Aberrant landscapes of maternal meiotic crossovers contribute to aneuploidies in human embryos

Daniel Ariad[1*], Svetlana Madjunkova[2,3], Mitko Madjunkov[2], Siwei Chen[2], Rina Abramov[2], Clifford Librach[2,4,5,6], Rajiv C. McCoy[1*]

[1]Department of Biology, Johns Hopkins University, Baltimore, MD, USA; [2]CReATe Fertility Centre, Toronto, Canada; [3]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada; [4]Department of Obstetrics and Gynecology, University of Toronto, Toronto, Canada; [5]Institute of Medical Sciences, University of Toronto, Toronto, Canada; [6]Department of Physiology, University of Toronto, Toronto, Canada

Abstract Meiotic recombination is a fundamental source of human genetic diversity and is also critical for ensuring the accuracy of chromosome segregation. Understanding the landscape of meiotic recombination, its variation across individuals, and the processes by which it goes awry are long-standing goals in human genetics. Current approaches for inferring the landscape of recombination either rely on population genetic patterns of linkage disequilibrium (LD)—capturing a time-averaged view—or direct detection of crossovers in gametes or multi-generation pedigrees, limiting the scale and availability of relevant datasets. Here, we introduce an approach for inferring sex-specific landscapes of recombination from retrospective analysis of data from preimplantation genetic testing for aneuploidy (PGT-A) based on low-coverage (<0.05×) whole-genome sequencing of biopsies from *in vitro* fertilized (IVF) embryos. To overcome the sparsity of these data, our method exploits its inherent relatedness structure, knowledge of haplotypes from external population reference panels, as well as the frequent occurrence of chromosome loss in embryos, whereby the remaining chromosome is phased by default. Based on extensive simulation, we show that our method retains high accuracy down to coverages as low as 0.02×. Applying this method to low-coverage PGT-A data from 18,967 embryos, we mapped 70,660 recombination events at an average resolution of ∼150 kbp, replicating key features of sex-specific recombination maps from the literature. We observed that the total length of the female genetic map is reduced for trisomies compared to disomies, while the genomic distribution of crossovers is also altered in a chromosome-specific manner. Based on haplotype configurations detected in regions surrounding the centromeres, our data additionally indicate that individual chromosomes possess unique propensities for different mechanisms of meiotic error. Together, our results provide a detailed view of the role of aberrant meiotic recombination in the origins of human aneuploidies, as well as a flexible tool for mapping crossovers in low-coverage sequencing data from multiple siblings.

Keywords: meiosis; recombination; trisomy; monosomy; in vitro fertilization; preimplantation genetic testing

Introduction

Recombination between homologous chromosomes is a key source of human genetic diversity [1, 2]. The crossovers that mediate such genetic exchanges during meiosis are also important for ensuring the accuracy of chromosome segregation [3, 4]. Notably, female meiosis initiates during fetal development, when homologs pair, acquire double-strand breaks, and establish crossovers that form physical linkages (chiasmata) to stabilize the chromosomes. Such chiasmata must then be maintained over decades-long meiotic arrest, until meiosis resumes at ovulation. Abnormal number and/or location of crossovers may predispose oocytes to gains or losses of whole chromosomes (aneuploidies), which are the leading cause of human pregnancy loss and congenital disorders [5]. Hypotheses about the role of recombination in aneuploidy formation largely originated from studies of model organisms [4, 6, 7]. Meanwhile, the smaller number of studies in humans have primarily focused on the subset of trisomies that are compatible with *in utero* development [8, 9, 10, 11], with less focus on the most common trisomies (chr15, chr16, and chr22) observed in oocytes and preimplantation embryos (though see [12, 13, 14]). To overcome this limitation, several previous studies have analyzed all products of meiosis (i.e., the first and second polar body, as well as a biopsy of the corresponding embryo) [15, 16]. While insightful, such sampling is technically demanding, limiting sample sizes and in turn limiting power and resolution for comparing genetic maps.

Over the last decade, several studies on human embryos have been conducted within the framework of preimplantation genetic testing for monogenic disorders (PGT-M) using methods such as Karyomapping [17], siCHILD/haplarithmisis [18], OnePGT [19], and GENType [20], whereby parental DNA is assayed along with that of the embryos, and unaffected embryos are prioritized for transfer. These genome-wide haplotyping methods allow mapping of maternal and paternal crossovers

*For correspondence: daniel@ariad.org (Daniel Ariad); rajiv.mccoy@jhu.edu (Rajiv C. McCoy)

along chromosomes. However, the number of patients that undergo PGT-M is small compared to the number of patients that undergo preimplantation genetic testing for aneuploidy (PGT-A), again limiting answers to broader questions about the crossover landscape. For example, a recent PGT-M study by Tšuiko et al. [21] inferred the parental and mechanistic origin of chromosome abnormalities in 2706 embryos from PGT-M patients and found 269 trisomies in total. A similar PGT-M study of recombination by Ma et al. [22] analyzed 1519 embryos and 353 autosomal aneuploidies.

Other current approaches for inferring the landscape of recombination rely either on population patterns of linkage disequilibrium (LD)—capturing a time-averaged view of historical recombination events—or direct detection of crossovers based on genotyping of haploid gametes or multi-generation pedigrees (e.g., parent-offspring trios), again limiting the scale and availability of relevant datasets [23, 24, 25, 26]. Moreover, most of these methods are designed for discovering recombination using data from normal, disomic chromosomes. Mapping meiotic crossovers in large samples of both normal and aneuploid embryos using a unified statistical framework would allow a robust test of the role of recombination in the genesis of aneuploidy.

To this end, we introduce a statistical approach tailored to sequencing data from PGT-A, which is based on low-coverage ($<0.05\times$ per homolog) whole-genome sequencing of biopsies from *in vitro* fertilized (IVF) embryos. We retrospectively apply our method to normal disomic chromosomes identified in existing low-coverage PGT-A data from 18,967 embryos and replicate features of sex-specific recombination maps that were previously described based on large prospective studies of living populations. We then extend our method to trisomies, testing the extent to which the landscape of recombination differs between normal and aneuploid chromosomes. Together, our study sheds light on the dual function of meiotic recombination in generating genetic diversity while ensuring fidelity of human meiosis.

Results

A method for inferring crossovers based on low-coverage sequencing data from multiple siblings

One general approach for discovering the genomic locations of meiotic recombination events is to compare genotype data from related individuals. Such data can be scanned to identify regions where haplotypes match (i.e., are identical by descent [IBD]). The boundaries of the matched haplotypes reflect the locations of meiotic crossovers in the history of the sample. The information gained by comparing haplotypes among relatives serves as the foundation for several different approaches for PGT-M [17, 18, 19]. However, directly calling diploid genotypes from sequencing data requires a minimum coverage of $2\times$ (to sample both alleles) and in practice requires coverage several fold higher to overcome technical challenges such as coverage variability, ambiguous alignments due to repetitive sequences, and other sequencing and analytic artifacts. Because data from PGT-A typically fall well below these coverage requirements, they are generally assumed unsuitable for applications that demand genotypes, including the study of recombination landscapes in embryos. However, as exemplified by common methods such as genotype imputation [27], knowledge of patterns of LD from external population genetic reference panels may facilitate the extraction of meaningful signal from sparse, low-coverage datasets, including in the context of prenatal genetics [28, 29].

Building on this logic, we introduce a haplotype matching approach, named Linkage Disequilibrium-Informed Comparison of Haplotypes Among Sibling Embryos (LD-CHASE), tailored to DNA sequencing data from PGT-A. Most current implementations of PGT-A involve low-coverage next-generation sequencing (NGS) of trophectoderm biopsies from IVF embryos at day 5 or 6 post-fertilization, with the goal of prioritizing chromosomally normal (i.e., euploid) embryos for transfer to improve IVF outcomes [30]. PGT-A offers a unique source of genomic data from large numbers of sibling samples, as each IVF cycle typically produces multiple embryos, and often multiple IVF cycles are necessary in infertility treatment.

Disomic chromosomes of any two sibling embryos will possess discrete genomic intervals with different counts of matching haplotypes, and transition points between these intervals reflect the locations of meiotic crossovers. The occurrence of monosomy (or uniparental isodisomy [isoUPD], isolated to individual chromosomes or genome-wide [GW-isoUPD]) among a set of sibling embryos greatly simplifies this comparison, as the remaining chromosome is phased by default, facilitating discovery of sex-specific crossovers (i.e., originating during gamete formation in one of the two parents). Here, we leverage the common occurrence of chromosome loss to reveal the sex-specific landscapes of meiotic crossovers among a large sample of IVF embryos.

Briefly, LD-CHASE uses sparse genotypes obtained from low-coverage sequencing data to identify the locations of meiotic crossovers (Figure 1 as well as Figure S1). At such coverages, direct comparison of haplotypes is not possible, as a small minority of the genome is covered by any sequencing reads, and positions of aligned reads from samples under comparison rarely overlap. We circumvent this challenge based on patterns of LD, whereby observations of a set of alleles from one sequencing read may provide indirect information about the probabilities of alleles at nearby, unobserved variant sites. This in turn informs the relative probability that a given pair of reads originated from identical homologous chromosomes versus from distinct

