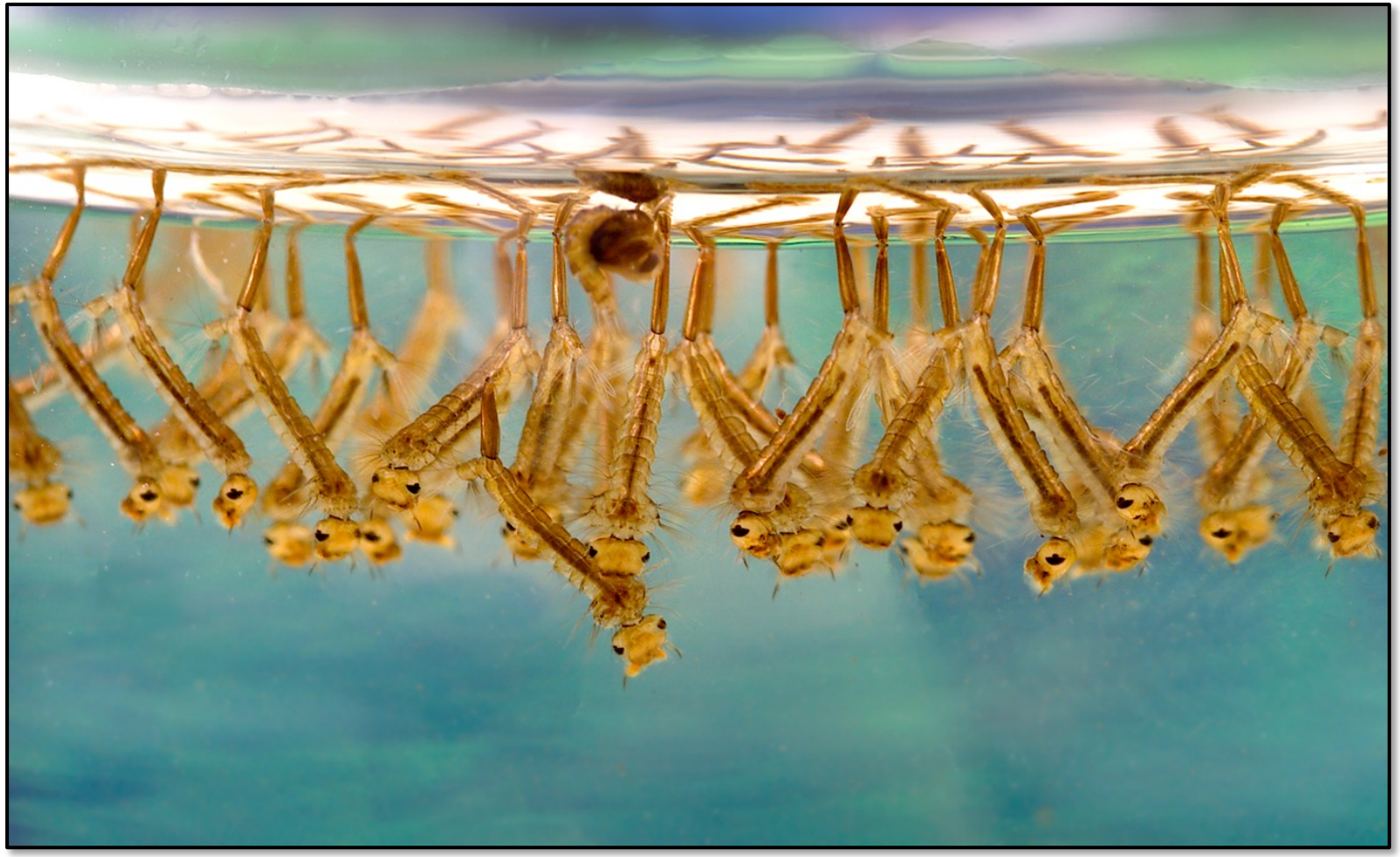
Agenda

1. Intros
2. AI
3. GMM
4. **GMM + AI**



Next-Generation Surveillance

to predict and prevent mosquito-borne diseases on a global scale

user action: | targets: | analysis: | outputs: | outcomes:

photo

report

tip/toss

mosquitoes
(adult, larval)

hosts

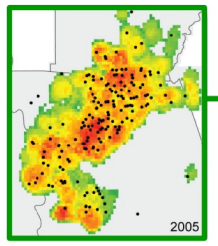
habitats

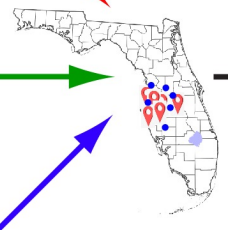
AI
(Thrust 1)

DYCAST
(Thrust 2)

RS, GIS
(Thrust 3)

- species ID
- distribution maps
- disease risk maps
- habitat maps





real-time maps used
by districts for targeting:

- mosquito control
- public education
- app promotion

reduced breeding

- FEWER MOSQUITOES & LESS DISEASE
- EDUCATED & ENGAGED CITIZENS
- SCIENTIFIC RESEARCH

Citizen Scientists,
Mosquito Control Personnel

global implementation and engagement

GLOBE mosquito habitat mapper

Global Mosquito
Alert Consortium

NASA

CDC
Centers for Disease
Control and Prevention

AI (Initial Results)

adults
AI
phones

Class Activation Map (CAM)

a). *Ae. aegypti*

a). *Ae. aegypti*

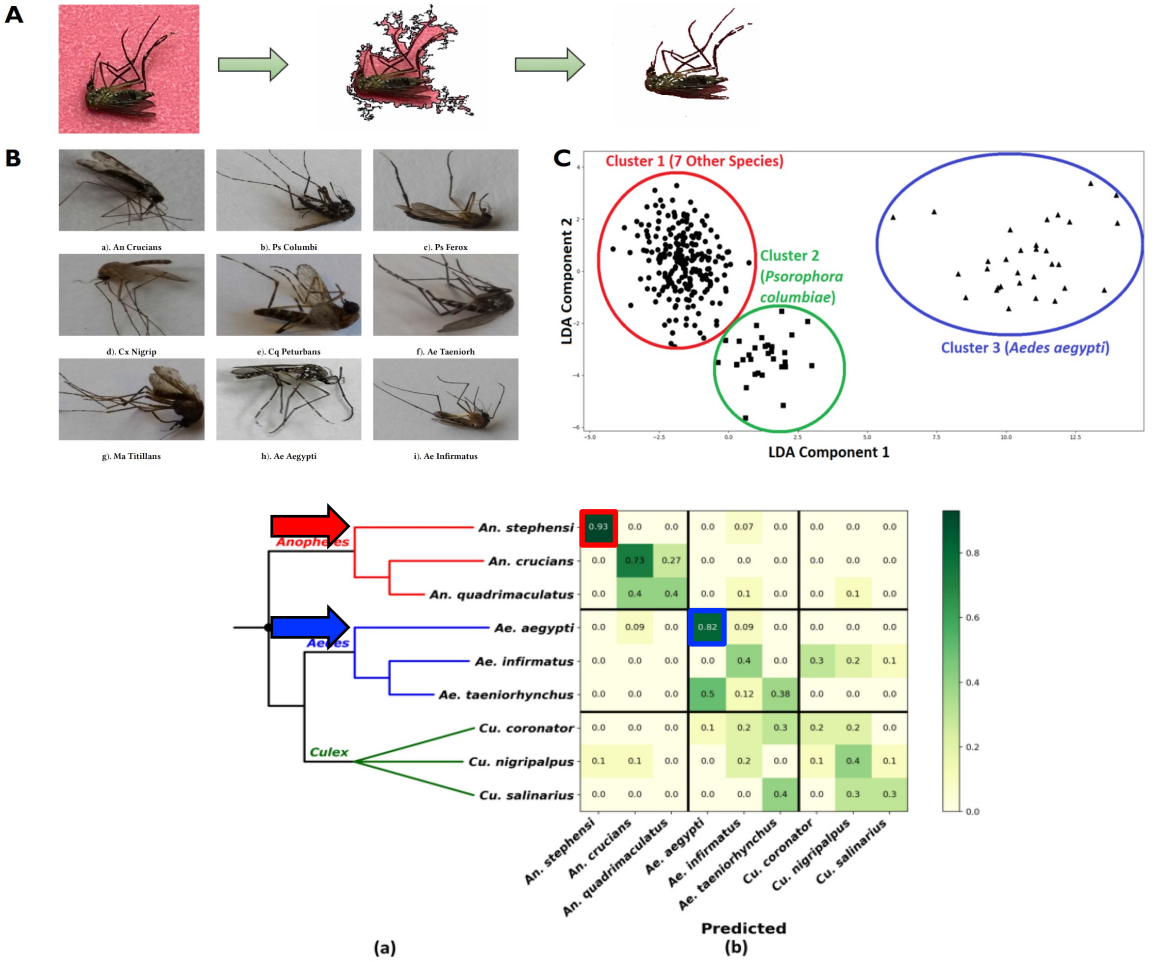
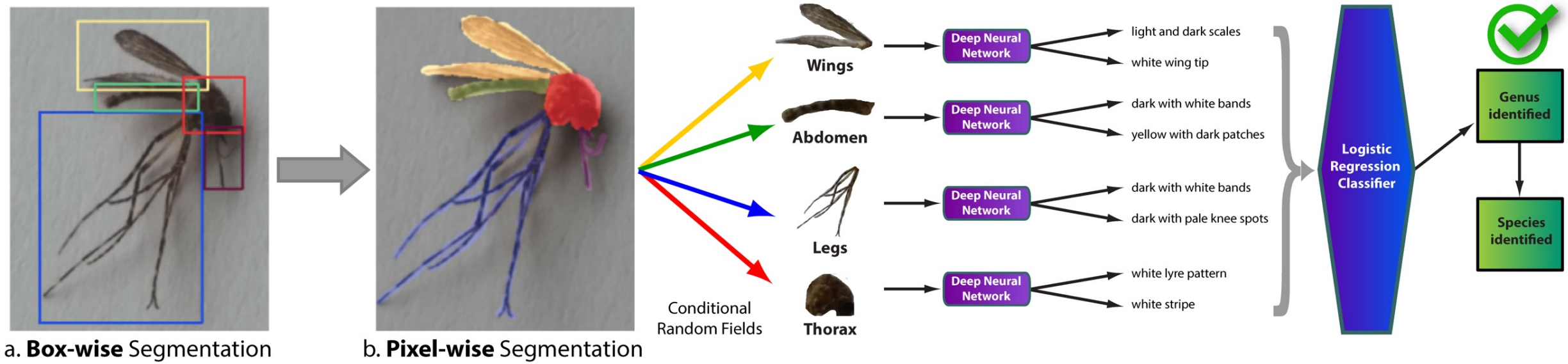
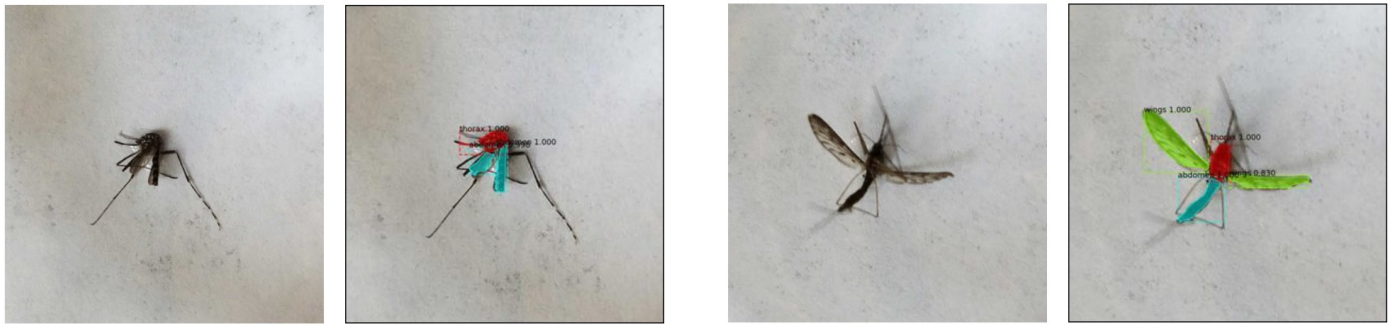


Figure 3: (a) Phlyogenetic tree showing evolutionary relationships of the mosquitoes [10] [11] [22]. Leftward is further back in geologic time. Black dot denotes the common ancestor, ~ 217 million years ago [22]. (b) Species testing set confusion matrix with classification accuracies. This Figure is best viewed in Color.

