

MASTER BIOINFORMATIQUE

Master Internship offer M2

Internship supervisors and host laboratory

Supervisors: Carina Farah Mugal (CR CNRS, carina.mugal@univ-lyon1.fr), Emmanuelle Lerat (CR CNRS, emmanuelle.lerat@univ-lyon1.fr)

Host laboratory: Laboratoire de Biométrie et Biologie Evolutive (LBBE), 43 boulevard du 11 novembre 1918, Villeurbanne, Team BPGÉ, <https://lbbe-web.univ-lyon1.fr/fr/equipe-bioinformatique-phylogenie-et-genomique-evolutive>

Research project title

The role of transposable elements during speciation in Ficedula flycatchers

Project description

Speciation is a central evolutionary process that can be described as the formation of reproductively isolated species, where the build-up of reproductive barriers can originate from different routes (Figure 1). Reproductive isolation can be initiated by divergent selection on ecological traits or mating preferences, or originate from incompatibilities between two or more loci that interact with each other, for example during the formation of the zygote or also in the genomic background of the hybrid (Presgraves 2010; Seehausen et al. 2014). In order to establish the route of reproductive isolation and improve our understanding of speciation, one of the major objectives in speciation research has therefore been directed towards pinpointing the genetic loci that encode the reproductive barriers. Two classical candidates that have received much attention are protein-coding sequences and their regulatory elements. The role of transposable elements (TEs) in speciation remains on the other hand less explored.

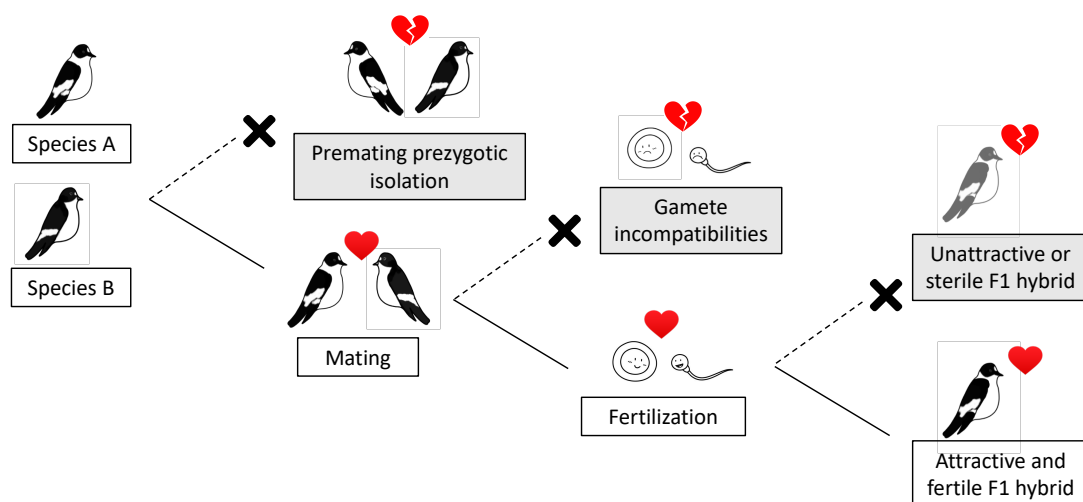


Figure 1: Schematic illustration of different sources of reproductive isolation, i) premating prezygotic isolation, which e.g. can result from habitat isolation or assortative mating, ii) postmating prezygotic isolation, such as gamete incompatibilities or cryptic sperm selection, and iii) reduced hybrid fitness due to either extrinsic environmental factors or intrinsic genomic incompatibilities.

The project will focus on the role of TE activity in the genomes of two closely related migratory birds, the collared flycatchers (*Ficedula albicollis*) and pied flycatchers (*Ficedula hypoleuca*). The collared and pied flycatchers represent a classic case of allopatric speciation with secondary contact that has proceeded to an advanced stage with many sources of reproductive isolation despite their recent divergence (Alund et al. 2024). Speciation genomic studies have revealed a role of recombination and chromosomal rearrangements in shaping the genomic differentiation landscape (Ellegren et al. 2012; Burri et al. 2015; Chase et al. 2021; Chase et al. 2023; Chase et al. 2024b, a), where recent TE activity might significantly contribute to chromosomal rearrangements observed in our study system. In order to test this hypothesis, we have generated long-read sequencing data across multiple individuals from both species (Chase et al. 2023), which permits the construction of a pangenome graph and the cartography of TE polymorphism within a given species and between the two species (Sierra and Durbin 2024). The student will perform annotation of TEs and use a pangenome approach to characterize recent TE activity. In the long run, the project can be continued as a PhD project and benefit from the availability of a rich resource of transcriptomic and epigenomic data to further investigate the role of TE activity during speciation in *Ficedula* flycatchers.

