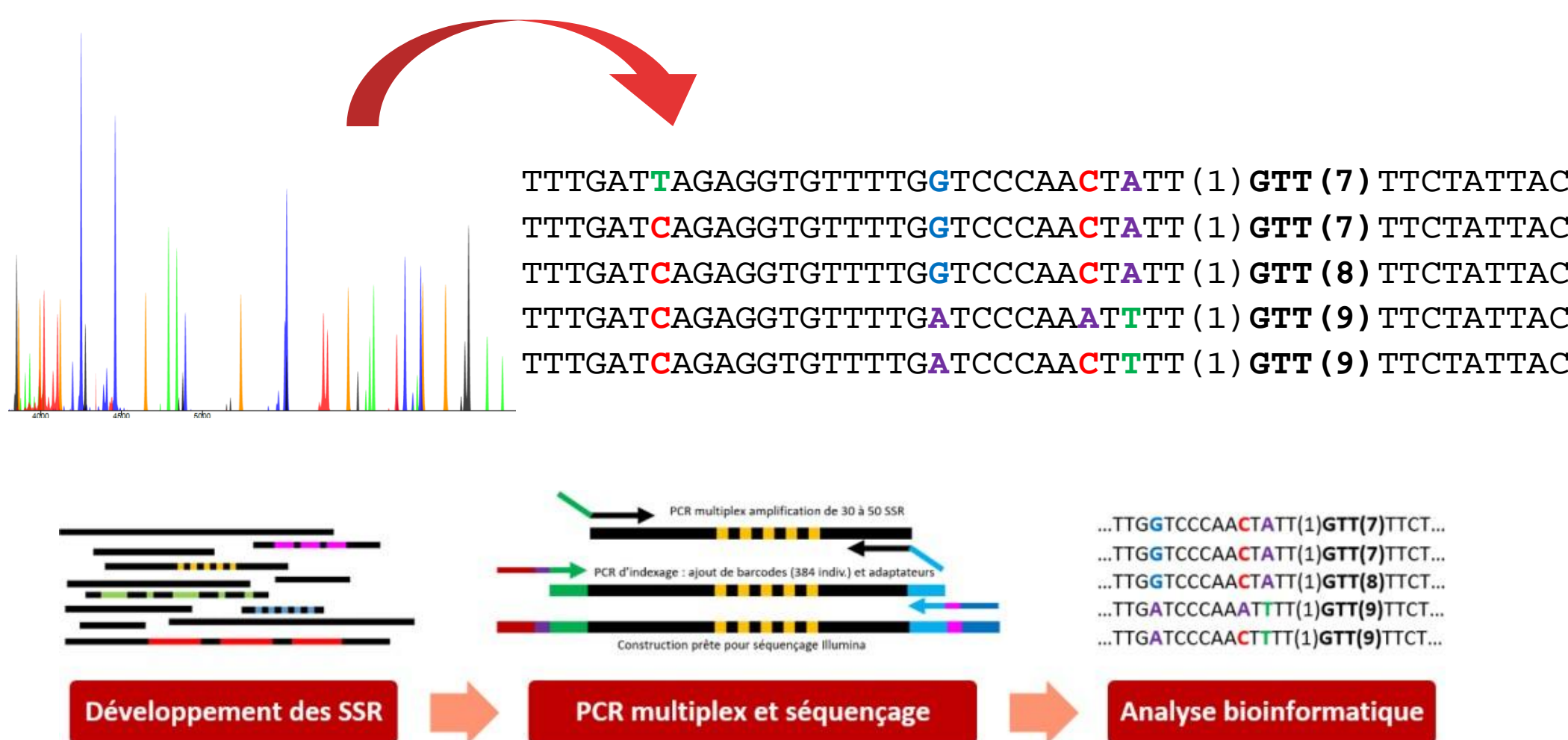


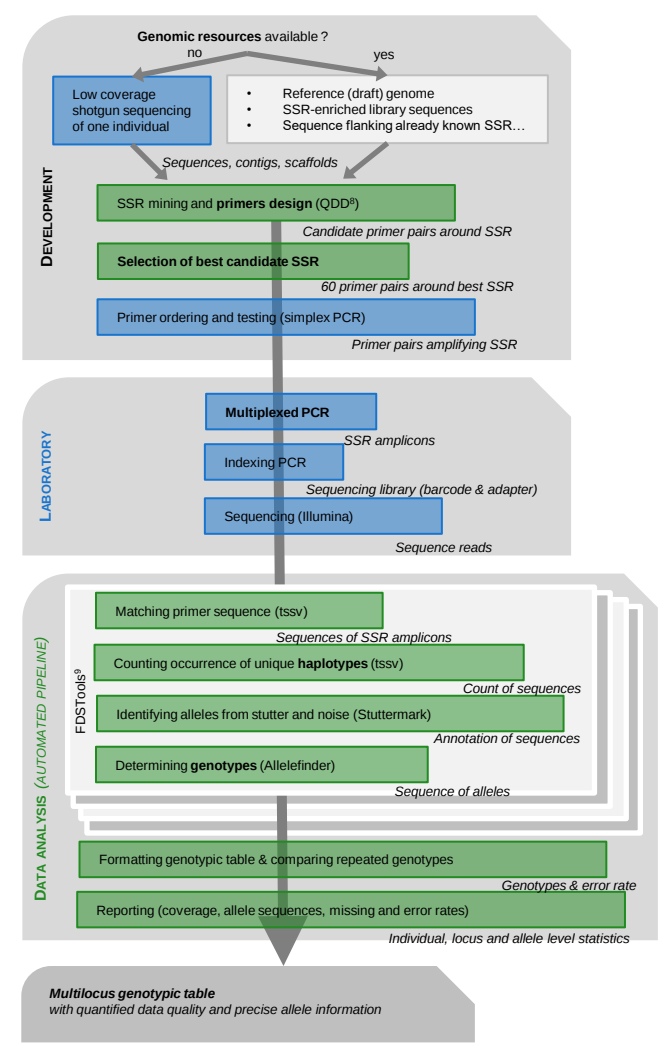
Feedback on sequence-based microsatellite genotyping (SSRseq) : easy access to high quality genetic data.

Olivier Lepais, Emilie Chancerel, Zoé Compagnie, Christophe Boury, Adline Delcamp, Erwan Guichoux
Plateforme Génome Transcriptome de Bordeaux (PGTB), BIOGECO, INRAE, Univ. Bordeaux