Minakshi et al 2020a *ACM COMPASS*





Genus: Aedes
Species: aegypti

Genus: Anopheles
Species: crucians

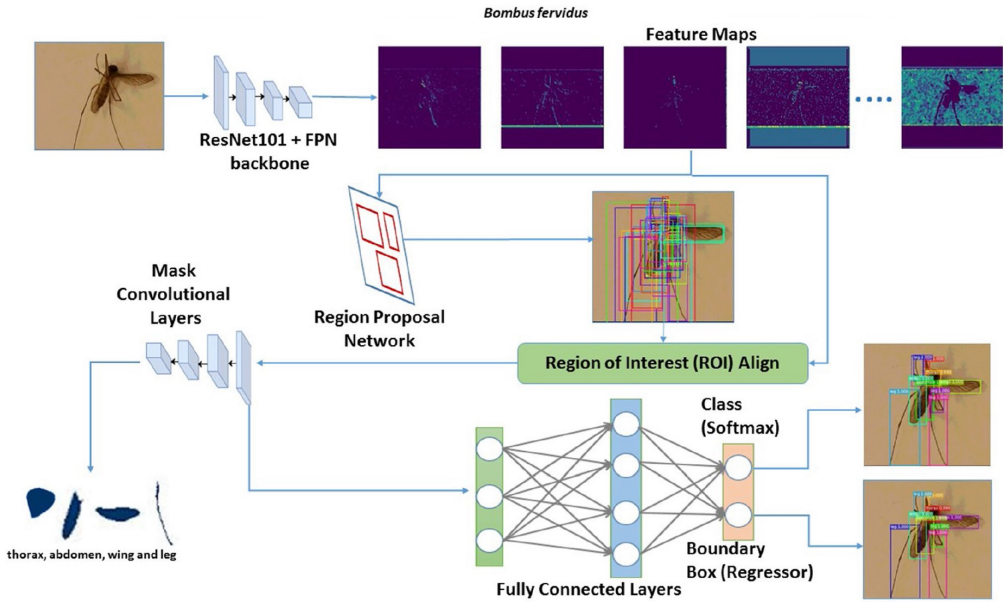


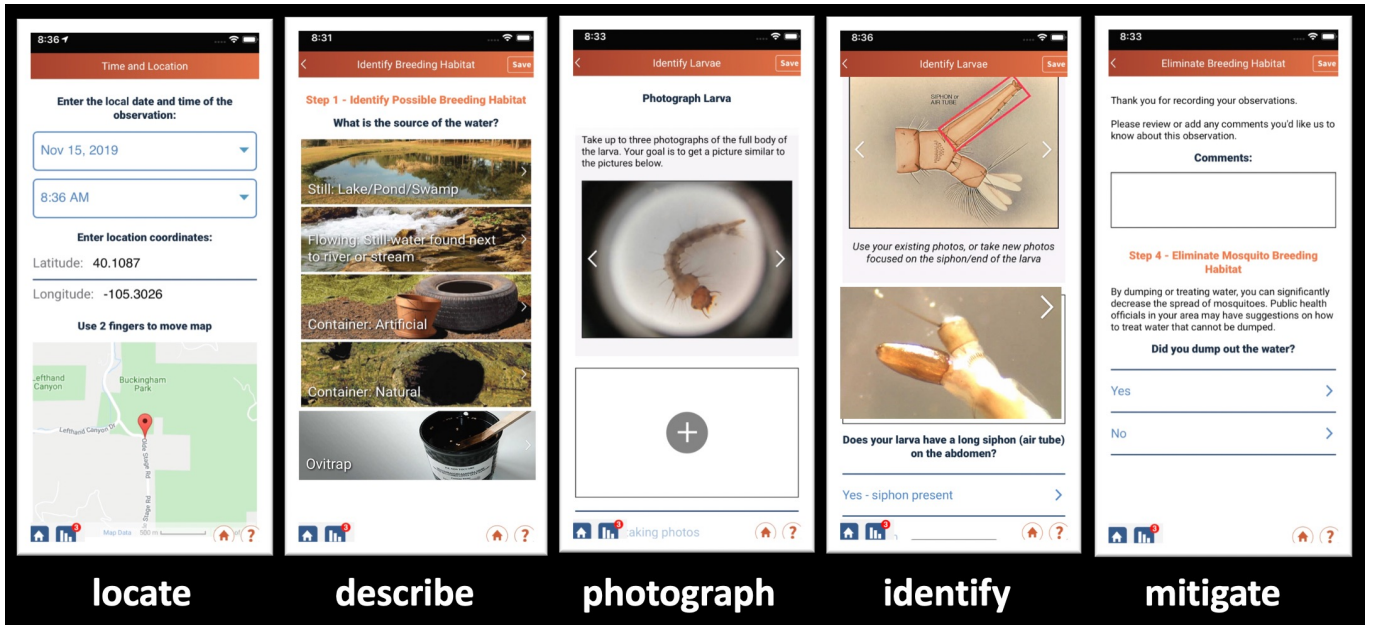
Figure 1. The workflow of our architecture based on Mask R-CNN.

Minakshi et al 2020b *Scientific Reports*

Mask Region-based Convolutional Neural Network

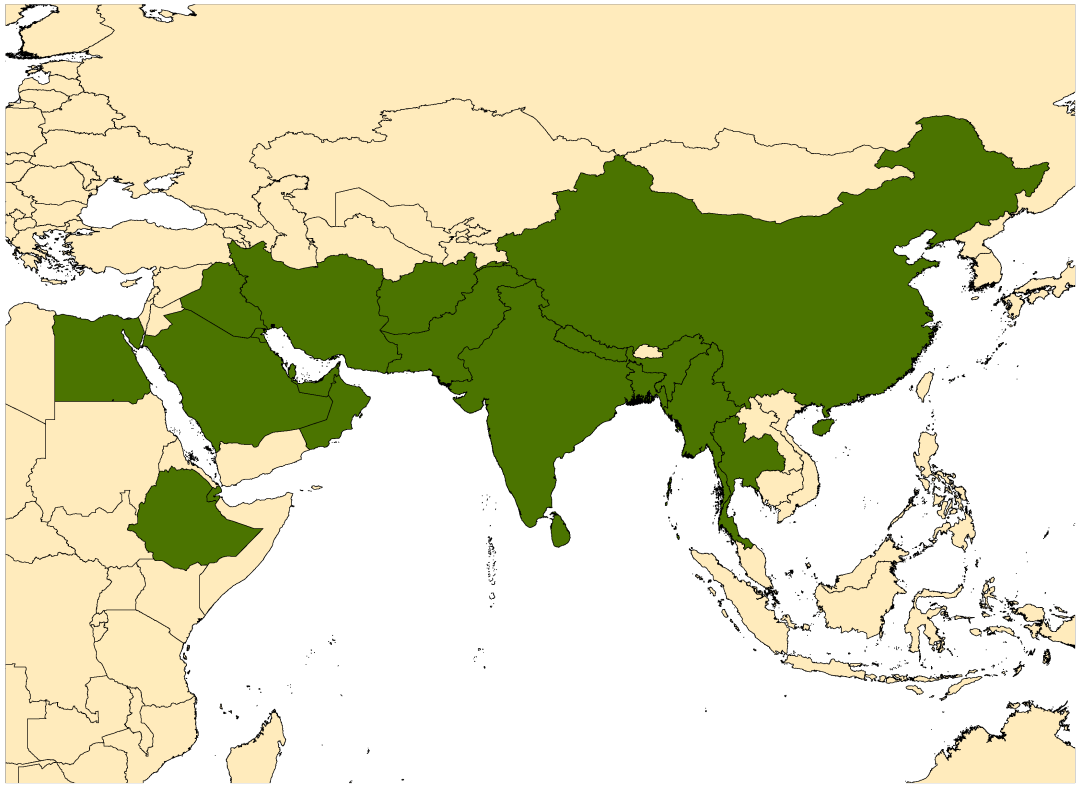
Mosquito Habitat Mapper app

larvae
AI
phones

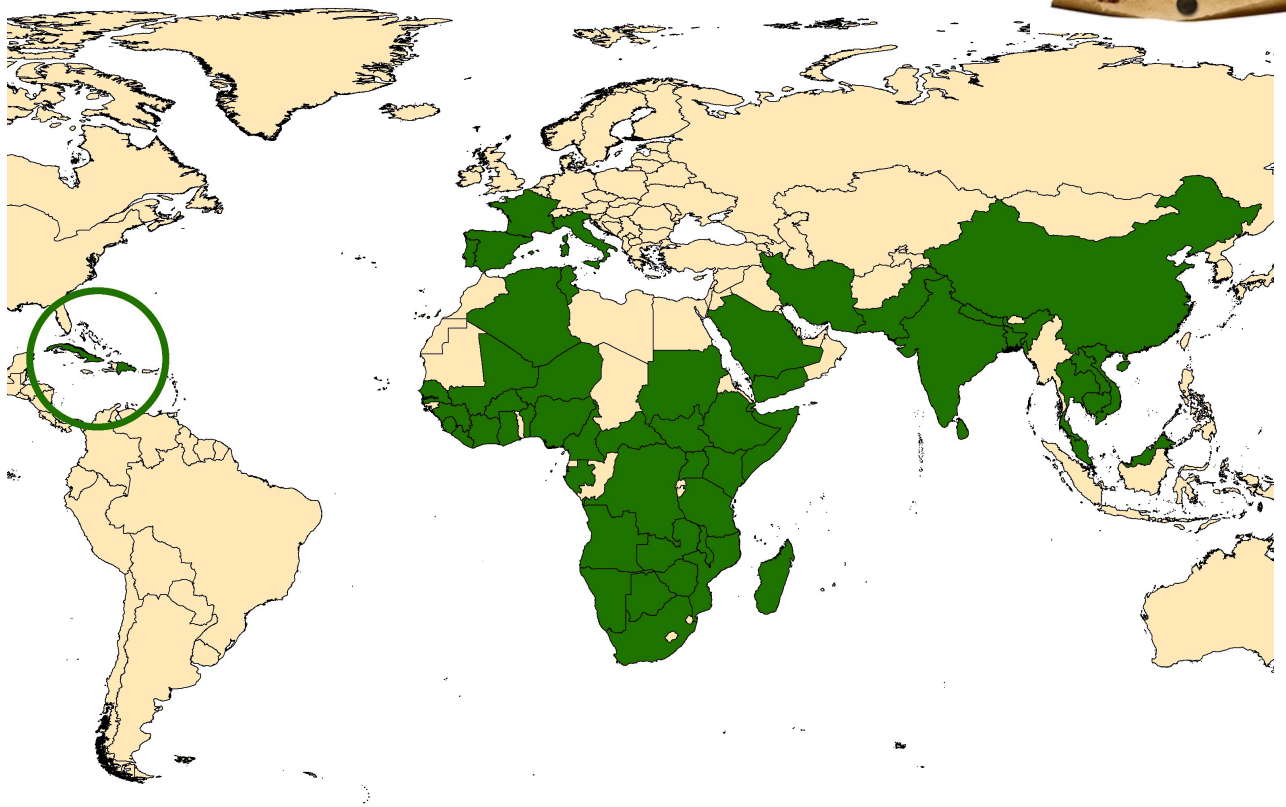


Artificial Containers

Anopheles stephensi

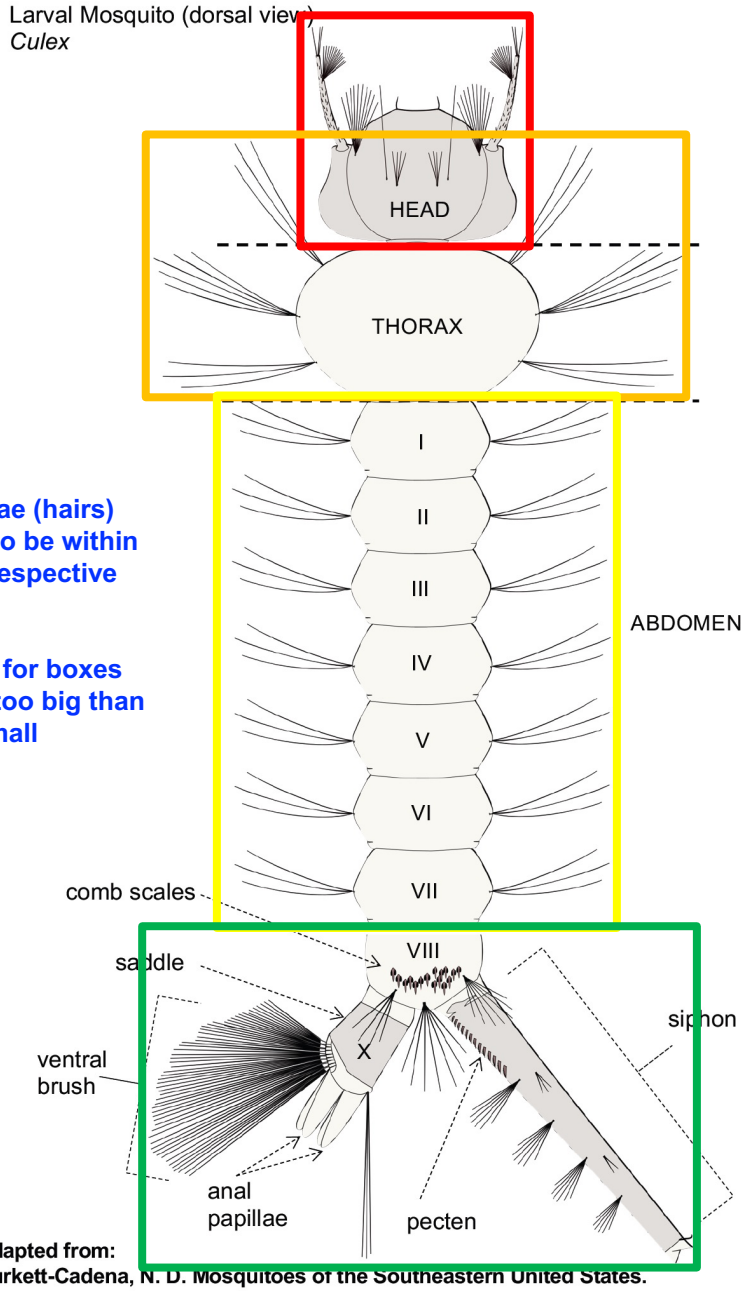


Aedes vittatus



+ Land Cover data!

