bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Unravelling the genetic architecture of musical rhythm: a large-scale genome-wide association study of beat synchronization

**Maria Niarchou[1,2*], Daniel E. Gustavson[1,2], J. Fah Sathirapongsasuti[3], Manuel Anglada-Tort[4], Else Eising[5], Eamonn Bell[6], Evonne McArthur[1], Peter Straub[1], The 23andMe Research Team[3^], J. Devin McAuley[7], John A. Capra[8], Fredrik Ullén[9], Nicole Creanza[10], Miriam A. Mosing[11,9], David Hinds[3], Lea K. Davis[2,1,13,14,15*†], Nori Jacoby[4†], Reyna L. Gordon[16,1,17,18*†]**

[1]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, TN, USA

[2]Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, TN, USA

[3]23andMe, Inc., Sunnyvale, CA, USA

[4]Computational Auditory Perception Research Group, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

[5]Department of Language and Genetics, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

[6]Department of Music, Trinity College Dublin, Dublin, Republic of Ireland

[7]Department of Psychology, Michigan State University, Michigan, USA

[8]Bakar Computational Health Sciences Institute and Department of Epidemiology & Biostatistics, University of California, San Francisco, CA, USA

[9]Department of Neuroscience, Karolinska Institutet, Sweden

[10]Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

[11]Melbourne School of Psychological Sciences, University of Melbourne, Australia

[13]Department of Biomedical Informatics, Vanderbilt University Medical Center, TN, USA

[14]Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, TN, USA

[15]Department of Molecular Physiology and Biophysics, Vanderbilt University, TN, USA

[16]Department of Otolaryngology – Head & Neck Surgery, Vanderbilt University Medical Center, TN, USA

[17]Department of Psychology, Vanderbilt University, TN, USA

[18]Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA

^full list of authors in Acknowledgements

*Correspondence to: Reyna L. Gordon (reyna.gordon@alumni.usc.edu), Lea Davis (lea.k.davis@gmail.com), Maria Niarchou (maria.niarchou@vumc.org)

†Co-senior authors

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.


## Abstract

Moving in synchrony to a musical beat is a fundamental component of musicality. Here, we conducted a genome-wide association study (GWAS) to identify common genetic variants associated with beat synchronization in 606,825 individuals. Beat synchronization exhibited a highly polygenic architecture, with sixty-seven loci reaching genome-wide significance ($p<5x10^{-8}$) and SNP-based heritability (on the liability scale) of 13%-16%. Heritability was enriched for genes expressed in brain tissues, and for fetal and adult brain-specific gene regulatory elements, underscoring the role of central nervous system biomarkers linked to the genetic basis of the trait. We performed validations of the self-report phenotype (through internet-based experiments) and of the GWAS (polygenic scores for beat synchronization were associated with patients algorithmically classified as musicians in medical records of a separate biobank). Genetic correlations with breathing function, motor function, processing speed, and chronotype suggest shared genetic architecture with beat synchronization and provide avenues for new phenotypic and genetic explorations.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Introduction

Our tendency to perceive, create, and appreciate rhythms in a variety of contexts (e.g., speech, music, movement) is a feature of the human experience[1–3]. Rhythmic patterns provide predictable and robust sensorimotor structure to human interactions[4,5], helping guide our attention to communicatively important moments in time[6,7]. Even very young children are sensitive to the social and linguistic signals carried by rhythm[8–10], thus it is not surprising that parents use rhythmic vocalizations and synchronous movement (e.g., lullabies and rocking) to interact with their infants from birth[11,12]. Rhythmic musical interactions in young children and across the lifespan[13] are structured around the percept of a stable periodic pulse, termed the "beat" in Western music and also present in music of other cultures[1,14] (though different musical cultures varies in the way they incorporate beats in musical structure; see[15,16]). While music in general and rhythmic structures in particular significantly vary from one culture to another[16–18], there is preliminary evidence that hierarchical beat structure of most music is robust to cultural transmission[2] and indeed common in many types of music[1].

*Beat perception and synchronization* (i.e. perceiving, predicting, and moving predictively in synchrony to a musical beat[19]) is an important feature of musical experiences across many human cultures and musical genres[1,20]. The predictive temporal mechanisms afforded by beat structure enhance general perceptual and learning processes in music, including melody perception and production, singing, and joint music-making[3,6]. Recent work showed that some features of rhythm perception and production (e.g., categorical rhythm perception) varies across listeners from different cultures[14,20–22], the same studies showed considerable consistencies across cultures for

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

other features (for example preference for beat-based isochrony). Beat synchronization, and musicality in general, appear to have broad implications for society by supporting pro-social behavior[11,23] and well-being[24]. Many have proposed that beat perception and synchronization evolved in humans to support communication and group cohesion[19,25–27].

Neuroimaging findings have highlighted auditory-motor networks in the brain underlying rhythm perception and production[28], during which there is precise entrainment of neural oscillatory activity to musical signals, primarily involving motor planning areas and auditory regions of the brain, even during simple passive listening to music[29]. Neural mechanisms of entrainment, prediction, and reward work in concert to coordinate the timing of beat-related expectancies to musical signals during listening, playing, singing, and dance[27].

Although most people are able to effortlessly detect and synchronize with the beat even without musical training[4,19], there is still substantial inter-individual variability, *within* cultures, in the extent to which individuals can perceive and produce musical rhythms accurately[30,31], including the temporal precision of coordinating movement with a musical beat[32]. These individual differences could be due in part to genetic variation, and thus genetic approaches can be used to gain a foothold on the biological basis of musicality in humans (with the definition of *musicality* broadly encompassing musical behavior, music engagement and musical skill per[25]). Indeed, twin-modelling and other family-based studies point to moderate heritability of rhythm-related traits such as duration discrimination[33,34], rhythm discrimination[35], isochronous sensori-motor production[36], and off-beat detection[37]. Much less is known at the molecular level about

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

human genome variation underlying musicality[38] which to date has been investigated in relatively small samples[34], due to the challenge of assessing musicality phenotypes in samples large enough to provide sufficient power to detect common variants with small effects (as expected for complex traits[39]). Large-scale genome-wide association studies (GWAS) of musicality traits are thus needed to advance this field.

Our understanding of the biological underpinnings of beat synchronization, from its genetic architecture to its neural instantiation and behavioral manifestation, requires complex multi-methodological approaches. For instance, post-GWAS approaches (i.e., heritability enrichment of gene expression in central nervous tissues) can eventually be used to illuminate the relationship between the genetic architecture of music-related traits and patterns of neural activity reported in neuroscience studies.

Human tendencies to engage with music are beyond recreational; individual differences in beat perception and synchronization abilities and the strength of neural resources associated with rhythm are predictive of language and literacy skills[40,41] and are more generally related to cognition, motor function, and social coordination[42]. The underlying genetic architecture of beat synchronization may have clinical-translational relevance, both for characterizing risk (i.e. atypical or impaired rhythm appears to be associated with developmental speech-language disorders[43]) and for rhythm-based rehabilitation (i.e. for stroke and Parkinson's[44]; with other promising benefits of music on health under investigation[45]). Applying advanced genetic epidemiology methods[46] to musicality phenotypes data collected in large population samples will generate new avenues of research, allowing us to examine to what extent beat synchronization might

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

share genetic architecture with other traits, thus also highlighting common risk or protective effects with other health and cognitive traits.

*Summary of Approach.*

Here, we report a novel genome-wide interrogation of beat synchronization. Our approach was as follows: 1) We validated a subjective self-reported beat synchronization item ("Can you clap in time with a musical beat?", referred to in this paper as the "target item"), in relation to measured beat synchronization and rhythm perception task performance, as a scalable and accurate phenotype for large-scale population-based studies. 2) We identified genomic loci and established a polygenic model of beat synchronization. We first conducted a genome-wide association study (GWAS) to identify common genetic variants associated with beat synchronization in 606,825 individuals participating in research with the personal genetics company 23andMe, Inc. We then validated the findings in a separate genetic sample associated with musicality in a health care context (in Vanderbilt's BioVU database), by testing whether a cumulative sum of the genetic effects for beat synchronization detected in our GWAS (i.e., polygenic score), was significantly associated with algorithmically identified musicianship. We also estimated how much of the total phenotypic variance of beat synchronization could be explained by all variation across the genome (i.e., SNP-based heritability). 3) We further investigated the genetic architecture of beat synchronization, using partitioned heritability and gene set enrichment analyses, with a particular focus on genes that play a role in the central nervous system. We also evaluated the contribution to beat synchronization of genomic regions that have experienced

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

150 significant human-specific evolutionary shifts (since the divergence of humans and
151 chimpanzees from their last common ancestor). 4) We explored shared genetic effects
152 (pleiotropy) on beat synchronization and other traits, performing exploratory genetic
153 correlation analysis including 764 complex traits for which a well-powered GWAS has
154 been performed and deposited in LDHub[47]. We then used genomic Structural Equation
155 Modelling (SEM) to characterize the relationship among the top associations between
156 beat synchronization and health-related traits. Findings were further investigated on the
157 phenotypic level in a separate experiment.
158

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Results

### *Overview. Validating the self-reported beat synchronization phenotype*

In light of prior work suggesting that musicality and rhythm skills are complex traits that can be quantified with both objective (experiment-derived) assessment and subjective self-reported data[48,49], we performed a series of validations of the self-report beat synchronization item (i.e., the single item “Can you clap in time with a music beat?” that was used in the genetic study), in relation to measured rhythm perception and beat production tasks. Both studies were administered in English for consistency. We also explored the relationship between task-based beat synchronization ability, a self-reported rhythm scale (from additional questionnaire items), and musicality. Study overviews and key results are summarized in Figure 1.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

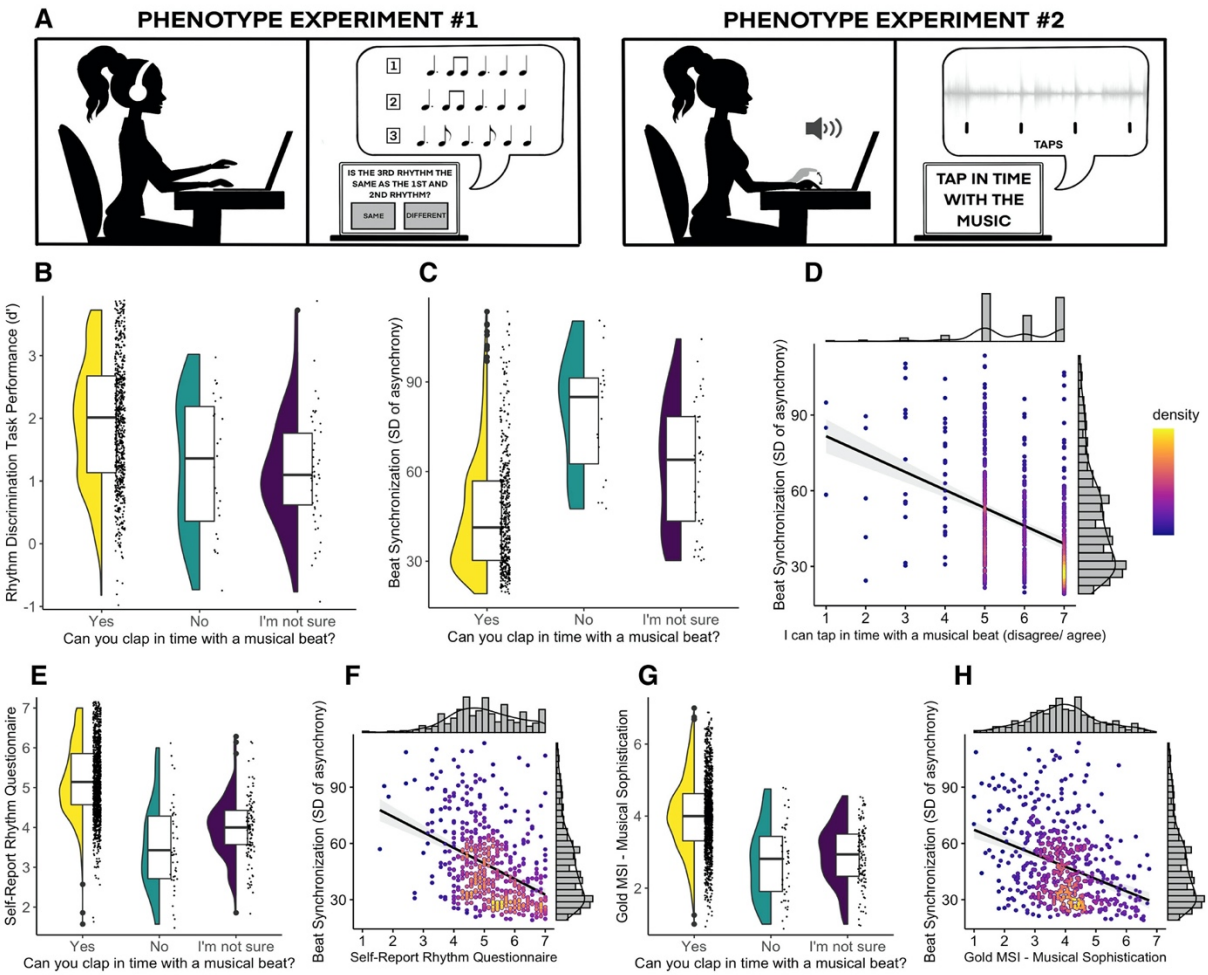



**Figure 1. Phenotype validation studies overview and results.** *A) Schema of internet-based phenotype validation studies.* In phenotype experiment #1, participants performed a musical rhythm perception test and provided self-report of the same target question in the GWAS study ("Can you clap in time with a musical beat"). In phenotype experiment #2, participants performed beat synchronization tasks (which involved tapping to the beat of musical excerpts) as well as responding to the same target question, in addition to a series of other questionnaires about their musical engagement/ability and health traits. B) Phenotype Experiment 1 results show rhythm perception task performance in association with responses to GWAS target question in N=724. C-H): Phenotype Experiment 2 results. C) Beat synchronization task performance in association with responses to the target question in n=542. D) Beat synchronization task performance in association with responses to a similar self-report question asked on a Likert scale, in n=542. E) Self-reported rhythm questionnaire (7-item scale) in association with responses to the target question in N=1,412. F) Beat synchronization task performance in association with Self-reported rhythm questionnaire in n=542. G) Gold-MSI (musical sophistication) in association with responses to the target question in N=1,412. H) Beat synchronization task performance in association with Gold-MSI in n=542.

*Phenotype Experiment 1: Rhythm perception task performance.*

Phenotype experiment #1 was conducted in N=724 participants recruited anonymously in Amazon's Mechanical Turk (see Table 1 for demographics).

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Participants were asked the target question “Can you clap in time with a musical beat?” and performed a musical rhythm perception test (Supplementary Figure 1). See Methods and Supplementary Notes for experiment details. Briefly, in each of the 32 trials of the task, participants had to judge whether a pair of rhythms were the same or different (see Figure 1A), following a standard procedure for assessing individual differences in musical perception ability[31] and utilizing rhythm sequences with simple (highly metrical) and complex (syncopated) rhythms[50]. The rhythm perception task yielded quantitative scores.

Individuals with better performance in the rhythm perception test (higher total *d'*) were more likely to answer Yes to the target item (i.e. that they can clap to the beat: OR(95%CI)=1.94 [1.28 to 3.01], p=0.002, McFadden's $R^2$=0.39), indicating there is approximately a 94% increase in the odds of answering 'Yes', per standard deviation increase in the rhythm perception test (see Figure 1B). In addition, individuals with higher scores on the subset of "simple" (i.e., more strongly beat-based) rhythm trials were more likely to answer that they can clap to the beat (OR(95%CI)=1.99[1.36-2.90], p<0.001, McFadden's $R^2$=0.40.

