Evolution and Human Behavior 39 (2018) 447–456

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens

More than skin deep: Major histocompatibility complex (MHC)-based attraction among Asian American speed-daters

Karen Wu[a,*], Chuansheng Chen[b], Robert K. Moyzis[c,d], Michelle Nuno[e], Zhaoxia Yu[e], Ellen Greenberger[b]

[a] Department of Psychology, 5151 State University Drive, California State University – Los Angeles, Los Angeles, CA 90032, United States
[b] Department of Psychology and Social Behavior, 4201 Social and Behavioral Sciences Gateway, University of California, Irvine, CA 92697, United States
[c] Department of Biological Chemistry, College of Medicine, University of California, Irvine, CA 92697, United States
[d] Institute of Genomics and Bioinformatics, University of California, Irvine, CA 92697, United States
[e] Department of Statistics, 2214 Bren Hall, University of California, Irvine, CA 92697, United States

ARTICLE INFO

Keywords:
Human mate selection
Major histocompatibility complex
Attraction
Speed-dating
Behavioral genetics

ABSTRACT

What might explain our instinctual attraction to certain individuals, aside from visible factors such as appearance? We examined possible biologically-driven selection for immunology genes, specifically preferences for Major Histocompatibility Complex (MHC)-dissimilarity, through the ecologically-valid method of speed-dating. Two-hundred-and-sixty-two single Asian Americans went on speed-dates (*N* observations = 2215) with participants of the other sex, making second date offers and rating each other on measures of mate desirability, facial attractiveness, and body scent attractiveness. Using a single nucleotide polymorphism (SNP) analysis, women, but not men, showed preferences for speed-dating partners based on MHC-complementarity. The direction of findings varied by single nucleotide polymorphism (SNP), such that SNPs closer to the major *HLA* (Human Leukocyte Antigen) genes supported dissimilarity preferences, whereas those farther away supported similarity preferences. The relative effects of MHC-based measures in comparison to an array of behavioral predictors were examined via random forests. Results indicated that for both men and women, the importance of MHC-based indices was comparable to that of a partner's self-reported personality attributes in predicting second date offers.

1. Introduction

"Choose your life's mate carefully. From this one decision will come 90 percent of all your happiness or misery." Although this piece of advice, by H. Jackson Brown Jr. (2012), may be exaggerated, it is well known that selecting a partner is one of the most consequential decisions in our lives. Our romantic partners have the potential to significantly help or hinder our long-term health and happiness. While satisfying marriages are linked to better health, including higher cancer survival rates (Goodwin, Hunt, & Samet, 1987), unhappy marriages are related to increased risks of negative health outcomes such as heart disease (Smith & Gallo, 1999) and shorter lifespans (Gottman & Silver, 1999). Thus, the process of mate selection, and all the factors that may influence our choice in partners, should be better understood.

Most previous research in the field of human mate selection has utilized self-report or hypothetical scenarios to measure individuals' preferences in romantic partners. Yet, recent speed-dating studies indicate that there is often little congruency between peoples' stated ideals or their hypothetical decisions and their attraction to potential partners after face-to-face interactions (Eastwick, Luchies, Finkel, & Hunt, 2014). One explanation for the incongruency between stated and revealed mate preferences is that people may lack awareness of the attributes to which they are attracted. It has been suggested that prior to in-person interactions, people may report partner preferences based on "cold" rational deliberation, whereas during in-person interactions, people may instead make decisions based upon their "hot" gut instinct and visceral reactions, thus resulting in the "hot-cold empathy gap" (Eastwick et al., 2014). Given the robust literature on the role of affect in decision-making (Peters, Västfjäll, Gärling, & Slovic, 2006), it is important to examine the basic instincts that humans may utilize when dating. Currently, it is still unclear what factors may play a role in instinctual attraction aside from visible attributes such as physical attractiveness (Eastwick & Finkel, 2008; Luo & Zhang, 2009) and race (Fisman, Iyengar, Kamenica, & Simonson, 2008). Perhaps detection systems for genetic compatibility (Yamazaki et al., 1988) of potential partners are at work. In this study, we examined possible major histocompatibility complex (MHC)-preferences through speed-dating.

The speed-dating method has several strengths: (1) high ecological

* Corresponding author.
E-mail address: Kwu25@calstatela.edu (K. Wu).

https://doi.org/10.1016/j.evolhumbehav.2018.04.001
Received 7 December 2017; Received in revised form 2 April 2018; Accepted 2 April 2018
1090-5138/ Published by Elsevier Inc.

validity because participants' behaviors have real-life consequences (Finkel & Eastwick, 2008); (2) the ability to capture subconscious factors such as "good genes" and revealed preferences (what people actually want), which often differ from stated preferences (what people say they want) (Finkel & Eastwick, 2008); (3) control of environmental variables (e.g., dating pool attributes such as age and race), and (4) the ability to obtain large amounts of data in a short time period, since each speed-dater evaluates numerous potential partners. With these strengths, speed-dating is the ideal method to use to examine a particularly enigmatic source of romantic attraction, that is, our genes. Below we will summarize previous literature on MHC-based attraction and present our hypotheses.

1.1. MHC-based attraction

Sexual strategies theory (Buss & Schmitt, 1993) posits that humans have evolved certain mechanisms that help them select appropriate mates, such as preferences for beauty, which is thought to be indicative of "good genes" or genetic fitness. However, another mechanism that has been well-documented in non-human species appears to be relevant to human mate selection as well: MHC-based attraction. Studies suggest that MHC-based attraction may operate through olfactory cues for both human and non-human animals, and additionally through facial attractiveness for humans (Havlicek & Roberts, 2009; Milinski, 2006). Mate-selection via MHC genes seems to be largely based upon the genetic-complementarity between two partners.

The MHC Class I region (containing the *HLA-A*, *HLA-B*, and *HLA-C* genes) and Class II region (containing the *HLA-DRB*, *HLA-DQA*, and *HLA-DQB* genes) encode proteins essential for pathogen defense (Unanue, Turk, & Neefjes, 2016). The *HLA* genes, while comprising a relatively small portion of the MHC region, are among the most polymorphic human genes known (Chaix, Cao, & Donnelly, 2008; Derti, Cenik, Kraft, & Roth, 2010; Qiao, Powell, & Evans, 2018; Zou et al., 2015), likely due to the constant onslaught of pathogens and/or parasites. It is believed that MHC-dissimilarity of parents increases the immunocompetence of offspring (i.e., heterozygote advantage), a phenomenon that is well-documented in other species (Milinski, 2006; Penn, Damjanovich, & Potts, 2002). Furthermore, preferences for MHC-dissimilarity may help to prevent inbreeding (Penn & Potts, 1999).

Studies suggest that women especially may use their perceptions of others' body odors as a guide for selecting MHC-dissimilar mates (Santos, Schinemann, Gabardo, & da Graça Bicalho, 2005). In several studies, normally-ovulating women rated the scents of MHC-dissimilar men (using t-shirts that the men slept in) as more pleasant smelling than those of MHC-similar men (e.g., Wedekind, Seebeck, Bettens, & Paepke, 1995). In another study, two subjects shared fewer MHC alleles than expected by chance when the odor of one reminded the other of his or her ex-romantic partner (Wedekind & Füri, 1997).

Furthermore, there is evidence that both European American and Hutterite spouses assort by MHC-dissimilarity (Chaix et al., 2008; Ober et al., 1997). Recent findings also suggest that MHC- dissimilarity of partners is important for relationship satisfaction and stability. Saphire-Bernstein et al. (2017) found that among Asian American couples, MHC-dissimilarity predicted greater attraction to one's partner (e.g., satisfaction with a partner's sexiness). Similarly, Garver-Apgar, Gangestad, Thornhill, Miller, and Olp (2006) found in a predominantly European American sample that women who shared fewer MHC alleles with their partners were less attracted to other men and less likely to cheat on their partners than women who shared more alleles. Accordingly, these women's self-reported and partner-reported sexual satisfaction and sexual responsiveness to their partners were higher than those of women who shared more MHC alleles with their partners. This may be explained in part by the high importance that women ascribe to the scent of a partner, placing scent as the most important attribute in a mate second to "pleasantness" in one study (Herz & Inzlicht, 2002). Women care the most about the smell of their mates when they are at the most fertile phase of their menstrual cycles, further supporting the possible adaptive significance of odor preferences in mates (Doty, 1981; Vierling & Rock, 1967). In sum, much evidence suggests that individuals, especially women, employ a detection system for MHC-dissimilarity in potential mates.