GLOBE Land Cover



1. **all setae (hairs) need to be within their respective boxes**

2. **better for boxes to be too big than too small**

Adapted from:
Burkett-Cadena, N. D. Mosquitoes of the Southeastern United States. Tuscaloosa: The University of Alabama Press. 2013. 202 pages.
https://fmel.ifas.ufl.edu/media/fmelifasufledu/workshop/Mosquito_Morphology.pdf

"**The shape of the head may be elongate (as in *Anopheles* and *Uranotaenia*) or broad (*Aedes* and *Culex*).**

The antennae are quite variable, and may be very short to quite long. One or more setae are borne usually along the length of the antenna and may be branched or unbranched.

Setae of the head are numerous and variable in length and form. **The arrangement, length, branching and shape of head setae are used in the identification of larvae**."

"**Thoracic setae are often useful in identification of mosquito larvae**."

"The abdomen is elongate, cylindrical and is made up of ten segments. Segments of the abdomen are denoted in Roman numerals, beginning with the most anterior segment (Segment I) and terminating in the anal segment (Segment X).

Segments I – VII are fairly uniform in size and shape and together constitute the bulk of the worm-like body.

"**Segment VIII** is usually smaller than the seven preceding segments, roughly pentagonal in shape, and bears the comb scales (when present) and the respiratory siphon (when present).

The comb scales are spine-like projections that occur in a row or patch, and are sometimes borne on a sclerotized plate, called the comb plate (as in *Uranotaenia*). **The number, shape and arrangement of comb scales are useful in identification of larvae, but often require high magnification (>50x) to examine in detail**. The respiratory siphon (or simply siphon) is a sclerotized dorsal breathing tube that bears the respiratory spiracles.

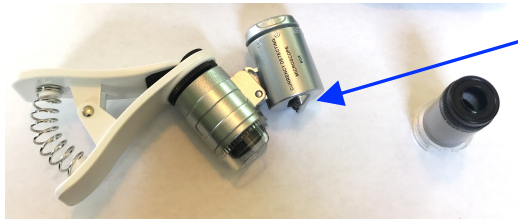
In most mosquito species of our region, the siphon bears a pecten, a row of spines (spicules) extending from the ventral base of the siphon to some point along its length. The size, shape and length of the siphon and the pecten vary from one species to the next and are very useful in genus and species-level mosquito larva identification.

Members of the genus *Anopheles* have no siphon, but breathe through a flattened spiracular apparatus on segment VIII.

Segment IX is reduced in mosquito larvae and is not discernible as a distinct segment.

The anal segment (**Segment X**) bears the anal papillae, saddle, and ventral brush. The anal papillae are bulbous, membranous protrusions of the exoskeleton that primarily function in osmoregulation. The saddle, a sclerotized plate, may cover only the dorsal portion of the anal segment, or may encircle it completely. The ventral brush is a row of paired setae extending along the ventral midline of the anal segment."

TAKE 3 PHOTOS WITH A 60X LENS:



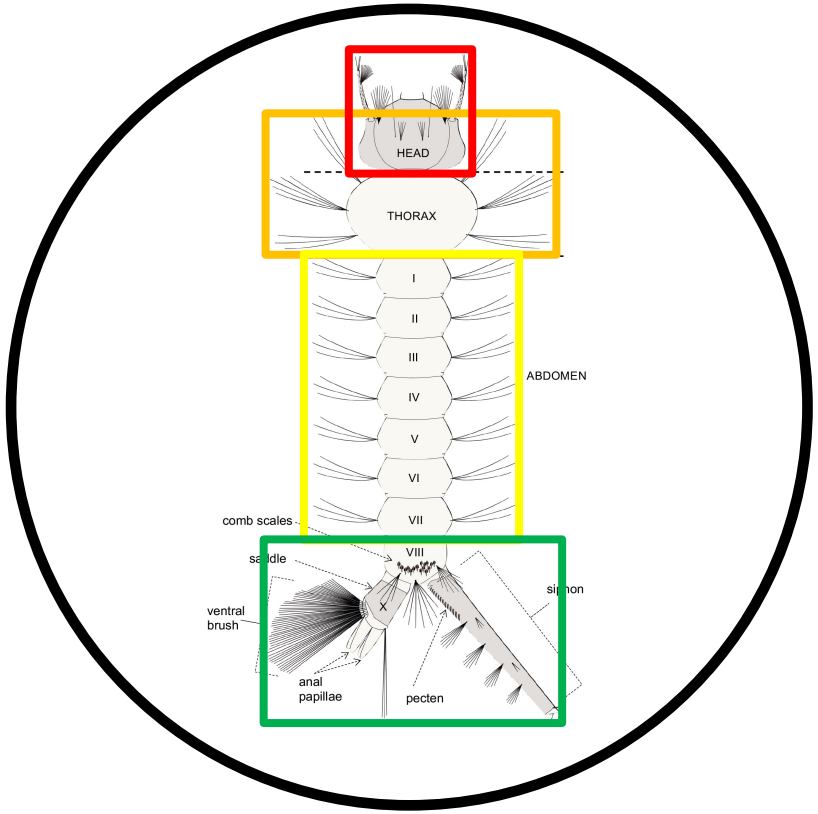
- illuminate larvae (white lights on clip-on, or sunlight or flashlight)
- DON'T extend the lenses; instead use digital zoom on phone

