



# Taxonomic knowledge in the Catalogue of Life

Yuri Roskov  
Workpackage 3

Interlinking the world's biodiversity catalogues  
i4Life Workshop at the Biodiversity Informatics Horizons 2013 conference  
Rome, 6th September 2013

Synthetic knowledge; hypotheses and concepts in taxonomy

## Foundation stones of the CoL

- Global Species Databases (GSDs)
- Regional Species Databases
- Thematic Databases
- Proto-GSDs

Expertise, based on validly published taxonomic concepts; author's view; editorial assembly; qualifiers and completeness

# Selection & quality control

Active and passive trawling

Assessment by Executive Editor

(incl. consultations with Editorial Panel and  
independent taxonomists)

Peer Review\*

three independent reviews

technical report (data vs CoL Standard Dataset)

pre-adoption by Editorial Panel

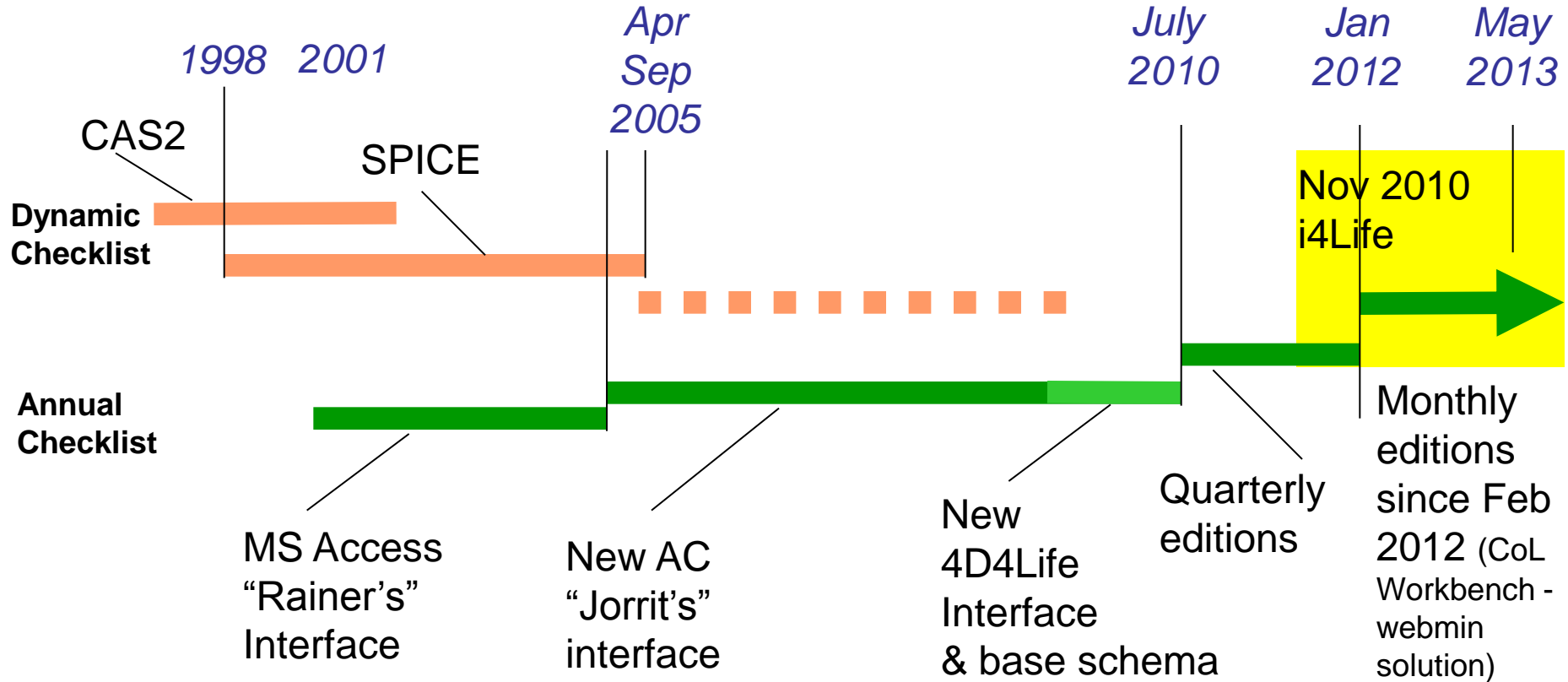
adoption by the Global Team

Data analysis and editorial transformations

(re-shaping, unification and simplification, assimilation  
in CoL classification, resolving nomenclatural and  
taxonomic issues; ref., LTS, metadata)

