BioSystems 187 (2020) 104040

Contents lists available at ScienceDirect

BioSystems

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

Assortative mating at loci under recent natural selection in humans

Akihiro Nishi[a,*], Marcus Alexander[b], James H. Fowler[c], Nicholas A. Christakis[b,d]

[a] Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, CA 90095, USA
[b] Yale Institute for Network Science, Yale University, CT 06520, USA
[c] Division of Medical Genetics and Department of Political Science, University of California, San Diego, La Jolla, CA, 92103, USA
[d] Department of Sociology, Ecology and Evolutionary Biology, Medicine, Biomedical Engineering, and Statistics & Data Science, Yale University, New Haven, CT, USA

ARTICLE INFO

Keywords:
Mate choice
Assortative mating
Positive selection
Humans

ABSTRACT

Genetic correlation between mates at specific loci can greatly alter the evolutionary trajectory of a species. Genetic assortative mating has been documented in humans, but its existence beyond population stratification (shared ancestry) has been a matter of controversy. Here, we develop a method to measure assortative mating across the genome at 1,044,854 single-nucleotide polymorphisms (SNPs), controlling for population stratification and cohort-specific cryptic relatedness. Using data on 1683 human couples from two data sources, we find evidence for both assortative and disassortative mating at specific, discernible loci throughout the entire genome. Then, using the composite of multiple signals (CMS) score, we also show that the group of SNPs exhibiting the most assortativity has been under stronger recent positive selection. Simulations using realistic inputs confirm that assortative mating might indeed affect changes in allele frequency over time. These results suggest that genetic assortative mating may be speeding up evolution in humans.

1. Introduction

Individuals actively choose with whom to mate in many species, including humans (Neff and Pitcher, 2005; Mays and Hill, 2004; Roberts and Little, 2008; Andersson and Simmons, 2006; Jones and Ratterman, 2009). One form of mate choice is assortative mating, which has been well described for specific traits such as body size, personality, and other apparent features (Jones and Ratterman, 2009; Redden and Allison, 2006; Jiang et al., 2013; Russell et al., 1985; Thiessen and Gregg, 1980). Disassortative mating has also been described for several traits, including immune function in humans (Mays and Hill, 2004; Roberts and Little, 2008; Andersson and Simmons, 2006; Jiang et al., 2013; Laurent and Chaix, 2012). A century ago, Fisher and Wright proposed that such positive and negative phenotypic correlations between mates would correspond to genetic correlations between them (Wright, 1920; Fisher, 1918). Hypothetically, considering the polygenic basis for most phenotypes, assortative or disassortative mating could thus result in marginal to strong genetic correlations at thousands of loci. Theory, and evidence from another taxon (i.e., *Drosophila pseudoobscura*), shows that assortative mating at even one locus can shift the adaptive landscape and yield sympatric speciation (Kirkpatrick and Nuismer, 2004; Otto et al., 2008; Williams and Sarkar, 1994; Ortiz-Barrientos and Noor, 2005).

However, previous studies have focused on whether assortative mating exists at the whole-genome level or genetic-region level rather than at specific loci, and the existence of "active" genetic assortative mating has been controversial (Sebro et al., 2010; Laurent et al., 2012; Domingue et al., 2014a; Abdellaoui et al., 2014; Domingue et al., 2014b; Sorokowska et al., 2019). Sebro and colleagues report evidence of assortative mating at principal components (n = 33 couples), suggesting that a certain type of population stratification (genetic similarity due to the sharing of similar ancestry) results in genetic assortativity (Sebro et al., 2010). Meanwhile, Laurent and colleagues report evidence of assortative and disassortative mating at multiple autosomal genetic regions (n = 50 couples of European descent) (Laurent et al., 2012). Domingue and colleagues used a kinship coefficient to provide evidence for genetic assortative mating at the whole-genome level in a larger sample (n = 825 couples) (Domingue et al., 2014a). Although the primary aim of that effort was to descriptively compare the degree of educational assortative mating (i.e., assortment by educational attainment) with that of genetic assortative mating (Domingue et al., 2014a, b), Abedellaoui and colleagues subsequently questioned whether the methods used by Domingue and colleagues could adequately control for population stratification (Domingue et al., 2014a; Abdellaoui et al., 2014). Since randomized controlled trials of genetic assortative mating in humans are clearly impossible, methodological

* Corresponding author at: Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, 650 Charles E Young Dr S, Los Angeles, CA 90095.
E-mail addresses: akihironishi@ucla.edu (A. Nishi), marcus.alexander@yale.edu (M. Alexander), fowler@ucsd.edu (J.H. Fowler), nicholas.christakis@yale.edu (N.A. Christakis).

https://doi.org/10.1016/j.biosystems.2019.104040
Received 24 August 2019; Received in revised form 23 September 2019; Accepted 24 September 2019
Available online 01 October 2019
0303-2647/

advances using observational data are required to delimit the potential influence of population stratification (as much as possible) and to more precisely understand the genetic architecture of human mating patterns.

