

Dragonfly genome project

Nadezda Pilshikova

Sofia Kolchanova

Nikolay Kochkin

Unscientific supervisor: Vladimir Suvorov

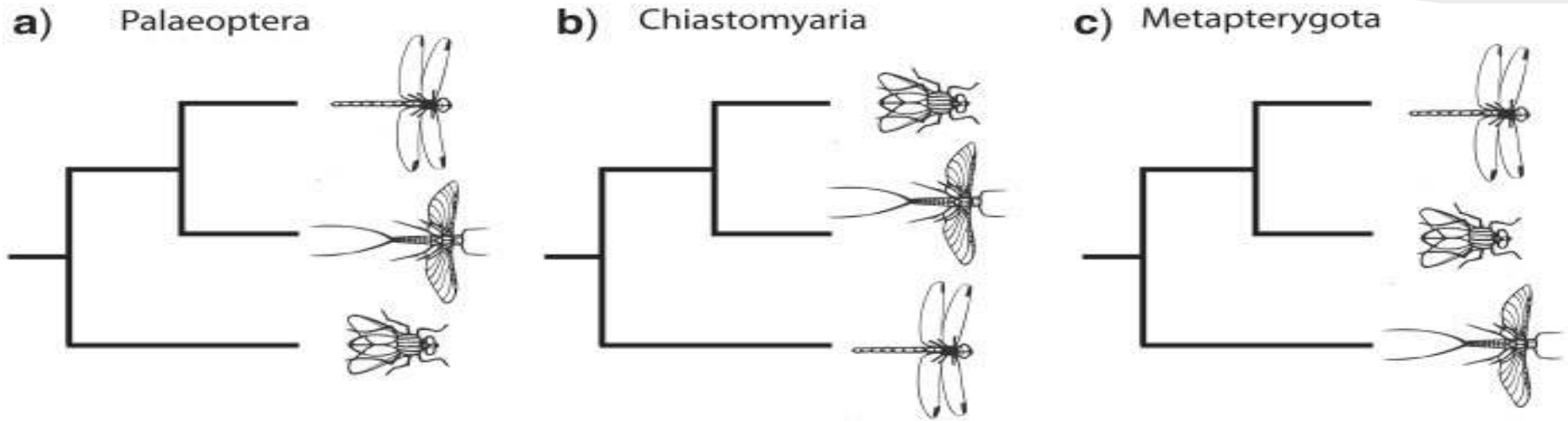
Scientific Support: Dobzhansky Lab



The Team

- Biotechnology track – Nadezda
 - Plan: Do some experiments using Dobzhansky Lab equipment
 - Results: All equipment used
- Biology track – Sofia
 - Plan: Find the explanation why should we study dragonflies
 - Result: “Because they are cool and they can fly”
- Computer Processing track – Nikolay
 - Plan: Try to get some genome pipeline soft work
 - Result: Consumed a lot of AU cluster, forced into “nice -n 19”

