

Major histocompatibility complex-associated odour preferences and human mate choice: near and far horizons

J. Havlíček*, J. Winternitz✉*, and S.C. Roberts

✉ Jamie C. Winternitz

Department of Animal Behaviour, Bielefeld University, Bielefeld 33615, Germany;

Email: jcwinternitz@gmail.com

* These authors contributed equally to this work

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TEXT	2
Data bibliography.....	2
Details of data extraction and effect size calculations for MHC mate choice studies	4
Details of data extraction and effect size averaging for neutral markers from MHC studies	11
List of relevant papers excluded from the present study.....	13
SUPPLEMENTARY TABLES	14
Table S1. Estimated effect of human genomic MHC similarity on mating using moderators.	14
Table S2. Estimated effect of human MHC similarity on relationship satisfaction using moderators.	16
Table S3. Estimated effect of human MHC similarity on odour preferences using moderators.	17
Table S4. Estimated effect of non-HC using human MHC similarity on odour preferences using moderators.....	18
Table S5. Estimated effect of human MHC similarity on mate selection using moderators.	19
Table S6. Testing for publication bias using Egger's regression test of funnel plot asymmetry.	21
SUPPLEMENTARY FIGURES.....	22
Figure S1. Forest plots of standardized effect sizes by study for different datasets.....	22
Figure S2. Trim and fill contour-enhanced funnel plots for the different datasets.....	24
Figure S3. Meta-regression plot for all mate selection studies indicates evidence of publication bias.....	26

Supplementary Text

This includes the following sections: (1) Data bibliography, (2) Details of data extraction and effect size calculation for MHC mate choice studies, (3) Details of data extraction and effect size averaging for neutral markers from MHC studies, and (4) List of relevant papers excluded from the present study. All data used for analyses are provided in a CSV file on figshare (doi: 10.6084/m9.figshare.8869505).

Data bibliography

- Chaix RI, Cao C, Donnelly P (2008) Is mate choice in humans MHC-dependent? *PLoS Genetics* **4**, e1000184.
- Cretu-Stancu M, Kloosterman WP, Pulit SL (2018) No evidence that mate choice in humans is dependent on the MHC. *bioRxiv*.
- Dandine-Roulland C, Laurent R, Dall'Ara I, Toupance B, Chaix R (2019) Genomic evidence for MHC disassortative mating in humans. *Proceedings of the Royal Society B* **286**, 20182664.
- Derti A, Cenik C, Kraft P, Roth FP (2010) Absence of evidence for MHC-dependent mate selection within HapMap populations. *PLoS Genetics* **6**, e1000925.
- Garver-Apgar CE, Gangestad SW, Thornhill R, Miller RD, Olp JJ (2006) Major histocompatibility complex alleles, sexual responsivity, and unfaithfulness in romantic couples. *Psychological Science* **17**, 830-835.
- Giphart M, D'Amato J (1983) HLA and reproduction? *Journal of immunogenetics* **10**, 25-29.
- Hedrick PW, Black FL (1997) HLA and mate selection: no evidence in South Amerindians. *American Journal of Human Genetics* **61**, 505-511.
- Ihara Y, Aoki K, Tokunaga K, Takahashi K, Juji T (2000) HLA and human mate choice: tests on Japanese couples. *Anthropological Science* **108**, 199-214.
- Israeli M, Kristt D, Nardi Y, Klein T (2014) Genetic considerations in human sex-mate selection: partners share Human Leukocyte Antigen but not Short-Tandem-Repeat identity markers. *American Journal of Reproductive Immunology* **71**, 467-471.
- Jacob S, McClintock MK, Zelano B, Ober C (2002) Paternally inherited HLA alleles are associated with women's choice of male odor. *Nature genetics* **30**, 175-179.
- Jin K, Speed T, Thomson G (1995) Tests of random mating for a highly polymorphic locus: application to HLA data. *Biometrics*, 1064-1076.
- Khankhanian P, Gourraud PA, Caillier SJ, Santaniello A, Hauser SL, Baranzini SE, Oksenberg JR (2010) Genetic variation in the odorant receptors family 13 and the mhc loci influence mate selection in a multiple sclerosis dataset. *BMC Genomics* **11**, 626.
- Kromer J, Hummel T, Pietrowski D, Giani A, Sauter J, Ehninger G, Schmidt A, Croy I (2016) Influence of HLA on human partnership and sexual satisfaction. *Scientific Reports* **6**, 32550.
- Laurent R, Toupance B, Chaix R (2012) Non-random mate choice in humans: insights from a genome scan. *Molecular Ecology* **21**, 587-596.
- Nordlander C, Hammarström L, Lindblom B, Smith CE (1983) No role of HLA in mate selection. *Immunogenetics* **18**, 429-431.
- Ober C, Weitkamp LR, Cox N, Dytch H, Kostyu D, Elias S (1997) HLA and mate choice in humans. *American Journal of Human Genetics* **61**, 497-504.
- Pollack MS, Wysocki CJ, Beauchamp GK, Braun Jr D, Callaway C, Dupont B (1982) Absence of HLA association or linkage for variations in sensitivity to the odor of androstenone. *Immunogenetics* **15**, 579-589.

- Probst F, Fischbacher U, Lobmaier JS, Wirthmüller U, Knoch D (2017) Men's preferences for women's body odours are not associated with human leucocyte antigen. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20171830.
- Qiao Z, Powell JE, Evans DM (2018) MHC-Dependent Mate Selection within 872 Spousal Pairs of European Ancestry from the Health and Retirement Study. *Genes* **9**, 53.
- Roberts SC, Gosling LM, Carter V, Petrie M (2008) MHC-correlated odour preferences in humans and the use of oral contraceptives. *Proceedings of the Royal Society B: Biological Sciences* **275**, 2715-2722.
- Rosenberg LT, Cooperman D, Payn R (1983) HLA and mate selection. *Immunogenetics* **17**, 89-93.
- Sans M, Alvarez I, Callegari-Jacques S, Salzano F (1994) Genetic similarity and mate selection in Uruguay. *Journal of biosocial science* **26**, 285-289.
- Santos PSC, Schinemann JA, Gabardo J, da Graça Bicalho M (2005) New evidence that the MHC influences odor perception in humans: a study with 58 Southern Brazilian students. *Hormones and Behavior* **47**, 384-388.
- Saphire-Bernstein S, Larson CM, Gildersleeve KA, Fales MR, Pillsworth EG, Haselton MG (2017) Genetic compatibility in long-term intimate relationships: partner similarity at major histocompatibility complex (MHC) genes may reduce in-pair attraction. *Evolution and Human Behavior* **38**, 190-196.
- Sorokowska A, Pietrowski D, Schäfer L, Kromer J, Schmidt AH, Sauter J, Hummel T, Croy I (2018) Human Leukocyte Antigen similarity decreases partners' and strangers' body odor attractiveness for women not using hormonal contraception. *Hormones and Behavior* **106**, 144-149.
- Thornhill R, Gangestad SW, Miller R, Scheyd G, McCollough JK, Franklin M (2003) Major histocompatibility complex genes, symmetry, and body scent attractiveness in men and women. *Behavioral Ecology* **14**, 668-678.
- Wedekind C, Furi S (1997) Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proceedings of the Royal Society of London B: Biological Sciences* **264**, 1471-1479.
- Wedekind C, Seebeck T, Bettens F, Paepke AJ (1995) MHC-dependent mate preferences in humans. *Proceedings of the Royal Society of London B: Biological Sciences* **260**, 245-249.
- Wu K, Chen C, Moyzis RK, Nuno M, Yu Z, Greenberger E (2018) More than skin deep: Major histocompatibility complex (MHC)-based attraction among Asian American speed-daters. *Evolution and Human Behavior* **39**, 447-456.
- Zaidi AA, White JD, Mattern BC, Liebowitz CR, Puts DA, Claes P, Shriver MD (2019) Facial masculinity does not appear to be a condition-dependent male ornament and does not reflect MHC heterozygosity in humans. *Proceedings of the National Academy of Sciences* **116**, 1633-1638.