Therefore, we address three key questions regarding genetic assortative mating in humans: whether genotypic assortative or disassortative mating happens at the whole-genome level (Study 1), confirming and extending prior work; whether genotypic assortative or disassortative mating happens at the SNP level for discernible loci (Study 2); and whether there is a relationship between genetic assortative and disassortative mating and natural selection (Studies 3 and 4), a topic that, to our knowledge, has not received prior empirical attention in humans.

To answer these questions, we use genome-wide data on 1,044,854 autosomal single-nucleotide polymorphisms (SNPs) with minor allele frequency (MAF) > 0.10 from 1683 unrelated heterosexual spousal pairs of European descent (3366 individuals total, restricted to first marriages) from two independent datasets (the Health and Retirement Study, and the Framingham Heart Study) (see Methods). For certain comparisons, we permuted the spousal relationships in the data to create five sets of randomly formed opposite-sex pairs (N = 8415 pairs); crucially, the only difference between the observed data and the permuted data is the set of relationships – the sample size, individual characteristics, and distribution of genotypes in the random samples are identical to the observed samples. The goal of our work here is not to identify specific genes guiding mate choice per se; rather, it is to measure patterns across the whole genome.

2. Results

2.1. Study 1 (whole-genome assortativity)

For the whole-genome-level analysis, we aim to discern whether the spousal pairs shared an ancestor in common, as compared with the randomly formed opposite-sex pairs drawn from the same sample (the stranger pairs) (see Supporting Information for the formulas and further details). Hence, we calculate the kinship coefficient as a metric of how similar spousal genomes are and also as a way of discerning whether the spousal pairs shared a recent ancestor (Domingue et al., 2014a; Abdellaoui et al., 2014; Christakis and Fowler, 2014). Overall, we find that the genomes of observed spousal pairs are significantly more similar than that of random opposite-sex pairs drawn from the same population (regression analysis, $\beta = 0.000898$, $P = 4.76 \times 10^{-5}$) (Fig. 1 & Supplementary Table 1). As a benchmark, the magnitude of this similarity roughly corresponds to the similarity between 4th and 5th cousins.

Although the procedures and sample inclusion criteria are different, these results (using a larger sample size from two datasets) generally reproduce the results of Domingue and colleagues (Domingue et al., 2014a). These results do not substantially change when we used another measure of genetic relatedness (identity by state [IBS] distance: $\beta = 0.000113$, $P = 8.03 \times 10^{-4}$). Moreover, these results do not substantially change even after controlling for the top ten principal components (PCs) of both spouses (a total of 20 PCs) (Supplementary Table 1). On the other hand, when we analyze the data of the Framingham Heart Study alone, the magnitude of spousal genetic similarity is not significant for IBS distance ($P = 0.467$), while that for kinship coefficient is significant ($P = 1.74 \times 10^{-7}$) (see Methods and Supplementary Table 1). Therefore, as Abdellaoui and colleagues recently stated (Abdellaoui et al., 2014), evidence for spousal genetic similarity may be less clearly identified when IBS distance is used, or this may reflect the intrinsic difference between what IBS distance measures and what the kinship coefficient measures.

Hence, we find whole-genome assortativity among spousal pairs. However, the question remains whether this provides sufficient evidence for "active" genetic assortative mating in humans. Investigating genetic assortative mating at the whole-genome level has two limitations. First, since the kinship coefficient and IBS distance capture the overall magnitude of shared ancestry (Manichaikul et al., 2010), examining if "active" genetic assortative mating (beyond population stratification) is occurring, using these genetic relatedness measures, can be problematic. Rather, the results of Study 1 can be interpreted descriptively as a mixture of potential genotypic assortative (and disassortative) mating in addition to any residual cryptic population stratification. Second, it is known that some forms of population stratification can be refractory to correction by PCs (McVean, 2009), though the inclusion of PCs in the analysis is still a recommended procedure in genetic epidemiology and evolutionary genetics.

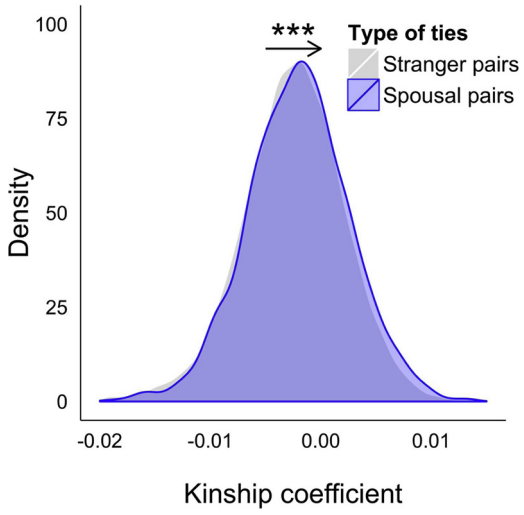

Fig. 1. Data from 1683 observed spousal pairs in the Health and Retirement Study (HRS) and the Framingham Heart Study (FHS) shows evidence of assortative mating at the whole-genome level. The mean kinship coefficient for spouse pairs is significantly higher than random opposite-sex pairs drawn from the same population. *P* values calculated by multiway clustered standard errors accounting for multiple observations of same individuals in observed and random pairs (see Methods).

