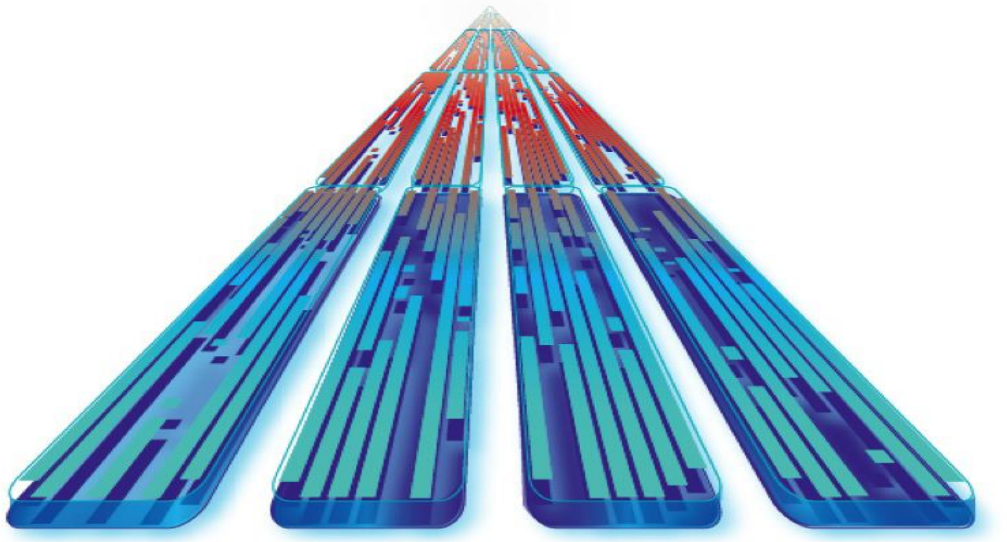
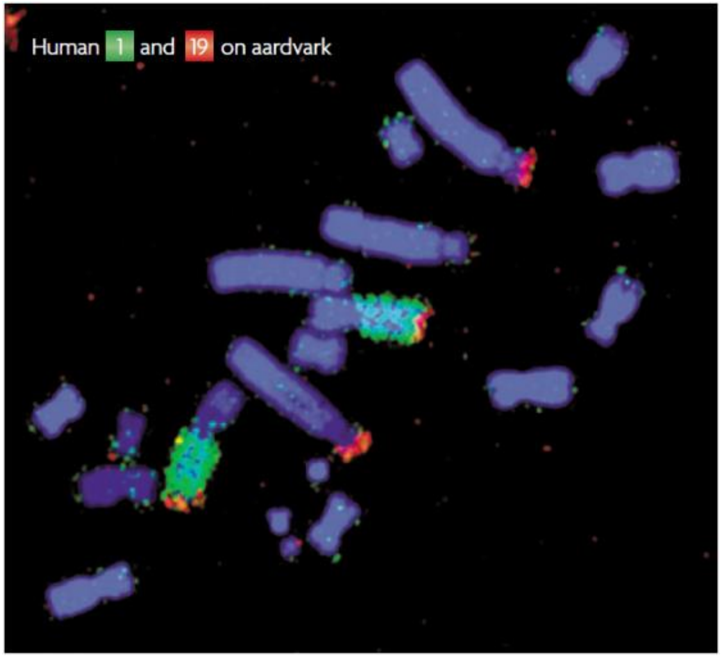


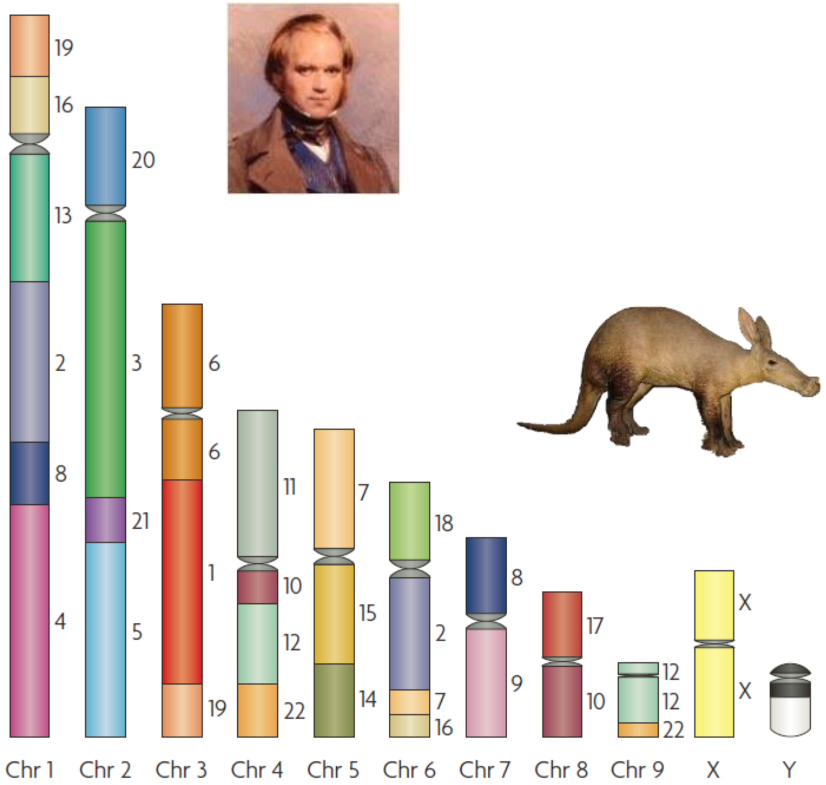
Сравнительная геномика того, чего никогда не было и что из этого получилось

Денис Ларкин

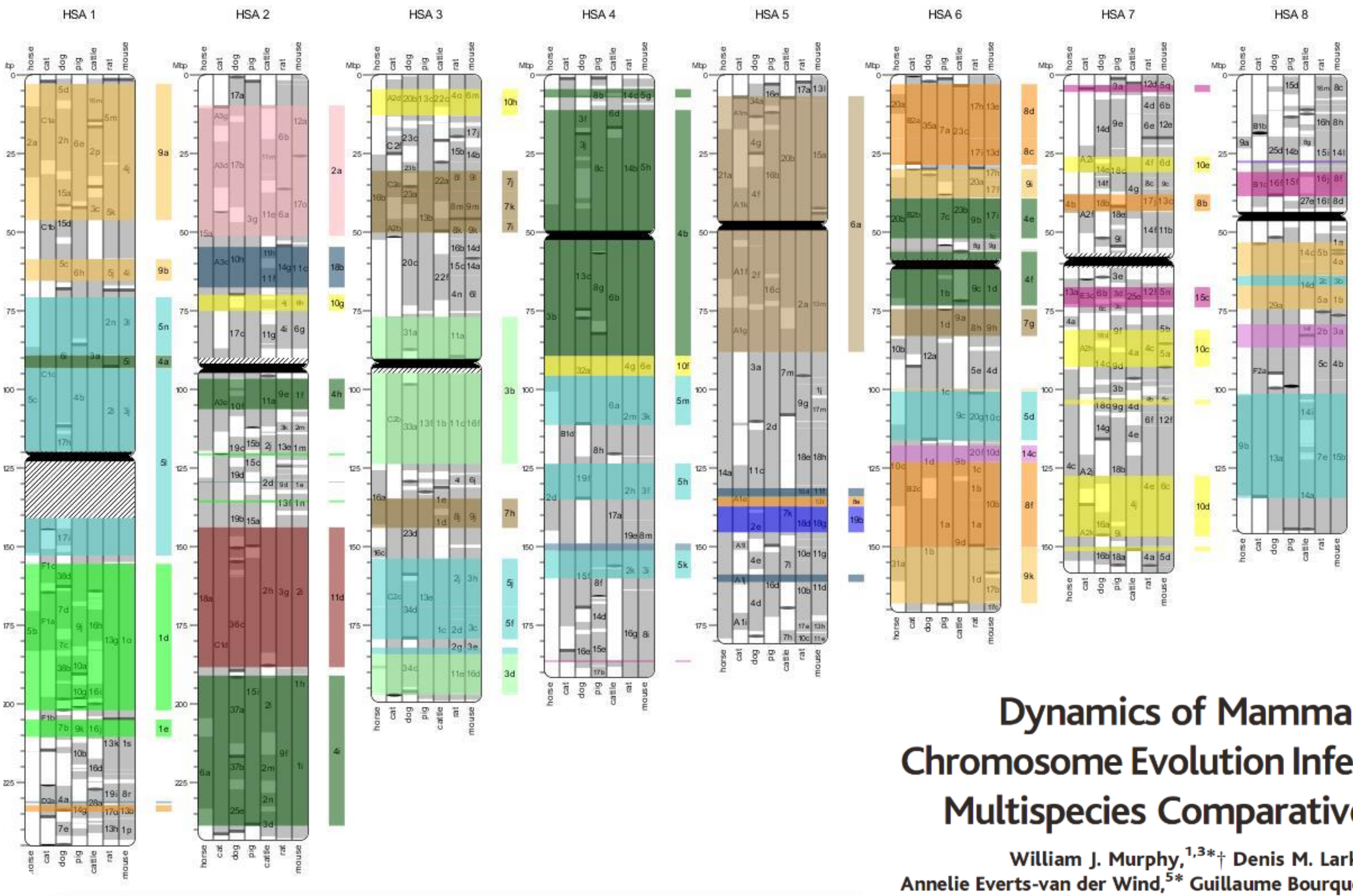
Synteny Shown by Chromosome Painting

Human on Aardvark





Ferguson-Smith & Trifonov *Nature Rev Genetics,* 2007



Science

AAAS

Science 309, 2005

Dynamics of Mammalian Chromosome Evolution Inferred from Multispecies Comparative Maps

William J. Murphy,[1,3*†] Denis M. Larkin,[5*]
Annelie Everts-van der Wind,[5*] Guillaume Bourque,[8] Glenn Tesler,[9]
Loretta Auvil,[6] Jonathan E. Beever,[5] Bhanu P. Chowdhary,[1]
Francis Galibert,[11] Lisa Gatzke,[6] Christophe Hitte,[11]
Stacey N. Meyers,[5] Denis Milan,[12] Elaine A. Ostrander,[13] Greg Pape,[6]
Heidi G. Parker,[13] Terje Raudsepp,[1] Margarita B. Rogatcheva,[5]
Lawrence B. Schook,[5,7] Loren C. Skow,[1] Michael Welge,[6]
James E. Womack,[2] Stephen J. O'Brien,[4]
Pavel A. Pevzner,[10] Harris A. Lewin[5,7†]

Murphy et al. *Science*. 2005

Royal Veterinary College
University of London

Lessons Learned from Multispecies Genome-Wide Comparisons of Chromosomal Organization

The distribution of evolutionary breakpoint regions in chromosomes is non-random (~20% reuse on a genome-wide basis)

Rates of chromosomal rearrangements differ in mammal lineages, and have sped up significantly after the Cretaceous-Tertiary boundary

High gene density, segmental duplications are associated with evolutionary breakpoint regions

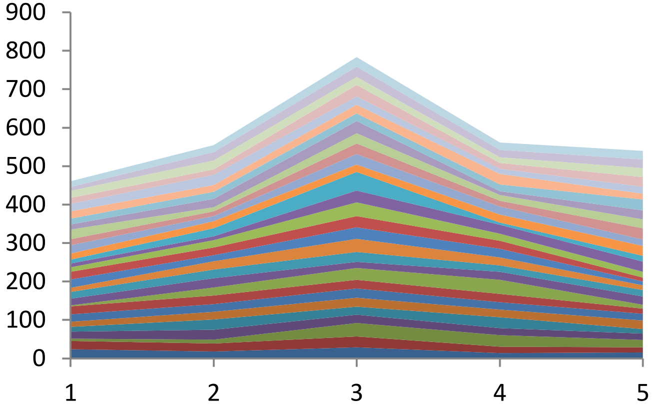
Evolutionary breakpoint regions affect genes related to clade or species-specific phenotypes and adaptations: e.g. defensin genes in cattle, taste receptors in pigs and forebrain development genes in budgerigar

Murphy et al. *Science*. 2005

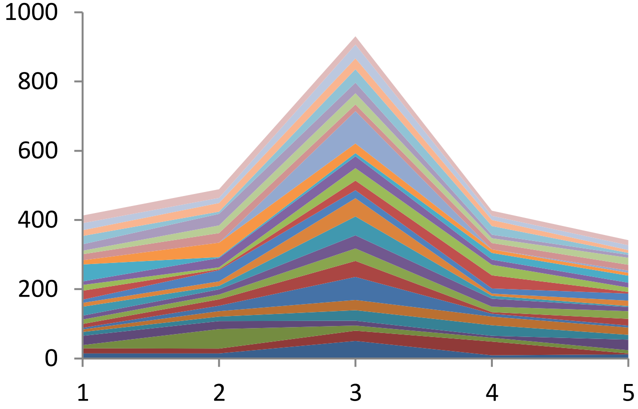
Royal
Veterinary
College
University of London

Density of Repetitive Elements in EBRs

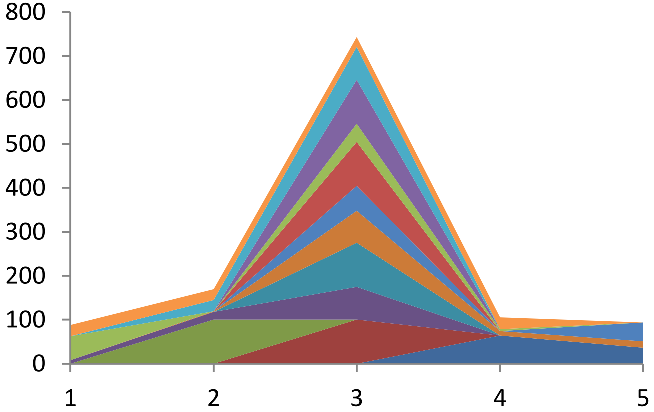
Cattle: **LINE-RTE**



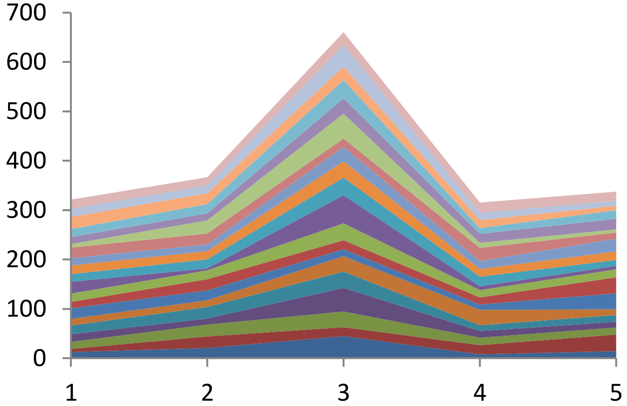
Dog: **LTR-ERV1**



Primates: **Satellite**



Rodents: **LTR-ERV1**

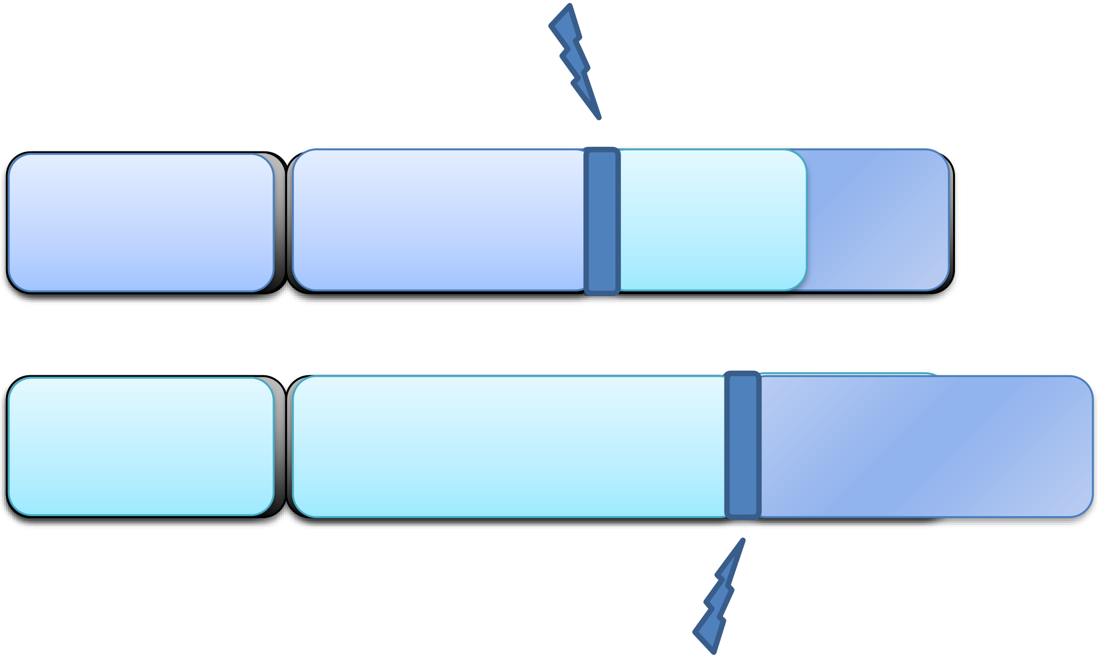


***Lineage-specific* transposable elements are enriched in *lineage-specific* EBRs**

Fragile Site Model of Chromosome Evolution

Invasion by retrotransposons
Segmental duplications
Hypomethylation

Chromosome Instability



NAHR

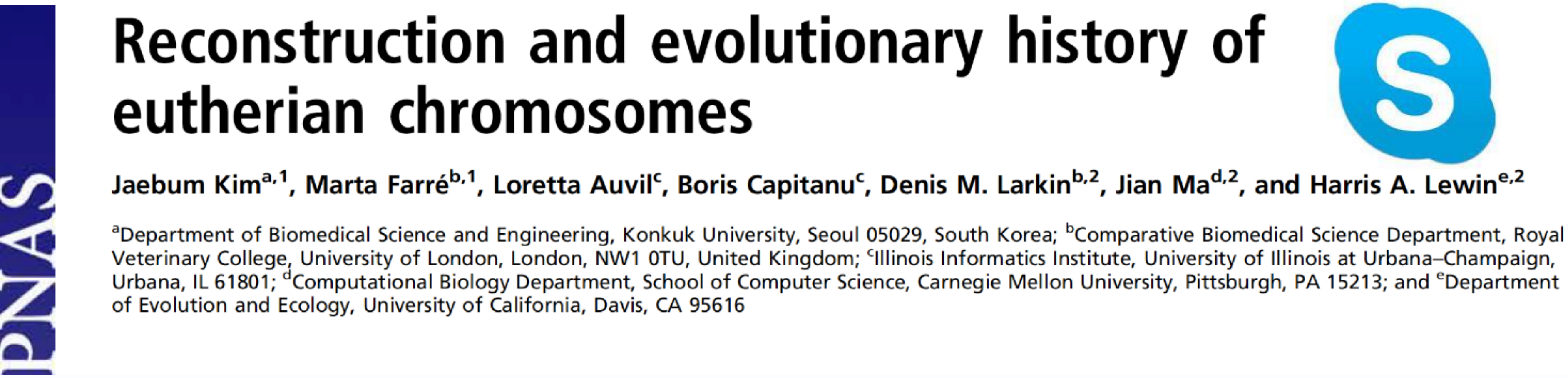
Inversions
Translocations
Fissions
Fusions

