AnnotatedFeature

- Simple ‘Processed’ Feature
  - *i.e.* a peak call from ChIP-Seq
  - Also inherits methods from Feature
- Used as evidence for Regulatory Features. Fetched via:
  - `$regulatory_feature->regulatory_attributes('annotated')`
- Some key attributes/methods:

<table>
<thead>
<tr>
<th>Method</th>
<th>Example value(s)</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>display_label</td>
<td>CTCF – IMR90 Enriched Site</td>
<td><code>$af-&gt;display_label</code></td>
</tr>
<tr>
<td>score</td>
<td>Value</td>
<td><code>$af-&gt;score</code></td>
</tr>
<tr>
<td>summit</td>
<td>Coordinate</td>
<td><code>$af-&gt;summit</code></td>
</tr>
<tr>
<td>get_associated_MotifFeatures</td>
<td>List of MotifFeature objects</td>
<td><code>$af-&gt;get_associated_MotifFeatures</code></td>
</tr>
</tbody>
</table>
Example: What AnnotatedFeatures support this RegulatoryFeature?

#This gets the 'MultiCell' Regulatory Feature
my $rf =
   $regfeat_adaptor->fetch_by_stable_id('ENSR00000165384');

my @annotated_features =
   @{$rf->regulatory_attributes('annotated')};

#Get annotated features supporting a regulatory feature
foreach my $annotated_feature (@annotated_features) {
   print $annotated_feature->display_label."\t";
   print $annotated_feature->score."\t";
   print $annotated_feature->summit."\t";
}
Example: What AnnotatedFeatures support this RegulatoryFeature?

```
#This gets the 'MultiCell' regulatory feature
my $rf = $regfeat_adaptor->fetch_by_stable_id('ENSR00000165384');

my @annotated_features = @{$rf->regulatory_attributes('annotated')};

#Get annotated features supporting a regulatory feature
foreach my $annotated_feature (@annotated_features) {
    print $annotated_feature->display_label.
    print $annotated_feature->score.
    print $annotated_feature->summit.
}
```

Example: What AnnotatedFeatures support this RegulatoryFeature?

<table>
<thead>
<tr>
<th>Annotated Feature</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNase1 - HUVEC Enriched Site</td>
<td>40.489637</td>
</tr>
<tr>
<td>DNase1 - K562 Enriched Site</td>
<td>27.862417</td>
</tr>
<tr>
<td>DNase1 - H1ESC Enriched Site</td>
<td>26.207104</td>
</tr>
<tr>
<td>CTCF - HUVEC Enriched Site</td>
<td>125.020915</td>
</tr>
<tr>
<td>DNase1 - H1ESC Enriched Site</td>
<td>107.692877</td>
</tr>
<tr>
<td>CTCF - HMEC Enriched Site</td>
<td>84.500112</td>
</tr>
<tr>
<td>DNase1 - HSMM Enriched Site</td>
<td>17.812017</td>
</tr>
<tr>
<td>DNase1 - H1ESC Enriched Site</td>
<td>107.692877</td>
</tr>
<tr>
<td>CTCF - CD4 Enriched Site</td>
<td>12.242441</td>
</tr>
<tr>
<td>CTCF - HepG2 Enriched Site</td>
<td>36.243751</td>
</tr>
<tr>
<td>CTCF - NHEK Enriched Site</td>
<td>156.68494</td>
</tr>
<tr>
<td>DNase1 - K562 Enriched Site</td>
<td>19.88997</td>
</tr>
<tr>
<td>CTCF - HUVEC Enriched Site</td>
<td>47.901836</td>
</tr>
<tr>
<td>Rad21 - K562 Enriched Site</td>
<td>33.054809</td>
</tr>
<tr>
<td>CTCF - K562 Enriched Site</td>
<td>29.488397</td>
</tr>
<tr>
<td>CTCF - NHEK Enriched Site</td>
<td>21.50665</td>
</tr>
<tr>
<td>CTCF - H1ESC Enriched Site</td>
<td>22.526458</td>
</tr>
<tr>
<td>CTCF - NHEK Enriched Site</td>
<td>84.547054</td>
</tr>
<tr>
<td>CTCF - K562 Enriched Site</td>
<td>85.822038</td>
</tr>
<tr>
<td>DNase1 - NHEK Enriched Site</td>
<td>26.878816</td>
</tr>
<tr>
<td>DNase1 - HepG2 Enriched Site</td>
<td>17.995201</td>
</tr>
</tbody>
</table>