2.2. Study 2 (SNP-level assortativity)

Next, we extend our analysis to the SNP level, developing an expanded method to examine genotypic assortative and disassortative mating beyond the potential influence of population stratification, and applying this method to each SNP across the whole genome. The goal of this analysis is not to identify specific genes involved in mate choice; rather, it is to measure patterns across the genome, for which this sample size is adequate, as in previous work (Christakis and Fowler, 2014; Ward and Kellis, 2012).

We construct a regression model to examine the degree of assortativity (correlation in dosage of 0, 1, or 2) at each SNP controlling for population stratification and cohort-specific cryptic relatedness (regressing the dosage of a male individual on that of a coupled female individual, along with covariates) (see Methods). In an ideal situation, we would test if a female individual with experimentally manipulated higher dosages (e.g., a dosage of 2) at a single locus is more likely to marry a male individual with experimentally manipulated higher dosages at the same locus, keeping the genetic background (i.e., the combination of SNP genotypes in all the other loci) unchanged for both the female and male individuals (Fig. 2a). To test this hypothesis using humans, where such gene manipulation is obviously not feasible, we implement the following procedures: (i) we calculate the top 10 PCs from the subject-SNP data matrix (Manichaikul et al., 2010) for both spouses and include them as covariates in regression models; (ii) we identify the genetically closest individual within a cohort for each

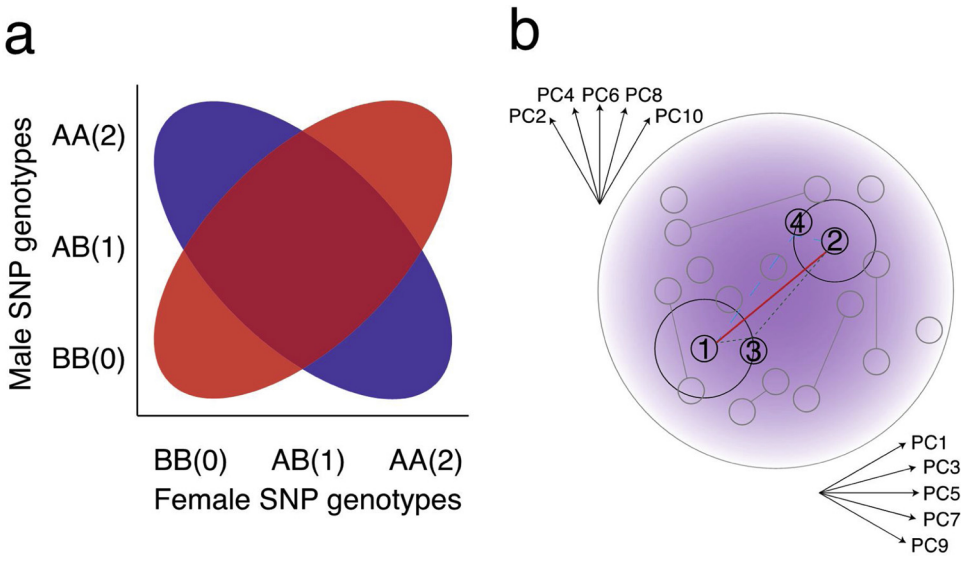

Fig. 2. Assortative and disassortative mating at the SNP level can be explained by population stratification or by active mate choice. **(a)** A and B are two alleles at a single locus. At a single locus, a couple composed of a female with AA (dosage = 2) and a male with AA reflects assortative mating (red), while a couple composed of a female with AA and a male with BB (dosage = 0) reflects disassortative mating (blue). The red and blue shades are illustrations of the rough probability densities of the two alleles of two individuals in couples under assortative mating and disassortative mating, respectively. **(b)** The genetic background of 20 illustrative individuals (indicated by circles) drawn at 10 dimensions of principal components (PCs) is shown. The top 10 axes of the principal components can show the variation of genetic background among individuals. When individual 1 marries individual 2, the mating can originate from active mate choice or from population stratification. Here, individuals 1 and 3 share similar genetic background, and individuals 2 and 4 share similar genetic background (the kinship coefficient between them is at a maximum). The assortativity or disassortativity of SNP genotypes in individuals 1 and 2 that cannot be explained by SNP genotypes of the genetically closest individuals (3 and 4) is the mating pattern that population stratification (genetic background) cannot explain. The red line represents the mating pattern not coming from population stratification, while the blue and green lines represent the mating pattern coming from population stratification. The grey lies represent the stranger pairs of non-married heterosexual individuals. PC1 to PC10 represent the 1st to 10th principal components, respectively.

subject (using one-for-one nearest matching on the kinship coefficient (Manichaikul et al., 2010); and (iii) we include the dosage of the genetically closest individual at the same locus as a covariate in regression models (Fig. 2b). These procedures jointly allow us to take into account *both* the global and local genetic background (population stratification) within a cohort (Fig. 2b). This novel procedure can control for the residual population stratification that cannot be taken into account solely by the inclusion of ten PCs for each spouse in the model.

