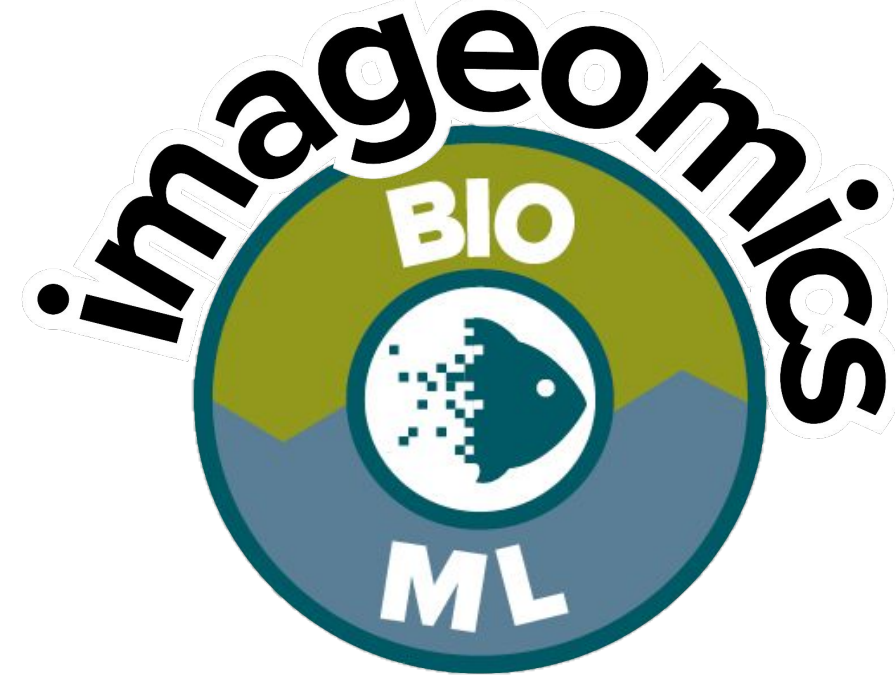


# Hierarchical Conditioning of Diffusion Models Using Tree-of-Life for Studying Species Evolution



Scan for Paper



Mridul Khurana<sup>1,\*</sup>, Arka Daw<sup>2</sup>, M. Maruf<sup>1</sup>, Josef Uyeda<sup>1</sup>, Wasila Dahdul<sup>3</sup>, Caleb Charpentier<sup>1</sup>, Yasin Bakis<sup>4</sup>, Henry Bart<sup>4</sup>, Paula Mabee<sup>5</sup>, Hilmar Lapp<sup>6</sup>, James Balhoff<sup>7</sup>, Wei-Lun Chao<sup>8</sup>, Charles Stewart<sup>9</sup>, Tanya Berger-Wolf<sup>8</sup>, and Anuj Karpatne<sup>1</sup>

<sup>1</sup>Virginia Tech, <sup>2</sup>Oak Ridge National Laboratory, <sup>3</sup>University of California Irvine, <sup>4</sup>Tulane University, <sup>5</sup>Battelle, <sup>6</sup>Duke University,

<sup>7</sup>University of North Carolina Chapel Hill, <sup>8</sup>The Ohio State University, <sup>9</sup>Rensselaer Polytechnic Institute

