The Multi-Species Coalescent (MSC) and its Application in Phylogenetics and Species Delimitation

L. Lacey Knowles
Dept. of Ecology and Evolutionary Biology
University of Michigan

Software: *Delineate*  
Jeet Sukumaran  
Dept. of Biology, Evolutionary Biology Program  
San Diego State University  
[https://github.com/jeetsukumaran/delineate](https://github.com/jeetsukumaran/delineate)

Software: *Decrypt*  
Arnaud Becheler  
Dept. of Ecology and Evolutionary Biology  
University of Michigan  
[https://becheler.github.io/pages/applications.html](https://becheler.github.io/pages/applications.html)
Delimitation models that bring speciation to the multispecies coalescent

Inference of species boundaries (beyond the MSC)

Software: *Delineate*  [https://github.com/jeetsukumaran/delineate](https://github.com/jeetsukumaran/delineate)

- Phylogenetic modeling approach that delineates species versus population lineages under a protracted speciation model


- Model of the geography of genetic divergence under a spatially explicit coalescent to evaluate competing hypotheses about cryptic diversity (inferred under the MSC)
Software: DELINEATE

- approach integrates an explicit model of speciation into the “censored” or “multispecies” coalescent model to organize a set of population lineages sampled from one or more species into mutually-exclusive and jointly-comprehensive subsets, where each subset of population lineages represents a distinct species.

(Sukumaran, Holder, and Knowles, 2020)
Seemingly fractal nature of species diversity

With more geographic sampling, get more and more species inferred under the MSC
Seemingly fractal nature of species diversity

With more geographic sampling, get more and more species inferred under the MSC

Heteronotia binoei complex – species tree from Moritz et al. 2016 (7-8 loci) All lineages diagnosable using BPP.

146 samples
10 candidate species
30 lineages

spelea – see below

planiceps – 7-8 lineages, 2 spp
Software: DELINEATE

Data Requirements

DELINEATE requires two items of data:

- A *population tree*
- A *species assignment table*

File: Dynastes_calib_GeoPop.tre
File: constraints1.txt

Input Data: Population Tree

This is rooted ultrametric tree where each tip lineage represents a population or deme. This tree is typically obtained through a classical “censored” or “multispecies” coalescent analysis, such as results from BP&P (either mode A01 or A10) or StarBeast. The tree can be be specified either in NEXUS or Newick format.
Software: DELINEATE

Data Requirements

- A population tree
- A species assignment table

Input Data: Species Assignment Table

This is a tab-delimited plain text file with at least three columns:

- "lineage"
- "species"
- "status"

There may be more than these three columns, but these columns are mandatory (and all other columns will be ignored). The order of columns does not matter, as the DELINEATE programs will use the labels specified in the header row (see below) to identify the columns.

The first row is the header row: i.e., column labels. Subsequent rows will map each tip in the population tree (the "lineage" column) to a species label ("species") as well as an indication of whether this specis assignment is known or not ("status"). Every tip in the population tree must be represented by a row (and no more than one row) in this table. If species assignments are not known for some population lineages (as would be expected in species discovery type analyses), then the "species" field can be left blank or populated with an arbitrary value, such as a "?" or "NewSp.?" etc., but the "status" field should be set to "0". Species assignments that are known will have the status field set to "1".
Software: DELINEATE   Running a Species Delimitation Analysis

To do a quick check of your files from your terminal:

(base) m-c02wt02ehh24:~ knowlesl$ delineate-check

delineate-check -c Downloads/constraints1.txt -t Downloads/Dynastes_calib_GeoPop.tre -f newick
Software: DELINEATE  Running a Species Delimitation Analysis

Basic Run

Given a population lineage tree file, “population-tree.nex”, and a species assignment table file “species-mappings.tsv”, then the following command will run a DELINEATE analysis on the data:

delineate-estimate partitions --tree-file population-tree.nex --config-file data1.tsv

or, using the short-form options:

delineate-estimate partitions -t population-tree.nex --t data1.tsv

This command has the following components:

- **delineate-estimate**: This is the name of the program to be run.
- **estimate**: This is the command or operation that the program will be running.
- **--tree-file population-tree.nex or -t population-tree.nex**: The **--tree-file** flag, or its short-form synonym, **-t**, specifies that the next element will be the path to tree file with data on the population lineage tree. In this example, the file is located in the current working directory, i.e., *population-tree.nex*. If it was in another directory, then the path could be /home/bilbo/projects/orc-species-delimitation/data1/population-tree.nex, for example.
- **--config-file data1.tsv or -c data1.tsv** The **--config-file** flag, or its short-form synonym, **-c**, specifies that the next element will be the path to species assignment configuration file. In this example, the file is located in the current working directory, i.e., *delineate-species.tsv*. Again, if it was in another directory, then the path could be /home/bilbo/projects/orc-species-delimitation/data1/data1.tsv, for example.

(base) m-c02wt02ehh24:~ knowlesl$ delineate-estimate partitions -c Downloads/constraints1.txt -t Downloads/Dynastes_calib_GeoPop.tre -f newick
Software: DELINEATE    Running a Species Delimitation Analysis

Basic Run Output

Executing this command will run the DELINEATE analysis and will produce the following output files:

- "data1.delimitation-results.json"    File: constraints1.delimitation-results.json
- "data1.delimitation-results.trees"    File: constraints1.delimitation-results.trees

More generally, unless the \texttt{-o} or \texttt{--output-prefix} flag (see below) is used to explicitly specify an alternate output prefix for all results generated, the files will take on a prefix given by the file stemname of the configuration file, "data1" in this example.
Software: DELINEATE  Running a Species Delimitation Analysis

Basic Run Output

Executing this command will run the DELINEATE analysis and will produce the following output files:

- "data1.delimitation-results.json" File: constraints1.delimitation-results.json
- "data1.delimitation-results.trees"

The first file, with the general name of "<output-prefix>.delimitation-results.json", is the primary results file. As can be inferred from its extension, it is a JSON format text file, and it consists of a single a dictionary. The dictionary provides information on the estimated speciation completion rate as well as the probabilities of all the possible partitions of the population lineage leafset into species sets, given the species assignment constraints, ranked by the probability of each partition.
Software: DELINEATE  Running a Species Delimitation Analysis

Basic Run Output

Executing this command will run the DELINEATE analysis and will produce the following output files:

- "data1.delimitation-results.json"
- "data1.delimitation-results.trees"  File: constraints1.delimitation-results.trees

The first file, with the general name of "<output-prefix>.delimitation-results.json", is the primary results file. As can be inferred from its extension, it is a JSON format text file, and it consists of a single a dictionary. The dictionary provides information on the estimated speciation completion rate as well as the probabilities of all the possible partitions of the population lineage leafset into species sets, given the species assignment constraints, ranked by the probability of each partition.
Basic Run Output

Executing this command will run the DELINEATE analysis and will produce the following output files:

- "data1.delimitation-results.json"
- "data1.delimitation-results.trees"  

File: constraints1.delimitation-results.trees

The second file, with the general name of "<output-prefix>.delimitation-results.trees", provides supporting results. Basically this is a collection of trees, with one tree for each partition considered. The topology of the trees are identical, corresponding to the topology of the input tree (i.e., the population lineage tree), as are the tip labels. However, the tips have associated with them some extra metadata that will be available for viewing in a program like FigTree. Most important of this is species, i.e., the label corresponding to the identity of the species assignment in that partition. In the case of species assignments that are constrained (i.e., status indicated by “1”), these will be identical to the assignment and invariant across all partitions, of course. However, in the case of population lineages of unknown species affinities (i.e., status indicated by “0”), this may be an existing species label (if the population lineage was assigned to an existing species in the partition under consideration) or a new, arbitrary species label (if the population lineage was assigned to a new distinct species in the partition under consideration). In addition, in FigTree you can also choose to have the branches colored by “status”, and this will highlight population lineages of (a priori) known vs unknown species affinities, and thus quickly identify the assigned species identities of the lineages of interest.
Software: DELINEATE  Running a Species Delimitation Analysis

