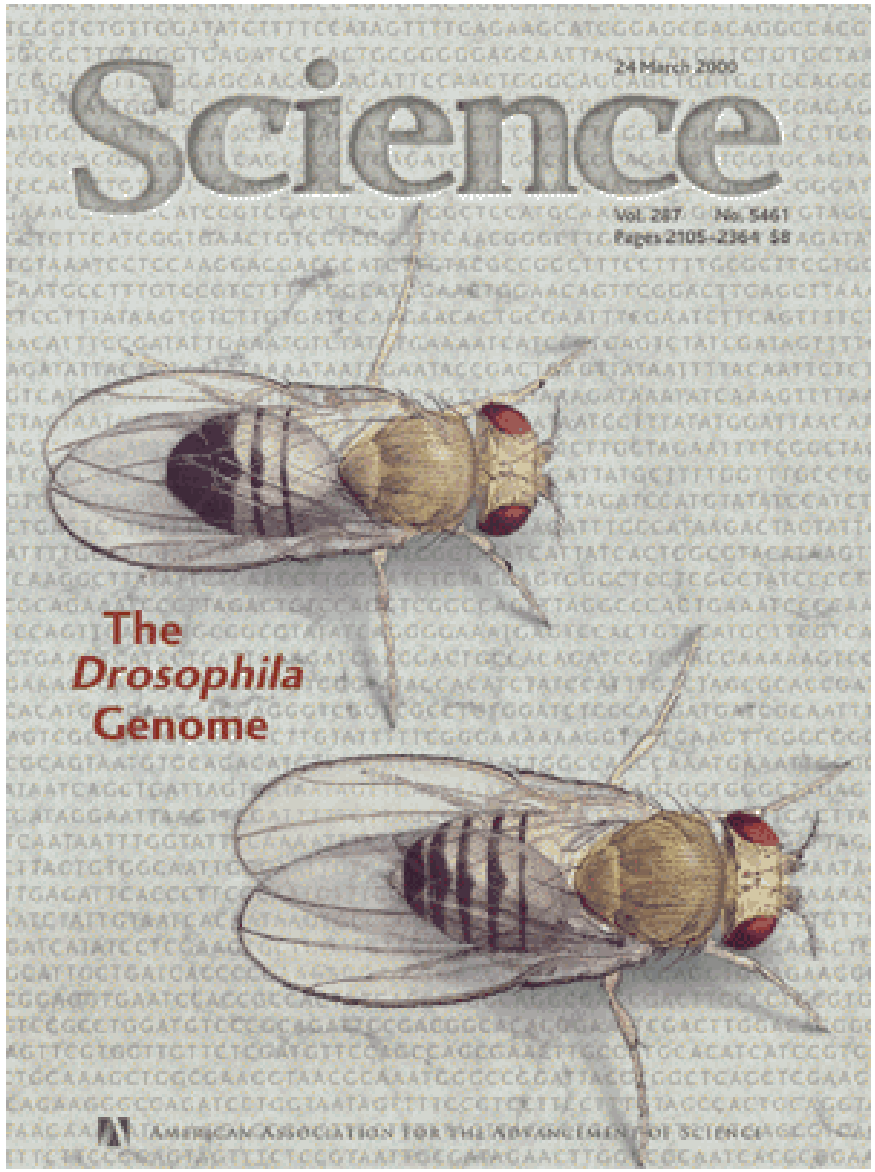
The place of *Drosophila* among the metazoans 2

Peter Holland

Zoology, Oxford

24 March 2000



24 March 2000

The Genome Sequence of *Drosophila melanogaster*

Mark D. Adams,[1*] Susan E. Celniker,[2] Robert A. Holt,[1] Cheryl A. Evans,[1] Jeannine D. Gocayne,[1] Peter G. Amanatides,[1] Steven E. Scherer,[3] Peter W. Li,[1] Roger A. Hoskins,[2] Richard F. Galle,[2] Reed A. George,[2] Suzanna E. Lewis,[4] Stephen Richards,[2] Michael Ashburner,[5] Scott N. Henderson,[1] Granger G. Sutton,[1] Jennifer R. Wortman,[1] Mark D. Yandell,[1] Qing Zhang,[1] Lin X. Chen,[1] Rhonda C. Brandon,[1] Yu-Hui C. Rogers,[1] Robert G. Blazej,[2] Mark Champe,[2] Barret D. Pfeiffer,[2] Kenneth H. Wan,[2] Clare Doyle,[2] Evan G. Baxter,[2] Gregg Helt,[6] Catherine R. Nelson,[4] George L. Gabor Miklos,[7] Josep F. Abril,[8] Anna Agbayani,[2] Hui-Jin An,[1] Cynthia Andrews-Pfannkoch,[1] Danita Baldwin,[1] Richard M. Ballew,[1] Anand Basu,[1] James Baxendale,[1] Leyla Bayraktaroglu,[9] Ellen M. Beasley,[1] Karen Y. Beeson,[1] P. V. Benos,[10] Benjamin P. Berman,[2] Deepali Bhandari,[1] Slava Bolshakov,[11] Dana Borkova,[12] Michael R. Botchan,[13] John Bouck,[3] Peter Brokstein,[4] Phillipe Brottier,[14] Kenneth C. Burtis,[15] Dana A. Busam,[1] Heather Butler,[16] Edouard Cadieu,[17] Angela Center,[1] Ishwar Chandra,[1] J. Michael Cherry,[18] Simon Cawley,[19] Carl Dahlke,[1] Lionel B. Davenport,[1] Peter Davies,[1] Beatriz de Pablos,[20] Arthur Delcher,[1] Zuoming Deng,[1] Anne Deslattes Mays,[1] Ian Dew,[1] Suzanne M. Dietz,[1] Kristina Dodson,[1] Lisa E. Doup,[1] Michael Downes,[21] Shannon Dugan-Rocha,[3] Boris C. Dunkov,[22] Patrick Dunn,[1] Kenneth J. Durbin,[3] Carlos C. Evangelista,[1] Concepcion Ferraz,[23] Steven Ferriera,[1] Wolfgang Fleischmann,[5] Carl Fosler,[1] Andrei E. Gabrielian,[1] Neha S. Garg,[1] William M. Gelbart,[9] Ken Glasser,[1] Anna Glodek,[1] Fangcheng Gong,[1] J. Harley Gorrell,[3] Zhiping Gu,[1] Ping Guan,[1] Michael Harris,[1] Nomi L. Harris,[2] Damon Harvey,[4] Thomas J. Heiman,[1] Judith R. Hernandez,[3] Jarrett Houck,[1] Damon Hostin,[1] Kathryn A. Houston,[2] Timothy J. Howland,[1] Ming-Hui Wei,[1] Chinyere Ibegwam,[1] Mena Jalali,[1] Francis Kalush,[1] Gary H. Karpen,[21] Zhaoxi Ke,[1] James A. Kennison,[24] Karen A. Ketchum,[1] Bruce E. Kimmel,[2] Chinnappa D. Kodira,[1] Cheryl Kraft,[1] Saul Kravitz,[1] David Kulp,[6] Zhongwu Lai,[1] Paul Lasko,[25] Yiding Lei,[1] Alexander A. Levitsky,[1] Jiayin Li,[1] Zhenya Li,[1] Yong Liang,[1] Xiaoying Lin,[26] Xiangjun Liu,[1] Bettina Mattei,[1] Tina C. McIntosh,[1] Michael P. McLeod,[3] Duncan McPherson,[1] Gennady Merkulov,[1] Natalia V. Milshina,[1] Clark Mobarry,[1] Joe Morris,[6] Ali Moshrefi,[2] Stephen M. Mount,[27] Mee Moy,[1] Brian Murphy,[1] Lee Murphy,[28] Donna M. Muzny,[3] David L. Nelson,[3] David R. Nelson,[29] Keith A. Nelson,[1] Katherine Nixon,[2] Deborah R. Nusskern,[1] Joanne M. Pacleb,[2] Michael Palazzolo,[2] Gjange S. Pittman,[1] Sue Pan,[1] John Pollard,[1] Vinita Puri,[1] Martin G. Reese,[4] Knut Reinert,[1] Karin Remington,[1] Robert D. C. Saunders,[30] Frederick Scheeler,[1] Hua Shen,[3] Bixiang Christopher Shue,[1] Inga Sidén-Kiamos,[11] Michael Simpson,[1] Marian P. Skupski,[1] Tom Smith,[1] Eugene Spier,[1] Allan C. Spradling,[31] Mark Stapleton,[2] Renee Strong,[1] Eric Sun,[1] Robert Svirskas,[32] Cyndee Tector,[1] Russell Turner,[1] Eli Venter,[1] Aihui H. Wang,[1] Xin Wang,[1] Zhen-Yuan Wang,[1] David A. Wassarman,[33] George M. Weinstock,[3] Jean Weissenbach,[14] Sherita M. Williams,[1] Trevor Woodage,[1] Kim C. Worley,[3] David Wu,[1] Song Yang,[2] Q. Alison Yao,[1] Jane Ye,[1] Ru-Fang Yeh,[19] Jayshree S. Zaveri,[1] Ming Zhan,[1] Guangren Zhang,[1] Qi Zhao,[1] Liansheng Zheng,[1] Xiangqun H. Zheng,[1] Fei N. Zhong,[1] Wenyan Zhong,[1] Xiaojun Zhou,[3] Shiaoping Zhu,[1] Xiaohong Zhu,[1] Hamilton O. Smith,[1] Richard A. Gibbs,[3] Eugene W. Myers,[1] Gerald M. Rubin,[34] J. Craig Venter[1]

Science 287, 2185 - 2195

Metazoan genomes published (up to 2006)

Pictures from here removed for copyright reasons.

Human Chimp Mouse Rat Dog Chicken *Takifugu* *Tetraodon*

Anopheles *Apis*

Ciona intestinalis Oikopleura dioica

Drosophila melanogaster & pseudoobscura

Caenorhabditis elegans & briggsae

Other essentially ‘complete’ Metazoan genomes

Pictures from here removed for copyright reasons.