Details of data extraction and effect size calculations for MHC mate choice studies

(-) indicates preference/choice for dissimilarity in the original study and (+) indicates preference/choice for similarity in the original study. Final effect sizes were calculated to reflect the direction of original preference (- for dissimilarity, + for similarity).

Publication	Original test statistic	Conversion method used	corr	N target	N rater	Details
Chaix et al. 2008	p=0.015 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.318	28	28	Calculated r from p-value. European pairs, SNP data.
Chaix et al. 2008	p=0.23 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.163	27	27	Calculated r from p-value. Yoruban pairs, SNP data.
Chaix et al. 2008	p=0.084 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.229	28	28	Calculated r from p-value for 6 HLA classical loci. European pairs.
Chaix et al. 2008	p=0.412 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.112	27	27	Calculated r from p-value for 6 HLA classical loci. Yoruban pairs.
Cretu-Stancu et al. 2018	p=0.702 (+)	calculated from pes(), program compute.es; 1-sided p-value	0.024	239	239	Calculated r from p-value. MHC region size = 3.6 Mb. Netherland pairs.
Cretu-Stancu et al. 2018	p=0.696 (+)	calculated from pes(), program compute.es; 1-sided p-value	0.023	239	239	Calculated r from p-value. MHC region size = 6.7 Mb. Netherland pairs.
Cretu-Stancu et al. 2018	p=0.740 (+)	Calculated from pes(), program compute.es; 1-sided p-value	0.029	239	239	Table 1 (only using classical HLA markers), calculated r from p-value. Netherland pairs.
Dandine-Roulland et al. 2019	p=0.343, p=.308 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.116	36	36	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. Belgium pairs.
Dandine-Roulland et al. 2019	p=0.068, p=0.055 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.143	86	86	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. Ireland pairs.
Dandine-Roulland et al. 2019	p=0.170, p=0.248 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.086	108	108	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. Germany pairs.
Dandine-Roulland et al. 2019	p=0.040, p=0.057 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.080	302	302	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. Netherland pairs.

Dandine-Roulland et al. 2019	p=0.157, p=0.124 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.093	126	126	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. UK pairs.
Dandine-Roulland et al. 2019	p=0.179, p=0.157 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.117	69	69	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. Spain pairs.
Dandine-Roulland et al. 2019	p=0.575, p=0.727 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.026	156	156	Calculated r from p-value(b) from Table 1 (between spouses and non-spouses) and Table 2 (between spouses and opposite-sex non-spouses) and then took the mean. Israel pairs.
Derti et al. 2010	p=0.351 (-)	Calculated from pes(), program compute.es; 1-sided p-value	-0.055	24	24	Calculated r from p-value. European pairs.
Garver-Apgar et al. 2006	t=1.26, n=48	$r = t/\sqrt{t^2 + df}$	-0.183	48	48	Calculated r from t-value. t(df) taken from p. 833. No HC USA pairs.
Garver-Apgar et al. 2006	r(37)=-0.12, r(42)=-0.13	NA	-0.125	48	48	Correlation between MHC similarity and male sexual responsiveness and satisfaction. Calculated the weighted mean of: male's sexual responsivity to partner and satisfaction with sexual responsiveness of partner. N-target and N-rater taken from p. 833. Male USA rater.
Garver-Apgar et al. 2006	r(41)=-0.35, r(43)=-0.30	NA	-0.324	48	48	Correlation between MHC similarity and female sexual responsiveness. Calculated the weighted mean of: female's sexual responsivity to partner and satisfaction with sexual responsiveness of partner. N-target and N-rater taken from p. 832. Female no HC USA rater.
Hedrick and Black 1997	HLA-A Chi-square=0.803; B Chi-square=0.60; n=194	weighted mean of r ($(.0647+.0556)/2$) calculated from chies(), program compute.es	0.06	194	194	For each locus, Chi-square values were calculated comparing sharing 0 to 1+2 alleles, r was computed from Chi-square values, and we took the mean of r as the effect size. No HC South Amerindian pairs.
Ihara et al. 2000	Tohoku Chi-square=0.9705; 8JW Chi-square=0.4495, n=138,131; weighted mean of r ($(138*.0839+131*.0586)/138+131$)	calculated from chies(), program compute.es	0.0716	269	269	Chi-square values were calculated comparing sharing 0 to 1+2 alleles, and the weighted mean r was used to calculate the mean effect size. Japan pairs.
Israeli et al. 2014	Chi-square=46.337	calculated from chies(), program compute.es	0.3879	308	308	Chi-square values were calculated comparing sharing 0 to 1+2 alleles. Data from Table 1 did not provide number of pairs with 0 antigens shared, but that was inferred from the sample sizes given for 1 and 2 matches and the total sample size. Israel pairs.

Israeli et al. 2014	Chi-square=228.472, 51.477; weighted mean of $r((.4775+.2267)/2)$	calculated from chies(), program compute.es	0.3521	1002	1002	Chi-square values were calculated comparing sharing 0 to 1+2 alleles. We calculated the weighted mean effect size for Class I and Class II loci. Israel pairs.
Jacob et al. 2002	paternal t=4.0, df=21; maternal t=-1.34, df=17	$(r=t/\sqrt{t^2+df})$	0.2226	6	22	Calculated r from t-values from paternally and maternally inherited allele effect sizes. Similarity at paternally inherited alleles was correlated with preference. No HC, Germany females.
Jin et al. 1995	weighted mean of A, B and DR loci p of Chi-squares: p=0.17, 0.48, 0.28 (n=542,521,466)	calculated from pes(), program compute.es	0.0467	510	510	We calculated r from p-values and then calculated the weighted mean effect size for A, B, and DR loci using results from corrected allele table test (recommended by authors for larger datasets). Sample size 510 (mean of 542,521,466). European pairs.
Kromer et al. 2016	r=-0.126	Calculated using mes() from compute.es.	-0.126	248	248	Sexual satisfaction raw mean scores from Table S1 were converted to r effect sizes. Then took the mean of Class I loci. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. The N of target and rater were taken from Table S1. Female raters.
Kromer et al. 2016	r=-0.024	Calculated using mes() from compute.es.	-0.024	248	248	Sexual satisfaction raw mean scores from Table S1 were converted to r effect sizes. Then took the mean of Class I loci. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Male raters.
Kromer et al. 2016	r=0.026	Calculated using mes() from compute.es.	0.026	248	248	Sexual satisfaction raw mean scores from Table S1 were converted to r effect sizes. Then took the mean of Class II loci. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Female raters.

Kromer et al. 2016	r=0.014	Calculated using mes() from compute.es.	0.014	248	248	Sexual satisfaction raw mean scores from Table S1 were converted to r effect sizes. Then took the mean of Class II loci. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Male raters.
Kromer et al. 2016	r=-0.050,-0.145,-0.139	Calculated using mes() from compute.es.	-0.111	248	248	Average effect size across loci from raw mean scores, from Table S1. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Odour preference for MHC class I female raters.
Kromer et al. 2016	r=-0.014,-0.095,-0.072	Calculated using mes() from compute.es.	-0.06	248	248	Average effect size across loci from raw mean scores, from Table S1. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Odour preference for MHC class I male raters.
Kromer et al. 2016	r=0.019,0.007,0.011	Calculated using mes() from compute.es.	0.0123	248	248	Average effect size across loci from raw mean scores, from Table S1. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Odour preference for MHC class II female raters.
Kromer et al. 2016	r=0.022,0.051,-0.007	Calculated using mes() from compute.es.	0.022	248	248	Average effect size across loci from raw mean scores, from Table S1. Pref for dissimilarity = (-) effect size. This study did not report methods to control for population stratification in the sample, but as they used the most common allele frequencies in German populations for data imputation, I assume the population is of German ancestry. Odour preference for MHC class II male raters.

