Performance and limitations of linkage-disequilibrium-based methods for inferring the genomic landscape of recombination and detecting hotspots: a simulation study

Marie Raynaud[1,*], Pierre-Alexandre Gagnaire[1], Nicolas Galtier[1]
[1]ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France
*To whom correspondence should be addressed: marie.raynaud@umontpellier.fr

Abstract

Knowledge of recombination rate variation along the genome provides important insights into genome and phenotypic evolution. Population genomic approaches offer an attractive way to infer the population-scaled recombination rate $\rho=4N_e r$ using the linkage disequilibrium information contained in DNA sequence polymorphism data. Such methods have been used in a broad range of plant and animal species to build genome-wide recombination maps. However, the reliability of these inferences has only been assessed under a restrictive set of conditions. Here, we evaluate the ability of one of the most widely used coalescent-based programs, *LDhelmet*, to infer a genomic landscape of recombination with the biological characteristics of a human-like landscape including hotspots. Using simulations, we specifically assessed the impact of methodological (sample size, phasing errors, block penalty) and evolutionary parameters (effective population size (N_e), demographic history, mutation to recombination rate ratio) on inferred map quality. We report reasonably good correlations between simulated and inferred landscapes, but point to limitations when it comes to detecting recombination hotspots. False positive and false negative hotspots considerably confound fine-scale patterns of inferred recombination under a wide range of conditions, particularly when N_e is small and the mutation/recombination rate ratio is low, to the extent that maps inferred from populations sharing the same recombination landscape appear uncorrelated. We thus address a message of caution for the users of these approaches, at least for genomes with complex recombination landscapes such as in humans.

Key words

Population-scaled recombination rate, LDhelmet, simulations, linkage disequilibrium, recombination hotspots

Introduction

Recombination is highly conserved among sexually reproducing species of eukaryotes. This fundamental mechanism of meiosis is essential for the proper segregation of homologous chromosomes during the reductional division. Recombination involves crossing over events (CO) that play a crucial evolutionary role by allowing genetic mixing and generating new combinations of alleles (Baudat and de Massy 2007; Cromie et al. 2001; Capilla et al. 2016). Measuring the rate at which recombination occurs and the magnitude of its variation along the genome has important implications for fundamental research in molecular biology and evolution (Stapley et al. 2017), but also for applied genomics such as genome-wide association studies (GWAS) (Morris et al. 2013; Hunter et al. 2016). Several approaches have been developed to reconstruct genome-wide recombination maps (reviewed in Peñalba and Wolf 2020). Cytological methods, like ChIP-seq, target protein-DNA complexes directly involved in the formation of double-strand breaks (DSB) and CO during meiosis (Pratto et al. 2014). Gamete typing methods analyse the meiotic products of a diploid individual (reviewed in (Carrington and Cullen 2004; Dréau et al. 2019; Sun et al. 2019). Methods based on pedigree analysis reconstruct the gametic phase from patterns of allele inheritance in bi-parental crosses (Lander et Green 1987; Kong et al. 2002; Kodama et al. 2014; Rastas 2017). All these approaches have the advantage of providing direct estimates of the recombination rate. However, by focusing on CO that occurred in a few individuals or families across one or a couple of generations, they remain intrinsically limited in resolution due to the small number of recombination events that occur per chromosome per generation (Clark et al. 2010; Peñalba and Wolf 2020).

Another type of approach uses genome sequence data from population samples to take advantage of the large number of recombination events that have occurred during the history of the considered population. Instead of directly observing crossover products, these methods detect the footprints left by historical recombination events on patterns of haplotype segregation and linkage disequilibrium (LD) (reviewed in Stumpf and McVean 2003). The recombination rate and its variation across the genome are inferred via coalescent-based analysis of DNA sequence polymorphism data (Chan et al. 2012; Kamm et al. 2016; Li and Stephens 2003.; McVean et al. 2004; Spence and Song 2019). The resulting LD maps have been widely used to evaluate the genomic impact of natural selection and admixture, and to perform genome-wide association studies (GWAS) (*e.g.* Chan et al. 2012; The International HapMap Consortium 2007). These approaches provide an accessible and attractive way of describing recombination landscapes - *i.e.* the variation of recombination rates along the

genome - particularly in non-model taxa where direct methods are often difficult to implement (Auton et al. 2012; 2013; Melamed-Bessudo et al. 2016; Shanfelter et al. 2019; Singhal et al. 2015; Shield et al., 2020).

Direct and indirect methods have revealed considerable variation in recombination rate at different scales along the genome, particularly in vertebrates. At a large scale (of the megabase order), recombination tends to be concentrated in subtelomeric regions compared to centromeric and centro-chromosomal regions, a pattern shared among many species of plants and animals (Auton et al. 2012; Melamed-Bessudo et al. 2016; Capilla et al. 2016; Danguy des Déserts et al. 2021; Haenel et al. 2018). At a finer scale (of the kilobase order), recombination events often cluster within small regions of about 2 kb, called recombination hotspots (Choi and Henderson 2015; Kim et al. 2007; Mancera et al. 2008; Myers et al. 2005; Singhal et al. 2015; Shanfelter et al. 2019; Schield et al., 2020). Two distinct regulatory systems of recombination hotspot location have been described to date, with major implications on the evolutionary dynamics of recombination landscapes. In passerine birds (Singhal et al. 2015), dogs (Axelsson et al. 2012; Auton et al. 2013) and some teleost fishes (Baker et al., 2017; Shanfelter et al., 2019), recombination hotspots tend to be found in CpG-islands / promoter-like regions, and are highly conserved between closely-related species (Singhal et al. 2015). In contrast, in humans (Myers et al. 2005; 2010), apes (Auton et al. 2012 ; Great Ape Genome Project 2016) and mice (Booker et al. 2017), hotspot location is directed by the PRDM9 protein, which binds specific DNA motifs and triggers DSBs (Oliver et al. 2009; Baudat et al. 2010; Myers et al. 2010; Parvanov et al. 2010; Grey et al. 2018). In these taxa, hotspots are mostly located away from genes (Auton et al., 2012; Baker et al. 2017), and show little or no conservation between closely related species (Myers et al. 2005, 2010; Auton et al. 2012; Booker et al. 2017) due to self-destruction by gene conversion and rapid turnover of PRDM9 alleles (Coop and Myers 2007; Lesecque et al. 2014; Latrille et al. 2017).

Population-based inference methods aim to infer the population recombination rate $\rho = 4N_e r$, r being the per generation, per bp recombination rate and N_e the effective population size (Stumpf et McVean 2003). The ρ parameter reflects the density of population recombination events that segregate in polymorphism data, integrated across time and lineages. Several programs have been developed for reconstructing LD-maps (reviewed in (Peñalba and Wolf 2020; including PHASE: Li and Stephens 2003, LDhat: McVean et al. 2004, LDhelmet: Chan et al. 2012, LDpop: Kamm et al. 2016, and pyrho: Spence and Song 2019), which use the theory of coalescence with recombination to model the complex genealogies of samples stored in the underlying ancestral recombination graph (Griffiths et al. 1997; Arenas,

2013). The most popular family of LD-based methods, comprising LDhat (McVean et al. 2004) and its improved version LDhelmet (Chan et al. 2012) (see a literature survey in Supplementary Figure S1), implement a pairwise composite likelihood method under a Bayesian framework using a reversible jump Markov Chain Monte Carlo (rjMCMC) algorithm. They have been used for building fine scale LD-based maps in a broad range of animal taxa including humans (McVean et al. 2004), dogs (Axselsson et al. 2012; Auton et al. 2013), fruit flies (Chan et al. 2012), finches (Singhal et al. 2015), honeybees (Wallberg et al. 2015), sticklebacks (Shanfelter et al. 2019) and rattlesnakes (Schield et al., 2020). In some species, inferred LD-based maps have been validated by comparison with recombination maps obtained using direct approaches, confirming their overall good quality (McVean et al. 2004; Chan et al. 2012; Singhal et al. 2015; Booker et al. 2017; Shanfelter et al. 2019). However, as genetic and LD-based maps greatly differ in their resolution (pedigree-based inference provide resolution of about 1 cM, while population-based methods can infer recombination events at the kilobase scale, Peñalba and Wolf 2020), such comparisons do not provide qualitative information on the reliability of inferred fine-scale variation and hotspots detection. Moreover, the heterogeneity of studies in terms of taxonomy, genetic diversity, demography, sample size, and software parameters, among other things makes it difficult to appreciate the performance and the possible weaknesses of LD-based methods. For these reasons, the reliability and conditions of application of LD-based methods are still poorly understood and need to be more thoroughly characterised, considering the growing importance of these tools.

The power and sensitivity of LDhat and LDhelmet have been tested by simulations aiming to evaluate the influence of switch error in haplotype phasing (Singhal et al. 2015; Booker et al. 2017), the amount of polymorphism, and the intensity of recombination hotspots (Singhal et al. 2015). These studies simulated simple recombination landscapes assuming either homogeneous recombination rates or a few, well-defined hotspots contrasting with a low-recombination background (McVean et al. 2004; Auton and McVean, 2007; Chan et al. 2012; Singhal et al. 2015; Booker et al. 2017; Shanfelter et al. 2019; Schield et al., 2020). Real recombination landscapes that were characterised with a fine-scale resolution such as in humans, however, appeared to be more complex and involve a continuous distribution of recombination hotspot density and intensity across genomic regions. This complexity has not been taken into account so far in benchmarking studies assessing the performance of LD-map reconstruction methods. We thus lack a comprehensive picture of the ability of these methods to properly recover the biological characteristics of human-like recombination landscapes interspersed with hotspots. In particular, the proportion of the inferred recombination hotspots that are correct, and the proportion of true hotpots that are missed,

141 have not yet been quantified under a biologically realistic scenario. These are crucial quantities
142 to properly interpret and use reconstructed LD-maps in genomic research.
143 In this work, we specifically assessed the performance of the LDhelmet program to
144 detect hotspots while assuming a biologically realistic recombination landscape. We evaluated
145 the influence of methodological parameters including sample size, phasing errors and block
146 penalty, the impact of the population demographic history including its long-term effective size
147 and the occurrence of bottleneck and admixture events, and finally the effect of the mutation
148 to recombination rate ratio. We also considered different definitions of a recombination hotspot
149 relative to its background recombination rate, with the aim of improving the sensitivity of the
150 analysis. We identified the conditions in which LD-based inferences can provide an accurate
151 mapping of hotspots, and the parameters that negatively affect the sensitivity and specificity
152 of their detection across biologically realistic recombination landscapes.
153
154
155
156 **Results**
157
158 ***Recombination landscape modelling***
159 Five realistic, heterogeneous recombination landscapes (referred to as "underlying
160 landscapes" throughout) of 1Mb length were built using the human genome high resolution
161 map of meiotic DSB from Pratto et al. (2014). In order to mimic both broad and fine-scale
162 variation in the recombination rate parameter "*r*", the first and second half of each landscape
163 were drawn from a gamma distribution with mean 1 cM/Mb and 3 cM/Mb, respectively, and
164 parameters fitted from Pratto et al. (2014) (1-500,000bp: shape=rate=0.1328; 500,001pb-
165 1Mb: shape=0.1598, rate=0.0532). Accordingly, the 5 recombination landscapes generated
166 (Supplementary Figure S2) showed broad-scale differences in recombination peak intensity,
167 with less elevated recombination peaks in the first half compared with the second half of each
168 chromosome. At a fine scale, recombination was concentrated in numerous peaks resembling
169 human recombination hotspots, with about 85% of the recombination concentrated in 15% of
170 the genome. The map lengths in recombination units were about 0.02 Morgan (Supplementary
171 Figure S2, S4).
172 Population-scaled recombination landscapes simulated under a constant effective
173 population size (hereafter called "simulated landscapes") were generated in 10 replicates for
174 the five underlying landscapes, using coalescent simulations with a mutation rate μ=10^-8
175 and 4 combinations of sample sizes (SS=10 or 20) and effective population sizes (N_e=25,000