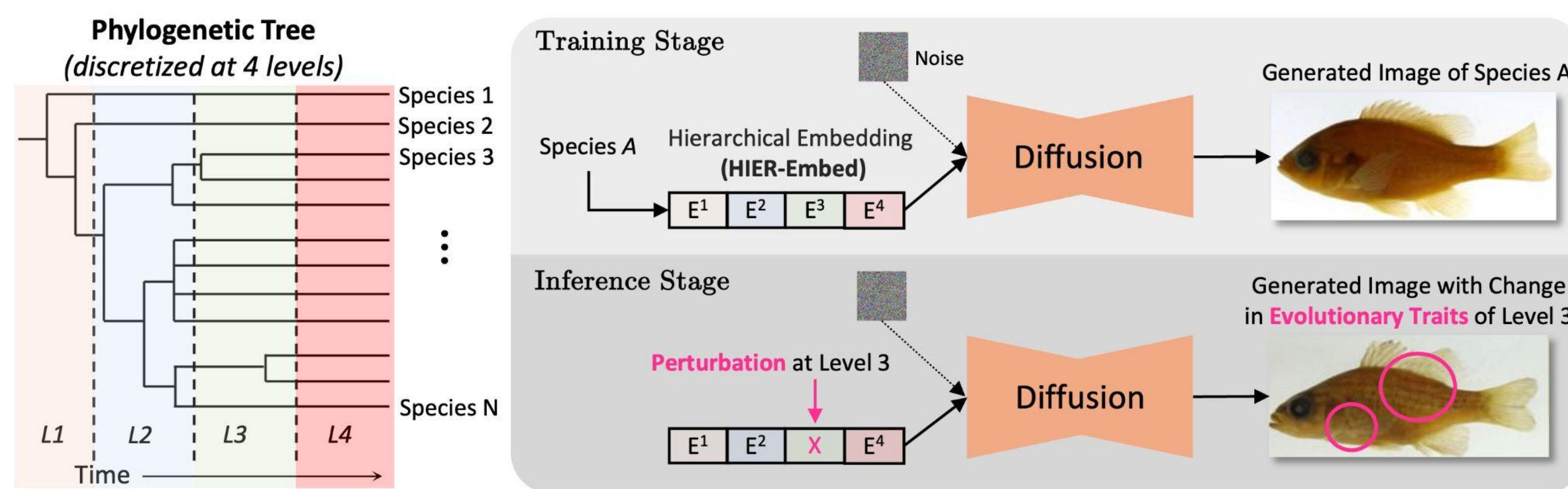
[ Accepted to ECCV 2024 ] \*presenting author

## Overview

- We aim to understand evolutionary traits in species directly from images without relying on expert-annotated trait labels.
- We introduce **Phyo-Diffusion**, a novel way of conditioning diffusion models with trees using hierarchical embeddings (HIER-Embeds)
- We introduce two novel operations for understanding evolutionary traits:
  - Trait Masking
  - Trait Swapping

## Methodology

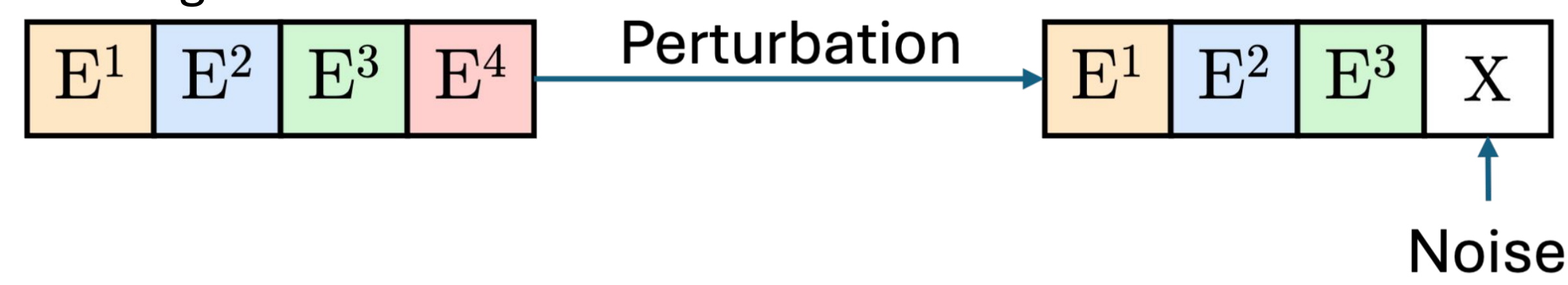
**Phylogeny:** “a tree of life” that encodes the evolutionary relationships between species. For computational simplicity, we quantize the tree into four ancestral levels.



