Approximate Bayesian computation supports a high incidence of chromosomal mosaicism in blastocyst-stage human embryos

Qingya Yang[1], Sara A. Carioscia[1], Matthew Isada[1], Rajiv C. McCoy[1*]

[1]Department of Biology, Johns Hopkins University, Baltimore, MD, USA 21218
*Correspondence to: rajiv.mccoy@jhu.edu

Running head: Simulation-based inference of embryonic aneuploidy
Keywords: aneuploidy, IVF, PGT-A, development, meiosis, mitosis

Summary

Chromosome errors are the main cause of human pregnancy loss. Researchers used computer simulations and statistics to estimate the rates of these errors based on published clinical data from *in vitro* fertilized embryos. To explain data, the researchers estimated that 40-58% of cell divisions that produced the egg or sperm were affected by an error, along with 1.5-6.3% of cell divisions after fertilization. Their results also showed that fewer than 1% of embryos are completely normal in all cells and that many embryos contain small groups of abnormal cells that testing often misses. These findings improve understanding of early development.

Abstract

Chromosome mis-segregation is common in human meiosis and mitosis, and the resulting aneuploidies are the leading cause of pregnancy loss. Preimplantation genetic testing for aneuploidy (PGT-A) prioritizes chromosomally normal embryos for transfer based on analysis of a biopsy of approximately five trophectoderm cells from blastocyst-stage *in vitro* fertilized (IVF) embryos. While modern PGT-A platforms classify these biopsies as aneuploid, euploid, or mosaic (a mixture of normal and aneuploid cells), the underlying incidences of aneuploid, euploid, and mosaic embryos and the rates of meiotic and mitotic error that produced them remain largely unknown. To address this knowledge gap, we paired a method for embryo simulation with approximate Bayesian computation (ABC) to infer rates of meiotic and mitotic error that explain published PGT-A data. Using simulation, we also evaluated the chromosomal status of entire embryos. For a published clinical sample, we estimated a 40-58% probability of meiotic error per meiosis and a 1.5-6.3% probability of mitotic error per mitosis, depending on assumptions about spatial organization. In addition, our analyses suggest that less than 1% of blastocysts are fully euploid and that many embryos possess low-level mosaic clones that are not captured during biopsy. These conclusions were relatively insensitive to misclassification of mosaic biopsies. Together, our findings imply that low-level mosaicism is a normal feature of embryogenesis and are consistent with clinical data demonstrating the developmental potential of mosaic-testing embryos. More broadly, our work helps overcome the limitations of embryo biopsies to estimate fundamental rates of chromosome mis-segregation in human development.

Introduction

It is estimated that only approximately half of human conceptions survive to birth. The leading cause of human pregnancy losses—most of which occur very early in development—is abnormality in chromosome number (aneuploidy) (Boklage 1990; Jarvis 2016; Macklon et al. 2002). Chromosomes frequently mis-segregate during (predominantly maternal) meiosis, producing embryos with homogeneous forms of aneuploidy that affect all cells (Gruhn et al. 2019; Hassold and Hunt 2001). However, errors may also arise during postzygotic cell divisions, producing "mosaic" embryos with two or more karyotypically distinct cell lineages (McCoy 2017). The first two cell divisions appear especially susceptible to such mitotic errors (Currie et al. 2022).

While many forms of aneuploidy are lethal during cleavage-stage development, other aneuploidies, especially in mosaic form, can survive to the blastocyst stage and beyond (Capalbo et al. 2021; McCoy et al. 2023; Viotti, Victor et al. 2021). Considering the potential fitness consequences of aneuploidy, preimplantation genetic testing for aneuploidy (PGT-A) seeks to improve IVF outcomes by prioritizing chromosomally normal (i.e., euploid) embryos for transfer based on genetic analysis of embryo biopsies (Cimadomo et al. 2020), although its clinical efficacy is the subject of long-standing debate (Cornelisse et al. 2020; Mastenbroek et al. 2007; Munne et al. 2007; Munné et al. 2019; Rubio et al. 2017; Verpoest et al. 2018). Current implementations of PGT-A involve sequencing or single nucleotide polymorphism (SNP) microarray-based genotyping of DNA extracted from trophectoderm biopsies of blastocyst-stage embryos, 5-6 days after fertilization (Vermeesch et al. 2016). These modern, sensitive PGT-A platforms have revealed evidence of mosaic aneuploidy within 2-26% of embryo biopsies, posing a dilemma for clinical diagnosis and management (Popovic et al. 2020).

One notable limitation of PGT-A is its dependence on a single, spatially restricted biopsy of 5-10 trophectoderm cells, which may or may not represent the constitution of the remaining >100 cells of the blastocyst-stage embryo, including the inner cell mass (Esfandiari et al. 2016; Vera-Rodriguez and Rubio 2017). While studies employing multiple biopsies of individual embryos (Kim et al. 2022; Victor, Griffin et al. 2019) or single-cell genomic analysis (Gallardo et al. 2023; Starostik et al. 2020) have offered additional insight, they are typically limited in sample size of embryos and/or availability of ground-truth data for comparison. Thus, the clinical literature is replete with reports of the relative incidences of euploid, mosaic, and aneuploid embryo biopsies as diagnosed by PGT-A, but the underlying incidence of mosaic embryos and the fundamental rates of meiotic and mitotic errors that produced them remain largely obscure.

Approximate Bayesian computation (ABC) offers a statistical framework for estimating unknown parameters given summary statistics from observed data and a generative model for simulating data under a specified set of parameters (Sunnaker et al. 2013). ABC is most useful in situations such as ours where the likelihood function is complex or intractable, but efficient simulation is possible. In addition to point estimates of the parameters, the posterior distributions produced by ABC provide measurements of uncertainty that can be summarized, for example, with credible intervals (CIs). ABC has found growing applications in the fields of

population genetics, ecology, and epidemiology, in concert with the development of powerful computational simulations (Beaumont 2010).

Skinner et al. (2024) recently developed an R package, Tessera, to model the growth and distribution of aneuploid cells in a three-dimensional early embryo, followed by a trophectoderm biopsy to mimic PGT-A. In our study, we used Tessera to simulate embryos over a range of meiotic and mitotic error rates. We then applied ABC to identify the subset of embryos whose biopsy results best reflect published PGT-A data from IVF clinics, in turn obtaining posterior probability estimates of meiotic and mitotic error rates and incidences of euploid, mosaic, and aneuploid embryos. Together, our study helps overcome the limited focus on embryo biopsies to reveal knowledge of fundamental parameters shaping the chromosomal landscape of human preimplantation embryos.

Results

ABC enables inference of rates of meiotic and mitotic error

We used Tessera to simulate large samples of human embryos and biopsies within the ABC framework (see **Methods**; **Figure 1**) and compare to published data from PGT-A. For simplicity, we focused our analysis on published reports of aneuploidies affecting whole chromosomes, as segmental aneuploidies (duplication or deletion of sub-chromosomal segments) originate from distinct mechanisms, and criteria for their reporting vary substantially across clinics (Girardi et al. 2020).

Given a rate of meiotic error per meiosis (i.e., the probability of producing an aneuploid zygote, ignoring the distinction between maternal and paternal meiosis) drawn from a uniform prior distribution, a rate of mitotic error per mitosis (i.e., the probability of a postzygotic cell division producing aneuploid daughter cells) drawn from a uniform prior distribution, and a rate of dispersal (i.e., the spatial organization of aneuploid cells within mosaic embryos) fixed at 0, 0.5, or 1, we simulated trials of 1,000 embryos each. By definition, a dispersal rate of 0 denotes complete spatial clustering, whereby aneuploid cells are adjacent to one another, a dispersal rate of 1 denotes maximum spatial separation of aneuploid cells, and intermediate values represent more moderate patterns of spatial clustering. We selected dispersal values of 0, 0.5, and 1 to span the entire parameter range, as the spatial distribution of aneuploid cells within mosaic embryos remains largely unknown (see **Discussion**).