RATIONALE & OBJECTIVE

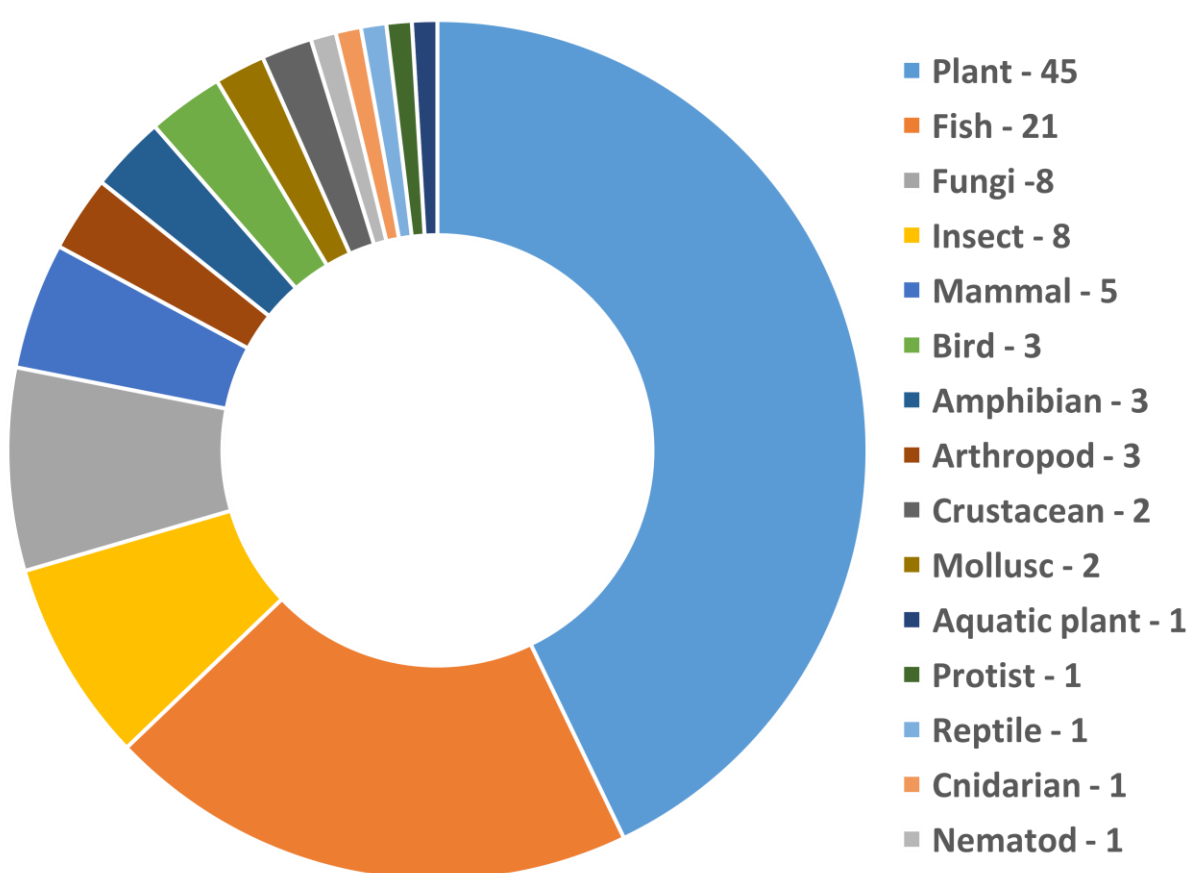


- **Microsatellite (SSR) is a unique kind of polymorphism**
 - 10-15 highly polymorphic SSR powerful enough for many applications in ecology, evol. biol., agronomy
 - Ease of development^{1,2} specially in orphan species, less subject to ascertainment bias than SNP
 - Linked to phenotypic traits^{3,4} (through gene expression regulation, DNA & protein conformation...)
 - But many drawbacks due to the use of electrophoreses for genotyping⁵
 - **Challenge:** transitioning to sequence-based genotyping^{6,7}
 - We developed a **workflow** for SSR development, multiplexed amplification and sequencing and automated bioinformatics data analysis (Lepais et al. 2020)⁸
 - **The objective is to summarize the characteristics of the developed genotypic data and illustrate the gain of information provided by the sequence information.**



EASE OF DEVELOPMENTS

- **Workflow applied to develop and apply SSRseq approach at PGTB**
 - 60 pairs of primer developed and amplified in a multiplex PCR
 - 95 blind-repeated genotyped individuals to identify unreliable loci and **estimate error**
 - 50% of success in **developing reliable genetic marker** (mean:30 high-quality loci developed)