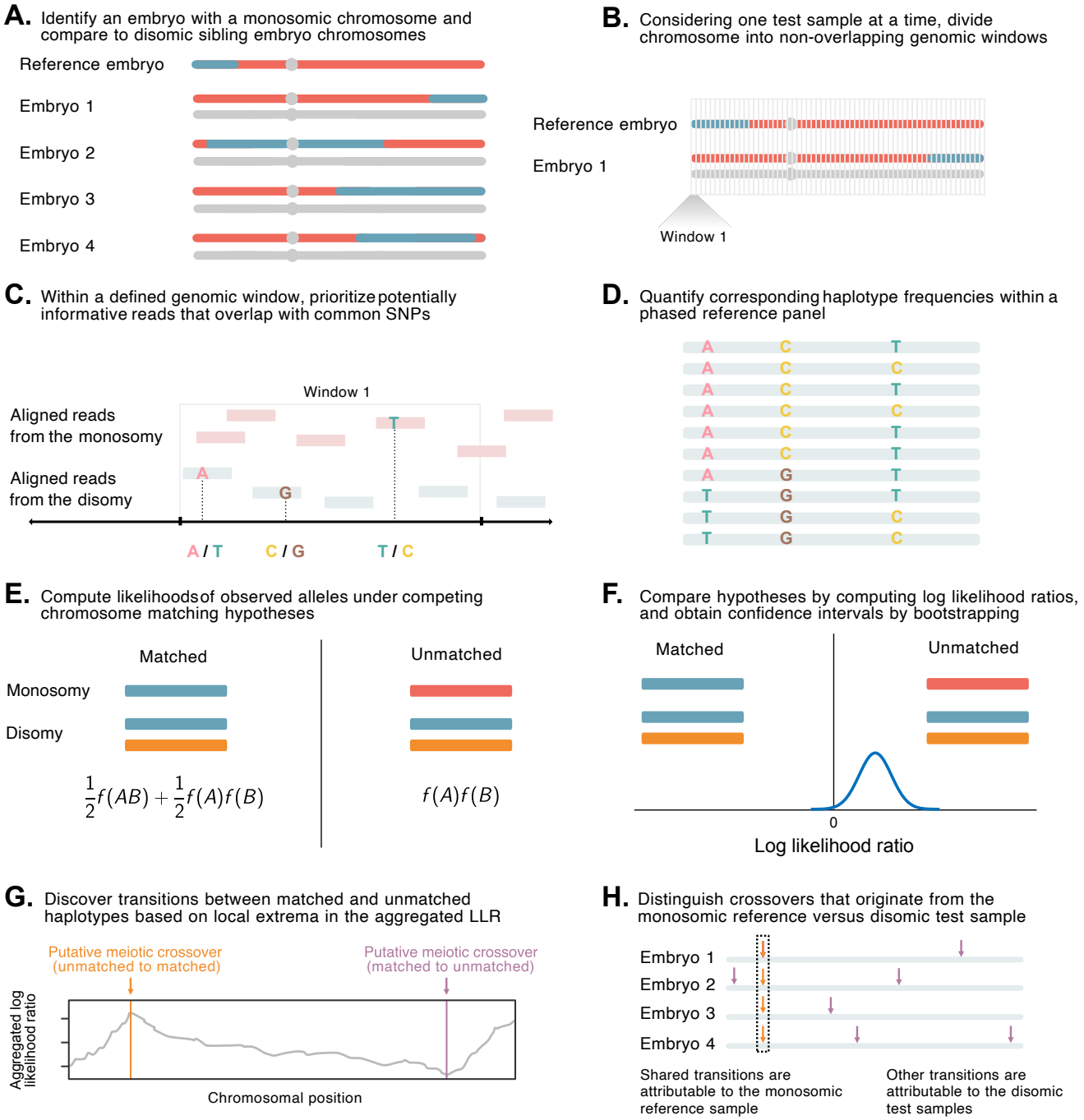

Figure 1. *A statistical approach for meiotic crossover discovery based on low-coverage sequencing data from preimplantation genetic testing.* **A.** Crossover detection is based on haplotype matching between a monosomic chromosome (which is phased by default) and the disomic chromosomes of sibling embryos from the same IVF case. **B.** Analysis is conducted within non-overlapping genomic windows on the scale of 10 - 100 kbp, defined by the length of typical human haplotypes. **C.** Within each window, 2-18 reads are re-sampled, prioritizing potentially informative reads that overlap common polymorphisms in the population. **D.** Frequencies and joint frequencies (i.e., haplotype frequencies) of these SNPs are quantified within an external phased genetic reference panel. **E.** Based on these frequencies, the likelihoods of the observed reads are computed under both the matched and unmatched haplotype hypotheses. **F.** The hypotheses are compared by computing a likelihood ratio, with variance estimated by bootstrapping. **G.** Local extrema in the aggregated log likelihood ratio indicate the locations of meiotic crossovers. **H.** Putative crossovers observed in the majority of sibling embryos can be attributed to the monosomic reference chromosome, while the remaining crossovers are attributed to the test samples.

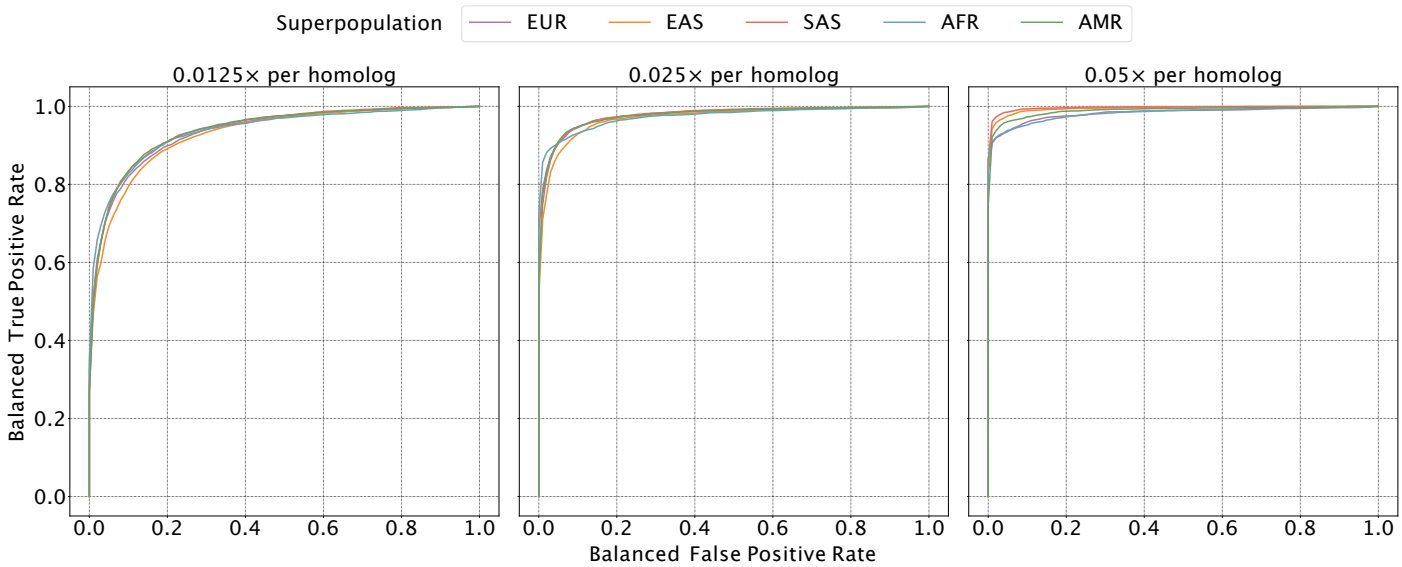

Figure 2. *Evaluating the sensitivity and specificity of meiotic crossover detection based on simulation.* Using data from the 1000 Genomes Project, we simulated pairs of Monosomy 16 and Disomy 16, where half of the pairs possessed matched haplotypes and the other half possessed unmatched haplotypes. Then we divided chromosome 16 into 45 bins (of $\sim$ 2 Mbp) and calculated a balanced ROC curve (see Methods) for each bin, averaging over all the balanced ROC curves to obtain a mean balanced ROC curve. We repeated this procedure over a range of depths of coverage and across sets of samples from all superpopulations of the 1000 Genomes Project.

homologous chromosomes, which we formalize using a likelihood framework (see Methods). Transitions between these matched and unmatched states indicate the locations of meiotic crossovers.

Evaluating method performance via simulation

In order to assess the performance of LD-CHASE, we simulated chromosomes from pairs of embryos consisting of a monosomy (i.e., reference sample) and disomy (i.e., test sample) that either shared a matching haplotype or were unrelated. A meiotic monosomy can occur due to errors at several distinct stages of oogenesis, as appears in Figure S2. We generated these pairs by mixing phased chromosomes from the 1000 Genomes Project [31], as described in the Methods. We focused our simulations on chromosome 16, which is the chromosome most frequently affected by aneuploidy in preimplantation embryos, using a bin size of 2 Mbp and varying the sample ancestries across all superpopulations from the 1000 Genomes Project.

We allowed our classifier to assign bins as "matched", "unmatched", or "ambiguous" to denote uncertainty, and we used a balanced ROC curve to evaluate performance (see Methods; Figures 2 and S3). Our results demonstrated high sensitivity and specificity across all ancestries at a coverage of 0.05× per homolog (average area under the curve [AUC] of 0.989). As coverage was reduced to 0.025× and 0.013×, the AUC decreased by 0.014 and 0.053 on average, respectively, though performance was more or less affected in certain regions of the genome (Figure 2). Notably, performance of the classifier is affected by the local density of SNPs and sequencing reads, as well as ancestry matching between the reference panel and the target samples (Figures S3 and S4).

Application to a large PGT-A dataset

Encouraged by the performance of LD-CHASE on simulated data, we proceeded to apply it to a large dataset from the CReATe Fertility Centre (Toronto, Canada). The dataset consists of low-coverage sequencing data from 18,967 embryos from 2558 IVF patients, collected between April 2020 and August 2022. In order to select appropriate ancestry-matched reference panels, we first inferred the genetic similarity of each embryo to reference samples from the 1000 Genomes Project [31] using LASER [32] (see Methods). The results reflect the diverse ancestry composition of the patient population, with 68.71% (13,104) of embryos exhibiting the greatest genetic similarity to European reference samples, 6.76% (1290) of embryos exhibiting the greatest genetic similarity to South Asian reference samples, 5.74% (1094) of embryos exhibiting the greatest genetic similarity to East Asian reference samples, and the remaining embryos exhibiting lower genetic similarity to reference samples, for example due to recent admixture (Figure S5).

Previous studies have demonstrated that the vast majority of monosomies observed in blastocyst-stage embryos with PGT-A are of maternal meiotic origin, such that only the paternal chromosome remains [33, 21]. LD-CHASE uses such

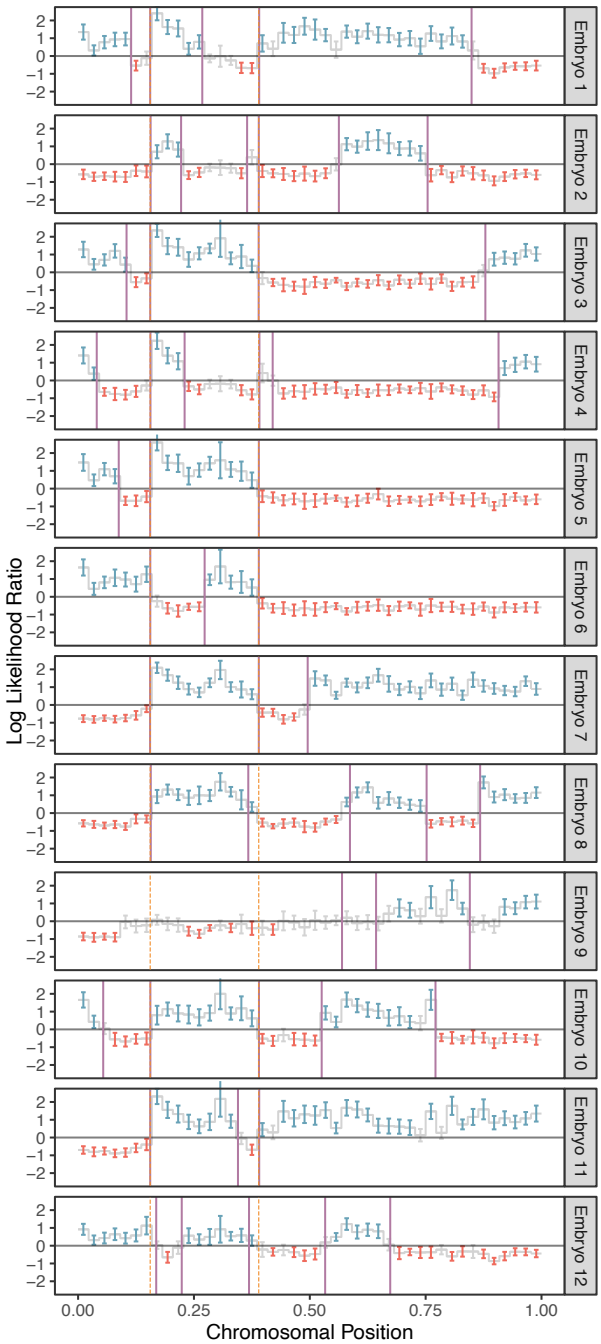

Figure 3. *A representative example of crossover discovery based on haplotype matching among sibling IVF embryos.* In this example, chromosome 10 of each test embryo is compared to a single sibling reference embryo with monosomy of chromosome 10. Evidence of haplotype non-matching is indicated by positive log-likelihood ratios, while evidence of haplotype matching is indicated by negative log-likelihood ratios. Crossovers are identified as transitions from positive (blue) to negative (red) log likelihood ratios or vice-versa. Error bars denote 95% confidence intervals. Crossovers attributed to the monosomic chromosome are indicated with dashed orange lines (defined as those observed in more than half of test samples), while crossovers attributed to the test samples themselves are indicated with purple lines.

monosomies to map paternal crossovers in sibling disomic embryos. Meanwhile, haploidy or genome-wide uniparental isodisomy (GW-isoUPD) observed at the blastocyst stage nearly exclusively involves the sole presence of the maternal genome [33, 21, 34], allowing us to map genome-wide maternal crossovers in sibling disomic embryos. LD-CHASE thus requires preliminary analysis to identify chromosome abnormalities based on signatures of altered depth of coverage and/or genotype observations.

