17006	1	gene	NaN	16960

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>21</b>	b0019	17489	18655	1	gene	NaN	17489
<b>22</b>	b0020	18715	19620	1	gene	NaN	18715
<b>23</b>	b0021	19811	20314	-1	gene	NaN	19811
<b>24</b>	b0022	20233	20508	-1	gene	NaN	20314
<b>25</b>	b0023	20815	21078	-1	gene	NaN	20815
<b>26</b>	PM0-8222	21079	21210	-1	promoter	[b0023]	21079
<b>27</b>	b0024	21181	21399	1	gene	NaN	21210
<b>28</b>	PM0-8866	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399
<b>29</b>	b0025	21407	22348	1	gene	NaN	21407
...	...	...	...	...	...	...	...
<b>6051</b>	b4383	4619603	4620826	1	gene	NaN	4619603
<b>6052</b>	b4384	4620883	4621602	1	gene	NaN	4620883
<b>6053</b>	b4385	4621769	4623100	1	gene	NaN	4621769
<b>6054</b>	b4386	4623101	4624117	-1	gene	NaN	4623101
<b>6055</b>	b4387	4624145	4624789	-1	gene	NaN	4624145
<b>6056</b>	PM00508	4624790	4624799	-1	promoter	[b4386, b4387]	4624790
<b>6057</b>	PM00509	4624856	4624894	1	promoter	[b4389, b4388]	4624856
<b>6058</b>	b4388	4624895	4625863	1	gene	NaN	4624895
<b>6059</b>	b4389	4625912	4627294	1	gene	NaN	4625912

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>6060</b>	b4390	4627315	4628547	1	gene	NaN	4627315
<b>6061</b>	b4391	4628855	4630522	-1	gene	NaN	4628855
<b>6062</b>	PM0-46112	4630700	4630732	1	promoter	[b4392]	4630700
<b>6063</b>	b4392	4630733	4632670	1	gene	NaN	4630733
<b>6064</b>	PM00095	4632704	4632759	1	promoter	[b4393]	4632704
<b>6065</b>	b4393	4632760	4633086	1	gene	NaN	4632760
<b>6066</b>	b4394	4633233	4633745	-1	gene	NaN	4633233
<b>6067</b>	b4395	4633797	4634444	1	gene	NaN	4633797
<b>6068</b>	b4396	4634441	4635310	-1	gene	NaN	4634444
<b>6069</b>	PM0-9728	4635243	4635520	1	promoter	[b4399, b4398, b4397]	4635310
<b>6070</b>	PM0-8184	4635311	4635353	-1	promoter	[b4396]	4635520
<b>6071</b>	PM00603	4635477	4635520	1	promoter	[b4397, b4398, b4399, b4400]	4635520
<b>6072</b>	b4397	4635521	4635994	1	gene	NaN	4635521
<b>6073</b>	b4398	4636007	4636696	1	gene	NaN	4636007
<b>6074</b>	b4399	4636696	4638120	1	gene	NaN	4636696
<b>6075</b>	PM0-10215	4638160	4638177	1	promoter	[b4400]	4638160
<b>6076</b>	b4400	4638178	4639530	1	gene	NaN	4638178
<b>6077</b>	b4401	4639590	4640306	-1	gene	NaN	4639590
<b>6078</b>	PM205	4640307	4640801	-1	promoter	[b4401]	4640307
<b>6079</b>	b4402	4640402	4640542	1	gene	NaN	4640801
<b>6080</b>	b4403	4640942	4641628	1	gene	NaN	4640942

6082 rows × 16 columns



```
In [55]: ecoli['cumulativeMass'] = np.insert(np.cumsum(ecoli[bw_glucose[0]]), 0,
0)
```

```
-----
-----
ValueError                                Traceback (most recent call l
ast)
<ipython-input-55-46adffff166bb> in <module>()
----> 1 ecoli['cumulativeMass'] = np.insert(np.cumsum(ecoli[bw_glucose[
0]]), 0, 0)

~/pyenv/versions/anaconda3-4.3.1/lib/python3.6/site-packages/numpy/li
b/function_base.py in insert(arr, obj, values, axis)
    4439         new[tuple(slobj)] = arr[tuple(slobj2)]
    4440         if wrap:
-> 4441             return wrap(new)
    4442         return new
    4443     elif indices.size == 0 and not isinstance(obj, np.ndarray):

~/pyenv/versions/anaconda3-4.3.1/lib/python3.6/site-packages/pandas/co
re/series.py in __array_wrap__(self, result, context)
    646         """
    647         return self._constructor(result, index=self.index,
--> 648                                 copy=False).__finalize__(self)
    649
    650     def __array_prepare__(self, result, context=None):

~/pyenv/versions/anaconda3-4.3.1/lib/python3.6/site-packages/pandas/co
re/series.py in __init__(self, data, index, dtype, name, copy, fastpat
h)
    260                                     'Length of passed values is {val},
    ,
    261                                     'index implies {ind}'
--> 262                                     .format(val=len(data), ind=len(ind
x)))
    263         except TypeError:
    264             pass
```

ValueError: Length of passed values is 6083, index implies 6082

## Mass per protein per cell calculated from Schmidt et al Table S9

[Mass of protein per cell][protein Molecular weight (kDa)][Copies/Cell]

```
In [56]: proteins = pd.read_excel('data/Ecoli/Schmidt/nbt.3418-S2.xlsx', sheet_na
me='Table S9', skiprows=2)
mg1655 = proteins[['Uniprot Accession', 'Description', 'Gene', 'proteinM
W', 'Copies/Cell_MG1655.Glucose' ]]
```

```
In [57]: mg1655 = proteins[['Uniprot Accession',
                           'Description',
                           'Gene',
                           'proteinMW',
                           'Copies/Cell_MG1655.Glucose' ]].\
                           iloc[:2019].\
                           set_index('Uniprot Accession')
mg1655['kDa/Protein'] = mg1655['proteinMW'] * mg1655['Copies/Cell_MG165
5.Glucose']
mg1655.sort_values('kDa/Protein', ascending=False)
```

Out[57]:

	Description	Gene	proteinMW	Copies/Cell_MG1655
Uniprot Accession				
<b>P0CE47</b>	Elongation factor Tu 1 OS=Escherichia coli (st...	tufA	43238.30279	279772.591884
<b>P0A910</b>	Outer membrane protein A OS=Escherichia coli (...)	ompA	37159.64679	108747.026756
<b>P25665</b>	5-methyltetrahydropteroyltriglutamate-homocys...	metE	84602.80905	37939.368305
<b>P0A6M8</b>	Elongation factor G OS=Escherichia coli (strai...	fusA	77514.47523	40158.649597
<b>P0ABK5</b>	Cysteine synthase A OS=Escherichia coli (strai...	cysK	34450.28634	65474.668058
<b>P05793</b>	Ketol-acid reductoisomerase OS=Escherichia col...	ilvC	54016.26705	34730.063498
<b>P04949</b>	Flagellin OS=Escherichia coli (strain K12) GN=...	fliC	51246.56095	34899.631529
<b>P08200</b>	Isocitrate dehydrogenase [NADP] OS=Escherichia...	icd	45709.46787	37682.155126
<b>P0A6P9</b>	Enolase OS=Escherichia coli (strain K12) GN=en...	eno	45608.41069	35656.217111
<b>P0ACF0</b>	DNA-binding protein HU-alpha OS=Escherichia co...	hupA	9511.17791	167069.666730
<b>P0A9B2</b>	Glyceraldehyde-3-phosphate dehydrogenase A OS=...	gapA	35492.23090	43004.950198
<b>P0AG67</b>	30S ribosomal protein S1 OS=Escherichia coli (...)	rpsA	61102.70430	24938.169144
<b>P61889</b>	Malate dehydrogenase OS=Escherichia coli (stra...	mdh	32299.20422	45044.312719
<b>P02359</b>	30S ribosomal protein S7 OS=Escherichia coli (...)	rpsG	19988.68708	65096.779696
<b>P0A6F5</b>	60 kDa chaperonin OS=Escherichia coli (strain ...)	groL	57274.67483	22304.785259
<b>P23843</b>	Periplasmic oligopeptide-binding protein OS=Es...	oppA	60842.91364	20761.120549
<b>P0A6Y8</b>	Chaperone protein dnaK OS=Escherichia coli (st...	dnaK	69054.47444	18132.697884

	Description	Gene	proteinMW	Copies/Cell_MG1655
Uniprot Accession				
<b>P0A799</b>	Phosphoglycerate kinase OS=Escherichia coli (s...	pgk	41074.62371	30363.426050
<b>P0A850</b>	Trigger factor OS=Escherichia coli (strain K12...	tig	48144.85133	25146.877882
<b>P0ABB4</b>	ATP synthase subunit beta OS=Escherichia coli ...	atpD	50275.81931	21539.317195
<b>P0AE08</b>	Alkyl hydroperoxide reductase subunit C OS=Esc...	ahpC	20730.40145	51209.990558
<b>P02931</b>	Outer membrane protein F OS=Escherichia coli (...)	ompF	39291.06024	25214.322876
<b>P0A825</b>	Serine hydroxymethyltransferase OS=Escherichia...	glyA	45269.96251	21854.316444
<b>P0AD96</b>	Leu/Ile/Val-binding protein OS=Escherichia col...	livJ	39033.98190	25126.451889
<b>P36683</b>	Aconitate hydratase 2 OS=Escherichia coli (str...	acnB	93420.94570	9625.639731
<b>P0A7V0</b>	30S ribosomal protein S2 OS=Escherichia coli (...)	rpsB	26708.79452	33317.431924
<b>P0A9C5</b>	Glutamine synthetase OS=Escherichia coli (stra...	glnA	51852.71686	17062.281156
<b>P28635</b>	D-methionine-binding lipoprotein metQ OS=Esche...	metQ	29395.54573	29248.962928
<b>P06996</b>	Outer membrane protein C OS=Escherichia coli (...)	ompC	40325.36434	20763.206184
<b>P0A7L0</b>	50S ribosomal protein L1 OS=Escherichia coli (...)	rplA	24696.28600	31336.657644
...	...	...	...	...
<b>P64585</b>	Inner membrane protein yqjE OS=Escherichia col...	yqjE	15119.20843	0.432701
<b>P0AFH0</b>	Methylated-DNA--protein-cysteine methyltransfe...	ogt	19148.91391	0.268607
<b>P52644</b>	Heat shock protein hslJ OS=Escherichia coli (s...	hslJ	15137.58295	0.268141
<b>P0AEP9</b>	Glycolate oxidase subunit glcD OS=Escherichia ...	glcD	53759.59300	0.074524

	Description	Gene	proteinMW	Copies/Cell_MG1655
Uniprot Accession				
<b>P37047</b>	Carbohydrate diacid regulator OS=Escherichia c...	cdaR	43641.75186	0.074410
<b>P0ACL7</b>	Putative L-lactate dehydrogenase operon regula...	lldR	29129.87064	0.076603
<b>P63340</b>	Inner membrane transport protein YqeG OS=Esche...	yqeG	45051.36046	0.006444
<b>P08555</b>	DsdX permease OS=Escherichia coli (strain K12)...	dsdX	47113.73234	0.005528
<b>P0ABJ3</b>	Cytochrome o ubiquinol oxidase subunit 3 OS=Es...	cyoC	22589.45441	0.011453
<b>P09129</b>	Protein traS OS=Escherichia coli (strain K12) ...	traS	19486.78142	0.012894
<b>P52614</b>	Flagellar hook-length control protein OS=Esche...	fliK	39270.01569	0.006011
<b>P24177</b>	Probable aminoglycoside efflux pump OS=Escheri...	acrD	112957.15190	0.001755
<b>P62672</b>	Protein ApaG OS=Escherichia coli (strain K12) ...	apaG	13840.01621	0.013194
<b>P76077</b>	Phenylacetic acid degradation protein paaA OS=...	paaA	35458.53694	0.005077
<b>P77148</b>	Uncharacterized protein ydhS OS=Escherichia co...	ydhS	60698.94360	0.002923
<b>Q46802</b>	Uncharacterized sigma-54- dependent transcripti...	ygeV	66036.49819	0.002557
<b>P37339</b>	L-2-hydroxyglutarate oxidase LhgO OS=Escherich...	lhgO	46034.50872	0.003448
<b>P0AF26</b>	Nitrate reductase molybdenum cofactor assembly...	narJ	26414.47486	0.005945
<b>P13959</b>	Replication regulatory protein RepA2 OS=Escher...	repA	9580.07437	0.016085
<b>P06615</b>	Resolvase OS=Escherichia coli (strain K12) GN=...	resD	29590.47830	0.005157
<b>P76081</b>	Probable phenylacetic acid degradation NADH ox...	paaE	39276.80969	0.003585
<b>P52073</b>	Glycolate oxidase subunit glcE OS=Escherichia ...	glcE	38318.46988	0.003664

	Description	Gene	proteinMW	Copies/Cell_MG1655
Uniprot Accession				
<b>Q00042</b>	Putative transposon gamma-delta 80.3 kDa prote...	tnpX	80261.47359	0.001727
<b>P0AGL2</b>	Protein tdcF OS=Escherichia coli (strain K12) ...	tdcF	13980.29107	0.009733
<b>P0AEL3</b>	Ferrous iron transport protein A OS=Escherichi...	feoA	8347.49547	0.016024
<b>P03856</b>	Replication initiation protein OS=Escherichia ...	repE	29322.03016	0.004496
<b>P0AFI5</b>	D-alanyl-D-alanine endopeptidase OS=Escherichi...	pbpG	33847.68281	0.003756
<b>P64599</b>	Uncharacterized protein yhbT OS=Escherichia co...	yhbT	19641.28483	0.006418
<b>P76078</b>	Phenylacetic acid degradation protein paaB OS=...	paaB	10917.31304	0.011180
<b>P0A772</b>	Protein nrdI OS=Escherichia coli (strain K12) ...	nrdI	15312.88748	0.007231

2019 rows × 5 columns

## Calculate Cumulative abundance

Note that we do not calculate min copy number and multiply by molecular weight, we just fill genes by the min g/Protein. The next version will include molecular weights for all proteins so we can do it by min copy number instead.