Nordlander et al. 1983	Chi-square for A and B loci=1.64, 0.39, n=826; weighted mean of $r((.0446+.0217)/2)$	calculated from chies(), program compute.es	0.0132	826	826	Calculated the weighted mean across A and B loci after converting Chi-square values to r. Sweden pairs.
Ober et al. 1997	Chi-square=7.90, n=411	calculated from chies(), program compute.es	-0.139	411	411	Calculated r from Chi-square using the program compute.es. European pairs.
Pollack et al. 1982	Chi-square: 0.0002, 0.0112, 0.0474, 1.0276; n: 47.4,32.7,8.2,38.2; weighted mean of r=0.0021, 0.0185, 0.076, 0.164	calculated from chies(), program compute.es	0.06	61	61	We took the sum of the shared allele frequencies at each antigen for each locus (A,B,C,DR) and calculated Chi-square (df1) values for obs vs exp sharing frequencies. From Chi-squared values we computed r values and then calculated the weighted mean of r. USA pairs.
Probst et al. 2017	t=-0.23, df=93	r = $t/\sqrt{(tsquared+df)}$; paired t-test	0.0238	42	94	Calculated r from t-value. Odour preferences of European males.
Qiao et al. 2018	p=0.829, p=0.218 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.017	872	872	Table 1. mean effect size of SNPs, from mean and median p-values. European pairs.
Qiao et al. 2018	p=0.258, p=0.500, p=0.211, p=0.419, p=0.040, p=0.235 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.028	872	872	Table 1. Calculated effect sizes of HLA gene and amino acid p-values, from mean and median p-values. Then took the mean effect size. European pairs.
Roberts et al. 2008	similar mean (sd) = 3.68(0.949); dissimilar mean = 3.80(0.885), n=40	http://www.lyonsmorris.com/ma1/	-0.065	79	40	From Table S3, for HC-using women including all shirts, we used the "desirability" measure, as this was more directly related to mate preferences than "intensity" or "pleasantness". We converted means and their standard deviations to r. UK female raters.
Roberts et al. 2008	similar mean (sd)=3.47(1.049); dissimilar mean=3.42(1.154), n=110	http://www.lyonsmorris.com/ma1/	0.0226	110	110	From Table 1, we used the mean of all women for all shirts (non-HC users) for desirability (we didn't include data from session 2 non-HC users because the same women were tested). We converted means and their standard deviations to r. UK female raters.
Rosenberg et al. 1983	Chi-square values: A:12.3,6.3; B:16.4,43.0, r: 0.2328, 0.1134, 0.2688, 0.2962	calculated from chies(), program compute.es	0.2193	1017	1017	We computed r from Chi-squared values across for HLA-A and for HLA-B for the two collection centers, and then calculated the weighted mean effect size. USA pairs.

Sans et al. 1994	Chi-square values for A, B, C: 0.7107, 9.783, 0.0739; r: 0.0441, 0.1635, 0.0142; n=366	calculated from chies(), program compute.es	0.0739	183	183	Chi-square values for real couples vs random pairs for HLA-A, B and C were calculated and then r effect sizes were derived and the average was taken for the 3 loci. Chi Square values compared groups sharing 0 vs 1/2 alleles where Total is Obs and Noncouples is Exp. Since the couples were selected based on paternity testing, indicating a degree of uncertainty in the genetic father, we chose to use the full dataset and not exclude couples where the putative father was not the genetic father of the child being tested. Uruguay pairs.
Santos et al. 2005	Chi-square=0.5341, n=369	http://www.lyonsmoris.com/ma1/	-0.038	29	29	We used Chi-squared goodness of fit test for the "pleasant" category for sweat olfactory session and only for females because detailed effect sizes for male smellers was not presented. 2x2 table (0 matching vs 2/3, pleasant vs unpleasant). Brazil females.
Saphire-Bernstein et al. 2017	r=-0.240, n=44; r=-0.136, n=43; r=-0.590, n=43; r=-0.399, n=33	Beta coefficients	-0.337	43	43	Weighted average of effect sizes taken from sexual responsivity, sexual adventurousness, satisfaction with partner as a turn-on, and in-pair attraction. N-target and N_rater were taken from Table 1. Couldn't find a way to extract effect sizes for males and females separately, as only contrasts were provided. Asian pairs.
Saphire-Bernstein et al. 2017	r=0.025, n=165; r=0.074, n=165; r=-0.008, n=114	Beta coefficients	0.0347	165	165	Weighted average of effect sizes taken from sexual responsivity, satisfaction with partner as a turn-on, and in-pair attraction. N-target and N_rater were taken from Table 1. USA pairs.
Sorokowska et al. 2018	d=0.5, n=24	Calculated r from Cohen's D (http://www.lyonsmoris.com/ma1/)	-0.243	47	24	Data from authors. Preference for odours from individuals with fewer shared alleles. No HC using female German raters.
Sorokowska et al. 2018	d=0.04, n=28	Calculated r from Cohen's D (http://www.lyonsmoris.com/ma1/)	-0.02	47	28	Data from authors. Preference for odours from individuals with fewer shared alleles. HC using female German raters.
Sorokowska et al. 2018	F(2,122)=0.03	Calculated r from F-statistic (http://www.lyonsmoris.com/ma1/)	0.022	52	47	Data from authors. Preference for odours from individuals with fewer shared alleles. Male German raters.
Thornhill et al. 2003	r=-0.033	NA	-0.033	48	77	The correlation coefficient r was extracted directly. Male USA raters.
Thornhill et al. 2003	r=0.025	NA	0.025	55	65	The correlation coefficient -r- was extracted directly from total number of alleles shared across 3 loci (A,B,DRB). No HC using female USA raters.

Wedekind and Furi 1997	r=-0.051	NA	-0.051	6	63	Pearson's r for men was estimated directly from Figure 2b using DataThief (B. Tummars, DataThief III. 2006 < http://datathief.org/ >). Male Switzerland raters.
Wedekind and Furi 1997	r=-0.11	NA	-0.11	6	25	Pearson's r for women not on birth control was estimated directly from Figure 2b using DataThief (B. Tummars, DataThief III. 2006 < http://datathief.org/ >). No HC using female Switzerland raters.
Wedekind and Furi 1997	r=0.0778	NA	0.0778	6	23	Pearson's r for women on birth control was estimated directly from Figure 2b using DataThief (B. Tummars, DataThief III. 2006 < http://datathief.org/ >). HC using female Switzerland raters.
Wedekind et al. 1995	p=0.02	calculated from pes(), program compute.es	0.4815	23	18	We calculated r from p-values for women taking hormonal contraception. Switzerland raters.
Wedekind et al. 1995	p=0.04	calculated from pes(), program compute.es	-0.335	38	31	We calculated r from p-values for women not taking HC. Switzerland raters.
Wu et al. 2018	-0.003	Mean of beta coefficients	-0.003	132	130	Table 4. Female mean effect sizes from beta coefficients for Date offer and Overall mate desirability for Class I. Positive coefficients indicate preferences for MHC-dissimilarity. No HC using female Asian raters.
Wu et al. 2018	-0.023	Mean of beta coefficients	-0.023	132	130	Table 4. Female mean effect sizes from beta coefficients for Date offer and Overall mate desirability for Class II. Positive coefficients (B) indicate preferences for MHC-dissimilarity. No HC using female Asian raters.
Wu et al. 2018	0.002	Mean of beta coefficients	0.002	130	132	Table 4. Male mean effect sizes from beta coefficients for Date offer and Overall mate desirability for Class I. Positive coefficients (B) indicate preferences for MHC-dissimilarity. Male Asian raters.
Wu et al. 2018	0.002	Mean of beta coefficients	0.002	130	132	Table 4. Male mean effect sizes from beta coefficients for Date offer and Overall mate desirability for Class II. Positive coefficients (B) indicate preferences for MHC-dissimilarity. Male Asian raters.
Wu et al. 2018	0.0353	Mean of beta coefficients	0.0353	132	130	Table 4. Female mean effect sizes from beta coefficients for body scent attractiveness for Class I. Positive coefficients (B) indicate preferences for MHC-dissimilarity. No HC using Asian female raters.
Wu et al. 2018	0.0115	Mean of beta coefficients	0.0115	132	130	Table 4. Female mean effect sizes from beta coefficients for body scent attractiveness for Class II. Positive coefficients (B) indicate preferences for MHC-dissimilarity. No HC using Asian female raters.
Wu et al. 2018	-0.0233	Mean of beta coefficients	-0.023	130	132	Table 4. Male mean effect sizes from beta coefficients for body scent attractiveness for Class I. Positive coefficients (B) indicate preferences for MHC-dissimilarity. Male Asian raters.

