scientific reports

OPEN

Using a polygenic score in a family design to understand genetic influences on musicality

Laura W. Wesseldijk[1,2,3,7✉], Abdel Abdellaoui[2], Reyna L. Gordon[4,5,6], 23andMe Research Team*, Fredrik Ullén[1,7,11] & Miriam A. Mosing[1,3,7,8,11]

To further our understanding of the genetics of musicality, we explored associations between a polygenic score for self-reported beat synchronization ability (PGS_{rhythm}) and objectively measured rhythm discrimination, as well as other validated music skills and music-related traits. Using family data, we were able to further explore potential pathways of direct genetic, indirect genetic (through passive gene–environment correlation) and confounding effects (such as population structure and assortative mating). In 5648 Swedish twins, we found PGS_{rhythm} to predict not only rhythm discrimination, but also melody and pitch discrimination (betas between 0.11 and 0.16, $p < 0.001$), as well as other music-related outcomes ($p < 0.05$). In contrast, PGS_{rhythm} was not associated with control phenotypes not directly related to music. Associations did not deteriorate within families (N = 243), implying that indirect genetic or confounding effects did not inflate PGS_{rhythm} effects. A correlation ($r = 0.05$, $p < 0.001$) between musical enrichment of the family childhood environment and individuals' PGS_{rhythm}, suggests gene–environment correlation. We conclude that the PGS_{rhythm} captures individuals' general genetic musical propensity, affecting musical behavior more likely direct than through indirect or confounding effects.

Listening to music or engaging actively in music by singing or playing an instrument, is an important part of life for many. Twin research on music-related behavior shows that there is a substantial genetic component to variation in musical engagement, as well as to musical ear, i.e. pitch, melody and rhythm discrimination skills, with heritability estimates ranging between 13 and 86% with an average around 40–50%[1–7]. Twin modelling studies have also shown that there is considerable overlap between the genetic variants underlying such music-related phenotypes (so-called 'genetic pleiotropy'), not only between different specific music skills, but also between music skills and other music-related traits, like music practice or listening to music[3,8,9]. This suggests that there are generalist genes as well as specialist genes underlying variation in music-related traits.

Which particular genes are relevant for music skills is not yet well understood, with the limited existing research largely being based on candidate gene and linkage studies (see Gingras et al.[10] for review), which are known to suffer from replication problems, and are hampered by small sample size[11]. Recently, a genome-wide association study (GWAS) in a well-powered sample of 606,825 individuals identified sixty-nine genetic variants that influence variation in self-reported beat synchronization ability, assessed with the question: 'Can you clap in time with a musical beat?' Yes/No[12]. The variance explained by the aggregate of these genetic effects (i.e., the SNP-based heritability) was ~ 13–16%. It is not surprising that this estimate is lower than heritability estimates of musical rhythm abilities reported in twin studies (~ 50%), as many common genetic variants with very small effects influence human behavior, requiring extremely large GWAS samples to detect all. In addition, potentially influential rare genetic variants are not yet well captured by current microarray chips[13]. Therefore, up-to-date GWASs capture only a fraction of the expected genetic variation underlying behavioural traits (a phenomenon

[1]Department of Neuroscience, Karolinska Institutet, Solnavägen 9, 171 77 Stockholm, Sweden. [2]Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. [3]Melbourne School of Psychological Sciences, Faculty of Medicine, Dentistry, and Health Sciences, University of Melbourne, Melbourne, Australia. [4]Department of Otolaryngology–Head and Neck Surgery, Vanderbilt University Medical Center, Nashville, TN, USA. [5]Department of Psychology, Vanderbilt University, Nashville, TN, USA. [6]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA. [7]Department of Cognitive Neuropsychology, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany. [8]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. *A list of authors and their affiliations appears at the end of the paper. [11]These authors jointly supervised this work: Fredrik Ullén and Miriam A. Mosing. ✉email: laura.wesseldijk@ki.se

referred to as 'missing heritability'). Furthermore, although the single item in this GWAS is obviously limited as a measure of musical aptitude, the authors performed extensive validations of this item, showing that it was well-correlated with accuracy in tapping to the beat of musical excerpts and with other related phenotypes such as rhythm perception, a multi-item rhythm questionnaire, and the Goldsmiths test of musical sophistication (see Niarchou et al.[12] for further details on validation procedures and findings). Most importantly, their GWAS was the first well-powered gene-finding investigation of a music-related trait, and thus allows us to estimate and investigate individuals' genetic rhythmic propensity using a polygenic score (PGS) approach. This approach can provide additional insight into the genetic architecture underlying musical ability, directing future research when disentangling gene–environment interplay underlying musicality and music acquisition.

Over the past decade, post-GWAS analyses were developed to further study genetic variants discovered by GWASs and their associations. In the PGS approach, the estimated effect sizes of genetic variants identified in the discovery sample used in the GWAS, are used as weights in an independent sample, to calculate a score for each individual, representing his or her genetic propensity for the trait under study[14]. It is important to bear in mind that this approach will capture only a fraction of the genetic variation underlying the trait and should not be used to estimate the heritability[13]. We here use the PGS approach to estimate the relationship between the genetics underlying self-reported rhythm ability, as estimated in the beat synchronization ability GWAS (PGS_{rhythm}), and objectively measured rhythm discrimination ability, as measured by a well-validated music discrimination test[1]. We will further compare this relationship to the effect the PGS_{rhythm} has on other music skills and phenotypes such as pitch and melody discrimination, motor timing, music practice, music listening, the experience of flow during musical activities and achievement in music. As a control, we investigate associations between PGS_{rhythm} and phenotypes not directly related to music, including general intelligence, sport practice, flow in non-musical activities, and achievement in non-musical domains.