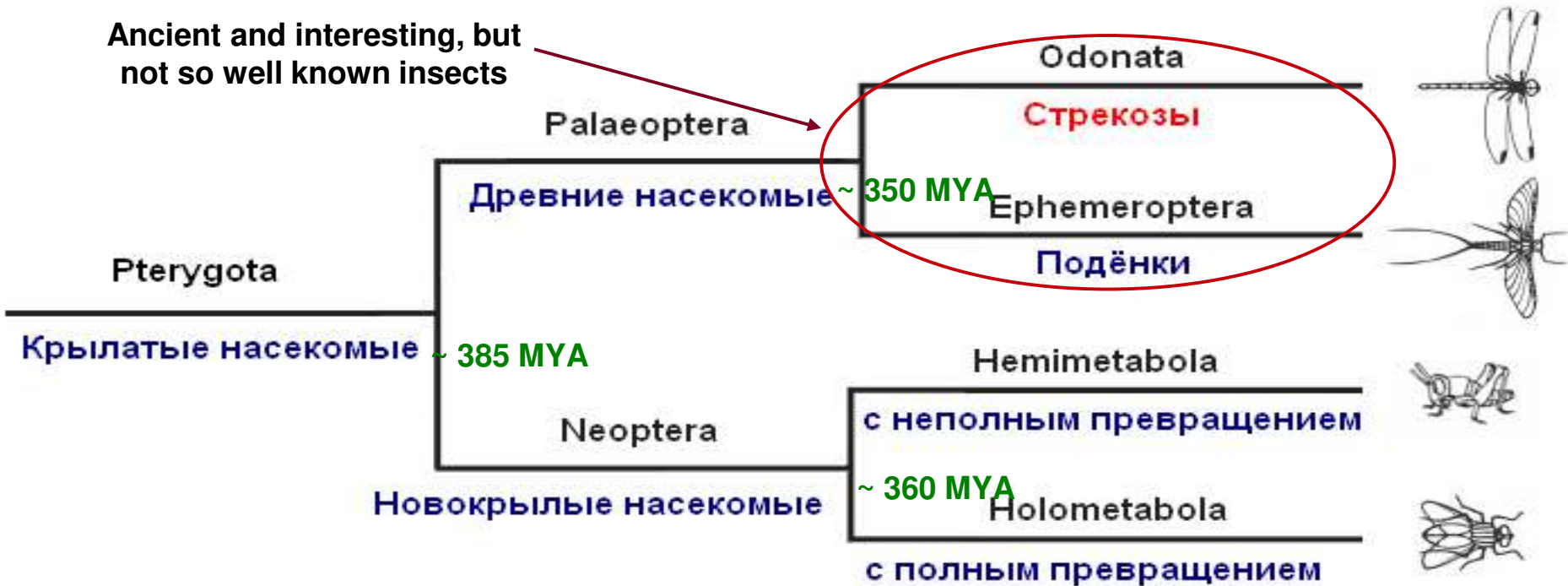
Phylogeny: the 3 hypothesized divisions within the Pterygota



Different researchers have argued for all 3 possible branching patterns between the Odonata (dragonflies), Ephemeroptera (mayflies), and Neoptera

Phylogeny: dragonfly among another insects

Ancient and interesting, but not so well known insects



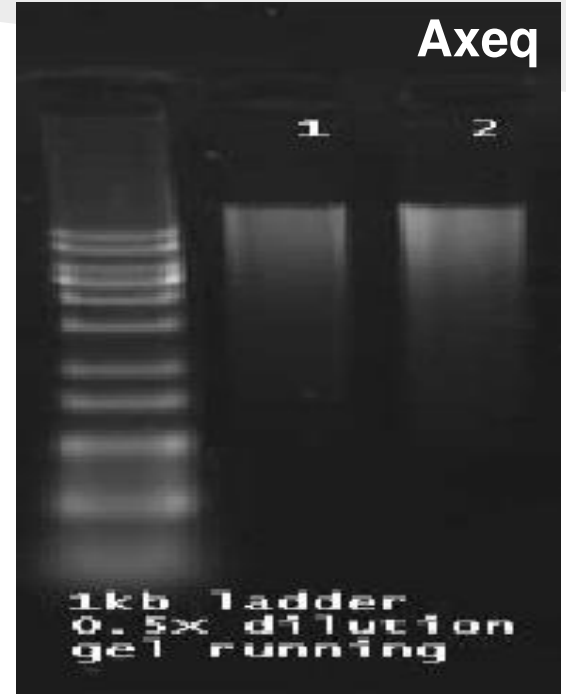
Biotechnology

DNA extraction results (Electrophoresis + Nanodrop + Qubit)

**Total amount
(A260/A280)**

**In the Lab
9.45 ug (1.92)**

**At Axeq (after 4 days)
4.61 ug (1.91)**



So, now we are waiting for summer

Genome size measurement

Why is it necessary?



