Published in final edited form as:

Science. 2015 April 10; 348(6231): 235–238. doi:10.1126/science.aaa3337.

Common variants spanning *PLK4* are associated with mitotic-origin aneuploidy in human embryos

Rajiv C. McCoy¹, Zachary Demko², Allison Ryan², Milena Banjevic², Matthew Hill², Styrmir Sigurjonsson², Matthew Rabinowitz², Hunter B. Fraser¹, and Dmitri A. Petrov¹

¹Department of Biology, Stanford University, Stanford, California, USA

²Natera, Inc., San Carlos, California, USA

Abstract

Aneuploidy, the inheritance of an atypical chromosome complement, is common in early human development and is the primary cause of pregnancy loss. By screening day-3 embryos during *in vitro* fertilization cycles, we identified an association between aneuploidy of putative mitotic origin and linked genetic variants on chromosome 4 of maternal genomes. This associated region contains a candidate gene, Polo-like kinase 4 (*PLK4*), which plays a well-characterized role in centriole duplication and has the ability to alter mitotic fidelity upon minor dysregulation. Mothers with the high-risk genotypes contributed fewer embryos for testing at day 5, suggesting that their embryos are less likely to survive to blastulation. The associated region coincides with a signature of a selective sweep in ancient humans, suggesting that the causal variant was either the target of selection or hitchhiked to substantial frequency.

Deviation from a balanced chromosome complement, a phenomenon known as aneuploidy, is common in early human embryos, and often leads to embryonic mortality (1).

Approximately 75% of embryos are at least partially aneuploid by day 3, due to prevalent errors of both meiotic and post-zygotic origin (2, 3), and this proportion increases with maternal age (1). The propensity to produce aneuploid embryos varies substantially, however, even among mothers of a similar age (4). We therefore hypothesized that variation

Correspondence regarding the data should be addressed to Z.D. (zdemko@natera.com), while other correspondence should be addressed to R.C.M. (rmccoy@stanford.edu) or D.A.P. (dpetrov@stanford.edu).

Author Contributions

D.A.P. devised the project. D.A.P., H.B.F. and Z.D. provided guidance throughout the project. R.C.M. performed analyses and wrote the manuscript. Z.D., A.R., M.B., M.H., S.S., and M.R. helped generate the data and design the algorithm by which aneuploidies are detected. All authors read the manuscript and provided comments.

D.A.P. has received stock options in Natera, Inc. as consulting fees. Z.D., A.R., M.B., M.H., S.S., and M.R. are full time employees and hold stock or options to hold stock in Natera, Inc. De-identified aneuploidy outcome data for all blastomeres are uploaded as supplemental material, as are GWAS summary statistics. Patient genotype data for the associated region are archived at Natera, which will cooperate with qualified researchers attempting to replicate the findings of this study, conditional on IRB approval and adherence to a Material Transfer Agreement. Stanford University filed a provisional patent related to this work with the USPTO on Nov. 14, 2014 (USSN 62/080,251).

Supplementary Materials

Materials and Methods

Supplementary Text

Figures S1–S5

Table S1

References (27–47)

in parents' genomes may explain variation in aneuploidy incidence. We tested this hypothesis by performing a genome-wide association study of aneuploidy risk among patients undergoing preimplantation genetic screening (PGS) of embryos collected from *in vitro* fertilization (IVF) cycles.

Embryo (single-cell day-3 blastomere biopsies or multi-cell day-5 trophectoderm biopsies) and parent DNA was genotyped on a single nucleotide polymorphism (SNP) microarray (5). The Parental Support™ algorithm (6) was then applied to determine chromosome-level ploidy status of each embryo sample. This algorithm overcomes high rates of allelic dropout and other quality limitations of whole-genome amplification by supplementing these data with high-quality genotypes from parental chromosomes. The copy number of each embryonic chromosome can then be inferred by comparing microarray channel intensities from DNA amplified from the embryo biopsy to those expected given the parental genotypes at each marker. Combining these fine-scale observations across large chromosomal windows facilitates detection of particular forms of aneuploidy and assignment of copy number variations to specific parental homologs (6). Previous validation has been performed for individual blastomeres (6), so it is unknown how accuracy would be affected in the face of chromosomal mosaicism that could potentially affect multi-cell trophectoderm biopsies. We therefore performed an association study on 2,362 unrelated mothers (1,956 IVF patients and 406 oocyte donors) and 2,360 unrelated fathers meeting genotype quality-control thresholds (5) and from whom at least one day-3 biopsy was obtained with the blastomere providing a high-confidence result (a total of 20,798 blastomeres). We then separately analyzed the additional 15,388 trophectoderm biopsies to gain insight into selection occurring prior to this developmental stage.

