

# DNA barcoding and molecular systematics of fishes

*Сергей Туранов*

Лаборатория молекулярной систематики  
Института биологии моря им. А.В. Жирмунского

MolSy-2015

## Biological identifications through DNA barcodes

Paul D. N. Hebert\*, Alina Cywinska, Shelley L. Ball  
and Jeremy R. deWaard

Department of Zoology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Although much biological research depends upon species diagnoses, taxonomic expertise is collapsing. We are convinced that the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon 'barcodes'. We establish that the mitochondrial gene cytochrome *c* oxidase I (COI) can serve as the core of a global bioidentification system for animals. First, we demonstrate that COI profiles, derived from the low-density sampling of higher taxonomic categories, ordinarily assign newly analysed taxa to the appropriate phylum or order. Second, we demonstrate that species-level assignments can be obtained by creating comprehensive COI profiles. A model COI profile, based upon the analysis of a single individual from each of 200 closely allied species of lepidopterans, was 100% successful in correctly identifying subsequent specimens. When fully developed, a COI identification system will provide a reliable, cost-effective and accessible solution to the current problem of species identification. Its assembly will also generate important new insights into the diversification of life and the rules of molecular evolution.

**Keywords:** molecular taxonomy; mitochondrial DNA; animals; insects; sequence diversity; evolution



Sample ID: HBLPD7  
License: Copyright (2012)  
License Holder: Paul Hebert, Biodiversity Institute of Ontario

*Homo sapiens* Linnaeus, 1758.  
Пол Эбер, создатель термина и направления.

Over 5000 citations

## Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species

Paul D. N. Hebert\*, Sujeevan Ratnasingham  
and Jeremy R. deWaard

Department of Zoology, University of Guelph, Guelph,  
Ontario N1G 2W1, Canada

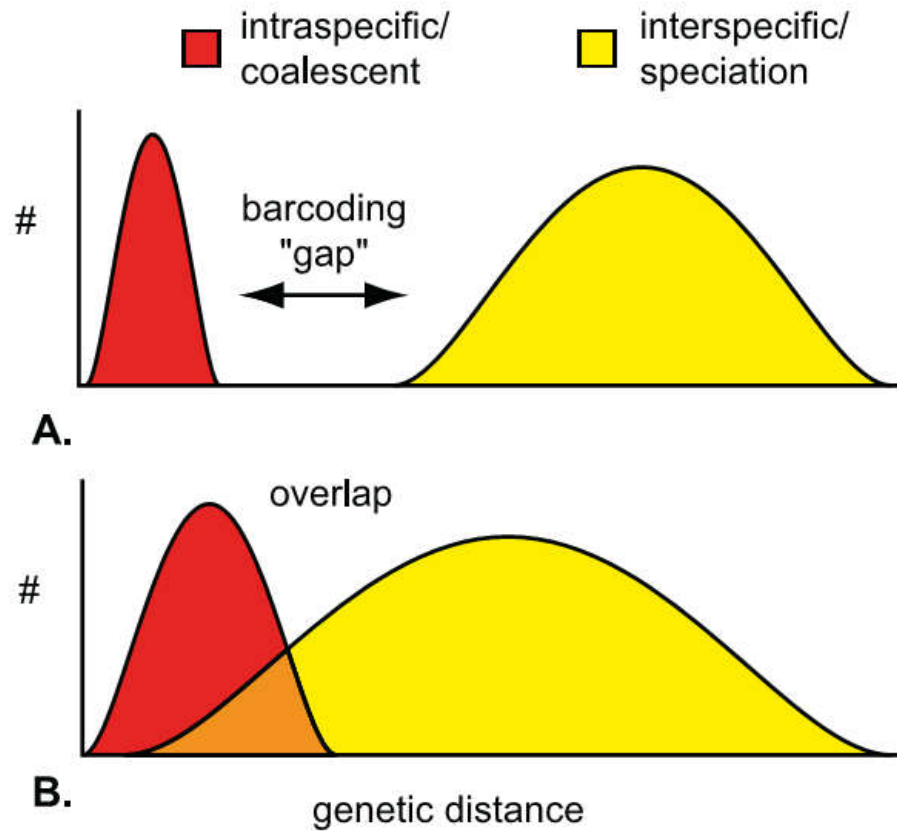
\* Author for correspondence (phebert@uoguelph.ca).

Recd 09.03.03; Acctd 28.03.03; Online 15.05.03

With millions of species and their life-stage transformations, the animal kingdom provides a challenging target for taxonomy. Recent work has suggested that a DNA-based identification system, founded on the mitochondrial gene, cytochrome *c* oxidase subunit 1 (COI), can aid the resolution of this diversity. While past work has validated the ability of COI sequences to diagnose species in certain taxonomic groups, the present study extends these analyses across the animal kingdom. The results indicate that sequence divergences at COI regularly enable the discrimination of closely allied species in all animal phyla except the Cnidaria. This success in species diagnosis reflects both the high rates of sequence change at COI in most animal groups and constraints on intraspecific mitochondrial DNA divergence arising, at least in part, through selective sweeps mediated via interactions with the nuclear genome.