Macaque Cow Opposum Xenopus Zebrafish Stickleback

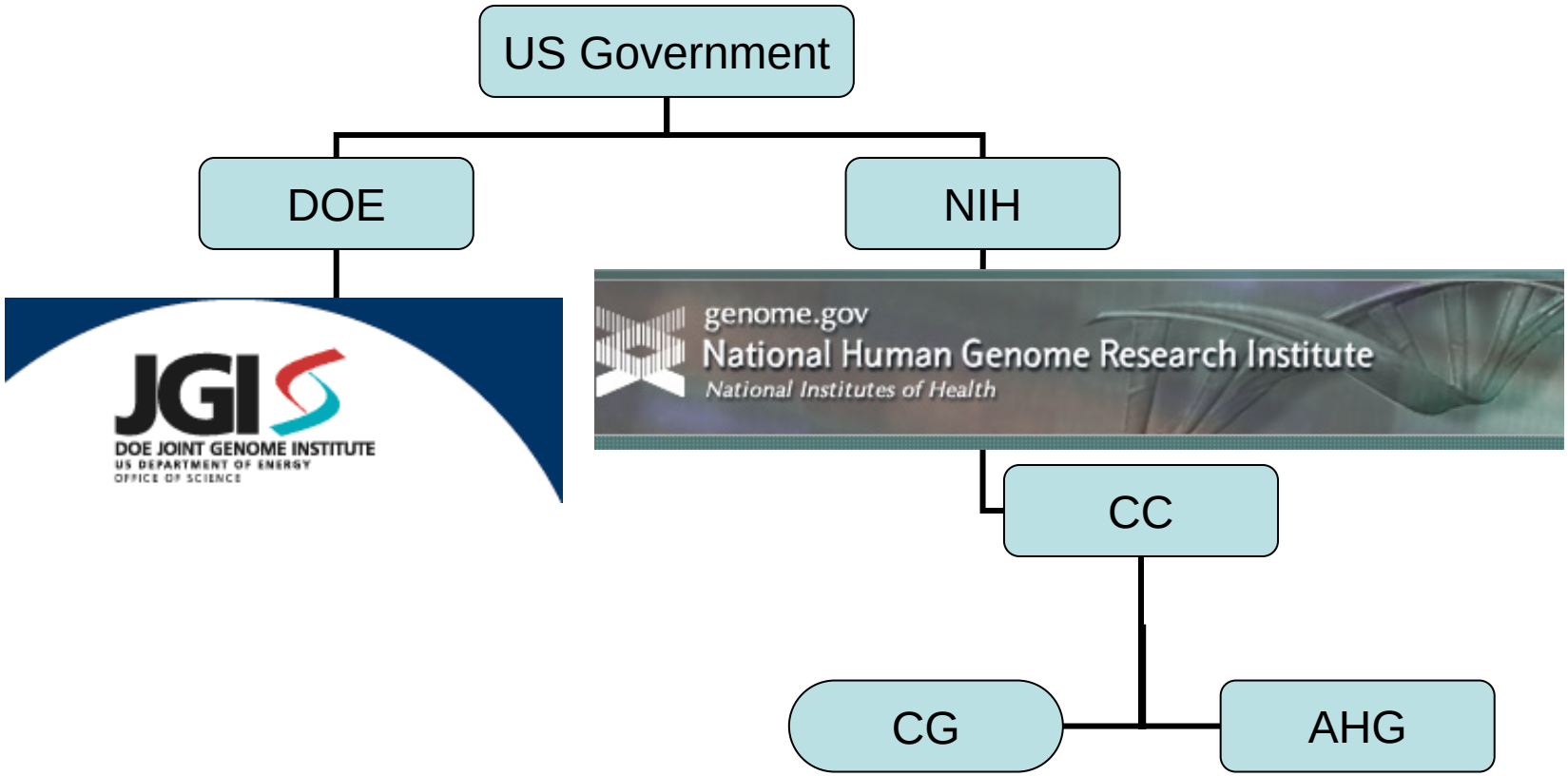
Sea urchin Schmidtea Lottia Amphioxus Ciona savignyi Medaka

Several nematodes,
notably Caenorhabditis remanei & Brugia

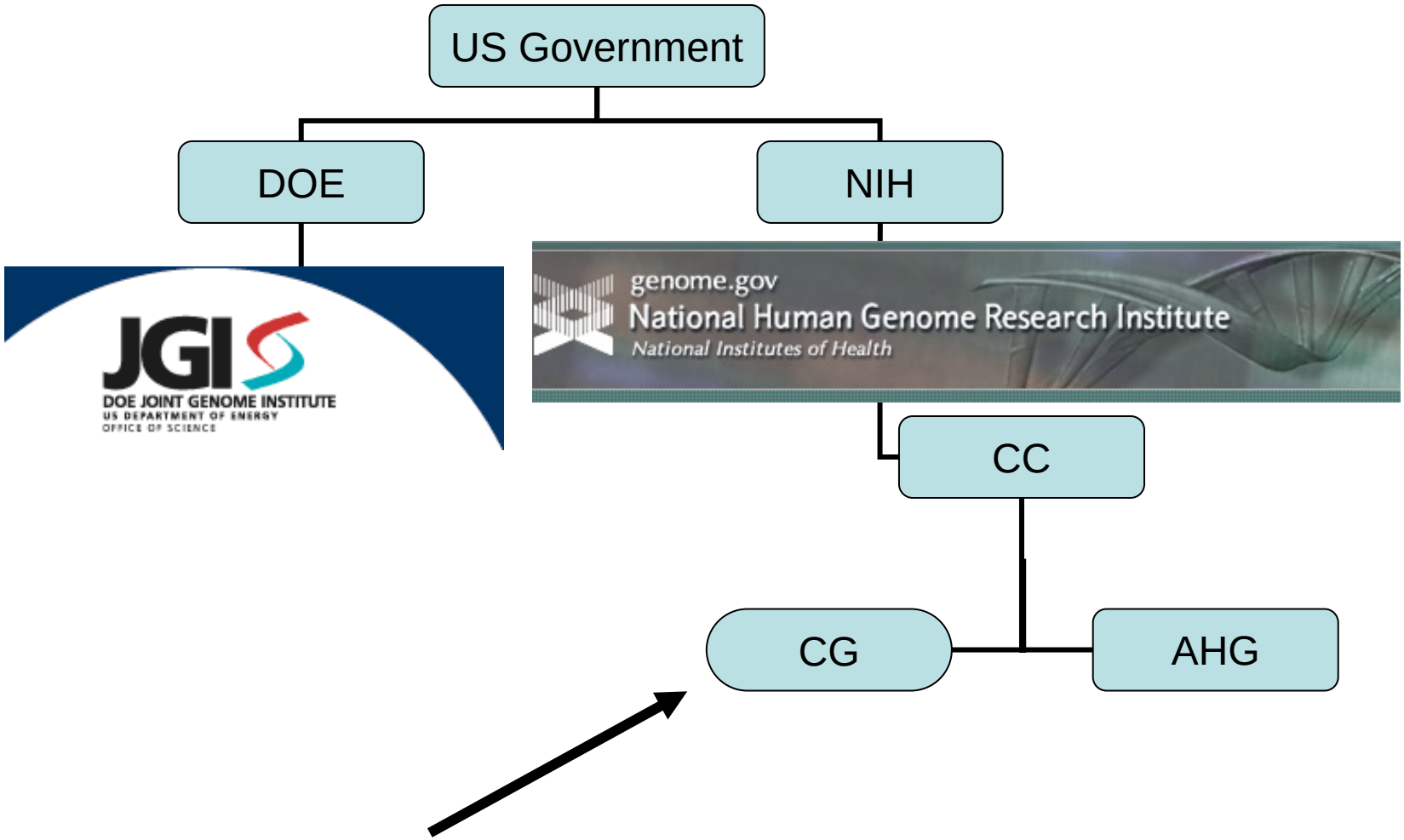
Nematostella

10 more Drosophila species Daphnia Tribolium Nasonia Sponge

Genome Sequencing



Genome Sequencing



Mammals – as of 2006

- **Placentals**
  - Human, chimp, mouse, rat, dog = published
  - Orang utan — NHGRI 7x in progress
  - Gibbon — NHGRI, approved ←
  - Macaque — draft assembly released
  - Marmaset — NHGRI 2x approved, not started
  - Bat — NHGRI, 7x approved ←
  - Cat — NHGRI 2x in progress, upgraded to 7x plan ←
  - Cow — draft assembly released
  - Elephant — NHGRI 2x approved, upgraded to 7x plan ←
  - Shrew — NHGRI 2x approved, not started
  - Hedgehog — NHGRI 2x approved, not started
  - Guinea pig — NHGRI 2x approved, upgraded to 7x plan ←
  - Tenrec — NHGRI 2x approved, not started
  - Armadillo — NHGRI 2x approved, upgraded to 7x plan ←
  - Rabbit — NHGRI 2x approved, upgraded to 7x plan ←
  - Tree shrew — NHGRI, 7x approved ←
- **Marsupials**
  - Opposum — draft assembly released
  - Wallaby — NHGRI 2x approved, not started
- **Monotremes**
  - Platypus — NHGRI draft assembly, started

Summer 2006 decisions

Other vertebrates

- **Birds**
  - Chicken – done 2005
  - Zebra finch – now approved, underway 2006
- **Reptiles**
  - Anolis lizard – NHGRI approved 2005
  - Freshwater turtle – NHGRI approved 2006 ←
- **Amphibians**
  - *Xenopus tropicalis* — done 2006
- **Basal sarcopterygians**
  - *Coelacanth* — approved 2006 ←
- **Ray-finned fish**
  - Three-spined stickleback — Almost done
  - Takifugu — Done
  - Tetraodon — Genoscope done 2005
  - Zebrafish — Sanger, almost done 2006
  - Medaka — almost done 2006
  - Gar — approved 2006 ←
- **Cartilagenous fish**
  - Skate *Raja erinacea* — NHGRI approved
  - Ratfish — NHGRI provisionally approved
- **Jawless vertebrates**
  - Sea lamprey — Approved, problems
  - Hagfish — Approved 2006 ←

Autumn 2006 decisions

Bilateria

Diploblasts

Ecdysozoans

Arthropods
Onycophorans
Nematomorphs
Nematodes
Kinorhynchs
Priapulids

Lophotrochozoans

Molluscs
Platyhelminths
Nemerteans
Echiurans
Annelids
Sipunculans
Brachiopods
Phoronids
Bryozoans

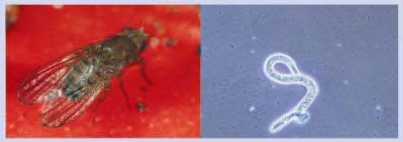
Deuterostomes

Echinoderms
Hemichordates
Cephalochordates
Tunicates
Vertebrates

Ctenophores
Cnidarians
Placozoa
Porifera

Bilateria

Diploblasts




Ecdysozoans

- Arthropods
- Onycophorans
- Nematomorphs
- Nematodes
- Kinorhynchs
- Priapulids