Arabidopsis thaliana

– the first plant to have its entire genome sequenced (in 2000)

Genome size:

– based on DNA sequencing data: **125 Mb**

– based on flow cytometry data: **157 Mb**

+25% => extra DNA in gaps is mainly composed of highly repeated (possibly junk) sequences but may also include some informational genes

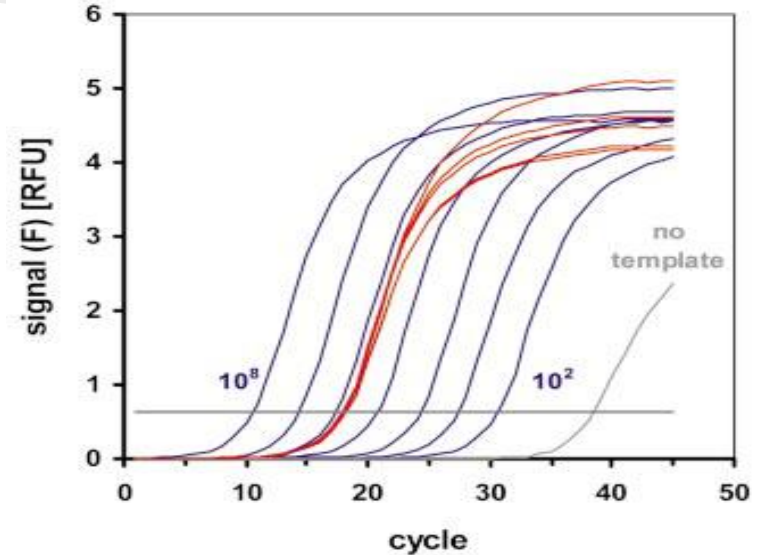
Genome size measurement

C-value, genome size in pikogramms

Amount of genome DNA in sample

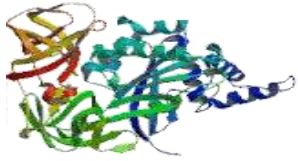
$$C = \frac{m}{N}$$

Number of copies of single copy gene in sample (measuring by real-time PCR)



Real-time PCR: absolute quantification of a single copy gene in a genomic DNA sample

Genome size measurement



Eukaryotic Translation Elongation Factor 1 Alpha 1 (EF1a) – single copy gene in human, mouse, fruit fly, but have two copies in some insects.

Measuring genome size of **Common Darter** (*Sympetrum striolatum*) by real-time PCR:

C = 0.74 pg

Γ = 0.67 Gb

=> There is one copy of EF1a in dragonfly genome

Our results are close to data in articles

<http://www.genomesize.com>

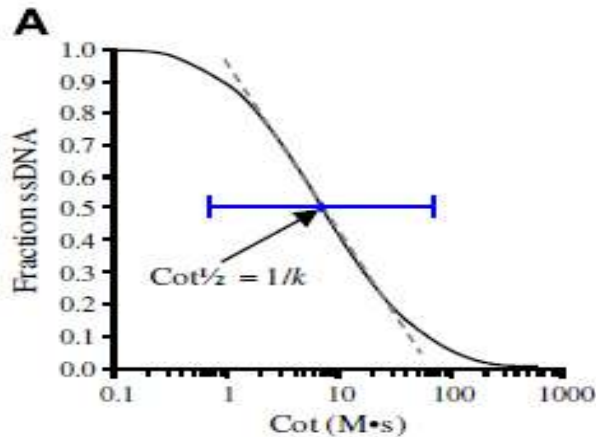
Genome size of genus <i>Sympetrum</i> dragonflies		
Species name	C-value, pg	Γ, Gb
<i>S. internum</i>	0.78	0.71
<i>S. meridionale</i>	0.65	0.59
<i>S. obtrusum</i>	0.82	0.75
<i>S. vicinum</i>	0.77	0.70

Cot analysis (DNA reassociation kinetics)