**Keywords:** molecular taxonomy; DNA barcode; cytochrome *c* oxidase subunit 1; DNA; mitochondrial

Over 2000 citations

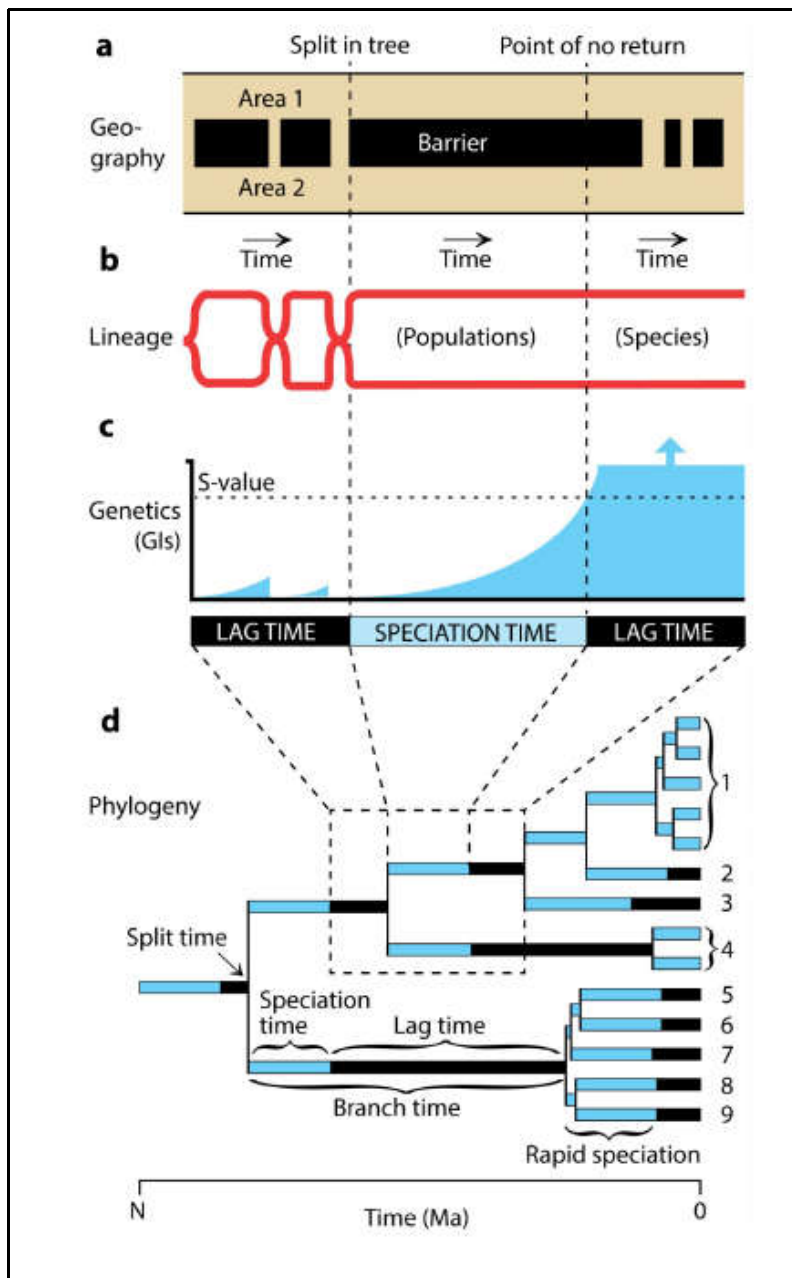


**Figure 2.** Schematic of the Inferred Barcoding Gap

The distribution of intraspecific variation is shown in red, and interspecific divergence in yellow. (A) Ideal world for barcoding, with discrete distributions and no overlap. (B) An alternative version of the world with significant overlap and no gap.

DOI: 10.1371/journal.pbio.0030422.g002

From: Meyer, C. P., & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS biology*, 3(12), 2229.



Summary model of speciation.

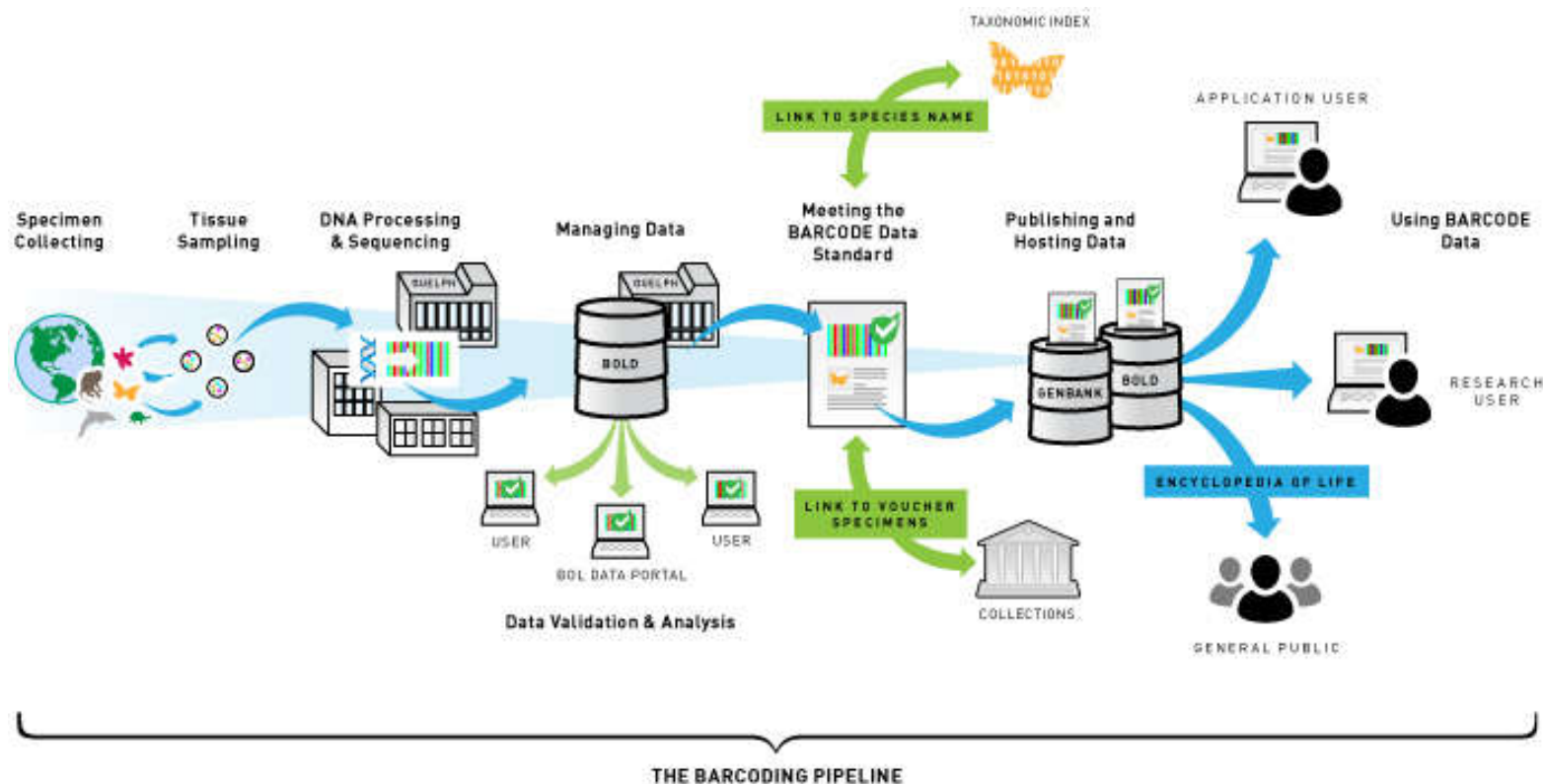
(a) Biogeographic history, showing the contact and isolation of areas occupied by the two populations.

(b) Phylogenetic lineages, showing times of independence (two lineages) and times of interbreeding (one lineage).

(c) Genic incompatibilities, GIs, between the two populations, showing how they accrue at a timedependent rate during geographic isolation, reset to zero during contact (interbreeding), increase to the S-value (the number of GIs that will cause speciation, the point of no return), and continue increasing beyond the S-value despite later contact of the newly-formed species.