Lophotrochozoans

- Molluscs
- Platyhelminths
- Nemerteans
- Echiurans
- Annelids
- Sipunculans
- Brachiopods
- Phoronids
- Bryozoans

Deuterostomes

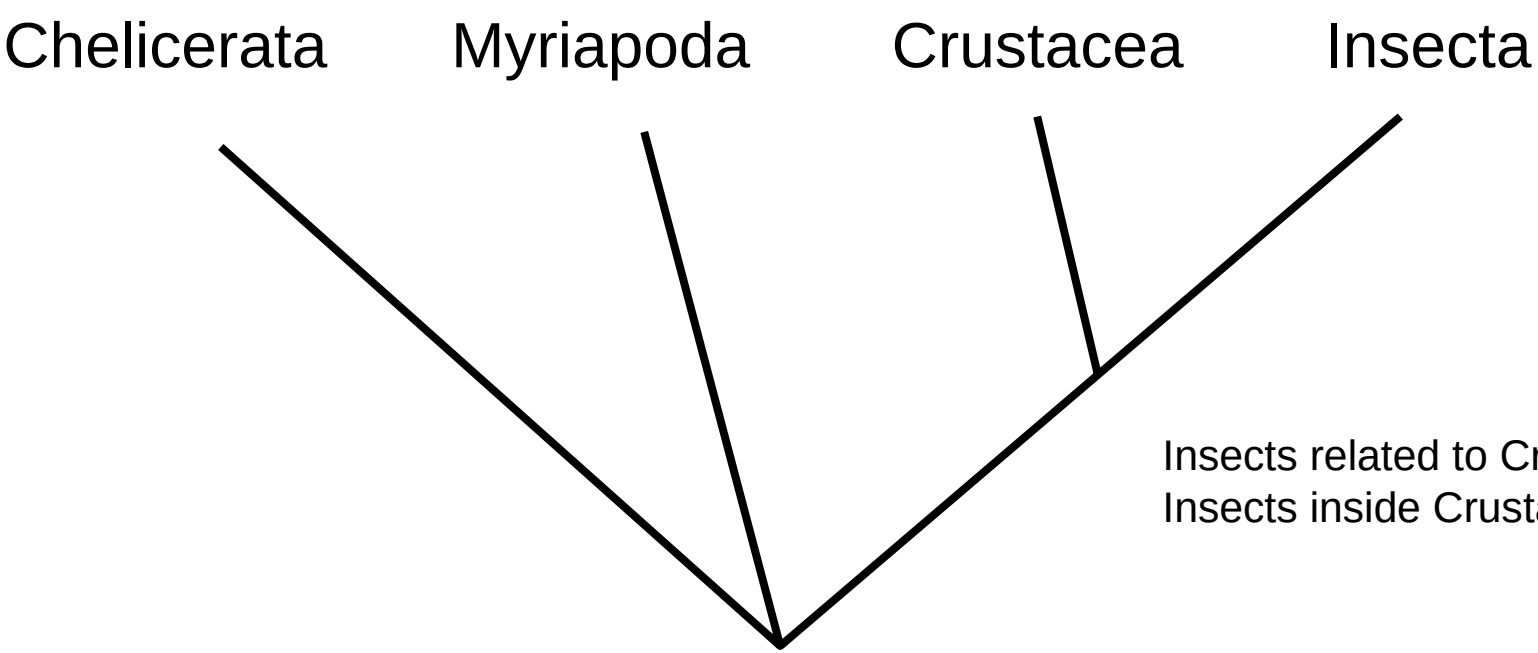
- Echinoderms
- Hemichordates
- Cephalochordates
- Tunicates
- Vertebrates