We also generate a comparison group of stranger pairs drawn from the same sample, each of which contains one male and one female who are not married to each other. Including this group in the analysis allows us to measure the degree to which assortativity differs from what we would expect due to cryptic relatedness generated by simply random mating within the same population structure (see Supplementary Methods for the formula and details).

Fig. 3a shows the Q-Q plot from the regression models across the genome. The λ statistic of this plot is 1.012, suggesting that the majority of SNPs do not exhibit assortativity or disassortativity. We found 23 SNPs exhibiting moderate-level assortative mating and 40 SNPs exhibiting moderate-level disassortative mating ($P < 5.0 \times 10^{-5}$, no SNPs with $P < 5.0 \times 10^{-8}$). However, this is what we would expect if there were widespread low-level genetic correlation (either positive or negative) in spouses across the genome, and it is consistent with recent work that shows that polygenic traits can generate inflation factors of these magnitudes (Yang et al., 2011) (please see an additional analysis regarding polygenic traits in the Supplementary Methods, Box. 1). Fig. 3b shows that the distributions of Z statistics of the estimates for the degree of assortativity in the regression models indeed differ from null distributions for both assortative loci (Kolmogorov-Smirnov test, $P = 0.0137$) and disassortative loci ($P = 2.71 \times 10^{-8}$) (see Methods). In contrast to previous work that focused exclusively on genetic regions involved in the major histocompatibility complex (MHC) regions (Roberts and Little, 2008; Laurent and Chaix, 2012; Chaix et al., 2008; Derti et al., 2010), these results suggest that assortative and disassortative mating is occurring at the allelic level across the human genome as a whole. We did not find substantial evidence for disassortative mating specifically in the MHC regions (see Supplementary Methods and Supplementary Fig. 4), in keeping with some past work (Chaix et al., 2008; Derti et al., 2010) (though this is a topic of ongoing debate (Laurent and Chaix, 2012; Derti and Roth, 2012)).

2.3. Study 3 (Role of assortative mating in natural selection)

To explore what role assortativity may be playing in human evolution, we measure its association with the Composite of Multiple Signals (CMS) score, an index that combines several signals of recent positive selection to identify which parts of the genome have been evolving the fastest over the last 30,000 years (Grossman et al., 2010). One caution here is that, since the CMS score is influenced by selective sweeps (Grossman et al., 2010), the high frequency of long haplotypes and large allele frequency changes that are observed around loci with selective sweeps may yield false positive results in further analyses. In our regression models, we control for minor allele frequency and differences between chromosomes, and we take into account potential serial correlation between SNPs due to linkage disequilibrium (LD) patterns using Newey-West standard errors (see Supplementary Methods). Fig. 3c shows that SNPs exhibiting moderate assortativity among spousal pairs have a significantly higher CMS score than SNPs exhibiting no assortative or disassortative mating ($P = 8.36 \times 10^{-4}$). In contrast, SNPs exhibiting moderate-level disassortative mating show no such pattern ($P = 0.992$). These results suggest that positively correlated genotypes in mates are under stronger positive selection in humans.

2.4. Study 4 (Assortative mating and allele frequency change)

Using simulations, we examine the influence of the observed level of genotypic assortative and disassortative mating on the potential change in allele frequency if the same mating patterns last for multiple generations (see Supplementary Methods). First, we prepare 1000 individuals (500 males and 500 females) as the 1st generation, and two alleles (with an initial frequency for the advantageous allele of 0.10) are randomly assigned to each of them (a single locus). A female mates with a male in the population under the rule of genotypic correlation (the degree of correlation of $r = -0.05$ to 0.05 [neither assortative nor disassortative], $r = 0.05$ to 0.125 [marginally/moderately assortative], or $r = -0.125$ to -0.05 [marginally/moderately disassortative]). Second, the generated couples produce children, and the number of

