

 Open access • Posted Content • DOI:10.1101/836197

Unravelling the genetic architecture of musical rhythm: a large-scale genome-wide association study of beat synchronization — [Source link](#)

Maria Niarchou, Daniel E. Gustavson, J. Fah Sathirapongsasuti, Manuel Anglada-Tort ...+14 more authors

Institutions: Vanderbilt University Medical Center, Max Planck Society, Trinity College, Dublin, Michigan State University ...+4 more institutions

Published on: 19 Jan 2021 - [bioRxiv](#) (Cold Spring Harbor Laboratory)

Related papers:

- [Genome-wide association study of musical beat synchronization demonstrates high polygenicity](#)
- [Genome-Wide Association Analysis for Oat Genetics Using Support Vector Machines](#)
- [Multi-ethnic genome-wide association study of decomposed cardioelectric phenotypes illustrates strategies to identify and characterize evidence of shared genetic effects for complex traits](#)
- [A global view of pleiotropy and genetic architecture in complex traits](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/unravelling-the-genetic-architecture-of-musical-rhythm-a-2gnzuy0hjk>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

Unravelling the genetic architecture of musical rhythm

**Maria Niarchou^{1,2*}, J. Fah Sathirapongsasuti³, Nori Jacoby^{4,5},
Eamonn Bell⁶, Evonne McArthur², Peter Straub², The 23andMe Research Team^{3^},
J. Devin McAuley⁷, John A. Capra^{8,9,10,2}, Fredrik Ullén¹¹, Nicole Creanza⁸,
Miriam A. Mosing^{11,12}, David Hinds³, Lea K. Davis^{2,13,10,14,15*†},
Reyna L. Gordon^{16,17,2*†}**

¹Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, UK

²Vanderbilt Genetics Institute, Vanderbilt University Medical Center, TN, USA

³23andMe, Inc., Sunnyvale, CA, USA

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

⁵Center for Science and Society, Columbia University, USA

⁶Music Department, Columbia University, USA

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

⁸Department of Biological Sciences, Vanderbilt University, TN, USA

⁹Department of Computer Science, Vanderbilt University, TN, USA

¹⁰Department of Biomedical Informatics, Vanderbilt University Medical Center, TN, USA

¹¹Department of Neuroscience, Karolinska Institutet, Sweden

¹²Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden

¹³Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, TN, USA

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

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

¹⁶Department of Otolaryngology, Vanderbilt University Medical Center, TN, USA

¹⁷Department of Psychology, Vanderbilt University, TN, USA

[^]full list of authors in Supplementary Note

^{*}Correspondence to: Reyna L. Gordon (reyna.gordon@Vanderbilt.Edu), Lea Davis (lea.k.davis@gmail.com), Maria Niarchou (maria.niarchou@vumc.org)

[†]Joint supervision

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

Abstract

While timing and rhythm-related phenotypes are heritable, the human genome variations underlying these traits are not yet well-understood. We conducted a genome-wide association study to identify common genetic variants associated with a self-reported musical rhythm phenotype in 606,825 individuals. Rhythm exhibited a highly polygenic architecture with sixty-eight loci reaching genome-wide significance ($p < 5 \times 10^{-8}$) and SNP-based heritability of 13%-16%. Polygenic scores for rhythm predicted the presence of musician-related keywords in the BioVU electronic health record biobank. Genetic associations with rhythm were enriched for genes expressed in brain tissues. Genetic correlation analyses revealed shared genetic architecture with several traits relevant to cognition, emotion, health, and circadian rhythms, paving the way to a better understanding of the neurobiological pathways of musicality.

51 **Introduction**

52 Rhythm is a fundamental aspect of music across cultures ^{1,2} and more broadly,
53 rhythmic patterns provide sensori-motor structure to human interactions. Our tendency
54 to perceive, create, manipulate, and appreciate rhythms in a variety of contexts (e.g.,
55 speech, music, movement) is part of what makes us human. Even very young children
56 are sensitive to the social and linguistic signals carried by rhythm ³, thus it is not
57 surprising that parents use rhythmic vocalizations and synchronous movement (e.g.,
58 lullabies and rocking) to interact with their infants from birth ⁴. Moving in synchrony to a
59 musical beat (“beat synchronization”) appears to be a key feature of human musical
60 experiences throughout the lifespan ⁵⁻⁷.

61 Although most people are able to effortlessly detect and synchronize with the
62 beat even without musical training ^{6,7}, there is substantial inter-individual variability
63 (within cultures) in the extent to which individuals can perceive and produce musical
64 rhythm accurately ⁸⁻¹⁰. While the neuroimaging literature points to auditory-motor
65 networks in the brain underlying rhythm perception and production ¹¹, less is known
66 about the genetic underpinnings that give rise to individual differences in these
67 networks. Heritability estimates from family-based studies, using a variety of measures
68 relevant for rhythmic ability, range from 21% ¹² to 50% ¹³. There is a gap in knowledge
69 about genomic loci underlying variation in rhythm ability ¹⁴, in part due to the challenge
70 of assessing the rhythm phenotype in a sample large enough to provide sufficient power
71 to detect common variants with small effects, as expected for complex traits ¹⁵.

72 *Summary of Approach.*

73 We conducted a genome-wide association study (GWAS) to identify common
74 genetic variants associated with a self-reported musical rhythm phenotype, i.e. “Can
75 you clap in time with a musical beat?”, collected from 606,825 individuals participating in
76 research with the personal genetics company 23andMe, Inc. We then validated this self-
77 reported phenotype in a separate internet-based behavioural study conducted in 724
78 individuals and found that it was significantly correlated with rhythm perception (Online
79 Methods). In the GWAS, a total of 68 independent SNPs surpassed the threshold for
80 genome-wide significance ($p < 5 \times 10^{-08}$). In addition to determining which genes were
81 implicated by these variants, we estimated how much of the total phenotypic variance
82 could be explained by all variation across the genome (i.e., SNP-based heritability). We
83 then further explored this heritability to test the hypothesis that variants associated with
84 rhythm were enriched among genes expressed in brain compared to genes expressed
85 in other tissues (e.g., muscle, adipose, etc.), and furthermore enriched in genes
86 expressed in neurons compared to other brain cell types (e.g., oligodendrocyte,
87 astrocyte).

88 Using an independent sample of 67,441 genotyped individuals from the
89 Vanderbilt University Medical Center biobank, BioVU, we tested whether a cumulative
90 sum of the genetic effects for rhythm detected in our GWAS (i.e., rhythm polygenic
91 score), was significantly associated with an indication of musician status in the
92 electronic health record (EHR). Because little is yet known about the relationship of the
93 genomics of rhythm to other traits, we also performed exploratory genetic correlation
94 analysis including 764 complex traits for which a well-powered GWAS has been
95 performed and deposited in LDHub¹⁶. Finally, we evaluated the contribution to rhythm

96 of regions of the genome that have experienced significant human-specific evolutionary
97 shifts (since the divergence of humans and chimpanzees from their last common
98 ancestor, ~6 million years ago).

99 **Results**

100 **Validating the self-reported rhythm phenotype**

101

102 The study population is N=606,825 participants of European ancestry (59% females,
103 mean age(SD)=52.09(18.5) years, who consented to participate in research with
104 23andMe, Inc. Data were available from individuals who answered the question “Can
105 you clap in time with a musical beat?” The majority of participants answered ‘Yes’
106 (91.57%) and 8.43% answered ‘No’, which is slightly higher than the estimated
107 population prevalence of poor rhythm at ~5%^{17,18} (Table 1 and Supplementary Note). In
108 light of prior work suggesting that human rhythm is a complex trait that can be quantified
109 with both objective and self-report measures¹⁰, we sought to validate the self-report
110 question against an objective measure of rhythm perception. We conducted a
111 phenotype validation study with a sample (N=724; mean age=36 years, SD=10.9; 46%
112 females) recruited anonymously from Amazon’s Mechanical Turk. Participants
113 performed an objective musical rhythm perception test and were asked “Can you clap in
114 time with a musical beat?” (details provided in Online Methods). In each of the 32 trials,
115 participants had to judge whether a pair of rhythms were the same or different, following
116 a standard procedure for assessing individual differences in musical perception ability⁹
117 and utilizing rhythm sequences with simple (highly metrical) and complex (syncopated)
118 rhythms¹⁹. Individuals who had better performance on discriminating musical rhythms
119 were more likely to answer ‘Yes’ to the self-report synchronization question than those
120 who answered ‘No’ (OR(95%CI)=1.94(1.28 to 3.01), p=0.002, McFadden’s R²=0.39 (i.e.,
121 we expect to see a 94% increase in the odds of answering ‘Yes’, for a standard

122 deviation increase in the rhythm discrimination test). In the remainder of the paper, the
123 “rhythm” trait in our study refers to the self-reported beat synchronization phenotype.

124

125 **GWAS results and heritability estimation**

126 GWAS was conducted using logistic regression under an additive genetic model, while
127 adjusting for age, sex, the top five principal components of ancestry in order to control
128 for population stratification, and indicators for genotype platforms to account for batch
129 effects. We excluded SNPs with Minor Allele Frequency (MAF) <0.01, low imputation
130 quality ($R^2 < 0.3$) and indels, resulting in a final set of 8,288,851 SNPs for all subsequent
131 analyses. Sixty-eight independent SNPs (after two rounds of LD pruning, first at $r^2 = 0.6$
132 and then at $r^2 = 0.1$, kb = 250) reached genome-wide significance ($p < 5 \times 10^{-8}$; Figure 1,
133 Supplementary Table 1, Supplementary Figure 1), from a total of 6,115 SNPs that
134 passed the significance threshold.

135

136 Linkage Disequilibrium Score Regression (LDSC)²⁰ analyses revealed that heritability
137 estimates on the liability scale ranged from 13% to 16% when adjusted for a range of
138 estimated population prevalence of rhythm deficits (from 3.5% to 6.5%^{17,18})
139 (Supplementary Table 2, Supplementary Note). The observed SNP-heritability explained
140 5% (se=0.0002) of the phenotypic variance in the rhythm trait, with an LD score
141 regression intercept of 1.02 (se=0.01).

142

143 **Gene-based analyses.** Gene-based association analyses performed with MAGMA
144 yielded 203 genes that surpassed the threshold of $p < 3 \times 10^{-6}$ (Supplementary Table 3).