*Phenotype Experiment 2: Beat synchronization task performance*

We conducted a second internet-based phenotype experiment (pre-registered in OSF) with N=1,412, to validate self-reported beat synchronization phenotype as a proxy for directly-measured beat synchronization ability (see Table 1 for demographics). Participants completed a questionnaire on musicality, health, and personality items, and were asked to tap in real time to the musical beat of 4 different musical excerpts (see

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Figure 2). Participants completed a questionnaire on musicality, health, and personality items, and were asked to tap in real time to the musical beat of 4 different musical excerpts (see Supplementary Figure 2). Beat synchronization tapping accuracy was assessed similarly to lab-based studies[32], but with a recently developed online-based technology that allowed us to precisely measure asynchrony of participants' taps along to music clips - i.e., REPP (Rhythm ExPeriment Platform; see[51]) for additional details on the experiment and pre-registered hypotheses (H1-H6), see Methods and Supplementary Notes. Key results of this study are summarized in Figure 1 and Supplementary Table 1. Note that better *tapping accuracy* is reflected in lower *tapping asynchrony* scores, i.e., more accurate timing of taps in relation to the beat.

First, we tested pre-registered H1, showing that individuals who respond Yes to the target question (i.e., "Can you clap in time with a musical beat") had lower tapping asynchrony, OR = 0.28, [0.18, 0.43], McFadden $R^2$ = .67 (Figure 1C). Tapping asynchrony was also negatively correlated with responses to a highly similar item ("I can tap in time to a musical beat") when asked on a seven-point Likert agreement scale (1= disagree; 7 = agree) instead of a "yes/no" answer, $r$ = -.40, 95% CI [-0.47, -0.33] (H1a; Figure 1D). Similarly, individuals with significantly better self-reported rhythmic ability (based on responses to a seven-item questionnaire) were much more likely respond "yes" to the target question, OR = 7.34, 95% CI [4.79, 11.23], McFadden $R^2$=.34 (Figure 1E), and had lower tapping asynchrony, r = -.41, 95% CI [-0.47, -0.33] (Figure 1F) (H2). The results for H5 show that Musical sophistication scores from the Gold-MSI were also positively associated with the target question, OR=4.16, 95% CI [2.86, 6.04] (Figure 1G) and negatively correlated with tapping asynchrony r= -.34, 95%

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CI [-0.42, -0.27] (Figure 1H). There was also no evidence that Musical Sophistication or prior/current musician status (measured with an additional item) interacted with the tapping asynchrony to predict responses to the target question (H6).

In addition, we demonstrated that although responses to the target question were associated with confidence judgements of one's own tapping performance assessed either immediately after the tapping trials, OR = 1.72, 95% CI [1.07, 2.76], or confidence assessed as a personality trait, OR = 1.75, 95% CI [1.02, 3.00], controlling for these confidence measures had no impact on the association with tapping asynchrony described above, OR = 0.28, 95% CI [0.17, 0.44], McFadden $R^2$=.69 (H3). These findings suggest that while the target question may encompass some self-reporting bias, the bias does not diminish its strong association with true beat synchronization ability. Moreover, all associations reported here were not altered when controlling for age, sex, and education (see Supplementary Table 1).

## Genomic study population

### Beat Synchronization GWAS sample.

The study population for the discovery GWAS consisted of N=606,825 unrelated participants of European ancestry (see Table 1 for demographics), who consented to participate in research with 23andMe, Inc. and answered Yes or No to the target question "Can you clap in time with a musical beat?" The majority of participants answered 'Yes' (91.57%) and 8.43% answered 'No'. See Methods for further sample details, genotyping, and quality control,.

### GWAS results and heritability estimation.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GWAS was conducted using logistic regression under an additive genetic model, while adjusting for age, sex, the first five principal components from genetic data, and genotype platforms (see Methods). Sixty-eight "sentinel" SNPs (after two rounds of LD pruning, first at $r^2$=0.6 and then at $r^2$=0.1, kb = 250) at 67 genomic loci (locus 28 contains two independent sentinel SNPs) reached genome-wide significance ($p<5x10^{-8}$; Figure 2, Table 2, and Supplementary Table 2), with a total of 6,115 SNPs passing the genome-wide significance threshold. QQ-plot is provided in Supplementary Figure 3. Local association plots at each locus are provided in the Regional Plots supplemental document.

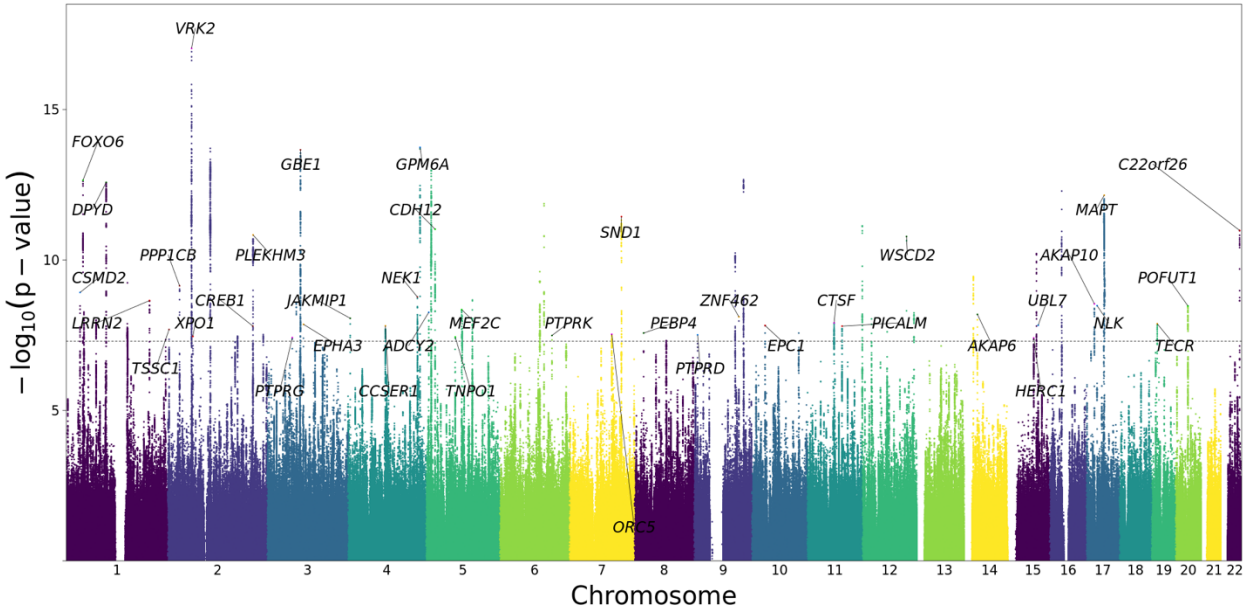



**Figure 2. Manhattan plot of GWAS results of beat synchronization.** Results are shown for 606,825 research participants with 23andMe who responded Yes (N=555,660) vs. No (N=51,165) to the question "Can you clap in time with a musical beat?", controlling for age, sex, top 5 PC's for ancestry, and genotype platform. The x-axis shows chromosomal position and the y-axis shows -log10 p-values). Sixty-seven loci (68 sentinel SNPs, with one locus containing two independent sentinel SNPs) surpassed the threshold for genome-wide significance of $p<5x10^{-8}$ (dotted horizontal line). For illustration purposes, only 500,000 SNPs with p<0.1 are shown; gene symbols for sentinel SNPs are notated when FUMA provided a gene mapped to nearest sentinel SNP.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The top associated locus was mapped at chromosome 2 close to *VRK2* (Vaccinia Serine/Threonine Kinase 2) and *FANCL* genes (rs848293, $p=9.2 \times 10^{-18}$), a region that has been linked to sleep, depression[52] and schizophrenia[53]. Another strongly associated locus was around a SNP on chromosome 17 (rs4792891, $p=7.07 \times 10^{-13}$) that maps to the Microtubule Associated Protein Tau (*MAPT*) gene, well-known for its associations with Parkinson's disease[54]. There was also a locus on the Mitogen-Activated Protein Kinase 3 (*MAPK3*) gene at 16p11.2, a region linked to risk for a number of neurodevelopmental disorders including autism and schizophrenia[55], and to verbal memory and language[56]. Another significantly associated gene was the chromosome 4 Glycoprotein M6A (*GPM6A*), whose gene promoter contains a transcription factor binding site for *GATA2*, a gene previously related to music phenotypes[34].

Linkage Disequilibrium Score Regression (LDSC)[57] analyses revealed that heritability estimates on the liability scale ranged from 13% to 16% when adjusted for a range of estimated population prevalence of poor rhythm skills (from 3.0% to 6.5%; Supplementary Table 3; see Supplementary Notes for explanation of prevalence estimates). The observed (unadjusted) genetic variance explained 5% (se=0.0002) of the phenotypic variance in the beat synchronization trait, with an LD score regression intercept of 1.02 (se=0.01).

**Gene-based analyses.**

Gene-based association analyses performed with MAGMA v1.08 yielded 125 genes that surpassed the threshold of $p<2.6 \times 10^{-6}$ (Supplementary Table 4). The top two

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

301 genes are: *CCSER1,* in proximity to genes previously associated with musicality[58], and
302 *VRK2* (converging with the top locus identified in our SNP-based association analyses).
303 We also examined potential replication of genetic associations with musicality in
304 humans from prior reports (29 genes were selected: 26 reported in[58] plus *GATA2* and
305 *PCDH7*[34] and *UGT8*[59]). None of the genes reached statistical significance after
306 genome-wide correction ($p<2.6x\ 10^{-6}$; Supplementary Table 5, Supplementary Notes),
307 neither independently, nor as a gene-set (p=0.30), however, several are located near
308 *CCSER1* in the 4q22-24 region.
309 **Gene set enrichment analyses**
310 We performed gene-set enrichment analyses using MAGMA, implemented in
311 FUMA[60]. As hypothesized, associations with beat synchronization were significantly
312 enriched among genes expressed in brain tissue compared to those expressed in non-
313 brain tissues (Figure 3A); tissue-expression gene sets were based on GTEx data on 53
314 different tissue types[61]. To further examine potential biological pathways associated with
315 beat synchronization, we performed GO term and pathway enrichment analyses[62,63].
316 Three gene-sets out of 15,496 achieved statistical significance after Bonferroni
317 correction (Supplementary Table 6). Results were associated with nervous system
318 function, specifically: gene sets for synaptic membrane adhesion ($p=2.55\ x\ 10^{-8}$) and
319 synaptic adhesion-like molecules ($p=1.90x10^{-6}$), and MeCP2 regulation of transcription
320 factors ($p=1.17x10^{-6}$).
321
322 **Heritability Partitioning**

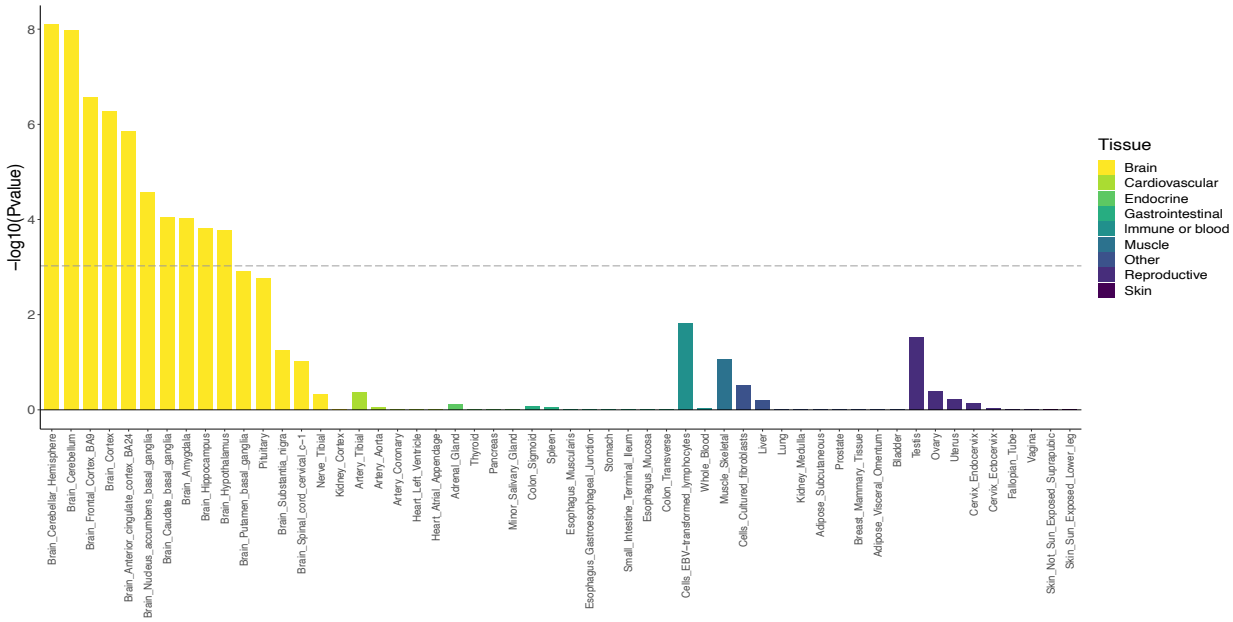
bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

One advantage to SNP-based heritability estimation is the ability to partition heritability according to genomic properties, which provides insight into the types of genetic variation that contribute most to beat synchronization. To determine whether the heritability of beat synchronization is enriched for specific functional categories, stratified LDSC[64] was used to partition heritability (Supplementary Table 7). We hypothesized that SNPs falling into open chromatin regulatory regions (i.e., accessible to transcriptional machinery), and regions with human-specific variation, would be enriched for beat synchronization-associated variation. When assessing broad SNP annotation categories[65], we found enrichment in regions conserved in mammals (regions of the genome identified under purifying selection[66]) (enrichment=15.8, $p=1.19 \times 10^{-12}$) and in regulatory regions marked by acetylation of histone H3 at lysine 9 (H3K9ac; generally considered a marker for active chromatin; enrichment=8.0, $p=1.85 \times 10^{-8}$) and monomethylation of histone H3 at lysine 4 (H3K4me1; generally considered a marker for enhancers; enrichment=1.29, $p=2.16 \times 10^{-5}$), supporting associations mediated by effects on gene regulation.