(d) Hypothetical phylogeny, with numbered species, illustrating parameters of speciation in (a–c) to splits and branches in a tree.

From: Hedges, S. B., Marin, J., Suleski, M., Paymer, M., & Kumar, S. (2015). Tree of life reveals clock-like speciation and diversification. *Molecular biology and evolution*, msv037.

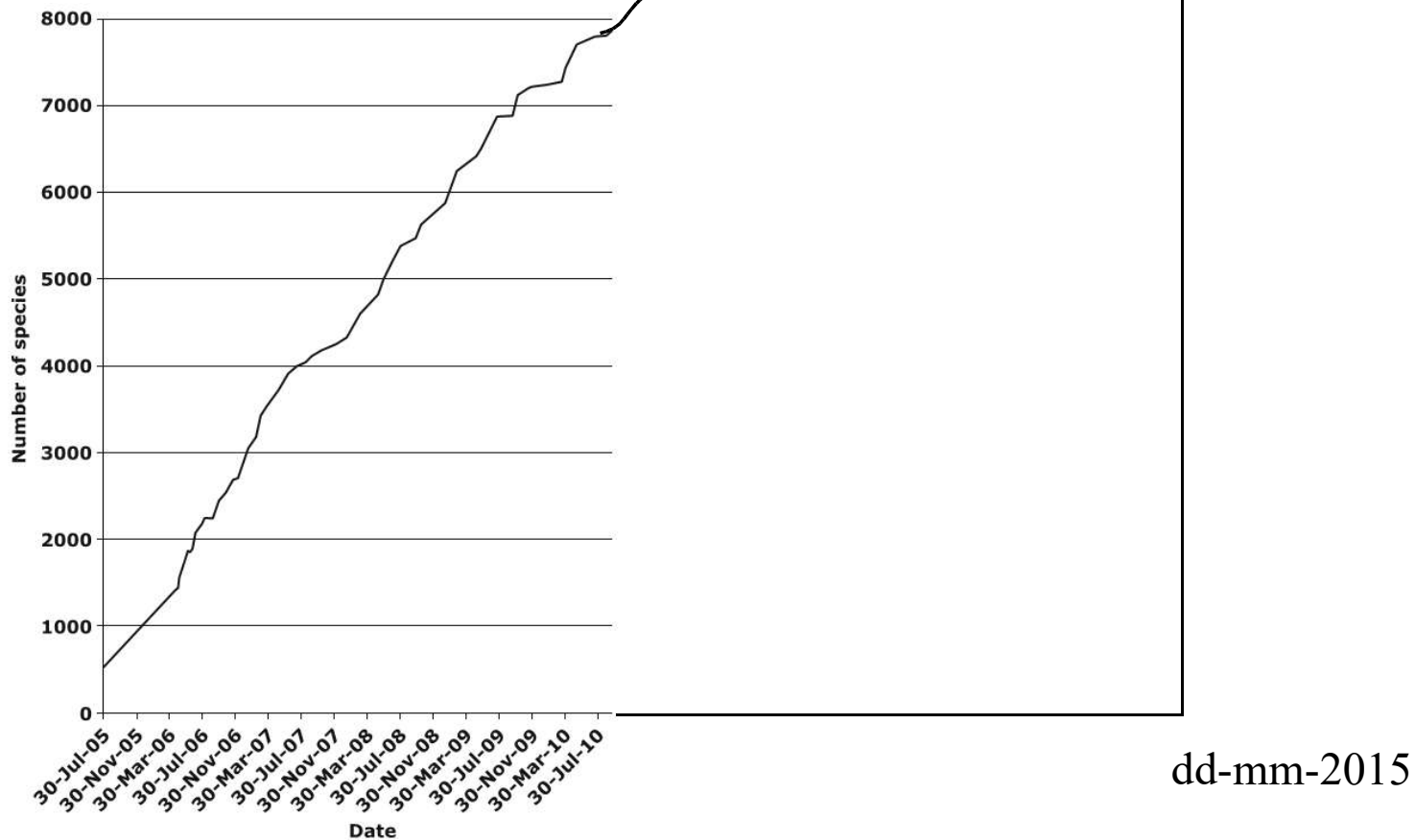


**ДНК-штрихкод** - одна или несколько коротких последовательностей, которые взяты из стандартизированного участка генома с целью использования для идентификации видов.

Для ДНК-штрихкодирования **наземных растений** в настоящее время в качестве стандарта используются гены хлоропластов **рибулозобисфосфаткарбоксилаза *rbcL*** и **матураза *matK***, для **грибов** – ***ITS***, для большинства **многоклеточных животных**, в том числе и рыб, лучше всего зарекомендовал себя 5'- участок митохондриального гена ***Co-1*** длиной ~ **650** п.о.