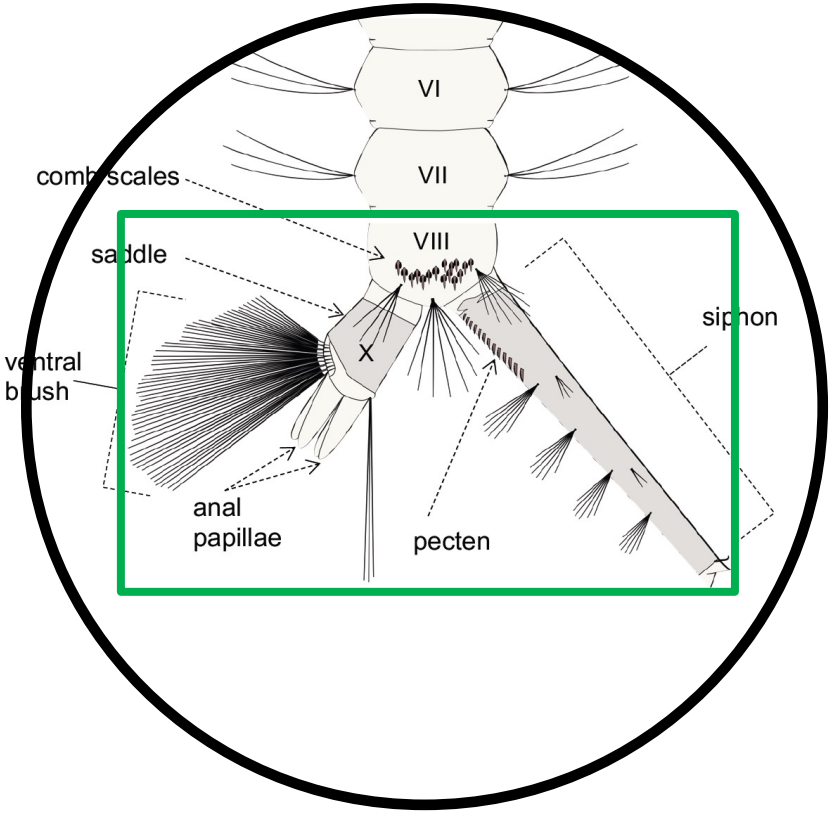
1. FULL BODY + all hairs

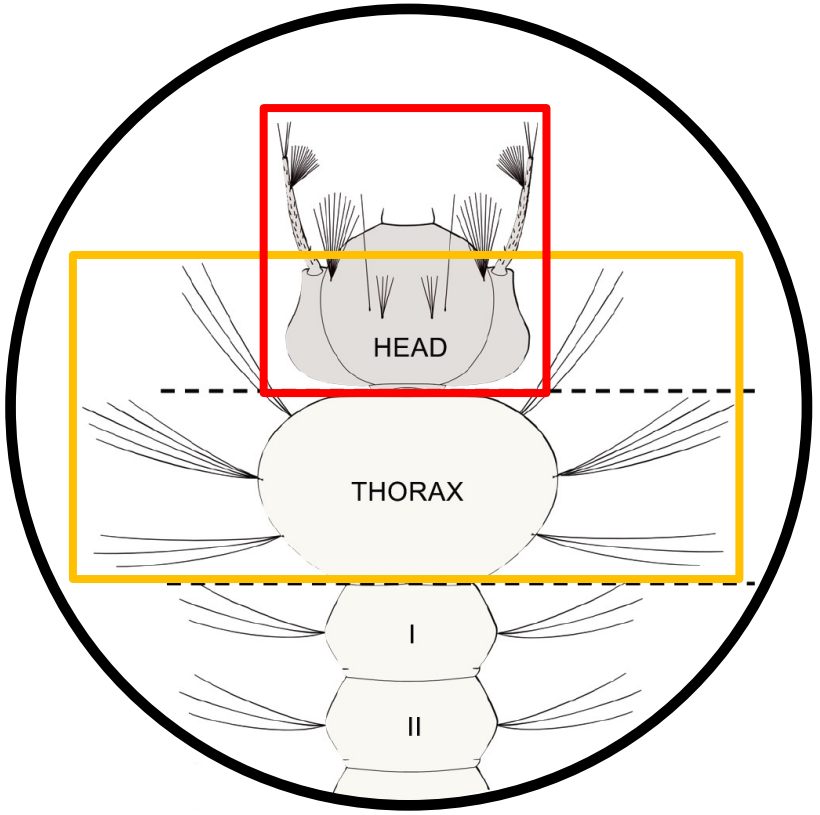
(use digital zoom) →

2. TAIL + all hairs

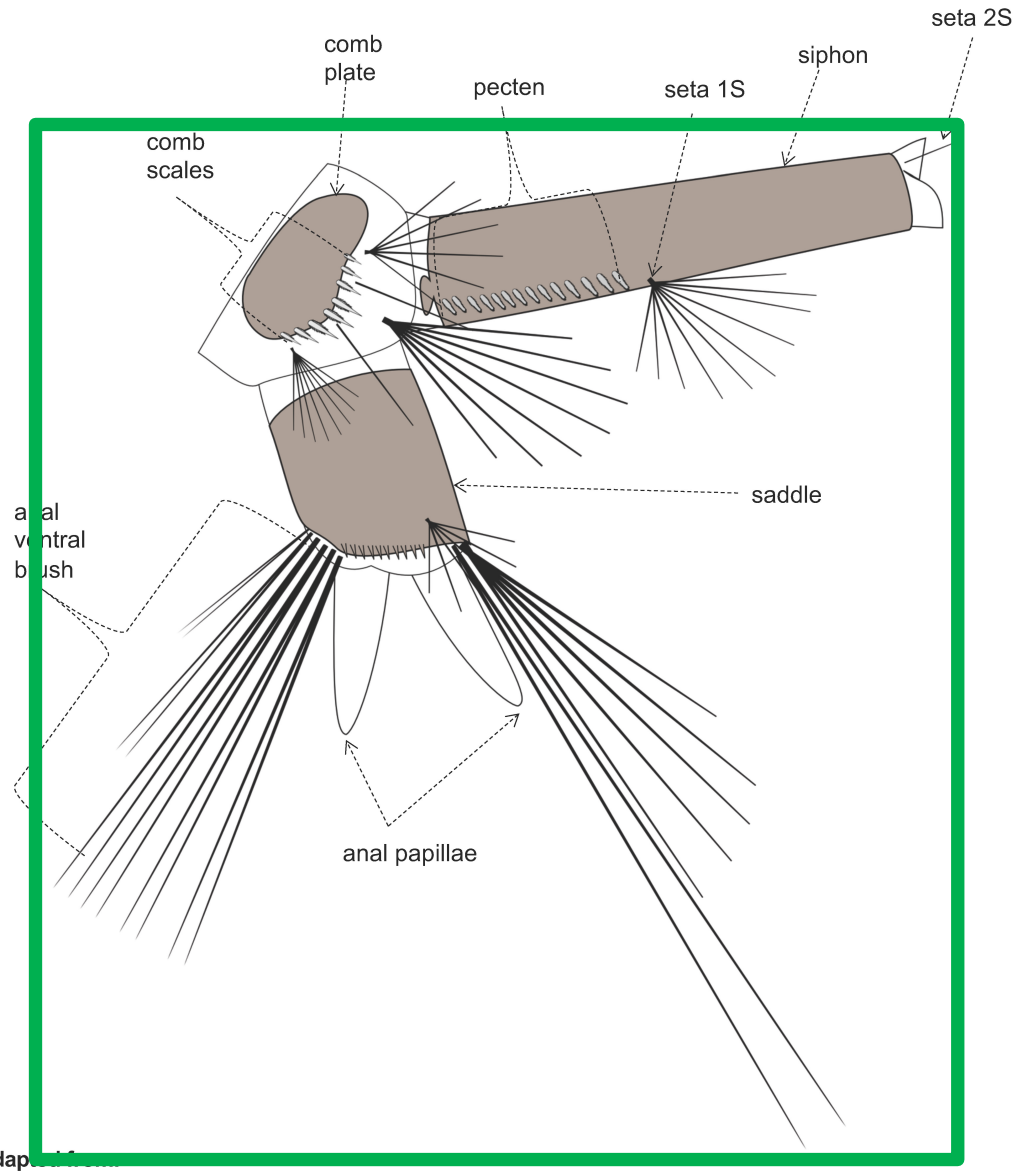
3. HEAD + THORAX + all hairs





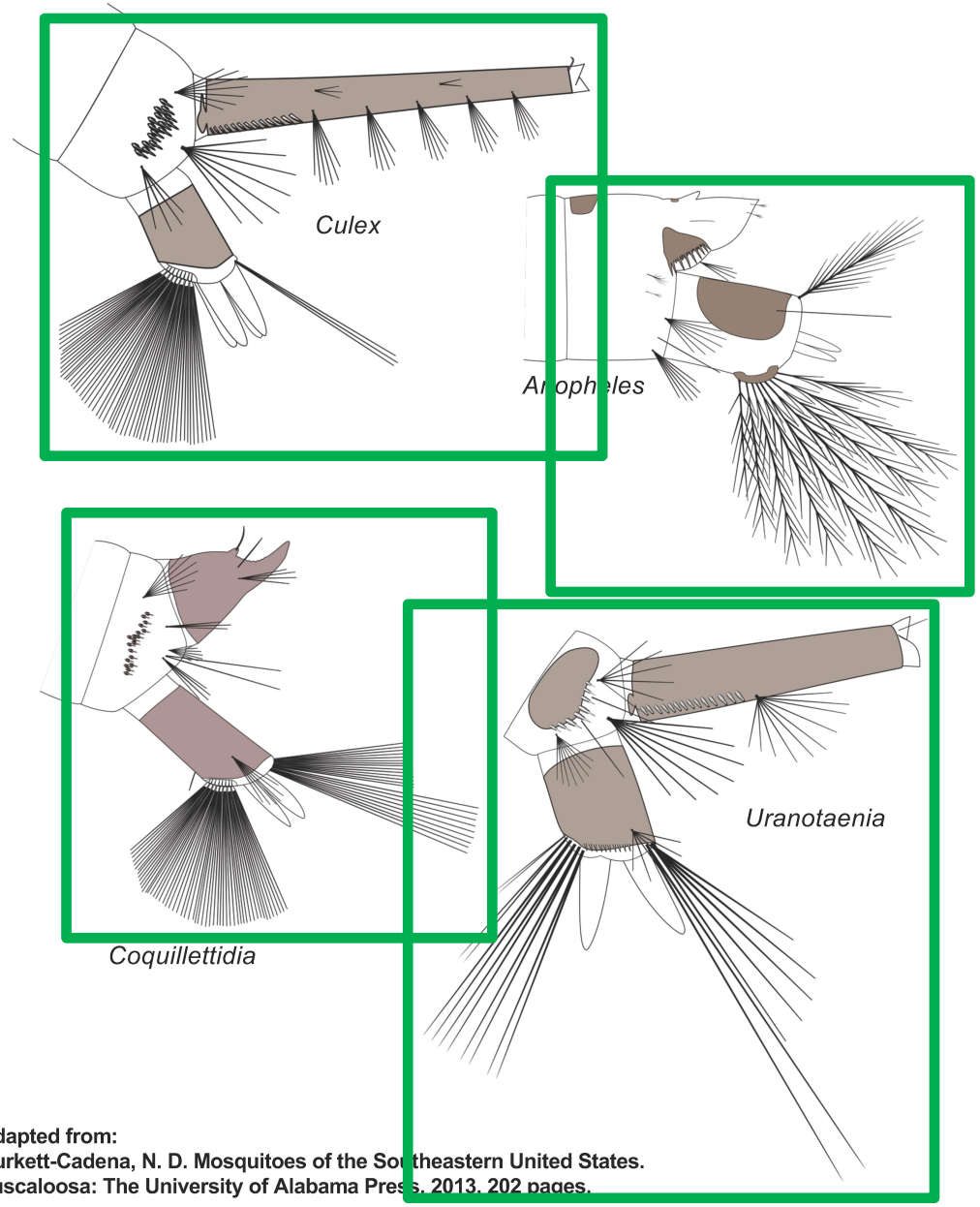


Larval Mosquito terminal abdominal segments (lateral view)
Uranotaenia



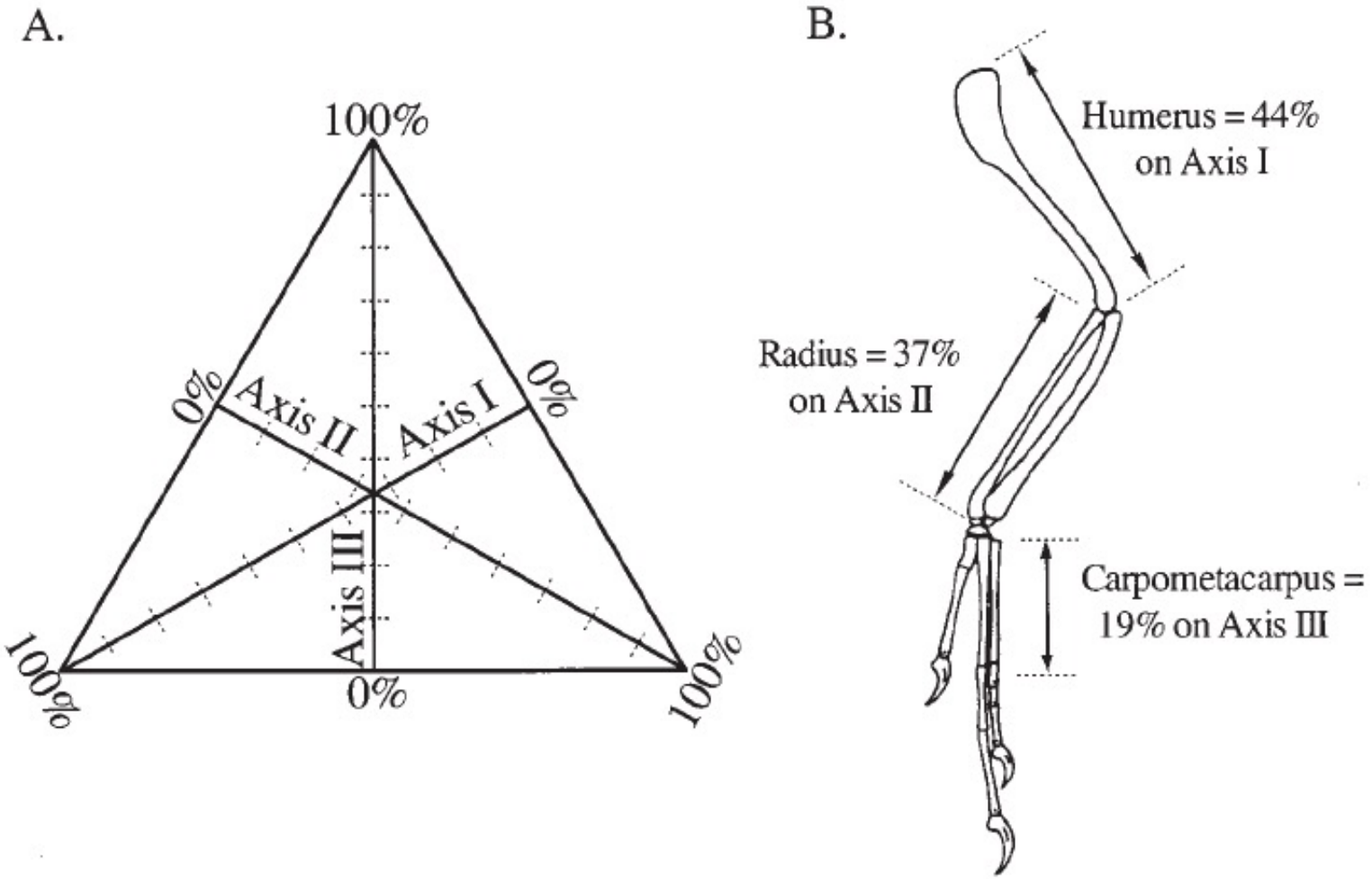
Adapted from:
Burkett-Cadena, N. D. Mosquitoes of the Southeastern United States. Tuscaloosa: The University of Alabama Press. 2013. 202 pages.

Larval Mosquito terminal abdominal segments (lateral view)



Adapted from:
Burkett-Cadena, N. D. Mosquitoes of the Southeastern United States. Tuscaloosa: The University of Alabama Press. 2013. 202 pages.