In [58]: df

Out[58]:

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
0	PM00249	148	189	1	promoter	[b0001, b0002, b0003, b0004]	148
1	b0001	190	255	1	gene	NaN	190
2	b0002	337	2799	1	gene	NaN	337
3	b0003	2801	3733	1	gene	NaN	2801
4	b0004	3734	5020	1	gene	NaN	3734
5	b0005	5234	5530	1	gene	NaN	5234
6	b0006	5683	6459	-1	gene	NaN	5683
7	b0007	6529	7959	-1	gene	NaN	6529
8	PM0-9956	8191	8237	1	promoter	[b0008]	8191
9	b0008	8238	9191	1	gene	NaN	8238
10	b0009	9306	9893	1	gene	NaN	9306
11	b0010	9928	10494	-1	gene	NaN	9928
12	b0011	10643	11356	-1	gene	NaN	10643
13	b0013	11382	11786	-1	gene	NaN	11382
14	PM00316	12048	12162	1	promoter	[b0014, b0015]	12048
15	b0014	12163	14079	1	gene	NaN	12163
16	b0015	14168	15298	1	gene	NaN	14168
17	b0016	15445	16557	1	gene	NaN	15445
18	b0018	16751	16960	-1	gene	NaN	16751
19	b4412	16751	16903	-1	gene	NaN	16960
20	b4413	16952	17006	1	gene	NaN	16960
21	b0019	17489	18655	1	gene	NaN	17489
22	b0020	18715	19620	1	gene	NaN	18715
23	b0021	19811	20314	-1	gene	NaN	19811
24	b0022	20233	20508	-1	gene	NaN	20314
25	b0023	20815	21078	-1	gene	NaN	20815
26	PM0-8222	21079	21210	-1	promoter	[b0023]	21079
27	b0024	21181	21399	1	gene	NaN	21210



	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>28</b>	PM0-8866	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399
<b>29</b>	b0025	21407	22348	1	gene	NaN	21407
...	...	...	...	...	...	...	...
<b>6051</b>	b4383	4619603	4620826	1	gene	NaN	4619603
<b>6052</b>	b4384	4620883	4621602	1	gene	NaN	4620883
<b>6053</b>	b4385	4621769	4623100	1	gene	NaN	4621769
<b>6054</b>	b4386	4623101	4624117	-1	gene	NaN	4623101
<b>6055</b>	b4387	4624145	4624789	-1	gene	NaN	4624145
<b>6056</b>	PM00508	4624790	4624799	-1	promoter	[b4386, b4387]	4624790
<b>6057</b>	PM00509	4624856	4624894	1	promoter	[b4389, b4388]	4624856
<b>6058</b>	b4388	4624895	4625863	1	gene	NaN	4624895
<b>6059</b>	b4389	4625912	4627294	1	gene	NaN	4625912
<b>6060</b>	b4390	4627315	4628547	1	gene	NaN	4627315
<b>6061</b>	b4391	4628855	4630522	-1	gene	NaN	4628855
<b>6062</b>	PM0-46112	4630700	4630732	1	promoter	[b4392]	4630700
<b>6063</b>	b4392	4630733	4632670	1	gene	NaN	4630733
<b>6064</b>	PM00095	4632704	4632759	1	promoter	[b4393]	4632704
<b>6065</b>	b4393	4632760	4633086	1	gene	NaN	4632760
<b>6066</b>	b4394	4633233	4633745	-1	gene	NaN	4633233
<b>6067</b>	b4395	4633797	4634444	1	gene	NaN	4633797
<b>6068</b>	b4396	4634441	4635310	-1	gene	NaN	4634444
<b>6069</b>	PM0-9728	4635243	4635520	1	promoter	[b4399, b4398, b4397]	4635310
<b>6070</b>	PM0-8184	4635311	4635353	-1	promoter	[b4396]	4635520
<b>6071</b>	PM00603	4635477	4635520	1	promoter	[b4397, b4398, b4399, b4400]	4635520
<b>6072</b>	b4397	4635521	4635994	1	gene	NaN	4635521

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>6073</b>	b4398	4636007	4636696	1	gene	NaN	4636007
<b>6074</b>	b4399	4636696	4638120	1	gene	NaN	4636696
<b>6075</b>	PM0-10215	4638160	4638177	1	promoter	[b4400]	4638160
<b>6076</b>	b4400	4638178	4639530	1	gene	NaN	4638178
<b>6077</b>	b4401	4639590	4640306	-1	gene	NaN	4639590
<b>6078</b>	PM205	4640307	4640801	-1	promoter	[b4401]	4640307
<b>6079</b>	b4402	4640402	4640542	1	gene	NaN	4640801
<b>6080</b>	b4403	4640942	4641628	1	gene	NaN	4640942

6081 rows × 8 columns

```
In [59]: import numpy as np
b2u = pd.read_table('data/Ecoli/blatter-to-uniprot.tab', index_col = 'Bl
attner')
ecoli = df.join(b2u, on='gene_or_promoter').join(BW25113_fg, on='Unipro
t')
ecoli
```

Out[59]:

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
0	PM00249	148	189	1	promoter	[b0001, b0002, b0003, b0004]	148
1	b0001	190	255	1	gene	NaN	190
2	b0002	337	2799	1	gene	NaN	337
3	b0003	2801	3733	1	gene	NaN	2801
4	b0004	3734	5020	1	gene	NaN	3734
5	b0005	5234	5530	1	gene	NaN	5234
6	b0006	5683	6459	-1	gene	NaN	5683
7	b0007	6529	7959	-1	gene	NaN	6529
8	PM0-9956	8191	8237	1	promoter	[b0008]	8191
9	b0008	8238	9191	1	gene	NaN	8238
10	b0009	9306	9893	1	gene	NaN	9306
11	b0010	9928	10494	-1	gene	NaN	9928
12	b0011	10643	11356	-1	gene	NaN	10643
13	b0013	11382	11786	-1	gene	NaN	11382
14	PM00316	12048	12162	1	promoter	[b0014, b0015]	12048
15	b0014	12163	14079	1	gene	NaN	12163
16	b0015	14168	15298	1	gene	NaN	14168
17	b0016	15445	16557	1	gene	NaN	15445
18	b0018	16751	16960	-1	gene	NaN	16751
19	b4412	16751	16903	-1	gene	NaN	16960
20	b4413	16952	17006	1	gene	NaN	16960

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>21</b>	b0019	17489	18655	1	gene	NaN	17489
<b>22</b>	b0020	18715	19620	1	gene	NaN	18715
<b>23</b>	b0021	19811	20314	-1	gene	NaN	19811
<b>24</b>	b0022	20233	20508	-1	gene	NaN	20314
<b>25</b>	b0023	20815	21078	-1	gene	NaN	20815
<b>26</b>	PM0-8222	21079	21210	-1	promoter	[b0023]	21079
<b>27</b>	b0024	21181	21399	1	gene	NaN	21210
<b>28</b>	PM0-8866	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399
<b>29</b>	b0025	21407	22348	1	gene	NaN	21407
...	...	...	...	...	...	...	...
<b>6051</b>	b4383	4619603	4620826	1	gene	NaN	4619603
<b>6052</b>	b4384	4620883	4621602	1	gene	NaN	4620883
<b>6053</b>	b4385	4621769	4623100	1	gene	NaN	4621769
<b>6054</b>	b4386	4623101	4624117	-1	gene	NaN	4623101
<b>6055</b>	b4387	4624145	4624789	-1	gene	NaN	4624145
<b>6056</b>	PM00508	4624790	4624799	-1	promoter	[b4386, b4387]	4624790
<b>6057</b>	PM00509	4624856	4624894	1	promoter	[b4389, b4388]	4624856
<b>6058</b>	b4388	4624895	4625863	1	gene	NaN	4624895
<b>6059</b>	b4389	4625912	4627294	1	gene	NaN	4625912

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>6060</b>	b4390	4627315	4628547	1	gene	NaN	4627315
<b>6061</b>	b4391	4628855	4630522	-1	gene	NaN	4628855
<b>6062</b>	PM0-46112	4630700	4630732	1	promoter	[b4392]	4630700
<b>6063</b>	b4392	4630733	4632670	1	gene	NaN	4630733
<b>6064</b>	PM00095	4632704	4632759	1	promoter	[b4393]	4632704
<b>6065</b>	b4393	4632760	4633086	1	gene	NaN	4632760
<b>6066</b>	b4394	4633233	4633745	-1	gene	NaN	4633233
<b>6067</b>	b4395	4633797	4634444	1	gene	NaN	4633797
<b>6068</b>	b4396	4634441	4635310	-1	gene	NaN	4634444
<b>6069</b>	PM0-9728	4635243	4635520	1	promoter	[b4399, b4398, b4397]	4635310
<b>6070</b>	PM0-8184	4635311	4635353	-1	promoter	[b4396]	4635520
<b>6071</b>	PM00603	4635477	4635520	1	promoter	[b4397, b4398, b4399, b4400]	4635520
<b>6072</b>	b4397	4635521	4635994	1	gene	NaN	4635521
<b>6073</b>	b4398	4636007	4636696	1	gene	NaN	4636007
<b>6074</b>	b4399	4636696	4638120	1	gene	NaN	4636696
<b>6075</b>	PM0-10215	4638160	4638177	1	promoter	[b4400]	4638160
<b>6076</b>	b4400	4638178	4639530	1	gene	NaN	4638178
<b>6077</b>	b4401	4639590	4640306	-1	gene	NaN	4639590
<b>6078</b>	PM205	4640307	4640801	-1	promoter	[b4401]	4640307
<b>6079</b>	b4402	4640402	4640542	1	gene	NaN	4640801
<b>6080</b>	b4403	4640942	4641628	1	gene	NaN	4640942

6082 rows × 15 columns

```
In [62]: minAbundance = mg1655['kDa/Protein'].dropna().min()  
minAbundance
```

```
Out[62]: 110.73090803705458
```

```
In [63]: ecoli.loc[ecoli['class']== 'promoter','kDa/Protein'] = 10
ecoli['kDa/Protein'].fillna(minAbundance,inplace=True)
ecoli['cumulativeAbundance'] = np.cumsum(ecoli['kDa/Protein'])

ecoli.to_csv('data/Ecoli/{}_cumulative_abundance.tab'.format('kDa_per_protein_per_cell'), sep='\t', index=False)
ecoli
```



Out[63]:

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
0	PM00249	148	189	1	promoter	[b0001, b0002, b0003, b0004]	148
1	b0001	190	255	1	gene	NaN	190
2	b0002	337	2799	1	gene	NaN	337
3	b0003	2801	3733	1	gene	NaN	2801
4	b0004	3734	5020	1	gene	NaN	3734
5	b0005	5234	5530	1	gene	NaN	5234
6	b0006	5683	6459	-1	gene	NaN	5683
7	b0007	6529	7959	-1	gene	NaN	6529
8	PM0-9956	8191	8237	1	promoter	[b0008]	8191
9	b0008	8238	9191	1	gene	NaN	8238
10	b0009	9306	9893	1	gene	NaN	9306
11	b0010	9928	10494	-1	gene	NaN	9928
12	b0011	10643	11356	-1	gene	NaN	10643
13	b0013	11382	11786	-1	gene	NaN	11382
14	PM00316	12048	12162	1	promoter	[b0014, b0015]	12048
15	b0014	12163	14079	1	gene	NaN	12163
16	b0015	14168	15298	1	gene	NaN	14168
17	b0016	15445	16557	1	gene	NaN	15445
18	b0018	16751	16960	-1	gene	NaN	16751
19	b4412	16751	16903	-1	gene	NaN	16960
20	b4413	16952	17006	1	gene	NaN	16960

