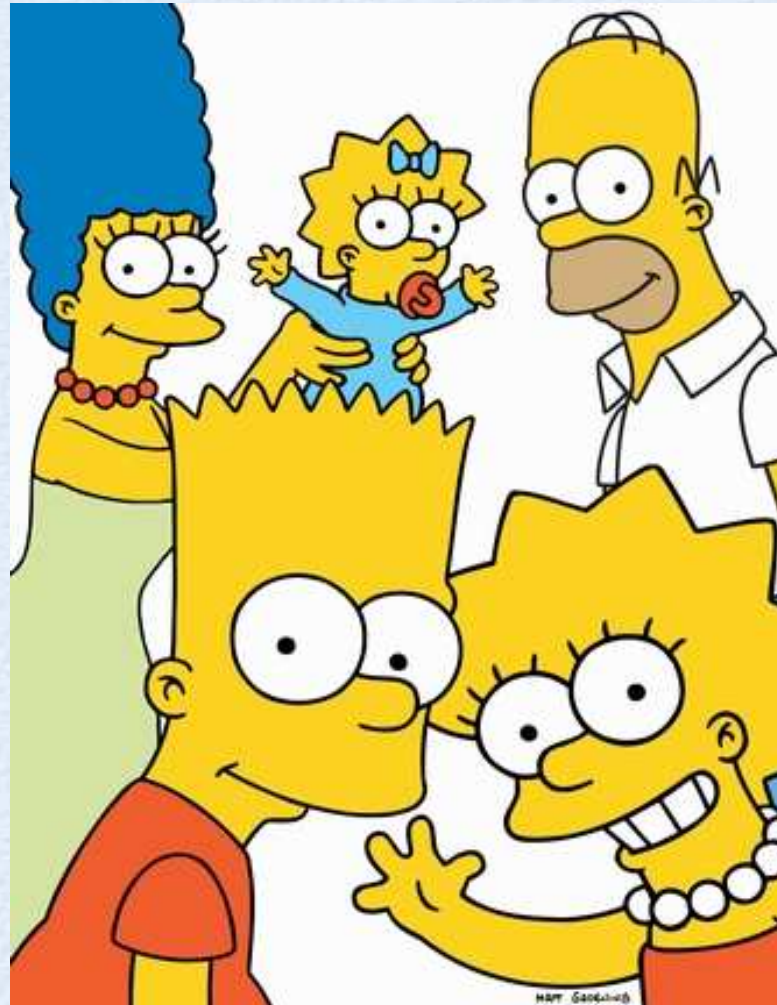# Identification of animal images based on DNA

Lior Wolf and Yoni Donner
The School of Computer Science
Tel-Aviv University

Visual appearance is largely determined by genetics…

...and appearance affecting genes have been identified

- Eye color
  Sturm & Frudakis (2004).

- Skeletel structure
  Chase, et al. (2002).

- Dog height
  Sutterm, (2007).

Our work is unique in the following ways:

High Dimensional working directly with images

Using very short 500bp sequences

Correlative links
Not causal ones

We employ mtDNA

- Has no causal link to appearance

- Easy to collect
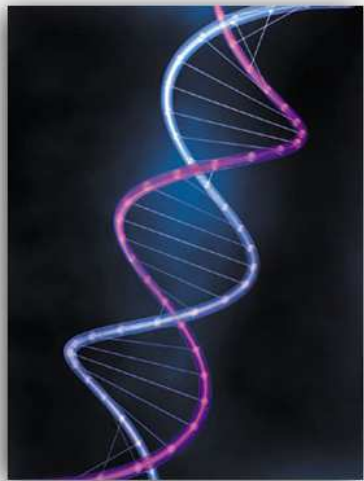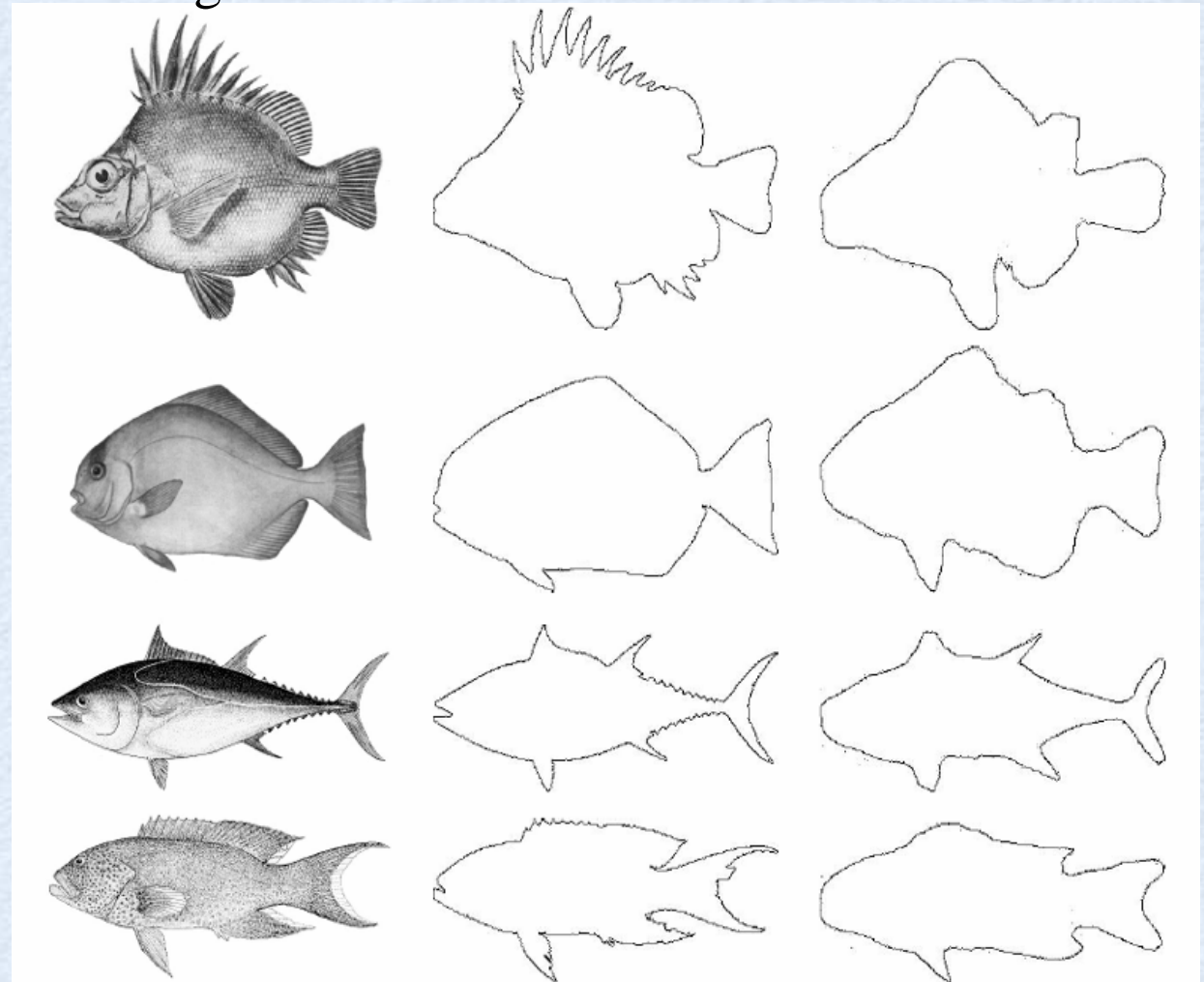
- Good marker for evolutionary history