Besides investigating the predictive value and specificity of the PGS_{rhythm}, we aim to identify the pathways through which the PGS_{rhythm} affects music skills. Genetic variants can relate to a trait through direct, indirect or confounding effects[15]. A *direct* genetic effect is when genetic variants directly affect the trait[16]. Importantly, this does not mean that the pathway necessarily needs to be simple or 'straight' in mechanistic terms[15]. A different scenario, however, is when a genetic variant affects a trait in an environmentally mediated way, for example, through its effect on parental behavior. This would be an example of an *indirect* genetic effect, i.e., the effect of a genetic variant in one individual on the trait of another individual through the environment[15]. In this case, genes that are shared between parents and children, tend to cause the parents to provide their children with a family environment of a particular type. This chain of events will result in a correlation between a genotype and the family environment, a phenomenon called passive gene–environment correlation (prGE)[17]. For example, parents might have provided a musically enriched family environment, due to their genetic predisposition for music; this environment may facilitate music skill acquisition in their offspring. Individuals with certain genotypes will therefore more often also be exposed to facilitating environments, which can lead to an overestimation of the direct effects of an individual's genetic variants in GWASs and therefore of the PGS. Evidence for this overestimation of effects has been found in previous within-family GWASs that found effects of genetic variants to be much smaller within families[18,19]. There are other *confounding* effects that can inflate the association between a PGS and the trait of interest, such as population stratification and assortative mating. Population stratification occurs when there are differences in allele frequencies between (sub)populations, due to systematic differences in ancestry. This can lead to spurious effects of genetic variants causing inflated estimates of the PGS prediction[20]. Assortative mating refers to a type of non-random mating, namely when partner bonds are established based on similarity on particular phenotypes, like for example, playing a musical instrument. This can also cause an overestimation of the direct effects of genetic variants[21]. In sum, over and above direct genetic effects, indirect genetic effects as well as population stratification and assortative mating can inflate the effect of individual's genetic variants in GWASs and therefore of the PGS on traits of interest[20,21].

With the use of genetically informative family data, it is possible to disentangle the PGS' direct and combined indirect genetic (referring to prGE) and confounding effects on a trait[15,16,20,22,23]. Dizygotic twins, just as regular siblings, differ at random in the genetic variants they receive from their parents during meiosis, while sharing on average 50% of their genes. Importantly, as the twins of a pair have the same parents, within-pair effects are not biased by population stratification or assortative mating[24]. Furthermore, dizygotic twins share their family environment and have the same age. Therefore, comparing the effect of a PGS on an outcome, between- and within-dizygotic twin pairs, is a way to control for both prGE through the shared family environment, and confounding effects of assortative mating and population stratification. Specifically, if the effect of the PGS on an outcome is larger in a sample of unrelated individuals than within dizygotic twin pairs, this indicates that the effects of the PGS might be inflated by indirect genetic or confounding effects[16,20]. This validated 'within-family PGS' method has recently gained popularity in the field of behavioural genetics. Even so, the idea of direct and indirect genetic effects is not new and has been described and studied with various different designs long before the development of PGSs, for example, in the field of evolutionary quantitative genetics (see for example[25,26]). Lastly, additional analyses can shed light on whether it is more likely that indirect genetic effects or confounding effects are at play. For example, a relationship between a measure of musical enrichment of the family childhood environment and the PGS_{rhythm} would suggest gene–environment correlation (while the absence of a relationship would not exclude its presence). A significant difference from the expected correlation of 0.50 for the PGS_{rhythm} between dizygotic twins could suggest confounding effects such as population structure and assortative mating[15,27].

In a large genetically informative sample of 5648 Swedish twins, we investigated the association between the PGS of self-reported beat synchronization ability (PGS_{rhythm}), and a well-validated measure of rhythm ability, other music skills (pitch and melody discrimination), motor timing and music-related phenotypes such as music practice and listening behaviour, musical flow experiences and lifetime achievement in music. As control phenotypes, we included general intelligence, sports practice, global and work-related flow experiences as well as

dance-, science- and writing-related achievements. We further explore the potential pathways of direct, indirect and confounding effects of the PGS$_{rhythm}$ on music skills using within-family analyses, and examine the possibility of gene–environment correlation and assortative mating.

Methods

Participants. The Study of Twin Adults: Genes and Environment (STAGE) is a cohort of the Swedish Twin Registry (STR) which includes approximately 32,000 adult twins born between 1959 and 1985[28,29]. In 2012 and 2013, 11,543 twins from this cohort completed a web survey on, among other things, musical engagement, musical aptitude, motor timing (finger tapping), flow proneness and achievements. More details on these phenotypes, their distributions and heritability estimates can be found in earlier published studies using this sample see[1–3,8,30–32]. In 2019 and 2020, individuals from the STR, who provided saliva samples between 2006 and 2008, were genotyped. After quality control, genotype data were available for 8,343 individuals from the STAGE cohort, of which 5648 had also completed the web survey in 2012/2013. Informed consent was obtained from all participants. The study and analyses of biomarkers were approved by the Regional Ethics Review Board in Stockholm (Dnr 2011/570-31/5, Dnr 2018/960-31/2, Dnr 2019-05879). The computations and data handling were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala partially, funded by the Swedish Research Council through grant agreement no. 2018-05973. All research methods were performed in accordance with relevant guidelines and regulations.

Measures. *Start age of playing music.* Participants answered whether they ever played a musical instrument (including singing) and, if so, at what age they started.

Amount of music practice. Total lifetime music practice of the participants was estimated based on questions about start and ending age of playing music and on the participant's indication of the number of hours per week they practiced (in 10 categories ranging from 0 h, via more than 6–9 h, to more than 40 h) during four age intervals (ages 0–5, 6–11, 12–17 and 18 years until time of measurement). From these answers and information on the numbers of years they played music, we calculated an estimate of total lifetime amount of music practice.

Hours of music listening. Individuals were asked how many hours per week they listened to music on average since they were 18 years old until the measurement.

Musical aptitude. Musical aptitude was measured using the Swedish Musical Discrimination Test (SMDT)[1]. The SMDT was administered online and includes three subtests: a pitch, melody and rhythm discrimination test. The pitch (27 items), melody (18 items) and rhythm (18 items) discrimination scores were standardized before calculating a mean overall musical aptitude score. The SMDT has been shown to have good reliability and internal consistency (reliabilities and alphas between 0.79 and 0.89) and inter-correlations between the discrimination scales ranged between 0.27 and 0.41[1].

Motor timing. Motor timing (or finger tapping) was measured using the isochronous serial interval production (ISIP) paradigm[33]. In the ISIP task, participants tap in synchrony with a regular, auditory metronome, followed by a phase where the participant has to continue to tap self-paced without a metronome. After an initial training trial, the participants completed six experimental trials with varying inter-onset intervals (in the order of 524, 819, 655, 1024, 655, and 524 ms). The motor timing score was calculated as the mean coefficient of variation during the self-paced phase across the six trials. We reversed the score so that a higher score indicated a more accurate performance. For detailed information on the task and scoring see Mosing, et al.[30].

