Why so pale and wan: Comparative genomics of a new cavefish in Europe

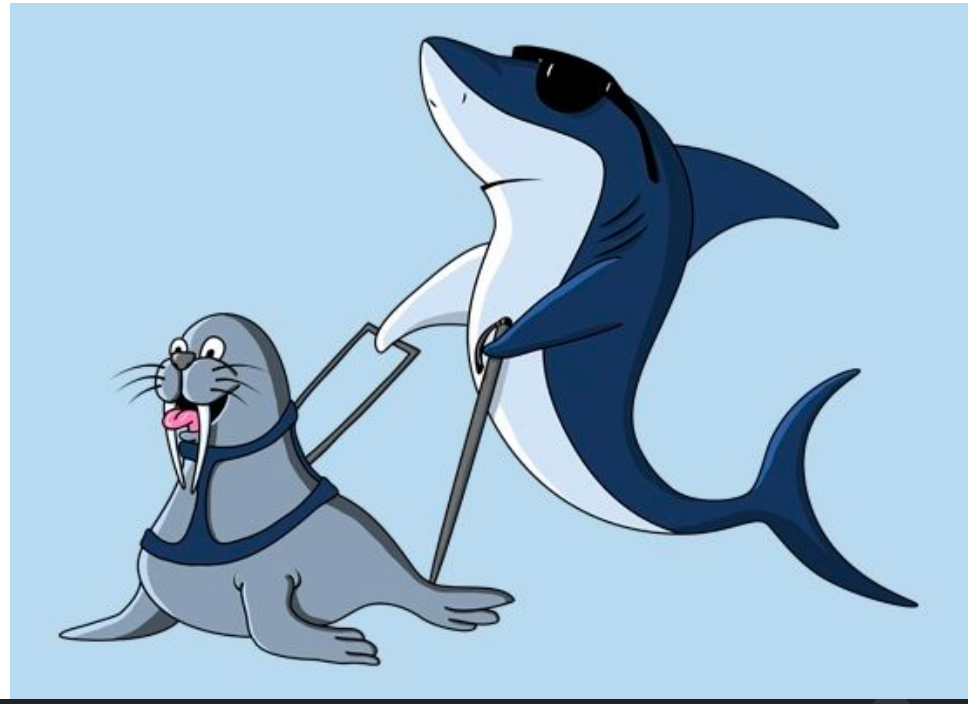
Fritz Sedlazeck

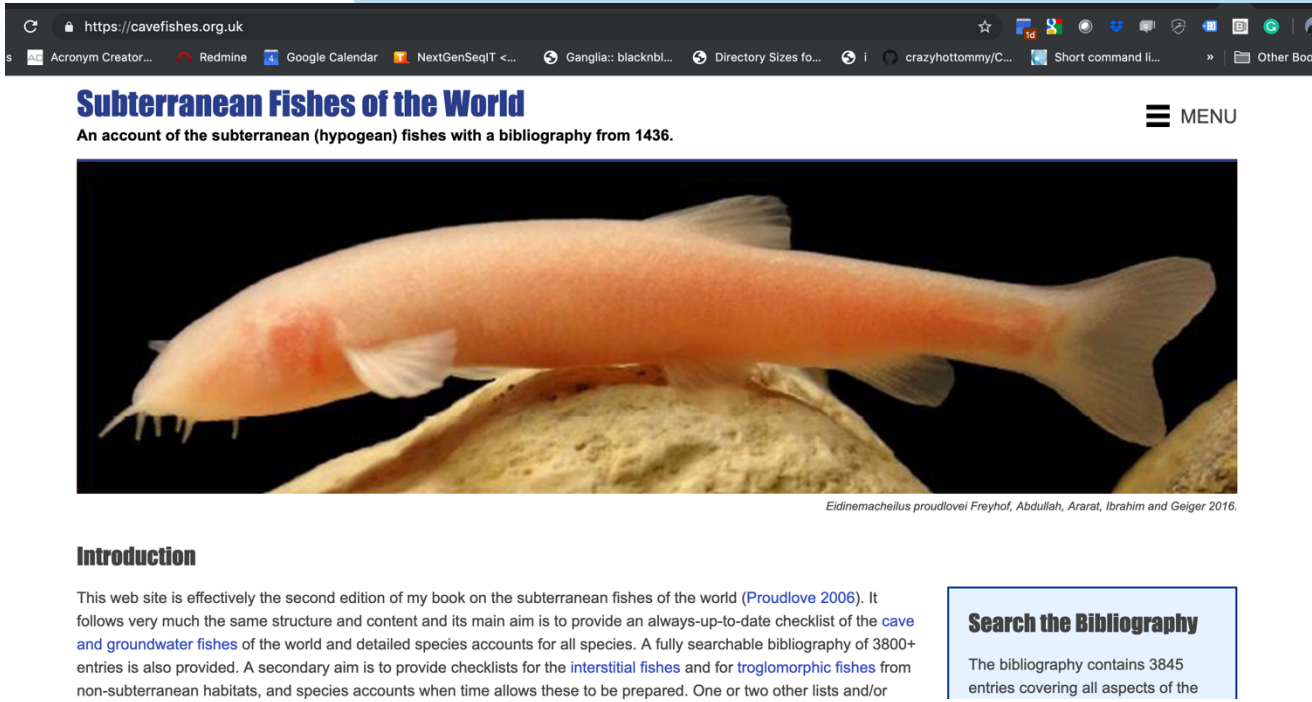
May, 7, 2019

Why cave fishes?