Aims

Development of an algorithm for reconstruction of chromosome structures for ancestral genomes from extant genome assemblies

Reconstruction of several ancestral genomes for the lineage leading to human

Revealing evolutionary history of each eutherian (and human) chromosome



Reconstruction and evolutionary history of eutherian chromosomes

Jaebum Kim[a,1], Marta Farré[b,1], Loretta Auvil[c], Boris Capitanu[c], Denis M. Larkin[b,2], Jian Ma[d,2], and Harris A. Lewin[e,2]

[a]Department of Biomedical Science and Engineering, Konkuk University, Seoul 05029, South Korea; [b]Comparative Biomedical Science Department, Royal Veterinary College, University of London, London, NW1 0TU, United Kingdom; [c]Illinois Informatics Institute, University of Illinois at Urbana–Champaign, Urbana, IL 61801; [d]Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213; and [e]Department of Evolution and Ecology, University of California, Davis, CA 95616

Dataset

19 placental mammal genomes
10 orders

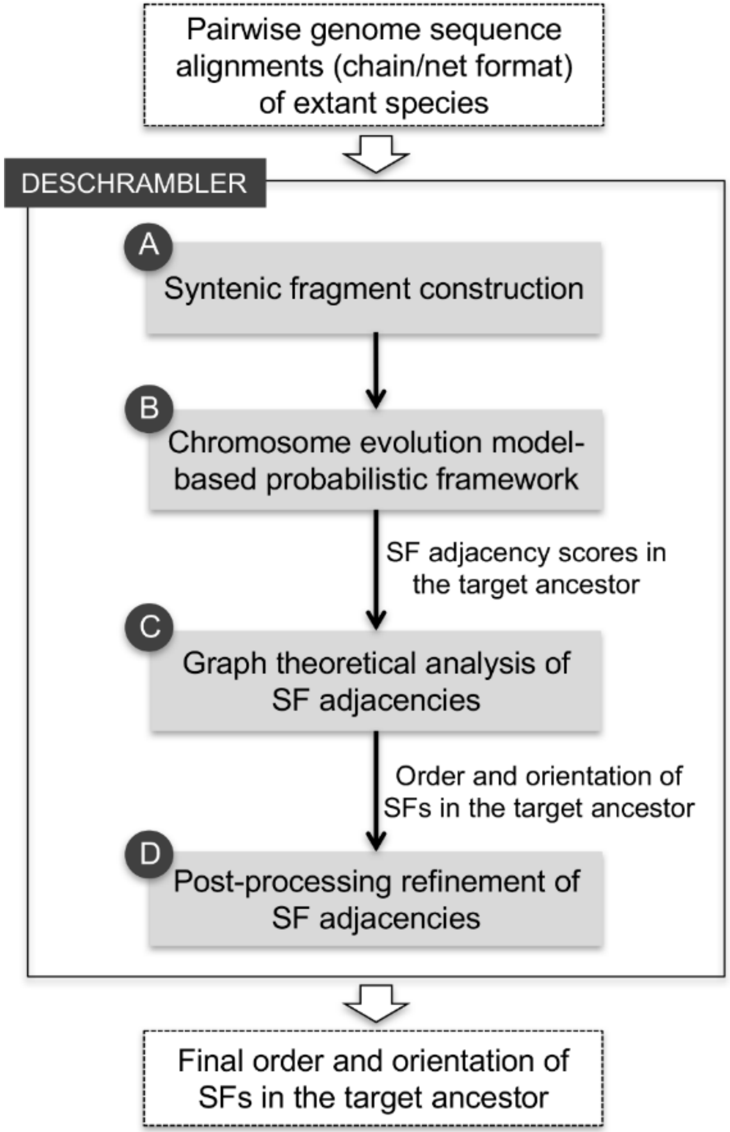
2 outgroups: opossum & chicken

14 chromosome assemblies
7 scaffold assemblies

Species	Order	Type
Human	Primates	Chr
Chimpanzee	Primates	Chr
Orangutan	Primates	Chr
Rhesus	Primates	Chr
Marmoset	Primates	Chr
Mouse	Rodentia	Chr
Rat	Rodentia	Chr
Guinea pig	Rodentia	Scf
Pika	Lagomorpha	Scf
Cattle	Cetartiodactyla	Chr
Goat	Cetartiodactyla	Chr
Pig	Cetartiodactyla	Chr
White rhinoceros	Perissodactyla	Scf
Horse	Perissodactyla	Chr
Dog	Carnivora	Chr
Shrew	Eulipotyphla	Scf
Elephant	Proboscidea	Scf
Manatee	Sirenia	Scf
Tenrec	Afrosoricida	Scf
Opossum	Didelphimorphia	Chr
Chicken	Galliformes	Chr

Kim et al. *PNAS*. 2017

Algorithm



Pairwise genome sequence alignments (chain/net format) of extant species

DESCHRAMBLER

A Syntenic fragment construction

B Chromosome evolution model-based probabilistic framework

SF adjacency scores in the target ancestor

C Graph theoretical analysis of SF adjacencies

Order and orientation of SFs in the target ancestor

D Post-processing refinement of SF adjacencies

Final order and orientation of SFs in the target ancestor

Kim et al. *PNAS*. 2017

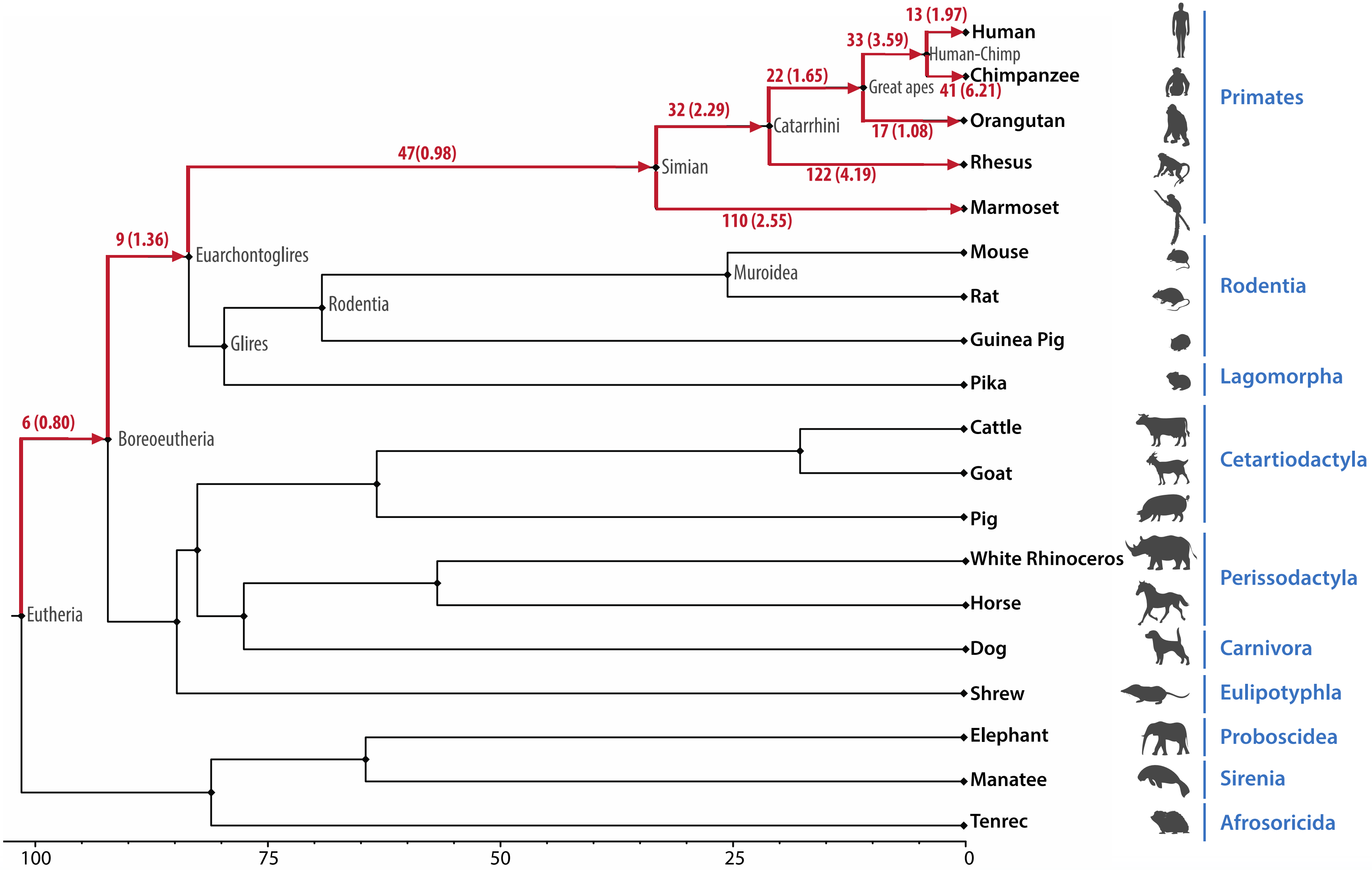
Ancestral Chromosome Reconstruction Statistics

Ancestor	No. APCFs	No. chromosomes	Total size (Kb)	Coverage (%)*
Eutherian	32	20	2,467,725	81
Boreoeutherian	34	23	2,536,880	84
Euarchontoglires	33	23	2,671,496	88
Simian	33	23	2,752,920	91
Catarrhini	33	23	2,767,322	91
Great apes	30	24	2,784,232	92
Human-chimpanzee	35	24	2,809,400	93

*reference genome (human)

Kim et al. *PNAS*. 2017

Rates of Chromosome Rearrangements



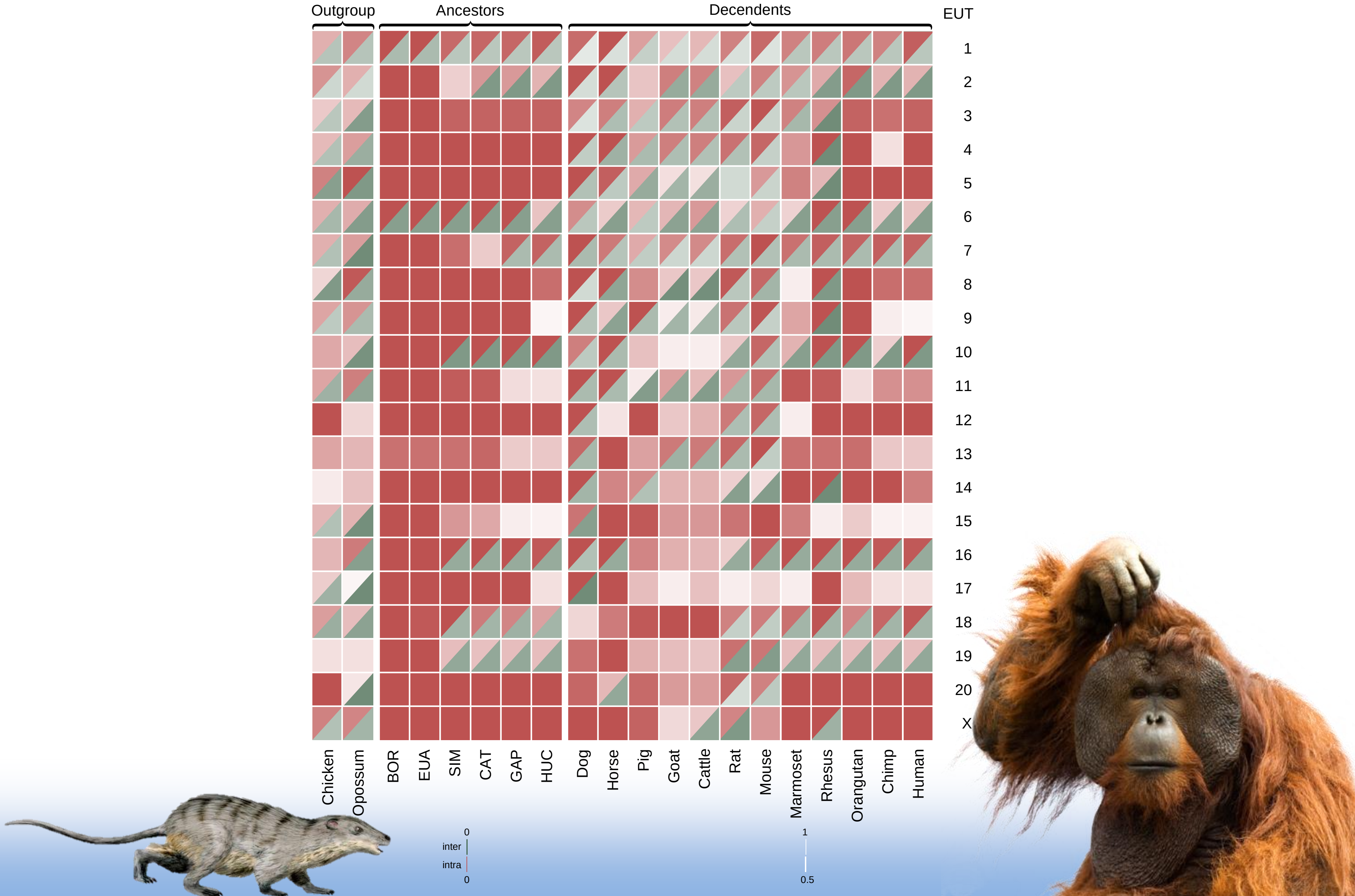
Kim et al. *PNAS*. 2017

RVC
Royal
Veterinary
College
University of London

Examples of Reconstructed Chromosomes

Kim et al. *PNAS*. 2017

Overall Picture



Summary

The eutherian ancestor had 42 chromosomes, four less than humans

The rates of evolution differ for different ancestral chromosomes in the different lineages

There were 117 chromosome rearrangements between the eutherian ancestor and the human genome

The rate of chromosome evolution accelerated along the primate lineage

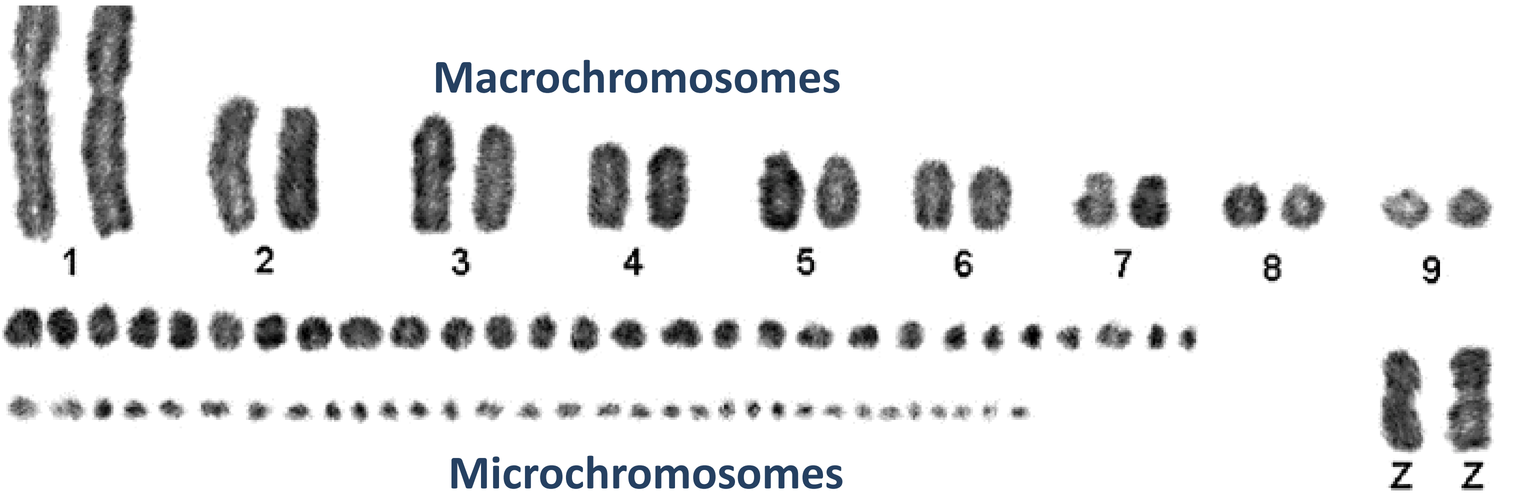
Avain Karyotypes are Different from Mammalian Ones

Stable no. of chromosomes (2n = ~80) in the majority of species

Presence of many micro-chromosomes

Low repeat content (~4% vs. 45% in mammals)

High fraction of conserved elements (15% vs. 7% in mammals)

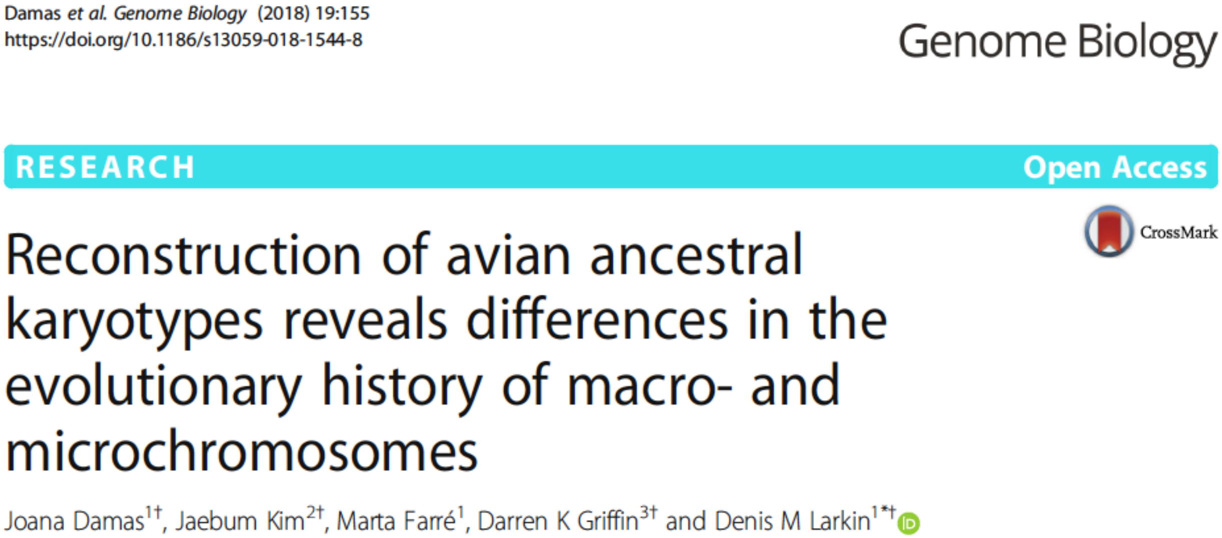


Aims

Reconstruct ancestral chromosome structures for 14 avain ancestors for the lineage leading to zebra finch

Compare evolutionary histories of micro- and macrochromosomes

Identify DNA features related to evolutionary stability of avian chromosomes in evolution



Damas *et al. Genome Biology* (2018) 19:155
https://doi.org/10.1186/s13059-018-1544-8