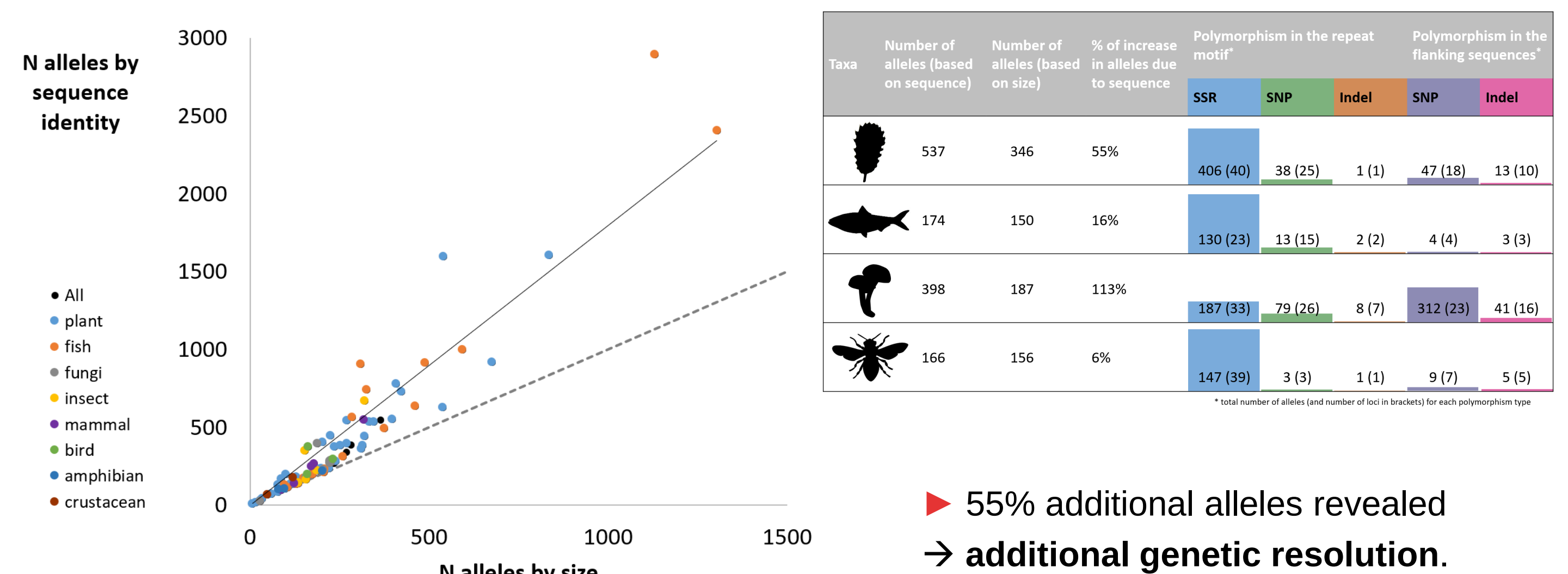


- Developments on 105 species so far (06/2026)
- 4,437 high-quality markers developed
- and genotyped in 48,000 samples
- 53% of studied species had no previous genetic resources: opening new research possibilities.

- **Bringing SSR genotyping into the digital area**
 - **Time and cost effective** workflow to develop SSRseq approaches for any species.
 - **Easy application:** Simple wet lab protocol and automated bioinformatics data analysis.
