13th of January 2022

Dear Editors,

With this letter we would like to submit our paper *On the origin and structure of haplotypes*, for consideration as opinion piece in Molecular Ecology. The article is intended for the upcoming **special issue:** *ecological and evolutionary inferences from long-read sequencing*.

As long and linked read sequencing technologies become more widely available, we will see a continued shift away from SNPs based inference to schemes that make use of the haplotypes.

In anticipation of this shift, we critically examine the fundamental definition of the ‘haplotype block’, which is widely used in reference to haplotype structure patterns. We argue that the term should be defined based on the structure of the Ancestral Recombination Graph (ARG), which contains complete information on the ancestry of a sample. We use simulated examples to demonstrate key features of the relation between haplotype blocks and ancestral structure, emphasising the stochasticity of the processes that generate them. We highlight a number of novel methods for inferring haplotype structure as full ARG or as a sequence of trees. While some of these new methods are computationally efficient, they still lack features to aid exploration of the haplotype blocks, as we define them, thus calling for the development of new methods.

Although we have submitted this as an opinion, we would like to highlight that the manuscript contains results from our original simulations.

Thank you for considering this manuscript, and we look forward to hearing from you.

Daria Shipilina

Sean Stankowski

Arka Pal

Yingguang Frank Chan

Nicholas Barton