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>21</b>	b0019	17489	18655	1	gene	NaN	17489
<b>22</b>	b0020	18715	19620	1	gene	NaN	18715
<b>23</b>	b0021	19811	20314	-1	gene	NaN	19811
<b>24</b>	b0022	20233	20508	-1	gene	NaN	20314
<b>25</b>	b0023	20815	21078	-1	gene	NaN	20815
<b>26</b>	PM0-8222	21079	21210	-1	promoter	[b0023]	21079
<b>27</b>	b0024	21181	21399	1	gene	NaN	21210
<b>28</b>	PM0-8866	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399
<b>29</b>	b0025	21407	22348	1	gene	NaN	21407
...	...	...	...	...	...	...	...
<b>6051</b>	b4383	4619603	4620826	1	gene	NaN	4619603
<b>6052</b>	b4384	4620883	4621602	1	gene	NaN	4620883
<b>6053</b>	b4385	4621769	4623100	1	gene	NaN	4621769
<b>6054</b>	b4386	4623101	4624117	-1	gene	NaN	4623101
<b>6055</b>	b4387	4624145	4624789	-1	gene	NaN	4624145
<b>6056</b>	PM00508	4624790	4624799	-1	promoter	[b4386, b4387]	4624790
<b>6057</b>	PM00509	4624856	4624894	1	promoter	[b4389, b4388]	4624856
<b>6058</b>	b4388	4624895	4625863	1	gene	NaN	4624895
<b>6059</b>	b4389	4625912	4627294	1	gene	NaN	4625912

	gene_or_promoter	start	end	strand	class	genes_in_TU	start_if_select
<b>6060</b>	b4390	4627315	4628547	1	gene	NaN	4627315
<b>6061</b>	b4391	4628855	4630522	-1	gene	NaN	4628855
<b>6062</b>	PM0-46112	4630700	4630732	1	promoter	[b4392]	4630700
<b>6063</b>	b4392	4630733	4632670	1	gene	NaN	4630733
<b>6064</b>	PM00095	4632704	4632759	1	promoter	[b4393]	4632704
<b>6065</b>	b4393	4632760	4633086	1	gene	NaN	4632760
<b>6066</b>	b4394	4633233	4633745	-1	gene	NaN	4633233
<b>6067</b>	b4395	4633797	4634444	1	gene	NaN	4633797
<b>6068</b>	b4396	4634441	4635310	-1	gene	NaN	4634444
<b>6069</b>	PM0-9728	4635243	4635520	1	promoter	[b4399, b4398, b4397]	4635310
<b>6070</b>	PM0-8184	4635311	4635353	-1	promoter	[b4396]	4635520
<b>6071</b>	PM00603	4635477	4635520	1	promoter	[b4397, b4398, b4399, b4400]	4635520
<b>6072</b>	b4397	4635521	4635994	1	gene	NaN	4635521
<b>6073</b>	b4398	4636007	4636696	1	gene	NaN	4636007
<b>6074</b>	b4399	4636696	4638120	1	gene	NaN	4636696
<b>6075</b>	PM0-10215	4638160	4638177	1	promoter	[b4400]	4638160
<b>6076</b>	b4400	4638178	4639530	1	gene	NaN	4638178
<b>6077</b>	b4401	4639590	4640306	-1	gene	NaN	4639590
<b>6078</b>	PM205	4640307	4640801	-1	promoter	[b4401]	4640307
<b>6079</b>	b4402	4640402	4640542	1	gene	NaN	4640801
<b>6080</b>	b4403	4640942	4641628	1	gene	NaN	4640942

6082 rows × 17 columns

```
In [61]: import pandas
abundance_f = 'data/Ecoli/{_cumulative_mass.tab'.format(bw_glucose[0])
cum_abundance = pd.read_table(abundance_f, usecols=['gene_or_promoter',
'cumulativeMass']).set_index('gene_or_promoter')
cum_abundance.loc['b3322':'b3340']
```

Out[61]:

	<b>cumulativeMass</b>
<b>gene_or_promoter</b>	
<b>b3322</b>	121.055606
<b>b3323</b>	121.055606
<b>PM-8819</b>	121.055607
<b>PM0-9514</b>	121.055607
<b>b3324</b>	121.055607
<b>b3325</b>	121.055608
<b>b3326</b>	121.055608
<b>b3327</b>	121.055609
<b>b3328</b>	121.055609
<b>b3329</b>	121.055610
<b>b3330</b>	121.055610
<b>b3331</b>	121.055611
<b>b3332</b>	121.055611
<b>b3333</b>	121.055612
<b>b3334</b>	121.055612
<b>b3335</b>	121.055613
<b>b3336</b>	121.055613
<b>b3337</b>	121.055613
<b>b3338</b>	121.160264
<b>PM0-3321</b>	121.160265
<b>b3339</b>	121.160265
<b>PM0-9799</b>	121.160266
<b>b3340</b>	140.151514

```
In [115]: from IPython.display import display, Image, HTML
solution = pd.read_table('out/local_result_essential.tab',index_col=0)
solution['end'] = solution['end'].str.lstrip('u_G_')
solution['start'] = solution['start'].str.lstrip('u_G_')
display( solution)
```

	end	start	status
0	b3340	b3322	Optimal
1	b0971	b0955	Optimal
2	b1636	b4493	Optimal
3	PM00633	PM0-10091	Optimal
4	b3559	b3465	Optimal
5	PM0-5862	PM00001	Optimal
6	b2779	b2765	Optimal
7	b0452	b0433	Optimal
8	PM324	b1289	Optimal
9	b2231	PM00652	Optimal
10	b0628	PM0-7163	Optimal

```
In [25]: from IPython.display import display, Image, HTML
solution = pd.read_table('out/local_result_essential.tab')
solution['end'] = solution['end'].str.lstrip('u_G_')
solution['start'] = solution['start'].str.lstrip('u_G_')
display( solution)
```

	end	start	status
0	b3340	b3322	Optimal
1	b0971	b0955	Optimal
2	PM0-7141	b1913	Optimal
3	b3833	PM00100	Optimal
4	PM00633	PM0-10091	Optimal
5	b0452	b0433	Optimal
6	b3559	b3465	Optimal
7	b1636	b4493	Optimal
8	b4005	b3998	Optimal
9	b2779	b2765	Optimal
10	PM0-8866	b0010	Optimal

In [ ]:

```
In [26]: from IPython.display import Markdown
genome_view = []
mass_reclaimed = []
for i in solution.index:
    start, end = solution.loc[i, 'start'], solution.loc[i, 'end']
    gv = ecoli.set_index('gene_or_promoter').loc[start:end].copy()
    mass_reclaimed.append(gv.loc[end, 'cumulativeMass'] - gv.loc[start,
'cumulativeMass'])
    gv[r'mass reclaimed $(fg/cell)$'] = mass_reclaimed[i]
    gv['SolutionOrder'] = len(solution.index) - i
    genome_view.append( gv )
genome_view = pd.concat(genome_view)
genome_view.index.name = '$gene_or_promoter'
genome_view.to_csv('out/genome_view.tab', sep='\t', columns=['SolutionOrder',
r'mass reclaimed $(fg/cell)$', bw_glucose[0], 'cumulativeMass'])
print('\n'.join([str(m) for m in mass_reclaimed]))
display(Markdown(r'Total mass reclaimed = %0.2g $fg/cell$' % sum(mass_reclaimed)))
HTML(genome_view.to_html())
```

24.1442705355  
6.88266638029  
0.910191488255  
0.0949279533945  
4.19366640413  
2.83457780068  
2.35459208769  
1.42959234843  
1.19512854724e-05  
2.9846500906  
2.63181406031

Total mass reclaimed = 48 fg/cell



Out[26]:

	start	end	strand	class	genes_in_TU	start_if_select_as_
<b>\$gene_or_promoter</b>						
<b>b3322</b>	3453508	3453927	-1	gene	NaN	3453508
<b>b3323</b>	3453929	3455398	-1	gene	NaN	3453929
<b>PM-8819</b>	3455399	3455634	-1	promoter	[b3322, b3323]	3455399
<b>PM0-9514</b>	3455482	3455577	1	promoter	[b3335, b3334, b3333, b3332, b3331, b3330, b33...	3455634
<b>b3324</b>	3455578	3456393	1	gene	NaN	3455634
<b>b3325</b>	3456377	3458329	1	gene	NaN	3456393
<b>b3326</b>	3458339	3459820	1	gene	NaN	3458339
<b>b3327</b>	3459817	3461013	1	gene	NaN	3459820
<b>b3328</b>	3461023	3461460	1	gene	NaN	3461023
<b>b3329</b>	3461468	3461977	1	gene	NaN	3461468
<b>b3330</b>	3461974	3462351	1	gene	NaN	3461977
<b>b3331</b>	3462344	3462931	1	gene	NaN	3462351
<b>b3332</b>	3462924	3463907	1	gene	NaN	3462931
<b>b3333</b>	3463922	3465085	1	gene	NaN	3463922
<b>b3334</b>	3465082	3465543	1	gene	NaN	3465085
<b>b3335</b>	3465543	3466220	1	gene	NaN	3465543
<b>b3336</b>	3466249	3466725	-1	gene	NaN	3466249
<b>b3337</b>	3466797	3466991	-1	gene	NaN	3466797
<b>b3338</b>	3467160	3469853	-1	gene	NaN	3467160
<b>PM0-3321</b>	3469854	3469896	-1	promoter	[b3338]	3469854
<b>b3339</b>	3470145	3471329	-1	gene	NaN	3470145
<b>PM0-9799</b>	3471330	3471724	-1	promoter	[b3339]	3471330

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3340</b>	3471400	3473514	-1	gene	NaN	3471724
<b>b0955</b>	1016539	1018299	-1	gene	NaN	1016539
<b>PM0-10011</b>	1018459	1018484	1	promoter	[b0956]	1018459
<b>b0956</b>	1018485	1018937	1	gene	NaN	1018485
<b>b0957</b>	1019013	1020053	-1	gene	NaN	1019013
<b>PM00064</b>	1020054	1020191	-1	promoter	[b0957]	1020054
<b>b0958</b>	1020410	1020919	-1	gene	NaN	1020410
<b>PM00092</b>	1020920	1020947	-1	promoter	[b0958]	1020920
<b>PM0-46336</b>	1021083	1021137	1	promoter	[b0959]	1021083
<b>b0959</b>	1021138	1021767	1	gene	NaN	1021138
<b>b0960</b>	1021730	1023883	-1	gene	NaN	1021767
<b>b0961</b>	1023902	1024348	-1	gene	NaN	1023902
<b>PM0-36232</b>	1024349	1024365	-1	promoter	[b0961, b0960]	1024349
<b>PM0-36233</b>	1024438	1024470	1	promoter	[b0962]	1024438
<b>b0962</b>	1024471	1026525	1	gene	NaN	1024471
<b>b0963</b>	1026557	1027015	-1	gene	NaN	1026557
<b>PM0-10012</b>	1027016	1027041	-1	promoter	[b0963]	1027016
<b>b0964</b>	1027111	1027773	-1	gene	NaN	1027111
<b>PM0-36218</b>	1027774	1027837	-1	promoter	[b0964]	1027774
<b>b0965</b>	1027946	1028359	1	gene	NaN	1027946
<b>b0966</b>	1028404	1028721	-1	gene	NaN	1028404
<b>b0967</b>	1028779	1029969	-1	gene	NaN	1028779
<b>b0968</b>	1030064	1030342	1	gene	NaN	1030064
<b>b0969</b>	1030339	1030668	-1	gene	NaN	1030342