Wu et al. 2018	0.0323	Mean of beta coefficients	0.0323	130	132	Table 4. Male mean effect sizes from beta coefficients for body scent attractiveness for Class II. Positive coefficients (B) indicate preferences for MHC-dissimilarity. Male Asian raters.
----------------	--------	---------------------------	--------	-----	-----	---

Details of data extraction and effect size averaging for neutral markers from MHC studies

(-) indicates preference/choice for dissimilarity in the original study and (+) indicates preference/choice for similarity in the original study. Final effect sizes were calculated to reflect the direction of original preference (- for dissimilarity, + for similarity).

Author	Genome-wide test statistic (control for socio-demographic factors influencing mate choice)	Conversion Method Used	corr neutral	N target	N rater	Details
Cretu-Stancu et al. 2018	North: N=95 pairs, p=0.48 (pairs less similar); Center: N=69 pairs, p=0.545 (pairs less similar); South: N=47 pairs, p=0.431 (pair less similar = neg)	Individual effect sizes calculated from pes(), program compute.es; 1-sided p-value. Weighted mean = $\text{Sum}(N_i \cdot r_i) / \text{Sum}(N_i)$	-0.00895	239	239	Took the weighted mean of 15 effect sizes by province.
Dandine-Roulland et al. 2019	p=0.033, p=0.975 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.126	36	36	Mean p-value(a) from Table 1 and Table 2. Belgium pairs.
Dandine-Roulland et al. 2019	p=0.032, p=0.372 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.1155	86	86	Mean p-value(a) from Table 1 and Table 2. Ireland pairs.
Dandine-Roulland et al. 2019	p=0.009 (+), p=0.030 (-)	Calculated from pes(), program compute.es; 2-sided p-value	0.015	108	108	Mean p-value(a) from Table 1 and Table 2. Germany pairs.
Dandine-Roulland et al. 2019	p=0.001, p=0.148 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.096	302	302	Mean p-value(a) from Table 1 and Table 2. Netherlands pairs.
Dandine-Roulland et al. 2019	p=0.696, p=0.991 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.013	126	126	Mean p-value(a) from Table 1 and Table 2. UK pairs.
Dandine-Roulland et al. 2019	p=0.734, p=0.863 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.022	69	69	Mean p-value(a) from Table 1 and Table 2. Spain pairs.
Dandine-Roulland et al. 2019	p=0.0001, p=0.178 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.147	156	156	Mean p-value(a) from Table 1 and Table 2. Israel pairs.
Qiao et al. 2018	p=0.0081 (-)	Calculated from pes(), program	-0.063	872	872	Calculated r from p-value. European pairs.

		compute.es; 2-sided p-value				
Chaix et al. 2008	p=0.739 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.045	28	28	Calculated r from p-value. European pairs.
Chaix et al. 2008	p=0.001 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.429	27	27	Calculated r from p-value. Yoruban pairs.
Author	Genome-wide test statistic (is MHC unique for real pairs?)	Conversion Method Used	corr MHC unique	N target	N rater	Details
Cretu-Stancu et al. 2018	genomic size: N=239, p=0.08 ; total number of markers: N=239, p=0.02 (MHC more dissimilar than genome-wide = neg)	Individual effect sizes calculated from pes(), program compute.es; 1-sided p-value. Weighted mean = $\text{Sum}(N_i * r_i) / \text{Sum}(N_i)$	-0.079	239	239	Weighted mean of regions based on genomic size and total number of markers in the region.
Dandine-Roulland et al. 2019	p=0.264, p=0.241 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.135	36	36	Mean p-value(c) from Table 1 and Table 2. Belgium pairs.
Dandine-Roulland et al. 2019	p=0.003, p=0.052 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.186	86	86	Mean p-value(c) from Table 1 and Table 2. Ireland pairs.
Dandine-Roulland et al. 2019	p=0.073, p=0.241 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.101	108	108	Mean p-value(c) from Table 1 and Table 2. Germany pairs.
Dandine-Roulland et al. 2019	p=0.010, p=0.051 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.092	302	302	Mean p-value(c) from Table 1 and Table 2. Netherlands pairs.
Dandine-Roulland et al. 2019	p=0.080, p=0.062 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.1135	126	126	Mean p-value(c) from Table 1 and Table 2. UK pairs.
Dandine-Roulland et al. 2019	p=0.120, p=0.079 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.1405	69	69	Mean p-value(c) from Table 1 and Table 2. Spain pairs.
Dandine-Roulland et al. 2019	p=0.625, p=0.657 (+)	Calculated from pes(), program compute.es; 2-sided p-value	0.0265	156	156	Mean p-value(c) from Table 1 and Table 2. Israel pairs.
Chaix et al. 2008	p=0.004, p=0.001 (-)	Calculated from pes(), program compute.es; 2-sided p-value.	-0.3975	28	28	Mean effect size for similar sized genomic window and similar recombination rate. Only 0.4% and 0.1% of the genomic windows had higher dissimilarity than MHC. European pairs.
Chaix et al. 2008	p=0.09, p=0.17 (+)	Calculated from pes(), program compute.es; 2-sided p-value.	0.2075	27	27	Mean effect size for similar sized genomic window and similar recombination rate. 9% and 17% of the genomic windows had higher similarity than MHC. Yoruban pairs.
Derti et al. 2010	p=0.016, p=0.035 (-)	Calculated from pes(), program compute.es; 2-sided p-value	-0.2815	24	24	From Table S4 for unphased Hap3 Europeans. Mean effect size for similar sized genomic window and similar recombination rate. 1.6% and 3.5% of the genomic windows had higher dissimilarity than MHC. European pairs.

List of relevant papers excluded from the present study

Publication	Details
Giphart and D'Amaro 1983	Some evidence for disassortative mating (different antigens at 4 genes are more frequent than expected out of 30 genes), but no way to extract effect size because the data for the pairwise tests is not available.
Khankhanian et al. 2010	This study was excluded because the sample population consisted of couples with a child affected by multiple sclerosis (MS). As MS is a complex genetic disease with strong associations with MHC class II genes, the sample population has higher frequencies of specific MHC risk alleles and is not a fair representation of the general population.
Laurent et al. 2012	This study did not present full test statistics for HLA region. It presented genes showing extreme similarity or dissimilarity based on 2 testing approaches (permutation and intergenic). Therefore, we could not calculate effect sizes.
Zaidi et al. 2019	This study tested if height and facial masculinity are correlated with MHC heterozygosity, but did not test if these 'male ornaments' were correlated with MHC dissimilarity.