- Study mutations in
  - Sight
  - Sensors
  - Metabolism
  - Pigmentation

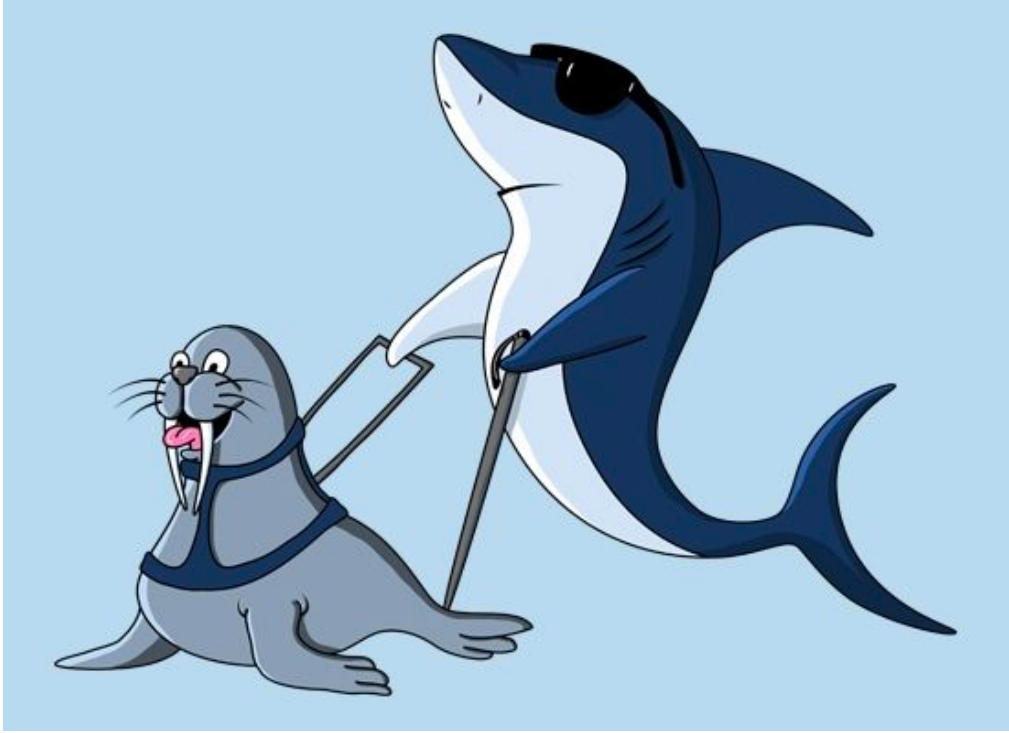
- ~326 cave fish species
  - https://cavefishes.org.uk/