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b0970</b>	1030759	1031418	-1	gene	NaN	1030759
<b>PM0-9395</b>	1031419	1031443	-1	promoter	[b0970]	1031419
<b>b0971</b>	1031625	1031712	-1	gene	NaN	1031625
<b>b1913</b>	1992874	1994706	-1	gene	NaN	1992874
<b>b1914</b>	1994703	1995359	-1	gene	NaN	1994706
<b>PM00388</b>	1994707	1994879	-1	promoter	[b1913]	1995359
<b>PM00389</b>	1995360	1995406	-1	promoter	[b1913, b1914]	1995360
<b>PM0-9377</b>	1995717	1995817	1	promoter	[b1915]	1995717
<b>b1915</b>	1995818	1996042	1	gene	NaN	1995818
<b>b1916</b>	1996110	1996832	-1	gene	NaN	1996110
<b>PM0-46164</b>	1996833	1996917	-1	promoter	[b1916]	1996833
<b>PM0-46369</b>	1996880	1996920	1	promoter	[sdsN]	1996917
<b>sdsN</b>	1996921	1997057	1	gene	NaN	1996921
<b>b1917</b>	1997062	1997814	-1	gene	NaN	1997062
<b>b1918</b>	1997811	1998479	-1	gene	NaN	1997814
<b>b1919</b>	1998494	1999480	-1	gene	NaN	1998494
<b>PM0-10069</b>	1999481	1999508	-1	promoter	[b1919, b1918, b1917]	1999481
<b>b1920</b>	1999585	2000385	-1	gene	NaN	1999585
<b>PM0-9795</b>	2000386	2000410	-1	promoter	[b1920]	2000386
<b>b1921</b>	2000473	2001024	-1	gene	NaN	2000473
<b>b1922</b>	2001070	2001789	-1	gene	NaN	2001070

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM00563</b>	2001790	2001819	-1	promoter	[b1920, b1921, b1922]	2001790
<b>b1923</b>	2002110	2003606	-1	gene	NaN	2002110
<b>PM0-1681</b>	2003607	2003676	-1	promoter	[b1923]	2003607
<b>PM00604</b>	2003817	2003871	1	promoter	[b1924, b1925, b1926]	2003817
<b>b1924</b>	2003872	2005278	1	gene	NaN	2003872
<b>b1925</b>	2005303	2005713	1	gene	NaN	2005303
<b>b1926</b>	2005713	2006078	1	gene	NaN	2005713
<b>PM0-7803</b>	2006115	2006155	1	promoter	[b1927]	2006115
<b>b1927</b>	2006156	2007643	1	gene	NaN	2006156
<b>b1928</b>	2007677	2008090	-1	gene	NaN	2007677
<b>b1929</b>	2008277	2009482	1	gene	NaN	2008277
<b>b1930</b>	2009479	2009712	1	gene	NaN	2009482
<b>b1931</b>	2009821	2010489	1	gene	NaN	2009821
<b>b1932</b>	2010600	2011079	1	gene	NaN	2010600
<b>b4495</b>	2011223	2012351	-1	gene	NaN	2011223
<b>b1936</b>	2012502	2012663	1	gene	NaN	2012502
<b>b1937</b>	2012700	2013014	-1	gene	NaN	2012700
<b>PM0-8720</b>	2013015	2013093	-1	promoter	[b1937]	2013015
<b>PM-8797</b>	2013146	2013228	1	promoter	[b1943, b1942, b1941, b1940, b1939, b1938]	2013146

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM-8793</b>	2013205	2013228	1	promoter	[b1938, b1939, b1940, b1941, b1942, b1943]	2013228
<b>b1938</b>	2013229	2014887	1	gene	NaN	2013229
<b>b1939</b>	2014880	2015875	1	gene	NaN	2014887
<b>b1940</b>	2015868	2016554	1	gene	NaN	2015875
<b>b1941</b>	2016554	2017927	1	gene	NaN	2016554
<b>b1942</b>	2017946	2018389	1	gene	NaN	2017946
<b>b1943</b>	2018386	2019513	1	gene	NaN	2018389
<b>PM164</b>	2019584	2019617	1	promoter	[b1944, b1945, b1946, b1947, b1948, b1949, b1950]	2019584
<b>PM165</b>	2019595	2019617	1	promoter	[b1950, b1949, b1948, b1947, b1946, b1945, b1944]	2019617
<b>b1944</b>	2019618	2020082	1	gene	NaN	2019618
<b>b1945</b>	2020087	2021091	1	gene	NaN	2020087
<b>b1946</b>	2021088	2021501	1	gene	NaN	2021091
<b>b1947</b>	2021504	2021869	1	gene	NaN	2021504
<b>b1948</b>	2021869	2022606	1	gene	NaN	2021869
<b>b1949</b>	2022616	2022885	1	gene	NaN	2022616

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1950</b>	2022893	2023678	1	gene	NaN	2022893
<b>PM00636</b>	2023836	2023967	1	promoter	[b1951]	2023836
<b>b1951</b>	2023968	2024591	1	gene	NaN	2023968
<b>b1952</b>	2024635	2024823	-1	gene	NaN	2024635
<b>b1953</b>	2024986	2025213	1	gene	NaN	2024986
<b>b1954</b>	2025227	2025313	-1	gene	NaN	2025227
<b>PM0-10501</b>	2025314	2025313	-1	promoter	[b1954]	2025314
<b>b1955</b>	2025511	2026326	1	gene	NaN	2025511
<b>b1956</b>	2026323	2028017	-1	gene	NaN	2026326
<b>b1957</b>	2028188	2028370	-1	gene	NaN	2028188
<b>b1958</b>	2028449	2029366	-1	gene	NaN	2028449
<b>b1959</b>	2029539	2030459	1	gene	NaN	2029539
<b>b1960</b>	2030448	2030918	-1	gene	NaN	2030459
<b>b1961</b>	2030899	2032317	-1	gene	NaN	2030918
<b>b1962</b>	2032384	2033079	-1	gene	NaN	2032384
<b>b1963</b>	2033119	2033484	-1	gene	NaN	2033119
<b>PM0-9468</b>	2033649	2033648	1	promoter	[b4603]	2033649
<b>b4603</b>	2033649	2033739	1	gene	NaN	2033649
<b>b4496</b>	2034051	2035243	1	gene	NaN	2034051
<b>PM0-9580</b>	2035630	2035834	1	promoter	[b1967]	2035630
<b>b1967</b>	2035835	2036686	1	gene	NaN	2035835
<b>b1968</b>	2036794	2038152	-1	gene	NaN	2036794
<b>b1969</b>	2038152	2038823	-1	gene	NaN	2038152
<b>PM0-45171</b>	2038840	2038955	1	promoter	[b1970]	2038840

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1970</b>	2038956	2039369	1	gene	NaN	2038956
<b>b1971</b>	2039478	2040482	1	gene	NaN	2039478
<b>b1972</b>	2040483	2041118	1	gene	NaN	2040483
<b>PM0-8710</b>	2041346	2041374	1	promoter	[b1973]	2041346
<b>b1973</b>	2041375	2042025	1	gene	NaN	2041375
<b>b1974</b>	2042368	2042898	1	gene	NaN	2042368
<b>b1975</b>	2043468	2043557	-1	gene	NaN	2043468
<b>PM00668</b>	2043558	2043562	-1	promoter	[b1975]	2043558
<b>b1976</b>	2043651	2044448	1	gene	NaN	2043651
<b>PM00649</b>	2044541	2044548	1	promoter	[b1977]	2044541
<b>b1977</b>	2044549	2044624	1	gene	NaN	2044549
<b>b1978</b>	2044938	2052014	1	gene	NaN	2044938
<b>b4497</b>	2052276	2053328	-1	gene	NaN	2052276
<b>PM0-46175</b>	2053566	2053642	1	promoter	[b1981]	2053566
<b>b1981</b>	2053643	2054959	1	gene	NaN	2053643
<b>PM00391</b>	2055031	2055060	1	promoter	[b1982]	2055031
<b>b1982</b>	2055061	2056515	1	gene	NaN	2055061
<b>b1983</b>	2056858	2057574	1	gene	NaN	2056858
<b>b1984</b>	2058027	2058102	-1	gene	NaN	2058027
<b>b1985</b>	2058203	2059690	-1	gene	NaN	2058203
<b>PM00650</b>	2059845	2059850	1	promoter	[b1986]	2059845
<b>b1986</b>	2059851	2059926	1	gene	NaN	2059851
<b>b1987</b>	2059964	2060914	-1	gene	NaN	2059964

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM326</b>	2060915	2060941	-1	promoter	[b1987]	2060915
<b>b1988</b>	2061016	2061933	-1	gene	NaN	2061016
<b>PM65</b>	2061934	2061978	-1	promoter	[b1988]	2061934
<b>PM00651</b>	2062252	2062259	1	promoter	[b1989]	2062252
<b>b1989</b>	2062260	2062335	1	gene	NaN	2062260
<b>b1990</b>	2062391	2063323	-1	gene	NaN	2062391
<b>b1991</b>	2063388	2064467	-1	gene	NaN	2063388
<b>b1992</b>	2064479	2065222	-1	gene	NaN	2064479
<b>b1993</b>	2065219	2065764	-1	gene	NaN	2065222
<b>PM0-8751</b>	2065765	2065972	-1	promoter	[b1991, b1992, b1993]	2065765
<b>b4639</b>	2066068	2066154	-1	gene	NaN	2066068
<b>b1994</b>	2066305	2067321	-1	gene	NaN	2066305
<b>b4640</b>	2067360	2067892	-1	gene	NaN	2067360
<b>b4641</b>	2068261	2068419	-1	gene	NaN	2068261
<b>b4582</b>	2068635	2070474	1	gene	NaN	2068635
<b>b1996</b>	2068952	2069857	-1	gene	NaN	2070474
<b>b1997</b>	2069815	2070180	-1	gene	NaN	2070474
<b>b1999</b>	2070660	2071211	1	gene	NaN	2070660
<b>PM0-2881</b>	2071317	2071316	1	promoter	[b4435, b2000]	2071317
<b>b4435</b>	2071317	2071511	1	gene	NaN	2071317
<b>PM0-10071</b>	2071474	2071538	1	promoter	[b2000]	2071511
<b>b2000</b>	2071539	2074658	1	gene	NaN	2071539
<b>b2001</b>	2074779	2076311	1	gene	NaN	2074779
<b>b2002</b>	2076308	2076754	1	gene	NaN	2076311
<b>b2003</b>	2076817	2077038	1	gene	NaN	2076817



	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b2004</b>	2077112	2077480	1	gene	NaN	2077112
<b>b2005</b>	2077569	2077943	1	gene	NaN	2077569
<b>b2006</b>	2077940	2078107	1	gene	NaN	2077943
<b>b4642</b>	2078549	2078677	-1	gene	NaN	2078549
<b>b4538</b>	2078746	2078931	1	gene	NaN	2078746
<b>b2007</b>	2079032	2079361	-1	gene	NaN	2079032
<b>b2008</b>	2079533	2080591	-1	gene	NaN	2079533
<b>b2009</b>	2080789	2081262	-1	gene	NaN	2080789
<b>PM0-10073</b>	2081263	2081325	-1	promoter	[b2009]	2081263
<b>b2010</b>	2081381	2082547	-1	gene	NaN	2081381
<b>PM00396</b>	2082728	2082755	1	promoter	[b2011]	2082728
<b>b2011</b>	2082756	2084183	1	gene	NaN	2082756
<b>b2012</b>	2084226	2084453	-1	gene	NaN	2084226
<b>b2013</b>	2084467	2085525	-1	gene	NaN	2084467
<b>PM0-10075</b>	2085526	2085552	-1	promoter	[b2013, b2012]	2085526
<b>b2014</b>	2085704	2087062	-1	gene	NaN	2085704
<b>b4678</b>	2087052	2087114	-1	gene	NaN	2087062
<b>b2015</b>	2087329	2088258	-1	gene	NaN	2087329
<b>b2016</b>	2088304	2089128	-1	gene	NaN	2088304
<b>b4539</b>	2089211	2089465	-1	gene	NaN	2089211
<b>b2017</b>	2089462	2089713	-1	gene	NaN	2089465
<b>PM0-10318</b>	2089714	2089733	-1	promoter	[b2017, b4539]	2089714
<b>PM0-7141</b>	2089965	2089995	1	promoter	[b2018, b2019, b2020, b2021, b2022, b2023, b20...	2089965
<b>PM00100</b>	3997907	3997982	1	promoter	[b3813]	3997907