176 or 250,000) (Figure 1A, Supplementary Figure S3A). The map lengths of simulated
177 landscapes were a little shorter than the underlying landscapes (about 0.015-0.018 Morgan),
178 reflecting the occasional occurrence of more than one recombination event between two
179 adjacent SNPs during the simulated coalescent histories (Supplementary Figure S4). These
180 simulated landscapes were also highly correlated with the underlying landscapes for each
181 combination of parameters (Spearman's rank correlation > 0.8 using a 500bp resolution level),
182 showing that the coalescent history has not resulted in a substantial loss of information about
183 recombination rate variation across the underlying landscape. As expected from the $\theta=4N_e\mu$
184 values used in our simulations (θ = 0.001 and 0.01 for N_e = 25,000 and 250,000, respectively),
185 the SNPs density of the large N_e populations was about one order of magnitude higher than
186 for smaller N_e populations (Supplementary Figure S5).
187

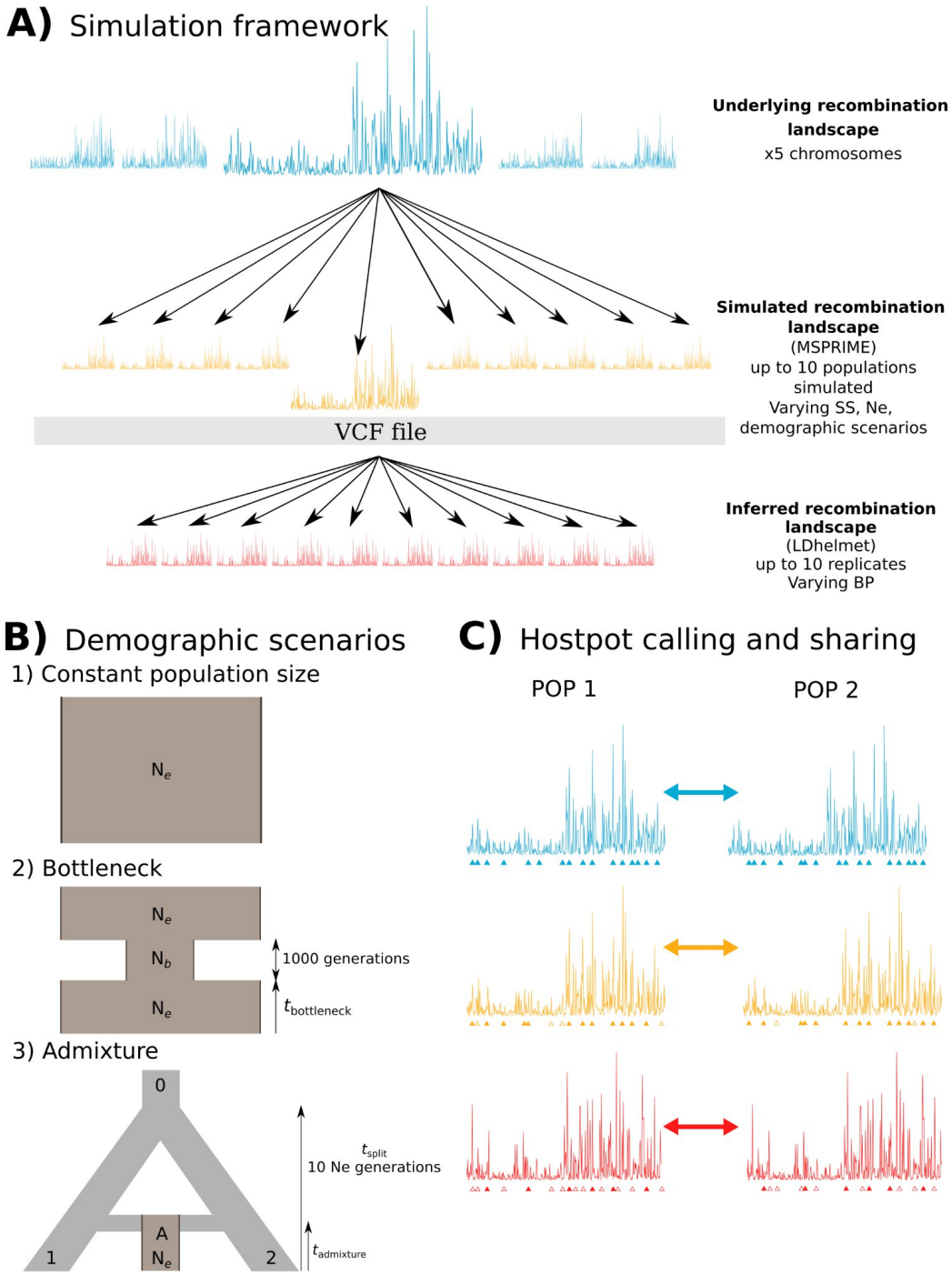

Figure 1. Simulation and hotspot calling protocols. **A)** Simulation framework. Five different underlying recombination landscapes were generated based on human empirical data. These five landscapes can either be considered as different regions from distinct chromosomes within a same species, or as orthologous regions of a same chromosome in different species. From these underlying landscapes, up to 10 recombination landscapes were simulated with MSPRIME 0.7.4 under various demographic scenarios, varying the effective population size (i.e. N_e=25,000, 100000, 250,000) and the sample size (i.e. SS=10, 20), to generate a VCF file for each simulated population. The VCF files were then used to infer the local population recombination rates using LDhelmet 1.19 setting the block penalty to BP = 5, 10 or 50. Up to

198 10 replicates per simulated population were analysed with LDhelmet. **B)** Demographic
199 scenarios. Three demographic scenarios were simulated with MSPRIME: 1) Constant
200 population size, varying the N_e of the simulated population; 2) Bottleneck event, varying the
201 age of the bottleneck (i.e. 500, 5000, 50000) and the N_e of the population during the bottleneck
202 (i.e. N_b=2500, 25,000), and setting the duration of the bottleneck to 1000 generations; 3)
203 Admixture event, varying the age of the admixture event (i.e. 500, 5000, 50000), and setting
204 the time of the split of the ancestral population into two populations 10*Ne generations ago.
205 **C)** Hotspot calling and sharing. Hotspots in the underlying (blue), simulated (orange) and
206 inferred (red) landscapes were defined as 2.5 kb-windows with a local recombination rate *X*
207 times as high as the averaged recombination rate of the 50 kb flanking regions. Several
208 threshold values were used to call the hotspots (i.e. *X*=2.5, 5, 10). The location of hotspots
209 was compared between populations (that share or not the same underlying landscape), to
210 compute the proportion of shared hotspots. Triangles below each landscape represent called
211 hotspots, with filled triangles indicating shared hotspots, and empty triangles hotspots that are
212 not shared with the compared landscape. In this example, the two compared populations
213 share the same underlying landscape, meaning that all the hotspots are shared between them
214 (all blue triangles are filled). The two simulated landscapes with independent coalescent
215 histories share most of these hotspots, while the two landscapes inferred with LDhelmet share
216 a smaller fraction of them, due to the additional inference step.
217

218 ## *Methodological parameters affecting LDhelmet performance*

219 Population-scaled recombination rates (ρ) were inferred from the simulated
220 polymorphism datasets using the program LDhelmet (Chan et al. 2012) (referred to as
221 "inferred landscapes" throughout). The effect of sample size and landscape resolution level
222 were assessed for a constant effective population size (N_e=25,000) using 10 or 20 diploid
223 individuals (SS=10 or 20) and three block penalty (BP) values (BP=5, 10 or 50), which
224 inversely determine the number of allowed changes in ρ value within windows of 50
225 consecutive SNPs (Figure 1A, Supplementary Figure S3A). Underlying and simulated
226 landscapes were converted into population-scaled recombination rates ($\rho=4N_er$), and each
227 underlying, simulated and inferred maps was smoothed using 500bp (*i.e.* the underlying
228 landscape resolution level) and 2500bp windows (*i.e.* a resolution level better suited to the
229 SNP density in our low-N_e simulations). The 10 simulated and inferred replicates of each
230 SS/BP condition were averaged to perform landscapes comparisons. Overall, local
231 recombination rates tended to be overestimated by LDhelmet, no matter the value used for
232 SS and BP, but this was especially observed when the local ρ was either very low (ρ < 10^-4)
233 or very high (ρ > 10^-2) (Supplementary Figure S6, panels A-F). The mean inferred map
234 lengths calculated across replicates varied substantially among tested conditions (0.017-
235 0.125 M), reaching up to 6 times the length of simulated maps in low SS and BP conditions
236 (*i.e.* SS = 10, BP = 5, 10, Supplementary Figure S4, upper panel). A BP value of 50 produced

very smooth recombination maps, which did not capture fine-scale variation in recombination rate. By contrast, maps inferred with BP=5 or BP=10 were visually similar and better reflected the fine-scale variation of the underlying landscapes (Supplementary Figure S7A). Spearman rank correlation coefficient between the mean simulated and inferred landscapes was reduced when SS were small (Figure 2A). Replicate runs of LDhelmet showed a strong consistency, as revealed by elevated correlations among the 10 replicate landscapes inferred from the same simulated landscape, whatever the SS and BP values being tested (Spearman's rho > 0.89, Figure 2B).