Creative achievement, including achievement in music. Achievement was measured with an adapted and translated version of the Creative Achievement Questionnaire (CAQ)[2,34–36]. This instrument is a self-report measure of achievement in the visual, music, dance, writing, invention, and theater domains using a seven-point scale. As an example, items to rate achievement in music range from 1 'I am not engaged in music at all' via 4 'I have played or sung, or my music has been played in public concerts in my home town, but I have not been paid for this' to 7 'I am professionally active as a musician and have been reviewed/featured in national or international media and/or have received an award for my musical activities'. Apart from music, we included dance, which also might be associated with rhythm ability, as well as writing and science as control domains.

Musical enrichment of family childhood environment. The level of musical enrichment of childhood environment was estimated as previously by Wesseldijk et al.[2], i.e. using the first principal component derived from the following family environment measures: (1) the number of records available in the family home, (2) how many individuals in the twin's environment played an instrument, (3) how often they visited concerts and (4) whether or not they were offered music education before the age of 12.

General intelligence. Intelligence was measured with the Wiener Matrizen Test (WMT), a visual matrix test similar to Raven's standard progressive matrices[37]. The test is 25-min long and consists of 24 multiple-choice items, which are scored either one (correct) or zero (incorrect or missing response). A sum of the 24 item scores was used as a measure of the participant's psychometric intelligence.

Flow, including flow proneness in the musical domain. Flow proneness was measured with the Swedish Flow Proneness Questionnaire (SFPQ)[38]. The SFPQ is a self-report measure of the frequency of psychological flow experiences in the domains work, leisure, maintenance and music. All four sub-scales consisted of seven items each rated on a 5-point Likert scale ('never' to 'every or almost every day'). A global flow proneness was calculated as the mean score of flow proneness in work, leisure and maintenance. We included besides flow proneness in the musical domain, global flow proneness and flow proneness in the domains of leisure and work for comparison. See Ullén et al.[38] for a more detailed description of the scale.

Genetic data processing. A total of 8442 twins from the STAGE cohort were genotyped using the Illumina Infinium assay (chip GSAMD-24v1-0_20011747_A1). Before standardized quality control, samples were excluded in case of discrepancies in observed sex and relatedness ($N = 86$). The remaining samples were processed using the Ricopoli pipeline for quality control[39], which led to the exclusion of 13 more samples. Quality control of the single nucleotide polymorphisms (SNPs) led to the exclusion of 581 SNPs, resulting in 491,839 included genotyped SNPs. These SNPs were imputed using the Haplotype Reference Consortium panel (HRC1.1)[40]. PLINK 2.0 sets were generated, and a light post-imputation quality filter was applied, leaving 9,141,508 markers. This resulted in a total sample size of 8343 twin individuals (2554 complete twin pairs including imputed monozygotic twin pairs). Of these, 5648 individuals (2303 monozygotic, 1581 same-sex dizygotic, 1663 opposite-sex dizygotic and 101 individuals of which zygosity was unknown) provided information on music variables. Of which 1184 same-sex and opposite-sex dizygotic complete twin pairs provided genetic information, and 243 dizygotic and opposite-sex dizygotic complete twin pairs provided both genetic and music information.

A principal component analysis (PCA) was performed in the full study sample (including 28 samples identified as non-European ancestral outliers as they remained in the genotype files as they passed all other filters) to generate ancestry covariates. PLINK 1.9 and 2.0 were used to extract the first 20 principal components (PCs) based on common independent genotyped markers (minor allele frequency (MAF) > 0.05, pairwise $R^2 < 0.1$). The first 10 PCs were included as covariates in further PGS analyses, to control for potential confounding by population stratification, i.e., when there are systematic differences in allele frequencies between subpopulations[41].

Polygenic score (PGS$_{rhythm}$) calculation. To create polygenic scores (PGSs) for the 8343 individuals, we extracted a restricted set of common, well-imputed 1,265,094 HapMap 3 SNPs from their genotype data[42,43]. We generated PGSs based on summary statistics from the large GWAS on self-reported beat synchronization ability in N = 606,825 individuals of European ancestry (Niarchou et al.[12] participating in research with the personal genetics company, 23andMe, Inc. Participants in the GWAS had provided informed consent and participated in the research via internet, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent Review Services (E&I Review). Polygenic scores are a weighted sum of each individual's trait-associated alleles at each SNP multiplied by that SNP's estimated effect size as detected by a GWAS[14]. The effect sizes were first re-estimated using the summary-data based on the best linear unbiased prediction (SBLUP) approach[44,45]. This approach computes effect sizes with best linear predictor properties that account for linkage disequilibrium between SNPs. As a reference sample for the linkage disequilibrium, a random sample of 11,064 unrelated individuals was extracted from a set of 1,246,531 HapMap 3 SNPs that passed quality control in the UK Biobank sample[46]. PGSs were generated, based on these re-estimated effect sizes, for the 8,343 samples that passed quality control, using PLINK 1.9. Scores were then imputed for monozygotic co-twins (N = 1350) and merged with the music data and the top 10 of the 20 obtained PCs from the PCA.

Statistical analyses. All analyses were performed in STATA. All variables, with the exception of sex and age and the PCs, were standardized.

Predictions of the PGS$_{rhythm}$. To investigate the relationship between the genetic variants underlying self-reported beat synchronization ability (PGS$_{rhythm}$) and objectively measured rhythmic ability, we performed a linear regression analysis, with the PGS$_{rhythm}$, the top 10 PCs to control for ancestry structures, sex and age as independent variables and the rhythm discrimination score as the dependent variable. Further linear regression analyses were performed to explore the effect of the PGS$_{rhythm}$ on other music-related outcomes, namely pitch and melody discrimination, general musical aptitude, motor timing, amount of music practice, achievement in music, flow proneness in the musical domain, and start age of playing music. In addition, to investigate how specific PGS$_{rhythm}$ predictions were to the music domain, we investigated associations between the genetic variants underlying rhythm discrimination and control phenotypes, i.e. general intelligence, amount of sport practice, flow proneness in non-musical domains and achievements in dance, writing, and science. To estimate an effect size of the PGS$_{rhythm}$ (variance explained R^2) for each dependent variable, we subtracted the total R^2 of the model without the PGS$_{rhythm}$ from the total R^2 of the full model (with PGS$_{rhythm}$). All regression analyses included sex, age and the top 10 PCs as covariates. We always corrected for relatedness in the twin sample by using the robust standard error estimator for clustered observations[47,48]. All analyses were first performed including imputed genetic data for monozygotic twins (N = 5648) and then repeated with only one of the monozygotic twin members included (N = 4787).