We first tested for associations between the rates of errors of putative maternal meiotic origin (Fig. S1) (5) and maternal genotypes, identifying no association achieving genome-wide significance (Logistic GLM, P-value threshold = 5×10^{-8}). We next tested for associations between the rates of errors of putative mitotic origin and parental genotypes. The first mitotic divisions of the developing embryo take place under the control of maternal gene products provided to the oocyte (7) and are substantially error-prone (2, 3). We hypothesized that variation in maternal gene products may thus contribute to variation in rates of post-zygotic error among embryos from different mothers. To encode the mitotic error phenotype, we designated all blastomeres with aneuploidies affecting a paternal chromosome copy (excluding paternal trisomies of putative meiotic origin) as cases, and all other blastomere samples as controls (Fig. 1A). Because aneuploidy has been estimated to affect fewer than 5% of sperm (8) and because paternal meiotic trisomies were detected for fewer than 1% of blastomeres in our data, this set of aneuploid cases should be nearly exclusively mitotic in origin.

The 5,438 putative mitotic-origin aneuploidies were predominantly characterized by a distinct error profile involving multiple chromosome losses (Fig. 1B, C), and their incidence was not associated with maternal age (Fig. 1D). This excess of chromosome losses is consistent with previous studies that identified anaphase lag as the primary mechanism contributing to mosaicism in preimplantation embryos (9, 10). Anaphase lag refers to the delayed migration of a chromosome during anaphase such that the lagging chromosome fails

to be incorporated into the reforming nucleus, resulting in chromosome loss with no corresponding chromosome gain (Fig. 1A). This error commonly arises as a consequence of merotelic kinetochore attachment: the attachment of a single kinetochore to microtubules emanating from both spindle poles (11). Merotelic attachment can in turn occur due to the presence of extra centrosomes or other centrosome aberrations (12, 13).

From our genome-wide analysis, we identified a peak on chromosome 4, regions q28.1–q28.2, of maternal genomes associated with this mitotic-error phenotype (Fig. 2C–E). The SNP rs2305957 was most strongly associated, with the minor allele conferring a significantly increased rate of mitotic error (Logistic GLM, $\beta = 0.218$, $SE = 0.0270$, $P = 8.68 \times 10^{-16}$). The minor allele is present in diverse human populations at frequencies of 20%–45% (Fig. S2) (14). Importantly, we observed no significant associations between paternal genotype and the same mitotic-error phenotype (Logistic GLM, $P = 0.389$), which demonstrates that population stratification did not drive the significant association with maternal genotype (Fig. 2A, B) (5). We also found that the association was robust when separately tested for mothers of European and East Asian ancestries (Table 1; Fig. S3).

The observed effect was characterized by means of 24.6%, 27.0%, and 31.7% of blastomeres affected with paternal-chromosome aneuploidies for the ‘GG’, ‘AG’, and ‘AA’ maternal genotypic classes, respectively (Fig. 3A), and was consistent across age classes (Fig. 3D). We additionally note that the effect size from individual blastomeres may underestimate the overall effect on aneuploidy, as diploid blastomeres will be sampled by chance from some diploid-aneuploid mosaics. The frequencies of the three genotypes were not significantly different between mothers and fathers ($\chi^2(2, N = 9, 418) = 1.17$, $P = 0.557$) or between egg-donors and non-donors ($\chi^2(2, N = 4, 712) = 2.49$, $P = 0.288$), together suggesting that this set of IVF patients is not enriched for the mitotic-error-associated genotypes.

For validation, genotypes from 34 additional unrelated mothers, representing new cases since the initial database pull, were tested for association with the same phenotype. Despite the small sample size ($N_{\text{patients}} = 34$, $N_{\text{blastomeres}} = 283$), the association was replicated, with 25.3%, 35.7%, and 51.3% of blastomeres with errors affecting paternal chromosomes among the three respective maternal genotypic classes (Logistic GLM, $\beta = 0.589$, $SE = 0.219$, $P = 0.0112$; Fig. 3B).