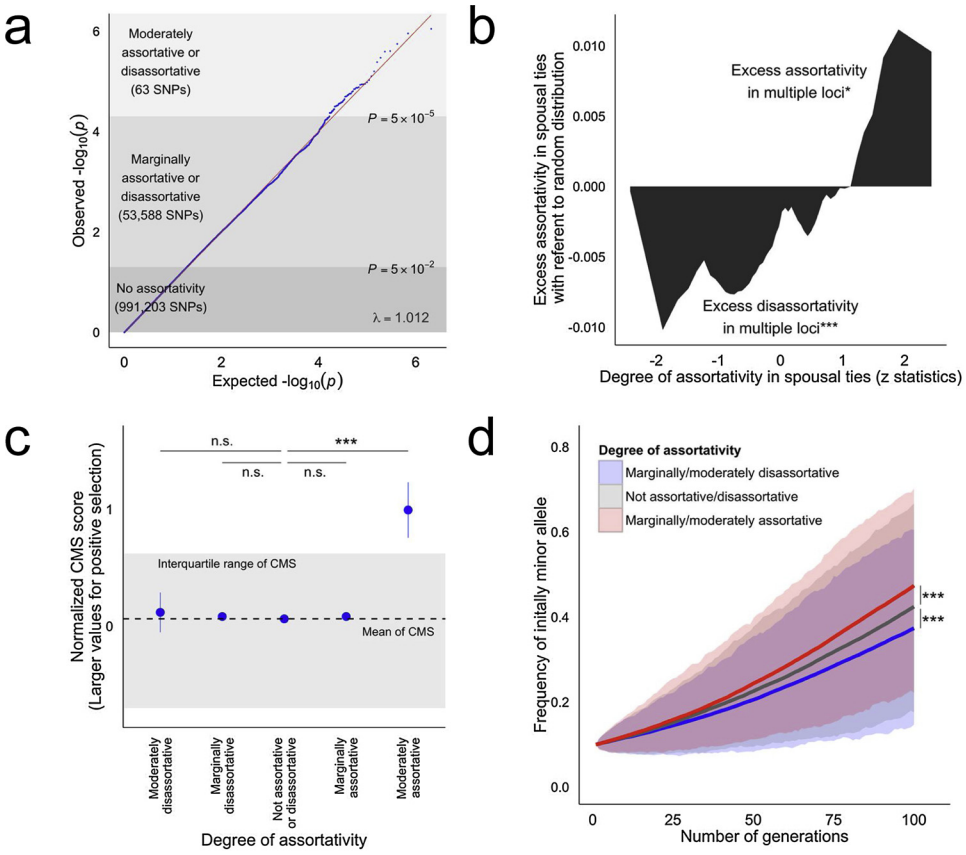

Fig. 3. The human genome evinces excess genotypic assortative mating and disassortative mating at multiple loci, and the loci exhibiting moderate assortativity among spousal pairs have higher signals of recent positive selection, which can accelerate the change in allele frequency. **(a)** Q-Q plot of the degree of assortativity at 1,044,854 SNPs shows more outliers of the observed $-\log_{10}(p)$ when the observed P is less than around 5×10^{-5}. P values are obtained from regression models. **(b)** Separate analysis for assortative and disassortative mating loci reveals excess assortativity and disassortativity at different groups of SNPs across the genome. The difference of z statistics of degree of assortativity in spousal ties to that from the standard normal distribution (null distribution) is calculated at each SNP. We made 50 bins (20,898 SNPs with similar degree of assortativity per bin). The shaded areas represent excess assortativity (positive) and excess disassortativity (negative). P values are obtained by Kolmogorov-Smirnov test. **(c)** SNPs exhibiting moderate assortativity ($P < 5.0 \times 10^{-5}$ for the genotypic correlation level, see Methods for other category definitions), have stronger signals of recent positive selection (higher composite of multiple signals [CMS] score) than SNPs exhibiting no assortativity or disassortativity. The plot shows mean CMS score by each category (blue). Vertical lines represent standard errors of the mean. For reference, the horizontal black dotted line shows the mean CMS score of all the SNPs, and the shaded area shows the interquartile range (IQR) of the CMS score. P values are obtained from regression models. **(d)** Simulations show that assortative mating (red) at a single locus over 100 generations accelerates the increase in the frequency of minor but advantageous alleles as compared with random mating (grey) or disassortative mating (blue). The lines show the 2.5 percentile (bottom), median (middle), and 97.5 percentile (top) of the 1000 iterations. P values are obtained by t-test. For all the panels, n.s. for ≥ 0.05, * for $P < 0.05$, and *** for $P < 0.001$.

children is proportional to the number of advantageous alleles in the couples. This cycle of mating and reproduction was repeated up to the 100th generation. This simulation is repeated 1000 times for each of the three settings, and a t test is performed to examine the difference of the advantageous allele frequency (see Supplementary Methods for details and R code).

Fig. 3d shows that, under reasonable assumptions about reproductive advantage, the speed of allele frequency change over 100 generations is 15.3% faster when the degree of assortative mating is at the marginal-to-moderate level (range of dosage correlation = 0.05 to 0.125 vs –0.05 to 0.05; t-test, $P < 2.2 \times 10^{-16}$), and 15.7% slower when the degree of disassortative mating is at the marginal-to-moderate level (range of dosage correlation = –0.125 to –0.05 vs –0.05 to 0.05; t-test, $P < 2.2 \times 10^{-16}$). These ranges of dosage correlation are empirically obtained from the regression analyses described earlier. While the magnitude of the change is dependent on many model assumptions, the direction of change is not. Assortative mating can speed up positive selection, which helps to explain why we find a relationship between assortativity and the CMS score at the allelic level.

3. Discussion

In sum, the results suggest that genotypic assortative and disassortative mating (here measured, in humans, as marriage) occurs at disparate loci across the human genome; that positively correlated loci tend to have higher signals of recent positive selection; and that the observed level of assortative mating might feasibly accelerate an increase in advantageous allele frequency. These results have implications for several distinct hypotheses and theories of mate choice in evolutionary biology, several of which may potentially be operating in parallel in humans.

