Check for updates

From molecules to populations: appreciating and estimating recombination rate variation

Joshua V. Peñalba ✉ *and Jochen B. W. Wolf* ✉

Abstract | Recombination is a central biological process with implications for many areas in the life sciences. Yet we are only beginning to appreciate variation in the recombination rate along the genome and among individuals, populations and species. Spurred by technological advances, we are now able to bring variation in this key biological parameter to centre stage. Here, we review the conceptual implications of recombination rate variation and guide the reader through the assumptions, strengths and weaknesses of genomic inference methods, including population-based, pedigree-based and gamete-based approaches. Appreciation of the differences and commonalities of these approaches is a prerequisite to formulate a unifying and comparative framework for understanding the molecular and evolutionary mechanisms shaping, and being shaped by, recombination.

Genetic drift
A stochastic change in allele frequency from one generation to the next due to random sampling in finite populations.

Meiotic recombination is a key cellular process shared across eukaryotic life (BOX 1). Alongside other DNA modifiers, such as mutation, it has had a marked impact on the evolutionary history of life. Building on molecular machinery that ensured DNA replication fidelity in prokaryotes, meiotic recombination became intrinsically linked to sex involving cellular fusion, fertilization and meiosis in the evolution of eukaryotes[1]. While still guarding against replication hazards, meiotic recombination has contributed substantially to evolutionary innovation, shaping central aspects of the living world today. Meiotic recombination plays a pivotal role in generating phenotypic novelty[2], is a central component of genetic trait architecture[3], facilitates adaptation[4], is key to the evolution of new species[5] and can prime the success of sexual reproduction[6]. It is thus not surprising that besides mutation, selection, genetic drift and migration, recombination is one of the five central parameters considered by evolutionary theory. Yet despite the theoretical recognition and detailed knowledge of the molecular machinery[7,8], empirical insight into recombination rate variation has been slowed by technological limitations.

This Review is broadly targeting geneticists, breeders, ecologists and evolutionary biologists working on systems for which accessing the variation in recombination rate has previously been challenging. Its objective is twofold. First, as we advance in the development of sequencing technology and analytical methods across many disciplines, we feel that it is timely to bring the variation in recombination rate to centre stage and incorporate it into our working hypotheses and models. Although recent reviews exist delving into the topic of recombination variation across organismal, spatial and temporal scales[9,10], here we discuss the recombination rate as a trait with an underlying variation that is subject to evolutionary change. Second, we aim to provide guidance to empiricists, as rapid advances in genomic technology open unprecedented access to quantify this central parameter. We review three approaches to infer recombination rate variation from genetic variation derived from high-throughput sequencing data: population-based, pedigree-based and gamete-based. We compare the underlying assumptions, typical study design and available software, and we detail the advantages and pitfalls of each. Current technology allows inferences to be expanded beyond human studies[11] and genetic model systems, thus holding promise to capture the processes interacting with recombination in captive or natural populations of organisms for which ample genetic resources may not yet be established. This opportunity to obtain insight into recombination rate variation across a broad taxonomic scale is reflected in our choice of literature, which is purposely focused towards genetic non-model systems. Considering the increasing ease of estimating recombination rates, we end by discussing the opportunities that arise by quantifying recombination variation across levels of biological organization (FIG. 1) and by combining the strengths of the various inference approaches in a unified framework (TABLE 1).

Division of Evolutionary Biology, LMU Munich, Planegg-Martinsried, Germany.

✉*e-mail: josh.penalba@gmail.com; j.wolf@bio.lmu.de*

https://doi.org/10.1038/s41576-020-0240-1

Box 1 | **What is meiotic recombination?**

Recombination, in the broad sense, refers to the new arrangement of genetic information from one generation to the next. In principle, this includes random Mendelian segregation of entire chromosomes, as well as genetic exchange between homologous chromosomes. In this Review, we focus on the latter definition in the narrow sense; that is, the reciprocal exchange of genetic information between homologous chromosomes during meiosis. Meiotic recombination is initiated by the formation of double-strand breaks (DSBs) across the length of the chromosome[8]. Although most of these DSBs are simply repaired and resolve in non-crossover events (including gene conversion), some DSBs result in crossover events with reciprocal exchange between the homologous parental haplotypes[8,9]. Along with any de novo mutations and gene conversion, this resulting new combination of haplotypes is packaged into the gametes and passed down to the next generation. As a consequence, genotypic combinations in the offspring are different from those in the parents.
In many diploid, sexual organisms, at least one recombination event per chromosome is necessary for proper segregation, although many more can occur depending on the species[38]. Additionally, not every location in the genome has an equal probability of forming DSBs, not every DSB has an equal probability of resulting in a crossover and not every recombinant haplotype is passed on to the next generation. Characterizing this variation within the genome, among individuals and among species is the first step in fully realizing the impact of recombination.

Why study recombination?

Mutation and recombination are the main sources of heritable phenotypic variation that selection can act upon[12]. Mutation generates new genetic variants, which recombination shuffles into novel allelic combinations, or haplotypes, while maintaining gene order[13–15] (BOX 1). Recombination can bring alleles from different genomic backgrounds into novel, and potentially advantageous, combinations, which allow for the evolution of complex phenotypes or behaviours requiring the concerted action of several genes. Similarly, recombination governs the association of meiotic drivers and their antidotes with direct relevance to reproductive isolation[16]. Additionally, recombination can influence the efficacy of selection[6,17,18] and prevent the accumulation of deleterious alleles[19]. It can facilitate adaptation to a changing and dynamic environment[20,21] but likewise can break up advantageous combinations that already exist[22]. Recombination may, in addition, exert mutagenic effects[23] and induce a shift in base pair composition[24,25]. All these processes are of immediate relevance to evolution, making recombination one of the key parameters influencing the evolutionary trajectory of organisms.

Importantly, recombination does not act in a simple cause-and-effect manner. It is involved in complex dynamics with other evolutionary forces, feeding back on the recombination rate itself[9]. In concert with selection and migration, recombination influences the degree of gene flow between diverging populations[26,27] and is of central importance to speciation[28,29]. It modulates the effects of selection on linked genetic variants and thereby influences the distribution of genetic variation across the genome[30–32] and the variation in genetic diversity across organisms[33]. In turn, the recombination rate itself is subject to evolution by means of natural and sexual selection. Examples are diverse and include selection for increased recombination in the human leukocyte antigen gene complex for effective immune response[34], as well as selection for decreased recombination to resolve sexual antagonism in the context of sex chromosome evolution[35,36], or favouring the formation of adaptive gene complexes in supergenes[37]. Given the interplay of recombination with a broad variety of biological processes, it is important to understand to what extent variation in recombination shapes, and is itself shaped by, evolution.

Gene flow
The movement of chromosomes across genetically structured populations.

Genetic diversity
The total number of non-identical genetic elements in a population.

Supergenes
A set of co-inherited, physically linked genes often contributing to complex phenotypes.

Heterochiasmy
A variation in recombination rates and, accordingly, genetic map length between sexes.

Heterogametic sex
The sex in which the sex chromosomes differ.

Structural variants
Any variations in the structure of a chromosome, including insertions, deletions, duplications, inversions or translocations.

How does recombination vary?

Despite being an evolutionarily conserved process that is central to meiosis, the recombination rate varies across multiple levels[38] (FIG. 1). Recombination rates can vary genome-wide in the total number of homologous crossovers per genome, or locally in the distribution of recombination frequency along the genome. The genome-wide recombination rate varies among the deepest branches of the evolutionary tree[38] (that is, between fungi and animals or plants), among recently diverged species[39,40], among subspecies, populations or domestic breeds within the same species[41–43] and among individuals[44,45]. Even within the same individual, rates of recombination can be influenced by a multitude of ontogenetic and environmental factors, such as age, temperature and pathogen infection[46,47]. Finally, the degree by which recombination rates differ between sexes (heterochiasmy) is considerable[48,49]. Whereas females generally tend to have higher recombination rates than males, this variation ranges all the way from equal rates of recombination to one sex having no recombination at all (achiasmy[49]), usually the heterogametic sex (Haldane–Huxley rule[50,51]).

There are a multitude of ways that recombination can vary within a single genome (FIG. 1). When considering chromosome types, the net recombination rate in a sex chromosome is generally lower than an autosome of similar size[52,53]. Within each chromosome, recombination tends to be lower around the centromeres[54,55] and higher towards the telomeres[56]. The recombination rate also varies locally in relation to structural variants[57,58], GC content[59,60], gene or exon density[61] and transposable element content[62–64]. Furthermore, there is a whole body of literature screening for recombination hotspots, which are highly localized, short regions (1–2 kb) where most recombination events occur[65–67]. We will not be focusing on recombination hotspot inference here and defer the interested reader to previous reviews (REFS[68–70]). Studies on the mechanisms underlying meiotic recombination may help shed light on whether these correlations are causative, a direct consequence of or a by-product of recombination variation.

Why does recombination vary?

Known causes underlying recombination rate variation can be classified into two types: ultimate (that is, referring to fitness consequences and, hence, the effects of selection) and proximate (that is, referring to proximate molecular mechanisms and constraints).

Ultimate causes

Recombination has both a heritable genetic component and a plastic component[71–73]. Therefore, recombination rate variation is subject to long-term selection[45,74,75] but is likewise influenced by environmental factors experienced by the individual[76,77]. These environmental

Hill–Robertson interference
The reduction in the efficacy of selection when acting on alleles in physically linked genes.

Double-strand breaks
(DSBs). Lesions in the DNA double helix induced by a wide range of DNA-damaging agents. Programmed enzymatic induction during meiosis triggers homologous recombination.

Crossover interference
The non-random occurrence of multiple crossover events relative to each other during meiosis.

factors include temperature extremes[76,78], food deprivation[79] and behavioural stress[80]. Heritability estimates of genome-wide recombination rates range from 0.16–0.17 in chickens[71] to 0.23 in Soay sheep[45] and 0.3 in humans[81]. The heritable component of recombination rates can be influenced by selection acting directly on the individual, where suboptimal recombination rates result in improper segregation and are detrimental to the individual, its gametes or its offspring[82,83]. Selection can also modulate recombination rates in response to a range of intrinsic factors, such as transposable element silencing in heterochromatin[84], or extrinsic factors, such as parasite load[85]. Indirect selection acts on the genomic level by modifying the degree of association between genes, such as by favouring reduced recombination between two or more advantageous alleles[86] or increased recombination to mitigate Hill–Robertson interference[17]. A similar form of indirect selection plays a role in recombination differences between sexes[49,87,88]. Selection is also responsible for the initial reduction in recombination in the autosomal ancestors of sex chromosomes to resolve sexual antagonism[35,36]. Theoretical expectations of variation in recombination due to ultimate causes are reviewed more extensively in REF.[10].

Proximate causes

The molecular correlates and proximate mechanisms of recombination rate variation have been reviewed extensively in REFS[38,89]. In the following discussion, we briefly consider proximate associations with genomic features and potential mechanical constraints mediating the variation described above. For example, lower recombination in heteromorphic sex chromosomes can be attributed to an absence of a homologous sequence to pair with during meiosis in the heterogametic sex[53]. Higher recombination at chromosome ends may be partially due to chromosome homology pairing being initiated at the telomeres, resulting in higher numbers of double-strand breaks (DSBs)[56]. Crossover interference has also been implicated for the higher levels of recombination at telomeres[90,91] and broadly for variation in the recombination rate across the chromosome[92,93], as well as between sexes owing to the influence of sex-specific compaction of chromosomes during meiosis[91]. Chromosomal

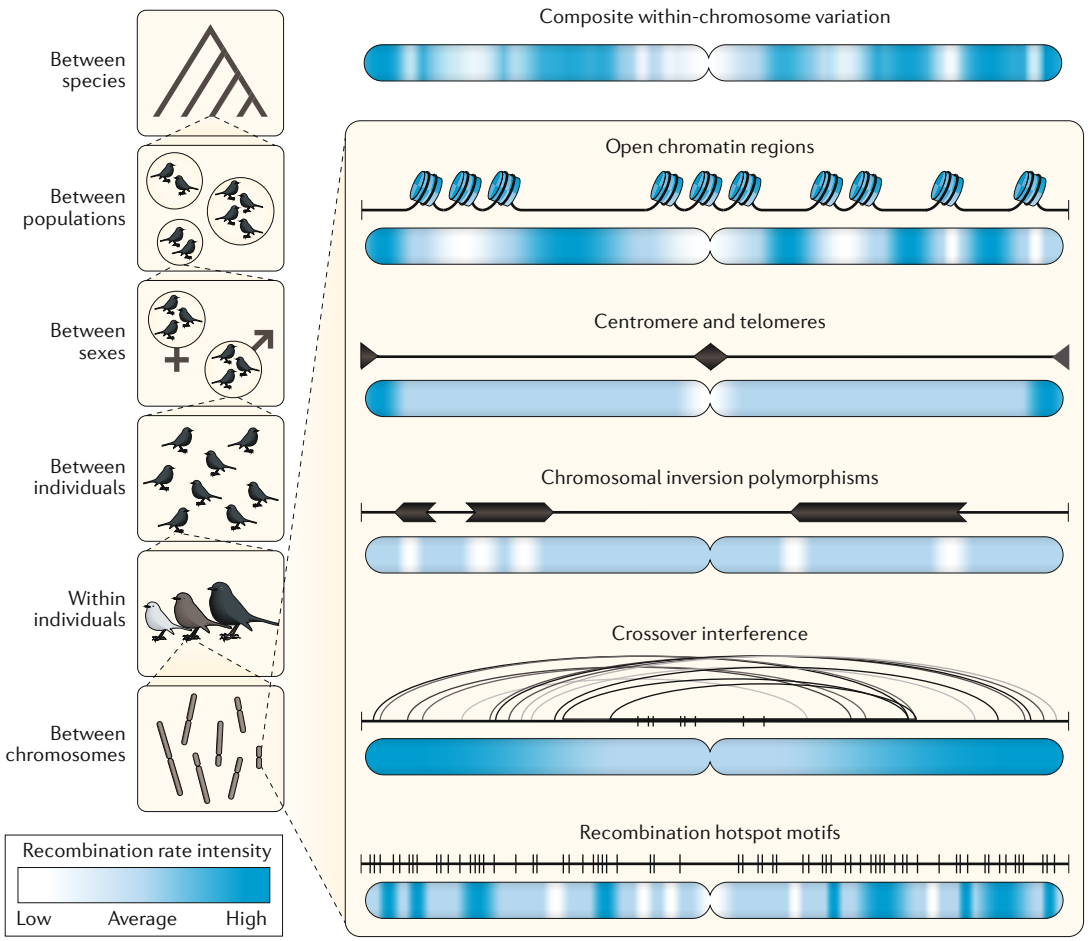


Fig. 1 | **Recombination variation.** The left part of the figure shows the different levels of biological organization within and between which recombination can vary. The right part shows some of the molecular mechanisms that affect within-chromosome variation. The chromosome below each mechanism depicts how that mechanism would affect the direction of recombination rate modification. The schematic on top depicts the overlaid recombination rate across the chromosome resulting from all processes.