- Ctenophores
- Cnidarians
- Placozoa
- Porifera

Phylum Arthropoda

Pictures from here removed for copyright reasons

e.g. spider millipede shrimp, Daphnia fly

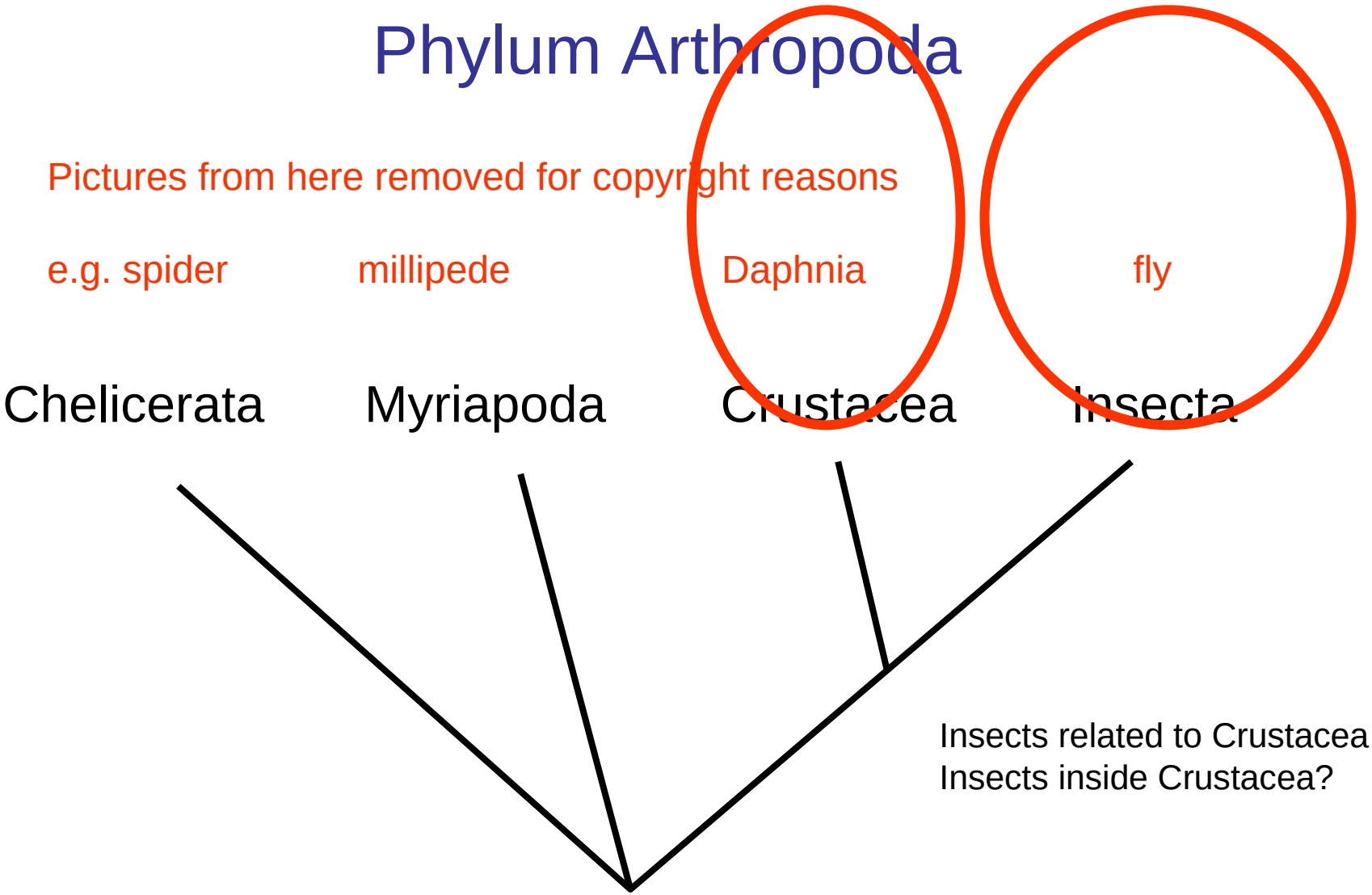


Revised (molecular) phylogeny

Phylum Arthropoda

Pictures from here removed for copyright reasons

e.g. spider millipede Daphnia fly

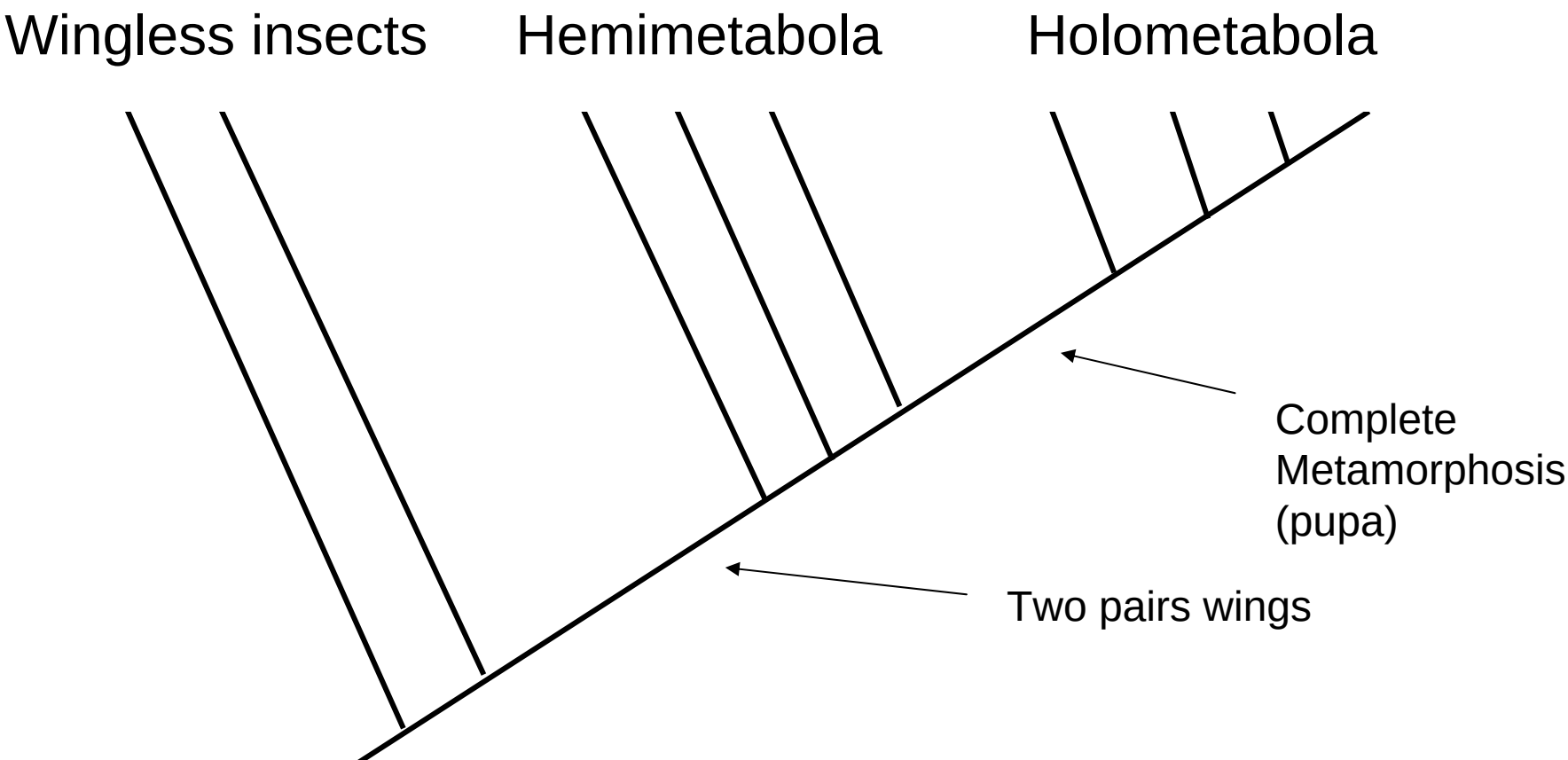


Revised (molecular) phylogeny

Class Insecta

Pictures from here removed for copyright reasons

e.g. bristletail pea aphid fly



Class Insecta

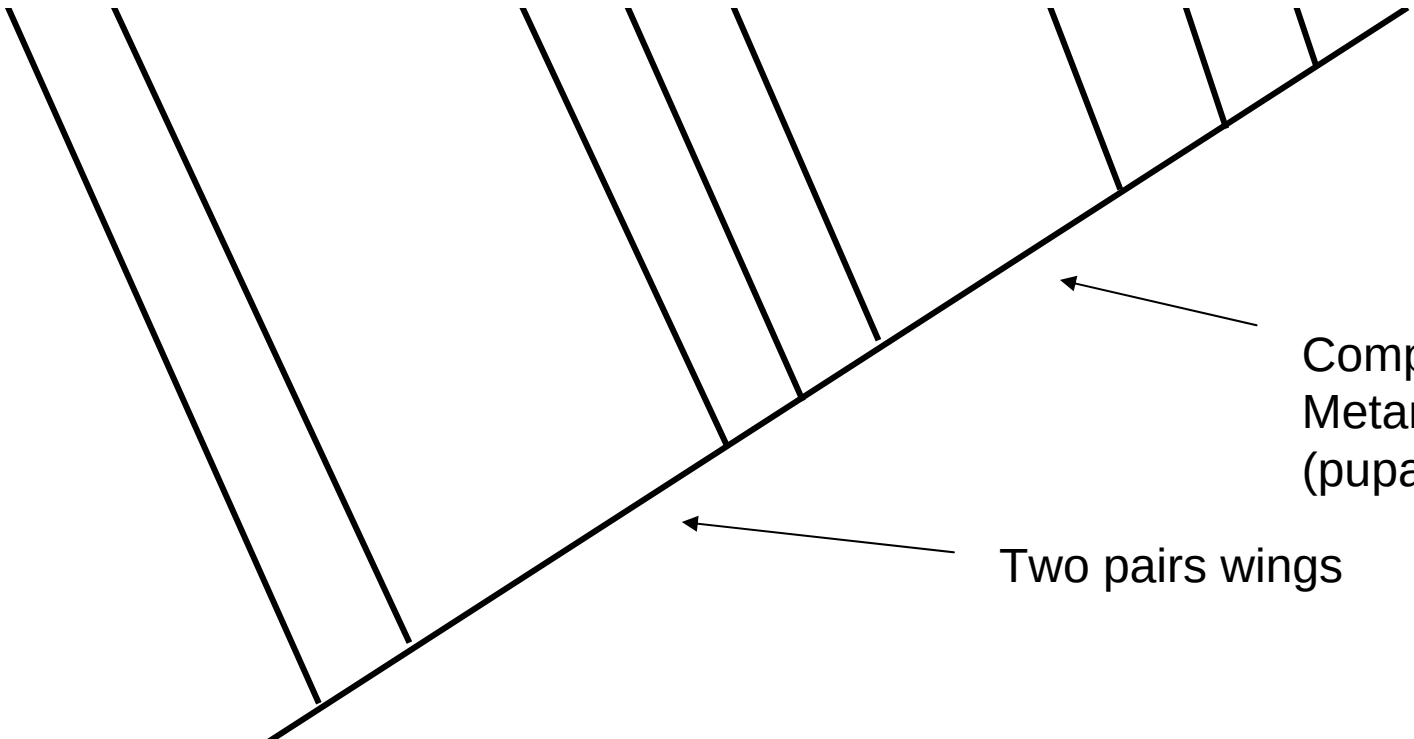
Pea aphid – in progress

Pictures from here removed for copyright reasons

e.g. bristletail

pea aphid

fly

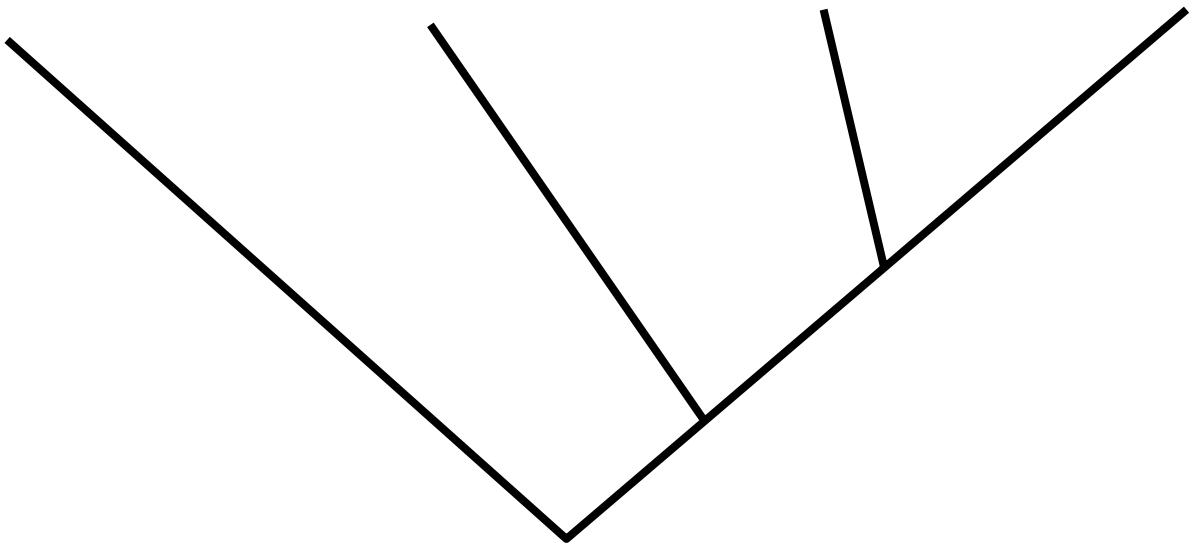


Holometabola

Pictures from here removed for copyright reasons

e.g. beetle honey bee butterfly fly

Coleoptera Hymenoptera Lepidoptera Diptera



Holometabola

Pictures from here removed for copyright reasons