Basic Run Output

Executing this command will run the DELINEATE analysis and will produce the following output files:

- "data1.delimitation-results.json"
- "data1.delimitation-results.trees"

The second file, with the general name of "<output-prefix>.delimitation-results.trees", provides supporting results. Basically this is a collection of trees, with one tree for each partition considered. The topology of the trees are identical, corresponding to the topology of the input tree (i.e., the population lineage tree), as are the tip labels. However, the tips have associated with them some extra metadata that will be available for viewing in a program like FigTree. Most important of this is species, i.e., the label corresponding to the identity of the species assignment in that partition. In the case of species assignments that are constrained (i.e., status indicated by “1”), these will be identical to the assignment and invariant across all partitions, of course. However, in the case of population lineages of unknown species affinities (i.e., status indicated by “0”), this may be an existing species label (if the population lineage was assigned to an existing species in the partition under consideration) or a new, arbitrary species label (if the population lineage was assigned to a new distinct species in the partition under consideration). In addition, in FigTree you can also choose to have the branches colored by “status”, and this will highlight population lineages of (a priori) known vs unknown species affinities, and thus quickly identify the assigned species identities of the lineages of interest.
Software: DELINEATE  Running a Species Delimitation Analysis

Open in Text editor: constraints1.delimitation-results.trees
Software: DELINEATE  Running a Species Delimitation Analysis
Open in FigTree; File: constraints1.delimitation-results.trees

First tree reflects constraints in configuration file (yellow are unknowns)
Software: DELINEATE  Running a Species Delimitation Analysis

Open in FigTree; File: constraints1.delimitation-results.trees

To quickly visualize assigned species identities of the lineages of interests (i.e., the unknowns)
Software: DELINEATE  Running a Species Delimitation Analysis

Open in FigTree; File: constraints1.delimitation-results.trees

unknown lineage assigned to pre-existing species

unknown lineages assigned to new species
### Software: DELINEATE

#### Original Alpha Taxonomy

<table>
<thead>
<tr>
<th>lineage</th>
<th>species</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGRP1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>DGRR1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>Dhy6</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>Dhy3Br</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>DtTN1</td>
<td>ty</td>
<td>1</td>
</tr>
<tr>
<td>Dhym5</td>
<td>mo</td>
<td>1</td>
</tr>
<tr>
<td>Dma2</td>
<td>ma</td>
<td>1</td>
</tr>
<tr>
<td>CCVO1</td>
<td>occ</td>
<td>1</td>
</tr>
<tr>
<td>DhsP1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsH3</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsG1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>Dhlc01</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>ERVL2</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhlP7</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhmB2Br</td>
<td>mor</td>
<td>0</td>
</tr>
<tr>
<td>DhpB1Br</td>
<td>pas</td>
<td>0</td>
</tr>
<tr>
<td>DheP8</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DheC06</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>Ysne1</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DhtT9</td>
<td>tri</td>
<td>0</td>
</tr>
<tr>
<td>DhbV2</td>
<td>blu</td>
<td>0</td>
</tr>
<tr>
<td>DhrSL5</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DhrM1</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DHHG2</td>
<td>her</td>
<td>0</td>
</tr>
<tr>
<td>DhhD1</td>
<td>her</td>
<td>0</td>
</tr>
</tbody>
</table>

File: constraints1.txt

### Data Requirements

- A population tree
- A species assignment table

![Image of beetles]
Hypotheses based on different species criteria:

Original Alpha Taxonomy

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Species</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGRP1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>DGRR1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>Dhy6</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>Dhy3Br</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>DtTN1</td>
<td>ty</td>
<td>1</td>
</tr>
<tr>
<td>Dhy5</td>
<td>mo</td>
<td>1</td>
</tr>
<tr>
<td>Dma2</td>
<td>ma</td>
<td>1</td>
</tr>
<tr>
<td>CCOV1</td>
<td>occ</td>
<td>1</td>
</tr>
<tr>
<td>DhsP1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsH3</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsG1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhcC1</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>ERVL2</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhtP7</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhmB2Br</td>
<td>mor</td>
<td>0</td>
</tr>
<tr>
<td>DhpB1Br</td>
<td>pas</td>
<td>0</td>
</tr>
<tr>
<td>DheP8</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DheC06</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>YSNE1</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DhtT9</td>
<td>tri</td>
<td>0</td>
</tr>
<tr>
<td>DhbV2</td>
<td>blu</td>
<td>0</td>
</tr>
<tr>
<td>DhrSL5</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DhrM1</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DHHG2</td>
<td>her</td>
<td>0</td>
</tr>
<tr>
<td>DhhD1</td>
<td>her</td>
<td>0</td>
</tr>
</tbody>
</table>

- Taxonomic practices vs diversity differences

Data Requirements

- A population tree
- A species assignment table
Could test different hypotheses about species boundaries based on different species criteria

Huang & Knowles (2016) *Syst. Biol.*
Software: DELINEATE

Hypotheses based on different species criteria:

Original Alpha Taxonomy

<table>
<thead>
<tr>
<th>lineage</th>
<th>species</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGRP1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>DGRR1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>Dhy6</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>Dhy3Br</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>DtTN1</td>
<td>ty</td>
<td>1</td>
</tr>
<tr>
<td>Dhym5</td>
<td>mo</td>
<td>1</td>
</tr>
<tr>
<td>Dma2</td>
<td>ma</td>
<td>1</td>
</tr>
<tr>
<td>CCVO1</td>
<td>occ</td>
<td>1</td>
</tr>
<tr>
<td>DhsP1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsH3</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsG1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>Dhlc01</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>ERVL2</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhlP7</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhmB2Br</td>
<td>mor</td>
<td>0</td>
</tr>
<tr>
<td>DhpB1Br</td>
<td>pas</td>
<td>0</td>
</tr>
<tr>
<td>DheP8</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DheC06</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>YSNE1</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DhtT9</td>
<td>tri</td>
<td>0</td>
</tr>
<tr>
<td>DhbV2</td>
<td>blu</td>
<td>0</td>
</tr>
<tr>
<td>DhrSL5</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DhrM1</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DHHG2</td>
<td>her</td>
<td>0</td>
</tr>
<tr>
<td>DhhD1</td>
<td>her</td>
<td>0</td>
</tr>
</tbody>
</table>