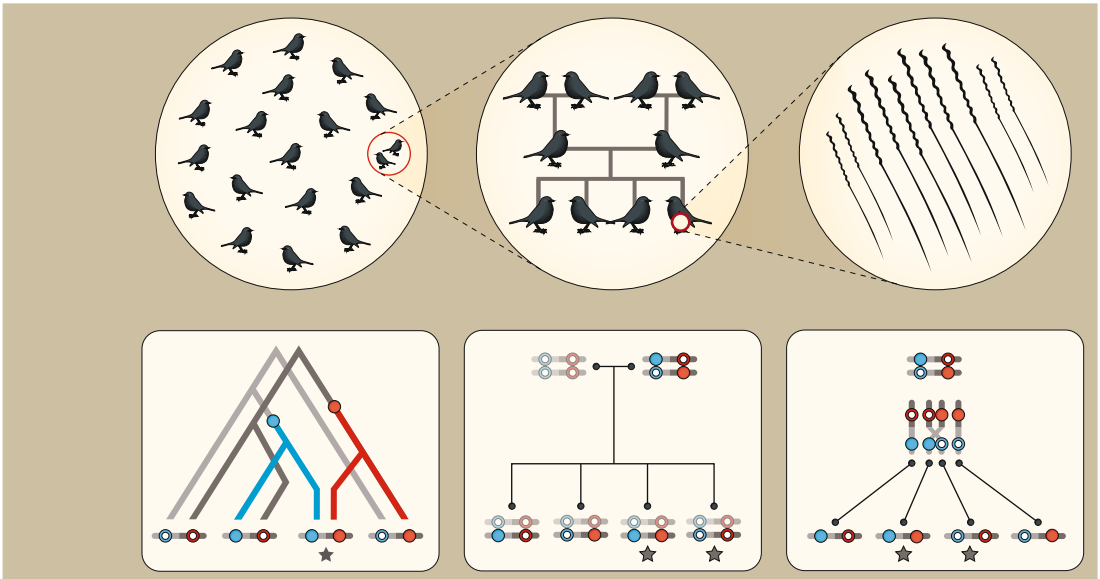
Gene conversion
The process whereby a fragment of DNA sequence is replaced by its homologue.

Crossover event
When a double-strand break (DSB) during meiosis results in exchange of homologous chromosomal regions. By contrast, a non-crossover event is when a DSB is repaired without homologous exchange of DNA material.

inversions have also been associated with reducing recombination locally within the inverted region in the short term via incorrect meiotic segregation in heterozygous individuals or in the long term by mutational loss of motifs associated with DSBs[94,95]. Conversely, this reduced recombination in the inverted region can result in an increase in local recombination in other regions of the same chromosome[96]. In addition, recombination rate variation associated with high GC content is a result of GC-biased gene conversion following DSB repair[25]. The common association of recombination with exon density, however, is less direct as both features covary with GC content[61]. Recombination also tends to be higher in regions of open chromatin, such as in promoter regions of genes, resulting in an association between recombination, transcription and epigenetic modifications[97–99]. The causality of such covariation remains to be determined.

Various proteins also substantially affect local recombination rate variation. Proteins such as HEI10 and RNF212 govern the recombination rate by determining whether a DSB will result in a crossover event introducing variation across species[100–102]. Primed by specific sequence motifs, the protein PRDM9 mediates the formation of recombination hotspots along the genomes of many mammals[103–105]. Certain DNA motifs, likely to be where some of these proteins bind, also tend to be associated with higher recombination rates[106–108]. The association of the recombination rate with proteins (such as PRDM9) or certain genomic features, such as telomeres or the transposable element content, tends to differ between species or groups[38,56,109], which provides opportunities for comparative genomics research to investigate the mechanistic differences in recombination control.

Table 1 | **Summary of the three genomic-based approaches to infer the recombination landscape**



Parameter	Population-based	Pedigree-based	Gamete-based
Estimate	Population recombination rate	Genetic distance	Crossover frequency
Sample size	10–30	>100	≥1
Pros	Low to moderate sampling Moderate to high resolution Simple study design	Sex-specific recombination maps Unbiased by population level processes Variation between individuals and heritability can be measured Can improve genome assemblies	High resolution Low sampling required Variation between individuals can be compared
Cons	Average between sexes and through time Biased by effective population size, demographic history, selection and mutation Dependent on genome assembly quality	Large sampling required Complex and time-intensive study design required Low resolution (~1 Mb, 0.5–2 cM) Biased by SNP density in mapping population Only contemporary snapshot of recombination	Single-sex recombination map (usually males) Biased by individual SNP density Dependent on genome assembly quality Only contemporary snapshot of recombination
Example refs	65,128,134,233	142,156–158	169,178,234,235

In the schematic figure, grey stars indicate recombination events. cM, centiMorgans; SNP, single-nucleotide polymorphism.

Recombination landscape
The variation in the local recombination rate plotted against the position along the chromosome.

Genetic mapping
A marker-based method to identify the order and genetic distance between loci.

Linkage disequilibrium
The non-random association of alleles at different loci.

Coalescent
A mathematical model describing the stochastic process of random reproduction backwards in time until all gene copies share a common ancestor. It predicts the distribution of gene genealogies of freely recombining segments of the genome.

Effective population size
(N_e). An abstract population genetic parameter describing the number of individuals in an idealized population in which the effect of genetic drift is representative of that in the real population.

Ancestral alleles
The allelic state of a locus that originated in the ancestral population. It is generally contrasted to the derived allele that arose by mutation in the evolutionary lineage or population under consideration.

Infinite sites model
A model in molecular evolution that assumes that there are an infinite number of sites where mutations can occur and that new mutations must occur in a novel site.

Four-gamete test
A test to detect historic recombination events by locating allelic combinations that could only have arisen as a result of recombination.

Ancestral recombination graph
A generalization of the coalescence tracing gene genealogies, integrating the recombination history of a population of samples.

Phased
The haplotype is inferred from genotype data.

To fully realize the extent of variation in recombination and to investigate its causes and consequences, we must first accurately characterize the recombination landscape across the genome. We advocate that the focus should not be restricted to a small subset of strains or a subset of genetic models or domesticated species but should be expanded across a large variety of taxa, populations within taxa and individuals within populations.

Methods for measuring recombination rates

Recombination rate quantification requires either direct observation of crossover events during meiosis in individual cells (cytological methods) or indirect inference of crossover events from haplotype shifts across generations using DNA sequence information (genomic methods; TABLE 1). Before high-throughput sequencing, the recombination rate was estimated using either cytological methods or genetic mapping with a few highly variable loci. These methods started to reveal the variation between organisms and within the genome but mostly at a coarse scale with only a few loci per chromosome. Here, we briefly present the cytological methods to then focus on sequence-based genomic methods.

Cytological methods

The cytological methods measure the recombination rate at the level of individual cells by bioimaging the meiotic process in different stages. One method relies on counting the recombination nodules during the pachytene stage[110,111], whereas another counts the chiasmata during diakinesis[112,113]. Although the resulting rates between these two methods are not identical, they tend to be strongly correlated[114–116]. Another method that has been used more recently is the immunolocalization of the mismatch repair protein MLH1 in the synaptonemal complex[117–119]. As these methods require visualization and counting, they provide a relatively coarse resolution of recombination rate variation across the genome. Despite this, the cytological methods can directly access information that is often hidden or indirectly inferred from the genomic methods. The direct observation of chromosome pairing provides insight into the mechanical controls of crossover frequency and interference[113], effects of inversion polymorphisms[120] or effects of chromosomal translocations[114]. As these methods allow different stages of the recombination process to be visualized, comparisons can be made between, for example, the formation of DSBs and subsequent crossover events; such analyses can reveal meiotic controls of crossover frequencies, which helps to unravel the mechanisms underlying the associations between the recombination rate and various genomic features[121,122]. In addition to methods investigating recombination in its native cellular context, single-molecule optical techniques are powerful tools to investigate the mechanistic underpinnings of variation in recombination rates (reviewed by REF.[7]).

Genomic methods

Population-based approach. Also referred to as either linkage disequilibrium or coalescent-based, the population-based approach leverages information from the association of alleles to estimate the population recombination rate (ρ). The rationale of this approach rests on the assumption that recombination is a main force in determining the statistical association between alleles. Consider a new neutral mutation arising on a given haplotype that is then increasing in frequency in the population. This may be due to stochastic drift or selection on a neighbouring site. In the absence of recombination, the new allele will always be associated with the allelic states of the surrounding genetic variants of the original haplotype. Linkage disequilibrium describes the magnitude of this non-random association of alleles in a population[123,124].

The population recombination rate can be derived from the statistical association described above as $\rho = 2cN_e r$, where c refers to the ploidy of the genome, N_e is the effective population size and r is the recombination rate in units of meiosis per generation. Population-based recombination rate estimation (TABLE 1) gathers information from haplotypes as follows. Consider the ancestral alleles of two biallelic loci *a* and *b*. If the mutations that result in the derived alleles *A* and *B* occurred only once and in different genealogical branches of a population, there are three possible haplotypes: *ab*, *Ab* and *aB*. Under the infinite sites model, not allowing for more than one mutation per site, the only process that leads to haplotype *AB* is recombination (four-gamete test). Thus, the prevalence of this fourth haplotype in the population is a function of the age of the mutations and the recombination rate between the loci. Therefore, a population of haplotypes has an underlying history of mutation and recombination events (stored in the ancestral recombination graph[125,126]). The framework of this inference is to either count these shifts or model the process that generated the underlying recombination graph and estimate the recombination rate along a genealogy of samples[127]. The resulting population recombination rate is then an average between both sexes integrated over hundreds to thousands of generations.

There are several estimators and software tools for population-based inference, which vary in accuracy and computation time (TABLE 2, and reviewed by REFS[128,129]). The input for this type of inference is simply an alignment of multiple genomes, preferably phased, which is often limited to ~30 diploid individuals (~60 chromosomes) because of computational costs. The most straightforward method counts the recombination events using the four-gamete test (that is, counting the *AB* haplotypes from the example above). Simulations show that this method will underestimate the recombination events and is most useful for providing a minimum recombination rate[130,131]. State-of-the-art approaches use coalescent theory based on the ancestral recombination graph to model the process that resulted in the underlying tree of the sequences and, subsequently, estimate the ρ value[125,126,132]. This can be conducted in various ways differing in algorithmic rationale and computational effort. The available software and underlying inference methods are described in TABLE 2. As the statistical association of alleles decays with genetic distance, estimates of ρ are computed for small subsections of a chromosome and are generally