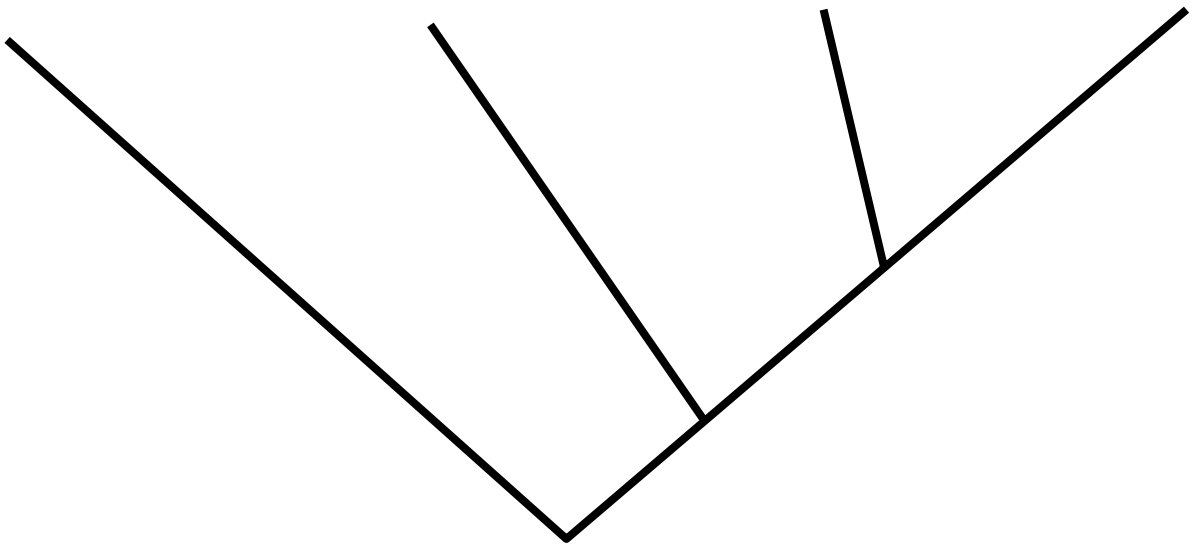
e.g. beetle

honey bee

butterfly

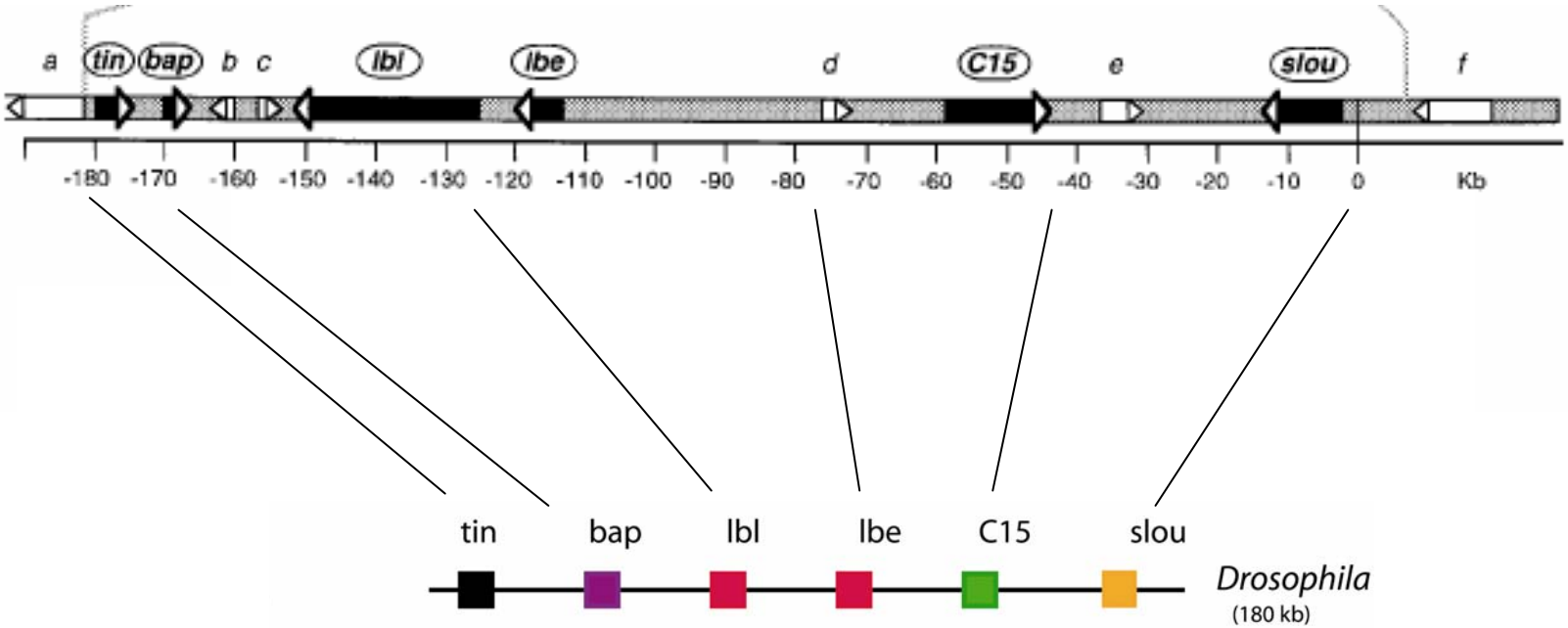
fly

Coleoptera Hymenoptera Lepidoptera Diptera

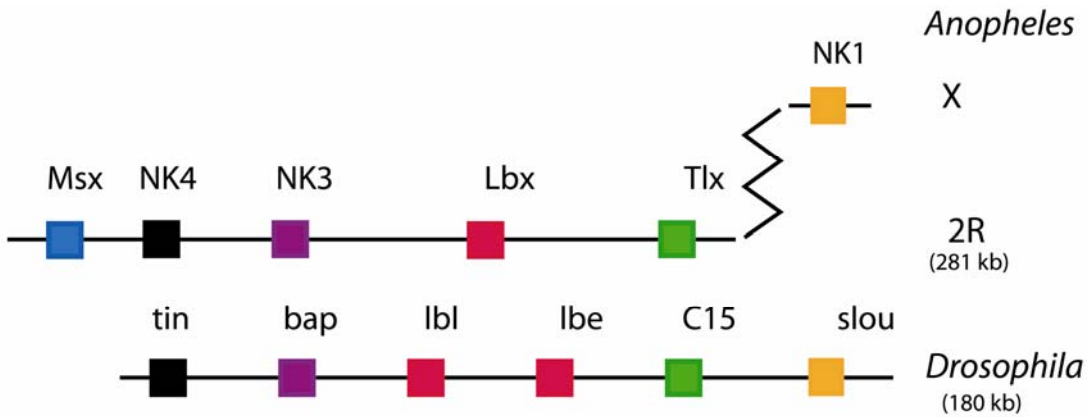


Now for an example of how comparative genomics can be used to understand a set of *Drosophila* genes

An unusual homeobox gene cluster in *Drosophila* : the NK-like genes



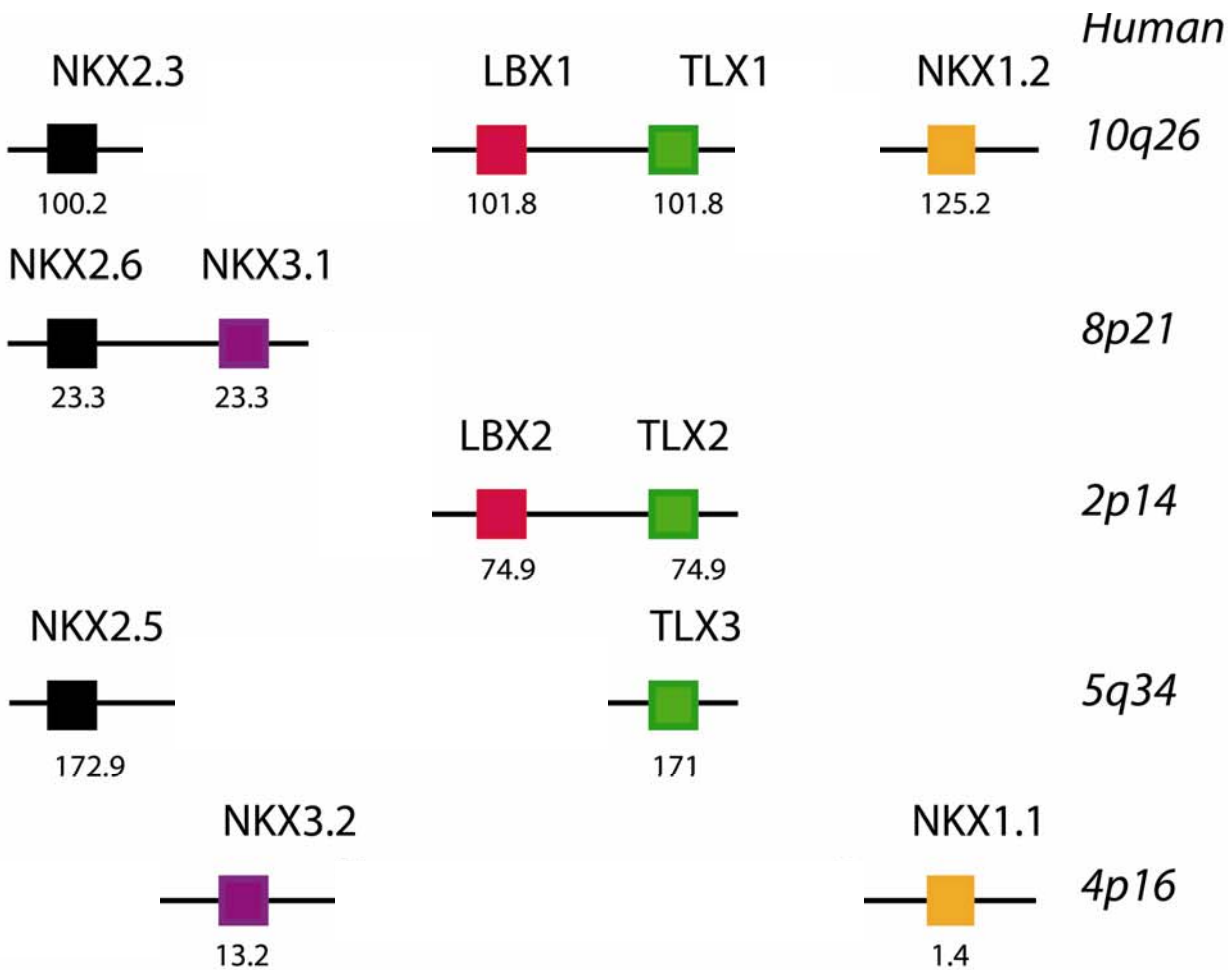
Gene name	Gene family	Expression/function
tinman	NK4	all mesoderm; later dorsal muscle (e.g. heart)
bagpipe	NK3	visceral mesoderm (e.g. around midgut)
ladybird	Lbx	heart cell fate; segmental border muscles
C15	Tlx	alary muscles
slouch	NK1	somatic muscle: differentiation of muscle fibres


Anopheles
NK1
X
Msx
NK4
NK3
Lbx
Tlx
2R
(281 kb)
tin
bap
lbl
lbe
C15
slou
Drosophila
(180 kb)

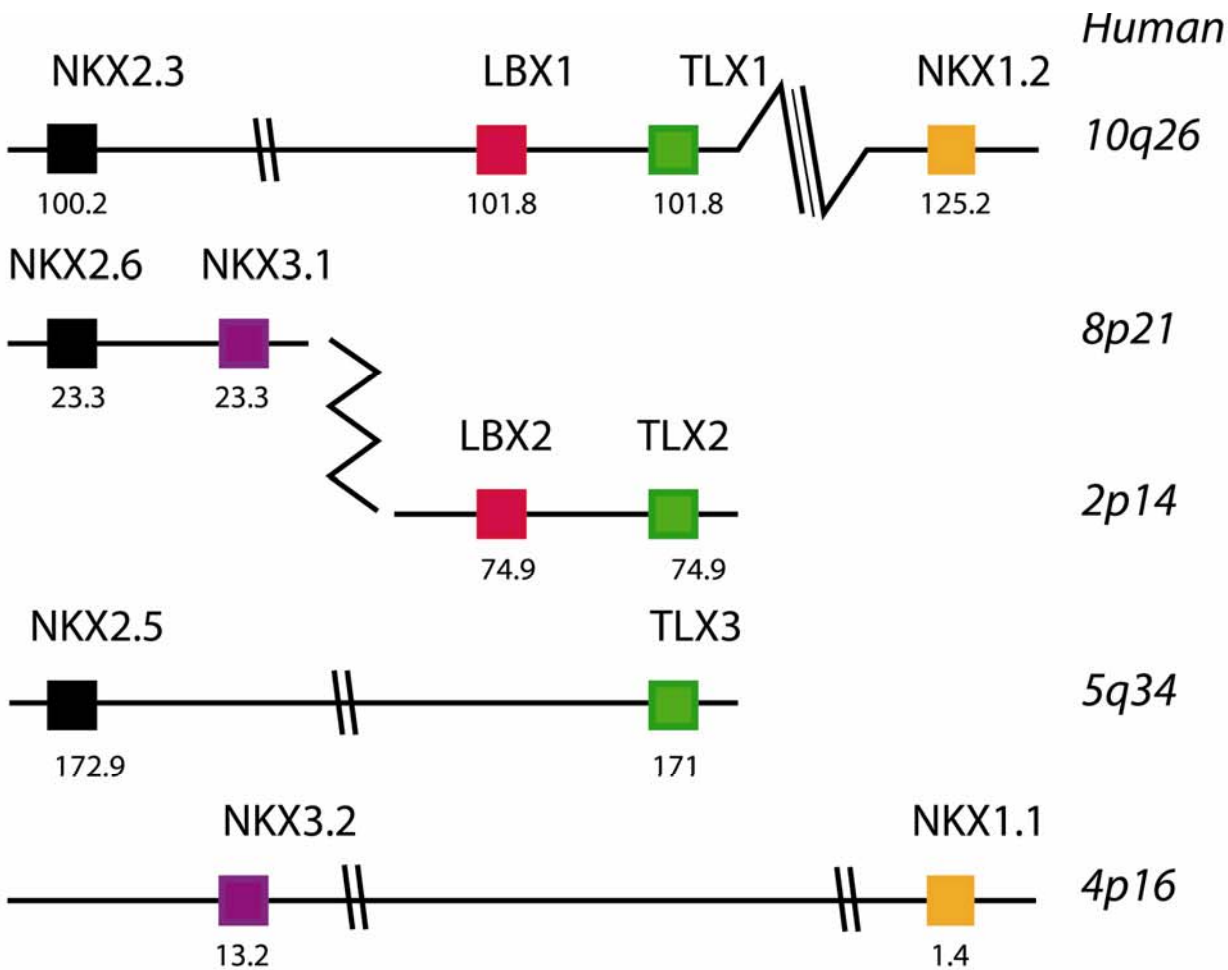
These fly genes have human orthologues;
therefore, it is an ancient gene cluster

