

# Bioinformatics study of convertases

Molly Novy<sup>1,2</sup> and Sandra Rodriguez-Zas<sup>1</sup>

<sup>1</sup>Department of Animal Sciences, University of Illinois Urbana-Champaign; <sup>2</sup>ACES James Scholar Honors Program

## Introduction

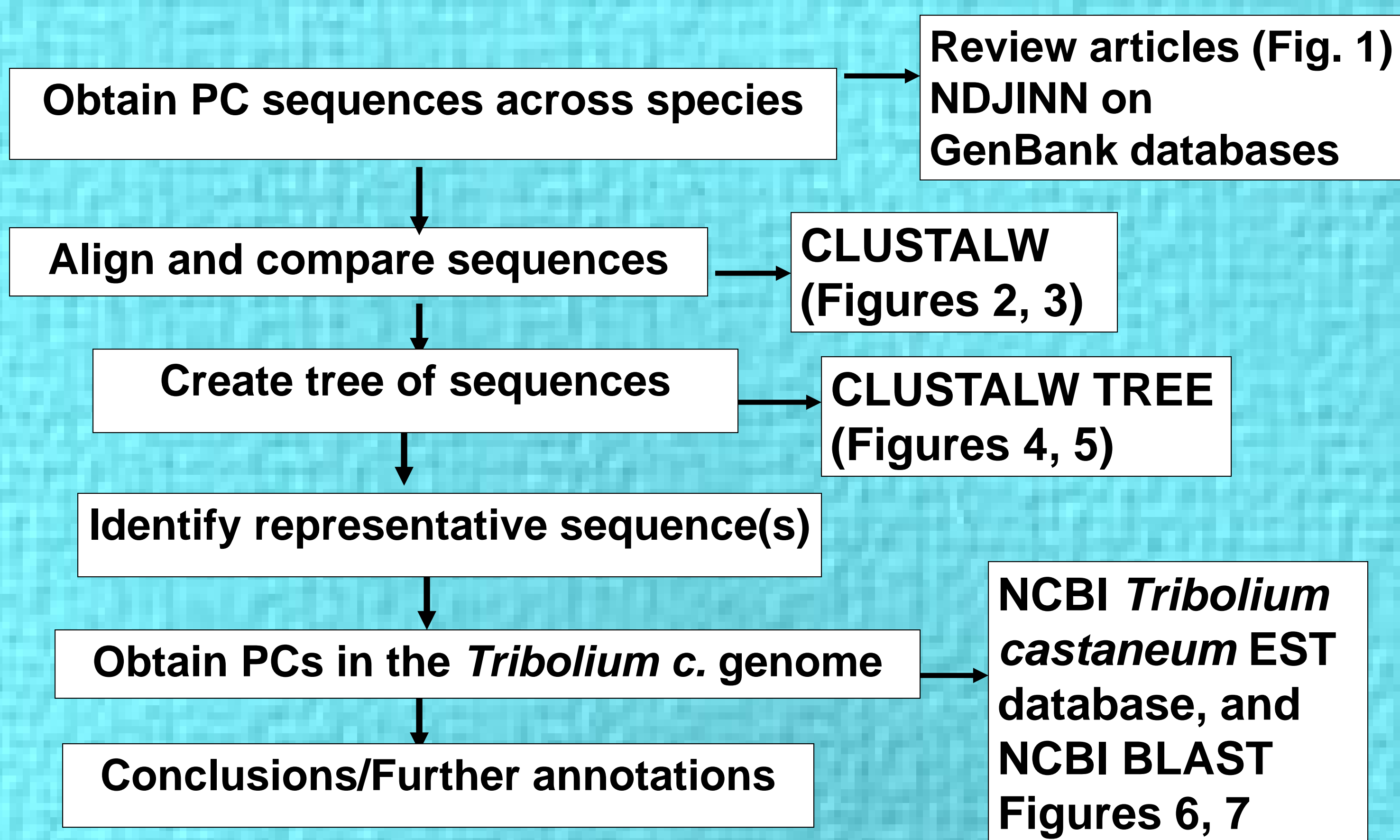
- Prohormone convertases (PCs) are enzymes that cleave the prohormone precursors to produce neuropeptides.
- Neuropeptides play an important role in learning, behavior, memory, and perception of stimuli and directly affect neural transmission and nerve activity. Therefore, PCs are critical for the health and well being of organisms.
- Seven subfamilies of PCs have been identified: PC1/PC3, PC2, PC4, PACE4, PC5, Furin, and PC7. The subfamilies may be active in different conditions (e.g. pH level) and organs (e.g. PC7 can be found in the liver, brain and gastrointestinal tract).
- The study of neuropeptides is challenging because of the complexity of the cleavage process and neuropeptide activity.
- Our objective was to study the PC subfamilies and their relationship using genomic and proteomic sequences and bioinformatics tools.