Table 2 | **Notable software for recombination rate inference**

Software	Description	Link	Ref.
Population-based inference			
ABLE	Composite likelihood method to jointly estimate recombination and demographic history using blockwise site frequency spectra	https://github.com/champost/ABLE	177
CodABC	Approximate Bayesian computation to jointly estimate recombination and substitution rates in coding sequences	https://github.com/MiguelArenas/codabc	175
FastEPRR	Machine learning method (boosting) to select the best regression model using a compact folded mutation frequency spectrum as the summary statistic	https://www.picb.ac.cn/evolgen/softwares/FastEPRR.html	236
InferRho	Full-likelihood under Bayesian framework using a Markov Chain Monte Carlo method	https://github.com/sisiwying/InferRho	237
iSMC	Combines sequentially Markovian coalescent process with Markov model of recombination variation — Markov-modulated Markov model; uses a hidden Markov model to estimate likelihood function	https://github.com/gvbarroso/iSMC	238
LDhat	Pairwise composite likelihood under a Bayesian framework using a reversible jump Markov Chain Monte Carlo method; importance sampling to compute two-locus likelihoods	https://github.com/auton1/Ldhat	239
LDhelmet	Pairwise composite likelihood under a Bayesian framework using a reversible jump Markov Chain Monte Carlo method; system of recursive relations to compute two-locus likelihoods	https://sourceforge.net/projects/ldhelmet/	134
LDjump	Regression model fitted to many summary statistics followed by multiscale change point estimator for recombination rate shifts; considers demographic history	https://github.com/PhHermann/LDJump	176
LDpop	Pairwise composite likelihood using Moran model to compute two-locus likelihoods; considers demographic history	https://github.com/popgenmethods/ldpop	240
omegaMap	Approximate likelihood through "product of approximate conditionals" under a Bayesian framework using a reversible jump Markov Chain Monte Carlo method; co-estimates selection parameter	http://www.danielwilson.me.uk/omegaMap.html	241
TREE	Uses summary statistics reflecting topological data analysis and exponential regression	https://github.com/MelissaMcguirl/TREE	242
Pedigree-based inference			
CRI-MAP	Computes pairwise 'two-point' estimates of LOD scores between SNPs and uses maximum likelihood to build and test the order of markers on genetic map; reports RFs and Kosambi's mapping distances	https://www.animalgenome.org/tools/share/crimap/	156
HighMap	Computes pairwise 'two-point' estimates of LOD scores between SNPs and uses a *k*-nearest neighbour algorithm and Monte Carlo multipoint maximum likelihood to build and test the order of markers on a genetic map	http://highmap.biomarker.com.cn/[a]	243
JoinMap	Computes pairwise 'two-point' estimates of RFs between SNPs for mapping ordering and multipoint estimates using Monte Carlo multipoint maximum likelihood for recombination; reports RFs and Haldane's mapping distances	https://www.kyazma.nl/index.php/JoinMap/	244
Lep-MAP	Assumes achiasmy; computes pairwise RFs; the initial gene order is built in a greedy function, and a hidden Markov model is employed to optimize ordering and estimate recombination	https://sourceforge.net/projects/lep-map/	159
Lep-MAP2	Similar to Lep-MAP but can compute sex-specific recombination and has improved efficiency	https://sourceforge.net/projects/lepmap2/	245
Lep-MAP3	Similar to Lep-MAP and Lep-MAP2 but uses genotype likelihoods (can accommodate whole-genome sequencing), faster ordering and a modified haplotype model; can report genetic distances from different mapping functions	https://sourceforge.net/projects/lep-map3/	158
MSTmap	Computes pairwise probability of recombination (similar to RFs); uses minimum spanning tree for ordering markers	http://mstmap.org/	157
Gamete-based inference			
DrLink	Estimates the relative crossover frequency based on the number of recombinant molecules per million molecules; recombinant molecules are identified using phased genomes provided by the user	https://github.com/schneebergerlab/DrLink	169
MRLR	Reconstructs the gamete haplotypes and detects crossover events using family trio data sets	https://github.com/ChongLab/MRLR	246
ReMIX	Provides putative recombinant linked-read molecules, phase-shift locations and likelihood score; recombinant molecules are identified by phasing of the input bulk sperm data	https://github.com/adreau/ReMIX	178

Software that only compute a genome-wide recombination rate and those that only focus on hotspots of recombination were omitted from this table. See REF.[68] for a review of detecting hotspots of recombination. The provided URLs were accessed on 6 March 2020. LOD, logarithm of the odds; RF, recombination fraction; SNP, single-nucleotide polymorphism. [a]Website unavailable as of 4 May 2020.

presented as sliding windows (defined by the total number of bases or the number of segregating sites)[133]. A normalized genome-wide median ρ value over a predefined window metric — such as ρ per kilobase or ρ per chosen number of single-nucleotide polymorphisms (SNPs) — can then be used to broadly compare populations[109]. We recommend use of the median over the mean for these broad comparisons as it is less distorted by variation in the presence and detection of recombination hotspots.

It is important to note that all methods of inference based on linkage disequilibrium as described above rely on a set of assumptions relating to constant population size[124,134], mutation rate[135,136], selective neutrality[137], random mating including the absence of population structure and migration[138], and genetic drift[139]. Deviations from model assumptions may distort linkage disequilibrium[140] and bias the estimates of the population recombination rate[141].

Pedigree-based approach. Whereas population-based inference integrates recombination acting across evolutionary times, the pedigree-based approach observes recombination across a few generations. The resulting recombination estimate, therefore, is not affected by long-term population size changes and is affected to a much lesser extent by selection. Pedigree-based inference provides a direct estimate of a genetic linkage map by observing the inheritance of alleles between parents and offspring and estimating the recombination fraction among the loci[142,143]. Although this analysis can be performed with two generations, a third (grandparents) is always preferred to phase the parental generation, resulting in more accurate characterization of crossovers in the offspring. Although deeper pedigrees contain more meiotic events, modern software tend to primarily consider nuclear families. Large, complex pedigrees with many generations, close relatives and extensive extra-pair mating can still be accommodated by the available software, but requires breaking large pedigrees down to three-generation nuclear families[45,144,145]. The pedigree used is referred to as the 'mapping population' and is either established using experimental crosses (particularly pertinent in domestic breeds[146,147]) or wild populations tracked through multiple generations. Pedigree-based inference can distinguish recombination rates between the sexes and even among individuals of the parental generation[45,71,148].

The required input data for this inference are genotypes, or genotype likelihoods and the underlying pedigree. Traditionally, the genetic markers of choice in mapping studies were sequence tags[149,150], DNA fragment length polymorphisms (restriction fragment length polymorphisms (RFLPs) or amplified fragment length polymorphisms (AFLPs))[151] or microsatellites[148,152]. The onset of high-throughput sequencing shifted inference to genome-wide SNPs, although the general theoretical framework stayed the same. No longer limited by marker density, the resolution of recombination fraction still primarily scales with the number of meioses. It is thus advised to use large pedigrees in the order of hundreds to thousands of individuals and dozens of families, although coarser resolution of recombination rates gleaned from smaller pedigrees can still provide useful insights into recombination rate variation[45,144]. Although larger sample sizes will increase precision, correct inference of the relationships in the pedigree is crucial for accurate inference of genetic distances. Various methods can be used to test the reliability of the pedigree using the genotype data or to assign paternity if unknown[153,154].

If a high-quality reference genome is not already available, the first step involves identifying linkage groups that are often assumed to correspond to different chromosomes[155]. Within each linkage group or chromosome, the markers are then ordered along the genetic map based on pairwise co-inheritance of alleles measured as logarithm of the odds scores (LOD scores), consequently minimizing the recombination fraction (r) between adjacent markers of a linkage group. The ordering step is often the most time-consuming and laborious, as it often requires manual curation to try different reference markers to build the map, trim off incorrectly mapped loci and run multiple iterations of building the map and determining the order. Many software packages, such as CRI-MAP and MSTmap, that were used to accommodate a few loci have also been shown to handle genome-wide SNP data[156,157]. Methods such as Lep-MAP3 were specifically designed for genomic data, such that they can handle larger amounts of data efficiently and incorporate uncertainty by using genotype likelihoods[158,159]. Alternatively, software can be simply used to phase the offspring haplotypes, and the crossover frequencies across the genome are subsequently counted[160,161].

The resulting genetic map is then used for the recombination inference. It is at the discretion of the researcher whether they prefer the physical map or the genetic map, as these maps may not necessarily be concordant and it is often difficult to determine the ground truth[162]. Alternatively, many studies filter genetic markers to only use those that are concordant between both maps. The measure of recombination between pairs of SNPs in the genetic map is in units of recombination fraction ranging from 0 (completely linked) to 0.5 (unlinked). Mapping functions are used to convert these fractions into additive measures of genetic distance: centiMorgans[163]. There are three commonly used mapping functions in recent pedigree-based analyses: Morgan's, Haldane's and Kosambi's mapping functions (FIG. 2). The main difference between the three is how each incorporates crossover interference. This difference tends to only affect adjacent pairs of loci with a relatively high recombination fraction. With genome-wide SNP sampling and mapping, the difference between the three becomes negligible except in regions of extraordinarily high recombination, such as hotspots or chromosome ends.

The map of each linkage group, ideally representing a chromosome, starts from zero and progressively increases with each variant. The final task is to convert this map into a recombination landscape. One of the main challenges is the variation in SNP density and the reliability of inferring the recombination rate at a

Genotype likelihoods
The probabilities of genotypes, accounting for potential errors in the sequencing data that occur during sequencing and processing.

Linkage groups
Genetic markers that are inherited together as a unit, usually representing a chromosome.

Genetic map
Also known as linkage map. A representation of the order of genetic markers and inter-marker distance derived from the frequency of meiotic recombination.

Logarithm of the odds scores
(LOD scores). A statistical estimate of the likelihood that two entities are co-inherited, referring to the association of phenotypic and allelic variation or to the association between genotypes.

Physical map
The physical order of genetic markers along a chromosome.

Mapping functions
Algorithms to infer the additive genetic distance between two loci from measurable recombination fractions between them.

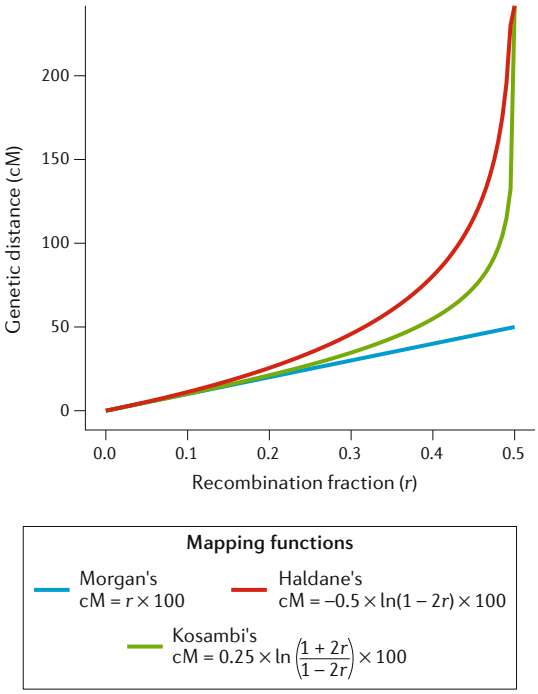


Fig. 2 | **Mapping functions.** Different mapping functions result in different conversion from recombination fraction to centiMorgans (cM) depending on how crossover interference is taken into account. Denser single-nucleotide polymorphism (SNP) sampling across the genome, resulting in recombination fractions <0.2 between adjacent SNPs, would converge on near-identical total genetic map distance.

given point. This is mitigated first by plotting the genetic distance along the physical distance (Marey Map)[164,165]. The local recombination rate can then be read out at any location along the chromosome as the slope of the Marey Map at that point, after it has been smoothed using a local weighted regression[166,167]. The final result is a sliding-window centiMorgans-per-megabase measure of recombination fraction along the genome.

Gamete-based approach. The gamete-based approach of recombination inference measures the frequency of crossovers between a diploid individual through its haploid gametes (FIG. 3). Although usually referred to as sperm typing[168] or sperm-based inference, because it is more often performed on male gametes, we generalize the term to gamete-based inference. This allows for accommodation of plant systems using pollen data[169] and the use of female gametes, such as ova or ovules[111]. These can be extracted with relative ease from systems that exhibit external fertilization, such as corals, fish or dioecious plants. More generally, the approach pertains to any system (including spores in fungi) where gametes can be unequivocally related to the source individual. It assumes that two loci for which the parental haplotype is known (or can be reliably inferred) can be scored (or 'typed') in a multitude of independent gametes. In the absence of any bias in inferring the observed diplotype, the recombination fraction (0–0.5) is then expected to equal the number of inferred crossover events divided by the total number of sampled gametes. Early sperm typing analysis achieved extremely high resolution, whereby hundreds to thousands of sperm could be typed for a pair or a handful of loci[170,171]. When scaling up to genome-wide recombination landscape inference, budget limitations will reallocate the sequencing effort to other parts of the genome and consequently reduce the resolution at any given point.

Whereas most of the inference on recombination landscapes to date is either population-based or pedigree-based, gamete-based inference is gaining traction with further advancement in high-throughput sequencing and library preparation methods. Gamete-based recombination inference can be achieved in two ways: single-gamete or bulk-gamete sequencing. The specific differences between the two methods are discussed in detail in BOX 2. In the following discussion, we outline the general framework of the inference pipeline. The only required input data are sequenced gametes from a single source individual. Sequence data from the somatic tissue of the source individual (referred to as the 'donor') can be useful as a control or to phase the genome but is not strictly required. In principle, the donor's phase can be estimated by the consensus of allelic combinations of the gametes. The vast majority of gametes (excepting those experiencing crossover events) should agree on either of the two haplotype blocks present at equal frequency in the donor. Incorrectly phased regions will artificially inflate recombination rates in downstream inference. Although phasing would ideally be done for the entire length of the chromosome, it can be restricted to haplotype blocks within the chromosome. Regions that are difficult to phase due to low SNP density will consequently lack the power to detect recombination events.

After inferring the phase of the donor, the next step is to detect crossovers in the gamete data. Crossover events are characterized by a phase shift from one haplotype to the other. In principle, a minimum of two loci are informative about a haplotype shift. In practice, however, a run of several alleles on both sides of the presumed crossover event that can be unequivocally attributed to the donor's haplotypes increases the confidence. Gene conversion can also result in very localized shifts in phase but will often shift back to the other haplotype. The closer the shifts, the less likely they are true double recombinants impaired by crossover interference. The main challenge is to properly identify true crossover events from gene conversion and genotyping errors.

As with the pedigree-based inference, what is being inferred in gamete-based inference is the recombination fraction. This fraction then needs to be converted into the additive genetic distance in centiMorgans using one of the mapping functions and taking the specific technical biases into account. However, the biases of the mapping functions are different to pedigree-based inference[172]. Lastly, inferred recombination fractions can be smoothed along the chromosome using LOESS regression or similar approaches to remove biases due to variation in SNP density. The resulting recombination rate is specific to the sex of the donor and may not reflect the net recombination rate of the species, particularly if there is strong heterochiasmy.

Haplotype blocks
Discrete stretches along the chromosome for which the phase can be unequivocally determined.

Demographic history
The history of a population with regard to change in size, structure and gene flow.

Comparison between approaches

The choice between the three approaches to quantify recombination variation depends on the available resources, the budget and, most importantly, the research questions (see Utility of different approaches below). General advantages and disadvantages together with typical sample sizes are summarized in TABLE 1.

