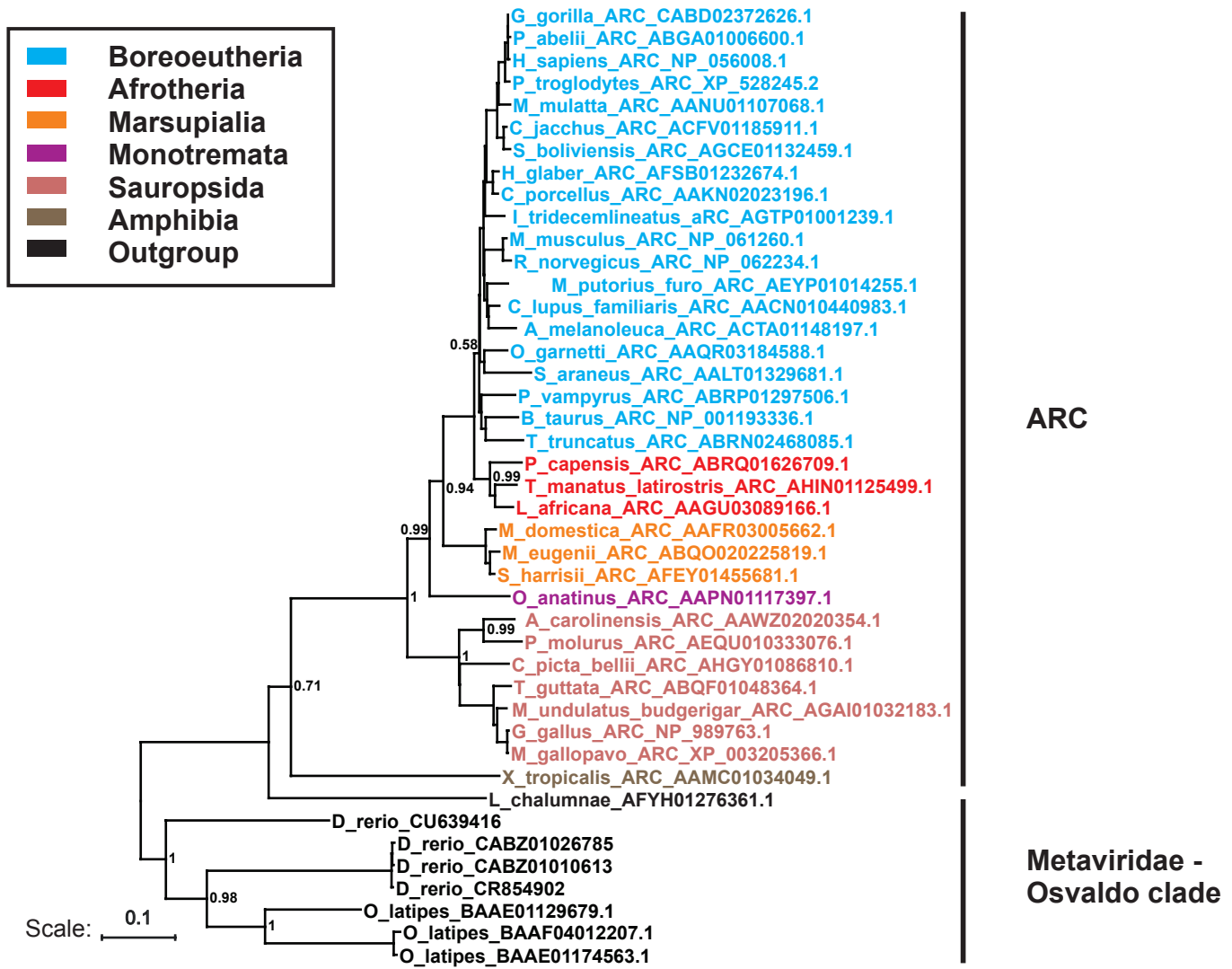