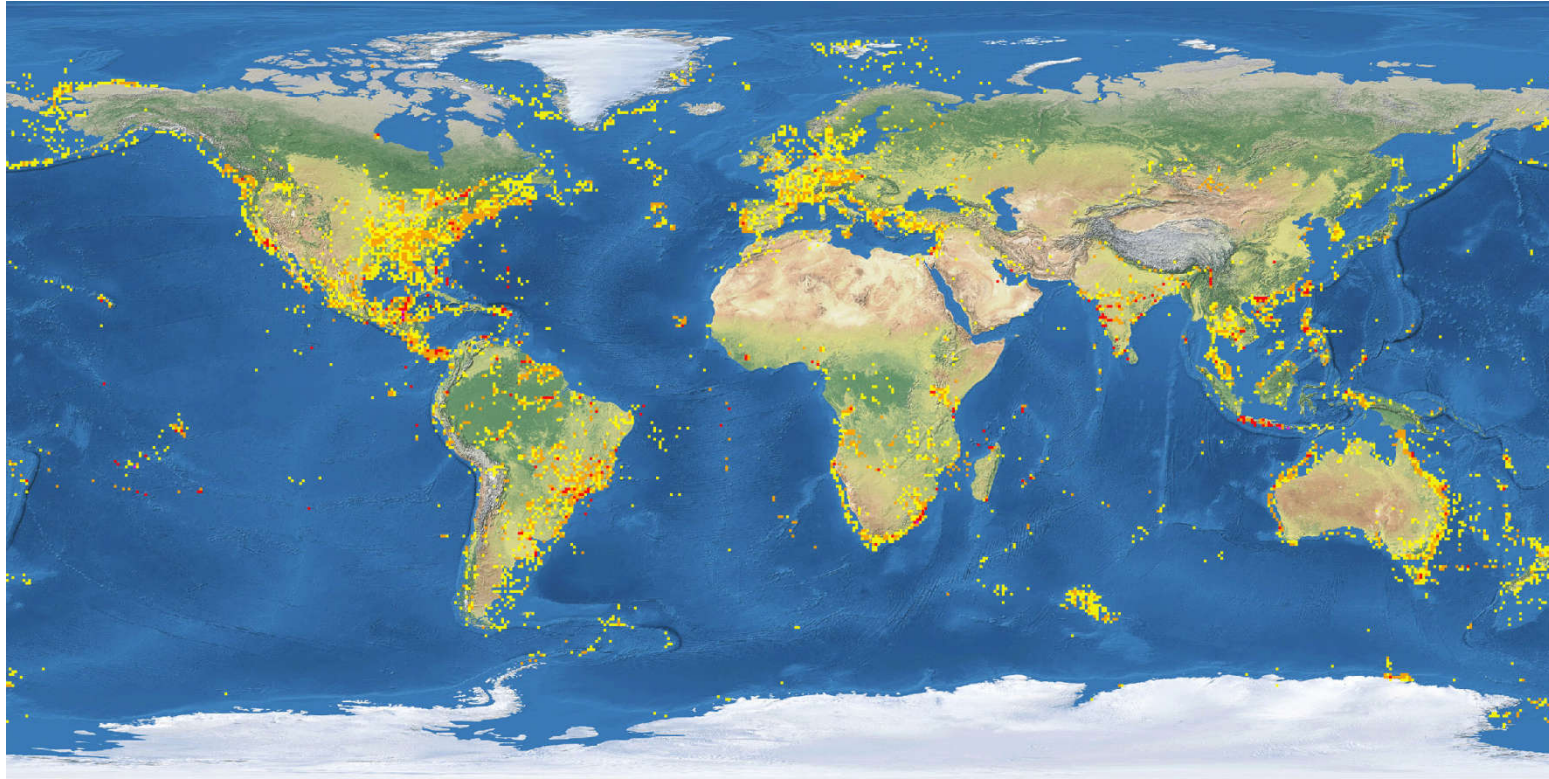


In total, there are likely to be around **40000** extant fish species



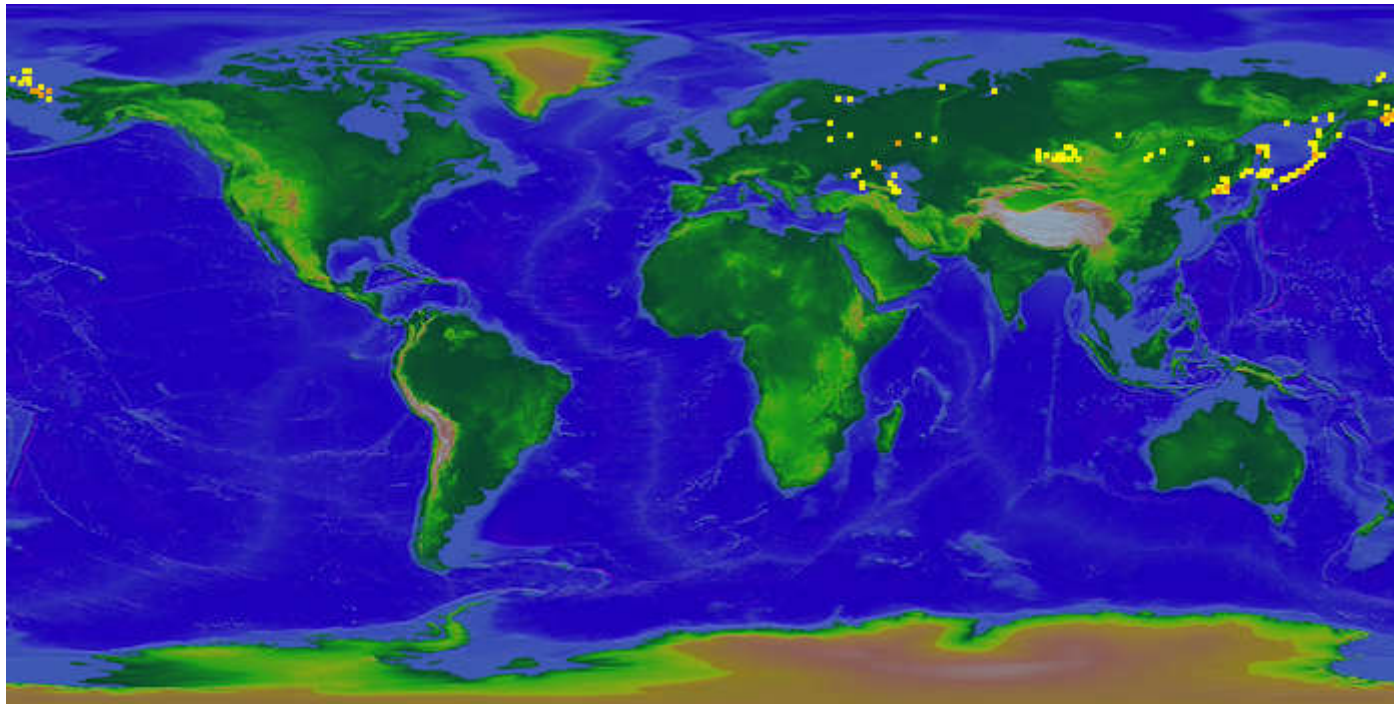
Progress of FISH-BOL, showing numbers of species barcoded by date.





Actinopterygii: Specimens distribution. ~ 226000 points.

Mined from **Barcode of Life Data System**.