To this end, copy number of each autosome of each sample was inferred using WisecondorX [35] based on within-sample normalized depth of coverage. Across the entire dataset, we identified 388,366 disomies, 3307 trisomies, 4294 monosomies, 332 segmental gains, and 685 segmental losses (see Figure S6 as well as Table S1). The monosomic chromosomes traced to embryos obtained from 1506 (58.85%) unique patients, facilitating mapping of paternal crossovers among 30,645 disomic chromosomes of 12,348 total embryos. Because maternal meiotic monosomies observed in blastocyst-stage embryos are highly enriched for chromosomes 15, 16, 21, and 22, the mapping of paternal crossovers was largely relegated to these chromosomes, with much lower resolution for the rest of the genome.

Importantly, methods such as WisecondorX compare coverage across chromosomes within a sample and may therefore fail to detect aneuploidies that simultaneously affect many chromosomes. In extreme cases such as triploidy and haploidy/GW-isoUPD, where coverage is uniform across the genome despite the ploidy aberration, embryos may be erroneously classified as euploid. To overcome this limitation, we applied our published haplotype-aware method, LD-PGTA [28], to reclassify all chromosomes that were initially identified as disomic by WisecondorX. LD-PGTA identified 155 (1.65%) samples as triploid and 395 (4.20%) samples as haploid/GW-isoUPD. Importantly, such haploid/GW-isoUPD embryos were distributed across 184 (7.19%) patients, facilitating mapping of genome-wide maternal crossovers among 40,015 disomic chromosomes of 1898 total embryos.

Sex-specific maps of crossovers on disomic chromosomes

Considering only chromosomes with informative genomic windows that covered at least 50% of their total length (see Methods), we identified 54,503 maternal crossovers across 27,026 chromosomes and 27,982 paternal crossovers across 20,936 chromosomes. The rate of maternal crossovers per chromosome exhibited a positive, but non-significant association with maternal age ($\hat{\beta}$ = 0.003, SE = 0.010, p-value = 0.756). An example of crossovers mapped in a single set of sibling embryos is provided in Figure 3, where transitions from intervals that do not match (blue) and match (red) the reference monosomic

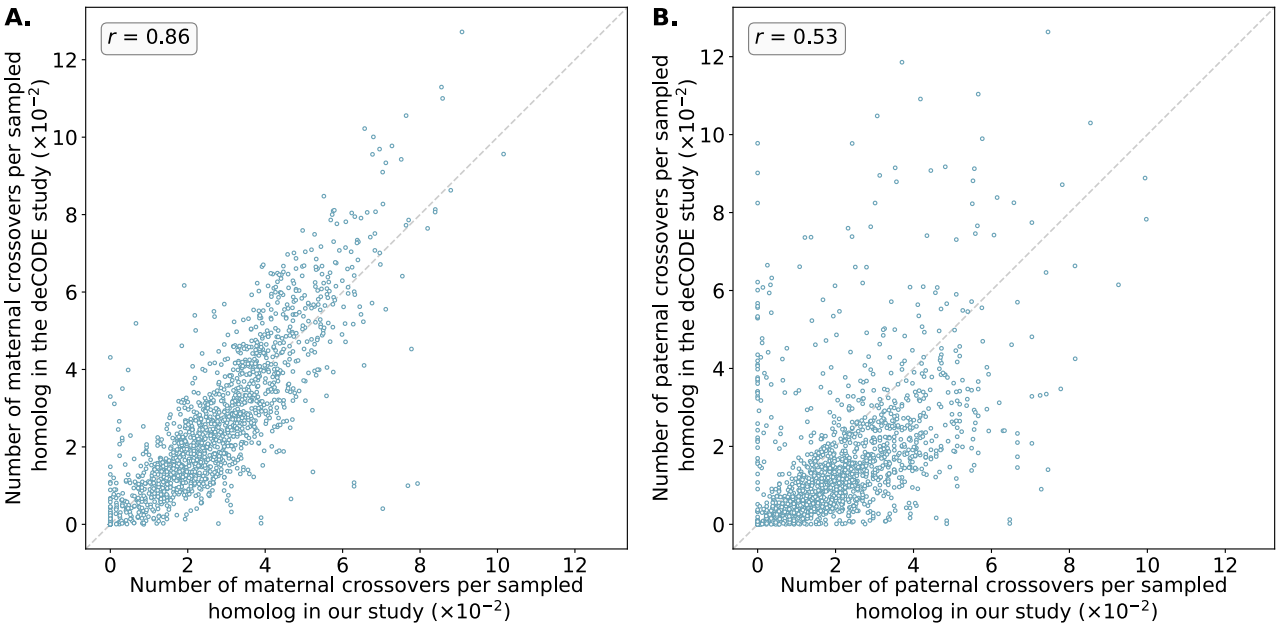

Figure 4. *Number of maternal (*A*) and paternal (*B*) crossovers per sampled homolog in genomic bins compared between our study and published genetic maps from deCODE [25].* Crossovers were identified as transitions between regions of "matched" and "unmatched" haplotypes, where each region included at least 15 genomic windows and a z-score of the at least 1.96. Pearson correlation coefficients (r) between the deCODE map and our map across genomic bins are reported for each panel (See Figures S9 and S10 for chromosome-specific comparison).

chromosome indicate the locations of meiotic crossovers (purple lines; see Methods). The exception to this interpretation involves transitions that are shared across all (or nearly all) sibling embryos, which instead reflect crossovers attributable the reference monosomic chromosome itself (orange dashed lines). The genome-wide distributions of crossovers on disomic chromosomes are provided in Figures S7 and S8.

The crossover distributions were strongly correlated with sex-specific genetic maps published by deCODE, which were based on whole-genome sequencing of living parent-offspring trios [25], in broad support of the accuracy of our method (Figure 4). The observed correlation was particularly strong for putative maternal crossovers ($r = 0.86$) compared to putative paternal crossovers ($r = 0.53$), though the latter correlation increased modestly ($r = 0.71$) when restricting to chromosomes 15, 16, 21 and 22, which are enriched for maternal meiotic aneuploidies and thus offer the greatest resolution for mapping paternal crossovers (see Figure S9 and Figure S10 for chromosome-specific correlations). Notably, the correlations substantially declined when comparing our inferred female map to the deCODE male map or vice versa, supporting our assumptions about the parental origins of various forms of chromosome abnormalities (Figures S11 and S12).

Chromosome-specific propensities for various mechanisms of trisomy formation

Previous studies have suggested that chromosomes may vary in their susceptibility to segregation errors occurring during meiosis I (MI), meiosis II (MII), and mitosis [12, 13, 14, 36, 37, 38]. MI and MII errors can be roughly identified based on tracts of distinct (i.e., "both parental homologs" or BPH) or identical (i.e., "single parental homolog" or SPH) haplotypes, respectively, inherited from a single parent in regions spanning the centromere. Meanwhile, patterns of SPH chromosome-wide indicate a potential mitotic origin of trisomy (or MII error without recombination). While previous studies have noted that the attribution of centeromere-spanning BPH and SPH patterns to MI and MII errors is imperfect due to alternative mechanisms by which the signatures may originate [39], our results support the hypothesis that chromosomes possess unique propensities for various forms of segregation error (χ^2 [2, N = 2540] = 489.2, P = 1.6×10^{-77}; Figure 5; Table S2). The vast majority (> 81%) of trisomies of chromosomes 15, 16, 19, 21, and 22 exhibited haplotypic patterns consistent with errors in MI, while chromosomes 11, 13, and 14 exhibited more modest excesses (∼70%) of MI errors (binomial test, Bonferroni adjusted-P < 0.05 for all noted chromosomes). Meanwhile, the remainder of chromosomes were characterized by a roughly equal number of MI and MII errors (36-61%; binomial test, Bonferroni adjusted-P > 0.05 for all noted chromosomes).

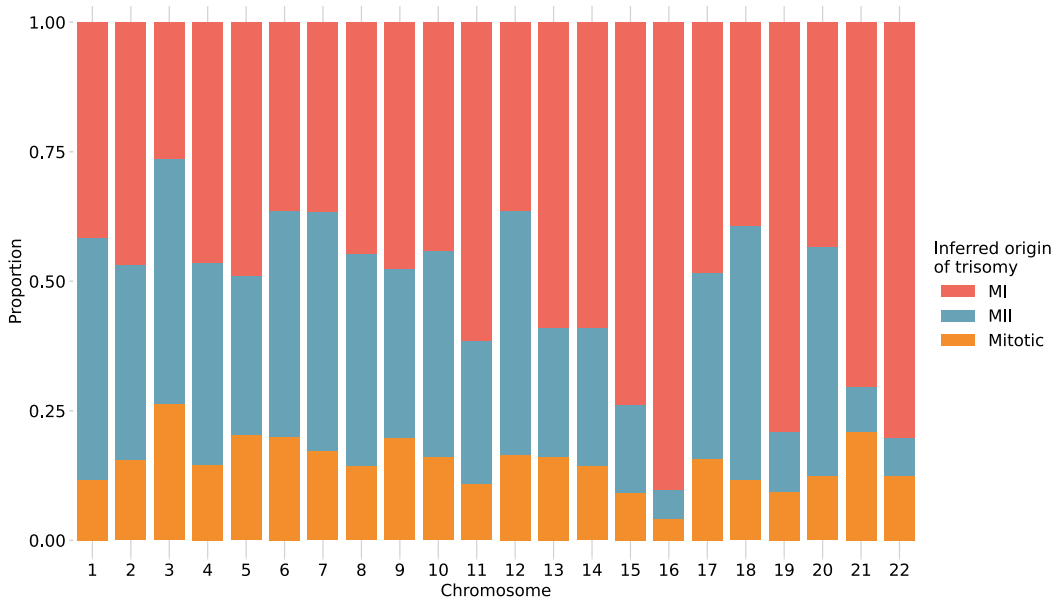

Figure 5. *Stratification of trisomies by autosome and inferred source of error.* Chromosome-wide patterns of SPH were designated as potentially mitotic in origin (yellow). Samples with tracts of BPH in regions surrounding the centromere were classified as putative MI errors (red), while tracts of BPH elsewhere on the chromosome were classified as putative MII errors (blue).