 - **Multi-species** analyses for comparative population genetics and community genetics.

IMPROVED RESOLUTION

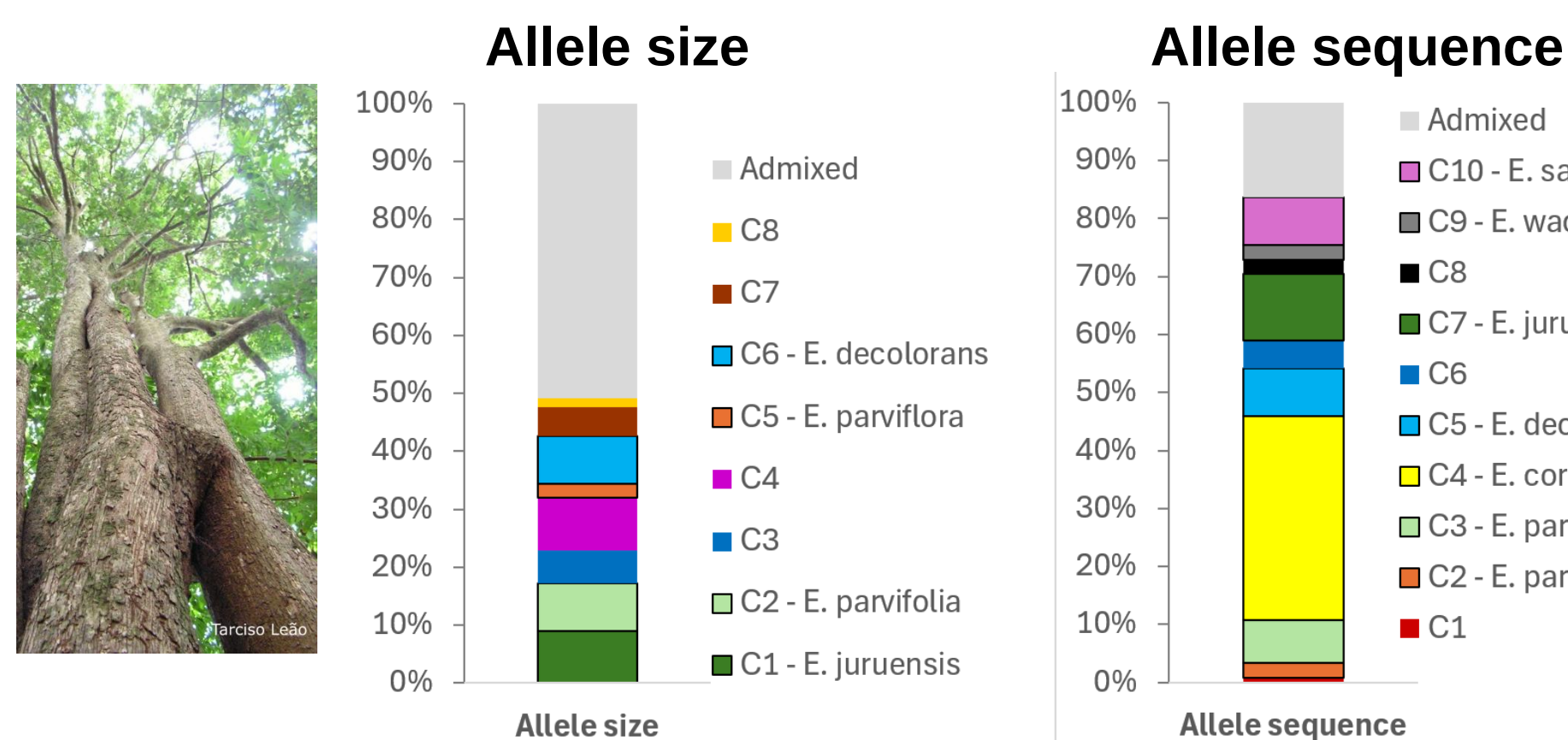
- **Information gained from sequence data**
 - Microhaplotype accounting for **all types of polymorphism** across the amplicon
 - SNP and Indel within the SSR or in the flanking region are common
 - Sequence data **reduces homoplasy** and increase the number of detected alleles
 - The proportion of the different type of polymorphisms varies between taxa