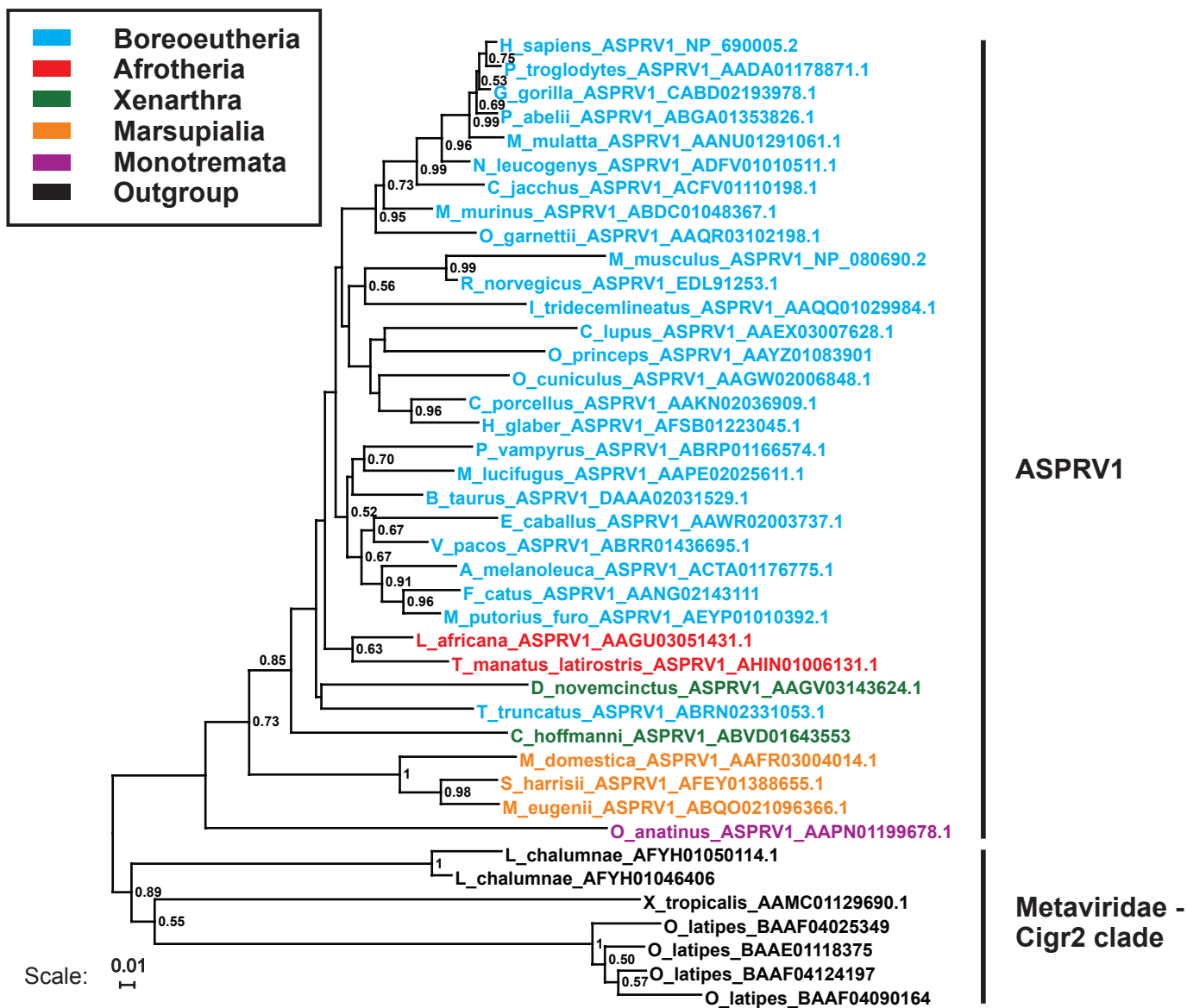