We then simulated biopsies of five spatially adjacent cells from each embryo and summarized each trial based on the proportion of euploid, mosaic, and aneuploid biopsies. A biopsy was categorized as mosaic if any but not all of the five sampled cells were aneuploid. Within the ABC framework, we compared these proportions to those from published PGT-A results from 6,766 clinical trophectoderm biopsies (23% euploid, 19% mosaic, and 58% aneuploid) by Capalbo et al. (2021) and rejected simulations whose distance from the published data

exceeded a given tolerance threshold. For the remaining selected simulations, we used the corresponding values of the meiotic and mitotic error rate parameters to approximate posterior probability distributions. The sequence of tolerance levels, stopping criteria, and bias correction were obtained using the adaptive method described by Lenormand et al. (2013) and implemented with the EasyABC package in R (Jabot et al. 2013).

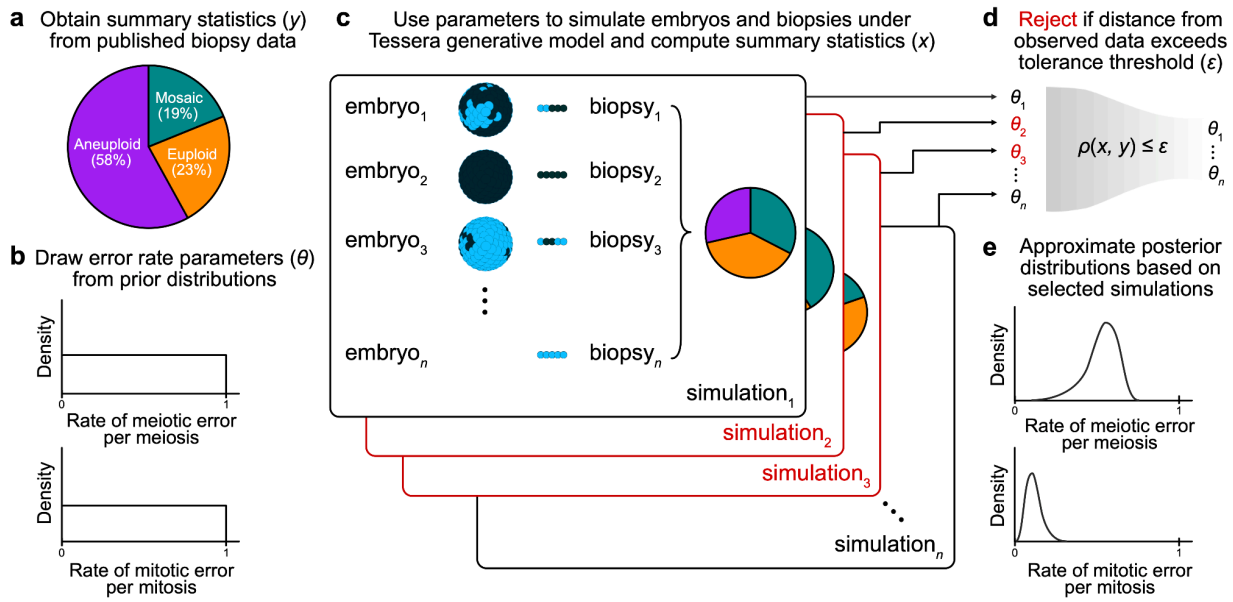

Figure 1: Schematic of the ABC approach for inferring meiotic and mitotic error rates. a) Observed proportions of euploid, aneuploid, and mosaic biopsies from PGT-A were obtained from the literature (Capalbo et al. 2021) and used as summary statistics for ABC. b) For each simulation trial, probabilities of meiotic and mitotic error (θ) were randomly drawn from uniform prior distributions ranging from 0 to 1. c) These values were used to determine the number of cells that would be aneuploid (represented as dark blue in the schematic) versus normal disomic (light blue) in the final embryo. This number was converted to a proportion and then used as input to the Tessera package. For each combination of error parameters and each of three fixed values for dispersal (0, 0.5, 1), 1,000 embryos were generated and biopsied. For each embryo, 8 rounds of mitotic division were simulated (producing a total of 256 cells), and one 5-cell biopsy was obtained. d) Using the adaptive method described by Lenormand et al. (2013), ABC selected the set of simulations that produced summary statistics (proportions of euploid, mosaic, and aneuploid biopsies) that best reflected the observed biopsy data. e) The error rates that produced the selected simulations were used to approximate the posterior distributions of these parameters.

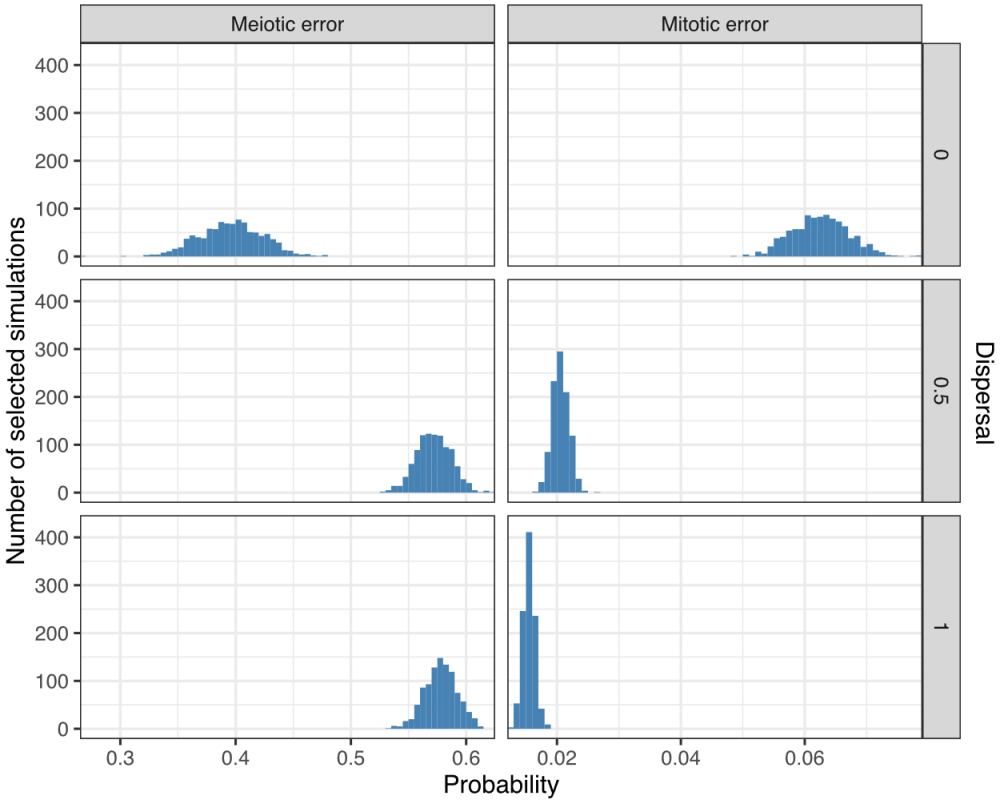

Figure 2: Posterior probabilities of meiotic and mitotic error. Posterior distributions of meiotic and mitotic error probabilities (per meiosis and mitosis, respectively) based on 3,000 simulations (1,000 simulations per level of dispersal) selected by ABC as best matching published data.

At a dispersal level of 0, the posterior mean estimates of the meiotic and mitotic error rates were 40% (95% CI [34%, 45%]) and 6.3% (95% CI [5.4%, 7.1%]), respectively; at a dispersal level of 0.5, the posterior mean estimates of the meiotic and mitotic error rates were 57% (95% CI [54%, 60%]) and 2.1% (95% CI [1.8%, 2.3%]), respectively; and at a dispersal level of 1, the posterior mean estimates of the meiotic and mitotic error rates were 58% (95% CI [55%, 60%]) and 1.5% (95% CI [1.4%, 1.7%]), respectively. These estimates are summarized in **Figure 2** and **Table S1** and indicate that higher dispersal levels (i.e., less spatial clustering aneuploid cells) require lower estimates of mitotic aneuploidy rates and modestly higher estimates of meiotic aneuploidy rates to explain the observed data. This result is intuitive, as higher dispersal levels make it more likely to detect mosaic aneuploidy within a given biopsy, such that relatively lower mitotic error rates would be required to explain the observed data.