User's feedback

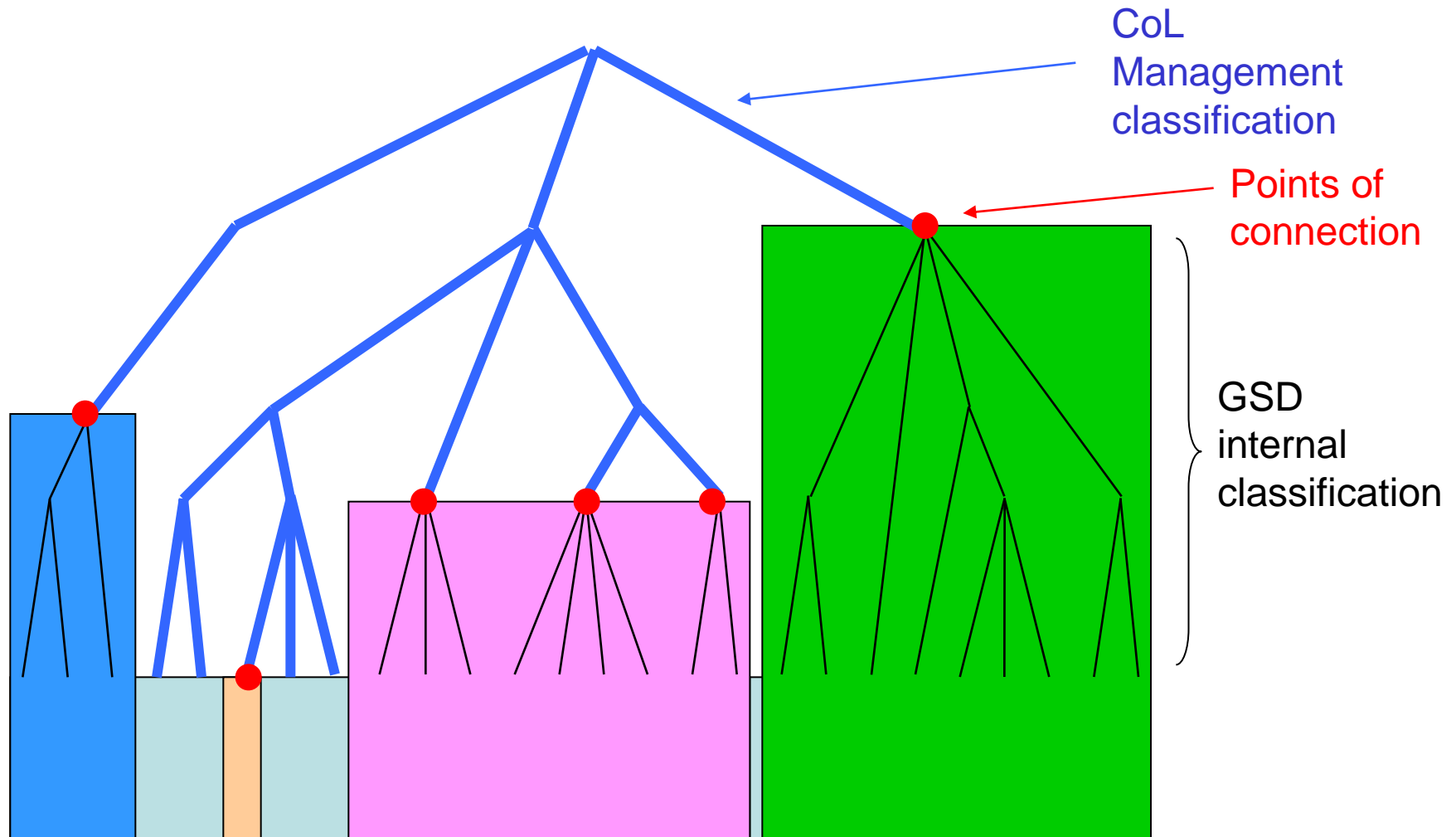
# Dynamic and Annual checklists: unified process



- 18 monthly editions have been released in 2012-13 from March 2012 to August 2013
- 31 WoRMS databases are automated through CoL Workbench templates
- SysMyr, Scarabs, FishBase, ILDIS, LIAS, GCC, EmbiopteraSF, OrthopteraSF, PlecopteraSF had more than one update in a year cycle