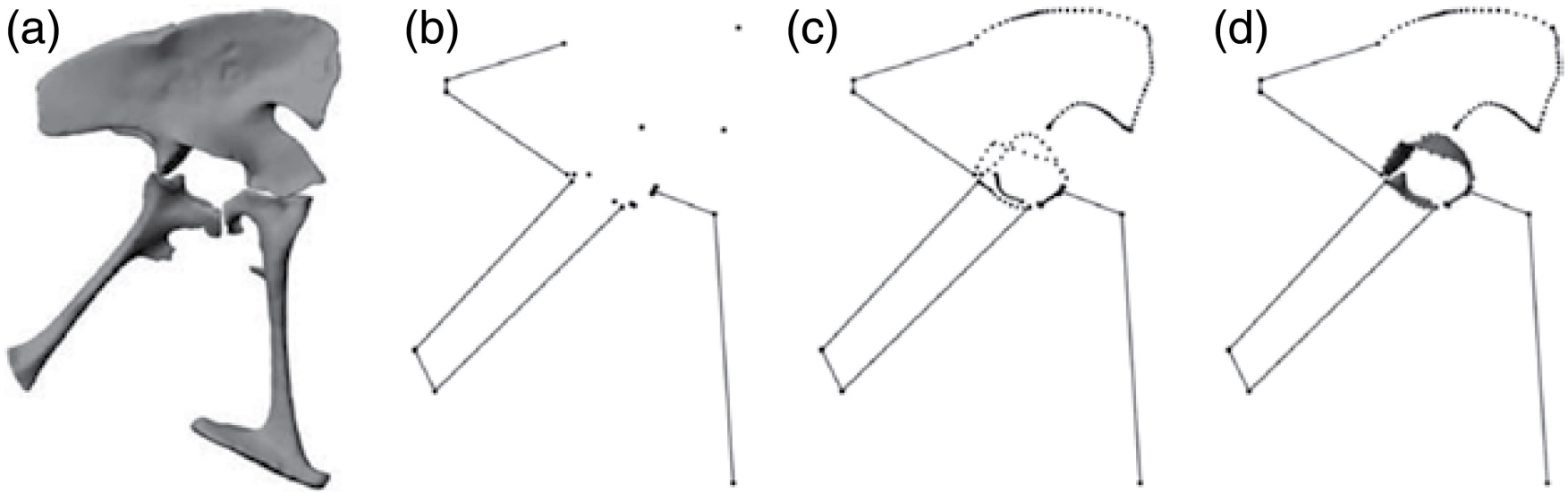
Morphometrics



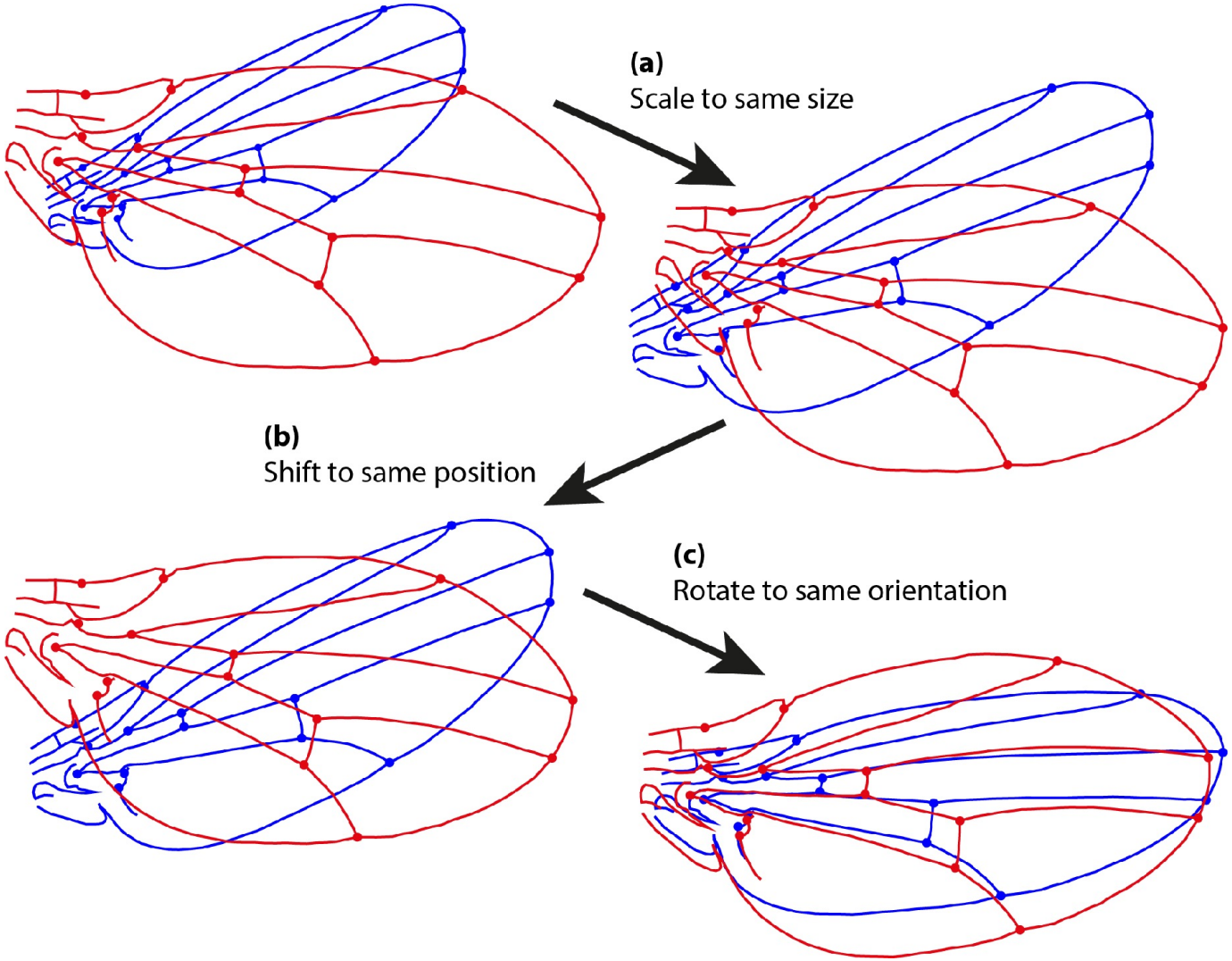
Middleton & Gatesy 2000

GMM: Geometric Morphometrics

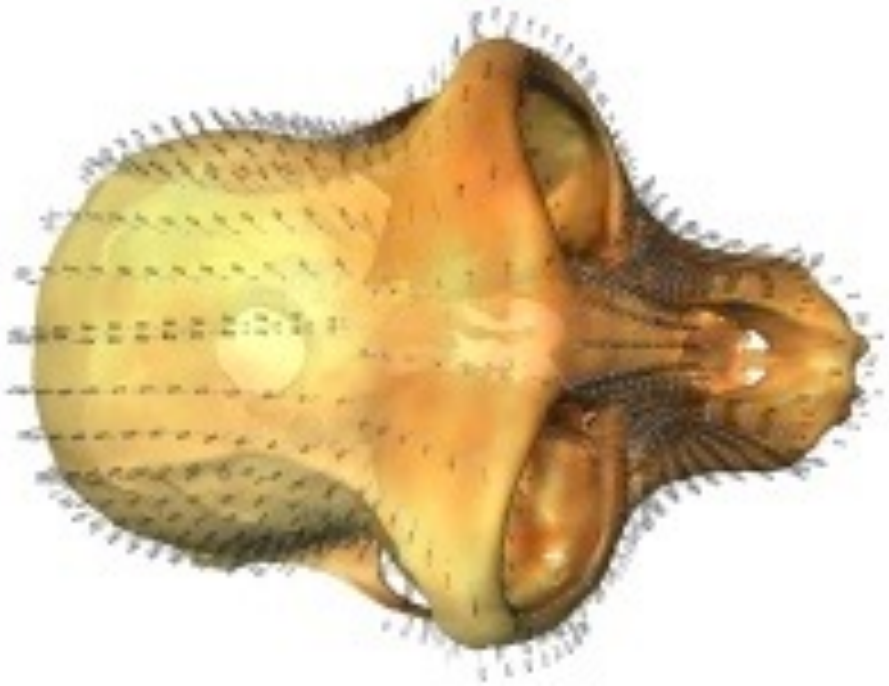
- **MorphoJ, R (geomorph)**

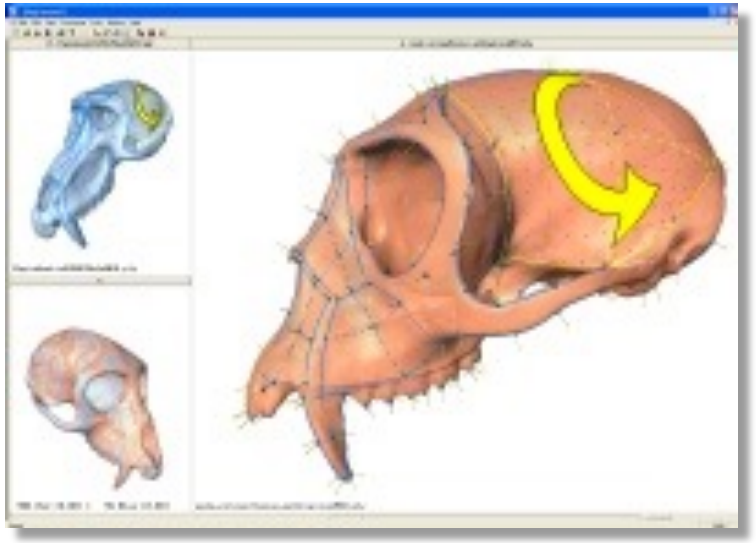


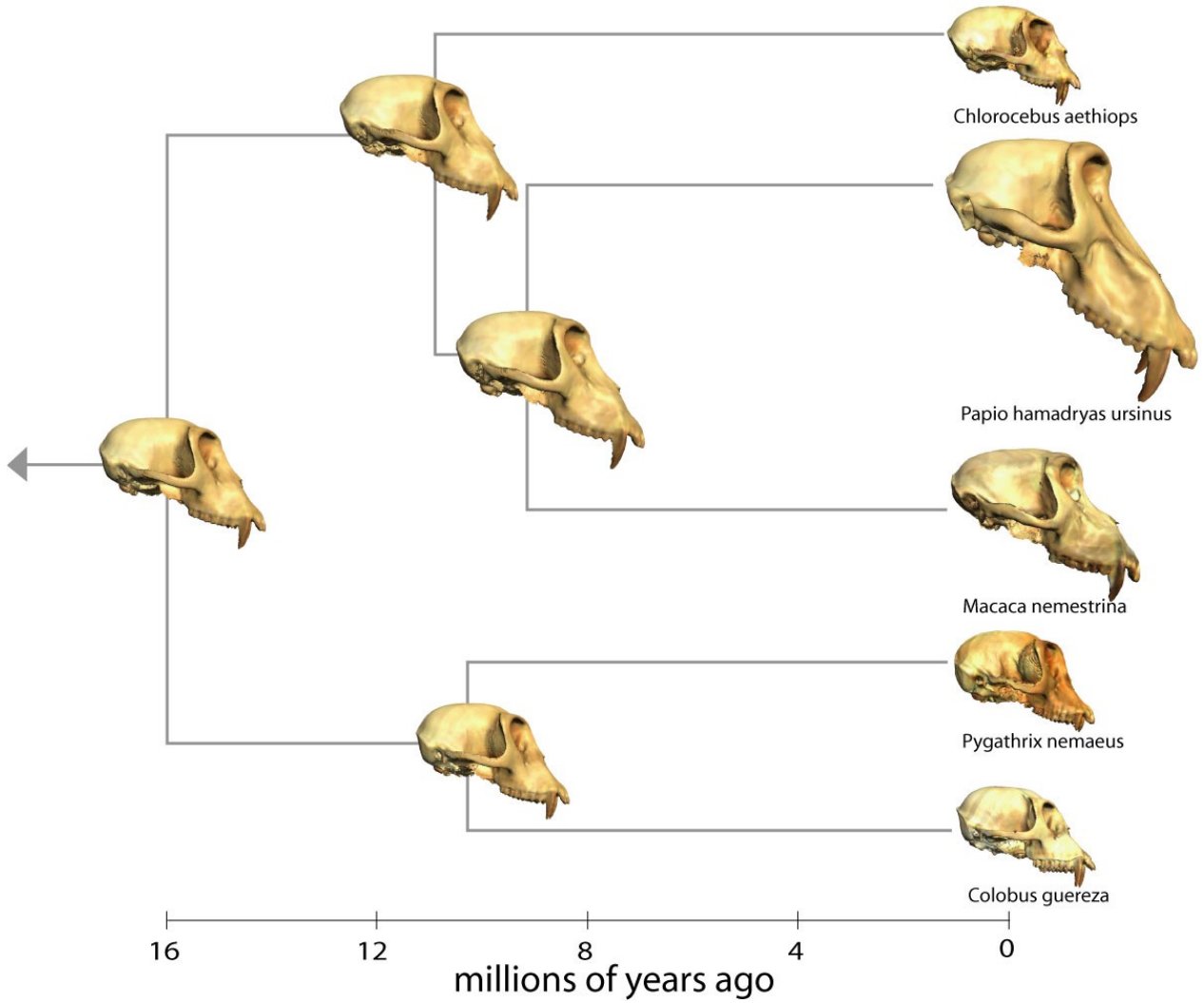
Procrustes