Supplementary Tables

Table S1. Estimated effect of human genomic MHC similarity on mating using moderators.

The estimates are derived from the univariate model containing the moderator as a fixed effect and study ID and sampling error (variance) as random effects. The intercept has been removed from the model, so each trait is tested against zero effect. Significant terms are in bold.

Fixed factor	Category	N	Estimate	l-95% CI	u-95% CI	p-value	R ² %
Year	(Intercept)	17	-0.027	-0.077	0.024	0.302	0.00%
	Year		0.000	-0.023	0.024	0.981	
Class	Both	13	-0.031	-0.080	0.019	0.222	0.00%
	Class 1	2	-0.001	-0.132	0.131	0.994	
	Class 2	2	-0.011	-0.142	0.121	0.875	
Rater	Female	2	-0.013	-0.144	0.118	0.846	0.00%
	Male	2	0.002	-0.130	0.134	0.976	
	Pairs	13	-0.031	-0.080	0.019	0.222	
Number MHC loci	(Intercept)	17	0.201	-0.530	0.933	0.589	0.00%
	Number MHC loci		-0.027	-0.113	0.059	0.541	
Hormonal Contraception	Male	2	0.002	-0.130	0.134	0.976	0.00%
	No HC	2	-0.013	-0.144	0.118	0.846	
	Unknown	13	-0.031	-0.080	0.019	0.222	
Population	Asian	4	-0.006	-0.145	0.134	0.938	0.00%
	Belgium	1	-0.089	-0.443	0.265	0.622	
	European	3	-0.046	-0.163	0.072	0.445	
	Germany	1	-0.059	-0.272	0.154	0.589	
	Ireland	1	-0.116	-0.351	0.118	0.330	
	Israel	1	0.054	-0.130	0.237	0.568	
	Netherlands	3	-0.023	-0.128	0.082	0.670	
	Spain	1	-0.090	-0.349	0.169	0.495	
	UK	1	-0.066	-0.266	0.134	0.519	
	Yoruban	1	0.185	-0.229	0.599	0.381	
Region code	EUR	11	-0.041	-0.106	0.024	0.216	0.00%
	NAF	1	0.178	-0.231	0.586	0.394	
	SEA	4	-0.006	-0.128	0.116	0.929	
	WAS	1	0.058	-0.114	0.230	0.511	

Background	(Intercept)	12	-0.033	-0.108	0.041	0.384	0.00%
	Background		0.276	-0.373	0.926	0.405	
MHC background	(Intercept)	12	0.035	-0.062	0.131	0.484	19.00%
	MHC background		0.778	0.102	1.453	0.024	
ln(N of rater)	(Intercept)	17	-0.114	-0.365	0.138	0.375	0.00%
	ln(N of rater)		0.016	-0.029	0.061	0.487	
Permutations	(Intercept)	17	-0.093	-0.347	0.161	0.473	0.00%
	Permutations		0.017	-0.046	0.080	0.599	
Phi-hat cryptic relatedness	(Intercept)	13	0.012	-0.057	0.082	0.734	100.00%
	Phi-hat cryptic relatedness		-0.686	-1.554	0.183	0.122	
Span of MHC (Mb)	(Intercept)	17	-0.052	-0.185	0.082	0.448	0.00%
	Span of MHC (Mb)		0.006	-0.025	0.038	0.693	
Intercept-only model	(Intercept)	17	-0.027	-0.067	0.013	0.191	

Table S2. Estimated effect of human MHC similarity on relationship satisfaction using moderators.

The estimates are derived from the univariate model containing the moderator as a fixed effect and study ID and sampling error (variance) as random effects. The intercept has been removed from the model, so each trait is tested against zero effect. Significant terms are in bold.

Fixed factor	Category	N	Estimate	l-95% CI	u-95% CI	p-value	R ² %
Year	(Intercept)	9	-0.090	-0.159	-0.021	0.011	89.45%
	Year		0.018	0.000	0.035	0.051	
Class	Both	5	-0.136	-0.284	0.012	0.072	0.00%
	Class 1	2	-0.075	-0.243	0.093	0.379	
	Class 2	2	0.020	-0.148	0.188	0.815	
Rater	Female	3	-0.124	-0.306	0.059	0.184	0.00%
	Male	3	-0.054	-0.236	0.129	0.564	
	Pairs	3	-0.093	-0.297	0.110	0.369	
Hormonal Contraception	Male	1	-0.126	-0.424	0.172	0.408	85.02%
	No HC	2	-0.261	-0.472	-0.050	0.015	
	Unknown	6	-0.031	-0.094	0.031	0.326	
Population	Asian	1	-0.448	-0.841	-0.054	0.026	0.00%
	Germany	4	-0.028	-0.340	0.285	0.862	
	USA	4	-0.089	-0.334	0.157	0.479	
Region code	EUR	4	-0.028	-0.340	0.285	0.862	0.00%
	NAM	4	-0.089	-0.334	0.157	0.479	
	SEA	1	-0.448	-0.841	-0.054	0.026	
ln(N of rater)	(Intercept)	9	-0.720	-1.242	-0.199	0.007	84.81%
	ln(N of rater)		0.128	0.029	0.228	0.012	
Intercept-only model	(Intercept)	9	-0.078	-0.180	0.023	0.131	

Table S3. Estimated effect of human MHC similarity on odour preferences using moderators.

The estimates are derived from the univariate model containing the moderator as a fixed effect and study ID and sampling error (variance) as random effects. The intercept has been removed from the model, so each trait is tested against zero effect.

Fixed factor	Category	N	Estimate	l-95% CI	u-95% CI	p-value	R ² %
Year	(Intercept)	23	-0.020	-0.080	0.040	0.513	10.23%
	Year		0.000	-0.008	0.008	0.989	
Class	Both	11	-0.008	-0.108	0.092	0.877	0.00%
	Class 1	8	-0.058	-0.123	0.008	0.083	
	Class 2	4	0.019	-0.053	0.091	0.608	
Rater	Female	15	-0.029	-0.086	0.029	0.327	15.78%
	Male	8	-0.009	-0.075	0.057	0.790	
Number MHC loci	(Intercept)	23	-0.063	-0.156	0.030	0.185	0.00%
	Number MHC loci		0.010	-0.009	0.029	0.309	
Hormonal Contraception	Male	6	0.004	-0.095	0.103	0.939	0.00%
	No HC	8	-0.026	-0.113	0.062	0.563	
	HC	4	0.034	-0.130	0.198	0.682	
	Unknown	5	-0.035	-0.096	0.027	0.273	
Population	Asian	4	0.014	-0.073	0.101	0.751	0.00%
	Brazil	1	-0.038	-0.422	0.346	0.846	
	European	1	0.024	-0.290	0.338	0.882	
	Germany	8	-0.039	-0.098	0.019	0.191	
	Switzerland	5	-0.030	-0.275	0.215	0.809	
	UK	2	-0.014	-0.159	0.131	0.852	
	USA	2	-0.002	-0.201	0.197	0.985	
Region code	EUR	16	-0.034	-0.086	0.019	0.207	0.00%
	NAM	2	-0.002	-0.201	0.197	0.985	
	SAM	1	-0.038	-0.422	0.346	0.846	
	SEA	4	0.014	-0.073	0.101	0.751	
ln(N of rater)	(Intercept)	23	-0.027	-0.326	0.271	0.859	0.00%
	ln(N of rater)		0.001	-0.059	0.061	0.963	
Intercept-only model	(Intercept)	23	-0.020	-0.064	0.023	0.360	

Table S4. Estimated effect of non-HC using human MHC similarity on odour preferences using moderators.