## Proposed Experiments

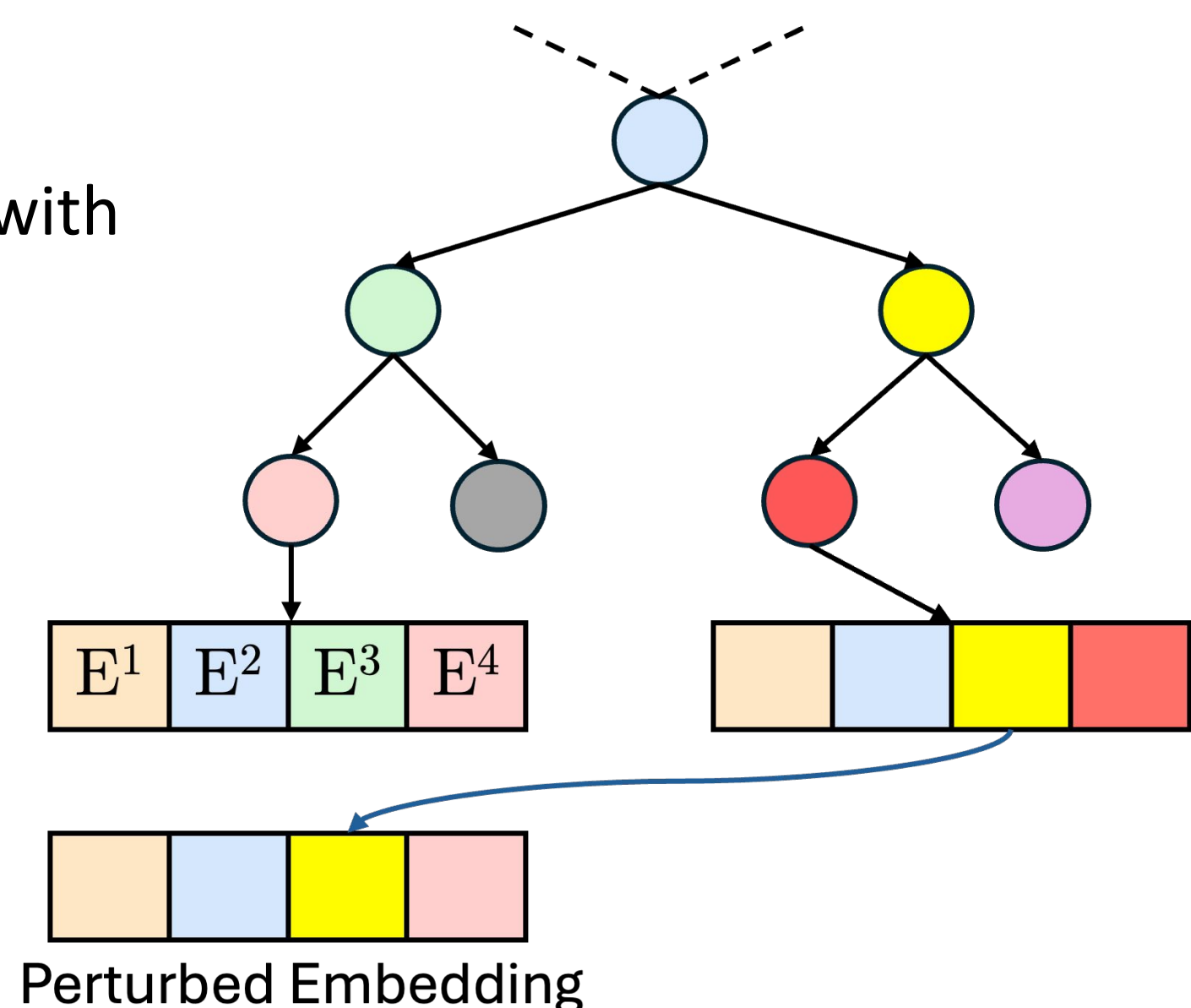
### Trait Masking:

- Replace components with noise
- Analogous to “Gene Knockout”



### Trait Swapping:

- Replace components with siblings
- Analogous to “CRISPR experiments”



## Results

- Qualitative evaluation of generated images for different models**
  - Order of species from top to bottom is Cyprinus Carpio, Notropis Hudsonius, Lepomis Auritus, Noturus Exilis, Gambusia Affinis