Yet, it should be mentioned that there are contrary findings as well. Various studies have found non-significance or preferences for MHC-similarity, perhaps due to differences in the specific loci typed, sample size, population, environmental or cultural factors, unit (i.e., couples, women, or men) examined, or ethnic heterogeneity (for reviews, see Havlicek & Roberts, 2009; Winternitz, Abbate, Huchard, Havlíček, & Garamszegi, 2017). For example, Thornhill et al. (2003) found that men preferred scents of women who were MHC-dissimilar, while women did not show preferences, which could be due to the use of a multi-ethnic sample instead of the ethnically homogenous samples in Wedekind et al.'s (1995, 1997) studies. In these studies, ethnicity, which is likely related to both degree of MHC-similarity and scent preferences, may have confounded results. In another study modeled after Wedekind et al.'s design, women who were single preferred the scents of MHC-similar men, whereas women who were in a relationship preferred the scents of MHC-dissimilar men (Roberts, Gosling, Carter, & Petrie, 2008). Other researchers found that women preferred the scent of men with an intermediate MHC-match, based on HLA alleles inherited from their fathers (Jacob, McClintock, Zelano, & Ober, 2002). Furthermore, evidence for assortment for MHC-dissimilarity has not been found among married couples in many populations including South Amerindians (Hedrick & Black, 1997) and Japanese (Ihara, Aoki, Tokunaga, Takahashi, & Juji, 2000) (for a review, see Havlicek & Roberts, 2009). More recently, a number of studies using full genome sequencing have failed to find support for selection for MHC-dissimilarity among married couples (Derti et al., 2010; Qiao et al., 2018). In a recent meta-analysis, Winternitz et al. (2017) noted the inconsistency of findings regarding human preferences for MHC-dissimilarity and showed that ethnic heterogeneity may help to explain these inconsistencies. In their analyses, preferences for MHC-similarity were found among ethnically diverse samples, but not homogeneous samples.

Recent studies, using single nucleotide polymorphism (SNP) typing of the MHC, have further suggested that married couples are significantly more dissimilar at this region than other genomic regions (Chaix et al., 2008). These results have been challenged, however, suggesting that an avoidance of extreme MHC similarity, rather than preference for dissimilarity, may drive the observed heterogeneity (Derti et al., 2010). Given that these SNP based approaches target a much wider region of the MHC than prior serological approaches, it is difficult to directly compare results, since the overall effects of assortative mating and demographics in humans are still unresolved (Domingue, Fletcher, Conley, & Boardman, 2014; Abdellaoui, Verweij, & Zietsch, 2014).

These past studies have one major limitation in common. Although much research has looked at MHC and indices of attraction such as ratings of scent, no prior study has established a direct link between MHC and initial mate choice. Currently, it is unclear when and to what extent MHC-based attraction is influential in a real life scenario, when many other salient factors (e.g., appearance, personality) may be at play. To examine this question, we conducted a speed-dating study in which participants were genotyped along their MHC genomic region.

1.2. Hypotheses

Based on the aforementioned literature, we predicted that individuals would show preferences based on MHC-complementarity, specifically rating MHC-dissimilar partners higher on overall, short-term, and long-term mate desirability and being more likely to offer them second dates (hypothesis 1). We also predicted that individuals would rate the body scents and faces of MHC-dissimilar partners as more attractive (body scent, hypothesis 2; face, hypothesis 3). Finally,

we examined the magnitude of the effects of the MHC in relation to those of other attributes known to be important in initial attraction, such as physical attractiveness, on date offers (Eastwick & Finkel, 2008; Luo & Zhang, 2009), and tested the role of MHC-complementarity in the presence of these other factors.

2. Methods

2.1. Participants

Participants were 262 single Asian Americans (men, $n = 132$, women, $n = 130$) who were romantically interested in the opposite sex. To reduce ethnic stratification of genetic effects (Gomes et al., 2017) and documented racial preferences in romantic partners (Fisman et al., 2008), we recruited only individuals of Asian descent, who comprise about 50% of the undergraduate population at our institution. To help control for possible effects of generational status on mate preferences, participation was limited to individuals who were born in the US or moved here before age 16.

Most participants attended or had recently graduated from our institution, which is a large, public university on the West Coast. We utilized a variety of recruitment methods, including advertisements in the Social Sciences Human Subjects Pool, in-person announcements in classes, emails to classes, flyers in various campus buildings, recruitment booths within the campus, social networking sites (i.e., Facebook.com, Reddit.com, Meetup.com), and word-of-mouth. 51% of participants were eligible for extra credit at the university and were credited with 2.5 extra credit points for their participation. The remaining participants were not compensated.

Participants ranged from 18 to 30 years of age ($M = 20.72$, $SD = 2.29$). Events were separated by age groups to help avoid the potential confounding effect of age on mate preferences. Specifically, sessions were separated into the following age groups: ages 18–20, 19–21, 21–23, and 24–30. Of the participants, 83% self-identified as mono-ethnic, 13% self-identified as mixed-ethnicity (but only Asian in their racial heritage), and 4% self-identified as mixed-race. Of the mono-ethnic group, the most common ethnicities were Chinese (40%), Vietnamese (21%), Korean (18%), Filipino (13%), and Taiwanese (7%). Most of the participants (73%) were born in the United States.

After data-cleaning (e.g., removing observations from people who were acquainted prior to the speed-dating session), the total number of observations, or dates, was $n = 2215$.

2.2. Measures

2.2.1. Date offers

Date offers was measured through the question: “Would you like to make a second date offer to the person you just interacted with?” Participants could respond “Yes” or “No” (Finkel & Eastwick, 2008; Fisman et al., 2008). We explained to the participants that they would receive each other as a match, and receive each other's email address if both they and their partner responded “Yes.”

2.2.2. Mate desirability

Participants rated, on a 1–11 scale (1 = “Extremely undesirable”, 6 = “Neither desirable/undesirable”, 11 = “Extremely desirable”), their partner's overall desirability, desirability as a short-term partner, and desirability as a long-term partner.

2.2.3. Body scent attractiveness

Body scent attractiveness was measured with the question, rated on a 1–11 scale, “How attractive/sexy was this person's natural body scent?” where 1 = “Extremely unattractive/sexy”, 6 = “Neither unattractive nor attractive”, 11 = “Extremely attractive”.

2.2.4. Facial attractiveness

To measure facial attractiveness, participants rated their partners on the item “physically attractive (face)” on a 1–11 scale (1 = “Not at all descriptive”, 6 = “Moderately descriptive”, 11 = “Extremely descriptive”).

2.2.5. MHC-dissimilarity

MHC-dissimilarity was computed through the absolute value of the difference score between speed-dating partners for each single nucleotide polymorphism (SNP). For example, a participant with the genotype AA would have a score of 0 with a partner with the genotype AA, 1 with a partner with the genotype AB, and 2 with the genotype BB.

2.2.6. Other measures

We tested effects of MHC-genes in date offers in relation to the effects of many other attributes that might be important based on prior literature. The measures included in this analysis were *subject's self-report* and *partner's self-report* on age, GPA, height, year in school, subjective socioeconomic status (Adler, Epel, Castellazzo, & Ickovis, 2000), political leaning, whether they were born in the US, whether they (women) were on birth control, whether they were sick, self-esteem (Rosenberg, 1965), social desirability (Reynolds, 1982), narcissism (Ames, Rose, & Anderson, 2006), depression (Radloff, 1977; Reynolds, 1997), attractiveness attributes, relational attributes, vibrancy attributes, cerebral attributes (Wu et al., 2015); *subject's rating of partner*, *partner's rating of subject*, *average ratings of partner by speed-dating partners*, and *average ratings of subject by speed-dating partners* on overall desirability as a partner, desirability as a short-term partner, desirability as a long-term partner, sexual chemistry, sexiness of body scent (Wedekind et al., 1995), intensity of body scent (Wedekind et al., 1995), masculinity-femininity, attractiveness attributes, relational attributes, vibrancy attributes, cerebral attributes (Wu et al., 2015); *similarity in attractiveness between the subject and partner*, measured through the absolute value of the difference between their average ratings on attractiveness (by speed-dating partners) as well as the absolute value of the difference between their self-ratings on attractiveness (Murstein, 1972); *genetic measures* comprised of MHC dissimilarity between the subject and partner across all 28 SNPs, MHC rareness (Coetzee et al., 2007) of the subject and partner across the same 28 SNPs, MHC heterozygosity (Coetzee et al., 2007) of the subject and partner across the same 28 SNPs, rs6311 genotype of subject and partner (previously related to leadership and social dominance in men, and found to predict greater speed-dating success for men compared to women in the current sample; see Wu et al., 2016), rs1799971 genotype of subject and partner (previously related to submissiveness and social sensitivity, and found to predict greater speed-dating success for women compared to men in the current sample; see Wu et al., 2016); and *other measures* comprised of the sex that rotated in the speed-dating study (Finkel & Eastwick, 2009), shared religion between the subject and partner, and shared ethnicity between the subject and partner. Details on these measures are presented in the Supplementary online materials (Table S1).

2.3. Procedure

2.3.1. DNA treatment

DNA was collected and purified using the Gentra Puregene Buccal Cell Kit from Qiagen, and shipped to the UC Davis Genome Center for genotyping using the Illumina GoldenGate Assay. Thirty-eight tag SNPs for the MHC region were selected using De Bakker et al.'s (2006) *HLA* SNP haplotype map based on the Chinese (CHB) population. To capture variation in this region where most alleles are rare, we chose SNPs that tagged the most common alleles according to studies of various Asian populations (Han Chinese in Beijing, De Bakker et al., 2006; Han Chinese in southern China, Trachtenberg et al., 2007; Asian Americans, Cao et al., 2001; Asian Americans, Maiers, Gragert, & Klitz, 2007;

Table 1
Genotype and allele frequencies for MHC SNPs.