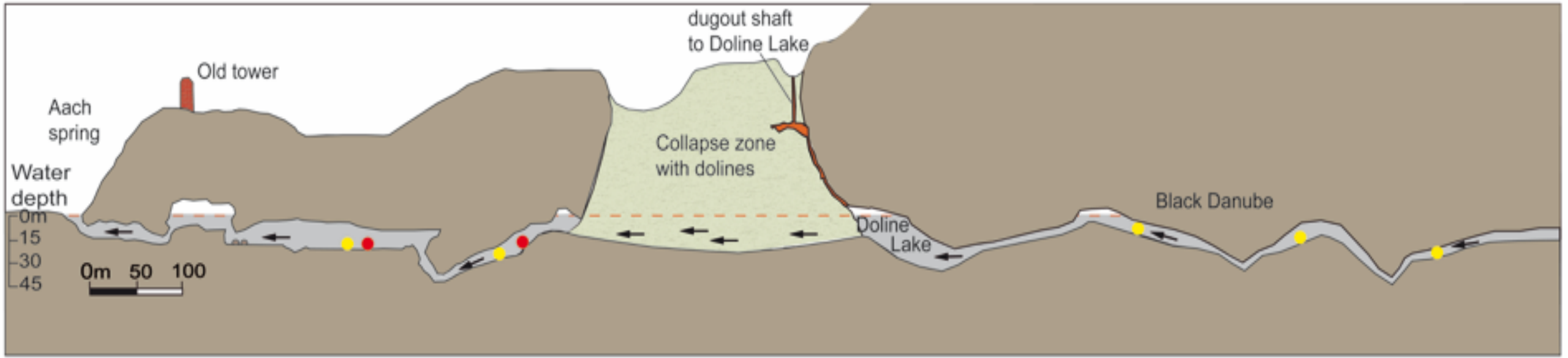




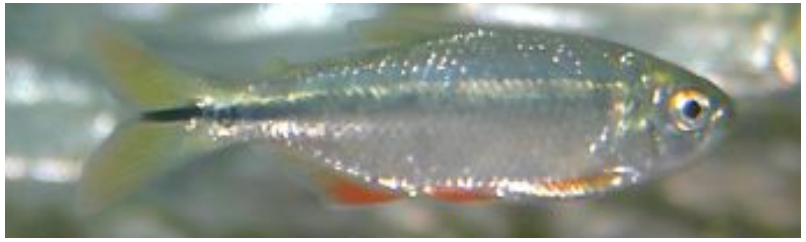
The first European cave fish

- *Barbatula* (stone loach)
- Discovered in 2015 in Southern Germany





Our cavefish



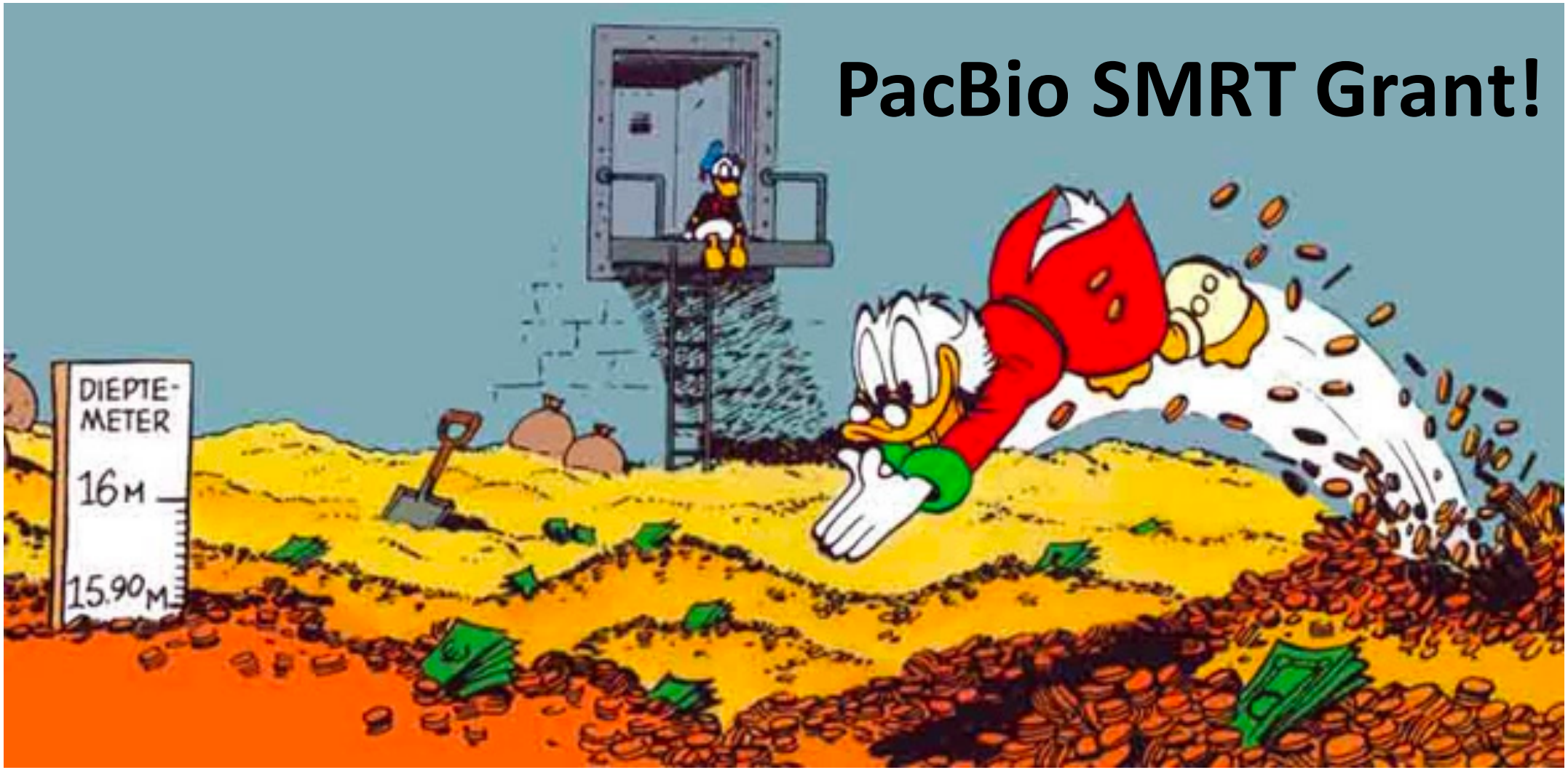
4.6-1.12 million years

~ 12,000 years

Our cavefish (~12,000 years)

