Received: 21 June 2020 | Revised: 13 November 2020 | Accepted: 18 November 2020

DOI: 10.1111/jeb.13745

RESEARCH PAPER

JOURNAL OF Evolutionary Biology ESEB WILEY

The role of balancing selection in maintaining human RhD blood group polymorphism: A preregistered cross-sectional study

Jaroslav Flegr[1] | Jan Toman[1] | Martin Hůla[1] | Šárka Kaňková[1,2]

[1]Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic

[2]Department of Applied Neuroscience and Neuroimaging, National Institute of Mental Health, Klecany, Czech Republic

Correspondence
Jaroslav Flegr, Faculty of Science, Viničná 7, 128 44, Prague, Czech Republic.
Email: flegr@cesnet.cz

Funding information
Grantová Agentura, Univerzita Karlova, Grant/Award Number: 204056; Grantová Agentura České Republiky, Grant/Award Number: 18-13692S

Abstract

Maintenance of genetic polymorphism remains one of the big questions of evolutionary biology, which for a long time tended to be explained by balancing selection. This explanation was later criticized, but now is again accepted as an important mechanism in evolution. Human blood group systems seem affected by balancing selection especially strongly. In this preregistered study, we focused on stable coexistence of RhD-positive and RhD-negative subjects in a population. This is an evolutionary conundrum, because carriers of the less frequent negative allele suffer from lower fecundity due to haemolytic disease of the newborn affecting RhD-positive infants born to RhD-negative women. One explanation of persisting stability of RhD polymorphism points to heterozygote advantage. Over the past decade, numerous studies demonstrated that RhD-positive subjects score better than RhD-negative homozygotes in psychomotor tests and physical health-related variables. Still, evidence of better health and performance of heterozygotes is scarce and merely indirect. We compared the physical and mental health of 2,539 subjects whose RhD genotype was estimated based on their and their parents' RhD phenotype. We confirmed that RhD-negative homozygotes fare worse in terms of physical and mental health than subjects with RhD-positive phenotype and that RhD-positive heterozygotes enjoy better health than both homozygotes. For the first time, we demonstrated that RhD-positive homozygotes might suffer from worse health than RhD-negative homozygotes. Our results strongly support the hypothesis that RhD polymorphism is maintained by heterozygote advantage and that balancing selection may have played an important role in human evolution in this context and in general.

KEYWORDS

balancing selection, frequency-dependent selection, heterozygote advantage, human evolution, polymorphism, viability

1 | INTRODUCTION

The origin, maintenance and fate of genetic diversity remain one of the central questions of evolutionary biology. Putting aside the dispute between proponents of the neutral theory of evolution (Kimura, 1968) and advocates of the selectionist, 'balanced' approach (Dobzhansky, 1955; Lewontin, 1974; Lewontin & Hubby, 1966), it is now clear that both random drift and natural selection play an important role in maintaining genetic diversity (Andrés et al., 2010; Croze et al., 2016; Key et al., 2014). Among all forms of selection, *balancing selection* is a factor that contributes to the maintenance of genetic variability most directly (Croze et al., 2016; Hedrick, 2012; Key et al., 2014). One of its components is frequency-dependent selection, a mechanism by which the fitness value of alleles changes with their frequency in the gene pool. In cases where fitness negatively correlates with the frequency of the focal allele, the whole system may stabilize and maintain genetic polymorphism for a long time (e.g. Raffini et al., 2017; Raffini & Meyer, 2019).

Selection that varies over time and space, between the sexes, life stages or fitness components may be balancing in the long term (balancing selection sensu lato: Hedrick, 2012). Nonetheless, heterozygote advantage—also known as overdominance or heterosis—is usually recognized as the most immediate mechanism of balancing selection. Under this regime, heterozygotes for a particular gene locus exhibit higher fitness than either of the homozygotes. This results in an equilibrium such that both alleles stabilize at a frequency that maximizes the mean fitness of the population. Such systems are usually recognizable by a relatively higher proportion of few minority alleles with intermediate frequencies (e.g. Allison, 1954).

Despite expectations of some classic work of the modern evolutionary synthesis (Dobzhansky, 1955; Lewontin, 1974; Lewontin & Hubby, 1966), balancing selection was until recently considered a rare phenomenon (e.g. Croze et al., 2016; Hedrick, 2012). Numerous new studies, however, show that it is an important way of maintaining genetic polymorphism, both in general and in human populations (Andrés et al., 2010; Croze et al., 2016; Hedrick, 2012; Key et al., 2014). Examples of systems maintained by balancing selection include warfarin drug resistance in rats (*Rattus norvegicus*) (Greaves et al., 1977), self-incompatibility systems of flowering plants (Magnoliophyta) (Roux et al., 2013), the ABO blood system of primates (Segurel et al., 2012) or the major histocompatibility complex (MHC) of vertebrates (Andres et al., 2009; Spurgin & Richardson, 2010). A majority of known targets of balancing selection are related to immunity (Andres et al., 2009; Ferrer-Admetlla et al., 2008; Fumagalli et al., 2009; Hellgren & Sheldon, 2011; Single et al., 2007). These concern mainly genes associated with antigen recognition (Croze et al., 2016; Key et al., 2014), but also genes with much looser relation to immunity, such as olfactory receptors (Alonso et al., 2008; Croze et al., 2016), extracellular cell components (Andres et al., 2009) or pattern recognition receptors (Těšický & Vinkler, 2015). This is not surprising because it has long been suspected that the arms race between parasites and their hosts produces strong negative frequency-dependent selective pressures (Haldane, 1949).

In humans, many of these polymorphisms are associated with congenital diseases. Examples include the recessive 'sickle-cell disease' allele of the HBB gene conferring resistance against malaria (Allison, 1954), or recessive 'cystic fibrosis' alleles of the *CFTR* gene that improve resistance to certain infections (Cuthbert et al., 1995). Andres et al. (2009) identified at least 60 human genes that show signs of ancient balancing selection, of which five are clearly related to immunity, and seven are associated with various other diseases. Blood groups are another prominent set of human traits affected by balancing selection. It has been proposed that particular ABO genotypes confer differential resistance to parasites and diseases (Segurel et al., 2012). Fumagalli et al. (2009) demonstrated that various types of balancing selection are also responsible for the maintenance of polymorphisms in the blood group systems CD55, CD151, SLC14A1, and possibly also BSG and FUT2. This is probably for immunological reasons, as various blood system minority alleles correlate with resistance to pathogens. Moreover, Fumagalli et al. (2009) showed that pathogen richness correlates with minority allele frequencies in several blood systems. Other potential advantages of rarer alleles in these systems include faster healing or augmented water retention in dry climates, usually at the cost of increased risk of autoimmune disorders.

Another prominent blood group polymorphism traditionally suspected of being maintained by balancing selection is the polymorphism in Rhesus factor, that is the *RHD* gene. The immunodominant antigen on the surface of human erythrocytes is the D epitope on the RhD protein, a product of the *RHD* gene, which is part of the transmembrane protein–glycoprotein complex (Rh complex) of human erythrocytes. The main role of the Rh complex is to transport the NH_3 or CO_2 gases and their ions (Kustu & Inwood, 2006; Nakhoul & Hamm, 2013), although the physiological function of this transport remains unknown. Moreover, in mammal species with biconcave discoid erythrocytes, such as humans, this complex also plays an important role in maintaining the integrity of plasma membrane and cell shape (Le Van Kim et al., 2006; Nakhoul & Hamm, 2013). The human *RHD* gene is highly polymorphic: over 200 alleles have been described in various populations around the world (Flegel, 2011).

A large proportion of people of European origin carry the recessive d allele of the *RHD* gene (its frequency is about 40%), which represents a large deletion covering almost the entire *RHD* gene (Wagner & Flegel, 2000). The spread of the d allele and stable coexistence of RhD-positive and RhD-negative individuals in European populations are well-known evolutionary enigmas (Fisher et al., 1944; Haldane, 1942). Carriers of the rarer d phenotype suffer from impaired fecundity. Before the advent of modern medicine, most RhD-positive children born to RhD-negative mothers who had been immunized by a previous birth of RhD-positive child suffered from severe haemolytic disease as newborns. Despite various efforts to explain RhD polymorphism by a combination of demographic factors (Feldman et al., 1969; Nei et al., 1981), it is clear that without the intervention of some form of balancing selection, the coexistence of d and D alleles would be in the

long run evolutionarily unstable: the less fit allele would be selected against, and all individuals in a population would eventually become either RhD negative or RhD positive.

The hypothesis that selection favouring heterozygotes (heterosis) plays a key role in maintaining *RHD* polymorphism was probably first explicitly formulated in late 1960s by Feldman et al. (1969), though supporting empirical evidence started to accumulate only about a decade ago. Several papers showed that subjects with the RhD-positive phenotype—and especially heterozygotes with one D and one d allele—are protected against certain adverse factors. Most prominent is protection against the health impact of toxoplasmosis, a common chronic infection by a protist parasite, but various studies also found indications of protection against the effects of fatigue, smoking and ageing (Flegr et al., 2010, 2012, 2018; Holub et al., 2011; Kaňková et al., 2010; Novotná et al., 2008). More recently, an ecological correlative study performed in 66 countries for which allele frequencies and WHO data on the incidence of 125 disorders were available (WHO, 2008) showed that the incidence of many disorders correlates positively with the frequency of RhD-negative subjects and negatively with the frequency of RhD-positive heterozygotes in the population (Flegr, 2016). A large-scale cross-sectional study had further shown that RhD-negative subjects exhibit many signs of impaired health and increased incidence of many diseases and disorders (Flegr et al., 2015): RhD-negative subjects reported more frequent allergic, digestive, heart, haematological, immunological, mental health-related and neurological problems. Another recent study demonstrated that RhD-negative subjects scored worse than RhD-positive subjects on well-being, mental and physical health (Flegr et al., 2020). Over the past 12 years, at least 14 published studies demonstrated effects of RhD phenotype on human psychomotor performance and health, whereas to the best of our knowledge only two studies showed no such effect (Flegr & Dama, 2014; Halmin et al., 2017). Most studies, however, compared the health or psychomotor performance of individuals with RhD-negative and RhD-positive *phenotype*, and not RhD *genotype*, because RhD-positive homozygotes and heterozygotes cannot be easily distinguished by a routine immunoagglutination test.

In our present study, we took advantage of the fact that an RhD heterozygote can be easily identified based on the subject's RhD-positive phenotype plus one RhD-negative biological parent (Kaňková et al., 2020). In relatively rich, low fecundity countries like the Czech Republic the vast majority of people know their biological parents' and their own RhD phenotype (Voracek et al., 2008). We can therefore easily establish the RhD genotype of not only the RhD-negative subjects (dd) but also of RhD-positive heterozygotes (Dd). Naturally, people who have two RhD-positive parents could be either DD or Dd (with higher probability of DD than in the general population). This makes this method less precise than molecular genotyping, but it enables researchers to obtain large amounts of data relatively easily and cheaply using, for instance, clinical records or internet questionnaires.

We here present the results of such a study performed with 2,539 participants. Our study was preregistered (see Section 2) and based on questionnaires about the RhD phenotype of respondents and their parents, their physical health, mental health, sexual desire and fecundity. The main aim was to test two hypotheses: first, whether RhD heterozygotes enjoy better health than both RhD-positive and RhD-negative homozygotes; and second, that subjects with RhD-negative phenotype enjoy worse health than subjects with RhD-positive phenotype. We thus directly test the role of balancing selection in maintaining human RHD gene polymorphism and ultimately contribute to a better understanding of the role of balancing selection in biological evolution.