Given the results of the broad categories showing enrichment in markers of enhancers and active chromatin, we further used LDSC-Specifically Enriched Genes (LDSC-SEG) heritability partitioning as described in[67] to examine whether genes expressed in specific cell- or tissue-types (conditional to the other annotations) would show enrichment for beat synchronization-associated variants. When assessing tissue-specific annotations of active chromatin and enhancers (marked by H3K9ac, H3K27ac, DNase hypersensitivity sites and H3K4me1), we found that heritability was enriched in central-nervous-system- and skeletal muscle-specific regulatory regions; see

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

346 Supplementary Table 8 and Supplementary Figures 4, 5 and 6, for all cell-type specific,
347 multi-tissue chromatin, and multi-tissue gene expression results, respectively.
348 Enrichment in brain-specific regulatory elements, in several fetal and adult tissue-
349 specific elements as well as CNS-specific cell cultures, are shown in Figure 3B.
350
351




352
353 **Figure 3A. Genes associated with beat synchronization are enriched for expression in brain tissue.** Results of
354 MAGMA gene-property analysis are based on gene expression levels from GTEx in 53 tissues. Associations with
355 beat synchronization were significantly enriched in brain-expressed genes (-log-10 p-values are on the y-axis, with
356 tissue type on the x-axis). Dotted line shows p-value threshold for significant enrichment after Bonferroni correction
357 for testing 53 tissues.
358

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

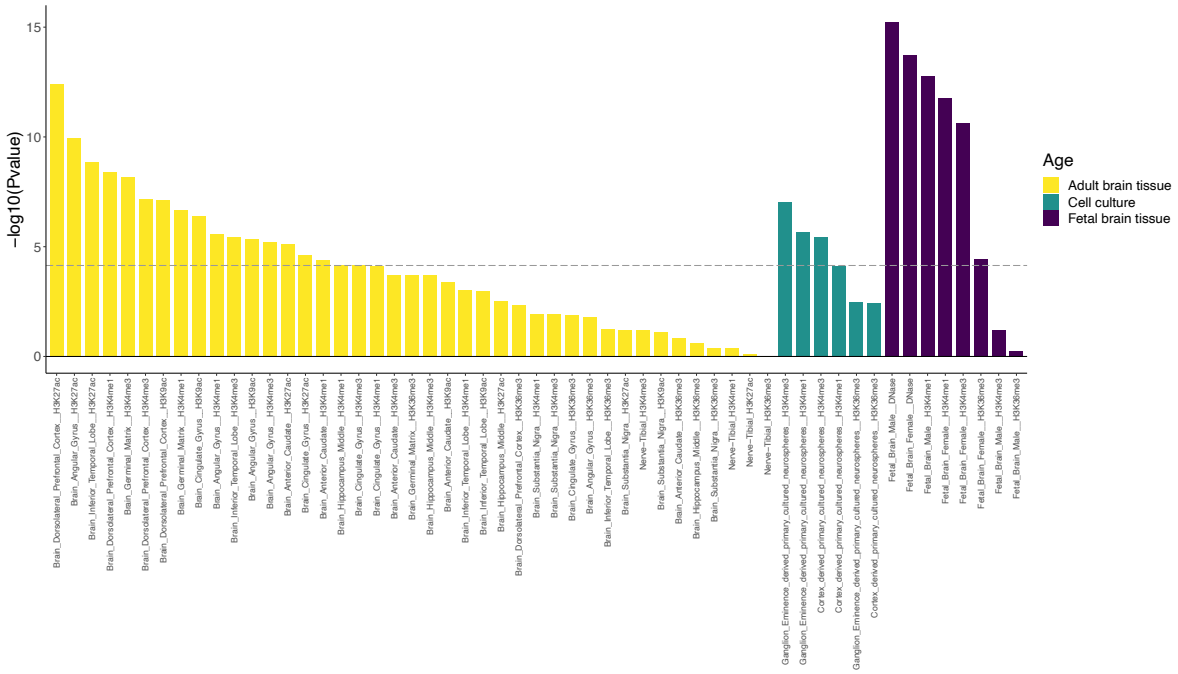



**Figure 3B. Partitioned heritability shows enrichment in brain-specific regulatory regions of the genome**. Partitioned heritability analysis was performed with LDSC-SEG. Tissue-specific regulatory elements are marked by histone 3 acetylation or DNase hypersensitivity (for open chromatin) and H3K4me1 (for enhancers). Regulatory regions in adult brain tissues are shown in yellow, with regulatory elements in in cell cultures in teal, and in fetal brain tissue shown in dark purple. The graph shows -log-10 p-values are on y-axis, with tissue and marker type on x-axis. The dotted line shows p-value threshold for significant enrichment after Bonferroni correction for number of gene sets tested.

## Human Accelerated Region and Neanderthal Introgression Stratified Heritability Analyses

Given evolutionary hypotheses about the origins of rhythm[4,19,68], we evaluated the contribution of regions of the human genome that have experienced significant human-specific shifts in evolutionary pressure, using stratified LDSC[57,64]. In particular, we analyzed the contribution to beat synchronization heritability from variants in genomic loci that are conserved across non-human species, but have elevated substitution rate on the human lineage[69]. Many of these human accelerated regions (HARs) play roles in human-enriched traits[70], including cognition[71]. Two of the variants significantly ($p < 5x10^{-8}$) associated with beat synchronization (rs14316, rs1464791) fall

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

within HARs (occurring within our locus 66 and locus 20, respectively). This is 11.2 times more overlap than expected by chance (μ = 0.178 overlaps; $p$ = 0.017, based on 10,000 permutations). The rs1464791 variant is near *GBE1*, a gene associated with neuromuscular disease[72], reaction time[73] and cognitive deficits[74]. Applying LDSC to consider the full set of association statistics, we find that genetic variants in HARs contribute 2.26 times more to the observed heritability of beat synchronization than would be expected if heritability were distributed uniformly across variants ($p$ = 0.14). Given the small number of common variants within HARs, this stratified heritability analysis is substantially underpowered (0.17% of variants considered are in HARs). The general agreement of these two approaches supports the enrichment of functional variation relevant to beat synchronization in HARs.

We also evaluated the contribution of genetic variants detected in the Neanderthal genome present in modern Eurasians due to interbreeding (hereafter "Neanderthal variants") to the heritability of the beat synchronization phenotype. Eurasian genomes contain ~1.5-4% of DNA as a result from interbreeding with Neanderthals around 50,000 years ago. Heritability of beat synchronization was significantly depleted among Neanderthal variants (1.97-fold depletion, $p$ = 0.001). However, Neanderthal ancestry is significantly depleted in functional genomic regions overall[75]; therefore, the depletion of beat synchronization heritability in these regions is likely the result of the overall depletion for Neanderthal ancestry in functional regions of the genome. This is supported by a non-significant $\tau_c^*$, illustrating that Neanderthal vs. human variants do not provide unique heritability when conditioned on a broad set of regulatory elements (Supplementary Table 9**,** Methods).

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

401

402 **Polygenic scores for beat synchronization are related to musicality reported in a**
403 **health care context**

404 We investigated whether the polygenic model of beat synchronization would be
405 associated with musicality, by examining whether a weighted sum of the common
406 alleles associated with beat synchronization, based on the GWAS results (also known
407 as polygenic scores [PGS]), differentiated musicians from non-musicians.
408 Musicians were drawn from recent study[76] that had algorithmically identified musically
409 active patients by the presence of specific musician-related keywords/regular
410 expressions in clinical documentation collected in the electronic health record (see
411 Methods and Supplementary Notes for details of automated search terms and chart
412 review validation details). Briefly, 1,753 musician cases (who also had linked genetic
413 data on file in the BioVU biobank) were compared with 65,147 genotyped controls. The
414 PGS for beat synchronization was significantly associated with musician status (OR per
415 SD increase in PGS, 1.26, 95%CI:1.20-1.33, $p<2 \times 10^{-16}$, Nagelkerke's $R^2$=1%)
416 (Supplementary Table 10, Figure 4), consistent with our hypothesis that the beat
417 synchronization phenotype captures a dimension of musicality. As expected, we did not
418 find evidence for associations of the beat synchronization PGS with the negative control
419 phenotypes (i.e., iron deficiency anemias, carcinoma in situ of skin, and cancer of the
420 brain).

421

422
423
424

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

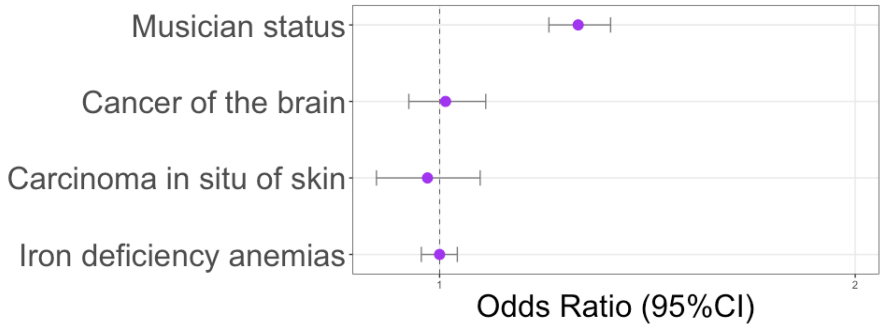



**Figure 4. Polygenic score for beat synchronization predicts musicianship in electronic health record (EHR) biobank.** With weights derived from the 23andMe beat synchronization GWAS, we applied polygenic scores (PGS) for beat synchronization to genomic data from N=1,753 musicians identified in Vanderbilt's BioVU compared to a control sample of N=65,147 (See Supplementary Notes). PGS-beat synchronization was associated with musician status (OR=1.26) but not with the three negative-control traits.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Cross-trait analyses

*Genetic correlations.* To determine if beat synchronization shares genetic architecture with other traits, we tested genetic correlations[57] between beat synchronization and all 764 available traits in LDHub (v.1.9.2) using LDscore regression. This method is designed to show whether there is shared genetic variation linked to a particular trait (here, our beat synchronization trait) and traits measured in other GWAS studies. There were 31 statistically significant genetic correlations ($p<6.5 \times 10^{-5}$) between beat synchronization and other traits after adjusting for multiple comparisons (Figure 5A, Supplementary Table 11). Briefly, top associations with beat synchronization included grip strength (rg(left)=0.18, $p=3.6 \times 10^{-16}$, rg(right)=0.16, $p=6.91 \times 10^{-15}$); peak expiratory flow from both the UKBiobank (rg=0.15, $p=2.11 \times 10^{-9}$) and a second independent GWA study (rg=0.11, $p=6.6 \times 10^{-8}$), shortness of breath when walking on level ground (rg = -0.16, $p=0.43 \times 10^{-5}$), and several other breathing/lung function phenotypes (Supplementary Table 11). Processing speed measured as 'mean time to correctly identify matches' was negatively genetically correlated with beat synchronization ability (rg=-0.16, $p=3.22 \times 10^{-13}$) (i.e., faster processing speed was associated with better beat synchronization). Additionally, smoking including 'ever smoked' (rg=0.16, $p=2.5 \times 10^{-11}$) and 'past tobacco smoking' (rg=-0.15, $p=4.6 \times 10^{-10}$), educational qualifications (O' levels/GCSEs or equivalent) (rg=0.16, $p=4.6 \times 10^{-7}$), evening chronotype (rg=0.09, $p=3.8 \times 10^{-5}$) and tinnitus (rg=0.20, $p=6.7 \times 10^{-6}$) were all positively associated with beat synchronization. While falling short of the correction for multiple testing, exposure to loud music was also correlated with a similar point estimate

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

458 (rg=0.20, p=2.0 x $10^{-4}$) and could be due to a relationship between tinnitus and loud
459 music exposure in the UKBB ($r_g$=0.30, p=4.8 x $10^{-6}$)[77,78].

460 *Genomic Structural Equation Modeling (SEM).* Next, we conducted Genomic
461 SEM[79] to examine whether the primary associations resulting from the exploratory
462 analyses represented distinct genetic correlations with beat synchronization and/or a
463 common set of genetic influences between beat synchronization and its top-associated
464 traits, some of which are also known to be related among each other in prior research
465 (e.g., musculoskeletal strength, lung function, and processing speed[80–82]). Analysis
466 details are reported in the Supplementary Notes. Briefly, we included the four most
467 significant traits from the LDHub analyses (grip strength, processing speed, smoking,
468 lung function), creating latent factors where possible using other GWAS from these
469 categories that were also significant in LDHub analyses (e.g., “Ever smoked”, “Smoking
470 status: Previous”, “Light smokers: At least 100 smokes in lifetime”). The final model,
471 displayed in Figure 5B, suggested that associations between beat synchronization,
472 reaction time, grip strength, and lung function were explained by a common set of
473 genetic influences, accounting for 16% of the total variance in the beat synchronization
474 GWAS (the squared factor loading of 0.40 on the Common factor). Beat synchronization
475 was genetically associated with smoking through a separate genetic association
476 ($r$=0.16; 95% CI [0.10, 0.23]), as smoking was uncorrelated with the other factors (see
477 Supplementary Notes for more information).
478

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

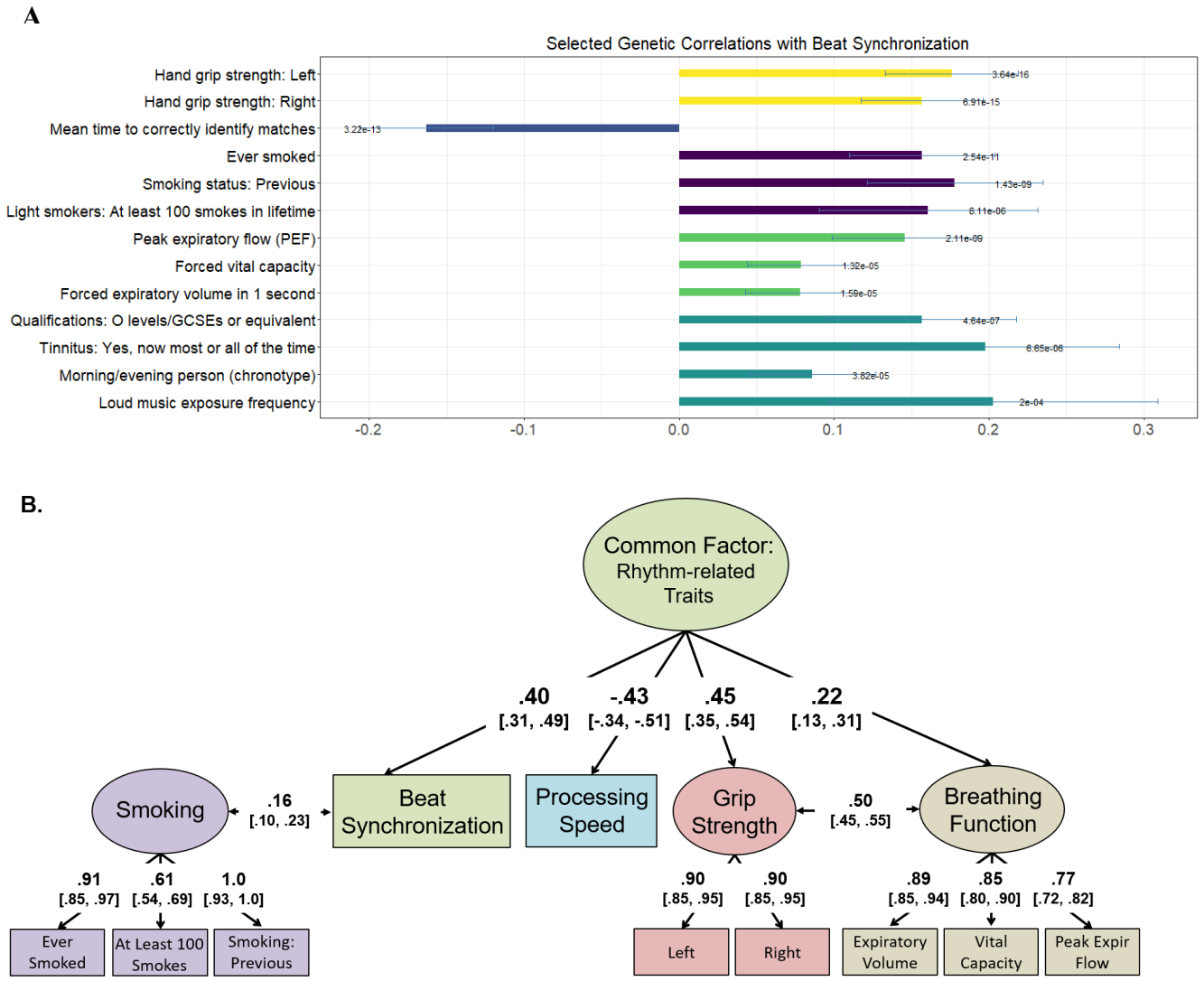



**Figure 5. Selected results from the cross-trait genetic correlations analysis.** Beat synchronization summary statistics were significantly genetically correlated with several health, cognitive, emotion, and circadian phenotypes in our analysis comparing beat synchronization to traits available in LDHub. **(A)** The most significant genetic associations from exploratory analyses of LDHub traits, along with other selected traits. The y-axis is the rg correlation, the bars represent standard errors and the p-values are also indicated. Full results are presented in Supplementary Table 11. **(B)** The best-fitting genomic structure equation model of beat synchronization with GWAS of grip strength, breathing, smoking, and processing speed from LDHub. 95% confidence intervals of factor loadings and correlations are displayed in brackets. Results suggest that beat synchronization was associated with the other traits (except smoking) through a set of common genetic influences. Model fit: $\chi^2$ (31) = 7136.92, $p$ < .001, CFI = .982, SRMR = .041.