145 The top two genes are: *CCSER1*, in proximity to genes previously associated with
146 musicality²¹, and *VRK2* (converging with the top locus identified in our SNP-based
147 association analyses).

148
149 We also examined potential replication of genetic associations with musicality in
150 humans from prior reports (28 genes were selected including 26 reported in meta-
151 analysis by²¹, and additionally, *GATA2* and *PCDH7*²² and *UGT8*²³. Although none of
152 the genes reached statistical significance (Supplementary Table 4, Supplementary
153 Note), several are located near *CCSER1* in the 4q22-24 region.

154 155 **Heritability Partitioning**

156 One advantage to SNP-based heritability estimation is the ability to partition heritability
157 according to SNP-annotations, which provides insight into the types of genetic variation
158 that contribute most to rhythm. To determine whether heritability is enriched for specific
159 functional categories of SNP annotations, stratified LDSC²⁴ was used to partition
160 heritability (Supplementary Table 5). We hypothesized that SNPs falling into regions of
161 open chromatin (i.e., accessible to transcriptional machinery), and regions with human-
162 specific variation, would be enriched for rhythm-associated variation. We found
163 enrichment in regions conserved in mammals (regions of the genome identified by
164 Lindblad-Toh et al. 2013 as being under purifying selection) (enrichment=15.8, $p=1.19 \times$
165 10^{-12}) and in functional categories involved in acetylation of histone H3 at lysine 9
166 (H3K9ac) (enrichment=8.0, $p=1.85 \times 10^{-8}$) and monomethylation of histone H3 at lysine
167 4 (H3K4me3) (enrichment=1.29, $p=2.16 \times 10^{-5}$), supporting associations mediated by

168 effects on gene regulation. Enrichment was also found in the ‘Repressed’ category of
169 chromatin states (enrichment=0.87, $p=0.0002$), and for introns. We also examined
170 whether genes expressed in specific cell-types show enrichment among rhythm-
171 associated variants as described in ²⁴: and found that genes expressed specifically in
172 neurons contributed significantly to trait heritability (coefficient= 1.19×10^{-9} , $p=0.037$)
173 conditional to the other annotations (Supplementary Table 6).

174

175 **Gene set analyses**

176 Using FUMA ²⁵, we performed a gene-property analysis where the average expression
177 of the genes per tissue type (using GTEx gene expression panels in 53 tissue types ²⁶
178 was added as a covariate in the model. As predicted, gene associations were
179 significantly enriched in brain tissue compared to non-brain tissues (Figure 2). To further
180 examine potential biological pathways associated with rhythm, we performed MAGMA
181 gene-set analyses as implemented by FUMA ²⁵. Two gene-sets out of 10,678 achieved
182 statistical significance after Bonferroni correction (Supplementary Table 7). The top
183 associated gene-sets with rhythm were: Negative regulation of transcription from RNA
184 polymerase II promoter ($p=8.6 \times 10^{-07}$), gene-set from the Gene Ontology project
185 ^{27,28}(i.e., any process that includes glucose and decreases the rate or frequency of
186 transcription from an RNA polymerase II promoter) and Negative regulation of gene
187 expression (2.9×10^{-6}).

188

189 **Human Accelerated Region and Neanderthal Introgression Stratified Heritability**

190 **Analyses**

191 Given previous hypotheses about the origins of rhythm^{6,7,29}, we evaluated the
192 contribution of regions of the human genome that have experienced significant human-
193 specific shifts in evolutionary pressure using stratified LDSC^{20,24}. In particular, we
194 analyzed the contribution to rhythm heritability from variants in genomic loci that are
195 conserved across non-human species, but have elevated substitution rate on the
196 human lineage³⁰. Many of these human accelerated regions (HARs) play roles in
197 human-specific traits³¹, including cognition³². The heritability of rhythm is enriched 2.26-
198 fold in variants in perfect linkage disequilibrium with HARs ($p = 0.14$). However, given
199 the small number of variants in these regions and the enrichment of HARs in functional
200 regions of the genome, it is difficult to explicitly link these shifts to rhythm. Nonetheless,
201 two of the variants most strongly associated with rhythm (rs14316, rs1464791) fall
202 within HARs, and the rs1464791 variant is near *GBE1*, a gene associated with a range
203 of traits including body-mass index (BMI)³³ and cognitive deficits³⁴.

204 We also evaluated the contribution of genetic variants detected in the
205 Neanderthal genome present in modern Eurasians due to interbreeding (hereafter
206 “Neanderthal variants”) to the heritability of the rhythm phenotype. Eurasian genomes
207 contain ~1.5-4% of DNA as a result from interbreeding with Neanderthals around
208 50,000 years ago. Heritability of rhythm was significantly depleted among Neanderthal
209 variants (1.97-fold depletion, $P = 0.001$). However, Neanderthal ancestry is significantly
210 depleted in functional genomic regions overall³⁵, therefore, the depletion of rhythm
211 heritability in these regions is likely the result of the overall depletion for Neanderthal
212 ancestry in functional regions of the genome. This is supported by a non-significant τ_c^* ,
213 illustrating that Neanderthal vs. human variants do not provide unique heritability when

214 conditioned on a broad set of regulatory elements ³⁶(Supplementary Table 8, Online
215 Methods).

216

217 **Proof-of-concept of the genetics of musicality in a health care context**

218 As a proof-of-concept that genetics of rhythm are more widely tied to the biology of
219 musicality, we further examined whether the contribution of the common alleles
220 associated with rhythm en masse (also known as polygenic scores (PGS)) predict the
221 presence of keywords indicating “musician” status in clinical documentation collected in
222 the electronic health record (see Supplementary Note for details). In a sample of 67,441
223 individuals in Vanderbilt’s BioVU, we identified 864 individuals with the keyword
224 “musician” (or other closely related keywords for musical instruments) present in the
225 EHR that we compared with 66,577 without any mention of “musician” keywords in their
226 EHR. We found evidence that the PGS for rhythm was significantly higher among
227 individuals with the “musician” keywords in their chart (OR per SD increase in PGS,
228 1.30, 95%CI:1.20-1.38, $p < 2.5 \times 10^{-13}$, Nagelkerke’s $R^2=1\%$) (Supplementary Table 9,
229 Figure 3), confirming our hypothesis that the rhythm phenotype assessed in our study
230 captures a dimension of musicality.

231

232

233

234 **Rhythm beyond the contribution of intelligence**

235 In light of previous work linking rhythm and IQ ^{17,37}, we used multi-trait conditional
236 joint analysis ³⁸ (mtcojo) to remove shared genetic effects between intelligence and

237 rhythm. This analysis generated a new set of summary statistics of rhythm in which
238 betas, standard errors and p-values were adjusted based on the intelligence summary
239 statistics from ³⁹. Using FUMA as described above, we identified 66 independent,
240 genome-wide significant loci in the conditioned GWAS, all of which were within 5kb of
241 the loci in the unadjusted rhythm summary statistics (Supplementary Table 10,
242 Supplementary Figure 2). We also compared effect estimates in 47 independent,
243 genome-wide significant SNPs available from both the unadjusted rhythm and IQ
244 GWAS datasets; all that were in common between these two datasets remained
245 significant at the GWAS threshold, and their effect estimates were not changed
246 (Supplementary Table 11). Also, the genetic correlation between the IQ GWAS dataset
247 and rhythm was not significant ($r_g = -0.003$ (standard error = 0.02), $p = 0.88$). Similarly, the
248 estimates of the heritability in the liability scale remained the same (13% to 16%). These
249 findings indicate that our results are largely driven by associations with rhythm rather
250 than cognitive ability.

251 Table 2 shows the rhythm-related loci that are also present in the GWAS
252 catalogue after adjusting for genetic effects shared with IQ (for a full list of loci see
253 Supplementary Table 12).

254

255 **Cross-trait analyses**

256 To determine if rhythm shares genetic architecture with other traits, we tested
257 genetic correlations ²⁰ between rhythm and all 764 available traits in LDHub (v.1.9.2)
258 using LDscore regression. This method is designed to show whether there is shared
259 genetic variation linked to a particular trait (here, our rhythm trait) and traits measured in

260 other samples/studies. There were 31 statistically significant genetic correlations ($p < 6.5$
261 $\times 10^{-5}$) between rhythm and other traits after adjusting for multiple comparisons (Figure
262 4, Supplementary Table 13).

263
264 As expected, processing speed measured as 'mean time to correctly identify
265 matches' was negatively correlated with rhythmic ability ($r_g = -0.16$, $p = 3.22 \times 10^{-13}$) (i.e.,
266 faster processing speed was associated with having rhythm). Educational qualifications
267 (O' levels/GCSEs or equivalent) ($r_g = 0.16$, $p = 4.6 \times 10^{-7}$), evening chronotype ($r_g = 0.09$,
268 $p = 3.8 \times 10^{-5}$) and tinnitus ($r_g = 0.20$, $p = 6.7 \times 10^{-6}$) were all positively associated with
269 rhythm. While falling short of the correction for multiple testing, exposure to loud music
270 was also correlated with a similar point estimate ($r_g = 0.20$, $p = 2.0 \times 10^{-4}$) and could be
271 due to a relationship between tinnitus and loud music exposure in the UKBB ($r_g = 0.30$,
272 $p = 4.8 \times 10^{-6}$)^{36,40}.

273 Additionally, we identified significant genetic correlations between rhythm and
274 hand grip strength ($r_g(\text{left}) = 0.18$, $se = 0.02$, $p = 3.6 \times 10^{-16}$, $r_g(\text{right}) = 0.16$, $se = 0.02$, $p = 6.91$
275 $\times 10^{-15}$), smoking including 'ever smoked' ($r_g = 0.16$, $p = 2.5 \times 10^{-11}$) and 'past tobacco
276 smoking' ($r_g = -0.15$, $p = 4.6 \times 10^{-10}$) as well as with peak expiratory flow from both the
277 UKBiobank ($r_g = 0.15$, $p = 2.11 \times 10^{-9}$) and a second independent GWA study ($r_g = 0.11$,
278 $p = 6.6 \times 10^{-8}$) and several other lung-related phenotypes (Supplementary Table 13).
279 Given that the majority of these traits come from the UKBiobank, it is also possible that
280 their genetic correlations with rhythm, may be a function of their correlation with each
281 other, as some degree of phenotypic correlation is also expected.