2 | MATERIALS AND METHODS

2.1 | Subjects

Most subjects were members of the 'Labbunnies' group, a community of currently over 20,000 Czechs and Slovaks who voluntarily participate in various experiments related to evolutionary psychology. Using several posts on the timeline of the Labbunnies Facebook group (administered by our research team), we invited group members and their Facebook friends to participate in an anonymous study that would 'investigate certain philosophical problems of personal identity, test evolutionary hypotheses regarding, for example the beauty of flowers, and study moral dilemmas related to autonomous cars'. Respondents received no remuneration for their participation in the study, but at the end of our 70-min questionnaire, they found a reward in the form of their own results for three tests that were part of the online questionnaire as well as the results of other similar studies.

At the beginning of the questionnaire, we asked participants for their informed consent. We informed them that 'The study is anonymous and obtained data will be used exclusively for scientific purposes. Your cooperation in the project is voluntary and you can terminate it at any time by closing this website. You can also skip any questions you may find uncomfortable, but complete data are the most valuable. If you agree to participate in this research, press the "Next" button'. Only subjects who gave us their informed consent by pressing the button were allowed to participate in the study. Between 9 May and 31 December 2017, we collected data from 4,650 respondents. Not all, however, answered questions pertaining to their and their parents' RhD phenotypes. The project, including the method used to obtain electronic informed consent with participation in this anonymous study, was approved by the Institutional Review Board of the Faculty of Science, Charles University (Komise pro práci s lidmi a lidským materiálem Přírodovědecké Fakulty University Karlovy)—No. 2017/08. The study was preregistered at Open Science Framework (OSF) prior to the start of data collection (https://osf.io/ykxpf).

2.2 | The questionnaire

In the anamnestic part of the questionnaire, which was run on the Qualtrics platform (https://www.qualtrics.com/), respondents

provided information about their age, body height and weight, and the size of the communities they currently reside in (ordinal variable *urbanization*—0: less than 1,000 inhabitants, 1: 1,000–5,000 inhabitants, 2: 5,000–50,000 inhabitants, 3: 50,000–100,000 inhabitants, 4: 100,000–500,000 inhabitants, 5: over 500,000 inhabitants, 6: Bratislava, 7: Praha), and the number of their children. Respondents were also asked about their ABO blood group (not used in the present study) and, in the following question, about their Rh factor. Possible answers were 'I do not know, I am not sure', 'negative (the rarer variant)' or 'positive (the more frequent variant)'. The answer 'I do not know, I am not sure' was marked in advance. In the following two questions, we used the same range of options to ask respondents about the RhD phenotype of their biological mother and father. Based on the RhD phenotype of respondents and their parents, we were able to identify RhD-positive heterozygotes (i.e. RhD-positive subjects with one RhD-negative parent). The remaining RhD-positive subjects, that is subjects with RhD-positive mother and father and those who did not answer questions on the RhD phenotype of their parents, represented a mixed group of homozygotes and heterozygotes, enriched by RhD-positive homozygotes in comparison with the general RhD-positive population.

In order to compare the relative importance of the influence of RhD phenotype on health, well-being and biological fitness with other important factors, we looked for associations between health, fecundity and sexual desire using three unrelated but well-known risk factors: *body mass index* (BMI) calculated from body height and body weight, frequency of *smoking*, that is how many cigarettes they smoke a day (ordinal scale 0: 0, 1: 0–0.1, 2: 0.1–1, 3: 1.1–3, 4: 3.1–10, 5: 11–20, 6: 21–40, 7: over 40), and frequency of *alcohol consumption* in a volume that would make it illegal for the person to drive a car (ordinal scale: 0: never, 1: at most once a month, 2: at most twice a month, 3: at most four times a month, 4: at most twice a week, 5: every other day, 6: every day, 7: nearly all the time). In contrast to our previous studies, we did not ask respondents about their consumption of illegal drugs or drugs/vitamins/supplements not prescribed by a medical professional.

Using scales of 0–100, participants rated how strongly they are sexually aroused by persons of the same and of the opposite sex. The higher of these two arousal-related variables was then designated as our variable *sexual desire*. The main outcome variables used here, physical health and mental health, were obtained in another section of the questionnaire where the respondents rated how they have been feeling physically and (separately) mentally during the past two years using six-point ordinal scales ranging from 1 (very bad) to 6 (very well). Next, respondents listed how many kinds of drugs prescribed by a medical professional for mental health problems they used over the past month and how many kinds of prescription drugs for other health problems they used over the same period of time. Respondents also expressed, using scales of 0 (definitively not) to 100 (definitively yes), whether and to what extent they suffer from anxieties, depressions, manias, obsessive/compulsive behaviour, auditory hallucinations, visual hallucinations, burnout syndrome and headaches. For example, they were asked 'Do you suffer from obsessive (compulsive) behaviour?' In the heading of this section of the questionnaire, participants were reminded that they ought to indicate what issues ('feelings, moods, or behaviours') they are struggling with right now, that is not problems they were diagnosed with in the past. They were also requested to try and estimate how much they would struggle with a particular problem if they stopped using medication prescribed for the issue. Using a list of 25 mental health disorders and epilepsy, they also noted which they suffer from (i.e. which 'had been diagnosed by medical professionals'). Based on their responses, we computed the variable *number of mental health disorders*. For English translation of relevant parts of the questionnaire, see Appendix S1.

2.3 | Statistics

Before analyses, we filtered out <3% of questionnaires submitted by subjects who failed our two attention tests ('Now kindly check the answer number two (four).') or provided a suspect combination of answers to other questions, for instance excessively high/low height, weight, age, an unrealistically high number of neuropsychiatric disorders, or answered all or nearly all questions by the same code, etc. In accordance with preregistration, the data of all five subjects who indicated that they had been diagnosed with either Parkinson's or Alzheimer's disease were also removed. Descriptive statistics (preparation of graphs and frequency tables), computation of simple univariate tests (two-sided, unpaired *t* tests, χ^2 test) and ordinal regressions with sex, RhD phenotype/genotype and their interaction as explanatory factors plus age and size of the place of living as covariates were performed using Statistica v. 8. Partial Kendall correlation tests were performed in R v. 3.3.1 (R Core Team, 2018), package ppcor v. 1.1 (Kim, 2015).

In the confirmatory part of the study, we controlled for age and/or sex (when analysing a mixed set of men and women). In the exploratory part, we also controlled for age, tobacco smoking and alcohol consumption. In accordance with preregistration, two preregistered one-sided hypotheses were tested with one-sided tests and all other hypotheses with two-sided tests. The level of significance was always set to 0.05. Correction for multiple tests was done using the Benjamini–Hochberg procedure with false discovery rate preset to 0.25 (Benjamini & Hochberg, 1995), using the Excel sheet published by McDonald (2014). All relevant data are available at figshare.com/s/d199dec0ba0fc5efde29.

2.4 | Differences between the preregistered and implemented protocol

We hoped to collect data from about 10,000 subjects before the preregistered end of the study (31 December 2017). It turned out, however, that the main topics of our questionnaire—those pertaining to philosophical problems of personal identity, beauty of flowers and autonomous cars—were apparently less attractive than the topics of

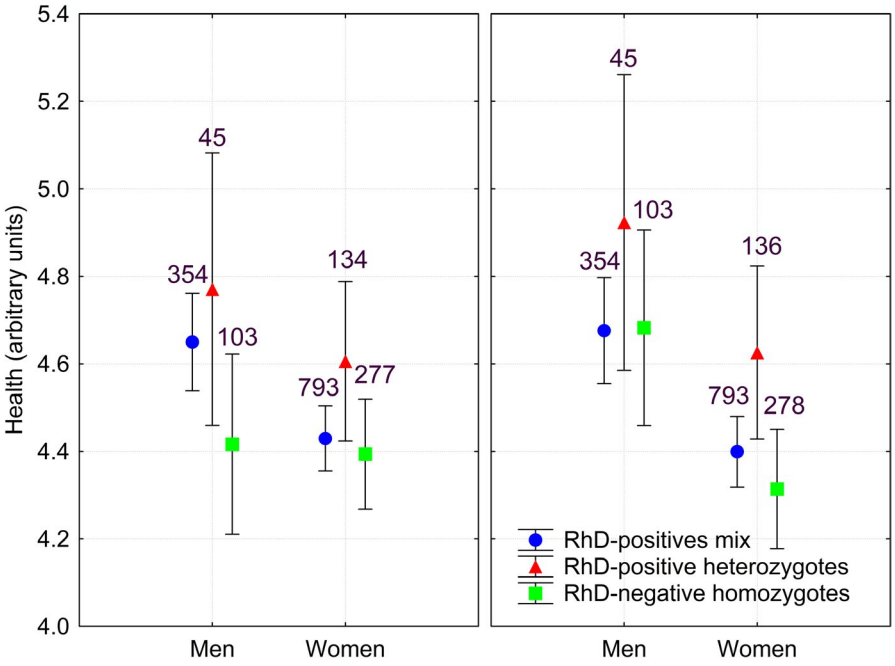


FIGURE 1 Differences in overall physical (left) and mental (right) health among RhD-negative homozygotes, RhD-positive heterozygotes, and the remaining mixed group of RhD-positive homozygotes and heterozygotes. The RhD-positive mix contains all RhD-positive homozygotes and part of RhD-positive heterozygotes (see Sections 2 and 4). High values on the y axis denote better health. Vertical bars denote 0.95 confidence intervals and numbers above the bars show the number of subjects in particular groups

our previous questionnaires. In the end, we thus obtained data from less than half of the originally planned respondent number, so the number of subjects tested for association between RhD-negative phenotype and health was about 50% lower than anticipated in the preregistration. On the other hand, we had anticipated that only about 15% of respondents would provide the RhD phenotype data of their parents, whereas that percentage turned out to be much higher. In the end, the number of subjects tested for the effect of RhD heterozygosity on health was thus slightly higher than expected in the preregistered protocol.

In the preregistration, the output variable was called *health problems*, but the questionnaire included no such variable. To make the results of the present study comparable to those of previous studies, we thus replaced it by two other outcome variables from the questionnaire, namely physical health and mental health (both ranging 1 to 6), performed a correction for multiple (two) tests, and reported the corresponding results. We additionally performed equivalent multivariate analyses using the various health parameters described above as outcome variables.

3 | RESULTS

3.1 | Data description

The sample (n = 2,539) consisted of 756 men (173, i.e. 22.88% of whom were RhD negative) of mean age 37.4 years (*SD* 12.8) and 1,783 women (405, i.e. 22.71% of whom were RhD negative) of mean age 35.5 years (*SD* 12.6). The men were thus significantly older than the women (t = 4.07, p < .00001, Cohen's d = 0.153), although the frequency of RhD-negative subjects did not differ between the sexes (χ^2 = 0.009, p = .929). Of the men, 314 provided information about the RhD phenotype of their mother and 259 about the RhD phenotype of their father. According to these reports, 23.6% of mothers and 23.17% of fathers were RhD negative. Of the women, 1,087 provided information about the RhD phenotype of their mother and 870 about the RhD phenotype of their father; 26.0% of mothers and 17.9% of fathers were RhD negative. Interaction between the RhD and sex was not significant for the mothers (χ^2 = 0.782, p = .377) or the fathers (χ^2 = 3.540, p = .060).

3.2 | Confirmatory analyses

We tested two predictions derived from two preregistered hypotheses (Figure 1):

1. RhD-positive heterozygotes, that is RhD-positive subjects with one RhD-negative parent, will report fewer health problems than other study participants, that is, fewer problems than RhD-negative subjects and subjects who report having two RhD-positive parents (and can be therefore either heterozygotes or homozygotes).
2. RhD-negative subjects (homozygotes) will report more health problems than RhD-positive study participants.