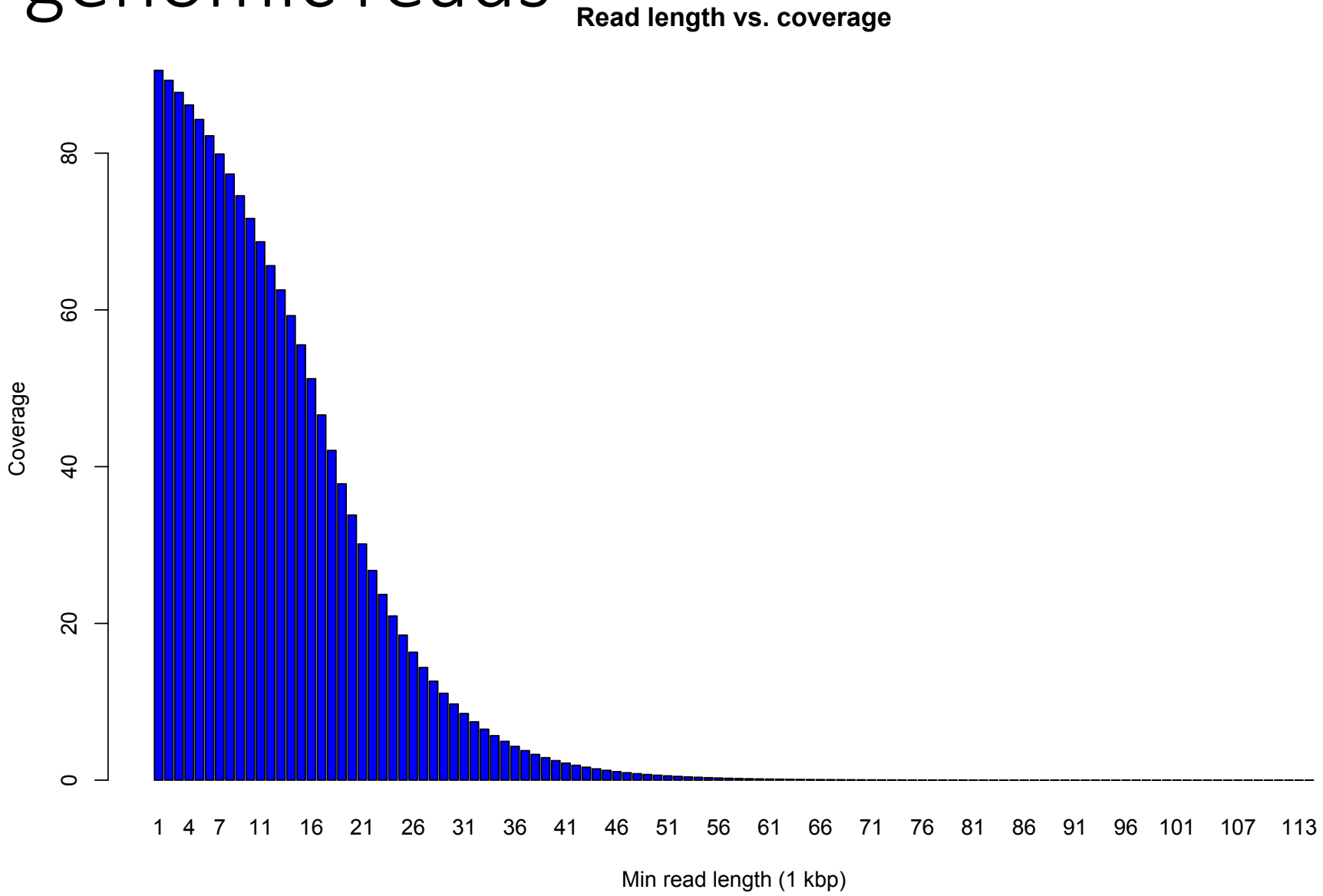
- 400k – 40k Neanderthal
- 13k Giant Sloth extinct.
- 13k Aleppo was founded
- 12k our cave fish
- 11k Jericho was founded
- 10-12k domestication of goat and cattle.
- Pyramids: ~6-7k

What do we need?

