Genetic and genomic analysis of early abortions in Israeli dairy cattle

Genetic analysis of early abortions in dairy cattle

Moran Gershoni[1], Ephraim Ezra[2] and Joel Ira Weller[1,2*]

[1]Department of Ruminant Science, Institute of Animal Sciences, Agricultural Research Organization, the Volcani Center, Rishon LeZion, Israel

[2] Israeli Cattle Breeders Association, Caesarea Industrial Park, Israel

* Corresponding author

E-mail: joel.weller@mail.huji.ac.il

Abstract

Female infertility accounts for at least 50% of all human infertility cases. One of the causes contributing for female infertility is embryo loss after fertilization. Previous findings suggested that more than half of fertilizations results in embryo loss before pregnancy is detected. Dairy cattle may be a useful model for study of the genetic architecture of this trait. In advanced commercial populations, all breeding is by artificial insemination, and extensive records of the cows' estrus, insemination and pregnancies are available. We proposed re-insemination between 49 and 100 days after the first insemination as an indicator trait for early abortion in dairy cattle, based on the mean estrus interval of 21 days. Israeli Holstein cows scored as early abortion were compare to cows recorded as pregnant from the first insemination. This trait was compare to conception rate from first insemination. Animal model variance components were estimated by REML, including parents and grandparents of cows with records. First parity heritability for conception rate was 3%. In the multi-trait analysis of parities 1-3 for abortion rate heritabilities ranged from 8.9% for first parity to 10.4% for second parity. The variance component for the service sire effect for abortion rate were less than half the variance component for conception rate. Thus genetic control of the two traits is clearly different. Genome wide association study were performed based on the genetic evaluations of ~1200 sires with reliabilities >50%. The markers with the lowest probabilities for early abortion were also included among the markers with the lowest probabilities for conception rate, but not vice versa. The marker explaining the most variance for abortion rate is located within the ABCA9 gene, which is found within an ABC genes cluster. The ATP-binding cassette family is the major class of primary active transporters in the placenta.

41

Author summary

43

44 Approximately 70% of human conceptions fail to achieve viability. Almost 50% of all
45 pregnancies end in miscarriage before the clinical recognition of a missed period. Cattle are a
46 useful model for human female reproductive processes, because of the similarities in the
47 reproductive cycles, and the extensive documentation in commercial cattle populations,
48 including estrus and insemination records. In addition to the expected benefits from cow fertility
49 research for human biomedical applications, fertility is an economically important trait in dairy
50 cattle with very low heritability. The mean estrous interval for cattle is 21 days. We therefore
51 proposed re-insemination between 49 and 100 days after the first insemination as an indicator
52 trait for early abortion. Israeli Holstein cows scored as having early abortion based on first
53 insemination after parturition were compare to cows recorded as pregnant from the first
54 insemination. Heritability for early abortion rate was three-fold the heritability for conception
55 rate. In a genome wide association study based on 1200 dairy bulls genotyped for 41,000
56 markers, six markers were found with nominal probabilities of $< 10^{-12}$ to reject the null
57 hypothesis of no effect on early abortion rate. Early abortion rate may be a useful indicator trait
58 for improvement of fertility in dairy cattle.

59

Introduction

61

62 Genetic factors that reduce the ability of an individual to reproduce are expected to be
63 under intensive negative selection, and therefore to remain rare in the population. It is thus

64 surprising that in humans about 15% of the reproductive age couples cannot achieve successful
65 pregnancy without medical assistance [1, 2]. This is partially due to the polygenic and sexually
66 differentiate nature of this trait [3 ,4]. Female infertility apparently accounts for more than half
67 of the cases [5]. Previous studies have suggested that the chances of a woman to achieve a
68 successful pregnancy per menstrual cycle is approximately 25% [6, 7]. Among other causes, the
69 lack of detected pregnancy could be the result of early abortion (EA). Approximately 70% of
70 human conceptions fail to achieve viability, with almost 50% of all pregnancies ending in
71 miscarriage before the clinical recognition of a missed period, or the presence of embryonal heart
72 activity [8, 9]. This complicates the study of EA in humans, although it is likely to have a large
73 impact on human fertility.

74 Cattle were found to be a useful model for human female reproductive processes, mainly
75 because of the similarities in the reproductive cycles [10]. Within this context, cows were used in
76 studies on ovarian function [11], effects of ageing on fertility [12], embryo-maternal
77 communication [13, 14], pregnancy maintenance and irregularities associated with assisted
78 reproduction techniques [15, 16]. The extensive documentation in many commercial cattle
79 populations, including estrus and insemination records, provides a good opportunity to
80 investigate the genetics of EA. In addition, each ejaculation is evaluated by AI labs, and
81 defective semen is rejected. Thus, the service sire has only a very minor effect on conception
82 rate [17]. Furthermore, fertility is an economically important trait in dairy cattle [18].

83 Like all economic traits in dairy cattle, genetic evaluation of fertility is based on field
84 records. Unlike milk production traits and somatic cell concentration, there is no accepted
85 consensus on how fertility should be scored. Traditionally the most common criterion was "non-
86 return rate," the fraction of cows that were not re-inseminated within a specific time interval.

This criterion ignored cows that were culled after the first insemination [19]. More recent measures have considered the time laps from first insemination to pregnancy or some function of the number of times a cow is inseminated during the lactation [20]. Nearly all measures of female fertility have very low heritability, in the range of 1 to 4% [18, 20].

Although abortions are recorded in Israel, very few first trimester abortions are noted either by the herd manager or the attending veterinarian. Indeed, previous studies in other populations estimated fertilization rate as greater than 75%, while the conception rate (CR) was approximately 35% [21]. The differences between these two observations are likely due to pregnancy loss that occur in more than 50%

Since genotyping of large numbers of animals with high density SNP chips has become routine, a number of recessive lethal alleles have been detected in commercial dairy cattle populations, which result in early term abortions [25]. Detection was originally based on the lack of homozygotes for the haplotype harboring the lethal allele, and a reduction in fertility rate for daughters of sires that received this haplotype. In several cases the causative polymorphism has been determined, e. g. [26]. Since the abortion is generally not observed, pregnancies of fetuses homozygous for the lethal alleles are recorded as "non-conception."

Many studies have shown that the average estrous interval in dairy cattle is 21 days [27]. Thus, most cows that do not conceive in the first service should be re-inseminated approximately 21 days later. The number of days between first and second insemination for first parity were previously shown with a peak at 21 days, and a secondary peak around 42 days, which corresponds to 2 estrus cycles. Despite these two prominent peaks, a significant number of cows were re-inseminated at >45 days after the first insemination [28]. Although these late inseminations may be due to non-conception at the first insemination, and lack of observed

110 estrous at the expected interval; another explanation is: conception at the first insemination, and
111 early term abortion due to embryonic lethality or female factors increasing the predisposition for
112 embryonic death.

113 One way to test this hypothesis is to demonstrate that the genetic factors that control long
114 intervals between first and second inseminations are different from the genetic factors that
115 control conception, and to evaluate the genetic contributions of the cow and the service sire. The
116 objectives of this study were a genetic and genomic analysis of cows with long intervals between
117 first and second insemination, as compared to cows recorded as conceiving at first insemination;
118 and comparison of this trait to CR at first insemination.

119

120 ## Results

121

122 To define long interval between inseminations as indication for EA, we first analyzed the
123 distribution of the insemination interval in Israeli-Holstein cows that were inseminated more than
124 once (Fig 1). The distribution of the insemination interval was similar to previously reported
125 [28]. Therefore, we defined EA as occurring when the cow was re-inseminated between 49-100
126 days after the first insemination (see also in the material and methods section). This interval was
127 previously suggested to represent embryonic death in most instances [29].

128 Effects of insemination month on conception and abortion rate as computed from the
129 REML analyses of data sets one and two are shown in Fig 2. Effects were set to zero for
130 December. Generally, the effects were similar for both traits, with major reductions in the late
131 summer, August and September. These results correspond to previous results for the effect of
132 insemination month on CR of Israeli Holsteins [17].

Estimates of variance components from the REML analyses of data sets one through four, and the heritabilities are given in Table 3. First parity heritabilities were 3.0% for CR from first insemination, but 7.7% for EA. For CR with assumed abortions deleted from the analysis, heritability for CR decreased to 2.6%. In the multi-trait analysis of parities 1-3 for EA heritabilities ranges from 8.9% for first parity to 10.4% for second parity, but differences among the parities were not significant. The variance components for the service sire were 8.6 and 10.1 for the two analyses of conception, but ≤ 3.5 for all the EA analyses. Although the service sire can effect EA by transfer of recessive lethal alleles to the fetus [25], the service sire factor apparently does not explain a major proportion of the genetic variance for EA. The additive genetic variance of the inseminated cow for EA is ~50 times greater than the variance of the service sire, as opposed to 8-fold for CR. Thus genetic control for the two traits is clearly different.

