

**4 April, 2016:**

**Genomic alignment, ancestral alleles, lift-over of coordinates**

# What have we all the time studied?



nine-spined stickleback

*Pungitius pungitius*

- genome unpublished
- genome unstudied



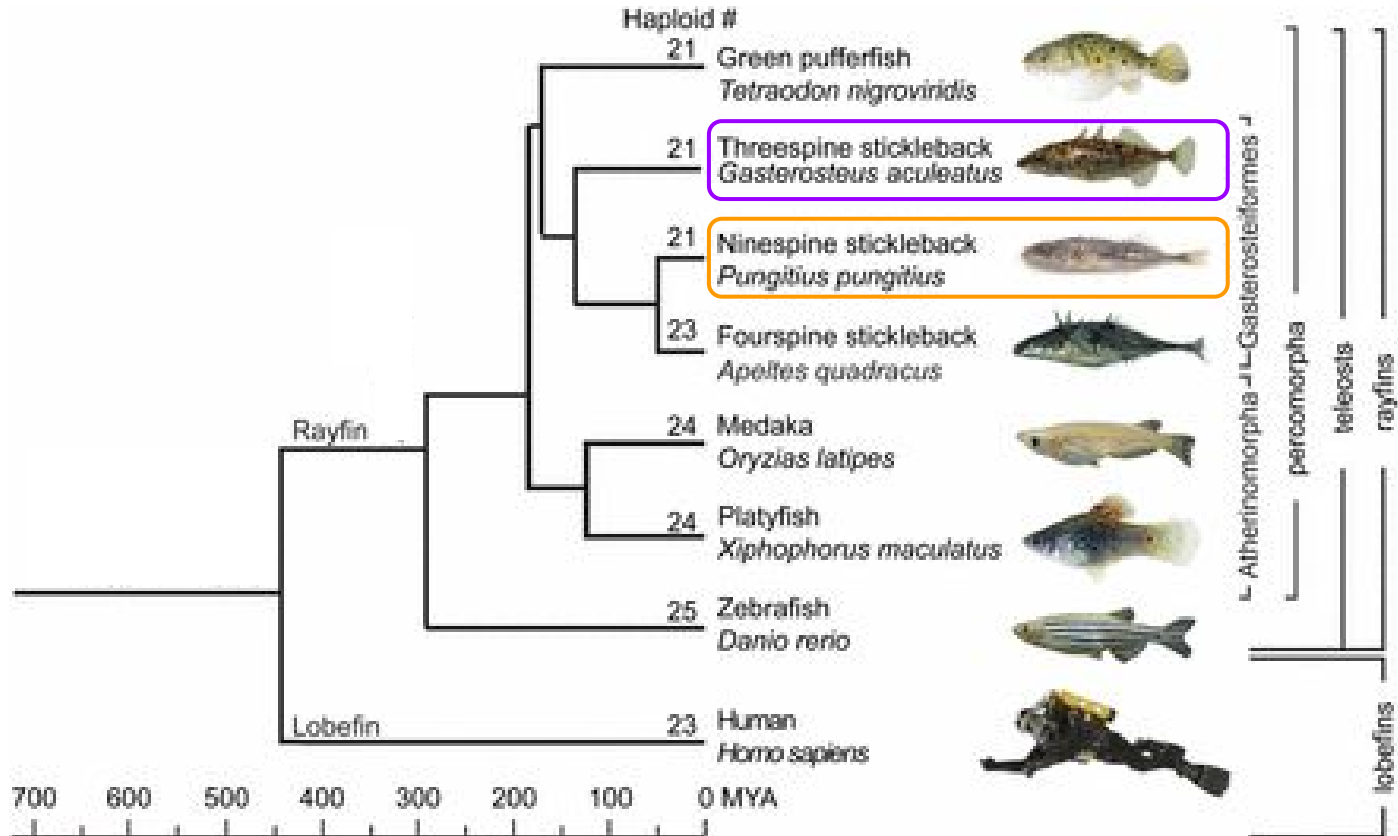
three-spined stickleback

*Gasterosteus aculeatus*

- model species
- genome in Ensembl



# What have we all the time studied?

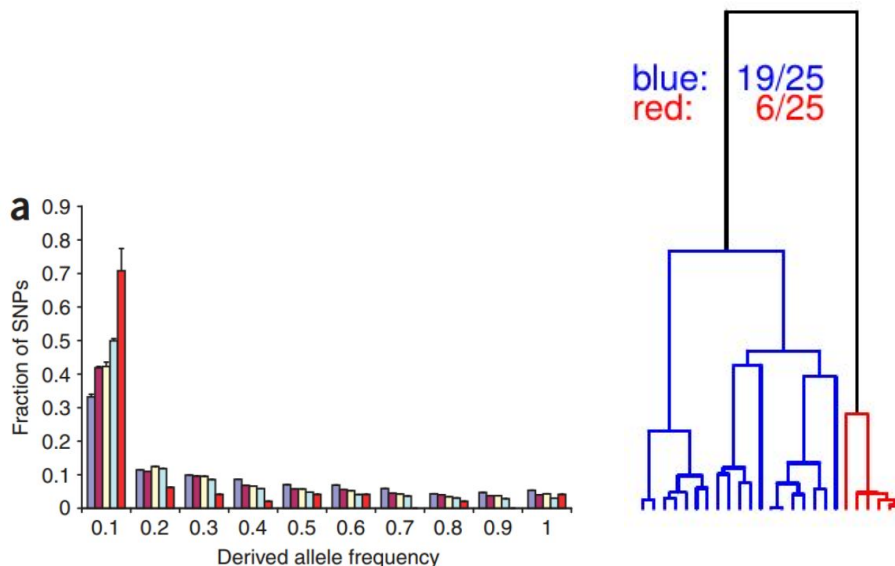


# How to use three-spined to study nine-spined?

## Inference of ancestral alleles

- we only see alternative alleles, not their evolutionary history