Highlighting its importance, genotype at rs2305957 was also a significant predictor of overall aneuploidy (Logistic GLM, $\beta = 0.139$, $SE = 0.0271$, $P = 3.05 \times 10^{-7}$; Fig. 3E), especially for complex aneuploidies affecting greater than two chromosomes (Logistic GLM, $\beta = 0.234$, $SE = 0.0329$, $P = 1.72 \times 10^{-12}$; Fig. S4). Means of 65.2%, 68.3%, and 71.4% of blastomeres per case were determined to be aneuploid for mothers with the ‘GG’, ‘AG’, and ‘AA’ genotypes, respectively. This 6.2% difference in proportion of aneuploid blastomeres between the two homozygous maternal genotype classes is roughly equivalent to the average effect of 1.8 years of age for mothers 35 years old (Fig. S5).

Given that the association in our study was driven by complex aneuploidies affecting many chromosomes, and that complex and mosaic aneuploidies are more likely to be inviable (15),

we hypothesized that arrest of aneuploid embryos would bias the genotypic ratios at associated SNPs for 15,388 embryos sampled at the day-5 blastocyst stage from 2,998 unrelated mothers. Patients with the mitotic error-associated genotypes at rs2305957 contributed significantly fewer trophoctoderm biopsies for testing (Poisson GLM, $\beta = -0.0619$, $SE = 0.0204$, $P = 0.00247$; Fig. 3C), consistent with an increased proportion of inviable aneuploidies. Together, these findings suggest that the mitotic-error association may affect fertility such that it may take longer, on average, for women with the associated genotypes to achieve successful pregnancies.

In order to characterize the extent of the associated region, we performed genotype imputation for a subset of 1,332 patients of European ancestry (5). The associated haplotype lies in a region of low recombination and spans over 600 Kbp of chromosome 4, regions q28.1–q28.2 (Fig. 2E), including the genes *INTU*, *SLC25A31*, *HSPA4L*, *PLK4*, *MFSD8*, *LARPIB*, and *PGRMC2*. While none of these candidates can yet be ruled out, we focused on *PLK4* on the basis of its well-characterized role as the master regulator of centriole duplication, a key component of the centrosome cycle (16,17). In addition, it was recently demonstrated that *PLK4* is essential for mediating bipolar spindle formation during the first cell divisions in mouse embryos, which take place in the absence of centrioles (18).

Due in part to the observation that centrosome aberrations and aneuploidies are common in human cancers, the role of *PLK4* and its orthologs in mediating the centrosome cycle has been investigated in several model systems. *PLK4* is a tightly regulated, low-abundance kinase with a short half-life (19). Overexpression of *PLK4* results in centriole overduplication, thereby increasing the frequency of multipolar spindle formation and subsequent anaphase lag (12). Reduced expression of *PLK4* results in centriole loss (17), which also leads to multipolar spindle formation, as well as the formation of monopolar spindles. Both up- and down-regulation of *PLK4* therefore have the potential to induce chromosome instability, and altered *PLK4* expression is commonly observed in several forms of cancer, consistent with a tumor-suppressor function (20, 21).

Along with hundreds of variants upstream and downstream of *PLK4*, the associated region contains two nonsynonymous SNPs within the *PLK4* coding sequence: rs3811740 (S232T) and rs17012739 (E830D), the former occurring in the protein's kinase domain and the latter occurring in the cryptic Polo-box domain (22). Neither site exhibits strong conservation over deep evolutionary time, and both SNPs were predicted as benign on the basis of sequence conservation, amino acid similarity, and mapping to three-dimensional protein structure (5).

Prompted by the observation that the minor allele of rs2305957 is derived and segregates at intermediate frequencies in diverse human populations, yet is absent from Neanderthal and Denisovan genomes (5), we investigated whether the region showed evidence of positive selection in humans. Unfortunately, classic frequency spectrum-based tests have sensitivity over the order of N_e generations, capturing only relatively recent human evolutionary history (~10,000 generations). We thus examined results of the selection scan from (23), which has resolution to detect signatures of ancient selective sweeps in the human lineage by identifying regions of aligned Neanderthal genomes that are deficient in high-frequency human derived alleles. The mitotic-error associated region identified in our study is among

the 212 previously-identified regions displaying such a signature (23). This finding suggests that either this seemingly deleterious allele hitchhiked with a linked adaptive variant or that the causal variant was adaptive in a context that is not currently understood.