- 

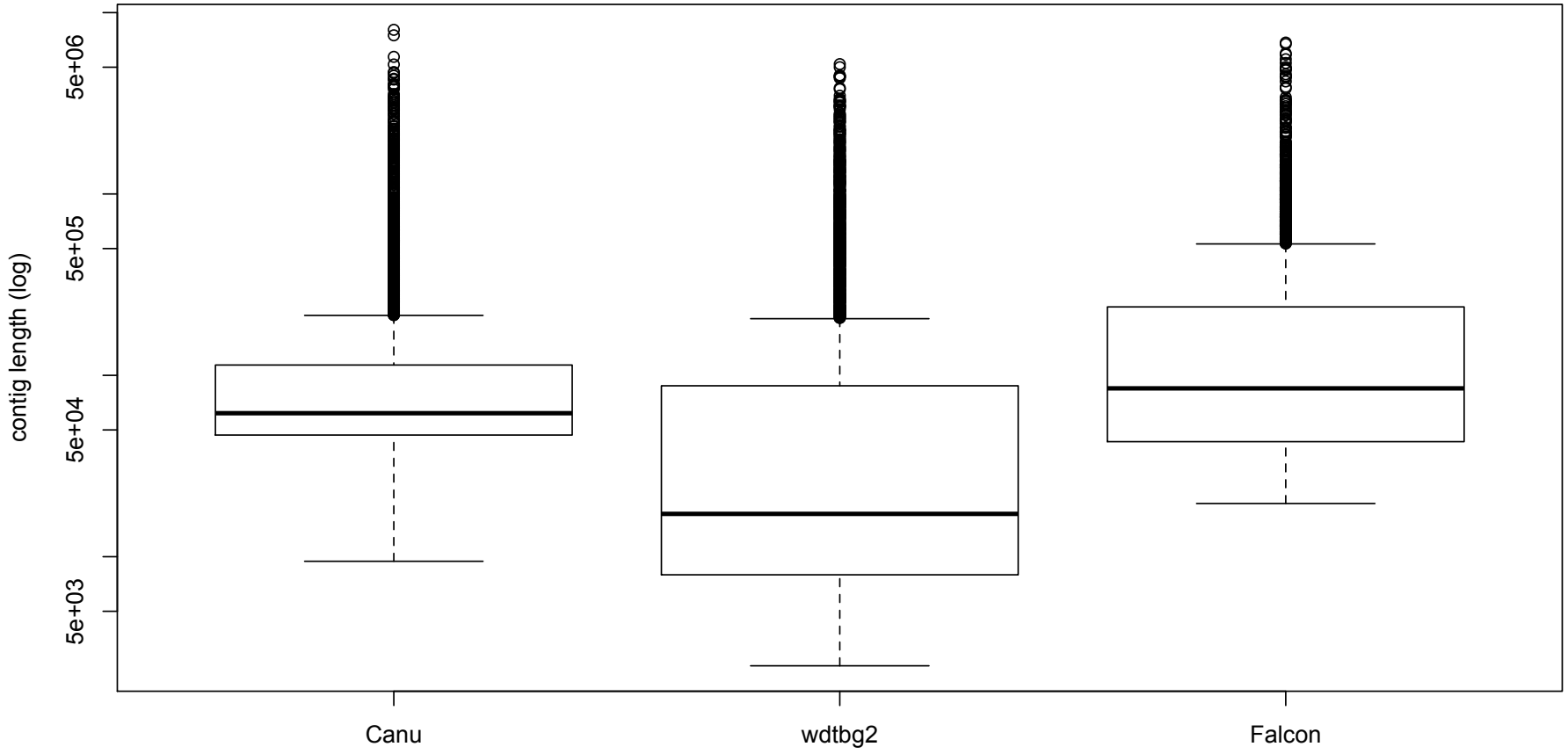
-
-

Raw genomic reads



De novo assembly

- Falcon
  - N50: 992,732
  - Size: 599 Mbp
- Wdtbg2
  - N50: 960,356
  - Size: 562 Mbp
- Canu
  - N50: 870,313
  - Size: 689 Mbp

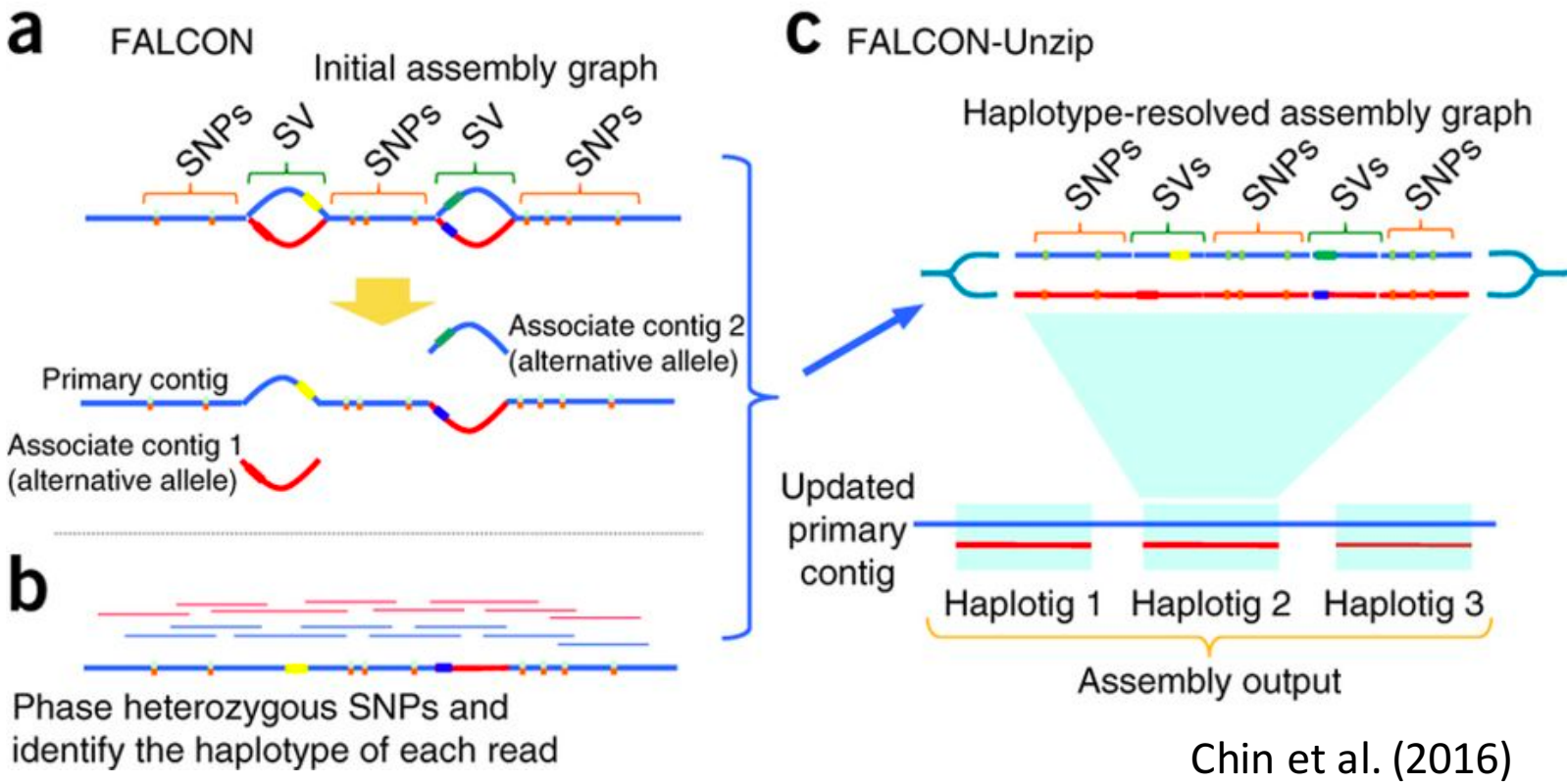


Falcon Unzip: phased contigs

Falcon Unzip

- Primary contigs:
  - N50: 992,732
  - Size: 599 Mbp
- Haplotigs:
  - N50: 98,429
  - Size: 414 Mbp