Pathways of direct, indirect and confounding effects of the PGS$_{rhythm}$. To explore the potential pathways of direct, indirect and confounding effects of the PGS$_{rhythm}$ on rhythm discrimination, we utilized the family structure (twins) of the sample. We applied a recently developed and validated random intercept mixed-effects model on the dizygotic twin data[16]. In this model, two fixed effects are fitted to partition the total effect of the PGS$_{rhythm}$

on the outcome into a between-family (β_B) and within-family effect (β_W) (see Equation adapted from Selzam et al.[16])

$$Y_{ij} = \propto_0 + \beta_W \left(PGS_{ij} - \overline{PGS}_j \right) + \beta_B \overline{PGS}_j + \gamma_j + \varepsilon_{ij}$$

Y denotes the outcome and PGS the polygenic score, $i = \{1,2\}$ corresponds to the individual twins that are clustered within family j, and $\overline{PGS}$ refers to the mean PGS value in family j. Therefore, the between-family effect represents the expected change in outcome Y given a one unit change in the family PGS average, while the within-effect represents the expected change in outcome given a one unit change in the difference between the individual PGS and the family average PGS. If the between-effect of the PGS on the outcome is significantly larger than the within-effect, indirect genetic or confounding effects may be inflating the effect of the PGS[16,20]. If the within- and between-effect are similar in size this provides evidence for direct effects of the measured genetic variants. Including both the between- and within-family effect in the same model, causes adjustment and independence of the two individual estimates, and offers the possibility to compare the confidence intervals for significance. See Selzam et al.[16] for more details about the model. Only complete dizygotic twin pairs were included in these analyses ($N = 243$). The analysis included sex and 10 PCs as covariates. We repeated this analyses with musical aptitude as an outcome and we repeated the analysis only in same-sex dizygotic twin pairs ($N = 128$).

Additional testing for gene–environment correlation and potential confounding. To further investigate possible indirect or confounding effects of the PGS$_{rhythm}$ on outcomes, namely gene–environment correlation and population structure confounding or assortative mating, we performed two additional analyses. We calculated the correlation between an index of the level of musical enrichment of the family childhood environment and the PGS$_{rhythm}$. A significant correlation would indicate that an individual's genetic make-up (in terms of rhythmic ability) is related to their (perceived) childhood environment, which could be interpreted as further support for gene–environment correlation. We also tested for the potential presence of confounding, such as assortative mating or other types of population structure confounding, by calculating a correlation between the dizygotic twins for the PGS$_{rhythm}$. If 0.50 is not included in the confidence interval of the correlation, this would indicate a significant difference from the expected genetic correlation of 0.50 for the PGS$_{rhythm}$ between siblings.

Results

Predictions of the PGS$_{rhythm}$.
The polygenic score for self-reported beat synchronization ability significantly predicted rhythm discrimination skill ($\beta = 0.11$, $p < 0.001$, $R^2 = 0.01$). Notably, effects for melody discrimination ($\beta = 0.13$, $p < 0.001$, $R^2 = 0.02$), pitch discrimination ($\beta = 0.13$, $p < 0.001$, $R^2 = 0.02$), and the overall musical aptitude score ($\beta = 0.16$, $p < 0.001$, $R^2 = 0.03$) were also significant, and nominally higher than for rhythm discrimination. The PGS$_{rhythm}$ also significantly predicted motor timing ($\beta = 0.17$, $p < 0.001$, $R^2 = 0.03$) and all other music related outcomes (Fig. 1, upper part).

The PGS$_{rhythm}$ did not predict individual differences in general intelligence, total sport practice, and flow proneness or achievements in domains other than music (Fig. 1, lower part). A small but significant effect was seen on achievement in dance ($\beta = 0.04$, $p = 0.02$, $R^2 = 0.001$). Analyses in the sample excluding imputed monozygotic twin data were nearly identical (see Supplementary Table 1).

Pathways of direct, indirect and confounding effects of the PGS$_{rhythm}$.
Within- and between-family PGS$_{rhythm}$ prediction estimates and their confidence intervals from the within-family analyses are displayed in Table 1. Within-family effects were certainly not lower—in fact, nominally higher—than between-family effects, but the partly overlapping confidence intervals indicate no significant difference between within- and between-family effects. This suggests that indirect genetic or confounding effects did not significantly inflate the effect of the PGS$_{rhythm}$. Estimating within- and between-family PGS$_{rhythm}$ effects only within same-sex twins did not change these results.

Additional testing for gene–environment correlation and potential confounding.
There was a small but significant correlation between the level of musical enrichment of family childhood environment and individuals' PGS$_{rhythm}$ ($r = 0.05$, $p < 0.001$, CI 0.02–0.08), supporting the presence of gene–environment correlation. In addition, the PGS$_{rhythm}$ correlation between dizygotic twin members was slightly higher than expected at 0.54 ($p < 0.001$, CI 0.50–0.58). Note that this trend was only borderline significant, with the confidence interval including 0.50.