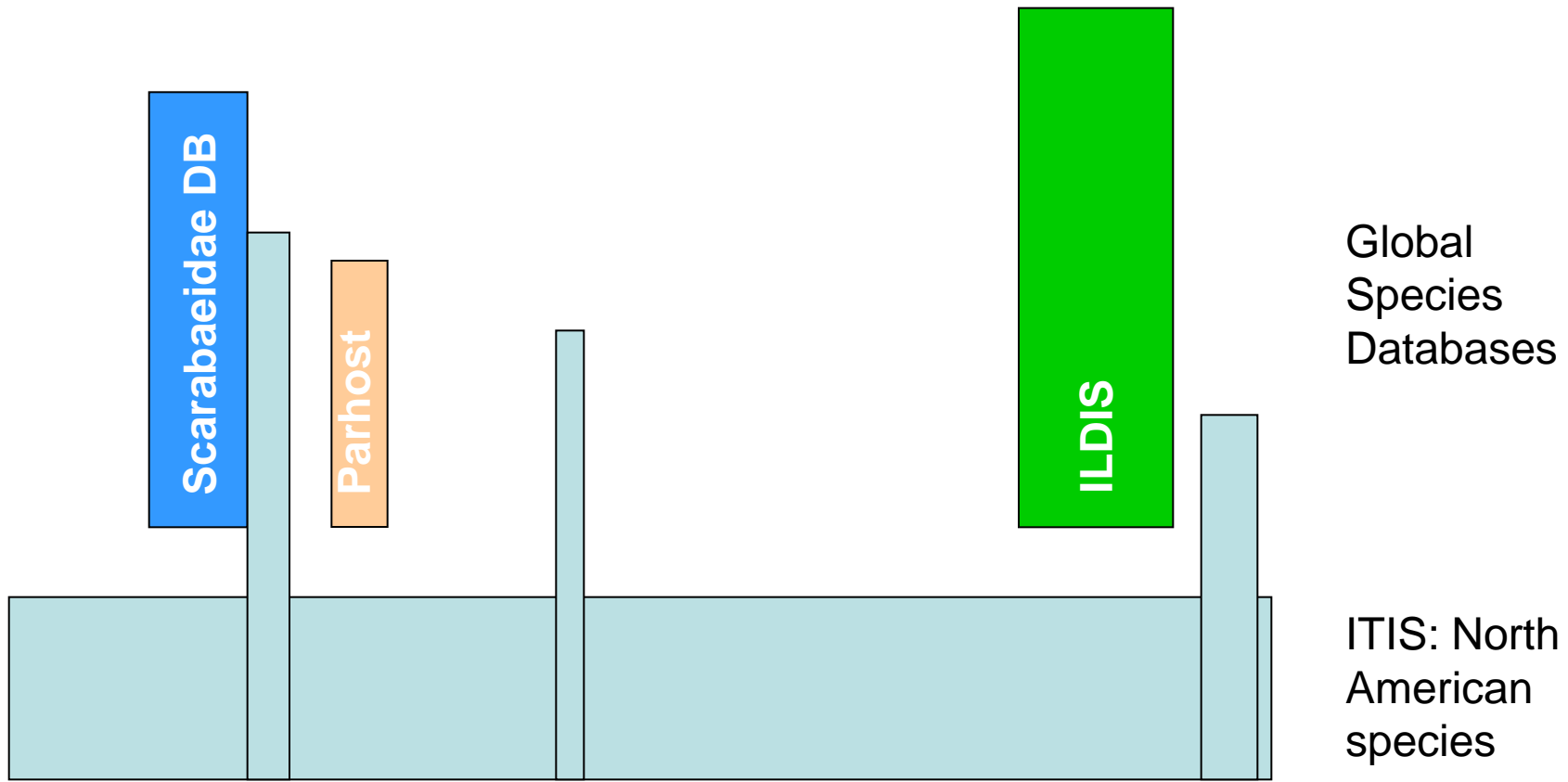
# CoL assembly procedure

## *management classification*



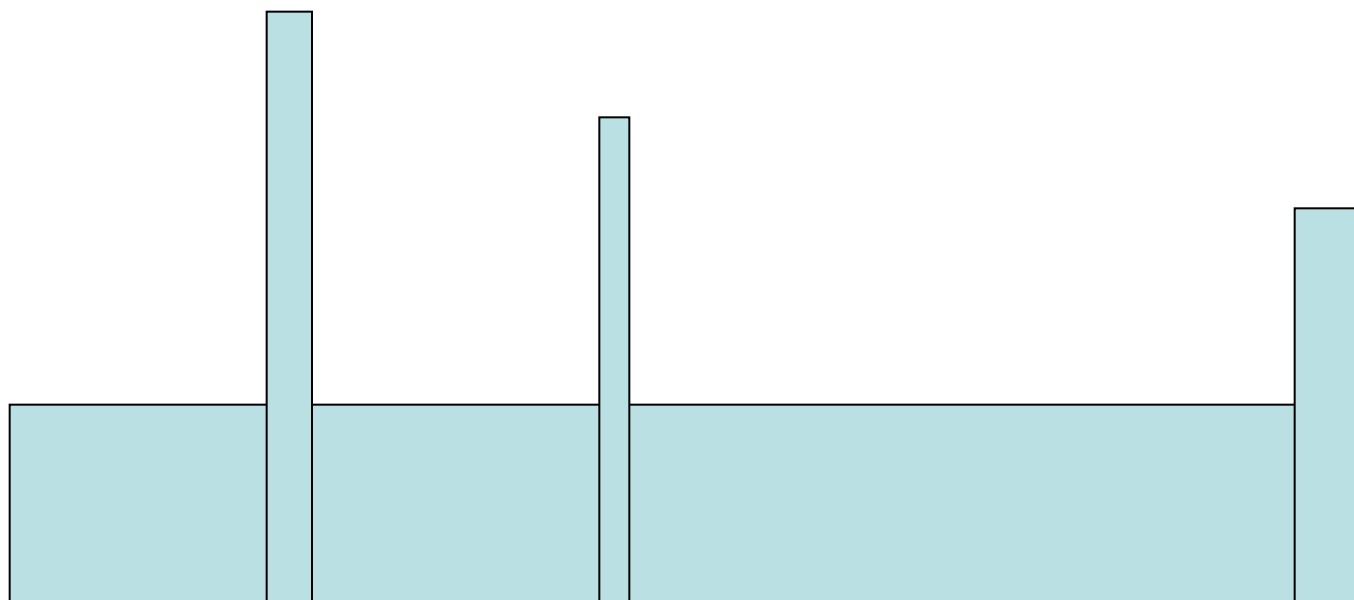


# CoL assembly





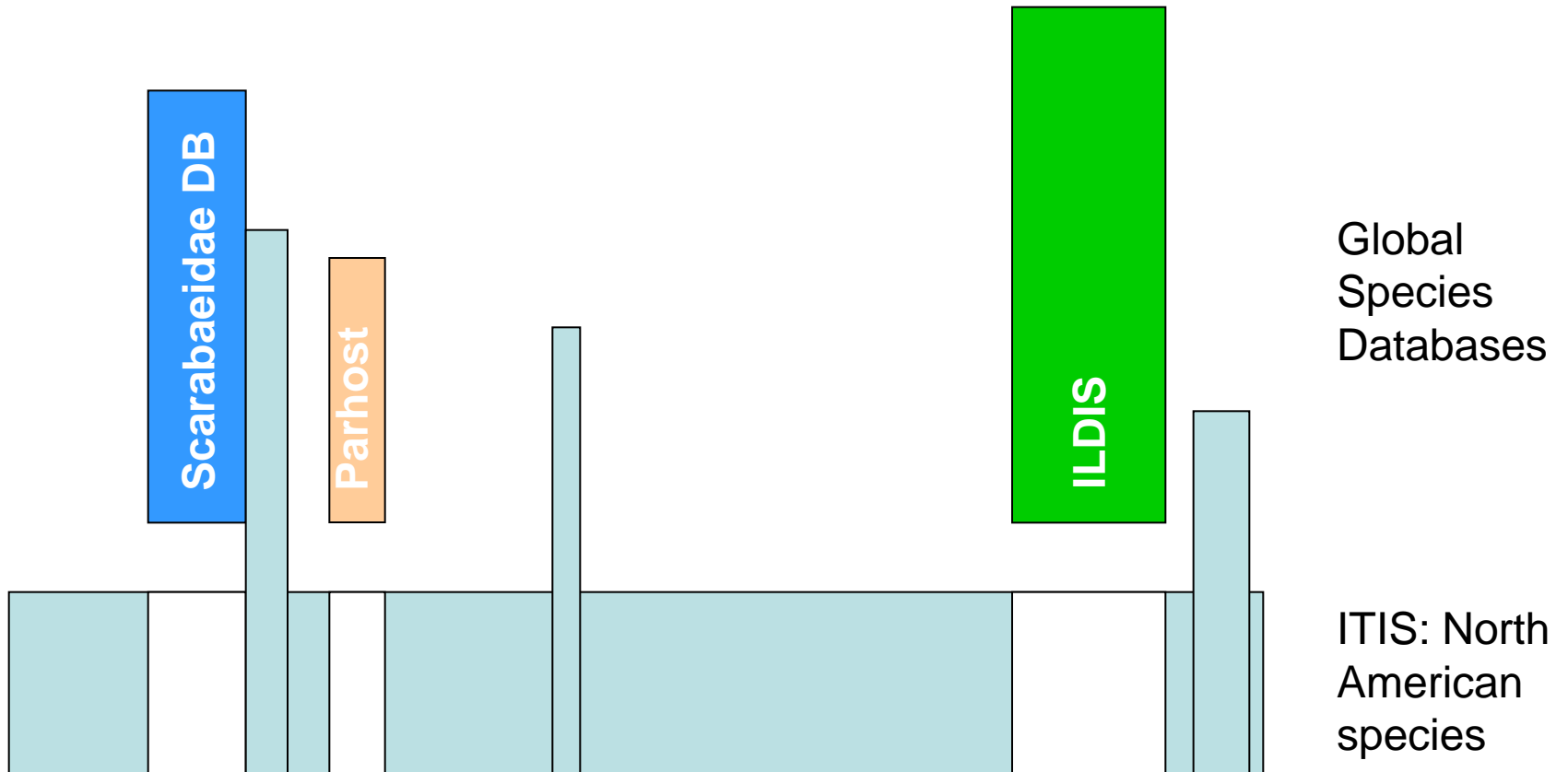
# CoL assembly



ITIS: North  
American  
species



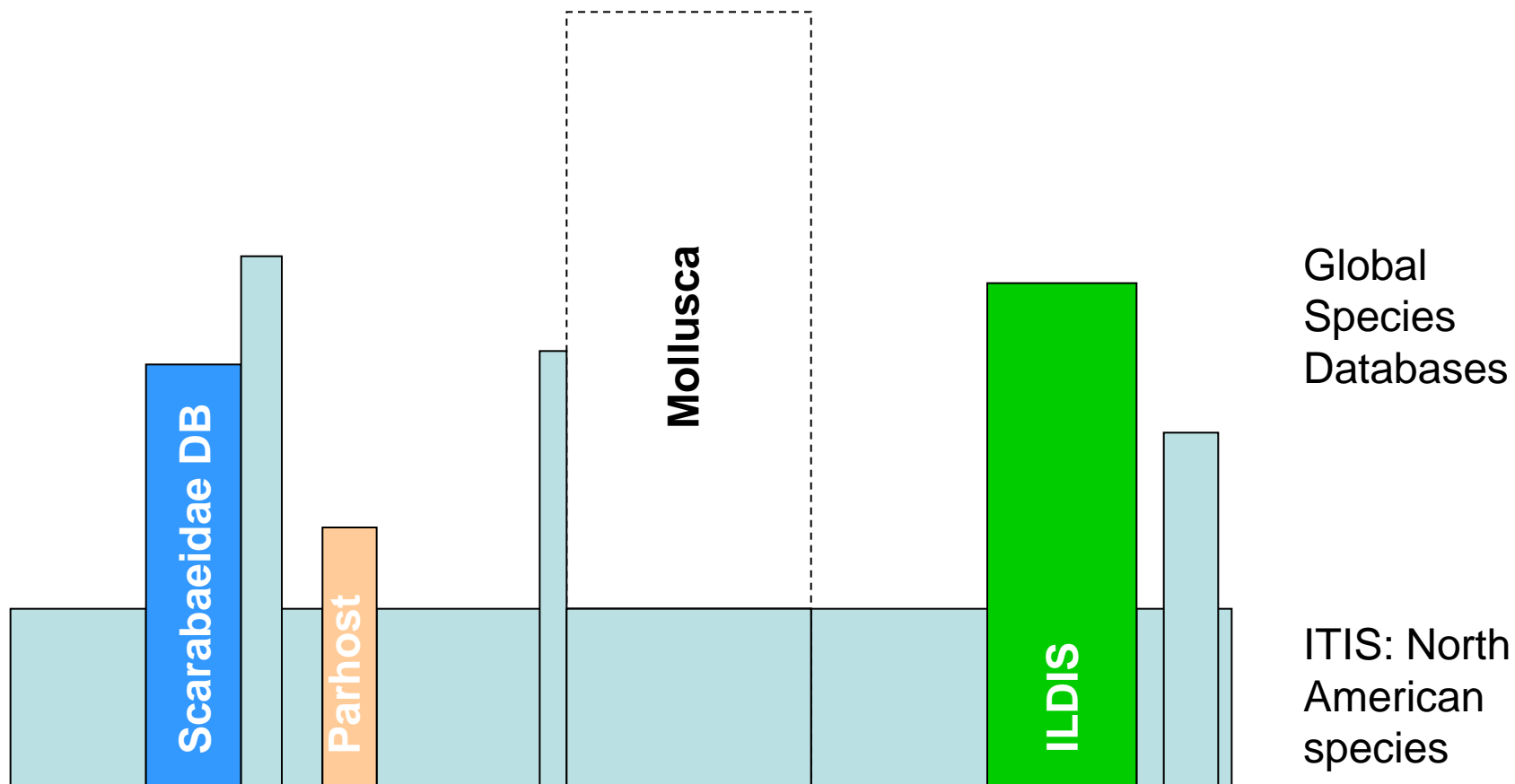