An altered landscape of crossovers among aneuploid versus disomic chromosomes

Abnormal number or location of meiotic crossovers between homologous chromosomes may predispose oocytes to aneuploidy, as demonstrated by several previous studies [reviewed in 40]. Our published method, LD-PGTA [28], facilitates the mapping of crossovers on trisomic chromosomes, which can then be compared to the crossover map for disomic chromosomes obtained via LD-CHASE. One caveat of this comparison is that LD-PGTA and LD-CHASE possess different sensitivities and specificities, which also vary along the genome, as evident from our simulation-based benchmarking analyses (Figures S3 and S4 and Figures S13 and S14). To ensure that observed differences between the crossover distributions were not driven by these technical differences, we merged together monosomies and disomies to create artificial trisomies (see Methods) such that the disomies and trisomies could both be analyzed with LD-PGTA in a standardized manner. The number ($r = 0.99$) and genome-wide distribution ($r = 0.90$) of disomic crossovers inferred by LD-CHASE and LD-PGTA exhibited strong agreement, suggesting that the methods are robust to their technical differences and supporting the use of LD-CHASE in downstream comparisons between trisomies and disomies (see Figure S15).

Across all chromosomes, these comparisons revealed a 35% depletion of crossovers for trisomies relative to disomies (Figure 6). On a per-chromosome basis, the depletion was observed across all chromosomes, but was largest for chromosome 16 (54%) and smallest for chromosome 4 (17%). While these observations are consistent with the hypothesis that a reduced rate (or absence; i.e., "exchangeless homologs") of recombination contributes to aneuploidy risk [41, 40], we note that they may be partially driven by a failure of our method to detect crossovers on reciprocal recombinant chromosomes—a limitation that uniquely applies to trisomies but not disomies and affects nearly all previous genotype-based studies (see Discussion).

In addition to these global differences in numbers of crossovers, several chromosome-specific alterations in the landscapes of crossovers were evident from our results (Figure 6 as well as Figures S16 to S19). These include the observation that on chromosome 16, trisomies lacked a pair of hotspots in the vicinity of the centromere—one on each arm—broadly consistent with the previous observation that distal crossovers were enriched among a smaller sample of 62 cases of Trisomy 16 [13, 40]. Meanwhile, trisomies of chromosome 13 were depleted of crossovers near the center of the q-arm compared to disomies, whereas trisomies of chromosome 22 appeared relatively enriched for crossovers near the center of the q-arm.

Discussion

Crossovers are established during meiotic prophase I and are necessary for ensuring accurate pairing and subsequent segregation of chromosomes following decades-long dictyate arrest in human females [42]. Previous studies have demonstrated chromosome-specific associations between

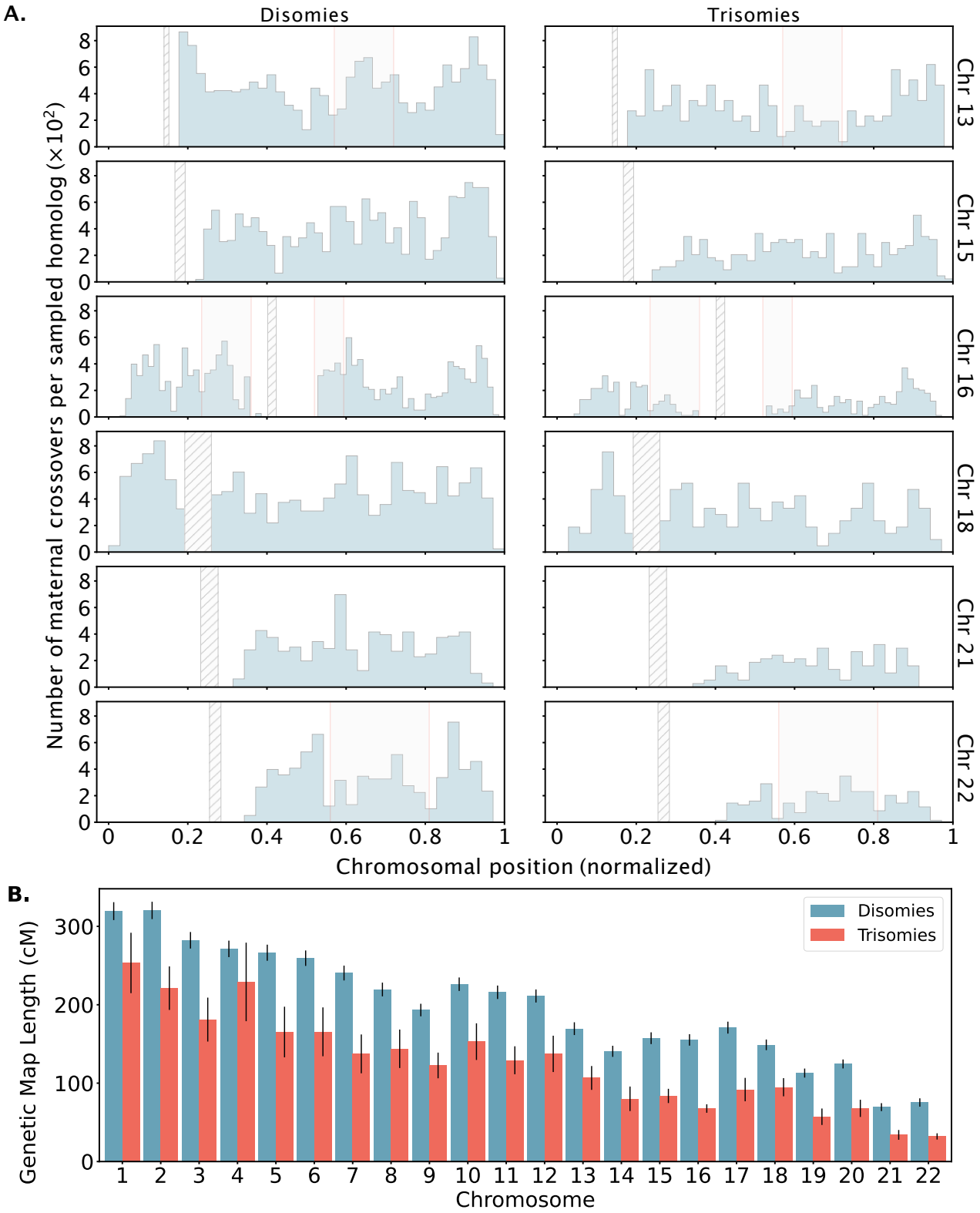

Figure 6. *Differences in number and location of inferred crossovers between disomic and trisomic samples.* **A.** Spatial distributions of inferred meiotic crossovers in disomies (left; mapped with LD-CHASE) and trisomies (right; mapped with LD-PGTA) across chromosomes 13, 15, 16, 18, 21, and 22. Regions with qualitative differences are highlighted in light gray, while centromeres are indicated with diagonal shading. **B.** Comparisons of female genetic map length for disomies (blue) versus trisomies (red) for each autosome. Observed rates of meiotic crossover were consistently lower for trisomies compared to disomies. Error bars denote 95% confidence intervals as estimated by bootstrapping.

patterns of recombination and various forms of meiotic aneuploidy [37, 12, 13, 36, 38, 14]. For example, early studies of Trisomy 16 (the most common autosomal trisomy detected in human preimplantation embryos) suggested a depletion of recombination in pericentromeric regions relative to euploid control samples [43, 44]. However, the accuracy, genomic resolution, and statistical power of many such studies have been limited by the genetic assays they employed (e.g., Southern blot, PCR, etc.), as well as the challenge of achieving large samples from living trisomic individuals or products of conception. Due to its inherent relatedness structure, PGT-A offers a natural source of retrospective data for crossover mapping—including both viable and nonviable embryos—but current implementations based on low-coverage whole-genome sequencing pose a technical challenge for recovering relevant genotype information.

To address this challenge, our haplotype-aware method (LD-CHASE) uses known LD structure from an external reference panel as well as the frequent occurrence of monosomies (which are phased by default) to map crossovers based on comparisons of haplotypes among samples of sibling embryos. The resulting sex-specific maps of crossovers generated by our method were broadly consistent with those generated in previous prospective studies. The sex-specific nature of these patterns also supports our assumption that most monosomies observed in blastocyst-stage IVF embryos (of adequate morphology to be candidates for transfer and thus tested with PGT-A) originate during maternal meiosis, whereas most cases of haploidy/GW-isoUPD solely possess a maternally inherited set of chromosomes. The latter phenomenon may arise when sperm cells trigger egg activation but fail to fuse with the ovum, after which the maternal genome may duplicate to produce two identical complements [45]. Our observations about the sex-specific origins of various forms of aneuploidy thus independently replicate previous studies that directly assayed parental genomes [33, 21]. More broadly, the ability to generate sex-specific crossover maps from PGT-A data may facilitate future studies to understand the quantitative genetic basis of recombination phenotypes, as well as more detailed understanding of how extreme deviations in these phenotypes contribute to meiotic errors and potential infertility.

As a first step toward this goal, we investigated the association between crossover phenotypes and chromosome abnormalities observed in preimplantation embryos. While broadly consistent with previous studies that used smaller samples or were restricted to living individuals with viable trisomies, our data provide unified views of these phenomena across all autosomes and support the hypothesis that the number and chromosome-specific locations of meiotic crossovers influence risk of aneuploidy. However, the observation that the length of the genetic map is shorter for trisomies versus disomies could be partially driven by the inability to detect crossovers on reciprocal chromosomes that derive from a single crossover event and were both transmitted to the oocyte to produce a trisomy. Future studies may employ linked-read or long-read sequencing to achieve direct read-based phasing and overcome this limitation [46, 47].

Despite the methodological advances we report here, crossover mapping from sequencing-based PGT-A data possesses several other technical limitations, including the modest resolution per sample ($\sim$150 kbp). This limitation is driven by the combination of the low coverage (and thus sparsity) of the aligned reads ($<0.05\times$), the low rates of heterozygosity of human genomes (<0.001), as well as the extent of LD in the reference panel (<1 Mb). Our benchmarking analyses additionally demonstrate that performance degrades with genetic distance between the test sample and reference panel due to differences in allele frequencies and LD structure, similar to challenges encountered when transferring polygenic scores across populations [48]. To overcome these sources of error, we aggregated signal across consecutive genomic windows, thereby increasing the classification accuracy at the cost of resolution. Even with such an approach, it is important to note that performance is not uniform across the genome, as regions near centromeres, telomeres, or other repetitive regions are enriched for false positives and negatives relative to the genomic background.