Discussion

Validity and specificity of the polygenic score for self-reported rhythm ability.
We investigated the predictive value of a polygenic risk score (PGS$_{rhythm}$) for self-rated beat synchronization ability on objective rhythm ability, other musical skills and music-related outcomes, as well as on a set of control outcomes, not directly related to music. Our results support the validity of PGS$_{rhythm}$, in that we indeed found an association between the genetic variants underlying self-reported beat synchronization ability and objectively measured rhythm discrimination, aligning with phenotypic associations reported in Niarchou et al.[12]. Moreover, PGS$_{rhythm}$ did not only predict rhythm skill but tapped broadly into music-related phenotypes. Interestingly, PGS$_{rhythm}$ thus predicted melody discrimination, pitch discrimination and the overall musical aptitude score as well as it predicted rhythm discrimination, if not better. Further, as expected, the PGS$_{rhythm}$ most strongly predicted individual's motor timing skills, converging with neuroscience findings on the role of the motor system in beat synchronization[49]. In addition, associations with PGS$_{rhythm}$ were also significant for other music-related traits,
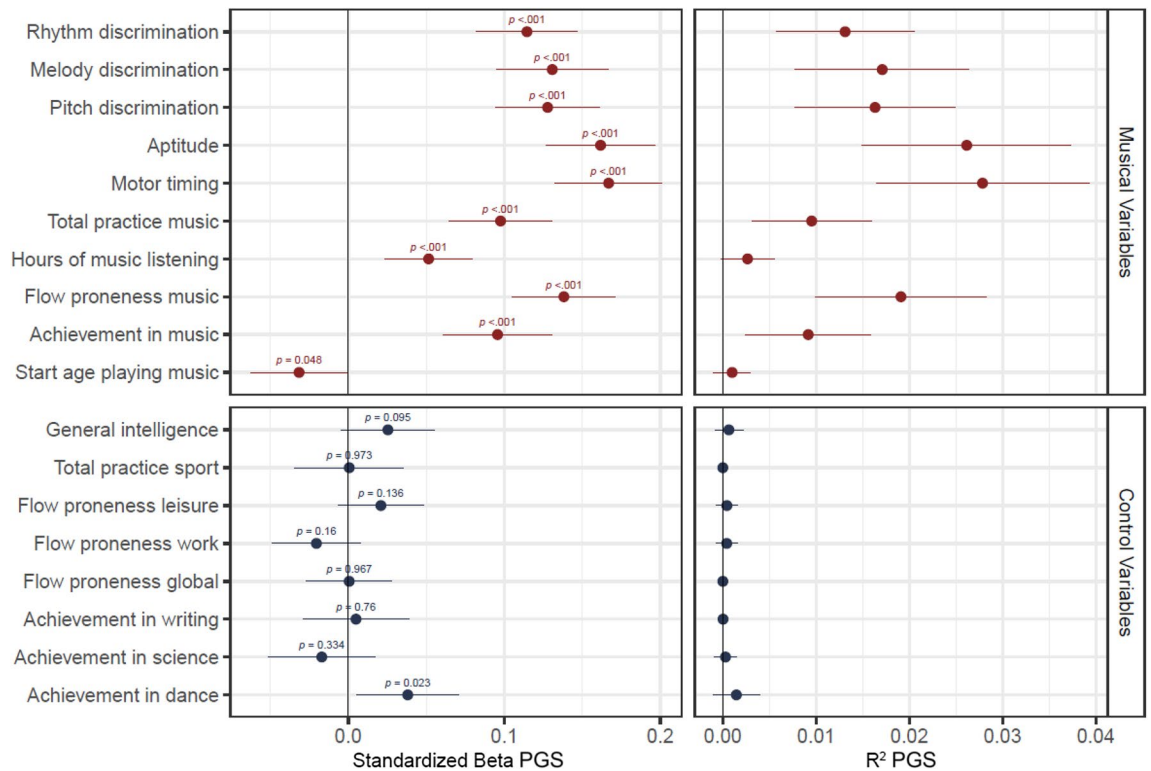
Figure 1. Effects of the PGS$_{rhythm}$ on music-related variables and control phenotypes. The association between PGS$_{rhythm}$ and start age playing music indicates that a higher PGS$_{rhythm}$ is associated with starting to play music at a younger age.

	Beta	Std. error	P-value	CI lower 95%	CI upper 95%
Rhythm discrimination					
Between	0.04	0.04	0.28	− 0.03	0.12
Within	0.15	0.07	0.04	0.02	0.28
Aptitude					
Between	0.06	0.03	0.28	− 0.04	0.16
Within	0.19	0.09	0.04	0.01	0.37

Table 1. Within- and between-family estimates of the effect of the PGS$_{rhythm}$ on levels of rhythm discrimination and musical aptitude in a sample of 243 complete dizygotic twin pairs.

such as total amount of music practice, music listening behavior, achievement in music, flow proneness in the music domain, as well as achievement in dance, a domain in which one would expect enhanced sensitivity to rhythm, and to music more generally.

Furthermore, discriminative validity of the PGS$_{rhythm}$ for music-related traits was demonstrated by the fact that we found no associations with traits of relevance for domain-general expertise (intelligence, general flow proneness) and creativity (in writing and science); nor with engagement in another control domain (sports). Overall, the data thus indicate that the PGS$_{rhythm}$ to a large part captures generalist genes underlying the music domain, rather than specialist genes only relevant to rhythmic ability. In other words, the underlying genetic variants essentially appear to tap into a complex of music-related traits that include overall musicality as well as a tendency to enjoy and engage with music, rather than only specifically reflecting a genetic predisposition to rhythmic propensity.

As already touched upon in the introduction, the discovery GWAS on which the PGS$_{rhythm}$ was based, identified genetic variants that are associated with whether individuals think they can clap in time with a musical beat or not. Although carefully validated, this binary self-report of beat synchronization ability is a very simplistic and minimal phenotype, as it is a broad 1-item yes/no question that is also likely to capture variation in general musical ability (e.g. whether an individual thinks that they are musical in general) and we cannot rule out some degree of rater bias in the GWAS sample. Niarchou and co-authors (2021) demonstrated that their derived polygenic score for beat synchronization was related to broad musician status in a health care context (OR = 1.33); however, no other musical phenotypes were available in their cohort to compare and determine specificity of the PRS[12]. It

would be expected, and has been shown for other traits like depression[50], that a risk score derived from a GWAS on a more specific, reliable, and objective measure of music skill would show higher discriminant validity for different music abilities (i.e. capture more specialist genes). Indeed, as mentioned in the introduction, using twin modelling in the same sample, we have previously found evidence for both specific genetic influences (specialist genes) on rhythm, pitch, and melody discrimination[51], and on perceptual and motor aspects of rhythmic skill[31], as well as factors shared between the different musical traits (generalist genes). This lends further support for the notion that the PGS$_{rhythm}$ captures more music generalist genes, rather than the genes specific to rhythmic ability.