At a dispersal level of 0, the posterior mean estimates of the meiotic and mitotic error rates were 40% (95% CI [34%, 45%]) and 6.3% (95% CI [5.4%, 7.1%]), respectively; at a dispersal level of 0.5, the posterior mean estimates of the meiotic and mitotic error rates were 57% (95% CI [54%, 60%]) and 2.1% (95% CI [1.8%, 2.3%]), respectively; and at a dispersal level of 1, the posterior mean estimates of the meiotic and mitotic error rates were 58% (95% CI [55%, 60%]) and 1.5% (95% CI [1.4%, 1.7%]), respectively. These estimates are summarized in **Figure 2** and **Table S1**

and indicate that higher dispersal levels (i.e., less spatial clustering aneuploid cells) require lower estimates of mitotic aneuploidy rates and modestly higher estimates of meiotic aneuploidy rates to explain the observed data. This result is intuitive, as higher dispersal levels make it more likely to detect mosaic aneuploidy within a given biopsy, such that relatively lower mitotic error rates would be required to explain the observed data.

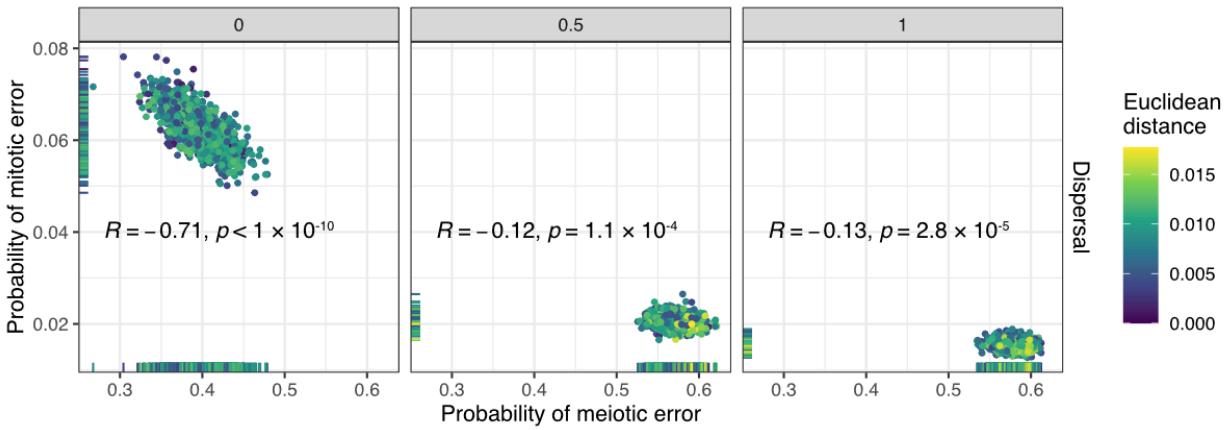

Figure 3: Relationship between inferred meiotic and mitotic error rates. Meiotic versus mitotic error rate, stratified by dispersal level, across 3,000 simulations selected by ABC. The coefficient (*R*) and p-values of Pearson correlations between inferred meiotic and mitotic error rates are displayed in each panel. The Euclidean distances between the vectors of summary statistics of selected simulations and the target summary statistics from the literature are indicated with colors.

For the selected simulations, we observed a negative correlation between the inferred meiotic and mitotic error rates at each dispersal level (**Figure 3**). The correlation was much more pronounced when the dispersal level was zero, reflecting the fact that spatial clustering of cells makes it more likely that a mitotic aneuploidy could mimic a meiotic aneuploidy (or euploid embryo) by affecting all cells (or no cells) within a biopsy. Higher assumed rates of dispersal make these forms of aneuploidy more distinguishable in biopsy data, reducing the correlation.

Evidence that few embryos are fully euploid

We next used simulations to examine the cellular content of embryos generated under the posterior distributions of meiotic and mitotic error rates inferred with ABC from Capalbo et al. (2021) (**Table S2**). To this end, we simulated new samples of embryos by drawing meiotic and mitotic error rates from the posterior distributions (at each of the three assumed levels of dispersal) and examined their entire cellular content. In contrast to the proportions of euploid, aneuploid, and mosaic biopsies, this allowed us to quantify the underlying proportions of fully euploid, fully aneuploid, and mosaic embryos (and their corresponding proportions of aneuploid cells) that produced these data—quantities that remain largely unknown within the field. An

embryo was considered mosaic if any but not all cells within it were aneuploid, but the entire distribution of aneuploid cell proportions was also visualized.

Strikingly, we found that regardless of assumptions about dispersal, fewer than 1% of embryos from the posterior predictive samples were entirely euploid (**Figure 4A**; **Table S2**). Specifically, for simulated dispersal levels of 0, 0.5, and 1, we inferred that 0, 1.5×10^{-5}, and 1.6×10^{-4} proportions of embryos were fully euploid, respectively. A large proportion of simulated embryos exhibited mosaicism, especially low-level mosaicism affecting <25% of cells when allowing for non-zero dispersal. Higher levels of dispersal were consistent with lower levels of aneuploidy within mosaic embryos and higher proportions of fully aneuploid embryos (**Figure 4A**).