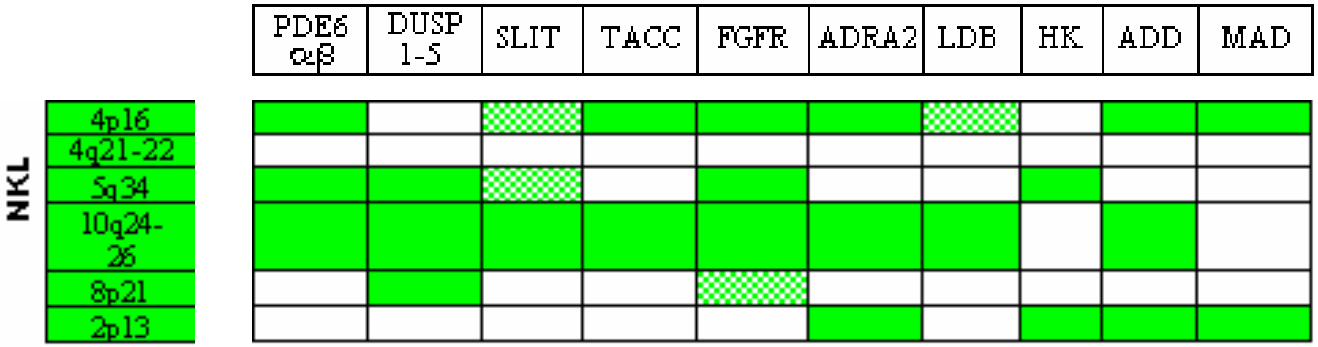
Drosophila homeobox		Human homeobox
NK4 (tinman)		~ NKX-2.5, -2.6, -2.3
NK3 (bagpipe)		NKX-3.1, -3.2
lbe & lbl		LBX-1, -2
C15 (93Bal)		TLX-1, -2, -3
slouch (S59)		NKX-1.1, -1.2

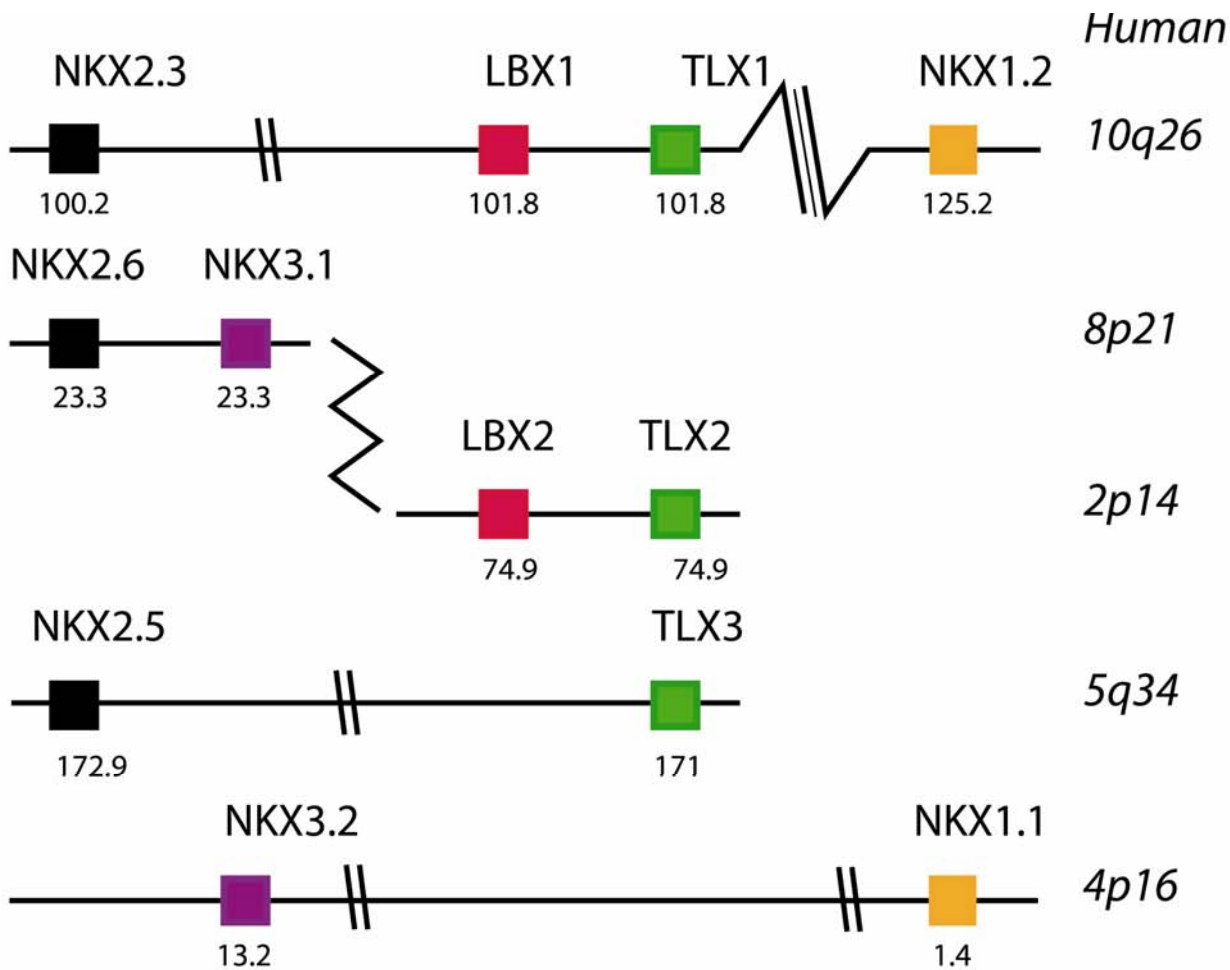
‘NK’ class homeobox genes are scattered on 9 locations on 5 human chromosomes



Maybe they are remnants of 4 gene clusters that have split and dispersed ?