*Common Factor GWAS: Rhythm-Related Traits.* Using genomic SEM, we conducted a multivariate GWAS on the latent genetic factor from the Common variance in the model presented above and portrayed in Figure 5B, after excluding smoking from

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the model and restricting to one set of summary statistics per domain (beat synchronization, processing speed, peak expiratory flow, grip strength [left]). The heritability of this latent genetic factor was .1068 (s.e.=0.0038). There were 270 independent genome-wide significant sentinel SNPs, 97 lead SNPs, and 80 genomic risk loci (Supplementary Table 12). Heritability was enriched for genes expressed in cerebellum. See Supplementary Figures 7 and 8 for the Manhattan plot and tissue expression plots, and Supplementary Notes for more detail.

**Sensitivity analyses.**

We conducted sensitivity analyses to examine whether 1) the GWAS beat synchronization results are due to shared genetic effects with cognitive ability; 2) the GWAS genetic correlation results are driven by subtle residual population substructure; 3) the inversion on chromosome 17q21 is associated with local ancestry in our study sample; and 4) the MAPT association is due to the presence of patients with Parkinson's in the sample. These analyses demonstrated that our results are robust to each of these potential biases (Supplementary Notes).

**Cross-trait phenotypic extension of genetic correlations.**

Data from Phenotype Experiment 2 was analyzed to examine whether the genetic associations between beat synchronization we uncovered from LDSC regression would be reflected in true phenotypic associations (H4 of the pre-registration). Poor tapping accuracy was weakly associated with a morningness preference (r=-.10), more shortness of breath (r=-.16), and smoking 20 or more

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(lifetime) cigarettes (r=-.11) (Supplementary Table 13). Association with tinnitus and loud music exposure were nonsignificant. In other words, individuals who had better beat synchronization task performance were more likely to report eveningness chronotype, less likely to report shortness of breath when walking on level ground (these associations go in the same direction of the genetic study). However, they were less likely to report ever smoking, and less likely to report tinnitus (these associations were opposite of what was found in the genetic study). The association with loud music exposure was nonsignificant. These associations with chronotype, shortness of breath, and smoking remained significant after controlling for age, sex and education, and/or removing professional musicians from the sample. Self-reported rhythm (assessed using the seven-item Rhythm scale) was only associated with smoking status ($r$=-.08) and loud music exposure ($r$=-.13), even when controlling for covariates or focusing on non-musicians; however, these associations appeared in the opposite direction of the corresponding genetic associations. There was no evidence of interactions with musical sophistication or prior/current musician status for the H4 constructs, except that the association between loud music exposure and self-reported rhythm was weaker in individuals with who more actively performed music ($p$=.022), though this effect would not survive a strict multiple test correction.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Discussion

This study demonstrates that common genetic variation plays a role in synchronizing to a musical beat, complementing prior evidence of innate human rhythm sensitivity[4,19]. We conducted a GWAS of beat synchronization in more than half a million individuals surveyed within 23andMe, Inc., with the resulting summary statistics and post-GWAS analyses representing a significant advancement of our understanding of the genomic basis of a musicality phenotype. Heritability of beat synchronization is enriched for functions of the central nervous system on a number of dimensions (gene expression in brain tissue, genes involved in synaptic function, and more generally, genes involved in neural development and brain-specific regulatory regions of the genome). We successfully applied polygenic scores for beat synchronization to a separate cohort to predict musicianship, showing that the GWAS taps into the larger construct of musicality. Our findings point to pleiotropy between beat synchronization and several other types of biological functions (breathing function, motor function, processing speed, and chronotype), paving the way to a better understanding of the biological underpinnings of musicality and its health relevance.

Of course, the interpretation of these results depends upon a valid phenotype. In a series of phenotypic experiments, we demonstrate that self-reported beat synchronization/rhythm measures can be used in large-scale population-based studies as suitable proxies for measuring individual differences in beat synchronization ability. Our findings indicate that the “target question” phenotype used in the genetic study (*Can you clap in time with a musical beat?*) was highly related to beat synchronization task performance (i.e., accuracy in tapping along to musical excerpts). We show that

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

562 the GWAS phenotype is also significantly associated with the following: rhythm
563 perception task performance[50], a self-report Rhythm scale (generated from an additional
564 questionnaire), and a well-established assessment of musical sophistication[48]. These
565 results also converge with findings from prior work in small samples that show shared
566 variance among task performance of beat synchronization, rhythm perception, and
567 musical engagement/training[49,83–86]. The phenotypic associations were robust to
568 demographic factors (age, sex, level of education) and confidence (both as a
569 personality trait and as confidence in assessing one's own tapping accuracy), and were
570 not driven by the presence of professional musicians in the sample. More generally, our
571 findings indicate that people were able to self-report their rhythm abilities accurately
572 using other similar questionnaire items (i.e., scale from a seven-item Rhythm
573 questionnaire and a Likert scale single item both predict tapping accuracy). These
574 phenotype validation studies represent critical groundwork enabling self-reported
575 rhythm traits to be deployed online in large-scale population genetic cohorts where
576 participant assessment time and technology are constrained.

577 Beat synchronization exhibited a highly polygenic architecture, with sixty-seven
578 loci surpassing the threshold for genome-wide significance. The top-associated locus
579 (at sentinel SNP rs848293) mapped to *VRK2*, a protein kinase with multiple spliced
580 isoforms expressed in the brain that was previously associated with behavioral and
581 psychiatric phenotypes (i.e. depression[52], schizophrenia[53] and developmental delay[87]),
582 suggesting a biological connection between rhythm and neurodevelopment. The SNP-
583 based heritability of beat synchronization on the liability scale was moderate, ranging
584 from 13 to 16%, similar to heritability estimates of other complex traits (e.g., chronotype

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

585 GWAS[88]) and consistent with moderate heritability estimates of musical rhythm abilities
586 reported in twin studies[35–37].

587 We used complementary methods (tissue-specific GTex-based gene set
588 enrichment analysis in MAGMA, and partitioned heritability in tissue-specific expressed
589 genes and regulatory regions with LDSC-SEG) to examine potential mechanisms linking
590 genetic variation to neural architecture of the beat synchronization trait. Collectively,
591 results showed enrichment of the heritability of beat synchronization in many brain
592 tissues including cerebellum, dorso-lateral prefrontal cortex, inferior temporal lobe, and
593 several basal ganglia nuclei (including putamen, caudate, and nucleus accumbens).
594 This pattern of results likely reflects a genetic contribution to subcortical-cortical
595 networks underlying musical rhythm perception and production[28]. It is important to note
596 that because of the highly overlapping gene expression and chromatin signatures
597 between related tissues, with differences reflecting both function and cell type
598 composition, these results do not yet allow us to directly compare the respective
599 contributions different enriched brain regions. Rather, enrichment of brain-tissue-
600 specific enhancers and active-regulatory regions, and enrichment of expression in brain
601 tissue, suggest that regions of the genome involved in regulation of gene expression
602 play a role in the beat synchronization trait. Moreover, partitioning heritability chromatin
603 results showed an enrichment in both fetal and adult brain tissues, suggesting that beat
604 synchronization may be the result of neurodevelopmental or basic brain processes.
605 Gene set enrichments were also observed for synaptic function and MeCP2
606 transcription regulation in the nervous system. Taken together, these results are a
607 building block towards understanding how genes influence neural processes during

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

beat perception and production, complementing results obtained with neuroimaging methods[89–94].

Insights about the evolution of rhythm traits are suggested by the occurrence of two of the beat-synchronization-associated loci in human-accelerated regions (HARS) of the genome. In particular, rs1464791 is an expressive quantitative trait locus (eQTL) that regulates expression of *GBE1* in multiple tissues, including adrenal gland and muscle[61]; *GBE1* is also linked to neuromuscular disease[72] and reaction time[73]. HARs are involved in many functions, so it is difficult to explicitly link their accelerated evolution to beat synchronization. It is too early to tell whether the overlap between beat synchronization-associated loci and those two HARS supports evolutionary theories about music (e.g., groups moving to a beat in synchrony during joint music-making and/or temporal coordination of movement between parents and young children have been posited to exert selective pressures in early humans by enhancing group social cohesion and family bonding[27,95]). The contribution of the genetic architecture of musculo-skeletal systems and motor function to beat synchronization is further suggested by enriched heritability of SNPs that are enhancers located in musculoskeletal-tissue-specific regulatory regions of the genome, as well as our findings of genetic correlations between grip strength and beat synchronization.

We derived polygenic scores (PGS) generated from the beat synchronization GWAS to a separate genetically informative sample (in a healthcare biobank), demonstrating that beat synchronization PGS's are significantly associated with musicianship. These findings suggest that the genetic signature of musical beat synchronization is more widely tied to the biology of musicality, a finding corroborated in

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

631 Phenotype Experiment 2 by phenotypic associations between musical engagement,
632 measured beat synchronization (tapping accuracy), and self-reported rhythm ability.
633 These results also align with literature on other related biological bases of musicality
634 and rhythm, i.e., brain network activity during rhythmic tasks and the effects of plasticity-
635 related effects of musical expertise[28].

636 Moreover, our findings are promising for future large-scale genomic
637 interrogations using comprehensive music phenotyping yielding continuous musicality
638 variables (whether questionnaire-based[48,96] or measured aptitude-based variables[35]).
639 When this new field examines GWAS results on other heritable musicality traits such as
640 pitch discrimination and music training, future work can examine potential genetic
641 correlations between beat synchronization and other musical traits as predicted by
642 family-based studies[33,34,38,97]. While the current data show a clear connection between
643 the beat synchronization and broader musicality at the phenotypic and genetic levels,
644 further genomic investigation in well-powered samples is needed to disentangle the
645 *specificity* of genetic influences on rhythm from more general genetic influences on
646 musical ability.

647 The ability to move in synchrony to a musical beat encompasses beat perception
648 and extraction, motor periodicity, meter perception, and auditory-motor entrainment (see
649 Glossary in Supplementary Notes and[4,28,98]). Despite this complexity, beat is a highly
650 frequent feature of many musical systems[1,3,27]. For Western participants, beat
651 perception and production does not depend on musical training or a particular genre of
652 music (note that deficits in beat synchronization are not linked to lack of music
653 exposure[99]). A limitation of the current work is the restriction of the genetic sample to a

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

European ancestry (due to GWAS methodology constraints); investigating beat synchronization, musicality, and cross-trait correlations in populations of non-European ancestry should be a future priority for capturing the spectra of musicality in a wider range of ethnic, cultural and socio-economic contexts (see[100]). It is important to note that early research on individual differences in music ability in the early 1900's was pursued not only using what we now recognize as highly culturally biased assessments, but also explicitly through the lens of eugenics (see[101]), similar to early research on individual differences in cognition. We strongly condemn the design and intent of those studies, and emphasize that even a robust polygenic model for beat synchronization (or musicality in general) cannot make deterministic predictions about a specific individual's actual ability; rather, these models are probabilistic at the group level[102], and explain only a small part of the etiology of complex traits[103]. Furthermore, new knowledge on the genetic basis of musicality must be used ethically and fairly for research discovery and never for harm (e.g., preventing children's access to musical activities).

We replicated previous findings implicating location 4q22.1 in musicality-related traits[33,59] (*CCSER1* was the top-associated gene in our MAGMA analysis), but did not find support for previous gene associations from a set of genes that was drawn from prior candidate-gene, linkage, and GWAS studies with relatively small samples[58]. This is potentially due to well-known methodological problems with these methods particularly when applied to complex traits in small samples[104]. Without a second comparably sized GWAS available within which to conduct replication of the loci discovered in the primary GWAS, we were still able to demonstrate generalizability of these results by showing that PGS for beat synchronization predicts a musical trait in a separate biobank sample.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The genetic architecture of beat synchronization also remained virtually unchanged after conditioning the analyses on known GWAS markers of intelligence, in line with twin studies showing specific genetic effects of rhythmic aptitude, over and above any common genetic influences on rhythm and intelligence[36,105].

Our multi-pronged genetic cross-trait explorations revealed pleiotropic effects between beat synchronization and several traits, including types of biological rhythms (breathing-related phenotypes and circadian chronotypes); these were verified by phenotypic replication of selected associations in an additional experiment. We initially discovered novel genetic correlations between beat synchronization and specific health and cognitive traits (i.e., increased breathing function, greater grip strength, and faster processing speed) consistent with prior phenotypic and behavioral genetic studies of cognition, sensory processing, and musicality[97,106–108]. We further explored these novel associations with genomic SEM, which revealed common genetic variance among beat synchronization and breathing function, musculoskeletal function, and cognitive function (with the latter three traits previously shown to be genetically interrelated during the aging process[80,81]). Poor beat synchronization could be tied to certain health risks during aging, in light of other genetic and epidemiological work showing that lung function decline predicts later declines in motor function and psychomotor speed in older adults[109–112]. We replicated the positive genetic correlation between better beat synchronization ability (accuracy in tapping to the beat of musical excerpts) and lung capacity in our phenotype validation study, where we also found that better beat synchronization task performance was related to lower likelihood of shortness of breath.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Thus, the LDSC and genomic SEM results suggest that beat synchronization shares common biology with a constellation of related health traits, converging with the growing literature on the overlapping biomechanical and perceptual mechanisms of rhythms harnessed during synchronization, communication, muscle tensioning, and breathing. The cerebellum in particular governs certain neural mechanisms shared among these processes and plays important roles in the control of coordinated movement, balance, respiration, dance, and even rhythm perception during passive listening to music[29]. The rhythm-related traits multi-variate GWAS obtained with Genomic SEM resulted in 80 genome-wide significant loci and enriched heritability of genes expressed in Cerebellar tissue. Such phenomenon reflecting potentially shared biology across a broad range of physiological and developmental processes has been the focus of recent frameworks in which the ontogeny of rhythm, and more specifically of sensitivity to beat synchronization, are been hypothesized to play a role very early in life in its potential connections with maternal breathing and locomotion[113]. Experiments in infants demonstrate that parental singing and rocking help to regulate newborn breathing[114] and that infants reap the socio-emotional benefits of their caregivers' beat-synchronous interactions even before they possess the motor coordination to synchronize precisely to music[115]. Respiratory and upper limb movements have been found to be functionally synchronized during vocalization such that listeners can detect nuances of oscillatory body movement in an unseen speaker's vocalizations even when highly constrained[5]. "Beat gestures" in speech involve the cerebellum[116] and are inextricably linked to respiration, upper limb movement, and postural control, all of which may be biomechanically related to tapping or clapping to music.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In addition, we identified a new relationship between chronotype and beat synchronization (genetic correlation between eveningness and beat synchronization GWAS) and replicated this phenotypically in our pre-registered Phenotype Experiment in which we found that individuals who self-identified as 'evening people' tended to tap more accurately to music. This particular association was not moderated by musical sophistication scores, and survived sensitivity analysis in which professional musicians were removed. These results complement recent evidence of the increased prevalence of eveningness in musicians[117], indicating that the relationship between chronotype and musicianship cannot solely be explained by environment (i.e., nocturnal job demands of professional musicians), but that also other shared biological factors may play a role.

