

8 Pathway Overview of the pan-hif-1 transcriptome

January 31, 2018

1 Table of Contents

- 1 Defining a gene compactifier for easy printing
 - 2 Effects of HIF-1 on mitochondrial proteins
 - 3 HIF-1 effects on the ribosome
 - 4 Bioenergetics of HIF-1
 - 5 Effects of HIF-1 on the Proteasome and Mediator
 - 5.1 Effect of HIF-1 on proteins involved in 'protein catabolic process'
 - 6 Proteins annotated as involved in protein folding
 - 7 Immune Involvement

In []:

An important question that I also wanted to address was the cell-wide effects of HIF-1. Although the hypoxia response by itself is informative as to what HIF-1 actually turns on and off in *C. elegans*, enrichment analyses are not the only way to get information out of transcriptomes.

Another way to get information about these effects is to change what biological units we are studying. In this paper, we have focused a lot on single genes. However, we could also ask what pathways, or what entities, are represented in our dataset.

The way I will look at pathways is by identifying the genes that are in a 'pathway' or biological process of interest. I will extract the genes within this process that are differentially expressed in each mutant. Then, I will look at how the pathway changes *overall*. If a pathway is being down-regulated in a given set of mutants, we would expect that all of the genes that are D.E. in this pathway would show up as down-regulated. However, we no longer require that ALL of the genes in this pathway be D.E. in our dataset.

When a pathway is mainly changing in one direction, with the exception of a single gene that is changing in the opposite direction, I only consider that gene to be informative if and only if it was represented in 2 samples or more. Why? Because false positives exist, but we also need to take into consideration that pathways are human constructs that are likely to be incomplete. Branching may be occurring, and there could be specific reasons for why a single node changes in opposite direction to the rest of the pathway.

```
In [1]: # important stuff:
import os
import pandas as pd
import numpy as np
```

```

# morgan
import morgan as morgan
import tissue_enrichment_analysis as tea

import matplotlib.pyplot as plt
import seaborn as sns

# Magic function to make matplotlib inline;
# other style specs must come AFTER
%matplotlib inline

# This enables SVG graphics inline.
# There is a bug, so uncomment if it works.
%config InlineBackend.figure_formats = {'png', 'retina'}

import genpy
import seqplotter
import gvars
import epistasis as epi

q = 0.1
genvar = gvars.genvars()

```

I will load the respiratory complexes and central dogma complexes, which I obtained from a manual curation of Wormbase using WormMine

```

In [2]: respiratory_complexes = pd.read_excel('../input/respiratory_complexes.xlsx')
        central_dogma = pd.read_excel('../input/central_dogma.xlsx')

In [3]: tissue_df = tea.fetch_dictionary()
        phenotype_df = pd.read_csv('../input/phenotype_ontology.csv')
        go_df = pd.read_csv('../input/go_dictionary.csv')

In [4]: melted_tissue = pd.melt(tissue_df, id_vars='wbid',
                                var_name='term', value_name='expressed')
        melted_tissue = melted_tissue[melted_tissue.expressed == 1]

        melted_phenotype = pd.melt(phenotype_df, id_vars='wbid',
                                    var_name='term', value_name='expressed')
        melted_phenotype = melted_phenotype[melted_phenotype.expressed == 1]

        melted_go = pd.melt(go_df, id_vars='wbid',
                             var_name='term', value_name='expressed')
        melted_go = melted_go[melted_go.expressed == 1]

In [5]: tidy_data = pd.read_csv('../output/temp_files/DE_genes.csv')
        tidy_data.sort_values('target_id', inplace=True)
        tidy_data.dropna(subset=['ens_gene'], inplace=True)
        # tidy_data.sort_values('sort_order', inplace=True)

```

```
# drop the fog-2 data:
tidy_data = tidy_data[tidy_data.genotype != 'fog-2']
# tidy_data = tidy_data[tidy_data.qual < q] # keep only sig data.
```

2 Defining a gene compactifier for easy printing

Before we start, I will define a function, called `gene_compactifier` which will make visualization of gene representation much easier. How does it work?

Given a gene list, it: 1. Finds all the genes that have the same WORM family name. In other words, find all the *unc* genes, all the *rpl* genes. 2. If there's more than one gene in a given family, print the number of genes that have that family name. 3. Print a list of all the suffixes.

So if a gene list contains *unc-119*, *unc-15* and *unc-1*, the program will output:

Gene "Family", Number Found unc, 3 ['1', '15', '119']

Moreover, if a gene list contains *unc-119*, *unc-119* and *unc-119* (the same gene repeated *n* times), the program will output:

Gene "Family", Number Found unc, 3 ['119', '119', '119']