One promising future direction is the extension of our haplotype-aware approach to PGT-M, combining knowledge of population genetic patterns of LD in a reference panel with information from sibling embryos (or alternative data sources such as gametes) to infer transmission or non-transmission of pathogenic haplotypes from parents to offspring. While potentially lowering costs and increasing efficiency, such an approach will require extensive validation and benchmarking to determine its feasibility and accuracy given the probabilistic nature of LD and the high stakes of PGT-M. In the meantime, LD-CHASE offers a flexible tool for mapping crossovers in low-coverage sequencing data from multiple sibling embryos, toward a better understanding of the factors that modulate the meiotic crossover landscape and the role of recombination in the origins of aneuploidies.

Methods

Prioritizing informative reads

Our method seeks to overcome the sparse nature of low-coverage sequencing data by leveraging LD structure of an ancestry-matched reference panel, consisting phased haplotypes from high-coverage sequencing data. Measurements of LD

require pairwise and higher order comparisons and may thus grow intractable when applied to large genomic regions. To ensure computational efficiency, we developed a scoring algorithm to prioritize reads based on their potential information content, as determined by measuring haplotype diversity within a reference panel at sites that they overlap. We emphasize that the priority score of a read only depends on variation within the reference panel and not on the alleles that the read possesses. The score of a read is calculated as follows:

1. Given the the ancestry composition of the target sample (e.g., 30% ancestry with genetic similarity to European reference samples and 70% ancestry with genetic similarity to East Asian reference samples), a suitable reference panel is chosen (see subsequent section titled "Assembling an ancestry-matched reference panel").
2. Based on this reference panel, we list all biallelic SNPs that overlap with the read and their reference and alternative alleles.
3. Using the former list, we enumerate all the possible haplotypes. In a region that contains n biallelic SNPs, there are 2^n possible haplotypes.
4. The effective frequency of each haplotype is estimated from the reference panel:

$$f_{\text{eff}}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \sum_i \alpha_i f_i(\mathbf{A}, \mathbf{B}, \mathbf{C}), \tag{1}$$

   where α_i is the ancestry proportion from the ith population, and $\sum_i \alpha_i = 1$ (e.g., $\alpha_1 = 0.3$ and $\alpha_2 = 0.7$). Moreover, $f_i(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is the joint frequency of the SNP alleles $\mathbf{A}$,$\mathbf{B}$ and $\mathbf{C}$ in the ith population. Here we assumed $n = 3$, but the formula is applicable to any n.
5. We increment the priority score of a read by one for every haplotype with a frequency between f_0 and $1 - f_0$.

An example of scoring a read that overlaps with three SNPs appears in the Supplementary Methods. Our scoring metric is based on the principle that reads that overlap SNPs with intermediate allele frequencies should receive high priority, as the inclusion of such sites will increase our ability to discern between two haplotypes. In the simplest case, where a read overlaps with only a single SNP, the score of the read would be two when the minor allele frequency (MAF) is at least f_0 and otherwise zero. We note that all observed alleles from the same read are considered as originating from the same underlying molecule. Hence, our score metric reflects the number of common haplotypes existing in the population in the chromosomal region that overlaps with the read. For a reference panel on the scale of the 1000 Genomes Project (∼2500 unrelated individuals), 25% - 45% of common SNPs have a nearest neighbor within 35 bp. Hence, even for short reads, it is beneficial to use a scoring metric that accounts for reads that span multiple SNPs.

Comparing haplotype matching hypotheses

By virtue of LD, observations of a set of alleles from one read may provide information about the probabilities of allelic states in another read that originated from the same DNA molecule (i.e., chromosome). In contrast, when comparing reads originating from distinct homologous chromosomes, allelic states observed in one read will be uninformative of allelic states observed in the other read. As two siblings are characterized by chromosomal regions with different counts matched haplotypes, a change in the count along the chromosome indicates the position of a meiotic crossover, i.e., an exchange of DNA segments between non-sister chromatids during prophase I of meiosis. Similarly, a pair of half-siblings allows us to contrast the crossovers in either the paternal or the maternal homologs. Extending this logic, we note that sequences from either a monosomic sibling-embryo, a sibling embryo with GW-isoUPD or individual parental gametes would similarly allow us to isolate crossovers in embryo genomes that arose during oogensis or spermatogenesis.

We consider two sibling embryos, one has a monosomy and the second is a healthy diploid, consisting two copies of the genome—a maternal and a paternal copy. For a set of reads aligned to a defined genomic region, we compare the likelihoods of the observed alleles under two competing hypotheses:

1. The monosomic embryo and the healthy diploid have matched haplotypes, denoted as matched haplotypes hypothesis.
2. The monosomic embryo and the healthy diploid have unmatched haplotypes, denoted as unmatched haplotypes hypothesis.

A transition between regions with matched and unmatched haplotypes indicates the location of a meiotic crossover.

Our statistical models consider a situation where one read is drawn from the monosomic chromosome and the second from the disomic chromosome of a sibling. The odds of two reads being drawn from identical versus distinct haplotypes differ under the matched and unmatched haplotypes hypotheses. Specifically, for the matched haplotypes hypothesis, the odds are 1 : 1, and for the unmatched haplotypes hypothesis, the odds are 0 : 1. If a pair of reads is drawn from identical haplotypes, the probability of observing the two alleles is given by the joint frequency of these two alleles (i.e., the frequency of the haplotype that they define) in the reference panel. In contrast, if a pair of reads is drawn from distinct haplotypes, then the probability of observing the two alleles is simply the product of the frequencies of the two

alleles in the reference panel:

$$P_{\text{matched}}(A \wedge B) = \frac{1}{2} f(A, B) + \frac{1}{2} f(A) f(B), \tag{2}$$

$$P_{\text{unmatched}}(A \wedge B) = f(A) f(B). \tag{3}$$

where $f(A)$ and $f(B)$ are the frequencies of alleles A and B in the population.

These statistical models can then be generalized to arbitrary admixture scenarios by a simple substitution. We assume that each distinct parental haplotype is drawn from an ancestral population with a probability equal to the ancestry proportion of the tested individual that is associated with that population. In accordance with the assumption, we replace each allele frequency distribution, f by the combination $\alpha_1 f_1 + \alpha_2 f_2$. Here α_i is the probability that the alleles originated from the i-th population, f_i is the allele frequency distribution for the i-th population and $\alpha_1 + \alpha_2$. For example, under this substitution $P_{\text{unmatched}}(A \wedge B) = \alpha_1 f_1(A, B) + (1 - \alpha_2) f_2(A, B)$.

Likelihoods of the two hypotheses are compared by computing a log-likelihood ratio:

$$\gamma(A, B) = \log \frac{P_{\text{unmatched}}(A \wedge B)}{P_{\text{matched}}(A \wedge B)}. \tag{4}$$

When a read overlaps with multiple SNPs, $f(A)$ should be interpreted as the joint frequency of all SNP alleles that occur in read A (i.e., the frequency of haplotype A). Similarly, $f(A, B)$ would denote the joint frequency of all SNP alleles occurring in reads A and B. The equations above were extended to consider up 6 reads per window and homolog, as described in the later subsection titled "Generalization to arbitrary number of reads". Estimates of allele and haplotype frequencies from a reference panel do not depend on theoretical assumptions, but rely on the idea that the sample is randomly drawn from a population with similar ancestry. One limitation, which we consider, is that reliable estimates of probabilities near zero or one require large reference panels, such as the 1000 Genomes Project [31].

Determining optimal size of a genomic window

Because pairwise LD in human genomes decays on average to a quarter of its maximal value over physical distances of 100 kbp [31], the length of the chromosomes is divided into genomic windows on a scale consistent with the length of typical human haplotypes (10^4 - 10^5 bp). While one library consists of two homologs and the second consists of a single homolog, we would like to sample an even number of reads from each homolog. Thus, for each DNA sequence, we only consider the depth of coverage per homolog (i.e., we divide the coverage by the ploidy for the chromosome of each sample under consideration).

We require a minimal number of reads per genomic window and homolog, as determined by the sample with the lowest average depth of coverage. We then scan the chromosome in a sliding window, using a window size that adjusts according to the local depth of coverage of the two different sequenced samples. This adaptive sliding window approach possesses advantages over a fixed length window in that it (a) accounts for GC-poor and GC-rich regions of a genome, which tend to be sequenced at lower depths of coverage using Illumina platforms [49] and (b) accounts for varying densities of SNPs across the genome [50].

The algorithm simultaneously scans aligned reads from the two samples in the forward direction of the reference genome and identifies informative reads (i.e., reads that reach the priority score threshold) from both samples within genomic windows. For each genomic window, the minimal number of required reads from the DNA sequence of the disomy is twice the minimal number from the monosomy. If (a) the distance between consecutive reads in one of the samples exceeds 100 kbp or (b) a genomic window extends to 350 kbp and does not meet the minimal number of reads per homolog, the window is dismissed.

Quantifying uncertainty by bootstrapping

To quantify uncertainty in our likelihood estimates, we performed m out of n bootstrapping by iteratively resampling reads within each window [51]. Resampling was performed without replacement to comply with the assumptions of the statistical models about the odds of drawing two reads from the same haplotype. Thus, in each iteration, only subsets of the available reads can be resampled. Specifically, within each genomic windows, up to 6 reads per homolog with a priority score exceeding a defined threshold are randomly sampled with equal probabilities. The likelihood of the observed combination of SNP alleles under each competing hypothesis is then calculated, and the hypotheses are compared by computing a log-likelihood ratio (LLR). The sample mean and the unbiased sample variance (i.e., with Bessel's correction) of the LLR in each window are calculated by repeating this process using a bootstrapping approach,

$$\bar{\gamma}^{(w)} = \frac{1}{m} \sum_{s \in w} \gamma_s, \tag{5}$$

$$\text{Var}(\gamma^{(w)}) = \frac{1}{m-1} \sum_{s \in w} \left(\gamma_s - \bar{\gamma}^{(w)} \right)^2 \tag{6}$$

where γ_s is the the log-likelihood ratio for s-th subsample of reads from the w-th genomic window and m is the number of subsamples. Because the number of terms in the statistical models grows exponentially, we subsample at most 6 reads per window and homolog. Moreover, accurate estimates of joint frequencies of many alleles requires a very large reference panel. Given the rate of heterozygosity in human populations and the size of the 1000 Genomes Project dataset, 6 reads per homolog is generally sufficient to capture one or more heterozygous SNPs that would inform our comparison of hypotheses.