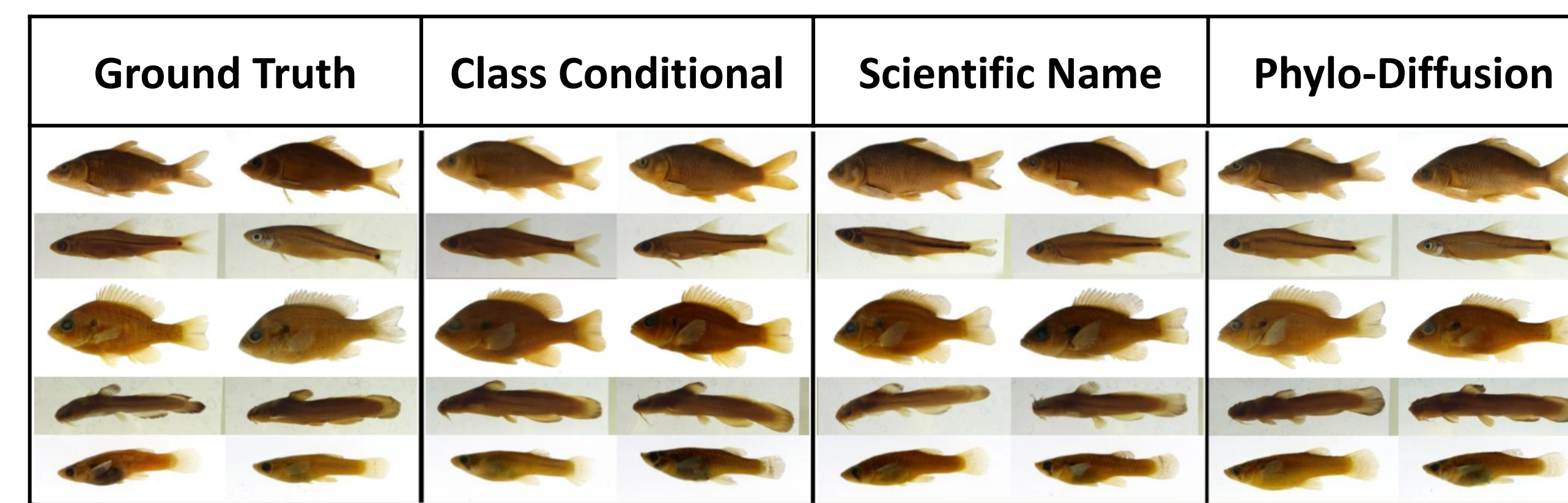


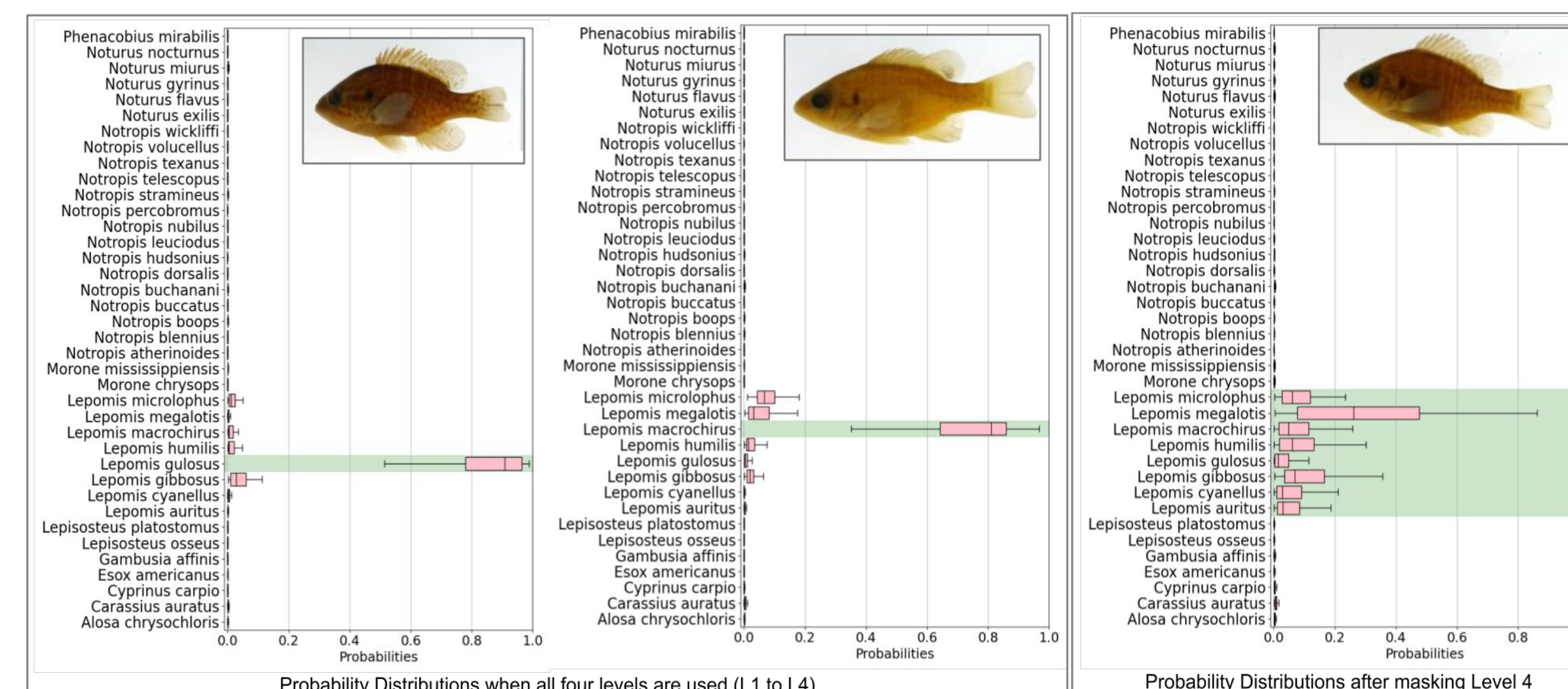
Table 1: Quantitative Comparison of generated images sampled using DDIM (100 samples per class).

Method	FID ↓	IS ↑	Prec. ↑	Recall ↑	F1 (%) ↑
PhyloNN	28.08	2.35	0.625	0.084	47.37
Class Conditional	11.46	2.47	0.679	0.359	81.99
Scientific Name	11.76	2.43	0.683	0.332	70.16
Phyo-Diffusion (ours)	11.38	2.53	0.654	0.367	82.21