The estimates are derived from the univariate model containing the moderator as a fixed effect and study ID and sampling error (variance) as random effects. The intercept has been removed from the model, so each trait is tested against zero effect. Significant terms are in bold.

Fixed factor	Category	N	Estimate	l-95% CI	u-95% CI	p-value	R ² %
Year	(Intercept)	19	-0.048	-0.116	0.019	0.163	0.00%
	Year		0.004	-0.005	0.014	0.356	
Class	Both	8	-0.031	-0.146	0.085	0.602	0.00%
	Class 1	7	-0.059	-0.126	0.007	0.081	
	Class 2	4	0.019	-0.053	0.091	0.608	
Rater	Female	11	-0.038	-0.099	0.024	0.230	0.00%
	Male	8	-0.009	-0.075	0.057	0.790	
Number MHC loci	(Intercept)	19	-0.075	-0.172	0.023	0.132	0.00%
	Number MHC loci		0.012	-0.008	0.031	0.253	
Hormonal Contraception	Male	6	0.004	-0.095	0.103	0.939	0.00%
	No HC	8	-0.026	-0.113	0.062	0.563	
	Unknown	5	-0.035	-0.096	0.027	0.273	
Population	Asian	4	0.014	-0.073	0.101	0.751	0.00%
	Brazil	1	-0.038	-0.422	0.346	0.846	
	European	1	0.024	-0.290	0.338	0.882	
	Germany	7	-0.040	-0.100	0.020	0.191	
	Switzerland	3	-0.309	-0.615	-0.003	0.048	
	UK	1	0.023	-0.167	0.212	0.815	
	USA	2	-0.002	-0.201	0.197	0.985	
Region code	EUR	12	-0.041	-0.097	0.014	0.142	0.00%
	NAM	2	-0.002	-0.201	0.197	0.985	
	SAM	1	-0.038	-0.422	0.346	0.846	
	SEA	4	0.014	-0.073	0.101	0.751	
ln(N of rater)	(Intercept)	19	-0.214	-0.615	0.187	0.295	0.00%
	ln(N of rater)		0.038	-0.042	0.119	0.351	
Intercept-only model	(Intercept)	19	-0.024	-0.069	0.021	0.289	

Table S5. Estimated effect of human MHC similarity on mate selection using moderators.

The estimates come from the univariate model containing the moderator as a fixed effect and study ID and sampling error (variance) as random effects. The intercept has been removed from the models with factor moderators, so each trait is tested against zero effect. Significant terms are in bold.

Fixed factor	Category	N	Estimate	l-95% CI	u-95% CI	p-value	R ² %
Year	(Intercept)	55	0.004	-0.060	0.067	0.913	0.00%
	Year		-0.002	-0.006	0.003	0.494	
Class	Both	31	-0.009	-0.076	0.059	0.795	0.00%
	Class 1	16	0.038	-0.044	0.119	0.364	
	Class 2	8	0.097	-0.006	0.200	0.064	
Rater	Female	16	-0.068	-0.171	0.035	0.197	7.17%
	Male	13	-0.033	-0.139	0.074	0.550	
	Pairs	26	0.042	-0.025	0.109	0.218	
Number MHC loci	(Intercept)	55	0.080	-0.030	0.190	0.153	8.10%
	Number MHC loci		-0.015	-0.036	0.006	0.163	
Hormonal Contraception	Male	9	-0.034	-0.160	0.093	0.602	3.15%
	No HC	13	-0.059	-0.171	0.053	0.301	
	Unknown	33	0.038	-0.030	0.107	0.272	
Population	Asian	9	-0.113	-0.312	0.087	0.268	2.77%
	Belgium	1	-0.009	-0.372	0.354	0.961	
	Brazil	1	-0.038	-0.489	0.413	0.869	
	European	6	-0.053	-0.178	0.073	0.411	
	Germany	12	-0.016	-0.163	0.131	0.833	
	Ireland	1	-0.037	-0.285	0.212	0.772	
	Israel	3	0.207	0.048	0.365	0.011	
	Japan	1	0.072	-0.193	0.336	0.595	
	Netherlands	3	0.027	-0.119	0.173	0.721	
	South Amerindian	1	0.060	-0.215	0.335	0.669	
	Spain	1	-0.010	-0.281	0.261	0.941	
	Sweden	1	0.013	-0.232	0.259	0.916	
	Switzerland	3	-0.297	-0.662	0.069	0.112	
	UK	2	0.016	-0.162	0.195	0.858	
	Uruguay	1	0.074	-0.203	0.351	0.601	
	USA	8	0.058	-0.068	0.184	0.368	
	Yoruban	1	0.238	-0.211	0.687	0.299	
Region code	EUR	30	-0.028	-0.097	0.040	0.413	30.14%

	NAF	1	0.227	-0.212	0.665	0.311	
	NAM	8	0.059	-0.055	0.172	0.311	
	NEA	1	0.072	-0.161	0.305	0.547	
	SAM	3	0.052	-0.109	0.214	0.525	
	SEA	9	-0.092	-0.270	0.086	0.312	
	WAS	3	0.205	0.068	0.343	0.003	
Background	(Intercept)	55	0.035	-0.158	0.228	0.722	0.00%
	Background		0.575	-0.219	1.369	0.156	
ln(N of rater)	(Intercept)	55	-0.293	-0.514	-0.072	0.009	28.96%
	ln(N of rater)		0.058	0.017	0.098	0.005	
Intercept-only model	(Intercept)	55	0.012	-0.046	0.070	0.685	

Table S6. Testing for publication bias using Egger's regression test of funnel plot asymmetry.

