Application of high-throughput sequencing to population differentiation in the cupped oysters *Crassostrea angulata/C. gigas*

P-A. Gagnaire†, F. Cornette§, S. Heurtebise§, M-T. Augé†, N. Bierne†, S. Lapègue§

† ISEM – CNRS – UM2, UMR5554, SMEL, 2 rue des Chantiers, 34200 Sète, France (http://www.isem.univ-montp2.fr)
§ IFREMER, Genetics and Pathology of Marine Molluscs, 17390 La Tremblade, France (http://www.ifremer.fr/sg2m)

SUMMARY

The cupped oysters *Crassostrea angulata* and *Crassostrea gigas* have been sequentially introduced in Europe from their native Pacific area for aquaculture purposes. Ongoing viral epidemic disease is responsible for large losses in oyster aquaculture production throughout Europe. In parallel with current efforts in developing selection programs for disease resistance, it becomes critically important to characterize genetic diversity patterns in both native and introduced areas. With the advent of next generation sequencing techniques, the power to resolve the fine-scale genetic structure at a genome-wide scale opens to new perspectives towards genetic resource management. Here, we present a genome-wide screen of genetic variation that combines SNP genotyping and RAD-sequencing approaches. Our results, based on more than 18,000 SNP markers, improve our comprehension of the population structure and provide a powerful subset of highly informative markers for stock assignment. As these markers were also used in QTL mapping approaches, these new genomic resources will orientate the choice of breeders in selection programs in order to maximize the natural genetic variance involved in pathogen resistance.

1 Sampling:

- **SNP genotyping:** 9 populations, 29-50 individuals per population, 384 SNPs
- **RAD sequencing:** 4 populations, 25 individuals per population, 18000 SNPs
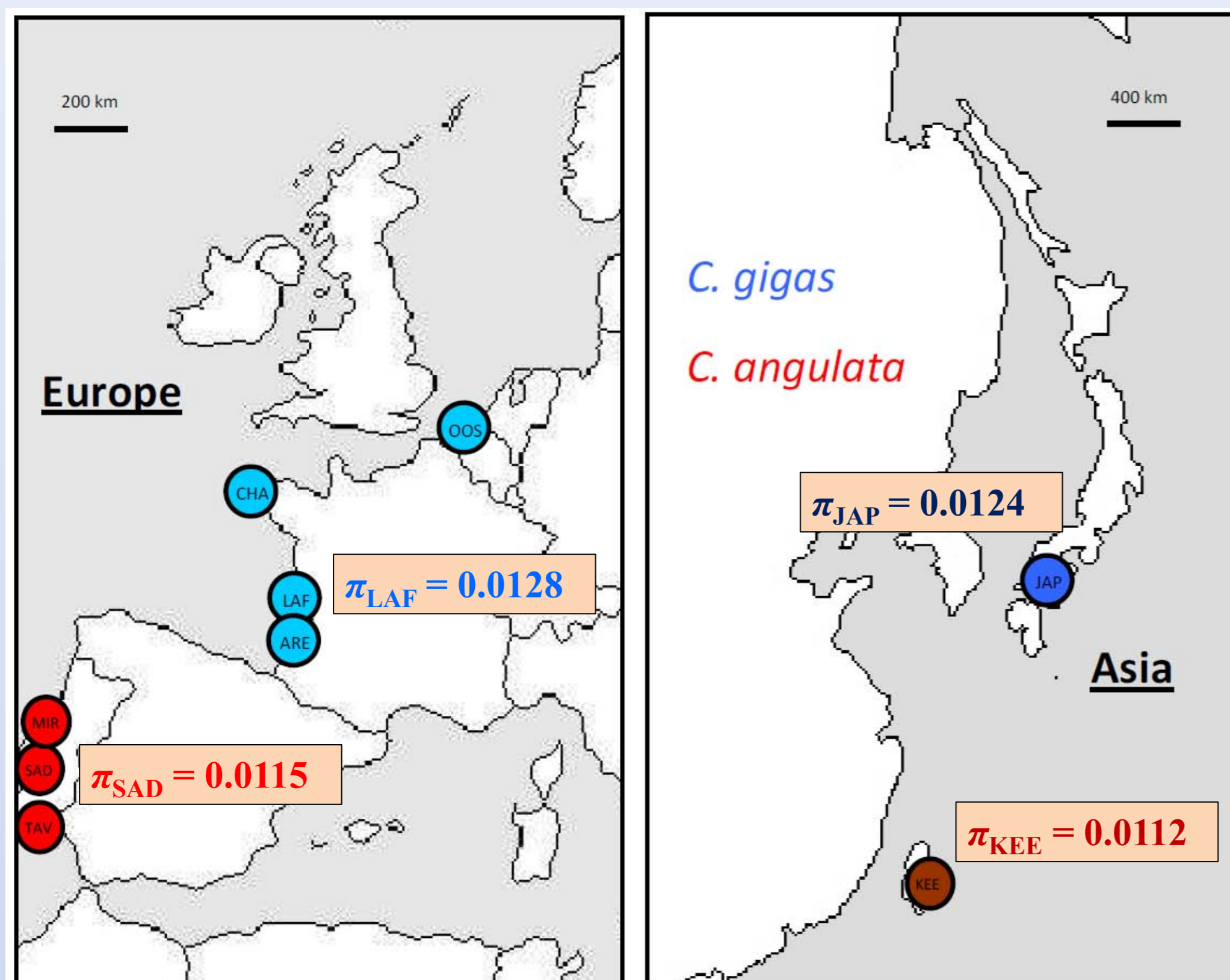