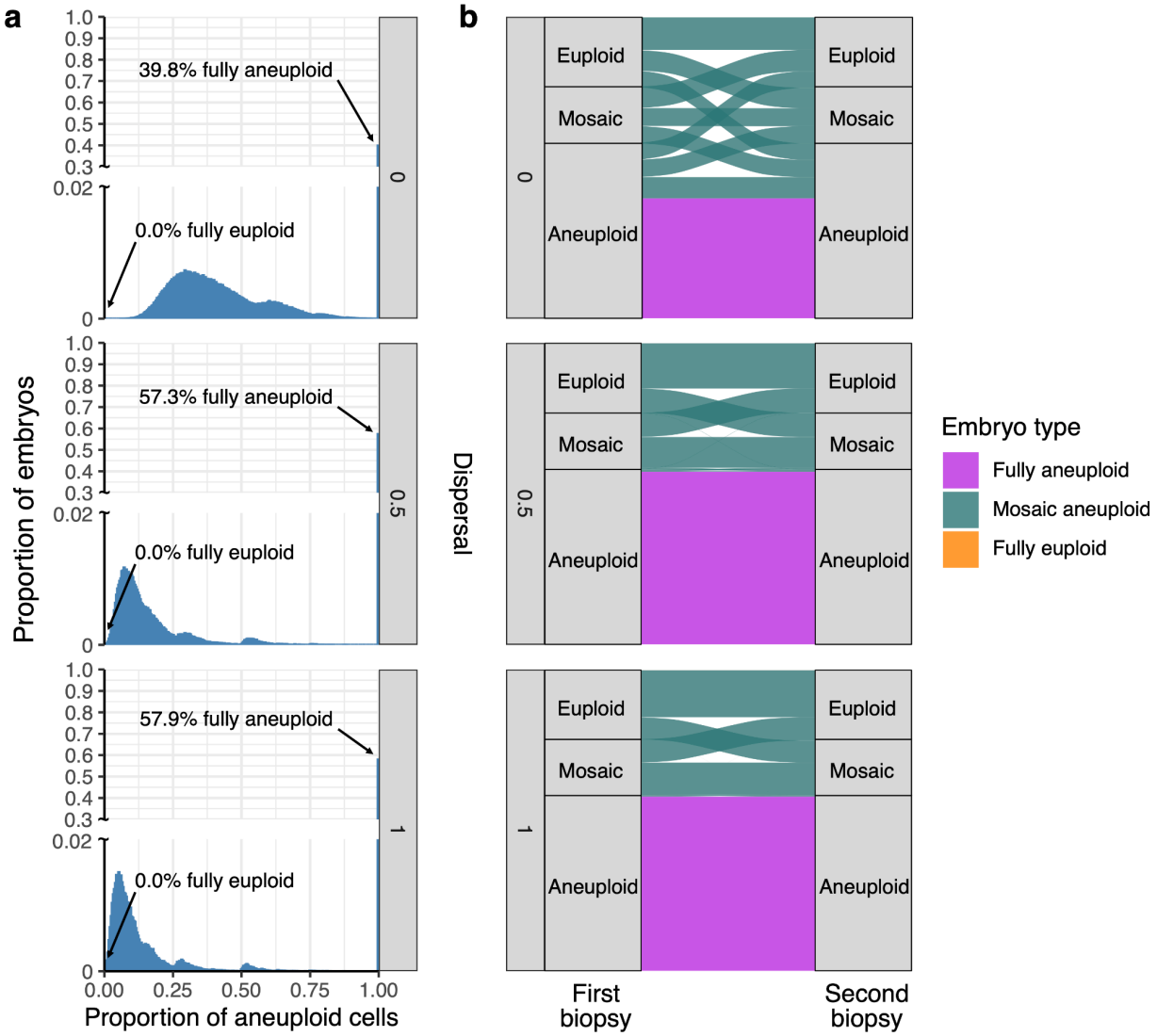

Figure 4: Posterior predictive samples of embryos that are fully aneuploid, mosaic aneuploid, and fully euploid. a) Distribution of proportions of aneuploid cells in embryos simulated based on meiotic and mitotic error rates drawn from the posterior distributions

(1,000,000 embryos for each level of dispersal; **Figure 2**). The arrow marks the percentage of fully euploid embryos (0 aneuploid cells) in the posterior predictive sample for each dispersal level (all 0.0%). b) Comparisons between ploidy statuses of first and second biopsies from embryos of different underlying types (indicated by ribbon colors), stratified by assumed levels of dispersal. Ribbons connecting the same versus different biopsy categories indicate concordant and discordant results, respectively, while the widths of the ribbons indicate proportions.

Evaluating the reliability and consistency of embryo biopsies

By simulating biopsies from these posterior predictive samples, we next evaluated the extent to which embryo biopsies represented the ploidy statuses of the embryos from which they originated. Intuitively, due to the small number of cells in biopsies, aneuploid cells of low-level mosaic embryos were frequently unsampled, and their embryo biopsies were more often classified as euploid than mosaic (**Figure 4B**). Within a mosaic embryo, higher levels of dispersal increased sensitivity for capturing one or more aneuploid cells within a biopsy. This effect is apparent from the skewness of the distribution of the proportion of aneuploid cells within mosaic embryos, where higher simulated levels of dispersal required smaller proportions of aneuploid cells within mosaic embryos to explain the observed data (**Figure 4A**).

Given reports of healthy births following the transfer of embryos categorized as mosaic (Capalbo et al. 2021; Greco et al. 2015; Viotti, Victor et al. 2021), one important clinical concern regards the possibility that mosaic embryos could produce a biopsy categorized as aneuploid, such that a potentially viable embryo would not be transferred. Within our posterior predictive samples, we observed that 31.5%, 1.4%, and 0.7% of aneuploid biopsies originated from mosaic embryos at assumed dispersal levels of 0, 0.5, and 1, respectively (**Figure 4B**; **Table S3**). This result implies that accurate interpretation of aneuploid embryo biopsies strongly depends on the assumed spatial organization of aneuploid cells within mosaic embryos, consistent with previous modeling using Tessera (Skinner et al. 2024).

In addition to cell tracking and other microscopy-based approaches, information about such spatial organization can be derived from previous studies that compared PGT-A results between multiple biopsies of individual blastocyst-stage embryos (Cascante et al. 2023; Kim et al. 2022; Marin et al. 2021; Navratil et al. 2020; Victor, Griffin et al. 2019). To mimic this study design, we simulated a second biopsy from each embryo in our posterior predictive samples and compared their results to the original biopsy (**Figure 4B**; **Table S4**). Across all dispersal levels, initial mosaic biopsies were often discordant with the second biopsy (31.7%, 54.5%, and 57.2% concordance at dispersal levels of 0, 0.5, and 1, respectively), consistent with the literature. However, discordance between the first and second biopsy was much more pronounced at the lowest dispersal level (dispersal = 0), whereas non-zero dispersal levels exhibited higher concordance—especially for the aneuploid class, which is of greatest clinical relevance (80.7%, 98.9%, and 99.5% concordance at dispersal levels of 0, 0.5, and 1, respectively).

Inferences are robust to moderate rates of misclassification

In various perspective articles, authors have argued that mosaic biopsies may be over-diagnosed, as technical variability in depth of coverage (e.g., due to amplification artifacts) may be falsely interpreted as evidence of mosaicism (Capalbo et al. 2017; Treff and Marin 2021). To address the possibility of biopsy misclassification and assess its impact on our results, we adjusted the reference data from Capalbo et al. (2021) by evenly reassigning a varying proportion of mosaic biopsies as either euploid or aneuploid (half to each category; **Figure S1**), and then applied the same ABC process to this adjusted reference data. We then sampled error rate pairs from each of the newly inferred posterior distributions and generated embryos for downstream analysis. While the inferred incidence of mosaic embryos quantitatively declined with higher rates of misclassification, qualitative conclusions about the incidence of mosaic embryos were relatively insensitive to the misdiagnosis of embryo biopsies (**Figure 5**). Notably, even when assuming a misdiagnosis rate as high as 90%, fewer than 10% of embryos were inferred as fully euploid, regardless of the assumed level of dispersal.

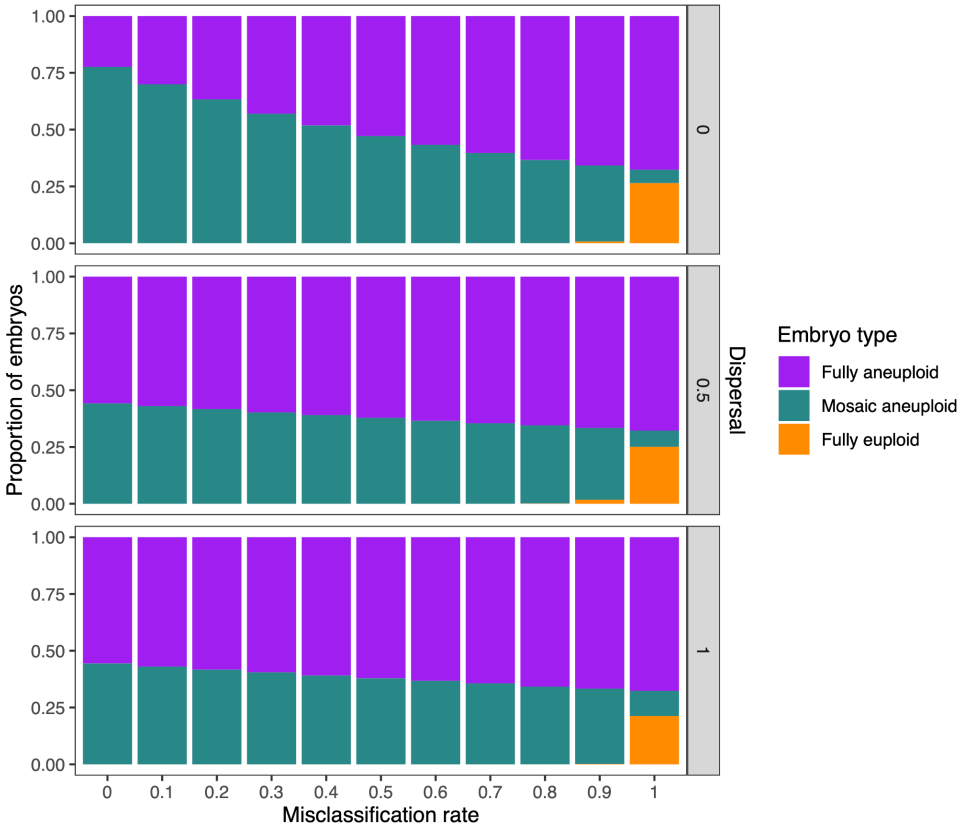

Figure 5: Inferred proportions of embryo types assuming varying levels of biopsy misclassification. 1,000 pairs of error rates were sampled from the posterior distributions generated by ABC with adjusted target values (based on the rate of mosaic biopsy misclassification; **Figure S1**). Each error rate pair was used to generate 1,000 embryos, and the proportions of each embryo type were plotted. The x-axis represents the proportion of published

mosaic biopsies that were assumed to be false positives and evenly reassigned to euploid and aneuploid categories.