Ordinal regressions with either physical or mental health (scale 1–6) as the dependent (outcome) variable, sex, RhD phenotype (positive or negative) and interaction between sex and RhD phenotype as explanatory factors, plus age and size of place of residence as covariates, showed that RhD-negative subjects were in worse physical health (Wald statistic = 6.33, p = .006, one-sided test; 0.012 after the Bonferroni correction for two tests) but not worse mental health (Wald statistic = 1.42, p = .117, one-sided test) than RhD-positive subjects (subsuming heterozygotes and homozygotes). Post hoc tests showed that in men the effect of RhD-negative phenotype on

physical health was very strong and on mental health very weak, whereas in women the effect on physical health was slightly weaker than on mental health (see Figure S1 and Table S1).

The same ordinal regression analyses with either physical or mental health as dependent (outcome) variable, sex, RhD heterozygosity (homozygote (+/+) or (−−) versus heterozygote (+/−)) and interaction between sex and RhD heterozygosity as explanatory factors, plus age and size of place of residence as covariates, showed that heterozygotes enjoy nonsignificantly better physical health (Wald statistic = 2.34, $p = .063$, one-sided test) and significantly better mental health (Wald statistic = 6.14, $p = .007$, one-sided test; $p = .014$ after the Bonferroni correction for two tests) than the remaining (homozygotic) subjects (see Figure S2 and Table S2).

We also detected a very strong negative effect of age (Table S2) and a significant negative effect of female sex on health in all of the above analyses, except for the analysis of the effect of RhD phenotype on physical health. We detected no effect of the size of the place of residence (all p-values > .60), and never any interactive effects between sex and RhD phenotype or sex and RhD heterozygosity on physical or mental health (all p-values > .12). Our ordinal regression models are shown in Tables S1 and S2, and the overall similar results of equivalent multivariate analyses using all 14 physical and mental health-related variables as dependents are given in Tables S3–S6.

3.3 | Exploratory analyses

We then investigated the physical and mental health of subjects with different RhD genotypes in more detail. Figure 1 shows differences between the physical and mental health of RhD-negative homozygotes, RhD-positive heterozygotes, and the remaining mixed group of RhD-positive homozygotes and heterozygotes (for determination of these three groups, see Materials and methods and Discussion). Repeated-measures ANCOVA with both physical health and mental health as dependent (outcome) variables, sex, RhD genotype and the sex by RhD genotype interaction as independent factors, plus age and size of place of residence as covariates, showed a significant effect of the RhD genotype ($F_{2,1697} = 3.851$, $p = .021$, $\eta^2 = 0.005$). Separate analyses showed a significant effect of RhD genotype on both physical health ($F_{2,1698} = 3.462$, $p = .032$, $\eta^2 = 0.004$) and mental health ($F_{2,1697} = 3.105$, $p = .045$, $\eta^2 = 0.004$; see Figure 1 for numbers of subjects in particular groups). No interactive effects between the genotype and other variables were observed in these three analyses.

For post hoc exploratory analyses of the effects of RhD genotype on particular health-related variables, the number of respondents' biological children, and level of sexual desire, we used nonparametric partial Kendall correlation tests with age as a covariate (dropping the effect of size of place of residence that was always nonsignificant in all previous analyses). To compare the relative strength of effects of RhD genotype or phenotype with the effects of known adverse factors, we used alcohol consumption, tobacco smoking and BMI as internal controls. Table 1 shows that RhD genotype had significant effects on several health-related variables, whereby the strength of these effects was comparable to other factors under investigation. For instance, the negative effect of RhD-negative phenotype was stronger than that of alcohol consumption, but weaker and less frequent than that of tobacco smoking, and of comparable strength to the effect of high BMI (with the exception of BMI's very strong negative effect on physical health). We detected no effect of RhD genotype on fecundity (number of biological children) or the level of sexual desire.

RhD-heterozygous men were on average 1 cm shorter (lower body height) ($F_{2,751} = 3.19$, $p = .042$, $\eta^2 = 0.008$) and 5 kg heavier (higher body mass) ($F_{2,750} = 2.12$, $p = .121$, $\eta^2 = 0.006$), and consequently had higher BMI ($F_{2,750} = 5.56$, $p = .004$, $\eta^2 = 0.015$) than men with other RhD genotypes, in particular RhD-negative homozygotes; no such differences were observed in women (Figure 2). Differences between RhD-positive heterozygotes and homozygotes might be actually even larger, as differences between heterozygotes and the mixed group of RhD-positive homozygotes and heterozygotes were of a similar magnitude as those between heterozygotes and RhD-negative homozygotes (see Figure 2). ANCOVA with genotype and sex as fixed factors and age as covariate showed that both genotype and the genotype–sex interaction had significant effects on the BMI (genotype: $F_{2,2,517} = 4.93$, $p = .007$, $\eta^2 = 0.004$; genotype–sex: $F_{2,2,517} = 4.72$, $p = .017$, $\eta^2 = 0.003$). We repeated all partial Kendall analyses while controlling for two additional potential confounding variables, namely smoking and alcohol consumption, but the results remained approximately the same (see Table S7).

4 | DISCUSSION

In accordance with our two preregistered hypotheses, RhD heterozygotic individuals reported better health than other, mostly homozygotic study participants, and RhD-negative subjects suffered from worse physical and mental health than RhD-positive participants (Figure 1, Figures S1 and S2). Unexpectedly, we also found the mental health (of primarily females) and physical health (of males) of RhD-positive homozygotes to be as bad or even worse than that of RhD-negative homozygotes. This seems to contradict the results of previous studies, as well as the results of our partial Kendall tests (see column 7 of Table 1). However, most previous studies compared the health of RhD-negative subjects with the health of both homozygous and heterozygous RhD-positive subjects (i.e. subjects with RhD-positive phenotype), which could effectively mask effects by averaging potential disadvantages of homozygotes with health advantages of heterozygotes. Merely three of the many recent studies of this topic specifically investigated the main prediction of the balancing selection hypothesis in trying to identify possible advantages of RhD heterozygotes: the correlational study of Flegr (2016), and a study of blood donors with completely known RhD genotypes by Novotná et al. (2008). Most recently, additional indirect evidence was offered by Kaňková et al. (2020), who showed that RhD heterozygous mothers, that is RhD-positive women who gave birth to at least one RhD-negative child, tend to deliver more sons than daughters. According to

TABLE 1 Correlation between RhD phenotype and genotype and health-related variables

	Age	Place	p−/p+	+−/p−	+−/+m	+m/p−	Alcohol	Smoking	BMI
All									
Physical health	**−0.05**	0.01	**−0.05**	**0.10**	**0.04**	**0.04**	**0.04**	**−0.06**	**−0.13**
Mental health	**0.10**	0.00	**−0.04**	**0.11**	**0.06**	**0.03**	0.03	**−0.10**	−0.01
Drugs mental health	**0.09**	0.00	−0.02	0.04	0.02	0.01	**−0.05**	**0.10**	**0.04**
Drugs other	**0.12**	0.00	−0.03	−0.01	**−0.03**	**0.03**	**−0.06**	**−0.04**	**0.10**
Number of disorders	−0.02	**0.05**	**0.05**	−0.03	0.02	**−0.05**	**−0.06**	**0.12**	0.01
Anxiety	**−0.11**	0.02	**0.04**	**−0.06**	−0.02	**−0.04**	0.00	**0.05**	−0.02
Phobias	**−0.15**	0.00	**0.05**	−0.03	0.02	**−0.06**	**−0.03**	0.02	0.01
Depression	**−0.11**	0.01	**0.04**	**−0.06**	−0.02	**−0.04**	0.00	**0.09**	0.00
Mania	**−0.12**	−0.02	**0.04**	−0.01	0.02	**−0.04**	**0.04**	**0.07**	0.02
Compulsive behaviour	**−0.19**	0.00	0.02	−0.03	0.00	−0.03	0.02	**0.04**	0.00
Hallucination auditory	**−0.04**	−0.01	**0.03**	−0.03	0.01	**−0.04**	0.00	0.02	**0.03**
Hallucination visual	−0.01	−0.01	**0.03**	−0.01	0.02	**−0.04**	0.00	**0.03**	**0.04**
Burnout	−0.01	0.00	0.02	**−0.06**	−0.03	−0.02	0.00	**0.06**	**0.03**
Headache	**−0.10**	0.00	0.02	0.01	0.02	−0.02	−0.02	0.01	0.01
Number of children	**0.52**	**−0.08**	−0.01	−0.01	−0.02	0.01	**−0.03**	**−0.06**	**0.03**
Sexual desire	**0.03**	**−0.04**	−0.01	0.00	−0.01	0.02	0.00	−0.02	0.00
BMI	**0.24**	**−0.05**	−0.01	0.03	0.02	0.01	−0.01	**0.04**	NA
Height	0.00	0.03	0.00	−0.03	−0.03	0.01	**0.09**	**0.05**	**0.08**
Weight	**0.19**	**−0.03**	−0.01	0.01	0.00	0.01	0.03	**0.05**	**0.69**
Men									
Physical health	**−0.11**	0.03	**−0.10**	**0.15**	0.02	**0.10**	0.00	**−0.10**	**−0.13**
Mental health	**0.09**	0.01	−0.02	**0.10**	**0.06**	0.01	−0.01	**−0.15**	−0.02
Drugs mental health	**0.06**	**0.08**	0.01	0.05	0.05	−0.02	−0.05	**0.18**	**0.09**
Drugs other	**0.22**	0.04	−0.01	−0.03	−0.03	0.01	**−0.06**	**−0.07**	**0.12**
Number of disorders	**−0.11**	**0.08**	**0.07**	0.04	**0.10**	**−0.09**	**−0.11**	**0.15**	**0.07**
Anxiety	**−0.11**	0.01	0.03	−0.06	−0.02	−0.03	0.01	**0.11**	0.02
Phobias	**−0.18**	0.03	**0.07**	−0.03	0.05	**−0.08**	−0.02	**0.12**	0.04
Depression	**−0.11**	−0.01	**0.06**	−0.03	0.03	**−0.07**	0.01	**0.18**	0.01
Mania	**−0.16**	−0.02	0.01	0.08	**0.06**	−0.02	0.04	**0.07**	0.00
Compulsive behaviour	**−0.21**	0.03	0.00	0.06	0.05	−0.01	0.00	**0.08**	0.00
Hallucination auditory	−0.04	−0.01	0.00	0.01	0.01	−0.01	0.00	0.05	0.00
Hallucination visual	0.01	−0.01	0.02	0.05	0.05	−0.03	0.01	0.04	0.00
Burnout	−0.04	0.01	0.02	0.02	0.03	−0.03	−0.01	**0.06**	0.02
Headache	**−0.09**	−0.01	0.02	**0.14**	**0.12**	−0.04	−0.01	0.01	−0.02
Number of children	**0.51**	**−0.05**	−0.04	0.06	0.01	0.04	0.04	−0.05	**0.09**
Sexual desire	**0.07**	**−0.07**	0.03	−0.01	0.02	−0.03	−0.01	−0.04	**0.07**
BMI	**0.27**	−0.01	−0.01	**0.11**	**0.08**	0.00	**−0.08**	0.05	NA
Height	**−0.05**	0.04	**−0.05**	−0.04	**−0.07**	**0.06**	0.04	0.02	−0.02
Weight	**0.23**	0.01	−0.02	**0.12**	**0.06**	0.02	−0.05	**0.06**	**0.71**
Women									
Physical health	**−0.04**	0.00	−0.02	**0.08**	**0.05**	0.02	**0.05**	**−0.05**	**−0.15**
Mental health	**0.10**	0.00	**−0.04**	**0.11**	**0.06**	0.03	0.03	**−0.08**	**−0.04**
Drugs mental health	**0.12**	−0.02	−0.03	0.03	0.00	0.03	**−0.04**	**0.09**	**0.05**
Drugs other	**0.09**	−0.01	**−0.03**	0.00	−0.04	**0.04**	**−0.04**	−0.02	**0.10**

(Continues)