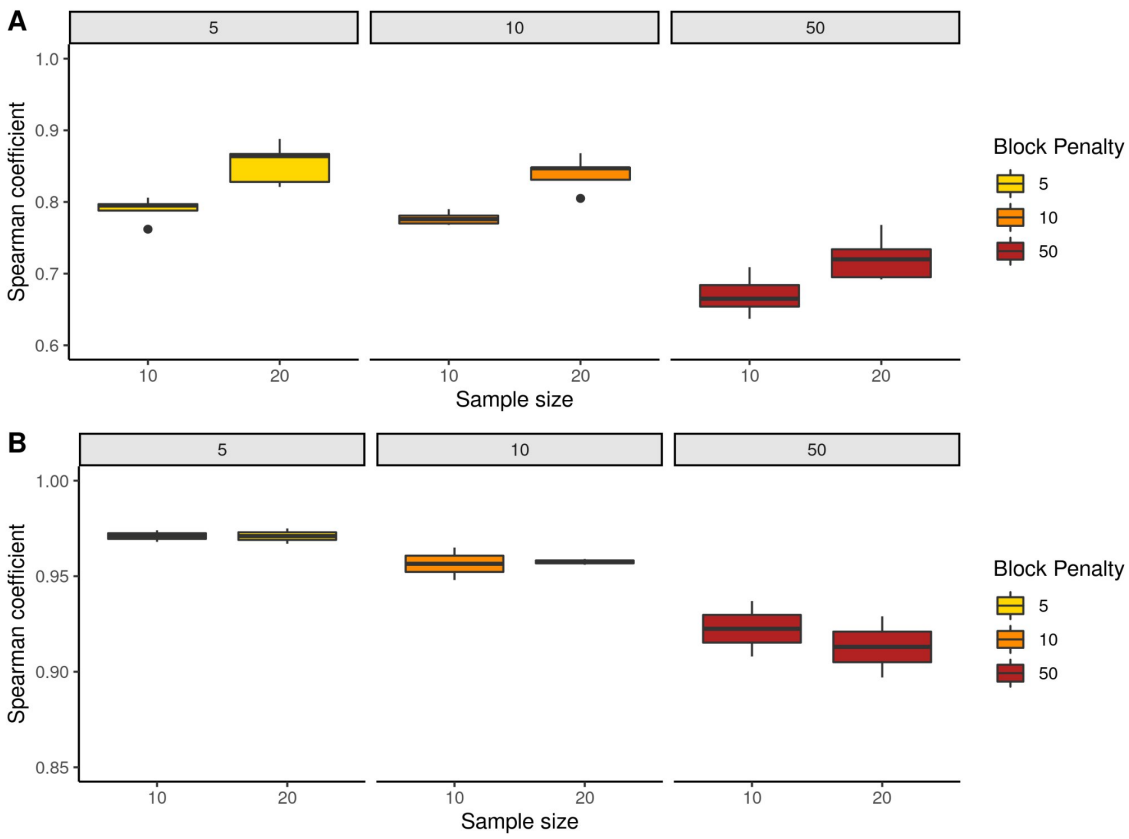

Figure 2. Performance (A), and repeatability (B) of LDhelmet as a function of the different parameters tested (i.e. N_e, SS, BP). The N_e of the simulated population is 25,000, the sample size is shown on the x axis (i.e. SS=10 or 20), and the LDhelmet BP values shown in color correspond to the different panels (i.e. BP=5, 10 or 50). **A)** Spearman correlation coefficients between the mean simulated and the mean inferred landscape calculated across ten replicates for each of five underlying landscapes (Supplementary Figure S3A). **B)** Mean pairwise Spearman correlation coefficients calculated between the 10 replicates of inferred landscapes from simulated populations sharing the same underlying landscape (i.e. Supplementary Figure S3B).

Recombination hotspots of the underlying, simulated and inferred landscapes were

called using three different threshold values commonly used in the literature (*i.e.* local recombination rate at least 2.5, 5 or 10 times higher than the background rate). True/False positives/negatives rates and discovery rates (TPR, FPR, TDR, FDR, TNR, FNR) were computed under each tested condition. The hotspot detection threshold ratio of 10 between the focal and flanking positions appeared too stringent and yielded a very small number of called hotpots (Supplementary Figure S8). Using a less conservative threshold ratio of 5, we detected 4 to 8 hotspots per Mb in the simulated landscapes, and 5 to 20 per Mb in the inferred landscapes. These numbers reached 40-50 and 20-50 per Mb, respectively, when a threshold of 2.5 was used. Irrespective of the chosen threshold, the number of inferred hotspots tended to be overestimated, notably when SS was small (Supplementary Figure S8). The 2.5 threshold was used for the remaining analysis as it reduced the variance in the number of called hotspots due to a higher call rate. The sensitivity (or TPR) of LDhelmet was medium, since depending on the SS and the BP used, between 29.4% and 52.7% of the simulated hotpots were inferred as such. The TPR was higher for small BP values (*i.e.* 5 or 10), but relatively insensitive to the SS value (Figure 3 left panel, Supplementary Figure S9A). The proportion of false hotspot calls (FDR, *i.e.* inferred hotspots corresponding to non-hotspot windows in the simulated maps) ranged between 25.6% and 52.9%, and was higher for SS = 10, without major differences between the BP values tested (Figure 3 right panel, Supplementary Figure S9B and C). No significant difference in the correlation between simulated and inferred landscapes was found between the first half of the chromosome with a mean r of 1 cM/Mb (referred to as the “cold” region) and the last half with a mean *r* of 3 cM/Mb (the “hot” region). This was also true for the TPR and the FDR, whatever the hotspot detection threshold used (*i.e.* 2.5 or 5) (Student test, $p > 0.05$).

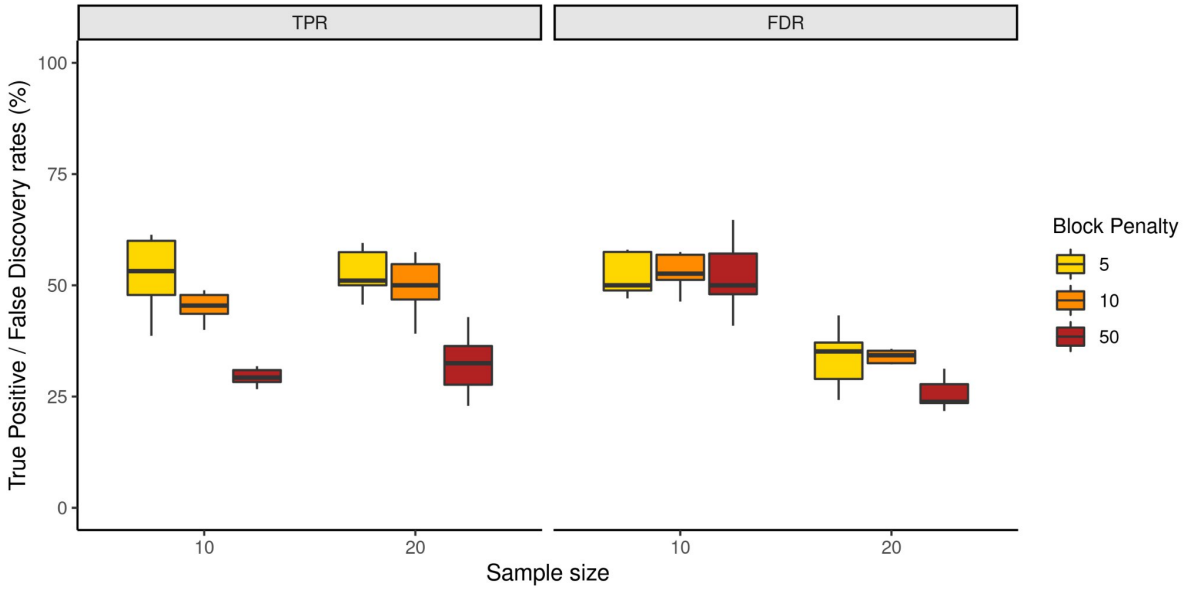

Figure 3. Hotspot detection. True positive (TPR, left panel) and false discovery (FDR, right panel) rates of inferred, as compared to simulated hotspots, called using a detection threshold of 2.5, for different sets of methodological parameters. The sample size parameter is shown on the x axis (i.e. 10 or 20), the block penalty (i.e. 5, 10, 50) is shown in color, and the N_e of the populations simulated is 25,000. See full results in Supplementary Figure S9.

To assess the impact of phasing errors on LDhelmet performance, the phase information was first removed from the whole VCF of the simulated landscapes for two replicate populations simulated with a constant Ne of 250,000 and a SS of 20. Statistical phasing performed with Shapeit 4.2.2 (Delaneau et al. 2019) resulted in a 6.7% average phasing error. We then introduced random phasing errors in the simulated VCFs with rates ranging from 2 to 10%, before inferring ρ with a BP value of 5. As the phasing error rate increased, Spearman rank correlation coefficients between simulated and inferred landscapes slightly decreased below the value obtained with perfectly phased data (Supplementary Figure S10A). Hotspot calling performance assessed with TPR and FDR was also negatively impacted by phasing errors (Supplementary Figure S10B).

Demographic and evolutionary parameters affecting LDhelmet performance

Methodological parameters were then set to SS=20 individuals and BP=5 - a trade-off optimising the balance between TPR and FDR - to focus on the effect of N_e on the quality of the LD inferences. When the simulated effective population size was large (i.e. 250,000, as compared to 25,000), the inferred map length was closer to the expected value of 0.02 M (Supplementary Figure S4, lower panels) and the local recombination rate tended to be less

overestimated (Supplementary Figure S6, panels G-L). A larger N_e also significantly increased the correlation between simulated and inferred landscapes (Figure 4A, Supplementary Figure S11), increased the TPR and decreased the FDR (Figure 4B, C).

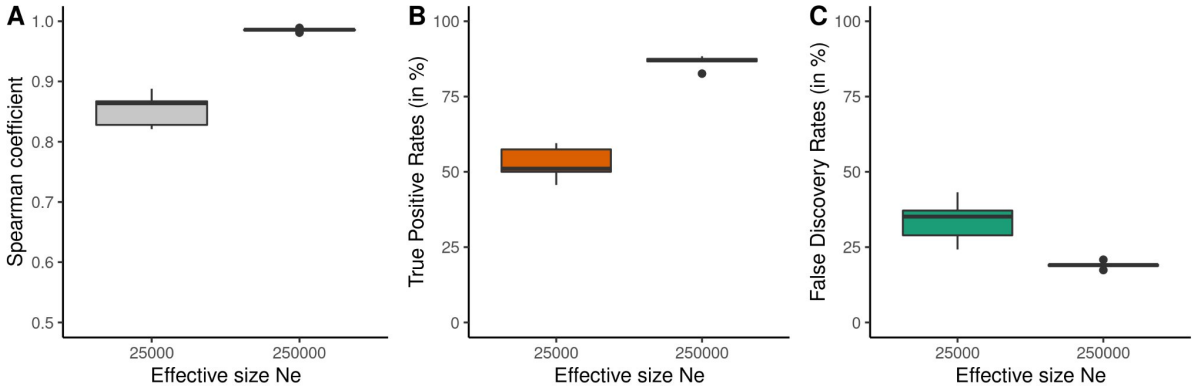

Figure 4. Effect of the effective population size parameter (Ne) on recombination rate inference and hotspot detection. **A)** Spearman correlation coefficients between simulated and inferred landscapes as a function of N_e (SS = 20, BP = 5). **B)** and **C)** True positive (TP) and false discovery (FD) rates of inferred, as compared to simulated hotspots, called using a detection threshold of 2.5, as a function of N_e (SS = 20, BP = 5). See full results in Supplementary Figure S9 and S11.

We then assessed the impact of non-equilibrium demographic histories. Populations undergoing bottleneck or admixture events of various ages were simulated, with the N_e of the ancestral and present-day populations set to 250,000, the SS to 20, and making other demographic parameters vary (*i.e.* t_b, N_b t_a, Figure 1B, see Materials and Methods). These demographic scenarios generally had a negative impact on the quality of the reconstructions (using a BP of 5), as compared to constant-size populations (Figure 5). The correlation between simulated and inferred maps decreased with the strength of the bottleneck (i.e. with lower N_b), and with the recentness of bottleneck and admixture events (Figure 4A, B). In the same way, TPR and FDR were degraded as compared to constant-size population scenarios (Figure 4C, D), particularly for young events.