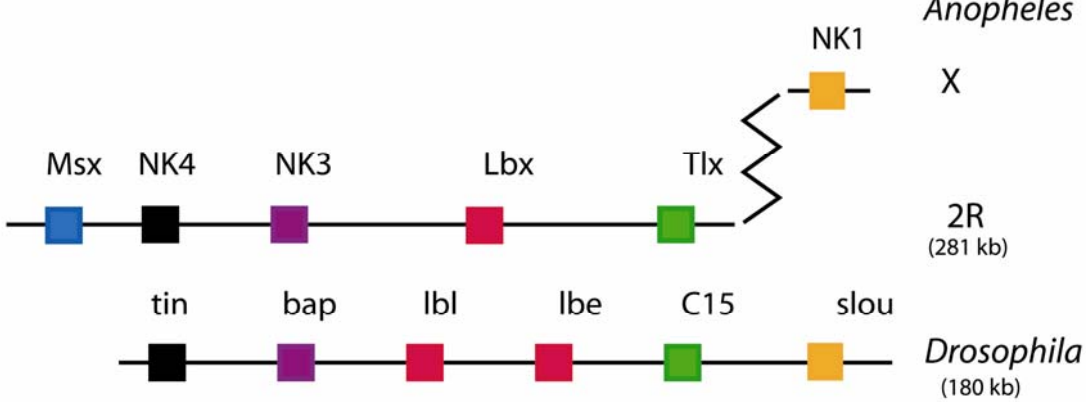


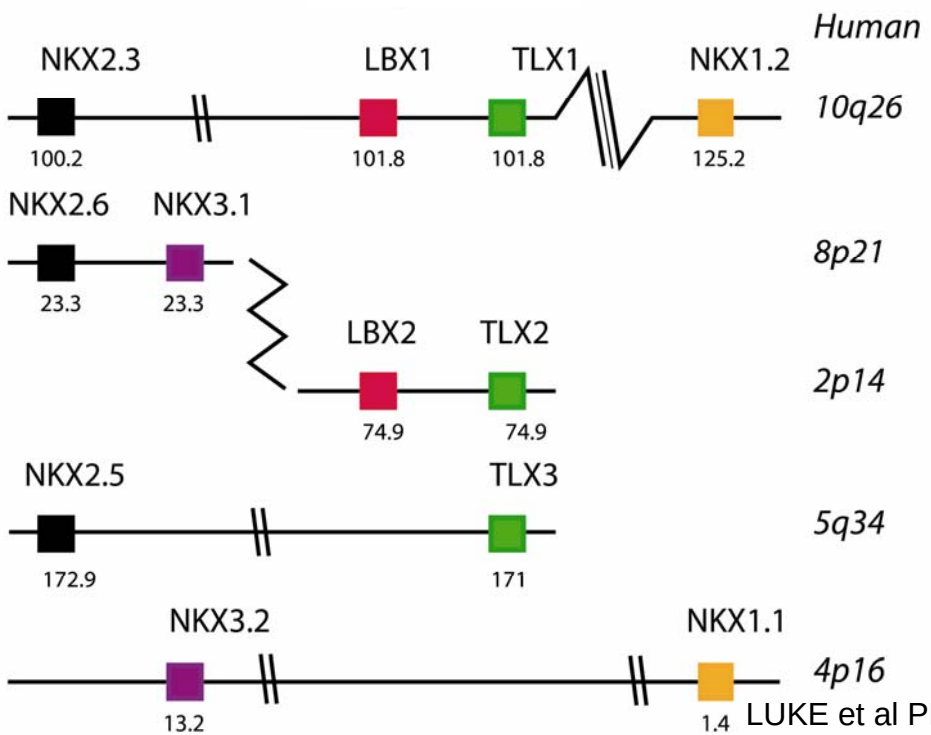




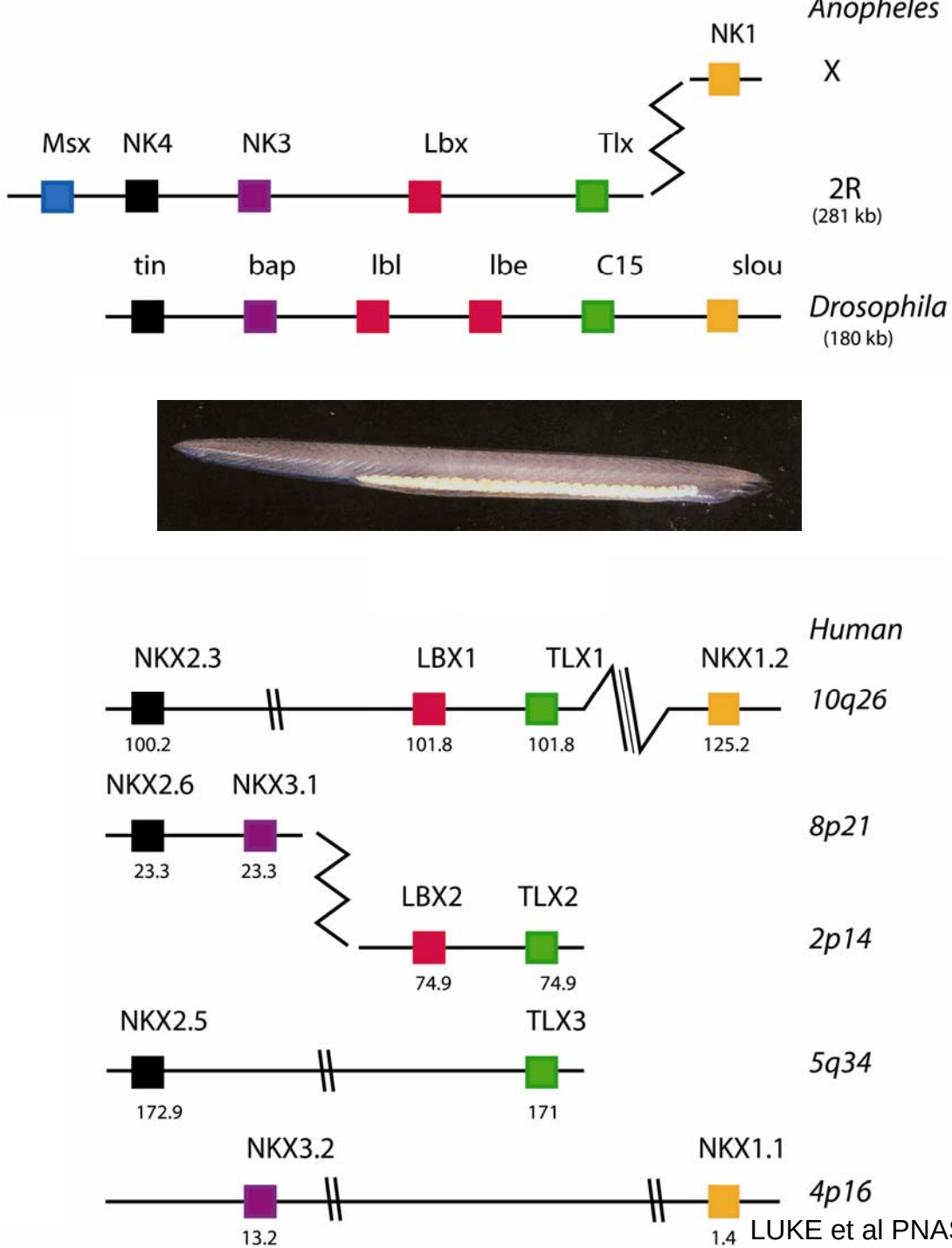
LUKE et al PNAS 100, 5292-5295

POLLARD & HOLLAND Current Biol 10: 1059-1062

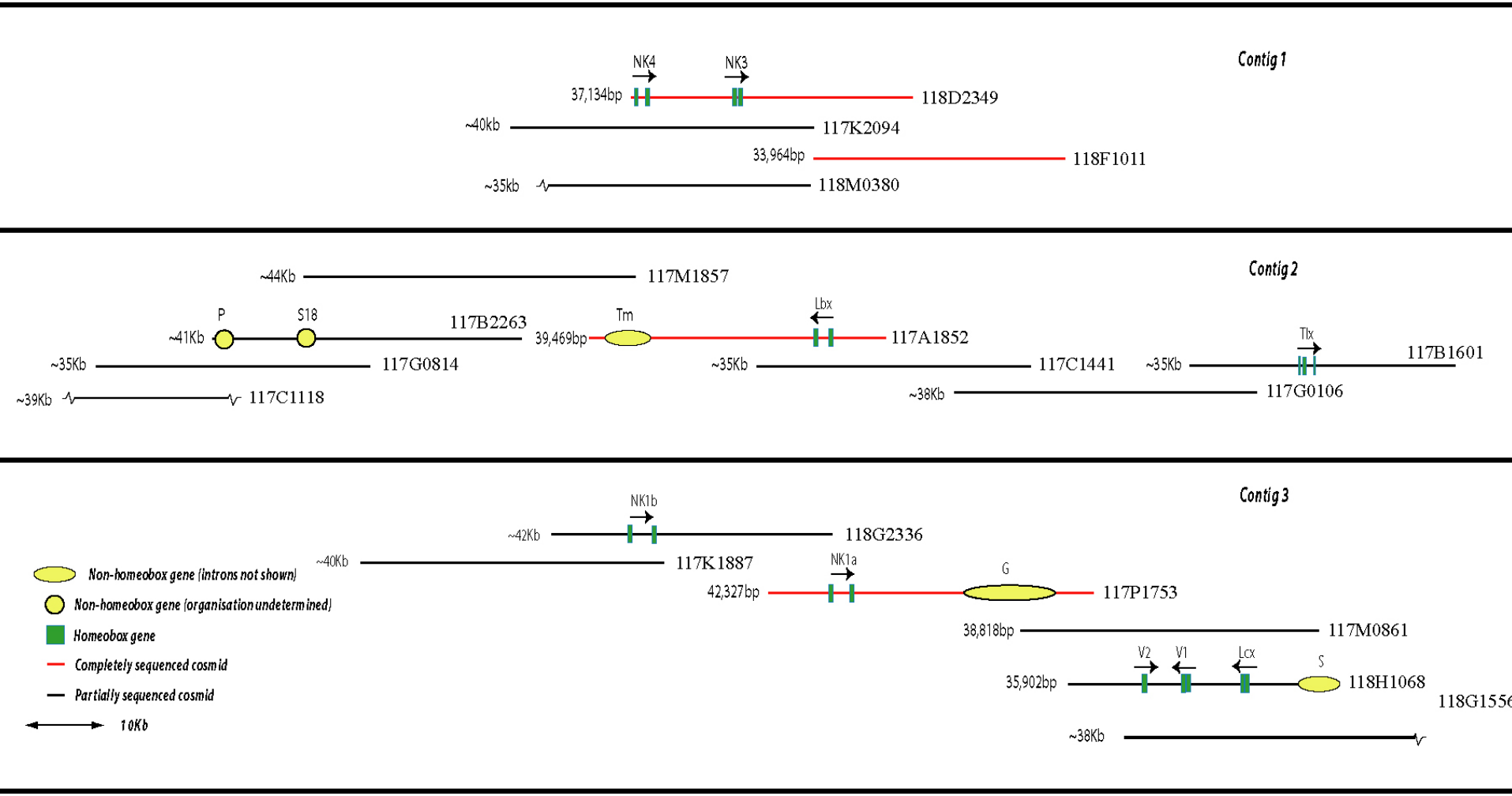




LUKE et al PNAS 100, 5292-5295



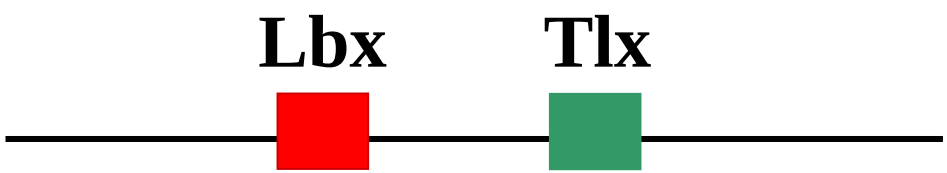
LUKE et al PNAS 100, 5292-5295



LUKE et al PNAS 100, 5292-5295

- Scaffold 1

- Scaffold 2

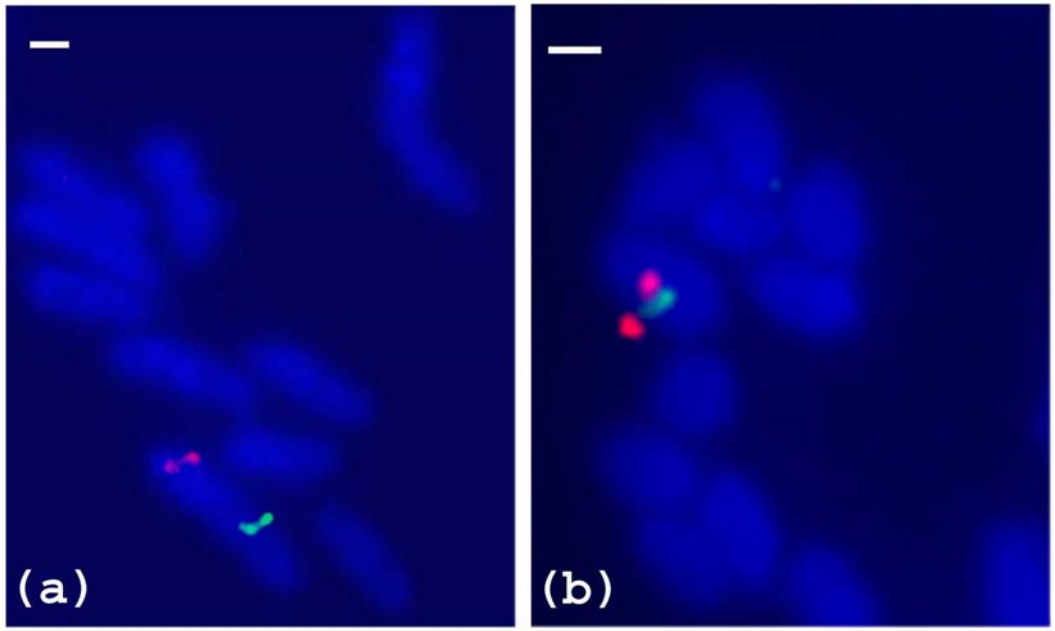


- Scaffold 3

DOUBLE PROBES

a. Contig 1, contig 2

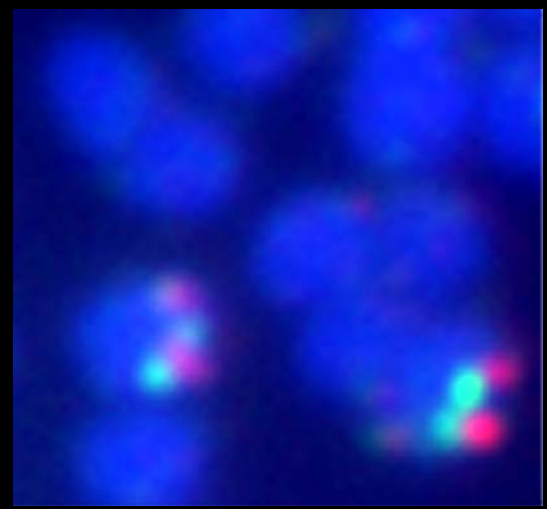
b. Contig 3, contig 2



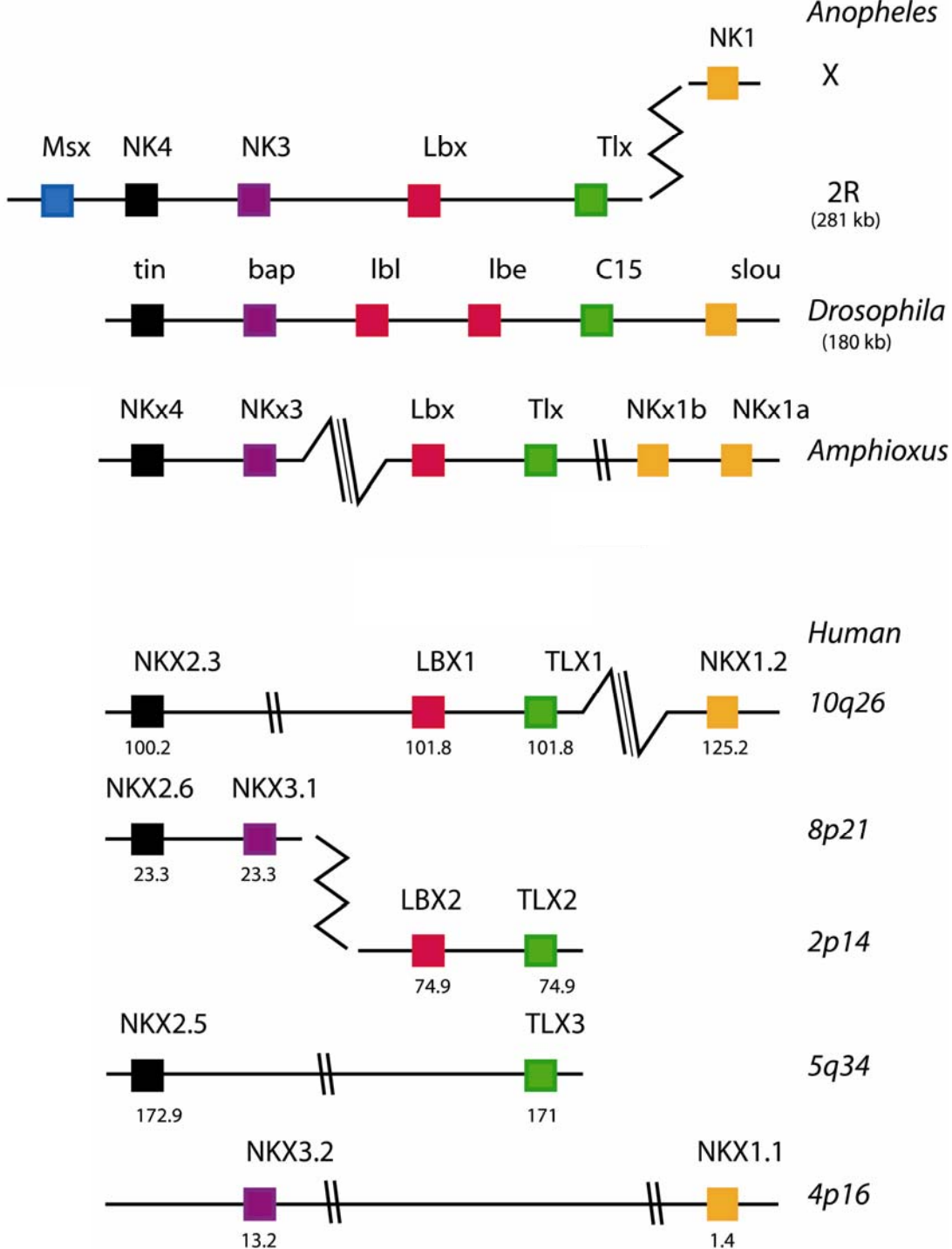
TRIPLE PROBE

Contig 1 (green) and contig 2 vs contig 3 (red)

Filipe CASTRO