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3813</b>	3997983	4000145	1	gene	NaN	3997983
<b>b4482</b>	4000292	4001056	-1	gene	NaN	4000292
<b>PM00602</b>	4001194	4001425	1	promoter	[b3816]	4001194
<b>b3816</b>	4001426	4002376	1	gene	NaN	4001426
<b>b3817</b>	4002419	4002799	-1	gene	NaN	4002419
<b>b3818</b>	4002813	4003193	-1	gene	NaN	4002813
<b>b3819</b>	4003288	4004178	-1	gene	NaN	4003288
<b>b3820</b>	4004230	4004697	-1	gene	NaN	4004230
<b>PM0-10162</b>	4004698	4004724	-1	promoter	[b3820]	4004698
<b>PM0-10360</b>	4004814	4004861	1	promoter	[b3821]	4004814
<b>b3821</b>	4004862	4005731	1	gene	NaN	4004862
<b>b3822</b>	4005864	4007693	1	gene	NaN	4005864
<b>b3823</b>	4007757	4008377	1	gene	NaN	4007757
<b>b3824</b>	4008439	4009059	-1	gene	NaN	4008439
<b>b3825</b>	4009170	4010192	1	gene	NaN	4009170
<b>b3826</b>	4010200	4011000	1	gene	NaN	4010200
<b>b3827</b>	4011076	4011975	1	gene	NaN	4011076
<b>b3828</b>	4011863	4012816	-1	gene	NaN	4011975
<b>PM0-8603</b>	4012817	4012855	-1	promoter	[b3828]	4012817
<b>PM0-8601</b>	4012884	4013052	1	promoter	[b3829]	4012884
<b>b3829</b>	4013053	4015314	1	gene	NaN	4013053
<b>b3830</b>	4015354	4016169	-1	gene	NaN	4015354
<b>PM0-9097</b>	4016170	4016343	-1	promoter	[b3830]	4016170
<b>PM00179</b>	4016391	4016430	1	promoter	[b3831]	4016391
<b>b3831</b>	4016431	4017192	1	gene	NaN	4016431

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3832</b>	4017333	4018760	1	gene	NaN	4017333
<b>b3833</b>	4018855	4019610	1	gene	NaN	4018855
<b>PM0-10091</b>	2532344	2532408	1	promoter	[b2414]	2532344
<b>b2414</b>	2532409	2533380	1	gene	NaN	2532409
<b>PM00633</b>	2533498	2533763	1	promoter	[b2415, b2416, b2417]	2533498
<b>b0433</b>	452070	453545	-1	gene	NaN	452070
<b>b0434</b>	453589	454167	-1	gene	NaN	453589
<b>PM0-46359</b>	454168	454210	-1	promoter	[b0433, b0434]	454168
<b>PM0-9325</b>	454349	454471	1	promoter	[b0435]	454349
<b>b0435</b>	454472	454789	1	gene	NaN	454472
<b>PM0-9980</b>	454993	455132	1	promoter	[b0436]	454993
<b>b0436</b>	455133	456431	1	gene	NaN	455133
<b>PM00339</b>	456577	456676	1	promoter	[b0437, b0438]	456577
<b>PM0-9744</b>	456605	456676	1	promoter	[b0438, b0437]	456676
<b>PM00569</b>	456646	456676	1	promoter	[b0437, b0438]	456676
<b>b0437</b>	456677	457300	1	gene	NaN	456677
<b>PM0-9696</b>	457202	457425	1	promoter	[b0439, b0438]	457300
<b>b0438</b>	457426	458700	1	gene	NaN	457426
<b>sraA</b>	458728	458784	-1	gene	NaN	458728
<b>PM00511</b>	458815	458887	1	promoter	[b0439]	458815

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b0439</b>	458888	461242	1	gene	NaN	458888
<b>PM590</b>	461305	461450	1	promoter	[b0440]	461305
<b>b0440</b>	461451	461723	1	gene	NaN	461451
<b>PM00615</b>	461819	461914	1	promoter	[b0441]	461819
<b>b0441</b>	461915	463786	1	gene	NaN	461915
<b>b0442</b>	463937	464308	1	gene	NaN	463937
<b>PM0-44483</b>	464360	464401	1	promoter	[b0443]	464360
<b>b0443</b>	464402	464800	1	gene	NaN	464402
<b>b0444</b>	464852	465547	-1	gene	NaN	464852
<b>b0445</b>	465612	467312	-1	gene	NaN	465612
<b>b0446</b>	467412	468230	1	gene	NaN	467412
<b>PM0-10221</b>	468352	468382	1	promoter	[b0447]	468352
<b>b0447</b>	468383	468841	1	gene	NaN	468383
<b>b0448</b>	468871	470643	1	gene	NaN	468871
<b>b0449</b>	470636	472417	1	gene	NaN	470643
<b>PM687</b>	472556	472597	1	promoter	[b0451, b0450]	472556
<b>b0450</b>	472598	472936	1	gene	NaN	472598
<b>b0451</b>	472966	474252	1	gene	NaN	472966
<b>b0452</b>	474301	475161	-1	gene	NaN	474301
<b>b3465</b>	3604393	3604989	1	gene	NaN	3604393
<b>b3466</b>	3604979	3605248	1	gene	NaN	3604989
<b>b3467</b>	3605251	3605610	-1	gene	NaN	3605251
<b>b3468</b>	3605751	3606377	1	gene	NaN	3605751
<b>PM0-4863</b>	3606422	3606450	1	promoter	[b3469]	3606422

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3469</b>	3606451	3608649	1	gene	NaN	3606451
<b>b3470</b>	3608751	3608996	-1	gene	NaN	3608751
<b>b3471</b>	3609217	3609882	1	gene	NaN	3609217
<b>b3472</b>	3609955	3610512	1	gene	NaN	3609955
<b>b3473</b>	3610516	3611733	-1	gene	NaN	3610516
<b>b3474</b>	3611865	3612914	1	gene	NaN	3611865
<b>b3475</b>	3612969	3613556	1	gene	NaN	3612969
<b>PM0-10466</b>	3613439	3613666	1	promoter	[b3480, b3481, b3479, b3478, b3477, b3476]	3613556
<b>PM881</b>	3613628	3613666	1	promoter	[b3481, b3476, b3477, b3478, b3479, b3480]	3613666
<b>b3476</b>	3613667	3615241	1	gene	NaN	3613667
<b>b3477</b>	3615241	3616185	1	gene	NaN	3615241
<b>b3478</b>	3616182	3617015	1	gene	NaN	3616185
<b>b3479</b>	3617015	3617779	1	gene	NaN	3617015
<b>b3480</b>	3617776	3618582	1	gene	NaN	3617779
<b>PM786</b>	3618537	3618587	1	promoter	[b3481]	3618582
<b>b3481</b>	3618588	3618989	1	gene	NaN	3618588
<b>b3482</b>	3619192	3623427	1	gene	NaN	3619192
<b>b3483</b>	3623399	3623782	1	gene	NaN	3623427

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b4552</b>	3623887	3624237	1	gene	NaN	3623887
<b>b3484</b>	3624378	3625514	1	gene	NaN	3624378
<b>b3485</b>	3625679	3626803	-1	gene	NaN	3625679
<b>b3486</b>	3626803	3629538	-1	gene	NaN	3626803
<b>b3487</b>	3629535	3630602	-1	gene	NaN	3629538
<b>b3488</b>	3630968	3632590	-1	gene	NaN	3630968
<b>b4660</b>	3632852	3634458	-1	gene	NaN	3632852
<b>PM0-36192</b>	3634743	3634840	1	promoter	[b3491]	3634743
<b>b3491</b>	3634841	3635893	1	gene	NaN	3634841
<b>b3492</b>	3636208	3637410	-1	gene	NaN	3636208
<b>PM0-8927</b>	3637612	3637641	1	promoter	[b3493]	3637612
<b>b3493</b>	3637642	3639141	1	gene	NaN	3637642
<b>b3494</b>	3639385	3639720	-1	gene	NaN	3639385
<b>PM0-1358</b>	3639721	3639848	-1	promoter	[b3494]	3639721
<b>PM00451</b>	3639983	3640110	1	promoter	[b3495]	3639983
<b>b3495</b>	3640111	3640545	1	gene	NaN	3640111
<b>b3496</b>	3640862	3642331	1	gene	NaN	3640862
<b>b3497</b>	3642380	3643132	-1	gene	NaN	3642380
<b>b3498</b>	3643140	3645182	-1	gene	NaN	3643140
<b>PM0-9720</b>	3645183	3645238	-1	promoter	[b3498, b3497]	3645183
<b>PM0-10146</b>	3645350	3645384	1	promoter	[b3499]	3645350
<b>b3499</b>	3645385	3646227	1	gene	NaN	3645385
<b>PM0-15</b>	3646112	3646298	1	promoter	[b3500]	3646227
<b>b3500</b>	3646299	3647651	1	gene	NaN	3646299
<b>b4613</b>	3647705	3647788	-1	gene	NaN	3647705
<b>PM0-8643</b>	3647789	3647982	-1	promoter	[b4613]	3647789

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-46267</b>	3648063	3648062	1	promoter	[b4712]	3648063
<b>b4712</b>	3648063	3648146	1	gene	NaN	3648063
<b>PM0-46268</b>	3648294	3648293	1	promoter	[b4713]	3648294
<b>b4713</b>	3648294	3648377	1	gene	NaN	3648294
<b>PM336</b>	3648512	3648527	1	promoter	[b3501, b3502, b3503]	3648512
<b>b3501</b>	3648528	3648881	1	gene	NaN	3648528
<b>b3502</b>	3648935	3650224	1	gene	NaN	3648935
<b>b3503</b>	3650237	3650662	1	gene	NaN	3650237
<b>b3504</b>	3651291	3653713	1	gene	NaN	3651291
<b>b3505</b>	3652182	3653198	-1	gene	NaN	3653713
<b>PM00137</b>	3653936	3653960	1	promoter	[b3507, b3506]	3653936
<b>b3506</b>	3653961	3654527	1	gene	NaN	3653961
<b>b3507</b>	3654683	3655213	1	gene	NaN	3654683
<b>b3508</b>	3655255	3655902	-1	gene	NaN	3655255
<b>b3509</b>	3655966	3656292	-1	gene	NaN	3655966
<b>b3510</b>	3656408	3656740	-1	gene	NaN	3656408
<b>PM928</b>	3656741	3656791	-1	promoter	[b3508, b3509, b3510]	3656741
<b>PM0-5143</b>	3656960	3656994	1	promoter	[b3511]	3656960
<b>b3511</b>	3656995	3657567	1	gene	NaN	3656995
<b>PM0-10349</b>	3657800	3658365	1	promoter	[b3512]	3657800
<b>b4704</b>	3657986	3658054	-1	gene	NaN	3658365
<b>PM0-36220</b>	3658055	3658054	-1	promoter	[b4704]	3658365

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-10366</b>	3658062	3658365	1	promoter	[b3514, b3513, b3512]	3658365
<b>PM0-5141</b>	3658242	3658365	1	promoter	[b3512]	3658365
<b>PM0-5982</b>	3658345	3658365	1	promoter	[b3514, b3513, b3512]	3658365
<b>b3512</b>	3658366	3658893	1	gene	NaN	3658366
<b>PM0-10251</b>	3659084	3659231	1	promoter	[b3514, b3513]	3659084
<b>b3513</b>	3659232	3660389	1	gene	NaN	3659232
<b>b3514</b>	3660414	3663527	1	gene	NaN	3660414
<b>b3515</b>	3663890	3664618	-1	gene	NaN	3663890
<b>PM0-10288</b>	3664619	3664781	-1	promoter	[b3515]	3664619
<b>PM0-6061</b>	3664864	3664863	1	promoter	[b4452]	3664864
<b>b4452</b>	3664864	3664968	1	gene	NaN	3664864
<b>b3516</b>	3664986	3665810	-1	gene	NaN	3664986
<b>PM0-2441</b>	3665811	3665839	-1	promoter	[b3515, b3516]	3665811
<b>b3517</b>	3666180	3667580	-1	gene	NaN	3666180
<b>PM461</b>	3667581	3667607	-1	promoter	[b3516, b3517]	3667581
<b>b3518</b>	3667791	3669188	-1	gene	NaN	3667791
<b>PM0-9588</b>	3669189	3669261	-1	promoter	[b3518]	3669189
<b>PM0-46122</b>	3669551	3669591	1	promoter	[b3519]	3669551
<b>b3519</b>	3669592	3671241	1	gene	NaN	3669592
<b>b3520</b>	3671292	3671894	-1	gene	NaN	3671292
<b>PM0-10468</b>	3672136	3672413	1	promoter	[b3521]	3672136
<b>b3521</b>	3672414	3673313	1	gene	NaN	3672414



