Errors in Visual Identification of Juvenile Steelhead, Cutthroat Trout and Their Hybrids

Project Funding through a grant from NOAA Fisheries

Primary Investigators: Hans Voight, Dr. David Hankin, Dr. Eric Loudenslager
Department of Fisheries Biology, Humboldt State University
Project Background

- ESA Listing of Steelhead in Northern California “ESU”
- Coastal Populations of *O. mykiss* and *O. clarki clarki* sympatric
- Diver based surveys unable to reliably ID juveniles
In sympatric settings... Field ID's of Juveniles = Error Prone

- Phenotypic Plasticity
- Presence of Hybrids

Hybrid physical appearance varies... phenotype not always intermediate
Field Protocols

Two-phase field design coinciding w/ “Hankin-Reeves”*
survey

Phase 1:
Collect large random sample of fish through electrofishing and visually classify into one of three categories: Steelhead, Unknown trout, Cutthroat trout.

Phase 2:
A systematic subsample of the first phase fish are selected for genetic analysis and phenotypic observations (correlate field ID with genetic identity)

*Modified Hankin and Reeves (1988)/ Hankin and Mohr (...2005...)*
Phenotypic Observations

1) Categorical index scores
   A. Slash Intensity (1, 2, or 3)
   B. Maxillary Extension (1, 2, or 3)

2) Morphometric relationships
   A. Maxillary Length : Forklength
   B. Maxillary Length : Head Length
   C. Head Length : Forklength
Visually Classify—All fish in hand
Phenotypic Index Scores—Same
Measurements—Systematic
Genetic samples—Systematic
**Study Sites**

**McGarvey Creek**, Lower Klamath River

*O. clarki clarki* abundant vs. *O. mykiss*

- 2002: n=188 sampled trout
- 2003: n=399 sampled trout

**Freshwater Creek**, Humboldt Bay tributary

*O. mykiss* and *O. clarki clarki* both “abundant”

- 2002: n=341 sampled trout
- 2003: n~700 sampled trout; n=362 genetics
Genetic Analyses

Baker et al. (2002) - Developed Oligonucleotide Primers

- Seven nuclear DNA loci and one mitochondrial DNA locus exhibit "fixed species specific" differences for steelhead vs. cutthroat
### Visual ID vs Genetic ID:
Freshwater Tributaries 2003 < 80 mm Fork Length

<table>
<thead>
<tr>
<th>VISUAL CATEGORIES</th>
<th>GENETIC CATEGORIES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pure Steelhead</td>
</tr>
<tr>
<td>Steelhead</td>
<td>9</td>
</tr>
<tr>
<td>Hybrid/Unknown</td>
<td>0</td>
</tr>
<tr>
<td>Cutthroat</td>
<td>0</td>
</tr>
</tbody>
</table>
Visual ID vs Genetic ID:  
Freshwater Tributaries 2003 >= 80 mm Fork Length

<table>
<thead>
<tr>
<th>VISUAL CATEGORIES</th>
<th>GENETIC CATEGORIES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pure Steelhead</td>
</tr>
<tr>
<td>Steelhead</td>
<td>35</td>
</tr>
<tr>
<td>Hybrid/Unknown</td>
<td>1</td>
</tr>
<tr>
<td>Cutthroat</td>
<td>0</td>
</tr>
</tbody>
</table>
Overlap of qualitative phenotypic characteristics
Overlap of quantitative phenotypic characteristics

Means of Morphometric Relationships, Age 0+ Trout, Freshwater Creek 2003
Means of Morphometric Relationships, Age 1+ Trout, Freshwater Creek 2003

- **CT 1+ (n=82)**
- **HY 1+ (n=40)**
- **SH 1+ (n=58)**

- mean Maxillary_Head
- mean Maxillary_FL
- mean Head_FL
Muddy Waters

Given that misclassifications DO occur
Can field methods be developed to replace or minimize the need for costly genetics???
Assess quantitative phenotypic data as potential solution

1. Perform simple (binary) logistic regression with target variable = “steelhead” or “non-steelhead”. Use Beta Pr. Chi. Sq values to identify significant predictor variables (e.g.: slash scores).

2. Develop Classification Tree models to evaluate performance of selected predictor variables in various combinations (Weigel et al. 2001).