The present findings are in line with recent empirical findings and models of expertise, which propose that a high level of domain-specific skill is associated with clusters of relevant traits across different modalities of individual differences, including abilities, as well as e.g. personality, physical trait, and interests[52]. Conceivably, the PGS$_{rhythm}$ thus taps into genes that pleiotropically influence a whole complex of music-related traits, which facilitate expertise acquisition within music and music-related domains, and include abilities as well motivational factors relevant to musical expertise (music flow, music-practicing behavior).

Pathways of the polygenic score for self-reported rhythm ability on music skills.
We further explored potential pathways through which the PGS$_{rhythm}$ affects music skills—direct, indirect and confounding effects—using a within-family design. If a PGS between unrelated individuals has a larger impact on a trait than the PGS has within dizygotic twin pairs, indirect genetic or confounding effects may be inflating the PGS effect[16,20]. As mentioned in the introduction, among indirect genetic effects are PGS$_{rhythm}$ effects that are mediated by the family environment due to passive GE correlation (prGE). However, we did not find evidence for that indirect (that is prGE) or confounding effects (e.g., population stratification or assortative mating) inflated of the association between PGS$_{rhythm}$ and musical aptitude and rhythm discrimination; specifically, the effects of the PGS$_{rhythm}$ on music skills were not significantly lower within dizygotic twin pairs. Previous studies have found that for several traits related to socio-economic status (mainly educational attainment), indirect genetic effects and passive gene–environment correlations do inflate the effects of PGSs, whereas this appears not to be the case for traits unrelated to educational attainment[16,18,19]. It could be that rhythm discrimination skills are less associated with educational attainment than other music-related behavior, like musical engagement.

Gene–environment correlation and potential confounding.
We found the level of musical enrichment of the family childhood environment to be associated with individuals' PGS$_{rhythm}$. This provides evidence for a role of gene–environment correlation for music skills, as individuals' genotypes correlate with their family environment. Although our within-family analyses suggest that prGE may not be relevant for the effect of the PGS$_{rhythm}$ on music skill, this correlation could reflect reactive or active gene–environment correlations. Reactive gene–environment correlation is when different conditions are evoked from the environment depending on genetically influenced phenotypes of the individual (e.g. musicality), and active gene–environment correlation will occur when genetic factors influence an individual's tendency to actively search out and create specific environmental conditions. Neither of these forms of gene–environment correlation can be disentangled with a within-family design.

We found a borderline significant trend for potential confounding effects, such as assortative mating or residual population structure effects, with a PGS$_{rhythm}$ correlation of exactly 0.50 between dizygotic twin members being the lower limit of the confidence interval (r = 0.54, CI 0.50–0.58). Assortative mating has been established for traits like height, body mass index[53], educational attainment[54] and psychiatric problems[55,56], but has also been indicated for personality traits and values[57]. Preferences for similarities with a mate, including music and art preferences has also been reported[58]. Although it seems plausible that assortative mating could occur based on levels of musicality or a tendency to enjoy and engage with music, Okbay et al.[27] recently showed for educational attainment that an inflated mate-pair PGS correlation most likely reflects other residual population structure effects or possibly more complex mating patterns rather than assortative mating. As the within-family analyses did not indicate an inflation due to prGE or confounding effects (which would include assortative mating) of PGS$_{rhythm}$ on music skills and the PGS$_{rhythm}$ has a relatively low prediction accuracy compared to the PGS of education attainment (which is based on a much larger discovery GWAS), the observed inflation of PGS$_{rhythm}$ correlation between dizygotic twins does not lend strong support for the presence of assortative mating effects.

Strengths, limitations and conclusion.
This is the first study to investigate the validity and specificity of the self-reported rhythm genotype in a large target sample of individuals that had completed well-validated music discrimination tests and a rich array of questionnaires about musical engagement. Moreover, it is the first study that investigates the genetic basis of music skills using a polygenic score approach, which is a substantial improvement to candidate gene approaches. However, despite using a powerful discovery GWAS and a large genetically informative sample, variance explained by the PGS$_{rhythm}$ for music outcomes is small. This is in line with PGS studies on other phenotypes. When the sample size of the GWASs increase, prediction accuracy of PGSs is expected to also increase with an expected r^2 of $h^2_{snp}/(1+(M/N \times h^2_{snp}))$, where h^2_{snp} is the heritability captured by the GWAS, M is the number of independent genetic variants and N is the discovery sample size[13]. Therefore, even larger sample sizes and better measures of (objective) music skills will surely be needed to increase the predictive value and accuracy of a musicality genotype in the future. A methodological challenge is that this requires harmonization of music related phenotypes across multiple genetically informative samples. Further, the availability of family data allowed us to explore potential pathways of direct genetic, indirect genetic and confounding effects of the genetic variants associated with self-reported rhythm. It is important to note that one should be careful when trying to isolate and interpret 'pure' direct genetic effects with sibling analyses, as measurement error due to genetic nurture effects can attenuate the associations[59]. Additionally, it is yet unknown

whether our results apply to rare genetic variants, since the GWAS and our polygenic scores are based on common genetic variants only. Lastly, findings from the within-family analyses are based on only complete dizygotic twin pairs, which resulted in a much smaller sample of 243 twin pairs. Lack of power could therefore cause a failure to detect a difference between within- and between-family effects and therefore we may have missed possible indirect and confounding effects of the PGS_{rhythm}.

To conclude, we investigated the validity and specificity of a self-reported rhythm genotype using the PGS approach. Our findings showed that the PGS_{rhythm} predicts overall musicality as well as a tendency to enjoy and engage in music including achievement in dance. The PGS_{rhythm} affects music skills and music-related behavior more likely through direct genetic effects rather than through genetic effects mediated by the family environment or confounding effects. Further, we found evidence for gene–environment correlation for music-related traits and the potential presence of assortative mating for self-reported beat synchronization ability. In this study, we showed that a polygenic score based on a GWAS that had been conducted on a minimalistic and broad music-related item can serve as a valid proxy for individuals' general genetic musical propensity, tapping into the genetic architecture underlying the broader music domain. Overall, our findings show that this PGS_{rhtyhm} can reliably be used to further disentangle gene–environment interplay underlying the broader musical trait complex, but it also lends hope that, as more genetically informative samples with minimal musical phenotyping become available, researchers in this emerging field may be able to join forces for a GWA-meta-analysis to identify additional generalist music genes and enhance our understanding of the genetic architecture of musicality.