# CoL assembly



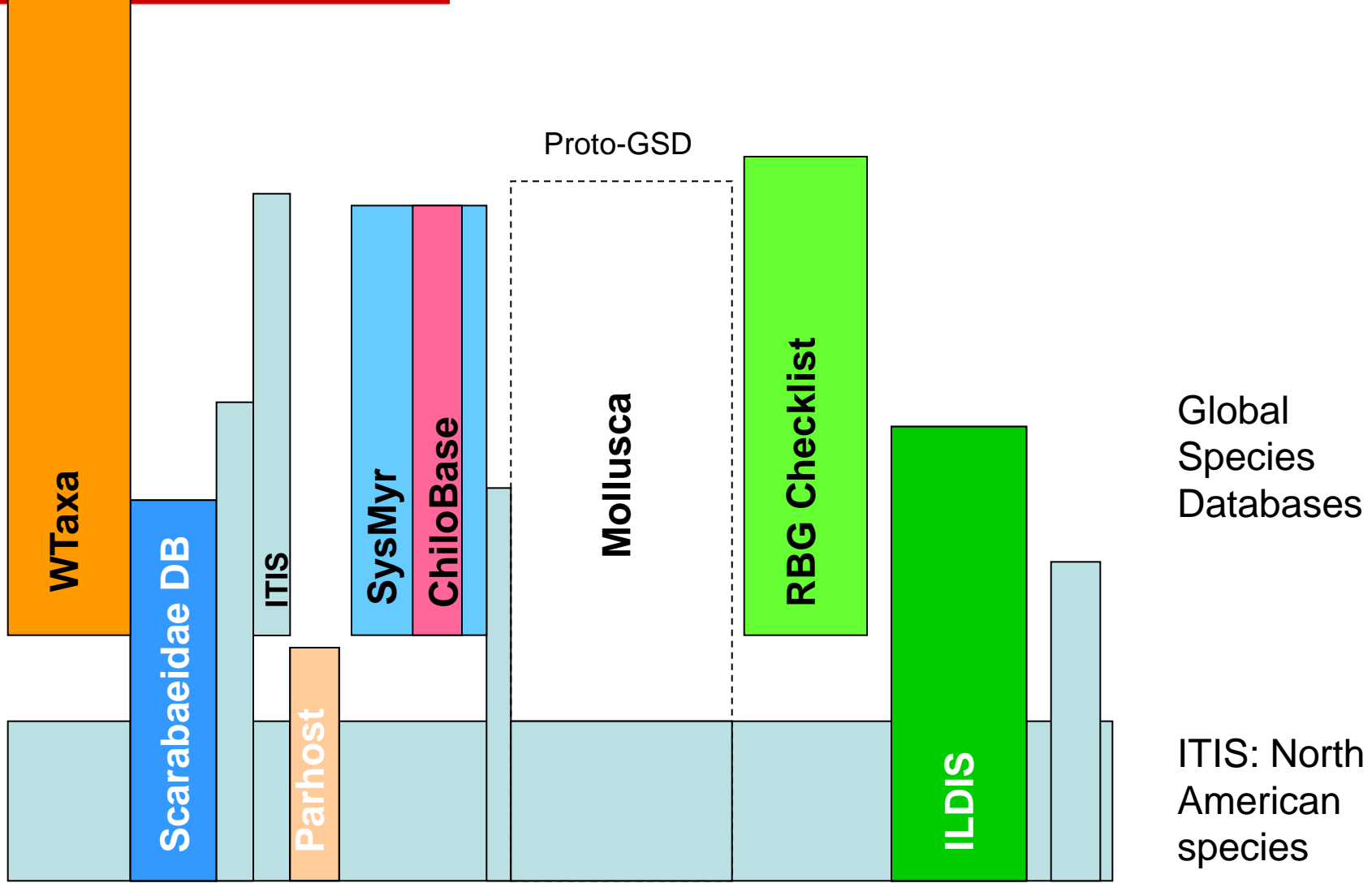




# CoL assembly

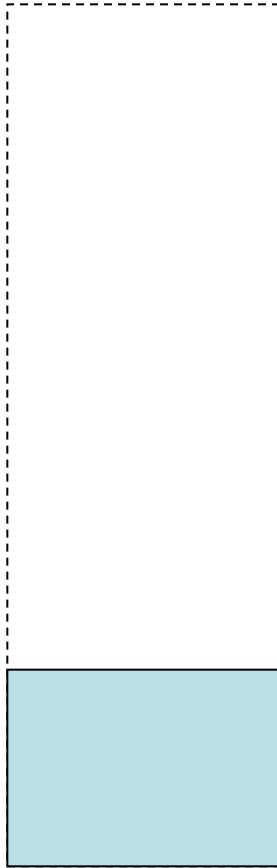


# CoL assembly

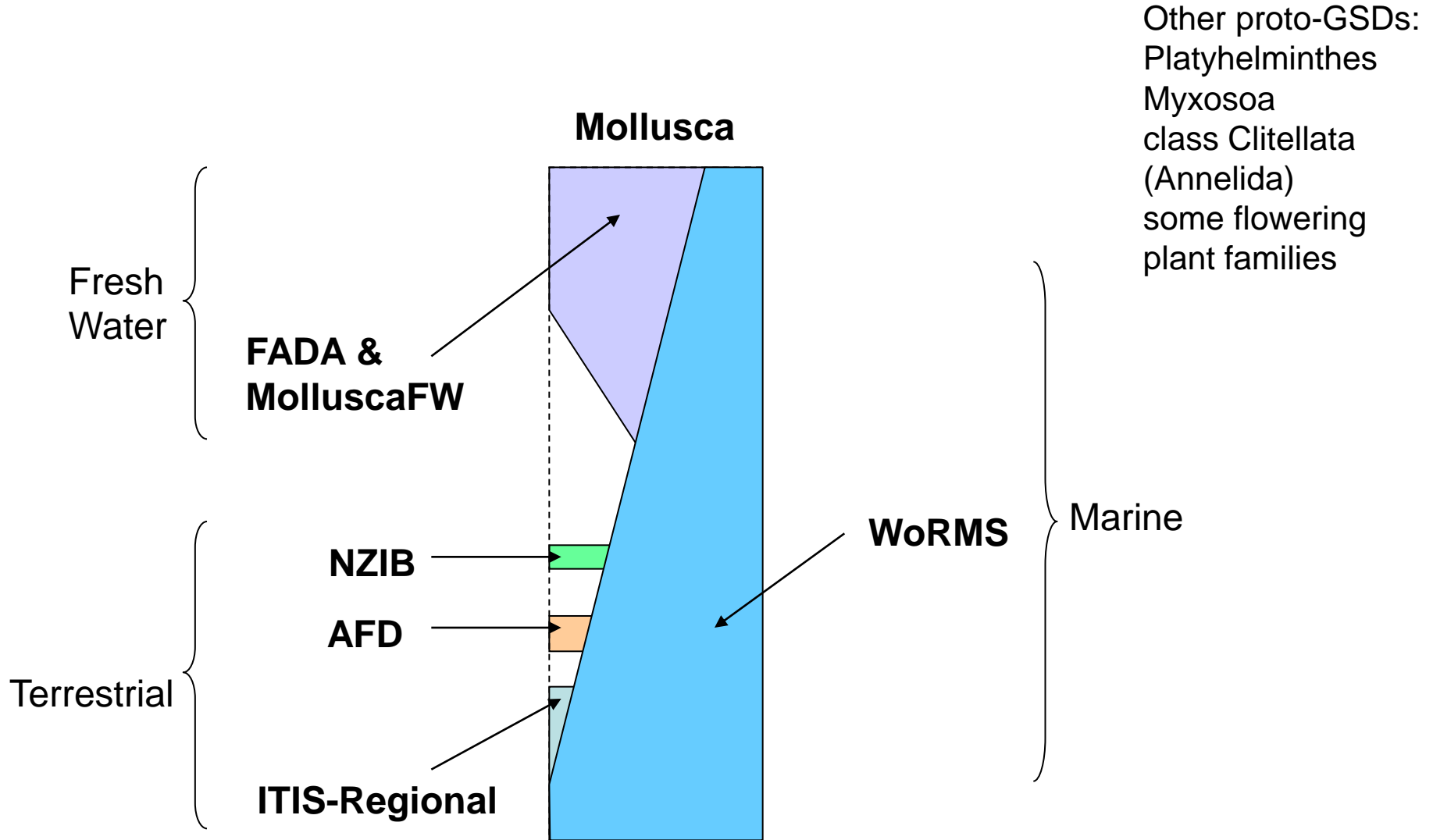


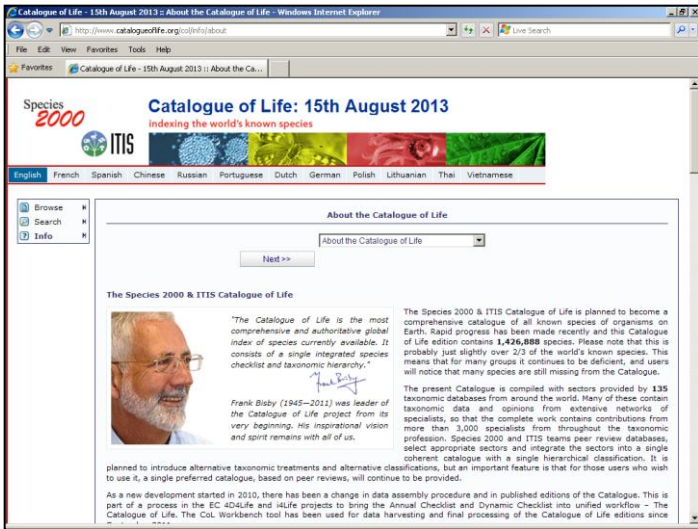
# Proto-GSD example

**Mollusca**



# Proto-GSD example





# The Catalogue of Life, 15<sup>th</sup> August 2013

Monthly editions