282 Recent studies illustrate the potential for very subtle residual population
283 substructure to influence some polygenic analyses⁴¹ including genetic correlations.
284 Therefore, we also adjusted the rhythm associations for SNP-loadings on the first
285 principal component of ancestry estimated from 1KG European populations. We then
286 used these SNP estimates of ancestry to adjust the rhythm GWAS results which yielded
287 no change in the genetic correlations results⁴¹ (Supplementary Table 14 and
288 Supplementary Note).

289 Although we cannot determine potential causality, we conducted MR analyses
290 using the GSMR³⁸ to examine whether there are significant bi-directional relationships
291 between rhythm and processing speed, handgrip strength, and chronotype
292 (Supplementary Note). We found significant bidirectional relationships for all traits in
293 the analysis (Supplementary Table 15).

294

295 **Sensitivity Analysis of Chromosome 17 locus for chromosomal inversions and** 296 **Parkinson's Disease**

297 Given that the genome-wide significant locus (lead SNP rs4792891) on
298 chromosome 17q21 is located within a well-established inversion region that may also
299 be associated with local population substructure⁴², we conducted additional analyses
300 focused on the region. The inversion was not associated with local ancestry within our
301 study sample (Supplementary Table 16), suggesting that the association between this
302 locus and rhythm is not likely to be due to local population confounding.

303 In addition, we sought to explore the potential effect of Parkinson's disease (PD)
304 phenotype on this Microtubule Associated Protein Tau (*MAPT*) locus (17q21). Taking

305 into account that PD patients may have difficulty discriminating beat-based rhythms⁴³,
306 and also that PD patients are over-represented in the 23andMe database, it was
307 possible that the inclusion of PD patients in the sample may account for these
308 associations. The associations between the independent SNP in the locus, rs4792891,
309 and rhythm remained after removing PD patients from the sample, indicating that this
310 MAPT association with rhythm is not driven by PD cases (Supplementary Note,
311 Supplementary Table 17).
312

313

Discussion

314 This study demonstrates that common genetic variation plays a role in a musical
315 rhythm trait, complementing prior evidence of innate human rhythm sensitivity^{6,7}. Based
316 on a self-reported beat synchronization phenotype that was validated with an objective
317 measure of rhythm perception, the present large-scale study (606,825 participants from
318 23andMe) is a significant first step towards well-powered genomic evidence of a
319 musicality phenotype. Sixty-eight independent SNPs (Supplementary Table 1)
320 surpassed the threshold for genome-wide significance, with the top-associated locus
321 mapped to *VRK2-FANCL* (rs848293, $p=9.2 \times 10^{-18}$), a protein kinase with multiple
322 spliced isoforms expressed in brain that was previously associated with behavioural and
323 psychiatric phenotypes (i.e., depression, neuroticism and schizophrenia⁴⁴⁻⁴⁶
324 developmental delay)⁴⁷, indicating a biological connection between rhythm and
325 neurodevelopment.

326 The total SNP-based heritability of our rhythm trait on the liability scale ranged
327 from 13 to 16%, in line with both estimates of other complex traits (e.g., asthma⁴⁸) and
328 previously reported heritability estimates of musical rhythm abilities reported in twins¹³.
329 Enrichment of heritability of rhythm in multiple brain tissues, notably cerebellum, basal
330 ganglia, and cortex, likely reflects the genetic contribution to subcortical-cortical
331 networks underlying musical rhythm perception and production¹¹. Indeed, brain
332 structures associated with rhythm include basal ganglia⁴⁹⁻⁵¹, cerebellum^{52,53} and
333 thalamus⁵⁴. Furthermore, we found heritability enriched in genes expressed in neuronal
334 cell types and in SNPs and genes responsible for expression regulation; taken together,

335 these results suggest that genomic loci that influence rhythm are enriched for effects on
336 the brain and mediated by regulation of gene expression.

337 Initial clues about the evolution of rhythm traits in humans may be indicated by
338 the occurrence of two of the rhythm-associated loci in human-accelerated regions
339 (HARS) of the genome. In particular, rs1464791 is an eQTL that regulates expression of
340 *GBE1* in multiple tissues including adrenal gland and muscle²⁶. It is too early to tell
341 whether the overlap between rhythm-associated loci and those two HARS support
342 evolutionary theories about music (e.g., moving to a beat in synchrony during joint
343 music-making and temporally coordination movements has been posited to have a
344 selection effect in modern humans by enhancing group social cohesion and mother-
345 infant bonding^{1,55}.

346 The genetic architecture of rhythm remained virtually unchanged after
347 conditioning the analyses on known GWAS markers of intelligence, in line with twin
348 studies showing specific genetic effects of rhythmic aptitude, over and above common
349 genetic influences on rhythm and intelligence^{17,56}. Furthermore, 30 loci do not appear to
350 have existing genome-wide significant associations with other traits in the current
351 literature, and thus may represent genomic regions newly associated to some aspect of
352 musicality. At the same time, the other 36 loci coincided with robust associations in the
353 GWAS catalogue for a variety of cognitive, neuropsychological, and health traits (Table
354 2, Supplementary Table 12), indicating that rhythm shares genetic architecture with
355 many other traits. We replicated previous findings implicating location 4q22.1 in
356 musicality-related traits^{12,23} (*CCSER1* was the top-associated gene in our MAGMA
357 analysis) but did not find support for previous gene associations from prior candidate-

358 gene, linkage, and GWAS studies with relatively small samples ²¹, potentially due to
359 well-known methodological problems with these methods particularly when applied to
360 small samples ⁵⁷.

361 Positive genetic correlations between rhythm and faster processing speed
362 aligned with prior phenotypic and behavioural genetic studies of cognition, sensory
363 processing, and musicality ^{17,56,58,59}. The correlation between rhythm and chronotype
364 opened up the possibility of a relationship between musical traits and evening
365 chronotype, complimenting evidence of insomnia in musicians ⁶⁰.

366 We found positive genetic correlations with tinnitus, which could be driven by
367 exposure to loud music (this latter correlation with rhythm was just above the
368 significance threshold after multiple-test corrections); both commonly occur among
369 musicians and may lead to hearing loss ⁶¹ highlighting the importance of estimating the
370 prevalence of professional musicians within the study sample in future GWAS of rhythm
371 (this information was not available in the current sample). Unexpected genetic
372 correlations included associations of rhythm with better lung capacity, previous smoker
373 phenotypes, and greater handgrip strength. In light of recent evidence that lung function
374 is genetically related to motor function, processing speed, and cognition in older adults
375 ⁶², it is possible that rhythm shares common biology with a constellation of traits. These
376 lines of research may have clinical-translational implications: for example, a recent
377 intervention study found that music listening improved handgrip strength in older adults
378 ⁶³. We also uncovered shared genetic effects between musical rhythm and biological
379 rhythms including circadian chronotypes and breathing-related phenotypes.

380 More broadly, the genetic correlations between rhythm and other complex traits
381 were relatively modest, suggesting that the present phenotype is not primarily
382 confounded/co-occurring with any particular trait we examined. There are no large-
383 sample GWAS data for major processes fundamental to beat synchronization ^{7,11}:
384 auditory processing, sensori-motor synchronization, locomotion, or temporal processing
385 as a component of general timing abilities ⁶⁴, for which we may expect greater genetic
386 correlations with rhythm in future studies.

387 The primary limitation of our study is the self-reported assessment of rhythm.
388 Although our independent phenotypic validation study indicated that an individual's self-
389 assessment of beat synchronization is related to their objectively-measured rhythm
390 perception abilities, the self-report itself is not an objective assessment of rhythm.
391 Nevertheless, previous studies of other health traits based on self-report have
392 effectively replicated associations from studies using validated assessments, indicating
393 that a powerful sample size can overcome limitations arising from phenotyping error ⁶⁵.
394 The selection of the self-report beat synchronization phenotype was made because it
395 theoretically relates to fundamental components of rhythm including motor periodicity,
396 beat extraction, meter perception, and auditory-motor entrainment (see ⁷ and Glossary
397 in Supplementary Note). Nevertheless, the phenotype available in our GWAS dataset
398 did not allow us to separate the rhythm phenotype into those component factors, and
399 the prevalence of individuals with musical training in the sample was not established.
400 However, given the result that polygenic score for rhythm predicted the presence of
401 musician keywords in an electronic health record-linked biobank, it is likely that we have
402 indeed captured a robust aspect of musicality. These results are promising for future

403 large-scale genomic interrogations using comprehensive music phenotyping yielding
404 continuous musicality variables (whether questionnaire-based ^{10,66} or objective aptitude-
405 based ¹³). Even without continuous measures of rhythm, here we have identified biology
406 potentially differentiating rhythm deficits ⁶⁷ from typical rhythm development. Once
407 GWAS results are available from other heritable musicality traits such as pitch
408 discrimination and music training ¹⁴, the field will be able to test for moderate genetic
409 correlations between rhythm and other musical traits as predicted by family-based
410 studies ^{12,13,68}. Another important area of inquiry will be to investigate musicality and
411 cross-trait correlations in populations of non-European ancestry, hence capturing the
412 spectra of musicality, a human universal, in a wider range of ethnic, cultural and socio-
413 economic contexts.

414 **Online Methods**

415 **Study sample**

416 We obtained genome-wide association study summary statistics from the personal
417 genetics company 23andMe, Inc. Phenotypic status was based on responses to online
418 surveys in which individuals self-reported “Yes” (cases) or “No” (controls) to the
419 question ‘Can you clap in time with a musical beat?’. Individuals who responded “I’m
420 not sure” were excluded from our genomic study. The GWAS included a total of
421 555,660 cases and 51,165 controls (total N=606,825, mean age(SD)=52.09(18.5),
422 prevalence=92%). Specifically, 10.4% of the individuals were 30 years old or younger,
423 24.4% were between 30 and 45 years old, 27.1% were between 45 and 60 years old
424 and 38.1% were older than 60 years old (Table 1). All individuals provided informed
425 consent according to 23andMe’s human subject protocol, which is reviewed and
426 approved by Ethical & Independent Review Services, a private institutional review board
427 (<http://www.eandireview.com>).