The design behind population-based recombination inference merely requires a random sample of unrelated individuals from the populations of interest. The samples are typically already available in respective laboratories, are available from museum collections or can be procured. Systems or species where obtaining more samples is difficult are going to pose a similar challenge for the other two alternative approaches. The main concerns for the population-based approach are the factors that confound linkage disequilibrium as discussed previously[173,174]. For certain questions, joint estimators of recombination and other parameters, such as demographic history, may be employed to alleviate the biases[175–177]. It can often be more appropriate to choose other recombination inference methods that are less susceptible to these biases until more algorithms to jointly estimate these parameters are available[26,27].

Pedigree-based approaches circumvent biases introduced by the long-term effects of selection and demography, and the recombination rates of both sexes and individuals can be estimated separately. Biases may arise because the inferred genetic map is strongly dependent on the variation present in the founder individuals of the pedigree. If these individuals deviate from the population norm for any reason or there is high heterogeneity in recombination rates across the population, then a single pedigree may not fully capture the true recombination landscape of the focal species. Moreover, similar to the population-based approach, the pedigree-based approach will generally only consider recombination events in gametes that led to offspring (that is, it does not characterize recombination in unsuccessful gametes). The largest limiting factor, however, is the prerequisite of large pedigrees. Even large pedigrees still result in coarse resolution owing to the limited numbers of meioses. This design is most feasible for species that can be reared in captivity and have relatively short generation times, or in situations where a multigeneration sampling of a wild pedigree has already been established. The scale of the sampling ranges from hundreds to thousands of individuals for dozens of families, which requires a considerable amount of time to establish from scratch.

Gamete-based inference presents a promising, relatively inexpensive, alternative as it is largely confined to sequencing costs and does not depend upon extensive breeding resources. The initial bottleneck is the ease of extracting the gametes from the individual. Naturally, this is one of the reasons why this analysis tends to be biased towards male-specific recombination inference. Even for sperm, the difficulty of extraction varies considerably between species. The next crucial consideration refers to the choice of the donor individual for inference. Similar to the pedigree-based approach, heritable variation in the recombination rate between individuals can directly be scored but depends on the examined individuals. Inbred individuals exhibit runs of homozygosity, which preclude the detection of recombination events occurring in these regions; collectively, these local downward biases result in a genome-wide underestimate of recombination. Furthermore, care must be taken to ensure low false-positive rates. Currently, neither of the bulk-sequencing methods have converted the recombination landscape estimates into standard centiMorgans per megabase units but rather used the crossover frequency as a measure. Although the recombination landscapes tend to mirror that of pedigree-based inference[169,178], further investigation is necessary to fully realize the potential of these new methods.

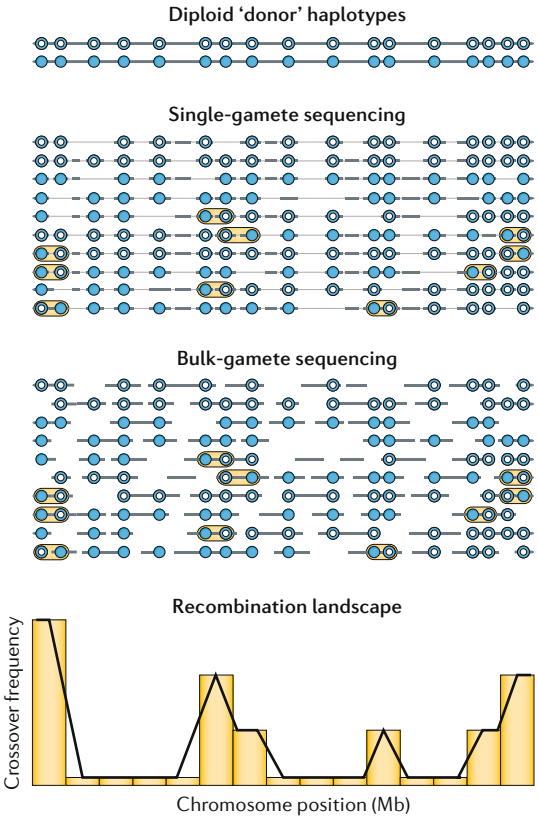


Fig. 3 | **Single-gamete versus bulk-gamete sequencing.** A schematic of the expected data when performing either single-gamete or bulk-gamete sequencing. Open and filled circles represent alleles of segregating genetic variation from each of the two haplotypes of the (diploid) donor individual. Each row in the single-gamete sequencing represents genetic data from one library constructed from one cell and the thick grey bars represent sequenced reads within each gamete cell. The single and double crossover events, marked in yellow, can be identified within each library. In the bulk sequencing, each thick grey bar represents a single DNA molecule from a large pool of gametes. The cell of origin and, therefore, associations among molecules are unknown. Crossover events can only be inferred within each molecule. For both methods, counts of crossover events yield information on the recombination landscape.

Sources of discrepancy between approaches

Although the different approaches broadly correlate in their recombination estimates, they differ in important aspects[72,134,179]. The sources of discrepancies can either

Box 2 | **Gamete-based recombination inference**

Single-gamete sequencing

Single-gamete sequencing refers to the isolation and amplification of individual gamete genomes before sequencing. To date, this method has been successfully conducted in human, mouse, *Daphnia pulex* and *Arabidopsis thaliana*[217,221,234,235]. Once the first hurdle of collecting gametes from an individual is overcome, cells are sorted using, for example, flow cytometry, which distributes individual cells into different wells[247]. Sorting is followed by single-cell, whole-genome amplification (reviewed in REFS[248,249]). The amplified genome can then either be sequenced as low-coverage whole-genome sequencing or with reduced representation using double digest restriction-site associated DNA (ddRAD) sequencing or similar methods. The studies to date have opted for whole-genome sequencing of ~100–200 individually barcoded sperm libraries sequenced on an Illumina platform.

The recombination inference proceeds generally as described in the main text. A high-quality reference genome is typically required but single-cell data can also be used to build a genetic map and upgrade the genome assembly (for proof of concept, see REF.[234]). The first step of phasing the donor genome can theoretically be done directly from the sperm data dependent on the degree of shared sequence variants among libraries. If the segregation follows as shown in FIG. 3, half of the sperm should contain the correct phase and most adjacent variants should be in phase. For example, a large genetic distance of 1 centiMorgan between two adjacent single-nucleotide polymorphism (SNPs) translates to only 1% of the sperm not being in phase. Once phasing of the donor genome is accomplished, algorithms to search for phase shifts in the sperm are implemented to detect crossover events. Note, however, that local phase shifts can be indicative of gene conversion rather than true crossover events. The underlying recombination landscape is represented by the frequency of crossover events in a given genomic window provided each window is covered by an equal number of gametes (FIG. 3). Currently, there are no publicly available software to perform this inference.

Bulk-gamete sequencing

Recent developments in sequencing technology and analytical software provide an arguably more efficient way of inferring the recombination landscape through the sequencing of a pool of gametes. Long-read sequencing, such as SMRT (Pacific Biosciences) or Nanopore (Oxford Nanopore), or linked-read sequencing (such as that provided by 10x Genomics or TELL-seq by Universal Sequencing) allows direct readout of long genomic segments from individual gametes. This approach has been performed successfully in humans, stickleback, mouse and *A. thaliana* using linked-read sequencing[169,178,246]. Reads need to be long enough to capture a minimum of two alleles that are informative about the donor's haplotype. The required length depends on the organism's level of polymorphism and the heterozygosity of the chosen donor. Inbred individuals, for example, will be uninformative for recombination in regions that are identical by descent (runs of homozygosity[250]). These uninformative regions can often be larger than the library insert sizes, leading to inability to detect recombination in large regions of the genome.

Regarding the practical steps, before library preparation, the cells can be sorted to remove potential diploid cells to ensure only gamete cells are sequenced[178]. This step is particularly important if the gametes are to be extracted from heterogeneous gonad tissue containing both haploid and diploid cells. It is then pertinent to extract only high molecular weight DNA, as the longer DNA molecules are more informative. The balance is to efficiently lyse the cells and unravel the tightly-packed DNA without significantly fragmenting it[251,252]. Although some DNA is required for quality control of the extraction, the actual input DNA for library preparation for linked-read sequencing is minimal and ranged from 0.8 to 1.2 ng in the examined studies, depending on the size of the genome. If opting for long-read sequencing technology, larger quantities of DNA would be required. The initial studies showed that increasing the input DNA per library tends to increase the false-positive rate when using linked-read technology. It is thus recommended to have multiple linked-read libraries of lower DNA input rather than a single library[169,178].

Three bioinformatic pipelines have been recently developed to process the linked reads and identify molecules with crossovers: ReMIX[169], DrLink[178] and MRLR[246]. The reported studies have used different ways of representing the recombination landscape: REF.[178] used the number of recombinant molecules in a given sliding window, whereas REF.[169] used the ratio of crossover molecules in a given window relative to the entire chromosome. The choice of the reference genome is particularly important as structural variation between the donor individual and reference genome can lead to incorrectly inferred recombinant molecules.

Whole-genome amplification
Genome-wide amplification of DNA, usually performed from DNA extracted from only one or a few cells.

Long-read sequencing
A class of DNA sequencing technologies and platforms that currently allows for sequencing of long (>20 kb) stretches of DNA.

Linked-read sequencing
A DNA library preparation method that incorporates unique barcodes to reads derived from a longer DNA molecule, such that the reads can be bioinformatically reassociated to the original DNA fragment.

be due to biological processes or technical differences. Biological discrepancies reflect true variation between recombination rates at different levels of organization, and their comparison can contain valuable information. Technical sources of discrepancy result from variation in resolution, false-positive rates or false-negative rates and need to be carefully considered before the interpretation of results.

Biological sources

The biological sources of discrepancy are due to the differences in timescales each inference is measuring and the selection on effective recombination rates at different biologically relevant levels. The population-based approach infers recombination in a much deeper timescale, averaged across hundreds to thousands of generations in the ancestors of the present population. Conversely, the pedigree-based and gamete-based approaches offer a snapshot of contemporary recombination, in a subset of individuals, in a given environment. Depending on selection during gametogenesis or acting on gametes, on individuals, on different sexes and on the population through time, the approaches will differ in their recombination estimate. Moreover, the impact of environmental contributions to recombination rate variation will differ between approaches. The population-based approach is expected to integrate the influence of environmental factors across much longer time frames (hundreds to thousands of generations) than the pedigree-based approach (a few generations) or the

gamete-based approach (gametogenesis). This means that both the pedigree-based and gamete-based approaches are amenable for experiments testing environmental effects on recombination[77] but may be more problematic for generalizing the recombination rates to the rest of the population if the mapping population or donor individual has experienced unusual environmental conditions.

Technical sources

First and foremost, the number of sampled meioses determines the resolution at which we can estimate the recombination distance between two genetic variants. Using the simple Haldane mapping function, 1 recombinant in 100 only provides resolution to 1 centiMorgan in pedigree-based and gamete-based recombination inference. The population-based inference can reach finer resolution as recombination is inferred across many past generations. The density of marker sampling also contributes to the resolution that can be achieved. Sparse sampling on chromosome ends, where recombination rates are generally higher, would reduce the overall genetic map length and genome-wide recombination rate. This is one reason why recombination rates inferred from pedigrees and from using fewer SNPs may result in shorter genetic maps[109,180]. A final source of technical variation between studies is the spatial genomic representation. The choice of window size, window overlap and smoothing parameters will all influence the resulting landscape. Larger windows (megabase scale) may be more amenable for comparison but lead to a lower-resolution map than for smaller windows (kilobase scale). In megabase-scale comparisons between population-level and pedigree-level recombination rates, for instance, we see a 0.98 and 0.90 correlation (r^2) in humans[65] and zebra finches[109], respectively. Too-narrow windows, on the other hand, are affected by variation in SNP density, in genotyping quality and in noise introduced by stochasticity. In *Drosophila pseudoobscura* and *Drosophila miranda*, for example, the Spearman's correlation was 0.6 for windows of 100–200 kb (REF.[180]). It is important that studies report the direct estimates that each inference approach provides (either ρ, genetic distances or crossover frequencies), as well as sample sizes, marker density and window sizes used for representation of the recombination landscape.

Utility of different approaches

Not all genomic approaches are equally suited to address the question at hand. In the following discussion, we describe several aspects where these differences may matter.

Population-based recombination inference is the only measure that allows for direct population genetic inference, for example, comparisons of effective population size between human populations[181]. More generally, it is suited to furthering our understanding of the evolutionary processes acting in natural populations. Consequently, population-based recombination inference has been used to assess the effect of linked selection on population genetic diversity[39,133,182], to identify genomic features associated with recombination variation[183,184] and to identify genes contributing to adaptation and speciation[185–187]. As mentioned above, the population-based approach is highly sensitive to processes that can distort linkage disequilibrium, which may bias recombination estimates[187,188] and warrants caution regarding comparison among populations at face value. Being less susceptible to these issues, pedigree-based inference may therefore better serve the purpose of establishing a baseline recombination rate across the genome for analyses exploring the effect of long-term selection on natural populations[189,190]. Yet this comes at the cost of reduced, often insufficient, resolution for inference around single genes and entails the constraint that pedigree-based recombination snapshots may not be reflective of the variation that is relevant at evolutionary timescales. Ideally, population-based and pedigree (or gamete)-based approaches would be combined to gather how contemporary recombination rates translate into long-term rates[72,180] and to confirm that the population-based inference is unbiased[26,27].

Biases that are inherent in population-based inference similarly apply to the characterization of recombination suppressors, such as chromosomal inversions, for which pedigree-based or gamete-based estimates are, in principle, also suited[191]. They may, however, go undetected if the donor individuals are monomorphic for an inversion type. The resolution of the pedigree-based approach may also be insufficient to infer double-recombinants and hence to assess the role of inversions in sheltering against (heterotypic) recombination[192]. Resolution limits of the pedigree-based approach similarly apply to the characterization of the forces underlying genomic hotspots and coldspots of recombination, for which a better resolution will generally be achieved using population-based[109,193] or gamete-based[194] inference (but see REF.[72] for fine-scale inference using parent–offspring pairs).