Our case/control GWAS has allowed us to effectively identify genetic alleles differentially associated with normative beat synchronization vs. beat impairment, complementing neural evidence of beat synchronization networks[83,85,86,99]. Future genetic studies could also examine beat synchronization task performance as a continuous trait, either through self-report or online-based methods directly measuring participants' tapping accuracy, such as demonstrated in Phenotype Experiment 2 using REPP[51]. Prior literature on liability threshold models has shown that case-control GWAS of complex traits yield similar results to those obtained through continuous phenotypic measures (for example, the genetic architecture of continuous measures of psychiatric symptoms is highly similar to the genetic architecture of cases versus controls[118]). Finally, although our GWAS was based on self-report, previous studies of other health traits based on self-report have effectively replicated associations from studies using

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

744 validated assessments, indicating that a powerful sample size can overcome limitations
745 arising from phenotyping error[119].

746 Taken together, our results advance knowledge of the biological basis of beat
747 synchronization by identifying genomic regions associated with individual differences in
748 beat synchronization, estimating its cumulative SNP-based heritability, successfully
749 applying a polygenic score model in a separate genetic sample, and exploring the
750 enrichment of heritability in genes tied to central nervous system function. Movement in
751 synchrony with a musical beat is a fundamental feature of music, and sensitivity to the
752 beat emerges early in development, supporting childhood development in numerous
753 ways[3,11,40,43] and with importance over the lifespan[120]. By elucidating the genetic
754 architecture of beat synchronization, we were able to identify the source of some of the
755 phenotypic variation observed in the general population and reveal health relevance
756 through cross-trait analyses. This study also provides a solid foundation for future
757 exploration of how specific genetic variants contribute to neural mechanisms of
758 entrainment, prediction, and reward harnessed during musical interactions.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Methods

## Phenotype validation studies

### *Phenotype Validation Experiment 1.*

*Overview.*

Phenotype Validation Experiment 1 was designed to determine if self-reported rhythm abilities measured with the question used in the GWAS (i.e., 'Can you clap in time with a musical beat?') would be associated with task-based rhythm perception performance. The study was conducted in Amazon's Mechanical Turk and received ethical approval from the Columbia University Institutional Review Board; participants gave their written informed consent. We selected the Beat-based Advantage paradigm as a rhythm discrimination (perception) test due to its design of stimuli with simple and complex meter[121] and prior history investigating individual differences in rhythm perception in a variety of brain and behavioural studies in adults and children with typical and atypical development[30,50,122,123] as well as feasibility for internet-based adaptation. A questionnaire (self-report questions) was administered prior to the perception task, to avoid biasing participant self-report responses by how they perceived their own task performance. See Supplementary Notes for additional details on procedure and self-report questionnaire.

*Participants*

The study sample was N=724 individuals who consented and passed a common headphone check[124] that guarantees good listening conditions and the ability to follow basic instructions; this test also effectively filters out bots. Participants (333 females;

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

387 males; 4 "other") were 18-73 years old (mean = 36.1 years, SD=10.9) with 0-45 years of self-reported musical experience (mean 3.7 years, SD=5.8), representing an average degree of musical experience (see norms in[48]); demographics are reported in Table 1 (note that n=2 did not report their age).

*Rhythm Perception Task*

Stimuli for the rhythm perception task consisted of 32 rhythms drawn from prior work[50,121]. For each participant, we randomized with probability of one half the occurrence of "simple" rhythms (strong beat-based metrical structure and generally easier to discriminate) and "complex" rhythms (weaker metrical structure due to syncopation and generally more challenging to discriminate). Each rhythm was presented using pure tone stimuli in one of 6 frequencies (294, 353, 411, 470, 528, and 587 Hz, selected at random), and one of 4 durations (ISI of 220, 230, 240, and 250 ms). Each trial consisted of 3 rhythms separated by 1500 ms of silence. The two first presentations were always identical, and in half of the trials (counterbalanced) the third rhythm was also identical (standard condition); in the other half of the trials, the rhythm differed by having one interval swapped (deviant condition). The pairings and structure of standard and deviant trials were taken from[50]. Participants were instructed that in each trial, they would listen to the series of three rhythms (the first two were always identical, and the third could be the same or different), and they had to indicate if the third rhythm was the same or different (see Supplementary Figure 1). Additional technical details are provided in the Supplementary Notes.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*Data analysis.*

*Self-report.* Responses to the target question 'Can you clap in time with a musical beat?' were as follows: n=654 (90.3%) participants answered 'Yes', n=25 (3.5%) answered 'No' and n=45 (6.2%) answered "I'm not sure." Regarding an additional self-report question 'Do you have a good sense of rhythm?', n=503(67%) answered 'Yes', 102(14%) answered 'No' and n=117(16%) answered 'I don't know'. n=488 answered 'Yes' to both questions; the tetrachoric correlation between these two self-report questions was $r$=0.73.

*Rhythm perception test.* Responses to the rhythm perception test were analysed using signal detection theory[50,125]; this method is appropriate for discrimination tasks where the participant has to categorize stimuli along some dimension with the resulting *d'* values the strength of detection of the signal relative to noise. *d'* values were calculated on the 32 test trials. As expected from prior work[50,126], individuals performed better at discriminating simple rhythms (mean d'= 1.98, SD =0.91) than complex rhythms (mean d'=1.43, SD =0.97) ($t(724)=11.11$, $p<2.2 \times 10^{-16}$, Cohen's d=0.58).

To examine whether the target question was related to the objective performance on the rhythm perception test, we performed a logistic regression analysis in which the clap-beat target question (Yes vs. No) was the outcome and quantitative scores on the rhythm perception test performance (standardized *d'* scores mean = 0, SD = 1) were the predictor. Covariates included age, education, and sex. McFadden's $R^2$ was also computed. We did not include 'I'm not sure' in the regressions, because this response was not available for data analysis in the GWAS. Given that the simple rhythms have a strong metrical structure that is known to facilitate detection and synchronization of the

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

828 beat[50], we also tested whether performance on the simple rhythm trials predicted self-
829 reported beat synchronization (i.e., those who responded Yes to the clap-beat
830 question).
831
832 ***Phenotype Experiment 2.***
833 *Overview.*
834 The aims of Phenotype Experiment 2 were two-fold: 1) to validate self-reported
835 beat synchronization phenotype as a proxy for objectively measured beat
836 synchronization ability, and 2) to explore phenotypic associations between rhythm/beat
837 synchronization and assorted traits found to be genetically correlated with beat
838 synchronization. Phenotype Experiment 2 was pre-registered in OSF prior to data
839 collection. This internet-based study consisted of a beat synchronization task to assess
840 the accuracy of participants' tapping in time with musical excerpts, and a series of
841 questionnaires assessing self-reported rhythm, musicality/music engagement, selected
842 health traits, confidence as a personality trait, and demographics. We used REPP[51] to
843 measure participants' tapping responses online with high temporal fidelity. The item
844 from the GWAS study, "Can you clap in time with a musical beat?" with possible
845 responses: Yes/No/I'm not sure, is referred to as the "target question."
846 We tested the following hypotheses: *H1*: Self-report responses to the target
847 question will be correlated with beat synchronization task performance (i.e., accuracy of
848 tapping to the beat of music), such that individuals who respond Yes to the "target
849 question" are predicted to tap more accurately to the beat of musical excerpts (i.e., they
850 will have lower standard deviation of asynchrony than individuals who respond No to the

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

target question). *H1*a: Self-report on a highly similar self-report question ("I can tap in time with a musical beat") with responses on a 7-point agreement Likert scale are predicted to be correlated with tapping accuracy. *H2a:* The target question will be associated with broader rhythm ability/engagement (measured with a rhythm scale from other self-report questions). *H2b:* Beat synchronization task performance reflects broader self-reported rhythm ability/engagement. *H3:* To examine whether confidence (either as a personality trait or sureness in one's own task performance) affects the reliability of self-reported beat synchronization. *H4:* Selected traits found to be genetically correlated with beat synchronization in the GWAS will be phenotypically correlated with beat synchronization task performance and the Rhythm Scale. Specifically: better beat/rhythm is correlated with evening chronotype (H4a), less shortness of breath (H4b), more tinnitus and loud music exposure (H4c), and more smoking (H4d); and that these associations would survive controlling for age, sex, and education (H4e). *H5.* Responses to the target question will be positively correlated with musical engagement measured with the Gold-MSI. *H6*. The associations in H4 would interact with being a musician, or more generally, with musical engagement.

*Participants.*

A total of N=1,412 individuals met participation criteria outlined in the pre-registration (including passing the attention check item and not abandoning the study before completion). The study took place in Amazon Mechanical Turk and all participants provided informed consent in accordance with the Max Planck Society Ethics Council's approved protocol. Participants (728 females; 678 males; 6 prefer not

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

874 answer) were 18-77 years old (mean=36.3 years, SD=11.9) and had of 1-2 years of
875 self-reported musical experience. To ensure that the tapping technology measured beat
876 synchronization with high temporal fidelity, it was crucial that participants complied with
877 instructions to perform the tapping task (e.g., using the laptop speakers instead of
878 headphones, with minimal background noise, etc.), and also used hardware and
879 software without any technical issues that would preclude the recording signal (e.g.,
880 malfunctioning speakers or microphones, or the use of strong noise cancellation
881 technology; see[51]). Thus, several precautions, including calibration tests and practice
882 trials, were taken to make sure the tapping technology would work effectively, excluding
883 cases that did not meet the requirements (see Supplementary materials for details). A
884 subset of n=542 had appropriate hardware to complete all parts of the study (including
885 the tapping tests). Questionnaires were administered in the full sample of participants.
886 Sample demographics are reported in Table 1. Demographics of the participants that
887 completed the tapping experiment was highly similar to the full sample, as shown in the
888 table; furthermore, 65.3% of the full sample and 64.9% of tapping sample had a
889 Bachelor's degree or higher.
890
891 *Data collection for Phenotype Experiment 2.*
892 The first questionnaire included self-report items, including the "target question,"
893 and also covering a variety of musical, health, and interest phenotypes. The health
894 phenotype questions were chosen from phenotypes (chronotype, smoking, shortness of
895 breath, tinnitus, and loud music exposure) found to be genetically correlated with beat
896 synchronization in our genetic analyses. Rhythm questions were selected for their

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

particular relevance to various aspects of interacting/engaging with musical rhythm. The order of the questions was fixed for all participants. In addition, we used an attention check item[127] between item 10 and 11, in order to exclude fraudulent responders, such as computer bots or disengaged participants responding randomly to the experiments. The end-questionnaire consisted of items covering the following additional self-report topics: another question about being a musician, a task confidence rating question, a Confidence scale, a 16-item short version of the Gold-MSI[48] (items were chosen due to their high reliability scores: reliability omega = 0.92), and a Demographic questionnaire. Questionnaire items for Phenotype Experiment 2 are listed in the Appendix of the Supplementary Notes.

*Tapping technology.* Beat synchronization is particularly challenging to study with online research, where variability in participants' hardware and software can introduce all kinds of delay in latency and jitter into the recorded time stamps[128,129]. Here we used REPP (see[51] for full details and a validation study of the technology), a robust cross-platform solution for measuring sensorimotor synchronization in online experiments that has high temporal fidelity and can work efficiently using hardware and software available to most participants online. To address core issues related to latency and jitter, REPP uses a free-field recording approach: specifically, the audio stimulus is played through the laptop speakers and the original signal is simultaneously recorded with participants' tapping responses using the built-in microphone. The resulting recording is then analyzed using signal processing techniques to extract and align timing cues with high temporal accuracy.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*Beat synchronization task.* The beat synchronization task procedure consisted of three parts: calibration tests, practice phase, and main tapping phase. Participants started with the calibration tests, including a volume test to calibrate the volume of the laptop speakers to a level sufficient for detection by the microphone, a background noise test to make sure participants were in a quiet environment, and a tapping test to help participants practice how to tap on the surface of their laptop in the right level and location to be detected by the microphone. Participants were then presented with the practice phase, which consisted of four 15-second trials of isochronous tapping to a metronome beat (two with inter-onset interval of 500 msec and two with inter-onset interval of 600 msec). Following the practice phase, participants were presented with the main tapping task consisting of eight trials (4 musical excerpts, each played twice), with each trial 30 seconds long. The order of presentation of the practice trials and test trials was randomized for each participant.

The musical excerpts were drawn from the MIREX 2006 Audio Beat Tracking database in which musical excerpts had been annotated for beat locations by 30 listeners who tapped along to the music[130]. We chose these four MIREX clips that represent different music genres with different tempos and tapping difficulty: track 1 (“You're the First, the Last, My Everything” by Barry White), track 3 (“El Contrapunto” by Los Mensajeros de La Libertad), track 7 (“Le Sacre du Printemps” by Stravinsky), and track 19 (“Possessed to Skate” by Suicidal Tendencies) of the MIREX training set (respectively). Based on the annotations in[130], we identified the target beat locations from those consistently produced by the annotators. We performed kernel density estimation with a kernel width of 20 msec; this provided an estimate of the probability of

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

producing a response in any given time. The peaks of the probability density were located using Matlab's findpeaks function with the following parameters: 'MinPeakHeight', 0.11/ts, 'MinPeakProminence',0.11/ts, 'MinPeakDistance',100 msec, where ts is the number of responses in the clip. Beat locations were extracted from the entire 30 seconds of the clip and used as the reference location for computing the asynchrony. To help participants find the beat and eliminate potential ambiguity of tapping at half- or double-time the tempo, a metronome marking the beats in the first 11 seconds of the clip were added to the stimulus (as commonly used in this type of tapping paradigm). Additional technical details are provided in the Supplementary Notes, and Supplementary Figure 2 illustrates the instructions for participants.

*Data Analysis.*

**Beat synchronization task performance: Tapping accuracy analysis**

Let $S_t$ and $R_t$ be the stimulus and response onsets, respectively. In case of the metronome $S_t$ are the metronome onset (practice phase) and for music clips $S_t$ is the target beat location based on the annotations. We define the asynchrony as $a_t=R_t - R_t$. Based on prior work[131], we chose the standard deviation of the asynchrony (std($a_t$)) as our main target interest variable, as this appears to be a robust measure of individual performance and tightly linked to musical abilities[132]. We used metronome onsets to mark the beat metric level in an unambiguous way[133]. We emphasize that the metronome onsets were only physically present during the beginning and end of each clip. We used only the participant-produced asynchronies during the epoch at beats in which the guiding metronome was *not* present, in order to test the ability of the

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

965 participants to synchronize to music without the metronome sounds (results were nearly
966 identical when we included all onsets including the one were physical metronome
967 onsets were present). For the main test scores, we used the asynchronies computed
968 relative to the virtual beat locations computed from prior human annotators in MIREX.

969 **Regression analyses**

970 In accordance with the OSF preregistration, we examined whether responses to
971 self-reported beat synchronization phenotype were associated with objectively-
972 measured tapping accuracy, other self-reported measures of rhythm ability, confidence,
973 and/or musical sophistication using logistic regression and McFadden's $R^2$ (for H1, H2a,
974 H3, and H5) and linear regression (for H1a and H2b). Likewise, we used linear
975 regression to examine potential replication of cross-trait associations uncovered by
976 genetic analyses (H4a-d), to examine whether musical background interacted with the
977 above associations (H6). Analyses were conducted in R version 3.5.1[134]. As described
978 in our preregistration, individuals were recruited using MTurk and were included unless
979 they failed and attention check item or abandoned the experiment before completing the
980 study (N=1,412). Usable tapping data was available for n=542 individuals. The majority
981 of exclusions were due to technical reasons detected by REPP's signal processing
982 pipeline during the practice trials (e.g., poor signal, noisy environment, wearing
983 headphone, issues with laptop microphone, or people not tapping at all), but some
984 additional subjects (n=19) were excluded for not having enough usable trials during data
985 analysis. Missing covariates were handled using pair-wise deletion. Exclusion criteria
986 are detailed in the Supplementary Notes.
987

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## GWAS of beat synchronization.