First, these results comport with the "good genes hypothesis," which posits that individuals choose mates with an allele that increases fitness independent of the architecture of the remaining genome (Neff and Pitcher, 2005; Mays and Hill, 2004; Roberts and Little, 2008; Andersson and Simmons, 2006; Jones and Ratterman, 2009). In monogamous species with mutual mate choice (de Waal and Gavrilets, 2013; Lukas and Clutton-Brock, 2013; Hooper and Miller, 2008; Stulp et al., 2013; Baldauf et al., 2009; Lovejoy, 2009), assortative mating can result because those with the higher fitness allele choose to mate with one another, leaving those without the allele to similarly assort with their own type (Neff and Pitcher, 2005; Stulp et al., 2013; Baldauf et al., 2009). In this scenario, we would expect to find a relationship between assortative mating and positive selection at the allelic level, which can increase the frequency of advantageous alleles more rapidly, as shown in the simulations (Fig. 3d).

Second, these results could also be partly explained by "genetic similarity theory," which posits that people prefer a genetically similar individual as an optimal mate (Russell et al., 1985). Under this hypothesis, genetic similarity in a mate pair at a certain locus might confer an advantage that is not available to a solitary individual (e.g., living with a similar spouse might be efficient with respect to food choice, or could lead to higher levels of cooperation (Nowak, 2006; Fu et al., 2012; Antal et al., 2009)). In this case, genotypic assortative mating might turn an otherwise-neutral mutant genotype into a kind of "good gene." However, finding a mate with a minor allele might be costly when the mutant allele is not common, and especially when it is not initially advantageous. Therefore, it is less clear how genetic similarity theory could drive the relationship between assortativity and positive selection that we see across the genome in humans.

Third, the finding of some genotypic disassortative mating could be explained by the "compatible genes hypothesis" at such loci, which posits that individuals choose mates with an allele that increases fitness when paired with a specific homologue or allele at a locus (Neff and Pitcher, 2005; Mays and Hill, 2004; Roberts and Little, 2008; Andersson and Simmons, 2006; Jones and Ratterman, 2009). Under this hypothesis, such loci exhibit heterozygote advantage, where an individual with the two different alleles has the greatest advantage in the given environment. Therefore, mating with an individual with the different allele can be the most advantageous, which results in disassortative mating at the locus. In this scenario, genotypic disassortative mating maintains both genotypes at the locus, though it does not lead to linkage disequilibrium (LD) patterns or a higher signal of positive selection (incomplete selective sweep) (Sellis et al., 2011).

It remains an open question whether the level of whole-genome assortativity that we observe optimizes overall fitness in humans. Intriguingly, Helgason and colleagues report a significant parabolic association between fertility and actual kinship in a large Icelandic population, with optimal fertility seen in couples related at the 3rd to 4th cousin level (Helgason et al., 2008). Our data show a mixture of genotypic assortative and disassortative mating in addition to a potential residual of population stratification, but, overall, we find a whole-genome assortativity with a magnitude similar to that between 4th and 5th cousins, and the relationship between assortativity and positive selection also suggests that some degree of similarity increases fitness.

A recent paper (Christakis and Fowler, 2014) that explored the genetic similarity of friends reported evidence for genetic homophily (assortative friendship ties) and heterophily (disassortative friendship ties) in humans, and showed a similar pattern regarding signals of recent positive selection, namely, a significant association of loci exhibiting assortativity with higher CMS score, and no association of loci exhibiting disassortativity with CMS scores. These results suggest that at least a part of any higher fitness arising due to assortative tie formation (whether in friends or spouses) might indeed arise *without* reproduction; and homophily/heterophily and assortative/disassortative mating, which have been of central interest in the social sciences, may have implications for evolutionary theory (Skyrms et al., 2014).

Finally, these results illustrate that, although humans have recently experienced the agricultural and technological revolutions of the Holocene, they may still be influenced by ongoing evolutionary forces in which sexual selection interacts with natural selection (Jones and Ratterman, 2009; Courtiol et al., 2012; Stearns et al., 2012). Moreover, patterns of mate choice, like friendship choice (Christakis and Fowler, 2014; Fowler et al., 2011), may in turn help to partly explain the acceleration of human evolution seen in the last 30,000 years (Hawks et al., 2007), especially at candidate good genes loci, and may play a role in the maintenance of genome diversity at compatible genes loci. The underlying biology of genotypic assortative and disassortative mating, which might be related to "greenbeard" phenotypes (West and Gardner, 2010), kin recognition mechanisms (e.g., maternal perinatal association and co-residence duration) (Lieberman et al., 2007), social familiarity (Okuyama et al., 2014), similarity in facial morphology (DeBruine et al., 2011), or multiple chemical signals (Roberts and Little, 2008; Laurent and Chaix, 2012; Johansson and Jones, 2007), warrants further research. Finally, and of course, mate choice and even monogamy itself are partially the product of cultural factors, and, therefore, the effect of cultural and ecological inheritance surely also plays a role, in a process of gene-culture coevolution (Richerson et al., 2010).