Aggregating signal across consecutive windows

Even when sequences are generated according to one of the hypotheses, a fraction of genomic windows will emit alleles that do not support that hypothesis and may even provide modest support for an alternative hypothesis. This phenomenon is largely driven by the sparsity of the data, as well as the low rates of heterozygosity in human genomes, which together contribute to random noise. Another possible source of error is a local mismatch between the ancestry of the reference panel and the tested sequence. Moreover, technical errors such as spurious alignment and genotyping could contribute to poor results within certain genomic regions (e.g., near the centromeres). To overcome this noise, we binned LLRs across consecutive genomic windows, thereby reducing biases and increasing the classification accuracy at the cost of a resolution. Specifically, we aggregated the mean LLRs of genomic windows within a larger bin,

$$\Gamma_{\text{bin}} = \sum_{w \in \text{bin}} \bar{\gamma}^{(w)}, \tag{7}$$

where $\bar{\gamma}^{(w)}$ is the mean of the LLRs associated with the w-th genomic window. In addition, using the Bienaymé formula, we calculated the variance of the aggregated LLRs,

$$\text{Var}(\Gamma_{\text{bin}}) = \sum_{w \in \text{bin}} \text{Var}(\gamma^{(w)}), \tag{8}$$

where $\text{Var}(\gamma^{(w)})$ is the variance of the LLRs associated with the w-th window. For a sufficiently large bin, the confidence interval for the aggregated LLR is $\Gamma_{\text{bin}} \pm z\sqrt{\text{Var}(\Gamma_{\text{bin}})}$, where $z = \Phi^{-1}(1-\alpha)$ is the z-score, Φ is the cumulative distribution function of the standard normal distribution and $C = 100(1-2\alpha)\%$ is the confidence level. The confidence level is chosen based on the desired sensitivity vs. specificity. We normalized the aggregated LLRs by the number of genomic windows that compose each bin, $\bar{\gamma}_{\text{bin}} = \Gamma_{\text{bin}}/g$. Thus, the variance of the mean LLR per window is $\text{Var}(\bar{\gamma}_{\text{bin}}) = \text{Var}(\Gamma_{\text{bin}})/g^2$. These normalized quantities can be compared across different regions of the genome, as long as the size of the genomic window is the same on average.

Simulating parental haplotypes and diploid offspring

Using a generative model, we simulated trios of a (a) parental chromosome, (b) diploid offspring with haplotypes matching the parental chromosome and (c) an unrelated chromosome. This allows us to evaluate the classifier performance in each genomic window along the human genome. To this end, we constructed synthetic samples comprising combinations of phased haplotypes from the 1000 Genomes Project [31]. These phased haplotypes are extracted from variant call sets to effectively form a pool of haploid sequences.

We first consider non-admixed offspring by drawing 3 effective haploid sequences from the same superpopulation. The first two haploid sequences are used to simulate the diploid offspring, while the first and third sequences are used to simulate the parental chromosome and the unrelated chromosome, respectively. We then simulate reads by selecting a random position along the chromosome from a uniform distribution to represent the midpoint of an aligned read with a given length. Based on the selected position, one out of the three haplotypes is drawn from a discrete distribution,

$$f(h,x) = \begin{cases} p_1(x) & , h = 1 \\ p_2(x) & , h = 2 \\ p_3(x) & , h = 3 \end{cases}, \tag{9}$$

where, in general, the probability of haplotype h depends on the position of the read, x. When simulating a non-admixed diploid offspring, the first haplotype is just as likely as the second haplotype ($p_2 = p_1$), and the third haplotype is absent ($p_3 = 0$). Similarly, for the parental chromosome, $p_1 = 1$ and $p_2 = p_3 = 0$, while for the unrelated chromosome, $p_3 = 1$ and $p_1 = p_2 = 0$. Then, from the selected haplotype, h, a segment of length l that is centered at the selected chromosomal position, x, is added to simulated data, mimicking the process of short-read sequencing. This process of simulating sequencing data is repeated until the desired depth of coverage is attained.

In order simulate an offspring descended from parents from distinct superpopulations (hereafter termed "recent admixed ancestry"), we draw two effective haploid sequences from different superpopulations of the 1000 Genomes Project. A third haploid sequence is then drawn from one of the two former pools to simulate the unrelated chromosome. Finally, we use the generative model with these three effective haploid sequences to simulate reads. A procedure for simulating diploid offspring under a scenario involving more distant admixture is discussed in the Supplementary Methods.

Evaluating model performance on simulated data

We developed a classification scheme to determine whether a bin supports one of two competing hypotheses. To this end, we performed m out of n bootstrapping by iteratively resampling reads within each window pairs and computed log likelihood ratios (LLRs) of competing statistical models, as described in the subsection titled "Quantifying uncertainty by bootstrapping".

The confidence interval for the mean LLR is $\bar{\gamma}_{\text{bin}} \pm z\sqrt{\text{Var}(\bar{\gamma}_{\text{bin}})}$, and z is referred to as the z-score. Thus, we classify a bin as exhibiting support for the matched haplotypes hypothesis when

$$\bar{\gamma}_{\text{bin}} - z\sqrt{\text{Var}(\bar{\gamma}_{\text{bin}})} > 0, \tag{10}$$

and for the unmatched haplotypes hypothesis when

$$\bar{\gamma}_{\text{bin}} + z\sqrt{\text{Var}(\bar{\gamma}_{\text{bin}})} < 0, \tag{11}$$

where the first (second) criterion is equivalent to requiring that the bounds of the confidence interval lie on the positive (negative) side of the number line. When a confidence interval crosses the origin of the number line, we classify the bin as ambiguous (see Figure S20 for a diagram of these classes).

For a given depth of coverage and read length, we simulate an equal number of sequences generated according to both hypotheses, as explained in the previous subsection. We define true positives (negatives) as simulations where sequences generated under the "unmatched" ("matched") haplotypes hypothesis are correctly classified. Based on these simulations, we generate balanced receiver operating characteristic (ROC) curves for each bin [28]. The balanced true and false positive rates for a bin are defined as

$$\mathrm{BTPR} = \frac{1}{2}\left(\mathrm{TPR} + \mathrm{TNR}\right), \tag{12}$$

$$\mathrm{BFPR} = \frac{1}{2}\left(\mathrm{FPR} + \mathrm{FNR}\right), \tag{13}$$

where TPR, TNR, FPR and FNR the true positive rate, true negative rate, false positive rate and false negative rate, respectively.

The "balanced ROC curve" is tailored for trinomial classification tasks. Here the three possible classes are "unmatched", "matched" and "ambiguous". The ambiguous class contains all instances that do not fulfill the criteria in Equations (10) and (11) (i.e., instances where the boundaries of the confidence interval span zero). This classification scheme allows us to optimize the classification of both "unmatched" and "matched" instances at the expense of leaving ambiguous instances. The advantage of this optimization is a reduction in the rate of spurious classification. To generate each curve, we varied the confidence level and the number of bins.

Generalization to an arbitrary number of reads

The model of the unmatched haplotypes hypothesis for non-admixed samples and an arbitrary number of reads is based on the disomy model for non-admixed samples that was derived for LD-PGTA [28]. More specifically, the "umatched" statistical model for $m+n$ reads is merely the joint frequency $f(A_1, A_2, \ldots, A_m)$ multiplied by the disomy model $P_{\mathrm{disomy}}(B_1 \wedge B_2 \wedge \ldots, B_n)$ of n-reads for non-admixed samples. Here A_i are reads drawn from the monosomic reference sample, while B_i are reads drawn from the disomic test sample.

The model of the matched haplotypes hypothesis for non-admixed samples and an arbitrary number of reads is based on the disomy model for recent-admixtures, which was previously introduced in Ariad et al. [28]. More specifically, modeling the matched haplotypes hypothesis for $m+n$ reads can be accomplished by substituting an effective joint frequency distributions in the disomy model of n-reads for recent-admixture. The adjusted model involves two distributions $f(X)$ and $g(X)$. The distribution f is derived from a reference panel of a population as before, while g is an effective distribution that is defined as $g(X) \equiv f(A_1 A_2 \ldots A_m, X)$. The reads A_i with $i = 1, 2, \ldots, m$ are drawn only from the monosomic reference sample, while the rest of the reads are drawn from the disomic test sample. Also, each term in the linear model should take into account the presence of the reads $A_1, A_2, \ldots, A_m$ and thus terms in the model that involve only the distribution f should multiplied by $g(\varnothing) \equiv f(A_1, \ldots, A_m)$. Derivations and explicit statistical models for non-admixed ancestry, recent admixture and more distant admixture scenarios for $m+2$, $m+3$ and $m+4$ reads can be found in the Supplementary Methods.

Identifying meiotic crossovers

In order to identify the locations of meiotic crossovers, we analyze the cumulative sums of log-likelihoods ratios (LLRs) from individual genomic windows as we move along the chromosome. Since we performed m out of n bootstrapping by iteratively resampling reads within each window and calculating LLRs, we define two quantities:

$$y_n = \sum_{w=0}^{n} \tilde{\gamma}^{(w)}, \qquad v_n = \sum_{w=0}^{n} \mathrm{Var}(\gamma^{(w)}), \tag{14}$$

where we assume negligible linkage disequilibrium between alleles in different genomic windows, and hence according to the Bienaymé formula $v_n = \mathrm{Var}\left(\sum_{w=0}^{n} \gamma^{(w)}\right)$. Local maxima of y_n indicate potential transitions from "unmatched" to "matched" regions, while local minima of y_n indicate potential transitions from "matched" to "unmatched" regions. Thus, a crossover occurred within the jth genomic window if there exists a region in which either $\arg\max_{i<n<k}(y_n) = j$ or $\arg\min_{i<n<k}(y_n) = j$ and in addition

$$|y_j - y_i| \geq z\sqrt{|v_j - v_i|}, \qquad |j - i| \geq \delta, \tag{15}$$

$$|y_k - y_j| \geq z\sqrt{|v_k - v_j|}, \qquad |k - j| \geq \delta, \tag{16}$$

where z is the threshold for calling a crossover. In addition, we define a metric that describes the confidence in calling the jth crossover:

$$\kappa_j = \min\left(\frac{|y_k - y_j|}{\sqrt{|v_k - v_j|}}, \frac{|y_j - y_i|}{\sqrt{|v_j - v_i|}}\right), \tag{17}$$

which fulfills $\kappa_j \geq z_{\mathrm{score}}$. When, based on the criteria above, two consecutive crossovers are identified as local maxima or minima of the accumulated LLR, y_n, this implies that a crossover was skipped. Failure to detect a crossover occurs when the values of z and/or δ are too restrictive. Given two consecutive maxima in the ith and kth genomic windows, we identify the genomic window with skipped minimum as

$$j = \arg\max_{i<n<k}\left(\frac{y_k - y_n}{\sqrt{v_k - v_n}} - \frac{y_n - y_i}{\sqrt{v_n - v_i}}\right), \tag{18}$$

where this condition means that z_{score} of the "matched" and "unmatched" intervals are maximized simultaneously. In case of two consecutive minima, we replace the function arg max with arg min.

Assigning crossovers to reference versus test samples

Because we scan for crossovers by comparing sequences from two sibling embryos, assigning each crossover to a single embryo requires additional information. When sequencing data from three or more sibling embryos is available, we contrast the crossovers of sibling embryos with one common monosomic sibling (see Figure 1). This, in turn, produces a repeated pattern in each sequence of crossovers that can be attributed to the common sibling (see Figure 3). Once the repeated pattern is identified, we subtract it to recover the crossovers in the rest of the embryos.

More specifically, we consider a set of $n+1$ sibling embryos, where one embryo is used to contrast the crossovers in the rest of the sibling embryos. All crossovers are then combined to form a sorted list, and we scan the list for n sequential crossovers within a region of size l. For each cluster, we calculate the average position of the crossovers. The average position from each cluster is associated with the common sibling. Finally, the union of all the clusters is subtracted from each of the n sequences of crossovers, and remaining crossovers are traced back to the sibling embryos.