- 55% additional alleles revealed
→ **additional genetic resolution.**

MORE PHYLOGENETIC DEPTH

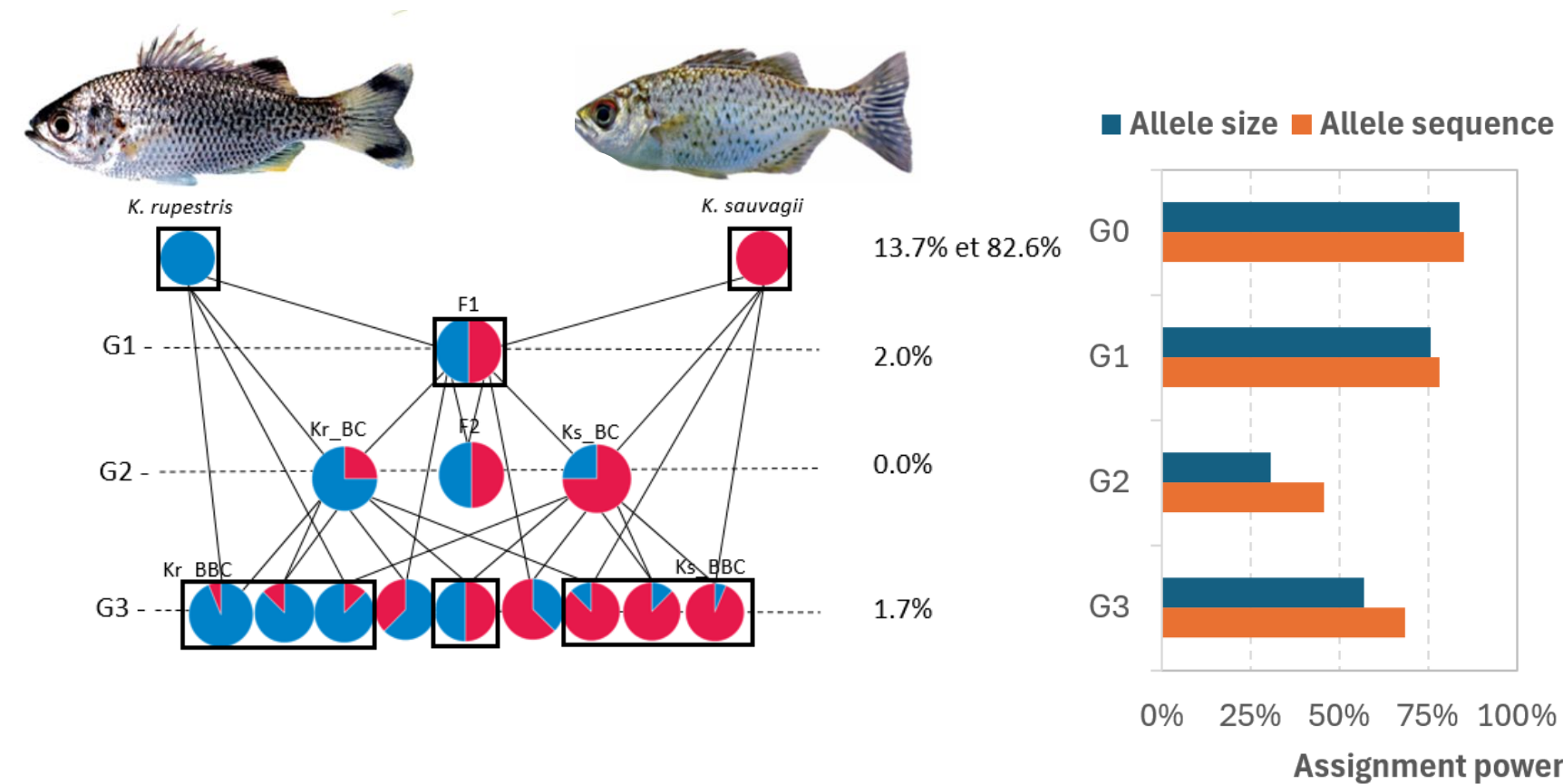
- **Improved species delineation in species complexes**