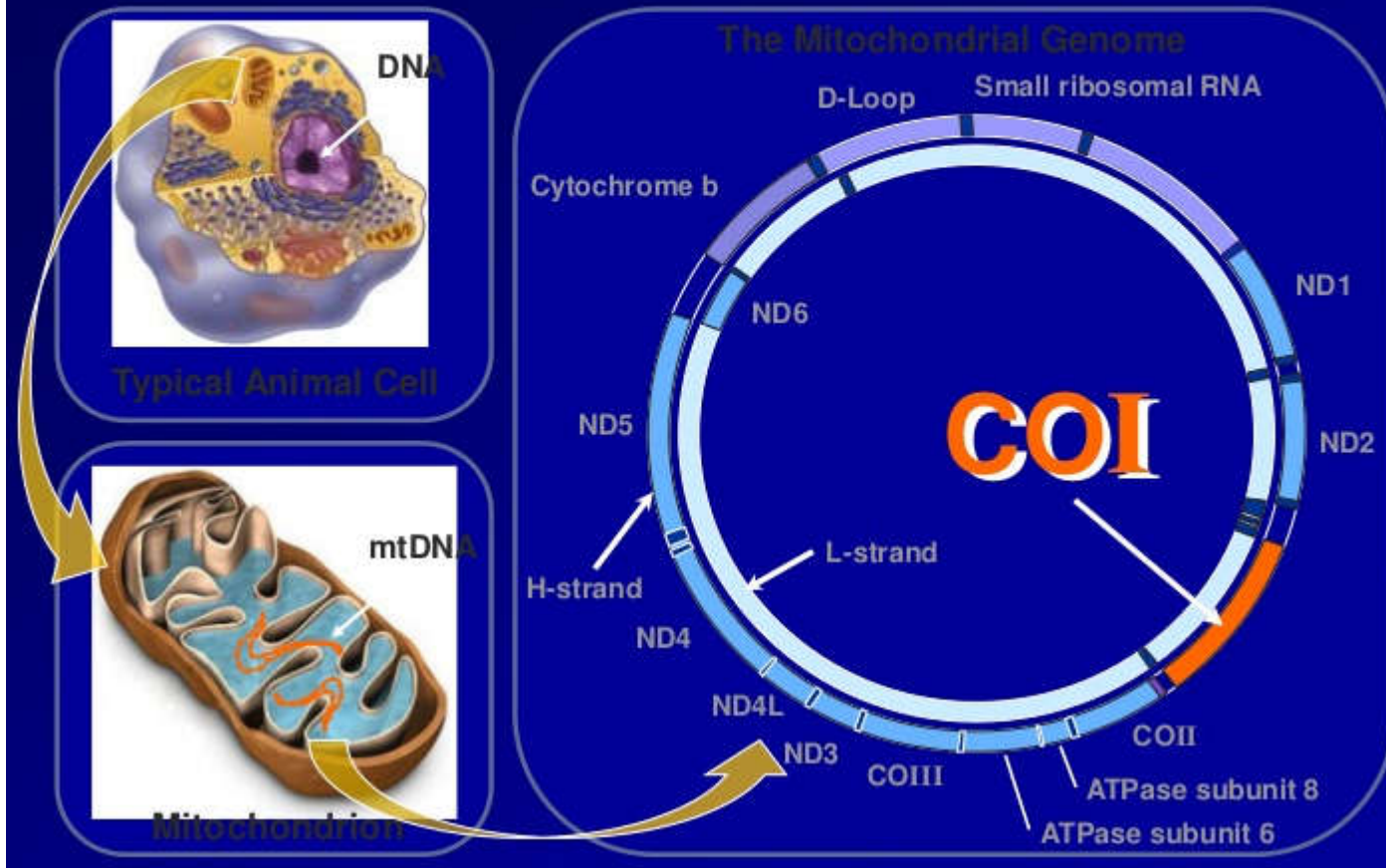
#### INSTITUTIONS [15]

Mined from GenBank, NCBI	481
<u>Institute of Marine Biology, Vladivostok</u>	<u>250</u>
Russian Academy of Sciences, Zoologic...	199
Research Collection of Vadim Birstein	124
Ulyanovsk State Pedagogical University	94
California Academy of Sciences	91
University of Alaska Museum	59
University of Washington, Fish Collec...	29
A.V. Zhirmunsky Institute of Marine B...	29
Biodiversity Institute of Ontario	7

Actinopterygii: Specimens distribution. Search restricted to Russian territory. 893 points.  
Mined from **Barcode of Life Data System**.



# An Internal ID System for All Animals



DNA barcode for fishes – mitochondrial gene COI (655 bp).

Наиболее быстрый и часто используемый способ взглянуть на свои данные в **BOLD\*** - анализ генетических расстояний с последующей их визуализацией с помощью дистанционного метода **NJ** (neighbor-joining).

Генетические расстояния:

*p*-расстояние – отношение количества сайтов, в которых две последовательности различаются, к общему количеству сайтов.

**K2P** – коррекция простого генетического расстояния, - гипотеза, предполагающая, что наиболее вероятный сценарий эволюции анализируемых последовательностей подразумевает эквивалентные частоты оснований, но разную скорость для транзиций и трансверсий.

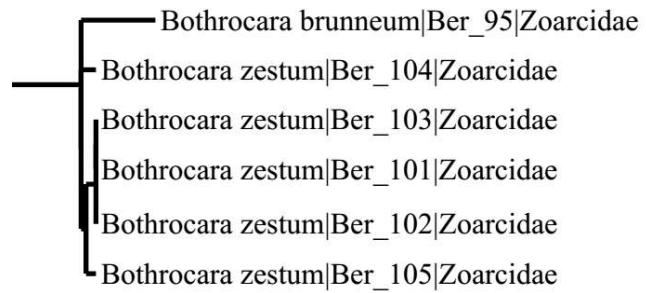
$$\hat{d} = -\frac{1}{2} \ln(1 - 2p - q) - \frac{1}{4} \ln(1 - 2q)$$



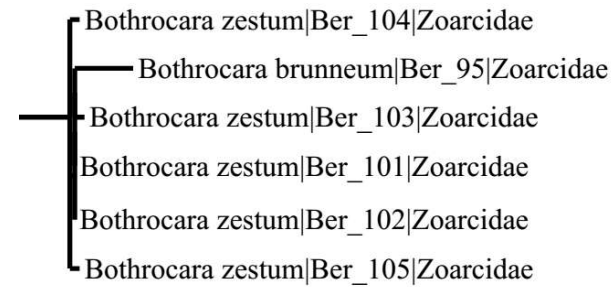
\*Barcode of Life Data System (<http://barcodinglife.com>)

Виды *B. zestum* и *B. brunneum* (Perciformes, Zoarcales, Zoarcidae)

*p*-расстояние



K2P



From:

Виды с незначительной дивергенцией на уровне мтДНК:

Представители рода *Sebastes* северо-восточной части Тихого океана (DNA barcoding of Pacific Canada's Fishes)

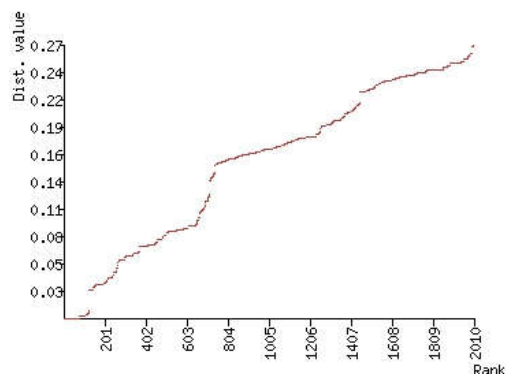
*Sebastes*, *Thunnus* и *Ammodytes* в северо-западной части Атлантики (Barcoding Atlantic Canada's commonly encountered marine fishes)