The genetic and environmental correlations among the three parities for the data set 4 analysis are given in Table 4. All genetic correlations were > 0.9, while all environmental correlations were < 0.12. Thus analysis by the single trait animal model is justified, with the variance ratios given in the methods section.

Correlations between breeding values for non-EA rate from data set 5 and the Israeli breeding index, PD16, and the other economic traits computed for Israeli Holsteins are given in Table 5. All correlations were significantly different from zero, except for the correlations with fat and protein production. The correlation with PD16 was 0.115. Thus selection for the index should have resulted in a decrease in abortion rate. The regression of the breeding value for non-EA rate on the cows' birth year was 0.083% per year ($P<0.001$), that is a decrease of close to 0.1% per year. The highest correlation was with female fertility, 0.75; but the correlation with

herd-life was also 0.3. Correlations among breeding values tend to slightly underestimate the actual genetic correlations, due to incomplete reliabilities of the evaluations.

Mean phenotypic and breeding values of non-EA rate for first parity cows by birth year are given in Fig 3. Although the overall regression of breeding value was positive for non-EA rate, EA rate increased until 1994, and then decreased beginning in 2002. These changes correspond to changes in the Israeli breeding index. Until 1996 the index included only milk production traits. Somatic cell score was added in 1996, and female fertility in 2000; which now accounts for 14% of the index. A lag of ~2 years between inclusion of a trait in the index and a change in the effective direction of selection is expected. The regression of breeding value on birth date for cows born since 2002 was 0.53% per year, as opposed to 0.083% since 1983. No clear trend is evident for the phenotypic means of first parity EA rate.

The "Manhattan Plot" for the genome wide association study (GWAS) results for EA rate are given in Fig 4, and the markers with the lowest probability values are presented in Table 6. There were eight markers with nominal probabilities $< 10^{-11}$. All of these markers has probabilities $<10^{-6}$ after permutation analysis and correction for multiple testing. Of the 8 markers listed, the 2 markers on chromosome 7 and the 2 markers on chromosome 17 are clearly due to a single quantitative trait locus segregating on each chromosome, since the distance between the two markers on each chromosome is <100,000 base pairs. Each of these 8 markers explained between 5 and 4% of the variance for the genetic evaluations for EA rate, and > 2.5% of the variance for CR. None of these chromosomal regions were found to have significant effects on cow conception rate or daughter pregnancy rate in the analysis of the US Holstein population by the posteriori granddaughter design [30].

178 While all of the 30 markers with probabilities < 10^{-8} for EA are also significantly
179 associated with CR, only subset of the markers with probabilities < 10^{-12} for CR are also
180 significantly associated with EA (Fig 5). The effects of these 30 markers on EA on their effects
181 on CR is plotted in Fig 6. The regression was 2.0 with a coefficient of determination of 0.97.
182 Thus the substitution effect for EA was generally twice the effect for CR.
183

184 ## Discussion

185

186 Reduce fertility is a major concern in humans, and a highly important economic trait in
187 dairy cattle. Previous studies suggested that a large portion of recorded non-conceptions in
188 human and cattle are apparently the result of unrecognized EA [8, 9, 21]. In this study we used
189 the extensive records of the Israeli dairy cattle population to study the genetics and the genomics
190 of EA and to assess its association with CR.
191 CR can be affected by the service sire and/or by the dam inseminated. In cattle, the service
192 sire can effect CR either via the quality of semen, or via genes that reduce embryo survival rate.
193 With respect to EA, the effect of the sire is limited to genes affection embryo survival rate. The
194 variance of the service sire effect on CR is approximately three-fold the variance of the service
195 sire effect on EA rate (Table 3). This suggests that most of the service sire effect on CR is due to
196 semen quality, otherwise the service sire variance component should have been of similar
197 magnitude to the variance component for EA. Although several recessive lethal genes have been
198 detected that cause EA [25], their effects on the genetic variance of EA rate appears to be
199 minimal. This is not surprising, considering that the lethal allele is always quite rare.
200 Generation of homozygotes for rare alleles is generally due to inbreeding, but inbreeding is

carefully monitored in the Israeli dairy cattle population. The inbreeding from each potential mating is checked by the inseminator, and matings that result in >3.125% inbreeding are generally rejected [31].

Only 2 previous studies attempted to estimate heritability of EA in dairy cattle, and both are somewhat problematic. In Bamber et al. [32] pregnancy loss was determined by an initial pregnancy diagnosis 26 to 33 d after AI followed by determination of loss of that embryo at a subsequent diagnosis 14 to 39 d later, but only 3,775 cows were included in the study. Due to the relatively small sample, confidence intervals for the genetic parameters were so wide, as to render the results virtually meaningless. They found a heritability of 17%, which was not significantly different from the value of ~10% in the current study. However, they found that the service sire variance was 16% of the total variance. This is clearly at variance with the current study, and was also considered difficult to explain by [32]. In Carthy et al. [33] EA was assumed to have occurred if pregnancy was determined by ultrasound examination, and the embryo was later deemed to be unviable by a later examination. However, ultrasound examinations were performed at various time points postpartum at the discretion of the producer. On a sample of 43,473 lactations they found heritability of only 2%, but repeatability of 66%. Part of the discrepancy to the current study may be due to the fact that the time periods for determination of both pregnancy and abortion were not consistent across records. Mean embryo loss was only 8%, as compared to 22% on first parity in the current study.

The genetic trend for EA rate corresponds to changes in the Israeli breeding index, as noted previously. However, the phenotypic trend is not similar to the genetic trend. Several factors can possibly explain this, including climatic variation, and the fact that the phenotypic means were computed only for first parity cows.

The GWAS results show that the markers with the lowest probability values for EA are also included among the markers with the lowest probability values for CR, but not vice versa (Fig 5). In addition, the regression of the substitution effects of the significant EA markers on their substitution effects for CR is ~2 (Fig 6). This suggests that the genetic factors that affect EA are likely a subgroup of factors affecting the CR. The markers that explain the most variance for EA suggest possible new insights on the polygenic architecture of this trait. For instance, by investigating the genomic area flanking these markers (Table 6), we found that the marker explaining the most variance (Table 6, Fig 7a) is located within the ABCA9 gene, that is found within an ABC genes cluster (Fig 7 a). The ATP-binding cassette (ABC) family are the major class of primary active transporters in the placenta. ABC proteins are reported to be important in efflux of xenobiotics and endogenous substrates like lipids, sterols and nucleotides. Recent studies provided evidence that ABC genes protect placental tissue by preventing accumulation of cytotoxic compounds, which is important in complicated pregnancies, such as in inflammatory or oxidative stress [34]. Thus, our finding of genetic variation within ABC genes cluster suggest that differences in EA predisposition might involve different sensitivity for oxidative stresses during the first trimester, mediated by ABC genes.

The significant marker on chromosome 21 is adjacent to the genes PAX9 and NKX2-1 (Fig 7c). Previous works suggested that PAX9 is required for the chondrogenic differentiation of sclerotomal cells during embryogenesis [35, 36], and that NKX2 is required for the embryonic development of cholinergic septohippocampal projection neurons [37]. Moreover, according to PathCards, NKX2-1 is part of the embryo pre-implementation path [38]. The fact that this marker is associated with a major effect on EA suggests that altered predisposition for EA might be mediated by genes that regulating embryo development.

Since fertility is a major economical trait in dairy cattle [18], our results suggest that EA, as defined in this study, should be considered for inclusion in the commercial selection index. Shook [39] listed the criteria that a potential trait must meet in order to be included in the selection objective. First, it should have an economic value. Second, the trait must have sufficiently large genetic variation in relation to its economic value and heritability. Third, the trait should be measurable at a low cost and consistently recorded. An indicator trait may be favored if it has a high genetic correlation with the economically important trait, is easier to record, has a higher heritability than the economic trait, or can be measured earlier in life [40]. The classic example of selection on an indicator trait in dairy cattle is somatic cell score as an indicator trait for mastitis [39]. The two traits have a relatively high genetic correlation, but somatic cell score has heritability of 10-20%, as opposed to 2-6% for clinical mastitis [39, 41]. Selection on EA rate, as defined in the current analysis is even more attractive as an indicator trait for female fertility; in that there is no requirement to generate new data. Considering the apparent high genetic correlation, and the fact that the heritability for EA rate is three-fold the heritability for most measures of fertility, genetic progress for fertility will be higher via selection for EA rate.

Materials and methods

Data sets analyzed

Eight data sets were analyzed. A basic description of the data sets is given in Table 1, and the numbers of animals and levels of effects included in each data set are given in Table 2. The first data set included conception status of first parity cows Israeli-Holstein cows with

calving dates from Jan. 1, 2007, through Dec. 31, 2016, and at least one insemination. Fertility data in Israel is unique in that cows that are not re-inseminated within 60 days are checked for pregnancy by a veterinarian [19]. The following records were deleted:

1. Cows that were daughters of foreign bulls.
2. Cows with first insemination by foreign bulls.
3. Cows with first insemination ≤30 and ≥135 days after parturition.
4. Cows for which pregnancy could not be determined for the first insemination. This chiefly included cows that were not re-inseminated, and were culled prior to determination of pregnancy by veterinary inspection.