- For F1-Score, we use a separately trained ResNet-18 classifier. The base model has a test accuracy of 0.85.
- Goal: To verify the generated images capture essential phylogenetic information but also enable accurate classification within their respective classes.

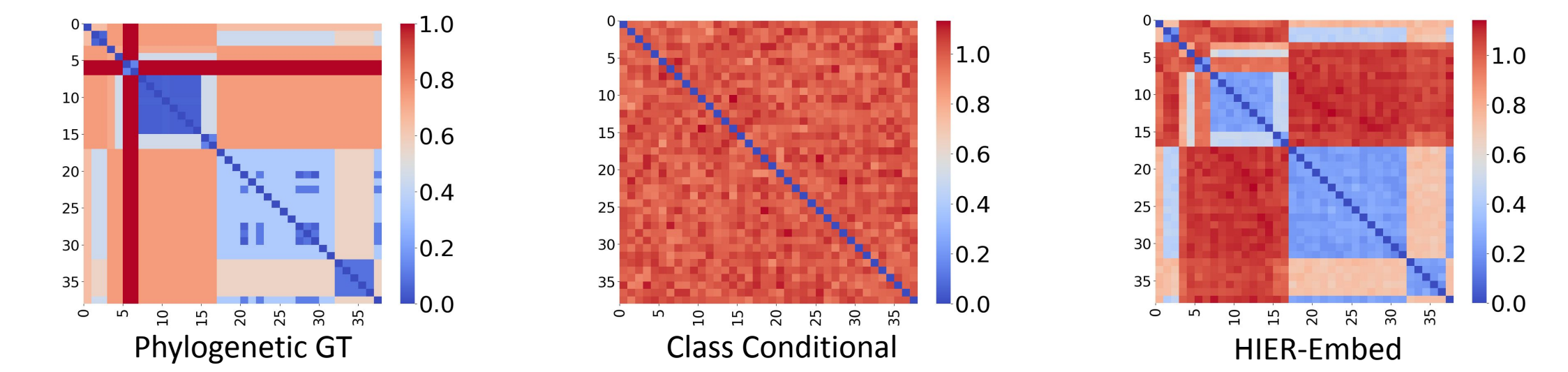
### Trait Masking: Perturbing Level embeddings with Noise