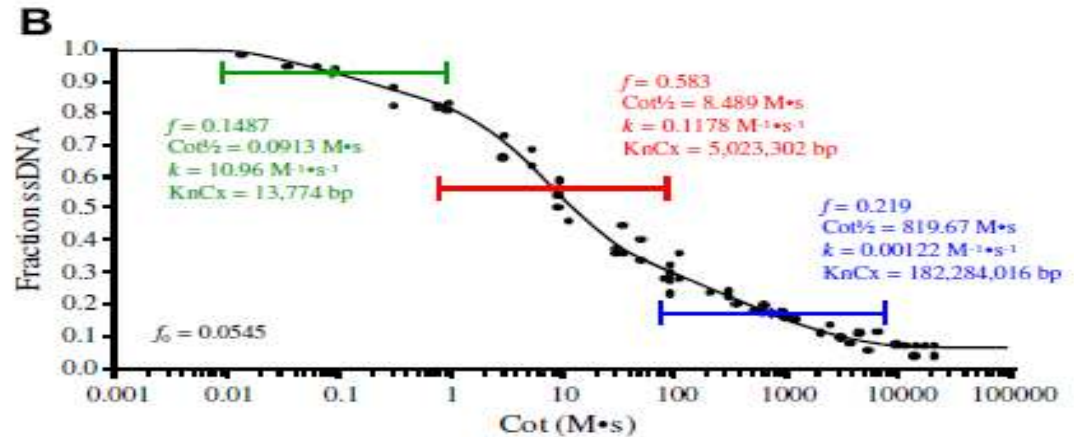
Genome structure examination in individual species:

- proportion of repetitive sequences
- genome size measurement

Reassociation kinetics curves:



ideal curve



real curve for eukaryotes

Biology in details

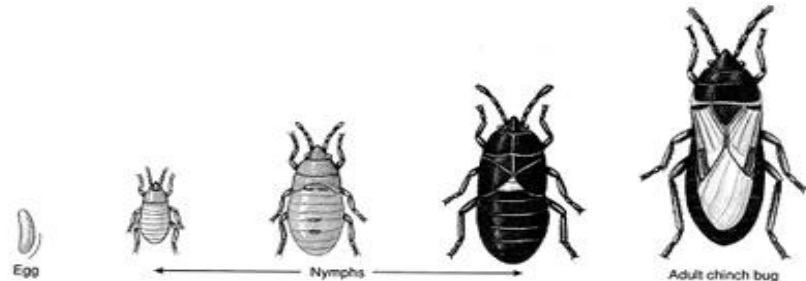
Major insect development genes

It is worth regarding them because:

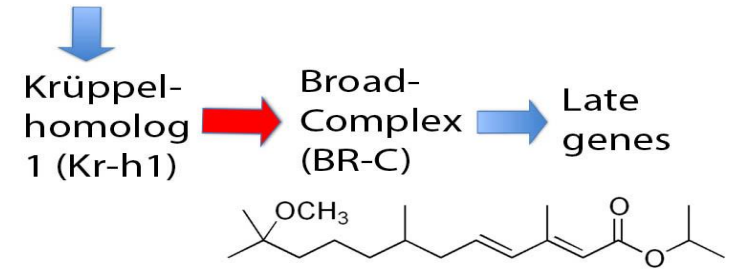
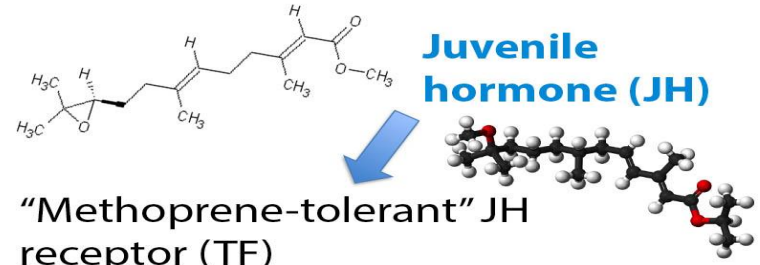
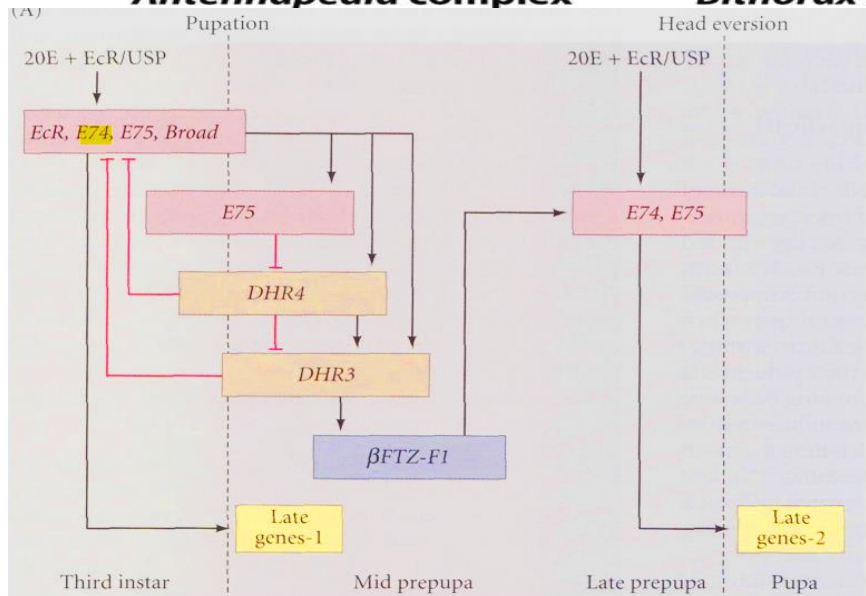
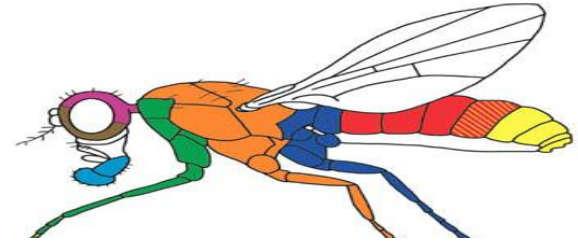
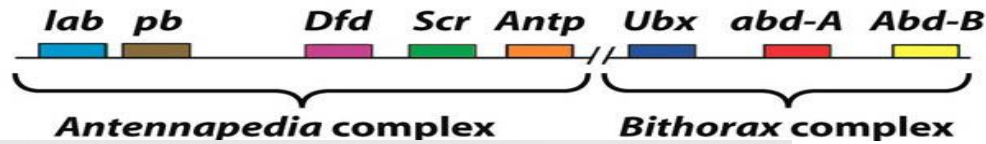
- Dragonflies constitute an incredibly old group
- These genes are important in studies of developmental regulatory mechanisms

We chose:

- Homeobox-containing genes (HomC)
- Ecdysone-responsive genes
- Juvenile hormone-responsive genes



Early development & metamorphosis regulatory genes

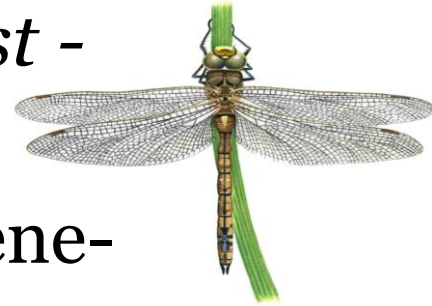


Early development & metamorphosis regulatory genes

- EcR
- USP
- Br-C
- Kr-h1
- E74
- E75
- DHR₃, DHR₄
- beta-FTZ-F1
- Met
- Lab
- pb
- Dfd
- Scr
- Antp
- Ubx
- AbdA
- AbdB

BLASTing Ladona fulva genome

- *Drosophila melanogaster* development genes cds used (FlyBase, NCBI)
- BLAST search (*blastn -task dc-megablast -num_threads 40*)
- All of them were found, except Methoprene-tolerant (Met)....