## Hypotheses

- The degree of conservation of the PC sequences varies with organism, phyla and subfamily.
- The detection of the specific amino acid positions or sequence regions that differ among species can help to explain major and minor differences in PC and neuropeptide activity among species.
- The conservation of PC sequences among species can help discover and annotate this family of proteins in in-progress or unavailable genomes sequencing projects.

## Objectives and activities

- The diagram below describes the activities and bioinformatic resources (databases and programs) used at each stage of the study. We used the Biology Workbench ([workbench.sdsc.edu](http://workbench.sdsc.edu)) for all sequence searches and alignments with the exception of the red flour beetle *Tribolium castaneum*. The discovery of PCs in this non-sequenced beetle genome was based on the (NCBI ([www.ncbi.nlm.nih.org](http://www.ncbi.nlm.nih.org)) Expressed Sequence Tagged (EST) database and BLAST.



## Results

Figure 1. List of species by PC subfamily

	Human	Mouse	Rat	Cow	Pig	Fish	Frog	Chicken	Dog	Chimp	Fruit Fly	Bee
PC1/3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes
PC2	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
PC4	Yes	Yes	Yes	Yes	No	No	Yes	No	Yes	No	No	No
PACE4	Yes	Partial	Yes	No	No	No	Yes	No	Yes	No	No	No
PC5	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	No
PC7	Yes	Yes	Yes	No	No	No	No	No	No	Yes	No	No
Furin	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Partial	Yes	Yes

Figure 2. Partial alignment of Furin sequences

```

Honeybee IYSASWGPDDDGKTVDPGPELATRAF IEGITKGRNGRGSIFVWASGNGGR
Fruit Fly IYSASWGPDDDGKTVDPGPELASRAF IEGTTKGRGGKGSIFVWASGNGGR
Domestic Silkworm IYSASWGPDDDGKTVDPGPELATRAF IEGVTKGRNGKGSIFVWASGNGGK
Dog IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Cow IYSASWGPEDDGKTVDPGPAHLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Rat IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Mouse IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Chinese Hamster IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Human IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Chimpanzee IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
Chicken IYSASWGPEDDGKTVDPGPARLAEEAF FRGVSQGRGGLGSIFVWASGNGGR
African Clawed Frog IYSASWGPEDDGKTVDPGPAKLAEEAF YRRVTQGRGGLGSIFVWASGNGGR
Snail IYSASWGPDDDGKTVDPGATLARKAF YDGIKGRGGLGSIFVWASGNGGR
Cabbage Looper IYSASWGPEDDGKTVDPGGLARRAF IYGVTSGRRGKGSIFVWASGNGGR
California Sea Hare IYSASWGPDDDGKTVDPGPKLAKEAF IKGIENGRGKGSIFVWASGNGGS
*****:***:****. ** .* * ** * **::*****
    
```

