Research in the Lawton-Rauh Lab: Plant Population Genetics

*Arabidopsis* sister species and the Hawaiian silversword alliance

**Brief overview:**

‘Comparative Molecular Population Genetics & Molecular Evolution’

...tools to investigate population dynamics and gene-specific processes

Amy Lawton-Rauh
Dept. Genetics & Biochemistry
Clemson University
AmyLR@clemson.edu
www.clemson.edu/~amylr
Research in the Lawton-Rauh Lab: Plant Population Genetics

*Arabidopsis* sister species and the Hawaiian silversword alliance

**Brief overview:**
‘Comparative Molecular Population Genetics & Molecular Evolution’
*tools to investigate population dynamics and gene-specifc processes*
European *Arabidopsis lyrata* (A. petraea):

- Glacial tills in Scandinavia: Vatnajokull glacier Iceland (photo by Ramos-Onsins)
- On isolated Dolomitic rock outcrops in Central Europe: Harz Mtns, Germany, Franconian Alps, Germany
European *Arabidopsis lyrata* (*A. petraea*):

- **glacial tills in Scandinavia**
  - Vatnajokull glacier
  - Iceland
  - (photo by Ramos-Onsins)

- **on isolated Dolomitic rock outcrops in Central Europe**
  - Harz Mtns, Germany
  - Franconian Alps, Germany
Relative Roles of Demography vs. Selection

*in shaping genetic variation: Arabidopsis petraea & Arabidopsis arenosa*

1. Frequency of unique polymorphisms increases with geographical distance between populations

2. Frequency distributions of sequence diversity differ between populations with different postglacial histories

3. Between-species isolation by distance is correlated with higher levels of linkage disequilibrium than within-species isolation by distance

- Population, *A. petraea*
- Population, *A. arenosa*
- Population, hybrids of *A. petraea & A. arenosa*
Hawaiian Silversword Alliance: origin & radiation

Ancestor of Carlquistia muirii, Madia nutans

Ancestor of Anisocarpus scabridus

Silversword Alliance Ancestor

allopolyploidization => duplicate gene copies
What are the relative roles of genome dynamics (heredity), morphology (form), and physiology (function) in shaping the adaptive evolution in the Hawaiian silversword alliance?

What are the relative contributions of demography & selection in shaping genetic and genomic variation?

How many genes are involved in the morphological and physiological differentiation among Hawaiian silversword alliance species?
Mahalo nui loa (und Danke schön)!

Lawton-Rauh Lab Folks:
Cindy Climer, Ph.D. Student
HSA, Arabidopsis petraea-A.arenosa
Brad Rauh, Research Associate, Manager
HSA, Madia, Arabidopsis petraea
Margaret Beaudrot
Clemson, HHMI/SCLife Honors Intern
Katie Horton
Clemson, Research Assistant
Chelsea Reighard
Clemson, HHMI/SCLife & Calhoun Intern
Jennifer Smith
Francis Marion Uni, NSF-REU Intern, S ‘07
Sarah Walker
Clemson, GEN-491 Intern 2006-07

Collaborators, HSA:
Bruce Baldwin
UC-Berkeley
Elizabeth Friar
Rancho Santa Ana Botanic Garden
David Remington
UNC-Greensboro
Rob Robichaux
University of Arizona

Collaborators, Arabidopsis sp.:
Juliette deMeaux
Max Planck Institute, Cologne
Marcus Koch
Universität Heidelberg
Research in the Lawton-Rauh Lab: Plant Population Genetics

*Arabidopsis* sister species and the Hawaiian silversword alliance

**Brief overview:**
‘Comparative Molecular Population Genetics & Molecular Evolution’
...tools to investigate population dynamics and gene-specific processes
‘Population Genomics’: Brief Introduction

relevant to:

association genetics
functional genomics
(human) disease mapping
crop science, breeding conservation genetics

Luikart et al. Nature Genetics. 2003
Expect some rare polymorphisms, some singletons

“singleton”: single nucleotide polymorphisms found in 1 individual
Allele/haplotype segregation patterns differ among populations
Demography

- Population expansion
- Population contraction (bottleneck)
- Population subdivision
- Gene flow, migration

...affects the entire genome (all loci at the same time)
Genetic variation is shaped by many factors

Nucleotide Diversity ($\theta$) = 4 $N_e \mu$

- Genetic Drift
- Gene Flow
- Inbreeding Depression
- Population Subdivision
- Population Bottleneck
- Population Expansion

Habitat Fragmentation
Interspecific Competition
Large-scale Events

Effective Population Size ($N_e$)
Non random association of alleles...

Linkage disequilibrium and selection for linked loci?