Tasks

1. Image synthesis given DNA (!)

2. Identification of the correct image given DNA
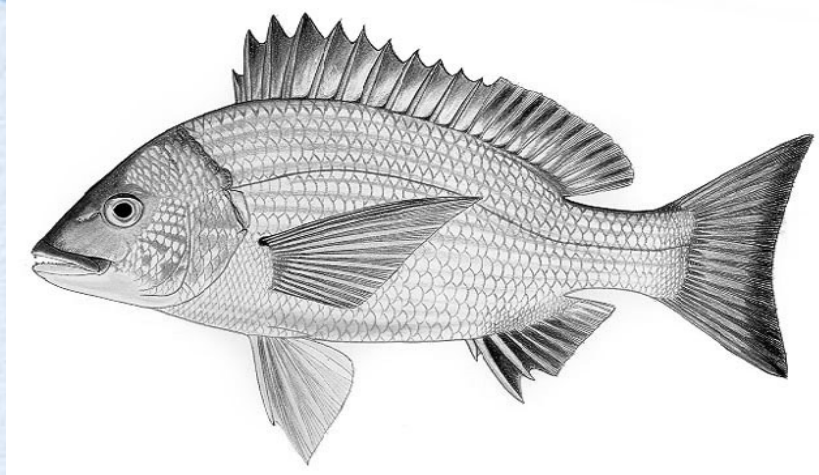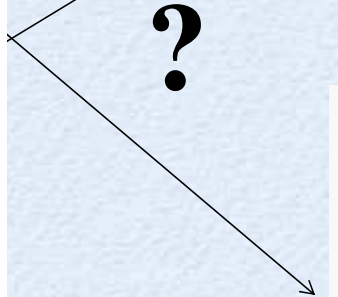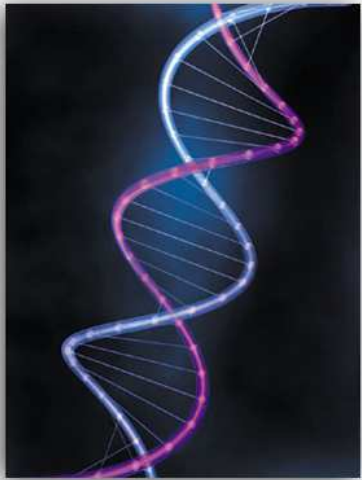
Image synthesis

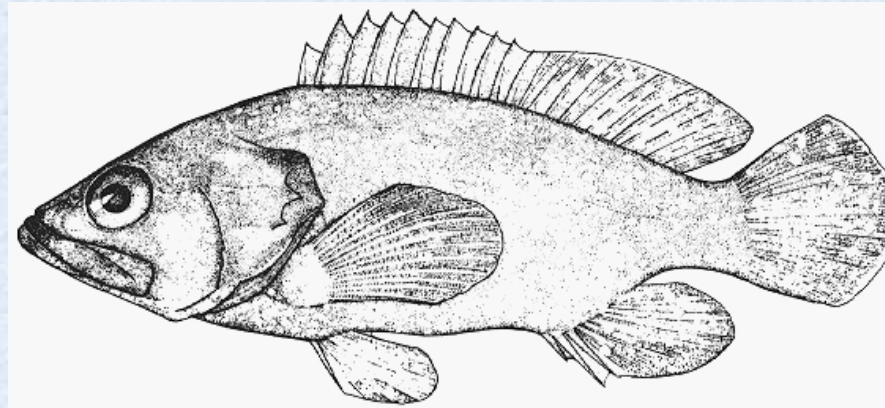Original image | Extracted contour | Recovered contour

The algorithm predicts the contour of an unseen species based on its mtDNA

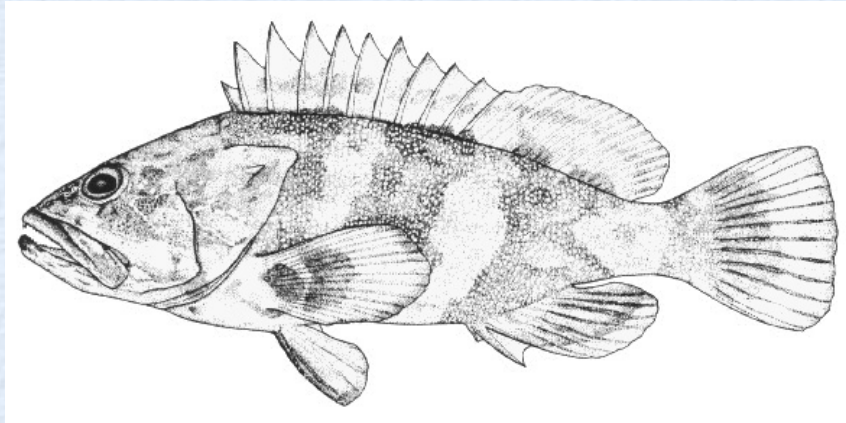# Identification

Acanthopagrus Butcheri

Epinephelus ongus

Given mtDNA, the algorithm identifies the correct unseen image

Data sets

Fishes of Australia

Birds of North America

Dorsal view

Head view

Profile view

Ants of Madagascar

Methods

Three technical questions:

1. How do we represent the genetic sequences?

   Simple intuitive vector representation

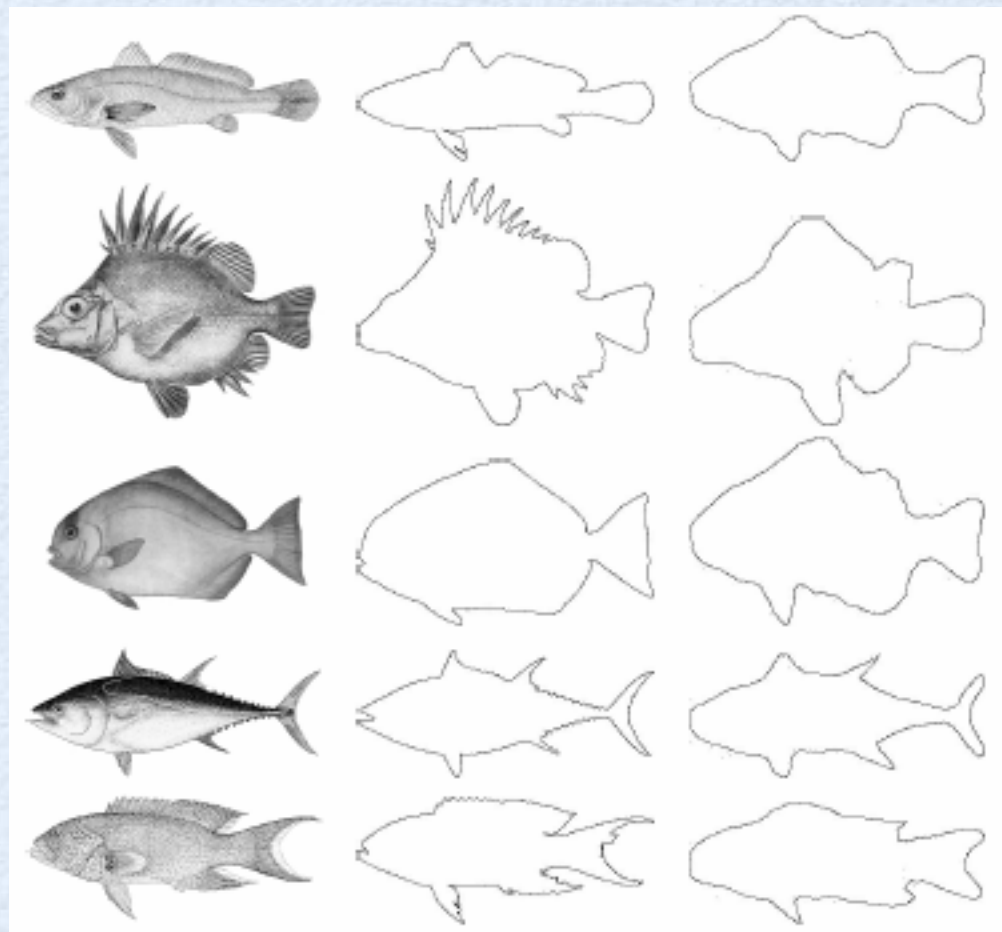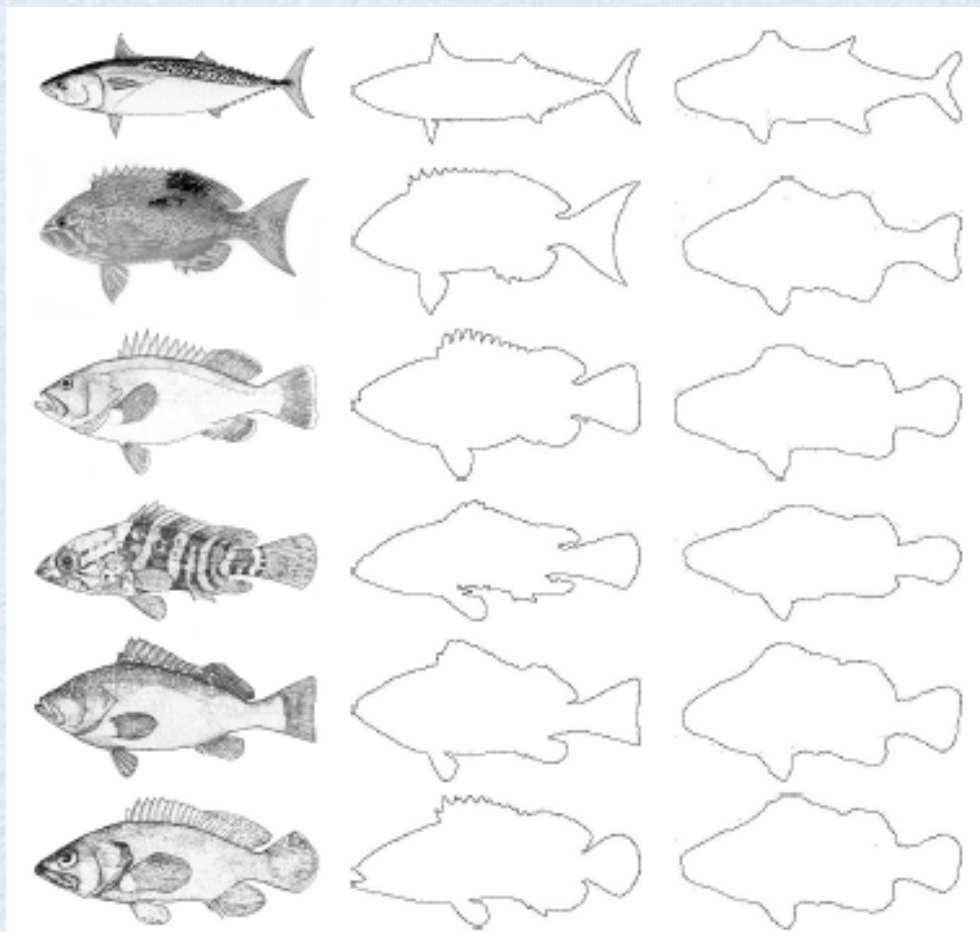2. How do we represent the images?