# Алгоритмы для выделения операционных таксономических единиц (OTU)

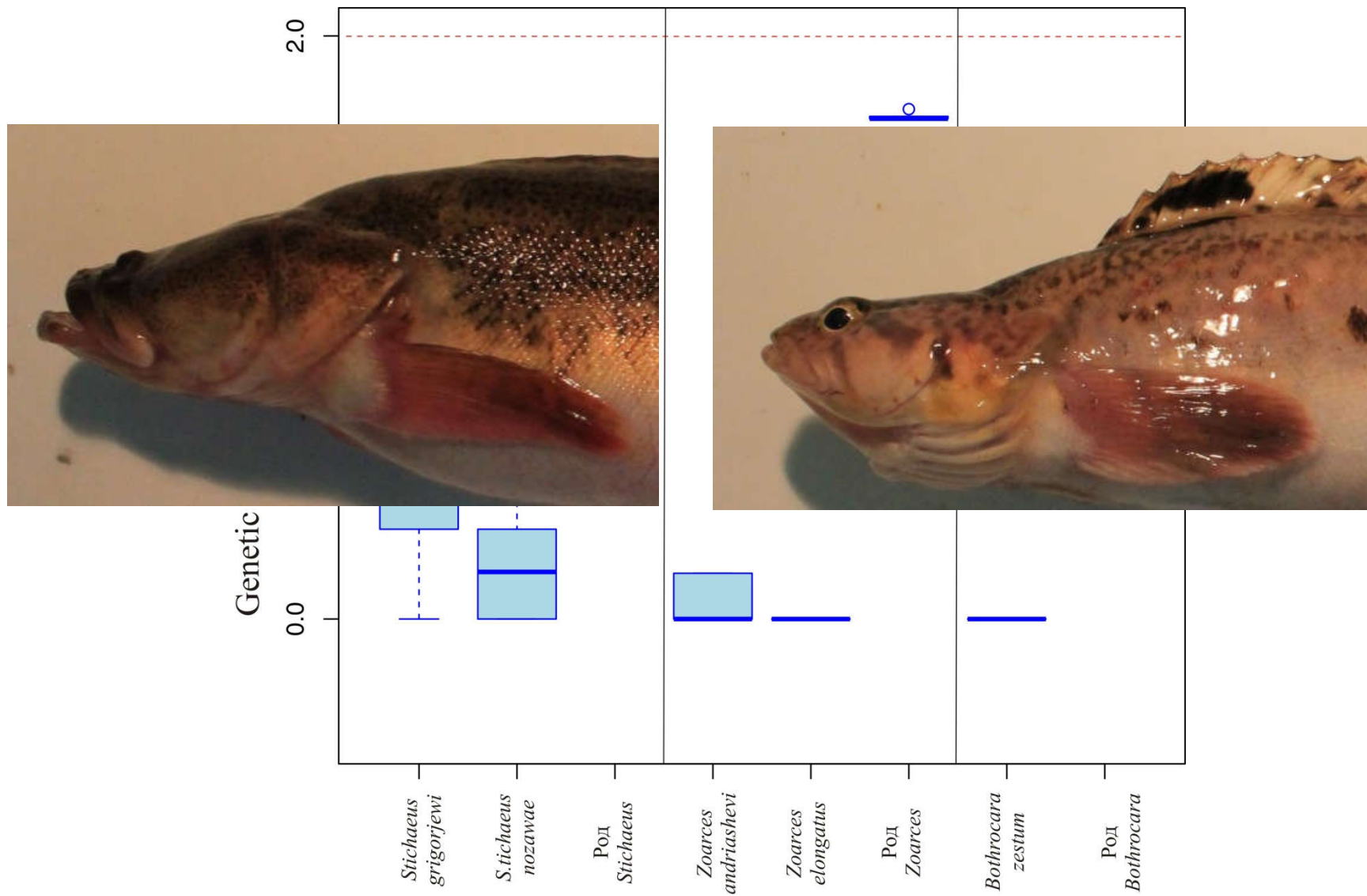


Automatic Barcode Gap Discovery for primary species delimitation (Pulliandre et al., 2011).



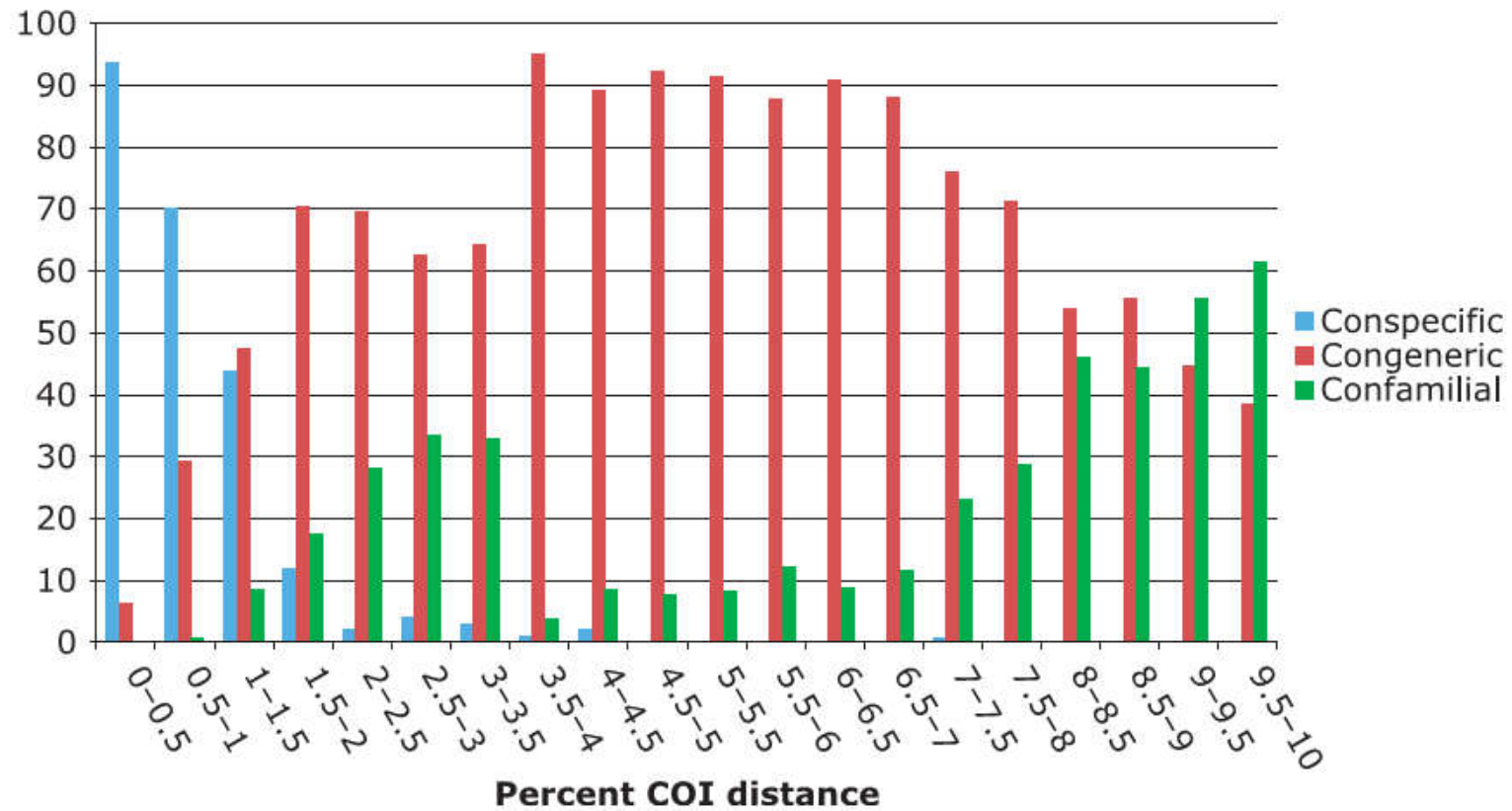
General Mixed Yule Coalescent model (GMYC).

