Working with GONE

Introduction to GONE
Genetic Optimization for N_e Estimation

- Originally published in 2020
- New version, GONE2, released last week
- Not to be confused with GONe (Coombs et al., 2011. Molecular Ecology Resources)

Introduction to GONE
Genetic Optimization for N_e Estimation

- Originally published in 2020

- New version, GONE2, released last week

- Not to be confused with GONe (Coombs et al., 2011. Molecular Ecology Resources)

**Accounting for population structure and data quality in demographic inference with linkage disequilibrium methods**

Enrique Santiago ✉, Carlos Köpke & Armando Caballero

986 Accesses | **20** Altmetric | Metrics

Introduction to GONE

- **Linkage Disequilibrium-based method**

- **Uses genome-wide data (SNPs) to infer recent demographic history**

- **Highest resolution within last 100-200 generations**

Figure from Nadachowska-Brzyska et al., 2021. *Methods in Ecology and Evolution*
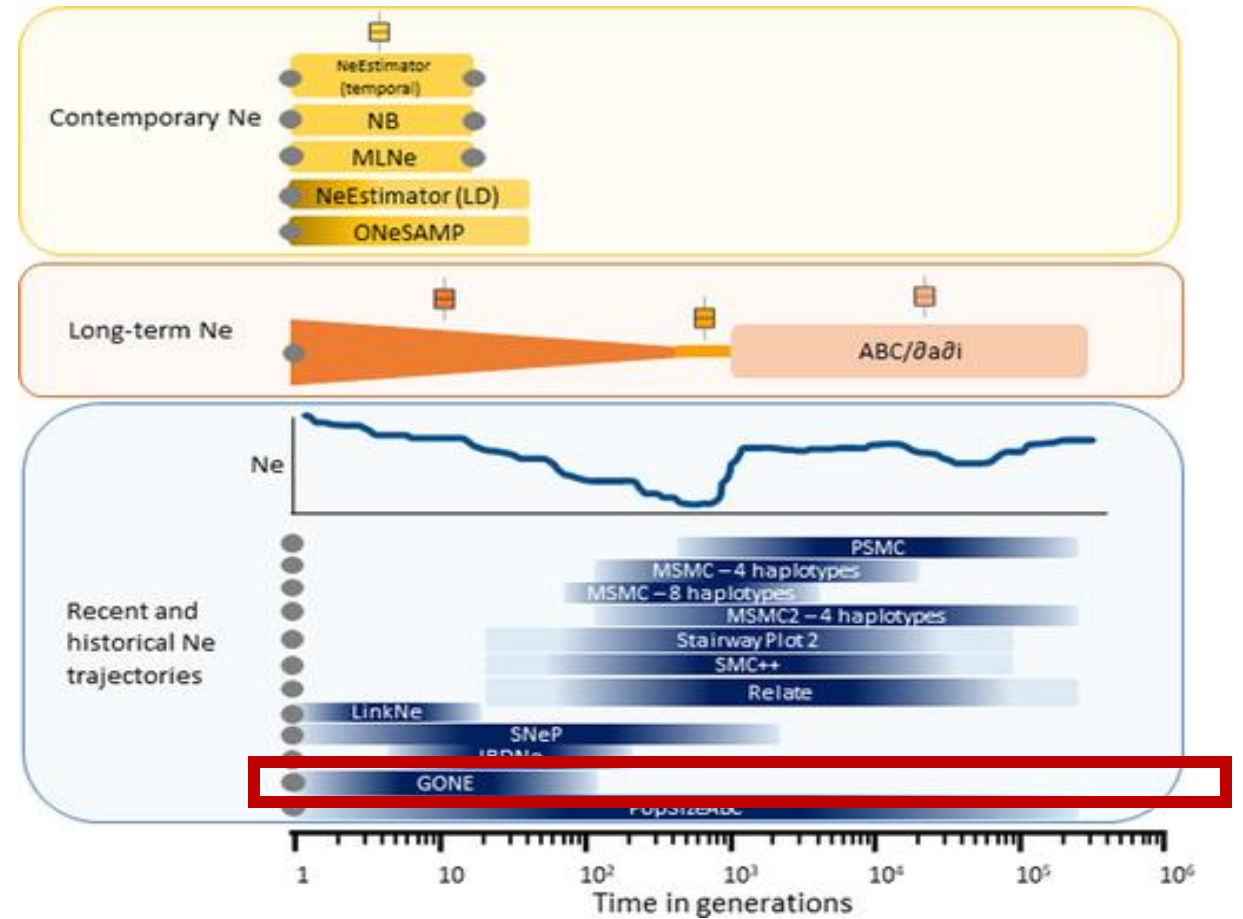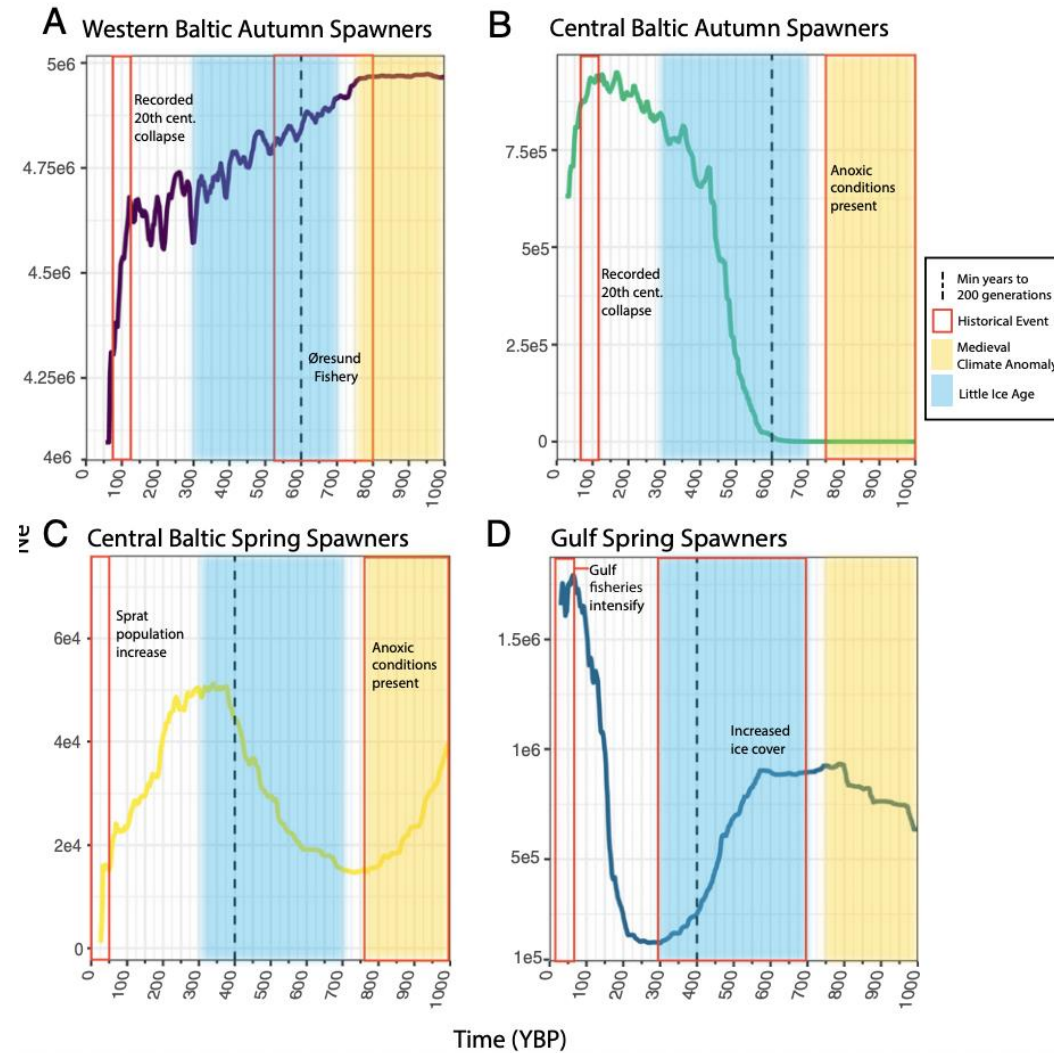
Introduction to GONE