   Well known techniques: Bag of SIFT or C_1

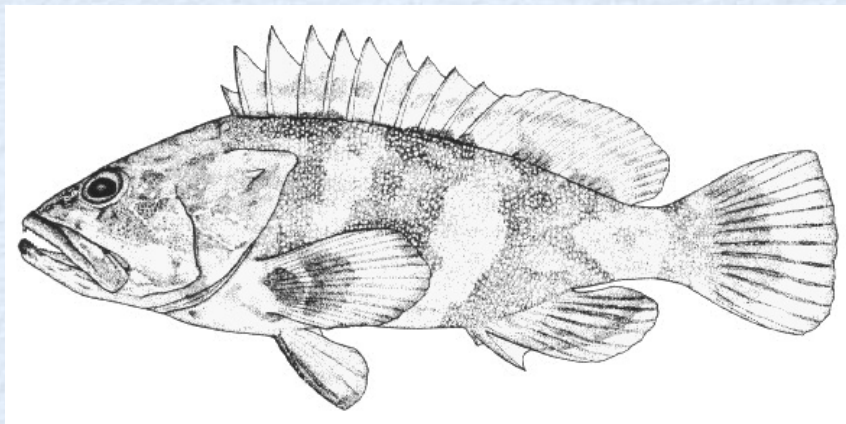3. How do we learn the connections?

   Regularized CCA

Results

- 93 fish species, 82 for training, 11 for testing

- Significantly better than chance or NN

Results

Fish: 90% correct

Birds: 72% correct

Dorsal: 59% correct

Head: 56% correct

Profile: 64% correct

Conclusions

- Visual identification and image synthesis based on DNA sequences is a reality

- Employing the power of correlation rather than causality

- Future applications?

    Virtual line-ups

    Personalized medication

    Prediction of appearance of extinct animals