SNP	HLA Tag	Loci	Chr6 position	Allele frequency		Genotype frequency			1000 genomes allele frequency (EAS)		1000 genomes genotype frequency (EAS)		
				p	q	p^2	pq	q^2	p	q	p^2	pq	q^2
rs2517754	HLA-A*1101	HLA-A 5′	29928653	0.70 (G)	0.30 (A)	131	92	30	0.69 (G)	0.31 (A)	284	128	92
rs6457109	HLA-A*0206	HLA-A 3′	29965234	0.93 (T)	0.07 (C)	222	32	2	0.94 (T)	0.05 (C)	444	56	4
rs4959037	HLA-A*0201	HLA-A 3′	29968280	0.71 (A)	0.29 (T)	122	121	14	0.69 (A)	0.31 (T)	244	206	54
rs2844713	HLA-A*0203	HLA-E 3′	30551231	0.66 (G)	0.34 (A)	106	130	23	0.60 (G)	40 (A)	182	241	81
rs3094663	HLA-C*1202	HLA-C 3′	31139060	0.67 (C)	0.33 (T)	108	132	20	0.70 (C)	0.30 (T)	248	206	50
rs2073716	HLA-C*1402	HLA-C 3′	31154970	0.81 (C)	0.19 (G)	165	88	5	0.82 (C)	0.18 (G)	340	144	20
rs4713438	HLA-C*0302	HLA-C 3′	31178819	0.87 (G)	0.13 (A)	197	58	4	0.86 (G)	0.14 (A)	371	125	8
rs3130457	HLA-C*0602	HLA-C 3′	31179167	0.94 (T)	0.06 (C)	228	30	1	0.96 (T)	0.04 (C)	468	35	1
rs4122190	HLA-C*0102	HLA-C 3′	31199825	0.86 (A)	0.14 (G)	186	73	1	0.82 (A)	0.18 (G)	336	155	13
rs2001181	HLA-C*0702	HLA-C intron	31268971	0.65 (T)	0.35 (C)	151	26	77	0.83 (T)	0.17 (C)	347	142	15
rs3905495	HLA-B*5101	HLA-C 5′ (HLA-B 3′)	31297512	0.64 (G)	0.36 (A)	101	125	31	0.59 (G)	0.41 (A)	178	243	83
rs4947248	HLA-B*6701	HLA-B 3′	31333998	0.74 (T)	0.26 (C)	136	111	11	0.79 (T)	0.21 (C)	314	168	22
rs2844586	HLA-B*5201	HLA-B 3′	31349997	0.92 (G)	0.08 (A)	218	42	0	0.91 (G)	0.09 (A)	422	77	5
rs3819294	HLA-B*1501	HLA-B intron	31354460	0.74 (G)	0.26 (A)	135	112	12	0.72 (G)	0.28 (A)	264	201	39
rs2844580	HLA-B*4001	HLA-B 5′	31365276	0.83 (T)	0.17 (C)	176	76	7	0.85 (T)	0.14 (C)	366	124	14
rs4713518	HLA-B*5801	HLA-DRA 5′	32289310	0.63 (A)	0.37 (G)	98	126	33	0.64 (A)	0.36 (G)	200	244	60
rs3129859	HLA-DRB*0803;*0701	HLA-DRA 5′	32432912	0.74 (G)	0.26 (C)	137	109	12	0.80 (G)	0.20 (C)	323	156	25
rs7773756	HLA-DRB*1501	HLA-DRA 5′	32434437	0.55 (C)	0.45 (T)	66	152	41	0.52 (C)	0.48 (T)	138	247	119
rs2395185	HLA-DRB*0901	HLA-DRA 3′	32465140	0.62 (G)	0.38 (T)	98	125	35	0.69 (G)	0.31 (T)	236	223	45
rs532098	HLA-DQA*0101	HLA-DRB1 5′	32610025	0.55 (G)	0.45 (A)	60	163	36	0.61 (G)	0.39 (A)	194	227	83
rs7744001	HLA-DQB*0601;*0502	HLA-DQB1 3′ noncoding	32658059	0.62 (G)	0.38 (A)	91	141	28	0.64 (G)	0.36 (A)	217	215	72
rs6928482	HLA-DQB*0301	HLA-DQB1 3′	32658222	0.60 (C)	0.40 (T)	87	134	37	0.51 (C)	0.49 (T)	133	249	122
rs1063355	HLA-DQA*0103	HLA-DQB1 3′	32659687	0.64 (G)	0.36 (T)	100	132	28	0.54 (G)	0.46 (T)	151	245	108
rs9275184	HLA-DQB*0302	HLA-DQB1 5′	32686687	0.91 (T)	0.09 (C)	213	44	1	0.93 (T)	0.07 (C)	436	64	4
rs3129718	HLA-DQA*0601	HLA-DQB1 5′	32691811	0.84 (C)	0.16 (T)	178	69	7	0.90 (C)	0.10 (T)	410	89	5
rs2856718	HLA-DRB*1201	HLA-DQB1 5′	32702228	0.52 (C)	0.48 (T)	69	131	60	0.53 (T)	0.47 (C)	133	266	105
rs7769979	HLA-DRB*0301	HLA-DQB2 3′	32755545	0.80 (A)	0.20 (G)	160	92	7	0.81 (A)	0.19 (G)	332	148	24
rs6903433	HLA-DQA*0301	HLA-DQB2 5′	32874948	0.87 (C)	0.13 (T)	194	62	2	0.89 (C)	0.11 (T)	395	104	5

Note. HLA Tag is the HLA type identified as associated with the given SNP from De Bakker et al. (2006). Loci and Chr 6 position is the location of the given SNP with respect to the nearest HLA gene in the MHC Complex (Trowsdale & Knight, 2013), using Assembly Dec. 2013 (GRCh38/hg38) in the UCSC Genome Browser (Kent et al., 2002; http://genome.ucsc.edu/). p = major allele frequency; q = minor allele frequency.

Japanese, Saito, Ota, Yamada, Inoko, & Ota, 2001). It should be noted that while De Bakker et al. (2006) chose SNPs that "tagged" particular HLA serotypes, many of these SNPs are, in fact, many hundreds of kilobases from the "tagged" gene. Table 1 indicates the original serotype associated with a particular SNP, as well as the actual location of the SNP in the MHC region.

After quality control (for call frequency, heterozygous genomes, cluster separation, ABT mean, minor allele frequency, replicability), 28 SNPs remained. The targeted SNPs span approximately 2.1 Mb of the MHC region (avg. spacing 75 kb), with clustering near the major Class I and Class II HLA genes (*HLA-A*, *HLA-C*, *HLA-B*, *HLA-DRB*, *HLA-DQA*, and *HLA-DQB*, Table 1). While little linkage disequilibrium was detected between the SNPs, this is difficult to confirm in the MHC region due to ongoing selection and gene conversion (De Bakker et al., 2006; Trowsdale & Knight, 2013). Noticeably, the allele frequencies for some SNPs in our sample, which was similar to those for East Asians in the 1000 Genomes Project Consortium (2015), appeared to reflect heterozygote advantage (i.e., an overrepresentation of heterozygotes; Chaix et al., 2008; Zou et al., 2015). This could be a consequence of various processes including non-random mate selection, disease resistance, and maternal-fetal interactions (Hedrick & Thomson, 1988; Trowsdale & Knight, 2013). The final 28 SNPs utilized in this study, then, were a sampling of the common diversity at the MHC region found in Asian populations.

2.3.2. Data collection

Fifteen speed-dating sessions, each of two hour duration, were held in on-campus reception-style rooms. Fifteen to 22 individuals participated in each session (M = 17.6, SD = 2.06). To select participants, we first asked interested individuals to complete prescreening surveys which included age, ethnicity, when they were born in the US/when they moved to the US, dating status, self-ratings of different attribute types, whether they were willing to participate in DNA collection, and whether they were willing to refrain during the day of the speed-dating session from "eating pungent foods, smoking, using perfumes/colognes, using scented lotions, and using deodorants". Only participants who fit the demographic criteria and were willing to participate in DNA collection and refrain from scent usage were selected. In the one or two days leading up to the speed-dating session, participants were again instructed (through both email and phone) to refrain from scent-usage. Prior to the speed-dating event, participants also provided their demographic information through an online survey.

Upon arrival to the session, males and females were separated and gave informed consent. Participants then went on 3-min speed-dates with each member of the opposite sex. Between each date, participants completed an interaction questionnaire (which included the second-date offer, ratings of mate desirability, body scent attractiveness, and facial attractiveness) for the person with whom they had just met. After the speed-dates ended, cheek swabs were collected from participants.

3. Analyses

3.1. MHC-dissimilarity in mate preferences

We removed speed-dates involving participants of mixed-race ancestry (n = 10) or women on hormonal contraception (n = 12) from these analyses due to possible confounding effects resulting in n = 2062 observations. Men and women were analyzed separately. There was

Table 2
Descriptive statistics of and correlations among dependent variables.

	M(*SD*)	*Range*	Correlation coefficient				
			Overall mate desirability	Short-term mate desirability	Long-term mate desirability	Body scent attractiveness	Facial attractiveness
Date offer	0.43 (0.50)	0–1 (binary)	0.53***	0.51***	0.49***	0.34***	0.48***
Overall mate desirability	5.95 (2.05)	1–11		0.82***	0.79***	0.55***	0.75***
Short-term mate desirability	5.65 (2.28)	1–11			0.75***	0.48***	0.66***
Long-term mate desirability	5.21 (2.32)	1–11				0.48***	0.65***
Body scent attractiveness	5.91 (1.48)	1–11					0.53***
Facial attractiveness	5.95 (2.18)	1–11					

Note. Number of observations (speed-dates) ranged from 2184 to 2214.
*** $p \leq 0.001$, two-tailed.