**135 databases**

**1,426,888** species belong to

**133,225** genera

**7,826** families,

**1,004** orders

**227** classes

**85** phyla and **7** kingdoms plus viruses;

it also includes **1,036,570** species synonyms  
and **411,816** common names

# CoL taxonomic coverage dynamics

*CoL 3<sup>rd</sup> Jan 2011 – 15<sup>th</sup> Aug 2013*

	<b>CoL 03 Jan 2011</b> 99 databases	<b>CoL 15 Aug 2013</b> 135 databases
<b>Animalia</b>	<b>1,026,635 (72%)</b>	<b>1,034,944 (73%)</b>
<b>Protozoa</b>	<b>6,092 (20% )</b>	<b>11,573 (40%)</b>
<b>Chromista</b>	<b>6,578 (26%)</b>	<b>2,056 (8%)</b>
<b>Fungi</b>	<b>42,767 (43%)</b>	<b>123,129 (98%)</b>
<b>Plantae</b>	<b>227,133 (50%)</b>	<b>245,957 (53%)</b>
<b>Bacteria</b>	<b>9,072 (95%)</b>	<b>6,468 (70%)</b>
<b>Archaea</b>	<b>281 (98%)</b>	<b>281 (98%)</b>
<b>Viruses</b>	<b>2,083 (100% )</b>	<b>2,480 (100%)</b>
<b><i>All biota</i></b>	<b>1,320,641</b>	<b>1,426,888</b>