The fact that the haplotype bearing the derived allele did not sweep to fixation and is present at similar frequencies across human populations is consistent with the action of long-term balancing selection. We speculate that the mitotic error phenotype may be maintained by conferring both a deleterious effect on maternal fecundity and a possible beneficial effect of obscured paternity via a reduction in the probability of successful pregnancy per intercourse. This hypothesis is based on the fact that humans possess a suite of traits (e.g. concealed ovulation, constant receptivity) that obscure paternity and may have evolved to increase paternal investment in offspring (24). Such a scenario could result in balancing selection by rewarding evolutionary ‘free riders’ who do not possess the risk allele—and thus do not suffer fecundity costs—but benefit from paternity confusion in the population as a whole.

In summary, mitotic fidelity is affected by variation in maternal gene products controlling the initial cell divisions of preimplantation embryos. This finding is important in the context of IVF where selection of euploid embryos may improve the success rate of implantation and ongoing pregnancy (25). More broadly, factors influencing variation in rates of aneuploidy may help explain variation in fertility status among the general population. Fewer than ~30% of conceptions result in successful pregnancy, mostly due to high rates of inviable aneuploidy in early development (26). By altering this rate, the associated locus described in our study likely influences the average time required to achieve successful pregnancy, which could be especially important for couples with already-reduced fertility. The identification of genetic variation influencing rates of aneuploidy is an important step in the understanding of aneuploidy risk and may assist the future development of diagnostic or therapeutic technologies targeting certain forms of infertility.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Thanks to C. Hin and J. Layne for computing support, E. Sharon for advice regarding analyses, R. Taylor and other members of the Petrov lab for helpful input, J. Sage, J. Arand, T. Stearns, and O. Cormier for advice regarding functional annotation, and C. Boggs for comments on the manuscript.

References and Notes

1. Hassold T, Hunt P. To err (meiotically) is human: the genesis of human aneuploidy. *Nature Rev Genet.* 2001; 2:280–291. [PubMed: 11283700]
2. Voullaire L, Slater H, Williamson R, Wilton L. Chromosome analysis of blastomeres from human embryos by using comparative genomic hybridization. *Hum Genet.* 2000; 106:210–217. [PubMed: 10746563]
3. Wells D, Delhanty JD. Comprehensive chromosomal analysis of human preimplantation embryos using whole genome amplification and single cell comparative genomic hybridization. *Mol Hum Reprod.* 2000; 6:1055–1062. [PubMed: 11044470]