The second data set was a subset of the first set, and included all cows that were either recorded as pregnant on the first insemination, and with pregnancy length ≤ 290 days; or cows that were recorded as "open" on the first insemination, and re-inseminated between 49 and 100 days after the first insemination. The latter group was assumed to represents cows with EA.

The third data set included all cows in the first data set recorded as pregnant, and cows that were recorded as open and re-inseminated prior to 49 days after the first insemination. This data set was assumed to represent cows that either became pregnant or remained open after the first insemination. Although the objective was to eliminate cows with EA, the latter group of cows clearly include a small fraction of cows that aborted, and still displayed estrus prior to 49 days after the first insemination.

Cows were included in the fourth data based on the same criteria as the second data set, but cows in second and third parities were also included. Since cows with second inseminations < 49 days or >100 days after first insemination were deleted for each parity, some cows had records on one, two or three parities, and for any specific parity a record could be included or

293 deleted. Variance components were computed for all four data sets for a binary trait. The
294 analyses included all parents and grandparents of cows with valid records.
295 The fifth data set was similar to the fourth data set, except that it included all first through
296 third parity cows with freshening dates from Jan. 1, 1985 through Dec. 31, 2016, which either
297 became pregnant on first insemination or were re-inseminated between 49 and 100 days after the
298 first insemination. This data set also included cows that were daughters of foreign bulls, and
299 even bulls of breeds other than Holsteins, although these cows were only about 1% of all the
300 cows. This data set was analyzed by the individual animal model in order to compute genetic
301 evaluations for all bulls with genotypes for medium or high density SNP-chips. As in the
302 previous data sets, the animal model analysis included all parents and grandparents of cows with
303 valid records.
304 The sixth data set included all valid first through fifth parity records for female fertility
305 with freshening dated from Jan 1 1985 through May 31, 2018. All cows that were inseminated at
306 least once were included. As described previously [42], fertility was scored as the inverse of the
307 number of inseminations to conception in percent. For cows that were culled prior to
308 conception, the expect number of inseminations to conception was computed. As in data set 5,
309 cows that were daughters of foreign Holstein bulls and bulls from breeds other than Holsteins
310 were included, and the data set included all parent and grandparents of cows with valid records.
311 This data set was analyzed by the multi-trait animal model, with each parity considered a
312 separate trait, as described by [43]. The separate parity evaluations were combined into a multi-
313 parity index, based on the economic value of each parity.
314 The seventh data set included all bulls with genetic evaluations from the analysis of the
315 fifth data set with reliabilities > 50% and genotypes for one of the medium or high density SNP-

316 chips. This data set was used to compute the GWAS analysis for frequency of EA. The final
317 data set included all bulls with genetic evaluations from the sixth data set with reliabilities > 50%
318 and genotypes. A GWAS analyses was also computed on this data set, and compared to the
319 GWAS for EA rate.
320
321 **Statistical analyses**
322
323 Variance components were estimated for data sets one through four using the
324 AIREMLf90 program [44]. The trait analyzed in data sets one and three was CR for the first
325 insemination, and for data sets two and four the trait analyzed was EA rate, under the assumption
326 that re-insemination between 49 and 100 days indicates an early abortion. Both traits were
327 scored dichotomously, with non-conception or abortion scored as zero, and pregnancy scored as
328 100. A single trait animal model was assumed for data sets one through three, and a multi-trait
329 animal model was assumed for data set four. For data set four, the three parities were considered
330 three separate traits. In addition to the random additive genetic effect of the cow calving, and
331 service sire for the first insemination, all models included the effect of insemination month and
332 herd-year-season as fixed effects. Two seasons were defined for each herd-year beginning in
333 April and October of each year. For data sets one through four, two genetic groups were defined
334 for animals with unlisted parents, one for males and one for females.
335 Heritability was defined as the ratio of the additive genetic variance to the sum of
336 additive genetic, service sire and the random residual variances. Genetic and environmental
337 correlations among the parities were computed for data set four. Genetic correlations were the
338 correlations among the additive genetic effects, and the environmental correlations were the

339 correlations among the residual effects. The AIREMLf90 program also computes solutions for
340 all effects included in the analysis model and standard errors for all variance components and the
341 heritabilities and the correlations.

342 Data set five was analyzed by a single trait animal model with all three parities
343 considered the same trait, as described by [42]. Thus the model included a random permanent
344 environmental effect in addition to the additive genetic effect. These effects differ in their
345 variance structure in that only the additive genetic effect included the relationship matrix. The
346 assumed ratios of variance between the residual and the additive genetic and permanent
347 environmental effects were both 9. Variances of abortion rate were lower in first and second
348 parities, due to lower abortion rates. Therefore, first and second parity records were each
349 multiplied by a factor greater than unity to obtain equal phenotypic variances for all three
350 parities. The adjusted records were then adjusted for the mean effects of parity and insemination
351 month by subtracting the means of the parity-insemination month classes from each record. In
352 addition to the additive genetic and the permanent environmental effects, the model included the
353 herd-year-season effect, as described previously; a parity-by herd type effect and a genetic group
354 effect. Two herd-types were defined; “Moshav” (family farms) and “Kibbutz” (communal
355 herds). Although the records were pre-adjusted for parity effects, a residual effect could remain
356 after accounting for all the effects included in the model.

357 In the analysis of data set 5, 84 groups were defined based on the sex of the animal with
358 unknown parents, which parent was unknown, and the birth year. In addition, separate groups
359 were defined for sire of cows of breeds other than Holstein. Although only a very small fraction
360 of the cows was sired by bulls of other breeds, these bulls were a significant fraction of the total
361 number of bulls, and an even larger fraction of the bulls with unknown parents.

362 The overall genetic trend was computed as the regression of the cows' breeding values on
363 their birth dates, for all cows born since Jan. 1, 1983. Yearly means of first parity non-abortion
364 rate and the breeding values of cows by birth year, relative to cows born in 2010, were
365 computed. Reliabilities of the breeding values of all animals included in the analysis were
366 estimated by the method of [45]. There were 1701 sires with reliabilities > 0.5 for EA rate.
367 Correlations were computed between the breeding values for EA rate for these sires and the
368 current Israel breeding index, PD16 and 11 economic traits analyzed in Israel. Breeding values
369 for these traits other than female fertility were computed as described previously [43, 46-48].
370
371 **Genome-wide association studies**
372
373 A total of 1749 Israeli Holstein bulls were genotyped. Since genotyping of these sires
374 were performed using several SNP-chip platforms, we filtered-in only those markers that were
375 covered in more than 90% of the tested cohort. Approximately 41,000 SNPs were retained.
376 Genome-wide associations were computed for the sires' transmitting abilities (½ of the breeding
377 value) for EA rate and CR (Table 1 and 2, data sets 6 and 7). Of the genotyped bulls, there were
378 1179 and 1297 with genetic evaluations for EA rate and CR, with reliabilities > 50%,
379 respectively. The additive substitution effects, the coefficients of determination (R^2) and the
380 nominal probabilities for the hypothesis of no effect were computed by using plink software
381 [49]. Genome-wide probabilities were estimated by generating one million permutations of
382 genotype data against the genetic evaluations. Thus, the minimal genome-wide probability was
383 $<10^{-6}$ if the substitution effect obtained from the actual data was greater than all of the
384 permutation effects.

Acknowledgments

This research was supported by grant number 58-8042-5-063F from the U.S.-Israel Binational Agricultural Research and Development (BARD) Fund, and by a grant from the Israel Dairy Board. We thank Ignacy Misztal and Shogo Tsuruta for use of the AIREMLF90 program, and Michael van Straten for useful discussions.

References

1. Agarwal A, Mulgund A, Hamada A, Chyatte MR. A unique view on male infertility around the globe. Reprod Biol Endocrinol. 2015; 13:37.
2. de Kretser DM. Male infertility. Lancet. 1997; 349: 787-90
3. Gershoni M, Pietrokovski S. Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. Nat Commun. 2014; 5: 4438.
4. Gershoni M, Pietrokovski S. The landscape of sex-differential transcriptome and its consequent selection in human adults. BMC Biol. 2017;15: 7.
5. Poppe K, Velkeniers B. Female infertility and the thyroid. Best Pract Res Clin Endocrinol Metab. 2004; 18: 153-165.
6. Henry L. French statistical research in natural fertility. Public health and population change: current research issues. Pittsburgh, Pa. University of Pitsburgh Press. 1966; 333-350.