Methoprene-tolerant (JH receptor)

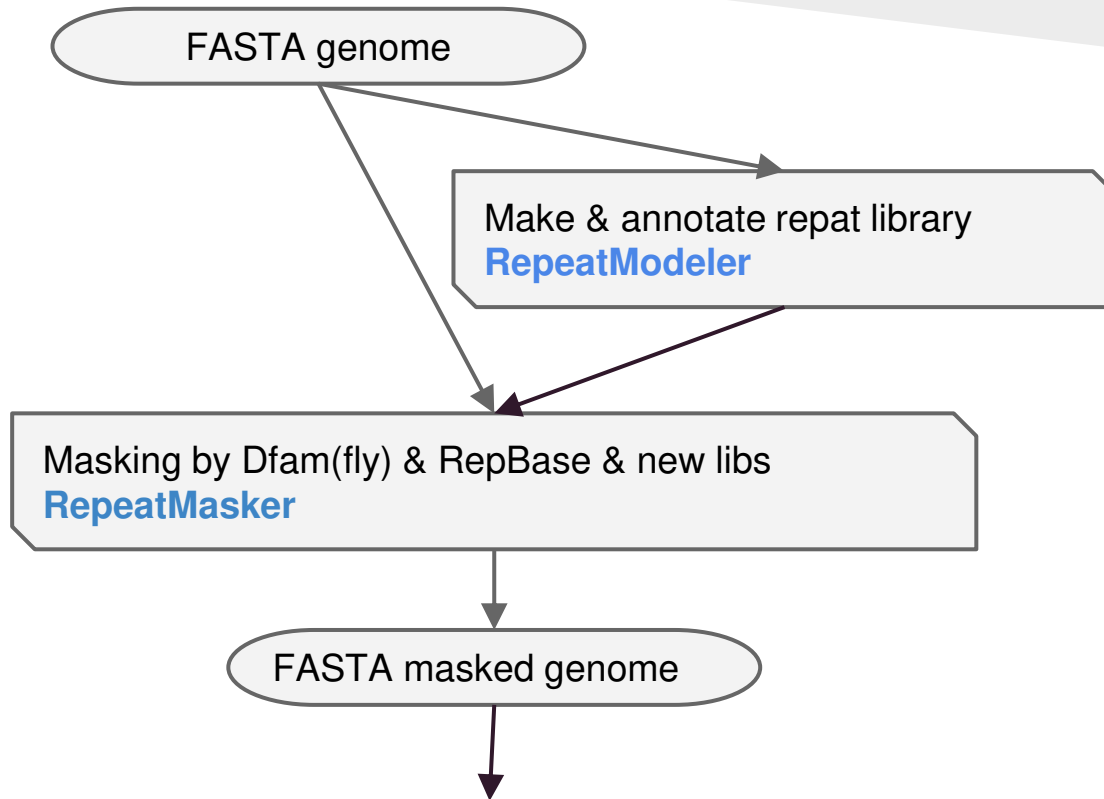
- Amino acid sequences of *D. melanogaster* bHLH - PAS and the **tblastn** search tool were employed to find the presumed location of Met-receptor gene in the dragonfly genome
- **Samtools** were used to retrieve the gene region (~7000 bp)
- **Python** script was employed to find ORFs and stops
- Some of the **ORFs** found proved to be the Met gene conservative domains, related to those of hemimetabolous insects