TABLE 1 (Continued)

	Age	Place	p−/p+	+−/p−	+−/+m	+m/p−	Alcohol	Smoking	BMI
Number of disorders	0.02	**0.03**	**0.03**	−0.05	−0.02	−0.03	−0.03	**0.12**	0.01
Anxiety	**−0.09**	0.02	**0.03**	**−0.06**	−0.03	−0.03	0.02	**0.04**	−0.01
Phobias	**−0.13**	−0.01	**0.03**	−0.03	0.00	**−0.04**	−0.01	−0.01	0.03
Depression	**−0.11**	0.01	0.03	**−0.07**	**−0.04**	−0.03	0.00	**0.05**	0.01
Mania	**−0.10**	−0.01	**0.05**	−0.04	0.01	**−0.05**	**0.04**	**0.06**	0.03
Compulsive behaviour	**−0.18**	0.00	**0.03**	−0.05	−0.02	−0.03	**0.03**	0.02	0.00
Hallucination auditory	**−0.04**	−0.01	**0.05**	−0.04	0.01	**−0.05**	0.01	0.01	**0.05**
Hallucination visual	−0.01	−0.02	**0.04**	−0.03	0.01	**−0.04**	0.00	**0.03**	**0.07**
Burnout	0.00	0.00	0.02	**−0.08**	**−0.05**	−0.01	0.01	**0.07**	**0.04**
Headache	**−0.09**	0.00	0.01	−0.03	−0.02	−0.01	0.00	0.02	**0.05**
Number of children	**0.53**	**−0.09**	0.01	−0.04	−0.03	0.00	**−0.06**	**−0.07**	0.02
Sexual desire	0.01	**−0.03**	−0.03	0.01	−0.01	0.03	0.00	−0.02	**−0.04**
BMI	**0.22**	**−0.07**	−0.01	0.03	0.02	0.01	−0.01	0.03	NA
Height	**−0.04**	**0.03**	0.00	0.04	0.03	0.00	**0.04**	0.04	−0.01
Weight	**0.18**	**−0.05**	−0.02	0.05	0.02	0.01	0.00	**0.04**	**0.74**

Note: The table shows the direction and strength of particular effects measured with a nonparametric partial Kendall Tau correlation test, controlling for age, place of living and sex (first third of the table). The column headings denote subjects with RhD-negative (p−) or RhD-positive phenotype (p+), RhD-positive heterozygotes (+−) or a mixed RhD-positive subset consisting of RhD-positive homozygotes plus -heterozygotes (+m) who reported no RhD-negative parent. Positive Tau indicates that the genotype/phenotype to the left of the slash has a higher value of the variable listed in the first column than the genotype/phenotype to the right of the slash. In accordance with preregistration, one-sided variants of tests were applied. Significant Taus are printed in bold. No correction for multiple tests was performed in this exploratory part of the study, although the majority of significant associations remained significant after such correction with Benjamini–Hochberg procedure when two-sided tests were performed. Still, some unexpected associations, for example, the higher physical health index of women consuming more alcoholic drinks, may indicate the presence of artefacts of multiple tests.

the Trivers–Willard hypothesis (Trivers & Willard, 1973), females in better physical (or psychological) condition should bear more sons. This is because in polygynous or serially monogamous species where the biological fitness of males depends on their status more strongly than it does in females (e.g. in humans), females in good condition should invest preferentially in their sons to increase their reproductive success. The more male-biased sex ratio of heterozygotic women was thus interpreted as indirect evidence of their better health.

Admittedly, our method also could not identify all heterozygotes, leading to the subset of RhD-positive homozygotes being diluted by a substantial fraction of RhD-positive heterozygotes (i.e. children of heterozygotic parents plus subjects who did not answer questions about their parents' RhD phenotype). In case Hardy–Weinberg equilibrium applies to our sample, these 'hidden heterozygotes' would comprise up to 60% of this mixed group (but see possible issues with reporting RhD phenotype and randomness of our sample discussed below). More specifically, the group of RhD-positive women originally consisted of about 35% homozygotes and 65% heterozygotes, which changed to about 41% share of homozygotes after the exclusion of self-identified heterozygotes; the proportion of male RhD-positive homozygotes thus changed from about 38% to 43%. As heterozygotes clearly had the highest health scores of all participants (filled triangles in Figure 1), they must have increased the mean health indices of this mixed group (open circles in Figure 1). Given that the subgroup of male RhD-positive homozygotes in Figure 1 enjoyed approximately the same mental health as the group of RhD-negative homozygotes despite being contaminated by an unknown fraction of heterozygotes, the actual group consisting solely of RhD-positive homozygotes therefore actually could be in worse health than the RhD-negative subjects. In this context, at least one earlier study found that RhD-negative subjects might actually be more resistant to certain viral infections than the rest of the population (Flegr et al., 2015). Nevertheless, the observation that (the dominant) RhD-positive homozygotes might suffer from equal or even greater health impairment than RhD-negative recessive homozygotes provides new and strong support for the hypothesis that the RhD polymorphism is maintained by selection in favour of heterozygotes in human populations.

From all proposed mechanisms of balancing selection, heterozygote advantage turns out to be one of the rarer ones at the intraspecific level (Hedrick, 2012). It is therefore especially interesting that the *RHD* gene is also present in several nonhuman primates (Blancher & Socha, 1997). Although the specific deletion (see Section 1) of the RhD-recessive allele has probably emerged after human speciation (Westhoff & Wylie, 1996), the Rh blood system is polymorphic also in nonhuman primates, supporting the possibility of heterosis being more a frequent phenomenon than previously thought, at least with respect to polymorphisms of blood or other immunity-related systems.

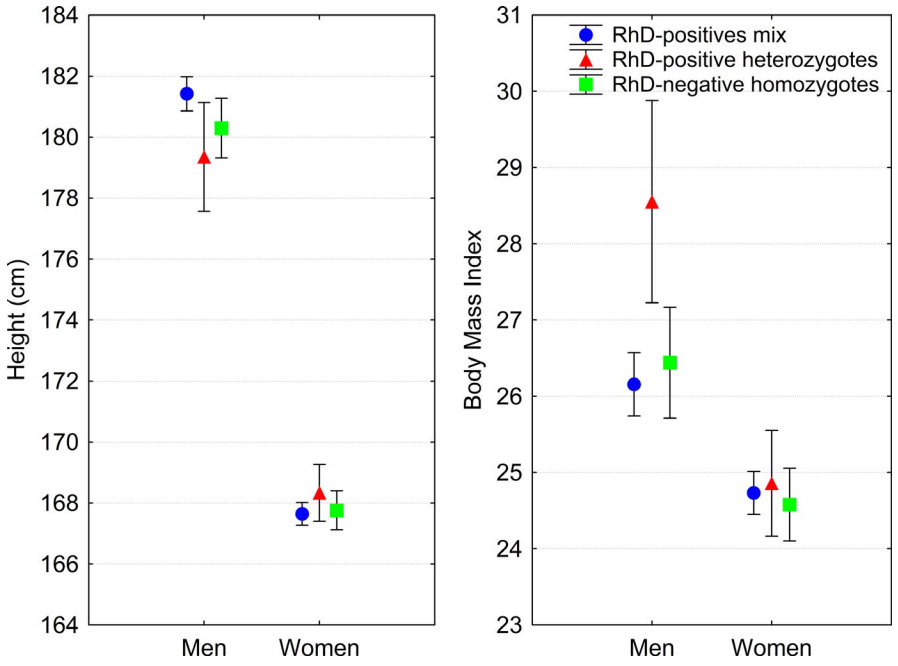


FIGURE 2 The effect of RhD genotype on height and body mass index (BMI)/for legend, see Figure 1.

One of the most important steps in search of a better understanding of balancing selection and the maintenance of genetic polymorphism in general is to link particular alleles of specific loci with identifiable phenotypic effects (Key et al., 2014). Aside from the generally known and expected decline in overall health with age (Table 1), we here managed to identify several associations of RhD genotype with particular physical or mental health-related variables (Table 1). Especially in women, RhD-negative subjects were generally less healthy and RhD-positive heterozygotes healthier, whereas in men we observed some notable exceptions. RhD-positive male heterozygotes reported a higher number of mental health disorders, more frequent or intensive mania, and more frequent or intense headaches than other male subjects. This suggests that, in males, heterozygote disadvantage might exist in a limited number of health-related traits. These three health-related traits should therefore possibly be excluded in overall health indices in future studies. Our analyses also showed that RhD-positive male heterozygotes had average BMIs over two points higher than men with homozygotic genotypes (Figure 2). It is thus possible that higher incidence of obesity in this group could be responsible for some of their health problems.

Most published studies on the maintenance of genetic polymorphism by balancing selection in humans (see Section 1) focus solely or primarily on variables related to physical health, such as the prevalence of certain diseases, resistance to infection or physiological deviations from the norm. Exceptions from this pattern are rare and only tangentially related to the study of genetic polymorphism (Lalovic et al., 2004; Prandota, 2011, 2012; Rossignol, 2007; Rossignol et al., 2007; Santangelo et al., 2018). The same holds for studies focused on blood group polymorphisms (Fumagalli et al., 2009). This is understandable, as the majority of studies investigate the human genome and the putative role of polymorphic genes. However, more than half of the human genome is expressed in the brain. Being highly polygenic, psychological and behavioural processes are admittedly difficult to reasonably infer from the study of individual alleles. In this sense, our research belongs to the minority of studies explicitly focused on the links between human mental states and maintenance of polymorphism in particular genes, in this case the *RHD* gene.

The existence of a correlation between RhD heterogeneity and physical health is not surprising. Physical health is substantially affected by pathogens and parasites, which still play a large role in human populations: even today, about 48% of deaths under 45 years of age are caused by infectious diseases (Kapp, 1999). This finding is therefore in line with other studies suggesting that genes under balancing selection are disproportionately often related to immunity (see Section 1). The correlation between RhD genotype and mental health is more intriguing, pointing to a strong role of physiology. It has been suggested that in RhD-negative subjects impaired function of the ion channel consisting of two molecules of RhAG (rhesus-associated glycoprotein) and the presence of two (instead of one) molecules of RhCE can lead to local hypoxia. This may enhance neuroinflammation in some brain areas in RhD-negative subjects, especially those who are also infected with *Toxoplasma* (Prandota, 2012). This could help explain the association of toxoplasmosis and RhD-negative phenotype with autism and epilepsy, as well as improvement in autistic symptoms after hyperbaric oxygen therapy (Rossignol, 2007; Rossignol et al., 2007). Prandota's (2012) hypotheses are further corroborated by a predicted association between toxoplasmosis and autism (Prandota, 2011), possibly the strongest association between toxoplasmosis and any mental disorder, many years before it was actually demonstrated in epidemiological surveys (Flegr & Horáček, 2018).

Impaired health of RhD-negative (homozygotic) subjects can be easily explained by the absence of the correct Rh complex on the surface of erythrocytes. We can only speculate why RhD-positive

heterozygotes perform better than RhD-positive homozygotes and why RhD-positive homozygotes score worse on some parameters than RhD-negative homozygotes (Table 1). It is known that erythrocytes of RhD-positive heterozygotic individuals have markedly fewer RhD proteins on their surface than those of RhD-positive homozygotes (17,720 vs. 33,560: Le Van Kim et al., 2006), which might present fewer targets for pathogens. Alternatively, it is possible that erythrocytes of heterozygotes with a generally lower number of Rh complexes, or with some proportion of aberrant (2 × RhAG + 2 × RhCE) plus some normal Rh complexes (2 × RhAG + RhCE + RhD; op. cit.), perform better than erythrocytes of RhD-positive homozygotes. For example, only 30% of subjects worldwide are currently infected with *Toxoplasma*, and in the past three decades, this number has been declining (Tenter et al., 2000). It is, however, highly probable that in the past, especially in tropical Africa, our evolutionary homeland, over 90% of humans would have been infected and therefore adapted to the infected condition. In accordance with this hypothesis, Novotná et al. (2008) showed that in a *Toxoplasma*-free people, RhD-positive homozygotes performed slightly worse on psychomotor tasks than RhD-positive heterozygotes but much worse than RhD-negative homozygotes. In *Toxoplasma*-infected subjects, on the other hand, RhD-positive homozygotes performed equally poorly as RhD-negative homozygotes, and much worse than RhD-positive heterozygotes (see also Flegr et al., 2018, for a confirmatory study).