With increasing sequencing power, population-based approaches can also be used for association mapping to link trait values to the underlying naturally segregating genetic variation[73,195,196]. They have thus complemented the more typical pedigree-based approaches for fine-mapping phenotype–genotype relationships. Despite their relatively coarse resolution, pedigree-based recombination rates are well suited for unravelling associations of recombination rate variation with various genomic features[197,198] or nucleotide diversity[199], and even for detecting recombination hotspots and coldspots[200,201]. Unlike population-based rates, pedigree-based recombination rates are informative on the extent of heterochiasmy[144,202,203], variation and heritability of individual recombination rates[45,204] and crossover interference[205]. Complemented by phenotypic data, the pedigree-based approach is still the method of choice in agriculture for marker-assisted selection[206,207], marker–trait associations[208,209] and quantitative trait locus (QTL) mapping[210,211].

The full extent of the utility of gamete-based recombination rates is yet to be realized as the methods are still developing. Gamete data enable the inference of recombination hotspots[212,213], with the added benefit of being able to characterize variation between individuals[214,215] and different conditions within the same individual[216]. Gamete-based inference is also suitable to support association between recombination and genomic features, as

Linked selection
A locus under selection causes corresponding changes in allele frequencies of other nearby loci owing to a lack of recombination between these loci.

Association mapping
A statistical approach to infer the association between phenotypes and genotypes.

Nucleotide diversity
A measure of genetic variation within a population or species, reflecting the average number of nucleotide differences between two chromosomes in a population.

Marker-assisted selection
Artificial selection whereby the founders for the next generation are selected on the basis of the genotype of a particular locus or set of loci underlying the phenotypic trait of interest.

Quantitative trait locus (QTL) mapping
A specific type of association mapping that focuses on a quantitative trait that is assumed to be encoded by multiple genes.

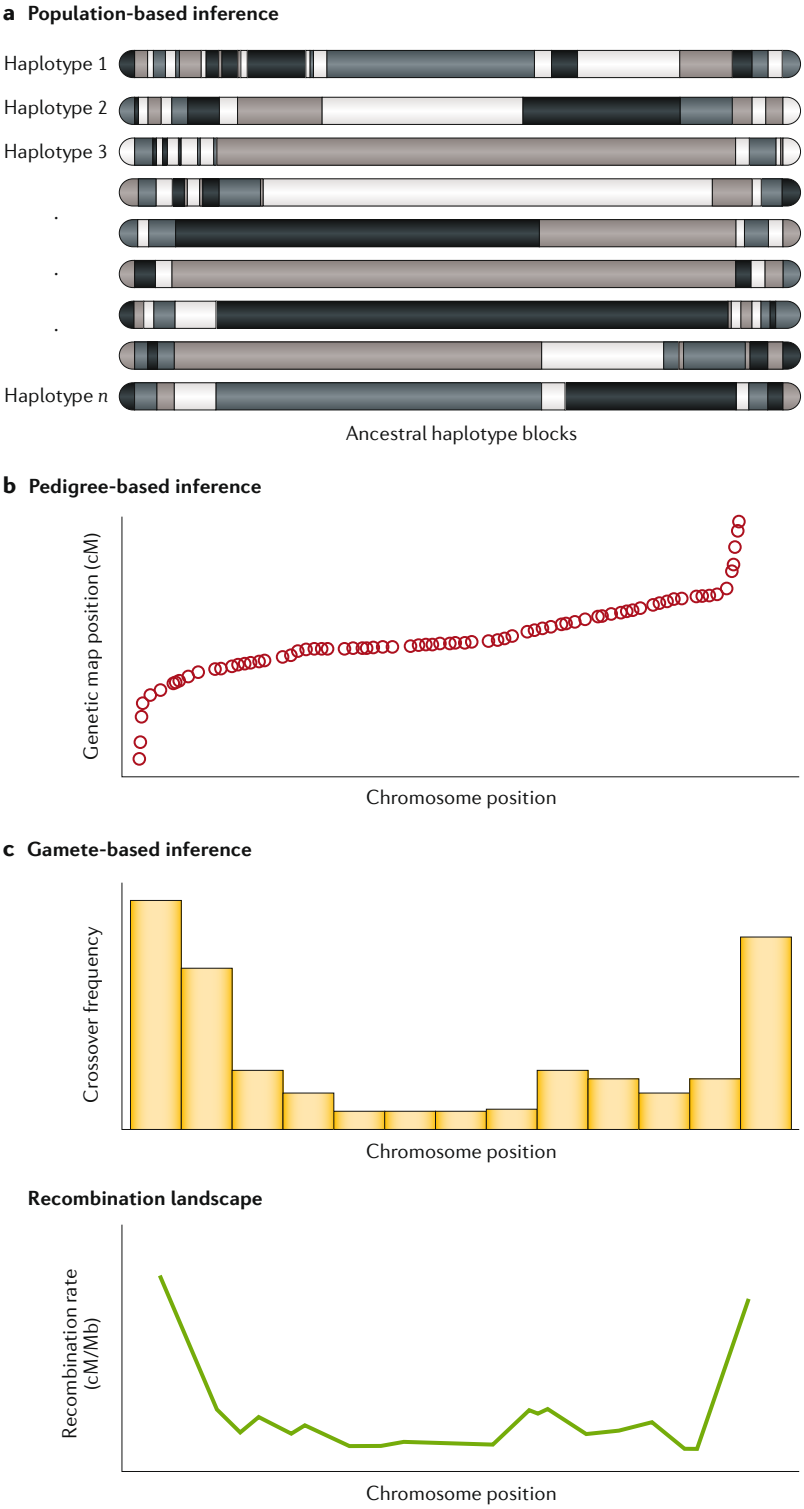


Fig. 4 | **From inference to landscape.** The actual result of each inference method and how it translates to the recombination landscape. **a** | Population-based inference involves direct analysis of haplotype structure along chromosomes. Contemporary haplotypes are composed of ancestral haplotypes (various shades) that arose at different points in the past. The identity and length of ancestral haplotype blocks are a function of the time at which the haplotype arose and recombination. **b** | Pedigree-based inference involves a comparative representation of genetic distance and physical distance where the local recombination rate is the slope at any given location. **c** | Gamete-based inference takes the crossover frequency of a given window and translates it into the recombination landscape. cM, centiMorgans.

illustrated by a study in mice associating recombination with PRDM9 binding and chromatin environment[217]. An important contribution of gamete-based methods is likely to be realized in situations where recombinant gamete types are less likely to develop properly, fertilize or develop into a normal zygote. This allows for studies on the impacts of chromosomal abnormalities during meiosis that can be hidden by selection on gametes or the resulting zygote[178,218], including meiotic drive[16,219,220], crossover interference[221], haploid selection[49] or fecundity[218,222]. However, in most cases, gamete-based inference is limited to males, for which the gametes are easier to access. Ignoring recombination in females may pose a problem for inference extended across the population when the study species exhibits some level of heterochiasmy. With the increasing ease of generating gamete-based recombination maps, we foresee more studies utilizing this type of data in various research studies in evolution, ecology and breeding.

Towards a unifying framework

Given the vast array of biological properties associated with recombination rate variation, foreseeing future directions for each of the research topics exceeds the scope of this Review. Instead, we wish to highlight the strength of combining the various approaches, encourage researchers to exploit comparative analyses of natural variation in recombination across scales and propose a unifying framework.

First and foremost, the integration of the different approaches offers high potential for a comprehensive understanding of the proximate and ultimate causes underlying recombination rate variation. Incorporating cytological techniques and the various methods exploiting genetic variation will provide information on how recombination interacts with selection: from the transmission of genetic variation from one generation to the next, all the way to evolutionary timescales of hundreds of thousands of generations. Cytological approaches are powerful for unravelling the molecular basis of recombination rate variation, whereas the genomic methods collectively provide insight into the degree of selection acting at the haploid gamete stage, between individuals, between sexes or across populations.

Additional information is to be gained from a comparative perspective integrating a wider range of study organisms. For example, studies comparing haplotype block structure across individuals, populations and species hold promise to jointly unravel the rate of recombination rate evolution and processes shaping genetic diversity, such as selection or demographic history[136,223]. In turn, a better grasp of the underlying mechanisms behind haplotype block structure will advance our ability to perform trait associations from SNP-based to haplotype-based[224,225], reconstruct complex evolutionary histories[136], infer recent demographic events[226] and disentangle the effects of selection on local ancestry of parental genomes in hybrid individuals[227]. Integration of many species may also help determine the generality for which PRDM9 or other recombination modifiers may be associated with recombination hotspots. The absence of PRDM9 in various taxa also exhibiting

Meiotic drive
A mechanism acting during meiosis or gametogenesis that distorts the equal transmission of alleles.

hotspots certainly requires alternative mechanisms (REF.[109] and references therein).

To fully realize the comparative potential between methods and biological levels of organization, we need to arrive at a single currency to which all recombination estimates across methods can be converted. All three sequence-based approaches obtain measures that are proportional to the recombination rate but do not measure the local recombination rate directly (FIG. 4). A common currency should, in principle, allow comparison of recombination estimates that are expected to differ between the approaches due to the difference in their assumptions, limitations, susceptibility to biases and the biological processes they capture. The recombination rate converted into a centiMorgans per megabase sequence arguably constitutes the most useful common measure.

A centiMorgans per megabase measure of recombination results naturally from the pedigree-based approach, and the slope of the Marey Map provides a direct estimate for any given window. Similarly, for the gamete-based approach, centiMorgans per megabase-based estimates should be straightforward to derive, although previous studies have not done so. We can only speculate that the conversion may be inflated by the false-positive rate owing to technical biases. In theory, the recombination rate r represents the frequency of gametes with crossovers for any given window. For a window size of 1 Mb, the recombination in centiMorgans would simply equal the number of molecules with crossovers per total number of genotyped molecules multiplied by 100. This relationship only holds in cases where molecules constitute independent draws from a large pool of gametes. It should be possible, however, to incorporate violations of this assumption into stochastic sampling models.

Although conversion of the pedigree-based and gamete-based approaches is, in principle, straightforward, conversion of the population-based measure ρ into r is conceptually more demanding. In theory, ρ can be translated into r (following $r=\rho/2cN_e$) and thereafter converted into centiMorgan units using one of the mapping functions. Estimation of the effective population size (N_e), however, constitutes a major challenge. An estimate can be obtained from the nucleotide diversity (θ_π) and the per-generation, per-base mutation rate (μ) following the relationship $N_e=\theta/4\mu$ under a restrictive set of assumptions, including neutrality and stability of the mutation rate across populations[228]. This assumption may not hold when comparing diverged populations or species where the mutation rate is more likely to differ. In addition, estimates of N_e may differ depending on the type of summary statistic considered (for example, inbreeding N_e or coalescent N_e), which is unlikely to capture all of the underlying evolutionary processes appropriately[229]. Moreover, θ_π is often estimated directly from the same population data used for the ρ estimate, which may pose a problem of circularity. Obtaining an estimate of the local mutation rate is likewise difficult to obtain. Whereas long-term averages can be inferred from putatively neutral substitutions[133,230], estimation of local mutation rates of remnant individuals itself relies on population-based, pedigree-based or gamete-based inference[219,231,232]. Research efforts in how to best convert the measures across these different approaches into a common parameter will be crucial for comparative studies utilizing these different estimates within and across studies.

Concluding remarks

Recombination is a central parameter in animal and plant breeding programmes, for human disease mapping and in evolutionary research. Yet fully characterizing its variation has proved difficult, and for a long time recombination inference has been limited to cytological or pedigree-based estimates in genetic model organisms. Parallel advancements in high-throughput sequencing, library preparation methods and analytical software have extended this capability to measure recombination at unprecedented resolution in essentially any organism of choice. Here, we reviewed several approaches and outlined methods to quantify recombination rate variation along genomes and among individuals, populations and species. With increasing ease of recombination inference comes increasing opportunities. Although we encourage empiricists and theoreticians alike to push the methodological boundaries further, in the near future we foresee important advances and insights from studies appreciating the variation in recombination rate and its biological corollaries. The integration of mechanistic knowledge and combined insight from various estimation methods and across taxonomic scales will start paving the way.

Published online: 29 May 2020

1. Cavalier-Smith, T. Origins of the machinery of recombination and sex. *Heredity* **88**, 125–141 (2002).
2. Ortiz-Barrientos, D., Engelstädter, J. & Rieseberg, L. H. Recombination rate evolution and the origin of species. *Trends Ecol. Evol.* **31**, 226–236 (2016).
3. Hansen, T. F. The evolution of genetic architecture. *Annu. Rev. Ecol. Evol. Syst.* **37**, 123–157 (2006).
4. Cooper, T. F. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS Biol.* **5**, e225 (2007).
5. Butlin, R. K. Recombination and speciation. *Mol. Ecol.* **14**, 2621–2635 (2005).
This influential perspective article discusses the variation in recombination, theoretical expectations and its importance for speciation.
6. Otto, S. P. & Lenormand, T. Evolution of sex: resolving the paradox of sex and recombination. *Nat. Rev. Genet.* **3**, 252 (2002).
7. Kaniecki, K., De Tullio, L. & Greene, E. C. A change of view: homologous recombination at single-molecule resolution. *Nat. Rev. Genet.* **19**, 191–207 (2017).
8. Cromie, G. A., Connelly, J. C. & Leach, D. R. Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. *Mol. Cell* **8**, 1163–1174 (2001).
9. Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W. & Smadja, C. M. Recombination: the good, the bad and the variable. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20170279 (2017).
10. Ritz, K. R., Noor, M. A. F. & Singh, N. D. Variation in recombination rate: adaptive or not? *Trends Genet.* **33**, 364–374 (2017).
This review emphasizes population-level variation in recombination rates and discusses the potential molecular constraints and evolutionary processes underlying this variation.
11. Clark, A. G., Wang, X. & Matise, T. Contrasting methods of quantifying fine structure of human recombination. *Annu. Rev. Genomics Hum. Genet.* **11**, 45–64 (2010).
This review compares the different recombination inference methods (linkage disequilibrium-based, pedigree-based, sperm-typing and hotspot detection) in humans where the most extensive data are available.
12. Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).
13. Andolfatto, P. & Przeworski, M. Regions of lower crossing over harbor more rare variants in African populations of *Drosophila melanogaster*. *Genetics* **158**, 657–665 (2001).
14. Ellegren, H. & Galtier, N. Determinants of genetic diversity. *Nat. Rev. Genet.* **17**, 422–433 (2016).
15. Comeron, J. M. Background selection as null hypothesis in population genomics: insights and challenges from *Drosophila* studies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160471 (2017).
16. Lindholm, A. K. et al. The ecology and evolutionary dynamics of meiotic drive. *Trends Ecol. Evol.* **31**, 315–326 (2016).
17. Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).
This article presents the theory describing the behaviour of two linked loci under selection.