3. Compare overall error rates of “Visual ID’s” vs. “Model classifications”
Using only “Maxillary Extension Score”
McGarvey Creek -mainstem and west fork 2003

Age 0+ Trout
N=247

Maxillary Extension =1
Steelhead
N=63
57.1 % correct

Maxillary Extension ={2,3}

Maxillary Extension=2
Cutthroat
N=124
82.3 % correct

Maxillary Extension=3
Cutthroat
N=60
98.3 % correct

Visual classification error rate: 27.1%
Model classification error rate: 20.2 %
Using only “Maxillary Extension Score”
Freshwater Creek - mainstem and tributaries 2003

Age 0+ Trout
N=168

Maxillary Extension =1
Steelhead
N=95
71.6% correct

Maxillary Extension ={2,3}
Cutthroat
N=73
67.1% correct

Visual classification error rate: 37.5%
Model classification error rate: 30.4%
Using All Five Phenotypic Variables
Freshwater Creek mainstem and tributaries 2003

**Age 1+ Trout**
N=188

- Slash Score=1
  - Steelhead
    - N = 71
    - 87.3 % Correct

- Slash Score={2,3}
  - Cutthroat
    - N=117
    - 70.9% Correct

**Visual Error Rate= 23.4%**
**Model Error Rate= 19.2%**

**Head_FL<0.2480912**
- Hybrids
  - N=58
  - 39.6 % correct

**Head_FL>0.2471249**
- Hybrids
  - N=48
  - 33.7% correct

**Max_Head<0.5208696**
- Cutthroat
  - N=24
  - 70.8 % correct

**Max_Head>0.5208696**
- Hybrids
  - N=5
  - 80 % correct
## Summary of *a posteriori* Classification Tree Models vs. Overall Visual Identification Errors (age class/location)

<table>
<thead>
<tr>
<th>Drainage</th>
<th>Sample Size</th>
<th>Decision Tree Model Specification</th>
<th>Model ID Error</th>
<th>Visual ID Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGarvey 0+ N = 247 trout</td>
<td>{max_extens}</td>
<td></td>
<td>20.2%</td>
<td>27.1%</td>
</tr>
<tr>
<td>McGarvey 1+ N = 152 trout</td>
<td>{slash_intensity}</td>
<td></td>
<td>29.6%</td>
<td>8.6%</td>
</tr>
<tr>
<td>McGarvey 1+ N = 152 trout</td>
<td>{max_extens}</td>
<td></td>
<td>23.0%</td>
<td>8.6%</td>
</tr>
<tr>
<td>McGarvey 1+ N = 152 trout</td>
<td>{slash_intensity, max_extens., max_head, max_FL, head_FL}</td>
<td></td>
<td>10.5%</td>
<td>8.6%</td>
</tr>
<tr>
<td>Freshwater 0+ N = 168 trout</td>
<td>{max_extens}</td>
<td></td>
<td>30.4%</td>
<td>38.4%</td>
</tr>
<tr>
<td>Freshwater 1+ N = 188 trout</td>
<td>{slash_intensity}</td>
<td></td>
<td>23.4%</td>
<td>23.4%</td>
</tr>
<tr>
<td>Freshwater 1+ N = 188 trout</td>
<td>{slash_intensity, max_extens.,}</td>
<td></td>
<td>24.5%</td>
<td>23.4%</td>
</tr>
<tr>
<td>Freshwater 1+ N = 188 trout</td>
<td>{slash_intensity, max_extens., max_head, max_FL, head_FL}</td>
<td></td>
<td>19.1%</td>
<td>23.4%</td>
</tr>
</tbody>
</table>
Summary and Conclusions

- Visual ID error rates complicate unbiased estimation in sympatric settings.
- Overlap of hybrid phenotypes with both parent species necessitates some level of genetics.
- Visual ID errors higher for age 0+ fish (vs. age 1+) in both streams, and higher overall when SH and CT are similarly abundant (Freshwater).
- Classification tree models formed with phenotypic data improved error rates for age 0+ trout in both streams (age 1+ less conclusive improvement).
- Additional analyses using ordinal regression techniques (proportional odds) may help to improve on existing models.
Special Thanks to:

Genetics Lab assistants Bryan Crouch and Michael Thomas, Yurok Tribal Fisheries Program, Seth Ricker (CDFG), Dana McCanne (Institute for River Ecosystems at HSU), Green Diamond Resource Company, and HSU Fisheries Dept