7. Sheps MC. An analysis of reproductive patterns in an American isolate. Popul Stud (NY) 1965; 21: 65.
8. Edmonds DK, Lindsay KS, Miller JF, Williamson E, Wood PJ. Early embryonic mortality in women. Fertil Steril. 1982; 38: 447-453.
9. Wilcox AJ, Weinberg CR, O'connor JF, Baird DD, Schlatterer JP, Canfield RE, Nisula BC. Incidence of early loss of pregnancy. N Engl J Med. 1988; 319: 189-194.
10. Evans AC. Characteristics of ovarian follicle development in domestic animals. Reprod Domest Anim. 2003; 38: 240–246.
11. Yapura J, Mapletoft RJ, Pierson R, Singh J, Naile J, Giesy JP, Adams GP. A bovine model for examining the effects of an aromatase inhibitor on ovarian function in women. Fertil Steril. 2011; 96: 434–438. e433.
12. Malhi PS, Adams GP, Singh J. Bovine model for the study of reproductive aging in women: follicular, luteal, and endocrine characteristics. Biol Reprod. 2005; 73: 45–53.
13. Bauersachs S, Mitko K, Ulbrich SE, Blum H, Wolf E. Transcriptome studies of bovine endometrium reveal molecular profiles characteristic for specific stages of estrous cycle and early pregnancy. Exp Clin Endocrinol Diabetes. 2008;116: 371–384.
14. Ostrup E, Hyttel P, Ostrup O, Embryo-maternal communication. signaling before and during placentation in cattle and pig. Reprod Fertil Dev. 2011; 23: 964–975.
15. Bauersachs S, Ulbrich SE, Zakhartchenko V, Minten M, Reichenbach M, Reichenbach HD, Blum H, Spencer TE, Wolf E. The endometrium responds differently to cloned versus fertilized embryos. Proc Natl Acad Sci USA. 2009; 106: 5681–5686.
16. Mansouri-Attia N, Sandra O, Aubert J, Degrelle S, Everts RE, Giraud-Delville C, Heyman Y, Galio L, Hue I, Yang X, Tian XC, Lewin HA, Renard JP. Endometrium as an

early sensor of in vitro embryo manipulation technologies. Proc Natl Acad Sci USA. 2009; 106: 5687–5692.

17. Weller JI, Ron M. Genetic analysis of fertility traits in Israeli Holsteins by linear and threshold models. J Dairy Sci. 1992; 75: 2541-2548.
18. Fleming A, Baes CF, Martin AAA, Chud TCS, Malchiodi F, Brito LF, Miglior F. Symposium review: The choice and collection of new relevant phenotypes for fertility selection. J Dairy Sci. 2019; 102:1–13.
19. Weller JI. Genetic analysis of fertility traits in the Israeli dairy cattle population. J Dairy Sci. 1989; 72: 2644-2650.
20. VanRaden PM, Sanders AH, Tooker ME, Miller RH, Norman HD, Kuhn MT, Wiggans GR. Development of a national genetic evaluation for cow fertility. J Dairy Sci. 2004; 87:2285–2292.
21. Santos, J. E. P., Thatcher, W. W., Chebel, R. C., Cerri, R. L. A., & Galvao, K. N. The effect of embryonic death rates in cattle on the efficacy of estrus synchronization programs. Anim Reprod Sci. 2004; 82: 513–535.
22. Moore DA, Overton MW, Chebel RC, Truscott ML, BonDurant RH. Evaluation of factors that affect embryonic loss in dairy cattle. J Am Vet Med Assoc. 2005; 226: 1112-1118.
23. Drost M, Ambrose JD, Thatcher MJ, Cantrell CK, Wolfsdorf KE, Hasler JF, Thatcher WW. Conception rates after artificial insemination or embryo transfer in lactating dairy cows during summer in Florida. Theriogenology. 1999; 52: 1161-1167.

24. Sartori R, Gümen A, Guenther JN, Souza AH, Caraviello DZ, Wiltbank MC. Comparison of artificial insemination versus embryo transfer in lactating dairy cows. Theriogenology. 2006; 65: 1311-1321.

25. Cole JB, Null DJ, VanRaden PM. Phenotypic and genetic effects of recessive haplotypes on yield, longevity, and fertility. J Dairy Sci. 2016; 99: 7274–7288.

26. Schütz E, Wehrhahn C, Wanjek M, Bortfeld R, Wemheuer WE, Beck J, et al. The Holstein Friesian lethal haplotype 5 (HH5) results from a complete deletion of TBF1M and cholesterol deficiency (CDH) from an ERV-(LTR) insertion into the coding region of APOB. Plos One. 2016; e0154602.

27. Remnant JG, Green MJ, Huxley JN, Hudson CD. Associations between dairy cow inter-service interval and probability of conception. Theriogenology. 2018; 114: 324-329

28. Greenham T, Oikonomou G, Grove-White DH. A description of interestrus and interservice intervals and associated fertility in 16 United Kingdom dairy herds. J Dairy Sci. 2019; 102: 824-832.

29. Markette KL, Seidel JR GE, Elsden RP. Estimation of embryonic losses in bovine embryo transfer recipients from progesterone profiles and returns to estrus. Theriogenology. 1985; 23.1: 45-62.

30. Weller JI, Bickhart DM, Wiggans GR, Tooker ME, O'Connell JR, Jiang J, Ron M, VanRaden PM. Determination of quantitative trait nucleotides by concordance analysis between quantitative trait loci and marker genotypes of US Holsteins. J Dairy Sci. 2018; 101: 9089-9107.

31. Weller JI, Ezra E. Analysis of inbreeding in the Israeli Holstein dairy cattle population. Proc. 2005 Interbull meeting, Uppsala, Sweden, June 2-4, pp. 85-87.

474 32. Bamber RL, Shook GE, Wiltbank MC, Santos JEP, Fricke PM, Genetic parameters for
475 anovulation and pregnancy loss in dairy cattle. J Dairy Sci. 2009; 92: 5739–5753.
476 33. Carthy TR, Ryan DP, Fitzgerald AM, Evans RD, Berry DP. Genetic parameters of
477 ovarian and uterine reproductive traits in dairy cows. J Dairy Sci. 2015; 98: 4095–4106.
478 34. Aye IL, Keelan JA. Placental ABC transporters, cellular toxicity and stress in pregnancy.
479 Chem Biol Interact. 2013; 203: 456-466.
480 35. Rodrigo I, Hill RE, Balling R, Münsterberg A, Imai K. Pax1 and Pax9 activate Bapx1 to
481 induce chondrogenic differentiation in the sclerotome. Development. 2003; 130: 473-482.
482 36. Müller TS, Ebensperger C, Neubüser A, Koseki H, Balling R, Christ B, Wilting J.
483 Expression of avian pax1 and pax9 is intrinsically regulated in the pharyngeal endoderm,
484 but depends on environmental influences in the paraxial mesoderm. Dev Biol. 1996; 178:
485 403-417.
486 37. Magno L, Barry C, Schmidt-Hieber C, Theodotou P, Häusser M, Kessaris N. NKX2-1 is
487 required in the embryonic septum for cholinergic system development, learning, and
488 memory. Cell reports. 2017; 20: 1572-1584.
489 38. Belinky F, Nativ N, Stelzer G, Zimmerman S, Iny Stein T, Safran M, Lancet D.
490 PathCards: multi-source consolidation of human biological pathways. Database, 2015;
491 doi.org/10.1093/database/bav006.
492 39. Shook GE. Selection for disease resistance. J Dairy Sci. 1989; 72: 1349–1362.
493 40. Miglior, F, Fleming A, Malchiodi F, Brito LF, Martin P, Baes CF. A 100-Year Review:
494 Identification and genetic selection of economically important traits in dairy cattle. J
495 Dairy Sci. 2017; 100: 10251–10271.

41. Shook GE. Major advances in determining appropriate selection goals. J Dairy Sci. 2006; 89: 1349–1361.

42. Weller JI, Ezra E. Genetic analysis of somatic cell concentration and female fertility of Israeli Holsteins by the individual animal model. J Dairy Sci. 1997; 80: 586-593.

43. Weller JI, Ezra E. Genetic analysis of the Israeli Holstein dairy cattle population for production and non-production traits with a multitrait animal model. J Dairy Sci. 2004; 87: 1519-1527.

44. Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z. Manual for BLUPF90 family of programs. University of Georgia, Athens, USA. 2014. (http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all2.pdf)

45. Misztal I, Wiggans GR. Approximation of prediction error variance in large-scale animal models. J Dairy. Sci 1988; 71(Suppl. 2): 27-32.

46. Settar P, Weller JI. Genetic analysis of cow survival in the Israeli dairy cattle population. J Dairy Sci. 1999; 82: 2170-2177.

47. Weller JI, Ezra E. Genetic analysis of calving traits by the multi-trait individual animal model. J Dairy Sci. 2016; 99: 427–442.

48. Weller JI, Ezra E, Leitner G. Genetic analysis of persistency in the Israeli Holstein population by the multitrait animal model. J Dairy Sci. 2006; 89: 2738-2746.

49. 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.

