Discovery of 36 loci significantly associated with stuttering

Jennifer Below (✉ jennifer.e.below@vanderbilt.edu)
  Vanderbilt University Medical Center   https://orcid.org/0000-0002-1346-1872
Hannah Polikowsky
  Vanderbilt University Medical Center
Alyssa Scartozzi
  Vanderbilt Genetics Institute / Vanderbilt University Medical Center
Douglas Shaw
  Vanderbilt Genetics Institute / Vanderbilt University Medical Center
Dillon Pruett
  Hearing and Speech Sciences / Vanderbilt University   https://orcid.org/0000-0002-9760-2133
Hung-Hsin Chen
  Vanderbilt University Medical Center   https://orcid.org/0000-0002-1921-2797
Lauren Petty
  Vanderbilt University Medical Center   https://orcid.org/0000-0003-1619-6303
Alexander Petty
  Vanderbilt Genetics Institute / Vanderbilt University Medical Center
Emily Lowther
  Curtin School of Allied Health / Curtin University
Yao Yu
  Department of Epidemiology / University of Texas MD Anderson Cancer Center
23andMe Research Team
  23andMe, Inc.
Heather Highland
  University of North Carolina at Chapel Hill   https://orcid.org/0000-0002-3583-8239
Christy Avery
  University of North Carolina at Chapel Hill   https://orcid.org/0000-0002-1044-8162
Kathleen Mullan Harris
  University of North Carolina at Chapel Hill   https://orcid.org/0000-0001-9757-1026
Reyna Gordon
  Vanderbilt University Medical Center   https://orcid.org/0000-0003-1643-6979
Janet Beilby
  Curtin School of Allied Health / Curtin University   https://orcid.org/0000-0002-7439-5514
Kathy Viljoen
  Curtin School of Allied Health / Curtin University
Robin Jones
  Department of Hearing and Speech / Vanderbilt University Medical Center   https://orcid.org/0000-0002-1150-0786
Chad Huff
  Department of Epidemiology, Division of Cancer Prevention and Population Sciences, MD Anderson Cancer Center,
Shelly Jo Kraft
  Communication Sciences and Disorders / Wayne State University

Abstract

Developmental stuttering is a common speech disorder (studies estimate at least a 5% lifetime prevalence) characterized by prolongations, blocks, and repetitions of speech sounds. In approximately 75–80% of cases in early childhood, stuttering will resolve within a few years (referred to as 'recovery'); the remaining cases will often experience stuttering into school-age years and adulthood (referred to as 'persistence'). In adults, the prevalence of stuttering is substantially higher in men compared to women, at a ratio of 4:1 or greater (compared to between 1:1 and 2:1 in young children); this has typically been explained by differences in likelihood of recovery by sex. Heritability studies have established that a genetic component for stuttering exists, with heritability estimates as high as 84%. However, genetic factors impacting stuttering risk remain largely uncharacterized. To date, only two prior genome-wide association studies (GWAS) of developmental stuttering have been published, both of which included less than 10,000 cases. Here, we performed eight self-reported stuttering GWAS that were stratified by sex and ancestries. These analyses included more than 1 million individuals (99,776 cases and 1,023,243 controls) and identified 36 unique genome-wide significant loci. We validated the self-reported stuttering phenotype using polygenic risk scores from two independent stuttering datasets. We examined genetic correlation of our GWAS results with published GWAS for other previously identified comorbid traits and found strong evidence of correlation with hearing loss, daytime sleepiness, depression, and poorer beat synchronization. We also performed Mendelian randomization analyses which revealed distinct causal relationships in males and females for genetically associated traits. These distinct causal relationships motivate continued research into sex-specific phenotypic differences, with emphasis on recovery status. Additionally, a high proportion of genes impacting stuttering risk were found to be associated with neurological traits from the GWAS catalog, supporting a neurological basis for stuttering. Our findings provide the first well-powered insight into genetic factors underlying stuttering, representing a major step forward in our understanding of this condition.