4. Delhanty JD, Harper JC, Ao A, Handyside AH, Winston RM. Multicolour FISH detects frequent chromosomal mosaicism and chaotic division in normal preimplantation embryos from fertile patients. *Hum Genet.* 1997; 99:755–760. [PubMed: 9187668]
5. Materials and methods are available as supplementary materials on *Science* Online.
6. Johnson D, et al. Preclinical validation of a microarray method for full molecular karyotyping of blastomeres in a 24-h protocol. *Hum Reprod.* 2010; 25:1066–1075. [PubMed: 20100701]
7. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development.* 2009; 136:3033–3042. [PubMed: 19700615]
8. Templado C, Vidal F, Estop A. Aneuploidy in human spermatozoa. *Cytogenet Genome Res.* 2011; 133:91–99. [PubMed: 21282942]
9. Coonen E, et al. Anaphase lagging mainly explains chromosomal mosaicism in human preimplantation embryos. *Hum Reprod.* 2004; 19:316–324. [PubMed: 14747173]
10. Daphnis D, et al. Detailed FISH analysis of day 5 human embryos reveals the mechanisms leading to mosaic aneuploidy. *Hum Reprod.* 2005; 20:129–137. [PubMed: 15567885]
11. Gregan J, Polakova S, Zhang L, Toli -NØrrelykke IM, Cimini D. Merotelic kinetochore attachment: causes and effects. *Trends Cell Biol.* 2011; 21:374–381. [PubMed: 21306900]
12. Ganem NJ, Godinho SA, Pellman D. A mechanism linking extra centrosomes to chromosomal instability. *Nature.* 2009; 460:278–282. [PubMed: 19506557]
13. Holland AJ, Cleveland DW. Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. *Nature Rev Mol Cell Biol.* 2009; 10:478–487. [PubMed: 19546858]
14. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature.* 2010; 467:1061–1073. [PubMed: 20981092]
15. Vega M, Breborowicz A, Moshier EL, McGovern PG, Keltz MD. Blastulation rates decline in a linear fashion from euploid to aneuploid embryos with single versus multiple chromosomal errors. *Fertil Steril.* 2014; 102:394–398. [PubMed: 24890269]
16. Habedanck R, Stierhof YD, Wilkinson CJ, Nigg EA. The polo kinase *Plk4* functions in centriole duplication. *Nature Cell Biol.* 2005; 7:1140–1146. [PubMed: 16244668]
17. Bettencourt-Dias M, et al. SAK/PLK4 is required for centriole duplication and flagella development. *Curr Biol.* 2005; 15:2199–2207. [PubMed: 16326102]
18. Coelho PA, et al. Spindle formation in the mouse embryo requires *Plk4* in the absence of centrioles. *Dev Cell.* 2013; 27:586–597. [PubMed: 24268700]
19. Firat-Karalar EN, Stearns T. The centriole duplication cycle. *Philos Trans R Soc London [Biol].* 2014; 369:20130460.
20. Ko MA, et al. Plk4 haploinsufficiency causes mitotic infidelity and carcinogenesis. *Nature Genet.* 2005; 37:883–888. [PubMed: 16025114]
21. Rosario CO, et al. Plk4 is required for cytokinesis and maintenance of chromosomal stability. *Proc Natl Acad Sci USA.* 2010; 107:6888–6893. [PubMed: 20348415]
22. Sillibourne JE, Bornens M. Polo-like kinase 4: the odd one out of the family. *Cell Div.* 2010; 5:1–9. [PubMed: 20157424]
23. Green RE, et al. A draft sequence of the Neandertal genome. *Science.* 2010; 328:710–722. [PubMed: 20448178]
24. Alexander, RD., Noonan, KM. *Evolutionary Biology and Human Social Organization.* Chagnon, NA., Irons, WG., editors. Duxbury Press; North Scituate: 1979. p. 436–453.
25. Scott RT Jr, et al. Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial. *Fertil Steril.* 2013; 100:697–703. [PubMed: 23731996]
26. Wilcox AJ, et al. Incidence of early loss of pregnancy. *N Engl J Med.* 1988; 319:189–194. [PubMed: 3393170]
27. Micallef L, Rodgers P. eulerAPE: Drawing area-proportional 3-Venn diagrams using ellipses. *PLoS ONE.* 2014; 9:e101717. [PubMed: 25032825]
28. Turner, SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv. (available at <http://biorxiv.org/content/early/2014/05/14/005165>)
29. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer; New York: 2009.

30. Manichaikul A, et al. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010; 26:2867–2873. [PubMed: 20926424]
31. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81:559–575. [PubMed: 17701901]
32. Turner S, et al. Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet*. 2011; 68:1–19.
33. Rabinowitz M, et al. Origins and rates of aneuploidy in human blastomeres. *Fertil Steril*. 2012; 97:395–401. [PubMed: 22195772]
34. Simerly C, et al. The paternal inheritance of the centrosome, the cell's microtubule-organizing center, in humans, and the implications for infertility. *Nature Med*. 1995; 1:47–52. [PubMed: 7584952]
35. Mantikou E, Wong KM, Repping S, Mastenbroek S. Molecular origin of mitotic aneuploidies in preimplantation embryos. *Biochim Biophys Acta*. 2012; 1822:1921–1930. [PubMed: 22771499]
36. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999; 55:997–1004. [PubMed: 11315092]
37. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009; 5:e1000529. [PubMed: 19543373]
38. Johnson AD, et al. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008; 24:2938–2939. [PubMed: 18974171]
39. Chelala C, Khan A, Lemoine NR. SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics*. 2009; 25:655–661. [PubMed: 19098027]
40. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protoc*. 2009; 4:1073–1081. [PubMed: 19561590]
41. Adzhubei IA, et al. A method and server for predicting damaging missense mutations. *Nature Methods*. 2010; 7:248–249. [PubMed: 20354512]
42. Simons YB, Turchin MC, Pritchard JK, Sella G. The deleterious mutation load is insensitive to recent population history. *Nature Genet*. 2014; 46:220–224. [PubMed: 24509481]
43. Castellano S, et al. Patterns of coding variation in the complete exomes of three Neandertals. *Proc Natl Acad Sci USA*. 2014; 111:6666–6671. [PubMed: 24753607]
44. Prüfer K, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. 2014; 505:43–49. [PubMed: 24352235]
45. Meyer M, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 2012; 338:222–226. [PubMed: 22936568]
46. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol*. 2001; 60:155–166. [PubMed: 11855950]
47. International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010; 467:52–58. [PubMed: 20811451]