Scale: 0.1

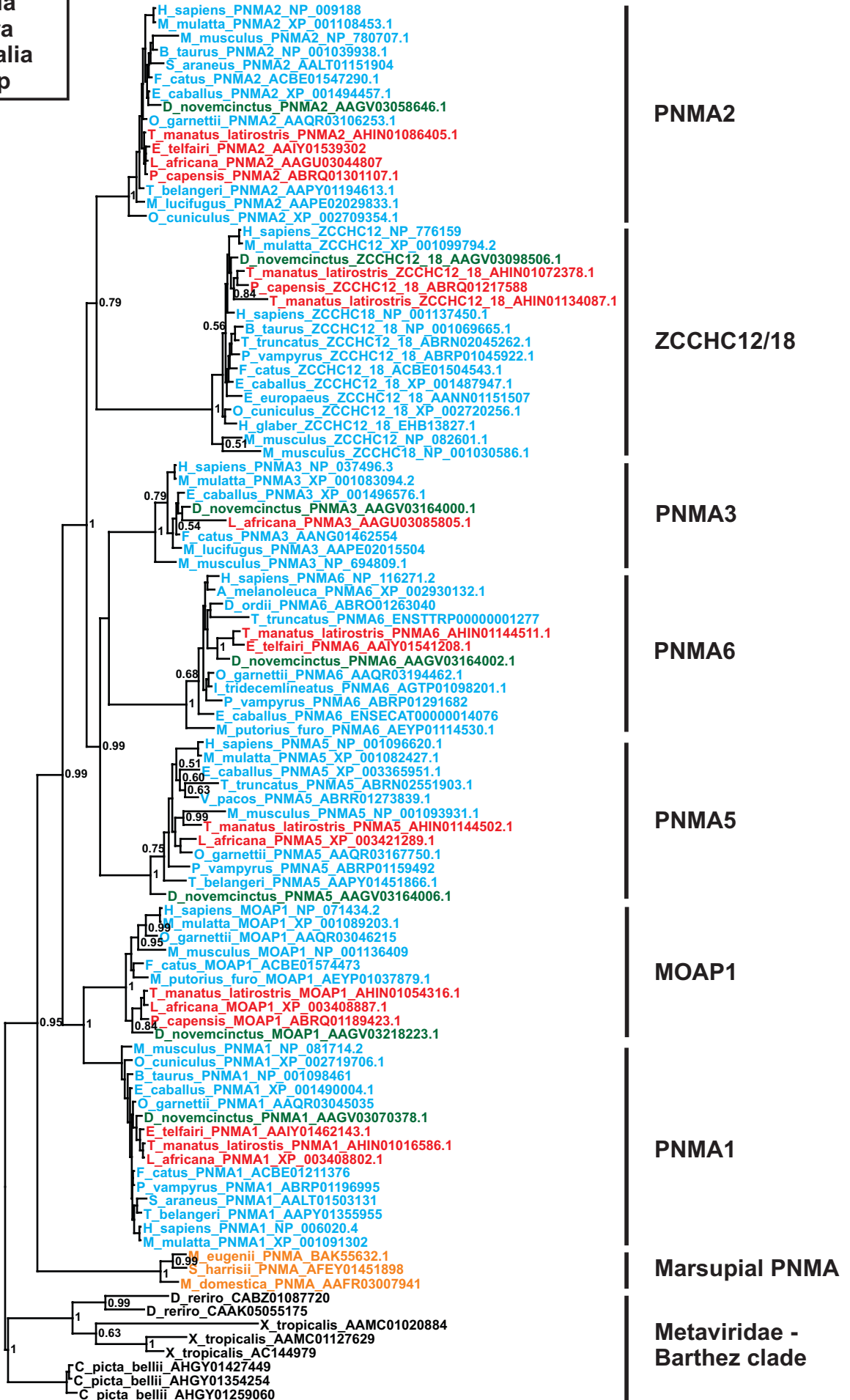
Supplementary Figure S1. ML phylogeny of the PNMA gene family. The maximum likelihood phylogenetic tree was inferred from the capsid part of the gag domain by PhyML 3.0 program under a Poisson+G4 model. Reliability for the internal branches was assessed using 100 bootstrap replications. Bootstrap support values larger than 0.5 are shown. The scale bar corresponds to 0.1 substitutions per site. Capsid part of the gag domain from diverse vertebrate Barthez retroelements was used to root this tree. PNMAL1, PNMAL2 and CCDC8 sequences were omitted from the alignment because they only retain the matrix part of the gag domain.



Supplementary Figure S2. NJ phylogeny of the ARC genes. The rooted neighbor-joining tree was inferred by MEGA 5.0 program under a *p*-distance+pairwise deletion model from the gag domain. Reliability for the internal branches was assessed using the 10000 bootstrap replications. Bootstrap support values larger than 0.5 are shown. The scale bar corresponds to 0.1 substitutions per site. Gag domains from diverse vertebrate Osvaldo retroelements were used to root this tree.