4. Methods

We briefly summarize the methods at the beginning of each study above (Studies 1–4). In the Supplementary Information, we explain the methods for each of these studies in detail.

Declaration of Competing Interest

The authors have no conflict of interest.

Acknowledgements

We thank Lauren Brent, Yukinori Okada, Richard Prum, and Stephen Stearns for helpful comments at earlier versions of the manuscript. A.N. was supported by the Japan Society for the Promotion of Science (JSPS) for his research at Yale University. This work was supported by the National Institute on Aging (NIA) (P01AG031093) and by the National Institute for General Medical Sciences (P41GM103504-03). The Health and Retirement Survey is supported by NIA (U01AG009740). The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01−HC-25195). This manuscript does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or NHLBI. Funding for SHARe Affymetrix genotyping was provided by NHLBI Contract N02-HL-64278. Data was downloaded from NIH dbGap, project #780, with accession numbers phs000153.SocialNetwork.v6.p5.c1.GRU - general research use

phs000153.SocialNetwork.v6.p5.c2.NPU.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.biosystems.2019.104040.

References

Abdellaoui, A., Verweij, K.J.H., Zietsch, B.P., 2014. No evidence for genetic assortative mating beyond that due to population stratification. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1410781111.

Andersson, M., Simmons, L.W., 2006. Sexual selection and mate choice. Trends Ecol. Evol. 21, 296–302.

Antal, T., Ohtsuki, H., Wakeley, J., Taylor, P.D., Nowak, M.A., 2009. Evolution of cooperation by phenotypic similarity. Proc Natl Acad Sci U S A 106, 8597–8600.

Baldauf, S.A., Kullmann, H., Schroth, S.H., Thunken, T., Bakker, T.C., 2009. You can't always get what you want: size assortative mating by mutual mate choice as a resolution of sexual conflict. BMC Evol. Biol. 9, 129.

Chaix, R., Cao, C., Donnelly, P., 2008. Is mate choice in humans MHC-dependent? PLoS Genet. 4, e1000184.

Christakis, N.A., Fowler, J.H., 2014. Friendship and natural selection. Proc. Natl. Acad. Sci. U. S. A. 111, 10796–10801.

Courtiol, A., Pettay, J.E., Jokela, M., 2012. Natural and sexual selection in a monogamous historical human population. Proc. Natl. Acad. Sci. U. S. A. 109, 8044–8049.

de Waal, F., Gavrilets, S., 2013. Monogamy with a purpose. Proc Natl Acad Sci U S A 110, 15167–15168.

DeBruine, L.M., et al., 2011. Opposite-sex siblings decrease attraction, but not prosocial attributions, to self-resembling opposite-sex faces. Proc. Natl. Acad. Sci. U. S. A. 108, 11710–11714.

Derti, A., Roth, F.P., 2012. Response to "MHC-dependent mate choice in humans: Why genomic patterns from the HapMap European American data set support the hypothesis". Bioessays 34, 576–577. https://doi.org/10.1002/bies.201100150.

Derti, A., Cenik, C., Kraft, P., Roth, F.P., 2010. Absence of evidence for MHC-dependent mate selection within HapMap populations. PLoS Genet. 6, e1000925.

Domingue, B.W., Fletcher, J., Conley, D., Boardman, J.D., 2014a. Genetic and educational assortative mating among US adults. Proc. Natl. Acad. Sci. U. S. A. 111, 7996–8000.

Domingue, B.W., Fletcher, J.M., Conley, D., Boardman, J.D., Reply to Abdellaoui, et al., 2014b. Interpreting GAM. Proc. Natl. Acad. Sci. U. S. A.

Fisher, B.A., 1918. XV. The correlation between relatives on the supposition of mendelian inheritance. Trans. Roy Soc. Edin. LII, 399–433.

Fowler, J.H., Settle, J.E., Christakis, N.A., 2011. Correlated genotypes in friendship networks. Proc. Natl. Acad. Sci. U. S. A. 108, 1993–1997.

Fu, F., Nowak, M.A., Christakis, N.A., Fowler, J.H., 2012. The evolution of homophily. Sci. Rep. 2, 845.

Grossman, S.R., et al., 2010. A composite of multiple signals distinguishes causal variants in regions of positive selection. Science 327, 883–886.

Hawks, J., Wang, E.T., Cochran, G.M., Harpending, H.C., Moyzis, R.K., 2007. Recent acceleration of human adaptive evolution. Proc. Natl. Acad. Sci. U. S. A. 104, 20753–20758.

Helgason, A., Palsson, S., Gudbjartsson, D.F., Kristjansson, T., Stefansson, K., 2008. An association between the kinship and fertility of human couples. Science 319, 813–816.