This makes it quite easy to visualize what genes within a pathway are represented in all of the mutants (coverage), as well as how many times each gene is represented in the dataset (coverage).

```
In [6]: def gene_compactifier(ext_gene):
        """Given a list of ext_gene names, compactify them and print"""
        d = {}
        ext_gene = sorted(ext_gene)
        for gene in ext_gene:
            ind = gene.find('-')
            if ind > 1:
                name = gene[:ind]
                number = gene[ind+1:]
            else:
                name = gene
                number = ''

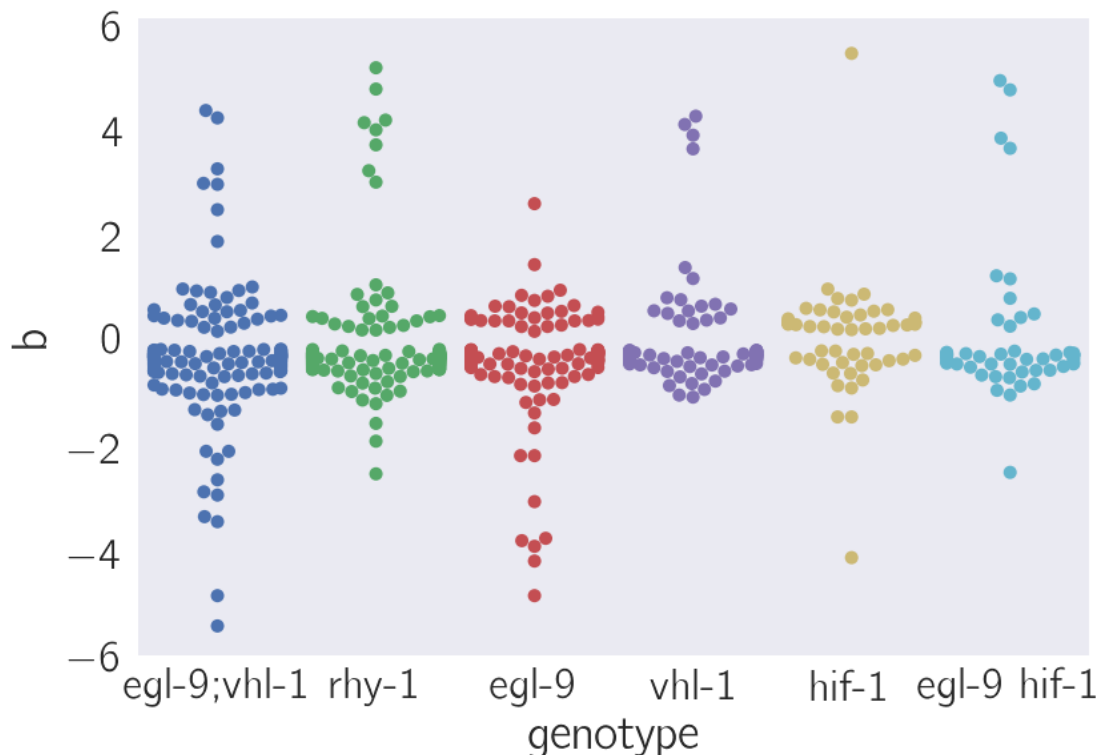
            if name in d.keys():
                d[name] += [number]
            else:
                d[name] = [number]

        print('Gene "Family", Number Found')
        for name, numbers in d.items():
            if len(numbers) > 1:
                print(name + ', ', len(numbers), sorted(numbers))
            else:
                if len(numbers[0]) > 0:
                    print(name + '-' + numbers[0])
                else:
                    print(name)
```

3 Effects of HIF-1 on mitochondrial proteins

First, let me extract all the genes that are overrepresented in mitochondria. The way I do this is via a function call `plot_by_term` which, given a string, a dataframe to search, and the kind of ontology that the string should be found in, plots for each genotype the perturbation values of the significantly altered genes and returns the axis that contains that plot, as well as the list of genes that are annotated with the desired string.

```
In [7]: ax, mito = seqplotter.plot_by_term('mitochondrion', df=tidy_data,
                                         kind='go', swarm=True)
```



Visually, it looks like maybe 1/3 of the mitochondrial genes go up, and the rest go down. What genes are most represented in this pathway? How many points consistently show up in mutants that have a constitutive HIF-1 response? Let's find out.

Next, I find out what genes that are annotated with the term 'mitochondria' go up across genotypes with a constitutive HIF-1 response:

```
In [8]: common = epi.find_overlap(['e', 'b', 'd', 'a'], tidy_data)
        trial = tidy_data[(tidy_data.ens_gene.isin(mito)) &
                           (tidy_data.target_id.isin(common)) &
                           (tidy_data.b > 0)].ext_gene
        gene_compactifier(trial)
```

Gene "Family", Number Found
F02A9.4

```

acl, 6 ['6', '6', '6', '6', '6', '6']
mdh-2
Y53G8AL.2
phb, 2 ['1', '2']
pcca, 5 ['1', '1', '1', '1', '1']
Y39E4A.3
F20D6.11
ZK669.4
wah-1
F53F4.10
fum, 6 ['1', '1', '1', '1', '1', '1']
nuo-1
B0272.3
suc1-1
timm-23
got-2.1
oxa-1
tomm-40
atp-5
sdha, 5 ['1', '1', '1', '1', '1']
mai, 5 ['1', '1', '1', '1', '1']

```

What about the genes that go DOWN in all genotypes with a constitutive HIF-1 response?

```

In [9]: trial = tidy_data[(tidy_data.ens_gene.isin(mito)) &
                           (tidy_data.target_id.isin(common)) &
                           (tidy_data.b < 0)].ext_gene
        gene_compactifier(trial)

```

```

Gene "Family", Number Found
F45H10.3, 6 ['', '', '', '', '', '']
C05C10.3, 6 ['', '', '', '', '', '']
ZK1320.9, 6 ['', '', '', '', '', '']
sco, 6 ['1', '1', '1', '1', '1', '1']
mdh, 5 ['2', '2', '2', '2', '2']
Y53G8AL.2, 5 ['', '', '', '', '']
acd, 6 ['1', '1', '1', '1', '1', '1']
Y39E4A.3, 5 ['', '', '', '', '']
wah, 5 ['1', '1', '1', '1', '1']
Y48A6B.3, 6 ['', '', '', '', '', '']
Y38F1A.6, 6 ['', '', '', '', '', '']
T27E9.2, 6 ['', '', '', '', '', '']
sucg, 6 ['1', '1', '1', '1', '1', '1']
phb, 10 ['1', '1', '1', '1', '1', '2', '2', '2', '2', '2']
pdhb, 6 ['1', '1', '1', '1', '1', '1']
sdha-1
cyc, 6 ['1', '1', '1', '1', '1', '1']