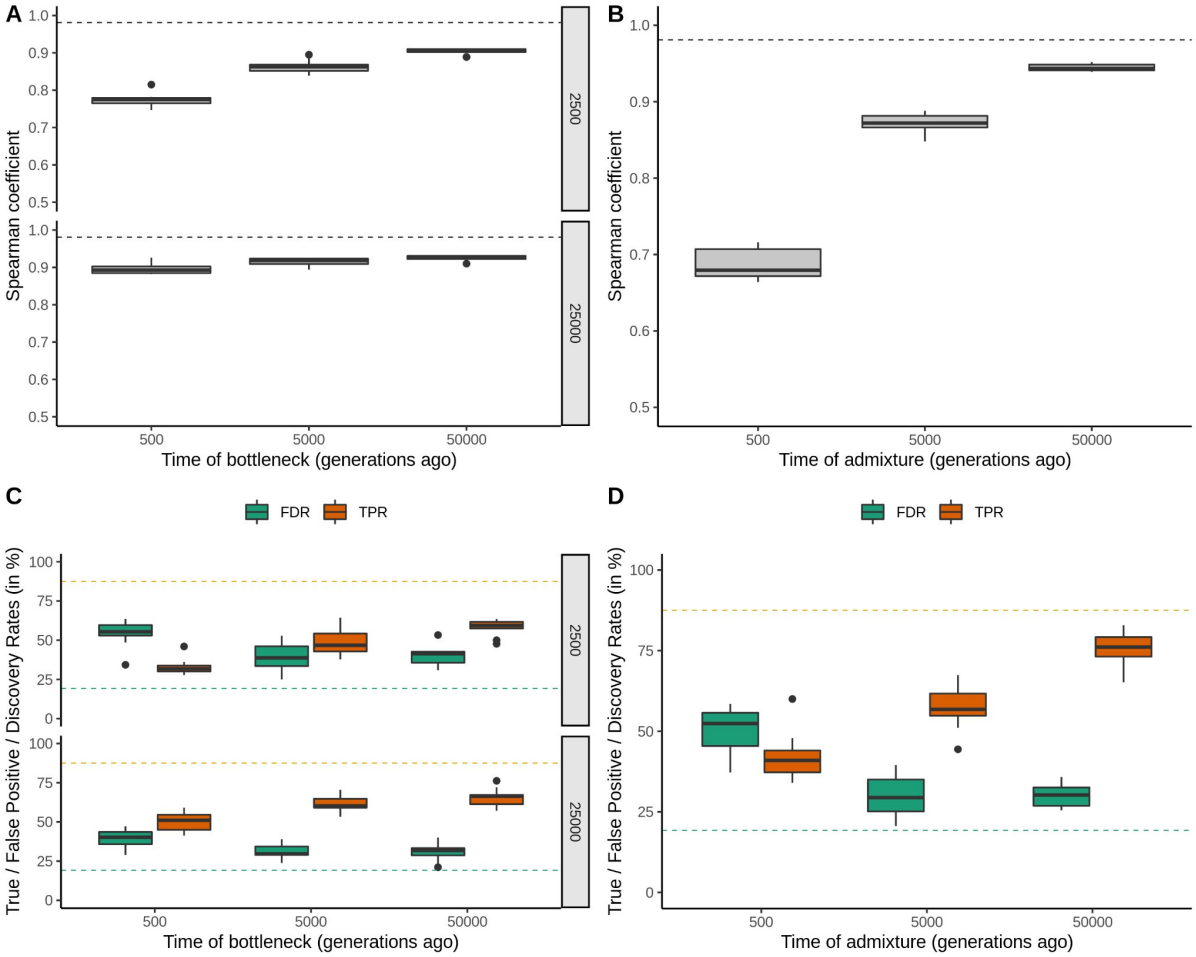

Figure 5. Influence of bottleneck (left panels A and C) and admixture (right panels B and D) events on recombination rate inference and hotspot detection. Spearman correlation coefficients between simulated and inferred landscapes are shown as a function of the age of the bottleneck event and the strength of the bottleneck (i.e. N_b, the N_e value during the 1000 generations of the bottleneck) (**A**), and as a function of the age of the admixture event (**B**). True positive (TP, in orange) and false discovery (FD, in green) rates of hotspots called using a detection threshold of 2.5, as a function of the time of the bottleneck event and the strength of the bottleneck (**C**), and as a function of the time of the admixture event (**D**). Dashed lines correspond to averaged Spearman's rho, TPR and FDR values in populations that did not experience bottleneck or admixture events.

The influence of species-specific evolutionary parameters such as the mutation and recombination rates was assessed by generating coalescent simulations under two additional underlying landscapes using a ten times higher (*i.e.* 20 cM/Mb) and a ten times lower (*i.e.* 0.2 cM/Mb) average recombination rate, and three different mutation rates (*i.e.* 10^-9, 10^-8 and 10^-7). The μ/r ratio under these 6 simulated conditions thus equalled 0.1, 1 or 10. For all conditions, Spearman's rank correlation between the mean simulated and the mean inferred landscapes was greater than 0.9, except when μ equalled 10^-9 (Spearman's rho ≃0.7,

Supplementary Table S1). An increased *μ*/*r* ratio improved the ability to detect hotspots when *r* was fixed to 10^-8, with a higher TPR (up to >80%) and a lower FDR (<5%) when *μ* increased (prop.test, p-value < 0.05, Figure 6, threshold 5 in Supplementary Figure S12). The *μ*/*r* ratio did not affect the performances the same way when *μ* was fixed to 10^-8: a *μ*/*r* ratio of 10 (r = 10^-9) yielded lower TPR (< 60%) and higher FDR (> 25%) than a ratio of 1 or 0.1, although these trends were not significant (prop.test, p-value > 0.05, Figure 6, threshold 5 in Supplementary Figure S12).

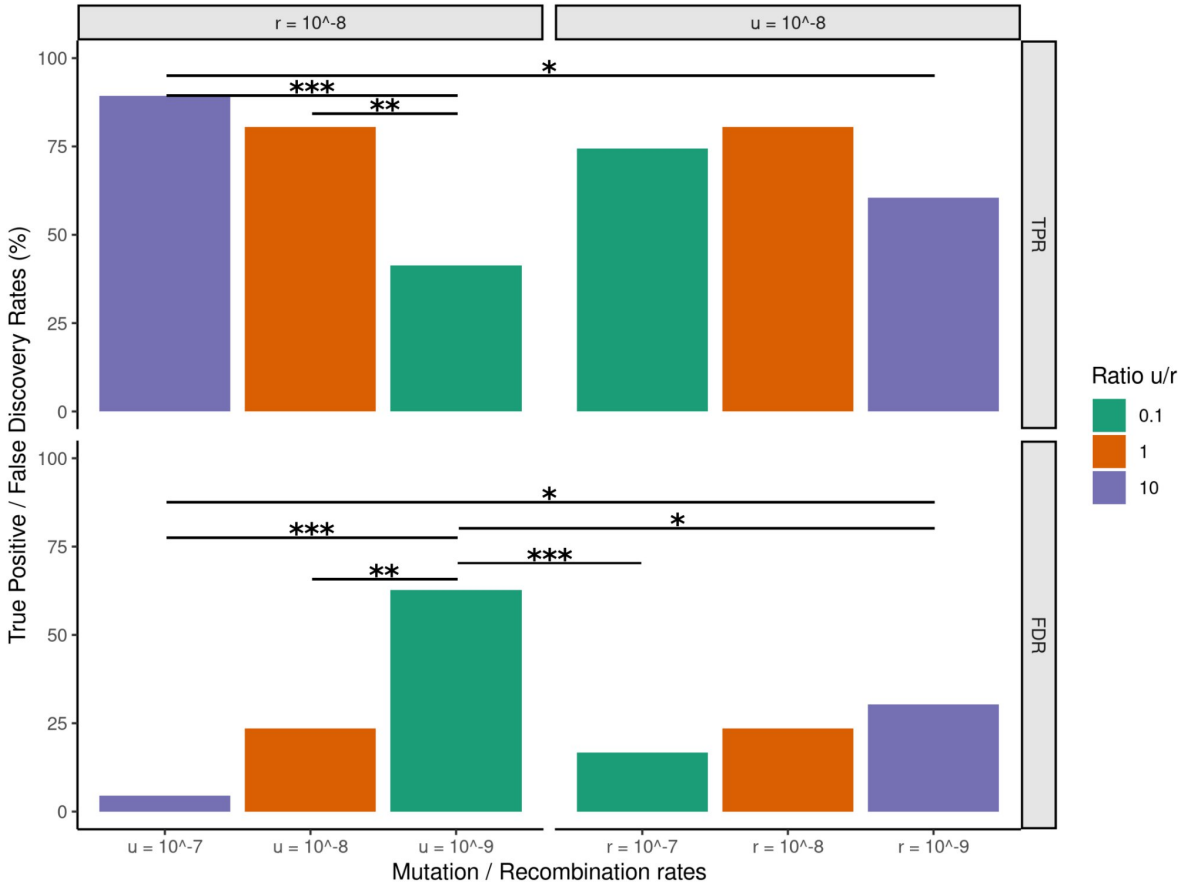

Figure 6. Influence of the u/r ratio on hotspot detection. True positive (TP, upper panel) and false discovery (FD, lower panel) rates of hotspots called using a detection threshold of 2.5. The x axis indicates values of u for *r*=10^-8 (left panels), and values of *r* for *μ*=10^-8 (right panels). Colors correspond to different values of the *μ*/*r* ratio. Asterisks show significant differences in percentages between comparisons (* prop.test p-value<0.05, **<0.01, ***<0.001).

Hotspot sharing between populations with different versus identical underlying recombination landscapes

As expected for a comparison between two populations simulated with different underlying recombination landscapes, the mean linear correlation (R^2 coefficient) between the corresponding inferred landscapes was low, between 0.012 and 0.084, and similar to the R^2 between the simulated landscapes (0.012-0.017) (Supplementary Table S2). A low percentage of shared hotspots (around 8% with a calling threshold of 2.5) occurred by chance between populations simulated with distinct underlying landscapes, with a SS of 20. Roughly similar proportions of shared hotspots were found between the corresponding inferred landscapes (with BP = 5), although these proportions were slightly overestimated (Figure 7, see Supplementary Table S2 to see all conditions). A minority of the shared inferred hotspots were TP, indicating that a non-zero fraction of truly shared hotspots is expected to be found between species with different biological recombination landscapes.

Then we compared simulated populations sharing the same underlying landscape, in order to check the ability of LDhelmet to recover similar recombination rates between populations with independent coalescence histories. The correlations between simulated landscapes were generally high for both low (R^2>0.7) and large N_e (R^2>0.9) conditions, but the correlations between inferred landscapes were much lower, with R^2< 0.3 and <0.75 for N_e = 25,000 and 250,000, respectively (Supplementary Table S2). The proportion of shared hotspots called with a threshold of 2.5 followed the same trend: it was high between simulated landscapes (>80%) and much lower between inferred landscapes (<50%) (Figure 7, Supplementary Table S2). Thus, one can expect LDhelmet to detect a moderate to low fraction of shared hotpots even between species truly sharing a common recombination landscape, depending on population size and sample size (and also hotpot definition).