Table 3
p-Values for joint analysis of all MHC SNPs and mate preferences of women and men.

	Date offer	Overall mate desirability	Short-term mate desirability	Long-term mate desirability	Body scent attractiveness	Facial attractiveness
Women	0.078	0.005	0.004	< 0.001	0.081	0.001
Men	0.612	0.443	0.182	0.647	0.064	0.300

little dependency in responses within the fifteen speed-dating sessions, ICC = 0.01–0.02 for date offers and measures of mate desirability, thus, we did not include session in the analyses. However, there was dependency within each speed-dater's responses, ICC = 0.23–0.37 for measures of mate desirability, 0.92 for date offers. We thus accounted for repeated measurements (arising from multiple speed-dates per participant) using both mixed-effect models and permutations.

First, given the likely small effects of individual SNPs, we tested the joint effects of all 28 genotyped SNPs in the MHC region using score tests in random-effects models. The effects of the SNPs were assumed to be a random sample from a normal distribution with a mean of zero (Goeman, Van De Geer, De Kort, & Van Houwelingen, 2004; Liu, Ghosh, & Lin, 2008). This model has an advantage over the traditional fixed-effects regression due to its statistical power in the presence of numerous predictor variables (in this case, SNPs). To account for the repeated measurements and to obtain accurate *p*-values, we conducted 10,000 permutations (Ludbrook & Dudley, 1998). In each permutation, the repeated measurements were shuffled within each participant while holding all genetic variables fixed.

To look at specific patterns of MHC-(dis)similarity (i.e., directionality), we then analyzed associations between individual SNPs and speed-dating outcomes using generalized linear mixed models for date offers and linear mixed models for ratings of partners' mate desirability, facial attractiveness, and body scent attractiveness. To adjust for false positives from multiple tests, we used the Benjamini and Hochberg (1995) procedure with a False Discovery Rate of 0.20, applying the procedure for each outcome variable.

3.2. Relative effects of MHC

In this analysis, we aimed to understand the effects of MHC relative to those of many other variables that might predict date offers based on prior literature. Thus, this analysis included all speed-daters regardless of their ancestry or birth control status. Random forest models were run with bootstrapping (1000 times) to assess the importance of different factors in predicting whether women offered men dates, whether men offered women dates, and whether a particular dyad resulted in a match (both partners offered each other second dates). Random forests were chosen over other methods such as logistic regression due to its ability to measure non-linear effects and to handle large numbers of predictor variables, as well as collinearity among them (Mendez, Buskirk, Lohr, & Haag, 2008; Strobl, Malley, & Tutz, 2009). This technique has recently been used to predict individuals' desire for their speed-dating partners and their desirability to their speed-dating partners using a large variety of behavioral measures (e.g., personality measures, self-attributes) relevant to human mate selection (Joel, Eastwick, & Finkel, 2017).

In our analysis, factors were grouped by the following categories: Subject Self-Report, Partner Self-report, Subject Report of Partner, Partner Report of Subject, Average Partner Report of Subject (across all of the subject's speed-dating partners), Average Partner Report of Partner (across all of the partner's speed-dating partners), Attractiveness Dis(similarity) between Subject and Partner, Genetics, and Other factors (i.e., same religion, same ethnicity, rotating sex in the session). Mean replacement was used for missing values for continuous variables, whereas mode replacement was used for missing values for categorical variables. Percentage of missing data was low, ranging from 0%–7% (see Table S1).

4. Results

All six outcomes measures (i.e., date offer, overall mate desirability, short-term mate desirability, body scent attractiveness, and facial attractiveness) were positively correlated with each other. Correlations ranged from medium to very large (see Table 2 for descriptive statistics of and correlations among the dependent variables).

4.1. MHC-dissimilarity in mate preferences

Joint analysis of all the genotyped SNPs in the MHC region indicated significant and consistent associations between MHC-complementarity and mate preferences for women but not for men, supporting Hypotheses 1–3 for women but not for men. Specifically, when examining all MHC SNPs together, women showed selection for men's MHC-complementarity across most measures. For men, all associations were non-significant (see Table 3).

Analysis of individual SNPs indicated that the association was

predominantly driven by particular regions of the genotyped MHC region (Table 4). For women, twenty-seven out of 168 associations, most found in ratings of mate desirability and facial attractiveness, passed the Benjamini and Hochberg (1995) procedure to adjust for multiple comparisons using a False Discovery Rate of 0.20. About 40% of the significant associations supported Hypotheses 1–3, indicating preferences for MHC-dissimilarity. For five SNPs, women's preferences for MHC- dissimilarity were supported (rs4959037: overall mate desirability; rs2073716: date offer, overall mate desirability, short-term mate desirability, long-term mate desirability, facial attractiveness; rs7744001: overall mate desirability; rs3129718: overall mate desirability; rs7769979: overall mate desirability, short-term mate desirability, long-term mate desirability, facial attractiveness). For six other SNPs, women's preferences for MHC-*similarity* were instead supported (rs2844713: overall mate desirability, short-term mate desirability, long-term mate desirability, body scent attractiveness; rs4713518: date offer, overall mate desirability, short-term mate desirability, long-term mate desirability, facial attractiveness; rs4713438: overall mate desirability, facial attractiveness; rs4947248: facial attractiveness; rs6903433: overall mate desirability; rs7773756: long-term mate desirability). There was little support for the role of the analyzed MHC SNPs in body scent preferences. See Table 4 for coefficients and exact (unadjusted) *p*-values.

Analyses of individual SNPs again indicated few MHC-based preferences of men for women. Using the Benjamini and Hochberg procedure with a False Discovery Rate of 0.20, six significant associations (out of 168 total comparisons) were found, three of which were for ratings of women's long-term mate desirability. Our hypotheses that men would prefer MHC-dissimilar women were supported for only one SNP (rs3129718: long-term mate desirability, facial attractiveness). For two SNPs, preference for MHC-*similarity* was instead supported (rs3130457: long-term mate desirability; rs532098: long-term mate desirability, body scent attractiveness, facial attractiveness) (see Table 4).

Interestingly, support for dissimilarity was predominantly found for SNPs in or near (< 25 kb) the *HLA* gene loci listed in Tables 1 and 2 (5 out of 6), while support for similarity was largely for SNPs farther from the *HLA* gene loci (> 25 kb, 7 out of 8, avg. distance 123.7 kb, $p = 0.026$, Fisher's Exact Test). This suggests that, consistent with prior work based on serological typing, selection for *HLA dissimilarity* was important. However, we also observed selection for *similarity* at other MHC genomic regions farther from the main *HLA* genes, which would not have been investigated with serological approaches.

4.2. Relative effects of MHC

In predicting women's date offers to men using random forests with bootstrapping, we obtained the following indices of model performance: Area under the Curve (AUC) = 0.98, Gini = 0.96, Accuracy = 0.93, Precision = 0.92, Recall = 0.97; Out of Bag Error (OOB) = 0.07. For men's date offers to women, indices were: AUC = 0.98, Gini = 0.96, Accuracy = 0.94, Precision = 0.94, Recall = 0.92, OOB = 0.06. For matches, indices were: AUC = 0.97, Gini = 0.94, Accuracy = 0.96, Precision = 0.95, Recall = 0.99, OOB = 0.04. These indices indicated that the models were excellent classifiers (AUC > 0.90; Gini > 0.60), high in accuracy (i.e., Accuracy, or proportion of predictions that were correct; Precision, or proportion of cases identified as positive that were truly positive; Recall, or proportion of true positives that were correctly identified as positive), and low in error (i.e., OOB, or error rate calculated using randomly re-sampled data) (e.g., Muchlinski, Siroky, He, & Kocher, 2015). Although these indices are widely used, it is advised to exercise caution when interpreting these metrics in the presence of class imbalance (i.e., when the proportion of one class is significantly higher than the proportion of the other class; Bekkar, Djemaa, & Alitouche, 2013). In our sample, class imbalance was a minor issue, particularly for date offers. The proportion of "Yeses" to making date offers was 0.32 for women and 0.54 for men, and the proportion of matches was 0.16.

Random forest analyses found that for both women and men (see Fig. 1A and B, respectively), subjects' ratings of partners were the most important predictors of date offers, followed by average ratings of partners by all speed-dating partners, average ratings of the subject by all speed-dating partners, (dis)similarity in attractiveness between the pair, the subject's self-reported characteristics, the partner's ratings of the subject, genetic measures, the partner's self-reported characteristics, and finally, other measures (i.e., shared religion, shared ethnicity, rotating sex). Attractiveness attributes, measures of desirability as a partner, and sexual chemistry stood out as especially important for both women and men. Of the genetic measures, MHC-related indices were more important than rs6311 or rs1799971 genotype, which previously had been shown to account for 2%–5% of the variance in speed-dating success (Wu et al., 2016). The importance of MHC-related indices (MHC-dissimilarity, MHC rareness, MHC heterozygosity) to making a date offer was comparable to that of a partner's self-reported personality characteristics (e.g., self-esteem, social desirability).