Inferences are robust to variation across PGT-A datasets

Given that patient cohorts, laboratory procedures, and details of data analysis and reporting may vary across published PGT-A studies, we extended our approach to three additional input datasets (Clarke et al. 2023; Munné et al. 2017; Rodrigo et al. 2020) to evaluate the robustness of our conclusions (**Table 1**). The results were qualitatively consistent across studies. While meiotic error rates vary between studies (e.g., due to differences in the maternal age distribution; **Table S5**), inferred rates of mitotic error remained relatively constant, and small proportions of embryos (<1%) simulated under these parameters were fully euploid.

Data set	Dispersal	Published biopsy data			Inferred probabilities		Inferred embryo types		
		Euploid	Mosaic	Aneuploid	Meiotic error	Mitotic error	Fully euploid	Mosaic aneuploid	Fully aneuploid
Capalbo et al. (2021)	0	0.23	0.19	0.58	0.39	0.064	0.00	0.60	0.40
Clarke et al. (2023)	0	0.49	0.18	0.33	0.30	0.024	0.00	0.70	0.30
Munné et al. (2017)	0	0.53	0.15	0.32	0.31	0.017	0.00	0.69	0.31
Rodrigo et al. (2020)	0	0.51	0.06	0.43	0.38	0.016	0.00	0.62	0.38
Capalbo et al.	0.5	0.23	0.19	0.58	0.59	0.021	0.00	0.43	0.57
Clarke et al.	0.5	0.49	0.18	0.33	0.34	0.011	0.00	0.67	0.33
Munné et al.	0.5	0.53	0.15	0.32	0.33	0.009	0.01	0.67	0.32
Rodrigo et al.	0.5	0.51	0.06	0.43	0.43	0.011	0.00	0.58	0.42

Data set	Dispersal	Published biopsy data			Inferred probabilities		Inferred embryo types		
Capalbo et al.	1	0.23	0.19	0.58	0.58	0.015	0.00	0.42	0.58
Clarke et al.	1	0.49	0.18	0.33	0.32	0.008	0.01	0.66	0.33
Munné et al.	1	0.53	0.15	0.32	0.32	0.007	0.02	0.65	0.33
Rodrigo et al.	1	0.51	0.06	0.43	0.42	0.017	0.00	0.58	0.42

Table 1: Posterior mean estimates of meiotic and mitotic error probabilities inferred with ABC across a range of target proportions of euploid, mosaic, and aneuploid embryo biopsies reported in the literature based on PGT-A. Estimates are shown across three levels of dispersal (0, 0.5, and 1). The posterior distributions of meiotic and mitotic error probabilities were used to simulate embryos, and the proportions of fully aneuploid, mosaic aneuploid, and fully euploid simulated embryos are also reported.

Discussion

Aneuploidies are prevalent in human development and are the leading cause of pregnancy loss (Hassold et al. 2007). Retrospective studies of PGT-A data typically report proportions of embryo biopsies classified as euploid, aneuploid, or mosaic. While potentially justified from a clinical perspective (as these are the classifications used to inform embryo prioritization), it is important to consider from a fundamental biological perspective that these results are based on a biopsy of ~5 spatially restricted trophectoderm cells from a blastocyst-stage embryo composed of >100 cells. The underlying incidences of fully euploid, fully aneuploid, and mosaic embryos remain largely unknown and contentious (Capalbo et al. 2017; Treff and Marin 2021; Vera-Rodriguez and Rubio 2017), and the fundamental rates of meiotic and mitotic chromosome mis-segregation that produce these patterns have not been estimated aside from live imaging studies of the first two postzygotic cell divisions (Currie et al. 2022). Here we used the statistical method ABC to address these limitations and achieve embryo-wide estimates of aneuploidy that best explain published clinical data.

Our most provocative finding is that the inferred proportion of mosaic embryos is much higher than the observed proportion of mosaic biopsies and that very few embryos are fully euploid at the blastocyst-stage of development. This conclusion, as well as the results of our re-biopsy simulation, is consistent with clinical studies performing multiple biopsies of individual embryos, which observed that "fully aneuploid" biopsies exhibited high concordance upon re-biopsy, suggesting their likely meiotic origins. Meanwhile, mosaic aneuploid biopsies exhibited low concordance, suggesting that they either originated from technical artifacts or reflected the

limited, localized sampling of aneuploid clones within patchy mosaic embryos (Cascante et al. 2023; Kim et al. 2022; Marin et al. 2021; Navratil et al. 2020; Victor, Griffin et al. 2019). Recent studies leveraging single-cell sequencing data also support the conclusion that low-level mosaicism is widespread at the blastocyst stage, although they were limited in sample size of embryos and ground-truth data for validation (Chavli et al. 2024; Gallardo et al. 2023; Starostik et al. 2020; Zhai et al. 2024).

As with all models, our results must be interpreted in light of several simplifying assumptions. One such assumption regards the spatial distribution of aneuploid cells within mosaic embryos, which remains largely unknown. Past work in mouse models of experimentally-induced mosaicism reported an even distribution of aneuploid cells throughout post-implantation conceptuses (day 13.5 of embryonic development), with no significant enrichment in the placenta versus the fetus (Bolton et al. 2016). To address this uncertainty, we performed simulations at varying dispersal levels ranging from complete spatial clustering (dispersal = 0) to maximum spatial separation (dispersal = 1) of aneuploid cells. Importantly, our major qualitative conclusions about the incidence of mosaic aneuploidy and underlying rates of meiotic and mitotic error were relatively insensitive to assumptions about dispersal levels. Meanwhile, the relationship between the ploidy statuses of embryo biopsies and underlying embryos, as well as the concordance between multiple embryo biopsies, was more sensitive to assumed dispersal, consistent with previous work (Skinner et al. 2024). Specifically, the dispersal level of zero exhibited the highest rate of discordance, including a substantial risk that mosaic embryos will produce aneuploid biopsies. Given that published studies have typically observed that embryos producing aneuploid biopsies exhibit high concordance upon re-biopsy (Cascante et al. 2023; Kim et al. 2022; Marin et al. 2021; Navratil et al. 2020; Victor, Griffin et al. 2019), it is likely that true levels of dispersal are substantially greater than zero. Further research based on imaging and/or spatial genomic methods will be required to detail the spatial distribution of daughter cells after each mitotic division, improving understanding of the location of aneuploid cells within mosaic embryos and how these patterns change throughout development (McDole et al. 2018).

For simplicity, our model represents the ploidy of a given cell as a binary variable (aneuploid vs. euploid) and does not consider the specific chromosome affected or the category of aneuploidy (gain vs. loss of maternal vs. paternal homologs). The main consequence of these simplifications is that we do not consider the possibility of "rescue" events where aneuploid lineages revert to euploidy due to subsequent mitotic errors. We believe that these simplifications are nevertheless justified to avoid model overparameterization, especially given current evidence that aneuploidy rescue is rare in blastocyst-stage human embryos (based on the paucity of uniparental disomy).[6] Because gains and losses are not distinguished, we do not account for the possibility that sampling of cells possessing a reciprocal gain and loss (e.g., trisomy and monosomy chromosome 1) within a single biopsy could evade detection by bulk sequencing. Although we expect this phenomenon to remain rare due to the requirement for balanced representation of both lineages, it would nonetheless cause us to underestimate the true incidence of mitotic error and mosaicism in human blastocysts. Similarly, our model does not account for potential changes in rates of mitotic error across cell divisions, and indeed, the first two cell divisions are known to be particularly error-prone (Currie et al. 2022). In the work

presented here, we estimated a single fixed probability of mitotic error across all cell divisions that produced an embryo. It is likely that this fixed estimate is an underestimate of the rate of error for initial divisions but an overestimate of the rate of error for later divisions. However, it is worth noting that early divisions may have a greater influence on our estimates, given their greater impact on observed patterns of mosaicism. Allowing for varying rates of mitotic error across rounds of cell division is an interesting future direction, but may again introduce too many additional parameters given the dimensionality of the target data.