Data availability
The datasets generated during the current study cannot be made publically as registry data and data from the company 23andme were used. Individuals are able to apply online at the Swedish Twin Registry to access the twin data. Our analyses code is available at https://osf.io/shqft/.

References
 1. Ullén, F., Mosing, M. A., Holm, L., Eriksson, H. & Madison, G. Psychometric properties and heritability of a new online test for musicality, the Swedish musical discrimination test. *Pers. Individ. Differ.* **63**, 87–93 (2014).
 2. Wesseldijk, L. W., Mosing, M. A. & Ullén, F. Gene–environment interaction in expertise: The importance of childhood environment for musical achievement. *Dev. Psychol.* **55**, 1473 (2019).
 3. Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R. & Ullen, F. Practice does not make perfect: No causal effect of music practice on music ability. *Psychol. Sci.* **25**, 1795–1803 (2014).
 4. Coon, H. & Carey, G. Genetic and environmental determinants of musical ability in twins. *Behav. Genet.* **19**, 183–193 (1989).
 5. Mosing, M. A. & Ullén, F. Genetic influences on musical specialization: A twin study on choice of instrument and music genre. *Ann. N. Y. Acad. Sci.* **1423**, 427 (2018).
 6. Seesjarvi, E. *et al.* The nature and nurture of melody: A twin study of musical pitch and rhythm perception. *Behav. Genet.* **46**, 506–515 (2016).
 7. Drayna, D., Manichaikul, A., de Lange, M., Snieder, H. & Spector, T. Genetic correlates of musical pitch recognition in humans. *Science* **291**, 1969–1972 (2001).
 8. Wesseldijk, L. W., Ullén, F. & Mosing, M. A. Does listening to music increase your ability to discriminate musical sounds?. *Pers. Individ. Differ* **161**, 110001 (2020).
 9. Wesseldijk, L. W., Mosing, M. A. & Ullen, F. Is there a sensitive period for musical skill acquisition during development? A genetically informative study. *Psychol. Sci.* **32**, 3–13 (2020).
10. Gingras, B., Honing, H., Peretz, I., Trainor, L. J. & Fisher, S. E. Defining the biological bases of individual differences in musicality. *Philos. Trans. R. Soc. Lond. B* **370**, 20140092 (2015).
11. Duncan, L. E. & Keller, M. C. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am. J. Psychiatry* **168**, 1041–1049 (2011).
12. Niarchou, M. *et al.* Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-022-01359-x (2022).
13. Abdellaoui, A. & Verweij, K. J. H. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nat. Hum. Behav.* **5**, 686–694 (2021).
14. Wray, N. R. *et al.* From basic science to clinical application of polygenic risk scores: A primer. *JAMA Psychiat.* https://doi.org/10.1001/jamapsychiatry.2020.3049 (2020).
15. Young, A. I., Benonisdottir, S., Przeworski, M. & Kong, A. Deconstructing the sources of genotype-phenotype associations in humans. *Science* **365**, 1396–1400 (2019).
16. Selzam, S. *et al.* Comparing within- and between-family polygenic score prediction. *Am. J. Hum. Genet.* **105**, 351–363 (2019).
17. Plomin, R., DeFries, J. C. & Loehlin, J. C. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychol. Bull.* **84**, 309–322 (1977).
18. Kong, A. *et al.* The nature of nurture: Effects of parental genotypes. *Science* **359**, 424–428 (2018).
19. Howe, L. J. *et al.* Within-sibship GWAS improve estimates of direct genetic effects. *Biorxiv* https://doi.org/10.1101/2021.03.05.433935 (2021).
20. Choi, S. W., Mak, T. S. & O'Reilly, P. F. Tutorial: A guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).
21. Kemper, K. E. *et al.* Phenotypic covariance across the entire spectrum of relatedness for 86 billion pairs of individuals. *Nat. Commun.* **12**, 1050 (2021).
22. Belsky, D. W. & Harden, K. P. Phenotypic annotation: Using polygenic scores to translate discoveries from genome-wide association studies from the top down. *Curr. Dir. Psychol. Sci.* **28**, 82–90 (2019).
23. Allegrini, A. G. *et al.* Multivariable G-E interplay in the prediction of educational achievement. *PLoS Genet.* **16**, e1009153 (2020).
24. Benyamin, B., Visscher, P. M. & McRae, A. F. Family-based genome-wide association studies. *Pharmacogenomics* **10**, 181–190. https://doi.org/10.2217/14622416.10.2.181 (2009).
25. Moore, A. J., Brodie, E. D. III. & Wolf, J. B. Interacting phenotypes and the evolutionary process: i. Direct and indirect genetic effects of social interactions. *Evolution* **51**, 1352–1362 (1997).