- **Eschweilera species complex:**
 - homoplasy blurs species delimitation
 - artificially increases admixture

Rakotoarisoa et al. (submitted)

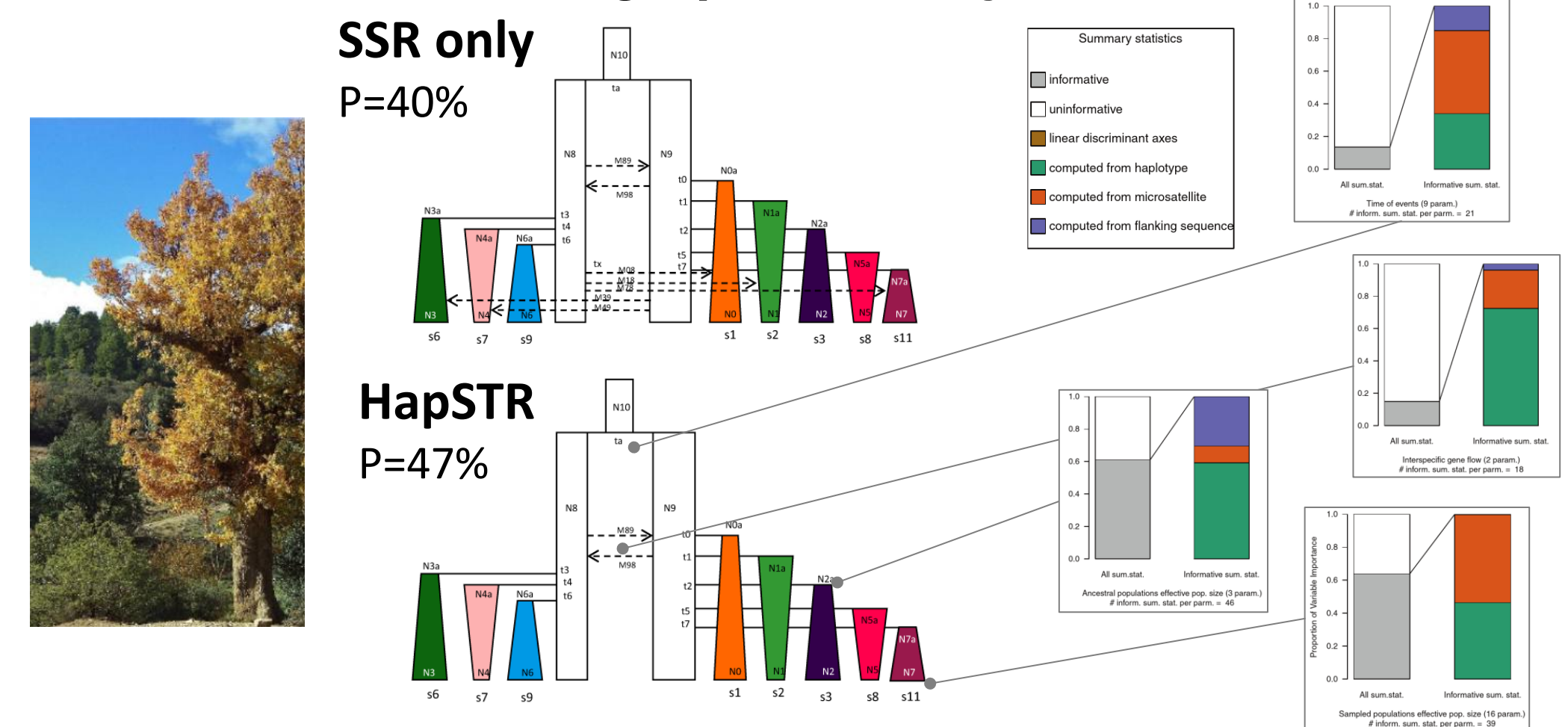
- **Increase power for hybrid assignment**