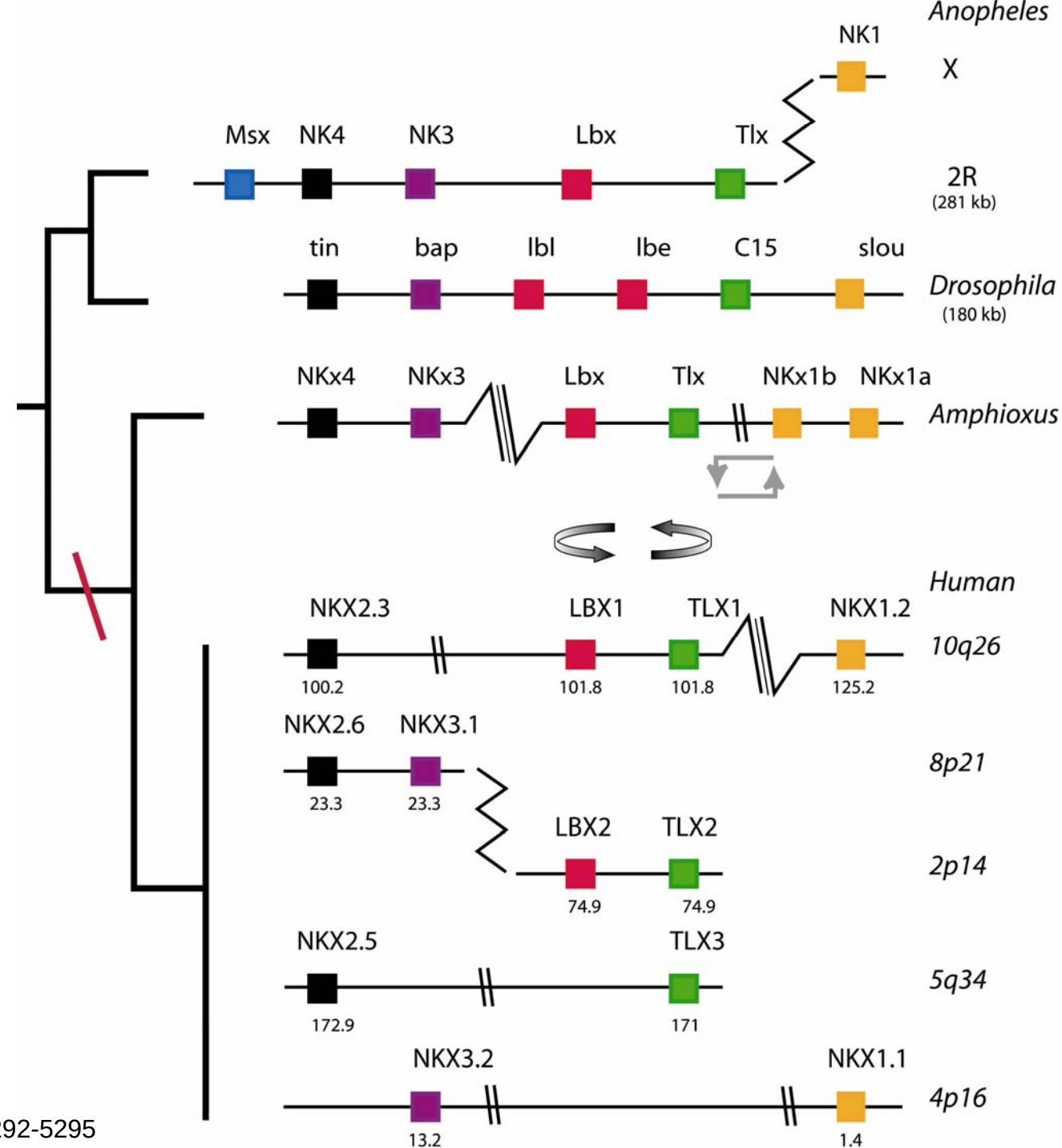


LUKE et al PNAS 100, 5292-5295



LUKE et al PNAS 100, 5292-5295

Ancient NK homeobox gene cluster



LUKE et al PNAS 100, 5292-5295

Conclusions

Dobzhansky *"Nothing in biology makes sense except in the light of evolution"*
Holland *"Nothing in genomics makes sense except in the light of evolution"*

- *Drosophila* is a metazoan (animal) – diversity of metazoan genomes
- *Drosophila* is a bilaterian – diversity of bilaterian genomes
- *Drosophila* is an ecdysozoan – insect and nematode genomes
- *Drosophila* is an arthropod – not very well sampled
- *Drosophila* is an insect – Holometabola well sampled, rest not

- Phylogeny plus genome sequences allows insight into genome evolution