The sequencing quality may vary from one embryo to another, and some crossovers might not be identified. Hence, we adjust the algorithm to allow clusters with various sizes. After seeking all clusters of size n, we continue seeking clusters of size $n-1$. This process of seeking smaller clusters is repeated iteratively for cluster sizes greater than $n/2$. Another issue that may arise is that when the regions size, l, is too large, two or more crossovers from the same embryo may overlap. In such cases, we only consider the crossover that is closest to the cluster mean.

Each crossover that is attributed to the common reference embryo is assigned two scores. The first score is the proportion of sibling embryos supporting the crossover: $\lambda_i = k_i/n$, where k_i and n are the size of cluster i and number of contrasted embryos, respectively. The second score is the minimal confidence score in the ith cluster: $\tilde{\kappa}_i = \min(\kappa_i^{\text{sibling } 1}, \kappa_i^{\text{sibling } 2}, \ldots, \kappa_i^{\text{sibling } k})$, where κ was defined in the subsection titled "Identifying chromosomal crossovers".

Isolating sex-specific meiotic crossovers

Here we discuss the possibility of identifying sex-specific crossovers by matching haplotypes between natural occurring monosomies and genome-wide isodisomies (GW-isoUPD) in IVF embryos and disomic sibling embryos from the same IVF cycles. This in turn allows us to leverage the high volume of aneuploidies observed in IVF to obtain sex-specific distributions of crossovers. We also review how maternal crossovers in trisomies can be identified by scanning for transitions between tracts of two versus three unique homologs.

Paternal crossovers (i.e., crossovers originating from spermatogenesis) can be identified by focusing on IVF cycles that yielded at least one embryo with a monosomy of a given chromosome, as well as at least one embryo that is disomic for that chromosome. The vast majority of monosomies are of maternal meiotic origin, such that the remaining chromosome is of paternal origin. The opposite scenario, whereby a haploid ovum is fertilized by an nullisomic sperm cell accounts for less than 10% of autosomal monosomies [5, 52, 53]. Moreover, when intracytoplasmic sperm injection (ICSI) is performed, sperm *in vitro* selection and capacitation may further reduce the rates of paternal-origin monosomies [54, 55, 56, 57]. Conversely, maternal crossovers (i.e., crossovers originating from oogenesis) can be identified by focusing on IVF cycles that yielded at least one case of haploidy/genome-wide isodisomy (which are indistinguishable with sequencing-based PGT-A), as well as at least one sibling that is disomic for one or more chromosomes. Such cases of GW-isoUPD may originate from oocytes that commence cleavage and early embryonic development without fertilization [34].

Aneuploidy (gain or loss of entire chromosomes) is the leading cause of IVF failure. We found that 351 (13.7%), 529 (20.7%), 332 (13.0%), and 470 (18.4%) of IVF patients in the CReATe dataset had at least one monosomy of chromosome 15, 16, 21, or 22, respectively, as well as 7.3 disomic siblings on average (see Table S3 for the rest of the autosomes). Thus, LD-CHASE can use these natural occurring monosomies to identify paternal crossovers along the chromosomes of sibling embryos. Similarly, we found that 201 (9.1%) of the IVF patients in the CReATe dataset had at least one haploid/GW-isoUPD embryo and an additional 7.8 euploid siblings on average (see Table S4), offering good resolution for mapping maternal crossovers genome-wide.

Coverage-based discovery of chromosome abnormalities

We used WisecondorX to deduce the chromosomal copy numbers of each of 20,160 embryos in the CReATe dataset [35, 58]. To this end, we first created four sets of reference samples for read counts: (a) 9202 sequences were obtained from biopsies of 4-6 cells and consist of single reads of 75 bp. (b) 9818 sequences were obtained from biopsies of 4-6 cells and consist of paired-end reads of 75 bp (c) 578 sequences were obtained from biopsies of 4-5 cells and consist of single reads of 75 bp, and (d) 562 sequences were obtained from biopsies of 4-5 cells and consist of paired-end reads of 75 bp. We used all the sequences in a

category (e.g., sequences associated with 4-5 cells and single reads) as reference samples. This approach is effective as long as aneuploid chromosomes are rare and as long as the rate of chromosome loss and chromosome gain are sufficiently similar to balance out one another in large samples. The first assumption is justified based on previous PGT-A studies, which showed that the rate of aneuploidy per chromosome is less than 10% (including monosomy, trisomy and mosaics)[28]. The second assumption is justified by noting that for each chromosome, the rate of trisomy should be similar to the rate of monosomy, as both are mainly caused nondisjunction [5]. To assess the robustness of this approach, we compared the results obtained from WisecondorX with those generated by NxClinical, a diagnostic tool utilized by the CReATe Fertility Centre, for select chromosomes that exhibited copy number variations, as shown in Figure Figure S21. In addition, we applied WisecondorX to a separate published PGT-A dataset from the Zouves Fertility Center consisting 8881 samples. The dataset was previously analyzed using BlueFuse Multi, and we found that the inferred copy numbers were in strong agreement with the WisecondorX. The number of sequences that were analyzed successfully via WisecondorX was 20,114. The number of relevant sequences was further reduced to 18,967 after filtering sequences without a corresponding record in the metadata table or when the genetic ancestry could not be inferred.

Haplotype-aware discovery of ploidy abnormalities

Coverage-based approaches for inferring chromosome copy numbers, such as WisecondorX, are based on relative differences in the depth of coverage across chromosomes within samples. As such approaches assume that the baseline coverage corresponds to disomy, scenarios when many chromosomes are simultaneously affected, such as haploidy and triploidy, violate this assumption and may elude detection. To overcome this limitation, we applied our haplotype-aware method, LD-PGTA, to scan for the number of unique haplotypes along the genome [28] and calculate chromosome-wide log-likelihood ratios (LLR) comparing hypotheses of monosomy, disomy, and various forms of trisomy. When the LLR for at least 15 chromosomes supported a common aneuploidy hypothesis (monosomy or trisomy), the sample was classified as haploid/GW-isoUPD or triploid, respectively.

Assembling an ancestry-matched reference panel

Given the aforementioned importance of the ancestry of the reference panel, we used LASER v2.04 [32, 59] to perform automated ancestry inference for each embryo sample from the low-coverage sequencing data. LASER applies principal components analysis (PCA) to genotypes of reference individuals with known ancestry. It then projects target samples onto the reference PCA space, using a Procrustes analysis to overcome the sparse nature of the data. We excluded markers with less than 0.01× depth of coverage and restricted analysis to the top 32 principal components, performing 5 replicate runs per sample.

Ancestry of each target sample was deduced using a k-nearest neighbors approach. Specifically, we identified the nearest 150 nearest neighbor reference sample to each target sample based on rectilinear distance. We then calculated the superpopulation ancestry composition of the 150 reference samples. When >95% of such samples derived from the same 1000 Genomes Project superpopulation, we employed downstream statistical models designed for non-admixed samples. In cases where two superpopulations were represented in roughly equal proportions (maximal difference of 10%), we employed downstream statistical models designed for recent admixture. For all other samples, we employed downstream statistical models designed for more distant admixture scenarios. For this latter group, superpopulations represented at levels of ≥5% among the nearest neighbors to each target sample were used in the construction of reference panels.

Testing robustness versus LD-PGTA

One aspect of our study involved comparing the landscapes of meiotic crossovers between disomic and trisomic chromosomes, as inferred by LD-CHASE and LD-PGTA, respectively. This comparison poses a statistical challenge, as the two methods possess distinct sensitivities and specificities, which also vary across the genome and as a function of coverage of the respective samples. To validate our findings, we therefore sought to compare these landscapes under a single statistical framework by combining disomic samples with separate monosomic samples to produce artificial trisomies. Such artificial trisomies can be analyzed uniformly with LD-PGTA and compared to true trisomies analyzed with the same method, as well as to the more direct analysis of disomies with LD-CHASE.

In principle, such artificial trisomies can be produced by combining monosomies and disomies from the same IVF patients in ratios of 2 : 1, such that they can be analyzed with LD-PGTA. In practice, naive merging based on genome-wide average depth of coverage yields poor results due to the low complexity of DNA libraries from PGT-A, which results in sample-specific non-uniformity in coverage across the genome. To overcome this challenge, we merely replaced the statistical models of LD-CHASE with those of LD-PGTA, thereby controlling the ratio of reads from each sample on a local (i.e., genomic window) as opposed to genome-wide scale. Thus, the construction of the genomic windows and the approach to sample reads from a genomic window is identical to LD-CHASE. However, the statistical models of LD-PGTA have no prior knowledge about the DNA sample from which a given read is drawn. Figures S16

to S19 displays the concordance in crossover distributions between disomic chromosomes as analyzed by both LD-CHASE and LD-PGTA.

The coverage of genomic windows

To further address the low complexity of DNA libraries prepared from few input cells, we introduced an additional metric that simultaneously captures both the depth of coverage and the complexity of the library. After we tile a chromosome with genomic windows, as described in "Determining optimal size of a genomic window", we calculate the coverage of genomic windows for a given sample as:

$$C = \frac{1}{L} \sum_i l_w, \tag{19}$$

where l_w is the length of the wth genomic window, L is the length of the chromosome and thus $0 < C < 1$. We then restricted our analysis to samples with $C \geq 0.5$ (i.e., genomic windows covering at least half of the chromosome). For the subset of patients with multiple monosomic embryo samples (affecting the same chromosome) from which to choose, we selected the monosomy that yielded the highest value of C in order to maximize resolution for identifying meiotic crossovers.

Availability of data and materials

Software for implementing our method is available at: https://github.com/mccoy-lab/LD-CHASE. Tables with chromosome abnormalities and crossovers detected for each sample in the study are available in the same repository.

Ethics approval and consent to participate

The Homewood Institutional Review Board of Johns Hopkins University determined that this work does not qualify as federally-regulated human subjects research (HIRB00011705).

Funding

Research reported in this publication was supported by National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM133747. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgements

Thank you to the staff of Advanced Research Computing at Hopkins.

References

[1] Lynn A, Ashley T, Hassold T. Variation in human meiotic recombination. Annu Rev Genomics Hum Genet. 2004;5:317-49.

[2] Peñalba JV, Wolf JB. From molecules to populations: appreciating and estimating recombination rate variation. Nature Reviews Genetics. 2020;21(8):476-92.

[3] Webster A, Schuh M. Mechanisms of aneuploidy in human eggs. Trends in cell biology. 2017;27(1):55-68.

[4] Lister LM, Kouznetsova A, Hyslop LA, Kalleas D, Pace SL, Barel JC, et al. Age-related meiotic segregation errors in mammalian oocytes are preceded by depletion of cohesin and Sgo2. Current Biology. 2010;20(17):1511-21.

[5] Hassold T, Hunt P. To err (meiotically) is human: the genesis of human aneuploidy. Nature Reviews Genetics. 2001;2(4):280-91.

[6] Lamb N, Sherman S, Hassold T. Effect of meiotic recombination on the production of aneuploid gametes in humans. Cytogenetic and genome research. 2005;111(3-4):250-5.