Phasing: 69.12%

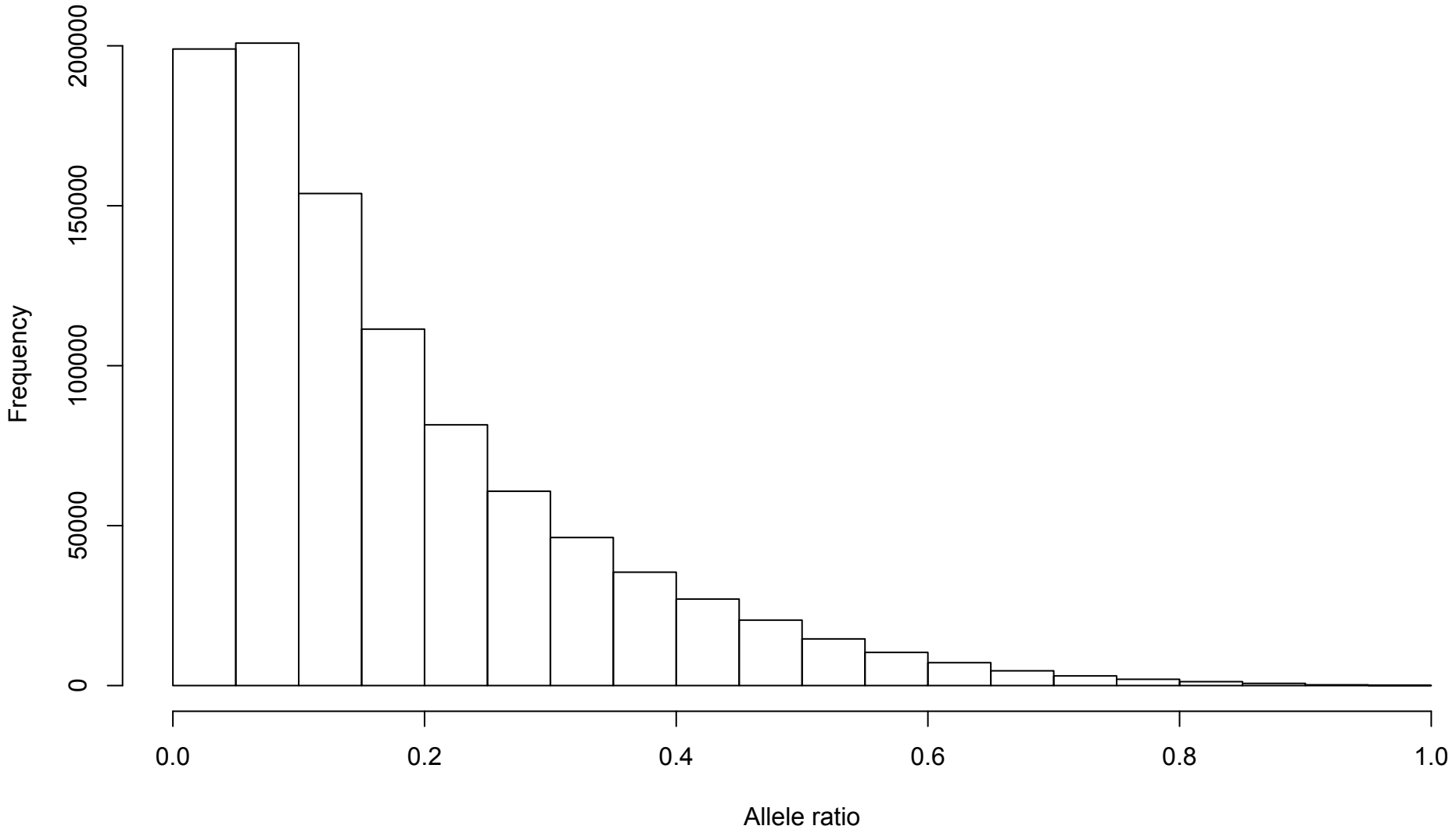


Chin et al. (2016)

What do we need?

- Genome
  - No reference genome yet for any loaches
- **Population information**
  - **Different mutations between cave and surface?**
  - **Allele frequencies?**
- Annotation
  - Which genes are impacted?
  - Common orthologues between cave fishes?