Figure 3. Partial alignment of PC1

```

Chimpanzee EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Human EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Dog EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Pig EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Mouse EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Rat EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANVLDL TURDMQHLT
Cow EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Chicken EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANLGLTURDMQHLT
Bullfrog EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANPGLTURDLQHLT
Marsh Frog EAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANPGLTURDLQHLT
Tunicate GAGVATTDLYGNCTLRHSGTSAAPAAAGVFLALALEANPSLTURDVQHLT
Fruit Fly NTGVATTDLYGKCTTTHSGTSAAPAAAGVFLALALEANPQLTURDIQHLT
Honeybee NIGVATTDLYGKCTTTHSGTSAAPAAAGVFLALALEANPQLTURDIQHLT
Nematode ETGVATTDLYGRCTRSHTSGTSAAPAAAGVFLALALEANPSLTURDLQHLT
California Sea Hare DAGVATTDLYNNCTASHSGTSAAPAAAGVFLALALEANKL TURDMQHLT
*****:***:****. ** .* * ** * **::*****
    
```

Figure 4. Furin tree

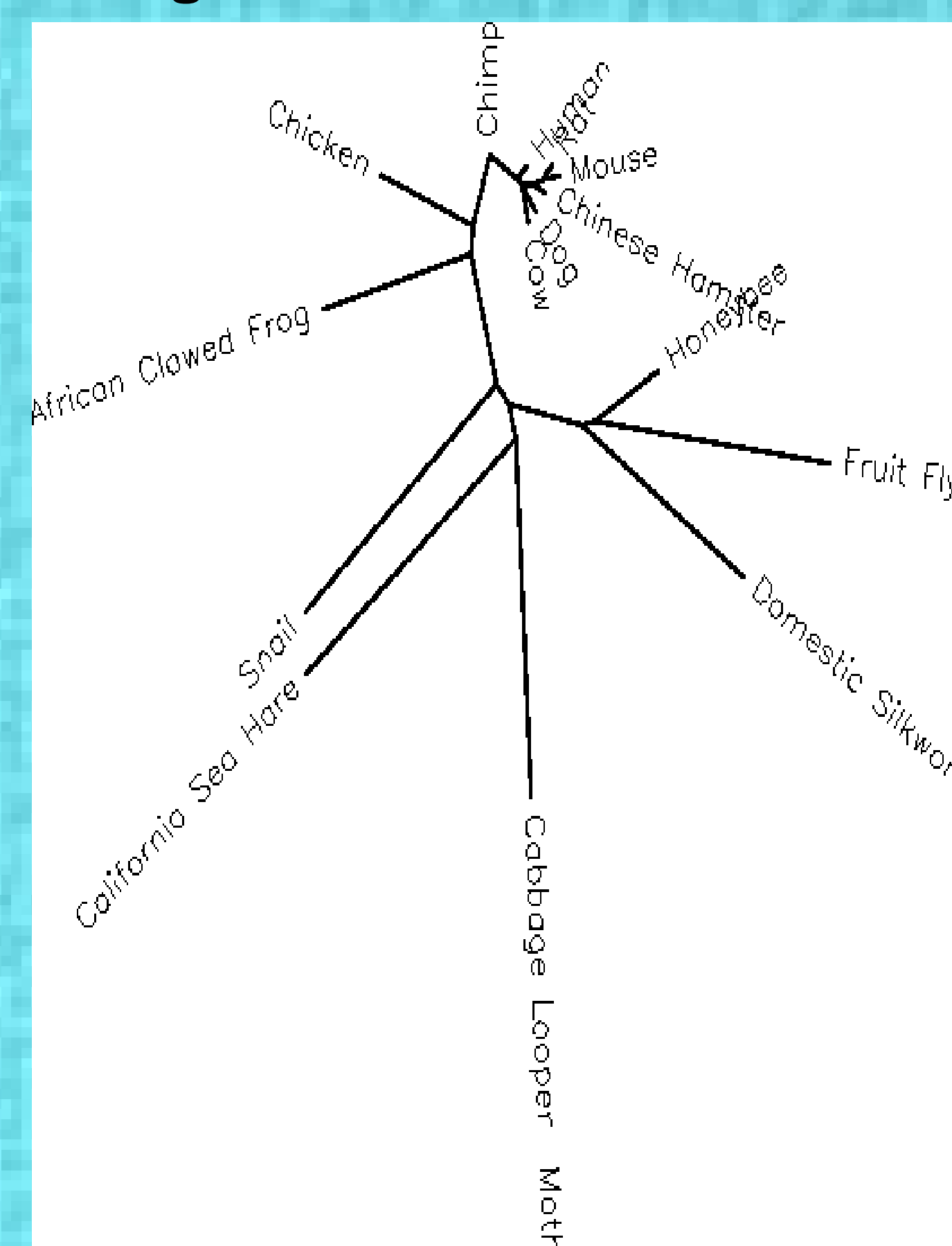


Figure 5. PC1 tree

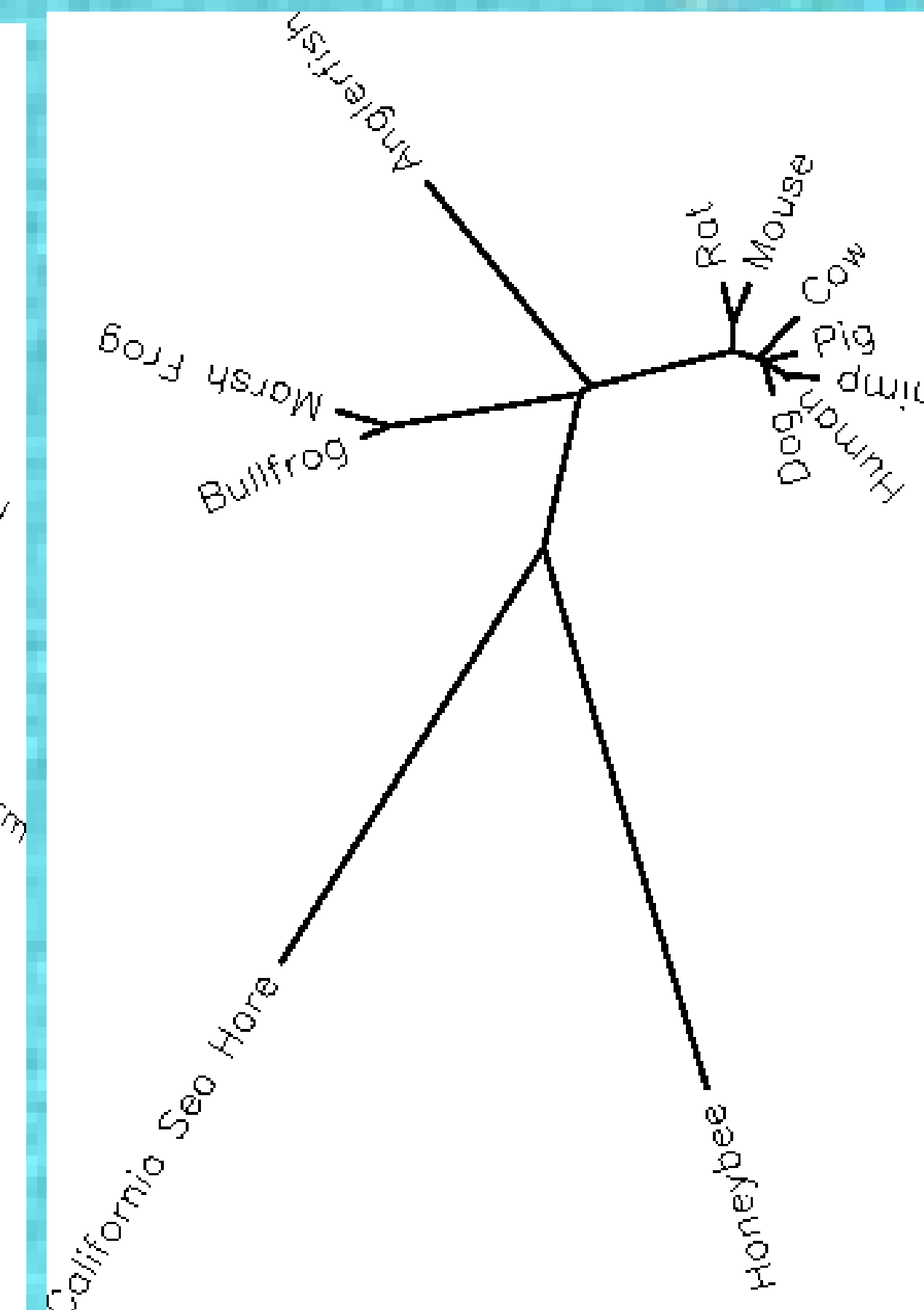


Figure 6. BLAST alignment of honey bee (query) flour beetle (sbjct) PC1/3 sequences

```

Score = 248 bits (634), Expect = 8e-156
Identities = 126/186 (67%), Positives = 152/186 (81%), Gaps = 0
Frame = +2
Query 231 DPKISYD VMEGDYD PFP RYELSGMNGHTRCGGE IAMEANNRRCG VGV
DP ISYD N+ D+DP PRY+ S N HGTRC GE+AM A+N KCGVG+A
Sbjct 49592 DPKISYD CNDEHDH P L PRYDHSR FN SHGTRCAGEVAMTADNGKCGV
Query 291 DGLVNDRVEGEALGYKPELVD IYTSWGPADDGKSLEAPGRLEALE
DG+V DR+EG ALGY LVDIY+ASWGP DDGK+++ PGRLE EAL F
Sbjct 49772 DGIVTDRIEGTALGYAHLVD IYSAWGPNDDGKTVDPGRLEA
Query 351 IYVWAXXXXXXXXXXXXXXXXXXVGSYTIYTIYAVGSASQTGRFPWY
I+VWASGNGG+K D+C CDGY+ S YTI++GSASQ G FPWYGE C +T
Sbjct 49952 IFVWASGNGGKGDNCNDGYLASPYTISI GSASQRKEFPWY
Query 411 DQMIVT 416
DQMIV+
    
```

Figure 7. BLAST alignment of fruit fly (query) flour beetle (Sbjct) Furin sequences

```

Score = 249 bits (635), Expect = 7e-112
Identities = 112/186 (60%), Positives = 138/186 (74%), Gaps = 0