```

```

acaa, 6 ['2', '2', '2', '2', '2', '2']
F02A9.4, 11 ['', '', '', '', '', '', '', '', '', '', '']
T02G5.7, 6 ['', '', '', '', '', '']
mrps, 6 ['6', '6', '6', '6', '6', '6']
F54D5.12, 6 ['', '', '', '', '', '']
mtss, 6 ['1', '1', '1', '1', '1', '1']
Y54F10AM.5, 6 ['', '', '', '', '', '']
pcca-1
C14B9.10, 6 ['', '', '', '', '', '']
F20D6.11, 5 ['', '', '', '', '']
ZK669.4, 5 ['', '', '', '', '']
F56B3.11, 6 ['', '', '', '', '', '']
nduf, 6 ['7', '7', '7', '7', '7', '7']
F53F4.10, 5 ['', '', '', '', '']
suc1, 5 ['1', '1', '1', '1', '1']
R07E5.13, 6 ['', '', '', '', '', '']
B0272.3, 5 ['', '', '', '', '']
nuo, 11 ['1', '1', '1', '1', '1', '1', '6', '6', '6', '6', '6', '6']
timm, 5 ['23', '23', '23', '23', '23']
got, 5 ['2.1', '2.1', '2.1', '2.1', '2.1']
oxa, 5 ['1', '1', '1', '1', '1']
tomm, 11 ['22', '22', '22', '22', '22', '22', '40', '40', '40', '40', '40']
atp, 5 ['5', '5', '5', '5', '5']
C25H3.9, 6 ['', '', '', '', '', '']
hsp, 6 ['60', '60', '60', '60', '60', '60']
mrpl, 24 ['2', '2', '2', '2', '2', '2', '2', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47', '47']
mai-1
F09F7.4, 6 ['', '', '', '', '', '']

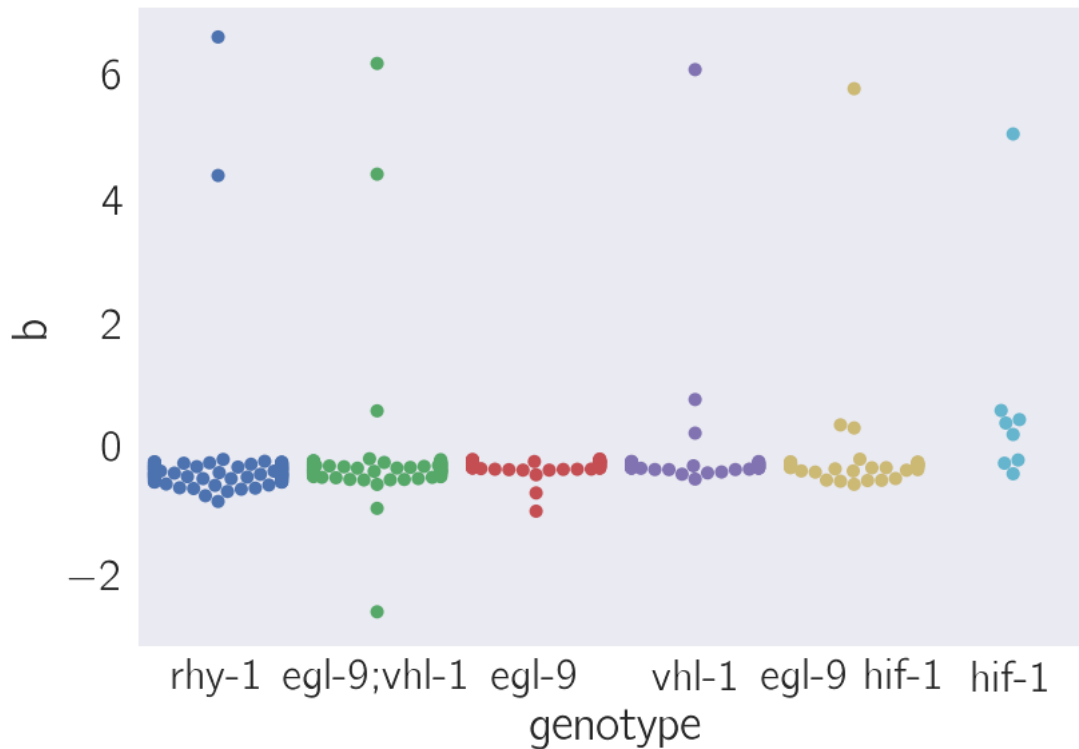
```

4 HIF-1 effects on the ribosome

```

In [10]: term = 'structural constituent of ribosome GO:0003735'
         ax, ribosome = seqplotter.plot_by_term(term, df=tidy_data, kind='go')

```



```
In [11]: trial = tidy_data[(tidy_data.ens_gene.isin(ribosome)) & (tidy_data.qval < q)].ext_gene
gene_compactifier(trial)
```

Gene "Family", Number Found

dap-3

F54D7.6

rps, 28 ['1', '10', '11', '12', '13', '14', '15', '16', '17', '18', '19', '20', '21', '22', '23', '24', '25', '26', '27', '28']

ubq-2

mrps, 16 ['12', '14', '15', '16', '17', '18B', '18C', '21', '22', '23', '24', '25', '34', '6']

C37A2.7

ubl-1

rpl, 41 ['1', '10', '11.1', '11.2', '12', '13', '14', '15', '16', '17', '18', '19', '2', '20', '21', '22', '23', '24', '25', '26', '27', '28', '29', '30', '31', '32', '33', '34', '35', '36', '37', '38', '39', '40', '41']

W01D2.1

F54D7.7

T07A9.14

rla, 3 ['0', '1', '2']

mrpl, 18 ['10', '11', '12', '13', '15', '16', '17', '19', '2', '23', '24', '32', '34', '41', '42', '43', '44', '45', '46', '47', '48', '49', '50', '51', '52', '53', '54', '55', '56', '57', '58', '59', '60', '61', '62', '63', '64', '65', '66', '67', '68', '69', '70', '71', '72', '73', '74', '75', '76', '77', '78', '79', '80', '81', '82', '83', '84', '85', '86', '87', '88', '89', '90', '91', '92', '93', '94', '95', '96', '97', '98', '99', '100']

Y37E3.8

5 Bioenergetics of HIF-1

What about the effects of HIF-1 on the Electron Transport Chain? Or the TCA cycle?