EMBL-EBI
MotifFeature

- BindingMatrix (PWM) match
  - Also inherits generic Feature methods (eg. seq)
- Fetched via:
  - $regulatory_feature->regulatory_attributes('motif')
  - $annotated_feature->get_associated_MotifFeatures()
- Some Key attributes / methods:

<table>
<thead>
<tr>
<th>Method</th>
<th>Example value(s)</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>display_label</td>
<td>Egr1:PB0010.1</td>
<td>$mf-&gt;display_label()</td>
</tr>
<tr>
<td>binding_matrix</td>
<td>MA0060.1</td>
<td>$mf-&gt;binding_matrix-&gt;name()</td>
</tr>
<tr>
<td>score</td>
<td>0.81</td>
<td>$mf-&gt;score()</td>
</tr>
<tr>
<td>associated_annotated_features</td>
<td>List of AnnotatedFeature objects</td>
<td>$mf-&gt;associated_annotated_features()</td>
</tr>
</tbody>
</table>
Example: What MotifFeatures support this RegulatoryFeature?

#This gets the 'MultiCell' Regulatory Feature
my $rf =
    $regfeat_adaptor->fetch_by_stable_id('ENSR00000165384');

my @motif_features = @{$rf->regulatory_attributes('motif')};

#Get motif features supporting a regulatory feature
foreach my $motif_feature (@motif_features) {
    print $motif_feature->display_label."\t";
    print $motif_feature->binding_matrix->name."\t";
    print $motif_feature->score."\n";

    foreach my $ann_feat (@{$motif_feature->associated_annotated_features()}) {
        print $ann_feat->display_label."\t";
        print $ann_feat->score."\t";
        print $ann_feat->summit."\n";
    }
}

Example: What MotifFeatures support this RegulatoryFeature?

#This gets the 'MultiCell' Regulatory Feature
my $rf = $regfeat_adaptor->fetch_by_stable_id('ENSR00000165384');

my @motif_features = @{$rf->regulatory_attributes('motif')};

#Get motif features supporting a regulatory feature
foreach my $motif_feature (@motif_features)
{
    print $motif_feature->display_label.
    print $motif_feature->binding_matrix->name.
    print $motif_feature->score.

    foreach my $ann_feat (@{$motif_feature->associated_annotated_features()})
    {
        print $ann_feat->display_label.
        print $ann_feat->score.
        print $ann_feat->summit.
    }

    print
}

Example: What MotifFeatures support this RegulatoryFeature?
Feature Exercises

1. Annotated Features
Annotated Features represents the results of an analysis of raw or processing signal data. These correspond to regions in the genome enriched for specific events (like TF binding or Histone Marks) i.e. they are 'peak calls'.

Compare the number of annotated features in the region Y:5000000-40000000 between the Human feature sets:

   K562_DNase1_ENCODE_Duke_SWEmbl_R0025_D150
   HepG2_DNase1_ENCODE_Duke_SWEmbl_R0025_D150

What are the differences and why?

2. Motif Features
Motif features represent putative binding sites based on alignments of PWMs from JASPAR. MotifFeatures are always associated to AnnotatedFeatures representing Transcription Factor (TF) Binding. More information about how we integrate these into the regulatory build process can be found here.

Get the 'motif' regulatory attributes associated to the Human Regulatory Feature 'ENSR00001227187'. Print their properties.

Hint: use 'motif' as a parameter for regulatory_attributes.

Print the properties of the annotated features associated to the motif feature.

3. Binding Matrices and motif strength
Each MotifFeature is associated with a PWM, which are represented by the 'BindingMatrix' class. The MotifFeature score represents the relative binding affinity with respect to the PWM defined in the BindingMatrix.

Using the Motif feature obtained in exercise 2, get the associated Binding Matrix and print some details.

Check potential effect of changes in the sequence of the motif feature on the relative strength of that motif feature.

Check the GERP conservation scores along the motif. Compare with the JASPAR matrix.