517 **Table 1. Basic description of the seven data sets analyzed.**

Data Set	Description	Analysis Type[a]	Parities	Freshening years included	
				Beginning	End
1	Conception rate, all cows	VC	1	2007	2016
2	Non-abortion rate	VC	1	2007	2016
3	Conception rate without abortions	VC	1	2007	2016
4	Non-abortion rate	VC	1-3	2007	2016
5	Non-abortion rate	SAM	1-3	1985	2016
6	Female fertility	MAM	1-5	1985	2018
7	Non-abortion rate, GWAS	GWAS	1-3	1985	2016
8	Female fertility, GWAS	GWAS	1-3	1985	2016

518

519 [a] VC = variance component analysis, SAM = single trait animal model, MAM = multi-trait

520 animal model, GWAS = genome-wide association study.

521

522 **Table 2. Number of animals with records, ancestors and levels of effects in the analysis**

523 **models.**

Data Set	Number of levels or animals					
	Insemination month	Service sire	Herd-year-seasons	Genetic groups	Animals with records	Ancestors
1	12	569	13699	2	260782	142297
2	12	554	12802	2	118039	123600
3	12	568	13552	2	233037	142816
4	-	-	13551	2	182758	136187
	12	554	-	-	118039	-
	12	542	-	-	77193	-
	12	431	-	-	51234	-
5	-	-	48757	84	571988	181478
6	-	-	50536	80	861939	143649
7	-	-	-	-	1179	-
8	-	-	-	-	1297	-

524

525 Data sets are defined in Table 1.

526

527 **Table 3. Variance components (±standard errors) and heritabilities for data sets 1-4.**

Data set	Parity	Frequency of abortion/ conception (%)	Variance components			Heritability
			Additive genetic	Service sire	Residual	
1	1	40.2	69.2±6.8	8.6±0.9	2207.4±7.9	0.030±0.003
2	1	21.7	119.7±11.1	2.8±0.7	1426.3±10.1	0.077±0.007
3	1	55.2	60.9±6.7	10.1±1.1	2273.5±8.4	0.026±0.003
4	1	21.7	138.0±10.2	2.9±0.7	1414.4±9.5	0.089±0.006
	2	29.6	201.2±15.6	3.5±1.0	1735.1±14.8	0.104±0.008
	3	31.0	188.6±18.6	3.5±1.5	1818.8±18.7	0.094±0.009

528 Heritability was estimated as the additive genetic variance component divided by the sum of all

529 three variance components.

530

531 **Table 4. Genetic and environmental correlations ± standard errors for abortion rate (data**

532 **set 4).**

Parities	Correlations	
	Genetic	Environmental
1, 2	0.966±0.012	0.096±0.008
1, 3	0.905±0.028	0.066±0.010
2, 3	0.963±0.017	0.116±0.010

533 Genetic correlations were the correlations among the additive genetic effects, and the

534 environmental correlations were the correlations among the residual effects.

535 **Table 5. Correlations between breeding values for non-abortion rate from data set 5 and**
536 **the Israeli breeding index (PD16) and the other economic traits computed for Israeli**
537 **Holsteins.**

Traits	% of index	Number of bulls	Correlation
PD16	100	1693	0.115***
Milk (kg)	0.00	1693	-0.077**
Fat (kg)	21.20	1693	-0.026
Protein (kg)	37.32	1693	-0.044
SCS[1]	10.98	1693	-0.196***
Female fertility (%)	14.38	1693	0.749***
Herd-life (days)	9.58	1693	0.296***
Persistency (%)	4.24	1693	0.078**
Dystocia, maternal (%)[a]	1.27	1693	-0.179***
Stillbirth, maternal (%)[a]	1.03	1693	-0.246***
Dystocia, direct (%)[1]	0.00	1592	-0.105***
Stillbirth, direct (%)[1]	0.00	1592	-0.112***

538 The relative contribution of each to the index and the numbers of bulls with evaluations with
539 reliabilities > 0.5 for each trait are also listed.
540 [a] Negative values are economically favorable.
541 * significant, p<0.05; **, significant, p<0.01, ***, significant, p<0.001.

Table 6. Markers with the lowest probability values for early abortion rate.

Location		SNP	Early abortion rate			Conception rate		
Chr.	Base pairs		Beta[a]	R^{2b}	P[c]	Beta	R^2	P
19	61503930	Hapmap43271-BTA-46356	1.56	0.051	4.96E-15	0.89	0.058	1.30E-18
8	24608595	Hapmap41408-BTA-103152	1.52	0.047	6.08E-14	0.83	0.048	1.34E-15
17	26712567	BTA-46662-no-rs	-1.95	0.047	6.60E-14	-1.02	0.045	1.19E-14
21	46984914	BTA-52458-no-rs	1.49	0.047	6.63E-14	0.60	0.026	4.51E-09
17	26735031	Hapmap41875-BTA-46663	-2.23	0.045	1.81E-13	-1.26	0.050	3.43E-16
24	53790841	BTA-58638-no-rs	1.77	0.043	4.99E-13	0.84	0.035	1.47E-11
7	22920391	BTB-01966013	1.49	0.042	1.01E-12	0.68	0.031	2.07E-10
7	22996615	BTB-01398686	1.47	0.041	1.95E-12	0.66	0.029	6.26E-10

Markers are sorted in descending order of the probability to reject the null hypothesis of no effect on early abortion rate. The substitution effects and coefficients of determination are given for each marker for early abortion and conception rate.

[a]The allele substitution effects in transmitting value trait units.

[b]Coefficient of determination.

[c]The nominal probability for the hypothesis of no effect. All genome-wide probabilities were $<10^{-6}$.

550 **Figure captions**

551

552 **Fig 1. Frequencies of days between first and second inseminations for first parity Israeli**
553 **Holstein cows.** Valid records of cows freshening between 2007 and 2016 are included.

554

555 **Fig 2. Effects of insemination month on conception and abortion rate.** Based on valid
556 records of cows freshening between 2007 and 2016. ▬, Conception rate; ▬ , early
557 abortion rate. A score of 0 = non-conception or abortion, and 100 = normal pregnancy. Effects
558 are computed relative to December.

559

560 **Fig 3. Mean breeding values of non-abortion rate and mean non-abortion rate of first**
561 **parity cows by birth year.** ▬, Mean breeding value for non-abortion rate; ▬ , Mean
562 non-abortion rate. A score of 0 = abortion, and 100 = normal pregnancy. Genetic evaluations
563 are computed relative to cows with born in 2010.

564

565 **Fig 4. GWAS Manhattan plot for EA rate**. Dots represent each marker. Chromosomal
566 positions are on the X-axis, and nominal –log10 P-values are on the Y-axis. The blue line
567 denotes the genome-wide significance threshold of 0.05, as derived from one million data
568 permutations and correction for multiple testing.

569

570 **Fig 5. Comparison between the GWAS results for early abortion rate and conception rate.**
571 a) The 30 markers with the lowest nominal probabilities for early abortion rate as a function of
572 the corresponding probabilities of these markers for conception rate. b) The 30 markers with the

573 lowest nominal probabilities for conception rate as a function of the corresponding probabilities
574 of these markers for early abortion rate. The red dashed line denotes the genome-wide
575 significance threshold of 0.05 for early abortion as derived from one million data permutations
576 and after correction for multiple testing.
577
578 **Fig 6.** The substitution effects for early abortion rate of the 30 markers with the lowest nominal
579 probabilities as a function of the substation effects for the same markers for conception rate.
580
581 **Fig 7.** Schematic representation of the genomic area flanking four of the markers with the lowest
582 nominal probability values (Table 6). The arrows represent the genes position and the strand
583 orientation. The red asterisk represents the marker positions for a) Hapmap43271-BTA-46356
584 located on chromosome 19 b) Hapmap41408-BTA-103152 located on chromosome 8 c) BTA-
585 52458-no-rs located on chromosome 21 and d) BTA-58638-no-rs located on chromosome 21.
586

