DOI 10.18699/SBB-2022-057

Comparative bioinformatic analysis of Hi-C chromosome-level scaffolds and chromosome painting maps of *Vulpes vulpes* and *Canis familiaris* (Canidae, Carnivora, Mammalia)

Fofanov M.V.[1]*, Prokopov D.Yu.[1], Trifonov V.A.[1], Beklemisheva V.R.[1], Khan R.[2], Aiden E.[2], Graphodatsky A.S.[1], Perelman P.L.[1]*

[1] *Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia*

[2] *Baylor College of Medicine, Houston, TX, USA*

** fofanov@mcb.nsc.ru, polina.perelman@gmail.com*

Key words: canids, rearrangements, chromosome evolution, genome assembly

Motivation and Aim: For the past decades the main source of information about the chromosome-level rearrangements underlying speciation was cross-species chromosome painting. Efficient comparison of chromosomes between species yielded a large bolus of data on chromosome evolution trends and rates for all mammalian orders. Recently, chromosome level genome assemblies became available for a variety of animal species, with the potential to broaden such comparisons including detecting "smaller" chromosomal events, identifying precise breakpoint coordinates, involving species that do not have painting data. There is a need for a translation pipeline so that the data coming from both methods can be used in a mutually beneficial way. Here we attempt to set up the pipeline to evaluate cross-species evolutionary chromosome rearrangements based on chromosome level assemblies.

We chose the family Canidae (wolves, foxes, raccoon dogs etc.) that is characterized by remarkably high rates of evolutionary chromosomal rearrangements compared to other mammalian groups. Among canids multiple breeds of domestic dog (*Canis familiaris*) now have their genomes assembled to chromosome level [1], and red fox (*Vulpes vulpes*) has a pre-publication of chromosome-length genome assembly has become recently available on DNA Zoo (www.dnazoo.org) [2, 3].

Methods and Algorithms: *Vulpes vulpes* (VulVul2.2_HiC) and *Canis familiaris*, golden retriever breed (canFamDis_HiC), assemblies were retrieved from the DNA Zoo website (https://www.dnazoo.org/assemblies). These genome assemblies were soft-masked using RepeatMasker [4] with the following options: -engine hmmer -s -nolow -norna -species canidae. Then chromosome-length scaffolds from these assemblies (17 scaffolds for fox and 39 for dog) were aligned using the LASTZ [5] with several non-default options (--format=maf --notransition --nogapped --seed=match15 --step=10 --hspthresh=100000 --ambiguous=iupac --allocate:traceback=200M). Next, chromosome-level alignments were visualized as dot-plots via D-GENIES [6]. Finally, we compared dot-plots and chromosome maps from the Atlas of Mammalian Chromosomes [7] and visualized them manually in one image.

Results: Chromosome-length scaffolds (17 comprising 96.54% of the fox genome assembly and 39 (89.51% of the dog genome assembly) were selected for the LASTZ. Alignments were visualized as dot-plots and used to build chromosome maps (Fig. 1).

All large genomic rearrangements observed during our bioinformatic analysis were previously detected using FISH with chromosome-specific probes. There were some

small genomic rearrangements involving genome fragments ranging from 1100 to 1600 bp which could represent an artifact of genome data preprocessing. On comparative dot-plots we observed a large number of gaps that could be caused by the use of only chromosome-length scaffolds, evolutionary distance between two species (9–10 million years), and soft-masking of genome assemblies prior to the analysis.

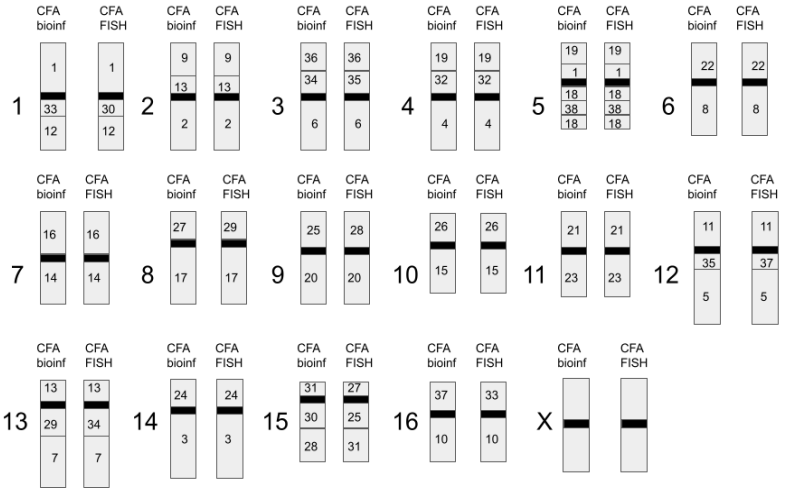

Fig. 1. Comparison of dog-fox chromosome maps derived from Hi-C scaffold alignments (bioinf (*left*)) and reciprocal chromosome painting (FISH (*right*)). Fox chromosome numbers are indicated on the left; corresponding dog (CFA) chromosome segments are indicated inside of chromosome outlines (note painting CFA nomenclature correspondence with the dog reference karyotype is in [8])

Conclusion: The approach of cross-species comparison of chromosome-level scaffolds can be reliably applied to species that do not have chromosome painting maps for the identification of the evolutionary chromosome rearrangements because here we show that all rearrangements previously detected by chromosome painting were also identified in the pairwise analysis of Hi-C scaffolds.

Acknowledgements: This study is supported by RSF grant No. 22-24-01076.

References

1. Field M.A. et al. The Australian dingo is an early offshoot of modern breed dogs. *Sci Adv*. 2022;8(16).
2. Dudchenko O. et al. The Juicebox Assembly Tools module facilitates *de novo* assembly of mammalian genomes with chromosome-length scaffolds for under $1000. *bioRxiv*. 254797.
3. Kukekova A.V. et al. Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. *Nat Ecol Evol*. 2018;2(9):1479-1491.
4. Tempel S. Using and Understanding RepeatMasker. *Methods Mol Biol.* 2012;859:29-51.
5. Schwartz S. et al. Human–Mouse Alignments with BLASTZ. *Genome Res*. 2003;13(1):103-107.
6. Cabanettes F., Klopp C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*. 2018;6:e4958.
7. Graphodatsky A.S., Perelman P.L., O'Brien S.J. Atlas of Mammalian Chromosomes. Wiley, 2020.
8. Breen M Canine cytogenetics – from band to basepair. *Cytogenet Gen Res*. 2008;120(1-2):50-60.