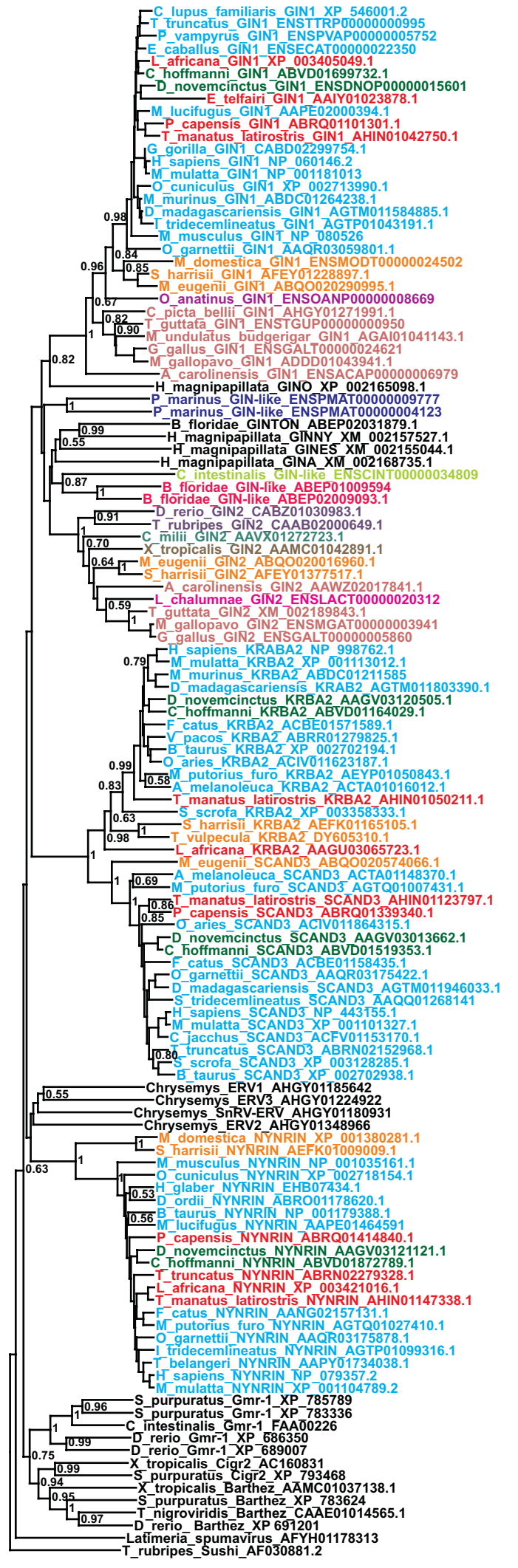
Supplementary Figure S3. NJ phylogeny of the ASPRV1 genes. The rooted neighbor-joining tree was inferred by MEGA 5.0 program under a *p*-distance+pairwise deletion model from the protease domain. Reliability for the internal branches was assessed using the 10000 bootstrap replications. Bootstrap support values larger than 0.5 are shown. The scale bar corresponds to 0.01 substitutions per site. Protease domains from diverse vertebrate Cigr2 retroelements were used to root this tree.