Hooper, P.L., Miller, G.F., 2008. Mutual mate choice can drive costly signaling even under perfect monogamy. Adapt. Behav. 16, 53–70.
Jiang, Y., Bolnick, D.I., Kirkpatrick, M., 2013. Assortative mating in animals. Am. Nat. 181, E125–138.
Johansson, B.G., Jones, T.M., 2007. The role of chemical comunication in mate choice. Biol Rev 82, 265–289.
Jones, A.G., Ratterman, N.L., 2009. Mate choice and sexual selection: what have we learned since Darwin? Proc. Natl. Acad. Sci. U. S. A. 106, 10001–10008.
Kirkpatrick, M., Nuismer, S.L., 2004. Sexual selection can constrain sympatric speciation. Proceedings of the Royal Society B-Biological Sciences 271, 687–693.
Laurent, R., Chaix, R., 2012. MHC-dependent mate choice in humans: why genomic patterns from the HapMap European American dataset support the hypothesis. Bioessays 34, 267–271.
Laurent, R., Toupance, B., Chaix, R., 2012. Non-random mate choice in humans: insights from a genome scan. Mol. Ecol. 21, 587–596.
Lieberman, D., Tooby, J., Cosmides, L., 2007. The architecture of human kin detection. Nature 445, 727–731.
Lovejoy, C.O., 2009. Reexamining human origins in light of Ardipithecus ramidus. Science 326 74e71-78.
Lukas, D., Clutton-Brock, T.H., 2013. The evolution of social monogamy in mammals. Science 341, 526–530.
Manichaikul, A., et al., 2010. Robust relationship inference in genome-wide association studies. Bioinformatics 26, 2867–2873.
Mays Jr., H.L., Hill, G.E., 2004. Choosing mates: good genes versus genes that are a good fit. Trends Ecol. Evol. 19, 554–559.
McVean, G., 2009. A genealogical interpretation of principal components analysis. PLoS Genet. 5, e1000686.
Neff, B.D., Pitcher, T.E., 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. Mol. Ecol. 14, 19–38.
Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314, 1560–1563.
Okuyama, T., et al., 2014. A neural mechanism underlying mating preferences for familiar individuals in medaka fish. Science 343, 91–94.
Ortiz-Barrientos, D., Noor, M.A., 2005. Evidence for a one-allele assortative mating locus. Science 310, 1467.
Otto, S.P., Servedio, M.R., Nuismer, S.L., 2008. Frequency-dependent selection and the evolution of assortative mating. Genetics 179, 2091–2112.
Redden, D.T., Allison, D.B., 2006. The effect of assortative mating upon genetic association studies: spurious associations and population substructure in the absence of admixture. Behav. Genet. 36, 678–686.
Richerson, P.J., Boyd, R., Henrich, J., 2010. Colloquium paper: gene-culture coevolution in the age of genomics. Proc. Natl. Acad. Sci. U. S. A. 107 (Suppl 2), 8985–8992.
Roberts, S.C., Little, A.C., 2008. Good genes, complementary genes and human mate preferences. Genetica 132, 309–321.
Russell, R.J.H., Wells, P.A., Rushton, J.P., 1985. Evidence for genetic similarity detection in human marriage. Ethol. Sociobiol. 6, 183–187.
Sebro, R., Hoffman, T.J., Lange, C., Rogus, J.J., Risch, N.J., 2010. Testing for non-random mating: evidence for ancestry-related assortative mating in the Framingham heart study. Genet. Epidemiol. 34, 674–679.
Sellis, D., Callahan, B.J., Petrov, D.A., Messer, P.W., 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. Proc Natl Acad Sci U S A 108, 20666–20671.
Skyrms, B., Avise, J.C., Ayala, F.J., 2014. In the light of evolution VIII: darwinian thinking in the social sciences. Proc. Natl. Acad. Sci. U. S. A. 111, 10781–10784.
Sorokowska, A., et al., 2019. Assortative mating and the evolution of desirability covariation. Evol. Hum. Behav. 40, 479–491.
Stearns, S.C., Govindaraju, D.R., Ewbank, D., Byars, S.G., 2012. Constraints on the coevolution of contemporary human males and females. Proceedings Biological sciences / The Royal Society 279, 4836–4844.
Stulp, G., Buunk, A.P., Kurzban, R., Verhulst, S., 2013. The height of choosiness: mutual mate choice for stature results in suboptimal pair formation for both sexes. Anim. Behav. 86, 37–46.
Thiessen, D., Gregg, B., 1980. Human assortative mating and genetic equilibrium - an evolutionary perspective. Ethol. Sociobiol. 1, 111–140.
Ward, L.D., Kellis, M., 2012. Evidence of abundant purifying selection in humans for recently acquired regulatory functions. Science 337, 1675–1678.
West, S.A., Gardner, A., 2010. Altruism, spite, and greenbeards. Science 327, 1341–1344.
Williams, S.M., Sarkar, S., 1994. Assortative Mating and the Adaptive Landscape. Evolution 48, 868–875.
Wright, S., 1920. Systems of mating. III. Assortative mating based on somatic resemblance. Genetics 6, 144–161.
Yang, J., et al., 2011. Genomic inflation factors under polygenic inheritance. Eur. J. Hum. Genet. 19, 807–812.