Main

Developmental stuttering is the most common fluency disorder with more than 400 million people affected worldwide and a lifetime prevalence of 5–8% among global populations.[2] Stuttering is characterized by syllable and word repetitions, sound prolongations, and involuntary breaks in words that disrupt the forward movement of speech. The onset of developmental stuttering typically occurs during childhood between ages 2 and 5; however, an estimated 80% of children who stutter will recover, with or without the aid of speech therapy. Developmental stuttering is also sexually dimorphic; at stuttering onset, the male-to-female ratio is approximately even (between 1:1 and 2:1).[4] Notably, stuttering is substantially higher in males compared to females in the adolescent population,[2,3] which has typically been explained by differences in likelihood of recovery by sex.[5,6]

Although stuttering is a relatively common disorder, it is often disguised by those affected via avoidance behaviors such as word substitutions.[7] Despite extensive research into treatment for stuttering, including speech motor interventions, behavior modification, cognitive interventions, and technology-based feedback interventions, many affected individuals receiving therapy experience only a modest reduction in stuttered syllables.[8] Once persistent, stuttering does not have a known cure and often involves a lifetime of therapy to help manage overt stuttering behaviors, covert psychological impact, identity and social impact, and secondary movement behaviors.[9] People who stutter often exhibit negative psychosocial attributes including avoidance of speaking situations, negative perceptions of identity and self-worth, and reduced overall quality of life.[10] For those who experience persistent stuttering, the impact can be profound and life-long. Young people who stutter experience increased bullying, decreased classroom participation, and report a more negative educational experience and stuttering in this population is associated with depression and suicide ideation.[10–13] For adults, stuttering can negatively impact employability, perceived job performance, socioeconomic status, and mental and social well-being.[10,11,14,15]

Studies of stuttering within families, twins, and population isolates provide overwhelming evidence for a strong genetic influence on stuttering risk with heritability estimates ranging from 0.42 to 0.84.[16–27] To date, family studies have identified six candidate causal stuttering genes: *GNPTAB, GNPTG*, and *NAGPA*,[20,28] *DRD2*,[21] *AP4E1*,[25] and *CYP17A1*.[23] However, efforts to replicate these findings in other families or global populations have not yet validated the observed effects.[29,30] To date, two population studies of stuttering have been published in the literature, reporting two genome-wide significant genomic loci that confer stuttering risk at the population-level.[31,32] Together, these prior investigations, leveraging both family data and global outbred populations, demonstrate that stuttering genetic risk factors are complex and involve both familial and population variation. Despite this progress, the biological mechanisms by which these variants impact stuttering are unknown. Larger sample sizes are urgently needed to elucidate genetic risk factors for this common complex trait, especially to examine relevant biological variables, such as sex and genetic ancestry. Furthermore, predictive models that leverage genetic risk markers (e.g. polygenic risk models, genetic correlation analysis, and causal inference models) may illuminate the broader clinical impact of the genetic risk of stuttering.

Here, we report the results of genome-wide association studies (GWAS) of stuttering from more than 1 million individuals (99,776 cases). The analysis is well-powered to detect stuttering risk alleles with modest effect size, and explores effects stratified by eight sex- and genetic ancestry-specific groups of self-reported stuttering among 23andMe, Inc. participants who answered the survey question: "Have you ever had a stammer or stutter?" This study reveals the complex genetic architecture of developmental stuttering, identifying 36 unique signals for self-reported stuttering. We further validated the observed genetic effects in two independent datasets including an international clinically ascertained stuttering cohort

called the International Stuttering Project (ISP)[31] and a cohort of self and parental reported stuttering in The National Longitudinal Study of Adolescent to Adult Health (Add Health).[33] These effects are further leveraged to explore genetic correlations between stuttering and its comorbidities. Together, these remarkable advances inform our understanding of the molecular etiology of stuttering and lay groundwork for the future of precision care in developmental speech disorders.

Results

Study Overview

We performed eight independent GWAS of self-reported stuttering that were stratified by sex and genetic ancestries in samples from 23andMe, Inc. From the resulting summary statistics, we estimated the genetic and partitioned heritability of developmental stuttering, developed polygenic risk models, and validated the predictive value of our derived polygenic risk score within the clinically ascertained ISP cohort and Add Health datasets. To better understand the causal relationships between previously identified comorbidities and stuttering, we performed genetic correlations and causal inference to establish sex-specific effects and the directionality of effect for each stuttering-associated trait. To assess the effects of non-coding variants in regulatory regions and better elucidate the underlying biological processes of stuttering, we computed credible variant sets, and performed partitioned heritability, colocalization analysis (See Supplemental Methods), and tissue-specific gene module enrichment analyses (See Supplemental Methods). Finally, we assessed replication of top signals within previously published family- and population-based genetic analysis, as well as associated GWAS Catalog traits of stuttering risk genes (See Supplemental Methods).