428 **Phenotype validation study**

429 *Overview.* To validate the rhythm phenotype used in the genetic study, we conducted a
430 separate internet-based study in N=724 participants from Amazon’s Mechanical Turk.
431 The experiment was designed to determine if self-reported rhythm abilities measured
432 with the question used in the GWAS (i.e., ‘Can you clap in time with a musical beat?’)
433 would be associated with objective performance on a task of rhythm abilities. The Beat-
434 based advantage paradigm was selected as a rhythm discrimination test due to its
435 design of stimuli with simple and complex meter⁶⁹ and prior history investigating

436 individual differences in rhythm perception in a variety of brain and behavioural studies
437 in adults and children with typical and atypical development ^{19,43,70,71}, as well as
438 feasibility for internet-based adaptation. The questionnaire (self-report questions) was
439 administered prior to the perception task, to avoid biasing participant self-report
440 responses by how they perceived they performed on the objective test.

441

442 *Participants*

443 We recruited 724 participants anonymously from Amazon Mechanical Turk. The study
444 received ethical approval from the Columbia University Institutional Review Board.
445 Participants (333 females) were 18-73 years old (mean = 36.1 years, SD=10.9) with 0-
446 45 years of self-reported musical experience (mean 3.7 years, SD=5.8).

447

448 *Stimuli*

449 Stimuli consisted of 32 rhythms drawn from prior work ^{19,69}; half were “simple” rhythms
450 (strong beat-based metrical structure and generally easier to discriminate) and half were
451 “complex” rhythms (weaker metrical structure due to syncopation and generally more
452 challenging to discriminate). Each rhythm was presented as a pure tone in one of 6
453 frequencies (294, 353, 411, 470, 528, and 587 Hz, selected at random), and one of 4
454 durations (ISI of 220, 230, 240, and 250 ms). Each trial consisted of 3 rhythms
455 separated by 1500 ms of silence. As in prior work, the two first presentations were
456 always identical, and in half of the trials (counterbalanced) the third rhythm was also
457 identical (standard condition); in the other trials the rhythm was slightly different (deviant
458 condition).

459

460 *Procedure*

461 Amazon Mechanical Turk (M-Turk) participants were invited to participate in an
462 experiment where they would “listen to sounds and answer questions”. To simulate the
463 user environment within 23andMe where research participants answer a series of
464 unrelated questions about health and other traits, we asked participants to provide
465 answers for a series of randomly presented questions on a variety of other topics
466 (presented at random order; see methods), such as “Do you have wisdom teeth?”.
467 Among these questions we embedded two rhythm-related questions: the target
468 question: “Can you clap in time with a musical beat?” and an additional question, “Do
469 you have a good sense of rhythm?”. After answering these questions, participants
470 passed a test for usage of headphones⁷². This test checks whether participants can
471 hear sounds that are presented through headphones, and guarantees good listening
472 conditions as well as the ability to follow instructions. Participants that passed the
473 headphone test were invited to perform the rhythm perception task (Supplementary
474 Figure 3).

475

476 Participants received 8 training trials that were selected from rhythms that were not part
477 of the test set, and then performed 32 rhythm perception task trials. In all trials (practice
478 and task) participants received feedback regarding their performance (“correct” and
479 “incorrect”), and each correct trial resulted in adding a small monetary bonus.
480 Participants were paid for their performance about \$1.60-\$2.00 depending on their
481 performance, and the duration of the test was about 16-18 minutes. Participants who

482 did not pass the headphone test received \$0.20 for about one minute of answering the
483 initial questions and performing the headphone test. Participant demographic data was
484 collected after the rhythm test.

485

486 **Phenotype Validation Results**

487 654 (90.3%), 25 (3.5%) and 45 (6.2%) participants answered “yes”, “no,” and “I’m not
488 sure” to the target question, “Can you clap in time to a musical beat”. Regarding the
489 self-report question ‘Do you have a good sense of rhythm?’, 503(67%) answered ‘Yes’,
490 102(14%) answered ‘No’ and 117(16%) answered ‘I don’t know’. N=488 answered Yes
491 to both questions, while 166 answered Yes to the Clap to Beat question and 15
492 answered Yes to the sense of rhythm question, resulting in a total tetrachoric correlation
493 between these two self-report questions of $r=0.73$.

494

495 Responses to the rhythm discrimination perception test were analysed using signal
496 detection theory⁷³, as in¹⁹; this method is appropriate for discrimination tasks where
497 the participant has to categorize stimuli along some dimension; the resulting d' values
498 the strength of detection of the signal relative to noise. d' values were calculated on the
499 32 test trials (16 simple rhythm trials and 16 complex rhythm trials) and are reported in
500 Supplementary Table 18. As expected from prior work^{19,70}, individuals scored better in
501 the simple rhythms than the complex rhythms ($t(724)=11.11$, $p<2.2 \times 10^{-16}$, Cohen’s
502 $d=0.58$ (Supplementary Figure 4).

503 To examine whether the self-report of rhythm ability was related to the objective
504 performance on the rhythm discrimination/perception test (see task performance in

505 relation to responses to self-report, shown in Supplementary Figure 5a), we performed
506 a logistic regression analysis in which the self-report rhythm question (Yes vs. No) was
507 the outcome and the rhythm discrimination test performance (standardized d' scores
508 mean = 0, SD = 1) was the predictor. Covariates included age at time of assessment,
509 education, and sex. Individuals with higher performance in the rhythm discrimination
510 test (total d') were more likely to answer that they can clap to the beat
511 (OR(95%CI)=1.94(1.28 to 3.01), $p=0.002$, McFadden's $R^2=0.39$), indicating there is
512 approximately a 94% increase in the odds of answering 'Yes', per standard deviation
513 increase in the rhythm discrimination test. We did not include 'I'm not sure' in the
514 regression, because this answer is not included in the phenotype assessment of the
515 genetic study. Because the simple rhythms have a strong metrical structure and are
516 known to facilitate detection and synchronization of the beat ¹⁹, we also tested whether
517 performance on the simple rhythm trials predicted self-reported beat synchronization
518 (i.e., those who responded Yes to the clap-to-beat question). As above, we found that
519 individuals with higher scores on the simple rhythm trials were more likely to answer
520 that they can clap to the beat (OR(95%CI)=1.99(1.36-2.90), $p<0.001$, McFadden's
521 $R^2=0.40$ (Supplementary Figure 5b). Taken together, these results suggest that the
522 "clap to the beat" self-report phenotype is a broad representation of musical rhythm
523 ability, potentially capturing aspects both of rhythm perception ability and of self-
524 perceived beat synchronization ability.

525

526 **Genotypes and QC**

527 **23andMe dataset**

528 The National Genetics Institute (NGI) performed the DNA extraction and genotyping on
529 saliva samples. Overall, there were five genotyping platforms and subjects were
530 genotyped on only one of them. The v1 and v2 platforms had variants of the Illumina
531 HumanHap550+ BeadChip, including approximately 25,000 custom SNPs selected by
532 23andMe, with a total of about 560,000 SNPs. The v3 platform had variants of the
533 Illumina OmniExpress+ BeadChip, with custom content to improve the overlap with the
534 v2 array, with a total of about 950,000 SNPs. The v4 platform covered about 570,000
535 SNPs, providing extra coverage of lower-frequency coding variation. The v5 platform, in
536 current use, is based on an Illumina Infinium Global Screening Array (~640,000 SNPs)
537 supplemented with ~50,000 SNPs of custom content. In cases where samples did not
538 reach the 98.5% call rate, the sample was re-genotyped. When analyses failed
539 repeatedly, then customers were re-contacted by 23andMe customer service to provide
540 additional samples.

541
542 23andMe restricted participants to a set of unrelated individuals of European ancestry
543 as determined through an analysis of local ancestry ⁷⁴. Relatedness was defined using
544 a segmental identity-by-descent (IBD) estimation algorithm ⁷⁵. Imputation was
545 conducted by combining the May 2015 release of 1000 Genomes Phase 3 haplotypes ⁷⁶
546 with the UK10K imputation reference panel ⁷⁷ to create a single unified imputation
547 reference panel. Phasing was conducted using an internally-developed tool, Finch,
548 which uses the Beagle graph-based haplotype phasing algorithm ⁷⁸ for platforms V1 to
549 V4 while for the V5 platform a similar approach was used with a new phasing algorithm,
550 Eagle2 ⁷⁹. SNPs with a Hardy-Weinberg $p < 10^{-20}$, or a call rate of <90% were flagged.

551 SNPs were also flagged if they were only genotyped on their 'V1' and/or 'V2' platforms
552 due to small sample size and also if SNPs had genotype date effects. Finally, SNPs
553 were also flagged if they had probes matching multiple genomic positions in the
554 reference genome ⁷⁵⁻⁷⁹.

555

556 **GWAS**

557 GWAS was conducted using logistic regression under an additive genetic model, while
558 adjusting for age, sex, the top five principal components of ancestry in order to control
559 for population stratification, and indicators for genotype platforms to account for batch
560 effects. We excluded SNPs with Minor Allele Frequency (MAF) <0.01, low imputation
561 quality ($R^2 < 0.3$) and indels, resulting in a final set of 8,288,851 SNPs for all subsequent
562 analyses.

563

564 **Statistical analyses**

565 *FUMA-based analyses.* The FUMA ²⁵ web application was used on the Genome-Wide
566 Association summary statistics to identify the SNPs that were independent in our
567 analysis with a genome-wide significant P-value ($< 5 \times 10^{-8}$) that are in approximate
568 linkage disequilibrium (LD) with each other at $r^2 < 0.1$ and to generate Manhattan and
569 Quantile-Quantile plots and the SNP functional annotations.

570 Gene analysis and gene-set analysis was performed with MAGMA (v1.07) using FUMA
571 (v1.3.4) and the association analysis summary statistics. Gene expression analysis was
572 obtained from GTEx v7 (<https://www.gtexportal.org/home/>) integrated by FUMA ⁸⁰. More
573 specifically, the gene expression values were log2 transformed average RPKM per

574 tissue type after winsorization at 50 based on GTEx RNA-seq data. Tissue expression
575 analysis was performed for 53 tissue types where the result of gene analysis was tested
576 for one side while conditioning on average expression across all tissue types.

577

578 *LD score regression and genetic correlations.* SNP-heritability was computed with LD
579 Score regression software ²⁴, and heritability estimates were adjusted to the liability
580 scale based on population prevalence of rhythm deficits of 3.5%-6.5% (Supplementary
581 Table 2, Supplementary Note). We then partitioned heritability of rhythm by functional
582 category and investigated cell-type-specific enrichments using stratified LD score
583 regression as per ²⁴. The Bonferroni-corrected p-value was $0.05/1015=4.9 \times 10^{-5}$.