- We replace each of the level embeddings with noise to examine the representations learned at internal nodes



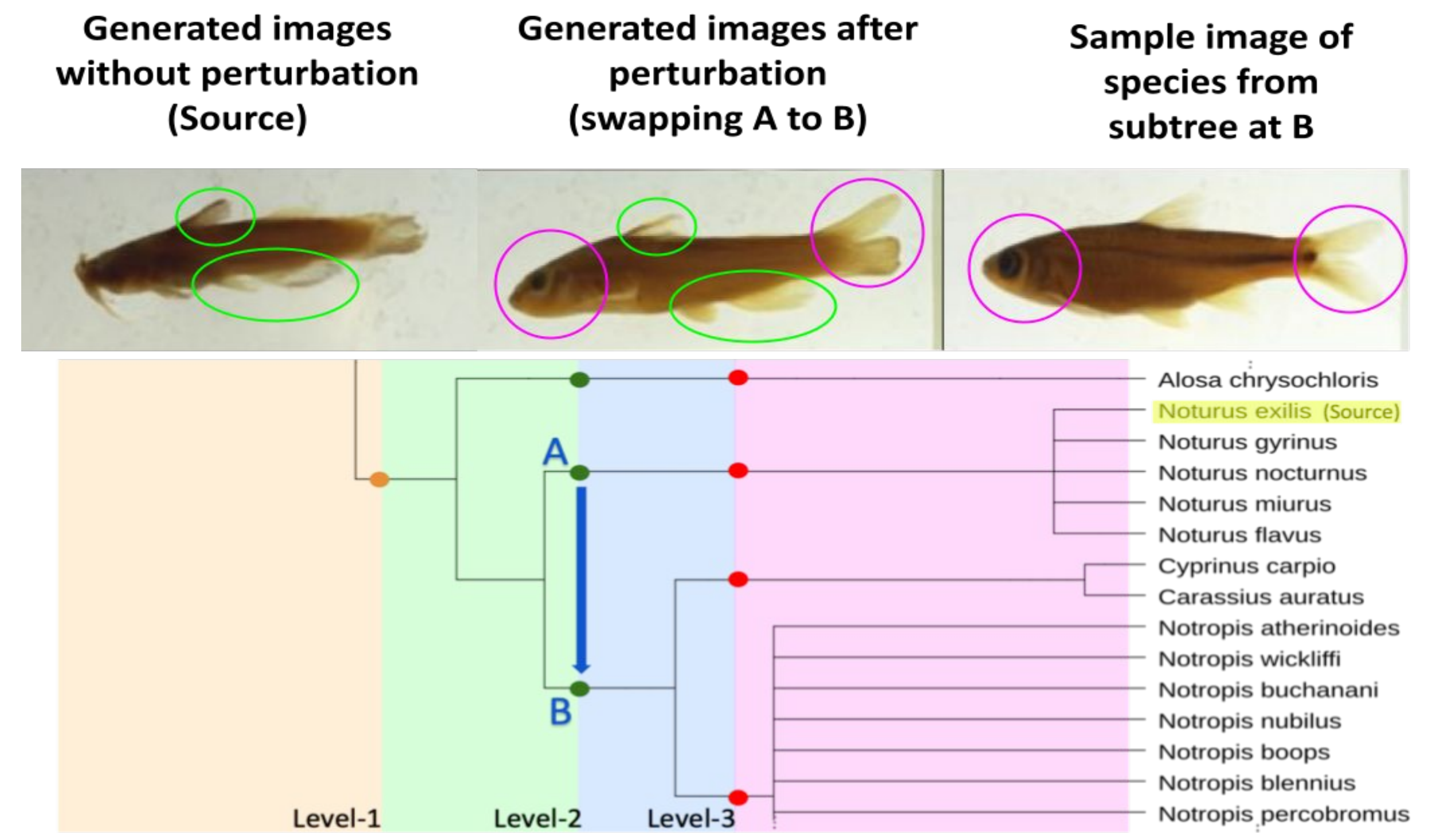
- Analysis of the quality of embeddings generated from conditioning encoders using cosine distance.**