For predicting matches (or mutual date offers), women's ratings of men were the most important, followed by average ratings of the men by all their speed-dating partners, (dis)similarity in attractiveness between the pair, average ratings of women by all their speed-dating partners, men's ratings of women, men's self-reported characteristics, genetic measures, women's self-reported characteristics, and other measures (i.e., shared religion, shared ethnicity, rotating sex). Again, ratings of attractiveness attributes, desirability as a partner, and sexual chemistry, were especially important. Genetic measures again had effects similar to men and women's self-reported personality characteristics, and MHC-based indices were more important than rs6311 or rs1799971 genotype. See Fig. 1 for the Gini importance of all variables in women's date offers to men (Fig. 1A), men's date offers to women (Fig. 1B), and matches, or mutual date offers (Fig. 1C).

5. Discussion

This study set out to test for the presence of MHC-based mate preferences in the ecologically-valid setting of speed-dating in which participants' behaviors had real world consequences. We had predicted based on prior literature that individuals would show preferences for MHC-dissimilar partners. We also examined the relative importance of genetic factors and often-studied behavioral factors in speed-dating outcomes.

Using this SNP based approach, we found support for MHC-based dating preferences. These patterns, however, were more complex than either an overall dissimilarity or similarity preference, instead varying by SNP. Over 260 genes are densely packed in the Human MHC region, with the classical HLA genes representing only a minor fraction of them (Trowsdale & Knight, 2013). While many of the SNPs utilized in this study are clustered near the HLA genes, others are present in adjacent genes (Table 1). Given the enormous genetic diversity at the MHC locus, it is likely that preferences for similarity at specific SNPs, and dissimilarity at others, is the expected result. Many allelic variants in the genome are expected to be disadvantageous, and hence selected against. For example, the two SNPs at which men showed MHC-similarity preferences in women have each been associated with diseases including vitiligo (minor allele of rs532098; Jin et al., 2011; Zhang & Xiang, 2014); and lung cancer, psoriasis, and vitiligo (minor allele of rs3130457, Niu et al., 2015; Zhu et al., 2011).

Consistent with this idea, SNPs that supported preferences for MHC-dissimilarity, as opposed to MHC-similarity, tended to be closer (< 25 kb) to *HLA*-genes. This specific observation is consistent with prior work supporting MHC-dissimilarity, largely conducted with serological typing of HLA proteins. Our observation of preferences for similarity at sites farther away from the *HLA* genes (often in or near other

Table 4
MHC-dissimilarity of individual SNPs on mate preferences of women and men.

SNP	Loci	Chr6 position	Date offer		Overall mate desirability		Short-term mate desirability		Long-term mate desirability		Body scent attractiveness		Facial attractiveness	
			B	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*	*B*	*p*
Women's preferences for men														
rs2517754	HLA-A 5′	29928653	0.06	0.628	0.05	0.605	0.15	0.148	0.08	0.469	−0.02	0.732	0.05	0.655
rs6457109	HLA-A 3′	29965234	−0.06	0.743	−0.08	0.613	−0.11	0.478	−0.06	0.696	0.01	0.937	−0.12	0.417
rs4959037	HLA-A 3′	29968280	0.12	0.388	0.22	0.046	0.12	0.291	0.13	0.259	0.06	0.426	0.11	0.324
rs2844713	HLA-E 3′	30551231	−0.20	0.141	−0.21	0.053	−0.27	0.012	−0.29	0.009	−0.28	0.000	−0.17	0.115
rs3094663	HLA-C 3′	31139060	0.09	0.514	−0.05	0.638	−0.21	0.067	0.07	0.519	−0.07	0.386	0.03	0.821
rs2073716	HLA-C 3′	31154970	0.47	0.002	0.26	0.030	0.35	0.006	0.30	0.020	0.20	0.023	0.38	0.002
rs4713438	HLA-C 3′	31178819	−0.31	0.057	−0.29	0.022	−0.34	0.012	−0.21	0.121	−0.01	0.924	−0.36	0.005
rs3130457	HLA-C 3′	31179167	−0.19	0.390	0.18	0.295	0.07	0.711	0.08	0.677	0.01	0.947	0.04	0.824
rs4122190	HLA-C 3′	31199825	0.07	0.677	−0.02	0.897	−0.03	0.830	0.06	0.643	0.04	0.647	0.02	0.851
rs2001181	HLA-C intron	31268971	−0.08	0.350	−0.04	0.610	−0.02	0.807	0.02	0.738	−0.07	0.178	−0.08	0.286
rs3905495	HLA-C 5′ (HLA-B 3′)	31297512	0.03	0.812	0.06	0.543	0.09	0.413	0.06	0.568	0.00	0.971	0.05	0.609
rs4947248	HLA-B 3′	31333998	−0.36	0.014	−0.20	0.082	−0.16	0.183	−0.22	0.067	−0.29	0.001	−0.31	0.007
rs2844586	HLA-B 3′	31349997	0.08	0.654	0.00	0.999	0.10	0.511	−0.07	0.668	−0.07	0.540	−0.01	0.969
rs3819294	HLA-B intron	31354460	0.19	0.163	0.13	0.249	0.11	0.329	0.12	0.281	−0.07	0.374	−0.01	0.929
rs2844580	HLA-B 5′	31365276	0.12	0.440	0.04	0.732	0.14	0.304	0.11	0.435	0.03	0.769	0.09	0.476
rs4713518	HLA-DRA 5′	32289310	−0.31	0.021	−0.24	0.019	−0.33	0.003	−0.25	0.022	−0.16	0.035	−0.30	0.005
rs3129859	HLA-DRA 5′	32432912	0.09	0.547	0.12	0.270	0.09	0.432	−0.01	0.949	0.03	0.730	0.16	0.166
rs7773756	HLA-DRA 5′	32434437	−0.07	0.609	−0.14	0.165	−0.13	0.252	−0.23	0.033	0.06	0.446	−0.05	0.613
rs2395185	HLA-DRA 3′	32465140	0.13	0.337	0.08	0.441	0.09	0.432	0.11	0.301	−0.09	0.257	0.06	0.608
rs532098	HLA-DRB1 5′	32610025	0.22	0.105	−0.01	0.903	−0.10	0.401	0.08	0.488	−0.06	0.448	−0.04	0.683
rs7744001	HLA-DQB1 3′ noncoding	32658059	0.26	0.044	0.24	0.018	0.11	0.302	0.19	0.076	0.06	0.447	0.12	0.270
rs6928482	HLA-DQB1 3′	32658222	−0.10	0.466	−0.18	0.093	−0.15	0.191	−0.12	0.271	−0.08	0.271	0.00	0.972
rs1063355	HLA-DQB1 3′	32659687	−0.01	0.926	−0.08	0.430	−0.05	0.657	−0.05	0.636	−0.07	0.340	0.08	0.451
rs9275184	HLA-DQB1 5′	32686687	−0.11	0.583	0.12	0.435	0.11	0.524	0.03	0.858	0.10	0.379	0.21	0.180
rs3129718	HLA-DQB1 5′	32691811	0.23	0.111	0.25	0.037	0.09	0.466	0.25	0.045	0.10	0.262	0.31	0.012
rs2856718	HLA-DQB1 5′	32702228	0.07	0.568	0.07	0.483	0.02	0.831	0.19	0.078	0.01	0.857	−0.07	0.514
rs7769979	HLA-DQB2 3′	32755545	0.17	0.249	0.28	0.015	0.41	0.001	0.27	0.025	0.05	0.587	0.21	0.086
rs6903433	HLA-DQB2 5′	32874948	−0.20	0.262	−0.28	0.046	−0.07	0.630	−0.01	0.960	−0.10	0.304	−0.14	0.337
Men's preferences for women														
rs2517754	HLA-A 5′	29928653	0.13	0.270	0.09	0.316	0.14	0.157	0.05	0.649	0.15	0.041	0.09	0.359
rs6457109	HLA-A 3′	29965234	0.08	0.707	−0.25	0.107	−0.26	0.127	−0.14	0.404	0.00	0.972	−0.08	0.636
rs4959037	HLA-A 3′	29968280	0.16	0.247	0.15	0.154	0.28	0.013	0.18	0.107	−0.07	0.356	0.21	0.067
rs2844713	HLA-E 3′	30551231	−0.17	0.185	0.09	0.350	0.05	0.634	0.10	0.369	0.06	0.417	0.13	0.206
rs3094663	HLA-C 3′	31139060	0.09	0.527	0.04	0.673	0.05	0.617	−0.07	0.547	0.00	0.967	0.10	0.360
rs2073716	HLA-C 3′	31154970	−0.10	0.508	−0.08	0.504	−0.10	0.417	−0.22	0.062	0.05	0.551	−0.03	0.804
rs4713438	HLA-C 3′	31178819	0.29	0.062	0.14	0.225	0.27	0.028	0.13	0.315	0.13	0.134	0.15	0.249
rs3130457	HLA-C 3′	31179167	−0.15	0.498	−0.28	0.086	−0.21	0.232	−0.44	0.012	−0.01	0.953	−0.05	0.786
rs4122190	HLA-C 3′	31199825	−0.25	0.139	−0.12	0.335	0.03	0.854	−0.10	0.460	−0.05	0.605	−0.02	0.902
rs2001181	HLA-C intron	31268971	0.01	0.938	0.01	0.854	−0.04	0.544	0.02	0.732	0.01	0.884	0.08	0.285
rs3905495	HLA-C 5′ (HLA-B 3′)	31297512	0.05	0.698	0.00	0.960	0.04	0.659	−0.01	0.887	0.04	0.590	−0.07	0.501
rs4947248	HLA-B 3′	31333998	0.02	0.889	0.13	0.245	0.15	0.187	0.06	0.615	0.18	0.028	0.05	0.704
rs2844586	HLA-B 3′	31349997	−0.03	0.881	−0.22	0.155	−0.16	0.348	−0.13	0.446	−0.05	0.678	−0.19	0.268
rs3819294	HLA-B intron	31354460	0.02	0.866	0.04	0.702	0.08	0.450	0.14	0.195	0.01	0.899	0.05	0.683
rs2844580	HLA-B 5′	31365276	0.10	0.505	−0.06	0.603	−0.10	0.404	−0.21	0.074	−0.10	0.234	−0.10	0.430
rs4713518	HLA-DRA 5′	32289310	−0.10	0.400	−0.15	0.127	−0.08	0.461	−0.09	0.405	−0.15	0.038	−0.23	0.028
rs3129859	HLA-DRA 5′	32432912	−0.04	0.794	−0.09	0.404	−0.10	0.398	−0.09	0.416	0.07	0.361	0.04	0.754
rs7773756	HLA-DRA 5′	32434437	−0.22	0.096	−0.18	0.070	−0.24	0.023	−0.12	0.252	0.01	0.919	−0.14	0.194
rs2395185	HLA-DRA 3′	32465140	−0.09	0.471	0.09	0.354	0.19	0.054	0.06	0.544	−0.09	0.202	0.06	0.565
rs532098	HLA-DRB1 5′	32610025	−0.20	0.157	−0.26	0.013	−0.14	0.201	−0.30	0.008	−0.23	0.004	−0.38	0.001
rs7744001	HLA-DQB1 3′ noncoding	32658059	0.22	0.086	0.09	0.374	0.21	0.047	0.03	0.762	−0.01	0.859	0.11	0.331
rs6928482	HLA-DQB1 3′	32658222	0.11	0.409	0.02	0.877	0.10	0.337	0.05	0.624	−0.11	0.151	.- 0.01	0.961
rs1063355	HLA-DQB1 3′	32659687	0.06	0.629	0.00	0.983	0.04	0.694	0.00	0.975	−0.05	0.472	−0.07	0.493
rs9275184	HLA-DQB1 5′	32686687	−0.14	0.458	0.00	0.998	−0.02	0.883	−0.07	0.625	−0.07	0.531	−0.21	0.160
rs3129718	HLA-DQB1 5′	32691811	0.28	0.063	0.23	0.044	0.18	0.144	0.36	0.004	0.14	0.117	0.36	0.004
rs2856718	HLA-DQB1 5′	32702228	0.02	0.906	0.06	0.559	0.08	0.428	0.14	0.180	0.02	0.757	0.00	0.997
rs7769979	HLA-DQB2 3′	32755545	0.05	0.736	0.02	0.850	0.04	0.708	0.17	0.165	−0.10	0.219	0.00	0.997
rs6903433	HLA-DQB2 5′	32874948	0.06	0.725	0.11	0.441	0.01	0.925	−0.08	0.585	0.15	0.553	0.13	0.376