584 The set of human accelerated regions (HARs) was taken from ³⁰. All variants in
585 perfect LD ($r^2 = 1.0$ in 1000 Genomes European individuals) with variants in HARs were
586 considered in the analysis. Similarly, variants tagging Neanderthal introgressed
587 haplotypes were defined as in ⁸¹. All variants in perfect LD with a Neanderthal tag SNP
588 were considered Neanderthal variants. For each set, we performed stratified LDSC
589 (v1.0.0) with European LD scores and the baseline LD-score annotations v2.1. The
590 heritability enrichment is defined as the proportion of heritability explained by SNPs in
591 the annotation divided by the proportion of SNPs in the annotation. Standard effect size
592 (τ_c^*), which quantifies the effects unique to the annotation, is the proportionate change in
593 per-SNP heritability associated with a one standard deviation increase in the value of
594 the annotation, conditional on other annotations in the baseline v2.1 model ⁸².

595 Genetic correlations between rhythm and other complex traits were estimated
596 using LDSC through LD Hub v1.9.0 (<http://ldsc.broadinstitute.org/ldhub/>) ¹⁶ and publicly

597 available GWAS summary statistics. 764 traits were examined and the Bonferroni
598 corrected p-value threshold for significance was $0.05/764=6.5 \times 10^{-5}$. To examine
599 whether the genetic correlations are influenced by residual population stratification, we
600 adjusted the rhythm GWAS summary statistics for the SNP PC-loadings of all top 10
601 PCs. PC loadings were generated from the 1000 Genomes Project because individual-
602 level genotype data was unavailable on the analysed sample⁸³, following⁴¹.

603

604 We used the gsmr R-package (gcta version:v1.92.1beta6) to implement Generalised
605 Summary-data-based Mendelian Randomization to test for causal genetic associations
606³⁸; see Supplementary Note.

607

608 *Conditional analyses.*

609 To control for pleiotropy between cognition and rhythm abilities(23) and identify genetic
610 effects of rhythm traits above and beyond those shared with IQ, we ran a multi-trait
611 conditional and joint analysis (mtCOJO)³⁸, conditioning on intelligence using GWAS
612 summary statistics from³⁹.

613

614

615
616 **Acknowledgments:** We are grateful to 23andMe participants for their contribution to
617 the study and to Nancy Cox for suggestions and insight throughout the process. We
618 would also like to thank Michaela Novakovic, Yune Lee, and Duane Watson for input
619 during previous stages of the project and Vanderbilt Trans-Institutional Programs for
620 sparking the initial collaboration. Members of the 23andMe Research Team are:
621 Michelle Agee, Stella Aslibekyan, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah
622 K. Clark, Sarah L. Elson, Kipper Fletez-Brant, Pierre Fontanillas, Nicholas A. Furlotte,
623 Pooja M. Gandhi, Karl Heilbron, Barry Hicks, Karen E. Huber, Ethan M. Jewett,
624 Yunxuan Jiang, Aaron Kleinman, Keng-Han Lin, Nadia K. Litterman, Jennifer C.
625 McCreight, Matthew H. McIntyre, Kimberly F. McManus, Joanna L. Mountain, Sahar V.
626 Mozaffari, Priyanka Nandakumar, Elizabeth S. Noblin, Carrie A.M. Northover, Jared
627 O'Connell, Steven J. Pitts, G. David Poznik, Anjali J. Shastri, Janie F. Shelton, Suyash
628 Shringarpure, Chao Tian, Joyce Y. Tung, Robert J. Tunney, Vladimir Vacic, and Xin
629 Wang.

630
631 **Funding:**
632 Research reported in this publication was supported by the Eunice Kennedy Shriver
633 National Institute of Child Health & Human Development of the National Institutes of
634 Health under Award Number DP2HD098859 and by the National Institute On Deafness
635 And Other Communication Disorders of the National Institutes of Health under Award
636 Number K18DC017383. JAC was supported by the National Institutes of Health
637 (R35GM127087). The content is solely the responsibility of the authors and does not

638 necessarily represent the official views of the National Institutes of Health. MN was
639 supported by the Wellcome Trust (110222/Z/15/Z).
640 The dataset used for the analyses described were obtained from Vanderbilt University
641 Medical Center's BioVU which is supported by numerous sources: institutional funding,
642 private agencies, and federal grants. These include the NIH funded Shared
643 Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445,
644 and UL1RR024975. Genomic data are also supported by investigator-led projects that
645 include U01HG004798, R01NS032830, RC2GM092618, P50GM115305,
646 U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed
647 at <https://victr.vanderbilt.edu/pub/biovu/>. Also, The Genotype-Tissue Expression (GTEx)
648 Project was supported by the Common Fund of the Office of the Director of the National
649 Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data
650 used for the analyses described in this manuscript were obtained from: the GTEx Portal
651 on 07/26/19 and dbGaP accession number phs000424.vN.pN on 07/26/19 .

652

653

654 **Author contributions**

655 *Conceptualization of study:* Reyna Gordon, Lea Davis

656 *Study design of GWAS and design of other genomic analyses:*

657 Lea Davis, Reyna Gordon, J. Fah Sathirapongsasuti, Maria Niarchou, Tony Capra, David Hinds

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

659 David Hinds

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

661

662 *Post-association QC and generation of figures:* Peter Straub and Maria Niarchou
663 *Post-GWAS analyses (heritability, gene-based analyses, gene set analyses, LD correlations,*
664 *GSMR, mtcojo, PGS in BioVU):*
665 Maria Niarchou, Reyna Gordon, Peter Straub, Lea Davis
666 *HARS and Neanderthal introgression analyses and interpretation:* Evonne McArthur, John A
667 Capra
668 *Sensitivity Analyses of Chromosome 17 inversion and Parkinson's Disease:* J. Fah
669 Sathirapongsasuti
670 *Estimation of population prevalence of rhythm deficits:* Miriam A Mosing and Reyna Gordon
671 *Phenotype validation study design and materials:* Reyna Gordon, J. Devin McAuley, Nori
672 Jacoby
673 *Phenotype validation data collection:* Nori Jacoby, Eamonn Bell
674 *Phenotype validation study data analysis:* Eamonn Bell, Nori Jacoby, Peter Straub, Maria
675 Niarchou, Reyna Gordon
676 *Interpretation of data, writing, editing, and reviewing drafts:* All authors
677 *Project Supervision:* Reyna Gordon, Lea Davis, David Hinds
678
679
680 **Competing interests:** JFS, DH, and members of the 23andMe Research Team are
681 employees of 23andMe, Inc., and hold stock or stock options in 23andMe. All other
682 authors declare no competing interests.
683
684 **Data and materials availability:** The full GWAS summary statistics for the 23andMe
685 dataset will be made available through 23andMe to qualified researchers under an
686 agreement that protects the privacy of the 23andMe participants. Please visit

687 research.23andme.com/collaborate/#publication for more information and to apply to
688 access the data. The data and code from the phenotype validation study, and all of the
689 code from the post-GWAS analyses, are also available upon reasonable request.
690

691 **References**

- 692
- 693 1 Savage, P. E., Brown, S., Sakai, E. & Currie, T. E. Statistical universals reveal
694 the structures and functions of human music. *Proc Natl Acad Sci U S A* **112**,
695 8987-8992, doi:10.1073/pnas.1414495112 (2015).
- 696 2 Trehub, S. E. Cross-cultural convergence of musical features. *Proceedings of the*
697 *National Academy of Sciences* **112**, 8809-8810, doi:10.1073/pnas.1510724112
698 (2015).
- 699 3 Nazzi, T., Bertoncini, J. & Mehler, J. Language discrimination by newborns:
700 toward an understanding of the role of rhythm. *J Exp Psychol Hum Percept*
701 *Perform* **24**, 756-766 (1998).
- 702 4 Cirelli, L. K., Trehub, S. E. & Trainor, L. J. Rhythm and melody as social signals
703 for infants. *Annals of the New York Academy of Sciences* **1423**, 66-72 (2018).
- 704 5 McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M. & Miller, N. S. The time
705 of our lives: life span development of timing and event tracking. *J Exp Psychol*
706 *Gen* **135**, 348-367, doi:10.1037/0096-3445.135.3.348 (2006).
- 707 6 Patel, A. D. & Iversen, J. R. The evolutionary neuroscience of musical beat
708 perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis.
709 *Frontiers in systems neuroscience* **8**, 57 (2014).
- 710 7 Kotz, S. A., Ravignani, A. & Fitch, W. T. The Evolution of Rhythm Processing.
711 *Trends in Cognitive Sciences* **22**, 896-910,
712 doi:<https://doi.org/10.1016/j.tics.2018.08.002> (2018).
- 713 8 Grahn, J. A. & McAuley, J. D. Neural bases of individual differences in beat
714 perception. *NeuroImage* **47**, 1894-1903 (2009).
- 715 9 Law, L. N. & Zentner, M. Assessing musical abilities objectively: Construction
716 and validation of the Profile of Music Perception Skills. *PloS one* **7**, e52508
717 (2012).
- 718 10 Müllensiefen, D., Gingras, B., Musil, J. & Stewart, L. The musicality of non-
719 musicians: an index for assessing musical sophistication in the general
720 population. *PloS one* **9**, e89642 (2014).
- 721 11 Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M. & Fitch, W. T. Finding the
722 beat: a neural perspective across humans and non-human primates. *Philos*
723 *Trans R Soc Lond B Biol Sci* **370**, 20140093, doi:10.1098/rstb.2014.0093 (2015).
- 724 12 Pulli, K. *et al.* Genome-wide linkage scan for loci of musical aptitude in Finnish
725 families: evidence for a major locus at 4q22. *J Med Genet* **45**, 451-456,
726 doi:10.1136/jmg.2007.056366 (2008).
- 727 13 Ullén, F., Mosing, M. A., Holm, L., Eriksson, H. & Madison, G. Psychometric
728 properties and heritability of a new online test for musicality, the Swedish Musical
729 Discrimination Test. *Personality and Individual Differences* **63**, 87-93 (2014).
- 730 14 Gingras, B., Honing, H., Peretz, I., Trainor, L. J. & Fisher, S. E. Defining the
731 biological bases of individual differences in musicality. *Philosophical*
732 *Transactions of the Royal Society B: Biological Sciences* **370**, 20140092,
733 doi:doi:10.1098/rstb.2014.0092 (2015).
- 734 15 Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic
735 risk of complex disease. *Current opinion in genetics & development* **18**, 257-263
736 (2008).