Frequencies of RhD-negative individuals (22.7%) and their parents (22.6%) in our sample were relatively high in comparison with the general prevalence in Europe (14%–19%) (Flegr & Dama, 2014). One could speculate that some small part of the respondents, both RhD-positive and RhD negative, say 5%, might have misreported their own or their parents' RhD phenotype. Such stochastic error would overestimate the frequency of the rarer variant, in this case the frequency of RhDnegatives, and underrate the frequency of the more common variant in the population. This is the most parsimonious explanation of the higher frequency of RhD negatives in our sample, but it is unlikely. Over the past six years, we immunoanalysed in our laboratory the RhD phenotype of 2,600 volunteer members of Labbunnies community. In that sample, we found 22.3% of RhD negatives, which is relatively close to the prevalence reported here. This close agreement thus independently supports the reliability of RhD phenotype-related data provided by respondents in the current study. The most likely explanation of this result is that the personality profile, and therefore also willingness to voluntarily participate in research, differs between RhD-negative and RhD-positive subjects, as discussed next.

4.1 | Strengths and limitations of the present study

The most important strength of our study is that the two tested hypotheses, methods of data collection (including a stopping rule), and methods of analysing the collected data were preregistered prior to commencement of the study. Another strength of our questionnaire-based study is its relatively high number of participants. It is also significant that the study was blinded: subjects were invited to participate in a study that would 'investigate certain philosophical problems of personal identity, test evolutionary hypotheses regarding, e.g., the beauty of flowers, and study moral dilemmas related to autonomous cars'. Subjects therefore did not know the study would include an investigation into the relation between RhD genotype and health and could therefore not intentionally or unintentionally bias their responses.

The main limitation of our study is that the participants themselves provided information about their health, their and their parents' RhD phenotype. It is thus highly probable that some part of that information was inaccurate or incomplete. It must be stressed, however, that while this kind of statistical noise can increase the risk of false-negative results, that is fail to confirm the existence of an existing association, it cannot produce false positive results; that is, it cannot indicate the existence of a nonexistent association. Our method is therefore conservative, and the strengths of the effects are in reality probably greater than those observed.

In comparison with other similar studies, the current questionnaire contained a relatively small number of questions related to physical health, which made it impossible for us to compute an index of physical health based on variables such as frequency of visits to medical specialists, frequency of antibiotic prescriptions, etc. Similarly, we monitored only a small subset of potentially important confounding variables, such as the use of alcohol or drugs, personality profile, etc. Future studies ought to collect such data to statistically control for such potentially confounding variables. The strength and novelty of our study lies in the inclusion of multiple mental health parameters in this context.

Another potential problem is that the participants were self-selected: they voluntarily and without any material compensation decided to complete a long anonymous questionnaire. One part of the questionnaire was dedicated to the rating of beauty of flowers and this topic was also included in the name of the questionnaire (see Section 2), which could be a reason for the overrepresentation of women (70%) among the participants. Earlier studies performed using the same method showed that with respect to prevalence of 24 neuropsychiatric disorders (Flegr & Horáček, 2020) or religiosity (Kopecky et al., 2019), the composition of the participant population is highly similar to the composition of the general internet population. Nevertheless, it is probable that participants of such studies consist mainly of curious, altruistic, internet-fluent members of the community. The observed results thus cannot be fully generalized to the Czech or even world population as a whole, so it will be necessary to repeat these analyses with other population samples in future.

The strengths of observed associations in Table 1 might seem relatively low. The magnitude of partial Kendall's Tau (here typically $<|0.15|$) suggests that RhD genotype explains merely a small part of total variability of the main health-related variables. This problem is commonly faced by studies performed on highly heterogeneous human populations living in a complex environment and exposed to a broad spectrum of internal and external factors (Flegr, 2013). Moreover, as shown also in our study, even such well-attested adverse factors as smoking, alcohol consumption and high BMI also

explain only a small and rather similar proportion of variability in health-related variables.

5 | CONCLUSIONS

Our study confirmed that RhD-negative homozygotes are characterized by worse and RhD-positive heterozygotes by better physical and mental health. This lends support to the hypothesis according to which balancing selection, or more precisely heterozygote advantage (heterosis), maintains RhD polymorphism in human populations. We further documented how various RhD genotypes differ in variables related to physical and mental health, pointing to *how* balancing selection might act and play important, medically and evolutionarily significant roles in (human) adaptation (Fumagalli et al., 2009; Key et al., 2014). Our study further shows that the health of heterozygotes and homozygotes differs considerably. For the first time, we documented that RhD-positive homozygotes might actually suffer from as bad or even worse health than RhD-negative homozygotes. Therefore, assessing only RhD phenotypes (as opposed to genotypes) to compare the state of health and psychomotor performance of RhD-negative and RhD-positive subjects—as is the case in nearly all published studies—can easily lead to false-negative results, as better health and/or performance of RhD-positive heterozygotes can be diluted by worse health and/or performance of RhD-positive homozygotes due to genetic dominance of the RhD + allele. Ultimately, therefore, the health and/or performance of RhD-positive subjects may turn out to be approximately the same or only slightly better than that of RhD-negative subjects. To avoid this source of error, it would be advisable for future studies to work with RhD genotypes rather than phenotypes.

ACKNOWLEDGMENTS

We would like to thank Anna Pilátová, Ph.D., for final revisions of our text. This work was supported by Czech Science Foundation (grant No. 18-13692S) and Charles University (Research Centre programme No. 204056).

CONFLICT OF INTERESTS

Authors have no conflict of interests.

AUTHOR CONTRIBUTIONS

JF designed the study and analysed data. All four authors participated in data collection, interpretation of results and preparation of the manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jeb.13745.

DATA AVAILABILITY STATEMENT

All relevant data are available at Figshare: figshare.com/s/d199dec0ba0fc5efde29.

ORCID

Jaroslav Flegr https://orcid.org/0000-0002-0822-0126
Jan Toman https://orcid.org/0000-0002-0776-2070
Martin Hůla https://orcid.org/0000-0003-2796-1226
Šárka Kaňková https://orcid.org/0000-0002-3087-7538

REFERENCES

Allison, A. C. (1954). The distribution of the sickle-cell trait in East Africa and elsewhere, and its apparent relationship to the incidence of subtertian malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *48*, 312–318.

Alonso, S., Lopez, S., Izagirre, N., & de la Rua, C. (2008). Overdominance in the human genome and olfactory receptor activity. *Molecular Biology and Evolution*, *25*, 997–1001.

Andrés, A., Dennis, M., Kretzschmar, W., Cannons, J., Lee-Lin, S., Hurle, B., Schwartzberg, P., Williamson, S., Bustamante, C., Nielsen, R., Clark, A., Green, E., & Progra, N. C. S. (2010). Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *Plos Genetics*, *6*, e1001157.

Andres, A., Hubisz, M., Indap, A., Torgerson, D., Degenhardt, J., Boyko, A., Gutenkunst, R., White, T., Green, E., Bustamante, C., Clark, A., & Nielsen, R. (2009). Targets of balancing selection in the human genome. *Molecular Biology and Evolution*, *26*, 2755–2764.

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 (Methodological)*, *57*, 289–300.

Blancher, A., & Socha, W. (1997). The rhesus system. In A. Blancher, J. Klein, & W. Socha (Eds.), *Molecular biology and evolution of blood group and MHC antigens in primates* (pp. 147–218). Springer.

Croze, M., Zivkovic, D., Stephan, W., & Hutter, S. (2016). Balancing selection on immunity genes: Review of the current literature and new analysis in *Drosophila melanogaster*. *Zoology*, *119*, 322–329.

Cuthbert, A., Halstead, J., Ratcliff, R., Colledge, W., & Evans, M. (1995). The genetic advantage hypothesis in cystic fibrosis heterozygotes: A murine study. *Journal of Physiology*, *482*, 449–454.

Dobzhansky, T. (1955). A review of some fundamental concepts and problems of population genetics. *Cold Spring Harbor Symposia on Quantitative Biology*, *20*, 1–15.

Feldman, M. W., Nabholz, M., & Bodmer, W. F. (1969). Evolution of the Rh polymorphism: A model for the interaction of incompatibility, reproductive compensation and heterozygote advantage. *American Journal of Human Genetics*, *21*, 171–193.

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marques-Bonet, T., Ramirez-Soriano, A., Muntasell, A., Navarro, A., Lazarus, R., Calafell, F., Bertranpetit, J., & Casals, F. (2008). Balancing selection is the main force shaping the evolution of innate immunity genes. *Journal of Immunology*, *181*, 1315–1322.

Fisher, R. A., Race, R. R., & Taylor, G. L. (1944). Mutation and the Rhesus reaction. *Nature*, *153*, 106.

Flegel, W. A. (2011). Molecular genetics and clinical applications for RH. *Transfusion and Apheresis Science*, *44*, 81–91.

Flegr, J. (2013). Influence of latent *Toxoplasma* infection on human personality, physiology and morphology: Pros and cons of the *Toxoplasma*-human model in studying the manipulation hypothesis. *Journal of Experimental Biology*, *216*, 127–133.

Flegr, J. (2016). Heterozygote advantage probably maintains Rhesus factor blood group polymorphism: Ecological regression study. *PLoS One*, *11*, e0147955.

Flegr, J., & Dama, M. (2014). Does the prevalence of latent toxoplasmosis and frequency of Rhesus-negative subjects correlate with the nationwide rate of traffic accidents? *Folia Parasitologica*, *61*, 485–494.

Flegr, J., Geryk, J., Volny, J., Klose, J., & Cernochova, D. (2012). Rhesus factor modulation of effects of smoking and age on psychomotor performance, intelligence, personality profile, and health in Czech soldiers. *PLoS One*, *7*, e49478.

Flegr, J., Hoffmann, R., & Dammann, M. (2015). Worse health status and higher incidence of health disorders in Rhesus negative subjects. *PLoS One*, *10*, e0141362.

Flegr, J., & Horáček, J. (2018). Toxoplasmosis, but not borreliosis, is associated with psychiatric disorders and symptoms. *Schizophrenia Research*, *197*, 603–604.

Flegr, J., & Horáček, J. (2020). Negative effects of latent toxoplasmosis on mental health. *Frontiers Psychiatry*, *10*, 1012.

Flegr, J., Kuba, R., & Kopecký, R. (2020). Rhesus-minus phenotype as a predictor of sexual desire and behavior, wellbeing, mental health, and fecundity. *PLoS One*, *15*, e0236134.

Flegr, J., Novotná, M., Fialová, A., Kolbeková, P., & Gašová, Z. (2010). The influence of RhD phenotype on toxoplasmosis- and age-associated changes in personality profile of blood donors. *Folia Parasitologica*, *57*, 143–150.

Flegr, J., Šebánková, B., Příplatová, L., Chvátalová, V., & Kaňková, Š. (2018). Lower performance of *Toxoplasma*-infected, Rh-negative subjects in the weight holding and hand-grip tests. *PLoS One*, *13*, e0200346.

Fumagalli, M., Cagliani, R., Pozzoli, U., Riva, S., Comi, G., Menozzi, G., Bresolin, N., & Sironi, M. (2009). Widespread balancing selection and pathogen-driven selection at blood group antigen genes. *Genome Research*, *19*, 199–212.