...Conclusions and plans

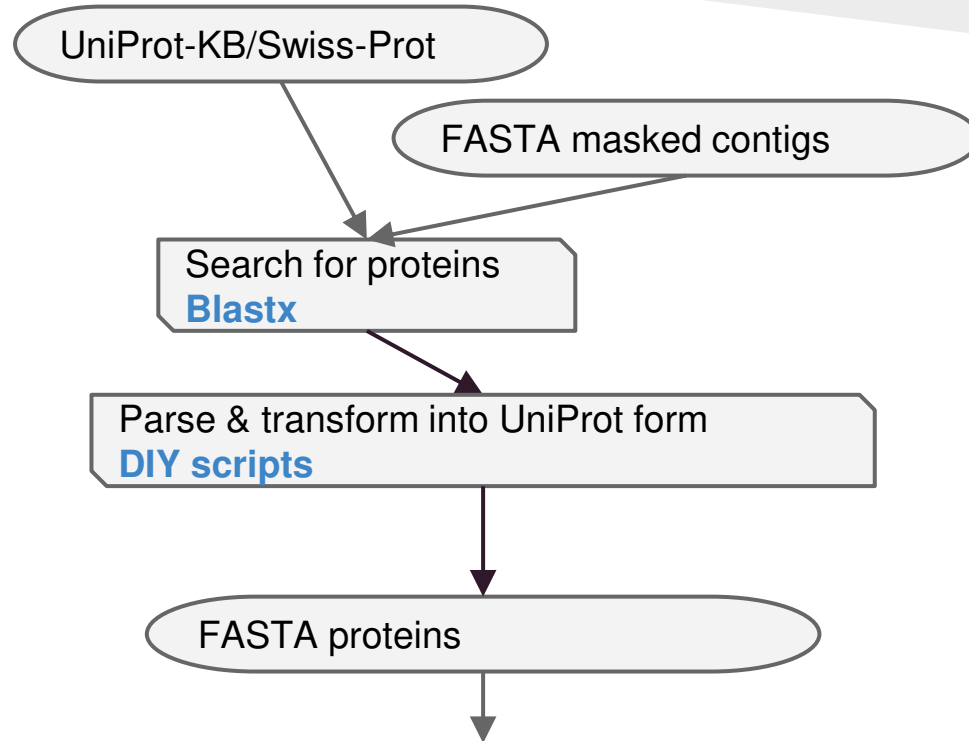
- The gene, which wasn't found with regular BLAST procedure, was finally located and compared to its homologs
- It would be interesting and useful to conduct further experiments involving an artificial Met gene construction and transformation of insects with the use of it. This could help to clarify its functions, which are still quite poorly understood.

Now some CS

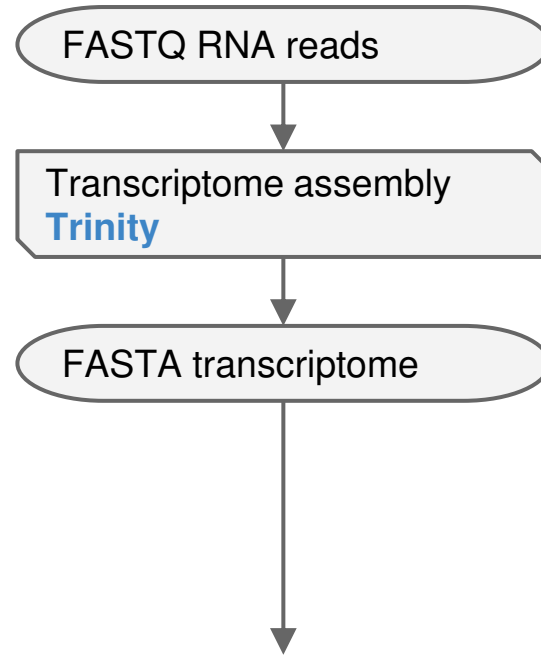
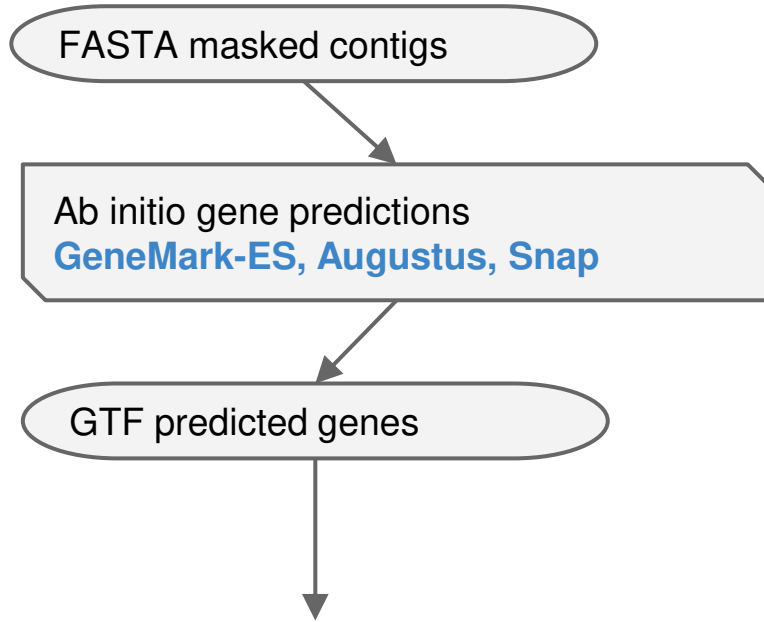
1. Repeat masking