- 737 16 Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD
738 score regression that maximizes the potential of summary level GWAS data for
739 SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272-279,
740 doi:10.1093/bioinformatics/btw613 (2017).
- 741 17 Mosing, M. A., Verweij, K. J., Madison, G. & Ullén, F. The genetic architecture of
742 correlations between perceptual timing, motor timing, and intelligence.
743 *Intelligence* **57**, 33-40 (2016).
- 744 18 Peretz, I. & Vuvan, D. T. Prevalence of congenital amusia. *European Journal of*
745 *Human Genetics* **25**, 625 (2017).
- 746 19 Grahn, J. A. & Brett, M. Rhythm and beat perception in motor areas of the brain.
747 *Journal of cognitive neuroscience* **19**, 893-906 (2007).
- 748 20 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from
749 polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-295,
750 doi:10.1038/ng.3211 (2015).
- 751 21 Oikkonen, J., Onkamo, P., Järvelä, I. & Kanduri, C. Convergent evidence for the
752 molecular basis of musical traits. *Scientific Reports* **6**, 39707,
753 doi:10.1038/srep39707
754 <https://www.nature.com/articles/srep39707 - supplementary-information> (2016).
- 755 22 Oikkonen, J. *et al.* A genome-wide linkage and association study of musical
756 aptitude identifies loci containing genes related to inner ear development and
757 neurocognitive functions. *Molecular psychiatry* **20**, 275 (2015).
- 758 23 Park, H. *et al.* Comprehensive genomic analyses associate UGT8 variants with
759 musical ability in a Mongolian population. *J Med Genet* **49**, 747-752,
760 doi:10.1136/jmedgenet-2012-101209 (2012).
- 761 24 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using
762 genome-wide association summary statistics. *Nature genetics* **47**, 1228 (2015).
- 763 25 Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional
764 mapping and annotation of genetic associations with FUMA. *Nature*
765 *Communications* **8**, 1826, doi:10.1038/s41467-017-01261-5 (2017).
- 766 26 Consortium, G. The Genotype-Tissue Expression (GTEx) pilot analysis:
767 multitissue gene regulation in humans. *Science* **348**, 648-660 (2015).
- 768 27 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature*
769 *genetics* **25**, 25 (2000).
- 770 28 The Gene Ontology, C. The Gene Ontology Resource: 20 years and still GOing
771 strong. *Nucleic Acids Res* **47**, D330-D338, doi:10.1093/nar/gky1055 (2019).
- 772 29 Mithen, S. J. *The Singing Neanderthals: The origins of music, language, mind,*
773 *and body.*, (Harvard University Press, 2005).
- 774 30 Capra, J. A., Erwin, G. D., McKinsey, G., Rubenstein, J. L. & Pollard, K. S. Many
775 human accelerated regions are developmental enhancers. *Philos Trans R Soc*
776 *Lond B Biol Sci* **368**, 20130025, doi:10.1098/rstb.2013.0025 (2013).
- 777 31 Hubisz, M. J. & Pollard, K. S. Exploring the genesis and functions of Human
778 Accelerated Regions sheds light on their role in human evolution. *Curr Opin*
779 *Genet Dev* **29**, 15-21, doi:10.1016/j.gde.2014.07.005 (2014).
- 780 32 Doan, R. N. *et al.* Mutations in Human Accelerated Regions Disrupt Cognition
781 and Social Behavior. *Cell* **167**, 341-354 e312, doi:10.1016/j.cell.2016.08.071
782 (2016).

- 783 33 Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve
784 GWAS Power. *Am J Hum Genet* **104**, 65-75, doi:10.1016/j.ajhg.2018.11.008
785 (2019).
- 786 34 Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic
787 loci influencing general cognitive function. *Nature communications* **9**, 2098
788 (2018).
- 789 35 Sankararaman, S. *et al.* The genomic landscape of Neanderthal ancestry in
790 present-day humans. *Nature* **507**, 354-357, doi:10.1038/nature12961 (2014).
- 791 36 Couth, S., Mazlan, N., Moore, D., Munro, K. & Dawes, P. Hearing difficulties and
792 tinnitus in construction, agricultural, music and finance industries: contributions of
793 demographic, health and lifestyle factors. *Trends in Hearing* (in press).
- 794 37 Mosing, M. A., Madison, G., Pedersen, N. L. & Ullén, F. Investigating cognitive
795 transfer within the framework of music practice: Genetic pleiotropy rather than
796 causality. *Developmental Science* **19**, 504-512 (2016).
- 797 38 Zhu, Z. *et al.* Causal associations between risk factors and common diseases
798 inferred from GWAS summary data. *Nature Communications* **9**, 224,
799 doi:10.1038/s41467-017-02317-2 (2018).
- 800 39 Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867
801 individuals identifies new genetic and functional links to intelligence. *Nature*
802 *genetics* **50**, 912 (2018).
- 803 40 Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in
804 complex traits. *Nature Genetics* **51**, 1339-1348, doi:10.1038/s41588-019-0481-0
805 (2019).
- 806 41 Bhatia, G. *et al.* Correcting subtle stratification in summary association statistics.
807 *bioRxiv*, 076133 (2016).
- 808 42 Steinberg, K. M. *et al.* Structural diversity and African origin of the 17q21.31
809 inversion polymorphism. *Nat Genet* **44**, 872-880, doi:10.1038/ng.2335 (2012).
- 810 43 Grahn, J. A. & Brett, M. Impairment of beat-based rhythm discrimination in
811 Parkinson's disease. *Cortex* **45**, 54-61 (2009).
- 812 44 Luciano, M. *et al.* Association analysis in over 329,000 individuals identifies 116
813 independent variants influencing neuroticism. *Nat Genet* **50**, 6-11,
814 doi:10.1038/s41588-017-0013-8 (2018).
- 815 45 Nagel, M. *et al.* Meta-analysis of genome-wide association studies for
816 neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat*
817 *Genet* **50**, 920-927, doi:10.1038/s41588-018-0151-7 (2018).
- 818 46 Pardini, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-
819 intolerant genes and in regions under strong background selection. *Nat Genet*
820 **50**, 381-389, doi:10.1038/s41588-018-0059-2 (2018).
- 821 47 Lévy, J. *et al.* Molecular and clinical delineation of 2p15p16. 1 microdeletion
822 syndrome. *American Journal of Medical Genetics Part A* **173**, 2081-2087 (2017).
- 823 48 Vicente, C. T., Revez, J. A. & Ferreira, M. A. R. Lessons from ten years of
824 genome-wide association studies of asthma. *Clinical & Translational Immunology*
825 **6**, e165, doi:10.1038/cti.2017.54 (2017).
- 826 49 Grahn, J. A. & Rowe, J. B. Feeling the Beat: Premotor and Striatal Interactions in
827 Musicians and Nonmusicians during Beat Perception. *The Journal of*
828 *Neuroscience* **29**, 7540-7548, doi:10.1523/jneurosci.2018-08.2009 (2009).

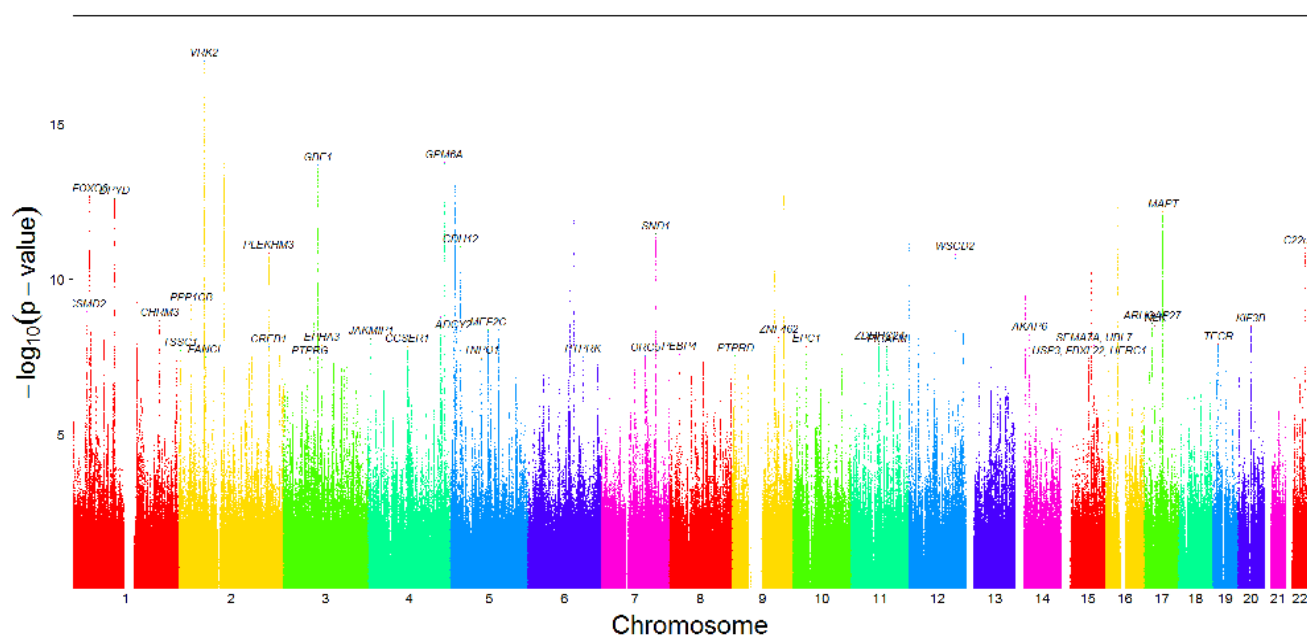
- 829 50 Grahn, J. A. & Rowe, J. B. Finding and feeling the musical beat: striatal
830 dissociations between detection and prediction of regularity. *Cerebral cortex* **23**,
831 913-921 (2012).
- 832 51 Kung, S.-J., Chen, J. L., Zatorre, R. J. & Penhune, V. B. Interacting cortical and
833 basal ganglia networks underlying finding and tapping to the musical beat.
834 *Journal of Cognitive Neuroscience* **25**, 401-420 (2013).
- 835 52 Bengtsson, S. L. *et al.* Listening to rhythms activates motor and premotor
836 cortices. *Cortex* **45**, 62-71, doi:<https://doi.org/10.1016/j.cortex.2008.07.002>
837 (2009).
- 838 53 Teki, S., Grube, M., Kumar, S. & Griffiths, T. D. Distinct Neural Substrates of
839 Duration-Based and Beat-Based Auditory Timing. *The Journal of Neuroscience*
840 **31**, 3805-3812, doi:10.1523/jneurosci.5561-10.2011 (2011).
- 841 54 McAuley, J. D., Henry, M. J. & Tkach, J. Tempo mediates the involvement of
842 motor areas in beat perception. *Annals of the New York Academy of Sciences*
843 **1252**, 77-84 (2012).
- 844 55 Dissanayake, E. If music is the food of love, what about survival and reproductive
845 success? *Musicae scientiae* **12**, 169-195 (2008).
- 846 56 Mosing, M. A., Pedersen, N. L., Madison, G. & Ullén, F. Genetic pleiotropy
847 explains associations between musical auditory discrimination and intelligence.
848 *PLoS One* **9**, e113874 (2014).
- 849 57 Border, R. *et al.* No support for historical candidate gene or candidate gene-by-
850 interaction hypotheses for major depression across multiple large samples.
851 *American Journal of Psychiatry* **176**, 376-387 (2019).
- 852 58 Zuk, J., Benjamin, C., Kenyon, A. & Gaab, N. Behavioral and neural correlates of
853 executive functioning in musicians and non-musicians. *PloS one* **9**, e99868
854 (2014).
- 855 59 Medina, D. & Barraza, P. Efficiency of attentional networks in musicians and non-
856 musicians. *Heliyon* **5**, e01315 (2019).
- 857 60 Vaag, J., Saksvik-Lehouillier, I., Bjørngaard, J. H. & Bjerkeset, O. Sleep
858 difficulties and insomnia symptoms in norwegian musicians compared to the
859 general population and workforce. *Behavioral sleep medicine* **14**, 325-342
860 (2016).
- 861 61 Jansen, E., Helleman, H., Dreschler, W. & de Laat, J. Noise induced hearing loss
862 and other hearing complaints among musicians of symphony orchestras.
863 *International archives of occupational and environmental health* **82**, 153-164
864 (2009).
- 865 62 Finkel, D. & Bravell, M. Genetic influences on lung function contribute to
866 subsequent age changes in motor and cognitive function. *Behavioral Genetics*
867 (2019).
- 868 63 van den Elzen, N. *et al.* The Power of Music: Enhancing Muscle Strength in
869 Older People. *Healthcare (Basel)* **7**, doi:10.3390/healthcare7030082 (2019).
- 870 64 Haegens, S. & Golumbic, E. Z. Rhythmic facilitation of sensory processing: a
871 critical review. *Neuroscience & Biobehavioral Reviews* **86**, 150-165 (2018).
- 872 65 Tung, J. Y. *et al.* Efficient replication of over 180 genetic associations with self-
873 reported medical data. *PloS one* **6**, e23473 (2011).