	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3522</b>	3673362	3674375	1	gene	NaN	3673362
<b>b3523</b>	3674786	3676108	1	gene	NaN	3674786
<b>b3524</b>	3676290	3678350	-1	gene	NaN	3676290
<b>PM0-9766</b>	3678351	3678412	-1	promoter	[b3524]	3678351
<b>b3525</b>	3678420	3679187	-1	gene	NaN	3678420
<b>PM0-1730</b>	3679188	3679229	-1	promoter	[b3525]	3679188
<b>b3526</b>	3679419	3680348	1	gene	NaN	3679419
<b>b3527</b>	3680444	3681940	-1	gene	NaN	3680444
<b>PM-8803</b>	3681941	3682016	-1	promoter	[b3527]	3681941
<b>b3528</b>	3682161	3683447	-1	gene	NaN	3682161
<b>PM0-1621</b>	3683448	3683498	-1	promoter	[b3528]	3683448
<b>b3529</b>	3683630	3685618	-1	gene	NaN	3683630
<b>b3530</b>	3685700	3689173	-1	gene	NaN	3685700
<b>b3531</b>	3689155	3690261	-1	gene	NaN	3689173
<b>b3532</b>	3690268	3692607	-1	gene	NaN	3690268
<b>PM0-8735</b>	3692608	3692673	-1	promoter	[b3531, b3532]	3692608
<b>b3533</b>	3692618	3695236	-1	gene	NaN	3692673
<b>b3534</b>	3695233	3695985	-1	gene	NaN	3695236
<b>b3535</b>	3695997	3696185	-1	gene	NaN	3695997
<b>PM0-46354</b>	3696186	3696364	-1	promoter	[b3535]	3696186
<b>PM0-46353</b>	3696443	3696457	1	promoter	[b3538, b3537, b3536]	3696443
<b>b3536</b>	3696458	3698029	1	gene	NaN	3696458

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3537</b>	3698026	3698217	1	gene	NaN	3698029
<b>b3538</b>	3698214	3699893	1	gene	NaN	3698217
<b>b4453</b>	3699980	3700087	-1	gene	NaN	3699980
<b>PM0-9604</b>	3700088	3700272	-1	promoter	[b4453]	3700088
<b>PM0-9605</b>	3700136	3700135	1	promoter	[b4454]	3700272
<b>b4454</b>	3700136	3700201	1	gene	NaN	3700272
<b>b3539</b>	3700563	3701834	1	gene	NaN	3700563
<b>b3540</b>	3701864	3702868	-1	gene	NaN	3701864
<b>b3541</b>	3702865	3703848	-1	gene	NaN	3702868
<b>b3542</b>	3703859	3704761	-1	gene	NaN	3703859
<b>b3543</b>	3704771	3705790	-1	gene	NaN	3704771
<b>b3544</b>	3706098	3707705	-1	gene	NaN	3706098
<b>PM00546</b>	3707706	3707870	-1	promoter	[b3540, b3541, b3542, b3543, b3544]	3707706
<b>b3545</b>	3708616	3708692	-1	gene	NaN	3708616
<b>PM00674</b>	3708693	3708694	-1	promoter	[b3545]	3708693
<b>b3546</b>	3708784	3710475	-1	gene	NaN	3708784
<b>PM0-10506</b>	3710476	3710581	-1	promoter	[b3546]	3710476
<b>b3547</b>	3710799	3712007	-1	gene	NaN	3710799
<b>PM0-46200</b>	3712008	3712043	-1	promoter	[b3547]	3712008
<b>b3548</b>	3712236	3712934	-1	gene	NaN	3712236
<b>b3549</b>	3713092	3713655	1	gene	NaN	3713092

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b3550</b>	3713652	3714092	1	gene	NaN	3713655
<b>b3551</b>	3714061	3716394	-1	gene	NaN	3714092
<b>b3552</b>	3716547	3717206	1	gene	NaN	3716547
<b>b3553</b>	3717310	3718284	1	gene	NaN	3717310
<b>b3554</b>	3718334	3719044	-1	gene	NaN	3718334
<b>PM0-10148</b>	3719045	3719097	-1	promoter	[b3554]	3719045
<b>PM0-8525</b>	3719431	3719477	1	promoter	[b3555]	3719431
<b>b3555</b>	3719478	3719768	1	gene	NaN	3719478
<b>PM00452</b>	3719889	3720048	1	promoter	[b3556]	3719889
<b>b3556</b>	3720049	3720261	1	gene	NaN	3720049
<b>b4647</b>	3720448	3720632	-1	gene	NaN	3720448
<b>b4455</b>	3720448	3720600	-1	gene	NaN	3720632
<b>b3557</b>	3720680	3721201	1	gene	NaN	3720680
<b>PM0-8642</b>	3721065	3721153	1	promoter	[b3558]	3721201
<b>b3558</b>	3721198	3722049	1	gene	NaN	3721201
<b>b4614</b>	3722076	3722105	1	gene	NaN	3722076
<b>b3559</b>	3722328	3724397	-1	gene	NaN	3722328
<b>b4493</b>	1489713	1490713	-1	gene	NaN	1489713
<b>b1418</b>	1490902	1491432	1	gene	NaN	1490902
<b>b4597</b>	1491443	1491506	-1	gene	NaN	1491443
<b>PM0-9545</b>	1491507	1491506	-1	promoter	[b4597]	1491507
<b>b1419</b>	1491677	1491850	1	gene	NaN	1491677
<b>b4428</b>	1491922	1492071	-1	gene	NaN	1491922
<b>b1420</b>	1491962	1492129	-1	gene	NaN	1492071
<b>PM0-46364</b>	1492119	1492118	1	promoter	[b4429]	1492129
<b>b4429</b>	1492119	1492174	1	gene	NaN	1492129

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-10039</b>	1492439	1492469	1	promoter	[b1421]	1492439
<b>b1421</b>	1492470	1494110	1	gene	NaN	1492470
<b>b1422</b>	1494148	1495071	-1	gene	NaN	1494148
<b>PM0-36238</b>	1495072	1495134	-1	promoter	[b1422]	1495072
<b>b1423</b>	1495288	1496631	1	gene	NaN	1495288
<b>b1424</b>	1496856	1498511	1	gene	NaN	1496856
<b>b1426</b>	1498651	1498875	1	gene	NaN	1498651
<b>PM00370</b>	1498743	1498937	1	promoter	[b1427]	1498875
<b>b1427</b>	1498938	1499477	1	gene	NaN	1498938
<b>b1428</b>	1499469	1500449	-1	gene	NaN	1499477
<b>PM-8830</b>	1500536	1500572	1	promoter	[b1429, b1430]	1500536
<b>b1429</b>	1500573	1501565	1	gene	NaN	1500573
<b>b1430</b>	1501562	1502155	1	gene	NaN	1501565
<b>b1431</b>	1502457	1503125	1	gene	NaN	1502457
<b>b4578</b>	1503218	1503649	-1	gene	NaN	1503218
<b>b1432</b>	1503717	1504865	1	gene	NaN	1503717
<b>b1433</b>	1504905	1506080	-1	gene	NaN	1504905
<b>PM0-46335</b>	1506150	1506171	1	promoter	[b1434]	1506150
<b>b1434</b>	1506172	1506708	1	gene	NaN	1506172
<b>b1435</b>	1506781	1508742	1	gene	NaN	1506781
<b>b1436</b>	1508834	1509064	-1	gene	NaN	1508834
<b>PM0-10389</b>	1509221	1509285	1	promoter	[b1438, b4532]	1509221
<b>b4532</b>	1509286	1509462	1	gene	NaN	1509286

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1438</b>	1509508	1509924	1	gene	NaN	1509508
<b>PM0-10040</b>	1509978	1510002	1	promoter	[b1439]	1509978
<b>b1439</b>	1510003	1511409	1	gene	NaN	1510003
<b>PM0-46279</b>	1511599	1511653	1	promoter	[b1444, b1443, b1442, b1441, b1440]	1511599
<b>b1440</b>	1511654	1512799	1	gene	NaN	1511654
<b>b1441</b>	1512817	1513830	1	gene	NaN	1512817
<b>b1442</b>	1513831	1514772	1	gene	NaN	1513831
<b>b1443</b>	1514762	1515556	1	gene	NaN	1514772
<b>b1444</b>	1515578	1517002	1	gene	NaN	1515578
<b>b4598</b>	1517099	1517194	-1	gene	NaN	1517099
<b>b1445</b>	1517389	1517562	1	gene	NaN	1517389
<b>b1446</b>	1517648	1517881	1	gene	NaN	1517648
<b>b1447</b>	1517882	1518331	-1	gene	NaN	1517882
<b>b1448</b>	1518328	1518846	-1	gene	NaN	1518331
<b>b1449</b>	1519027	1520064	1	gene	NaN	1519027
<b>b1450</b>	1520262	1520927	1	gene	NaN	1520262
<b>b1451</b>	1520963	1523065	-1	gene	NaN	1520963
<b>PM0-10041</b>	1523227	1523306	1	promoter	[b1452]	1523227
<b>b1452</b>	1523307	1524368	1	gene	NaN	1523307
<b>b1453</b>	1524481	1525980	-1	gene	NaN	1524481
<b>PM0-8528</b>	1525981	1526020	-1	promoter	[b1453]	1525981
<b>PM0-10042</b>	1526228	1526246	1	promoter	[b1454]	1526228
<b>b1454</b>	1526247	1526864	1	gene	NaN	1526247
<b>b1455</b>	1526940	1527152	1	gene	NaN	1526940

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1456</b>	1527902	1529938	1	gene	NaN	1527902
<b>b1457</b>	1529922	1530404	1	gene	NaN	1529938
<b>b1459</b>	1530586	1531639	1	gene	NaN	1530586
<b>b1460</b>	1531816	1532952	1	gene	NaN	1531816
<b>b1461</b>	1533052	1533285	1	gene	NaN	1533052
<b>b1462</b>	1533282	1533851	-1	gene	NaN	1533285
<b>b1463</b>	1534024	1534869	1	gene	NaN	1534024
<b>b1464</b>	1534965	1535858	-1	gene	NaN	1534965
<b>b1465</b>	1535937	1536617	-1	gene	NaN	1535937
<b>b1466</b>	1536614	1537309	-1	gene	NaN	1536617
<b>b1467</b>	1537309	1538853	-1	gene	NaN	1537309
<b>PM-8813</b>	1537310	1538687	-1	promoter	[b1465, b1466]	1538853
<b>b1468</b>	1538850	1542590	-1	gene	NaN	1538853
<b>b1469</b>	1542672	1544060	-1	gene	NaN	1542672
<b>PM0-2841</b>	1544061	1544163	-1	promoter	[b1469]	1544061
<b>b1471</b>	1544384	1545714	-1	gene	NaN	1544384
<b>b1472</b>	1545750	1546028	-1	gene	NaN	1545750
<b>b1473</b>	1546288	1547169	-1	gene	NaN	1546288
<b>PM0-10383</b>	1547170	1547205	-1	promoter	[b1473]	1547170
<b>PM00217</b>	1547372	1547400	1	promoter	[b1474, b1475, b1476]	1547372
<b>b1474</b>	1547401	1550448	1	gene	NaN	1547401
<b>b1475</b>	1550461	1551345	1	gene	NaN	1550461
<b>b1476</b>	1551338	1551991	1	gene	NaN	1551345
<b>b1477</b>	1552398	1552682	-1	gene	NaN	1552398
<b>b1478</b>	1552828	1553838	-1	gene	NaN	1552828
<b>b1479</b>	1553972	1555669	-1	gene	NaN	1553972
<b>b1480</b>	1555826	1555963	-1	gene	NaN	1555826

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-4141</b>	1555964	1556046	-1	promoter	[b1480]	1555964
<b>b1481</b>	1556065	1556280	-1	gene	NaN	1556065
<b>PM0-7481</b>	1556281	1556340	-1	promoter	[b1480, b1481]	1556281
<b>PM00544</b>	1556596	1556624	1	promoter	[b1482]	1556596
<b>b1482</b>	1556625	1557056	1	gene	NaN	1556625
<b>b1483</b>	1557112	1558038	-1	gene	NaN	1557112
<b>b1484</b>	1558031	1559017	-1	gene	NaN	1558038
<b>b1485</b>	1559014	1559910	-1	gene	NaN	1559017
<b>b1486</b>	1559907	1560929	-1	gene	NaN	1559910
<b>b1487</b>	1560931	1562481	-1	gene	NaN	1560931
<b>b1488</b>	1562495	1563076	-1	gene	NaN	1562495
<b>PM0-46160</b>	1563077	1563249	-1	promoter	[b1483, b1484, b1485, b1486, b1487, b1488]	1563077
<b>b1489</b>	1563334	1565733	-1	gene	NaN	1563334
<b>b1490</b>	1565758	1567140	-1	gene	NaN	1565758
<b>b1491</b>	1567504	1568823	-1	gene	NaN	1567504
<b>b1492</b>	1568954	1570489	-1	gene	NaN	1568954
<b>b1493</b>	1570645	1572045	-1	gene	NaN	1570645
<b>PM450</b>	1572046	1572072	-1	promoter	[b1492, b1493]	1572046
<b>b1494</b>	1572407	1575202	-1	gene	NaN	1572407
<b>b1495</b>	1575247	1577619	-1	gene	NaN	1575247