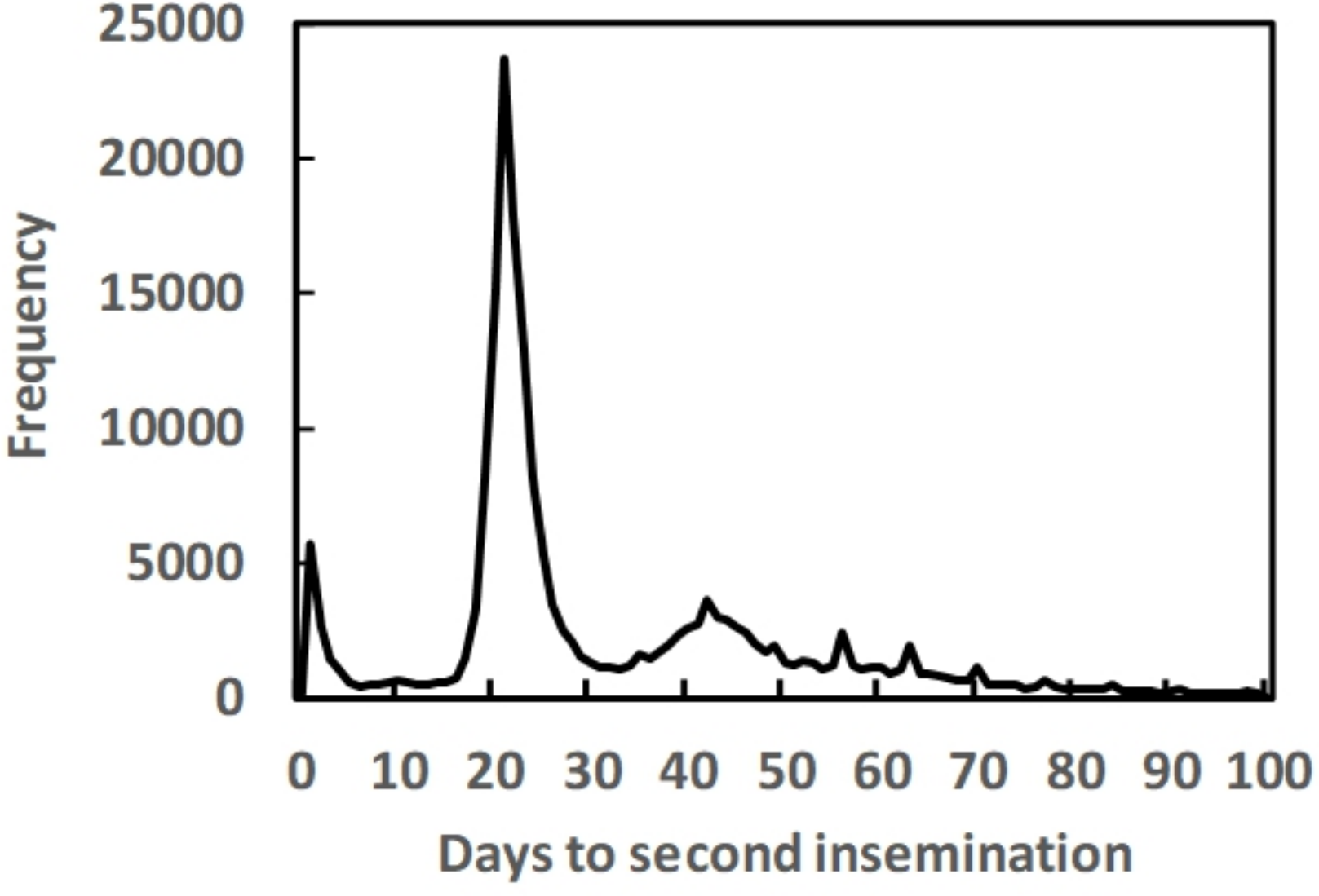
25000
20000
15000
10000
5000
0
Frequency
0 10 20 30 40 50 60 70 80 90 100
Days to second insemination

Figure 1

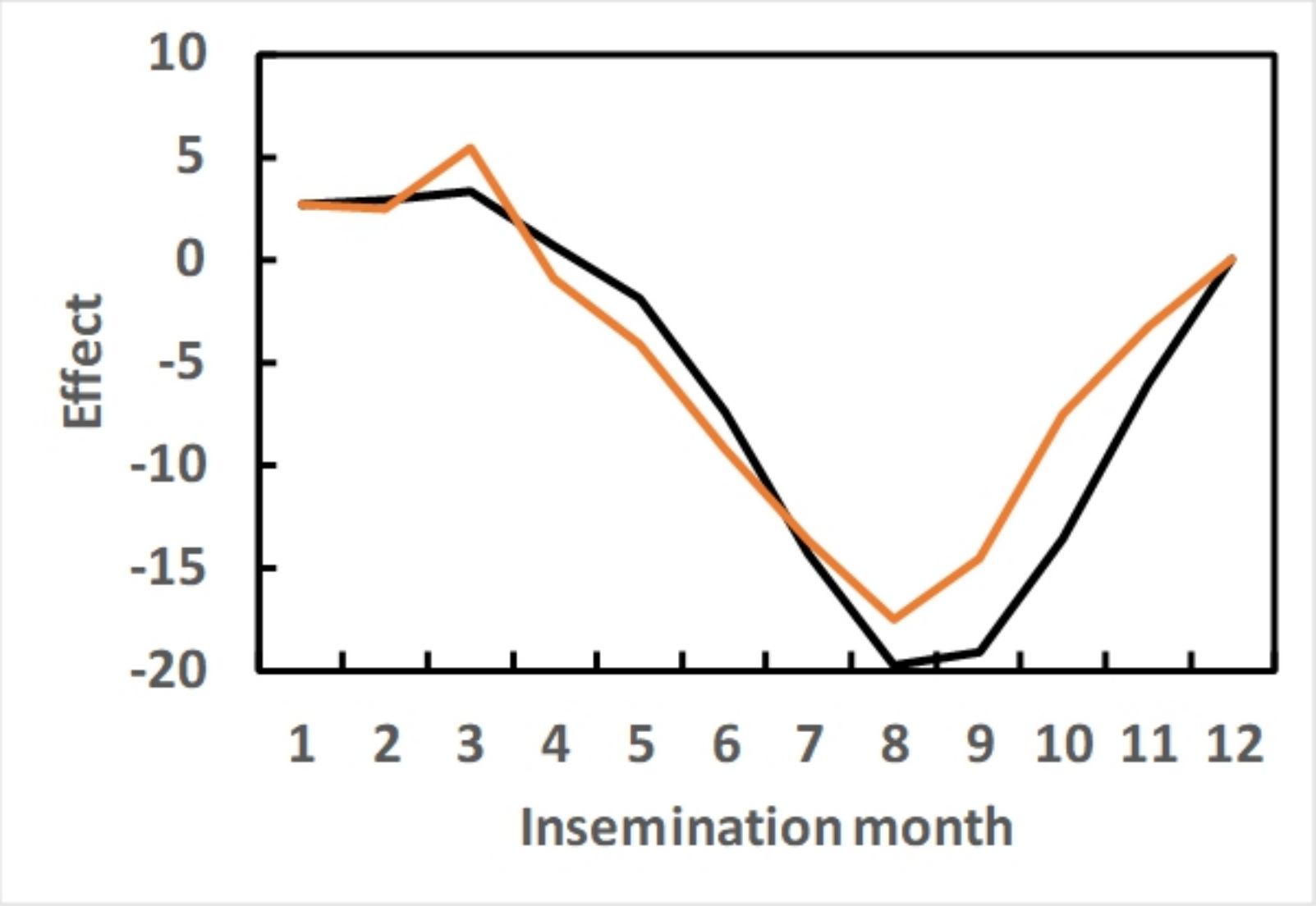

Figure 2

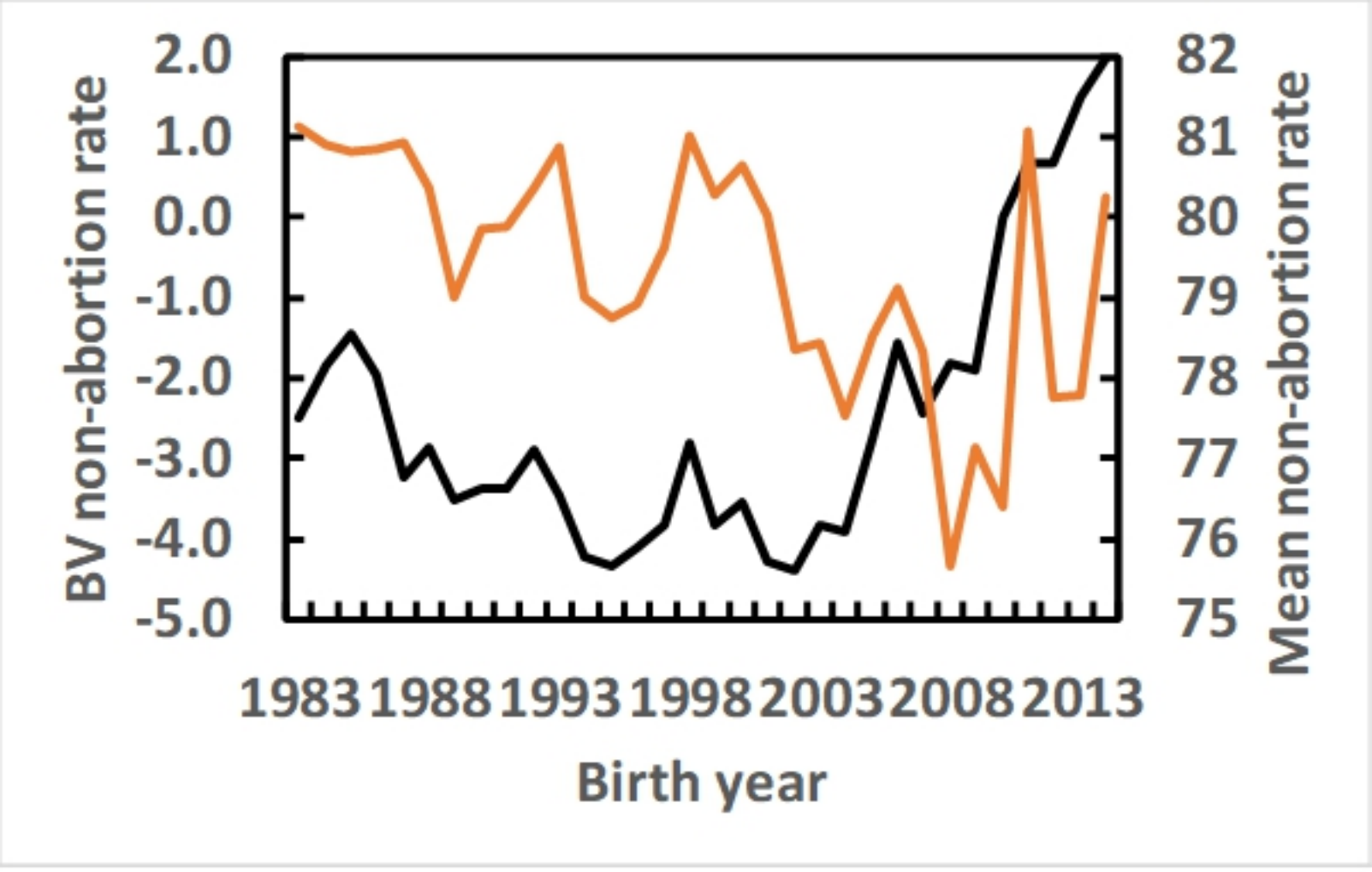
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0
-5.0
BV non-abortion rate
82
81
80
79
78
77
76
75
Mean non-abortion rate
1983
1988
1993
1998
2003
2008
2013
Birth year