Greaves, J., Redfern, R., Ayres, P., & Gill, J. (1977). Warfarin resistance: A balanced polymorphism in the Norway rat. *Genetics Research*, *30*, 257–263.

Haldane, J. B. S. (1942). Selection against heterozygosis in man. *Eugenics*, *11*, 333–340.

Haldane, J. B. S. (1949). Disease and evolution. *Supplement to La Ricerca Scientifica*, *19*, 68–76.

Halmin, M., Rostgaard, K., Lee, B. K., Wikman, A., Norda, R., Nielsen, K. R., Pedersen, O. B., Holmqvist, J., Hjalgrim, H., & Edgren, G. (2017). Length of storage of red blood cells and patient survival after blood transfusion a binational cohort study. *Annals of Internal Medicine*, *166*, 248–256.

Hedrick, P. (2012). What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution*, *27*, 698–704.

Hellgren, O., & Sheldon, B. (2011). Locus-specific protocol for nine different innate immune genes (antimicrobial peptides: Beta-defensins) across passerine bird species reveals within-species coding variation and a case of trans-species polymorphisms. *Molecular Ecology Resources*, *11*, 686–692.

Holub, D., Bankovská, M., Dragomirecká, E., Rodriguez, M., Preiss, M., Novák, T., Čermák, J., Horáček, J., Libiger, J., & Flegr, J. (2011). Possible protective function of Rh factor in schizophrenia. *Psychiatrie*, *15*, 37–42.

Kaňková, Š., Flegr, J., Toman, J., & Calda, P. (2020). Maternal RhD heterozygous genotype is associated with male biased secondary sex ratio. *Early Human Development*, *140*, 104864.

Kaňková, Š., Šulc, J., & Flegr, J. (2010). Increased pregnancy weight gain in women with latent toxoplasmosis and RhD-positivity protection against this effect. *Parasitology*, *137*, 1773–1779.

Kapp, C. (1999). WHO warns of microbial threat. *Lancet*, *353*, 2222.

Key, F., Teixeira, J., de Filippo, C., & Andres, A. (2014). Advantageous diversity maintained by balancing selection in humans. *Current Opinion in Genetics & Development*, *29*, 45–51.

Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, *22*, 665–674.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, *217*, 624–626.

Kopecky, R., Boschetti, S., & Flegr, J. (2019). Effect of being religious on wellbeing in a predominantly atheist country: Explorative study on wellbeing, fitness, physical and mental health. *PsyArXiv*, April 29.

Kustu, S., & Inwood, W. (2006). Biological gas channels for NH_3 and CO_2: Evidence that Rh (rhesus) proteins are CO_2 channels. *Transfusion Clinique Et Biologique*, *13*, 103–110.

Lalovic, A., Merkens, L., Russell, L., Arsenault-Lapierre, G., Nowaczyk, M., Porter, F., Steiner, R., & Turecki, G. (2004). Cholesterol metabolism and suicidality in smith-lemli-opitz syndrome carriers. *American Journal of Psychiatry*, *161*, 2123–2125.

Le Van Kim, C., Colin, Y., & Cartron, J. P. (2006). Rh proteins: Key structural and functional components of the red cell membrane. *Blood Reviews*, *20*, 93–110.

Lewontin, R. (1974). *The genetic basis of evolutionary change*. Columbia University Press.

Lewontin, R., & Hubby, J. (1966). A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, *54*, 595–609.

McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Sparky House Publishing.

Nakhoul, N. L., & Hamm, L. L. (2013). Characteristics of mammalian Rh glycoproteins (SLC42 transporters) and their role in acid-base transport. *Molecular Aspects of Medicine*, *34*, 629–637.

Nei, M., Li, W., Tajima, F., & Narain, P. (1981). Polymorphism and evolution of the Rh blood-groups. *Japanese Journal of Human Genetics*, *26*, 263–278.

Novotná, M., Havlíček, J., Smith, A. P., Kolbeková, P., Skallová, A., Klose, J., Gašová, Z., Písačka, M., Sechovská, M., & Flegr, J. (2008). *Toxoplasma* and reaction time: Role of toxoplasmosis in the origin, preservation and geographical distribution of Rh blood group polymorphism. *Parasitology*, *135*, 1253–1261.

Prandota, J. (2011). Metabolic, immune, epigenetic, endocrine and phenotypic abnormalities found in individuals with autism spectrum disorders, Down syndrome and Alzheimer disease may be caused by congenital and/or acquired chronic cerebral toxoplasmosis. *Research in Autism Spectrum Disorders*, *5*, 14–59.

Prandota, J. (2012). Rhesus-associated glycoprotein (RhAG) phenotype of the red blood cells modulates *T. gondii* infection-associated psychomotor performance reaction times and changes in the human personality profile. Impaired function of the CO_2, AQP1, and AQP4 gas channels may cause hypoxia and thus enhance neuroinflammation in autistic individuals. In Gemma, C. (Ed.), *Neuroinflammation: Pathogenesis, mechanisms and management* (pp. 423–439). : Nova Publishers.

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raffini, F., Fruciano, C., Franchini, P., & Meyer, A. (2017). Towards understanding the genetic basis of mouth asymmetry in the scale-eating cichlid *Perissodus microlepis*. *Molecular Ecology*, *26*, 77–91.

Raffini, F., & Meyer, A. (2019). A comprehensive overview of the developmental basis and adaptive significance of a textbook polymorphism: Head asymmetry in the cichlid fish *Perissodus microlepis*. *Hydrobiologia*, *832*, 65–84.

Rossignol, D. A. (2007). Hyperbaric oxygen therapy might improve certain pathophysiological findings in autism. *Medical Hypotheses*, *68*, 1208–1227.

Rossignol, D. A., Rossignol, L. W., James, S. J., Melnyk, S., & Mumper, E. (2007). The effects of hyperbaric oxygen therapy on oxidative stress, inflammation, and symptoms in children with autism: An open-label pilot study. *BMC Pediatrics*, *7*, 36.

Roux, C., Pauwels, M., Ruggiero, M., Charlesworth, D., Castric, V., & Vekemans, X. (2013). Recent and Ancient Signature of Balancing Selection around the S-Locus in *Arabidopsis halleri* and *A. lyrata*. *Molecular Biology and Evolution*, *30*, 435–447.

Santangelo, G., Piscopo, F., Santangelo, F., & Trojano, L. (2018). Neuropsychological profile in parents of adult phenylketonuria patients. *Neurological Sciences*, *39*, 161–164.

Segurel, L., Thompson, E., Flutre, T., Lovstad, J., Venkat, A., Margulis, S., Moyse, J., Ross, S., Gamble, K., Sella, G., Ober, C., & Przeworski, M. (2012). The ABO blood group is a trans-species polymorphism in primates. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 18493–18498.

Single, R., Martin, M., Gao, X., Meyer, D., Yeager, M., Kidd, J., Kidd, K., & Carrington, M. (2007). Global diversity and evidence for coevolution of KIR and HLA. *Nature Genetics*, *39*, 1114–1119.

Spurgin, L., & Richardson, D. (2010). How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B-Biological Sciences*, *277*, 979–988.

Tenter, A. M., Heckeroth, A. R., & Weiss, L. M. (2000). *Toxoplasma gondii*: From animals to humans. *International Journal for Parasitology*, *30*, 1217–1258.

Těšický, M., & Vinkler, M. (2015). Trans-species polymorphism in immune genes: General pattern or MHC-restricted phenomenon? *Journal of Immunology Research*, *2015*, 838035.

Trivers, R. L., & Willard, D. E. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science*, *179*, 90–92.

Voracek, M., Haubner, T., & Fisher, M. L. (2008). Recent decline in nonpaternity rates: A cross-temporal meta-analysis. *Psychological Reports*, *103*, 799–811.

Wagner, F. F., & Flegel, W. A. (2000). RHD gene deletion occurred in the Rhesus box. *Blood*, *95*, 3662–3668.

Westhoff, C., & Wylie, D. (1996). Investigation of the RH locus in gorillas and chimpanzees. *Journal of Molecular Evolution*, *42*, 658–668.

WHO (2008). *The global burden of disease: 2004 update*. World Health Organization.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Flegr J, Toman J, Hůla M, Kaňková Š. The role of balancing selection in maintaining human RhD blood group polymorphism: A preregistered cross-sectional study. *J Evol Biol*. 2021;34:426–438. https://doi.org/10.1111/jeb.13745

Supplement 1: Questionnaire

Questions used in this research

You were (officially) born as:

1 - Male

2 - Female

Your age:

dropdown list

1 - 10

2 - 11

3 - 12

…

89 - 99 or more

Your height (cm):

dropdown list

1 – 120 or less
2 – 121
3 – 122
…
100 – 220 or more

Your body weight (kg):

dropdown list

1 – 25 or less
2 – 26
3 – 27
…
155 – 180 or more

The highest level of education you reached:

1 - primary

2 - primary, but I am studying at a secondary school

3 - secondary vocational school (without A-level leaving certificate)

4 - complete secondary or higher educational (A-level leaving certificate)

5 - complete secondary or higher educational, I am studying for bachelor degree

6 - bachelor's

7 - master's

8 – doctoral

You currently live in a neighborhood:

1 - with less than 1000 residents

2 - with up to 5000 residents

3 - with up to 50,000 residents

4 - with up to 100,000 residents

5 - up to 500,000 residents

6 - more than 500,000 residents

7 - Bratislava

8 - Praha

By scrolling the bar indicate the number of your own and foster children:
(it is necessary to click the scale even for 0, 10 means 10 and more)

Own children:
0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more

Foster children:
0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more

Your blood type (AB0 blood group):

1 - I don't know / not sure

2 - A

3 - B

4 - AB

5 - 0

Your Rh factor:

1 - I don't know / I'm not sure

2 - Negative (the rarer one)

3 - Positive (the more common one)

Rh factor of your biological mother:

1 - I don't know / I'm not sure

2 - Negative (the rarer one)

3 - Positive (the more common one)

Rh factor of your biological father:

1 - I don't know / I'm not sure

2 - Negative (the rarer one)

3 - Positive (the more common one)

How many cigarettes do you smoke a day?

0 - 0
1 - 0–0.1
2 - 0.1–1
3 - 1.1–3
4 - 3.1–10
5 - 11–20
6 - 21–40
7 - over 40

How often do you drink alcohol
(in a volume that would make it illegal for you to drive a car for a while)?

0 – never
1 – at most once a month
2 – at most twice a month
3 – at most four times a month
4 - at most twice a week
5 – every other day
6 – every day
7 – nearly all the time

Sexual behavior and sexual preferences

(If you want to reply 0 - definitively no, you will have to at least click on the bar):

slider bar

Males sexually attract you:

0 – definitively no
10
20
30
40
50
60
70
80
90
100 – definitively yes

Females sexually attract you:

0 – definitively no
10
20
30
40
50
60
70
80
90
100 – definitively yes

How have you been feeling physically in the last two years?

scale

1 - Very poorly

2

3

4

5

6 - Very well

How have you been feeling mentally in the last two years?

scale

1 - Very poorly

2

3

4

5

6 - Very well

How many kinds of drugs prescribed by a medical doctor for mental health problems did you use over the past month?

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and more

How many kinds of other drugs prescribed by a medical doctor did you use over the past month?

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and more

Following questions refer to present (not past) moods, feelings and behavior, not only to medically diagnosed illnesses.
If you take medication for a given illness, try to estimate how would you feel without medication.
(if you want to answer 0 - not at all - you have to at least click on the scale)

Do you suffer from depression?

slider bar

0 – definitively no
10
20
30
40
50
60
70
80
90
100 - definitively yes

The same applies also for questions:

Do you suffer from any mania (opposite of depression)?