- 874 66 Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J. &
875 Rodriguez-Fornells, A. Individual differences in music reward experiences. *Music*
876 *Perception: An Interdisciplinary Journal* **31**, 118-138 (2013).
- 877 67 Sowiński, J. & Dalla Bella, S. Poor synchronization to the beat may result from
878 deficient auditory-motor mapping. *Neuropsychologia* **51**, 1952-1963 (2013).
- 879 68 Oikkonen, J. *et al.* A genome-wide linkage and association study of musical
880 aptitude identifies loci containing genes related to inner ear development and
881 neurocognitive functions. *Molecular Psychiatry* **20**, 275, doi:10.1038/mp.2014.8
882 <https://www.nature.com/articles/mp20148-supplementary-information> (2014).
- 883 69 Povel, D.-J. & Essens, P. Perception of temporal patterns. *Music Perception: An*
884 *Interdisciplinary Journal* **2**, 411-440 (1985).
- 885 70 Gordon, R. L. *et al.* Musical rhythm discrimination explains individual differences
886 in grammar skills in children. *Developmental Science* **18**, 635-644 (2015).
- 887 71 Wieland, E. A., McAuley, J. D., Dilley, L. C. & Chang, S.-E. Evidence for a rhythm
888 perception deficit in children who stutter. *Brain Lang* **144**, 26-34 (2015).
- 889 72 Woods, K. J., Siegel, M. H., Traer, J. & McDermott, J. H. Headphone screening
890 to facilitate web-based auditory experiments. *Attention, Perception, &*
891 *Psychophysics* **79**, 2064-2072 (2017).
- 892 73 Macmillan, N. A. & Creelman, C. D. *Detection theory: A user's guide.*
893 (Psychology press, 2004).
- 894 74 Durand, E. Y., Do, C. B., Mountain, J. L. & Macpherson, J. M. Ancestry
895 Composition: A Novel, Efficient Pipeline for Ancestry Deconvolution. *bioRxiv*,
896 010512, doi:10.1101/010512 (2014).
- 897 75 Henn, B. M. *et al.* Cryptic distant relatives are common in both isolated and
898 cosmopolitan genetic samples. *PLoS One* **7**, e34267 (2012).
- 899 76 Consortium, G. P. A global reference for human genetic variation. *Nature* **526**, 68
900 (2015).
- 901 77 consortium, U. K. The UK10K project identifies rare variants in health and
902 disease. *Nature* **526**, 82 (2015).
- 903 78 Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and
904 missing-data inference for whole-genome association studies by use of localized
905 haplotype clustering. *The American Journal of Human Genetics* **81**, 1084-1097
906 (2007).
- 907 79 Loh, P.-R., Palamara, P. F. & Price, A. L. Fast and accurate long-range phasing
908 in a UK Biobank cohort. *Nature genetics* **48**, 811 (2016).
- 909 80 Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. FUMA:
910 Functional mapping and annotation of genetic associations. *bioRxiv*,
911 doi:10.1101/110023 (2017).
- 912 81 Vernot, B. *et al.* Excavating Neandertal and Denisovan DNA from the genomes
913 of Melanesian individuals. *Science* **352**, 235-239, doi:10.1126/science.aad9416
914 (2016).
- 915 82 Hujoel, M. L. A., Gazal, S., Hormozdiari, F., van de Geijn, B. & Price, A. L.
916 Disease Heritability Enrichment of Regulatory Elements Is Concentrated in
917 Elements with Ancient Sequence Age and Conserved Function across Species.
918 *Am J Hum Genet* **104**, 611-624, doi:10.1016/j.ajhg.2019.02.008 (2019).

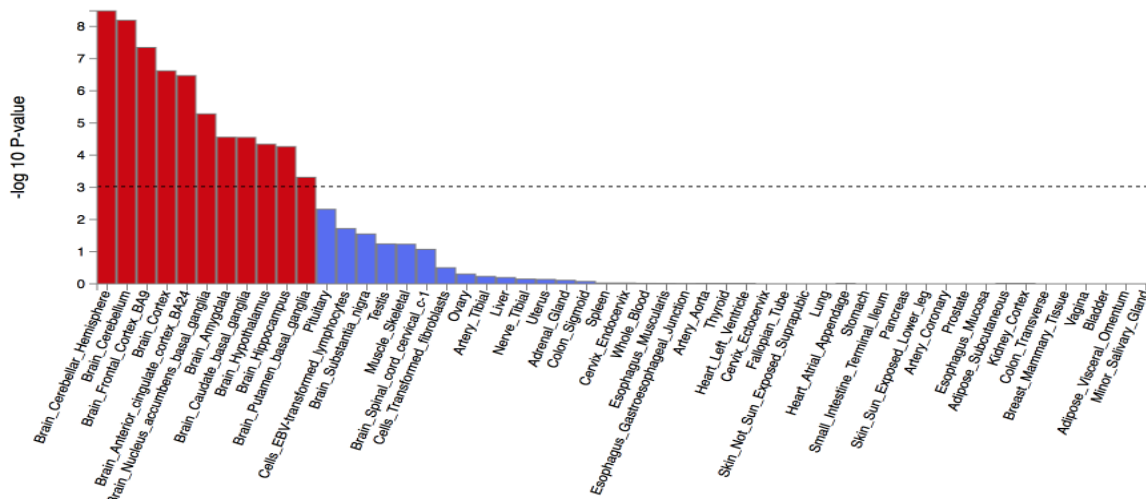
919 83 The Genomes Project, C. *et al.* A global reference for human genetic variation.
920 *Nature* **526**, 68, doi:10.1038/nature15393
921 [https://www.nature.com/articles/nature15393 - supplementary-information](https://www.nature.com/articles/nature15393-supplementary-information) (2015).
922

923
924 **FIGURES**

925
926 **Figure 1. Manhattan plot of GWAS results for rhythm.** Results are shown for
927 606,825 research participants with 23andMe who responded Yes (N=555,660) vs. No
928 (N=51,165) to the question “Can you clap in time to a musical beat?”, controlling for
929 age, sex, top 5 PC’s for ancestry, and genotype platform. The x-axis shows
930 chromosomal position and the y-axis shows $-\log_{10}$ p-values). 68 loci surpassed the
931 threshold for genome-wide significance of $p < 5 \times 10^{-8}$ (dotted horizontal line). For
932 illustration purposes, we only included 500,000 SNPs with $p < 0.1$
933

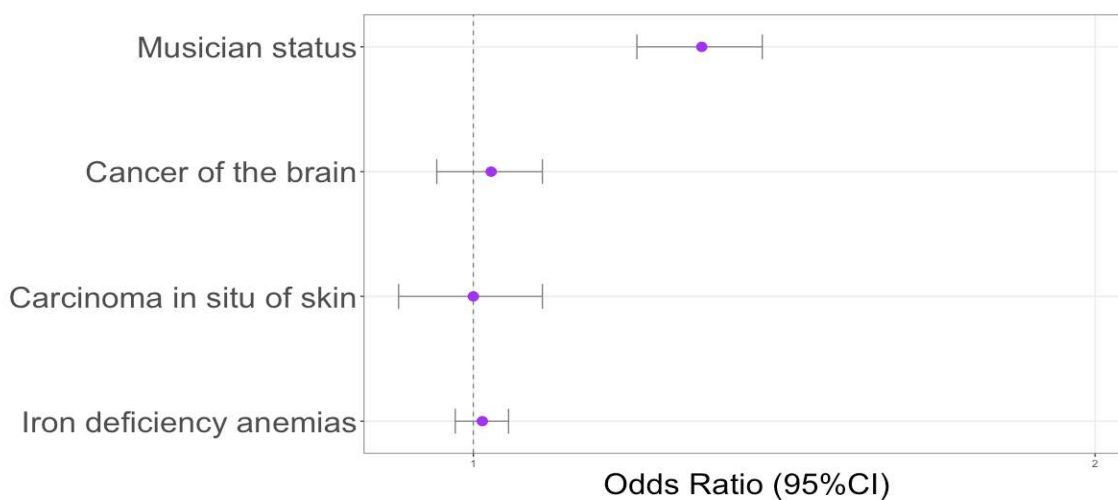


935 **Figure 2. Gene expression profiles for genes associated with rhythm in 53 tissue**
936 **types from the GTEx database.** Gene property analysis was conducted in FUMA on
937 the MAGMA gene-based results, in which the average expression per gene was added
938 as a co-variate to the model. Associations with rhythm were significantly enriched in
939 brain tissue compared to other tissues (-log₁₀ p-values are on y-axis, with type on x-
940 axis).