Specific tasks

- Construction of a repeat sequence library / transposable element (TE) library for *Ficedula* flycatchers
- Annotation of TEs in the pangenome of collared and pied flycatchers
- Cartography of TE polymorphism within a given species and between the two species

As part of the project work, the student will develop skills in handling high-throughput sequencing data, the implementation of bioinformatics workflows and statistical data analysis with R. The student will get familiar with computational methods commonly applied in the field (RepeatModeler2 (Flynn et al. 2020), RepeatMasker (<https://www.repeatmasker.org>), pantera (Sierra and Durbin 2024)), will gain a sound understanding of pangenome graphs and the role of TEs in assemblies, which are central to the field of evolutionary genetics.

Expected profile and skills of the candidate

The candidate should have a strong interest in evolutionary genomics and ecology; should be willing to acquire an extensive training in bioinformatics and statistical data analysis; should have good communication skills in English and enjoy collaborative work.

Relevant publications from the team

- Chase MA, Scofield DG, Kraft F, Segami JC, Alund M, Qvarnstrom A, Wheatcroft D, Mugal CF. 2023. The combination of HiFi and HiC sequencing technologies enables the investigation of structural variants in speciation of *Ficedula* flycatchers. *unpublished PhD thesis manuscript*.
- Chase MA, Vilcot M, Mugal CF. 2024b. The role of recombination dynamics in shaping signatures of direct and indirect selection across the *Ficedula* flycatcher genome. *Proc Biol Sci* 291:20232382. doi: 10.1098/rspb.2023.2382
- Lerat E, Fablet M, Modolo L, Lopez-Maestre H, and Vieira C. 2017. TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. *Nucleic Acids Res* 45: gkw953. doi:10.1093/nar/gkw953
- Lerat E. 2022. Recent bioinformatic progress to identify epigenetic changes associated to transposable elements. *Frontiers in Genetics* 13:891194.

References

- Alund M, Marzal JCS, Zhu Y, Menon PNK, Jones W, Qvarnstrom A. 2024. Tracking hybrid viability across life stages in a natural avian contact zone. *Evolution* 78:267-283. doi: 10.1093/evolut/qpad204
- Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, Smeds L, Suh A, Dutoit L, Bures S, Garamszegi LZ, et al. 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Res* 25:1656-1665. doi: 10.1101/gr.196485.115
- Chase MA, Ellegren H, Mugal CF. 2021. Positive selection plays a major role in shaping signatures of differentiation across the genomic landscape of two independent *Ficedula* flycatcher species pairs. *Evolution* 75:2179-2196. doi: 10.1111/evo.14234
- Chase MA, Scofield DG, Kraft F, Segami JC, Alund M, Qvarnstrom A, Wheatcroft D, Mugal CF. 2023. The combination of HiFi and HiC sequencing technologies enables the investigation of structural variants in speciation of *Ficedula* flycatchers. *unpublished PhD thesis manuscript*.
- Chase MA, Vilcot M, Mugal CF. 2024a. Evidence that genetic drift not adaptation drives fast-Z and large-Z effects in *Ficedula* flycatchers. *Mol Ecol*:e17262. doi: 10.1111/mec.17262
- Chase MA, Vilcot M, Mugal CF. 2024b. The role of recombination dynamics in shaping signatures of direct and indirect selection across the *Ficedula* flycatcher genome. *Proc Biol Sci* 291:20232382. doi: 10.1098/rspb.2023.2382
- Ellegren H, Smeds L, Burri R, Olason PI, Backstrom N, Kawakami T, Kunstner A, Makinen H, Nadachowska-Brzyska K, Qvarnstrom A, et al. 2012. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* 491:756-760. doi: 10.1038/nature11584
- Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A* 117:9451-9457. doi: 10.1073/pnas.1921046117
- Presgraves DC. 2010. The molecular evolutionary basis of species formation. *Nat Rev Genet* 11:175-180. doi: 10.1038/nrg2718
- Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, Peichel CL, Saetre GP, Bank C, Brannstrom A, et al. 2014. Genomics and the origin of species. *Nat Rev Genet* 15:176-192. doi: 10.1038/nrg3644
- Sierra P, Durbin R. 2024. Identification of transposable element families from pangenome polymorphisms. *Mob DNA* 15:13. doi: 10.1186/s13100-024-00323-y