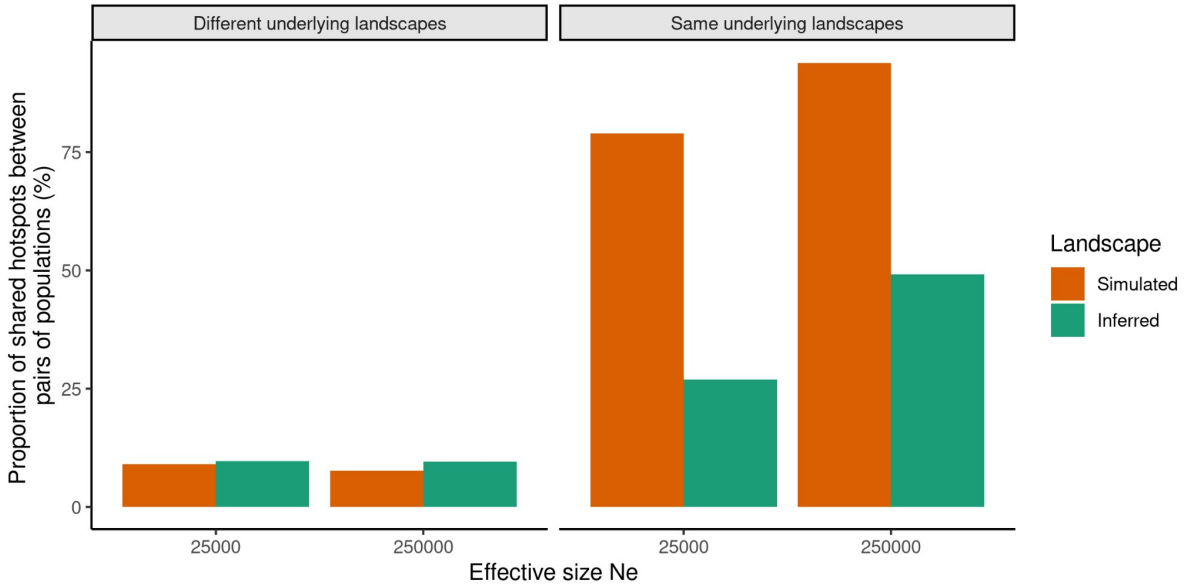

Figure 7. Expected and observed hotspot sharing between populations with different (left panel, following simulation framework from Supplementary Figure S3 A)) or identical (right panel, following simulation framework from Supplementary Figure S3 B) underlying landscapes. Mean proportion of shared hotspots between pairs of simulated (expected proportion, orange bars) and pairs of inferred (observed proportion, green bars) recombination landscapes as a function of N_e (i.e. 25,000 and 250,000, x axis). Only shown here are simulations with SS = 20, BP = 5, hotspots detection threshold = 2.5. The proportion of shared simulated and inferred hotspots for all combinations of parameters and for populations sharing or not the same underlying landscape are shown in Supplementary Table S2.

Populations simulated with the same underlying landscape that underwent a bottleneck or an admixture event also showed reduced hotspot sharing between inferred landscapes, as compared to truly shared hotspots between simulated landscapes. The proportion of shared hotspots between pairs of simulated landscapes was similar to constant-size populations and did not vary substantially according to the time or the strength of the demographic event, while the proportion of shared inferred hotspots decreased for younger events. This proportion was overall lower than for constant-size populations, but it sometimes reached, or even exceeded, the constant-N_e reference when events were ancient (Supplementary Figure S13 G, H).

Discussion

Inferred LD-maps should be interpreted with caution

Inference methods based on linkage disequilibrium provide an attractive way to characterise genomic recombination landscapes from sequence data. As such, they promise to become increasingly popular in empirical studies of eukaryotes. However, their ability to accurately reproduce real recombination landscapes has not been specifically evaluated. It should be recalled, however, that LD-based recombination maps are inferences, not observations; biases and uncertainty must be quantified and taken into account when it comes to interpreting the results. Here, we modelled the biological characteristics of a particularly well-documented recombination landscape, that of humans, as a basis for assessing the impact of methodological and species-specific demographic and evolutionary parameters on the performance of the LDhelmet method. Our results send a message of caution regarding the reliability of reconstructed recombination maps and hotspot location.

Indeed, we show that the recombination landscapes inferred with LDhelmet differ from real landscapes, sometimes substantially, with Spearman rank correlation between simulated and inferred 2.5 kb windows sometimes getting as low as 0.7 (Figure 2A, Supplementary Figure S11A). Hotspot detection is a particularly tricky and error-prone task: while up to 85% of true hotspots can be detected in the most favourable situations (N_e = 250,000, SS = 20, BP = 5, threshold = 2.5), the FDR ranged from 19% to 82% (Figure 3, Figure 4B and C, Supplementary Figure S9) according to the type of data and parameters used, meaning that in many cases a majority of the detected hotspots are incorrect calls. These discrepancies lead to a reduction in the apparent overlap in hotspot location between species/populations with identical recombination landscapes, while in turn inflating apparent hotspot sharing in populations with divergent landscapes (Figure 7, supplementary Table S2). These results were obtained with recombination landscapes of simulated populations with constant effective sizes, and with perfectly phased data. In reality, empirical data needs to be phased, and the phasing process can be prone to errors. Our analyses suggest that the typical error rate of statistical phasing methods such as Shapeit4 is relatively low (~ 6.7% in our simulations) and only marginally affects the performance of LDhelmet in terms of sensitivity and hotspot sharing (Supplementary Figure S10), which is quite reassuring. Moreover, studied populations often hide complex demographic histories that are known to impact the power to correctly infer recombination rates (Dapper and Payseur 2018, Samuk and Nook 2021). We showed that recent bottleneck and admixture events tend to decorrelate simulated and inferred landscapes, decrease the TPR and increase the FDR, thus increasing the difficulty to call

451 shared hotspots between populations sharing the same underlying landscape (Figure 5,
452 Supplementary Figure S13). The significant impact of non-equilibrium demographic histories
453 illustrated by our simulations provides additional motivation to characterise these histories in
454 comparative studies of recombination landscapes. These results also provide a qualitative
455 assessment of the impact of linked selection, whose effects may be similar to a local reduction
456 in N_e (i.e. purifying selection), or to the maintenance of anciently diverged alleles (i.e. balanced
457 selection, possibly involving structural variants such as chromosome inversions). If neglected,
458 these effects might mislead biological interpretations regarding the evolutionary conservation
459 of recombination maps.

460 In a study of the short time-scale dynamics of recombination landscapes based on
461 LDhelmet, Shanfelter et al. (2019) found only 15% of shared hotspots between two recently-
462 diverged populations of threespine stickleback. A greater overlap in hotspot location was *a*
463 *priori* expected given that this species lacks a functional PRDM9 protein, which is responsible
464 for the rapid turnover of recombination landscapes in mammals (Axelsson et al. 2012; Paigen
465 and Petkov 2018). The authors suggested that a new mechanism of recombination hotspot
466 regulation, different from the two already described in the literature, might be at play in this
467 teleost species. In the light of our results, however, one cannot exclude that the strong
468 divergence between the two reconstructed landscapes is due to a lack of power of the method
469 in the first place. While the sample size of both fish populations was at least 20 individuals, θ
470 was about 0.002, similar to our simulated conditions with a low N_e. Under these conditions, a
471 high FDR and a low proportion of shared hotspots can be expected even if the true underlying
472 maps are identical (Supplementary Figure S9, Supplementary Table S2).

473 It should be recalled that real data sets typically carry less signal and more noise than
474 simulated data sets, meaning that our assessment of the reliability of LDhelmet might be an
475 overoptimistic one. In particular, our data sets are immune from sequencing errors or mapping
476 errors, all of which presumably make the problem of recombination map inference an even
477 harder one.

478

479 ***Guidelines for population-based inference of recombination maps***

480 Our study revealed that whatever the parameters used, the inference of recombination
481 rates by LDhelmet is more reliable for species with large as compared to small effective
482 population size (Figure 4, 7, Supplementary Figure S9, S11). This might be expected since
483 long-term N_e determines the amount of nucleotide diversity ($\theta=4N_e\mu$, Watterson 1975), so that
484 a higher N_e results in a higher SNP density and a finer scale characterization of the
485 recombination rate variation along the genome. Moreover, a higher effective size greatly

corrects the general tendency of LDhelmet to overestimate the ρ value, especially for low and high recombination rates (Supplementary Figure S6, Singhal et al. 2015; Booker et al. 2017). Thus, when studying species with heterogeneous effective population sizes in nature, it is recommended to select populations with the largest N_e, for which genetic diversity is greater. The question is then: how to obtain a good-quality recombination map when dealing with low N_e species? The sampling effort also determines, to a lesser extent, the polymorphism level of the dataset (Supplementary Figure S5), improving the accuracy of the inference (Figure 2, 3 Supplementary Figure S9, S11). A sample size of 20 is recommended based on our simulations. Moreover, as previously mentioned (Chan et al. 2012; Singhal et al. 2015), the block penalty parameter of LDhelmet, which determines the resolution level of the inferred landscape, also influences the length of the inferred map (*i.e.* a higher BP tends to mitigate the tendency of LDhelmet to overestimate the map length) and the number of detected hotspots (Supplementary Figure S4, S8). A small BP, that allows more fine-scale changes in the inferred ρ value, should be used to detect recombination hotspots. The ability of LDhelmet to faithfully reflect the fine-scale variation of real recombination landscapes is of great importance when it comes to detecting recombination hotspots. To this purpose, the threshold used to decide which region is defined as a "hotspot" is a key parameter that determines the level of detection stringency. If the chosen value is not appropriate, LDhelmet will detect false positives while also missing true hotspots (Supplementary Figure S9). This threshold should thus be adapted to the species studied, using a less stringent threshold in species with lower genome-wide average recombination rate.

Other intrinsic biological variables influence the ability to produce a faithful recombination map, such as the μ/r ratio, which in part determines the power to measure ρ at a fine-scale. The among-species variations in genome-average recombination rate r is well documented, ranging from 0.01 to 100 cM/Mb in animals and plants, with vertebrate taxa displaying an average r around 1 cM/Mb (Stapley et al. 2017). As previously mentioned, high and low recombination rates tend to be overestimated by LDhelmet, thus the average r of the studied species is obviously a key parameter to account for. The mutation rate μ also has a key impact on the performance of LDhelmet, since ancestral recombination events can only be detected if properly tagged by flanking mutations. The variation in μ across taxa, and consequently the ratio of μ/r, are much less well known than the variation in r. This ratio, which does not depend on the effective size of the population, is about 1 in humans, which means that two recombination events are separated by one mutation on average. A ratio in favour of mutations ($\mu/r > 1$) will improve the signal, increasing the TP rate and reducing the FD rate (Figure 6, Supplementary Figure S12). But ultimately the performance of LDhelmet is

521 conditioned by *r*, as low *r* values provide less power to detect the recombination events, even
522 with μ/r = 10. Thus, the mutation to recombination rate ratio is crucial to build a non-biased
523 recombination map. When studying a species for which it appears that this ratio is not
524 favourable, a high rate of false positive hotspots is expected from the inferred population
525 recombination landscape (Figure 6), making it difficult to compare maps between closely
526 related species in a meaningful way.
527

528 ***Limitations***

529 The aim of our study was to determine the limits of LD-based methods in inferring
530 biologically realistic recombination landscapes. For this purpose, we used the Pratto et al.
531 (2014) ChIP-seq DMC1 data set to build human-like recombination landscapes including both
532 broad and fine scale variation, reflected by the presence of numerous recombination hotspots
533 of different intensities (Supplementary Figure S2, Myers et al. 2005, 2006; Pratto et al. 2014).
534 We therefore assumed that the distribution of DSB reflects the distribution of crossing overs,
535 which is not true for sure. For instance, hotspots were here placed without taking into account
536 the existence of genomic features that correlate with the recombination rate, such as genes
537 and promoter-like regions, GC-rich regions, CpG islands, and polymorphic regions, which can
538 explain why a very intense and narrow hotspot is never found within a region of near zero
539 recombination. The sensibility of LD-based methods with respect to this architecture was not
540 tested. We did not take into account the effect of gene conversion on the dissipation of LD in
541 high-recombining regions. While recent methods aim to distinguish between crossing-over
542 (CO) and non-crossing-over (NCO) events (heRho, Setter et al. 2022), they do not (yet)
543 account for the small-scale heterogeneity of recombination rates, and so are not really
544 applicable when it comes to differentiate hotspots and NCO. Our simulated data were perfectly
545 polarised, without missing or low-quality genotypes, which can't be the case when dealing
546 with empirical data. We simulated phasing errors in order to assess the robustness of
547 LDHelmet to this problem. However, we estimated the phasing error rate of Shapeit4 from our
548 simulated data which lack most of the biases found in empirical data, thus probably
549 underestimating the typical phasing error of this method.
550 Finally, we don't know if these simulated landscapes are representative of the diversity
551 of recombination landscapes that exist in the living world (*i.e.* PRDM9-dependent vs
552 independent landscapes, hotspot-free landscapes…). Indeed, it is likely that the high
553 complexity of the human recombination landscape is not a universal feature in the animal
554 kingdom. Singhal et al. (2015) used LDhelmet for building the recombination map in two
555 species of birds, the zebra finch and the long-tailed finch, that lack a full-length PRDM9 gene