Genome Biology

RESEARCH Open Access

CrossMark

Reconstruction of avian ancestral karyotypes reveals differences in the evolutionary history of macro- and microchromosomes

Joana Damas[1†], Jaebum Kim[2†], Marta Farré[1], Darren K Griffin[3†] and Denis M Larkin[1*†]

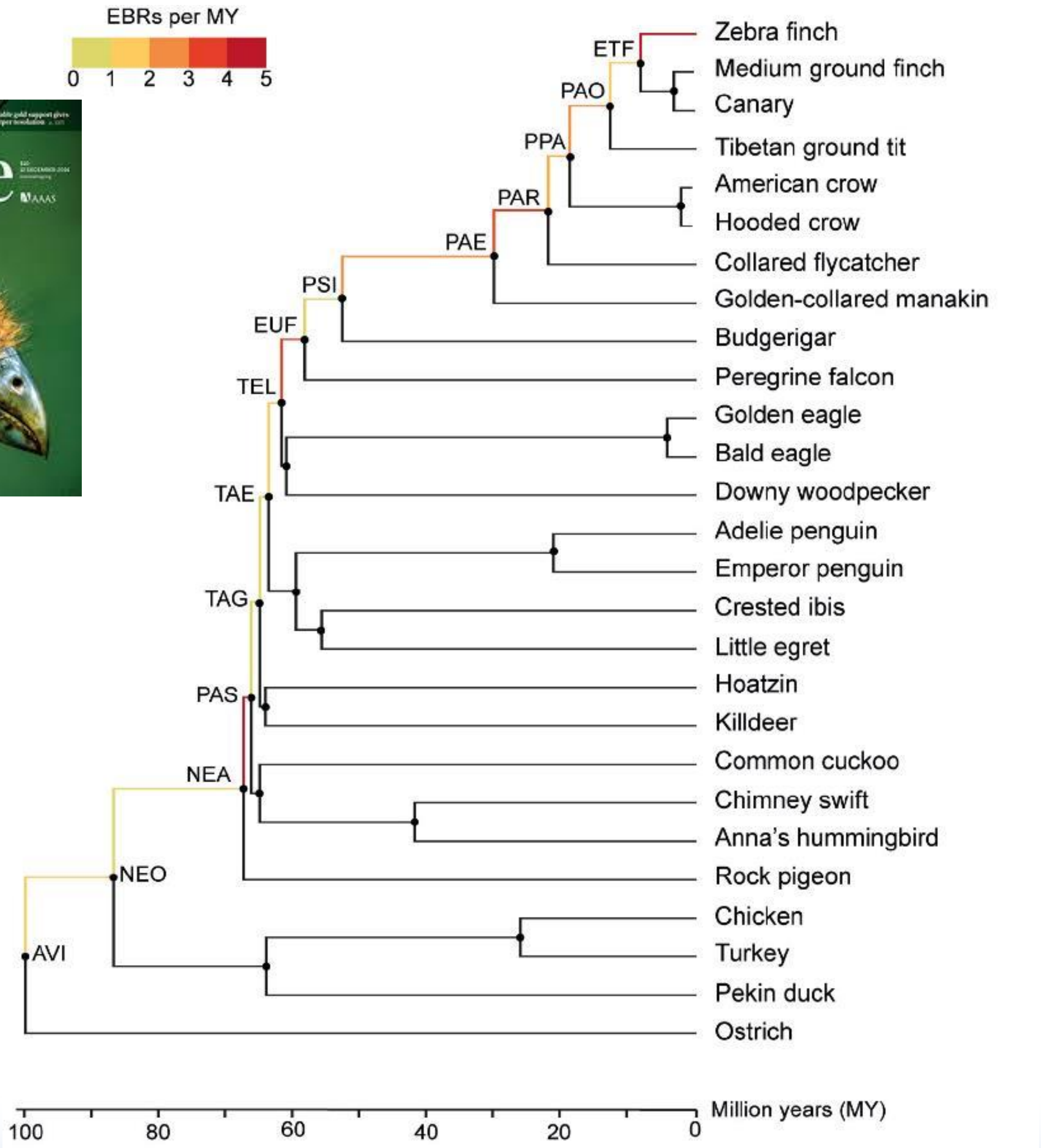
Damas et al. *Genome Biology*, 2018

Dataset

27 avian genomes
~17 Orders

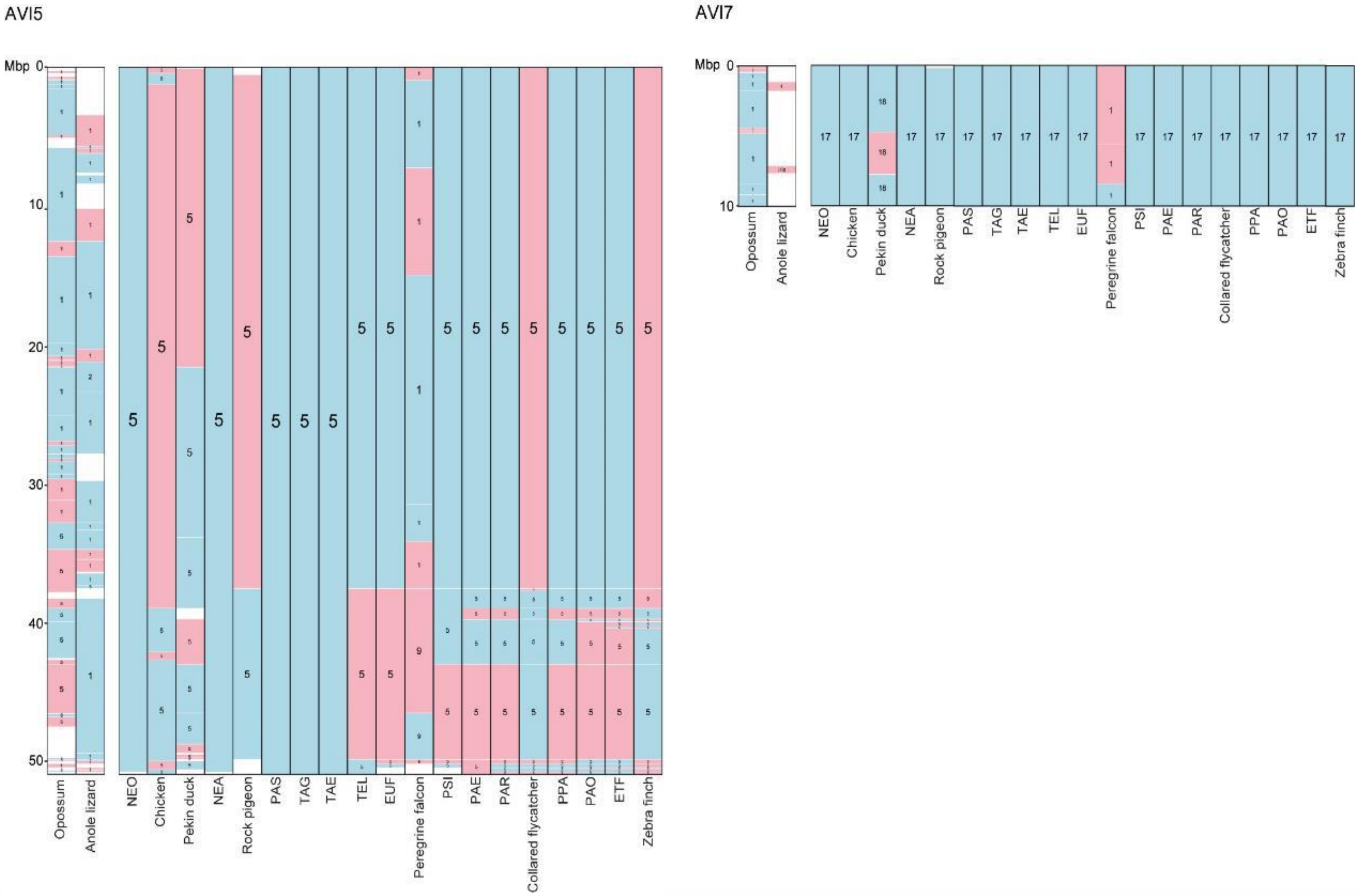
6 chromosome assemblies
21 scaffold assemblies

Outgroups:
opossum
Anole lizard



Damas et al. *Genome Biology*, 2018

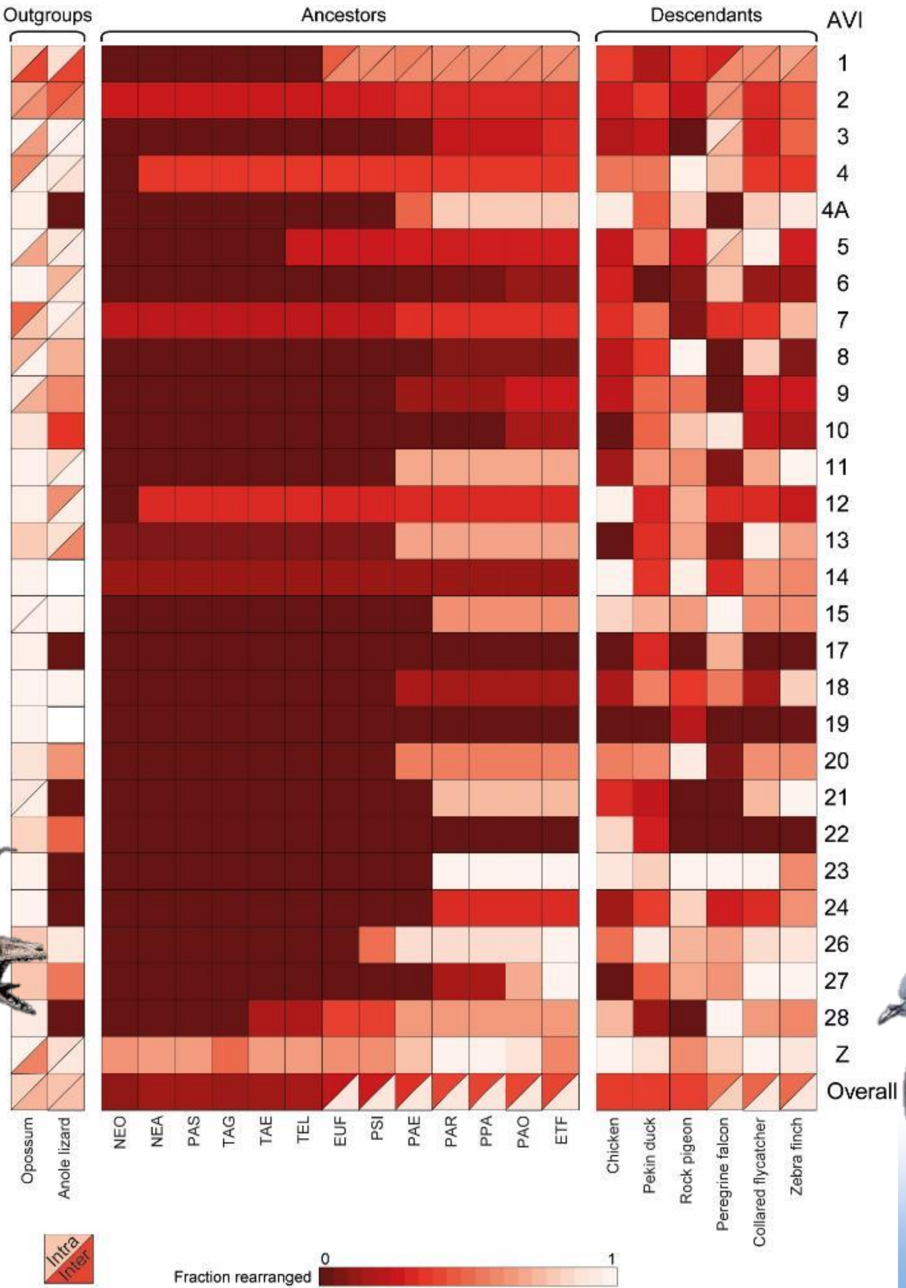
Example of a Reconstructed Macro- and Microchromosome