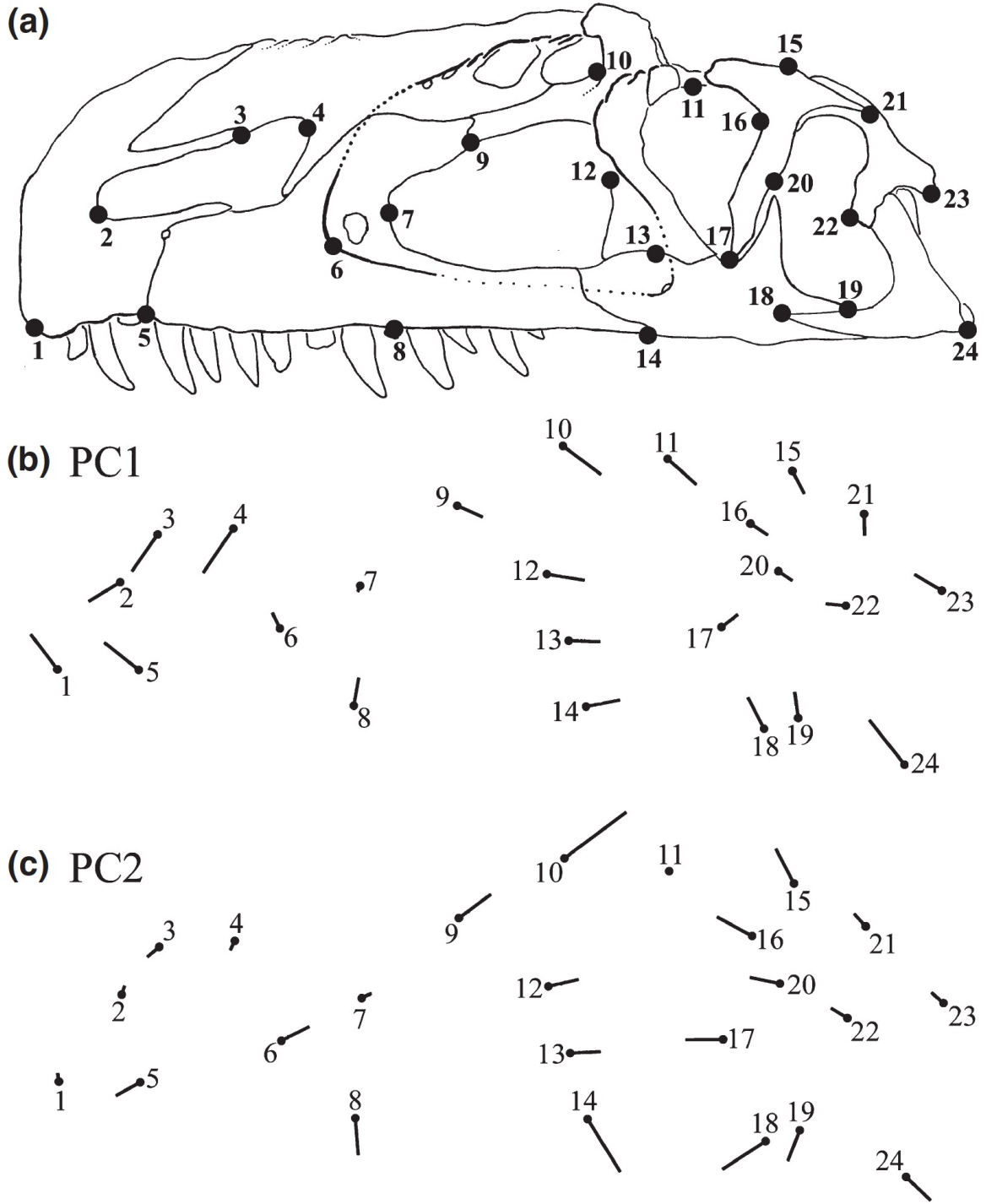

(a)
Scale to same size
(b)
Shift to same position
(c)
Rotate to same orientation

Landmark

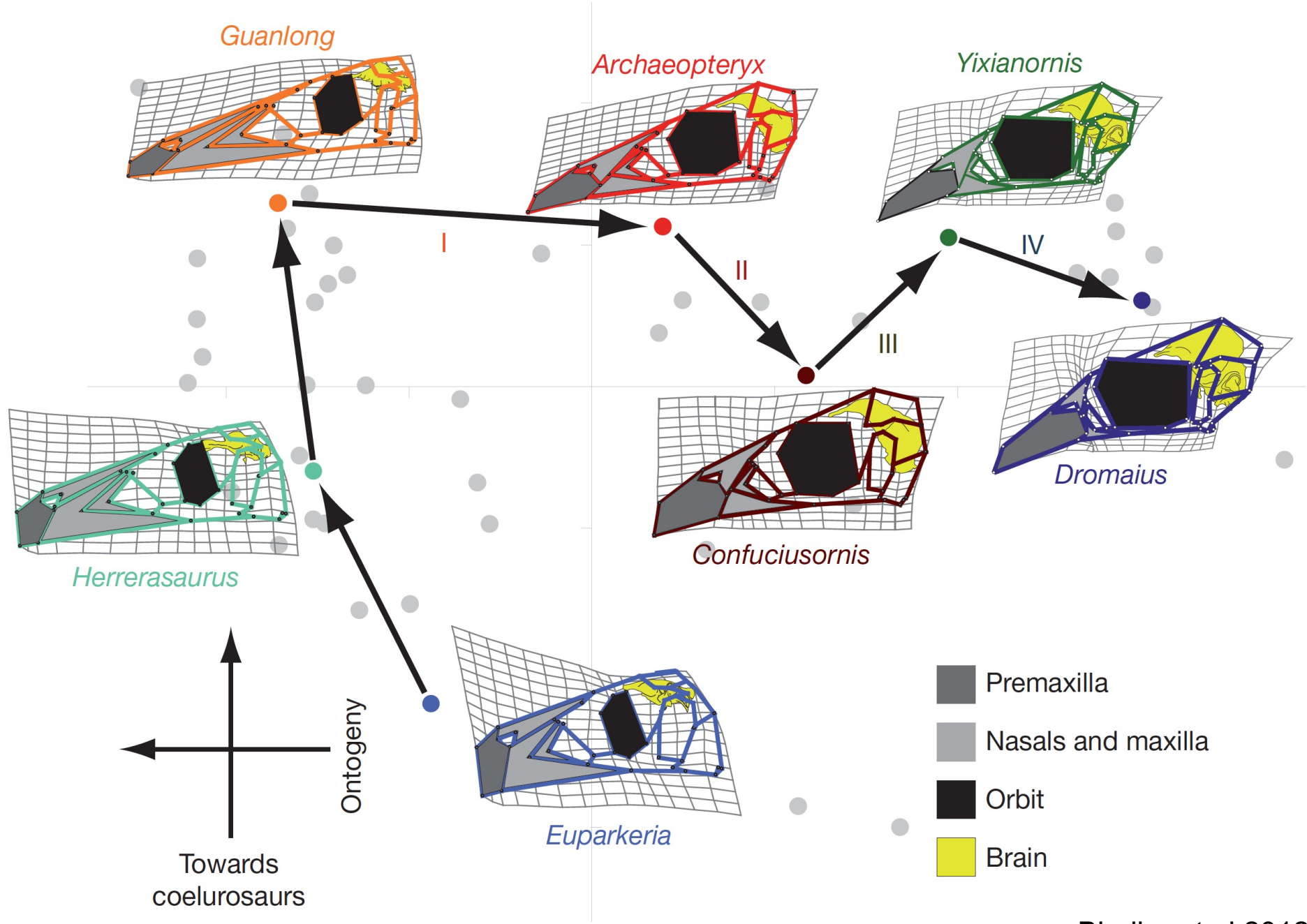





Chlorocebus aethiops
Papio hamadryas ursinus
Macaca nemestrina
Pygathrix nemaeus
Colobus guereza
16
12
8
4
0
millions of years ago