556 copy and diverged about 2.9 Myr. The sample size for both populations was about 20
557 individuals, and θ (~ 0.01) was about ten times higher than in apes or the threespine
558 stickleback (Shanfelter et al. 2019), thus corresponding to our high N_e simulation conditions.
559 Singhal et al. (2015) found 73% of shared hotspots between the two finch species, which is
560 a higher rate of hotspot sharing than in any of the scenarios we simulated. The median
561 estimated recombination rate was 0.14 cM/Mb in both species of finch, which is seven times
562 lower than the genome-wide average recombination rate in humans (about 1 cM/Mb, Jensen-
563 Seaman et al. 2004). Combined with the strong polymorphism in those species, we may
564 suppose that birds possess less complex recombination landscapes than humans or
565 compared to what we simulated, which might explain why LDhelmet recovered such a high
566 percentage of shared hotspots in this study.
567
568

569 Conclusion

570
571 In the past few years, we have seen a growing interest in recombination rate estimation
572 in functional and evolutionary genomics. Indirect, LD-based approaches raise methodological
573 challenges that are addressed by sophisticated methods such as LDhat or LDhelmet, the
574 reliability of which is still poorly characterised. Our study provides guidance to users of these
575 methods based on the characteristics of their species, and calls for caution when it comes to
576 interpreting fine-scale differences in recombination rates between species. Extending this
577 approach to a more diverse set of underlying recombination landscapes would help
578 characterise further the reliability of these methods and their range of applicability across data
579 sets and taxa.
580
581

582 Materials and Methods

583
584 Our approach separately considers three different layers of information that are
585 involved in the study of recombination landscapes (Figure 1). The first layer that we call the
586 "underlying" recombination landscape corresponds to the true biological distribution of
587 recombination rate (*r*) across the considered genome. We here used experimental
588 measurements from human studies to model and generate the "underlying" landscapes. The
589 second layer, the population recombination landscape, describes the genomic location of
590 recombination events that occurred during the history of the sample. We used coalescent

591 simulations to produce these population recombination landscapes, thereafter called
592 "simulated" landscapes. Simulated landscapes differ from the underlying landscape due to
593 the stochasticity of the coalescent process, which is inversely proportional to N_e. The third
594 layer, called the "inferred" landscape, corresponds to the output of LDhelmet, *i.e.* an estimate
595 of the population recombination rate between adjacent SNPs. In total we generated five
596 independent replicates of underlying landscapes, and for each of them up to 10 simulated and
597 inferred landscapes under various demographic scenarios (Figure 1, Supplementary Figure
598 S3A).
599

600 ***Underlying landscapes***

601 Underlying recombination landscapes were first generated to reproduce the features
602 of the human recombination landscape. These include large-scale variation in the mean
603 background recombination rate and fine-scale variation reflecting the presence of hotspots
604 with varied intensities. Meiotic DSB are the major determinant of crossing over (CO) location
605 along the genome (Li et al. 2019; Pratto et al. 2014). We used the high-resolution map of
606 meiotic DSB obtained using ChIP-seq DMC1 in 5 non-related human genomes (Pratto et al.
607 2014) to define the genome-wide distribution of recombination rates in our simulations. The
608 five individuals analysed in Pratto et al. (2014) carried different PRDM9 genotypes totalizing
609 about 40,000 hotspots per individual, with distinct genotypes having different sets of DSB
610 hotspots. For each individual, a gamma distribution was fitted to the empirical distribution of
611 hotspot intensity measured by ChIP-seq DMC1 with the R package *figdistribplus* (Delignette-
612 Muller and Dutang 2015). Extreme ChIP-Seq intensity values (>500) lying above the 97.5th
613 quantile and likely representing technical artefacts were removed. Remaining values were
614 rescaled to 0-100, so as to transform ChIP-Seq intensity values into quantities reflecting the
615 range of recombination rates reported in cM/Mb across the human genome (McVean et al.
616 2004, 2005). This conversion assumed a linear relation between DMC1 activity and CO
617 frequency (Pratto et al. 2014). We then removed null values and replaced them with small but
618 non-null values (0.001), so that the genome-wide mean recombination rate equaled a target
619 average (*e.g.* 1 cM/Mb). A Gamma distribution was fitted to these transformed empirical values
620 separately for each of the 5 individuals, before averaging shape and scale parameters across
621 individuals. Targeted genome-wide average value was set to either 1 cM/Mb or 3 cM/Mb,
622 respectively reflecting the average centro-chromosomal and subtelomeric rates in humans.
623 Underlying landscapes of 1 Mb length were built by randomly drawing independent
624 recombination rate values from the fitted distribution and assigning these to non-overlapping
625 windows of 500 pb. Values in the first 500 kb were drawn from a distribution of mean 1 cM/Mb,

626 while values in the last 500 kb were drawn from a distribution of mean 3 cM/Mb. Our approach
627 thus mimics both the large scale variation in recombination rate existing in humans (Nachman,
628 2002; Myers et al. 2005; Buard and de Massy 2007; Pratto et al. 2014) and the nearly absence
629 of recombination events outside hotspots (96% of CO occur in hotspots in mice, Pratto et al.
630 2014; Li et al. 2019). In total, 5 underlying landscapes were generated (mean r = 2 cM/Mb),
631 which can be considered as independent replicates driven from the same distribution (*i.e.*
632 regions from different chromosomes of the same species, or orthologous chromosome region
633 from closely related species).
634

635 ***Simulated landscapes***

636 For each of the 5 underlying landscapes, 10 simulated landscapes were generated via
637 coalescent simulations using the program MSPRIME (v0.7.4, Kelleher et al. 2016), varying
638 the constant effective population size (N_e= 25,000 or 250,000) and the sample size (SS=10
639 or 20) and setting the mutation rate to μ = 10^-8. These sets of simulation parameters were
640 combined with three values of the Block Penalty (BP) parameter of the LDhelmet program
641 (see below), resulting in twelve conditions tested. For each combination of parameters, ten
642 population samples were simulated, to generate independent replicates of the coalescent
643 history (Figure 1A, Supplementary Figure S3).
644 Populations undergoing bottleneck and admixture events were also simulated with
645 MSPRIME, using one of the five underlying landscapes. The sample size parameter was set
646 to 20, and the N_e of the ancestral and present-day population was set to 250,000. The
647 simulated bottleneck scenarios varied according to the timing of the bottleneck event (t_b = 500,
648 5,000, 50,000 generations ago) and the N_e of the population during the bottleneck (N_b = 2,500,
649 25,000). The duration of the bottleneck was fixed to 1,000 generations. The admixture
650 scenarios varied in terms of the timing of the admixture event between the derived populations
651 1 and 2 (t_a = 500, 5,000, 50,000 generations ago). The time of split of the ancestral population
652 (0) into two derived population 1 and 2 (t_{split}) was set to 10*N_e generations. For each bottleneck
653 and admixture scenario, ten replicated populations were simulated (Figure 1A,B).
654 A VCF file was generated with MSPRIME for each simulated population (Figure 1A,
655 Supplementary Figure S3), which contains the genotypes of variants that segregate in the
656 population sample consisting of 2n sequences (with n being the number of diploid samples)
657 following the given underlying recombination landscape.
658 The impact of phasing errors on the inference of recombination rates was also
659 assessed. Two replicate VCFs, simulated using the same underlying landscape, were
660 manually dephased and then phased using Shapeit 4.2.2 (Delaneau et al., 2019). The phasing

661 error rate was computed using the --switch-error option of VCFtools 0.1.17 (Danecek et al.
662 2011 ; using the original phased VCF as a reference). Phasing error rates of 2, 4, 6, 8 and
663 10% between heterozygous positions of the two original phased VCFs were then randomly
664 generated, to produce five independent replicate VCF for each phasing error rate value.
665

666 ***Inferred landscapes***

667 Recombination rates were estimated for each of the simulated samples with LDhelmet
668 (v1.19, Chan et al. 2012, Figure 1). Briefly, LDhemet uses phased sequence data to infer the
669 ρ parameter locally, using likelihood computation between pairs of SNPs and then averaging
670 over 50 consecutive variants to obtain a composite likelihood. The ρ parameter is inferred with
671 a reversible-jump Markov Chain Monte Carlo algorithm using a step function applied to every
672 window of 50 consecutive SNPs and determined by three parameters: the number of change-
673 points, the locations of changes, and the recombination rate value of each constant fragment
674 between two changes. We used VCFtools 0.1.17 (Danecek et al. 2011) and the vcf2fasta
675 function of vcflib (https://github.com/vcflib/vcflib) to convert the SNP data obtained from
676 MSPRIME simulations into the input format to LDhelmet, consisting of FASTA sequences of
677 each individual haplotype. Ancestral states defined as the reference allele of each variant
678 were also used as inputs. Each simulated replicate was analysed with LDhelmet using the
679 following parameters. The haplotype configuration files were created with the find_conf
680 function using the recommended window size of 50 SNPs. The likelihood look-up tables were
681 created with the table_gen function using the recommended grid for the population
682 recombination rate (ρ/pb) (*i.e.* ρ from 0 to 10 by increments of 0.1, then from 10 to 100 by
683 increments of 1), and with the Watterson $\theta = 4N_e\mu$ parameter corresponding to the condition
684 analysed. The Padé files were created using 11 Padé coefficients as recommended. The
685 Monte Carlo Markov chain was run for 1 million iterations with a burn-in period of 100,000 and
686 a window size of 50 SNPs. An important parameter to LDhelmet is the block penalty (BP),
687 which determines the number of change-points, and thus the variance of the inferred
688 recombination rates at a fine scale (*i.e.* smaller block penalty values correspond to a lower
689 penalty for background rate changes, and thus generate more heterogeneous recombination
690 landscapes). For each simulated combination of N_e and SS for populations with constant size,
691 the block penalty was set to either 5, 10 or 50, and for the simulated populations undergoing
692 bottleneck or admixture event, as for the ”phase-error” datasets, the block penalty used was
693 set to 5. Finally, the population recombination rates between each SNP pair were extracted
694 with the post_to_text function, and were reported in $\rho=4N_er$ per pb unit.
695 The reliability of the inferences was evaluated in various ways. For each combination

696 of N_e, SS, BP parameters and demographic scenarios simulated, the inferred, simulated and
697 underlying landscapes were compared, in order to assess the ability of LDhelmet to reliably
698 infer the true biological landscape (Figure 1, Supplementary Figure S3A, and see below
699 Hotspot detection and Statistical Analysis). In order to evaluate the convergence of LDhelmet
700 inferences across replicate runs, LDhelmet was run 10 times, for the 12 combinations of the
701 parameters N_e, SS and BP, on two independently simulated VCF files from constant-size
702 populations sharing the same underlying landscape (Figure 1A, Supplementary Figure S3B).
703 Finally, the inferred recombination landscapes of pairs of populations sharing the same
704 underlying landscape were compared in order to assess the reproducibility of the LDhelmet
705 inference, *i.e.*, the expected variance between inferred maps in the absence of underlying
706 biological variation (Figure 1, Supplementary Figure S3B).
707