“Lumper” Taxonomy

<table>
<thead>
<tr>
<th>lineage</th>
<th>species</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGRP1</td>
<td>gr</td>
<td>grhy</td>
</tr>
<tr>
<td>DGRR1</td>
<td>gr</td>
<td>grhy</td>
</tr>
<tr>
<td>Dhy6</td>
<td>hy</td>
<td>grhy</td>
</tr>
<tr>
<td>Dhy3Br</td>
<td>hy</td>
<td>grhy</td>
</tr>
<tr>
<td>DtTN1</td>
<td>ty</td>
<td>motyma</td>
</tr>
<tr>
<td>Dhym5</td>
<td>mo</td>
<td>motyma</td>
</tr>
<tr>
<td>Dma2</td>
<td>ma</td>
<td>motyma</td>
</tr>
<tr>
<td>CCVO1</td>
<td>occ</td>
<td>seppoc</td>
</tr>
<tr>
<td>DhsP1</td>
<td>sep</td>
<td>seppoc</td>
</tr>
<tr>
<td>DhsH3</td>
<td>sep</td>
<td>seppoc</td>
</tr>
<tr>
<td>DhsG1</td>
<td>sep</td>
<td>seppoc</td>
</tr>
<tr>
<td>Dhlc01</td>
<td>lic</td>
<td>lic</td>
</tr>
<tr>
<td>ERVL2</td>
<td>lic</td>
<td>lic</td>
</tr>
<tr>
<td>DhlP7</td>
<td>lic</td>
<td>lic</td>
</tr>
<tr>
<td>DhmB2Br</td>
<td>mor</td>
<td>mor</td>
</tr>
<tr>
<td>DhpB1Br</td>
<td>pas</td>
<td>pas</td>
</tr>
<tr>
<td>DheP8</td>
<td>ecu</td>
<td>ecu</td>
</tr>
<tr>
<td>DheC06</td>
<td>ecu</td>
<td>ecu</td>
</tr>
<tr>
<td>YSNE1</td>
<td>ecu</td>
<td>ecu</td>
</tr>
<tr>
<td>DhtT9</td>
<td>tri</td>
<td>tri</td>
</tr>
<tr>
<td>DhbV2</td>
<td>blu</td>
<td>blu</td>
</tr>
<tr>
<td>DhrSL5</td>
<td>rei</td>
<td>rei</td>
</tr>
<tr>
<td>DhrM1</td>
<td>rei</td>
<td>rei</td>
</tr>
<tr>
<td>DHHG2</td>
<td>her</td>
<td>her</td>
</tr>
<tr>
<td>DhhD1</td>
<td>her</td>
<td>her</td>
</tr>
</tbody>
</table>

“Splitter” Taxonomy

<table>
<thead>
<tr>
<th>lineage</th>
<th>species</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGRP1</td>
<td>gr1</td>
<td>1</td>
</tr>
<tr>
<td>DGRR1</td>
<td>gr2</td>
<td>1</td>
</tr>
<tr>
<td>Dhy6</td>
<td>hy1</td>
<td>1</td>
</tr>
<tr>
<td>Dhy3Br</td>
<td>hy2</td>
<td>1</td>
</tr>
<tr>
<td>DtTN1</td>
<td>ty1</td>
<td>1</td>
</tr>
<tr>
<td>Dhym5</td>
<td>mo1</td>
<td>1</td>
</tr>
<tr>
<td>Dma2</td>
<td>ma1</td>
<td>1</td>
</tr>
<tr>
<td>CCVO1</td>
<td>occ1</td>
<td>1</td>
</tr>
<tr>
<td>DhsP1</td>
<td>sep1</td>
<td>1</td>
</tr>
<tr>
<td>DhsH3</td>
<td>sep1</td>
<td>1</td>
</tr>
<tr>
<td>DhsG1</td>
<td>sep1</td>
<td>1</td>
</tr>
<tr>
<td>Dhlc01</td>
<td>lic1</td>
<td>1</td>
</tr>
<tr>
<td>ERVL2</td>
<td>lic1</td>
<td>1</td>
</tr>
<tr>
<td>DhlP7</td>
<td>lic1</td>
<td>1</td>
</tr>
<tr>
<td>DhmB2Br</td>
<td>mor0</td>
<td>0</td>
</tr>
<tr>
<td>DhpB1Br</td>
<td>pas0</td>
<td>0</td>
</tr>
<tr>
<td>DheP8</td>
<td>ecu0</td>
<td>0</td>
</tr>
<tr>
<td>DheC06</td>
<td>ecu0</td>
<td>0</td>
</tr>
<tr>
<td>YSNE1</td>
<td>ecu0</td>
<td>0</td>
</tr>
<tr>
<td>DhtT9</td>
<td>tri0</td>
<td>0</td>
</tr>
<tr>
<td>DhbV2</td>
<td>blu0</td>
<td>0</td>
</tr>
<tr>
<td>DhrSL5</td>
<td>rei0</td>
<td>0</td>
</tr>
<tr>
<td>DhrM1</td>
<td>rei0</td>
<td>0</td>
</tr>
<tr>
<td>DHHG2</td>
<td>her0</td>
<td>0</td>
</tr>
<tr>
<td>DhhD1</td>
<td>her0</td>
<td>0</td>
</tr>
</tbody>
</table>