18. Gossmann, T. I., Santure, A. W., Sheldon, B. C., Slate, J. & Zeng, K. Highly variable recombinational landscape modulates efficacy of natural selection in birds. *Genome Biol. Evol.* **6**, 2061–2075 (2014).
19. Barton, N. H. A general model for the evolution of recombination. *Genet. Res.* **65**, 123–145 (1995). **This article describes mathematical models of how recombination evolves via selection on recombination modifiers**.
20. Rice, W. R. Evolution of sex: experimental tests of the adaptive significance of sexual recombination. *Nat. Rev. Genet.* **3**, 241 (2002).
21. Charlesworth, B. Recombination modification in a fluctuating environment. *Genetics* **83**, 181–195 (1976).
22. Charlesworth, B. & Barton, N. H. Recombination load associated with selection for increased recombination. *Genet. Res.* **67**, 27–41 (1996).
23. Rattray, A., Santoyo, G., Shafer, B. & Strathern, J. N. Elevated mutation rate during meiosis in *Saccharomyces cerevisiae*. *PLoS Genet.* **11**, e1004910 (2015).
24. Bolívar, P. et al. Biased inference of selection due to GC-biased gene conversion and the rate of protein evolution in flycatchers when accounting for it. *Mol. Biol. Evol.* **35**, 2475–2486 (2018).
25. Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009).
26. Schumer, M. et al. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* **360**, 656–660 (2018).
27. Martin, S. H., Davey, J. W., Salazar, C. & Jiggins, C. D. Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol.* **17**, e2006288 (2019).
28. Barton, N. & Bengtsson, B. O. The barrier to genetic exchange between hybridising populations. *Heredity* **57**, 357–376 (1986).
29. Nachman, M. W. & Payseur, B. A. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 409–421 (2012).
30. Melamed-Bessudo, C., Shilo, S. & Levy, A. A. Meiotic recombination and genome evolution in plants. *Curr. Opin. Plant. Biol.* **30**, 82–87 (2016).
31. Capilla, L., Garcia Caldés, M. & Ruiz-Herrera, A. Mammalian meiotic recombination: a toolbox for genome evolution. *Cytogenet. Genome Res.* **150**, 1–16 (2016).
32. Nam, K. & Ellegren, H. Recombination drives vertebrate genome contraction. *PLoS Genet.* **8**, e1002680 (2012).
33. Corbett-Detig, R. B., Hartl, D. L. & Sackton, T. B. Natural selection constrains neutral diversity across a wide range of species. *PLoS Biol.* **13**, e1002112 (2015).
34. Vandiedonck, C. & Knight, J. C. The human major histocompatibility complex as a paradigm in genomics research. *Brief. Funct. Genomic. Proteomic.* **8**, 379–394 (2009).
35. Charlesworth, B. The evolution of sex chromosomes. *Science* **251**, 1030–1033 (1991).
36. Ellegren, H. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. *Nat. Rev. Genet.* **12**, 157–166 (2011).
37. Schwander, T., Libbrecht, R. & Keller, L. Supergenes and complex phenotypes. *Curr. Biol.* **24**, R288–R294 (2014).
38. Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W. & Smadja, C. M. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160455 (2017). **This review, part of the special issue 'Evolutionary Causes and Consequences of Recombination Rate Variation in Sexual Organisms', includes a meta-analysis characterizing the variation in recombination rates across the tree of life and outlines outstanding questions in the field**.
39. Wang, J., Street, N. R., Scofield, D. G. & Ingvarsson, P. K. Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related *Populus* species. *Genetics* **202**, 1185–1200 (2016).
40. Dumont, B. L. & Payseur, B. A. Evolution of the genomic rate of recombination in mammals. *Evolution* **62**, 276–294 (2008).
41. Dumont, B. L., White, M. A., Steffy, B., Wiltshire, T. & Payseur, B. A. Extensive recombination rate variation in the house mouse species complex inferred from genetic linkage maps. *Genome Res.* **21**, 114–125 (2011).
42. Graffelman, J., Balding, D. J., Gonzalez-Neira, A. & Bertranpetit, J. Variation in estimated recombination rates across human populations. *Hum. Genet.* **122**, 301–310 (2007).
43. Thomsen, H. et al. A whole genome scan for differences in recombination rates among three *Bos taurus* breeds. *Mamm. Genome* **12**, 724–728 (2001).
44. Cheung, V. G., Burdick, J. T., Hirschmann, D. & Morley, M. Polymorphic variation in human meiotic recombination. *Am. J. Hum. Genet.* **80**, 526–530 (2007).
45. Johnston, S. E., Bérénos, C., Slate, J. & Pemberton, J. M. Conserved genetic architecture underlying individual recombination rate variation in a wild population of Soay sheep (*Ovis aries*). *Genetics* **203**, 583–598 (2016). **This article examines recombination rate variation and isolation of candidate genetic modifier loci in a natural, pedigreed population**.
46. Hussin, J., Roy-Gagnon, M.-H., Gendron, R., Andelfinger, G. & Awadalla, P. Age-dependent recombination rates in human pedigrees. *PLoS Genet.* **7**, e1002251 (2011).
47. Singh, N. D. Wolbachia infection associated with increased recombination in *Drosophila*. *G3* **9**, 229–237 (2019).
48. Berset-Brändli, L., Jaquiéry, J., Broquet, T., Ulrich, Y. & Perrin, N. Extreme heterochiasmy and nascent sex chromosomes in European tree frogs. *Proc. R. Soc. B: Biol. Sci.* **275**, 1577–1585 (2008).
49. Lenormand, T. & Dutheil, J. Recombination difference between sexes: a role for haploid selection. *PLoS Biol.* **3**, e63 (2005).
50. Haldane, J. B. S. Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* **12**, 101–109 (1922).
51. Huxley, J. S. Sexual difference of linkage in *Gammarus chevreuxi*. *J. Genet.* **20**, 145–156 (1928).
52. Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160456 (2017).
53. Bergero, R. & Charlesworth, D. The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* **24**, 94–102 (2009).
54. Limborg, M. T., McKinney, G. J., Seeb, L. W. & Seeb, J. E. Recombination patterns reveal information about centromere location on linkage maps. *Mol. Ecol. Resour.* **16**, 655–661 (2016).
55. Vincenten, N. et al. The kinetochore prevents centromere-proximal crossover recombination during meiosis. *eLife* **4**, e10850 (2015).
56. Haenel, Q., Laurentino, T. G., Roesti, M. & Berner, D. Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Mol. Ecol.* **27**, 2477–2497 (2018).
57. Morgan, A. P. et al. Structural variation shapes the landscape of recombination in mouse. *Genetics* **206**, 603–619 (2017).
58. Völker, M. et al. Copy number variation, chromosome rearrangement, and their association with recombination during avian evolution. *Genome Res.* **20**, 503–511 (2010).
59. Fullerton, S. M., Bernardo Carvalho, A. & Clark, A. G. Local rates of recombination are positively correlated with GC content in the human genome. *Mol. Biol. Evol.* **18**, 1139–1142 (2001).
60. Marsolier-Kergoat, M.-C. & Yeramian, E. GC content and recombination: reassessing the causal effects for the *Saccharomyces cerevisiae* genome. *Genetics* **183**, 31–38 (2009).
61. Freudenberg, J., Wang, M., Yang, Y. & Li, W. Partial correlation analysis indicates causal relationships between GC-content, exon density and recombination rate in the human genome. *BMC Bioinforma.* **10** (Suppl 1), S66 (2009).
62. Kent, T. V., Uzunović, J. & Wright, S. I. Coevolution between transposable elements and recombination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160458 (2017).
63. Rizzon, C., Marais, G., Gouy, M. & Biémont, C. Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. *Genome Res.* **12**, 400–407 (2002).
64. Dolgin, E. S. & Charlesworth, B. The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* **178**, 2169–2177 (2008).
65. Myers, S., Bottolo, L., Freeman, C., McVean, G. & Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310**, 321–324 (2005).
66. Paigen, K. & Petkov, P. Mammalian recombination hot spots: properties, control and evolution. *Nat. Rev. Genet.* **11**, 221–233 (2010).
67. Lichten, M. & Goldman, A. S. Meiotic recombination hotspots. *Annu. Rev. Genet.* **29**, 423–444 (1995).
68. Paul, P., Nag, D. & Chakraborty, S. Recombination hotspots: models and tools for detection. *DNA Repair* **40**, 47–56 (2016). **This detailed review covers recombination hotspot evolution, molecular mechanisms underlying recombination and comparison of various inference methods of hotspot detection**.
69. Arnheim, N., Calabrese, P. & Tiemann-Boege, I. Mammalian meiotic recombination hot spots. *Annu. Rev. Genet.* **41**, 369–399 (2007).
70. Choi, K. & Henderson, I. R. Meiotic recombination hotspots — a comparative view. *Plant J.* **83**, 52–61 (2015).
71. Weng, Z. et al. Identification of recombination hotspots and quantitative trait loci for recombination rate in layer chickens. *J. Anim. Sci. Biotechnol.* **10**, 20 (2019).
72. Kong, A. et al. Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467**, 1099–1103 (2010).
73. Hunter, C. M., Huang, W., Mackay, T. F. C. & Singh, N. D. The genetic architecture of natural variation in recombination rate in *Drosophila melanogaster*. *PLoS Genet.* **12**, e1005951 (2016).
74. Chinnici, J. P. Modification of recombination frequency in *Drosophila*. I. Selection for increased and decreased crossing over. *Genetics* **69**, 71–83 (1971).
75. Shaw, D. D. Genetic and environmental components of chiasma control. *Chromosoma* **37**, 297–308 (1972).
76. Parsons, P. A. Evolutionary rates: effects of stress upon recombination. *Biol. J. Linn. Soc. Lond.* **35**, 49–68 (1988).
77. Stevison, L. S., Sefick, S., Rushton, C. & Graze, R. M. Recombination rate plasticity: revealing mechanisms by design. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160459 (2017).
78. Lloyd, A., Morgan, C., H Franklin, F. C. & Bomblies, K. Plasticity of meiotic recombination rates in response to temperature in *Arabidopsis*. *Genetics* **208**, 1409–1420 (2018).
79. Neel, J. V. A relation between larval nutrition and the frequency of crossing over in the third chromosome of *Drosophila melanogaster*. *Genetics* **26**, 506–516 (1941).
80. Belyaev, D. K. & Borodin, P. M. The influence of stress on variation and its role in evolution. *Biol. Zent. Bl.* **101**, 705–714 (1982).
81. Kong, A. et al. Recombination rate and reproductive success in humans. *Nat. Genet.* **36**, 1203–1206 (2004).
82. Lobkovsky, A. E., Wolf, Y. I. & Koonin, E. V. Evolvability of an optimal recombination rate. *Genome Biol. Evol.* **8**, 70–77 (2015).
83. Alves, I., Houle, A. A., Hussin, J. G. & Awadalla, P. The impact of recombination on human mutation load and disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160465 (2017).
84. Székvölgyi, L., Ohta, K. & Nicolas, A. Initiation of meiotic homologous recombination: flexibility, impact of histone modifications, and chromatin remodeling. *Cold Spring Harb. Perspect. Biol.* **7**, a016527 (2015).
85. Salathé, M., Kouyos, R. D., Regoes, R. R. & Bonhoeffer, S. Rapid parasite adaptation drives selection for high recombination rates. *Evolution* **62**, 295–300 (2008).
86. Lenormand, T. & Otto, S. P. The evolution of recombination in a heterogeneous environment. *Genetics* **156**, 423–438 (2000).
87. Mank, J. E. The evolution of heterochiasmy: the role of sexual selection and sperm competition in determining sex-specific recombination rates in eutherian mammals. *Genet. Res.* **91**, 355–363 (2009).
88. Brandvain, Y. & Coop, G. Scrambling eggs: meiotic drive and the evolution of female recombination rates. *Genetics* **190**, 709–723 (2012).
89. Zelkowski, M., Olson, M. A., Wang, M. & Pawlowski, W. Diversity and determinants of meiotic recombination landscapes. *Trends Genet.* **35**, 359–370 (2019). **This review considers the difference between patterns of DSBs and crossover events along the genome across many organisms and discusses the molecular determinants that govern variation in recombination hotspots and landscapes**.
90. Gray, S. & Cohen, P. E. Control of meiotic crossovers: from double-strand break formation to designation. *Annu. Rev. Genet.* **50**, 175–210 (2016).