Related to the above considerations, we emphasize that published PGT-A data are subject to several selection and ascertainment biases. One of the most important such biases is that only embryos that survive to the blastocyst stage and exhibit good morphology are typically biopsied and reported in retrospective clinical studies (ESHRE Working Group on Chromosomal Mosaicism et al. 2022; Forman et al. 2013; Viotti, Victor et al. 2021). Meanwhile, the approximately half of embryos that arrest prior to blastocyst formation are highly enriched for abnormal cell divisions and complex forms of chromosomal mosaicism, as well as meiotic aneuploidies to a lesser extent (Chatzimeletiou et al. 2005; Maurer et al. 2015; McCoy et al. 2023; Santos et al. 2010). Additional biases include characteristics of the patient population, which could be enriched for particular forms of aneuploidy associated with infertility diagnosis. However, our previous research showed that common clinical indications for PGT-A exhibit only modest associations with meiotic or mitotic aneuploidy (McCoy et al. 2015), suggesting that this clinical cohort bias may be negligible. Nevertheless, we emphasize that our inferred rates of both meiotic and mitotic error should be interpreted as estimates for the clinically tested samples and may not generalize to samples of *in vivo* conceived embryos from the broader human population.

Our research findings underscore the distinction between evidence of mosaicism within a biopsy, the existence of mosaicism within an embryo, and the clinical implications of mosaicism for pregnancy outcomes. While meiotic aneuploidies are generally harmful for development, mosaic aneuploidies that persist to the blastocyst stage are potentially compatible with healthy birth (Greco et al. 2015; Victor, Tyndall et al. 2019). This hypothesis was recently supported by a non-selection clinical trial which demonstrated that embryos exhibiting mosaicism as assessed by PGT-A had pregnancy outcomes equivalent to euploid embryos (Capalbo et al. 2021), though this may depend on the specific definition of mosaicism (Viotti, McCoy et al. 2021). Our findings also support this hypothesis, as they suggest that most embryos yielding euploid biopsies are low-level mosaics. While numerous forms of pathogenic mosaic aneuploidy have been reported (Biesecker and Spinner 2013), current evidence suggests that mosaic aneuploidies identified in blastocyst-stage embryos are rarely detected at later stages of pregnancy or at birth (Viotti, Victor et al. 2021) (although see (Greco et al. 2023)). This observation is consistent with a model of selection against aneuploid cells within mosaic embryos during peri- and post-implantation development, recently supported by data from mouse and human embryo models (Bolton et al. 2016; Singla et al. 2020; Yang et al. 2021), as well as analyses of single-cell genomic data from human embryos (Starostik et al. 2020; Yang et al. 2021). Future extensions of our model may consider such mechanisms of intra-embryo

negative selection, especially if supplemented by data from additional embryonic and fetal time points.

Together, our work helps expand beyond the limited snapshot of embryo biopsies by estimating the fundamental rates of cell division errors that best explain clinical data. Understanding these basic error mechanisms, the evolutionary forces that shape them, and their consequences for development enhance basic understanding of human development, while building the foundations for future clinical innovations.

Methods

For each sampled combination of meiotic and mitotic error rates, we simulated a single biopsy from each embryo using the Tessera R package (Skinner et al. 2024). We then applied approximate Bayesian computation (ABC) using the R package EasyABC (Jabot et al. 2013) to infer posterior distributions of the meiotic and mitotic error rates by comparing the simulated biopsy data to published clinical data (Capalbo et al. 2021). Our approach is summarized in **Figure 1**.

Embryo model abstraction

Meiotic errors were simulated as producing embryos with entirely aneuploid cells. Mitotic errors were simulated as a given division of a euploid cell producing two aneuploid daughter cells. Upon selecting a meiotic and mitotic error rate, we simulated 8 rounds of mitotic cell division for a total of 256 cells. We repeated each simulation trial across three fixed dispersal levels: 0 (complete clustering of aneuploid cells), 0.5 (intermediate clustering of aneuploid cells), and 1 (maximum spatial separation of aneuploid cells). For the purpose of abstraction, aneuploidy was defined as a binary variable, ignoring the identity of individual chromosomes (see **Discussion**).

Biopsy data reference

The clinical data used as the target for ABC are derived from Capalbo et al. (2021) and include 6766 clinical trophectoderm biopsies performed in 2020 and 2021. Of these biopsies, 23.2% were euploid, 18.7% were mosaic, and 58.1% were aneuploid. Embryos in that study were classified as euploid if, based on estimates from copy number analysis, fewer than 20% of the biopsied cells were aneuploid, as mosaic if 20-70% of the cells were aneuploid, and as aneuploid if more than 70% of the cells were aneuploid for any chromosome (Capalbo et al. 2021).

ABC algorithm

ABC involves simulating a large number of trials under a given model, comparing the simulation results to a set of target summary statistics, and identifying the parameters that best explain

those data (Jabot et al. 2013). ABC is useful in situations such as these, where explicit likelihoods are not available but a parameterized simulator is available. Specifically, we used the adaptive population Monte Carlo ABC method from Lenormand et al. (2013) to automatically select a series of tolerance levels and stopping criteria. This method was demonstrated by the authors to produce high-quality estimates of the posterior distributions with fewer simulations.

We implemented this procedure with the `ABC_sequential` function and the `methods="Lenormand"` argument in the EasyABC package in R. Meiotic and mitotic error rates were drawn from uniform distributions (ranging from 0 to 1, exclusive), and each selected error rate pair was used to simulate 1,000 embryos (Jabot et al. 2013). The embryos and their biopsy summaries were then compared with the published clinical PGT-A results (Capalbo et al. 2021), and the posterior distributions were obtained.

Posterior predictive distributions of embryos

To understand the characteristics of embryos given our inferences, we simulated 1,000,000 unique embryos by drawing meiotic and mitotic error rates from our posterior distributions (1,000 times), repeating this procedure for each of the three fixed dispersal levels. Proportions of fully euploid, fully aneuploid, and mosaic embryos were then tabulated for each dispersal level, as was the distribution of the proportion of aneuploid cells within mosaic embryos.

To understand the concordance between multiple biopsies of the same embryo, we simulated two consecutive biopsies from each of the above embryos using Tessera's biopsy functions. The first biopsy used the default index cell, while the second biopsy randomly selected one of the 256 cells as the index cell. Biopsy classification followed the thresholds from Capalbo et al. (2021), with copy number estimates fewer than 20% being euploid, between 20% and 70% being mosaic, and over 70% being aneuploid.

Misclassification simulation

To investigate a scenario in which some biopsies classified as mosaic were in fact technical false positives, we reclassified mosaic biopsies from the Capalbo et al. (2021) study as euploid and aneuploid (from 0% to 100%, incrementing in steps of 10%), with misclassified embryos evenly reassigned to the euploid and aneuploid categories. Using these adjusted values as the new target, we repeated our ABC and posterior predictive sampling procedures described above across each of the three fixed dispersal levels.

Acknowledgments

We thank members of the McCoy lab and Origins of Aneuploidy Research Consortium for helpful feedback and discussions. Thank you also to staff at the Advanced Research Computing at Hopkins (ARCH) core facility.