To explore this, I will make a new dataframe, that contains only the genes in the ETC. I will also annotate each gene with the complex it belongs to, and then I will add a column called `sort_order` so I can sort the dataframe at my pleasure.

```
In [12]: resp = tidy_data[tidy_data.ens_gene.isin(respiratory_complexes.ens_gene)
          & (~tidy_data.code.isin(['f', 'c']))].copy()

f = lambda x: respiratory_complexes[respiratory_complexes.ens_gene == x].complex.values
resp['complex'] = resp.ens_gene.map(f)

f = lambda x: respiratory_complexes[respiratory_complexes.ens_gene == x].sort_order.values
resp['sort_order'] = resp.ens_gene.map(f)
resp.sort_values('sort_order', inplace=True)
resp = resp[resp.complex != 'Ubiquinone Biosynthesis']
```

This is what the dataframe looks like:

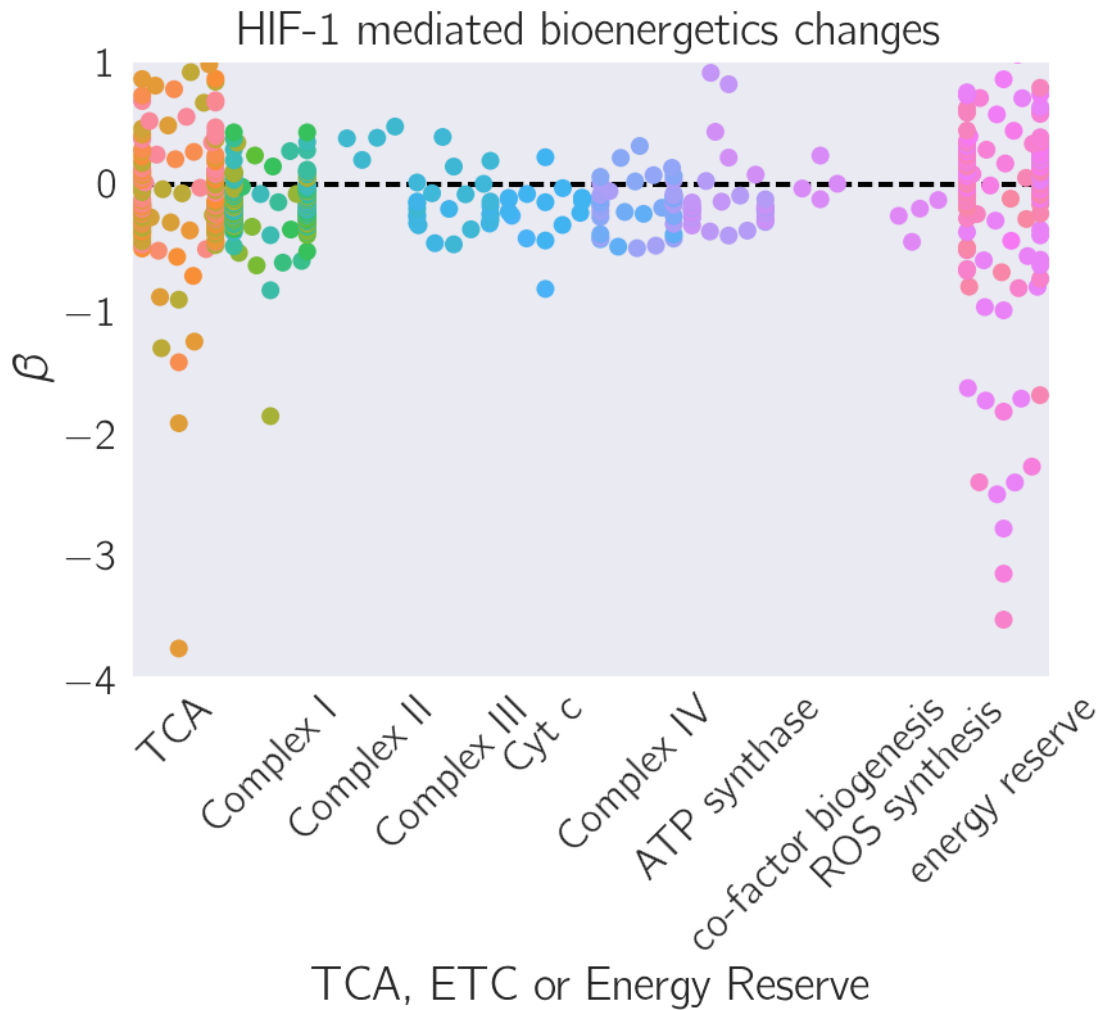
```
In [13]: resp[['ext_gene', 'genotype', 'complex', 'sort_order']].head()
```

```
Out[13]:
```

	ext_gene	genotype	complex	sort_order
88496	fum-1	vhl-1	TCA	0
124985	sdhd-1	egl-9;vhl-1	TCA	0
85633	sdhd-1	vhl-1	TCA	0
26605	sdhd-1	rhy-1	TCA	0
124986	sdhd-1	egl-9;vhl-1	TCA	0

Let's plot the dataframe, see what comes out. We would expect all genes in the ETC and TCA to go down:

```
In [14]: fig, ax = plt.subplots()
ax = sns.swarmplot(x='complex', y='b', hue='ens_gene', data=resp, size=7)
plt.xticks(rotation=45)
ax.legend_.remove()
plt.title('HIF-1 mediated bioenergetics changes')
plt.ylabel(r'\beta')
plt.xlabel('TCA, ETC or Energy Reserve')
ax.hlines(0, xmin=-2, xmax=10, lw=2, linestyle='--')
plt.ylim(-4, 1)
plt.savefig('../output/mito_function.pdf')
```