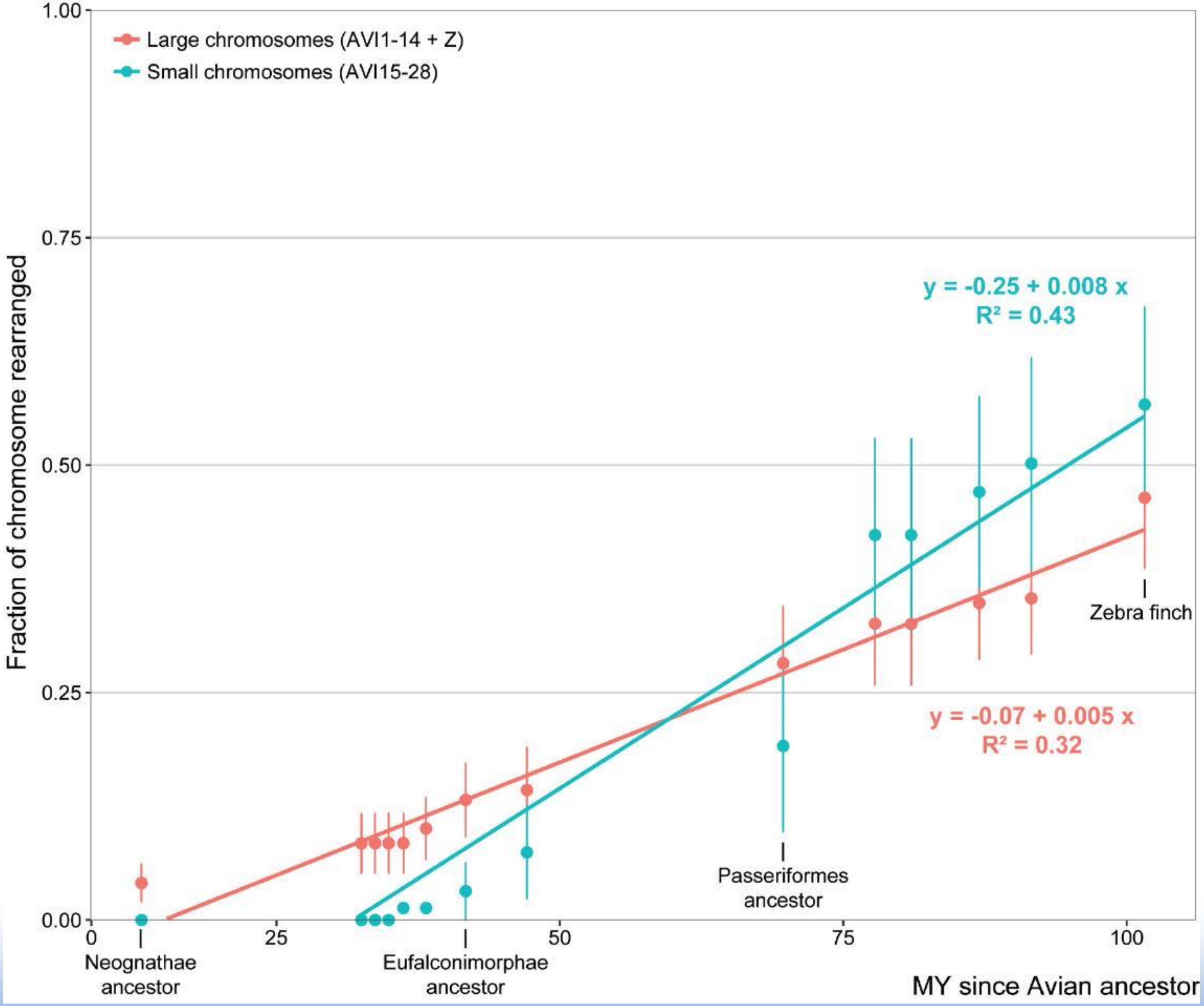
Damas et al. *Genome Biology*, 2018

Royal
Veterinary
College
University of London

Overall Picture

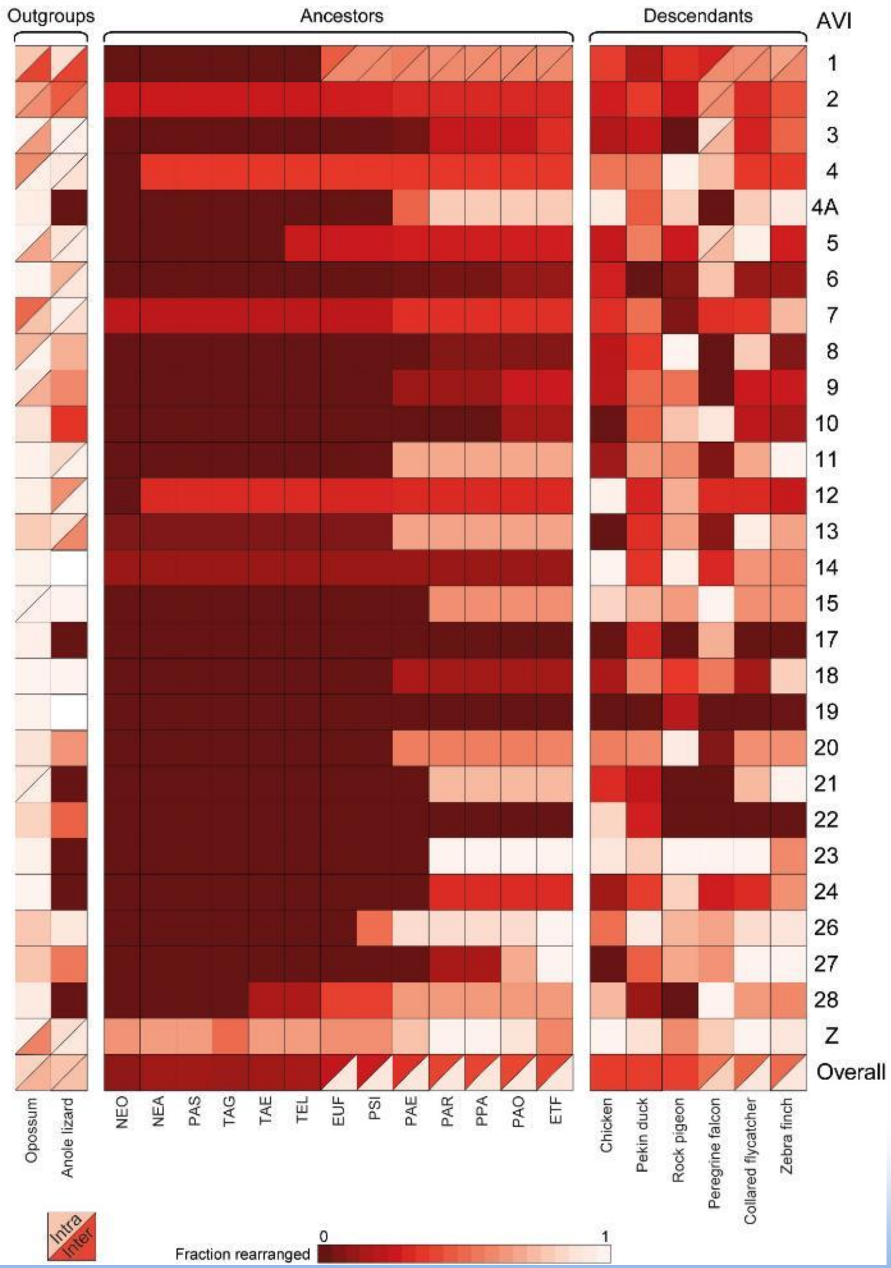


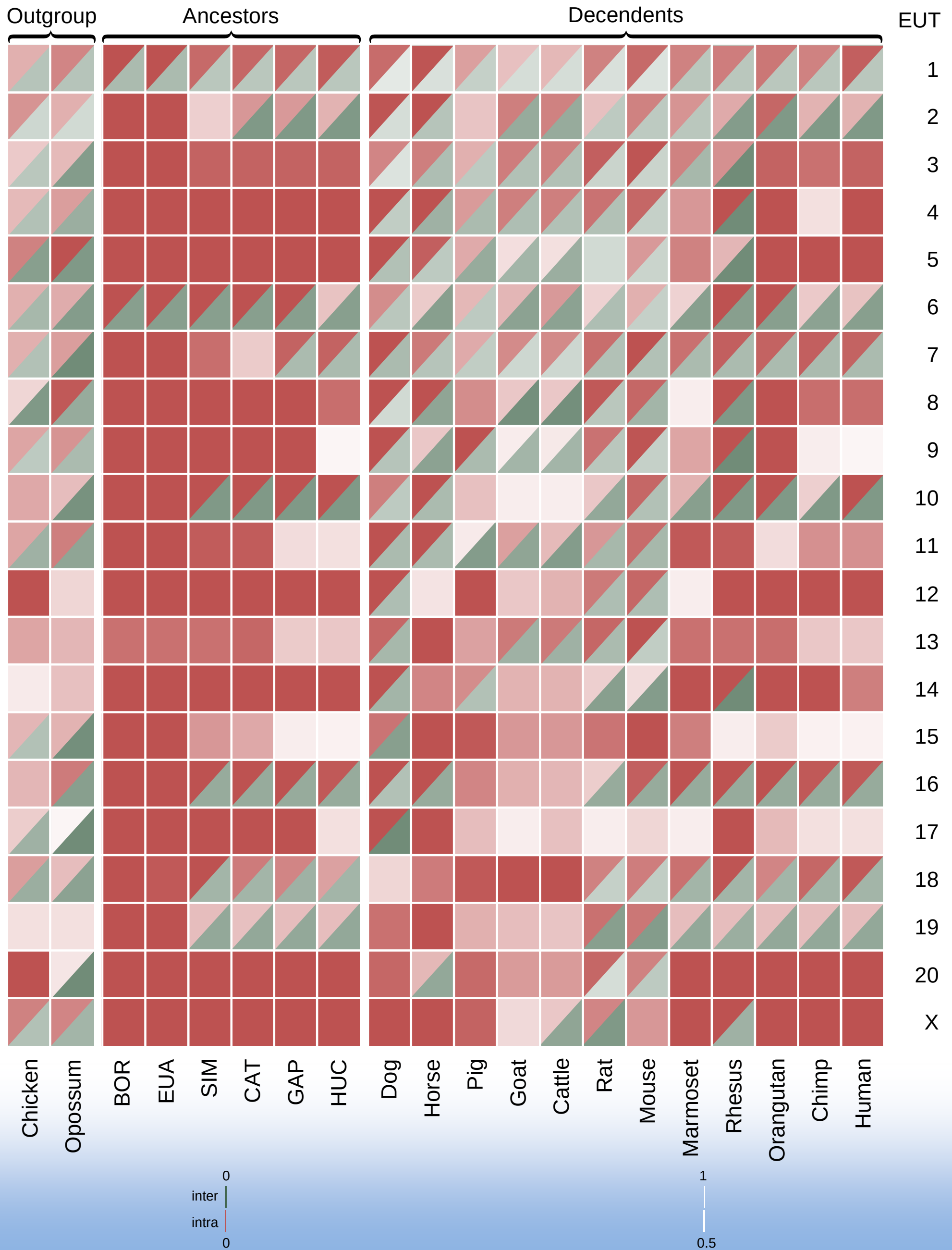
Evolutionary History of Micro- and Macrochromosomes



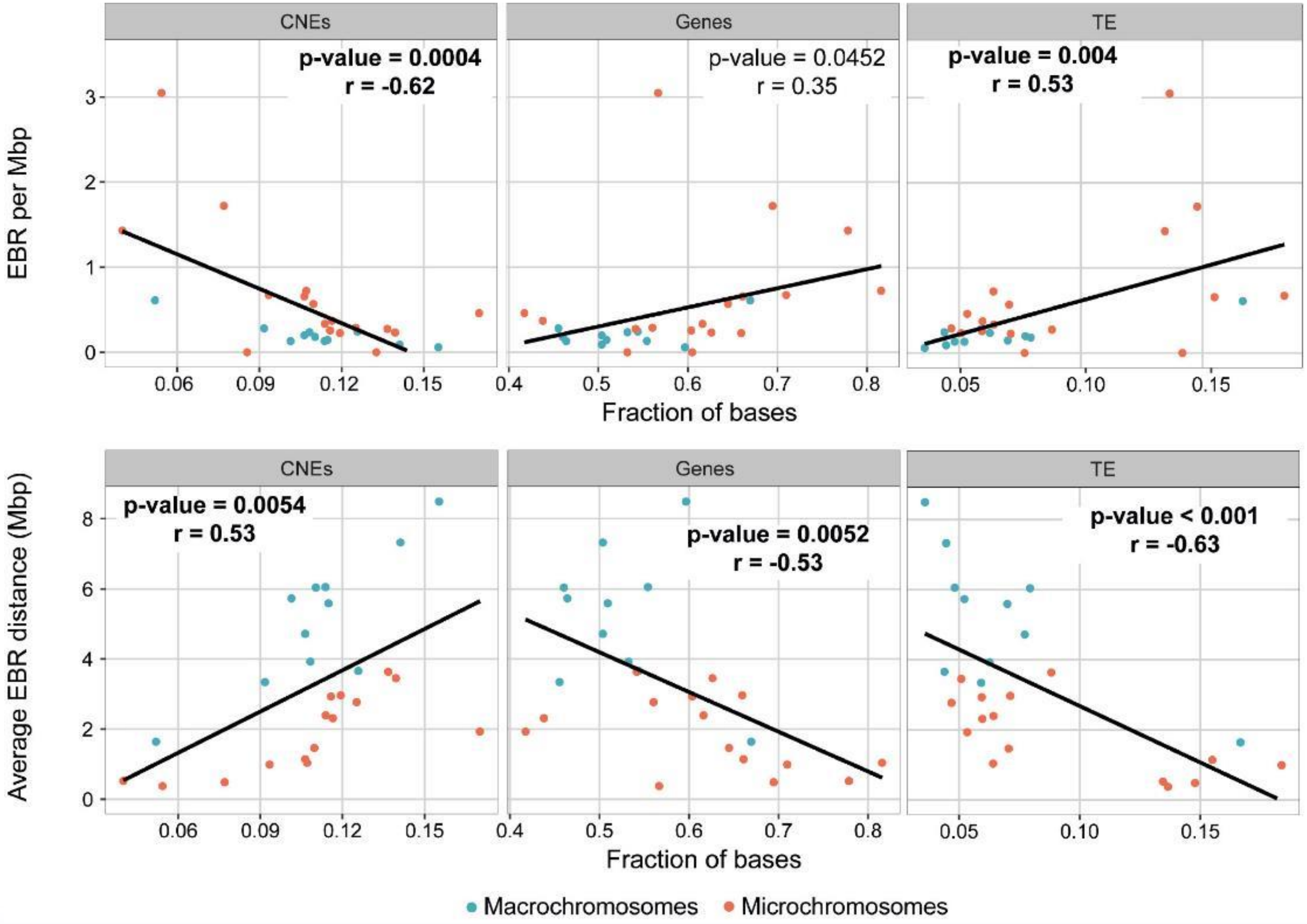
Damas et al. *Genome Biology*, 2018

Birds vs. Mammals





DNA Features and Density of EBRs in Avian Chromosomes



Damas et al. *Genome Biology*, 2018

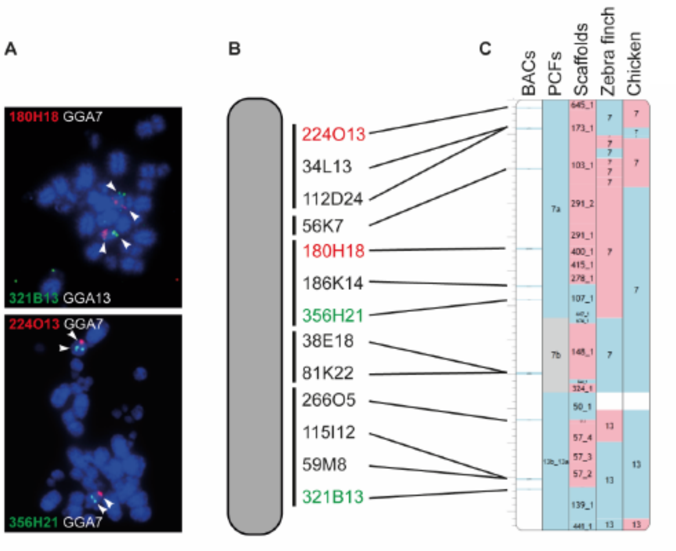
RVC Royal Veterinary College University of London

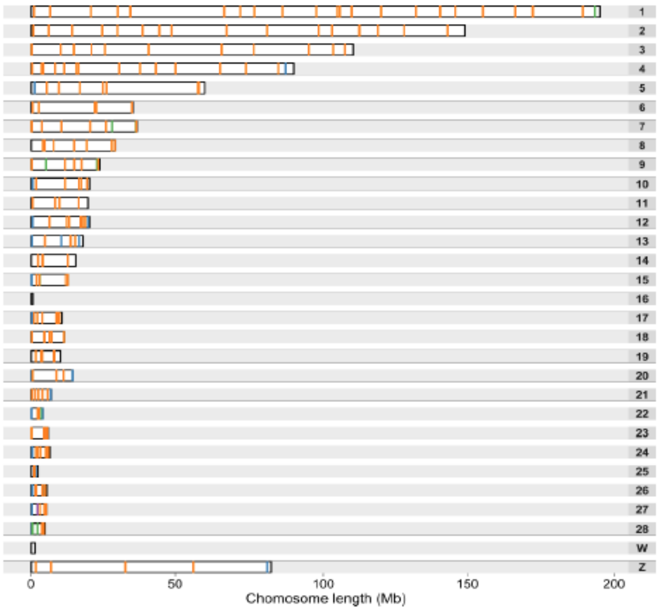
CSH PRESS

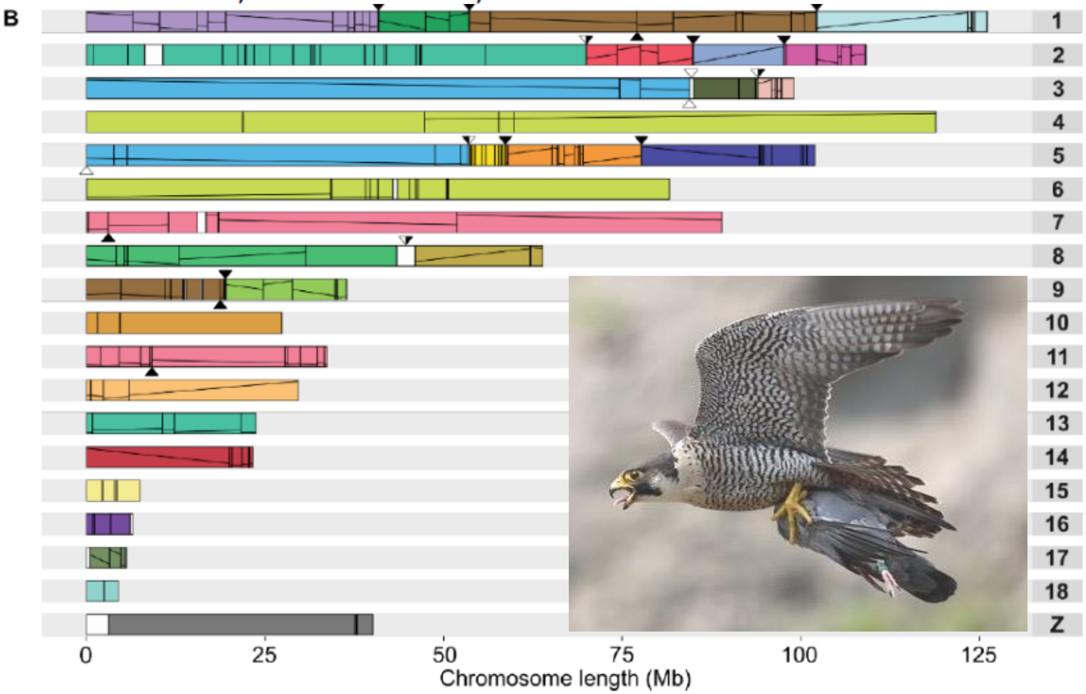
GENOME RESEARCH

Upgrading short-read animal genome assemblies to chromosome level using comparative genomics and a universal probe set

Joana Damas,[1,5] Rebecca O'Connor,[2,5] Marta Farré,[1] Vasileios Panagiotis E. Lenis,[3] Henry J. Martell,[2] Anjali Mandawala,[2,4] Katie Fowler,[4] Sunitha Joseph,[2] Martin T. Swain,[3] Darren K. Griffin,[2,6] and Denis M. Larkin[1,6]





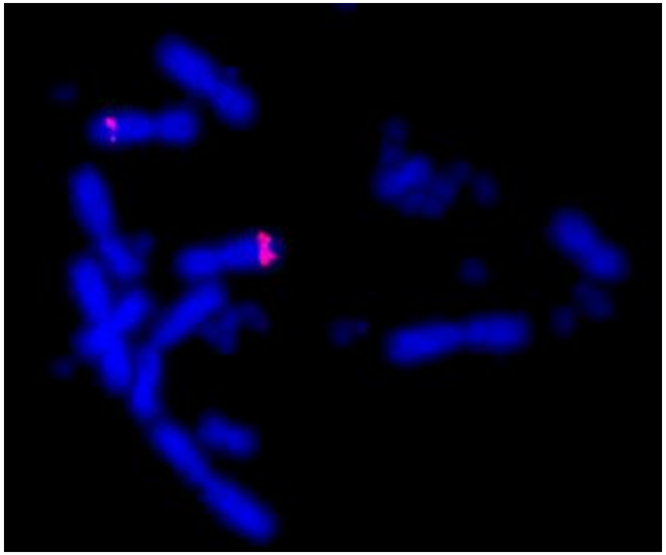


Mapping to chromosomes any good quality avian assembly

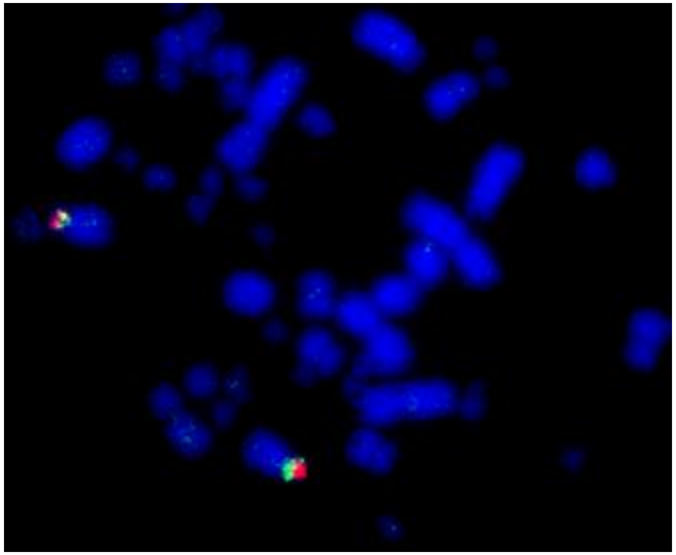
Royal Veterinary College
University of London

Damas et al., *Genome Research* 2017; O'Connor et al., *Genome Biology* 2018

Using Avian Conserved DNA Probes on Non-avian Reptile Chromosomes

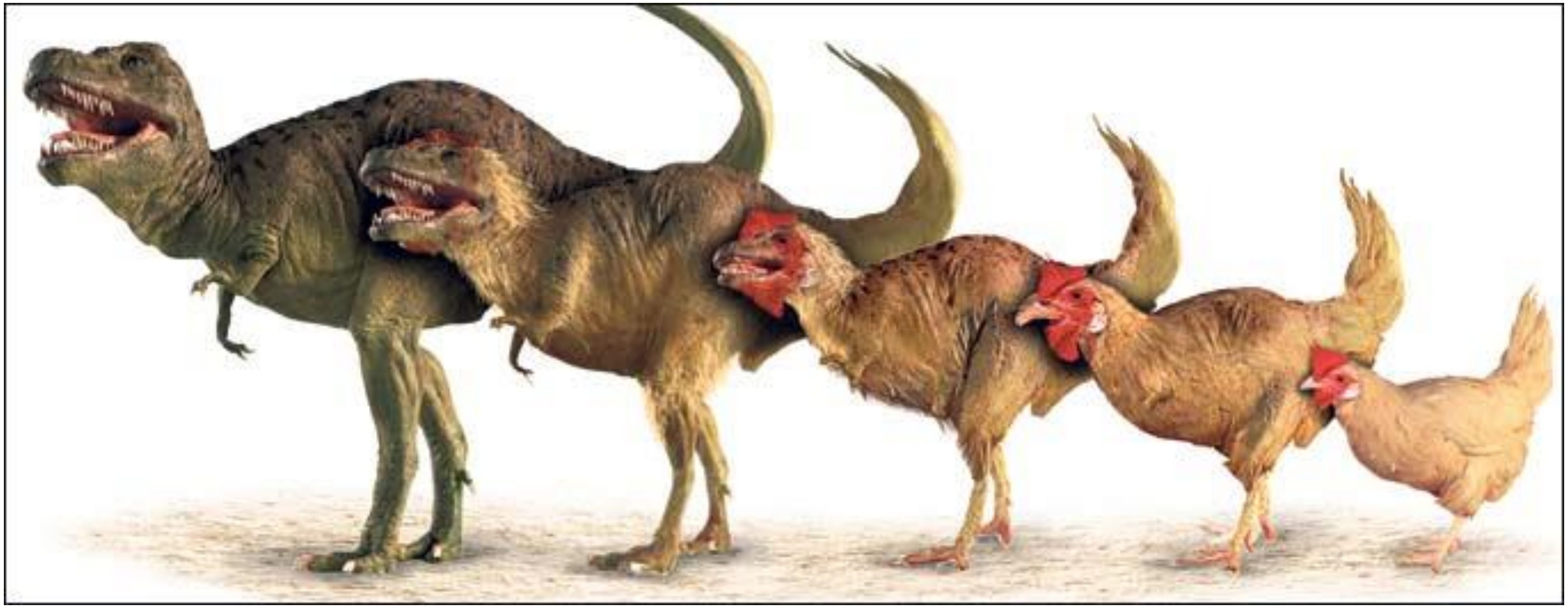


Lizard



Turtle

Reconstructing Dinosaur Chromosomes?