- We found that HIER-Embed distances correlated more closely with the GT phylogeny than any other model



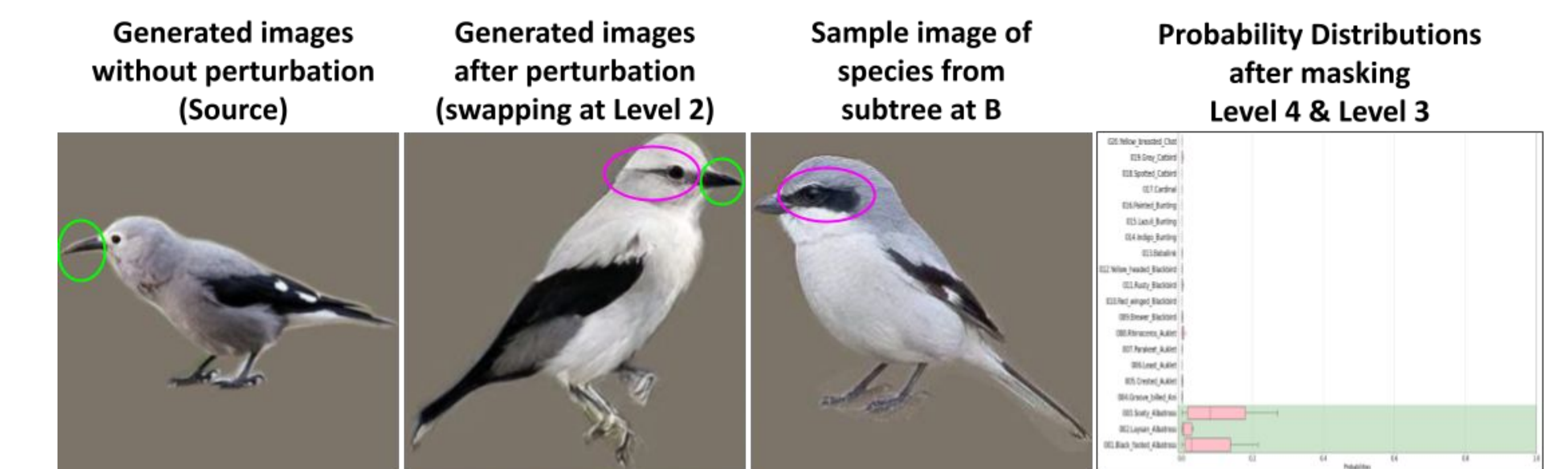
### Trait Swapping: Perturbing Species with Siblings

- We replace a specific level encoding of a reference species with that of a sibling node to observe the evolutionary trait changes occurred at that node.



### Results on the CUB dataset.

- Visualization of changes in traits after swapping information at Level 2 for *Clark nutcracker* (left) with its species from its sibling subtree (right) to generate perturbed species (center)



## On-going Work and Future Steps

- Extending the work to Butterflies (Meier-Jiggins dataset) which have well defined phylogenies and Birds-525
- Handling imbalanced trees and at different depths of species nodes in the phylogeny

This work was supported by the NSF OAC 2118240 Imageomics Institute award used computing resources from the Advanced Research Computing (ARC) Center at Virginia Tech