Well, we can definitely see that not all genes in the ETC and TCA go down. Let's figure out what genes go UP in each cycle/complex.

```
In [15]: gene_compactifier(resp[(resp.complex == 'TCA') & (resp.b > 0)].ext_gene)
```

```
Gene "Family", Number Found
fum, 13 ['1', '1', '1', '1', '1', '1', '1', '1', '1', '1', '1', '1', '1']
sdhd, 6 ['1', '1', '1', '1', '1', '1']
sdhb, 3 ['1', '1', '1']
mdh, 8 ['1', '1', '1', '1', '1', '1', '1', '1']
ZK836.2, 6 ['', '', '', '', '', '']
ogdh, 7 ['1', '1', '1', '1', '1', '1', '1']
idhb, 3 ['1', '1', '1']
idhg, 4 ['1', '1', '1', '1']
sdha, 4 ['2', '2', '2', '2']
idh, 12 ['1', '1', '1', '1', '1', '1', '1', '1', '1', '1', '2', '2']
suc1, 4 ['2', '2', '2', '2']
```

```
aco, 8 ['2', '2', '2', '2', '2', '2', '2', '2']
```

```
In [16]: gene_compactifier(resp[(resp.complex == 'Complex I') & (resp.b > 0)].ext_gene)
```

```
Gene "Family", Number Found
```

```
F44G4.2
```

```
C16A3.5, 4 ['', '', '', '']
```

```
lpd, 2 ['5', '5']
```

```
nuo, 14 ['2', '2', '2', '2', '3', '3', '3', '3', '3', '3', '3', '3', '3', '3']
```

```
C25H3.9, 2 ['', '']
```

```
nduo, 4 ['3', '4', '5', '5']
```

```
ndfl, 3 ['4', '4', '4']
```

```
Y51H1A.3, 5 ['', '', '', '', '']
```

```
In [17]: gene_compactifier(resp[(resp.complex == 'Complex II') & (resp.b > 0)].ext_gene)
```

```
Gene "Family", Number Found
```

```
sdha, 4 ['1', '1', '1', '1']
```

```
In [18]: gene_compactifier(resp[(resp.complex == 'energy reserve') & (resp.b > 0)].ext_gene)
```

```
Gene "Family", Number Found
```

```
Y67D8A.2, 15 ['', '', '', '', '', '', '', '', '', '', '', '', '', '']
```

```
agl, 4 ['1', '1', '1', '1']
```

```
T22F3.3, 16 ['', '', '', '', '', '', '', '', '', '', '', '', '', '']
```

```
T04A8.7, 6 ['', '', '', '', '']
```

```
gyg, 8 ['1', '1', '1', '1', '1', '1', '1', '1']
```

```
gsy-1
```

```
CC8.2, 8 ['', '', '', '', '', '', '']
```

```
aagr, 9 ['1', '1', '1', '1', '3', '3', '3', '3', '3']
```

```
Y50D7A.3, 2 ['', '']
```

```
oga, 4 ['1', '1', '1', '1']
```

```
H18N23.2, 5 ['', '', '', '']
```

```
ogt, 7 ['1', '1', '1', '1', '1', '1', '1']
```

Notice that for complex I, the genes *nuo-2* and *nduo-4* are up-regulated. But those exact same genes are also down-regulated (see below). Therefore, there is insufficient information to conclude whether these genes are going up, or down as a result of HIF-1. However, for other genes, namely *fum-1* and *sdha-1* we can conclude that those are significantly and consistently up-regulated in mutants that have a constitutive HIF-1 mutant.

```
In [19]: gene_compactifier(resp[(resp.complex == 'TCA') & (resp.b < 0)].ext_gene)
```

```
Gene "Family", Number Found
```

```
sucg, 4 ['1', '1', '1', '1']
```

```
fum, 3 ['1', '1', '1']
sdhb, 5 ['1', '1', '1', '1', '1']
suca, 4 ['1', '1', '1', '1']
sdhd, 2 ['1', '1']
mdh, 8 ['1', '1', '1', '1', '2', '2', '2', '2']
ZK836.2, 6 ['', '', '', '', '', '']
dlst, 4 ['1', '1', '1', '1']
ogdh, 5 ['1', '1', '1', '1', '1']
idha, 4 ['1', '1', '1', '1']
idhb, 5 ['1', '1', '1', '1', '1']
sdha, 4 ['2', '2', '2', '2']
icl, 4 ['1', '1', '1', '1']
idhg, 8 ['1', '1', '1', '1', '2', '2', '2', '2']
idh, 12 ['1', '1', '1', '1', '1', '1', '1', '1', '2', '2', '2', '2']
suc1, 8 ['1', '1', '1', '1', '2', '2', '2', '2']
cts, 4 ['1', '1', '1', '1']
aco, 12 ['1', '1', '1', '1', '2', '2', '2', '2', '2', '2', '2', '2']
```

```
In [20]: gene_compactifier(resp[(resp.complex == 'Complex I') & (resp.b < 0)].ext_gene)
```

Gene "Family", Number Found

```
djr, 4 ['1.1', '1.1', '1.1', '1.1']
F53F4.10, 4 ['', '', '', '']
C18E9.4, 4 ['', '', '', '']
F45H10.3, 4 ['', '', '', '']
nduo, 16 ['1', '1', '1', '1', '2', '2', '2', '2', '3', '3', '3', '4', '4', '4', '5', '5']
C16A3.5, 4 ['', '', '', '']
nuo, 30 ['1', '1', '1', '1', '2', '2', '2', '2', '3', '3', '3', '3', '3', '3', '3', '3', '3', '3', '3', '3']
Y69A2AR.3, 4 ['', '', '', '']
T20H4.5, 4 ['', '', '', '']
nduf, 20 ['5', '5', '5', '5', '6', '6', '6', '6', '6', '6', '6', '6', '7', '7', '7', '7', '7', '7']
Y53G8AL.2, 4 ['', '', '', '']
F44G4.2, 3 ['', '', '']
Y54F10AM.5, 4 ['', '', '', '']
lpd, 6 ['5', '5', '5', '5', '5', '5']
gas, 4 ['1', '1', '1', '1']
C25H3.9, 6 ['', '', '', '', '', '']
Y63D3A.7, 4 ['', '', '', '']
C33A12.1, 4 ['', '', '', '']
F59C6.5, 4 ['', '', '', '']
Y51H1A.3, 7 ['', '', '', '', '', '', '']
ndf1-4
```

```
In [21]: gene_compactifier(resp[(resp.complex == 'energy reserve') & (resp.b < 0)].ext_gene)
```

Gene "Family", Number Found

```
Y67D8A.2, 17 ['', '', '', '', '', '', '', '', '', '', '', '', '', '', '', '']
```

```

R05F9.6, 4 ['', '', '', '']
C14B9.8, 4 ['', '', '', '']
aagr, 19 ['1', '1', '1', '1', '1', '1', '1', '1', '1', '2', '2', '2', '2', '3', '3', '3', '3', '3']
Y50D7A.3, 2 ['', '']
oga, 4 ['1', '1', '1', '1']
T04A8.7, 2 ['', '']
H18N23.2, 3 ['', '', '']
gsy, 3 ['1', '1', '1']
ogt-1