nature COMMUNICATIONS

ARTICLE

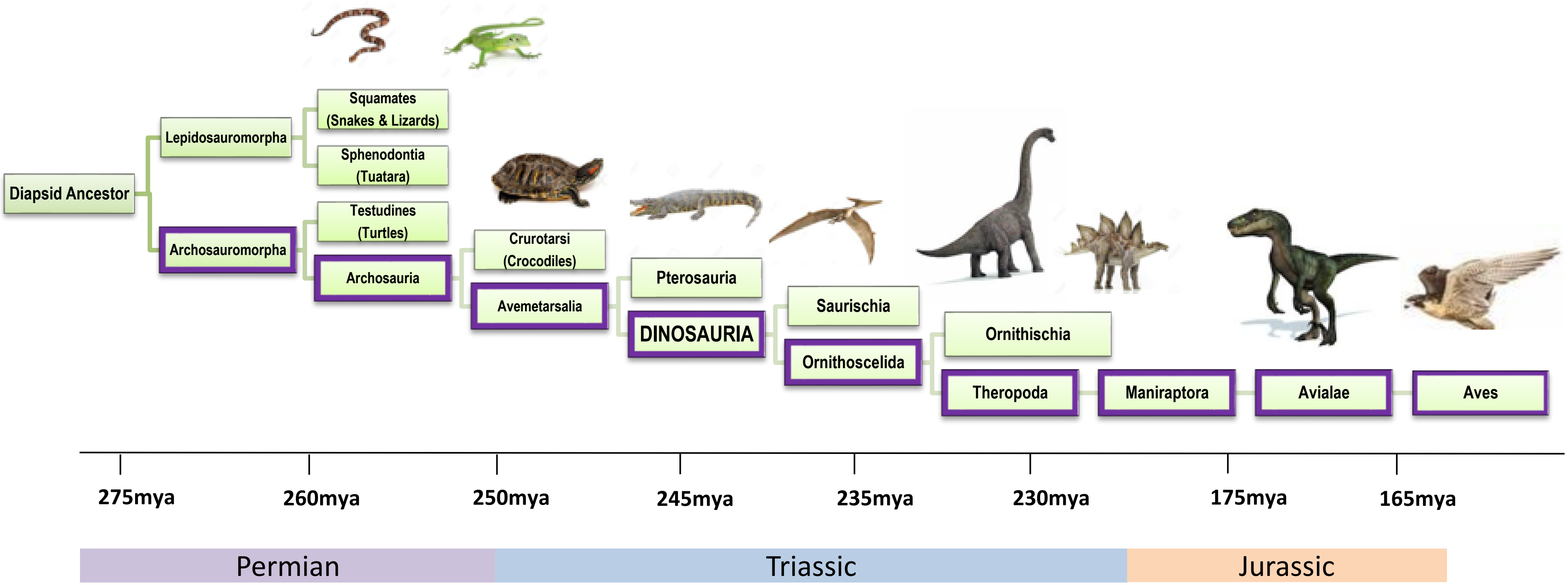
DOI: 10.1038/s41467-018-04267-9

OPEN

Reconstruction of the diapsid ancestral genome permits chromosome evolution tracing in avian and non-avian dinosaurs

Rebecca E. O'Connor[1], Michael N. Romanov[1], Lucas G. Kiazim[1], Paul M. Barrett[2], Marta Farré[3], Joana Damas[3], Malcolm Ferguson-Smith[4], Nicole Valenzuela[5], Denis M. Larkin[3] & Darren K. Griffin[1]

Theropod Dinosaur Lineage



Diapsid Ancestor

Lepidosauromorpha

Squamates (Snakes & Lizards)

Sphenodontia (Tuatara)

Archosauromorpha

Testudines (Turtles)

Archosauria

Crurotarsi (Crocodiles)

Avemetarsalia

Pterosauria

DINOSAURIA

Saurischia

Ornithoscelida

Ornithischia

Theropoda

Maniraptora

Avialae

Aves

275mya

260mya

250mya

245mya

235mya

230mya

175mya

165mya

Permian

Triassic

Jurassic

O’Connor et al., *Nature Communications,* 2018

Royal Veterinary College
University of London

Fluorescence *in situ* Hybridization

Eight microchromosomes unchanged in all three species

13 microchromosomes unchanged between chicken and turtle

Three 'protomicrochromosomes' – probes hybridise to a macrochromosome in the turtle and lizard but to a microchromosome in all avian species

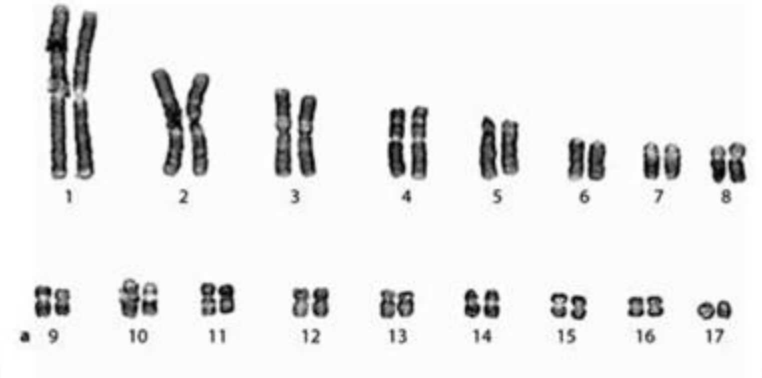
O'Connor et al., *Nature Communications,* 2018

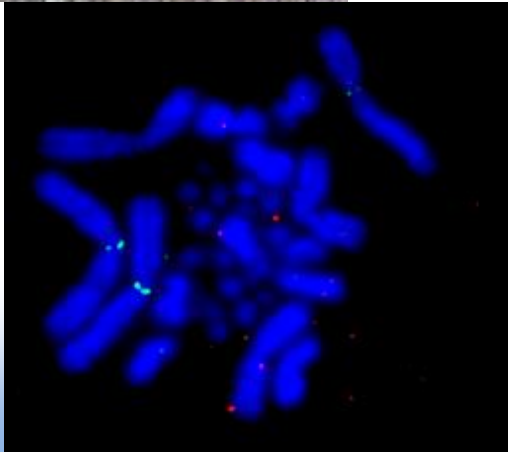
Crocs Are Bad

Atypical reptilian karyotype

- 2n=44 ish
- No microchromosomes
- ~23 species that underwent wholesale chromosome fusion

A shame because it might have given us Archosaurian ancestor ~240MYA

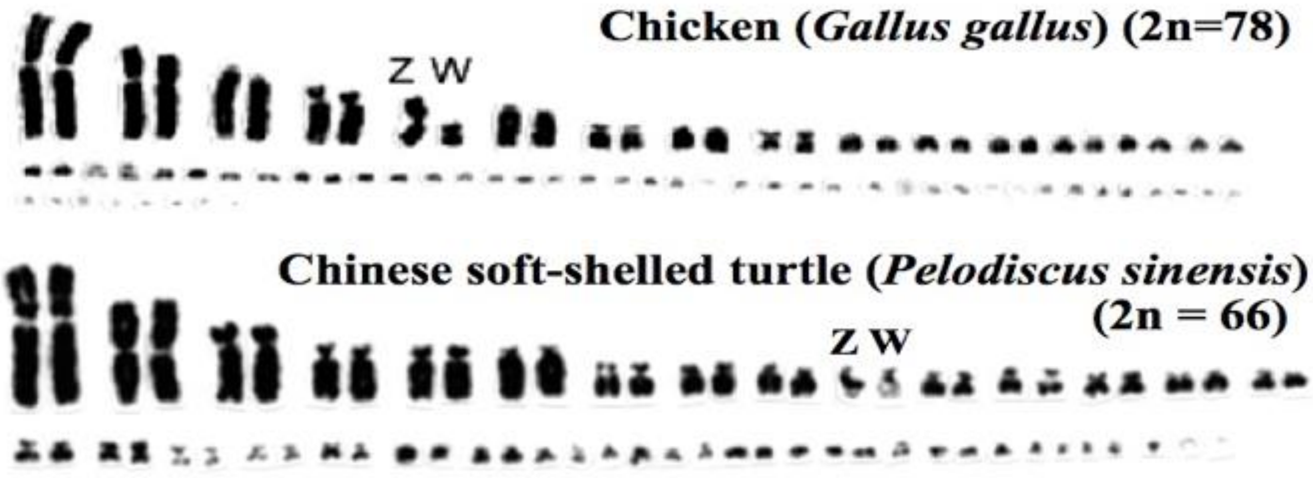




Combined Previous Karyotype and Chromosome Painting Data

Highly conserved avian FISH probes used on:

- Anole lizard (*Anolis carolinensis*)
- Red eared slider turtle (*Trachemys scripta*)
- Spiny soft-shelled turtle (*Apalone spinifera*)



O'Connor et al., *Nature Communications,* 2018

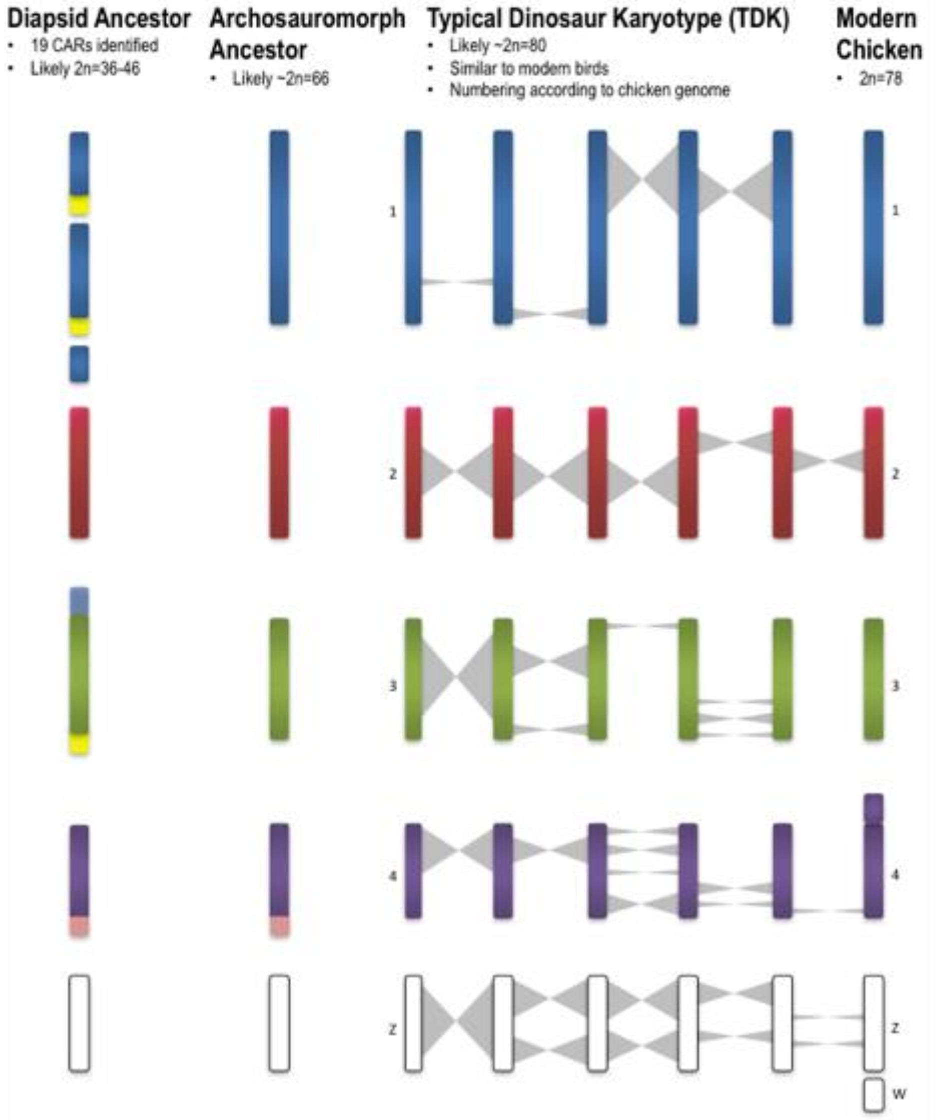
Ancestor to Avian Genome

Interchromosomal rearrangements

- 10 between diapsid ancestor and chicken
- 9 prior to turtle/avian ancestor

Intrachromosomal rearrangements

- 49 between diapsid ancestor and chicken
- Most occurred post turtle/avian ancestor



O'Connor et al., *Nature Communications* 2018

Ancestor to Avian Genome

Interchromosomal rearrangements

- 10 between diapsid ancestor and chicken
- 9 prior to turtle/avian ancestor

Intrachromosomal rearrangements

- 49 between diapsid ancestor and chicken
- Most occurred post turtle/avian ancestor

Diapsid Ancestor

- 19 CARs identified
- Likely 2n=36-46

Archosauromorph Ancestor

- Likely ~2n=66

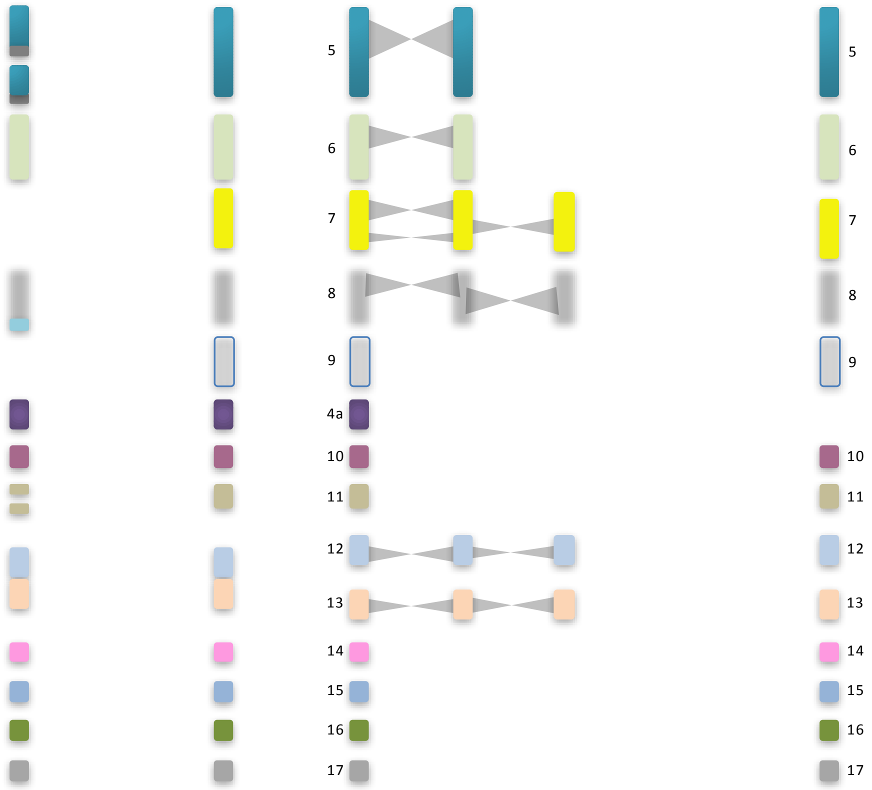
Typical Avian/Dinosaur Karyotype

Likely ~2n=80

- Similar to modern birds
- Numbering according to chicken genome

Modern Chicken

- 2n=78



O’Connor et al., *Nature Communications* 2018

Ancestor to Avian Genome

Interchromosomal rearrangements

- 10 between diapsid ancestor and chicken
- 9 prior to turtle/avian ancestor

Intrachromosomal rearrangements

- 49 between diapsid ancestor and chicken
- Most occurred post turtle/avian ancestor

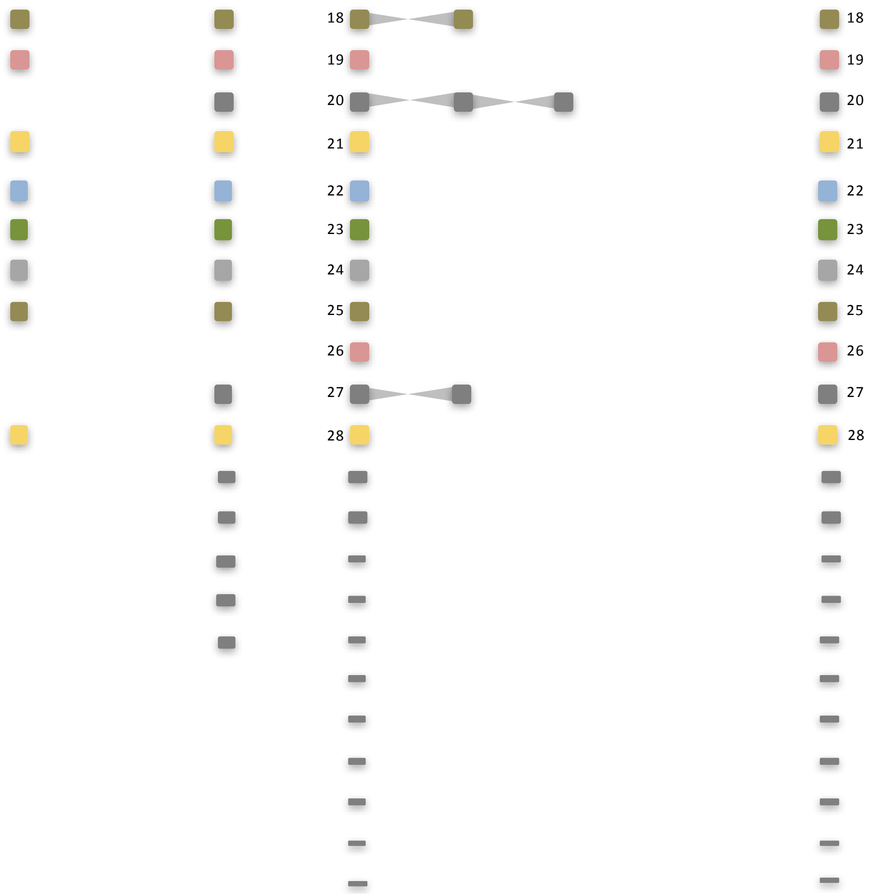
Diapsid Ancestor
- 19 CARs identified
- Likely 2n=36-46