708 ### *Variation in the μ/r ratio*

709 To explore the influence of variation in mutation and recombination rates on the
710 inference of recombination maps, two additional underlying landscapes were generated using
711 the same procedure, this time targeting a ten times higher (i.e. 20 cM/Mb) or ten times lower
712 (0.2 cM/Mb) mean recombination rate. Then, using one of the 5 underlying landscapes ($r \sim$
713 10^-8 M/pb) and the 2 newly generated landscapes with mean r = 10^-7 and 10^-9 M/pb,
714 respectively, sets of simulations were run with a μ/r ratio of 0.1, 1 and 10. This was achieved
715 by fixing μ to either 10^-9, 10^-8 or 10^-7, while keeping a fixed N_e = 100000 and SS = 20
716 (Supplementary Table S1). For each of the tested combinations of μ and r, 10 populations
717 were simulated. These simulated landscapes were inferred with LDhelmet, using a block
718 penalty of 5.
719

720 ### *Hotspot detection*

721 Underlying and simulated landscapes were first converted into population
722 recombination rate landscapes by scaling them by $4N_e$. Underlying, simulated and inferred
723 landscapes were then smoothed at a 500 bp and 2500 bp resolution using the Python package
724 *scipy.stats*. The former corresponds to the underlying landscape resolution, and the latter to
725 a trade off between the density of segregating sites and the resolution often used in the
726 literature. For the different combinations of N_e, SS and BP of the constant-size populations
727 simulated with the 5 underlying landscapes, a mean simulated landscape and a mean inferred
728 landscape were generated by averaging recombination rates across replicates.
729 Recombination hotspots of the underlying, simulated and inferred landscapes were
730 called by comparing local vs surrounding recombination rates at each genomic window. A

hotspot was defined as a window of 2.5 kb with an average recombination rate either 2.5, 5 or 10 times higher than the 50 kb flanking regions (excluding the focal window). Hotspot locations were then compared among landscapes using the same threshold values (Figure 1C)

Statistical analyses

Statistical analyses were run with R 4.0.3. The length of underlying, simulated and inferred maps (*L*) was calculated at the 2.5 kb resolution using the formula:

$$L = \sum \quad (\rho \times win)/4Ne,$$

with ρ the population-scaled recombination rate, *win* the window size resolution used to smooth the maps in bp, and N_e the effective size of the simulated population. Several measures of the sensitivity, specificity, reliability, and repeatability of LDhelmet were computed, using the mean simulated and inferred landscapes of the constant-size populations, and replicates of simulated and inferred landscapes of the populations that underwent bottleneck of admixture events. Spearman rank correlation coefficients were calculated between the underlying and the corresponding simulated landscapes, between the simulated and inferred landscapes, and pairwise Spearman coefficients among the 10 replicates inferred from the two simulated populations sharing the same underlying landscape. True/false positive rates (TPR = TP/(TP+FN) ; FPR = FP/(FP+TN)), true/false negative rates (TNR = TN/(TN+FP) ; FNR = FN/(FN/(TP+FN)), and true/false discovery rates (TDR = TP/(TP+FP) ; FDR = FP/(TP+FP)) were calculated by comparing the simulated and inferred landscapes. The mean pairwise linear correlation (R^2) and the proportion of shared hotspots was calculated between the 5 underlying landscapes, and for each condition simulated with a constant population size scenario, and for the three threshold values tested (*i.e.* 2.5, 5 and 10) between the simulated and inferred landscapes from the 5 different underlying landscapes, as well as between the pairs of populations sharing the same underlying landscape.

The statistical analyses were performed using home-made R scripts available upon request.

Data, script and code availability

The underlying landscapes, the main scripts used to generate the underlying landscapes, run the simulations under the various demographic scenarios, infer the simulated landscapes with LDhelmet, and call hotspots from the landscapes can be found at https://github.com/marie-raynaud/Performance_LD_methods_recombination. The Singularity

container recipe built to run the simulations is available at:https://github.com/marie-raynaud/Singularity-Recipes/tree/master/HotRec-Recipes. This recipe contains the installation command lines of the required programs, the scripts used for the simulations, and the five underlying landscapes used in our study. More scripts and data are available upon request.

Supplementary information

Supplementary Materials, containing Supplementary Figures and Tables are available at https://doi.org/10.5281/zenodo.7544555.

Acknowledgments

We are grateful to Julien Joseph, Nicolas Lartillot, Frédéric Baudat, Bernard de Massy and Laurent Duret for the helpful discussions and feedback. We also thank Khalid Belkhir, Jimmy Lopez and Mathieu Massaviol for their support to build the Singularity container used to run our simulations.

Funding

This project was funded by the ANR HotRec ANR-19-CE12-0019.

Conflict of interest disclosure

The authors declare no conflict of interest.

References

Arenas, Miguel. 2013. « The Importance and Application of the Ancestral Recombination Graph ». *Frontiers in Genetics* 4. https://doi.org/10.3389/fgene.2013.00206.

Auton, Adam, Adi Fledel-Alon, Susanne Pfeifer, Oliver Venn, Laure Ségurel, Teresa Street, Ellen M. Leffler, et al. 2012. « A Fine-Scale Chimpanzee Genetic Map from Population Sequencing ». *Science* 336 (6078): 193-98. https://doi.org/10.1126/science.1216872.

801 Auton, Adam, et Gil McVean. 2007. « Recombination Rate Estimation in the Presence of
802 Hotspots ». *Genome Research* 17 (8): 1219-27. https://doi.org/10.1101/gr.6386707.

803 Auton, Adam, Ying Rui Li, Jeffrey Kidd, Kyle Oliveira, Julie Nadel, J. Kim Holloway, Jessica
804 J. Hayward, et al. 2013. « Genetic Recombination Is Targeted towards Gene
805 Promoter Regions in Dogs ». Édité par Ian Henderson. *PLoS Genetics* 9 (12):
806 e1003984. https://doi.org/10.1371/journal.pgen.1003984.

807 Axelsson, Erik, Matthew T. Webster, Abhirami Ratnakumar, The LUPA Consortium, Chris P.
808 Ponting, et Kerstin Lindblad-Toh. 2012. « Death of *PRDM9* Coincides with
809 Stabilization of the Recombination Landscape in the Dog Genome ». *Genome*
810 *Research* 22 (1): 51-63. https://doi.org/10.1101/gr.124123.111.

811 Baker, Zachary, Molly Schumer, Yuki Haba, Lisa Bashkirova, Chris Holland, Gil G
812 Rosenthal, et Molly Przeworski. 2017. « Repeated Losses of PRDM9-Directed
813 Recombination despite the Conservation of PRDM9 across Vertebrates ». *ELife* 6
814 (juin): e24133. https://doi.org/10.7554/eLife.24133.

815 Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober, M. Przeworski, G. Coop, et B.
816 de Massy. 2010. « PRDM9 Is a Major Determinant of Meiotic Recombination
817 Hotspots in Humans and Mice ». *Science* 327 (5967): 836-40.
818 https://doi.org/10.1126/science.1183439.

819 Baudat, Frédéric, et Bernard de Massy. 2007. « Regulating Double-Stranded DNA Break
820 Repair towards Crossover or Non-Crossover during Mammalian Meiosis ».
821 *Chromosome Research* 15 (5): 565-77. https://doi.org/10.1007/s10577-007-1140-3.

822 Booker, Tom R, Rob W Ness, et Peter D Keightley. 2017. « The Recombination Landscape
823 in Wild House Mice Inferred Using Population Genomic Data ». *Genetics* 207 (1):
824 297-309. https://doi.org/10.1534/genetics.117.300063.

825 Buard, Jérôme, et Bernard de Massy. 2007. « Playing Hide and Seek with Mammalian
826 Meiotic Crossover Hotspots ». *Trends in Genetics* 23 (6): 301-9.
827 https://doi.org/10.1016/j.tig.2007.03.014.

828 Capilla, Laia, Montserrat Garcia Caldés, et Aurora Ruiz-Herrera. 2016. « Mammalian Meiotic
829 Recombination: A Toolbox for Genome Evolution ». *Cytogenetic and Genome*
830 *Research* 150 (1): 1-16. https://doi.org/10.1159/000452822.

831 Carrington, Mary, et Michael Cullen. 2004. « Justified Chauvinism: Advances in Defining
832 Meiotic Recombination through Sperm Typing ». *Trends in Genetics* 20 (4): 196-205.
833 https://doi.org/10.1016/j.tig.2004.02.006.

834 Chan, Andrew H., Paul A. Jenkins, et Yun S. Song. 2012. « Genome-Wide Fine-Scale
835 Recombination Rate Variation in Drosophila Melanogaster ». Édité par Gil McVean.
836 *PLoS Genetics* 8 (12): e1003090. https://doi.org/10.1371/journal.pgen.1003090.

837 Choi, Kyuha, et Ian R. Henderson. 2015. « Meiotic Recombination Hotspots - a Comparative
838 View ». *The Plant Journal* 83 (1): 52-61. https://doi.org/10.1111/tpj.12870.

839 Clark, Andrew G., Xu Wang, et Tara Matise. 2010. « Contrasting Methods of Quantifying
840 Fine Structure of Human Recombination ». *Annual Review of Genomics and Human*
841 *Genetics* 11 (1): 45-64. https://doi.org/10.1146/annurev-genom-082908-150031.

Cromie, Gareth A., John C. Connelly, et David R.F. Leach. 2001. « Recombination at Double-Strand Breaks and DNA Ends ». *Molecular Cell* 8 (6): 1163-74. https://doi.org/10.1016/S1097-2765(01)00419-1.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, et al. 2011. « The Variant Call Format and VCFtools ». *Bioinformatics* 27 (15): 2156-58. https://doi.org/10.1093/bioinformatics/btr330.

Danguy des Déserts, Alice, Sophie Bouchet, Pierre Sourdille, et Bertrand Servin. 2021. « Evolution of Recombination Landscapes in Diverging Populations of Bread Wheat ». Édité par Brandon Gaut. *Genome Biology and Evolution*, juin, evab152. https://doi.org/10.1093/gbe/evab152.

Dapper, Amy L, et Bret A Payseur. 2018. « Effects of Demographic History on the Detection of Recombination Hotspots from Linkage Disequilibrium ». *Molecular Biology and Evolution* 35 (2): 335-53. https://doi.org/10.1093/molbev/msx272.

Delaneau, O., Zagury, J. F., Robinson, M. R., Marchini, J. L., & Dermitzakis, E. T. 2019. « Accurate, scalable and integrative haplotype estimation ». *Nature communications* 10(1): 1-10. https://doi.org/10.1038/s41467-019-13225-y.

Delignette-Muller, Marie Laure, et Christophe Dutang. 2015. « **Fitdistrplus** : An *R* Package for Fitting Distributions ». *Journal of Statistical Software* 64 (4). https://doi.org/10.18637/jss.v064.i04.

Dréau, Andreea, Vrinda Venu, Elena Avdievich, Ludmila Gaspar, et Felicity C. Jones. 2019. « Genome-Wide Recombination Map Construction from Single Individuals Using Linked-Read Sequencing ». *Nature Communications* 10 (1): 4309. https://doi.org/10.1038/s41467-019-12210-9.

Ellegren, Hans, et Nicolas Galtier. 2016. « Determinants of Genetic Diversity ». *Nature Reviews Genetics* 17 (7): 422-33. https://doi.org/10.1038/nrg.2016.58.

Grey, Corinne, Frédéric Baudat, et Bernard de Massy. 2018. « PRDM9, a Driver of the Genetic Map ». Édité par Paula E. Cohen. *PLOS Genetics* 14 (8): e1007479. https://doi.org/10.1371/journal.pgen.1007479.