Do you suffer from anxiety?

Do you suffer from any phobia?

Do you suffer from obsessive (compulsive) behavior?

Do you suffer from auditory hallucinations?

Do you suffer from visual hallucinations?

Do you suffer from burnout syndrome (a feeling of emotional exhaustion, fatigue?)

Do you suffer from frequent headaches? (0- never, 25- once in six months, 50- once a month, 75- once a week, 100-practically every day)

What type of psychological or psychiatric disorder do you suffer from or have you suffered from (both diagnosed and undiagnosed)?

(note the age since when you've known or suspected it in the space provided below)

multiple answers allowed

1 – phobias
2 – Asperger syndrome
3 – autism
4 – unipolar depression
5 – bipolar disorder
6 – schizophrenia
7 – pathological anxiety
8 – alcoholism
9 – gambling addiction
10 – Parkinson's disease
11 – Alzheimer's disease
12 – epilepsy
13 – drug addiction (nicotine excepted)
14 – post-traumatic stress disorder
15 – obsessive-compulsive disorder
16 – panic attacks
17 – sleep disorder – insomnia
18 – learning disability (e.g. dyslexia)
19 – borderline personality disorder
20 – antisocial personality disorder
21 – hyperactivity personality disorder
22 – bulimia, anorexia
23 – burnout syndrome
24 – sexual disorder (e.g. paraphyly)
25 – suicide attempt
26 – other (please specify in the space below)

The same applies also for the question:

What type of psychological or psychiatric disorder do you suffer from or have you suffered from (only diagnosed)?

(note the age since when you've known or suspected it in the space provided below)

The questionnaire also contained 44-item Big-Five scale and 64-item CSIV scale

Big-Five:

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (Vol. 2, pp. 102–138). New York: Guilford Press.

CSIV:

Locke, K.D. (2000). Circumplex scales of interpersonal values: Reliability, validity, and applicability to interpersonal problems and personality disorders. *Journal of Personality Assessment, 75*, 249-267.

Supplement 2

Tables S1–S7 and Figures S1 and S2

Table S1 Effects of sex, RhD-negative phenotype, and interaction between sex and RhD-negative phenotype on physical or mental health

		Physical health		Mental health	
	Degrees of freedom	Wald	p	Wald	p
Intercept	5	1834.909	0.000	1738.525	0.000
Age	1	11.263	0.001	26.277	0.000
Size of place of residence	1	0.024	0.878	0.098	0.754
Sex	1	2.537	0.111	18.817	0.000
Sex*RhD-negative phenotype	1	6.330	0.012	1.419	0.234
Sex*RhD-negative phenotype	1	2.395	0.122	0.250	0.617

The table shows the results (Wald statistics and two-sided p) of ordinal regression.

Table S2 Effects of sex, RhD-heterozygosity, and interaction between sex and RhD-heterozygosity on physical or mental health

		Physical health		Mental health	
	Degrees. of freedom	Wald	p	Wald	p
Intercept	5	1834.909	0.000	1738.525	0.000
Age	1	11.263	0.001	26.277	0.000
Size of place of residence	1	0.024	0.878	0.098	0.754
Sex	1	2.537	0.111	18.817	0.000
RhD-heterozygosity	1	6.330	0.012	1.419	0.234
Sex* RhD-heterozygosity	1	2.395	0.122	0.250	0.617

The table shows the results (Wald statistics and two-sided p) of ordinal regression.

Table S3 Effects of age, sex, size of place of residence, RhD-negative phenotype, and interaction between sex and RhD-heterozygosity on physical and mental health

	Test	Value	F	Effect df	Error df	p	Partial eta-squared
Intercept	Wilks	0.183	475.537	14.000	1496.000	0.000	0.817
Age	Wilks	0.826	22.555	14.000	1496.000	0.000	0.174
Place of residence	Wilks	0.991	0.972	14.000	1496.000	0.480	0.009
Sex	Wilks	0.954	5.174	14.000	1496.000	0.000	0.046
RhD-negative phenotype	Wilks	0.987	1.431	14.000	1496.000	0.131	0.013
sex*RhD-negative phenotype	Wilks	0.986	1.467	14.000	1496.000	0.116	0.014

This table shows the results of MANCOVA with 14 health-related output variables (Physical health, Mental health, Drugs mental health, Drugs other, Number of disorders, Anxiety, Phobias, Depression, Mania, Compulsive behaviour, Hallucination auditory, Hallucination visual, Burnout, and Headache) and variables and interaction listed in the first column as independent variables. Two-sided test.

Table S4 Effects of age, sex, size of place of residence, RhD-negative phenotype, and sex-RhD-negative phenotype, RhD-negative phenotype-age, and sex-RhD-negative phenotype-age interactions on physical and mental health

	Test	Value	F	Effect - df	Error - df	p	Partial eta-squared
Intercept	Wilks	0.218	382.689	14.000	1494.000	0.000	0.782
Age	Wilks	0.863	16.975	14.000	1494.000	0.000	0.137
Sex	Wilks	0.953	5.255	14.000	1494.000	0.000	0.047
Place of residence	Wilks	0.991	0.970	14.000	1494.000	0.482	0.009
RhD-negative phenotype	Wilks	0.984	1.743	14.000	1494.000	0.042	0.016
Sex*RhD-negative phenotype	Wilks	0.984	1.776	14.000	1494.000	0.037	0.016
RhD-negative phenotype*age	Wilks	0.986	1.511	14.000	1494.000	0.099	0.014
Sex*RhD-negative phenotype*age	Wilks	0.985	1.673	14.000	1494.000	0.055	0.015

For legend, see Table S3.

Table S5 Effects of sex, RhD-heterozygosity, and interaction between sex and RhD-heterozygosity on physical and mental health

	Test	Value	F	Effect - df	Error - df	p	Partial eta-squared
Intercept	Wilks	0.192	448.843	14.000	1496.000	0.000	0.808
Age	Wilks	0.827	22.292	14.000	1496.000	0.000	0.173
Place of residence	Wilks	0.991	0.946	14.000	1496.000	0.508	0.009
Sex	Wilks	0.967	3.599	14.000	1496.000	0.000	0.033

RhD-heterozygosity	Wilks	0.983	1.866	14.000	1496.000	0.026	0.017
Sex*RhD-heterozygosity	Wilks	0.985	1.603	14.000	1496.000	0.072	0.015

For legend, see Table S3.

Table S6 Effects of age, sex, size of place of residence, RhD-heterozygosity, and sex-RhD-heterozygosity, RhD-heterozygosity-age, and sex-RhD-heterozygosity-age interactions on physical and mental health

	Test	Value	F	Effect - df	Error - df	p	Partial eta-squared
Intercept	Wilks	0.337	210.014	14.000	1494.000	0.000	0.663
Age	Wilks	0.950	5.624	14.000	1494.000	0.000	0.050
Place of residence	Wilks	0.991	0.940	14.000	1494.000	0.514	0.009
Sex	Wilks	0.967	3.675	14.000	1494.000	0.000	0.033
RhD-heterozygosity	Wilks	0.992	0.812	14.000	1494.000	0.657	0.008
Sex*RhD-heterozygosity	Wilks	0.989	1.218	14.000	1494.000	0.255	0.011
rhHeterozyg*age	Wilks	0.992	0.867	14.000	1494.000	0.595	0.008
Sex*RhD-heterozygosity*age	Wilks	0.981	2.080	14.000	1494.000	0.011	0.019

For legends see the table S3.

Table S7 Correlation between RhD phenotype and genotype and health-related variables.

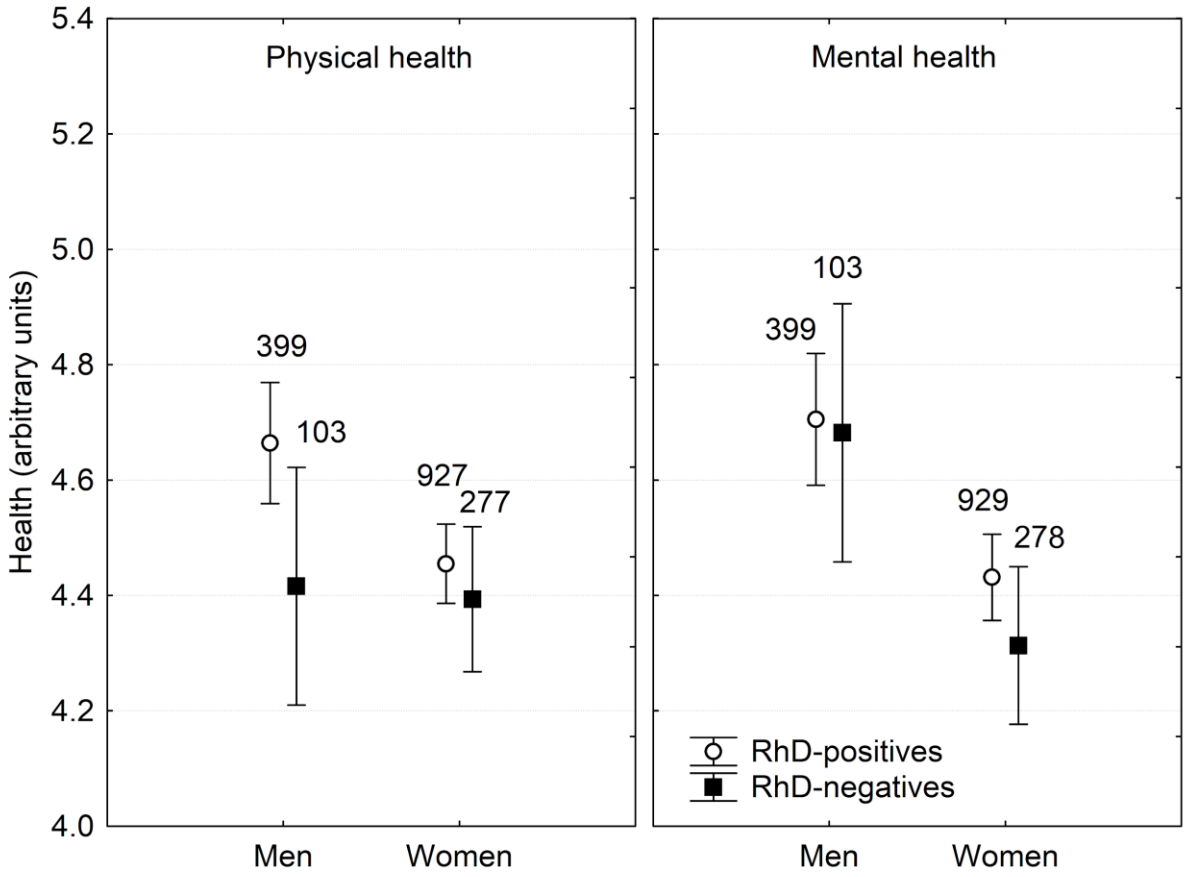
				All					
	Age	Place	p-/p+	+-/p-	+-/+m	+m/p-	Alcohol	Smoking	BMI
Physical health	**-0.05**	0.01	**-0.04**	**0.09**	**0.04**	**0.04**	**0.04**	**-0.06**	**-0.14**
Mental health	**0.10**	0.00	**-0.03**	**0.10**	**0.06**	0.02	0.03	**-0.10**	-0.02
Drugs mental health	**0.10**	0.00	-0.02	0.04	0.01	0.02	**-0.05**	**0.11**	**0.05**
Drugs other	**0.12**	0.01	-0.03	-0.01	**-0.04**	**0.04**	**-0.05**	-0.03	**0.10**
Number of disorders	-0.03	**0.05**	**0.04**	-0.02	0.01	**-0.04**	**-0.07**	**0.13**	0.01
Anxiety	**-0.10**	0.02	**0.03**	**-0.06**	-0.02	-0.03	0.00	**0.06**	0.00
Phobias	**-0.15**	0.00	**0.05**	-0.03	0.02	**-0.05**	-0.02	0.03	0.03
Depression	**-0.11**	0.01	**0.04**	**-0.06**	-0.02	**-0.04**	-0.01	**0.09**	0.00
Mania	**-0.12**	-0.02	**0.04**	-0.01	0.03	**-0.04**	**0.04**	**0.06**	0.02
Compulsive behaviour	**-0.19**	0.00	0.02	-0.02	0.00	-0.03	0.02	**0.04**	0.00
Hallucination auditory	**-0.04**	-0.01	**0.03**	-0.03	0.01	**-0.04**	0.00	0.02	**0.03**
Hallucination visual	-0.01	-0.01	0.03	-0.01	0.02	**-0.04**	0.00	**0.04**	**0.05**
Burnout	-0.01	0.00	0.02	-0.05	-0.03	-0.01	0.00	**0.07**	0.03
Headache	**-0.09**	0.00	0.01	0.01	0.01	-0.01	-0.01	0.01	**0.03**
Number of children	**0.52**	**-0.08**	-0.01	-0.01	-0.02	0.01	-0.02	-0.03	**0.04**
Sexual desire	0.02	**-0.04**	-0.01	0.00	-0.01	0.01	0.00	-0.02	-0.01
BMI	**0.23**	**-0.05**	-0.01	**0.05**	0.03	0.00	**-0.03**	0.02	NA
Height	**-0.03**	**0.03**	-0.01	0.02	0.01	0.01	**0.04**	0.01	-0.01
Weight	**0.18**	**-0.04**	-0.02	**0.05**	0.02	0.01	-0.01	0.02	**0.69**
				Men					
	Age	Place	p-/p+	+-/p-	+-/+m	+m/p-	Alcohol	Smoking	BMI
Physical health	-0.03	0.00	-0.02	**0.07**	**0.05**	0.01	**0.06**	**-0.05**	**-0.15**
Mental health	**0.10**	0.00	**-0.04**	**0.11**	**0.06**	0.03	0.04	**-0.08**	-0.03
Drugs mental health	**0.11**	-0.02	-0.03	0.03	0.00	0.03	**-0.05**	**0.09**	**0.04**
Drugs other	**0.09**	-0.01	-0.03	0.00	-0.04	**0.04**	**-0.04**	-0.01	**0.10**
Number of disorders	0.01	0.03	0.03	-0.05	-0.02	-0.03	**-0.04**	**0.12**	0.00
Anxiety	**-0.08**	0.02	0.03	-0.06	-0.02	-0.03	0.01	0.04	-0.01