91. Petkov, P. M., Broman, K. W., Szatkiewicz, J. P. & Paigen, K. Crossover interference underlies sex differences in recombination rates. *Trends Genet.* **23**, 539–542 (2007).
92. Zhang, L., Liang, Z., Hutchinson, J. & Kleckner, N. Crossover patterning by the beam-film model: analysis and implications. *PLoS Genet.* **10**, e1004042 (2014).
93. Kleckner, N. et al. A mechanical basis for chromosome function. *Proc. Natl Acad. Sci. USA* **101**, 12592–12597 (2004).
94. Kirkpatrick, M. How and why chromosome inversions evolve. *PLoS Biol.* **8**, e1000501 (2010).
95. Farré, M., Micheletti, D. & Ruiz-Herrera, A. Recombination rates and genomic shuffling in human and chimpanzee — a new twist in the chromosomal speciation theory. *Mol. Biol. Evol.* **30**, 853–864 (2013).
96. Crown, K. N., Miller, D. E., Sekelsky, J. & Hawley, R. S. Local inversion heterozygosity alters recombination throughout the genome. *Curr. Biol.* **28**, 2984–2990.e3 (2018).
97. Mirouze, M. et al. Loss of DNA methylation affects the recombination landscape in Arabidopsis. *Proc. Natl Acad. Sci. USA* **109**, 5880–5885 (2012).
98. Brachet, E., Sommermeyer, V. & Borde, V. Interplay between modifications of chromatin and meiotic recombination hotspots. *Biol. Cell* **104**, 51–69 (2012).
99. Marand, A. P. et al. Meiotic crossovers are associated with open chromatin and enriched with Stowaway transposons in potato. *Genome Biol.* **18**, 203 (2017).
100. Qiao, H. et al. Antagonistic roles of ubiquitin ligase HEI10 and SUMO ligase RNF212 regulate meiotic recombination. *Nat. Genet.* **46**, 194–199 (2014).
101. Ziolkowski, P. A. et al. Natural variation and dosage of the HEI10 meiotic E3 ligase control *Arabidopsis* crossover recombination. *Genes Dev.* **31**, 306–317 (2017).
102. Kong, A. et al. Sequence variants in the RNF212 gene associate with genome-wide recombination rate. *Science* **319**, 1398–1401 (2008).
103. Baudat, F. et al. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327**, 836–840 (2010).
104. Parvanov, E. D., Petkov, P. M. & Paigen, K. Prdm9 controls activation of mammalian recombination hotspots. *Science* **327**, 835 (2010).
105. Paigen, K. & Petkov, P. M. PRDM9 and its role in genetic recombination. *Trends Genet.* **34**, 291–300 (2018).
106. He, Y. et al. Genomic features shaping the landscape of meiotic double-strand-break hotspots in maize. *Proc. Natl Acad. Sci. USA* **114**, 12231–12236 (2017).
107. Shilo, S., Melamed-Bessudo, C., Dorone, Y., Barkai, N. & Levy, A. A. DNA crossover motifs associated with epigenetic modifications delineate open chromatin regions in *Arabidopsis*. *Plant Cell* **27**, 2427–2436 (2015).
108. Lange, J. et al. The landscape of mouse meiotic double-strand break formation, processing, and repair. *Cell* **167**, 695–708.e16 (2016).
109. Singhal, S. et al. Stable recombination hotspots in birds. *Science* **350**, 928–932 (2015).
This article presents a good example of the population-based approach identifying recombination hotpots, their evolutionary stability and the underlying genomic features in avian populations.
110. Anderson, L. K. et al. High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics* **165**, 849–865 (2003).
This article is one of the early studies to use MLH1 foci to estimate recombination frequency across the genome.
111. Rahn, M. I. & Solari, A. J. Recombination nodules in the oocytes of the chicken, *Gallus domesticus*. *Cytogenet. Cell Genet.* **43**, 187–193 (1986).
112. Pollock, D. L. & Fechheimer, N. S. The chromosomes of cockerels (*Gallus domesticus*) during meiosis. *Cytogenet. Cell Genet.* **21**, 267–281 (1978).
113. Lawrie, N. M., Tease, C. & Hultén, M. A. Chiasma frequency, distribution and interference maps of mouse autosomes. *Chromosoma* **104**, 308–314 (1995).
114. Herickhoff, L., Stack, S. & Sherman, J. The relationship between synapsis, recombination nodules and chiasmata in tomato translocation heterozygotes. *Heredity* **71**, 373–385 (1993).
115. Holm, P. B. & Rasmussen, S. W. Chromosome pairing, recombination nodules and chiasma formation in diploid *Bombyx* males. *Carlsberg Res. Commun.* **45**, 483 (1980).
116. Rasmussen, S. W. & Holm, P. B. The synaptonemal complex, recombination nodules and chiasmata in human spermatocytes. *Symp. Soc. Exp. Biol.* **38**, 271–292 (1984).
117. Calderón, P. L. & Pigozzi, M. I. MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Res.* **14**, 605–612 (2006).
118. Sun, F. et al. Variation in MLH1 distribution in recombination maps for individual chromosomes from human males. *Hum. Mol. Genet.* **15**, 2376–2391 (2006).
119. Anderson, L. K., Reeves, A., Webb, L. M. & Ashley, T. Distribution of crossing over on mouse synaptonemal complexes using immunofluorescent localization of MLH1 protein. *Genetics* **151**, 1569–1579 (1999).
120. del Priore, L. & Pigozzi, M. I. Heterologous synapsis and crossover suppression in heterozygotes for a pericentric inversion in the zebra finch. *Cytogenet. Genome Res.* **147**, 154–160 (2015).
121. Zickler, D., Moreau, P. J., Huynh, A. D. & Slezec, A. M. Correlation between pairing initiation sites, recombination nodules and meiotic recombination in *Sordaria macrospora*. *Genetics* **132**, 135–148 (1992).
122. Gruhn, J. R., Rubio, C., Broman, K. W., Hunt, P. A. & Hassold, T. Cytological studies of human meiosis: sex-specific differences in recombination originate at, or prior to, establishment of double-strand breaks. *PLoS One* **8**, e85075 (2013).
123. Golding, G. B. The sampling distribution of linkage disequilibrium. *Genetics* **108**, 257–274 (1984).
124. Pritchard, J. K. & Przeworski, M. Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* **69**, 1–14 (2001).
125. Griffiths, R. C. & Marjoram, P. in *Progress in Population Genetics and Human Evolution* (eds Donnelly P. & Tavaré S.) 257–270 (Springer, 1997).
This article presents a theoretical formalization of the evolutionary history of recombination through the ancestral recombination graph.
126. Arenas, M. The importance and application of the ancestral recombination graph. *Front. Genet.* **4**, 206 (2013).
127. McVean, G., Awadalla, P. & Fearnhead, P. A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics* **160**, 1231–1241 (2002).
128. Stumpf, M. P. H. & McVean, G. A. T. Estimating recombination rates from population-genetic data. *Nat. Rev. Genet.* **4**, 959–968 (2003).
This Review provides details on the models, assumptions and inference methods of the population recombination rate and compares applications in human populations.
129. Hellenthal, G. & Stephens, M. Insights into recombination from population genetic variation. *Curr. Opin. Genet. Dev.* **16**, 565–572 (2006).
130. Hudson, R. R. & Kaplan, N. L. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**, 147–164 (1985).
131. Myers, S. R. & Griffiths, R. C. Bounds on the minimum number of recombination events in a sample history. *Genetics* **163**, 375–394 (2003).
132. Hudson, R. R. Two-locus sampling distributions and their application. *Genetics* **159**, 1805–1817 (2001).
This article presents an influential description of how the population recombination rate can be estimated from data.
133. Vijay, N. et al. Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nat. Commun.* **7**, 13195 (2016).
134. Chan, A. H., Jenkins, P. A. & Song, Y. S. Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. *PLoS Genet.* **8**, e1003090 (2012).
135. Calafell, F., Grigorenko, E. L., Chikanian, A. A. & Kidd, K. K. Haplotype evolution and linkage disequilibrium: a simulation study. *Hum. Hered.* **51**, 85–96 (2001).
136. Wang, N., Akey, J. M., Zhang, K., Chakraborty, R. & Jin, L. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am. J. Hum. Genet.* **71**, 1227–1234 (2002).
137. Barton, N. H. Genetic hitchhiking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355**, 1553–1562 (2000).
138. Charlesworth, B., Nordborg, M. & Charlesworth, D. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* **70**, 155–174 (1997).
139. Chapman, N. H. & Thompson, E. A. Linkage disequilibrium mapping: the role of population history, size, and structure. *Adv. Genet.* **42**, 413–437 (2001).
140. Baird, S. J. E. Exploring linkage disequilibrium. *Mol. Ecol. Resour.* **15**, 1017–1019 (2015).
141. Nordborg, M. in *Handbook of Statistical Genetics* (eds Balding, D. J., Bishop, M. & Cannings, C.) 179–212 (Wiley, 2004).
142. Lander, E. S. & Green, P. Construction of multilocus genetic linkage maps in humans. *Proc. Natl Acad. Sci. USA* **84**, 2363–2367 (1987).
This article is the first description of inferring multilocus linkage maps using maximum likelihood in a three-generation human pedigree.
143. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**, 314–331 (1980).
144. Johnston, S. E., Huisman, J., Ellis, P. A. & Pemberton, J. M. A high-density linkage map reveals sexual dimorphism in recombination landscapes in red deer (*Cervus elaphus*). *G3* **7**, 2859–2870 (2017).
145. Peñalba, J. V. et al. Genome of an iconic Australian bird: high-quality assembly and linkage map of the superb fairy-wren (*Malurus cyaneus*). *Mol. Ecol. Resour.* **20**, 560–578 (2020).
146. Decker, J. E. et al. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet.* **10**, e1004254 (2014).
147. Fang, L. et al. Positive correlation between recombination rate and nucleotide diversity is shown under domestication selection in the chicken genome. *Chin. Sci. Bull.* **53**, 746–750 (2008).
148. Kong, A. et al. A high-resolution recombination map of the human genome. *Nat. Genet.* **31**, 241–247 (2002).
149. Lashermes, P. et al. Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome* **44**, 589–596 (2001).
150. Sun, Z. et al. An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. *Theor. Appl. Genet.* **114**, 1305–1317 (2007).
151. Bowers, J. E. et al. A high-density genetic recombination map of sequence-tagged sites for sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. *Genetics* **165**, 367–386 (2003).
152. Pombi, M., Stump, A. D., Della Torre, A. & Besansky, N. J. Variation in recombination rate across the X chromosome of *Anopheles gambiae*. *Am. J. Trop. Med. Hyg.* **75**, 901–903 (2006).
153. O'Connell, J. R. & Weeks, D. E. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.* **63**, 259–266 (1998).
154. Blouin, M. S. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends Ecol. Evol.* **18**, 503–511 (2003).
155. Fierst, J. L. Using linkage maps to correct and scaffold de novo genome assemblies: methods, challenges, and computational tools. *Front. Genet.* **6**, 220 (2015).
156. Green, P., Falls, K. & Crooks, S. *Documentation for CRI-MAP, version 2.4* (Washington Univ. School of Medicine, 1990).
157. Wu, Y., Bhat, P. R., Close, T. J. & Lonardi, S. Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet.* **4**, e1000212 (2008).
158. Rastas, P. Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics* **33**, 3726–3732 (2017).
159. Rastas, P., Paulin, L., Hanski, I., Lehtonen, R. & Auvinen, P. Lep-MAP: fast and accurate linkage map construction for large SNP datasets. *Bioinformatics* **29**, 3128–3134 (2013).
160. Miar, Y., Sargolzaei, M. & Schenkel, F. S. A comparison of different algorithms for phasing haplotypes using Holstein cattle genotypes and pedigree data. *J. Dairy. Sci.* **100**, 2837–2849 (2017).
161. Hickey, J. M. et al. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. *Genet. Sel. Evol.* **43**, 12 (2011).
162. DeWan, A. T., Parrado, A. R., Matise, T. C. & Leal, S. M. The map problem: a comparison of genetic and sequence-based physical maps. *Am. J. Hum. Genet.* **70**, 101–107 (2002).