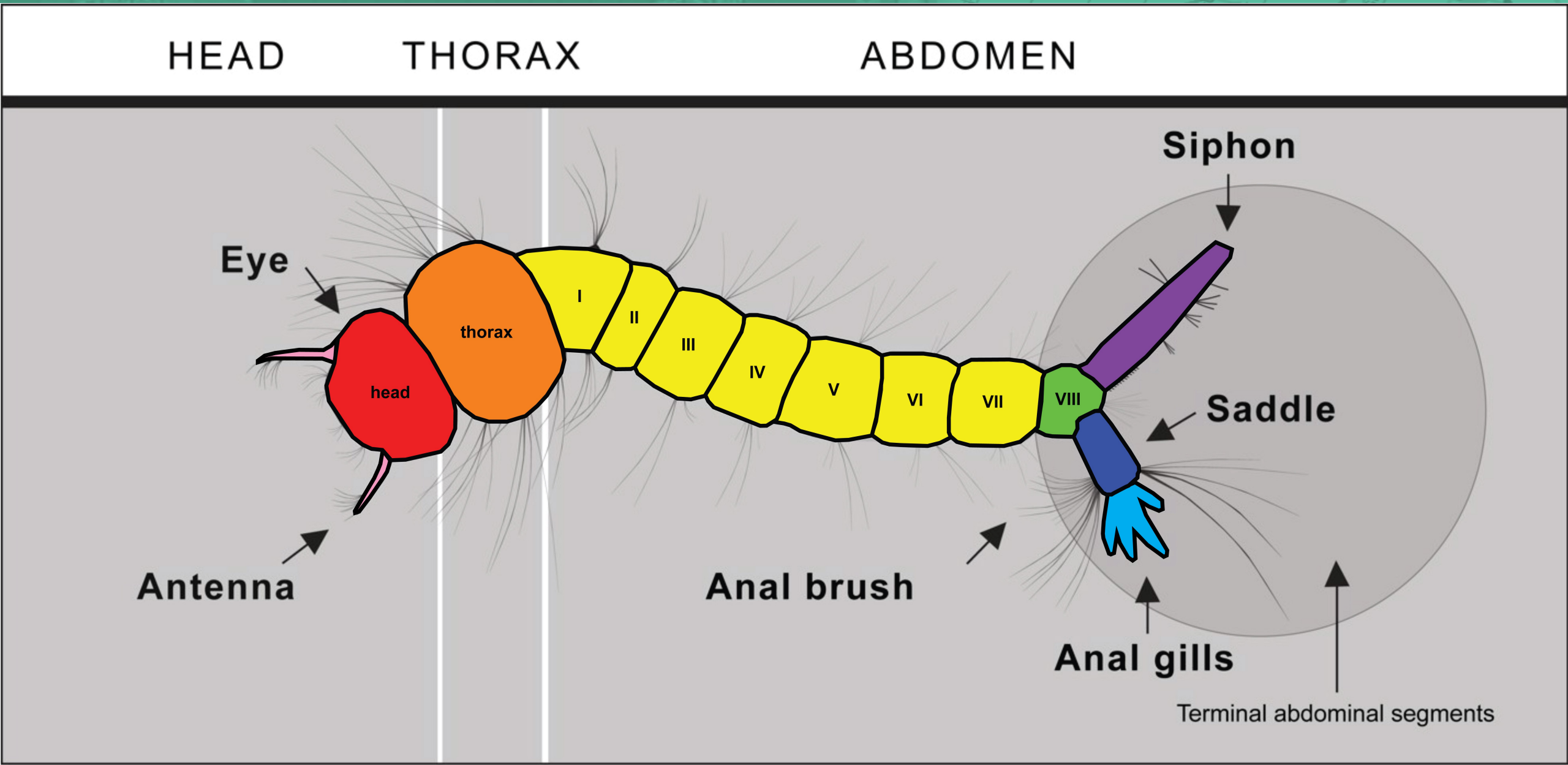


Brusatte et al 2012

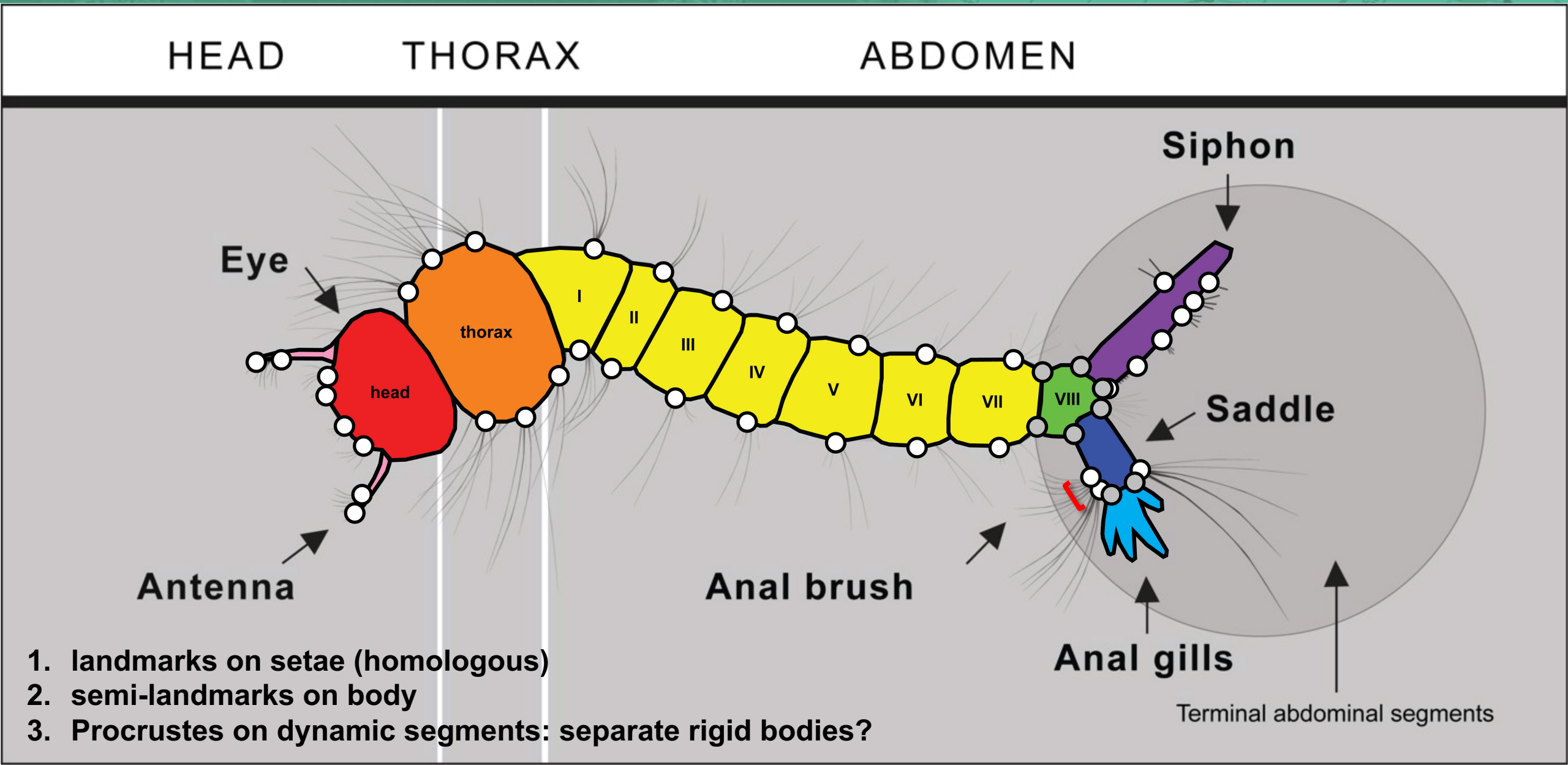


Bhullar et al 2012

Mosquito Larva



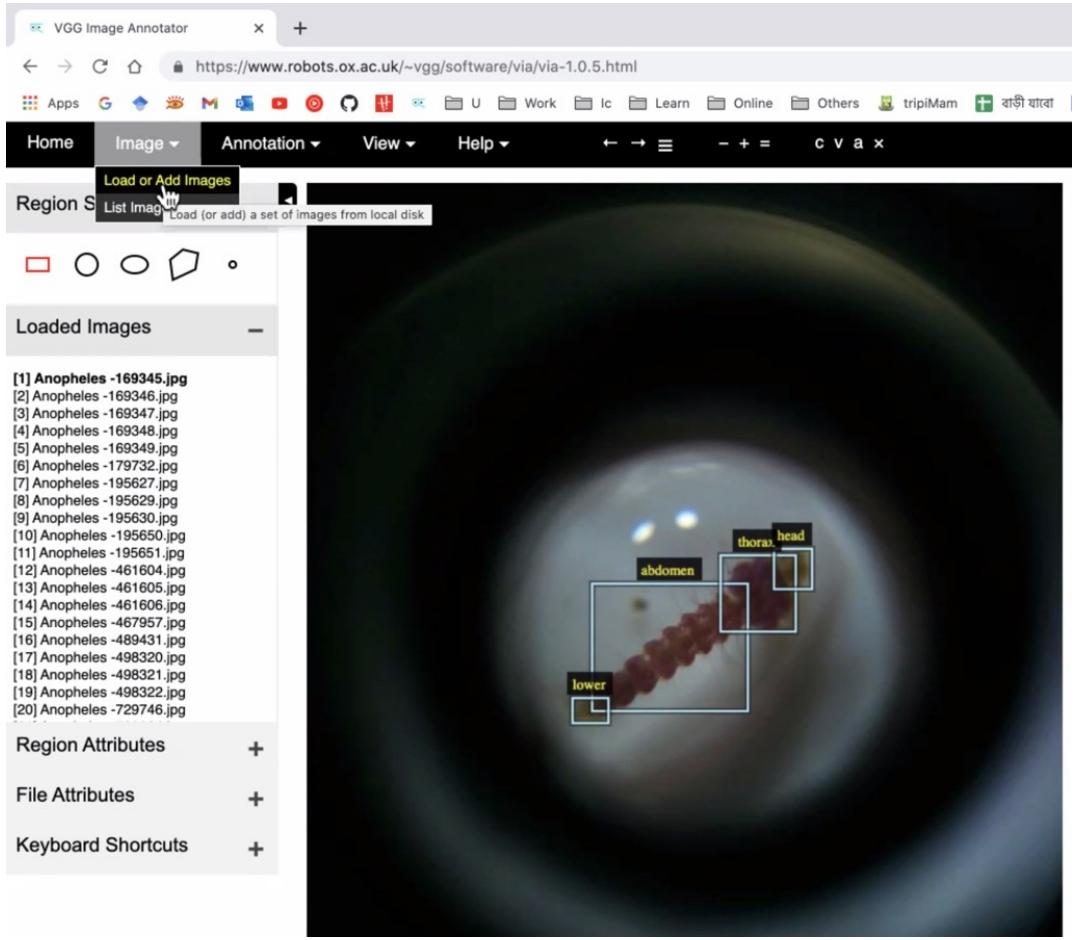
Landmarking



1. landmarks on setae (homologous)
2. semi-landmarks on body
3. Procrustes on dynamic segments: separate rigid bodies?

Landmarking: VGG Image Annotator

VIA homepage: https://www.robots.ox.ac.uk/~vgg/software/via/
browser annotator: https://www.robots.ox.ac.uk/~vgg/software/via/via.html



exports as **json** or **csv.** CSV results:

1	2	3	4	5	6	7
filename	file_size	file_attributes	region_count	region_id	region_shape_attributes	region_attributes
VGG.jpg	1184779	{}	19	0	{"name":"point","cx":1003,"cy":1243}	{}
VGG.jpg	1184779	{}	19	1	{"name":"point","cx":1012,"cy":1223}	{}
VGG.jpg	1184779	{}	19	2	{"name":"point","cx":1032,"cy":1223}	{}
VGG.jpg	1184779	{}	19	3	{"name":"point","cx":1034,"cy":1203}	{}
VGG.jpg	1184779	{}	19	4	{"name":"point","cx":1063,"cy":1197}	{}
VGG.jpg	1184779	{}	19	5	{"name":"point","cx":1077,"cy":1170}	{}
VGG.jpg	1184779	{}	19	6	{"name":"point","cx":1192,"cy":1074}	{}
VGG.jpg	1184779	{}	19	7	{"name":"point","cx":1246,"cy":1037}	{}
VGG.jpg	1184779	{}	19	8	{"name":"point","cx":1283,"cy":992}	{}
VGG.jpg	1184779	{}	19	9	{"name":"point","cx":1346,"cy":988}	{}
VGG.jpg	1184779	{}	19	10	{"name":"point","cx":1399,"cy":1041}	{}
VGG.jpg	1184779	{}	19	11	{"name":"point","cx":1341,"cy":1070}	{}
VGG.jpg	1184779	{}	19	12	{"name":"point","cx":1266,"cy":1101}	{}
VGG.jpg	1184779	{}	19	13	{"name":"point","cx":1206,"cy":1152}	{}
VGG.jpg	1184779	{}	19	14	{"name":"point","cx":1163,"cy":1161}	{}
VGG.jpg	1184779	{}	19	15	{"name":"point","cx":1154,"cy":1192}	{}
VGG.jpg	1184779	{}	19	16	{"name":"point","cx":1105,"cy":1192}	{}
VGG.jpg	1184779	{}	19	17	{"name":"point","cx":1101,"cy":1226}	{}
VGG.jpg	1184779	{}	19	18	{"name":"point","cx":1081,"cy":1228}	{}

Landmarking

...modify .csv manually in Excel or text editor (TextEdit, Notepad), or via script, to create the necessary Morphologika txt file for **MorphoJ:**

Cranial landmarks.txt

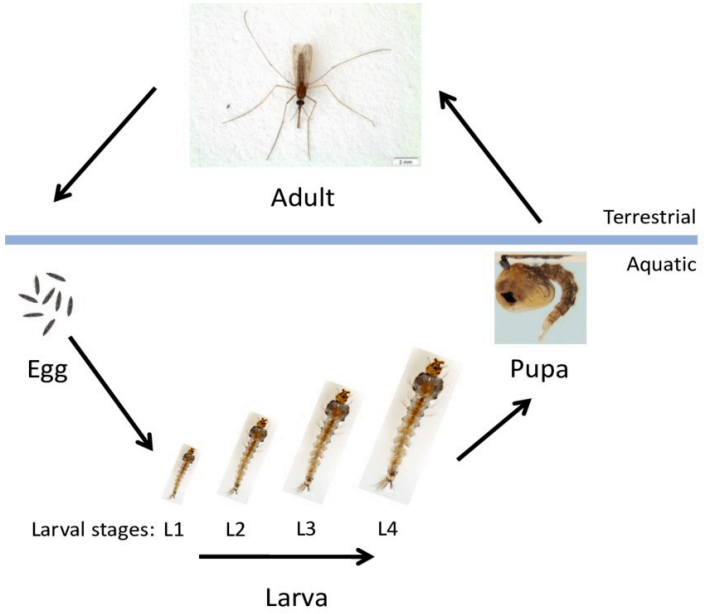
```
[individuals]
33
[landmarks]
24
[dimensions]
3
[names]
Acomys_cahirinus
Aplodontia_rufa
Bathyergus_suillus
Cannomys_badius
Capromys_pilorides
Castor_canadensis
Cavia_porcellus
Cricetomys_gambianus
Ctenomys_opimus
Dasyprocta_punctata
Daubentonia_madagascarensis
Dendrohyrax_arboreus
Dipus_sagitta
Erethizon_dorsata
Georychus_capensis
Gerbillus_watersi
Graphirus_nagtglasii
Hydrochoerus_hydrochaeris
Hystrix_cristata
Lagostomus_maximus
Laonastes_aenigmamus
Lepus_europaeus
Myocastor_coypus
Octodon_degus
Oryctolagus_cuniculus
Paralomys_gerbillus
Pedetes_capensis
Petaurista_petaurista
Procavia_capensis
Rattus_norvegicus
Sciurus_carolinensis
Thomomys_umbrinus
Vombatus_ursinus
[rawpoints]
'Acomys_cahirinus
5.648530483245850e+000 2.348586082458496e+000 1.072210025787354e+001
7.757369518280029e+000 1.122772312164307e+001 1.305764198303223e+001
8.207379341125488e+000 1.301984310150147e+001 1.358475494384766e+001
1.006982421875000e+001 2.086593055725098e+001 1.479658699035645e+001
1.232744598388672e+001 2.885864448547363e+001 1.279560089111328e+001
6.848373889923096e+000 2.814226388931274e+000 1.066898250579834e+001
1.026459217071533e+001 1.048984527587891e+001 1.202399444580078e+001
```

GMM

1. Homologous landmarking of the entire anatomy and setae (hair) origins

2. Conduct GMM analyses: Procrustes, Principal Components, **covariation of anatomical regions**, and ultimately the **differentiation/identification of genus/species**

3. Identify unique morphological variability and ontogenetic trends across **metamorphosis** (larval, pupal, and adult stages). "**Geometric** (meta)**morpho**(sis)**metrics**"



H: **pupa < L3 < L4 < adult**, with respect to morphological variation

GMM + AI

Combining machine learning algorithms and geometric morphometrics: A study of carnivore tooth marks

Lloyd A. Courtenay[a,b,c,*], José Yravedra[c,d], Rosa Huguet[b,a,e], Julia Aramendi[c,f], Miguel Ángel Maté-González[g], Diego González-Aguilera[g], Mari Carmen Arriaza[h,i]

ABSTRACT

Since the 1980s an intense scientific debate has revolved around the hunting capacities of early hominin populations and the behavioral patterns of carnivores sharing the same ecosystem, and thus competing for the same resources. This debate, commonly known as the hunter-scavenger debate, fostered the emergence of a new research line into the Bone Surface Modifications (BSMs) produced by both taphonomic agents. Throughout the following 20 years, multiple studies concerning the action of carnivores have been developed, with a particular focus on the oldest archaeological sites in East Africa. Recent technological advances applied to taphonomy have provided new insight into carnivore BSMs. A newly developed part of this work relies on Geometric Morphometrics (GMM) studies aimed at discerning carnivore agency through the morphologic characterization of tooth scores and pits. GMM studies have produced promising results, however methodological limitations are still present. This paper presents the first combined application of Machine Learning (ML) algorithms and GMM to the analysis of carnivore tooth marks, generating classification rates of 100% between carnivore species in some cases.