Archosauromorph Ancestor
- Likely ~2n=66

Typical Avian/Dinosaur Karyotype

Likely ~2n=80
- Similar to modern birds
- Numbering according to chicken genome

Modern Chicken
- 2n=78

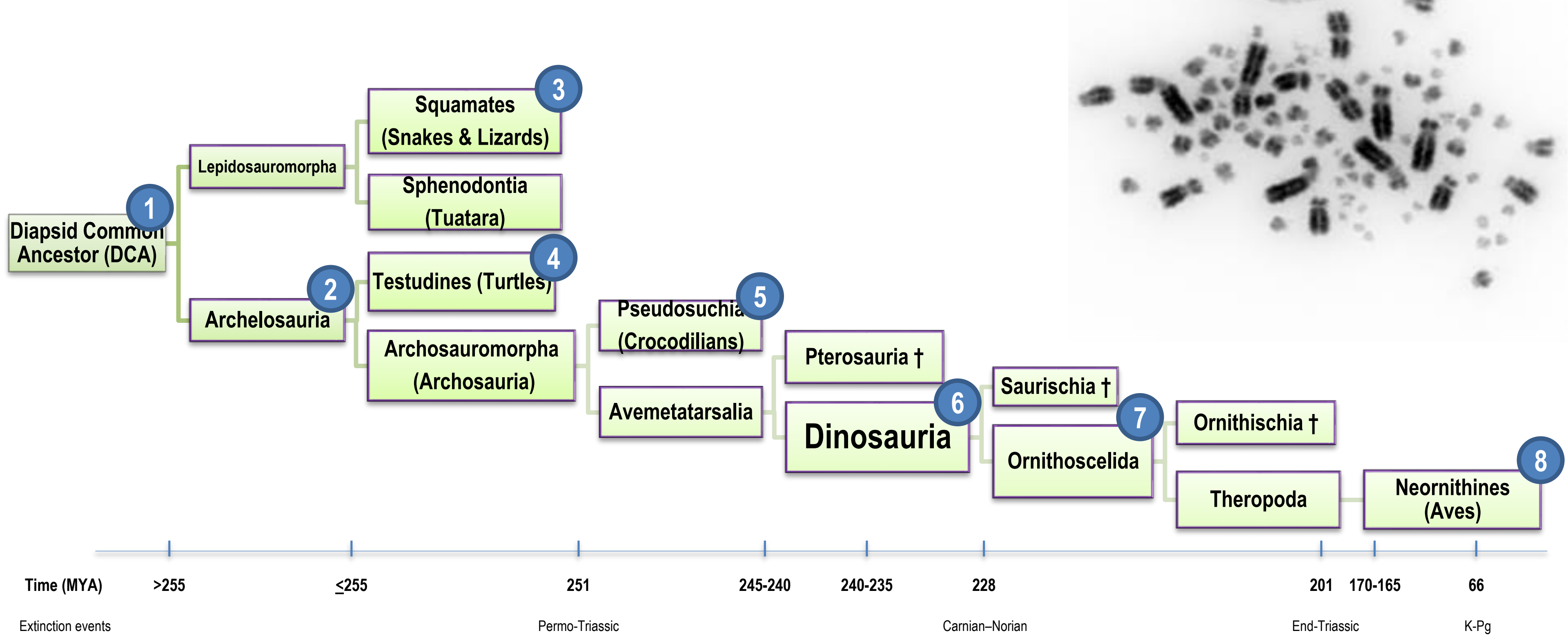


19 CARs identified
- Some CARs likely fused as single chromosomes
- Hence this diagram appears as apparently more chromosomes
- Some chromosomes not covered by sequence assembly

Most chromosomes syntenic to modern birds

Multiple fissions
- Karyotype similar to modern birds
- Numbering according to chicken genome

Known fusion of chromosome 4 (not present in most birds) is shown



1 Diapsid Common Ancestor (DCA)

Lepidosauromorpha

3 Squamates (Snakes & Lizards)

Sphenodontia (Tuatara)

2 Archelosauria

4 Testudines (Turtles)

Archosauromorpha (Archosauria)

5 Pseudosuchia (Crocodilians)

Avemetatarsalia

Pterosauria †

6 Dinosauria

Saurischia †

7 Ornithoscelida

Ornithischia †

Theropoda

8 Neornithines (Aves)

Time (MYA) >255 ≤255 251 245-240 240-235 228 201 170-165 66

Extinction events Permo-Triassic Carnian–Norian End-Triassic K-Pg

- Diapsid ancestor ~2*n*=40 (half microchromosomes)
- Predominantly fissions until turtle bird ancestor ~260 mya ~2*n*=64-68
- At a similar rate of fission avian pattern (2*n*=80) established around dinosaur/avian ancestor
- *Likely that dinosaur karyotypes looked like avian ones*

nature COMMUNICATIONS

ARTICLE

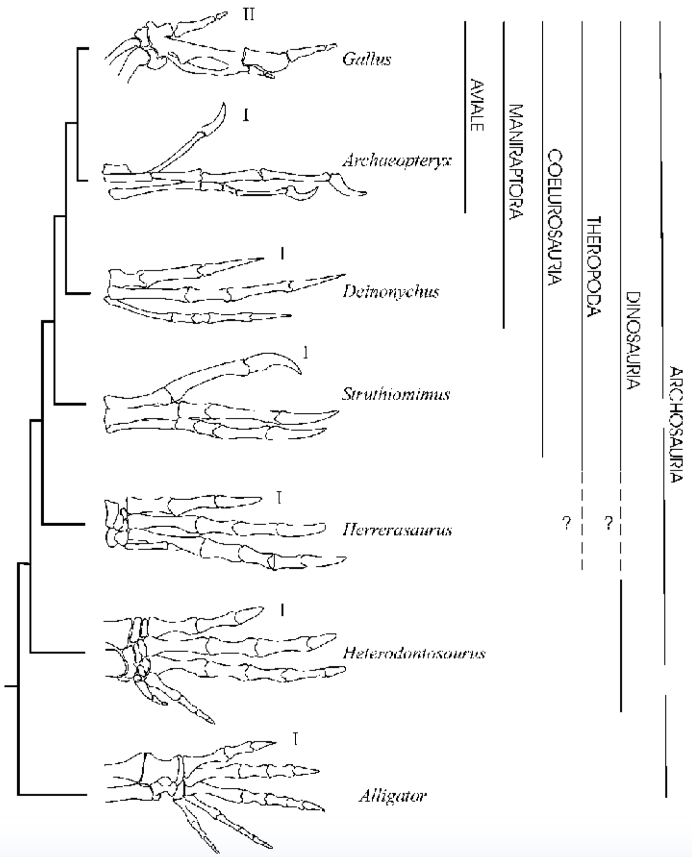
DOI: 10.1038/s41467-018-04267-9 OPEN

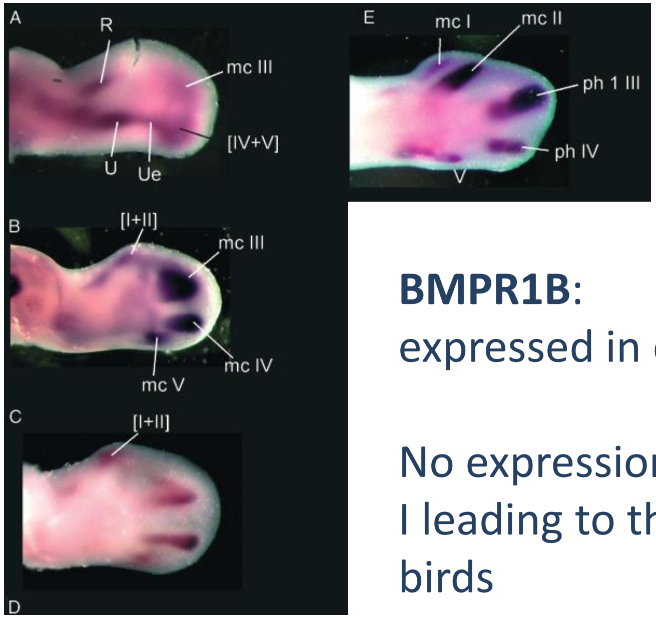
Reconstruction of the diapsid ancestral genome permits chromosome evolution tracing in avian and non-avian dinosaurs

Rebecca E. O'Connor[1], Michael N. Romanov[1], Lucas G. Kiazim[1], Paul M. Barrett[2], Marta Farré[3], Joana Damas[3], Malcolm Ferguson-Smith[4], Nicole Valenzuela[5], Denis M. Larkin[3] & Darren K. Griffin[1]

Avian-msHSBs are Enriched for Genes Related to Limb Development

Alligators and crocodiles have 5 digit limbs
Archaeopteryx had feathers and was able to fly, but it still had the digit I in limbs





DLX5:
responsible for feathered phenotypes

BMPR1B:
expressed in chicken wings

No expression in the area of digit I leading to the lack of digit I in birds

GBE

Novel Insights into Chromosome Evolution in Birds, Archosaurs, and Reptiles

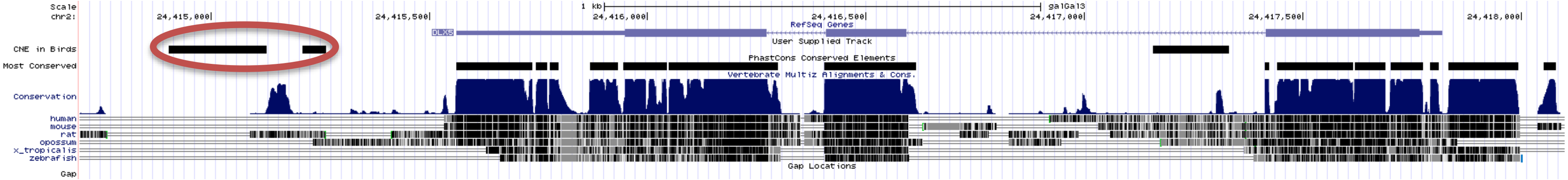
Marta Farré[1], Jitendra Narayan[2], Gancho T. Slavov[2], Joana Damas[1], Loretta Auvil[3], Cai Li[4,5], Erich D. Jarvis[6,7], David W. Burt[8], Darren K. Griffin[9,†], and Denis M. Larkin[1,†,*]

DLX5/6, BMPR1B, SHOX, HOXA11

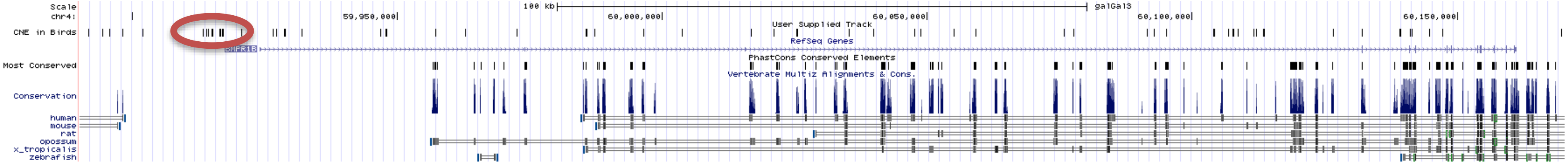
Farré et al, Genome Biology Evolution, 2016

Possible Changes in Regulation in Birds Due to New Conserved Non-coding Elements (CNEs) in HSBs

DLX5



BMPR1B



CNEs present in birds but not in other vertebrates

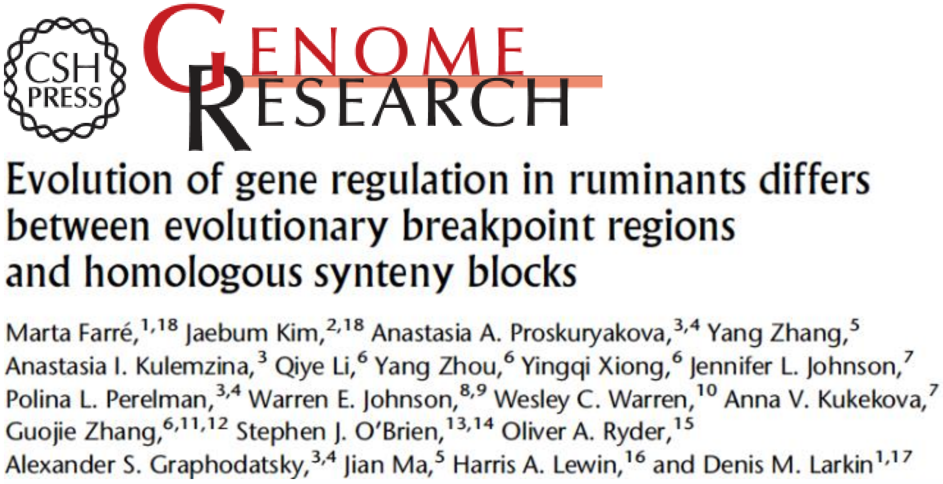
Farré et al., *Genome Biology Evolution*, 2016

Reconstruction of Ancestral Ruminant Chromosomes

Reconstruction of several ancestral genomes for the lineage leading to cattle

Revealing links between chromosome rearrangements and DNA features in the ruminant lineage

Finding effects of chromosome rearrangements on gene expression



CSH PRESS GENOME RESEARCH

Evolution of gene regulation in ruminants differs between evolutionary breakpoint regions and homologous synteny blocks

Marta Farré,[1,18] Jaebum Kim,[2,18] Anastasia A. Proskuryakova,[3,4] Yang Zhang,[5] Anastasia I. Kulemzina,[3] Qiye Li,[6] Yang Zhou,[6] Yingqi Xiong,[6] Jennifer L. Johnson,[7] Polina L. Perelman,[3,4] Warren E. Johnson,[8,9] Wesley C. Warren,[10] Anna V. Kukekova,[7] Guojie Zhang,[6,11,12] Stephen J. O'Brien,[13,14] Oliver A. Ryder,[15] Alexander S. Graphodatsky,[3,4] Jian Ma,[5] Harris A. Lewin,[16] and Denis M. Larkin[1,17]

Farré et al., *Genome Research* 2019

Dataset

18 mammal genomes

4 Orders

9 chromosome assemblies

9 scaffold assemblies

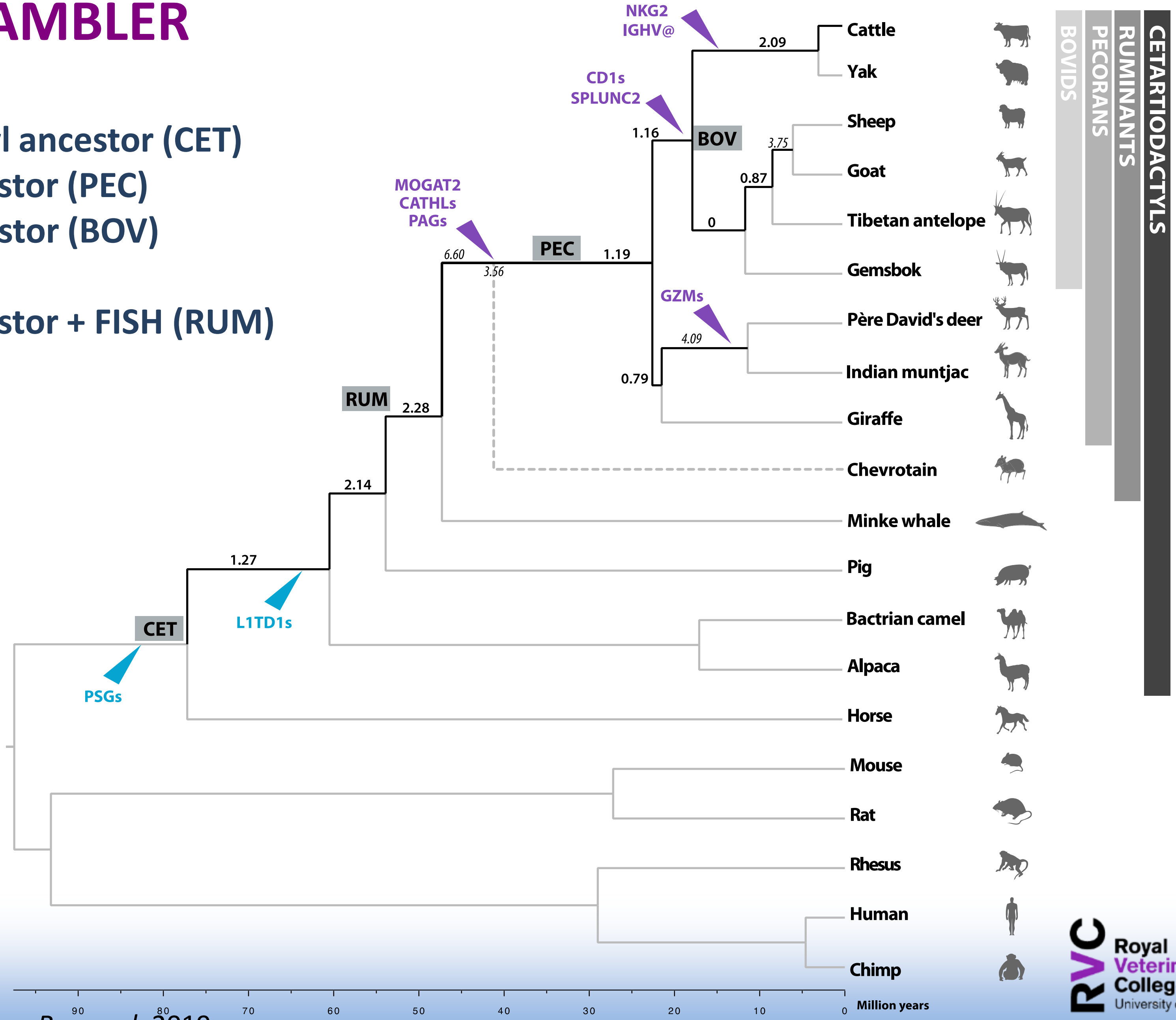
Species	Order	Type
Cattle	Cetartiodactyla	Chr
Yak	Cetartiodactyla	Scf
Sheep	Cetartiodactyla	Chr
Goat	Cetartiodactyla	Chr
Tibetan antelope	Cetartiodactyla	Scf
Gemsbok	Cetartiodactyla	Scf
Pere David's deer	Cetartiodactyla	Scf
Indian muntjac	Cetartiodactyla	Scf
Giraffe	Cetartiodactyla	Scf
Minke whale	Cetartiodactyla	Scf
Bactrian camel	Cetartiodactyla	Scf
Alpaca	Cetartiodactyla	Scf
Horse	Perissodactyla	Chr
Mouse	Rodentia	Chr
Rat	Rodentia	Chr
Rhesus	Primates	Chr
Human	Primates	Chr
Chimp	Primates	Chr

Farré et al., *Genome Research* 2019

DESCHRAMBLER

Cetartiodactyl ancestor (CET)
Pecoran ancestor (PEC)
Bovidae ancestor (BOV)

Pecoran ancestor + FISH (RUM)