```

6 Effects of HIF-1 on the Proteasome and Mediator

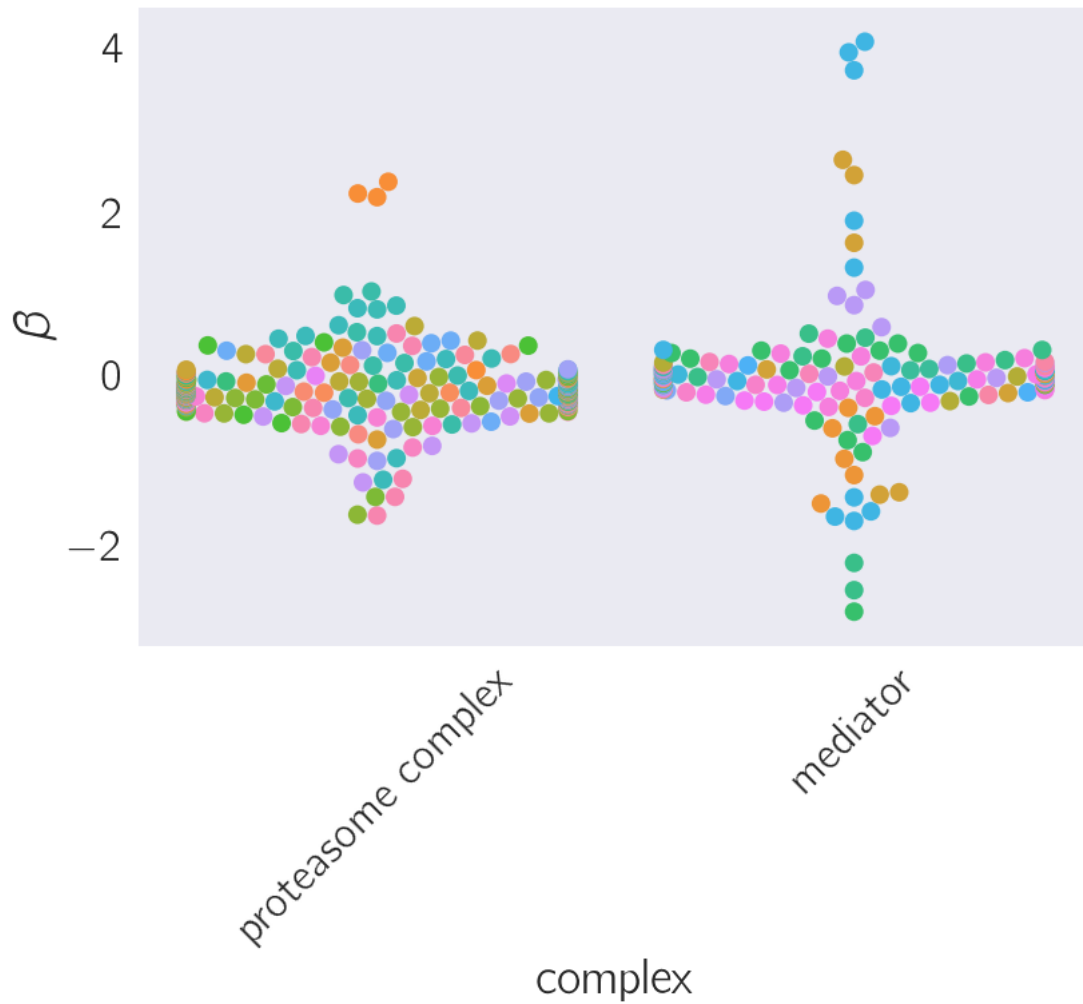
```

In [22]: prot = tidy_data[tidy_data.ens_gene.isin(central_dogma.ens_gene)].copy()
         prot['complex'] = prot.ens_gene.map(lambda x: central_dogma[central_dogma.ens_gene ==

In [23]: fig, ax = plt.subplots()
         ax = sns.swarmplot(x='complex', y='b', hue='ens_gene', data=prot, size=7)
         plt.xticks(rotation=45)
         ax.legend_.remove()
         # plt.title('HIF-1 mediated changes in ETC expression')
         plt.ylabel(r'\beta')
         # plt.xlabel('Electron Transport Chain Complexes')

Out[23]: <matplotlib.text.Text at 0x10bb912b0>

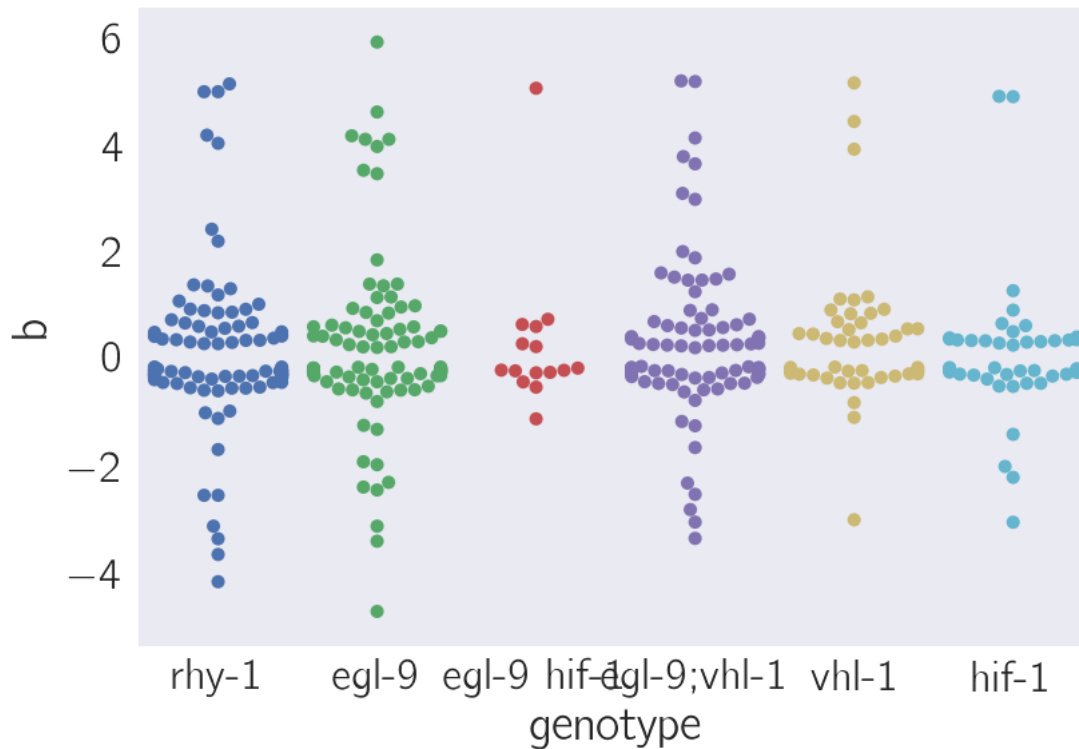
```