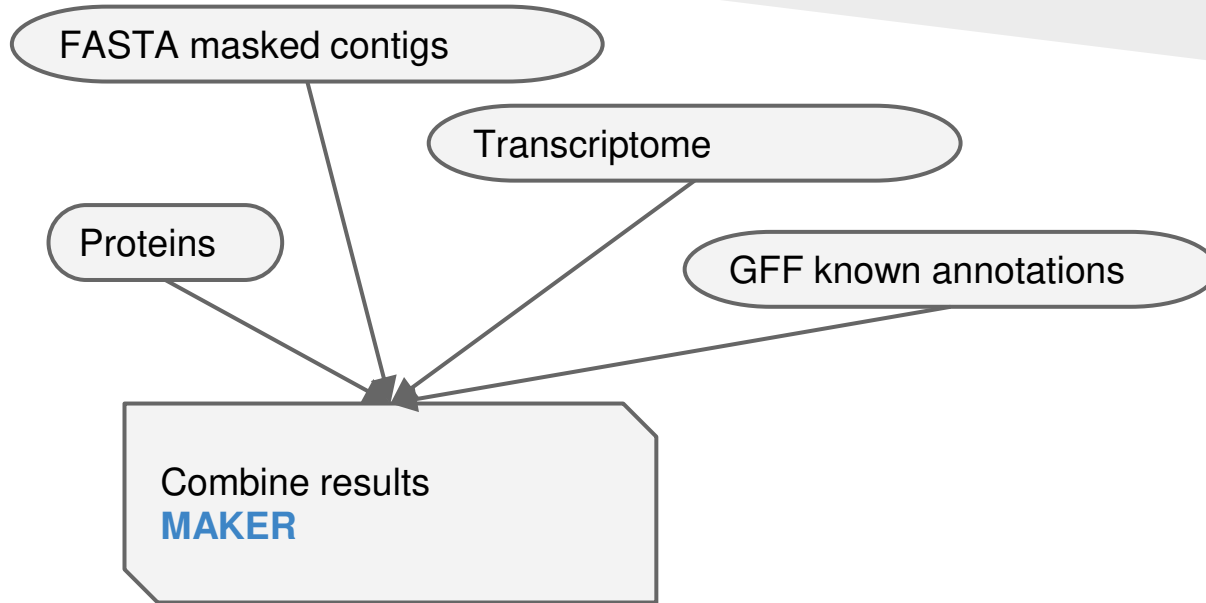
2. Search for proteins



3. Ab initio & RNA assembly



Final



... Finally we have **Big Data**

CEGMA validation

Dragonfly
contigs

Mapping Approach:
248 ultra-conserved Core
Eukaryotic Genes

Dragonfly
scaffolds

Results

Complete: 135 CEGs
(54.44%)
Partial: 208 CEGs
(83.87%)

Results

Complete: 160 CEGs
(64.52%)
Partial: 229 CEGs
(92.34%)

Results

We had

- 10 Dragonflies
- No experience
- No scientific plan

Now we have

- No Dragonflies
- Some experience
- A scientific plan