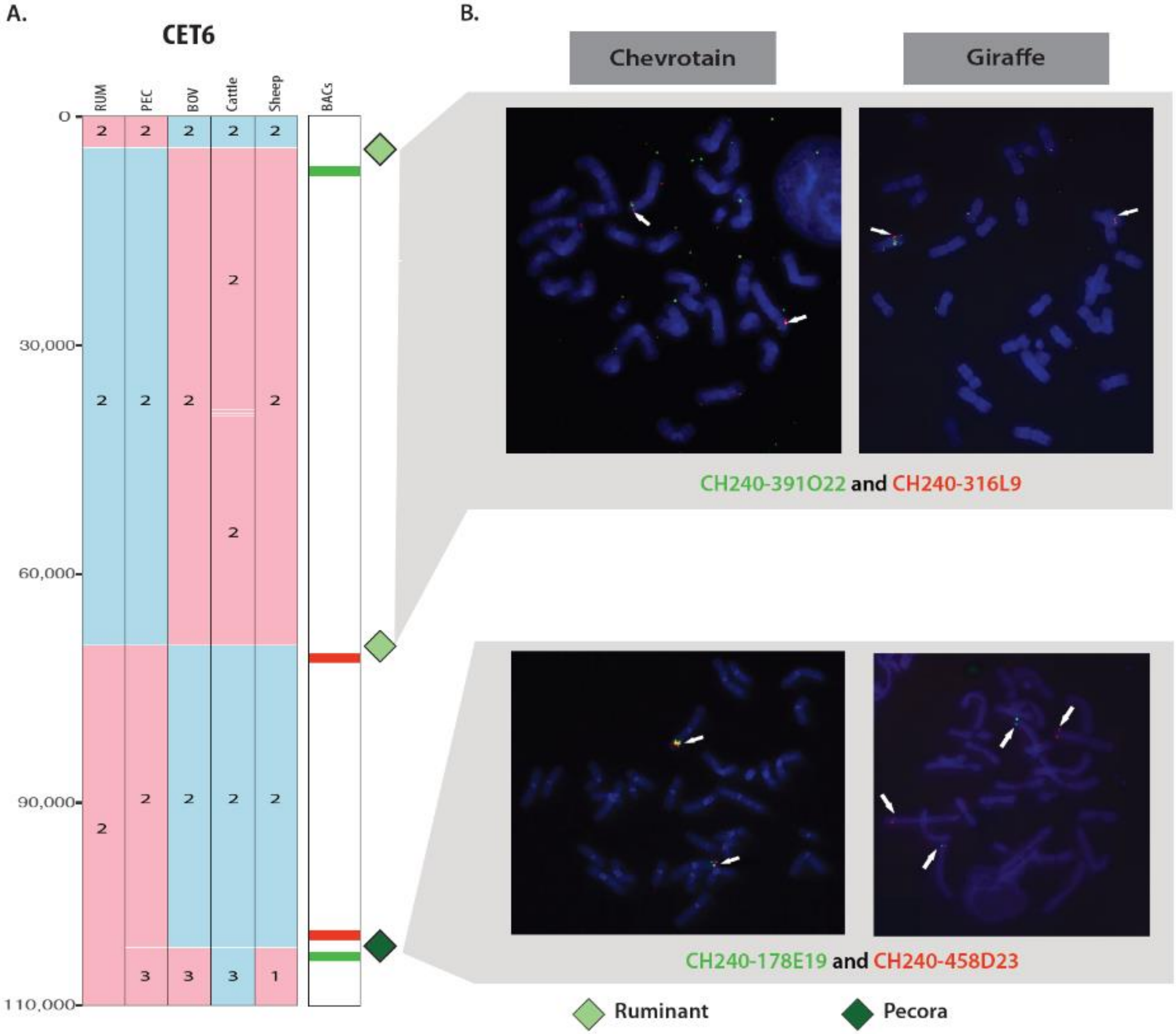
Farré et al., *Genome Research* 2019

RVC Royal Veterinary College University of London

Reconstruction of Ruminant Ancestor Chromosomes

126 cattle BACs were hybridized on chevrotain and giraffe chromosomes

Patten of hybridization allowed distinguishing Pecoran and Ruminant rearrangements



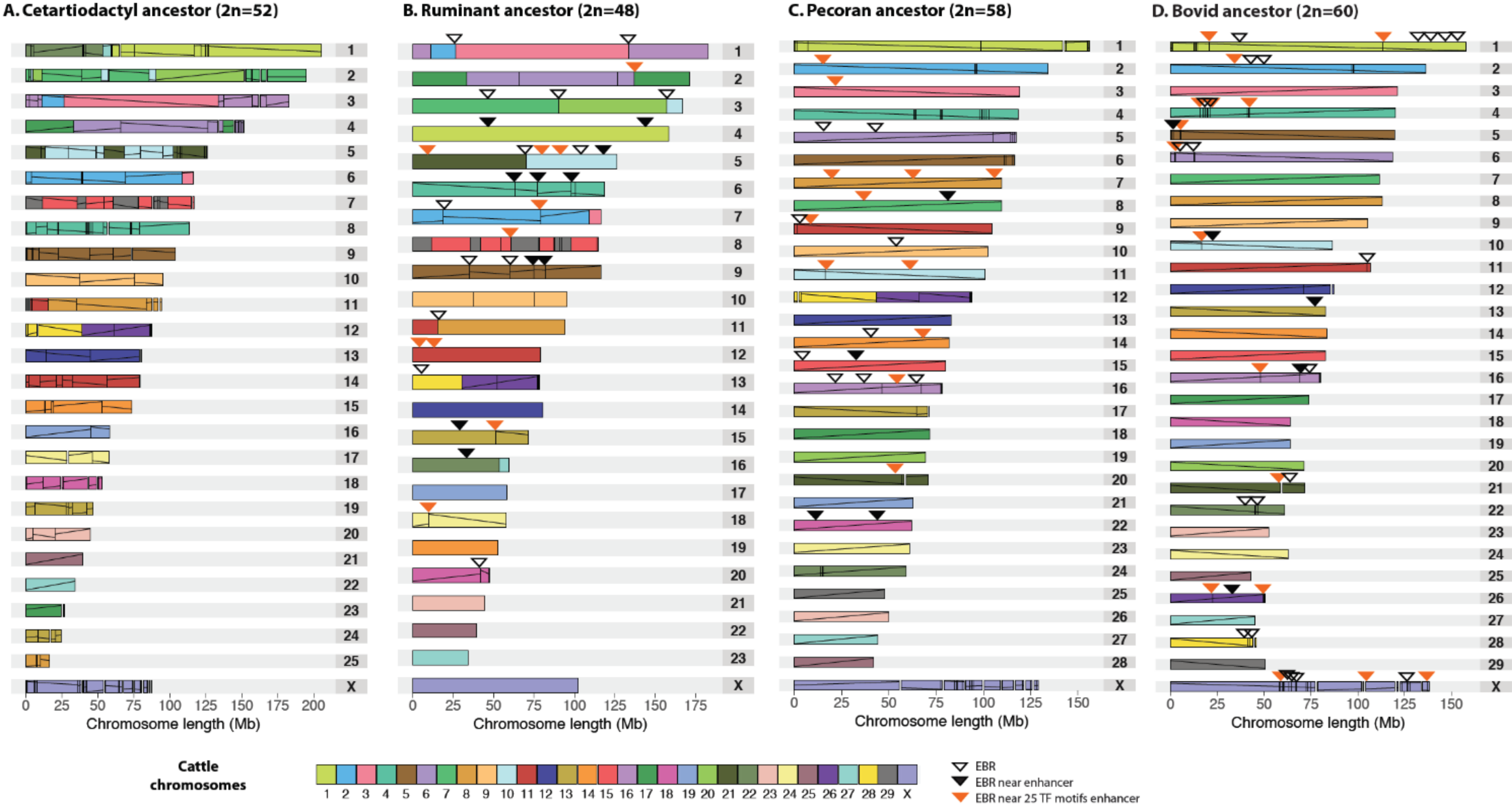
Farré et al., *Genome Research* 2019

Reconstructed Ancestors Statistics

Ancestor	Predicted no. chromosomes	No. of RACFs	Total size (Kbp)	Coverage of cattle genome (%)
Cetartiodactyla	26	57	2,551,909	95.90
Ruminant	24	NA	NA	NA
Pecora	29	40	2,637,874	99.13
Bovidae	30	35	2,644,132	99.37

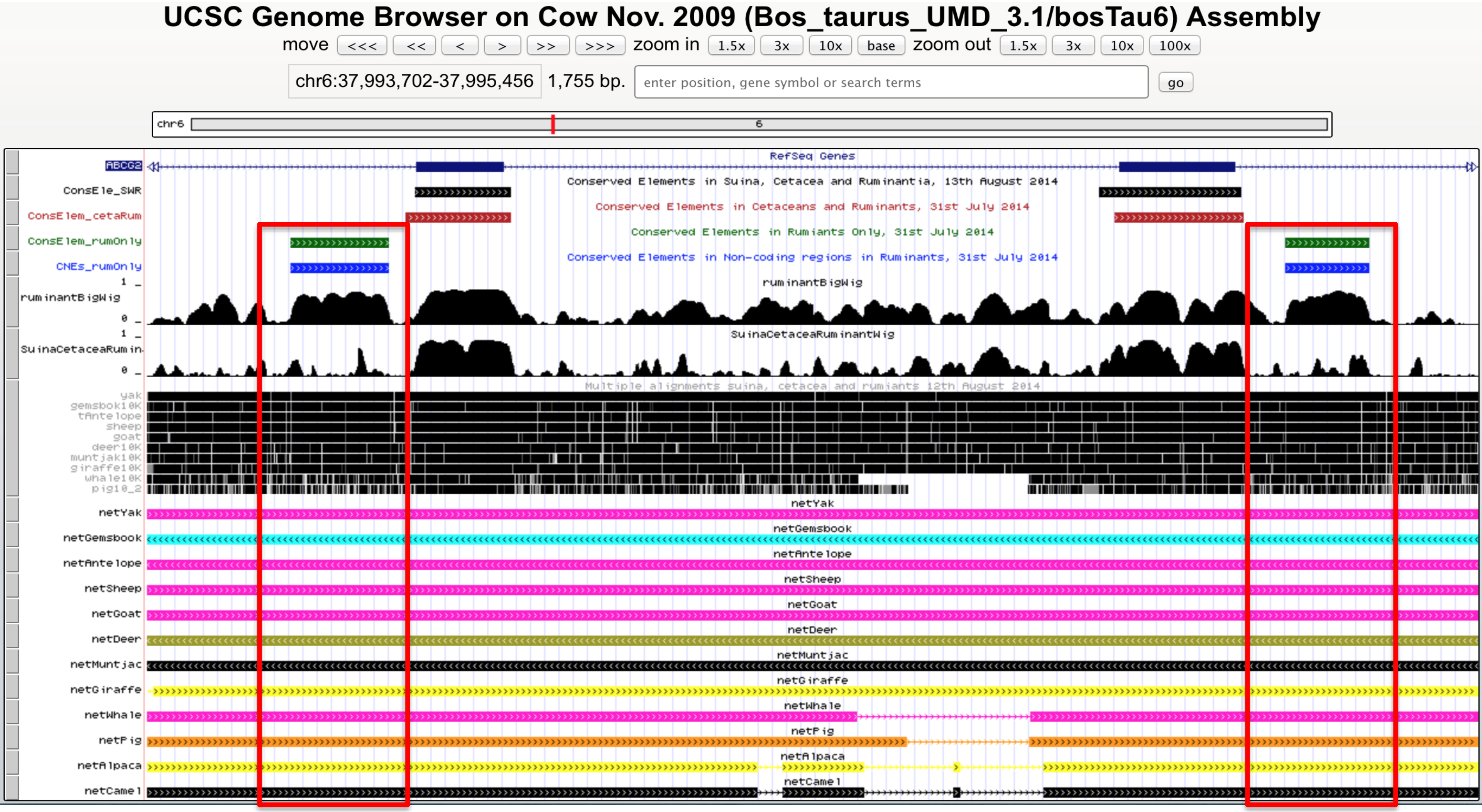
Farré et al., *Genome Research* 2019

Reconstructed Ancestral Karyotypes



Farré et al., *Genome Research* 2019

Ruminant Conserved Non-Coding Elements



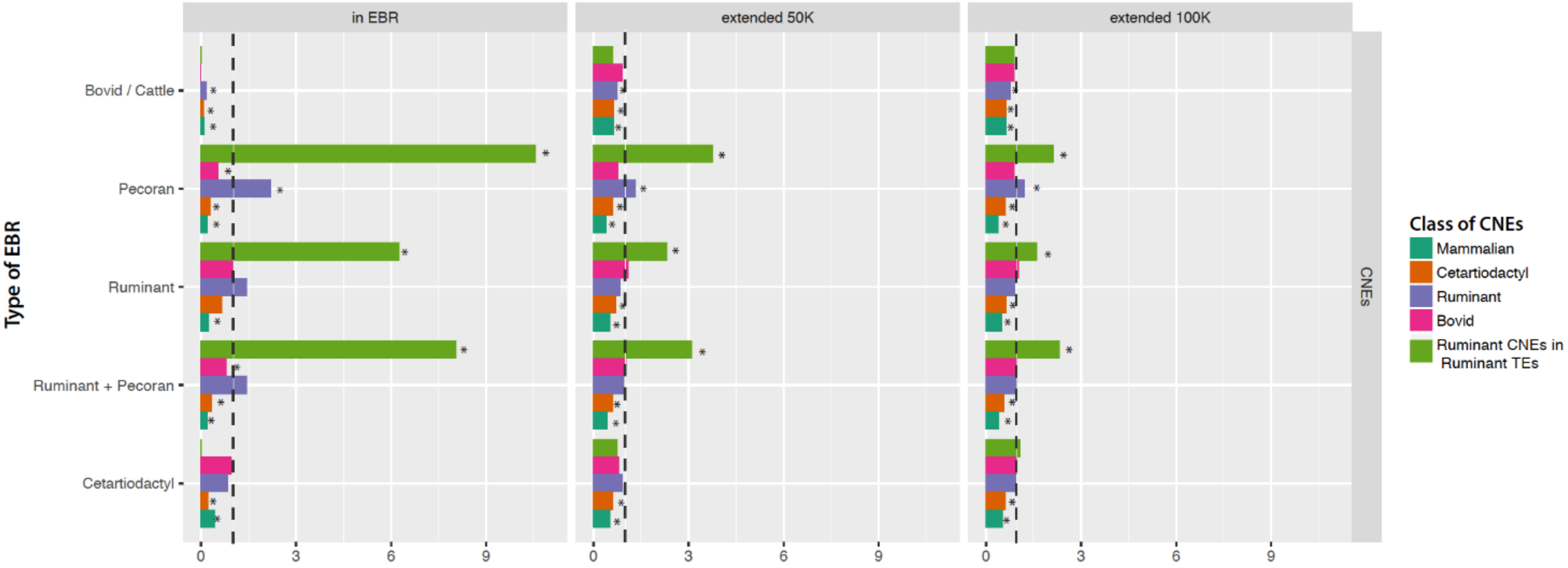
UCSC Genome Browser on Cow Nov. 2009 (Bos_taurus_UMD_3.1/bosTau6) Assembly

chr6:37,993,702-37,995,456 1,755 bp.

Farré et al., *Genome Research* 2019

Royal Veterinary College
University of London

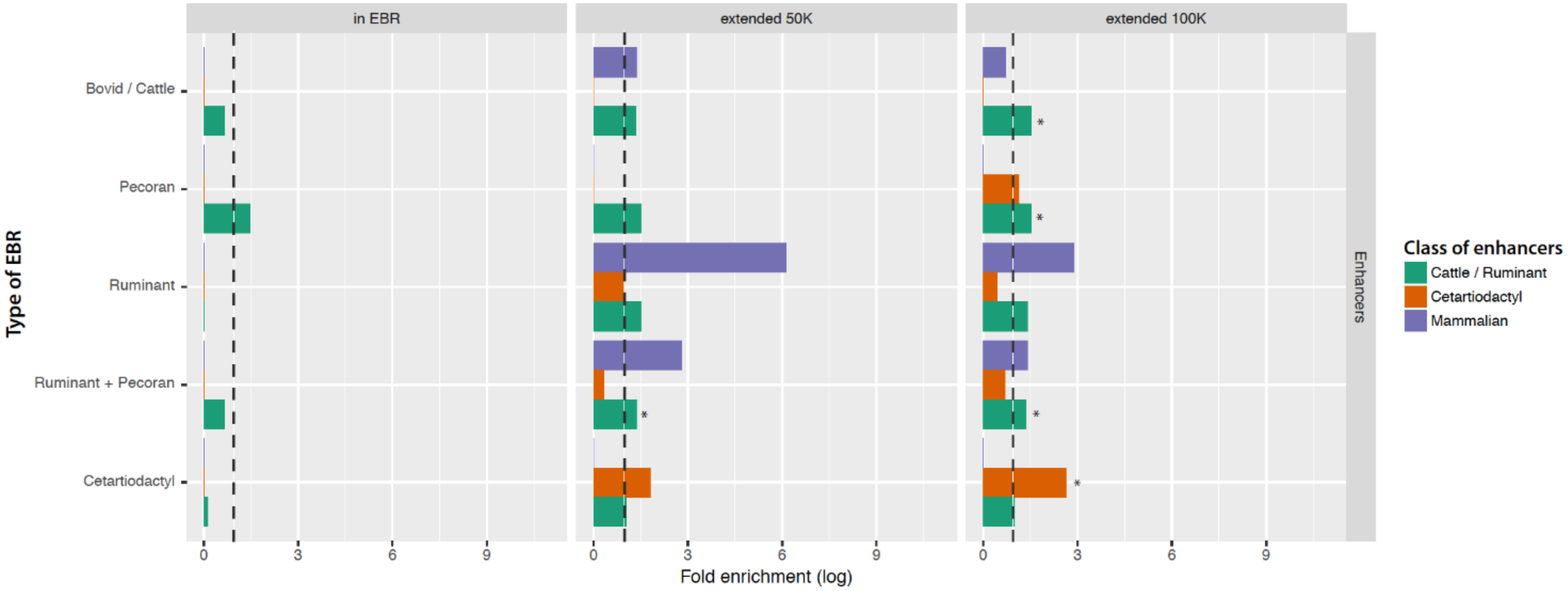
Ruminant Evolutionary Breakpoint Regions are Enriched in CNEs Found in Transposable Elements



Transposable element enrichment ~2 fold
Conserved non-coding element enrichment ~8 fold

Farré et al., *Genome Research* 2019

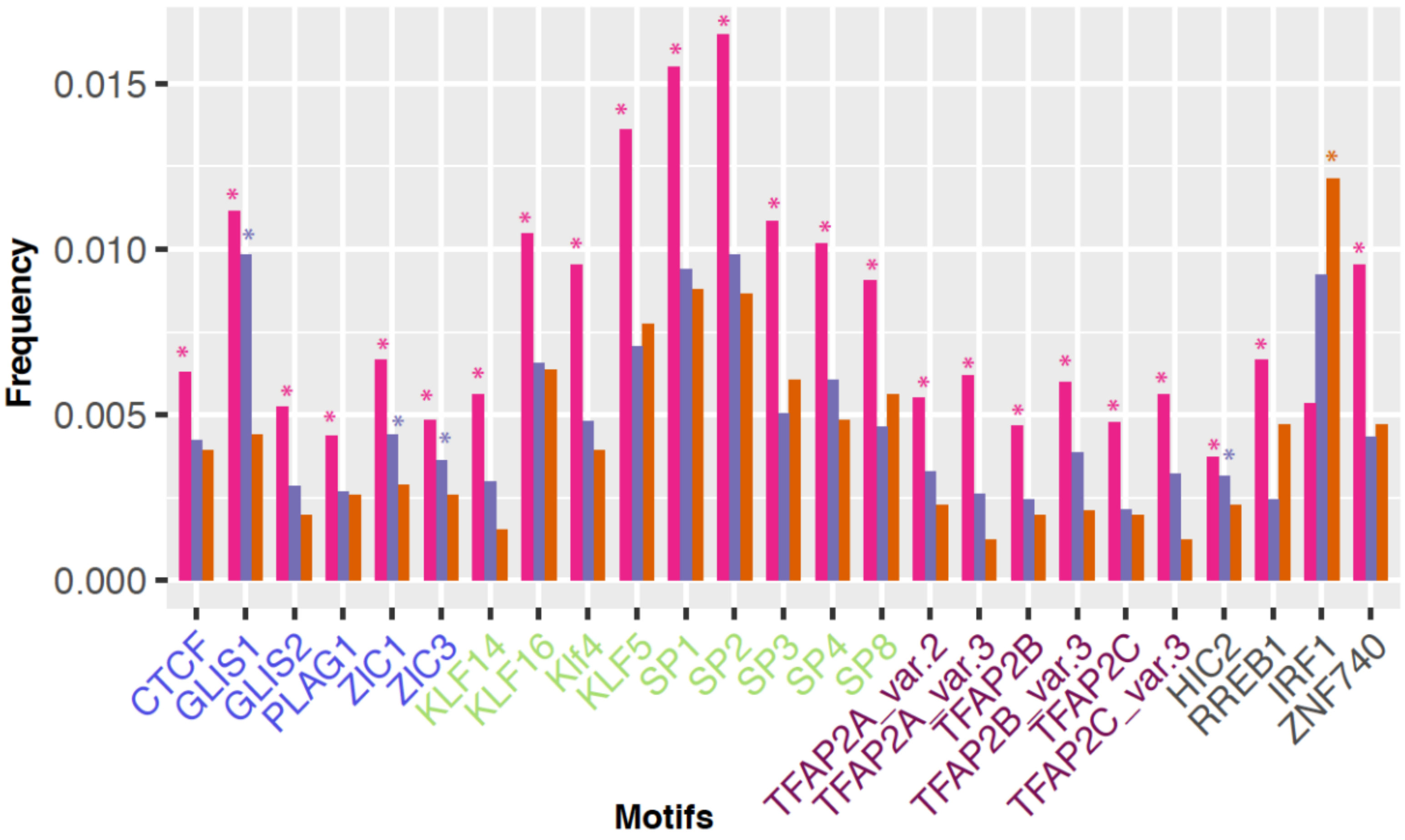
Gene Enhancers in Evolutionary Breakpoint Regions