6.1 Effect of HIF-1 on proteins involved in 'protein catabolic process'

This GO term includes proteins that are involved in protein degradation, including the proteasome, a variety of ubiquitin-related enzymes and proteases

```
In [24]: ax, negregproteolysis = seqplotter.plot_by_term('protein catabolic process GO:0030163
                                                         df=tidy_data, kind='go')
```



```
In [25]: temp = tidy_data[(tidy_data.ens_gene.isin(negregproteolysis)) &
                           (tidy_data.target_id.isin(common)) &
                           (tidy_data.b > 0)
                           ].ext_gene.unique()
gene_compactifier(temp)
```

```
Gene "Family", Number Found
asp,  3 ['14', '5', '8']
ctsa-2
Y119C1B.5
cpr,  3 ['1', '3', '6']
rpt-6
uev-3
K10C2.1
zyx-1
cpz-1
mans,  2 ['3', '4']
F57F5.1
aex-5
```

```
In [26]: temp = tidy_data[(tidy_data.ens_gene.isin(negregproteolysis)) &
                           (tidy_data.target_id.isin(common)) &
```

```

        (tidy_data.b < 0)
    ].ext_gene.unique()
gene_compactifier(temp)

```

Gene "Family", Number Found

asp, 2 ['14', '8']

pas-1

unc-60

ubq-2

ctsa-2

Y119C1B.5

cpr, 2 ['1', '6']

rpt-6

uev-3

ubh-3

pbs-4

rpn-3

cpz-1

mans-3

ubc-20

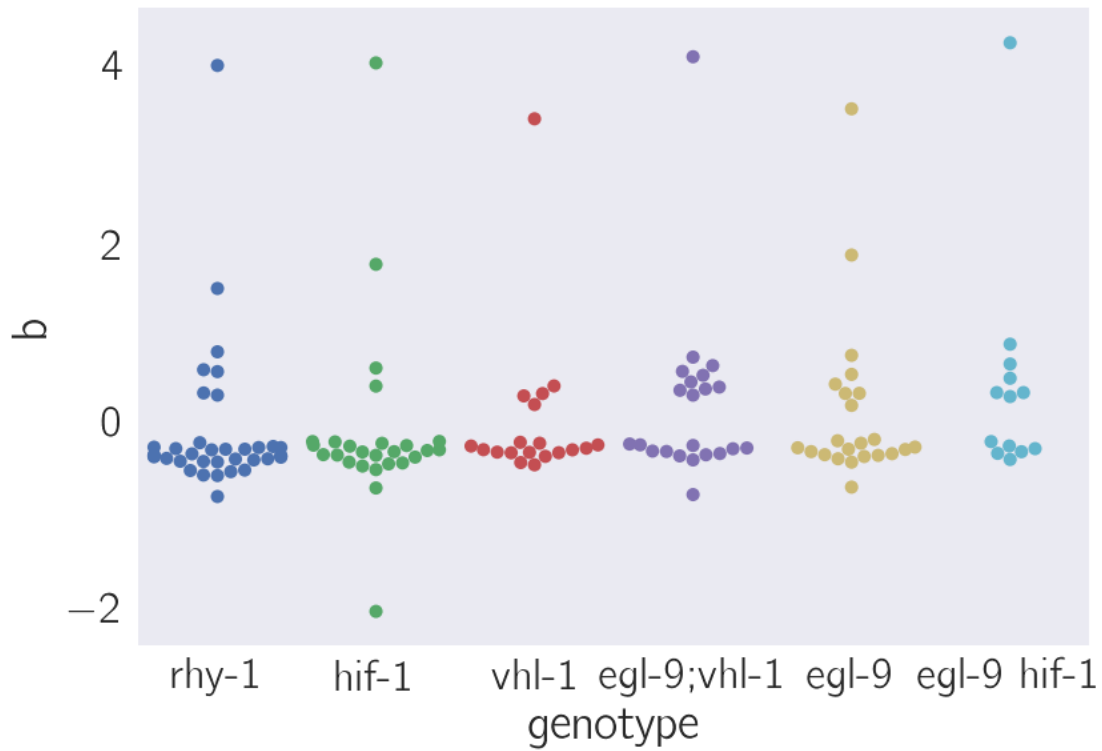
F57F5.1

Y66D12A.9

cpl-1

7 Proteins annotated as involved in protein folding

```
In [27]: ax, folding = seqplotter.plot_by_term('protein folding', df=tidy_data, kind='go')
```



```
In [28]: temp = tidy_data[(tidy_data.ens_gene.isin(folding)) & (tidy_data.b > 0)].ext_gene.unique()
gene_compactifier(temp)
```

Gene "Family", Number Found

pdi-6

dnj, 5 ['10', '12', '15', '20', '27']

C06A6.5

fkf, 6 ['1', '3', '4', '5', '7', '8']

C03H12.1

C34C12.8

F47B7.2

dpy-11

emc, 2 ['1', '3']

C30H7.2

ZC250.5

hsp, 3 ['6', '60', '75']