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1496</b>	1577657	1579342	-1	gene	NaN	1577657
<b>b1497</b>	1579633	1580790	-1	gene	NaN	1579633
<b>b1498</b>	1580842	1582524	-1	gene	NaN	1580842
<b>PM0-44963</b>	1582525	1582554	-1	promoter	[b1497, b1498]	1582525
<b>b1499</b>	1582926	1583687	-1	gene	NaN	1582926
<b>b1500</b>	1583762	1583959	-1	gene	NaN	1583762
<b>PM0-9828</b>	1583960	1584028	-1	promoter	[b1499, b1500]	1583960
<b>b1501</b>	1584207	1586486	-1	gene	NaN	1584207
<b>PM0-9849</b>	1586487	1586664	-1	promoter	[b1501]	1586487
<b>b1502</b>	1586820	1587734	-1	gene	NaN	1586820
<b>b1503</b>	1587793	1588296	-1	gene	NaN	1587793
<b>b1504</b>	1588309	1588839	-1	gene	NaN	1588309
<b>b1505</b>	1588853	1590079	-1	gene	NaN	1588853
<b>b1506</b>	1590334	1590536	-1	gene	NaN	1590334
<b>b1507</b>	1590854	1592176	-1	gene	NaN	1590854
<b>b1508</b>	1592176	1592442	-1	gene	NaN	1592176
<b>PM0-8864</b>	1592443	1592482	-1	promoter	[b1507, b1508]	1592443
<b>b4696</b>	1592665	1597987	-1	gene	NaN	1592665
<b>b1511</b>	1598617	1600209	-1	gene	NaN	1598617
<b>b1512</b>	1600288	1601241	-1	gene	NaN	1600288
<b>PM0-10233</b>	1601242	1601257	-1	promoter	[b1511, b1512]	1601242
<b>PM0-8241</b>	1601460	1601489	1	promoter	[b1519, b1518, b1517, b1516, b1515, b1514, b1513]	1601460



	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1513</b>	1601490	1603025	1	gene	NaN	1601490
<b>b1514</b>	1603019	1604047	1	gene	NaN	1603025
<b>b1515</b>	1604047	1605039	1	gene	NaN	1604047
<b>b1516</b>	1605051	1606073	1	gene	NaN	1605051
<b>b1517</b>	1606100	1606975	1	gene	NaN	1606100
<b>b1518</b>	1606999	1607289	1	gene	NaN	1606999
<b>PM0-9760</b>	1607221	1607345	1	promoter	[b1519]	1607289
<b>b1519</b>	1607346	1608104	1	gene	NaN	1607346
<b>b1520</b>	1608108	1609022	-1	gene	NaN	1608108
<b>b1521</b>	1609229	1610680	-1	gene	NaN	1609229
<b>PM0-36188</b>	1610681	1610720	-1	promoter	[b1521]	1610681
<b>b1522</b>	1610907	1611854	-1	gene	NaN	1610907
<b>b1523</b>	1611966	1612325	-1	gene	NaN	1611966
<b>b1524</b>	1612325	1613251	-1	gene	NaN	1612325
<b>b1525</b>	1613315	1614703	-1	gene	NaN	1613315
<b>b1526</b>	1614804	1615685	1	gene	NaN	1614804
<b>b1527</b>	1615763	1616878	1	gene	NaN	1615763
<b>PM0-1581</b>	1616955	1617027	1	promoter	[b1528]	1616955
<b>b1528</b>	1617028	1618218	1	gene	NaN	1617028
<b>b1529</b>	1618243	1618908	-1	gene	NaN	1618243
<b>PM0-9954</b>	1618909	1618941	-1	promoter	[b1529]	1618909
<b>PM00265</b>	1619093	1619119	1	promoter	[b1530, b1531, b1532]	1619093
<b>b1530</b>	1619120	1619554	1	gene	NaN	1619120
<b>b1531</b>	1619574	1619957	1	gene	NaN	1619574

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1532</b>	1619989	1620207	1	gene	NaN	1619989
<b>b1533</b>	1620238	1621137	-1	gene	NaN	1620238
<b>b1534</b>	1621332	1622519	1	gene	NaN	1621332
<b>PM0-10505</b>	1622588	1622645	1	promoter	[b4599]	1622588
<b>b4599</b>	1622646	1622741	1	gene	NaN	1622646
<b>b4698</b>	1622817	1622914	-1	gene	NaN	1622817
<b>PM0-10504</b>	1622915	1622914	-1	promoter	[b4698]	1622915
<b>b1535</b>	1622960	1623850	-1	gene	NaN	1622960
<b>PM0-9401</b>	1623851	1623879	-1	promoter	[b1535]	1623851
<b>b1536</b>	1624105	1624497	-1	gene	NaN	1624105
<b>PM0-44976</b>	1624498	1624559	-1	promoter	[b1536]	1624498
<b>PM0-10367</b>	1624688	1624772	1	promoter	[b1537]	1624688
<b>b1537</b>	1624773	1625291	1	gene	NaN	1624773
<b>b1538</b>	1625335	1627380	-1	gene	NaN	1625335
<b>PM0-10044</b>	1627491	1627516	1	promoter	[b1539]	1627491
<b>b1539</b>	1627517	1628263	1	gene	NaN	1627517
<b>PM0-10045</b>	1628326	1628351	1	promoter	[b1540]	1628326
<b>b1540</b>	1628352	1629038	1	gene	NaN	1628352
<b>b1541</b>	1629215	1629418	1	gene	NaN	1629215
<b>b1542</b>	1629453	1630913	-1	gene	NaN	1629453
<b>b4600</b>	1631002	1652838	-1	gene	NaN	1631002
<b>PM0-36239</b>	1632723	1633071	1	promoter	[b1544]	1652838
<b>b1544</b>	1633072	1633305	1	gene	NaN	1652838
<b>b1545</b>	1633622	1634212	1	gene	NaN	1652838
<b>b1546</b>	1634310	1634885	-1	gene	NaN	1652838
<b>b1547</b>	1634885	1635847	-1	gene	NaN	1652838
<b>b1548</b>	1635840	1636367	-1	gene	NaN	1652838

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-8061</b>	1636368	1636603	-1	promoter	[b1546, b1547, b1548]	1636368
<b>b4533</b>	1636756	1636989	1	gene	NaN	1636756
<b>b1549</b>	1637047	1637457	1	gene	NaN	1637047
<b>b1550</b>	1637609	1637782	-1	gene	NaN	1637609
<b>PM0-10046</b>	1637783	1637846	-1	promoter	[b1550]	1637783
<b>b1551</b>	1637954	1638109	-1	gene	NaN	1637954
<b>b1552</b>	1638455	1638667	-1	gene	NaN	1638455
<b>PM0-7321</b>	1638668	1638812	-1	promoter	[b1552]	1638668
<b>b1553</b>	1639030	1639527	-1	gene	NaN	1639030
<b>b4689</b>	1639080	1639334	-1	gene	NaN	1639527
<b>b1554</b>	1639524	1640057	-1	gene	NaN	1639527
<b>b1555</b>	1640054	1640365	-1	gene	NaN	1640057
<b>b1556</b>	1640370	1640585	-1	gene	NaN	1640370
<b>b1557</b>	1641339	1641554	-1	gene	NaN	1641339
<b>b1558</b>	1641855	1642067	1	gene	NaN	1641855
<b>b1559</b>	1642489	1643241	-1	gene	NaN	1642489
<b>b1560</b>	1643255	1644304	-1	gene	NaN	1643255
<b>b1561</b>	1644651	1644902	-1	gene	NaN	1644651
<b>b1562</b>	1645119	1645274	-1	gene	NaN	1645119
<b>b1563</b>	1645346	1645633	-1	gene	NaN	1645346
<b>b1564</b>	1645633	1645872	-1	gene	NaN	1645633
<b>PM0-9527</b>	1645873	1645903	-1	promoter	[b1562, b1563, b1564]	1645873
<b>b1565</b>	1645897	1646202	1	gene	NaN	1645903
<b>PM0-1661</b>	1646388	1646404	1	promoter	[b1566]	1646388
<b>b1566</b>	1646405	1646737	1	gene	NaN	1646405
<b>b1567</b>	1647174	1647356	-1	gene	NaN	1647174

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1568</b>	1647358	1647636	-1	gene	NaN	1647358
<b>b1569</b>	1647620	1647850	-1	gene	NaN	1647636
<b>PM0-46242</b>	1647851	1647876	-1	promoter	[b1569, b1568, b1567]	1647851
<b>b1570</b>	1647934	1648341	1	gene	NaN	1647934
<b>PM0-44991</b>	1648438	1648507	1	promoter	[b1573, b1572, b1571]	1648438
<b>b1571</b>	1648508	1648663	1	gene	NaN	1648508
<b>b1572</b>	1648665	1648793	1	gene	NaN	1648665
<b>b1573</b>	1648823	1649041	1	gene	NaN	1648823
<b>b1574</b>	1649382	1649434	1	gene	NaN	1649382
<b>PM0-9327</b>	1649597	1649608	1	promoter	[b1579, b1578, b1577, b1576, b1575]	1649597
<b>b1575</b>	1649609	1649797	1	gene	NaN	1649609
<b>b1576</b>	1649794	1649985	1	gene	NaN	1649797
<b>b1577</b>	1650078	1650842	1	gene	NaN	1650078
<b>b1578</b>	1650845	1651537	1	gene	NaN	1650845
<b>b1579</b>	1651551	1652708	1	gene	NaN	1651551
<b>b1580</b>	1652896	1653915	-1	gene	NaN	1652896
<b>b1581</b>	1653927	1655141	-1	gene	NaN	1653927
<b>b1582</b>	1655347	1655673	-1	gene	NaN	1655347
<b>PM0-10047</b>	1655709	1655807	1	promoter	[b1584, b1583]	1655709
<b>b1583</b>	1655808	1656149	1	gene	NaN	1655808
<b>b1584</b>	1656184	1656744	1	gene	NaN	1656184
<b>b1585</b>	1656747	1657457	-1	gene	NaN	1656747

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1586</b>	1657565	1657870	1	gene	NaN	1657565
<b>PM0-5721</b>	1658003	1658068	1	promoter	[b1591, b1590, b1589, b1588, b1587]	1658003
<b>b1587</b>	1658069	1660495	1	gene	NaN	1658069
<b>b1588</b>	1660556	1662979	1	gene	NaN	1660556
<b>b1589</b>	1662990	1663607	1	gene	NaN	1662990
<b>b1590</b>	1663609	1664463	1	gene	NaN	1663609
<b>b1591</b>	1664506	1665120	1	gene	NaN	1664506
<b>b1592</b>	1665315	1666571	1	gene	NaN	1665315
<b>b1593</b>	1666524	1667219	-1	gene	NaN	1666571
<b>b1594</b>	1667344	1668564	-1	gene	NaN	1667344
<b>PM935</b>	1668565	1668591	-1	promoter	[b1593, b1594]	1668565
<b>PM934</b>	1668565	1668603	-1	promoter	[b1594]	1668591
<b>b1595</b>	1668699	1669592	-1	gene	NaN	1668699
<b>PM0-36240</b>	1669672	1669698	1	promoter	[b1596]	1669672
<b>b1596</b>	1669699	1670952	1	gene	NaN	1669699
<b>PM0-1881</b>	1671327	1671375	1	promoter	[b1597]	1671327
<b>b1597</b>	1671376	1671684	1	gene	NaN	1671376
<b>b4601</b>	1671777	1671860	1	gene	NaN	1671777
<b>b1598</b>	1671960	1672781	1	gene	NaN	1671960
<b>b1599</b>	1672820	1673149	-1	gene	NaN	1672820
<b>b1600</b>	1673136	1673501	-1	gene	NaN	1673149

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-46378</b>	1673502	1673779	-1	promoter	[b1600, b1599]	1673502
<b>PM0-10048</b>	1673890	1673912	1	promoter	[b1601]	1673890
<b>b1601</b>	1673913	1674947	1	gene	NaN	1673913
<b>b1602</b>	1674972	1676360	-1	gene	NaN	1674972
<b>b1603</b>	1676371	1677903	-1	gene	NaN	1676371
<b>PM124</b>	1677904	1677975	-1	promoter	[b1602, b1603]	1677904
<b>b1604</b>	1678427	1679371	1	gene	NaN	1678427
<b>b1605</b>	1679557	1680939	1	gene	NaN	1679557
<b>b1606</b>	1680976	1681698	1	gene	NaN	1680976
<b>b1607</b>	1681695	1682030	-1	gene	NaN	1681698
<b>PM0-8164</b>	1682138	1682158	1	promoter	[b1609, b1608]	1682138
<b>b1608</b>	1682159	1682878	1	gene	NaN	1682159
<b>b1609</b>	1682882	1684183	1	gene	NaN	1682882
<b>b1610</b>	1684259	1685188	1	gene	NaN	1684259
<b>b1611</b>	1685185	1686588	-1	gene	NaN	1685188
<b>PM00147</b>	1686589	1686709	-1	promoter	[b1611]	1686589
<b>b1612</b>	1686731	1688377	-1	gene	NaN	1686731
<b>PM00146</b>	1688378	1688440	-1	promoter	[b1611, b1612]	1688378
<b>PM00371</b>	1688551	1688575	1	promoter	[b1613]	1688551
<b>b1613</b>	1688576	1689751	1	gene	NaN	1688576
<b>PM0-10049</b>	1689794	1689851	1	promoter	[b1614]	1689794