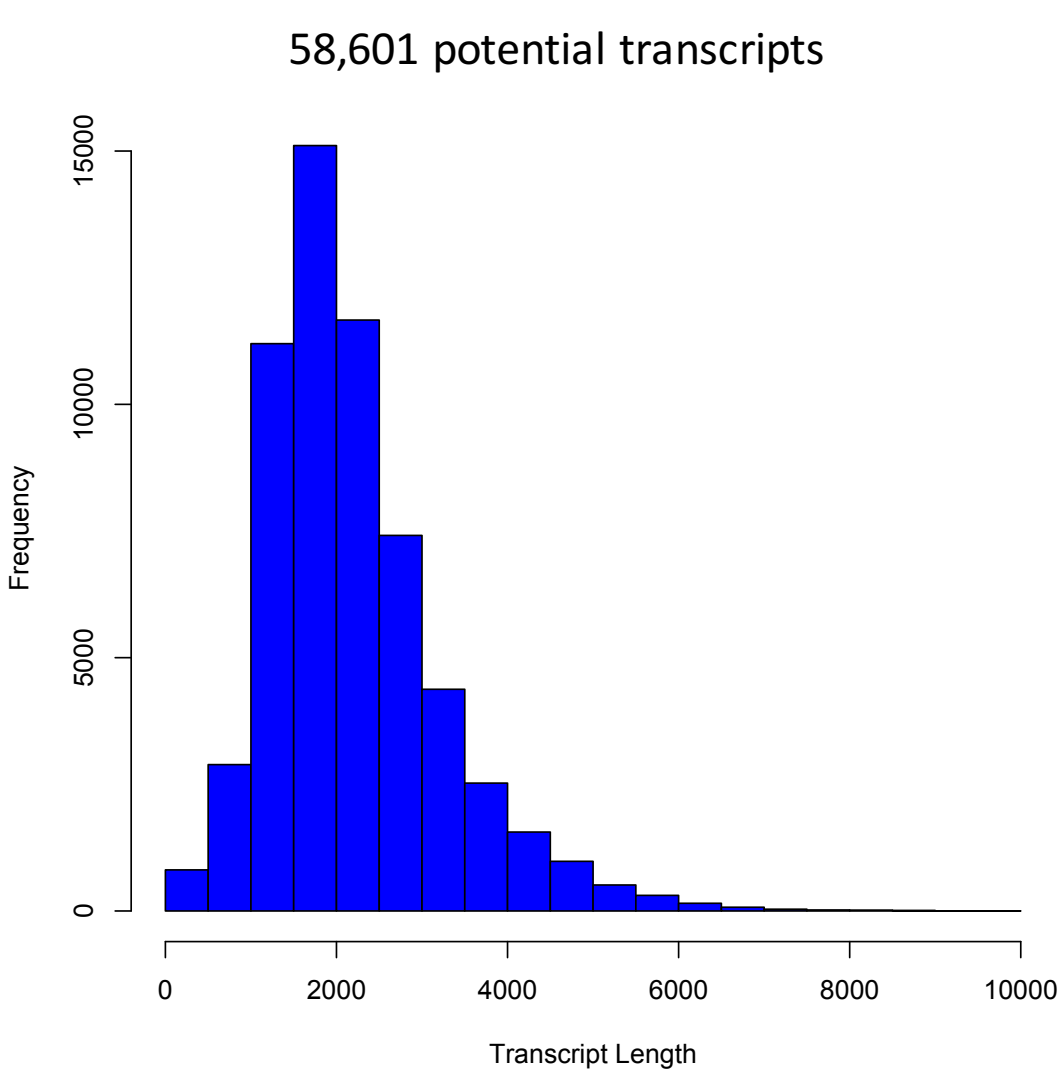
Pool sequencing of population

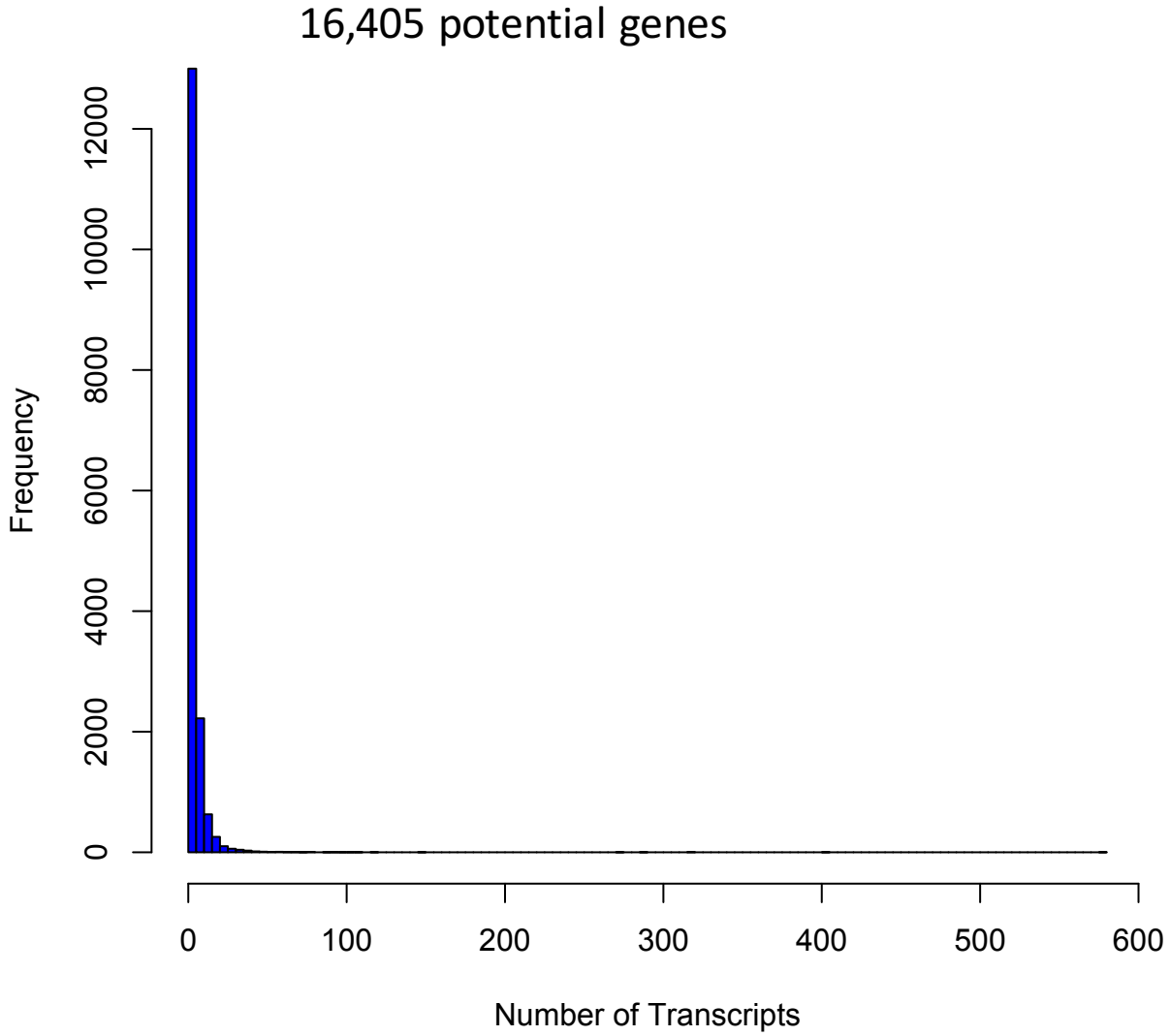


What do we need?

- Genome
  - No reference genome yet for any loaches
- Population information
  - Different mutations between cave and surface?
  - Allele frequencies?
- **Annotation**
  - **Which genes are impacted?**
  - **Common orthologues between cave fishes?**