[7] Herbert M, Kalleas D, Cooney D, Lamb M, Lister L. Meiosis and maternal aging: insights from aneuploid oocytes and trisomy births. Cold Spring Harbor perspectives in biology. 2015;7(4):a017970.

[8] Zaragoza MV, Jacobs PA, James RS, Rogan P, Sherman S, Hassold T. Nondisjunction of human acrocentric chromosomes: studies of 432 trisomic fetuses and liveborns. Human genetics. 1994;94(4):411-7.

[9] Oliver TR, Feingold E, Yu K, Cheung V, Tinker S, Yadav-Shah M, et al. New insights into human nondisjunction of chromosome 21 in oocytes. PLoS genetics. 2008;4(3):e1000033.

[10] Middlebrooks CD, Mukhopadhyay N, Tinker SW, Allen EG, Bean LJ, Begum F, et al. Evidence for dysregulation of genome-wide recombination in oocytes with nondisjoined chromosomes 21. Human molecular genetics. 2014;23(2):408-17.

[11] Kong A, Barnard J, Gudbjartsson DF, Thorleifsson G, Jonsdottir G, Sigurdardottir S, et al. Recombination rate and reproductive success in humans. Nature genetics. 2004;36(11):1203-6.

[12] Robinson W, Kuchinka B, Bernasconi F, Petersen M, Schulze A, Brøndum-Nielsen K, et al. Maternal meiosis I non-disjunction of chromosome 15: dependence of the maternal age effect on level of recombination. Human molecular genetics. 1998;7(6):1011-9.

[13] Hassold T, Merrill M, Adkins K, Freeman S, Sherman S. Recombination and maternal age-dependent nondisjunction: molecular studies of trisomy 16. American journal of human genetics. 1995;57(4):867.

[14] Hall HE, Surti U, Hoffner L, Shirley S, Feingold E, Hassold T. The origin of trisomy 22: evidence for acrocentric chromosome-specific patterns of nondisjunction. American Journal of Medical Genetics Part A. 2007;143(19):2249-55.

[15] Capalbo A, Bono S, Spizzichino L, Biricik A, Baldi M, Colamaria S, et al. Sequential comprehensive chromosome analysis on polar bodies, blastomeres and trophoblast: insights into female meiotic

errors and chromosomal segregation in the preimplantation window of embryo development. Human Reproduction. 2013;28(2):509-18.

[16] Ottolini CS, Newnham LJ, Capalbo A, Natesan SA, Joshi HA, Cimadomo D, et al. Genome-wide maps of recombination and chromosome segregation in human oocytes and embryos show selection for maternal recombination rates. Nature genetics. 2015;47(7):727-35.

[17] Handyside AH, Harton GL, Mariani B, Thornhill AR, Affara N, Shaw MA, et al. Karyomapping: a universal method for genome wide analysis of genetic disease based on mapping crossovers between parental haplotypes. Journal of medical genetics. 2010;47(10):651-8.

[18] Zamani Esteki M, Dimitriadou E, Mateiu L, Melotte C, Van der Aa N, Kumar P, et al. Concurrent Whole-Genome Haplotyping and Copy-Number Profiling of Single Cells. The American Journal of Human Genetics. 2015;96(6):894-912.

[19] Masset H, Zamani Esteki M, Dimitriadou E, Dreesen J, Debrock S, Derhaag J, et al. Multi-centre evaluation of a comprehensive preimplantation genetic test through haplotyping-by-sequencing. Human Reproduction. 2019;34(8):1608-19.

[20] De Witte L, Raman L, Baetens M, De Koker A, Callewaert N, Symoens S, et al. GENType: all-in-one preimplantation genetic testing by pedigree haplotyping and copy number profiling suitable for third-party reproduction. Human Reproduction. 2022.

[21] Tšuiko O, Vanneste M, Melotte C, Ding J, Debrock S, Masset H, et al. Haplotyping-based preimplantation genetic testing reveals parent-of-origin specific mechanisms of aneuploidy formation. NPJ genomic medicine. 2021;6(1):1-10.

[22] Ma Y, Wang J, Li R, Ding C, Xu Y, Zhou C, et al. Mapping of meiotic recombination in human preimplantation blastocysts. G3: Genes, Genomes, Genetics. 2023;13(4):jkad031.

[23] Auton A, McVean G. Recombination rate estimation in the presence of hotspots. Genome research. 2007;17(8):1219-27.

[24] Spence JP, Song YS. Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. Science Advances. 2019;5(10):eaaw9206.

[25] Halldorsson BV, Palsson G, Stefansson OA, Jonsson H, Hardarson MT, Eggertsson HP, et al. Characterizing mutagenic effects of recombination through a sequence-level genetic map. Science. 2019;363(6425):eaau1043.

[26] Adrion JR, Galloway JG, Kern AD. Predicting the landscape of recombination using deep learning. Molecular biology and evolution. 2020;37(6):1790-808.

[27] Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nature Reviews Genetics. 2010;11(7):499-511.

[28] Ariad D, Yan SM, Victor AR, Barnes FL, Zouves CG, Viotti M, et al. Haplotype-aware inference of human chromosome abnormalities. Proceedings of the National Academy of Sciences. 2021;118(46).

[29] Liu S, Huang S, Chen F, Zhao L, Yuan Y, Francis SS, et al. Genomic analyses from non-invasive prenatal testing reveal genetic associations, patterns of viral infections, and Chinese population history. Cell. 2018;175(2):347-59.

[30] Vermeesch JR, Voet T, Devriendt K. Prenatal and pre-implantation genetic diagnosis. Nature Reviews Genetics. 2016;17(10):643-56.

[31] The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

[32] Wang C, Zhan X, Bragg-Gresham J, Kang HM, Stambolian D, Chew EY, et al. Ancestry estimation and control of population stratification for sequence-based association studies. Nature Genetics. 2014;46(4):409-15.

[33] McCoy RC, Demko ZP, Ryan A, Banjevic M, Hill M, Sigurjonsson S, et al. Evidence of selection against complex mitotic-origin aneuploidy during preimplantation development. PLoS genetics. 2015;11(10):e1005601.

[34] Sagi I, Benvenisty N. Haploidy in Humans: An Evolutionary and Developmental Perspective. Developmental Cell. 2017;41(6):581-9.

[35] Raman L, Dheedene A, De Smet M, Van Dorpe J, Menten B. WisecondorX: improved copy number detection for routine shallow whole-genome sequencing. Nucleic acids research. 2019;47(4):1605-14.

[36] Bugge M, Collins A, Petersen MB, Fisher J, Brandt C, Michael Hertz J, et al. Non-disjunction of chromosome 18. Human molecular genetics. 1998;7(4):661-9.

[37] Bugge M, Collins A, Hertz JM, Eiberg H, Lundsteen C, Brandt CA, et al. Non-disjunction of chromosome 13. Human molecular genetics. 2007;16(16):2004-10.

[38] Lamb NE, Freeman SB, Savage-Austin A, Pettay D, Taft L, Hersey J, et al. Susceptible chiasmate configurations of chromosome 21 predispose to non–disjunction in both maternal meiosis I and meiosis II. Nature genetics. 1996;14(4):400-5.

[39] Chernus JM, Sherman SL, Feingold E. Analyses stratified by maternal age and recombination further characterize genes associated with maternal nondisjunction of chromosome 21. Prenatal Diagnosis. 2021;41(5):591-609.

[40] Hassold TJ, Hunt PA. Missed connections: recombination and human aneuploidy. Prenatal Diagnosis. 2021;41(5):584-90.

[41] Hassold T, Maylor-Hagen H, Wood A, Gruhn J, Hoffmann E, Broman KW, et al. Failure to recombine is a common feature of human oogenesis. The American Journal of Human Genetics. 2021;108(1):16-24.

[42] Gray S, Cohen PE. Control of meiotic crossovers: from double-strand break formation to designation. Annual review of genetics. 2016;50:175-210.

[43] Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of molecular biology. 1975;98(3):503-17.

[44] Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 1985;230(4732):1350-4.

[45] Xu J, Zhang M, Niu W, Yao G, Sun B, Bao X, et al. Genome-wide uniparental disomy screen in human discarded morphologically abnormal embryos. Scientific Reports. 2015;5(1):1-10.

[46] Sun H, Rowan BA, Flood PJ, Brandt R, Fuss J, Hancock AM, et al. Linked-read sequencing of gametes allows efficient genome-wide analysis of meiotic recombination. Nature Communications. 2019;10(1):4310.

[47] Bansal V. Integrating read-based and population-based phasing for dense and accurate haplotyping of individual genomes. Bioinformatics. 2019;35(14):i242-8.

[48] Wang Y, Tsuo K, Kanai M, Neale BM, Martin AR. Challenges and opportunities for developing more generalizable polygenic risk scores. Annual review of biomedical data science. 2022;5:293-320.

[49] Chen YC, Liu T, Yu CH, Chiang T, Hwang C. Effects of GC Bias in Next-Generation-Sequencing Data on De Novo Genome Assembly. PLoS One. 2013;8.

[50] Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature. 2001;409(6822):928-34.

[51] Chernick MR. Bootstrap methods: A guide for practitioners and researchers. vol. 619. 2nd ed. United States of America: John Wiley & Sons; 2007.

[52] Rabinowitz M, Ryan A, Gemelos G, Hill M, Baner J, Cinnioglu C, et al. Origins and rates of aneuploidy in human blastomeres. Fertility and Sterility. 2012;97(2):395-401.

[53] Kubicek D, Hornak M, Horak J, Navratil R, Tauwinklova G, Rubes J, et al. Incidence and origin of meiotic whole and segmental chromosomal aneuploidies detected by karyomapping. Reproductive BioMedicine Online. 2019;38(3):330-9.

[54] Cayli S, Jakab A, Ovari L, Delpiano E, Celik-Ozenci C, Sakkas D, et al. Biochemical markers of sperm function: male fertility and sperm selection for ICSI. Reproductive biomedicine online. 2003;7(4):462-8.

[55] Jakab A, Sakkas D, Delpiano E, Cayli S, Kovanci E, Ward D, et al. Intracytoplasmic sperm injection: a novel selection method for sperm with normal frequency of chromosomal aneuploidies. Fertility and Sterility. 2005;84(6):1665-73.

[56] Huszar G, Jakab A, Sakkas D, Ozenci CC, Cayli S, Delpiano E, et al. Fertility testing and ICSI sperm selection by hyaluronic acid binding: clinical and genetic aspects. Reproductive biomedicine online. 2007;14(5):650-63.

[57] Sakkas D, Ramalingam M, Garrido N, Barratt CL. Sperm selection in natural conception: what can we learn from Mother Nature to improve assisted reproduction outcomes? Human reproduction update. 2015;21(6):711-26.

[58] Straver R, Sistermans EA, Holstege H, Visser A, Oudejans CB, Reinders MJ. WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme. Nucleic acids research. 2014;42(5):e31-1.

[59] Wang C, Zhan X, Liang L, Abecasis GR, Lin X. Improved ancestry estimation for both genotyping and sequencing data using projection procrustes analysis and genotype imputation. The American Journal of Human Genetics. 2015;96(6):926-37.