Migration of unique alleles that appear to be linked (via gene flow)?
Linkage Disequilibrium (LD): Non-random association among alleles

**Diagram:**

(a) Disequilibrium

<table>
<thead>
<tr>
<th>Locus 2</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

(b) Equilibrium

<table>
<thead>
<tr>
<th>Locus 2</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

Linkage Disequilibrium Influences Nucleotide Sequence Diversity

Linkage Disequilibrium:
Non-random association among loci/sites
Depends on recombination rates & distance between loci/sites
Linkage disequilibrium breaks down at different rates, correlated with recombination rate (c).
Recombination Rate Impacts LD and Nucleotide Sequence Diversity

When a mutation arises,
   it is in complete linkage (LD) with entire genome.

Over generations, recombination breaks up LD in proportion to genetic distance between loci.
Recombination Rate Impacts LD and *Nucleotide Sequence Diversity*  

When a mutation arises,  
it is in complete linkage (LD) with entire genome.  

Over generations, recombination breaks up LD in proportion to genetic distance between loci.

Other factors ‘interrupt’ linkage breakdown  
*Selection*  
*Demography*
Linkage Disequilibrium (LD) breaks down in regions outside selection.

**Figure 2** Linkage disequilibrium (LD) decay plot of *shrunken 1 (sh1)* in maize. LD, measured as $r^2$, between pairs of polymorphic sites is plotted against the distance between the sites. For this particular gene, LD decayed within 1500 bp. Data from Reference 48.
Molecular population genetics

Genetic scales to consider:

**Single gene**: selection or neutrality at ‘favorite/candidate’ locus (YFG)

**Genomic neighborhood**: selection or neutrality in region containing YFG, related to effective recombination rate (linkage disequilibrium)

**Entire genome**: changes in population size, gene flow, population structure

...all relate to timeframe and organismal life history; scope of hypothesis determines analysis tools used

Molecular evolution, phylogenetics, population genetics
Molecular population genetics: Different ways to estimate and compare levels of genetic variation

**Different genetic markers:**
- AFLP
- Microsatellites
- Organelle Loci (cpDNA, mtDNA)
- Nuclear Loci (Coding regions, Introns)
  - General loci
  - Candidate genes and ‘your favorite gene (YFG)’

**Must consider:**
- Levels of genetic variation/polymorphism
- Mutational model limitations
- Analytical power for detecting specific processes
  - Allele/haplotype age
  - Gene divergence times (for molecular evolution studies)
Phylogenetic trees for the population level: haplotype network analysis

Statistical parsimony-based haplotype network: **ASAPETALA1-B**

- **A. sandwicense**
- **D. arborea**
- **D. ciliolata**
Shared alleles between taxon 1 and taxon 2:
Testing Isolation vs. Migration using the coalescent & Bayesian statistics.

Species 1 | Species 2
---|---
B | B
B | B
E | E
E | E

Migration, allele sharing

Isolation, incomplete lineage sorting

Shared alleles between taxon 1 and taxon 2: Testing *Isolation* vs. *Migration* using the coalescent & Bayesian statistics

\[
\text{TMRCA} = \frac{t}{2N_e}
\]

\[
M = 2N_e m
\]

Gene divergence
Species divergence
Migration rate

Testing “Isolation” vs. “Migration” using the coalescent & Bayesian statistics

Gene divergence    Species divergence    Migration rate

TMRCA = \frac{t}{2 \, N_e}    M = 2N_em

Gene divergence (credibility interval)    Species divergence (credibility interval)


Migration, allele sharing    Isolation, incomplete lineage sorting
Isolation with migration model: gene flow directionality

(Hey et al. 2004)

Present

$N_1$

$\theta_1 = 4N_1\mu$

$m_2 = m_2 / \mu$

$t = t\mu$

Past

$N_2$

$\theta_2 = 4N_2\mu$

$m_1 = m_1 / \mu$

$N_A$

$\theta_A = 4N_A\mu$
Distribution of estimates across loci:

coordinates of ‘favorite/candidate’ genes?

Population expansion  Demography-equilibrium

Luikart et al. Nature Genetics. 2003
Research in the Lawton-Rauh Lab: Plant Population Genetics

*Arabidopsis* sister species and the Hawaiian silversword alliance

**Brief overview:**

‘Comparative Molecular Population Genetics & Molecular Evolution’

...tools to investigate population dynamics and gene-specific processes

Amy Lawton-Rauh
Dept. Genetics & Biochemistry
Clemson University
AmyLR@clemson.edu
www.clemson.edu/~amylr
Estimating Migration depends upon $T$

($T$ = time to most recent common ancestor)

coalesscence @ common ancestor

species (entire genome)

alleles @ specific locus