Study funding

This work was supported by a Johns Hopkins University Provost's Undergraduate Research Award to Q.Y., a National Science Foundation Graduate Research Fellowship (1746891) to S.A.C., and a US National Institutes of Health/National Institute of General Medical Sciences (NIH/NIGMS) Award R35GM133747 to R.C.M. ARCH is supported by the National Science Foundation (OAC1920103).

Conflict of interest

R.C.M. is a co-inventor of a method for analysis of PGT-A data which is the subject of a U.S. patent (12,322,509) assigned to Johns Hopkins University.

Author contributions

Q.Y.: software, validation, formal analysis, investigation, data curation, writing - original draft, writing - review & editing, visualization. S.A.C.: methodology, software, investigation, data curation, writing - original draft, writing - review & editing, supervision, project administration, funding acquisition. M.I.: investigation, writing - review & editing. R.C.M.: conceptualization, methodology, resources, data curation, writing - review & editing, visualization, supervision, project administration, funding acquisition

Data and code availability

All scripts to execute the simulations described in the study are available at https://github.com/mccoy-lab/aneuploidyRates. Clinical data used for parameter inference were published in Capalbo et al. (2021) (see their Figure A), Clarke et al. (2023) (see their Table 3), Rodrigo et al. (2020) (see their Table 1), and Munne et al. (2017) (see their Table 3).

References

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. Annu Rev Ecol Evol Syst. 41(1):379–406.

Biesecker LG, Spinner NB. 2013. A genomic view of mosaicism and human disease. Nat Rev Genet. 14(5):307–20.

Boklage CE. 1990. Survival probability of human conceptions from fertilization to term. Int J Fertil. 35(2):75, 79–94.

Bolton H, Graham SJL, Van der Aa N, Kumar P, Theunis K, Fernandez Gallardo E, Voet T, Zernicka-Goetz M. 2016. Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. Nat Commun. 7:11165.

Capalbo A, Ubaldi FM, Rienzi L, Scott R, Treff N. 2017. Detecting mosaicism in trophectoderm biopsies: Current challenges and future possibilities. Hum Reprod. 32(3):492–8.

Capalbo A, Poli M, Rienzi L, Girardi L, Patassini C, Fabiani M, Cimadomo D, Benini F, Farcomeni A, Cuzzi J, et al. 2021. Mosaic human preimplantation embryos and their developmental potential in a prospective, non-selection clinical trial. Am J Hum Genet. 108(12):2238–47.

Cascante SD, Besser A, Lee H, Wang F, McCaffrey C, Grifo JA. 2023. Blinded rebiopsy and analysis of noneuploid embryos with 2 distinct preimplantation genetic testing platforms for aneuploidy. Fertil Steril. 120(6):1161–9.

Chatzimeletiou K, Morrison EE, Prapas N, Prapas Y, Handyside AH. 2005. Spindle abnormalities in normally developing and arrested human preimplantation embryos in vitro identified by confocal laser scanning microscopy. Hum Reprod. 20(4):1120.

Chavli EA, Klaasen SJ, Van Opstal D, Laven JS, Kops GJ, Baart EB. 2024. Single-cell DNA sequencing reveals a high incidence of chromosomal abnormalities in human blastocysts. J Clin Invest. 134(6).

Cimadomo D, Rienzi L, Capalbo A, Rubio C, Innocenti F, Garcia-Pascual CM, Ubaldi FM, Handyside A. 2020. The dawn of the future: 30 years from the first biopsy of a human embryo. the detailed history of an ongoing revolution. Hum Reprod Update. 26(4):453–73.

Clarke EA, Dahiya AK, Cascante SD, Blakemore JK. 2023. Evaluation of pre-implantation genetic testing for aneuploidy outcomes in patients without infertility undergoing in vitro fertilization compared to infertile controls. J Assist Reprod Genet. 40(11):2697–704.

Cornelisse S, Zagers M, Kostova E, Fleischer K, van Wely M, Mastenbroek S. 2020. Preimplantation genetic testing for aneuploidies (abnormal number of chromosomes) in in vitro fertilisation. Cochrane Database Syst Rev. 9(9):CD005291.

Currie CE, Ford E, Benham Whyte L, Taylor DM, Mihalas BP, Erent M, Marston AL, Hartshorne GM, McAinsh AD. 2022. The first mitotic division of human embryos is highly error prone. Nat Commun. 13(1):6755–6.

Esfandiari N, Bunnell ME, Casper RF. 2016. Human embryo mosaicism: Did we drop the ball on chromosomal testing? J Assist Reprod Genet. 33(11):1439–44.

ESHRE Working Group on Chromosomal Mosaicism, De Rycke M, Capalbo A, Coonen E, Coticchio G, Fiorentino F, Goossens V, Mcheik S, Rubio C, Sermon K, et al. 2022. ESHRE survey results and good practice recommendations on managing chromosomal mosaicism. Hum Reprod Open. 2022(4):hoac044.

Forman EJ, Upham KM, Cheng M, Zhao T, Hong KH, Treff NR, Scott RTJ. 2013. Comprehensive chromosome screening alters traditional morphology-based embryo selection: A prospective study of 100 consecutive cycles of planned fresh euploid blastocyst transfer. Fertil Steril. 100(3):718–24.

Gallardo EF, Sifrim A, Chappell J, Demeulemeester J, Herrmann JC, Vermote R, Kerremans A, Michiel Van Der Haegen, Jens Van Herck, Vanuytven S, et al. 2023. A multi-omics

genome-and-transcriptome single-cell atlas of human preimplantation embryogenesis reveals the cellular and molecular impact of chromosome instability. bioRxiv: 2023.03.08.530586v1. https://www.biorxiv.org/content/10.1101/2023.03.08.530586v1

Girardi L, Serdarogullari M, Patassini C, Poli M, Fabiani M, Caroselli S, Coban O, Findikli N, Boynukalin FK, Bahceci M, et al. 2020. Incidence, origin, and predictive model for the detection and clinical management of segmental aneuploidies in human embryos. Am J Hum Genet. 106(4):525–34.

Greco E, Minasi MG, Fiorentino F. 2015b. Healthy babies after intrauterine transfer of mosaic aneuploid blastocysts. N Engl J Med. 373(21):2089–90.

Greco E, Yakovlev P, Kornilov N, Vyatkina S, Bogdanova D, Ermakova M, Tarasova Y, Tikhonov A, Pendina A, Biricik A, et al. 2023. Two clinical case reports of embryonic mosaicism identified with PGT-A persisting during pregnancy as true fetal mosaicism. Hum Reprod. 38(2):315–23.

Gruhn JR, Zielinska AP, Shukla V, Blanshard R, Capalbo A, Cimadomo D, Nikiforov D, Chan AC, Newnham LJ, Vogel I, et al. 2019. Chromosome errors in human eggs shape natural fertility over reproductive life span. Science. 365(6460):1466–9.

Hassold T, Hunt P. 2001. To err (meiotically) is human: The genesis of human aneuploidy. Nat Rev Genet. 2(4):280–91.

Hassold T, Hall H, Hunt P. 2007. The origin of human aneuploidy: Where we have been, where we are going. Hum Mol Genet. 16(R2):R203–8.

Jabot F, Faure T, Dumoulin N, O'Hara RB. 2013. EasyABC: Performing efficient approximate Bayesian computation sampling schemes using R. Methods Ecol Evol. 4(7):684–7.

Jarvis GE. 2016. Early embryo mortality in natural human reproduction: What the data say. F1000Res. 5:2765.

Kim J, Tao X, Cheng M, Steward A, Guo V, Zhan Y, Scott RTJ, Jalas C. 2022. The concordance rates of an initial trophectoderm biopsy with the rest of the embryo using PGTseq, a targeted next-generation sequencing platform for preimplantation genetic testing-aneuploidy. Fertil Steril. 117(2):315–23.

Lenormand M, Jabot F, Deffuant G. 2013. Adaptive approximate Bayesian computation for complex models. Comput Stat. 28(6):2777–96.