Figure from Atmore et al., 2022. *PNAS*

Linkage Disequilibrium

- **Non-random association between alleles at two or more loci**
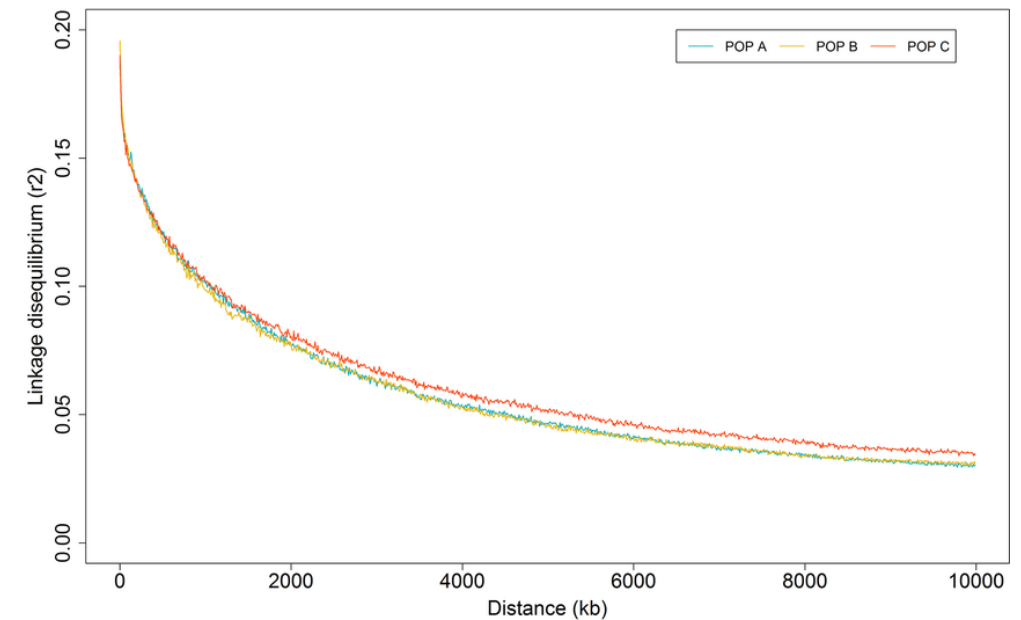- Function of both **population size** and **physical distance** between loci