Note. Positive coefficients (*B*) indicate preferences for MHC-dissimilarity. Unadjusted *p*-values are reported. *p*-Values that passed the Benjamini and Hochberg (1995) procedure using a False Discovery Rate of 0.20 are underlined.

MHC genes) may suggest a separate selection mechanism for consistent functionality rather than complementarity. Future research can test this question.

Overall, our findings are in line with the mixed literature that may be attributed to differences in methods (e.g., specific loci typed, population, ethnic heterogeneity; Havlicek & Roberts, 2009; Winternitz et al., 2017). There were a few notable differences between our study and prior research, however. In particular, participants in our study met

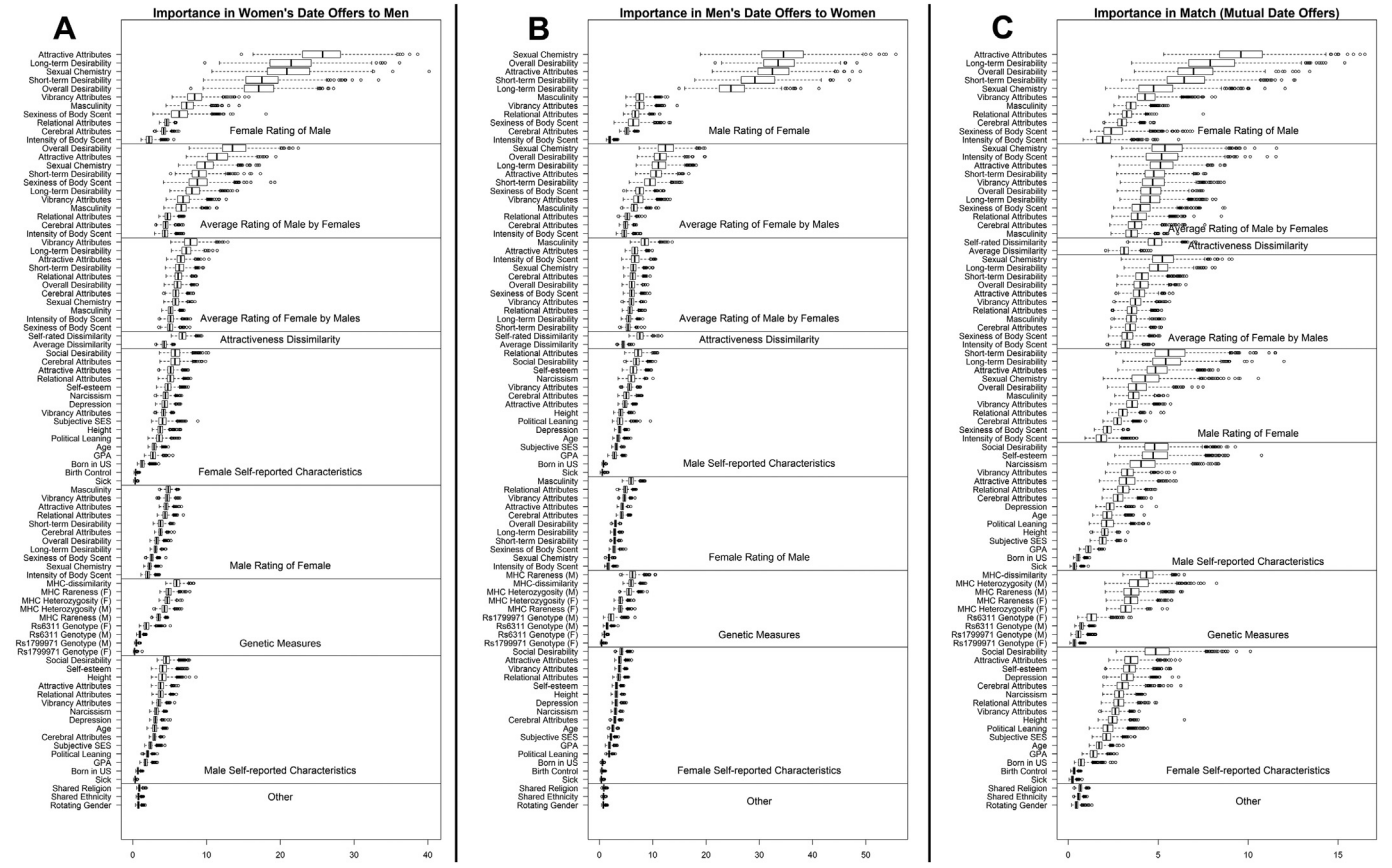

Fig. 1. A–C. Random forest results of predictors of date offers. The Gini importance of various attributes is depicted for women's date offers to men (panel A), men's date offers to women (panel B), and matches (mutual date offers; panel C). Groups of attributes are arranged from top to bottom by average importance. Attributes are also arranged by importance within each block.

and interacted with partners rather than smelling t-shirts or satchels of paper that others wore. Thus, effects of MHC-complementarity were evaluated in the presence of many other factors, such as appearance and personality. This setting may explain why we largely failed to replicate past findings that individuals prefer the scent of those who are more MHC-dissimilar (e.g., Wedekind et al., 1995). First, although participants sat very close to each other, they may not have been able to accurately pick up one another's body scents. Second, it is probable that, consistent with the halo effect (Nisbett & Wilson, 1977) and the medium-to-high correlations between our outcome measures, speed-daters' evaluations of their partners' physical attractiveness or overall desirability carried over to their assessments of body scent.