- minor allele is **not** always the derived one

- our data: A G  
0/0 x 1  
0/1 x 6  
1/1 x 3
- which allele is ancestral?  
8 x 0  
12 x 1



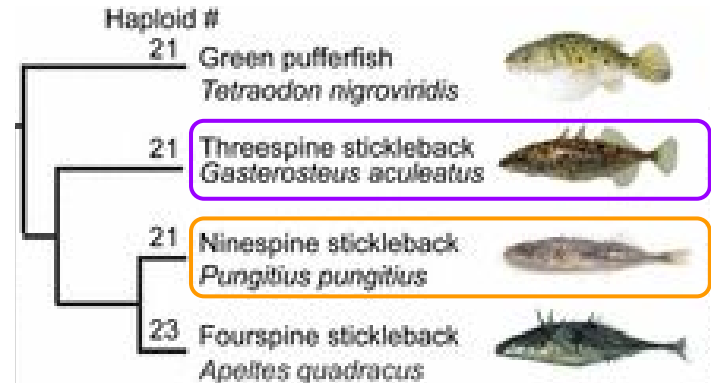
# How to use three-spined to study nine-spined?

Inference of ancestral alleles

nine-spined                      three-spined



8 x A  
12 x G



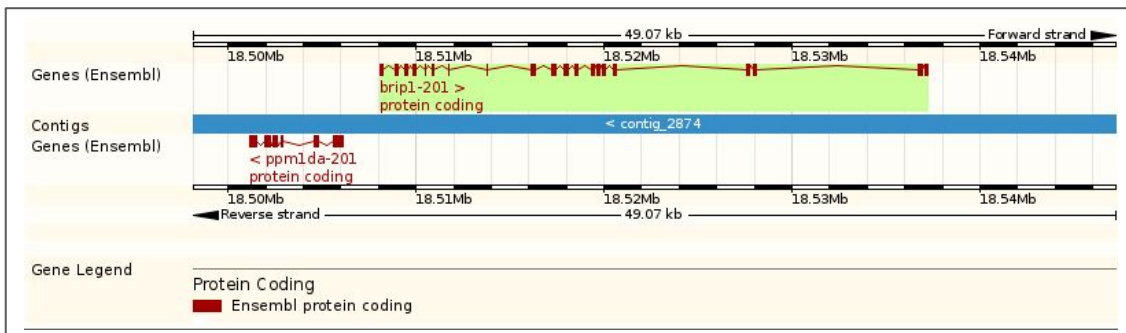
Assumption: shared, fixed allele in an outgroup species is ancestral