Population size

- Small population → strong drift → more LD
- Large population → weak drift → less LD

Physical distance

- Closely linked loci → LD decays slowly
- Loosely linked loci → LD decays quickly

LD decay in three Nile tilapia populations

Figure from Yoshida et al., 2019, *Frontiers in Genetics*

Linkage disequilibrium patterns reflect population history

Loosely linked loci
- Few generations to be shuffled → recent past

Closely linked loci: distant past
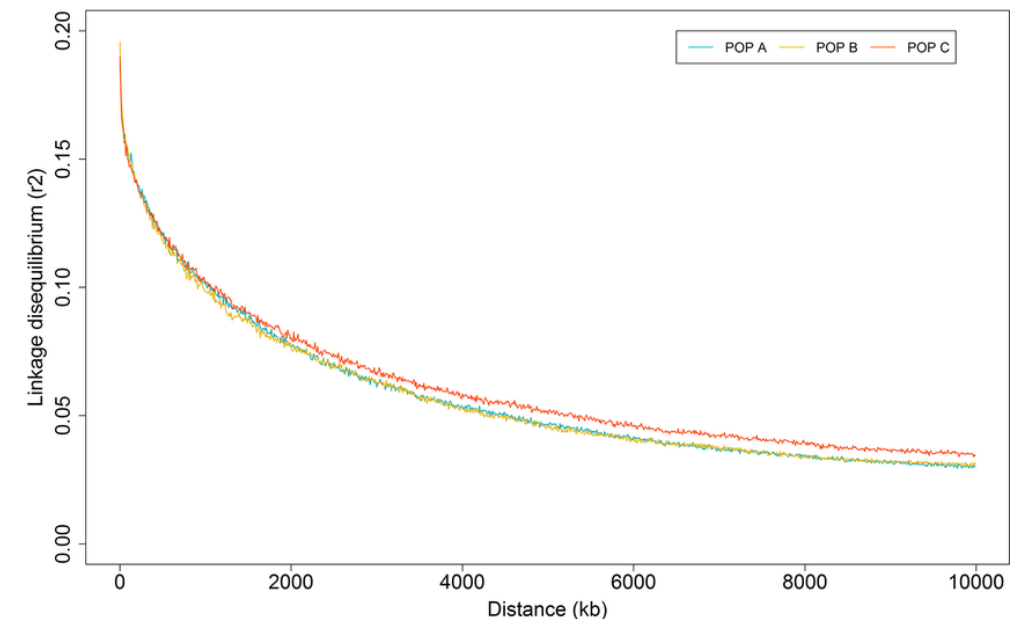- Many generations to be shuffled → distant past

GONE: linked loci
→Population size through time

NeEstimator: unlinked loci
→ Contemporary population size

LD decay in three Nile tilapia populations

Distant past   **Recent past**

Figure from Yoshida et al., 2019, *Frontiers in Genetics*

Working with GONE

INPUT

- SNP data from a population (ideally more than 10 individuals, number depends on population size)

- Map information
    - Physical distance: absolute distance between SNPs
    - Genetic distance: relative distance between SNPs (if available)

- Input parameters
    - Average recombination rate (if genetic map unavailable)
    - Phase
    - MAF
    - Maximum value of recombination rate
    - …

Working with GONE

ANALYSIS

1. LD measurement between many loci
2. Estimation of Ne using genetic optimization algorithm

OUTPUT

File containing generations and geometric means

Pitfalls and considerations

- Requires a contiguous reference genome

- Developed for an *ideal* population (no gene-flow, randomly mating, discrete generations)
    - → Usually not the case in wild populations

- Number of sampled individuals depends on population size
    - →Small population size → strong LD patterns → less individuals required for accurate estimates

- Increasing number of SNPs does not automatically increase accuracy

Improvements in GONE2

- Population structure (big issue in original GONE)
- Low-coverage sequencing data
- Genotyping errors
- Faster and more user-friendly

Further reading

- Santiago, Enrique, et al. "Recent demographic history inferred by high-resolution analysis of linkage disequilibrium." *Molecular Biology and Evolution* 37.12 (2020): 3642-3653.

- Santiago, Enrique, Carlos Köpke, and Armando Caballero. "Accounting for population structure and data quality in demographic inference with linkage disequilibrium methods." *Nature Communications* 16.1 (2025): 6054.

- Nadachowska-Brzyska, Krystyna, Mateusz Konczal, and Wieslaw Babik. "Navigating the temporal continuum of effective population size." *Methods in Ecology and Evolution* 13.1 (2022): 22-41.