Haenel, Quiterie, Telma G. Laurentino, Marius Roesti, et Daniel Berner. 2018. « Meta-Analysis of Chromosome-Scale Crossover Rate Variation in Eukaryotes and Its Significance to Evolutionary Genomics ». *Molecular Ecology* 27 (11): 2477-97. https://doi.org/10.1111/mec.14699.

Hunter, Chad M., Wen Huang, Trudy F. C. Mackay, et Nadia D. Singh. 2016. « The Genetic Architecture of Natural Variation in Recombination Rate in Drosophila Melanogaster ». Édité par Jeff Sekelsky. *PLOS Genetics* 12 (4): e1005951. https://doi.org/10.1371/journal.pgen.1005951.

Jensen-Seaman, Michael I., Terrence S. Furey, Bret A. Payseur, Yontao Lu, Krishna M. Roskin, Chin-Fu Chen, Michael A. Thomas, David Haussler, et Howard J. Jacob. 2004. « Comparative Recombination Rates in the Rat, Mouse, and Human Genomes ». *Genome Research* 14 (4): 528-38. https://doi.org/10.1101/gr.1970304.

Kamm, John A, Jeffrey P Spence, Jeffrey Chan, et Yun S Song. 2016. « Two-Locus

Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation ». *Genetics* 203 (3): 1381-99. https://doi.org/10.1534/genetics.115.184820.

Kelleher J, Etheridge AM, McVean G. 2016. “Efficient Coalescent Simulation and Genealogical Analysis for Large Sample Sizes”. PLoS Comput Biol 12(5): e1004842. https://doi.org/10.1371/journal.pcbi.1004842

Kim, Sung, Vincent Plagnol, Tina T Hu, Christopher Toomajian, Richard M Clark, Stephan Ossowski, Joseph R Ecker, Detlef Weigel, et Magnus Nordborg. 2007. « Recombination and Linkage Disequilibrium in Arabidopsis Thaliana ». *Nature Genetics* 39 (9): 1151-55. https://doi.org/10.1038/ng2115.

Kodama, Miyako, Marine S O Brieuc, Robert H Devlin, Jeffrey J Hard, et Kerry A Naish. 2014. « Comparative Mapping Between Coho Salmon (*Oncorhynchus Kisutch*) and Three Other Salmonids Suggests a Role for Chromosomal Rearrangements in the Retention of Duplicated Regions Following a Whole Genome Duplication Event ». *G3 Genes|Genomes|Genetics* 4 (9): 1717-30. https://doi.org/10.1534/g3.114.012294.

Kong, Augustine, Daniel F. Gudbjartsson, Jesus Sainz, Gudrun M. Jonsdottir, Sigurjon A. Gudjonsson, Bjorgvin Richardsson, Sigrun Sigurdardottir, et al. 2002. « A High-Resolution Recombination Map of the Human Genome ». *Nature Genetics* 31 (3): 241-47. https://doi.org/10.1038/ng917.

Lander, E. S., et P. Green. 1987. « Construction of Multilocus Genetic Linkage Maps in Humans. » *Proceedings of the National Academy of Sciences* 84 (8): 2363-67. https://doi.org/10.1073/pnas.84.8.2363.

Latrille, Thibault, Laurent Duret, et Nicolas Lartillot. 2017. « The Red Queen Model of Recombination Hot-Spot Evolution: A Theoretical Investigation ». *Philosophical Transactions of the Royal Society B: Biological Sciences* 372 (1736): 20160463. https://doi.org/10.1098/rstb.2016.0463.

Lesecque, Yann, Sylvain Glémin, Nicolas Lartillot, Dominique Mouchiroud, et Laurent Duret. 2014. « The Red Queen Model of Recombination Hotspots Evolution in the Light of Archaic and Modern Human Genomes ». Édité par Graham Coop. *PLoS Genetics* 10 (11): e1004790. https://doi.org/10.1371/journal.pgen.1004790.

Li, Na, et Matthew Stephens. s. d. « Modeling Linkage Disequilibrium and Identifying Recombination Hotspots Using Single-Nucleotide Polymorphism Data », 22.

Li, Ran, Emmanuelle Bitoun, Nicolas Altemose, Robert W. Davies, Benjamin Davies, et Simon R. Myers. 2019. « A High-Resolution Map of Non-Crossover Events Reveals Impacts of Genetic Diversity on Mammalian Meiotic Recombination ». *Nature Communications* 10 (1): 3900. https://doi.org/10.1038/s41467-019-11675-y.

Mancera, Eugenio, Richard Bourgon, Alessandro Brozzi, Wolfgang Huber, et Lars M. Steinmetz. 2008. « High-Resolution Mapping of Meiotic Crossovers and Non-Crossovers in Yeast ». *Nature* 454 (7203): 479-85. https://doi.org/10.1038/nature07135.

McVean, Gilean A. T., Simon R. Myers, Sarah Hunt, Panos Deloukas, David R. Bentley, et Peter Donnelly. 2004. « The Fine-Scale Structure of Recombination Rate Variation in

the Human Genome ». *Science* 304 (5670): 581-84. https://doi.org/10.1126/science.1092500.

Melamed-Bessudo, Cathy, Shay Shilo, et Avraham A Levy. 2016. « Meiotic Recombination and Genome Evolution in Plants ». *Current Opinion in Plant Biology* 30 (avril): 82-87. https://doi.org/10.1016/j.pbi.2016.02.003.

Morris, G. P., P. Ramu, S. P. Deshpande, C. T. Hash, T. Shah, H. D. Upadhyaya, O. Riera-Lizarazu, et al. 2013. « Population Genomic and Genome-Wide Association Studies of Agroclimatic Traits in Sorghum ». *Proceedings of the National Academy of Sciences* 110 (2): 453-58. https://doi.org/10.1073/pnas.1215985110.

Myers, S, C C A Spencer, A Auton, L Bottolo, C Freeman, P Donnelly, et G McVean. 2006. « The Distribution and Causes of Meiotic Recombination in the Human Genome ». *Biochemical Society Transactions* 34: 5.

Myers, Simon, Leonardo Bottolo, Colin Freeman, Gil McVean, et Peter Donnelly. 2005. « A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome ». *Science* 310 (5746): 321-24. https://doi.org/10.1126/science.1117196.

Myers, Simon, Rory Bowden, Afidalina Tumian, Ronald E. Bontrop, Colin Freeman, Tammie S. MacFie, Gil McVean, et Peter Donnelly. 2010. « Drive Against Hotspot Motifs in Primates Implicates the *PRDM9* Gene in Meiotic Recombination ». *Science* 327 (5967): 876-79. https://doi.org/10.1126/science.1182363.

Nachman, M. 2002. « Variation in Recombination Rate across the Genome: Evidence and Implications ». *Current Opinion in Genetics & Development* 12 (6): 657-63. https://doi.org/10.1016/S0959-437X(02)00358-1.

Oliver, Peter L., Leo Goodstadt, Joshua J. Bayes, Zoë Birtle, Kevin C. Roach, Nitin Phadnis, Scott A. Beatson, Gerton Lunter, Harmit S. Malik, et Chris P. Ponting. 2009. « Accelerated Evolution of the Prdm9 Speciation Gene across Diverse Metazoan Taxa ». Édité par Michael W. Nachman. *PLoS Genetics* 5 (12): e1000753. https://doi.org/10.1371/journal.pgen.1000753.

Paigen, Kenneth, et Petko M. Petkov. 2018. « PRDM9 and Its Role in Genetic Recombination ». *Trends in Genetics* 34 (4): 291-300. https://doi.org/10.1016/j.tig.2017.12.017.

Parvanov, Emil D., Petko M. Petkov, et Kenneth Paigen. 2010. « *Prdm9* Controls Activation of Mammalian Recombination Hotspots ». *Science* 327 (5967): 835-835. https://doi.org/10.1126/science.1181495.

Peñalba, Joshua V., et Jochen B. W. Wolf. 2020. « From Molecules to Populations: Appreciating and Estimating Recombination Rate Variation ». *Nature Reviews Genetics* 21 (8): 476-92. https://doi.org/10.1038/s41576-020-0240-1.

Pratto, Florencia, Kevin Brick, Pavel Khil, Fatima Smagulova, Galina V. Petukhova, et R. Daniel Camerini-Otero. 2014. « Recombination Initiation Maps of Individual Human Genomes ». *Science* 346 (6211): 1256442. https://doi.org/10.1126/science.1256442.

Rastas, Pasi. 2017. « Lep-MAP3: Robust Linkage Mapping Even for Low-Coverage Whole Genome Sequencing Data ». Édité par Bonnie Berger. *Bioinformatics* 33 (23):

3726-32. https://doi.org/10.1093/bioinformatics/btx494.

Samuk, K., & Noor, M. A. 2022. « Gene flow biases population genetic inference of recombination rate ». *G3* 12(11), jkac236. https://doi.org/10.1093/g3journal/jkac236.

Shanfelter, Alice F, Sophie L Archambeault, et Michael A White. 2019. « Divergent Fine-Scale Recombination Landscapes between a Freshwater and Marine Population of Threespine Stickleback Fish ». Édité par Laurence Hurst. *Genome Biology and Evolution* 11 (6): 1552-72. https://doi.org/10.1093/gbe/evz090.

Singhal, S., E. M. Leffler, K. Sannareddy, I. Turner, O. Venn, D. M. Hooper, A. I. Strand, et al. 2015. « Stable Recombination Hotspots in Birds ». *Science* 350 (6263): 928-32. https://doi.org/10.1126/science.aad0843.

Spence, Jeffrey P, et Yun S Song. 2019. « Inference and Analysis of Population-Specific Fine-Scale Recombination Maps across 26 Diverse Human Populations ». *SCIENCE ADVANCES*, 15.

Stapley, Jessica, Philine G. D. Feulner, Susan E. Johnston, Anna W. Santure, et Carole M. Smadja. 2017. « Variation in Recombination Frequency and Distribution across Eukaryotes: Patterns and Processes ». *Philosophical Transactions of the Royal Society B: Biological Sciences* 372 (1736): 20160455. https://doi.org/10.1098/rstb.2016.0455.

Stumpf, Michael P. H., et Gilean A. T. McVean. 2003. « Estimating Recombination Rates from Population-Genetic Data ». *Nature Reviews Genetics* 4 (12): 959-68. https://doi.org/10.1038/nrg1227.

Sun, Hequan, Beth A. Rowan, Pádraic J. Flood, Ronny Brandt, Janina Fuss, Angela M. Hancock, Richard W. Michelmore, Bruno Huettel, et Korbinian Schneeberger. 2019. « Linked-Read Sequencing of Gametes Allows Efficient Genome-Wide Analysis of Meiotic Recombination ». *Nature Communications* 10 (1): 4310. https://doi.org/10.1038/s41467-019-12209-2.

The International HapMap Consortium. 2007. « A Second Generation Human Haplotype Map of over 3.1 Million SNPs ». *Nature* 449 (7164): 851-61. https://doi.org/10.1038/nature06258.

Wallberg, Andreas, Sylvain Glémin, et Matthew T. Webster. 2015. « Extreme Recombination Frequencies Shape Genome Variation and Evolution in the Honeybee, Apis Mellifera ». Édité par Nick H. Barton. *PLOS Genetics* 11 (4): e1005189. https://doi.org/10.1371/journal.pgen.1005189.

Watterson, G.A. 1975. « On the Number of Segregating Sites in Genetical Models without Recombination ». *Theoretical Population Biology* 7 (2): 256-76. https://doi.org/10.1016/0040-5809(75)90020-9.