С использованием предварительно построенной филогении на основе критерия максимального правдоподобия находит на дереве порог, разграничивающий меж- и внутривидовые узлы.

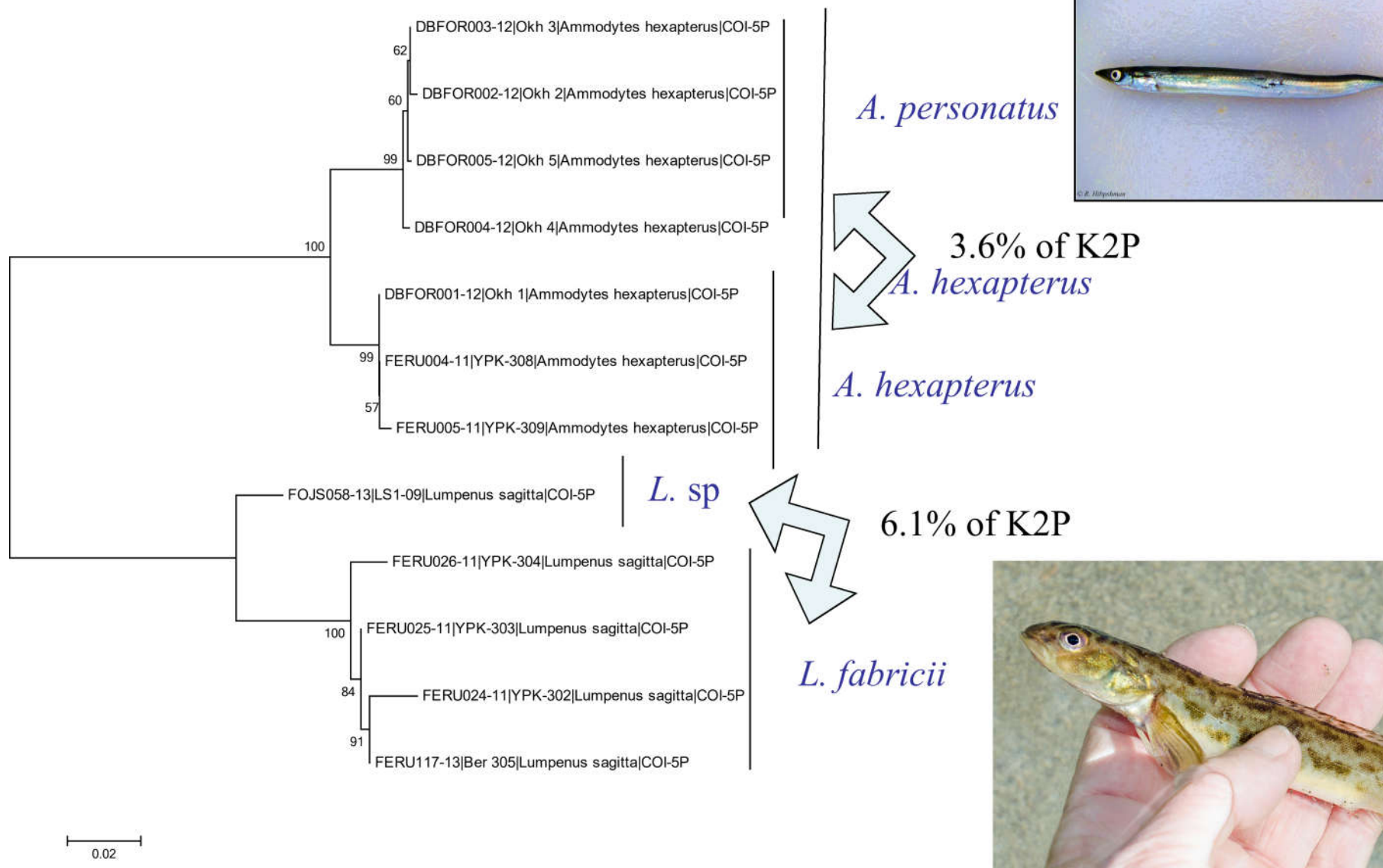


Box plot shows distribution of intraspecific and interspecific genetic distances for those genera whose nucleotide variation is concentrated below 2% threshold. Vertical lines are the boundaries showing the “Barcoding gap” condition.