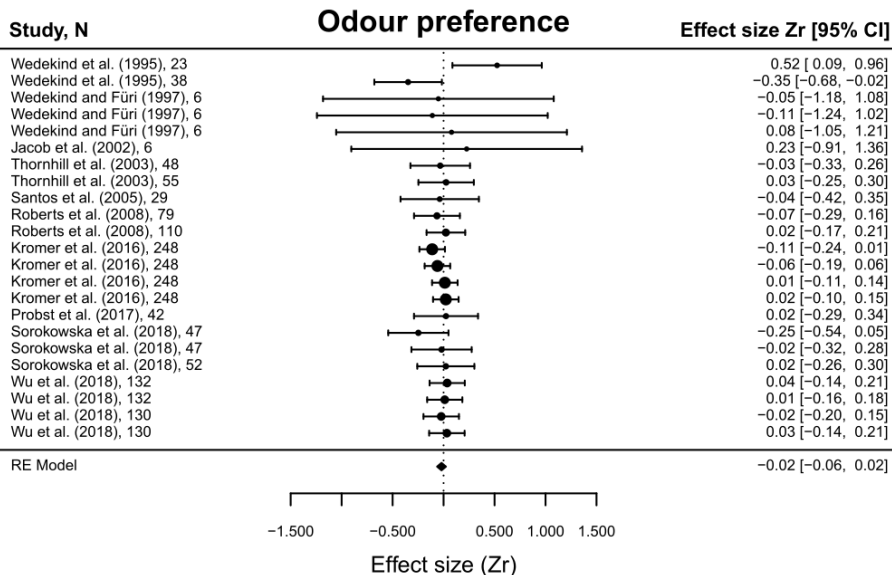
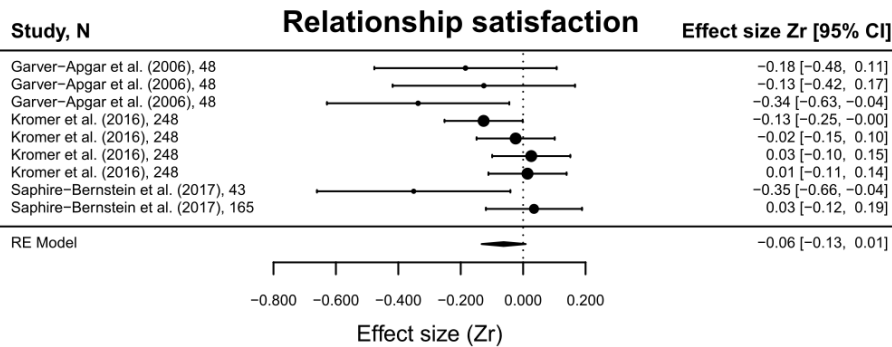
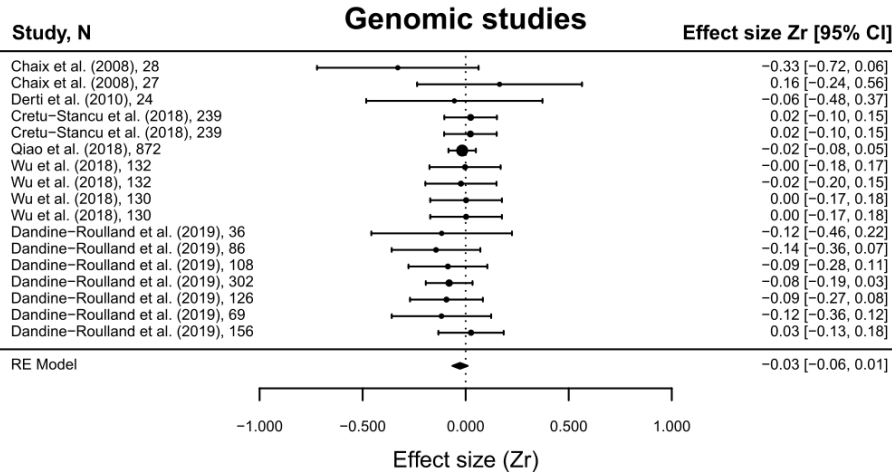
This tests for a linear regression between the standardized effect size (Z_r) and the standard error ($\sqrt{\text{variance}}$), using a random-effects meta-regression model. A significant value indicates that the standard error predicts the effect size, and this indicates asymmetry in the funnel plot and suggests publication bias may be present. N is number of effect sizes, z is the test statistic. Significant terms are in bold.

Egger's regression test for funnel plot asymmetry	N	z	p-value
Genomic studies	17	-0.892	0.373
Relationship satisfaction	9	-2.644	0.008
Odour preference	23	0.351	0.726
Odour preference (No HC)	19	-0.354	0.724
Mate selection (No HC)	55	-2.592	0.010

Supplementary Figures

Figure S1. Forest plots of standardized effect sizes by study for different datasets.

Points are proportional to inverse standard error. N, sample size.



Study, N

Mate selection

Effect size Zr [95% CI]