163. Zhao, H. & Speed, T. P. On genetic map functions. *Genetics* **142**, 1369–1377 (1996).
164. Chakravarti, A. A graphical representation of genetic and physical maps: the Marey map. *Genomics* **11**, 219–222 (1991).
165. Rezvoy, C., Charif, D., Guéguen, L. & Marais, G. A. B. MareyMap: an R-based tool with graphical interface for estimating recombination rates. *Bioinformatics* **23**, 2188–2189 (2007).
166. Berloff, N., Perola, M. & Lange, K. Spline methods for the comparison of physical and genetic maps. *J. Comput. Biol.* **9**, 465–475 (2002).
167. Yu, A. et al. Comparison of human genetic and sequence-based physical maps. *Nature* **409**, 951–953 (2001).
168. Carrington, M. & Cullen, M. Justified chauvinism: advances in defining meiotic recombination through sperm typing. *Trends Genet.* **20**, 196–205 (2004).
169. Sun, H. et al. Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. *Nat. Commun.* **10**, 4310 (2019). **This article is one of the first studies to perform bulk-gamete sequencing using linked-read technology to infer the recombination rate across the genome**.
170. Arbeithuber, B., Betancourt, A. J., Ebner, T. & Tiemann-Boege, I. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proc. Natl Acad. Sci. USA* **112**, 2109–2114 (2015).
171. Li, H. H. et al. Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature* **335**, 414–417 (1988).
172. Huehn, M. On the bias of recombination fractions, Kosambi's and Haldane's distances based on frequencies of gametes. *Genome* **54**, 196–201 (2011).
173. O'Reilly, P. F., Birney, E. & Balding, D. J. Confounding between recombination and selection, and the Ped/Pop method for detecting selection. *Genome Res.* **18**, 1304–1313 (2008).
174. Slotte, T. The impact of linked selection on plant genomic variation. *Brief. Funct. Genomics* **13**, 268–275 (2014).
175. Arenas, M., Lopes, J. S., Beaumont, M. A. & Posada, D. CodABC: a computational framework to coestimate recombination, substitution, and molecular adaptation rates by approximate Bayesian computation. *Mol. Biol. Evol.* **32**, 1109–1112 (2015).
176. Hermann, P., Heissl, A., Tiemann-Boege, I. & Futschik, A. LDJump: estimating variable recombination rates from population genetic data. *Mol. Ecol. Resour.* **19**, 623–638 (2019).
177. Beeravolu, C. R., Hickerson, M. J., Frantz, L. A. F. & Lohse, K. ABLE: blockwise site frequency spectra for inferring complex population histories and recombination. *Genome Biol.* **19**, 145 (2018).
178. Dréau, A., Venu, V., Avdievich, E., Gaspar, L. & Jones, F. C. Genome-wide recombination map construction from single individuals using linked-read sequencing. *Nat. Commun.* **10**, 4309 (2019). **This article is one of the first studies to perform bulk-gamete sequencing using linked-read technology to infer the recombination rate across the genome**.
179. Kawakami, T. et al. Whole-genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences of fine-scale recombination rate variation in birds. *Mol. Ecol.* **26**, 4158–4172 (2017).
180. Smukowski Heil, C. S., Ellison, C., Dubin, M. & Noor, M. A. F. Recombining without hotspots: a comprehensive evolutionary portrait of recombination in two closely related species of *Drosophila*. *Genome Biol. Evol.* **7**, 2829–2842 (2015).
181. Gabriel, S. B. et al. The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
182. Vijay, N. et al. Genomewide patterns of variation in genetic diversity are shared among populations, species and higher-order taxa. *Mol. Ecol.* **26**, 4284–4295 (2017).
183. Paape, T. et al. Fine-scale population recombination rates, hotspots, and correlates of recombination in the *Medicago truncatula* genome. *Genome Biol. Evol.* **4**, 726–737 (2012).
184. Weissensteiner, M. H. et al. Combination of short-read, long-read, and optical mapping assemblies reveals large-scale tandem repeat arrays with population genetic implications. *Genome Res.* **27**, 697–708 (2017).
185. Savolainen, O., Lascoux, M. & Merilä, J. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14**, 807–820 (2013).
186. Ravinet, M. et al. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* **30**, 1450–1477 (2017).
187. Wolf, J. B. W. & Ellegren, H. Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* **18**, 87–100 (2017).
188. Jeffreys, A. J. & Neumann, R. Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot. *Hum. Mol. Genet.* **14**, 2277–2287 (2005).
189. McGaugh, S. E. et al. Recombination modulates how selection affects linked sites in *Drosophila*. *PLoS Biol.* **10**, e1001422 (2012).
190. Burri, R. et al. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Res.* **25**, 1656–1665 (2015).
191. Rueppell, O., Kuster, R., Miller, K. & Fouks, B. A new Metazoan recombination rate record and consistently high recombination rates in the honey bee genus *Apis* accompanied by frequent inversions but not translocations. *Genome Biol. Evol.* **8**, 3653–3660 (2016).
192. Guerrero, R. F., Rousset, F. & Kirkpatrick, M. Coalescent patterns for chromosomal inversions in divergent populations. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 430–438 (2012).
193. Jeffreys, A. J., Ritchie, A. & Neumann, R. High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. *Hum. Mol. Genet.* **9**, 725–733 (2000).
194. Jeffreys, A. J., Neumann, R., Panayi, M., Myers, S. & Donnelly, P. Human recombination hot spots hidden in regions of strong marker association. *Nat. Genet.* **37**, 601–606 (2005).
195. Morris, G. P. et al. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl Acad. Sci. USA* **110**, 453–458 (2013).
196. Knief, U. et al. Epistatic mutations under divergent selection govern phenotypic variation in the crow hybrid zone. *Nat. Ecol. Evol.* **3**, 570–576 (2019).
197. Roesti, M., Moser, D. & Berner, D. Recombination in the threespine stickleback genome—patterns and consequences. *Mol. Ecol.* **22**, 3014–3027 (2013).
198. Liu, H. et al. Causes and consequences of crossing-over evidenced via a high-resolution recombinational landscape of the honey bee. *Genome Biol.* **16**, 15 (2015).
199. Reddy, U. K. et al. High-resolution genetic map for understanding the effect of genome-wide recombination rate on nucleotide diversity in watermelon. *G3* **4**, 2219–2230 (2014).
200. Palomar, G. et al. Comparative high-density linkage mapping reveals conserved genome structure but variation in levels of heterochiasmy and location of recombination cold spots in the common frog. *G3* **7**, 637–645 (2017).
201. Jiang, H. et al. High recombination rates and hotspots in a *Plasmodium falciparum* genetic cross. *Genome Biol.* **12**, R33 (2011).
202. van Oers, K. et al. Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. *Heredity* **112**, 307–316 (2014).
203. Tortereau, F. et al. A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC Genomics* **13**, 586 (2012).
204. Petit, M. et al. Variation in recombination rate and its genetic determinism in sheep populations. *Genetics* **207**, 767–784 (2017).
205. Limborg, M. T., Waples, R. K., Allendorf, F. W. & Seeb, J. E. Linkage mapping reveals strong chiasma interference in sockeye salmon: implications for interpreting genomic data. *G3* **5**, 2463–2473 (2015).
206. Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B. & Pang, E. C. K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* **142**, 169–196 (2005).
207. Yue, G. H. Recent advances of genome mapping and marker-assisted selection in aquaculture. *Fish. Fish.* **15**, 376–396 (2014).
208. Joron, M. et al. A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. *PLoS Biol.* **4**, e303 (2006).
209. Dixon, G. B. et al. Genomic determinants of coral heat tolerance across latitudes. *Science* **348**, 1460–1462 (2015).
210. Recoquillay, J. et al. A medium density genetic map and QTL for behavioral and production traits in Japanese quail. *BMC Genomics* **16**, 10 (2015).
211. Qiu, D. et al. A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theor. Appl. Genet.* **114**, 67–80 (2006).
212. Cullen, M., Perfetto, S. P., Klitz, W., Nelson, G. & Carrington, M. High-resolution patterns of meiotic recombination across the human major histocompatibility complex. *Am. J. Hum. Genet.* **71**, 759–776 (2002).
213. Hubert, R., MacDonald, M., Gusella, J. & Arnheim, N. High resolution localization of recombination hot spots using sperm typing. *Nat. Genet.* **7**, 420–424 (1994).
214. Yu, J. et al. Individual variation in recombination among human males. *Am. J. Hum. Genet.* **59**, 1186–1192 (1996).
215. Lien, S., Szyda, J., Schechinger, B., Rappold, G. & Arnheim, N. Evidence for heterogeneity in recombination in the human pseudoautosomal region: high resolution analysis by sperm typing and radiation-hybrid mapping. *Am. J. Hum. Genet.* **66**, 557–566 (2000).
216. Shi, Q. et al. Absence of age effect on meiotic recombination between human X and Y chromosomes. *Am. J. Hum. Genet.* **71**, 254–261 (2002).
217. Hinch, A. G. et al. Factors influencing meiotic recombination revealed by whole-genome sequencing of single sperm. *Science* **363**, eaau8861 (2019).
218. Shi, Q. et al. Single sperm typing demonstrates that reduced recombination is associated with the production of aneuploid 24,XY human sperm. *Am. J. Med. Genet.* **99**, 34–38 (2001).
219. Wang, J., Fan, H. C., Behr, B. & Quake, S. R. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell* **150**, 402–412 (2012). **This influential study describes a method to genotype single human sperm cells and estimate recombination and mutation rates using a genome-wide sequencing approach**.
220. Williams, C., Davies, D. & Williamson, R. Segregation of ΔF508 and normal CFTR alleles in human sperm. *Hum. Mol. Genet.* **2**, 445–448 (1993).
221. Dreissig, S., Fuchs, J., Himmelbach, A., Mascher, M. & Houben, A. Sequencing of single pollen nuclei reveals meiotic recombination events at megabase resolution and circumvents segregation distortion caused by postmeiotic processes. *Front. Plant. Sci.* **8**, 1620 (2017).
222. Ma, S., Ferguson, K. A., Arsovska, S., Moens, P. & Chow, V. Reduced recombination associated with the production of aneuploid sperm in an infertile man: a case report. *Hum. Reprod.* **21**, 980–985 (2006).
223. Guryev, V. et al. Haplotype block structure is conserved across mammals. *PLoS Genet.* **2**, e121 (2006).
224. Tishkoff, S. A. & Verrelli, B. C. Role of evolutionary history on haplotype block structure in the human genome: implications for disease mapping. *Curr. Opin. Genet. Dev.* **13**, 569–575 (2003).
225. Zhang, K., Calabrese, P., Nordborg, M. & Sun, F. Haplotype block structure and its applications to association studies: power and study designs. *Am. J. Hum. Genet.* **71**, 1386–1394 (2002).
226. Ringbauer, H., Coop, G. & Barton, N. H. Inferring recent demography from isolation by distance of long shared sequence blocks. *Genetics* **205**, 1335–1351 (2017).
227. Sedghifar, A., Brandvain, Y. & Ralph, P. Beyond clines: lineages and haplotype blocks in hybrid zones. *Mol. Ecol.* **25**, 2559–2576 (2016).
228. Laayouni, H. et al. Similarity in recombination rate estimates highly correlates with genetic differentiation in humans. *PLoS One* **6**, e17913 (2011).
229. Sjödin, P., Kaj, I., Krone, S., Lascoux, M. & Nordborg, M. On the meaning and existence of an effective population size. *Genetics* **169**, 1061–1070 (2005).
230. Nachman, M. W. & Crowell, S. L. Estimate of the mutation rate per nucleotide in humans. *Genetics* **156**, 297–304 (2000).
231. Smeds, L., Qvarnström, A. & Ellegren, H. Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **26**, 1211–1218 (2016).
232. Scally, A. The mutation rate in human evolution and demographic inference. *Curr. Opin. Genet. Dev.* **41**, 36–43 (2016).

233. Auton, A. & McVean, G. Recombination rate estimation in the presence of hotspots. *Genome Res.* **17**, 1219–1227 (2007).
234. Xu, S. et al. A male-specific genetic map of the microcrustacean *Daphnia pulex* based on single-sperm whole-genome sequencing. *Genetics* **201**, 31–38 (2015).
235. Lu, S. et al. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science* **338**, 1627–1630 (2012).
236. Gao, F., Ming, C., Hu, W. & Li, H. New software for the fast estimation of population recombination rates (FastEPRR) in the genomic era. *G3* **6**, 1563–1571 (2016).
237. Wang, Y. & Rannala, B. Population genomic inference of recombination rates and hotspots. *Proc. Natl Acad. Sci. USA* **106**, 6215–6219 (2009).
238. V Barroso, G., Puzović, N. & Dutheil, J. Y. Inference of recombination maps from a single pair of genomes and its application to ancient samples. *PLoS Genet.* **15**, e1008449 (2019).
239. McVean, G. A. T. et al. The fine-scale structure of recombination rate variation in the human genome. *Science* **304**, 581–584 (2004).
240. Kamm, J. A., Spence, J. P., Chan, J. & Song, Y. S. Two-locus likelihoods under variable population size and fine-scale recombination rate estimation. *Genetics* **203**, 1381–1399 (2016).
241. Wilson, D. J. & McVean, G. Estimating diversifying selection and functional constraint in the presence of recombination. *Genetics* **172**, 1411–1425 (2006).
242. Humphreys, D. P., McGuirl, M. R., Miyagi, M. & Blumberg, A. J. Fast estimation of recombination rates using topological data analysis. *Genetics* **211**, 1191–1204 (2019).
243. Liu, D. et al. Construction and analysis of high-density linkage map using high-throughput sequencing data. *PLoS One* **9**, e98855 (2014).
244. van Ooijen, J. W. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genet. Res.* **93**, 343–349 (2011).
245. Rastas, P., Calboli, F. C. F., Guo, B., Shikano, T. & Merilä, J. Construction of ultradense linkage maps with Lep-MAP2: stickleback F2 recombinant crosses as an example. *Genome Biol. Evol.* **8**, 78–93 (2015).
246. Xu, P. et al. MRLR: unraveling high-resolution meiotic recombination by linked reads. *Bioinformatics* **36**, 10–16 (2020).
247. Knowlton, S. M., Sadasivam, M. & Tasoglu, S. Microfluidics for sperm research. *Trends Biotechnol.* **33**, 221–229 (2015).
248. Chen, M. et al. Comparison of multiple displacement amplification (MDA) and multiple annealing and looping-based amplification cycles (MALBAC) in single-cell sequencing. *PLoS One* **9**, e114520 (2014).
249. Huang, L., Ma, F., Chapman, A., Lu, S. & Xie, X. S. Single-cell whole-genome amplification and sequencing: methodology and applications. *Annu. Rev. Genomics Hum. Genet.* **16**, 79–102 (2015).
250. Curik, I., Ferenčaković, M. & Sölkner, J. Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livest. Sci.* **166**, 26–34 (2014).
251. Bahnak, B. R. et al. A simple and efficient method for isolating high molecular weight DNA from mammalian sperm. *Nucleic Acids Res.* **16**, 1208 (1988).
252. Griffin, J. Methods of sperm DNA extraction for genetic and epigenetic studies. *Methods Mol. Biol.* **927**, 379–384 (2013).

Acknowledgements
The authors thank the Dobzhansky Journal Club in the Division of Evolutionary Biology, LMU, Munich, for useful discussion and comments on this review. Funding was provided by the European Research Council (ERCStG-336536 FuncSpecGen) and LMU Munich.

Author contributions
J.V.P. wrote the manuscript. J.B.W.W. edited the manuscript before submission. Both authors researched data for the article and substantially contributed to the discussion of content.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Genetics thanks J. McAuley, S. Johnston and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2020