- in this example, A is ancestral allele; G is derived allele

# How to use three-spined to study nine-spined?

## Annotation lift-over

- Ensembl provides data for three-spined



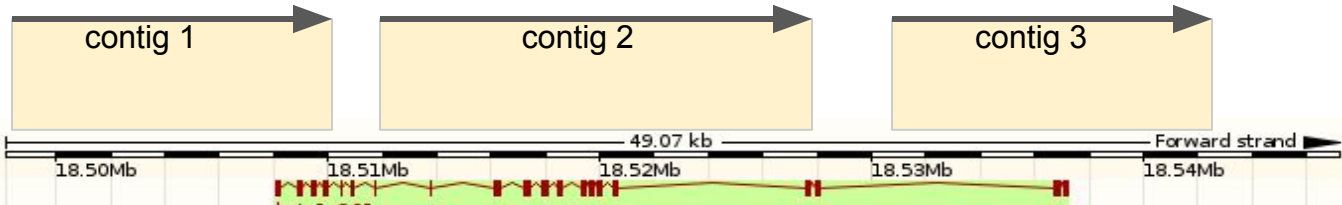
- data also available in computer-readable format

```
##gff-version 3
##sequence-region groupI 18508123 18537196
groupI ensembl gene 18508123 18537196 . + . ID=gene: ENSGACG00000013558;Name=brip1;assembly_
groupI ensembl exon 18508123 18508215 . + . Name=ENSGACE000000158395;Parent=transcript: ENSGA
groupI ensembl exon 18508895 18509006 . + . Name=ENSGACE000000158397;Parent=transcript: ENSGA
groupI ensembl exon 18509372 18509554 . + . Name=ENSGACE000000158399;Parent=transcript: ENSGA
groupI ensembl exon 18509844 18509986 . + . Name=ENSGACE000000158400;Parent=transcript: ENSGA
groupI ensembl exon 18510465 18510481 . + . Name=ENSGACE000000158404;Parent=transcript: ENSGA
groupI ensembl exon 18510873 18510958 . + . Name=ENSGACE000000158406;Parent=transcript: ENSGA
groupI ensembl exon 18511741 18511755 . + . Name=ENSGACE000000158408;Parent=transcript: ENSGA
groupI ensembl exon 18513764 18513774 . + . Name=ENSGACE000000158413;Parent=transcript: ENSGA
groupI ensembl exon 18516095 18516367 . + . Name=ENSGACE000000158416;Parent=transcript: ENSGA
groupI ensembl exon 18517215 18517436 . + . Name=ENSGACE000000158419;Parent=transcript: ENSGA
groupI ensembl exon 18517911 18518110 . + . Name=ENSGACE000000158421;Parent=transcript: ENSGA
groupI ensembl exon 18518467 18518599 . + . Name=ENSGACE000000158424;Parent=transcript: ENSGA
groupI ensembl exon 18519333 18519487 . + . Name=ENSGACE000000158427;Parent=transcript: ENSGA
groupI ensembl exon 18519609 18519786 . + . Name=ENSGACE000000158429;Parent=transcript: ENSGA
groupI ensembl exon 18519902 18520042 . + . Name=ENSGACE000000158432;Parent=transcript: ENSGA
```

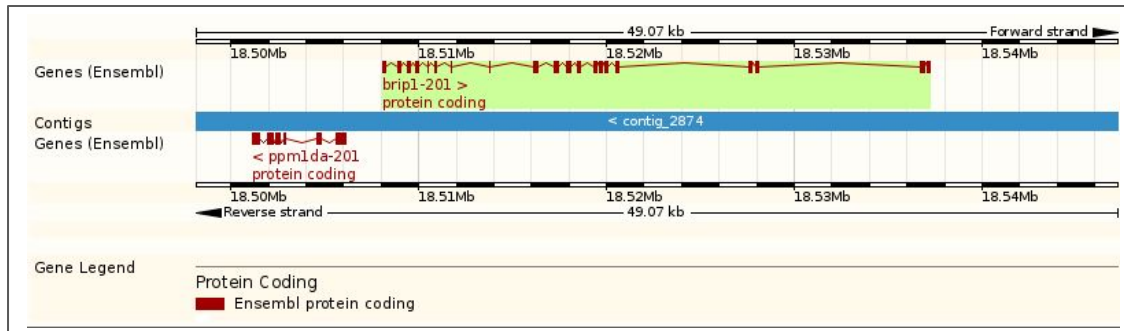
# How to use three-spined to study nine-spined?

## Annotation lift-over

nine-spined

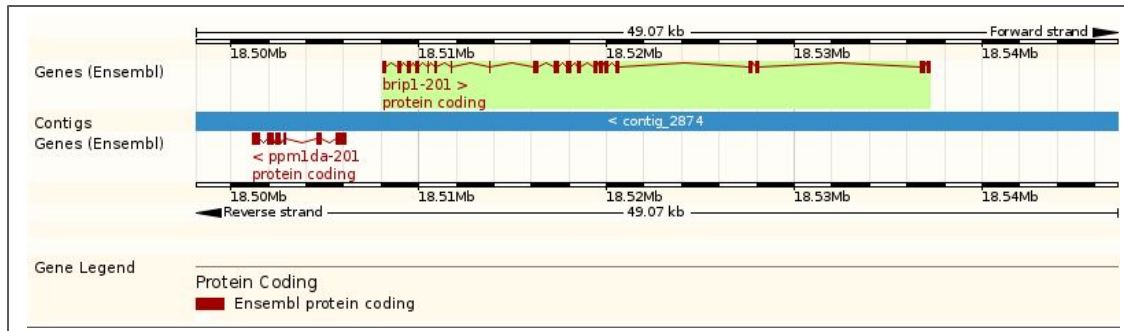


three-spined



# How to use three-spined to study nine-spined?

## Annotation lift-over





# How to use three-spined to study nine-spined?

**Inference of ancestral alleles** and **Annotation lift-over** require sequence alignment between the two genomes

- use inferred homology to transfer information between the two genomes