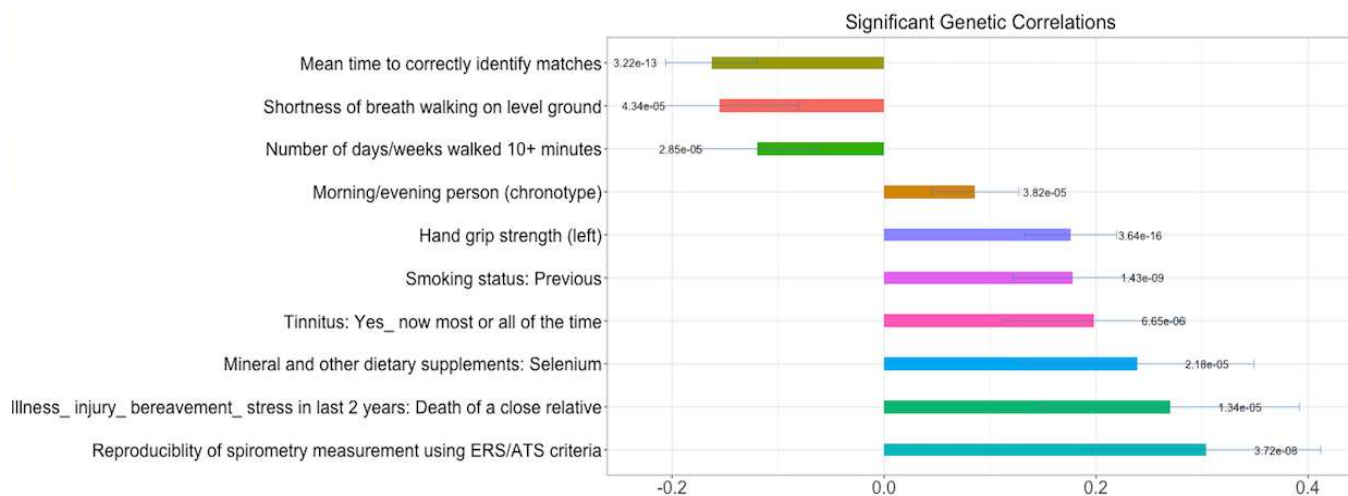
941
942
943
944

945 **Figure 3. Polygenic score of rhythm predicts presence of musician-related**
946 **keywords in electronic health record (EHR) biobank.** With weights derived from the
947 23andMe GWAS, we applied polygenic scores (PGS) for rhythm to genomic data from
948 N=864 individuals in Vanderbilt's BioVU whose EHR contained "musician" and related
949 keywords, and compared to a control sample of N=66,577 (See Supplementary Note).
950 PGS-rhythm were higher for musician vs. controls (OR=1.3) but did not predict 3
951 negative-control traits, shown here.
952



953
954

955 **Figure 4. Selected results from the cross-trait genetic correlations analysis.**
956 Rhythm summary statistics were significantly genetically correlated with several health,
957 cognitive, emotion, and circadian phenotypes in our analysis comparing rhythm to traits
958 available in LDHub. Full results are presented in Supplementary Table 13. The y-axis is
959 the rg correlation, the bars represent standard errors and the p-values are also
960 indicated.
961



962
963

964
965
966
967
968
969
970

Tables

Table 1. Number of participants in each rhythm phenotype group, broken down by sex and age range in the GWAS sample.

Phenotype group (response to Clap- to-beat question)	Total N	Males	Females	0 to 30 years old	30 to 45 years old	45 to 60 years old	60 years old and over
Yes (cases)	555660	226188	329472	57898	135168	150939	211655
No (controls)	51165	23998	27167	5186	12909	13312	19758

971
972
973
974

975 Table 2. Genomic loci associated with rhythm after adjusting for intelligence and traits
 976 with which they have been previously associated (loci in pink do not exist in GWAS
 977 catalog)
 978

SNPs after IQ cojo	chr	p-value	GWAS catalog (yes/no)	GWAS catalog
rs34762587	1	2.21584E-13	no	
rs2635634	5	9.96339E-12	no	
rs12913592	15	6.59873E-11	no	
rs2061843	1	1.08739E-09	no	
rs4443239	4	1.5082E-09	no	
rs1901739	5	2.15026E-09	no	
rs8079923	17	3.11892E-09	no	
rs7501911	17	3.99553E-09	no	
rs2284901	14	7.46306E-09	no	
rs10978661	9	9.75689E-09	no	
rs4263335	4	9.25137E-09	no	
rs7939759	11	1.45292E-08	no	
rs9710427	19	1.26321E-08	no	
rs12638746	3	1.56044E-08	no	
rs2505344	10	1.6844E-08	no	
rs6548147	2	2.0951E-08	no	
rs10885458	12	2.56522E-08	no	
rs10877461	10	2.61599E-08	no	
rs1996148	8	2.90685E-08	no	
rs12056186	7	2.93356E-08	no	
rs7856850	9	3.38293E-08	no	
rs4704043	5	3.49958E-08	no	
rs43182	3	4.50456E-08	no	
rs7875397	9	2.21728E-13	no	
rs2018545	14	3.58572E-10	no	
rs7002174	8	4.75329E-08	no	
rs34863893	5	8.60406E-10	no	
rs7297439	12	5.07484E-09	no	
rs72826882	17	2.58837E-09	no	
rs7625774	3	3.12812E-09	no	
rs848293	2	1.01057E-17	yes	depression, neuroticism, schizophrenia
rs62340585	4	2.2504E-14	yes	neuroticism, schizophrenia, highest math class taken
rs10168817	2	1.83859E-14	yes	reaction time, worry, pneumonia, phonic sneeze reflex, reaction time, HIV-1 viral set point
rs10779987	3	2.04673E-14	yes	hand grip strength
rs10875125	1	3.43636E-13	yes	educational attainment, schizophrenia, ASD, ADHD, MDD
rs4792891	17	7.06239E-13	yes	general cognitive ability, depressed affect, neuroticism
rs1468701	7	2.72542E-12	yes	intelligence
rs9626920	22	9.91369E-12	yes	self-reported risk-taking behaviour, smoking status, dupuytren's disease
rs764299	2	1.47E-11	yes	educational attainment, highest math class taken, cognitive performance
rs1426371	12	1.34602E-11	yes	extraversion, worry
rs7586405	2	7.19E-10	yes	heel bone mineral density
rs55678522	1	2.81431E-09	yes	glioma, non-glioblastoma glioma, highest math class taken, general cognitive ability, educational attainment, intelligence
rs6087848	20	3.75408E-09	yes	inflammatory bowel disease, Crohn's disease, cognitive performance,
rs13163173	5	6.08339E-09	yes	platelet count, intelligence, depressive symptoms, subjective well-being, depressed affect, neuroticism, BMI, conscientiousness, depression

rs67264739	5	5.89821E-09	yes	adolescent idiopathic scoliosis
rs1596431	5	5.77642E-09	yes	intelligence
rs12909047	15	1.46255E-08	yes	caffeine metabolism (plasma 1,3, 7, -trimethylxanthine (caffeine)level)
rs67816799	4	1.64433E-08	yes	educational attainment
rs10932201	2	1.55342E-08	yes	systolic blood pressure
rs526904	11	1.54544E-08	yes	Alzheimer's disease, family history of Alzheimer's disease,
rs13197257	6	2.77659E-08	yes	reaction time, educational attainment, general cognitive ability, highest math class taken
rs11692449	2	3.2277E-08	yes	immature fraction of reticulocytes
rs62014217	15	3.1548E-08	yes	atrial fibrillation, urinary albumin excretion
rs476141	1	4.68634E-08	yes	diabetic retinopathy
rs7715357	5	1.30965E-13	yes	smoking status
rs11865086	16	5.94369E-13	yes	blood protein levels, menarche(age at onset)
rs1536057	6	1.35469E-12	yes	platelet distribution width, smoking status (ever vs. never smokers), smoking initiation (ever regular vs. never regular), anxiety/tension(special factor of neuroticism)
rs1972582	12	7.90101E-12	yes	chronic kidney disease, glomerular filtration rate(creatinine), blood metabolite levels, chronotype
rs3780420	9	6.39377E-11	yes	quantitative traits, platelet distribution width, platelecrit, platelet count
rs9385269	6	1.96812E-10	yes	educational attainment, cognitive performance, intelligence, regular attendance at a pub or social club, Tourette's syndrome, educational attainment, bipolar disorder, highest math class taken, self-reported math ability, cognitive function, alcohol consumption (drinks per week), general risk tolerance, extremely high intelligence, autism spectrum disorder, QT interval, risk taking tendency
rs16837903	1	4.74528E-10	yes	monocyte chemoattractant protein-1 levels, lung function
rs778353	2	3.01299E-08	yes	heel bone mineral density
rs10497357	2	3.42377E-08	yes	longitudinal change in brain amyloid plaque burden
rs2467452	12	9.8837E-09	yes	automobile speeding propensity, chronotype, morning person
rs6684973	1	1.01877E-08	yes	schizophrenia, depression, smoking initiation, smoking status, alcohol consumption self-reported math ability, highest math class taken
rs2819336	1	4.59116E-09	yes	smoking cessation, smoking initiation, cognitive ability, intelligence, general cognitive ability, cognitive performance, age of smoking initiation, hypertension risk in short sleep duration, menarche (age of onset), red blood cell count, educational attainment, self-reported math ability, highest math class taken, attention deficit hyperactivity disorder, male-pattern baldness