We used random forests to examine the importance of MHC-based measures in predicting date offers and matches, in comparison to many other factors that have been previously linked to attraction. For both men and women, subjects' ratings of partners were the most important predictors of extending second date offers, followed by average ratings of partners by all speed-dating partners, with Attractiveness attributes, measures of desirability as a partner, and sexual chemistry standing out as particularly important. These findings mirror Joel et al.'s (2017) random forests analyses which found that subjects' ratings of chemistry with their partners and physical attractiveness of their partners were the most consistent predictors of romantic desire. In predicting matches, or mutual date offers, women's ratings of men were the most important, which is likely due to women's greater selectivity in partners (in both the current study and prior studies; Buss & Schmitt, 1993). The importance of MHC-based measures (dissimilarity, rareness, heterozygosity), was comparable to a partner's self-reported personality characteristics. It is notable that in this analysis, MHC-based measures were more important than two SNPs (rs6311, rs1799971) that we had previously found to predict speed-dating success in the same sample, accounting for 2–5% of the observed variance (Wu et al., 2016). Rs6311 and rs1799971 were found in the prior study to be associated with mate desirability measures, which were also used as predictor variables in the random forest analyses. Yet, in the random forests, MHC predicted unique variance in date offers even in the presence of these strong predictors (e.g., mate desirability), reinforcing that MHC-based attraction may operate as a subconscious mechanism, separate from traditional predictors of attraction.

Several limitations of our study must be mentioned. First, our sample consisted mainly of young Asian American college students on the West Coast, so our findings cannot be generalized to other populations. This homogeneity was part of the study design, to minimize the well-documented population specific heterogeneity in the MHC region (De Bakker et al., 2006; Trowsdale & Knight, 2013) and to control racial preferences in dating (Fisman et al., 2008).

Second, the SNP based method to determine similarity at the MHC region is significantly different from prior serological methods, making direct comparisons difficult. A strength of the SNP based approach, however, is that many other regions of the MHC region can be analyzed, not just the classical HLA gene coding regions (which represent approximately 1% of the region; De Bakker et al., 2006; Trowsdale & Knight, 2013). Further studies should expand the number of SNPs used in the current study, to refine and confirm the observed associations. Future studies should also be conducted using samples of different ages and ancestry, keeping in mind, however, that the population specific heterogeneity in the MHC region will likely require targeting different SNPs than used in the current study.

Third, because we did not conduct genome-wide tests of (dis)similarity preferences, we do not know whether our findings are specific to

the MHC region.

Finally, we only examined the role of MHC-based preferences in initial attraction. While the speed-dating format is ideal for measuring initial attraction, future studies should test whether selection may take place at later, more intimate stages of a developing relationship.

In summary, our study is the first that we know of to test for the presence of a detection system for MHC-based human mate preferences in a real-life setting and to directly compare the effects of MHC-based measures to that of other factors important in human attraction. Our study contributes to knowledge about the unseen factors in human attraction and indicates that genetic factors may be just as important as a partner's self-reported personality characteristics in predicting date offers to them.

Data availability

The data used in the above analyses are available to authorized researchers at *openICPSR* via the following link: https://doi.org/10.3886/E102340V1.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.evolhumbehav.2018.04.001.

Author contributions

Wu, Chen, Moyzis, and Greenberger designed research; Wu performed research; Wu, Chen, Moyzis, Nuno, and Yu analyzed data; Wu, Chen, Moyzis, and Yu wrote the paper.

Declarations of interest

None.

Acknowledgements

We thank Society of Research on Adolescence (SRA-53016) and Psi Chi National Honor Society (PSICHI-100461) for funding this study, and our research assistants for their work on this project: Jonathan B. Lim, Stephanie Nguyen, Blaise Lallathin, Paul Phandl, Melody Lim, Ronica Senores, Justin Huft, Gabriel Corpus, Jennifer Lai, and Marissa Tom.

References

1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature, 526*, 68–74. http://dx.doi.org/10.1038/nature15393.

Abdellaoui, A., Verweij, K. J., & Zietsch, B. P. (2014). No evidence for genetic assortative mating beyond that due to population stratification. *Proceedings of the National Academy of Sciences, 111* (E4137).

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovis, J. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology, 19*, 586–592. http://dx.doi.org/10.1037//0278-6133.19.6.586.

Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality, 40*, 440–450. http://dx.doi.org/10.1016/j.jrp.2005.03.002.

Bekkar, M., Djemaa, H. K., & Alitouche, T. A. (2013). Evaluation measures for models assessment over imbalanced datasets. *Journal of Information Engineering and Applications, 3*, 27–39.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57*, 289–300.

Brown, H. J. (2012). *Life's little instruction book: Simple wisdom and a little humor for living a happy and rewarding life.* Nashville, Tennessee: Thomas Nelson Inc.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100*, 204–232. http://dx.doi.org/10.1037/0033-295X.100.2.204.

Cao, K., Hollenbach, J., Shi, X., Shi, W., Chopek, M., & Fernández-Viña, M. A. (2001). Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Human Immunology, 62*, 1009–1030. http://dx.doi.org/10.1016/S0198-8859(01)00298-1.

Chaix, R., Cao, C., & Donnelly, P. (2008). Is mate choice in humans MHC-dependent? *PLoS Genetics, 4*, e1000184. http://dx.doi.org/10.1371/journal.pgen.1000184.

Coetzee, V., Barrett, L., Greeff, J. M., Henzi, S. P., Perrett, D. I., & Wadee, A. A. (2007). Common HLA alleles associated with health, but not with facial attractiveness. *PLoS One, 2*, e640. http://dx.doi.org/10.1371/journal.pone.0000640.

De Bakker, P. I., McVean, G., Sabeti, P. C., Miretti, M. M., Green, T., Marchini, J., et al. (2006). A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nature Genetics, 38*, 1166–1172. http://dx.doi.org/10.1038/ng1885.

Derti, A., Cenik, C., Kraft, P., & Roth, F. P. (2010). Absence of evidence for MHC-dependent mate selection within HapMap populations. *PLoS Genetics, 6*, e1000925. http://dx.doi.org/10.1371/journal.pgen.1000925.

Domingue, B. W., Fletcher, J., Conley, D., & Boardman, J. D. (2014). Genetic and educational assortative mating among US adults. *Proceedings of the National Academy of Sciences, 111*, 7996–8000. http://dx.doi.org/10.1073/pnas.1321426111.

Doty, R. L. (1981). Olfactory communication in humans. *Chemical Senses, 6*, 351–376.

Eastwick, P. W., & Finkel, E. J. (2008). Sex differences in mate preferences revisited: Do people know what they initially desire in a romantic partner? *Journal of Personality and Social Psychology, 94*, 245–264. http://dx.doi.org/10.1037/0022-3514.94.2.245.

Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin, 140*, 623–665. http://dx.doi.org/10.1037/a0032432.

Finkel, E. J., & Eastwick, P. W. (2008). Speed-dating. *Current Directions in Psychological Science, 17*, 193–197. http://dx.doi.org/10.1111/j.1467-8721.2008.00573.

Finkel, E. J., & Eastwick, P. W. (2009). Arbitrary social norms influence sex differences in romantic selectivity. *Psychological Science, 20*, 1290–1295. http://dx.doi.org/10.1111/j.1467-9280.2009.02439.

Fisman, R., Iyengar, S. S., Kamenica, E., & Simonson, I. (2008). Racial preferences in dating. *The Review of Economic Studies, 75*, 117–132. http://dx.doi.org/10.1111/j.1467-937X.2007.00465.x.

Garver-Apgar, C. E., Gangestad, S. W., Thornhill, R., Miller, R. D., & Olp, J. J. (2006). Major histocompatibility complex alleles, sexual responsivity, and unfaithfulness in romantic couples. *Psychological Science, 17*, 830–835.

Goeman, J. J., Van De Geer, S. A., De Kort, F., & Van Houwelingen, H. C. (2004). A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics, 20*, 93–99. http://dx.doi.org/10.1093/bioinformatics/btg382.

Gomes, K. F., Santos, A. S., Semzezem, C., Correia, M. R., Brito, L. A., Ruiz, M. O., Fukui, R. T., Matioli, S. R., Passos-Bueno, M. R., & Da Silva, M. E. (2017). The influence of population stratification on genetic markers associated with type 1 diabetes. *Scientific Reports, 7*, 43513. http://dx.doi.org/10.1038/srep43513.

Goodwin, J. S., Hunt, W. C., & Samet, J. M. (1987). The effect of marital status on stage, treatment, and survival of cancer patients. *Journal of the American Medical Association, 258*, 3125–3130. http://dx.doi.org/10.1001/jama.1987.0340021006702.

Gottman, J. M., & Silver, N. (1999). *The seven principles for making marriage work: A practical guide from the country's foremost relationship expert.* USA: Three Rivers Press.

Havlicek, J., & Roberts, S. C. (2009). MHC-correlated mate choice in humans: A review. *Psychoneuroendocrinology, 34*, 497–512. http://dx.doi.org/10.1016/j.psyneuen.2008.10.007.

Hedrick, P. W., & Black, F. L. (1997). HLA and mate selection: No evidence in South Amerindians. *The American Journal of Human Genetics, 61*, 505–511. http://dx.doi.org/10.1086/515519.

Hedrick, P. W., & Thomson, G. (1988). Maternal-fetal interactions and the maintenance of HLA polymorphism. *Genetics, 119*, 205–212.

Herz, R. S., & Inzlicht, M. (2002). Sex differences in response to physical and social factors involved in human mate selection: The importance of smell for women. *Evolution and Human Behavior, 23*, 359–364. http://dx.doi.org/10.1016/S1090-5138(02)00095-8.

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. http://dx.doi.org/10.1537/ase.108.199.

Jacob, S., McClintock, M. K., Zelano, B., & Ober, C. (2002). Paternally inherited HLA alleles are associated with women's choice of male odor. *Nature Genetics, 30*, 175–179. http://dx.doi.org/10.1038/ng830.

