Well known adaptive radiation and model system in evolutionary biology
The Challenge

Spatial information
(historical samples, coordinate information, habitat characteristics)

Phenotypic Diversity
[EAV terms, wing phenotypes (quantitative descriptions, images), pedigrees, other morphological traits]

Genomic/Genetic Diversity
(linkage/QTL maps, array data, *Ests*, *SNP*, BAC sequence)
(more BAC sequence, genome)

How to link information on genomic and phenotypic diversity
Morphological similarity is not always a good predictor of genomic similarity (and vica versa)!
What emerged for the mini-GMOD meeting in January

**Diversity rules:** Most of the MODs are tackling similar issues.

**We can play too:** The evolutionary community can help.

**We should play:** It is simply makes good sense to utilize existing resources and to help develop new ones.
Our task

Explore the utility of GMOD tools for non-genome enabled organisms.

Understand Chado (both its strengths and weaknesses).
Our task (needs analysis)

We understood Chado at a very general level but needed to understand it at a very detailed level.

We understood our schema needs at a very general level but needed to understand it at a very detailed level.
Our working model evolutionary schema

90% chadoized (?)

90% complete (?)
Why not just go with GDPDM?

Interpreting a schema can be a big job for a biologist that doesn’t understand relational databases and a computer scientist that doesn’t understand biology.
Major join tables that make associations genotype and phenotype associations.
Genotype_experiment

Snap shot of genotype experiment
**Individual_phenotype**

Record of phenotypic information on an individual

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>individual_phenotype_id</td>
<td>integer</td>
<td>[PK]</td>
</tr>
<tr>
<td>individual_id</td>
<td>integer</td>
<td>[FK]</td>
</tr>
<tr>
<td>ptassay_id</td>
<td>integer</td>
<td>[PK]</td>
</tr>
<tr>
<td>phenotype_id</td>
<td>integer</td>
<td></td>
</tr>
<tr>
<td>experimenter_id</td>
<td>integer</td>
<td>[FK]</td>
</tr>
<tr>
<td>assay_date</td>
<td>date</td>
<td></td>
</tr>
<tr>
<td>phenotype_published_id</td>
<td>integer</td>
<td></td>
</tr>
<tr>
<td>comment</td>
<td>varchar(255)</td>
<td></td>
</tr>
</tbody>
</table>
Stock and biotype phenotype

Deal with different aspects of phenotype and phenotype experiments
What next?
What next?

Is this an efficient strategy?

Or should we stick with individual, taxon specific schema and develop middleware and mapping tools to integrate them?