## Amniotic CNEs (Dataset iib)









log S

log S

Ż

# Amniotic CNEs on gene-masked genome

0<u></u>

log S



log S 



## Mammalian CNEs (Dataset iia)











## Mammalian CNEs on gene-masked genome







## EU100+ CNEs (Dataset i)













## EU100+ CNEs on gene-masked genome







## Worm CNEs (Dataset ix)



## Human-Zebrafish CNEs (Dataset v)









## Human-Zebrafish CNEs on gene-masked genome



## Human-Fugu CNEs (Dataset iv)

















*Human - Elephant Shark CNEs on hg17: chr17 E* = 2.3, μ = 0.25 2





Human - Elephant Shark CNEs on hg17: chr16  $E = 2.1, \mu = 0.7$ 











## Human-Elephant Shark CNEs on gene-masked genome







## Human-Chicken UCNEs (Dataset iii)







## Insect CNEs (Dataset viii)









log S

## D.rerio PCNEs (Dataset vii)











# Human-Rodents CNEs (Dataset x)







## **EU-FR CNEs (Dataset ia)**









## FR CNEs (Dataset ib)









## **XT-FR CNEs (Dataset ic)**



log S

## **GG-XT-FR CNEs (Dataset id)**



log S

## **EU-GG-XT-FR CNEs (Dataset ie)**



To obtain the figures from the merged and random datasets (adopted from Amniotic and Mammalian datasets as described in the text) shown below, we use the built-in function in R called sample (without replacement). Only the first five chromosomes are depicted :









#### Appendix



6

Additional Simulations using the genomic duplications – CNE loss model

#### Simulations' details

3

4 log S

02

Examples of simulations additional to the ones given in Figure 3 are shown. Initially, 1000 markers (representing CNEs) are randomly inserted in a sequence 2Mnt long. Then, in (a), 173 segmental duplication events occurred, with lengths sampled from a uniform distribution with maximum the 5% of the actual length of the simulated sequence. After each such event a number of CNEs equal to 90% of the number of the duplicated ones are eliminated. The length reached by the simulated chromosome was ~195 Mnt (steps *i* and *ii* of the model). In (b) the segmental duplications were 87 (~50% less) and the final sequence length ~22 Mnt. In (c) the number of segmental duplications is as in (a) and the finally reached length ~119Mnt. Here, 256 additional events of non-duplicated CNEs eliminations are also allowed (step *iii* of the model). These eliminations represent cases of loss of function of ancestral CNEs with a subsequent progressive decomposition. For genomic and for model-generated size distributions, circles and rhombs are used respectively, throughout.