Enhancer data from Villar et al. 2015
No. enhancers classified 16,100/31,372

Farré et al., *Genome Research* 2019

Transcription Factor Binding Sites Found Enriched in Ruminant Lineage Evolutionary Breakpoint Regions

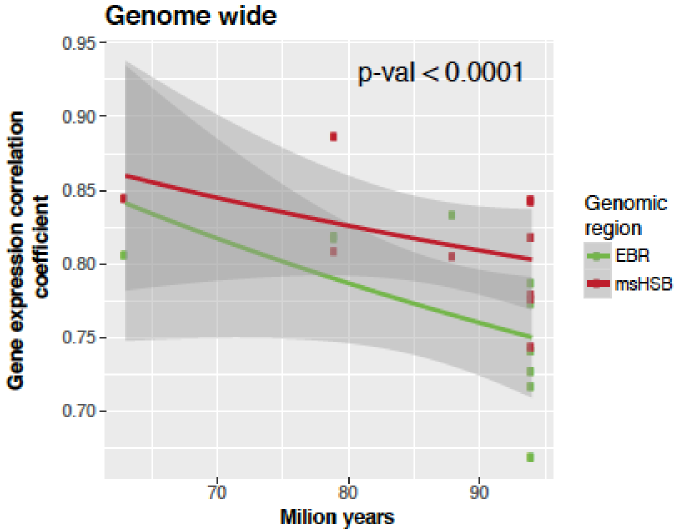


TFBS branch of origin was identified using birth and death model (Yokoyama et al. 2014)
25 TFs are enriched in ruminant-specific transposable elements

Farré et al., *Genome Research* 2019

Comparison of Gene Expression Profiles for Five Mammals

Liver expression data on ~11,000 genes, one2one orthologs (Berthelot et al. 2017)
Five species used: cattle, pig, human, mouse, cat

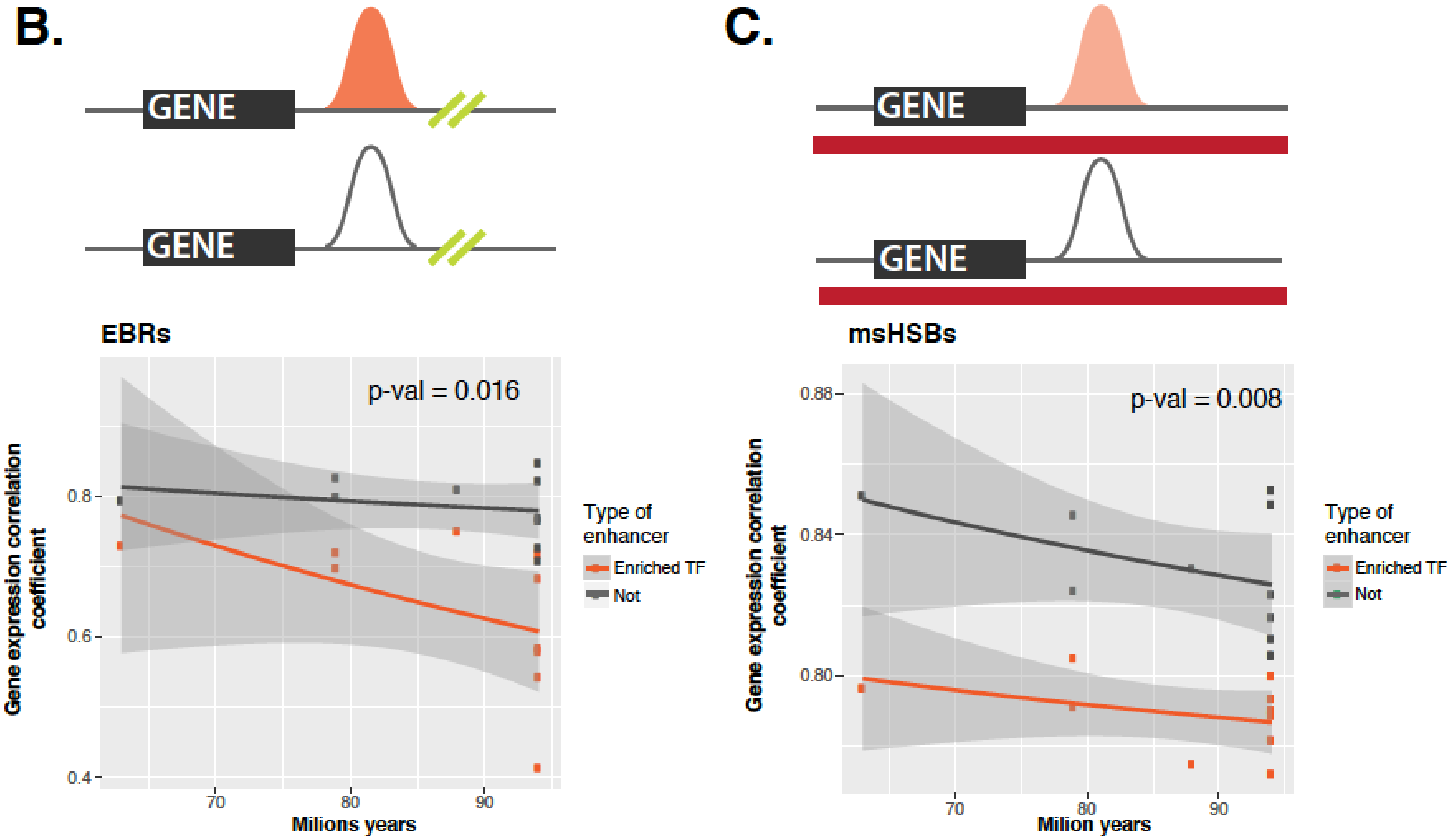


Genes near EBRs have more diverged expression than in HSBs

Farré et al. *Genome Research*, 2019

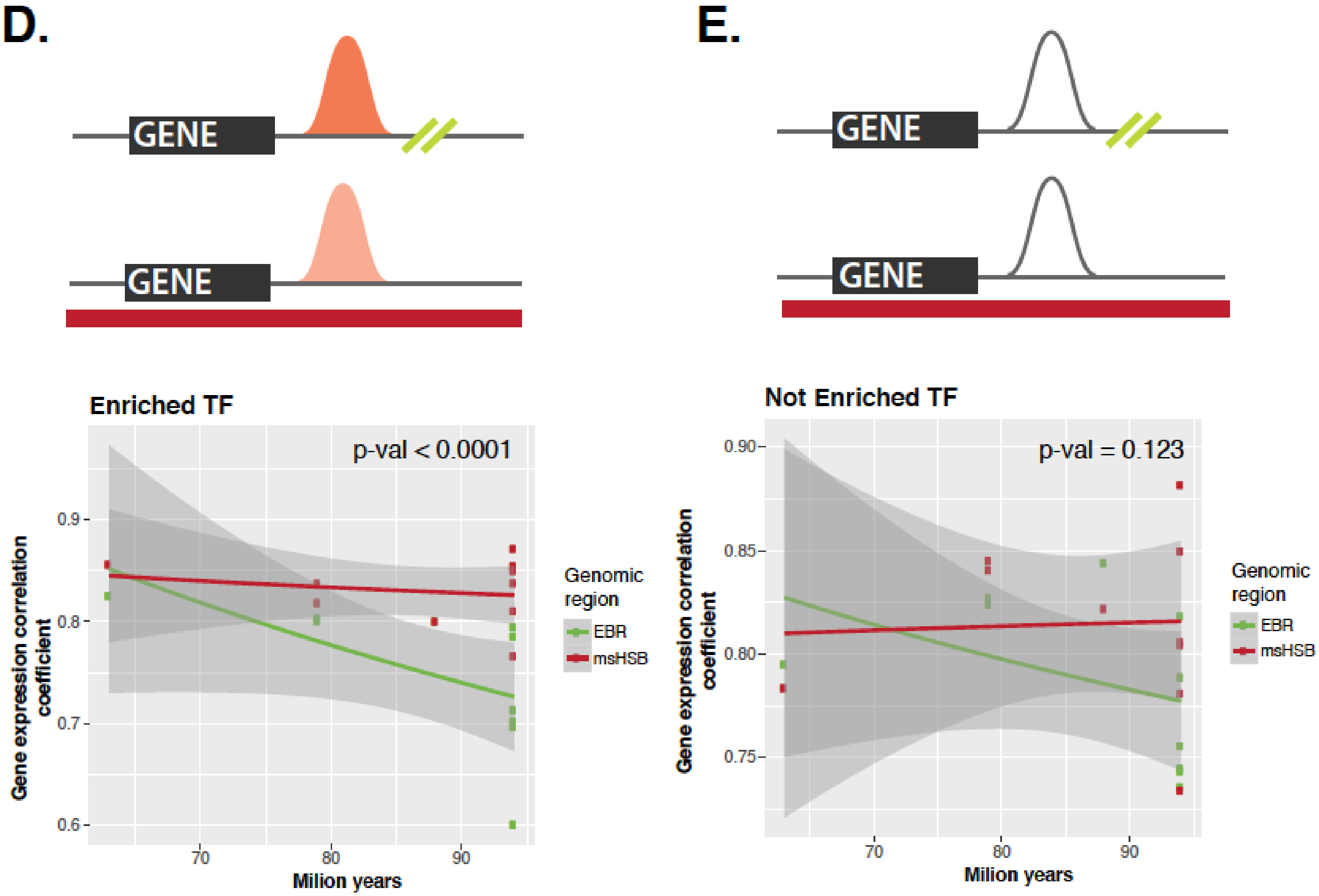
RVC Royal Veterinary College
University of London

Comparison of Gene Expression Profiles for Five Mammals



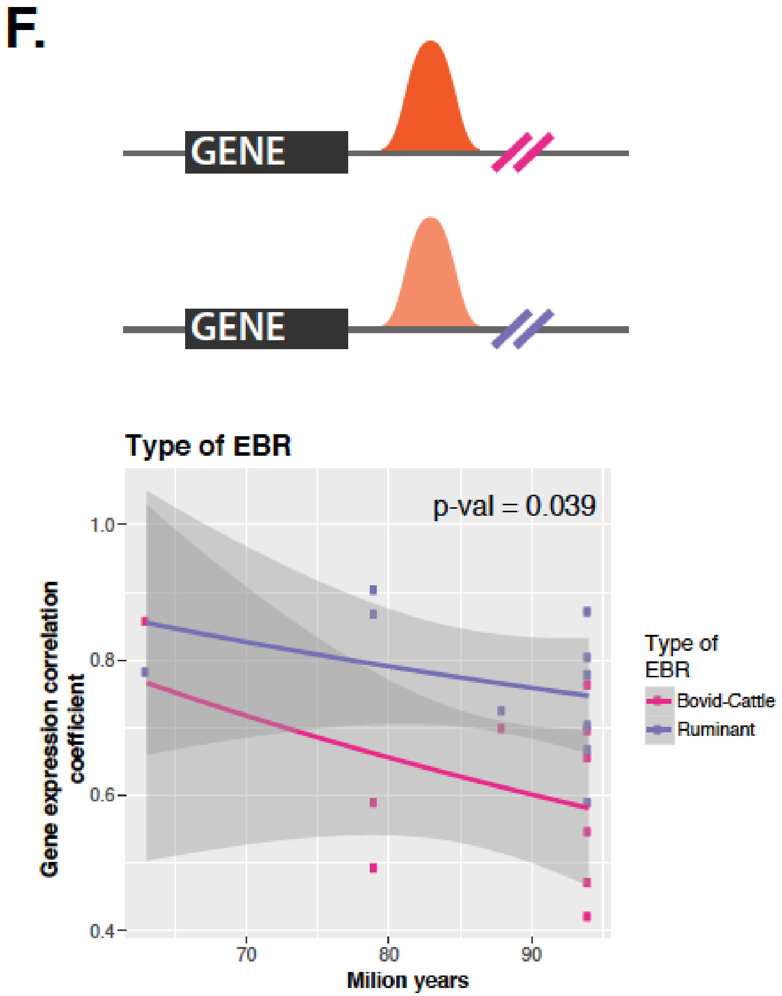
Farré et al. *Genome Research*, 2019

Comparison of Gene Expression Profiles for Five Mammals



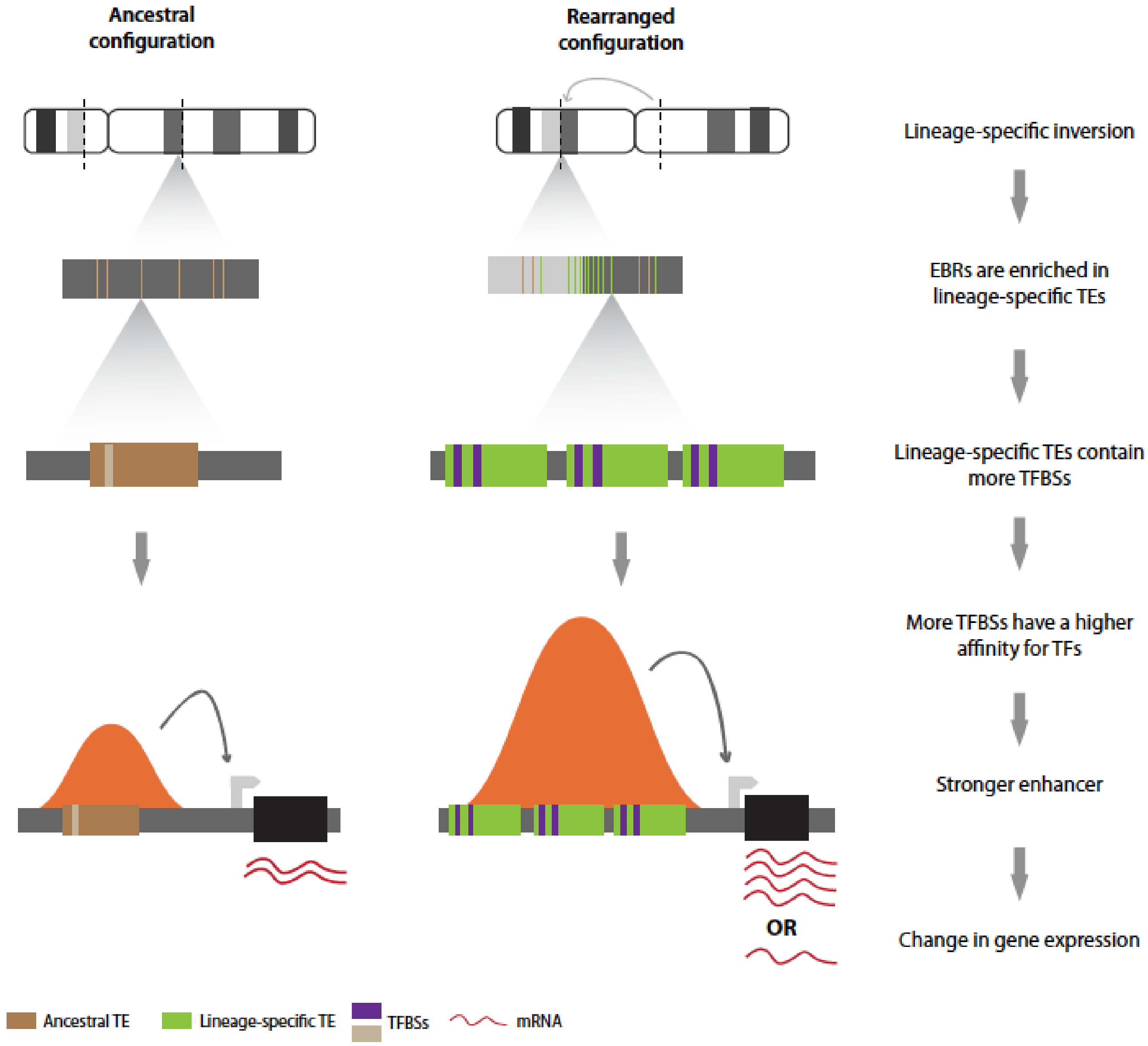
Farré et al. *Genome Research*, 2019

Comparison of Gene Expression Profiles for Five Mammals



Farré et al. *Genome Research*, 2019

Royal Veterinary College
University of London



Lineage-specific inversion

↓

EBRs are enriched in lineage-specific TEs

↓

Lineage-specific TEs contain more TFBSs

↓

More TFBSs have a higher affinity for TFs

↓

Stronger enhancer

↓

Change in gene expression

Farré et al., *Genome Research* 2019

Summary

The ruminant ancestor had 48 chromosomes, 12 less than the cattle genome

Ruminant lineage evolutionary breakpoint regions are enriched for CNEs and TFBSs enriched in ruminant-specific transposable elements

Enhancers containing TFs enriched in ruminant breakpoints are related to changes of gene expression in cattle with a more profound effect observed for the most recent evolutionary breakpoints

Collaborators & Acknowledgements

Royal Veterinary College
University of London

Marta Farré
Joana Damas

KONKUK UNIVERSITY KOREA

Jaebum Kim

ILLINOIS

Anna Kukekova
Jennifer Johnson

Институт молекулярной и клеточной биологии СО РАН

Anastasia Proskurjakova
Anastasia Kulemzina
Polina Perelman
Alexander Graphodatsky

UC DAVIS
UNIVERSITY OF CALIFORNIA

Harris A. Lewin

BBSRC
bioscience for the future

University of Kent

Darren Griffin
Rebecca O'Connor
Mikhail Romanov

CARNEGIE MELLON UNIVERSITY

Jian Ma