### GWAS sample.

Genome-wide association study summary statistics were generated from data acquired by personal genetics company 23andMe, Inc. Phenotypic status was based on responses to an English-language online questionnaire in which individuals self-reported “Yes” (cases) or “No” (controls) to the question ‘Can you clap in time with a musical beat?”. Individuals who responded “I’m not sure” were excluded from the genomic dataset as their data was not available. The GWAS included a total of 555,660 cases and 51,165 controls (total N=606,825, mean age(SD)=52.09(18.5), prevalence=92%), all of European ancestry; age range breakdown is provided in Table 1. All individuals provided informed consent according to 23andMe’s human subject protocol, which is reviewed and approved by Ethical & Independent Review Services, a private institutional review board (http://www.eandireview.com).

### *Genotypes and QC.*

The National Genetics Institute (NGI) performed the DNA extraction and genotyping on saliva samples for the 23andMe GWAS. Overall, there were five genotyping platforms and subjects were genotyped on only one of them. The v1 and v2 platforms had variants of the Illumina HumanHap550+ BeadChip, including approximately 25,000 custom SNPs selected by 23andMe, with a total of about 560,000 SNPs. The v3 platform had variants of the Illumina OmniExpress+ BeadChip, with custom content to improve the overlap with the v2 array, with a total of about 950,000 SNPs. The v4 platform covered about 570,000 SNPs, providing extra coverage of

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lower-frequency coding variation. The v5 platform, in current use, is based on an Illumina Infinium Global Screening Array (~640,000 SNPs) supplemented with ~50,000 SNPs of custom content. In cases where samples did not reach the 98.5% call rate, the sample was re-genotyped. When analyses failed repeatedly, then customers were re-contacted by 23andMe customer service to provide additional samples.

23andMe restricted participants to a set of unrelated individuals of European ancestry, determined through an analysis of local ancestry[135]. Relatedness was defined using a segmental identity-by-descent (IBD) estimation algorithm[136]. Imputation was conducted by combining the May 2015 release of 1000 Genomes Phase 3 haplotypes[137] with the UK10K imputation reference panel[138] to create a single unified imputation reference panel. Phasing was conducted using an internally-developed tool, Finch, which uses the Beagle graph-based haplotype phasing algorithm[139] for platforms V1 to V4 while for the V5 platform a similar approach was used with a new phasing algorithm, Eagle2[140]. SNPs with a Hardy-Weinberg $p<10^{-20}$, or a call rate of <90% were flagged. SNPs were also flagged if they were only genotyped on their 'V1' and/or 'V2' platforms due to small sample size and also if SNPs had genotype date effects. Finally, SNPs were also flagged if they had probes matching multiple genomic positions in the reference genome[136–140].

*GWAS.*

GWAS was conducted using logistic regression under an additive genetic model, while adjusting for age, sex, the top five principal components of ancestry in order to control for population stratification, and indicators for genotype platforms to account for

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

batch effects. We excluded SNPs with Minor Allele Frequency (MAF) <0.01, low imputation quality ($R^2$<0.3) and indels, resulting in 8,288,850 SNPs in the GWAS summary statistics. SNPs within the X chromosome were further excluded, resulting in 8,076,862 SNPs for subsequent analyses unless otherwise indicated.

**Statistical analyses**

*FUMA-based analyses.* The FUMA[60] web application was used on the Genome-Wide Association summary statistics to identify genomic loci along with the "sentinel" SNPs that were independent in our analysis with a genome-wide significant P-value (<5 x $10^{-8}$) that are in approximate linkage disequilibrium (LD) with each other at $r^2$<0.1 and to generate Manhattan plots and Quantile-Quantile plots.

Gene-based analysis and gene-set analysis was performed with MAGMA (v1.08) using FUMA (v1.3.4) and the association analysis summary statistics. Gene expression data analysis was obtained from GTEx v8 (https://www.gtexportal.org/home/) integrated by FUMA. More specifically, the gene expression values were log2 transformed average RPKM per tissue type after winsorization at 50 based on GTEx RNA-seq data. Tissue expression analysis was performed for 53 tissue types where the result of gene analysis was tested for one side while conditioning on average expression across all tissue types.

*LD score regression and genetic correlations.* SNP-heritability was computed with LD Score regression software[64], and heritability estimates were adjusted to the liability scale based on population prevalence of poor rhythm of 3.0%-6.5% (Supplementary

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3, Supplementary Notes). We partitioned heritability of beat synchronization by 52 broad functional categories (Supplementary Table 7), using stratified LD score regression[64,67] (Bonferroni-corrected significance level of $p=9.6x10^{-4}$).

We further investigated cell-type-specific and tissue-specific enrichments with LDSC-SEG (LDSC Specifically Expressed Genes)[67], using a total of 697 gene sets (3 Cahoy gene sets, 205 Multi-tissue gene expression sets and 489 Multi-tissue chromatin sets from the RoadMap Epigenomics and ENCODE datasets); the Bonferroni-corrected significance level for this analysis was $7.1x10^{-5}$ (Supplementary Table 8).

The set of human accelerated regions (HARs) was taken from[69]. All variants in perfect LD ($r^2 = 1.0$ in 1000 Genomes European individuals) with variants in HARs were considered in the analysis. Similarly, variants tagging Neanderthal introgressed haplotypes were defined as in[141]. All variants in perfect LD with a Neanderthal tag SNP were considered Neanderthal variants. For each set, we performed stratified LDSC (v1.0.0) with European LD scores and the baseline LD-score annotations v2.1. The heritability enrichment is defined as the proportion of heritability explained by SNPs in the annotation divided by the proportion of SNPs in the annotation. Standard effect size ($\tau_c^*$), which quantifies the effects unique to the annotation, is the proportionate change in per-SNP heritability associated with a one standard deviation increase in the value of the annotation, conditional on other annotations in the baseline v2.1 model[66]. To determine the expected number of overlaps between the N loci significantly associated with beat synchronization and HARs, we computed all overlaps between these sets of genomic regions (in hg19 coordinates) using bedtools2[142]. We then randomly shuffled the locations of HARs around the genome respecting their lengths and avoiding gaps in

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the genome assembly. We repeated this process 10,000 times and for each iteration computed the number of overlaps observed with the significantly associated loci. Based on this empirical distribution created with no association between the region sets, we computed the enrichment and p-value for the observed number of overlaps.

Genetic correlations between beat synchronization and other complex traits were estimated using LDSC through LDHub v1.9.0 (http://ldsc.broadinstitute.org/ldhub/)[47] using publicly available GWAS summary statistics therein. In total, 764 traits were examined (Bonferroni corrected threshold = $6.5 \times 10^{-5}$).

*Beat synchronization Polygenic Score (PGS) prediction of musicianship*

*Overview.* We examined whether beat synchronization polygenic scores (PGS) would be associated with musicianship in a health care context. Musicians cases were drawn from a recent phenome-wide study of 9,803 musicians[76] identified from keyword searches of patient electronic health records (EHRs) in Vanderbilt University Medical Center's de-identified research database (Synthetic Derivative). The phenotyping method was based on mining of clinical notes, utilizing 4 keywords and 449 regular expressions (i.e., "musician", "plays the piano"); see Supplementary Notes and[76] for details. Their method was validated with manually conducted chart review, with a positive predictive value (PPV) of 93%. Here we accessed the subset of n=1,753 musicians that were also part of the BioVU database and had genotyped data on file, to test the hypothesis that higher PGS for beat synchronization would be associated with musicianship (i.e., higher likelihood of having musician-related keywords/regular expressions recorded in an individual's electronic health record).

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We only selected individuals of European ancestry with genetic data that met standard quality control thresholds, resulting in n=1,753 individuals (965 (55%) males, mean median age of record (SD)=44.3(22.5)) as musician "cases" and 65,147 controls (28698(44%) males, mean median age of record (SD)=48.3(22.3)). See Supplementary Notes for details on the phenotyping, the samples, genotyping, and QC.

*Polygenic scores.* We used an IBD filter of 0.2 in order to include unrelated European samples of BioVU. PGS were generated using the beat synchronization GWAS summary statistics, using software PRS_CS[143]. Briefly, this method uses a Bayesian regression framework and places continuous shrinkage (CS) prior on SNP effect sizes; this method outperforms previous methods in terms of prediction accuracy especially when the training sample size is large[143], as is the case with the beat synchronization GWAS. The 1000genomes reference set was used. The PGS was standardized to have a mean of 0 and SD of 1.

*Control traits in BioVU.* As negative control phenotypes we selected the following traits: Iron deficiency anemias (phecode:280, cases=4,594, controls=62,306), carcinoma in situ of skin (phecode:172.3, cases=523, controls=66,377), and cancer of the brain (phecode:191.11, cases=970, controls=65,930).

*Data analysis.* We conducted a logistic regression where the outcome variable was keyword "musician" (yes vs no) and the predictor variable was PGS for beat synchronization, while also adjusting for median age, sex, 10 Principal Components and genotyping batch. The same process was followed when the outcome variables were iron deficiency anemias, carcinoma in situ of skin and cancer of the brain (cases vs. controls).

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Acknowledgments:** We are grateful to 23andMe participants for their contribution to the study, Nancy Cox and Simon Fisher for suggestions and insight throughout the process, Navya Thakkar for assistance with graphics, and Miriam Lense, Matthew D. Morrison, and Wim Pouw for thoughtful discussions. We would also like to thank Michaela Novakovic, Yune Lee, and Duane Watson for input during previous stages of the project, and Vanderbilt Trans-Institutional Programs for sparking the initial collaboration. Members of the 23andMe Research Team are: Michelle Agee, Stella Aslibekyan, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah K. Clark, Sarah L. Elson, Kipper Fletez-Brant, Pierre Fontanillas, Nicholas A. Furlotte, Pooja M. Gandhi, Karl Heilbron, Barry Hicks, Karen E. Huber, Ethan M. Jewett, Yunxuan Jiang, Aaron Kleinman, Keng-Han Lin, Nadia K. Litterman, Jennifer C. McCreight, Matthew H. McIntyre, Kimberly F. McManus, Joanna L. Mountain, Sahar V. Mozaffari, Priyanka Nandakumar, Elizabeth S. Noblin, Carrie A.M. Northover, Jared O'Connell, Steven J. Pitts, G. David Poznik, Anjali J. Shastri, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Joyce Y. Tung, Robert J. Tunney, Vladimir Vacic, and Xin Wang.

**Funding:**

Research reported in this publication was supported by the National Institutes of Health Common Fund through the Office of the NIH Director and the Eunice Kennedy Shriver National Institute of Child Health & Human Development under award number DP2HD098859, and by the National Institute On Deafness And Other Communication Disorders under award numbers K18DC017383 and R01DC016977. JAC was supported by the National Institutes of Health (R35GM127087). The content is solely

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the responsibility of the authors and does not necessarily represent the official views of the NIH.

The dataset used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU which is supported by numerous sources: institutional funding, private agencies, and federal grants. These include the NIH funded Shared Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed at https://victr.vumc.org/biovu-funding/. Also, The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. GTEx data used for the analyses described in this manuscript were obtained from: the GTEx Portal on 07/26/19 and dbGaP accession number phs000424.vN.pN on 07/26/19 .

## Author contributions

*Conceptualization of study:* Reyna Gordon, Lea Davis

*Study design of GWAS and design of other genomic analyses:*
Lea Davis, Reyna Gordon, J. Fah Sathirapongsasuti, Maria Niarchou, Tony Capra, David Hinds

*Data collection of genomic data:* J. Fah Sathirapongsasuti, The 23andMe Research Team, David Hinds

*Genome-wide Association analysis:* J. Fah Sathirapongsasuti, The 23andMe Research Team

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*Post-association QC and generation of figures:* Peter Straub and Maria Niarchou

*Post-GWAS analyses (heritability, gene-based analyses, gene set analyses, LD correlations, mtcojo, PGS in BioVU*):

Maria Niarchou, Reyna Gordon, Peter Straub, Else Eising, Lea Davis

*HARS and Neanderthal introgression analyses and interpretation:* Evonne McArthur, John A Capra

*Genomic SEM:* Daniel Gustavson

*Sensitivity Analyses of Chromosome 17 inversion and Parkinson's Disease:* J. Fah Sathirapongsasuti

*Estimation of population prevalence of rhythm deficits:* Miriam Mosing, Daniel Gustavson, and Reyna Gordon

*Phenotype validation studies - design and materials*: Reyna Gordon, J. Devin McAuley, Nori Jacoby, Daniel Gustavson, Manuel Anglada-Tort

*Phenotype validation data collection:* Nori Jacoby, Eamonn Bell, Manuel Anglada-Tort

*Phenotype validation study data analysis:* Manuel Anglada-Tort, Daniel Gustavson, Eamonn Bell, Nori Jacoby, Peter Straub, Maria Niarchou, Reyna Gordon

*Interpretation of data, writing, editing, and reviewing drafts:* All authors

*Project Supervision:* Reyna Gordon, Lea Davis, Nori Jacoby, David Hinds

**Competing interests:** JFS, DH, and members of the 23andMe Research Team are employees of 23andMe, Inc., and hold stock or stock options in 23andMe. All other authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Data and materials availability:** The full GWAS summary statistics for the 23andMe dataset will be made available through 23andMe to qualified researchers under an agreement that protects the privacy of the 23andMe participants. Please visit research.23andme.com/collaborate/#publication for more information and to apply to access the data. The data and code from the phenotype validation studies, and all of the code from the post-GWAS analyses, are also available upon reasonable request.

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## References


1. Savage, P. E., Brown, S., Sakai, E. & Currie, T. E. Statistical universals reveal the structures and functions of human music. *Proc Natl Acad Sci U S A* **112**, 8987–8992 (2015).
2. Ravignani, A., Delgado, T. & Kirby, S. Musical evolution in the lab exhibits rhythmic universals. *Nat. Hum. Behav.* (2017) doi:10.1038/s41562-016-0007.
3. Mehr, S. A. *et al.* Universality and diversity in human song. *Science (80-. ).* (2019) doi:10.1126/science.aax0868.
4. Kotz, S. A., Ravignani, A. & Fitch, W. T. The Evolution of Rhythm Processing. *Trends in Cognitive Sciences* (2018) doi:10.1016/j.tics.2018.08.002.
5. Pouw, W., Paxton, A., Harrison, S. J. & Dixon, J. A. Acoustic information about upper limb movement in voicing. *Proc. Natl. Acad. Sci. U. S. A.* (2020) doi:10.1073/pnas.2004163117.
6. Large, E. W. & Jones, M. R. The dynamics of attending: How we track time varying events. *Psychol. Rev.* **106**, 119–159 (1999).
7. Nobre, A. C. & Van Ede, F. Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience* (2018) doi:10.1038/nrn.2017.141.
8. Hannon, E. E. & Trehub, S. E. Tuning in to musical rhythms: Infants learn more readily than adults. *Proc. Natl. Acad. Sci. U. S. A.* (2005) doi:10.1073/pnas.0504254102.
9. Winkler, I., Haden, G. P., Ladinig, O., Sziller, I. & Honing, H. Newborn infants detect the beat in music. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2468–2471 (2009).
10. Zentner, M. & Eerola, T. Rhythmic engagement with music in infancy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5768–5773 (2010).
11. Cirelli, L. K., Trehub, S. E. & Trainor, L. J. Rhythm and melody as social signals for infants. *Ann. N. Y. Acad. Sci.* **1423**, 66–72 (2018).
12. Nazzi, T., Bertoncini, J. & Mehler, J. Language discrimination by newborns: toward an understanding of the role of rhythm. *J Exp Psychol Hum Percept Perform* **24**, 756–766 (1998).
13. McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M. & Miller, N. S. The time of our lives: Life span development of timing and event tracking. *J. Exp. Psychol. Gen.* (2006) doi:10.1037/0096-3445.135.3.348.
14. Polak, R. *et al.* Rhythmic Prototypes Across Cultures. *Music Percept. An Interdiscip. J.* (2018) doi:10.1525/mp.2018.36.1.1.
15. London, J., Polak, R. & Jacoby, N. Rhythm histograms and musical meter: A corpus study of Malian percussion music. *Psychon. Bull. Rev.* (2017) doi:10.3758/s13423-016-1093-7.
16. Clayton, M., Sager, R. & Will, U. In time with the music : the concept of entrainment and its significance for ethnomusicology. *Eur. Meet. Ethnomusicol.* (2005).
17. Polak, R. & London, J. Timing and Meter in Mande Drumming from Mali. *Music Theory Online* (2014) doi:10.30535/mto.20.1.1.
18. Polak, R., London, J. & Jacoby, N. Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

study of malian jembe drumming. *Front. Neurosci.* (2016) doi:10.3389/fnins.2016.00285.