Scale: 0.1

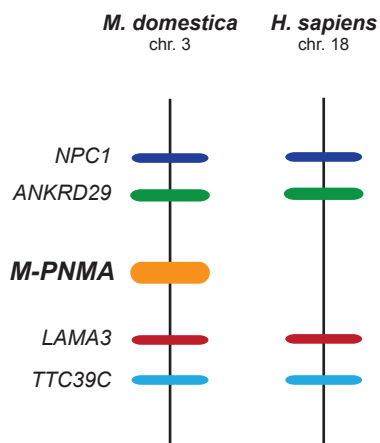
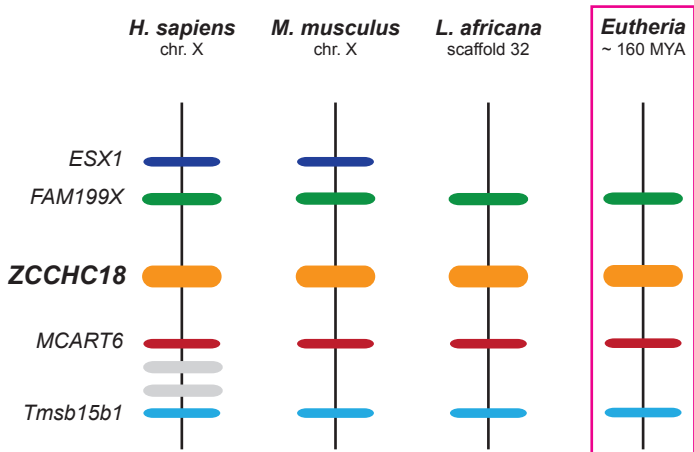
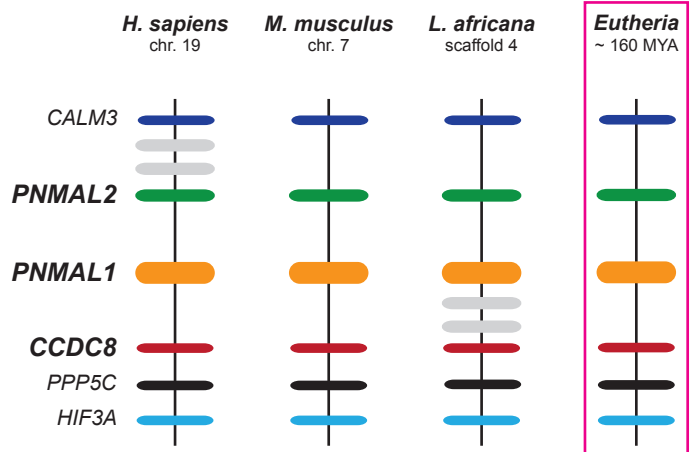
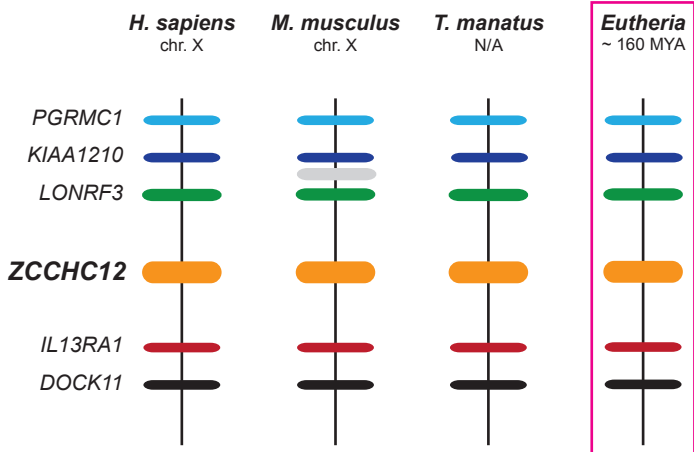
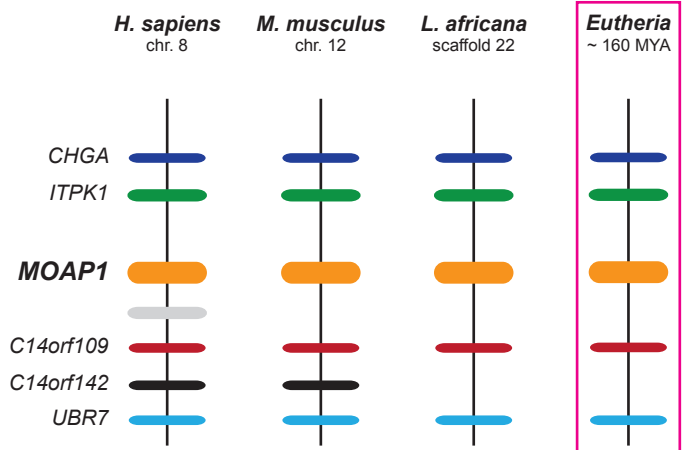
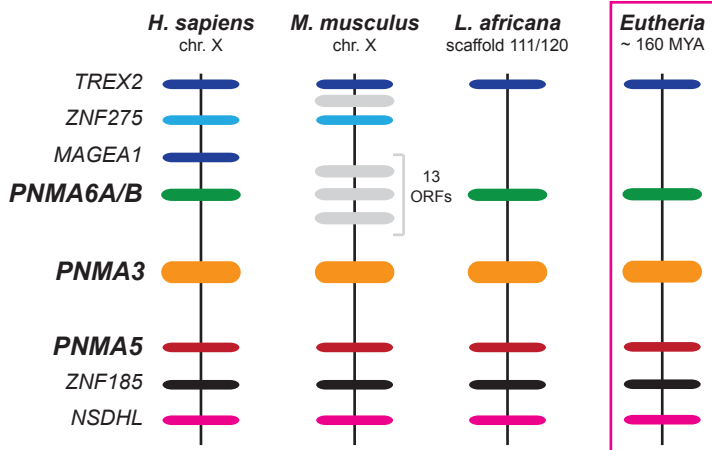
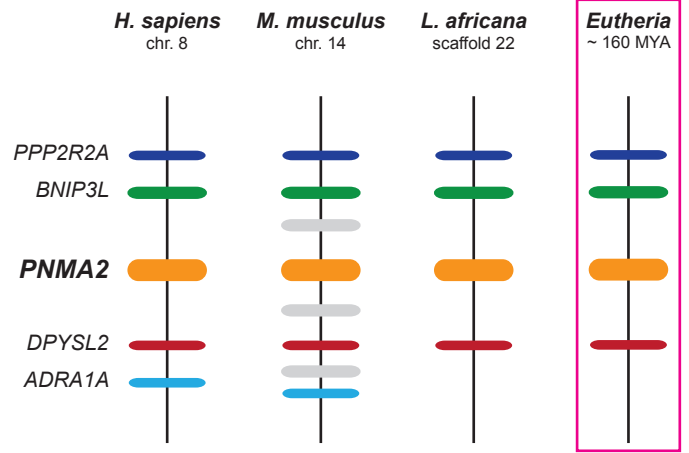
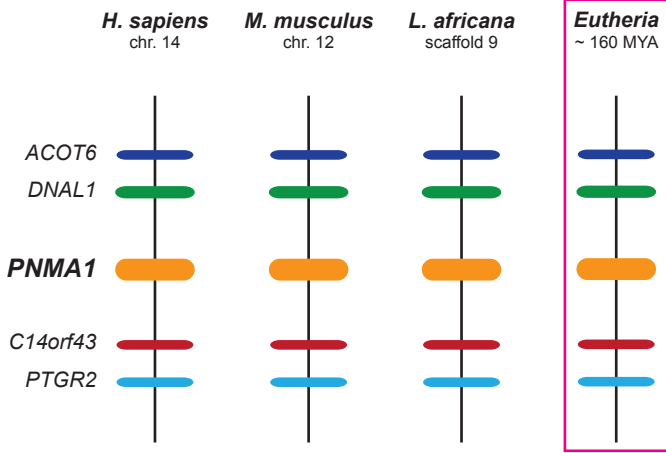
Supplementary Figure S4. Bayesian phylogeny of the PNMA gene family. This tree was inferred from the capsid part of the gag domain by MrBayes 3 program under a JTT+G4 model. Only posterior probabilities larger than 0.5 are shown. The scale bar corresponds to 0.1 substitutions per site. Capsid part of the gag domain from diverse vertebrate Barthez retroelements was used to root this tree. PNMAL1, PNMAL2 and CCDC8 sequences were omitted from the alignment because they only retain the matrix part of the gag domain.