Frame = +2
Query 387 DPKASYD VNSHDDDPH PHYDMDTSDNRHGTRCAGEVAATANN SFC
DPK SYD N D DP+P YD + N HGTRCAGEVA TA+N C VGIAY
Sbjct 49592 DPKISYD CNDEHDH P L PRYDHSR FN SHGTRCAGEVAMTADNGKCGV
Query 447 DGDVDAVEARSLSLNPQHID IYSAWGPDDDGKTVDPG GELASRA
DG VTD +E +L +DIYSAWGP+DDGKTVDPG LA+ A G
Sbjct 49772 DGIVTDRIEGTALGYAHLVD IYSAWGPNDDGKTVDPGRLEA
Query 507 IFVWASGNGGREGDNCNDGYNTSIWTLISI SSATEEGHV
IF+WASGNGG + DMCNDGY S +T+SI SA+++G PWY E+C+STL
Sbjct 49952 IFVWASGNGGKGDNCNDGYLASPYTISI GSASQRKEFPWY
Query 567 EKQVVT 572
++ +W+
Sbjct 50132 DQMIVS 50149
    
```

## Conclusions

- All PC sequences are well conserved within class and species.
- The differences between sequences from different species vary with PC subfamily.
- The homology between honey bee, fruit fly and beetle PC sequences was high.
- The alignment of PC sequences from well studied species allows to enhance the annotation of genomes of species not yet sequenced and can help uncover PC genes in the

## Acknowledgement

ACES James Scholar Honors Program  
This material is available at <http://bighorn.anim>