Jin, Y., Birlea, S. A., Fain, P. R., Gowan, K., Riccardi, S. L., Holland, P. J., ... Kemp, E. H. (2011). Genome-wide analysis identifies a quantitative trait locus in the MHC class II region associated with generalized vitiligo age of onset. *Journal of Investigative Dermatology, 131*, 1308–1312. http://dx.doi.org/10.1038/jid.2011.12.

Joel, S., Eastwick, P. W., & Finkel, E. J. (2017). Is romantic desire predictable? Machine learning applied to initial romantic attraction. *Psychological Science,* 1–12. http://dx.doi.org/10.1177/0956797617714580 Advance online edition.

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research, 12*, 996–1006. http://dx.doi.org/10.1101/gr.229102.

Liu, D., Ghosh, D., & Lin, X. (2008). Estimation and testing for the effect of a genetic pathway on a disease outcome using logistic kernel machine regression via logistic mixed models. *BMC Bioinformatics, 9*, 292. http://dx.doi.org/10.1186/1471-2105-9-292.

Ludbrook, J., & Dudley, H. (1998). Why permutation tests are superior to t and F tests in biomedical research. *The American Statistician, 52*, 127–132.

Luo, S., & Zhang, G. (2009). What leads to romantic attraction: Similarity, reciprocity, security, or beauty? Evidence from a speed-dating study. *Journal of Personality, 77*, 933–964. http://dx.doi.org/10.1111/j.1467-6494.2009.00570.x.

Maiers, M., Gragert, L., & Klitz, W. (2007). High-resolution HLA alleles and haplotypes in the United States population. *Human Immunology, 68*, 779–788. http://dx.doi.org/10.1016/j.humimm.2007.04.005.

Mendez, G., Buskirk, T. D., Lohr, S., & Haag, S. (2008). Factors associated with persistence in science and engineering majors: An exploratory study using classification trees and random forests. *Journal of Engineering Education, 97*, 57–70.

Milinski, M. (2006). The major histocompatibility complex, sexual selection, and mate choice. *Annual Review of Ecology, Evolution, and Systematics, 37*, 159–186. http://dx.doi.org/10.1146/annurev.ecolsys.37.091305.110242.

Muchlinski, D., Siroky, D., He, J., & Kocher, M. (2015). Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. *Political Analysis, 24*, 87–103.

Murstein, B. I. (1972). Physical attractiveness and marital choice. *Journal of Personality and Social Psychology, 22*, 8–12. http://dx.doi.org/10.1037/h0032394.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*, 250–256. http://dx.doi.org/10.1037/0022-3514.35.4.250.

Niu, R., Wang, Y., Zhu, M., Wen, Y., Sun, J., Shen, W., ... Hu, Z. (2015). Potentially functional polymorphisms in POU5F1 gene are associated with the risk of lung cancer in Han Chinese. *BioMed Research International, 2015*, 1–7. http://dx.doi.org/10.1155/2015/851320.

Ober, C., Weitkamp, L. R., Cox, N., Dytch, H., Kostyu, D., & Elias, S. (1997). HLA and mate choice in humans. *The American Journal of Human Genetics, 61*, 497–504. http://dx.doi.org/10.1086/515511.

Penn, D. J., Damjanovich, K., & Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences, 99*, 11260–11264. http://dx.doi.org/10.1073/pnas.162006499.

Penn, D. J., & Potts, W. K. (1999). The evolution of mating preferences and major histocompatibility complex genes. *The American Naturalist, 153*, 145–164.

Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision making: A "hot" topic. *Journal of Behavioral Decision Making, 19*, 79–85. http://dx.doi.org/10.1002/bdm.528.

Qiao, Z., Powell, J., & Evans, D. (2018). MHC-dependent mate selection within 872 spousal pairs of European ancestry from the health and retirement study. *Genes, 9*, 53.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401. http://dx.doi.org/10.1177/014662167700100306.

Reynolds, J. R. (1997). The effects of industrial employment conditions on job-related distress. *Journal of Health and Social Behavior, 38*, 105–116. http://dx.doi.org/10.2307/2955419.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38*, 119–125.

Roberts, S. C., Gosling, L. M., Carter, V., & Petrie, M. (2008). MHC-correlated odour preferences in humans and the use of oral contraceptives. *Proceedings of the National Academy of Sciences of the United States of America, 275*, 2715–2722. http://dx.doi.org/10.1098/rspb.2008.0825.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton university press.

Saito, S., Ota, S., Yamada, E., Inoko, H., & Ota, M. (2001). Allele frequencies and haplotypic associations defined by allelic DNA typing at HLA class I and class II loci in the Japanese population. *HLA, 56*, 522–529. http://dx.doi.org/10.1034/j.1399-0039.2000.560606.x.

Santos, P. S. C., Schinemann, J. A., 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. http://dx.doi.org/10.1016/j.yhbeh.2004.11.005.

Saphire-Bernstein, S., Larson, C. M., Gildersleeve, K. A., Fales, M. R., Pillsworth, E. G., & Haselton, M. G. (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. http://dx.doi.org/10.1016/j.evolhumbehav.2016.09.003.

Smith, T. W., & Gallo, L. C. (1999). Hostility and cardiovascular reactivity during marital interaction. *Psychosomatic Medicine, 61*, 436–445.

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods, 14*, 323–348. http://dx.doi.org/10.1037/a0016973.

Thornhill, R., Gangestad, S. W., Miller, R., Scheyd, G., McCollough, J. K., & Franklin, M. (2003). Major histocompatibility complex genes, symmetry, and body scent attractiveness in men and women. *Behavioral Ecology, 14*, 668–678. http://dx.doi.org/10.1093/beheco/arg043.

Trachtenberg, E., Vinson, M., Hayes, E., Hsu, Y. M., Houtchens, K., Erlich, H., ... Hollenbach, J. (2007). HLA class I (A, B, C) and class II (DRB1, DQA1, DQB1, DPB1) alleles and haplotypes in the Han from southern China. *HLA, 70*, 455–463. http://dx.doi.org/10.1111/j.1399-0039.2007.00932.x.

Trowsdale, J., & Knight, J. C. (2013). Major histocompatibility complex genomics and human disease. *Annual Review of Genomics and Human Genetics, 14*, 301–323. http://dx.doi.org/10.1146/annurev-genom-091212-153455.

Unanue, E. R., Turk, V., & Neefjes, J. (2016). Variations in MHC class II antigen processing and presentation in health and disease. *Annual Review of Immunology, 34*, 265–297. http://dx.doi.org/10.1146/annurev-immunol-041015-055420.

Vierling, J. S., & Rock, J. (1967). Variations in olfactory sensitivity to exaltolide during the menstrual cycle. *Journal of Applied Physiology, 22*, 311–315.

Wedekind, C., & Füri, 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. http://dx.doi.org/10.1098/rspb.1997.0204.

Wedekind, C., Seebeck, T., Bettens, F., & Paepke, A. J. (1995). MHC-dependent mate preferences in humans. *Proceedings of the Royal Society of London B: Biological Sciences, 260*, 245–249. http://dx.doi.org/10.1098/rspb.1995.0087.

Winternitz, J., Abbate, J. L., Huchard, E., Havlíček, J., & Garamszegi, L. Z. (2017). Patterns of MHC-dependent mate selection in humans and nonhuman primates: A meta-analysis. *Molecular Ecology, 26*, 668–688. http://dx.doi.org/10.1111/mec.13920.

Wu, K., Chen, C., & Greenberger, E. (2015). The sweetness of forbidden fruit: Interracial daters are more attractive than intraracial daters. *Journal of Social and Personal Relationships, 32*, 650–666. http://dx.doi.org/10.1177/0265407514541074.

Wu, K., Chen, C., Moyzis, R. K., Greenberger, E., & Yu, Z. (2016). Gender interacts with opioid receptor polymorphism A118G and serotonin receptor polymorphism- 1438 A/G on speed-dating success. *Human Nature, 27*, 244–260. http://dx.doi.org/10.1007/s12110-016-9257-8.

Yamazaki, K., Beauchamp, G. K., Kupniewski, D., Bard, J., Thomas, L., & Boyse, E. A. (1988). Familial imprinting determines H-2 selective mating preferences. *Science, 240*, 1331–1332. http://dx.doi.org/10.1126/science.3375818.

Zhang, Z., & Xiang, L. F. (2014). Genetic susceptibility to vitiligo: Recent progress from genome-wide association studies. *Dermatologica Sinica, 32*, 225–232. http://dx.doi.org/10.1016/j.dsi.2014.09.004.

Zhu, K. J., Lv, Y. M., Yin, X. Y., Wang, Z. X., Sun, L. D., He, S. M., ... Zuo, X. B. (2011). Psoriasis regression analysis of MHC loci identifies shared genetic variants with vitiligo. *PLoS One, 6*, e23089. http://dx.doi.org/10.1371/journal.pone.0023089.

Zou, J. Y., Park, D. S., Burchard, E. G., Torgerson, D. G., Pino-Yanes, M., Song, Y. S., ... Zaitlen, N. (2015). Genetic and socioeconomic study of mate choice in Latinos reveals novel assortment patterns. *Proceedings of the National Academy of Sciences, 112*, 13621–13626.