- Boreoeutheria
- Afrotheria
- Xenarthra
- Marsupialia
- Monotremata
- Sauropsida
- Amphibia
- Coelacanthimorpha
- Actinopterygii
- Chondrichthyes
- Cyclostomata
- Urochordata
- Cephalochordata
- Outgroup



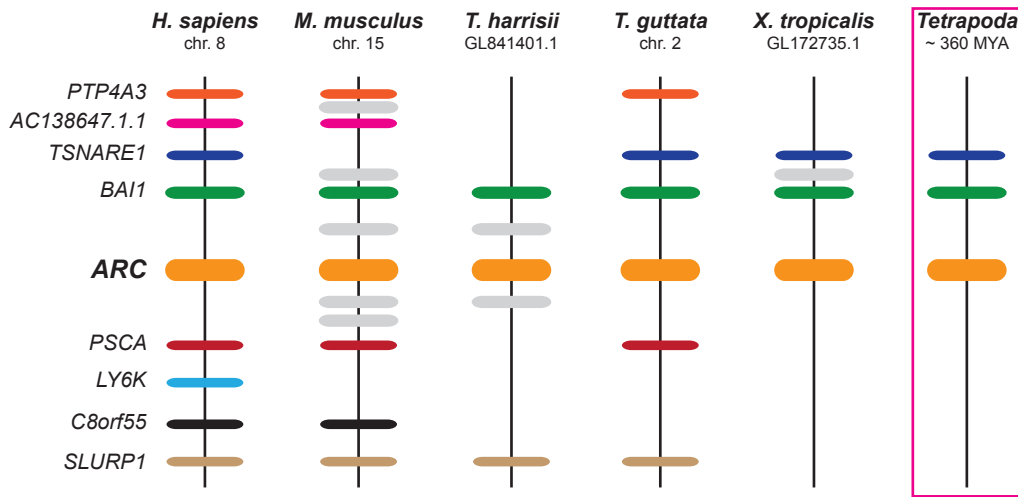
Scale: 0.1

Supplementary Figure S5. NJ phylogeny of the integrase-derived RDDGs. The rooted neighbor-joining tree was inferred by MEGA 5.0 program under a *p*-distance+pairwise deletion model from the integrase domain. Reliability for the internal branches was assessed using the 10000 bootstrap replications. Bootstrap support values larger than 0.5 are shown. The scale bar corresponds to 0.1 substitutions per site. Integrase domains from diverse deuterostome Metaviridae and ERV retroelements were used to root this tree.

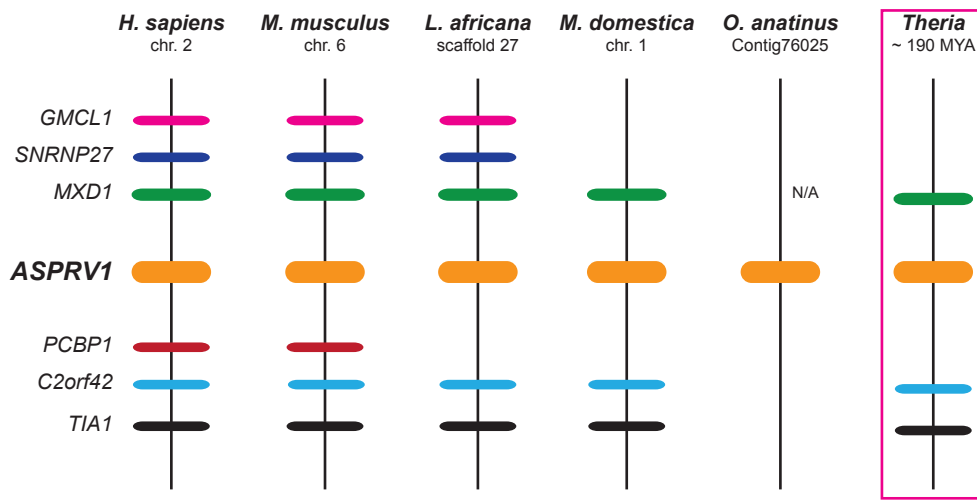


Supplementary Figure S6. Conserved synteny in the PNMA family of RDDGs.