## (b) Fishes



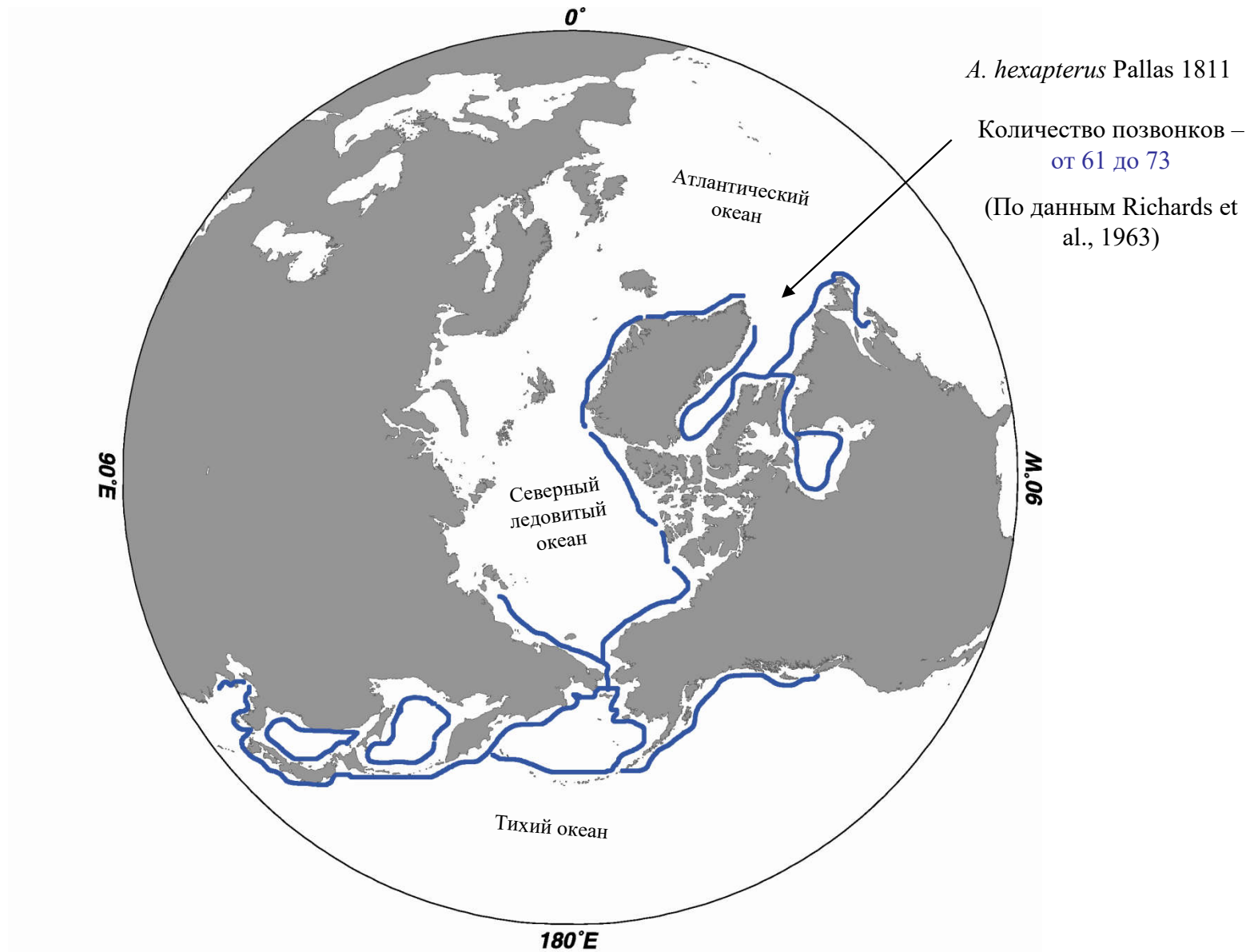
# Криптические виды



NJ-K2P-tree. Illustrates the relationships withing *Lumpenus* and *Ammodytes*.

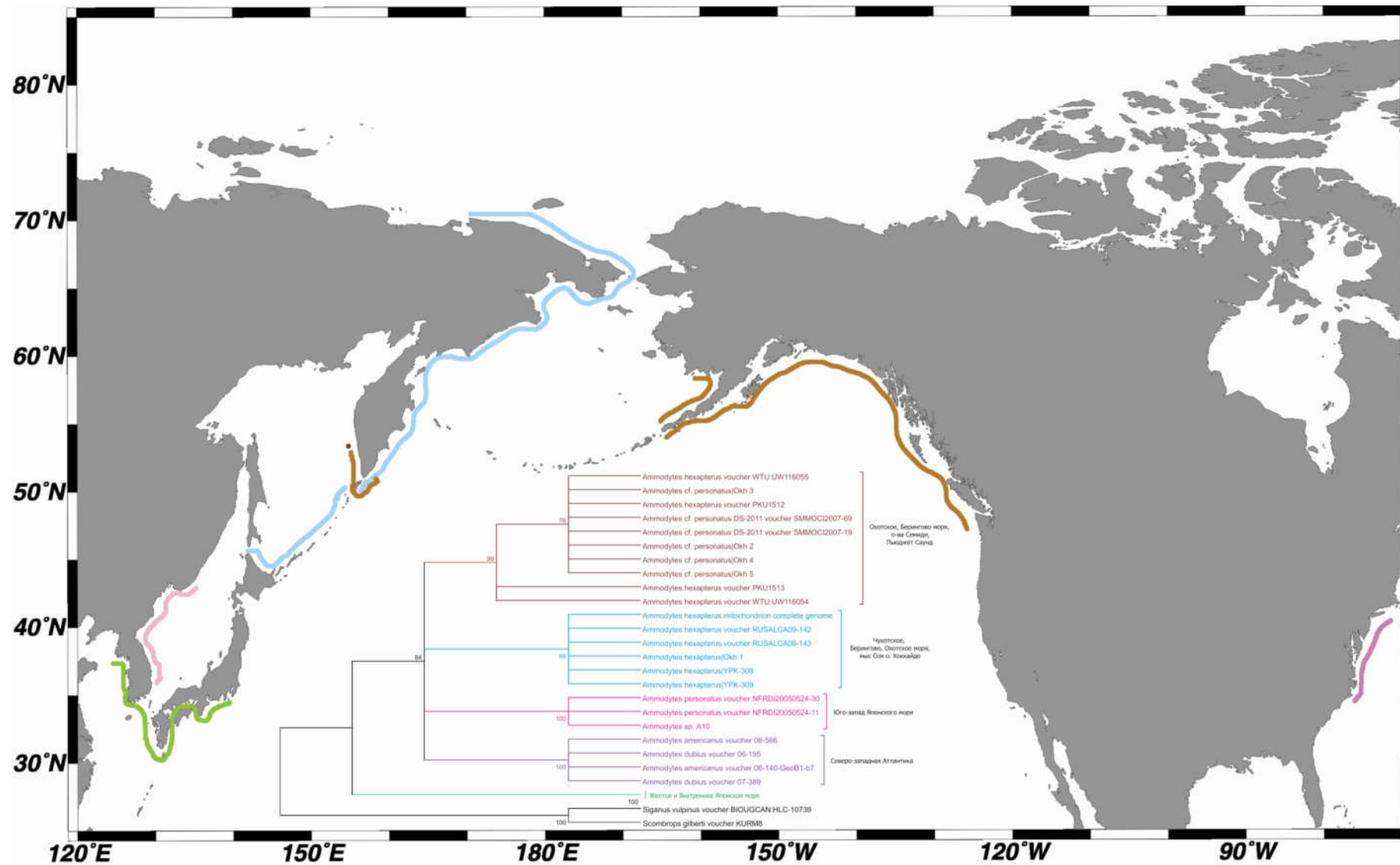


## Криптические виды. *Ammodytes*.



Предположения о распространении *A. hexapterus* (Линдберг, Красюкова, 1975)

# Криптические виды. *Ammodytes*.



*A. personatus*, *A. hexapterus*, *A. japonicus*, *A. heian*