26. McGlothlin, J. W. & Brodie, E. D. III. How to measure indirect genetic effects: The congruence of trait-based and variance-partitioning approaches. *JSTOR* **63**, 1785–1795 (2009).
27. Okbay, A. *et al.* Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat. Gen.* **54**, 437–449 (2022).
28. Lichtenstein, P. *et al.* The Swedish Twin Registry: A unique resource for clinical, epidemiological and genetic studies. *J. Intern. Med.* **252**, 184–205 (2002).
29. Lichtenstein, P. *et al.* The Swedish twin registry in the third millennium: an update. *Twin Res. Hum. Genet.* **9**, 875–882 (2006).
30. Mosing, M. A., Butkovic, A. & Ullén, F. Can flow experiences be protective of work-related depressive symptoms and burnout? A genetically informative approach. *J. Affect. Disord.* **226**, 6–11 (2018).
31. Mosing, M. A., Verweij, K. J. H., Madison, G. & Ullén, F. The genetic architecture of correlations between perceptual timing, motor timing, and intelligence. *Intelligence* **57**, 33–40 (2016).
32. Wesseldijk, L. W., Ullén, F. & Mosing, M. A. The effects of playing music on mental health outcomes. *Sci. Rep.* **9**, 12606 (2019).
33. Madison, G. Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. *J. Exp. Psychol. Hum. Percept. Perform.* **27**, 411–422 (2001).
34. Carson, S. H., Peterson, J. B. & Higgins, D. M. Reliability, validity, and factor structure of the creative achievement questionnaire. *Creat. Res. J.* **17**, 37–50 (2005).
35. Mosing, M. A., Verweij, K. J., Abe, C., de Manzano, O. & Ullen, F. On the relationship between domain-specific creative achievement and sexual orientation in Swedish twins. *Arch. Sex. Behav.* **45**, 1799–1806 (2016).
36. Mosing, M. A. *et al.* Did sexual selection shape human music? Testing predictions from the sexual selection hypothesis of music evolution using a large genetically informative sample of over 10,000 twins. *Evol. Hum. Behav.* **36**, 359–366 (2015).
37. Formann, A. K. & Piswanger, K. *Wiener Matrizen-Test (WMT)* (Beltz Test, London, 1979).
38. Ullén, F. *et al.* Proneness for psychological flow in everyday life: Associations with personality and intelligence. *Pers. Individ. Differ.* **52**, 167–172 (2012).
39. Lam, M. *et al.* RICOPILI: Rapid imputation for consortias pipeline. *Bioinformatics* **36**, 930–933 (2019).
40. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Gen.* **48**, 1279–1283 (2016).
41. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Gen.* **38**, 904–909 (2006).
42. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Gen.* **47**, 291–295 (2015).
43. Altshuler, D. M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
44. Robinson, M. R. *et al.* Genetic evidence of assortative mating in humans. *Nat. Hum. Behav.* **1**, 0016 (2017).
45. Ni, G. *et al.* A comprehensive evaluation of polygenic score methods across cohorts in psychiatric disorders. *Medrxiv* https://doi.org/10.1101/2020.09.10.20192310 (2020).
46. Abdellaoui, A. *et al.* Gene–environment correlations across geographic regions affect genome-wide association studies. *Nat. Gen.* https://doi.org/10.1038/s41588-022-01158-0 (2022).
47. Williams, R. L. J. B. A note on robust variance estimation for cluster-correlated data. *Biometrics* **56**, 645–646 (2000).
48. Rogers, W. Regression standard errors in clustered samples. *Stata Tech. Bull.* **3**, 13 (1994).
49. Cannon, J. J. & Patel, A. D. How beat perception co-opts motor neurophysiology. *Trends Cogn. Sci.* **25**, 137–150 (2021).
50. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Gen.* **52**, 437–447 (2020).
51. Mosing, M. A., Pedersen, N. L., Madison, G. & Ullén, F. Genetic pleiotropy explains associations between musical auditory discrimination and intelligence. *PLoS ONE* **9**, e113874 (2014).
52. Ullén, F., Hambrick, D. Z. & Mosing, M. A. Rethinking expertise: A multifactorial gene–environment interaction model of expert performance. *Psychol. Bull.* **142**, 427–446 (2016).
53. Robinson, M. R. *et al.* Population genetic differentiation of height and body mass index across Europe. *Nat. Gen.* **47**, 1357–1362 (2015).
54. Hugh-Jones, D., Verweij, K. J. H., St. Pourcain, B. & Abdellaoui, A. Assortative mating on educational attainment leads to genetic spousal resemblance for polygenic scores. *Intelligence* **59**, 103–108 (2016).
55. Wesseldijk, L. W. *et al.* Spousal resemblance in psychopathology: A comparison of parents of children with and without psychopathology. *Eur Psychiatry* **34**, 49–55 (2016).
56. Nordsletten, A. E. *et al.* Patterns of Nonrandom mating within and across 11 major psychiatric disorders. *JAMA Psychiat.* **73**, 354–361 (2016).
57. Parshikova, O., Chertkova, Y., Alekseeva, O. & Kozlova, I. Assortative mating: Personality and values. *Pers. Individ. Differ.* **60**, S51 (2014).
58. Lee, A. J., Jones, B. C. & DeBruine, L. M. Investigating the association between mating-relevant self-concepts and mate preferences through a data-driven analysis of online personal descriptions. *Evol. Hum. Behav.* **40**, 325–335 (2019).
59. Fletcher, J., Wu, Y., Li, T. & Lu, Q. Interpreting polygenic score effects in sibling analysis. *Biorxiv* https://doi.org/10.1101/2021.07.16.452740 (2021).

Acknowledgements

Author contributions

Conceptualization, L.W.W., M.A.M., F.U.; Methodology, L.W.W., M.A.M., A.A.; Investigation, L.W.W., M.A.M.; Visualization: A.A.; Writing—Original Draft, L.W.W., M.A.M.; Writing—Review & Editing, R.L.G., A.A., F.U.; Funding Acquisition, F.U., R.L.G.

Funding

Competing interests

The authors declare no competing interests.

Additional information

23andMe Research Team

Stella Aslibekyan[9], Adam Auton[9], Elizabeth Babalola[9], Robert K. Bell[9], Jessica Bielenberg[9], Katarzyna Bryc[9], Emily Bullis[9], Daniella Coker[9], Gabriel Cuellar Partida[9], Devika Dhamija[9], Sayantan Das[9], Sarah L. Elson[9], Teresa Filshtein[9], Kipper Fletez-Brant[9], Pierre Fontanillas[9], Will Freyman[9], Anna Faaborg[9], Shirin T. Fuller[9], Pooja M. Gandhi[9], Karl Heilbron[9], Barry Hicks[9], Ethan M. Jewett[9], Katelyn Kukar[9], Keng-Han Lin[9], Maya Lowe[9], Jey C. McCreight[9], Matthew H. McIntyre[9], Steven J. Micheletti[9], Meghan E. Moreno[9], Joanna L. Mountain[9], Priyanka Nandakumar[9], Elizabeth S. Noblin[9], Jared O'Connell[9], Yunru Huang[9], Aaron A. Petrakovitz[9], Vanessa Lane[9], Aaron Petrakovitz[9], Joanne S. Kim[9], G. David Poznik[9], Morgan Schumacher[9], Anjali J. Shastri[9], Janie F. Shelton[9], Jingchunzi Shi[9], Suyash Shringarpure[9], Vinh Tran[9], Joyce Y. Tung[9], Xin Wang[9], Wei Wang[9], Catherine H. Weldon[9], Peter Wilton[9], Alejandro Hernandez[9], Corinna Wong[9] & Christophe Toukam Tchakouté[9]

[9]23andMe, Inc, Sunnyvale, CA, USA.