Figure 1. Mitotic error phenotypes

A: Two mechanisms that frequently contribute to aneuploidy are depicted: maternal meiotic non-disjunction and mitotic anaphase lag. **B:** Aneuploidies where at least one paternal chromosome is affected are likely mitotic in origin and include an excess of chromosome losses compared to chromosome gains, consistent with the signature of anaphase lag. Paternal chromosome loss (paternal monosomy) commonly co-occurs with other forms of chromosome loss including maternal monosomy and nullisomy. **C:** Blastomeres with aneuploidies affecting at least one paternal chromosome (blue; putative mitotic-origin aneuploidies) often contain multiple aneuploid chromosomes in contrast to aneuploid blastomeres in which no paternal chromosome copies are affected (red; predominantly meiotic-origin aneuploidies). Heights of bars indicate densities (i.e., relative frequencies). **D:** Aneuploidies in which at least one paternal chromosome copy is affected do not increase in frequency with increasing maternal age, while maternal aneuploidies increase in frequency beginning in the mid-thirties. Error bars indicate standard errors of proportions.

Figure 2. Association test for aneuploidy

A–D: Manhattan and QQ plots depicting P-values of association tests of each genotyped SNP versus the rate of aneuploidy affecting paternal chromosomes (a proxy for mitotic aneuploidy). P-values are corrected using the genomic control method (5). **A:** Results for association with paternal genotypes, a negative control. **B:** QQ plot of the distribution of observed P-values versus those expected under the null. **C:** Association with maternal genotypes, with rs2305957 highlighted as the most significant genotyped SNP. **D:** QQ plots of P-values. For A & C, the red lines represent a standard genome-wide cutoff of 5×10^{-8} , while the gray dotted lines represent a less stringent P-value of 1×10^{-6} . For B & D, the gray shaded regions indicate probability bounds. **E:** Regional association plot for mothers of European ancestry, inferred by comparison to reference populations (Fig. S1). rs2305957 is indicated (purple point), whereas the colors of other variants represent linkage disequilibrium with rs2305957 (5).

Figure 3. Effect of genotype on mitotic-error-related phenotypes

For boxplots, we restricted figures to include only mothers for whom >2 embryos were tested. **A:** The proportion of blastomeres per mother with an error affecting a paternal chromosome (a proxy for mitotic aneuploidy) stratified by maternal genotype at rs2305957 for the discovery sample ($N_{\text{patients}} = 2,362$, $N_{\text{embryos}} = 20,798$; $P = 8.68 \times 10^{-16}$). **B:** The same phenotype as panel A, replicated in the validation sample ($N_{\text{patients}} = 34$, $N_{\text{embryos}} = 283$; $P = 0.0112$). **C:** Mean number of day-5 trophectoderm biopsies per mother, stratified by genotype at rs2305957 (Poisson GLM, $P = 0.00247$). Error bars represent standard error. **D:** The mean proportion of blastomeres with an aneuploidy affecting a paternal chromosome versus maternal age, stratified by genotype at rs2305957. Error bars represent standard error of the proportion. **E:** The mean proportion of aneuploid blastomeres versus maternal age, stratified by genotype at rs2305957. Error bars represent standard error of the proportion.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1
Association of SNP rs2305957 with the rate putative mitotic-origin aneuploidy

Sample size, regression coefficient (β), standard error (SE), odds-ratio (OR), genomic inflation factor (λ), and P-values are reported. The upper section gives results of association test of all female patients, including those falling outside of the European and East Asian principal component boundaries. The lower two rows control for potential population stratification by separating analyses of female patients with a high proportion of European and East Asian ancestry, respectively.

Sample size			Uncorrected			Genomic control		
	Patients	Embryos	β	SE	OR (95% CI)	λ	P	P
Discovery	2,362	20,798	0.218	0.0270	1.244 (1.179–1.311)	1.059	8.68×10^{-16}	5.99×10^{-15}
Europe	1,332	11,861	0.214	0.0353	1.238 (1.155–1.327)	1.066	1.91×10^{-9}	6.67×10^{-9}
East Asia	259	2,222	0.280	0.0788	1.323 (1.133–1.543)	1.088	4.58×10^{-4}	8.51×10^{-4}
Validation	34	283	0.589	0.219	1.802 (1.173–2.768)	NA	0.0112	NA