Figure 3

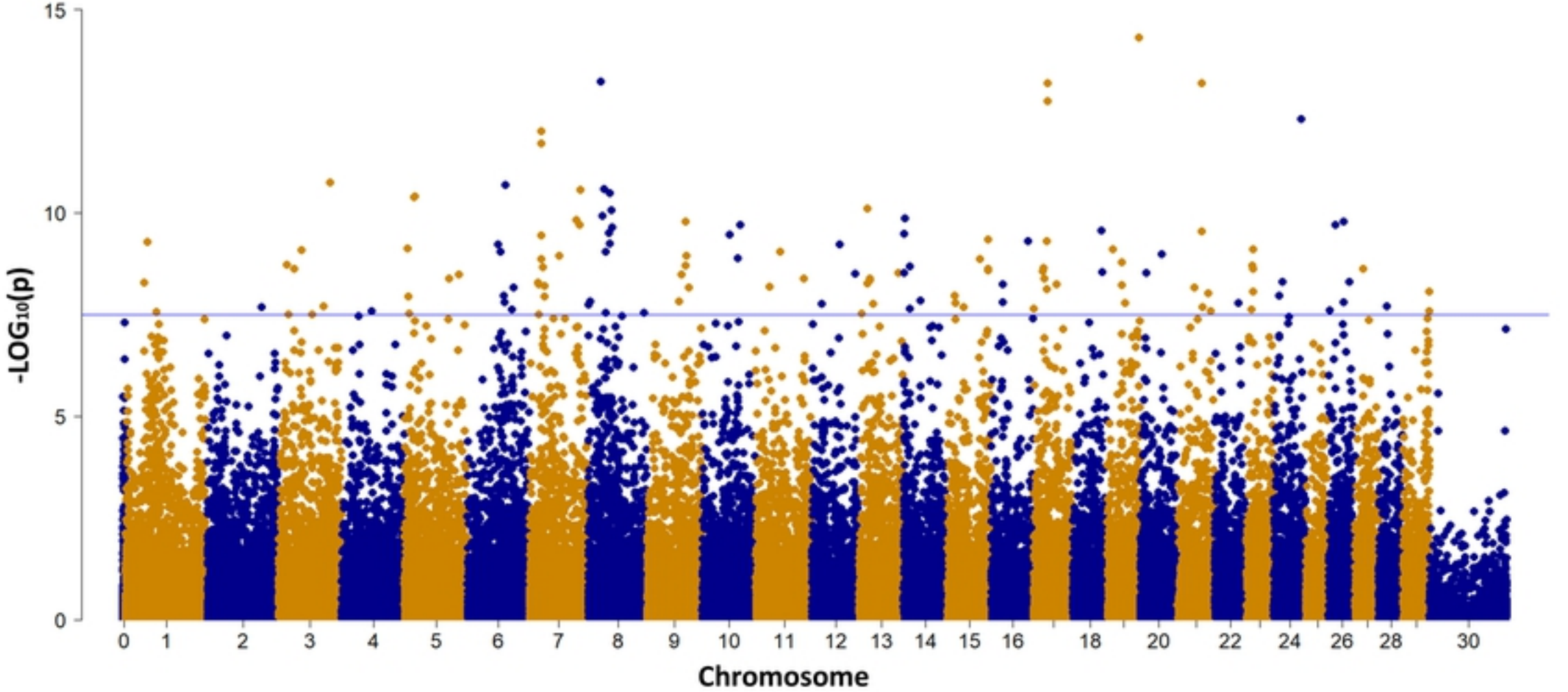
15
10
5
0
-LOG10(p)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 20 22 24 26 28 30
Chromosome

Figure 4

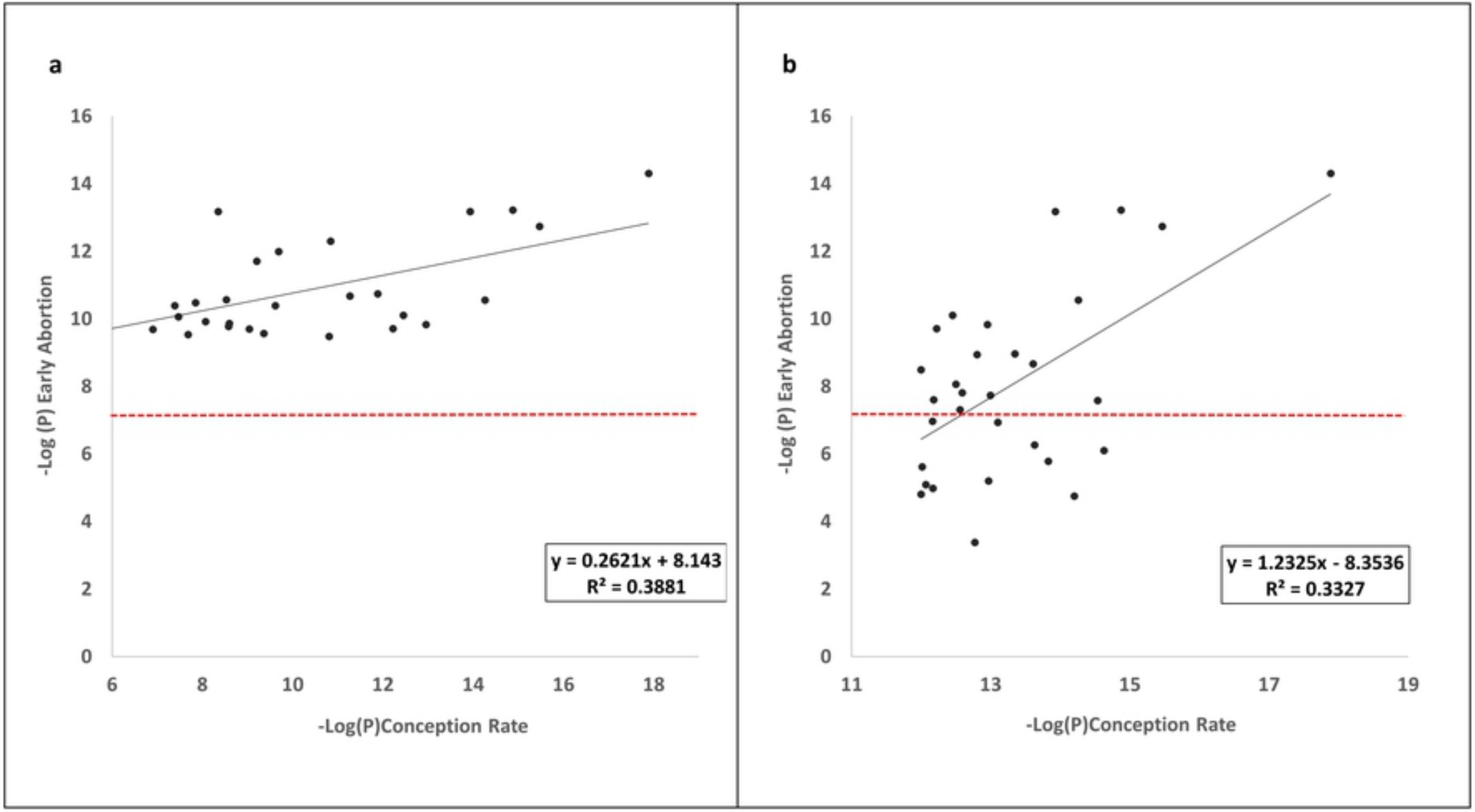
a
16
14
12
10
8
6
4
2
0
-Log (P) Early Abortion
y = 0.2621x + 8.143
R² = 0.3881
6
8
10
12
14
16
18
-Log(P)Conception Rate
b
16
14
12
10
8
6
4
2
0
-Log (P) Early Abortion
y = 1.2325x - 8.3536
R² = 0.3327
11
13
15
17
19
-Log(P)Conception Rate