IsoSeq Analysis



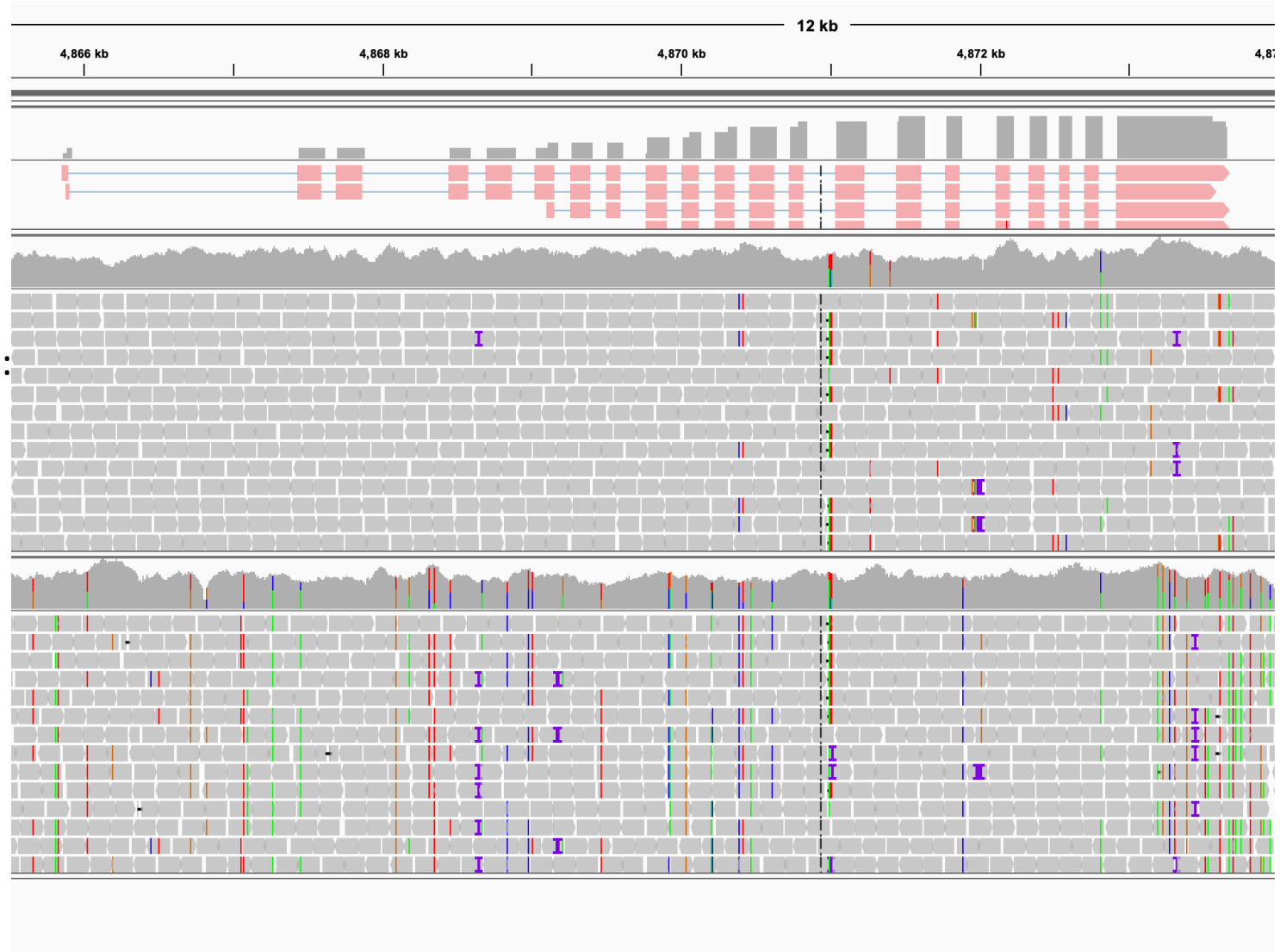


Comparison: So what is going on?

IsoSeq cave:

Pool seq cave:

Pool seq surface:



Ou

- Hi-

Summary

- First Loach genome
  - 70% phased contigs
- Isoform annotation
  - Close to estimated number of genes.
- …

Medhat Mahmoud
Richard Gibbs

Sarah Kingan

Arne Nolte

Jasminca Behrmann-Godel