Macklon NS, Geraedts JPM, Fauser BCJM. 2002. Conception to ongoing pregnancy: The 'black box' of early pregnancy loss. Hum Reprod Update. 8(4):333–43.

Marin D, Xu J, Treff NR. 2021. Preimplantation genetic testing for aneuploidy: A review of published blastocyst reanalysis concordance data. Prenat Diagn. 41(5):545–53.

Mastenbroek S, Twisk M, van Echten-Arends J, Sikkema-Raddatz B, Korevaar JC, Verhoeve HR, Vogel NEA, Arts EGJM, de Vries JWA, Bossuyt PM, et al. 2007. In vitro fertilization with preimplantation genetic screening. N Engl J Med. 357(1):9–17.

Maurer M, Ebner T, Puchner M, Mayer RB, Shebl O, Oppelt P, Duba H. 2015. Chromosomal aneuploidies and early embryonic developmental arrest. Int J Fertil Steril. 9(3):346–53.

McCoy RC. 2017. Mosaicism in preimplantation human embryos: When chromosomal abnormalities are the norm. Trends Genet. 33(7):448–63.

McCoy RC, Summers MC, McCollin A, Ottolini CS, Ahuja K, Handyside AH. 2023. Meiotic and mitotic aneuploidies drive arrest of in vitro fertilized human preimplantation embryos. Genome Med. 15(1):77–1.

McCoy RC, Demko ZP, Ryan A, Banjevic M, Hill M, Sigurjonsson S, Rabinowitz M, Petrov DA. 2015. Evidence of selection against complex mitotic-origin aneuploidy during preimplantation development. PLoS Genet. 11(10).

McDole K, Guignard L, Amat F, Berger A, Malandain G, Royer LA, Turaga SC, Branson K, Keller PJ. 2018. In toto imaging and reconstruction of post-implantation mouse development at the single-cell level. Cell. 175(3):859,876.e33.

Munne S, Cohen J, Simpson JL, Wilton LJ, Handyside AH, Thornhill AR, Mastenbroek S, van der Veen F, Repping S. 2007. In vitro fertilization with preimplantation genetic screening. N Engl J Med. 357(17):1769–71.

Munné S, Blazek J, Large M, Martinez-Ortiz PA, Nisson H, Liu E, Tarozzi N, Borini A, Becker A, Zhang J, et al. 2017. Detailed investigation into the cytogenetic constitution and pregnancy outcome of replacing mosaic blastocysts detected with the use of high-resolution next-generation sequencing. Fertil Steril. 108(1):62,71.e8.

Munné S, Kaplan B, Frattarelli JL, Child T, Nakhuda G, Shamma FN, Silverberg K, Kalista T, Handyside AH, Katz-Jaffe M, et al. 2019. Preimplantation genetic testing for aneuploidy versus morphology as selection criteria for single frozen-thawed embryo transfer in good-prognosis patients: A multicenter randomized clinical trial. Fertil Steril. 112(6):1071,1079.e7.

Navratil R, Horak J, Hornak M, Kubicek D, Balcova M, Tauwinklova G, Travnik P, Vesela K. 2020. Concordance of various chromosomal errors among different parts of the embryo and the value of re-biopsy in embryos with segmental aneuploidies. Mol Hum Reprod. 26(4):269–76.

Popovic M, Dhaenens L, Boel A, Menten B, Heindryckx B. 2020. Chromosomal mosaicism in human blastocysts: The ultimate diagnostic dilemma. Hum Reprod Update. 26(3):313–34.

Rodrigo L, Clemente-Císcar M, Campos-Galindo I, Peinado V, Simón C, Rubio C. 2020. Characteristics of the IVF cycle that contribute to the incidence of mosaicism. Genes. 11(10):1151.

Rubio C, Ph.D, Bellver J, M.D, Rodrigo L, Ph.D, Castillón G, M.D, Guillén A, M.D, Vidal C, M.D, Giles J, M.D, Ferrando M, M.D, Cabanillas S, M.D, Remohí, José, M.D., Ph.D, et al. 2017. In vitro fertilization with preimplantation genetic diagnosis for aneuploidies in advanced maternal age: A randomized, controlled study. Fertil Steril. 107(5):1122–9.

Santos MA, Teklenburg G, Macklon NS, Van Opstal D, Schuring-Blom GH, Krijtenburg P, de

Vreeden-Elbertse J, Fauser BC, Baart EB. 2010. The fate of the mosaic embryo: Chromosomal constitution and development of day 4, 5 and 8 human embryos. Hum Reprod. 25(8):1916–26.

Singla S, Iwamoto-Stohl LK, Zhu M, Zernicka-Goetz M. 2020. Autophagy-mediated apoptosis eliminates aneuploid cells in a mouse model of chromosome mosaicism. Nat Commun. 11(1):2958–3.

Skinner BM, Viotti M, Griffin DK, Ellis PJI. 2024. Explaining the counter-intuitive effectiveness of trophectoderm biopsy for PGT-A using computational modelling. eLife. 13:RP94506.

Starostik MR, Sosina OA, McCoy RC. 2020. Single-cell analysis of human embryos reveals diverse patterns of aneuploidy and mosaicism. Genome Res. 30(6):814–25.

Sunnaker M, Busetto AG, Numminen E, Corander J, Foll M, Dessimoz C. 2013. Approximate Bayesian computation. PLoS Comput Biol. 9(1):e1002803.

Treff NR, Marin D. 2021. The “mosaic” embryo: Misconceptions and misinterpretations in preimplantation genetic testing for aneuploidy. Fertil Steril. 116(5):1205–11.

Vera-Rodriguez M, Rubio C. 2017. Assessing the true incidence of mosaicism in preimplantation embryos. Fertil Steril. 107(5):1107–12.

Vermeesch JR, Voet T, Devriendt K. 2016. Prenatal and pre-implantation genetic diagnosis. Nat Rev Genet. 17(10):643–56.

Verpoest W, Staessen C, Bossuyt PM, Goossens V, Altarescu G, Bonduelle M, Devesa M, Eldar-Geva T, Gianaroli L, Griesinger G, et al. 2018. Preimplantation genetic testing for aneuploidy by microarray analysis of polar bodies in advanced maternal age: A randomized clinical trial. Hum Reprod. 33(9):1767–76.

Victor AR, Tyndall JC, Brake AJ, Lepkowsky LT, Murphy AE, Griffin DK, McCoy RC, Barnes FL, Zouves CG, Viotti M. 2019. One hundred mosaic embryos transferred prospectively in a single clinic: Exploring when and why they result in healthy pregnancies. Fertil Steril. 111(2):280–93.

Victor AR, Griffin DK, Brake AJ, Tyndall JC, Murphy AE, Lepkowsky LT, Lal A, Zouves CG, Barnes FL, McCoy RC, et al. 2019. Assessment of aneuploidy concordance between clinical trophectoderm biopsy and blastocyst. Hum Reprod. 34(1):181–92.

Viotti M, McCoy RC, Griffin DK, Spinella F, Greco E, Madjunkov M, Madjunkova S, Librach CL, Victor AR, Barnes FL, et al. 2021. Let the data do the talking: The need to consider mosaicism during embryo selection. Fertil Steril. 116(5):1212–9.

Viotti M, Victor AR, Barnes FL, Zouves CG, Besser AG, Grifo JA, Cheng E, Lee M, Horcajadas JA, Corti L, et al. 2021. Using outcome data from one thousand mosaic embryo transfers to formulate an embryo ranking system for clinical use. Fertil Steril. 115(5):1212–24.

Yang M, Rito T, Metzger J, Naftaly J, Soman R, Hu J, Albertini DF, Barad DH, Brivanlou AH, Gleicher N. 2021. Depletion of aneuploid cells in human embryos and gastruloids. Nat Cell Biol. 23(4):314–21.

Zhai F, Kong S, Song S, Guo Q, Ding L, Zhang J, Wang N, Kuo Y, Guan S, Yuan P, et al. 2024. Human embryos harbor complex mosaicism with broad presence of aneuploid cells during early development. Cell Discov. 10(1):98–9.