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b1614</b>	1689852	1691360	1	gene	NaN	1689852
<b>b1615</b>	1691586	1692851	-1	gene	NaN	1691586
<b>b1616</b>	1692890	1694263	-1	gene	NaN	1692890
<b>b1617</b>	1694260	1696071	-1	gene	NaN	1694263
<b>PM0-5321</b>	1696072	1696152	-1	promoter	[b1615, b1616, b1617]	1696072
<b>b1618</b>	1696462	1697052	-1	gene	NaN	1696462
<b>PM0-5322</b>	1697053	1697075	-1	promoter	[b1618]	1697053
<b>b1619</b>	1697273	1698040	-1	gene	NaN	1697273
<b>PM0-10050</b>	1698041	1698122	-1	promoter	[b1619]	1698041
<b>b1620</b>	1698152	1699180	-1	gene	NaN	1698152
<b>PM00260</b>	1699181	1699228	-1	promoter	[b1620]	1699181
<b>PM00177</b>	1699313	1699354	1	promoter	[b1621, b1622]	1699313
<b>b1621</b>	1699355	1700947	1	gene	NaN	1699355
<b>b1622</b>	1700957	1702129	1	gene	NaN	1700957
<b>PM0-45010</b>	1702202	1702232	1	promoter	[b1623]	1702202
<b>b1623</b>	1702233	1703234	1	gene	NaN	1702233
<b>b1624</b>	1703268	1704308	-1	gene	NaN	1703268
<b>PM0-10409</b>	1704512	1704550	1	promoter	[b4409]	1704512
<b>b4409</b>	1704551	1704676	1	gene	NaN	1704551
<b>b1625</b>	1704949	1705164	1	gene	NaN	1704949

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM0-8843</b>	1705207	1705249	1	promoter	[b1633, b1632, b1631, b1630, b1629, b1628, b16...	1705207
<b>b1626</b>	1705250	1705690	1	gene	NaN	1705250
<b>PM0-8845</b>	1705755	1705766	1	promoter	[b1633, b1632, b1631, b1630, b1629, b1628, b1627]	1705755
<b>b1627</b>	1705767	1706348	1	gene	NaN	1705767
<b>b1628</b>	1706348	1706926	1	gene	NaN	1706348
<b>b1629</b>	1706919	1709141	1	gene	NaN	1706926
<b>b1630</b>	1709142	1710200	1	gene	NaN	1709142
<b>b1631</b>	1710204	1710824	1	gene	NaN	1710204
<b>b1632</b>	1710828	1711523	1	gene	NaN	1710828
<b>b1633</b>	1711523	1712158	1	gene	NaN	1711523
<b>PM0-8909</b>	1712671	1712768	1	promoter	[b1634]	1712671
<b>b1634</b>	1712769	1714271	1	gene	NaN	1712769
<b>PM0-10051</b>	1714327	1714376	1	promoter	[b1635]	1714327
<b>b1635</b>	1714377	1714982	1	gene	NaN	1714377
<b>b1636</b>	1715026	1715889	-1	gene	NaN	1715026
<b>b3998</b>	4198790	4199461	1	gene	NaN	4198790
<b>b3999</b>	4199504	4200094	1	gene	NaN	4199504
<b>PM586</b>	4200176	4200280	1	promoter	[b4000]	4200176



	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>b4000</b>	4200281	4200553	1	gene	NaN	4200281
<b>b4001</b>	4200566	4201261	1	gene	NaN	4200566
<b>b4002</b>	4201263	4201688	-1	gene	NaN	4201263
<b>PM923</b>	4201689	4201712	-1	promoter	[b4002]	4201689
<b>PM924</b>	4201905	4201925	1	promoter	[b4004, b4003]	4201905
<b>b4003</b>	4201926	4203323	1	gene	NaN	4201926
<b>b4004</b>	4203320	4204645	1	gene	NaN	4203323
<b>b4005</b>	4204642	4205931	-1	gene	NaN	4204645
<b>b2765</b>	2892214	2892579	1	gene	NaN	2892214
<b>b2766</b>	2892657	2893928	1	gene	NaN	2892657
<b>b2767</b>	2893919	2894179	1	gene	NaN	2893928
<b>b2768</b>	2894196	2894771	1	gene	NaN	2894196
<b>b2769</b>	2894919	2895779	-1	gene	NaN	2894919
<b>b2770</b>	2895776	2896555	-1	gene	NaN	2895779
<b>b2771</b>	2896533	2897870	-1	gene	NaN	2896555
<b>b4463</b>	2897964	2899418	-1	gene	NaN	2897964
<b>b2774</b>	2899488	2900273	-1	gene	NaN	2899488
<b>b2775</b>	2900592	2901869	1	gene	NaN	2900592
<b>b2776</b>	2901896	2903374	1	gene	NaN	2901896
<b>b2777</b>	2904747	2905418	-1	gene	NaN	2904747
<b>b4682</b>	2905557	2905697	1	gene	NaN	2905557
<b>b2778</b>	2905711	2906583	1	gene	NaN	2905711
<b>b2779</b>	2906643	2907941	-1	gene	NaN	2906643
<b>b0010</b>	9928	10494	-1	gene	NaN	9928
<b>b0011</b>	10643	11356	-1	gene	NaN	10643
<b>b0013</b>	11382	11786	-1	gene	NaN	11382

	start	end	strand	class	genes_in_TU	start_if_select_as
<b>\$gene_or_promoter</b>						
<b>PM00316</b>	12048	12162	1	promoter	[b0014, b0015]	12048
<b>b0014</b>	12163	14079	1	gene	NaN	12163
<b>b0015</b>	14168	15298	1	gene	NaN	14168
<b>b0016</b>	15445	16557	1	gene	NaN	15445
<b>b0018</b>	16751	16960	-1	gene	NaN	16751
<b>b4412</b>	16751	16903	-1	gene	NaN	16960
<b>b4413</b>	16952	17006	1	gene	NaN	16960
<b>b0019</b>	17489	18655	1	gene	NaN	17489
<b>b0020</b>	18715	19620	1	gene	NaN	18715
<b>b0021</b>	19811	20314	-1	gene	NaN	19811
<b>b0022</b>	20233	20508	-1	gene	NaN	20314
<b>b0023</b>	20815	21078	-1	gene	NaN	20815
<b>PM0-8222</b>	21079	21210	-1	promoter	[b0023]	21079
<b>b0024</b>	21181	21399	1	gene	NaN	21210
<b>PM0-8866</b>	21383	21406	1	promoter	[b0025, b0026, b0027, b0028, b0029]	21399

```
In [31]: W3110toMG1655 = pd.read_table('data/Ecoli/lambdaRed_W3110_MG1655.locus.tab', na_values='NIL', index_col='Blattner')
W3110toMG1655 = W3110toMG1655.loc[sorted(W3110toMG1655.index.dropna())]
W3110toMG1655
```

Out[31]:

	E_coli_K12_W3110_lambdaRed	E_coli_K12_W3110	W3110-Nan
<b>Blattner</b>			
<b>b0002</b>	W3110_lambdaRed.CDS.1	JW0001	thrA
<b>b0003</b>	W3110_lambdaRed.CDS.2	JW0002	thrB
<b>b0004</b>	W3110_lambdaRed.CDS.3	JW0003	thrC
<b>b0006</b>	W3110_lambdaRed.CDS.5	JW0005	yaaA
<b>b0007</b>	W3110_lambdaRed.CDS.6	JW0006	yaaJ
<b>b0008</b>	W3110_lambdaRed.CDS.7	JW0007	talB
<b>b0009</b>	W3110_lambdaRed.CDS.8	JW0008	mog
<b>b0010</b>	W3110_lambdaRed.CDS.9	JW0009	satP
<b>b0011</b>	W3110_lambdaRed.CDS.10	JW0010	yaaW
<b>b0013</b>	W3110_lambdaRed.CDS.11	JW0012	yaal
<b>b0014</b>	W3110_lambdaRed.CDS.12	JW0013	dnaK
<b>b0015</b>	W3110_lambdaRed.CDS.13	JW0014	dnaJ
<b>b0019</b>	W3110_lambdaRed.CDS.15	JW0018	nhaA
<b>b0020</b>	W3110_lambdaRed.CDS.16	JW0019	nhaR
<b>b0021</b>	W3110_lambdaRed.CDS.3854	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.17	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.2294	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.3854	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.17	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.2294	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.3854	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.17	JW0020	insB-1
<b>b0021</b>	W3110_lambdaRed.CDS.2294	JW0020	insB-1
<b>b0023</b>	W3110_lambdaRed.CDS.18	JW0022	rpsT
<b>b0025</b>	W3110_lambdaRed.CDS.19	JW0023	ribF
<b>b0026</b>	W3110_lambdaRed.CDS.20	JW0024	ileS
<b>b0027</b>	W3110_lambdaRed.CDS.21	JW0025	lspA
<b>b0028</b>	W3110_lambdaRed.CDS.22	JW0026	fkpB
<b>b0029</b>	W3110_lambdaRed.CDS.23	JW0027	ispH
<b>b0030</b>	W3110_lambdaRed.CDS.24	JW0028	rihC

	E_coli_K12_W3110_lambdaRed	E_coli_K12_W3110	W3110-Nan
<b>Blattner</b>			
...	...	...	...
<b>b4511</b>	W3110_lambdaRed.CDS.589	JW0577	ybdZ
<b>b4512</b>	W3110_lambdaRed.CDS.604	JW0591	ybdD
<b>b4513</b>	W3110_lambdaRed.CDS.702	JW0687	kdpF
<b>b4514</b>	W3110_lambdaRed.CDS.707	JW0694	ybfQ
<b>b4515</b>	W3110_lambdaRed.CDS.738	JW0724	cydX
<b>b4520</b>	W3110_lambdaRed.CDS.1130	JW1156	ymgF
<b>b4522</b>	W3110_lambdaRed.CDS.1246	JW5198+JW5199	ECK1270:JW5198+JW5199:b4
<b>b4524</b>	W3110_lambdaRed.CDS.1295	JW5203	ymjB
<b>b4525</b>	W3110_lambdaRed.CDS.1304	JW5960	ymjC
<b>b4526</b>	W3110_lambdaRed.CDS.1328	JW1346	ydaE
<b>b4527</b>	W3110_lambdaRed.CDS.1331	JW1349	ydaF
<b>b4529</b>	W3110_lambdaRed.CDS.1357	JW5215	ydbJ
<b>b4533</b>	W3110_lambdaRed.CDS.1536	JW5251	ynfO
<b>b4535</b>	W3110_lambdaRed.CDS.1707	JW5911	yniD
<b>b4536</b>	W3110_lambdaRed.CDS.1817	JW5298	yobH
<b>b4537</b>	W3110_lambdaRed.CDS.1895	JW1891	yecJ
<b>b4539</b>	W3110_lambdaRed.CDS.2003	JW5331	yoeB
<b>b4544</b>	W3110_lambdaRed.CDS.2260	JW2252	arnE

	E_coli_K12_W3110_lambdaRed	E_coli_K12_W3110	W3110-Nan
<b>Blattner</b>			
<b>b4546</b>	W3110_lambdaRed.CDS.2415	JW5877	ypeB
<b>b4547</b>	W3110_lambdaRed.CDS.2469	JW2457	ypfN
<b>b4551</b>	W3110_lambdaRed.CDS.3949	JW3312	yheV
<b>b4553</b>	W3110_lambdaRed.CDS.3725	JW3532	ysaB
<b>b4554</b>	W3110_lambdaRed.CDS.3681	JW3576	yibT
<b>b4555</b>	W3110_lambdaRed.CDS.3620	JW5965	yicS
<b>b4558</b>	W3110_lambdaRed.CDS.3475	JW3781	yifL
<b>b4559</b>	W3110_lambdaRed.CDS.4156	JW5732	ghoT
<b>b4565</b>	W3110_lambdaRed.CDS.4332	JW5967	sgcB
<b>b4566</b>	W3110_lambdaRed.CDS.4336	JW5968	yjhX
<b>b4567</b>	W3110_lambdaRed.CDS.4405	JW5797	yjjZ
<b>b4568</b>	W3110_lambdaRed.CDS.4412	JW5891	ytjA

4278 rows × 12 columns