# Animalia: incomplete phyla (7 from 36)

Acanthocephala

Annelida 77% → 78%

Arthropoda 78% → 79%

Brachiopoda

Bryozoa

Cephalorhyncha

Chaetognatha

Chordata

Cnidaria

Ctenophora

Cycliophora

Echinodermata 70% → 96%

Echiura

Gastrotricha

Gnathostomulida

Hemichordata

Kamptozoa

Kinorhyncha

Loricifera

Micrognathozoa

Mollusca 16% → 45%

Myxozoa → 3% → 19%

Nematoda 14% → 14%

Nematomorpha

Nemertea

Onychophora

Orthonectida

Phoronida

Placozoa

Platyhelminthes 14% → 46%

Porifera

Rhombozoa

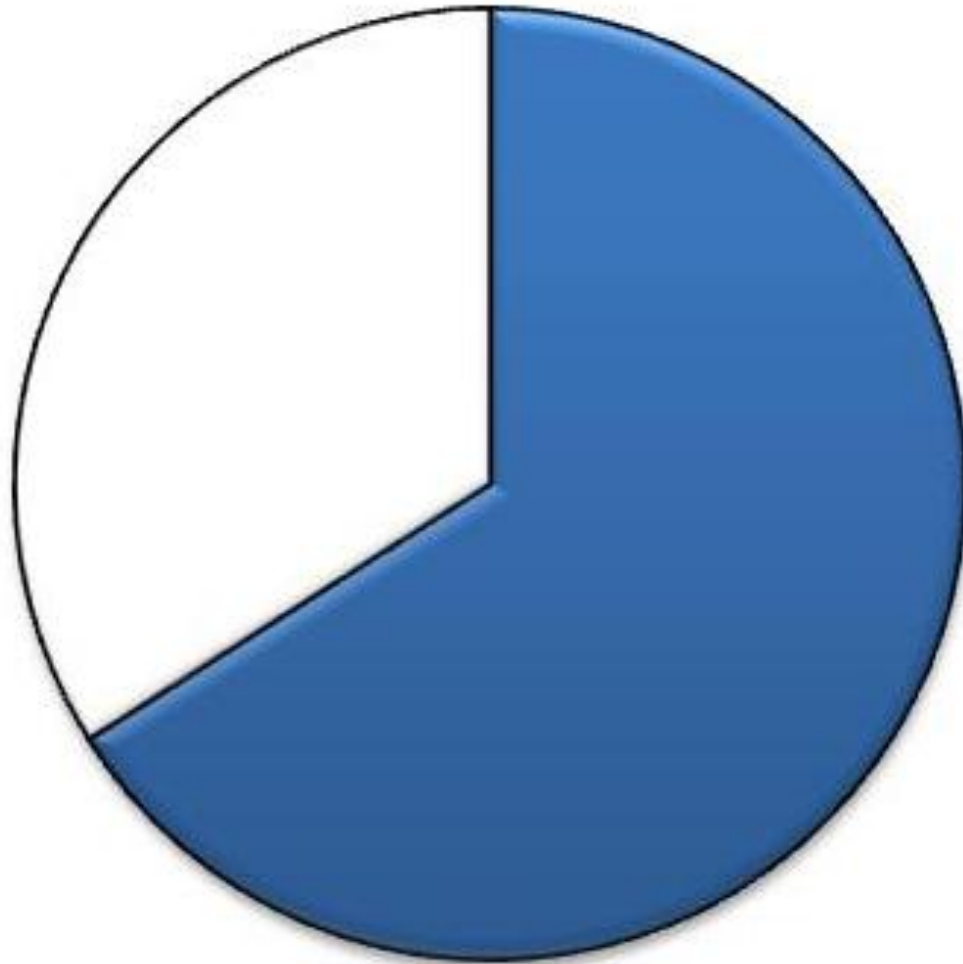
Rotifera

Sipuncula

Tardigrada

Xenacoelomorpha

## How complete is the Catalogue of Life?



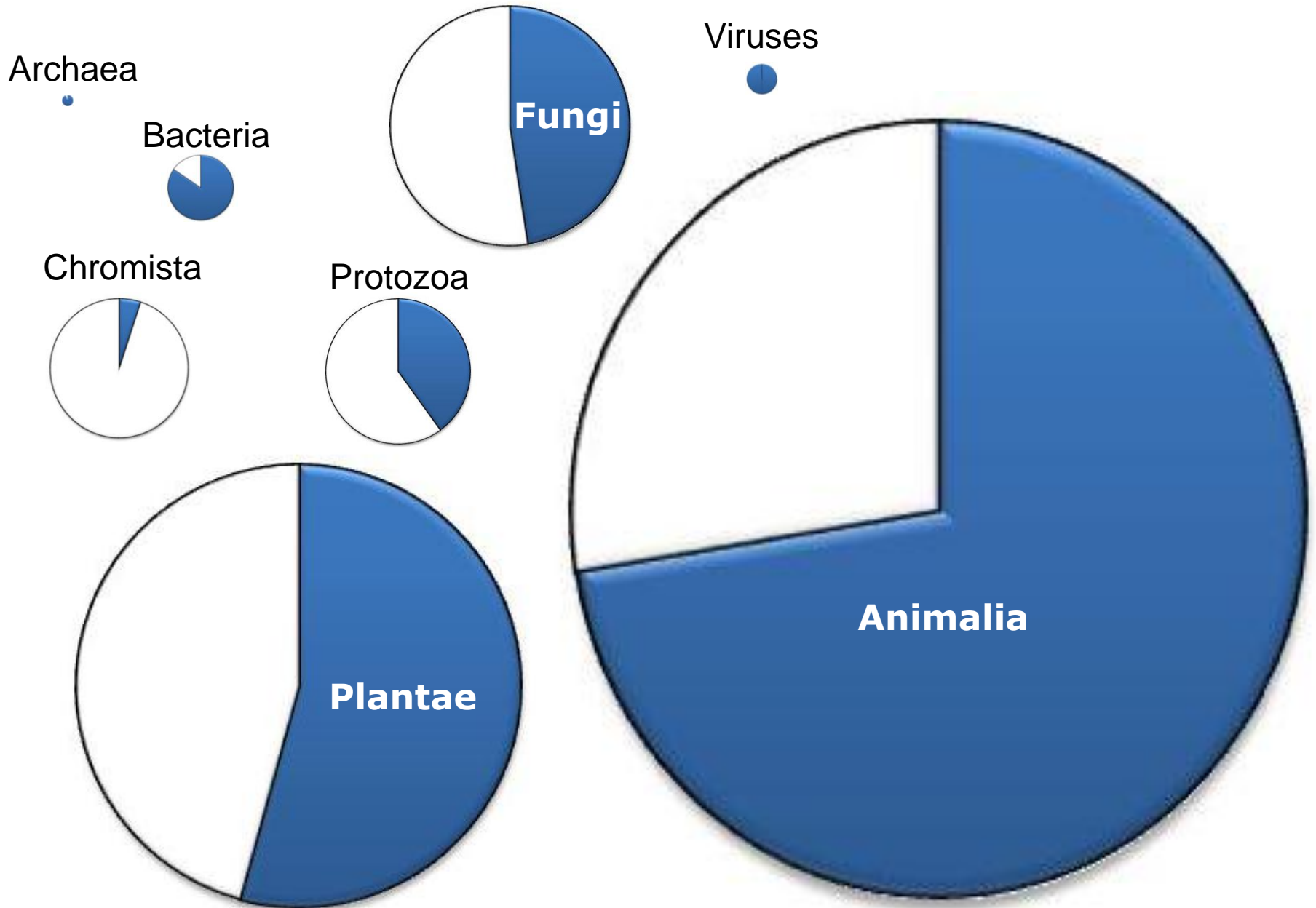
**2013 Annual Checklist**

**71%**

**Whole biota, 1.9 M spp  
(estimations based on  
Arthur Chapman, 2009)**



# How complete is the 2013 Annual Checklist?



**Thank you!**