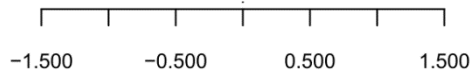
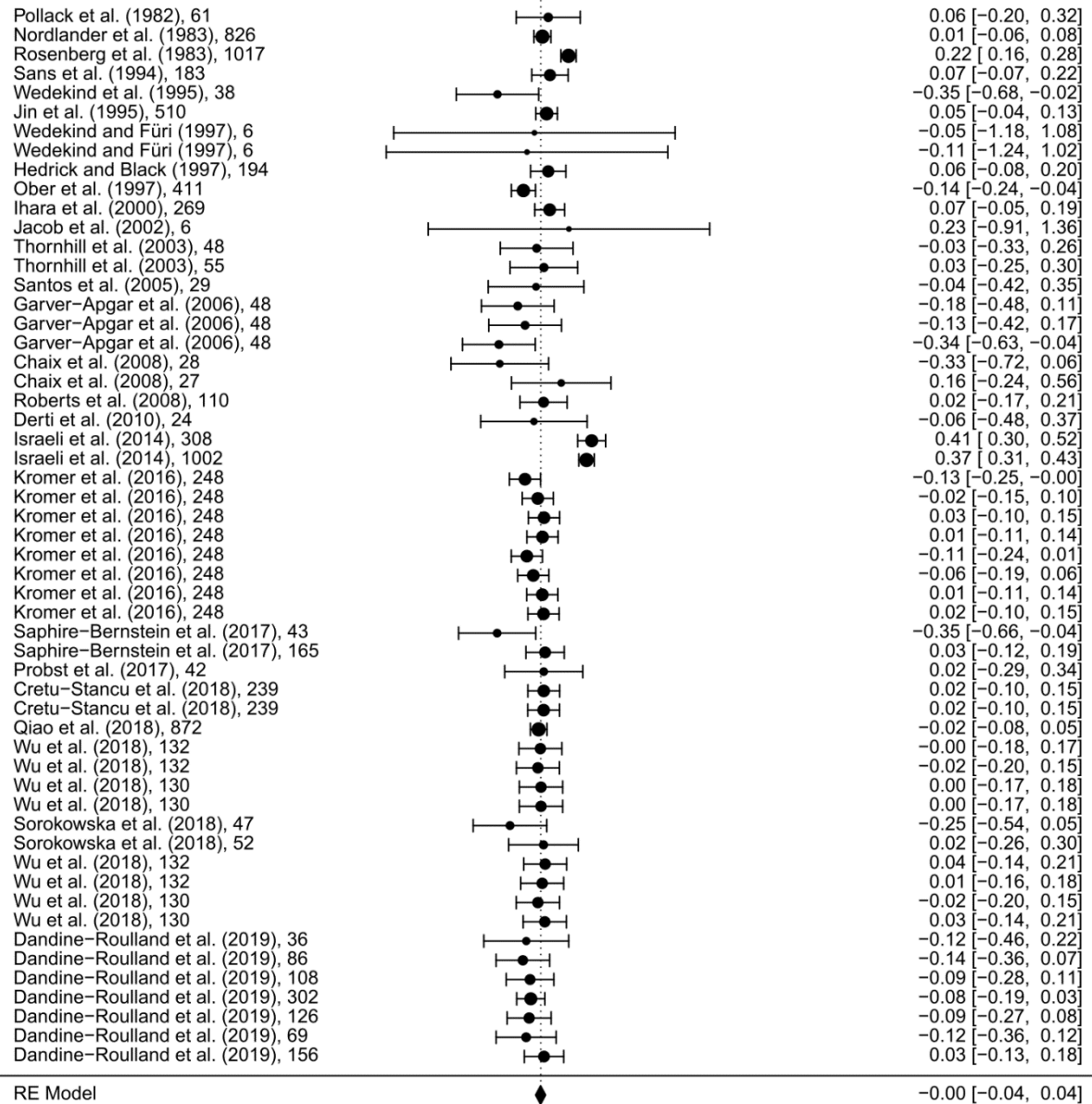
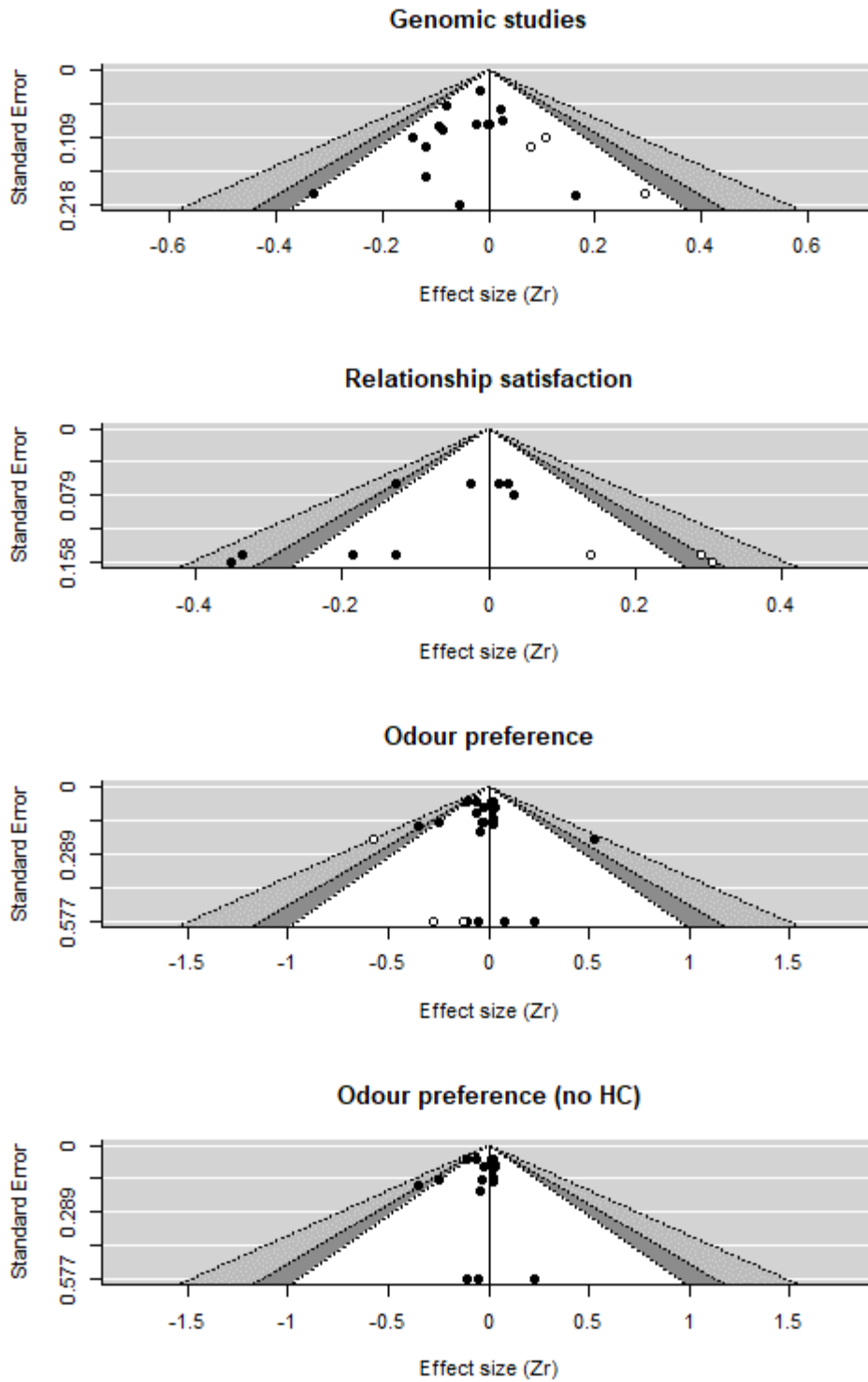
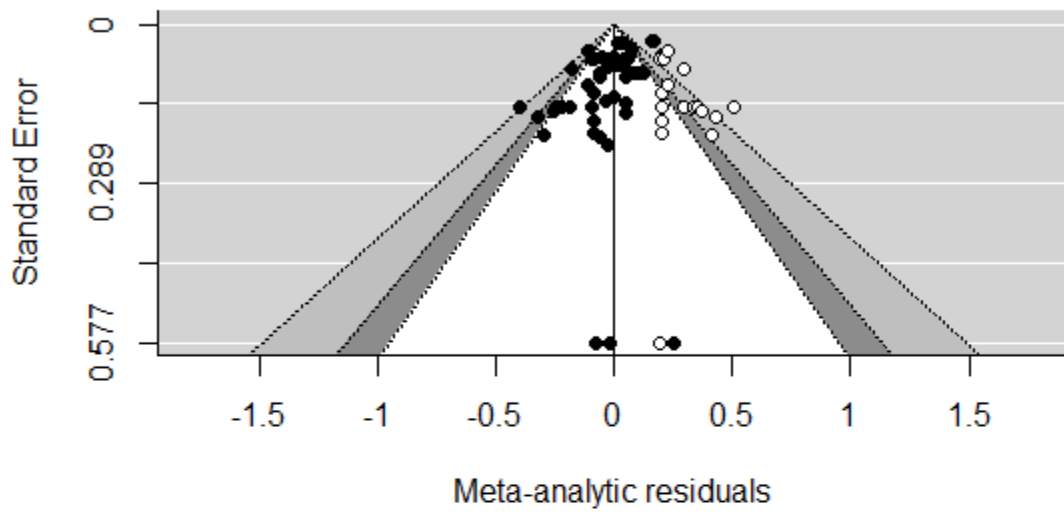


Figure S2. Trim and fill contour-enhanced funnel plots for the different datasets.

Black points are observed effect sizes and white points are filled in effects studies.



Mate selection using residuals



Mate selection

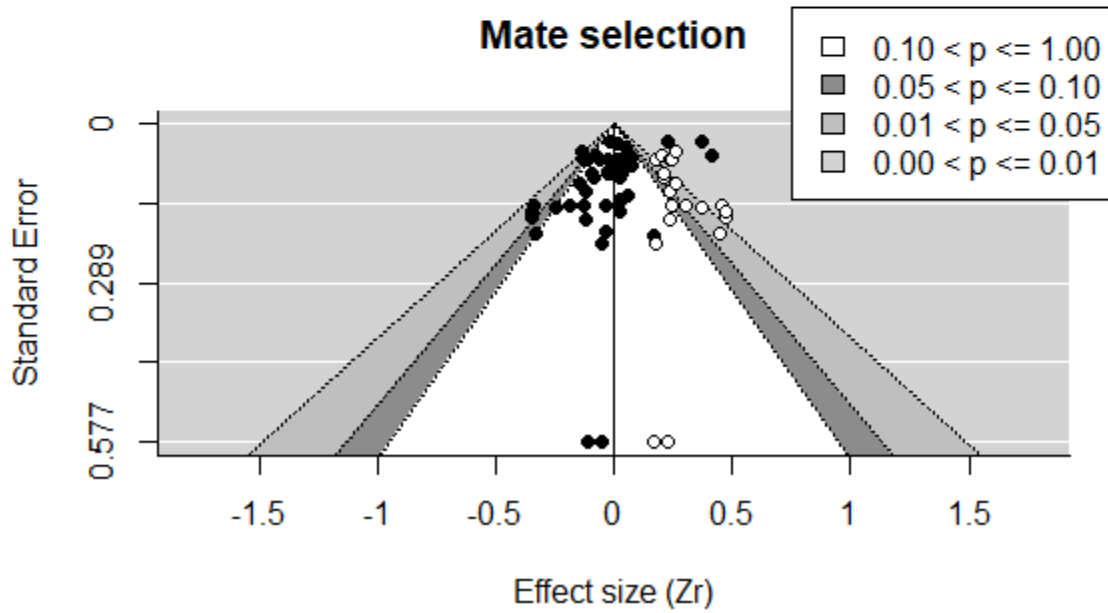


Figure S3. Meta-regression plot for all mate selection studies indicates evidence of publication bias.

The significant positive relationship between natural log number of raters and MHC effect size (Z_r) suggests that greater power diminishes the effect of MHC dissimilarity on mate selection ($N=55$, $\beta=0.058$, $p=0.005$). The colored line represent model predictions and grey regions represent 95% confidence intervals. Size of points is proportional to their weight (inverse SE).