2 Principle of RAD sequencing:
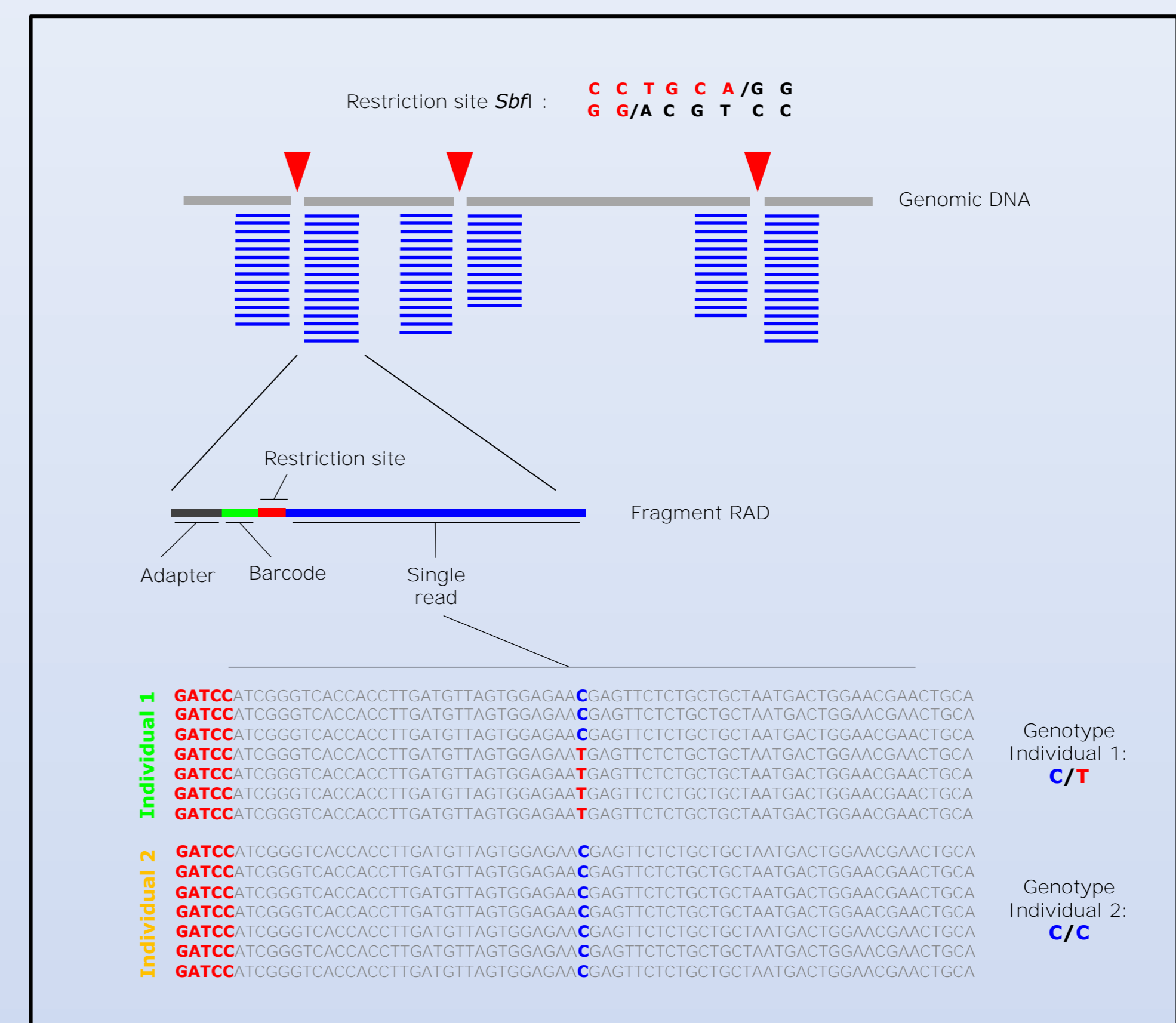
3 Genetic differentiation indices estimated from SNP and RAD datasets:

| Comparison	Type	No of RAD SNPs	average F_{ST} SNPs	average F_{ST} RADs	maximum F_{ST} RADs
KEE/SAD	within species *C. angulata*	8289	0.009	0.028	0.527
JAP/LAF	within species *C. gigas*	8499	0.004	0.019	0.272
KEE/JAP	between species native area	6479	0.177	0.058	0.741
SAD/LAF	between species introduced area	6148	0.177	0.071	0.775

- **Within species differentiation** between native and introduced areas is stronger in *C. angulata* for both SNP and RAD datasets. Average within species genetic differentiation is higher from RAD data.

- **Between species differentiation** is stronger in Europe than in Asia when assessed with RAD data. Average between species genetic differentiation estimated with the SNP panel dataset is higher that with the RAD dataset.

- Differences among datasets for the average differentiation values may be explained by the difficulty to assemble highly divergent RAD tags between species.

## 4 Distribution of RAD pairwise F_{ST} values:
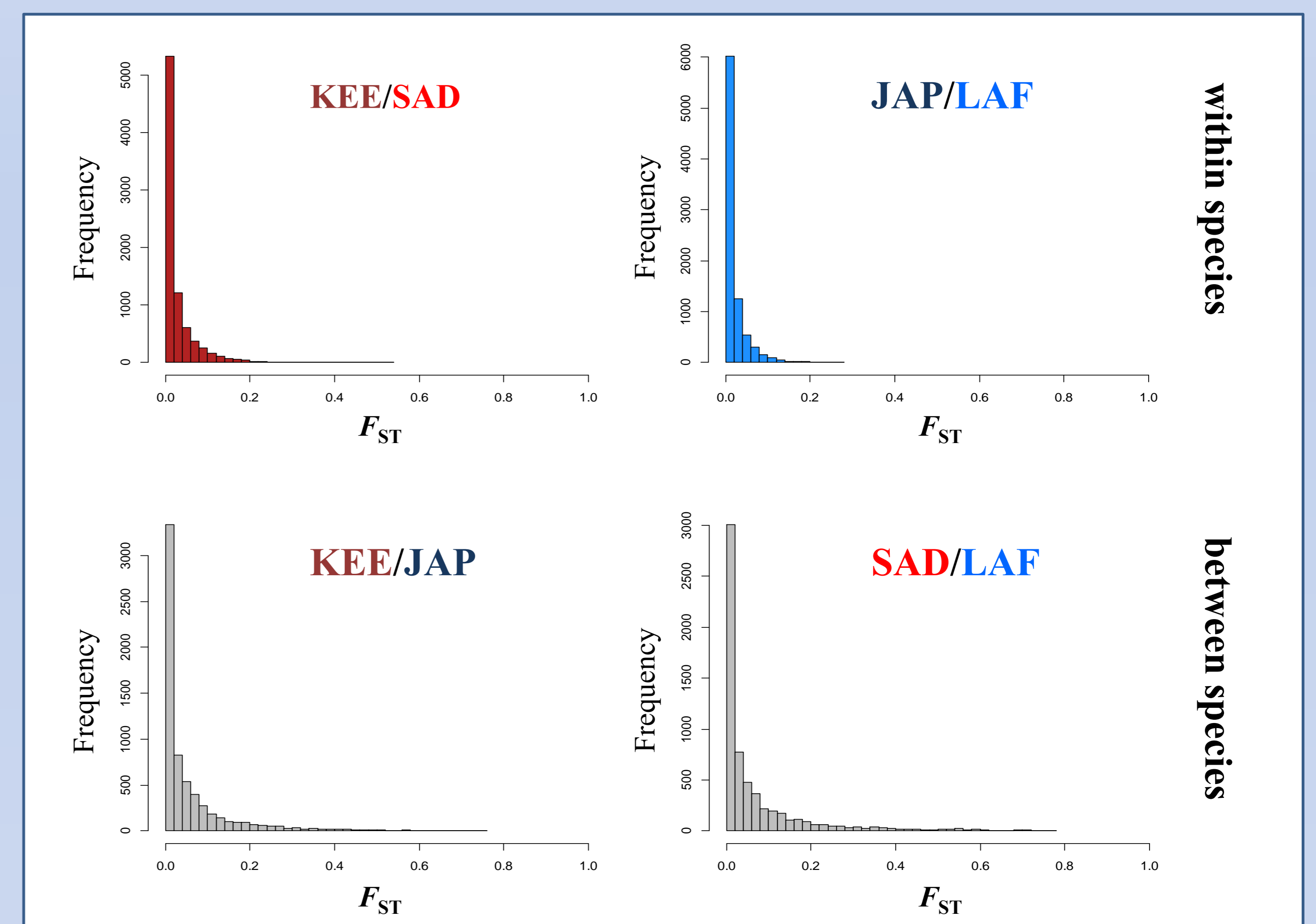
5 Spatial frequency patterns of between-species outlier SNPs:

- Highly differentiated SNPs may be used to **discriminate species** and understand spatial **introgression patterns**.
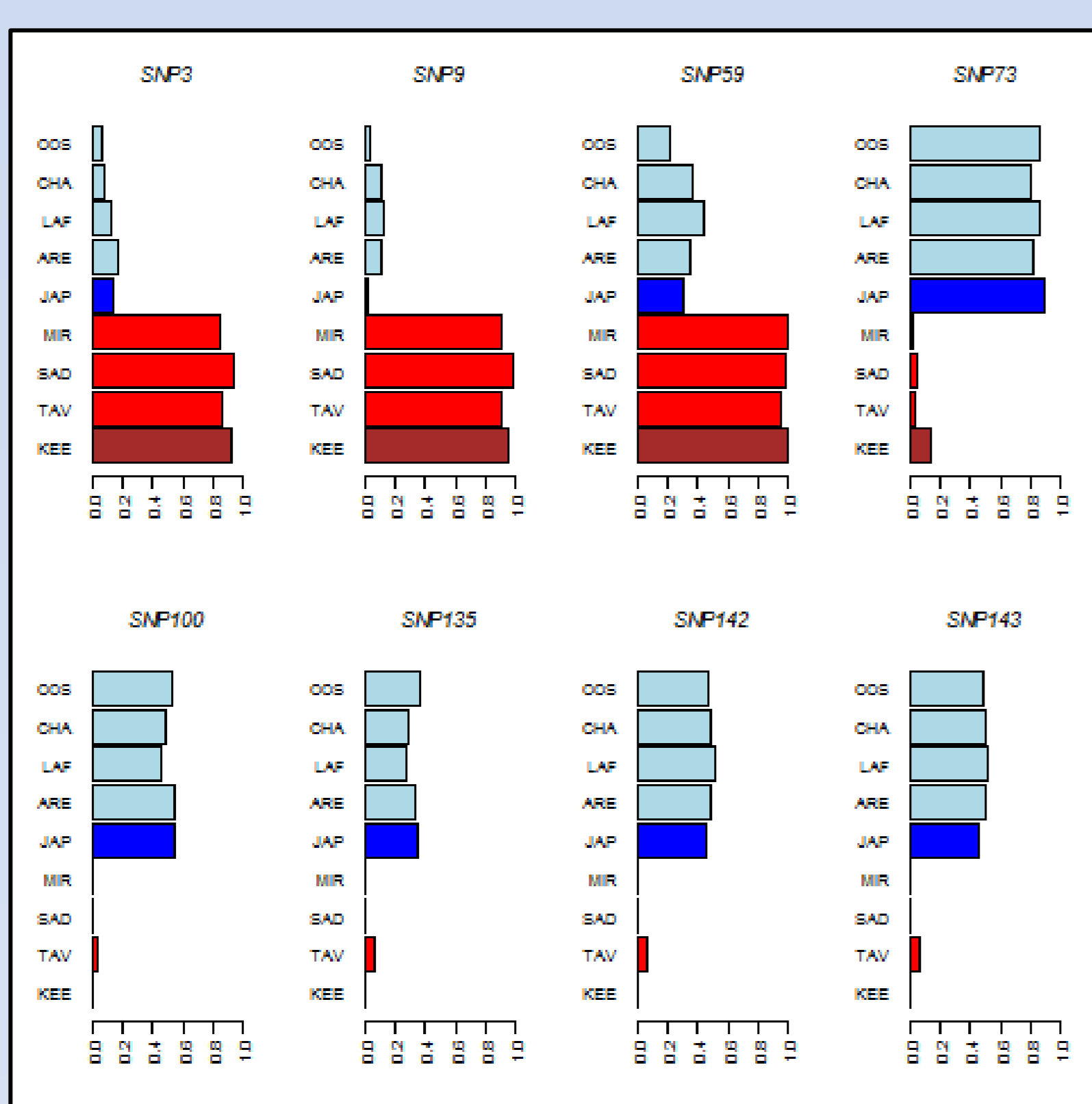
6 Comparison of PCA from SNP and RAD data:

- **RAD data** increase the power to **discriminate native and introduced populations**.
- Stronger differentiation in *C. angulata* may reflect the **older introduction time**.
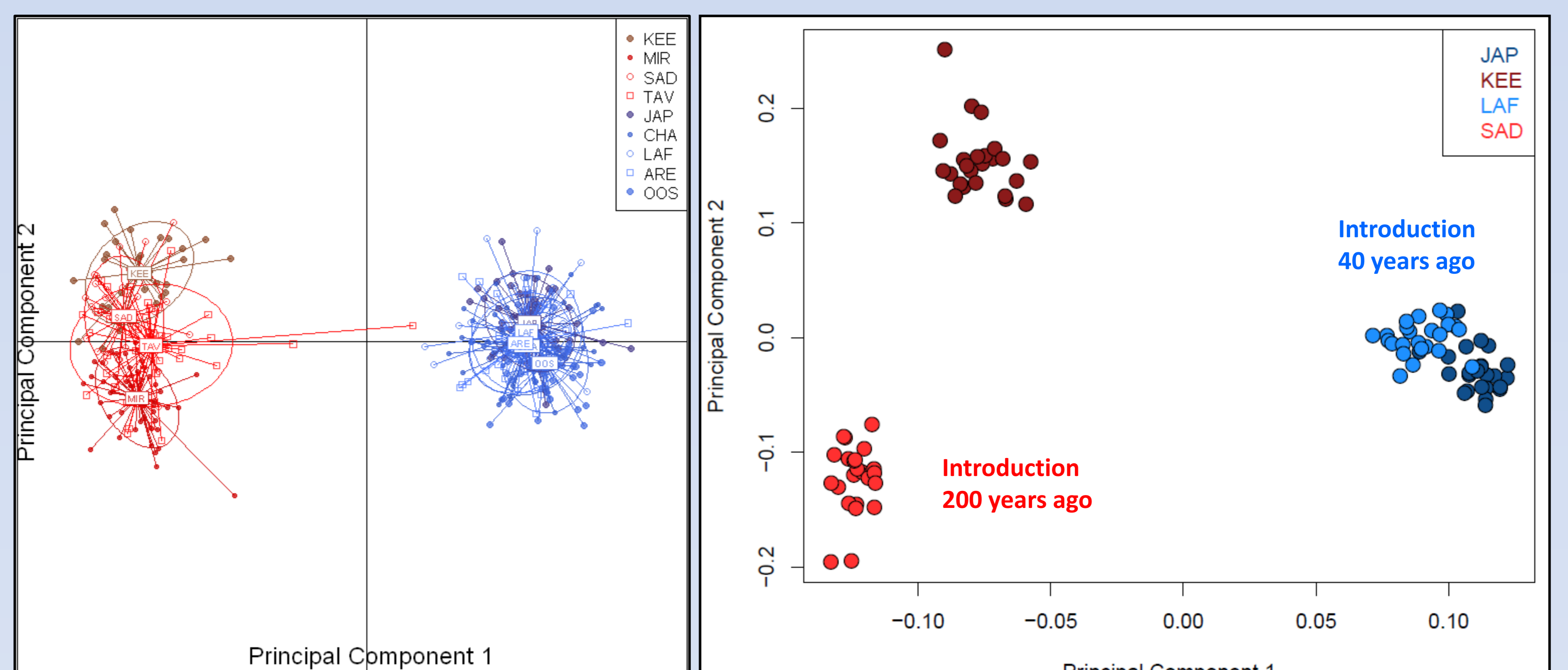
Conclusions

- ❑ **Genome-wide polymorphism** datasets significantly improve our ability to **discriminate closely related species** that still exchange genes in nature.
- ❑ The amount of information allows the **identification of native and introduced stocks**, even when genetic differentiation is weak and introduction is recent.
- ❑ **Mapping SNPs** on the high-density **linkage map** and the **reference genome** will reveal the location of the **genomic regions involved in reproductive isolation**.
- ❑ **Chromosomal patterns** of genetic differentiation will be compared with the map of **recombination rate** to test the effect of reduced recombination in maintaining the species barrier.