"PC scores produced through PCA were used for ML analysis"

Inferring locomotor behaviours in Miocene New World monkeys using finite element analysis, geometric morphometrics and machine-learning classification techniques applied to talar morphology

Thomas A. Püschel[1], Jordi Marcé-Nogué[2,3], Justin T. Gladman[4], René Bobe[5,6] and William I. Sellers[1]

The talus is one of the most commonly preserved post-cranial elements in the platyrrhine fossil record. Talar morphology can provide information about postural adaptations because it is the anatomical structure responsible for transmitting body mass forces from the leg to the foot. The aim of this study is to test whether the locomotor behaviour of fossil Miocene platyrrhines could be inferred from their talus morphology. The extant sample was classified into three different locomotor categories and then talar strength was compared using finite-element analysis. Geometric morphometrics were used to quantify talar shape and to assess its association with biomechanical strength. Finally, several machine-learning (ML) algorithms were trained using both the biomechanical and morphometric data from the extant taxa to infer the possible locomotor behaviour of the Miocene fossil sample. The obtained results show that the different locomotor categories are distinguishable using either biomechanical or morphometric data. The ML algorithms categorized most of the fossil sample as arboreal quadrupeds. This study has shown that a combined approach can contribute to the understanding of platyrrhine talar morphology and its relationship with locomotion. This approach is likely to be beneficial for determining the locomotor habits in other fossil taxa.

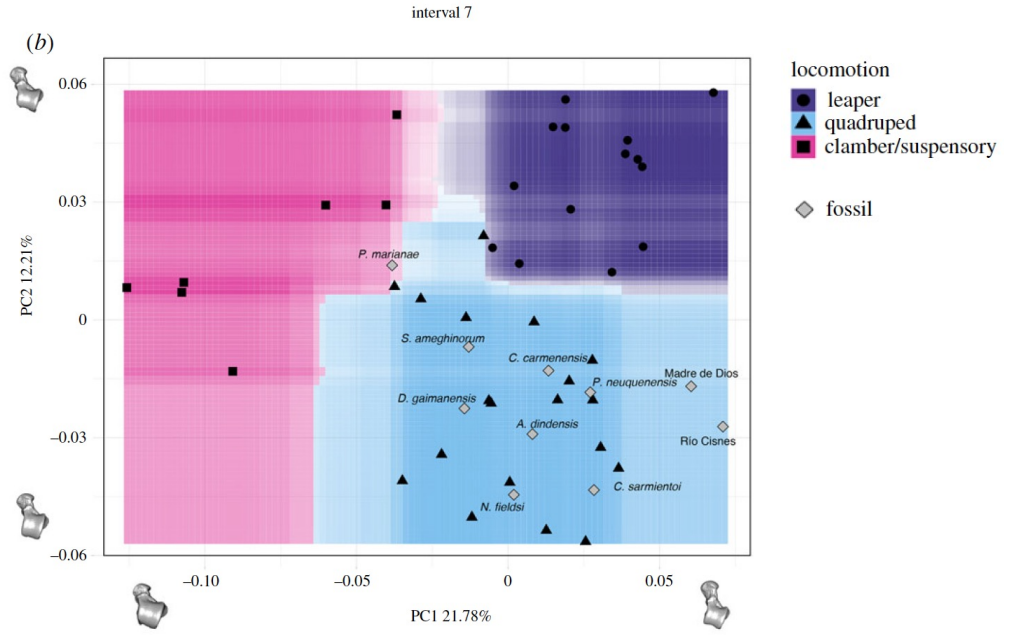
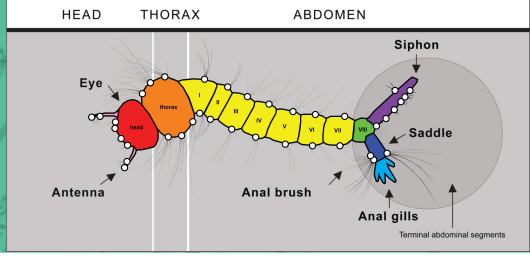


Figure 8. Decision boundary plots for (*a*) biomechanical and (*b*) morphometric data. In (*a*), only the seventh and 10th intervals are displayed because they contribute the most to class separation, while in (*b*) only the first two PCs are shown. The space is coloured depending on what locomotor category the (*a*) SVM or the (*b*) RF algorithm predict that region belongs to, whereas the lines between coloured areas represent the decision boundaries. Colour intensity indicates the certainty of the prediction in a particular graph area (i.e. darker colours imply a higher probability of belonging to a particular class). Symbols surrounded by a white rim represent misclassified specimens. In (*b*), one of the models closest to the mean shape was warped to match the multivariate mean using the thin-plate spline method, and then the obtained average model was warped to represent the variation along the two PC axes. (Online version in colour).

GMM + AI



1. **Explainable AI**. Following AI identification of the most critical bounding box and/or pixels (class activation map), conduct GMM analysis on the critical box/anatomical component in order to explain the anatomical findings of the AI box/class activation map results

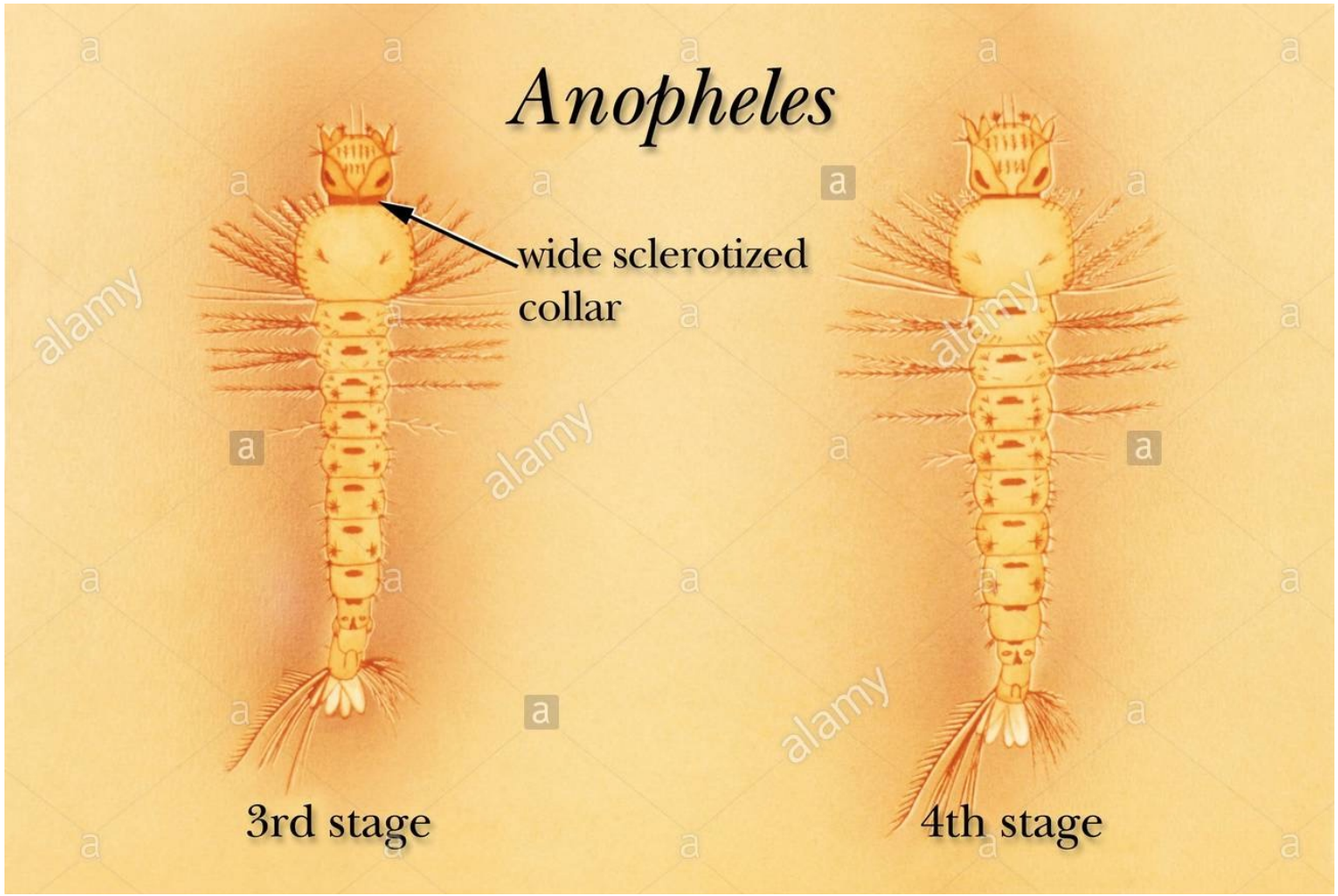
2. **GMM+AI**. Following placement of GMM landmarks (in the critical box/component or entire insect), develop an AI algorithm to analyze these landmarks in order to identify genus/species [AI on landmarks]

3. **AI+GMM**. Develop an AI algorithm to define, subselect, and/or optimize the locations of landmarks, for subsequent GMM-based identification of genus/species

4. Once the locations of landmarks are determined, develop an image recognition algorithm to **automate the placement** of these landmarks

5. **Comparison** and documentation of classification accuracies, and present use case scenarios in practice:

human eye vs **AI** vs **GMM** vs **GMM+AI** (*landmarks*) vs **GMM+AI** (*PCs*) vs **AI+GMM**



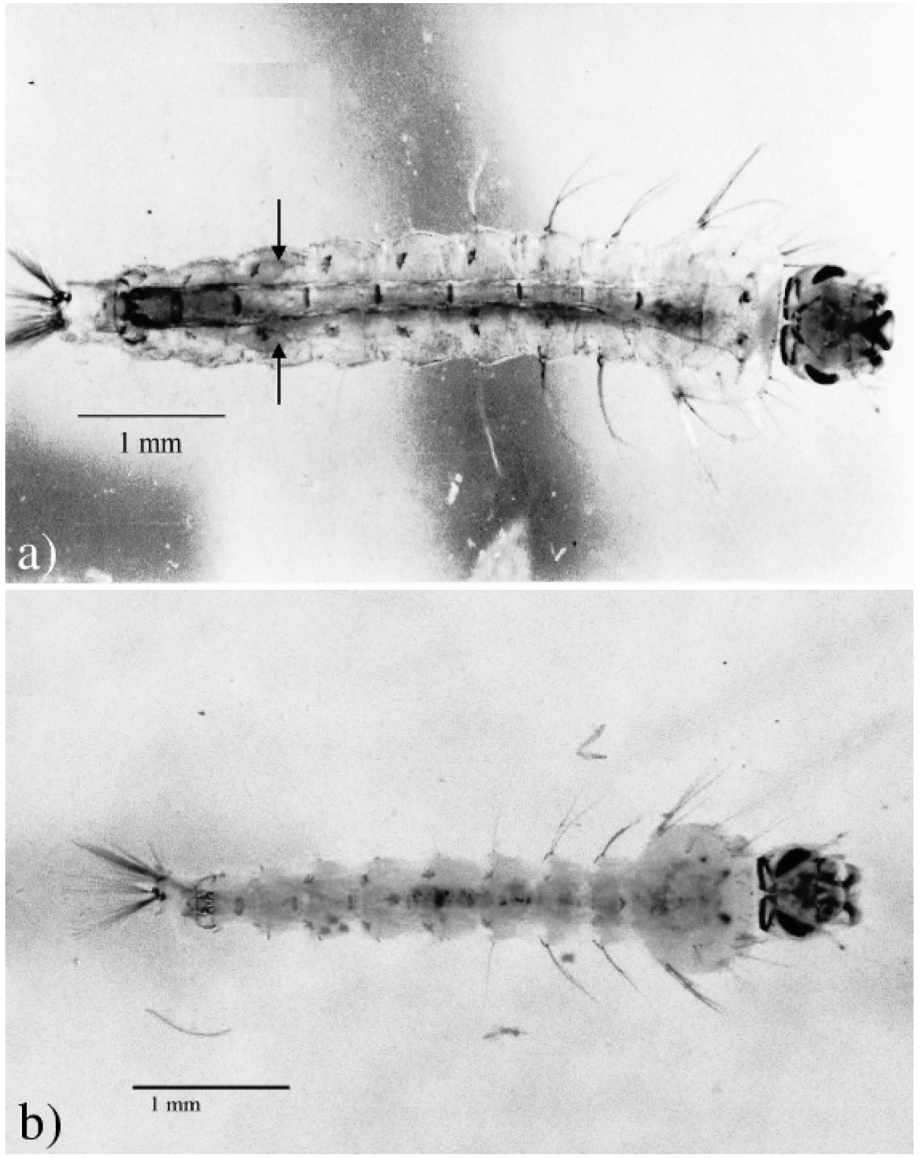


Fig. 1. Photomicrograph of *Anopheles stephensi* 4th instar, (a) male larva with abdominal spots indicated with arrows, (b) female larva without abdominal spots.

Suman et al 2008