Phobias	**-0.13**	-0.01	0.03	-0.03	0.00	-0.04	-0.01	-0.01	0.03
Depression	**-0.11**	0.01	0.03	**-0.07**	-0.04	-0.03	-0.01	**0.05**	0.00
Mania	**-0.10**	-0.02	**0.05**	-0.04	0.02	**-0.06**	0.03	**0.06**	0.03
Compulsive behaviour	**-0.17**	-0.01	0.03	-0.05	-0.02	-0.03	0.03	0.02	0.00
Hallucination auditory	**-0.04**	-0.01	**0.05**	-0.04	0.01	**-0.05**	0.01	0.01	**0.05**
Hallucination visual	-0.01	-0.02	0.04	-0.03	0.01	-0.04	0.00	0.03	**0.07**
Burnout	0.00	0.00	0.02	**-0.08**	**-0.05**	-0.01	0.01	**0.07**	0.04
Headache	**-0.09**	0.00	0.01	-0.03	-0.02	-0.01	0.00	0.02	**0.05**
Number of children	**0.52**	**-0.08**	0.00	-0.03	-0.03	0.00	**-0.05**	-0.02	0.01
Sexual desire	0.01	-0.03	-0.03	0.01	-0.02	0.03	0.00	-0.02	**-0.04**
BMI	**0.22**	**-0.06**	-0.01	0.03	0.02	0.01	-0.01	0.02	NA
Height	**-0.04**	**0.03**	0.00	0.04	0.03	-0.01	0.03	0.02	-0.01
Weight	**0.18**	**-0.05**	-0.02	0.05	0.02	0.01	0.00	0.03	**0.74**
				Women					
	Age	Place	p-/p+	+-/p-	+-/+m	+m/p-	Alcohol	Smoking	BMI
Physical health	-0.03	0.00	-0.02	**0.07**	**0.05**	0.01	**0.06**	**-0.05**	**-0.15**
Mental health	**0.10**	0.00	**-0.04**	**0.11**	**0.06**	0.03	0.04	**-0.08**	-0.03
Drugs mental health	**0.11**	-0.02	-0.03	0.03	0.00	0.03	**-0.05**	**0.09**	**0.04**
Drugs other	**0.09**	-0.01	-0.03	0.00	-0.04	**0.04**	**-0.04**	-0.01	**0.10**
Number of disorders	0.01	0.03	0.03	-0.05	-0.02	-0.03	**-0.04**	**0.12**	0.00
Anxiety	**-0.08**	0.02	0.03	-0.06	-0.02	-0.03	0.01	0.04	-0.01
Phobias	**-0.13**	-0.01	0.03	-0.03	0.00	-0.04	-0.01	-0.01	0.03
Depression	**-0.11**	0.01	0.03	**-0.07**	-0.04	-0.03	-0.01	**0.05**	0.00
Mania	**-0.10**	-0.02	**0.05**	-0.04	0.02	**-0.06**	0.03	**0.06**	0.03
Compulsive behaviour	**-0.17**	-0.01	0.03	-0.05	-0.02	-0.03	0.03	0.02	0.00
Hallucination auditory	**-0.04**	-0.01	**0.05**	-0.04	0.01	**-0.05**	0.01	0.01	**0.05**
Hallucination visual	-0.01	-0.02	0.04	-0.03	0.01	-0.04	0.00	0.03	**0.07**
Burnout	0.00	0.00	0.02	**-0.08**	**-0.05**	-0.01	0.01	**0.07**	0.04
Headache	**-0.09**	0.00	0.01	-0.03	-0.02	-0.01	0.00	0.02	**0.05**
Number of children	**0.52**	**-0.08**	0.00	-0.03	-0.03	0.00	**-0.05**	-0.02	0.01

Sexual desire	0.01	-0.03	-0.03	0.01	-0.02	0.03	0.00	-0.02	**-0.04**
BMI	**0.22**	**-0.06**	-0.01	0.03	0.02	0.01	-0.01	0.02	NA
Height	**-0.04**	**0.03**	0.00	0.04	0.03	-0.01	0.03	0.02	-0.01
Weight	**0.18**	**-0.05**	-0.02	0.05	0.02	0.01	0.00	0.03	**0.74**

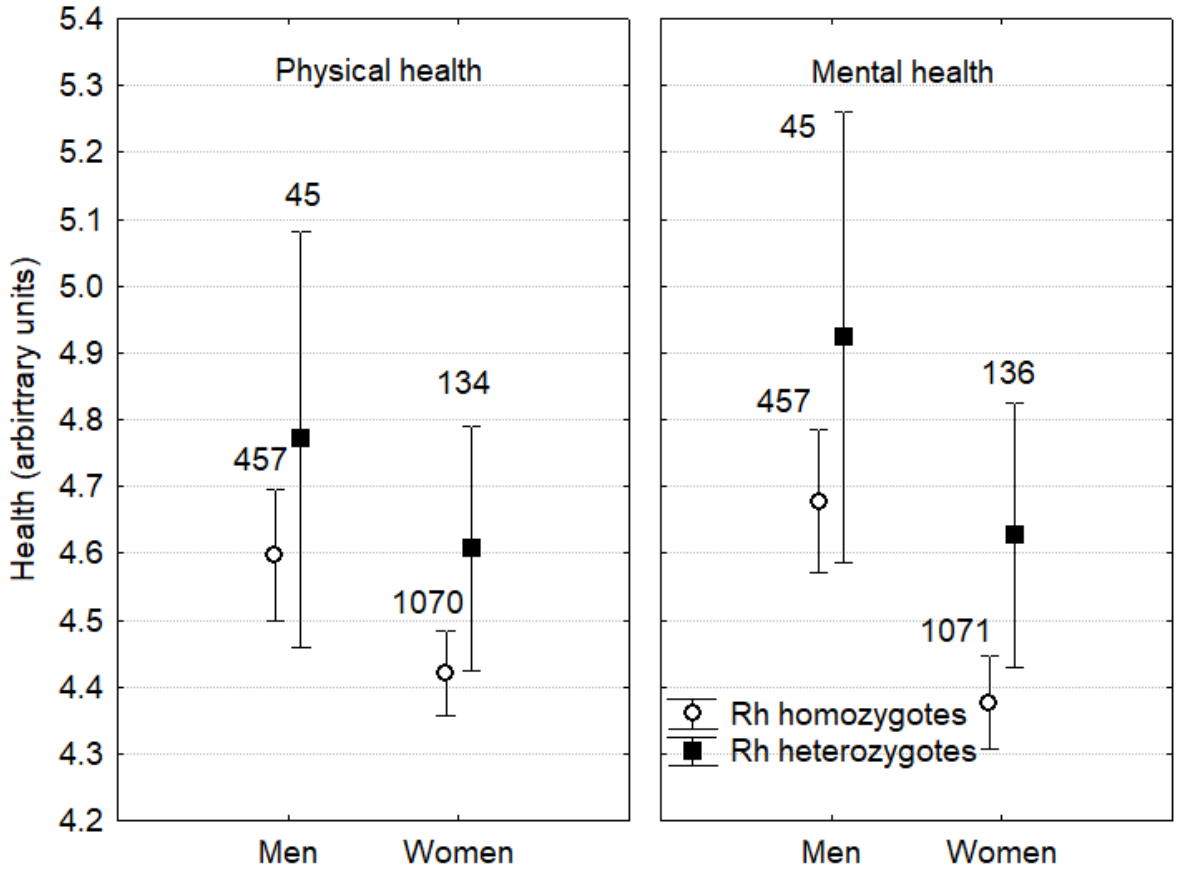
This table shows the direction and strength of particular effects (partial Taus) measured using a nonparametric test, namely partial Kendall correlation test controlled for age and sex (first third of the table), smoking, and alcohol consumption. The p-, p+, +-, and +m column headings denote subjects with RhD-negative homozygotes (p-), subjects with the positive phenotype (p+), RhD-positive heterozgotes (+-), and a mixed RhD-positive population enriched with RhD-positive homozygotes subjects (+m), respectively. Subpopulation +m consisted of those RhD-positive respondents who reported no RhD-negative parent, which means that this subpopulation includes a mixture of RhD-positive homozygotes and heterozygotes enriched with RhD positive homozygotes. Positive Tau indicates that the genotype/phenotype to the left of the slash has a higher value of the variable listed in the first column than the genotype/phenotype to the right of the slash. In accordance with preregistration, one-sided variant of tests were applied. Significant Taus are printed in bold. No correction for multiple tests was done in this exploratory part of the study, although the majority of significant associations remained significant after such correction by Benjamini–Hochberg procedure when two-sided tests were performed. Still, some unexpected associations, e.g. the higher physical health index of women consuming more alcoholic drinks, may indicate the presence of artefacts of multiple tests.

Fig. S1. The effect of RhD phenotype on physical and mental health



RhD-negative homozygotes (black squares) are compared with all other subjects (empty circles). Vertical bars denote 0.95 confidence intervals and numbers above the bars show the number of subjects in particular groups.

Fig. S2. The effect of RhD zygosity on physical and mental health



RhD positive heterozygotes (black squares) are compared with all other subjects (empty circles). Vertical bars denote 0.95 confidence intervals and numbers above the bars show the number of subjects in particular groups.