M04D5.1

enpl-1

uggt, 2 ['1', '2']

crt-1

C14B9.2

ZK973.11

K07E8.6

W01B11.6
 tbcc-1
 pfd, 3 ['1', '3', '6']
 unc-23
 sig-7
 catp-6
 cyn, 10 ['10', '12', '13', '15', '2', '4', '5', '6', '8', '9']
 cdc-37
 Y17G9B.4
 F53A3.7
 cnx-1
 ooc-5
 cct, 4 ['4', '5', '6', '8']
 Y71F9AL.11
 daf-21
 F42G8.7
 Y22D7AL.10
 T10H10.2
 ero-1
 trx, 2 ['2', '4']
 R05D3.9
 F35G2.1
 Y49E10.4

```
In [29]: temp = tidy_data[(tidy_data.ens_gene.isin(folding)) & (tidy_data.b < 0)].ext_gene.uniq
         gene_compactifier(temp)
```

Gene "Family", Number Found

pdi-6
 C06A6.5
 C03H12.1
 C34C12.8
 dpy-11
 emc, 3 ['1', '3', '6']
 C30H7.2
 enpl-1
 C05G5.3
 K07E8.6
 fkb, 6 ['1', '2', '3', '5', '6', '7']
 pfd, 6 ['1', '2', '3', '4', '5', '6']
 unc-23
 catp-6
 tbcc-1
 txl-1
 tbcd-1
 F53A3.7
 tbca-1

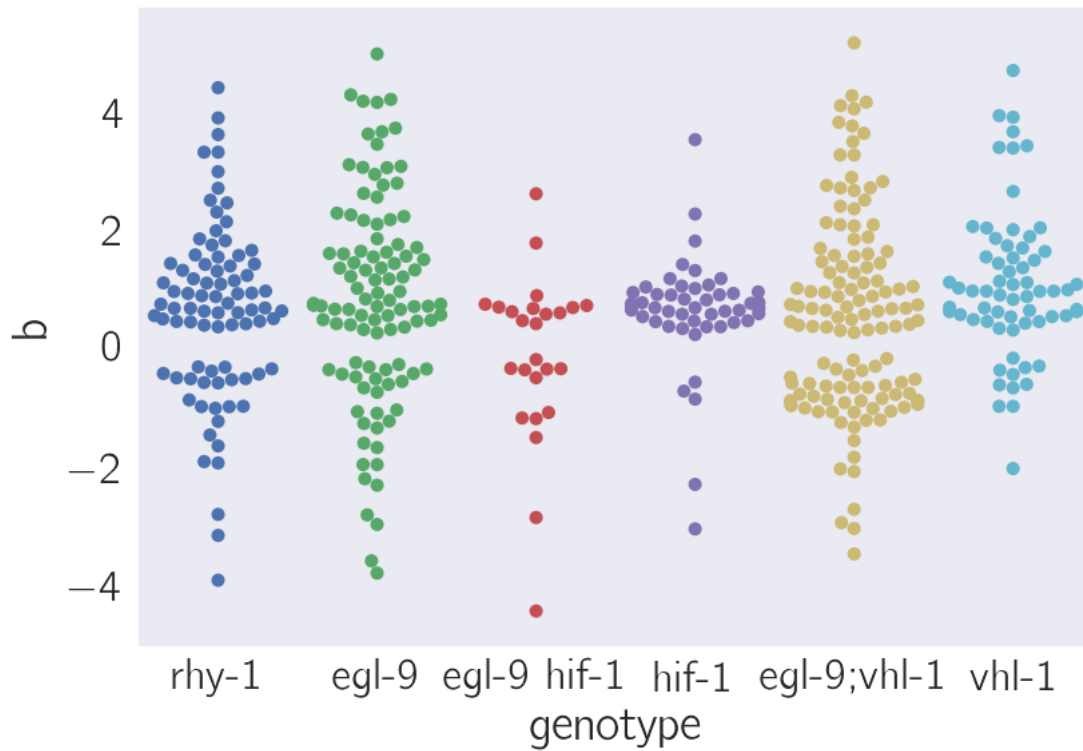
```

cct, 7 ['1', '3', '4', '5', '6', '7', '8']
Y71F9AL.11
daf-21
bag-1
Y22D7AL.10
trx, 2 ['2', '4']
dnj, 7 ['10', '12', '13', '15', '19', '20', '27']
nud-1
F47B7.2
ero-1
ZC250.5
crt-1
C14B9.2
ZK973.11
W01B11.6
Y55F3AR.2
sig-7
cdc-37
uggt, 2 ['1', '2']
Y17G9B.4
ooc-5
cyn, 15 ['1', '10', '11', '12', '13', '15', '16', '2', '3', '4', '5', '6', '7', '8', '9']
F42G8.7
hsp, 3 ['6', '60', '75']
cnx-1
R05D3.9
F35G2.1

```

8 Immune Involvement

```
In [30]: ax, immune = seqplotter.plot_by_term('immune system process', df=tidy_data, kind='go')
```



```
In [31]: temp = tidy_data[(tidy_data.ens_gene.isin(immune)) & (tidy_data.target_id.isin(common)
                        (tidy_data.b > 0)].ext_gene.unique()
gene_compactifier(temp)
```

Gene "Family", Number Found

```
asp-14
T24B8.5
aqp-10
C17H12.8
lec-11
C25D7.5
F35E12.9
clec, 4 ['210', '66', '70', '72']
lys, 2 ['2', '7']
fat-3
cyp-35A5
C49C3.9
tag-244
F55G11.8
nhr-57
dod, 2 ['22', '24']
cpr-3
Y41D4B.17
```

F01D5.1
F01D5.5
dct-17
his-10
C34H4.2
gst-7
F55G11.2
F53A9.6
K08D8.4

```
In [32]: temp = tidy_data[(tidy_data.ens_gene.isin(immune)) & (tidy_data.target_id.isin(common)
                           (tidy_data.b < 0)].ext_gene.unique()
        gene_compactifier(temp)
```

Gene "Family", Number Found
F55G11.8
acdh-1
asp-14
lys-7
aqp-10
clec, 2 ['210', '72']
F55G11.2
dod-24
nhr-57
cyp-35A5