Figure 5

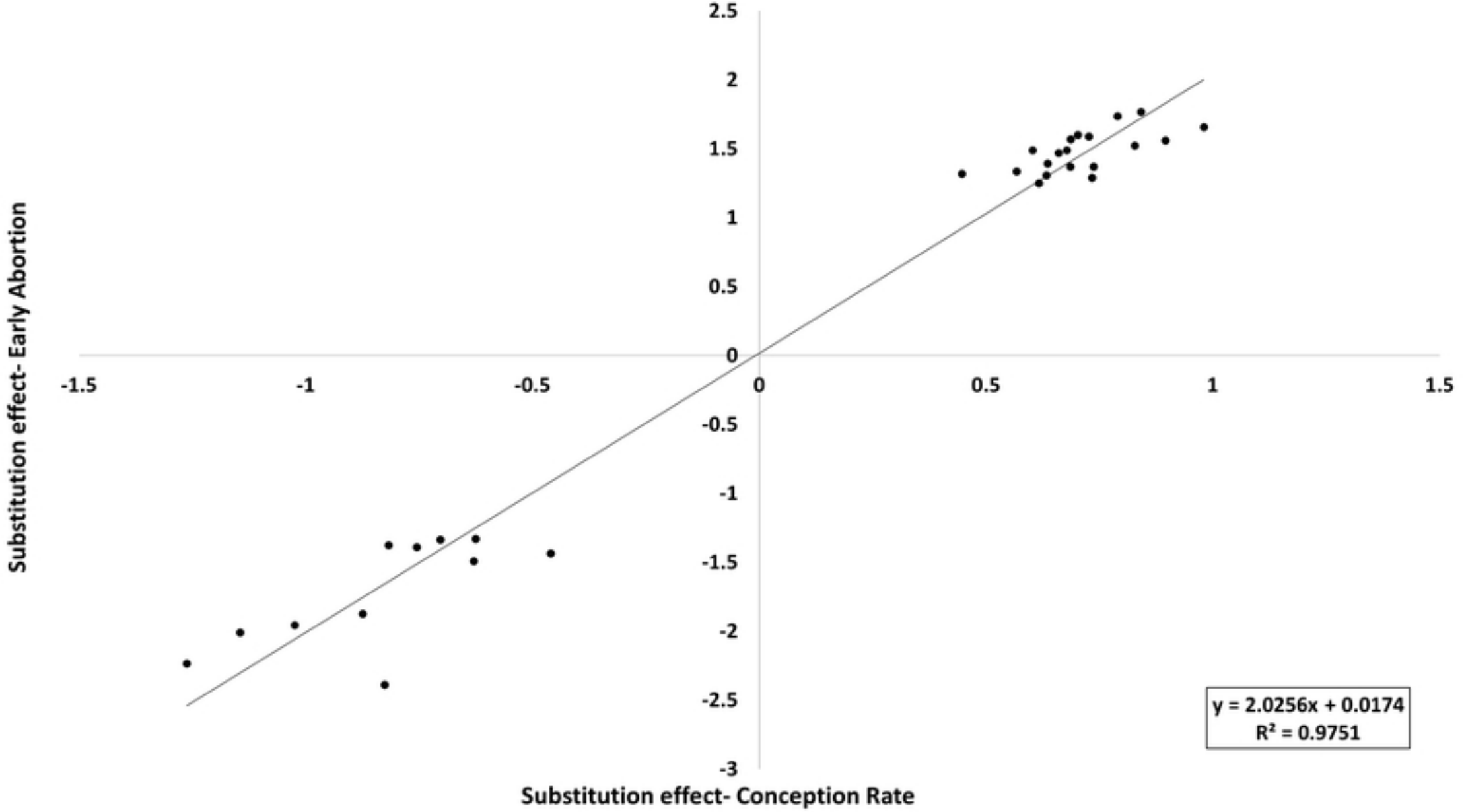
Substitution effect- Early Abortion
Substitution effect- Conception Rate
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-1.5
-1
-0.5
0
0.5
1
1.5
y = 2.0256x + 0.0174
R² = 0.9751

Figure 6

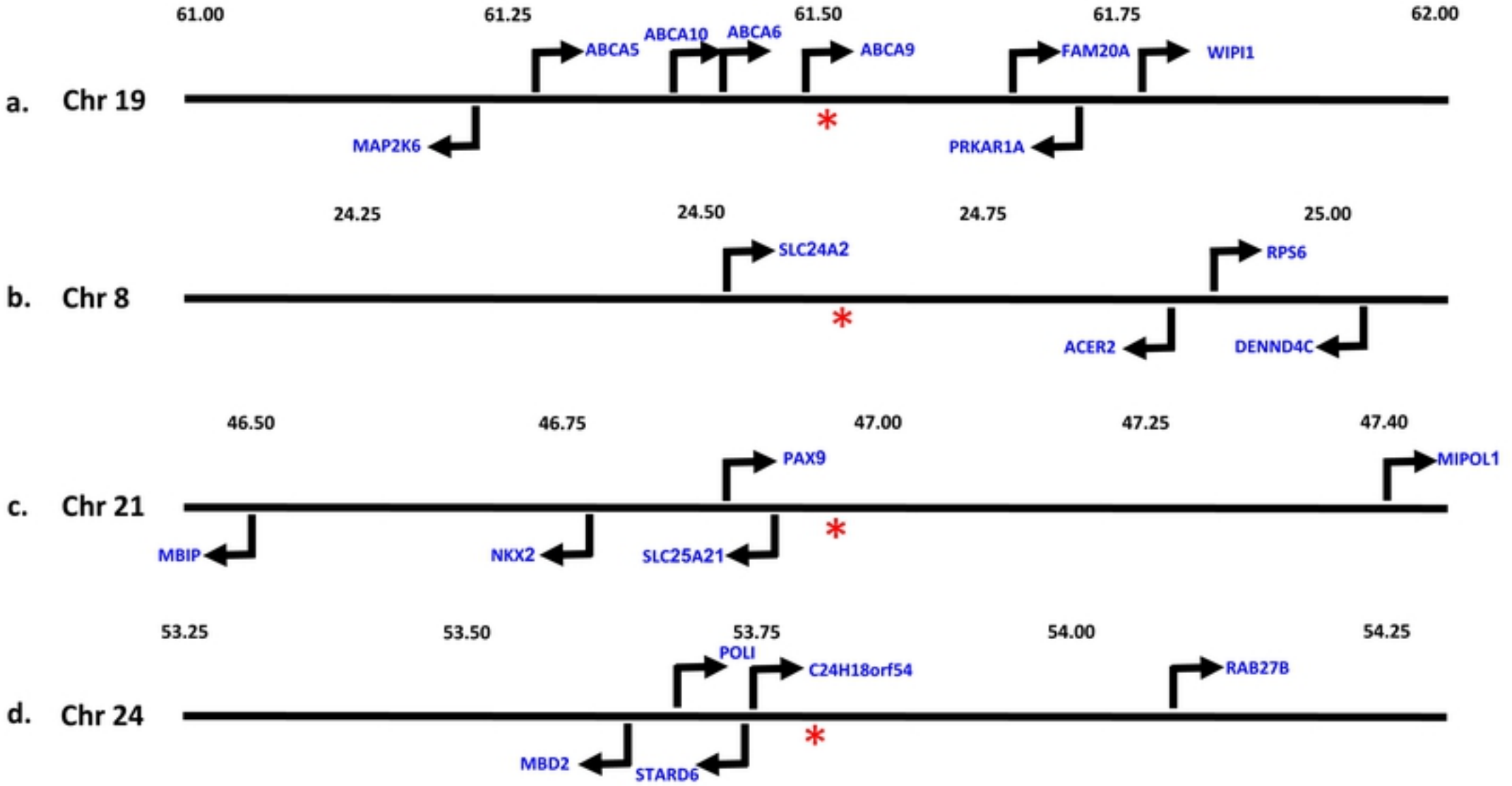
61.00
61.25
61.50
61.75
62.00
a. Chr 19
ABCA5
ABCA10
ABCA6
ABCA9
FAM20A
WIPI1
MAP2K6
PRKAR1A
24.25
24.50
24.75
25.00
b. Chr 8
SLC24A2
RPS6
ACER2
DENND4C
46.50
46.75
47.00
47.25
47.40
c. Chr 21
PAX9
MIPOL1
MBIP
NKX2
SLC25A21
53.25
53.50
53.75
54.00
54.25
d. Chr 24
POLI
C24H18orf54
RAB27B
MBD2
STARD6

Figure 7