19. Patel, A. D. & Iversen, J. R. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front. Syst. Neurosci.* **8**, 57 (2014).
20. Jacoby, N. & McDermott, J. H. Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Curr. Biol.* (2017) doi:10.1016/j.cub.2016.12.031.
21. Cameron, D. J., Bentley, J. & Grahn, J. A. Cross-cultural influences on rhythm processing: Reproduction, discrimination, and beat tapping. *Front. Psychol.* (2015) doi:10.3389/fpsyg.2015.00366.
22. Neuhoff, H., Polak, R. & Fischinger, T. Perception and evaluation of timing patterns in drum ensemble music from Mali. *Music Perception* (2017) doi:10.1525/MP.2017.34.4.438.
23. Tarr, B., Slater, M. & Cohen, E. Synchrony and social connection in immersive Virtual Reality. *Sci. Rep.* (2018) doi:10.1038/s41598-018-21765-4.
24. Lense, M. D. & Camarata, S. PRESS-Play: Musical Engagement as a Motivating Platform for Social Interaction and Social Play in Young Children with ASD. *Music Sci.* (2020) doi:10.1177/2059204320933080.
25. Honing, H. On the biological basis of musicality. *Ann. N. Y. Acad. Sci.* (2018) doi:10.1111/nyas.13638.
26. Fitch, W. T. Empirical approaches to the study of language evolution. *Psychon. Bull. Rev.* **24**, 3–33 (2017).
27. Savage, P. E. *et al.* Music as a coevolved system for social bonding. *Behav. Brain Sci.* 1–42 (2020) doi:10.1017/S0140525X20000333.
28. Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M. & Fitch, W. T. Finding the beat: a neural perspective across humans and non-human primates. *Philos Trans R Soc L. B Biol Sci* **370**, 20140093 (2015).
29. Gordon, C. L., Cobb, P. R. & Balasubramaniam, R. Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLoS One* (2018) doi:10.1371/journal.pone.0207213.
30. Grahn, J. A. & McAuley, J. D. Neural bases of individual differences in beat perception. *Neuroimage* **47**, 1894–1903 (2009).
31. Law, L. N. C. & Zentner, M. Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PLoS One* **7**, e52508 (2012).
32. Dalla Bella, S. *et al.* BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behav. Res. Methods* (2017) doi:10.3758/s13428-016-0773-6.
33. Pulli, K. *et al.* Genome-wide linkage scan for loci of musical aptitude in Finnish families: evidence for a major locus at 4q22. *J Med Genet* **45**, 451–456 (2008).
34. Oikkonen, J. *et al.* A genome-wide linkage and association study of musical aptitude identifies loci containing genes related to inner ear development and neurocognitive functions. *Mol. Psychiatry* **20**, 275 (2014).
35. 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

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discrimination Test. *Pers. Individ. Dif.* **63**, 87–93 (2014).
36. 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).
37. Seesjärvi, E. *et al.* The Nature and Nurture of Melody: A Twin Study of Musical Pitch and Rhythm Perception. *Behav. Genet.* (2016) doi:10.1007/s10519-015-9774-y.
38. 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. B Biol. Sci.* **370**, 20140092 (2015).
39. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk of complex disease. *Curr. Opin. Genet. Dev.* **18**, 257–263 (2008).
40. Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L. & Kraus, N. Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14559–14564 (2014).
41. Swaminathan, S. & Schellenberg, E. G. Musical Ability, Music Training, and Language Ability in Childhood. *J. Exp. Psychol. Learn. Mem. Cogn.* (2019) doi:10.1037/xlm0000798.
42. Keller, P. E., Novembre, G. & Hove, M. J. Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences* (2014) doi:10.1098/rstb.2013.0394.
43. Ladányi, E., Persici, V., Fiveash, A., Tillmann, B. & Gordon, R. L. Is atypical rhythm a risk factor for developmental speech and language disorders? *WIREs Cogn. Sci.* **11**, (2020).
44. Moumdjian, L., Sarkamo, T., Leone, C., Leman, M. & Feys, P. Effectiveness of music-based interventions on motricity or cognitive functioning in neurological populations: A systematic review. *European Journal of Physical and Rehabilitation Medicine* (2017) doi:10.23736/S1973-9087.16.04429-4.
45. Cheever, T. *et al.* NIH/Kennedy Center Workshop on Music and the Brain: Finding Harmony. in *Neuron* (2018). doi:10.1016/j.neuron.2018.02.004.
46. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. Supplementary Note with R Code. *Nat. Rev. Genet.* (2019) doi:10.1038/s41576-019-0137-z.
47. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
48. Müllensiefen, D., Gingras, B., Musil, J. & Stewart, L. The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS One* **9**, e89642 (2014).
49. Musil, J. J., Iversen, J. R. & Müllensiefen. Measuring individual differences in musical beat alignment perception. *Pers. Individ. Dif.* **60**, S35 (2014).
50. Grahn, J. A. & Brett, M. Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* **19**, 893–906 (2007).
51. Anglada-Tort, M., Harrison, P. & Jacoby, N. REPP: A robust cross-platform

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

solution for online sensorimotor synchronization experiments. *bioRxiv* (2021) doi:10.1101/2021.01.15.426897.

52. Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* **50**, 920–927 (2018).
53. Pardinas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet* **50**, 381–389 (2018).
54. Chang, D. *et al.* A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* (2017) doi:10.1038/ng.3955.
55. D'Angelo, D. *et al.* Defining the effect of the 16p11.2 duplication on cognition, behavior, and medical comorbidities. *JAMA Psychiatry* (2016) doi:10.1001/jamapsychiatry.2015.2123.
56. Hippolyte, L. *et al.* The Number of Genomic Copies at the 16p11.2 Locus Modulates Language, Verbal Memory, and Inhibition. *Biol. Psychiatry* (2016) doi:10.1016/j.biopsych.2015.10.021.
57. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291–295 (2015).
58. Oikkonen, J., Onkamo, P., Järvelä, I. & Kanduri, C. Convergent evidence for the molecular basis of musical traits. *Sci. Rep.* **6**, 39707 (2016).
59. Park, H. *et al.* Comprehensive genomic analyses associate UGT8 variants with musical ability in a Mongolian population. *J Med Genet* **49**, 747–752 (2012).
60. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
61. Consortium, Gte. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (80-. ).* **348**, 648–660 (2015).
62. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25 (2000).
63. The Gene Ontology, C. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* **47**, D330–D338 (2019).
64. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228 (2015).
65. Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* (2011) doi:10.1038/nature10530.
66. Hujoel, M. L. A., Gazal, S., Hormozdiari, F., van de Geijn, B. & Price, A. L. Disease Heritability Enrichment of Regulatory Elements Is Concentrated in Elements with Ancient Sequence Age and Conserved Function across Species. *Am J Hum Genet* **104**, 611–624 (2019).
67. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* (2018) doi:10.1038/s41588-018-0081-4.
68. Mithen, S. J. *The Singing Neanderthals: The origins of music, language, mind, and body.* (Harvard University Press, 2005).
69. Capra, J. A., Erwin, G. D., McKinsey, G., Rubenstein, J. L. & Pollard, K. S. Many human accelerated regions are developmental enhancers. *Philos Trans R Soc L.*

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*B Biol Sci* **368**, 20130025 (2013).
70. Hubisz, M. J. & Pollard, K. S. Exploring the genesis and functions of Human Accelerated Regions sheds light on their role in human evolution. *Curr Opin Genet Dev* **29**, 15–21 (2014).
71. Doan, R. N. *et al.* Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. *Cell* **167**, 341-354 e12 (2016).
72. Todd, E. J. *et al.* Next generation sequencing in a large cohort of patients presenting with neuromuscular disease before or at birth. *Orphanet J. Rare Dis.* (2015) doi:10.1186/s13023-015-0364-0.
73. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet* **104**, 65–75 (2019).
74. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* **9**, 2098 (2018).
75. Sankararaman, S. *et al.* The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).
76. Niarchou, M., Lin, G., Lense, M. D., Gordon, R. L. & Davis, L. K. The medical signature of musicians: A Phenome-wide association study using an Electronic Health Record database. *medRxiv* 2020.08.14.20175109 (2020) doi:10.1101/2020.08.14.20175109.
77. Couth, S., Mazlan, N., Moore, D. R., Munro, K. J. & Dawes, P. Hearing Difficulties and Tinnitus in Construction, Agricultural, Music, and Finance Industries: Contributions of Demographic, Health, and Lifestyle Factors. *Trends Hear.* (2019) doi:10.1177/2331216519885571.
78. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
79. Grotzinger, A. D. *et al.* Genomic SEM Provides Insights into the Multivariate Genetic Architecture of Complex Traits HHS Public Access. *Nat Hum Behav* (2019) doi:10.1038/s41562-019-0566-x.
80. Shrine, N. *et al.* New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat. Genet.* (2019) doi:10.1038/s41588-018-0321-7.
81. Willems, S. M. *et al.* Large-scale GWAS identifies multiple loci for hand grip strength providing biological insights into muscular fitness. *Nat. Commun.* (2017) doi:10.1038/ncomms16015.
82. Finkel, D., Ernsth-Bravell, M. & Pedersen, N. L. Temporal Dynamics of Motor Functioning and Cognitive Aging. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* (2015) doi:10.1093/gerona/glv110.
83. Bégel, V., Verga, L., Benoit, C. E., Kotz, S. A. & Dalla Bella, S. Test-retest reliability of the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA). *Ann. Phys. Rehabil. Med.* (2018) doi:10.1016/j.rehab.2018.04.001.
84. Bonacina, S., Krizman, J., White-Schwoch, T., Nicol, T. & Kraus, N. How Rhythmic Skills Relate and Develop in School-Age Children. *Glob. Pediatr. Heal.* (2019) doi:10.1177/2333794x19852045.
85. Tranchant, P., Vuvan, D. T. & Peretz, I. Keeping the beat: A large sample study of bouncing and clapping to music. *PLoS One* (2016)

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

doi:10.1371/journal.pone.0160178.
86. Tranchant, P. & Peretz, I. Basic timekeeping deficit in the Beat-based Form of Congenital Amusia. *Sci. Rep.* (2020) doi:10.1038/s41598-020-65034-9.
87. Lévy, J. *et al.* Molecular and clinical delineation of 2p15p16. 1 microdeletion syndrome. *Am. J. Med. Genet. Part A* **173**, 2081–2087 (2017).
88. Jones, S. E. *et al.* Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat. Commun.* (2019) doi:10.1038/s41467-018-08259-7.
89. Grahn, J. A. & Rowe, J. B. Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *J. Neurosci.* **29**, 7540–7548 (2009).
90. Grahn, J. A. & Rowe, J. B. Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. *Cereb. Cortex* **23**, 913–921 (2013).
91. Kung, S.-J., Chen, J. L., Zatorre, R. J. & Penhune, V. B. Interacting cortical and basal ganglia networks underlying finding and tapping to the musical beat. *J. Cogn. Neurosci.* **25**, 401–420 (2013).
92. Bengtsson, S. L. *et al.* Listening to rhythms activates motor and premotor cortices. *Cortex.* **45**, 62–71 (2009).
93. Teki, S., Grube, M., Kumar, S. & Griffiths, T. D. Distinct Neural Substrates of Duration-Based and Beat-Based Auditory Timing. *J. Neurosci.* **31**, 3805–3812 (2011).
94. McAuley, J. D., Henry, M. J. & Tkach, J. Tempo mediates the involvement of motor areas in beat perception. *Ann. N. Y. Acad. Sci.* **1252**, 77–84 (2012).
95. Dissanayake, E. If music is the food of love, what about survival and reproductive success? *Music. Sci.* **12**, 169–195 (2008).
96. Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J. & Rodriguez-Fornells, A. Individual differences in music reward experiences. *Music Percept. An Interdiscip. J.* **31**, 118–138 (2013).
97. 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).
98. Haegens, S. & Golumbic, E. Z. Rhythmic facilitation of sensory processing: a critical review. *Neurosci. Biobehav. Rev.* **86**, 150–165 (2018).
99. Sowiński, J. & Dalla Bella, S. Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia* **51**, 1952–1963 (2013).
100. Jacoby, N. *et al.* Cross-Cultural Work in Music Cognition. *Music Percept. An Interdiscip. J.* (2020) doi:10.1525/mp.2020.37.3.185.
101. Devaney, J. Eugenics and Musical Talent: Exploring Carl Seashore's Work on Talent Testing and Performance. *Am. Music Rev.* **48**, (2019).
102. Adam, D. The promise and peril of the new science of social genomics. *Nature* (2019) doi:10.1038/d41586-019-03171-6.
103. Wray, N. R. *et al.* Research Review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry and Allied Disciplines* (2014) doi:10.1111/jcpp.12295.
104. Border, R. *et al.* No support for historical candidate gene or candidate gene-by-

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interaction hypotheses for major depression across multiple large samples. *Am. J. Psychiatry* **176**, 376–387 (2019).
105. Mosing, M. A., Madison, G., Pedersen, N. L. & Ullén, F. Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Dev. Sci.* **19**, 504–512 (2016).
106. Zuk, J., Benjamin, C., Kenyon, A. & Gaab, N. Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One* **9**, e99868 (2014).
107. Ullen, F., Mosing, M. A. & Madison, G. Associations between motor timing, music practice, and intelligence studied in a large sample of twins. *Ann. N. Y. Acad. Sci.* **1337**, 125–129 (2015).
108. Medina, D. & Barraza, P. Efficiency of attentional networks in musicians and non-musicians. *Heliyon* **5**, e01315 (2019).
109. Emery, C. F., Finkel, D. & Pedersen, N. L. Pulmonary Function as a Cause of Cognitive Aging. *Psychol. Sci.* (2012) doi:10.1177/0956797612439422.
110. Finkel, D., Ernsth Bravell, M. & Pedersen, N. L. Role of motor function and lung function in pathways to ageing and decline. *Aging Clin. Exp. Res.* (2020) doi:10.1007/s40520-020-01494-3.
111. Duggan, E. C. *et al.* A Multi-study Coordinated Meta-analysis of Pulmonary Function and Cognition in Aging. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* (2019) doi:10.1093/gerona/glz057.
112. Clouston, S. A. P. *et al.* The dynamic relationship between physical function and cognition in longitudinal aging cohorts. *Epidemiol. Rev.* (2013) doi:10.1093/epirev/mxs004.
113. Larsson, M., Richter, J. & Ravignani, A. Bipedal Steps in the Development of Rhythmic Behavior in Humans. *Music Sci.* (2019) doi:10.1177/2059204319892617.
114. Provasi, J., Anderson, D. I. & Barbu-Roth, M. Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology* (2014) doi:10.3389/fpsyg.2014.01048.
115. Cirelli, L. K. How interpersonal synchrony facilitates early prosocial behavior. *Current Opinion in Psychology* (2018) doi:10.1016/j.copsyc.2017.08.009.
116. Bernard, J. A., Millman, Z. B. & Mittal, V. A. Beat and metaphoric gestures are differentially associated with regional cerebellar and cortical volumes. *Hum. Brain Mapp.* (2015) doi:10.1002/hbm.22894.
117. Gjermunds, N., Brechan, I., Johnsen, S. Å. K. & Watten, R. G. Musicians: Larks, Owls or Hummingbirds? *J. Circadian Rhythms* **17**, 4 (2019).
118. Martin, J., Taylor, M. J. & Lichtenstein, P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychological Medicine* (2018) doi:10.1017/S0033291717003440.
119. Tung, J. Y. *et al.* Efficient replication of over 180 genetic associations with self-reported medical data. *PLoS One* **6**, e23473 (2011).
120. Mansens, D., Deeg, D. J. H. & Comijs, H. C. The association between singing and/or playing a musical instrument and cognitive functions in older adults. *Aging Ment. Heal.* (2018) doi:10.1080/13607863.2017.1328481.
121. Povel, D.-J. & Essens, P. Perception of temporal patterns. *Music Percept. An*