- **Kuhlia rupestris and K. sauvagii hybridization:**
 - high accuracy of hybrid assignment
 - sequence increases power for G2 and G3 hybrids

Lepais, Teichert, Faivre, Valade et al. (in prep)

- **Refined demographic history inferences**



- **Demographic history of Algerian Quercus trees⁸:**
 - homoplasy caused overestimation of recent hybridization
 - flanking sequences inform about ancient events

ECO-OPTIMIZED PROTOCOL

- PCRs in **384-well plates** (low volumes), indexing up to 1536 samples per run.
- Multi-allelic & multi-polymorphism markers: **compact** source of genetic information, simple and **frugal** method for individual identification.
- **Cost and environment-effective** genotyping approach.

TOWARD COMPLEX DNA GENOTYPING

- Allelic dosage estimation from allele sequence coverage in **polyploidy species**.
- Optimized for **non-invasive sampling** (buccal swabs, hairs, faeces, feathers,...)
- Applicable to DNA mixtures or **environmental DNA¹⁰**.

RESOURCES: SSRseq development and genotyping protocols are described in Lepais et al. 2020 Fast sequence-based microsatellite genotyping development workflow. PeerJ 8:e9085 (<https://peerj.com/articles/9085/>). The documented automated bioinformatics pipeline to call genotypes from sequence reads is available in Data INRAE (<https://doi.org/10.15454/HBXKVA>), as well as an updated list of species for which the method was developed (<https://doi.org/10.15454/FOIFCN>).

ACKNOWLEDGMENTS: Technical developments and sequencing were performed at the Genome Transcriptome Facility of Bordeaux (doi:10.15454/1.5572396583599417E12).

REFERENCES: 1: Guichoux et al. 2011 Current trends in microsatellite genotyping. Mol. Ecol. Resour. 11: 591–611; 2: Lepais & Bacles 2011 Comparison of random and SSR-enriched shotgun pyrosequencing for microsatellite discovery and single multiplex PCR optimization in *Acacia harpophylla* F. Muell. Ex Benth. Mol. Ecol. Resour. 11: 711–724; 3: Verbiest et al. 2023 Mutation and selection processes regulating short tandem repeats give rise to genetic and phenotypic diversity across species. J. Evol. Biol. 36, 321–336; 4: Xie et al. 2019 DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. Science 363: 81–84; 5: De Barba et al. 2016 High-throughput microsatellite genotyping in ecology: Improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. Mol. Ecol. Resour. 17: 492–507; 6: Darby et al. 2016 Digital fragment analysis of short tandem repeats by high-throughput amplicon sequencing. Ecol. Evol. 6: 4502–4512; 7: Vartia et al. 2016 A novel method of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. R. Soc. Open Sci. 3: 150565; 8: Lepais et al. 2020 Fast sequence-based microsatellite genotyping development workflow. PeerJ 8:e9085; 9: Lepais et al. 2022 Joint analysis of microsatellites and flanking sequences enlightens complex demographic history of interspecific gene flow and vicariance in rear-edge oak populations. Heredity doi: <https://doi.org/10.1038/s41437-022-00550-0>; 10: Lepais & Paz-Vinas 2025 Counting the Invisible: New Tools to Estimate the Number of Contributors From Sequence-Based Microsatellite Genotyping of Environmental DNA Samples. Mol. Ecol. Resour. 25(8), e70051.



PGTB
69 route d'Arcachon
33610 CESTAS FRANCE
05.35.38.53.42
pgtb@inrae.fr
<https://pgtb.fr/>

université
de BORDEAUX

INRAE

IBISA
Infrastructures
en Biologie
Santé et
Agronomie

INRAE
GENOMICS

RÉGION
Nouvelle-Aquitaine

FRANCE
GENOMIQUE