Data Requirements

- A population tree
- A species assignment table
- Number of inferred species naturally increases as the known species assignments are made under more of a splitter, rather than a lumper, than perspective.
- However, naturally as well, the relationship is not a simple numerical one but is rather dependent on the branch lengths leading to the assigned species subtrees.

| Regime  | Constrained | Inferred | New |
|----------------+-------------+----------+-----|
| lumper2        | 3           | 1        | 0   |
| lumper1        | 4           | 1        | 0   |
| original       | 8           | 2        | 2   |
| splitter1      | 12          | 2        | 2   |
| splitter2      | 14          | 10       | 10  |

Data Requirements

- A population tree
- A species assignment table

<table>
<thead>
<tr>
<th>lineage</th>
<th>species</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGRP1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>DGRR1</td>
<td>gr</td>
<td>1</td>
</tr>
<tr>
<td>Dhy6</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>Dhy3Br</td>
<td>hy</td>
<td>1</td>
</tr>
<tr>
<td>DtTN1</td>
<td>ty</td>
<td>1</td>
</tr>
<tr>
<td>Dhym5</td>
<td>mo</td>
<td>1</td>
</tr>
<tr>
<td>Dma2</td>
<td>ma</td>
<td>1</td>
</tr>
<tr>
<td>CCV01</td>
<td>occ</td>
<td>1</td>
</tr>
<tr>
<td>DhsP1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DshH3</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhsG1</td>
<td>sep</td>
<td>1</td>
</tr>
<tr>
<td>DhlC01</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>ERVL2</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhlP7</td>
<td>lic</td>
<td>1</td>
</tr>
<tr>
<td>DhmB2Br</td>
<td>mor</td>
<td>0</td>
</tr>
<tr>
<td>DhpB1Br</td>
<td>pas</td>
<td>0</td>
</tr>
<tr>
<td>DheP8</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DhecO6</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>YSNE1</td>
<td>ecu</td>
<td>0</td>
</tr>
<tr>
<td>DhtT9</td>
<td>tri</td>
<td>0</td>
</tr>
<tr>
<td>DhbV2</td>
<td>blu</td>
<td>0</td>
</tr>
<tr>
<td>DhrSL5</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DhrM1</td>
<td>rei</td>
<td>0</td>
</tr>
<tr>
<td>DHHG2</td>
<td>her</td>
<td>0</td>
</tr>
<tr>
<td>DhhD1</td>
<td>her</td>
<td>0</td>
</tr>
</tbody>
</table>
A new era of species delimitation models that brings speciation models to the multispecies coalescent

- Erroneous species boundaries are inferred from current model-based genetic approaches based on MSC
- Delimitation under the MSC:
  - genetic structure = species

- Relying on heuristics to interpret results from current genetic methods (e.g., bpp) is not the answer

- Future of genetic-based species delimitation is with speciation-based delimitation models in which species criteria is explicitly incorporated into test of species status of unknowns (e.g., “splitter” vs “lumper”)
Our work pioneers a new age of species delimitation approaches that deal with the problem directly by modeling the issue instead of ignoring it! That is, by actually incorporating an explicit model of the speciation process — in particular, an extended or protracted speciation process — into species delimitation, we are able to discriminate between species and population (or other) boundaries in genomic data.
QUESTIONS?

Jeet Sukumaran
San Diego State Univ.

Mark Holder
Univ. of Kansas

knowlesl@umich.edu

support NSF & the UM
Software: *Decrypt*  
https://becheler.github.io/pages/applications.html

- Model of the geography of genetic divergence under a spatially explicit coalescent to evaluate competing hypotheses about cryptic diversity (inferred under the MSC)