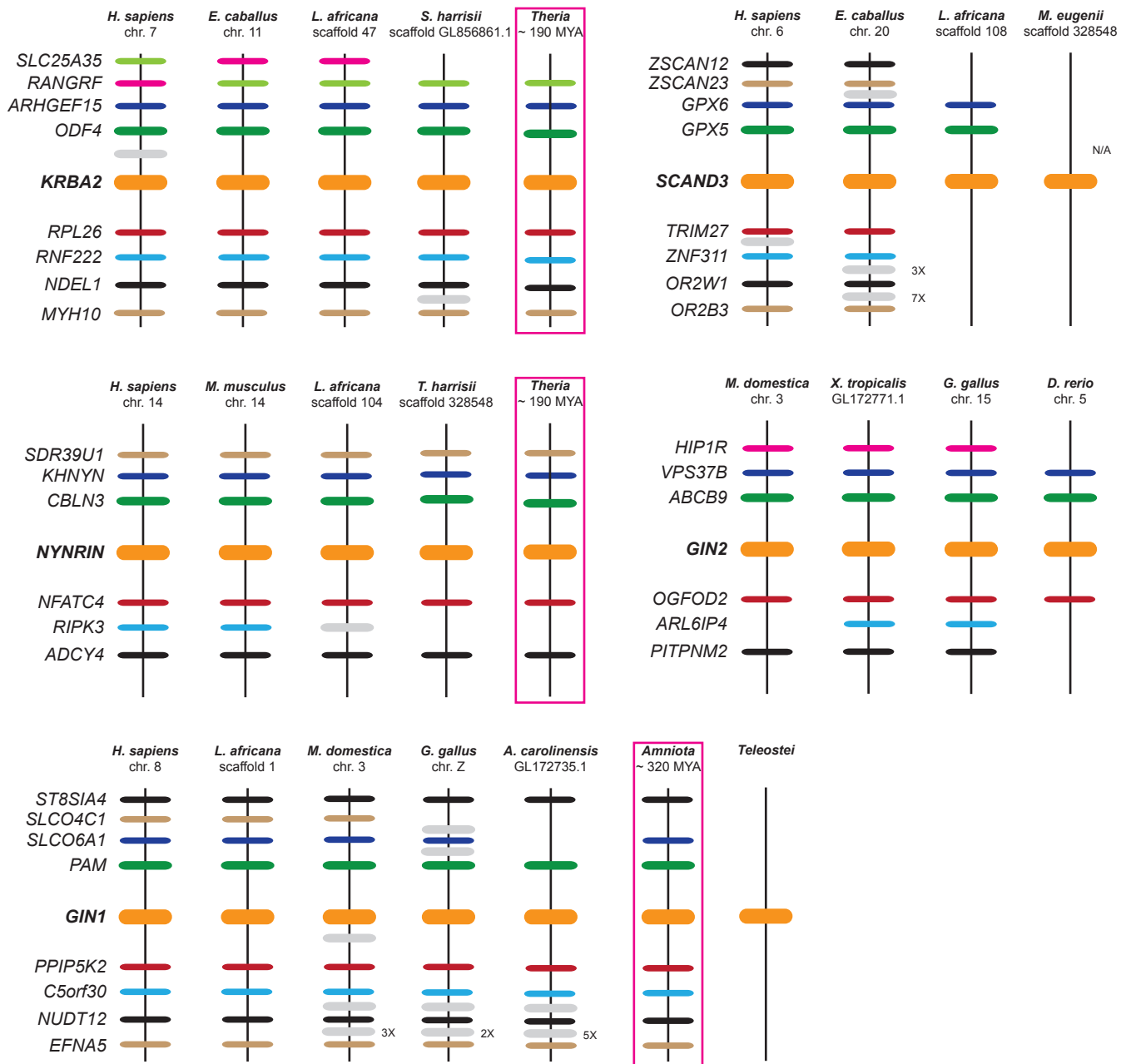
Chromosomal regions carrying all RDDGs in the species considered in this analysis were compared, and neighbouring genes with conserved synteny were identified. Horizontal lines denote orthologous relationships. Each RDDG gene is represented in bold as a horizontal orange line on the chromosome. Neighbouring genes that are in synteny are shown with a schematic indication of their distance (not to scale). The ancestral states of the RDDG chromosomal positions were reconstructed from a comparison of syntenic positions between multiple mammalian lineages.



Supplementary Figure S7. Conserved synteny of the ARC gene. Chromosomal regions carrying ARC gene in the species considered in this analysis were compared, and neighbouring genes with conserved synteny were identified. Horizontal lines denote orthologous relationships. ARC gene is represented in bold as an horizontal orange line on the chromosome. Neighbouring genes that are in synteny are shown with a schematic indication of their distance (not to scale). The ancestral state of the ARC gene chromosomal position was reconstructed from a comparison of syntenic positions between multiple tetrapod lineages.



Supplementary Figure S8. Conserved synteny of the ASPRV1 gene. Chromosomal regions carrying ASPRV1 gene in the species considered in this analysis were compared, and neighbouring genes with conserved synteny were identified. Horizontal lines denote orthologous relationships. ASPRV1 gene is represented in bold as an horizontal orange line on the chromosome. Neighbouring genes that are in synteny are shown with a schematic indication of their distance (not to scale). The ancestral state of the ASPRV1 gene chromosomal position was reconstructed from a comparison of syntenic positions between multiple mammalian lineages.



Supplementary Figure S9. Conserved synteny in the integrase-derived RDDGs.

Chromosomal regions carrying all RDDGs in the species considered in this analysis were compared, and neighbouring genes with conserved synteny were identified. Horizontal lines denote orthologous relationships. Each RDDG gene is represented in bold as a horizontal orange line on the chromosome. Neighbouring genes that are in synteny are shown with a schematic indication of their distance (not to scale). The ancestral states of the RDDG chromosomal positions were reconstructed from a comparison of syntenic positions between multiple mammalian or sauropsid lineages.