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*Interdiscip. J.* **2**, 411–440 (1985).
122. Gordon, R. L., Jacobs, M. S., Schuele, C. M. & Mcauley, J. D. *Perspectives on the rhythm-grammar link and its implications for typical and atypical language development*. *Annals of the New York Academy of Sciences* vol. 1337 (2015).
123. Wieland, E. A., McAuley, J. D., Dilley, L. C. & Chang, S.-E. Evidence for a rhythm perception deficit in children who stutter. *Brain Lang.* **144**, 26–34 (2015).
124. Woods, K. J. P., Siegel, M. H., Traer, J. & McDermott, J. H. Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, Psychophys.* **79**, 2064–2072 (2017).
125. Macmillan, N. A. & Creelman, C. D. *Detection theory: a user's guide*. (Cambridge University Press, 1991).
126. Gordon, R. L. *et al.* Musical rhythm discrimination explains individual differences in grammar skills in children. *Dev. Sci.* (2015) doi:10.1111/desc.12230.
127. Berinsky, A. J., Margolis, M. F. & Sances, M. W. Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *Am. J. Pol. Sci.* (2014) doi:10.1111/ajps.12081.
128. Anwyl-Irvine, A. L., Dalmaijer, E. S., Hodges, N., & Evershed, J. Online Timing Accuracy and Precision: A comparison of platforms, browsers, and participant's devices. *PsyArXiv*.
129. Bridges, D., Pitiot, A., MacAskill, M. R. & Peirce, J. W. The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ* (2020) doi:10.7717/peerj.9414.
130. McKinney, M. F., Moelants, D., Davies, M. E. P. & Klapuri, A. Evaluation of audio beat tracking and music tempo extraction algorithms. *J. New Music Res.* (2007) doi:10.1080/09298210701653252.
131. Repp, B. H. Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 1. Qualitative observations. *Music Percept.* (2005) doi:10.1525/mp.2005.22.3.479.
132. Repp, B. H. & Su, Y. H. Sensorimotor synchronization: A review of recent research (2006-2012). *Psychon. Bull. Rev.* (2013) doi:10.3758/s13423-012-0371-2.
133. London, J. *Hearing in Time: Psychological Aspects of Musical Meter*. *Hearing in Time: Psychological Aspects of Musical Meter* (2012). doi:10.1093/acprof:oso/9780199744374.001.0001.
134. 3.5.1., R. D. C. T. A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* (2018).
135. Durand, E. Y., Do, C. B., Mountain, J. L. & Macpherson, J. M. Ancestry Composition: A Novel, Efficient Pipeline for Ancestry Deconvolution. *bioRxiv* 10512 (2014) doi:10.1101/010512.
136. Henn, B. M. *et al.* Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS One* **7**, e34267 (2012).
137. The Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68 (2015).
138. consortium, U. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82 (2015).
139. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).

140. Loh, P.-R., Palamara, P. F. & Price, A. L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat. Genet.* **48**, 811 (2016).
141. Vernot, B. *et al.* Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science (80-. ).* **352**, 235–239 (2016).
142. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* (2010) doi:10.1093/bioinformatics/btq033.
143. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Tables

Table 1. Sample demographics for each of the three cohorts (GWAS, Phenotype Experiment 1, and Phenotype Experiment 2).

**GWAS Sample by Phenotype group (response to Clap-to-beat question)**

| | Cases (Yes) | Controls (No) |
|---|---|---|
| Total | 555660 | 51165 |
| Males | 226188 | 23998 |
| Females | 329472 | 27167 |
| 18 to 30 years old | 57898 | 5186 |
| 30 to 45 years old | 135168 | 12909 |
| 45 to 60 years old | 150939 | 13312 |
| 60 years old and over | 211655 | 19758 |

**Phenotype Validation Experiment 1 – rhythm perception**

| *Full sample who provided demographics* | N | Mean Age in years (for N=722 who reported demographics) | SD Age |
|---|---|---|---|
| Total | 722 | 36.03 | 10.98 |
| Males | 386 | 34.86 | 10.74 |
| Females | 332 | 37.49 | 11.07 |

**Phenotype Validation Experiment 2 – beat production and cross-trait**

| *Full sample* | N | Mean Age in years | SD Age |
|---|---|---|---|
| Total | 1412 | 36.34 | 11.93 |
| Males | 678 | 35.53 | 11.12 |
| Females | 728 | 37.15 | 12.61 |
| *Subset with valid tapping data* | n | Mean Age in years | SD Age |
| Total | 542 | 35.24 | 11.39 |
| Males | 241 | 35.02 | 10.62 |
| Females | 300 | 35.43 | 12.00 |

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Table 2. Genomic loci and sentinel SNPs significantly associated with beat synchronization in the primary GWAS.** Further details (e.g., chromosomal location) are provided in Supplementary Table 2.

| Genomic Locus | Sentinel SNP | chr | A1 | MAF | OR | SE | p-value | gene symbol |
|---|---|---|---|---|---|---|---|---|
| 11 | rs848293 | 2 | G | 0.42228 | 0.94 | 0.007 | 9.23E-18 | *VRK2* |
| 26 | rs62340585 | 4 | G | 0.20695 | 0.94 | 0.008 | 1.81E-14 | *GPM6A* |
| 13 | rs10168817 | 2 | G | 0.49299 | 0.95 | 0.007 | 1.94E-14 | *NA* |
| 20 | rs10779987 | 3 | T | 0.38175 | 0.95 | 0.007 | 2.21E-14 | *GBE1* |
| 28 | rs28392605 | 5 | G | 0.33904 | 0.95 | 0.007 | 8.93E-14 | *NA* |
| 45 | rs1832909 | 9 | T | 0.40687 | 1.05 | 0.007 | 1.78E-13 | *NA* |
| 2 | rs34762587 | 1 | T | 0.31379 | 1.05 | 0.007 | 2.25E-13 | *FOXO6* |
| 60 | rs7542 | 16 | G | 0.46184 | 0.95 | 0.007 | 2.41E-13 | *MAPK3* |
| 5 | rs10875125 | 1 | C | 0.15305 | 0.93 | 0.009 | 2.61E-13 | *DPYD* |
| 35 | rs9400241 | 6 | C | 0.28851 | 0.95 | 0.007 | 4.49E-13 | *FOXO3* |
| 64 | rs4792891 | 17 | T | 0.34013 | 1.05 | 0.007 | 7.07E-13 | *MAPT* |
| 39 | rs1468701 | 7 | G | 0.29172 | 0.95 | 0.007 | 3.62E-12 | *SND1* |
| 50 | rs10848650 | 12 | G | 0.42192 | 1.05 | 0.007 | 6.04E-12 | *SLC6A13* |
| 29 | rs2635634 | 5 | T | 0.45317 | 1.05 | 0.007 | 9.54E-12 | *CDH12* |
| 67 | rs9626920 | 22 | G | 0.41282 | 1.05 | 0.007 | 1.04E-11 | *MIRLET7BHG* |
| 16 | rs764299 | 2 | G | 0.26719 | 1.05 | 0.007 | 1.47E-11 | *PLEKHM3* |
| 43 | rs10984506 | 9 | T | 0.36558 | 1.05 | 0.007 | 1.66E-11 | *ANP32B* |
| 53 | rs1426371 | 12 | G | 0.25919 | 1.05 | 0.008 | 1.67E-11 | *WSCD2* |
| 58 | rs12913592 | 15 | T | 0.3596 | 1.05 | 0.007 | 6.13E-11 | *NA* |
| 6 | rs72700870 | 1 | G | 0.14377 | 0.94 | 0.009 | 1.42E-10 | *MCL1* |
| 34 | rs9388171 | 6 | G | 0.47595 | 0.96 | 0.007 | 2.16E-10 | *NA* |
| 55 | rs6572878 | 14 | T | 0.39477 | 0.96 | 0.007 | 3.48E-10 | *HAUS4* |
| 4 | rs11210206 | 1 | T | 0.31286 | 0.96 | 0.007 | 3.93E-10 | *NA* |
| 28 | rs72633496 | 5 | T | 0.43224 | 0.95 | 0.008 | 6.21E-10 | *NA* |
| 10 | rs7586405 | 2 | G | 0.30559 | 1.04 | 0.007 | 7.19E-10 | *PPP1CB* |
| 63 | rs3024293 | 17 | T | 0.23528 | 1.05 | 0.008 | 8.26E-10 | *C1QL1* |
| 1 | rs2061843 | 1 | G | 0.4001 | 0.96 | 0.007 | 1.19E-09 | *CSMD2* |
| 19 | rs1349028 | 3 | T | 0.25977 | 0.95 | 0.008 | 1.54E-09 | *EIF4E3* |
| 25 | rs4443239 | 4 | T | 0.2463 | 1.05 | 0.008 | 1.68E-09 | *C4orf27* |
| 33 | rs1901739 | 5 | T | 0.47772 | 1.04 | 0.007 | 2.14E-09 | *NA* |
| 7 | rs55678522 | 1 | G | 0.21629 | 0.95 | 0.008 | 2.25E-09 | *LRRN2* |
| 61 | rs8079923 | 17 | T | 0.25309 | 1.05 | 0.008 | 2.88E-09 | *AKAP10* |
| 62 | rs7501911 | 17 | T | 0.18191 | 0.95 | 0.009 | 3.34E-09 | *NLK* |
| 66 | rs6087848 | 20 | G | 0.44304 | 0.96 | 0.007 | 3.40E-09 | *POFUT1* |
| 54 | rs10744255 | 12 | G | 0.23229 | 0.96 | 0.008 | 4.24E-09 | *NA* |
| 31 | rs13163173 | 5 | C | 0.16597 | 0.95 | 0.009 | 4.51E-09 | *MEF2C* |
| 3 | rs2819333 | 1 | T | 0.37068 | 1.04 | 0.007 | 4.54E-09 | *PTPRF* |
| 51 | rs2453873 | 12 | G | 0.22254 | 0.95 | 0.008 | 5.17E-09 | *NA* |
| 27 | rs67264739 | 5 | G | 0.27395 | 0.96 | 0.007 | 5.54E-09 | *ADCY2* |
| 56 | rs2284901 | 14 | G | 0.37485 | 1.04 | 0.007 | 6.48E-09 | *AKAP6* |
| 32 | rs1596431 | 5 | T | 0.19182 | 1.05 | 0.008 | 7.42E-09 | *NA* |
| 44 | rs10978661 | 9 | T | 0.12006 | 0.94 | 0.01 | 7.74E-09 | *ZNF462* |
| 23 | rs4263335 | 4 | G | 0.49483 | 1.04 | 0.007 | 8.74E-09 | *JAKMIP1* |
| 48 | rs7939759 | 11 | T | 0.23981 | 1.05 | 0.008 | 1.23E-08 | *CTSF* |
| 65 | rs9710427 | 19 | G | 0.41536 | 1.04 | 0.007 | 1.32E-08 | *TECR* |
| 21 | rs12638746 | 3 | G | 0.33546 | 0.96 | 0.007 | 1.37E-08 | *EPHA3* |
| 59 | rs12909047 | 15 | G | 0.48251 | 1.04 | 0.007 | 1.49E-08 | *UBL7* |
| 46 | rs2505344 | 10 | G | 0.17674 | 0.95 | 0.009 | 1.51E-08 | *EPC1* |
| 24 | rs67816799 | 4 | C | 0.38188 | 1.04 | 0.007 | 1.56E-08 | *CCSER1* |

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 15 | rs10932201 | 2 | G | 0.46351 | 1.04 | 0.007 | 1.59E-08 | *CREB1* |
| 49 | rs526904 | 11 | T | 0.34865 | 0.96 | 0.007 | 1.60E-08 | *PICALM* |
| 9 | rs6548147 | 2 | T | 0.4402 | 1.04 | 0.007 | 2.05E-08 | *TSSC1* |
| 52 | rs10877461 | 12 | G | 0.29968 | 0.96 | 0.008 | 2.44E-08 | *NA* |
| 41 | rs11996434 | 8 | G | 0.27037 | 1.04 | 0.008 | 2.61E-08 | *NA* |
| 40 | rs1996148 | 8 | G | 0.31961 | 0.96 | 0.007 | 2.69E-08 | *PEBP4* |
| 47 | rs10885458 | 10 | G | 0.28314 | 0.96 | 0.007 | 2.69E-08 | *NA* |
| 17 | rs191373913 | 2 | T | 0.43899 | 0.96 | 0.007 | 2.74E-08 | *NGEF* |
| 38 | rs12056186 | 7 | C | 0.42875 | 0.96 | 0.007 | 2.93E-08 | *ORC5* |
| 42 | rs7856850 | 9 | C | 0.22184 | 0.96 | 0.008 | 3.07E-08 | *PTPRD* |
| 36 | rs13197257 | 6 | T | 0.27444 | 0.96 | 0.007 | 3.23E-08 | *PTPRK* |
| 14 | rs10497355 | 2 | T | 0.46078 | 1.04 | 0.007 | 3.43E-08 | *UBR3* |
| 12 | rs11692449 | 2 | T | 0.37522 | 1.04 | 0.007 | 3.45E-08 | *XPO1* |
| 30 | rs4704043 | 5 | T | 0.2827 | 1.04 | 0.007 | 3.65E-08 | *TNPO1* |
| 18 | rs43182 | 3 | T | 0.13443 | 1.06 | 0.01 | 3.80E-08 | *PTPRG* |
| 57 | rs62014217 | 15 | G | 0.20132 | 0.96 | 0.008 | 3.91E-08 | *HERC1* |
| 8 | rs476141 | 1 | T | 0.49868 | 1.04 | 0.007 | 4.49E-08 | *NA* |
| 37 | rs2849543 | 6 | G | 0.41591 | 1.04 | 0.007 | 4.60E-08 | *PARK2* |
| 22 | rs571760466 | 3 | C | 0.27511 | 0.96 | 0.007 | 4.81E-08 | *LSAMP* |

Abbreviations: SNP=Single Nucleotide Polymorphism, chr=Chromosome, A1=effect allele, MAF=Minor Allele Frequency, OR=Odds Ratio, S.E.=Standard Error, Notes: Gene symbol is based on HUGO (HGNC). These are all genes annotated to SNPs in r2> 0.1 with the lead SNP; sentinel SNP in a given locus refers to independent SNP from FUMA. The SNPs were mapped to genes based on ANNOVAR annotation and on being physically located inside a protein coding gene using 10kb window. NA=when the SNP is not within the 10kb window of a gene.

22

bioRxiv preprint doi: https://doi.org/10.1101/836197; this version posted January 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.