Genome-wide association studies

The full dataset included 99,776 (48,217 males, 51,559 females) participants responding 'yes' to the question 'Have you ever had a stammer or stutter?' (cases) and 1,023,243 (392,414 males, 630,829 females) participants responding 'no' (controls, Table S1). Four genetic ancestries, African Ancestry (AA), East Asian Ancestry (EAA), European Ancestry (EA), and Latino/Admixed American Ancestry (AdA) were defined through an analysis of local ancestry (see Table S1 for sample sizes by ancestry).[34] In the analysis of each sex- and ancestries- specific GWAS (eight total), we considered autosomal variants that were successfully imputed across all platforms and reached our quality control thresholds (see Methods for details).

From the eight genetic ancestries- and sex-specific GWAS, we identified 24 loci (Figs. S1-24) associated with stuttering at the conventional genome-wide significance threshold of p-value < 5 x 10^{-8}: nine loci were from the EA-female study, ten loci in the EA-male study, three loci in the AA-male study, one locus in the AdA-female and AdA-male study (Fig. 1, Tables 1 and 2, Extended Data Figs. 1−7). No loci reached genome-wide significance in the AA-female GWAS, the EAA-female GWAS, nor the EAA-male GWAS. The top two associations in the EA-female GWAS were independently replicated (p-value < 0.05) in another sex and ancestries-specific GWAS in the 23andMe dataset (Table 2, Table S2). In the EA-male GWAS, six out of 10 of the genome-wide significant loci were independently replicated in one or more sex and ancestries-specific GWAS (Table 1, Table S2). We did not observe any replication for the five genome-wide significant signals in our non-European studies (Table 3, Table S2).

Table 1

Sentinel hits from EA-male genome-wide association study. Sentinel variant reported for each locus with a genome-wide significant hit, p-value < 5 x 10^{-8}. The functional gene(s) represents the variant-to-gene predicted by Open Targets V2G pipeline, which integrates evidence from molecular quantitative trait loci, chromatin interactions, in silico functional predictions from Ensembl, and distance between the variant and gene canonical transcription start site. NA (not available) reported for variants where Open Targets did not identify a gene. Base-pair positions listed according to human genome reference build 19. Dashes in location column indicate distance, where [] is contained within the transcripts of the specified gene; " = <1kb; '‑' = < 10kb; '‑‑' = < 100kb; '‑‑‑' = < 1000kb to either upstream or downstream of the gene. Replicating Study indicates the study where replication was observed or 'NA' if replication was not observed in any of the seven tested independent studies.

rsid	chr	pos_b37	EA	NEA	EAF	OR [95% CI]	P-value	Functional Gene(s)	Location	Replicating Study
rs35609938	2	58756729	T	C	0.501	0.95 [0.93–0.96]	5.84E-12	VRK2	FANCL‑‑[]	EA-F
rs1040225	2	58139593	G	A	0.598	1.05 [1.04–1.07]	1.82E-11	VRK2	[VRK2]	EA-F
rs34394051	1	6853091	G	A	0.157	1.07 [1.04–1.09]	1.51E-09	CAMTA1	[CAMTA1]	EA-F
rs545889942	2	104116510	I	D	0.445	1.05 [1.03–1.06]	5.07E-09	NA	TMEM182‑‑[]	NA
rs72664949	13	109280508	G	A	0.245	1.06 [1.04–1.07]	7.42E-09	MYO16	[MYO16]	NA
rs10850379	12	110002777	T	C	0.445	1.04 [1.03–1.06]	1.77E-08	MMAB	[MMAB]	EA-F, AdA-M
rs62337988	5	12031700	T	A	0.317	1.05 [1.03–1.07]	2.04E-08	CTNND2	CTNND2‑‑[]	EA-F
rs11353659	15	48059138	I	D	0.637	0.96 [0.94–0.97]	2.56E-08	SEMA6D	[SEMA6D]	EA-F
rs58120907	13	110413514	G	A	0.666	0.96 [0.94–0.97]	4.81E-08	IRS2	[IRS2]	NA
rs558002155	8	121159409	G	A	0.999	3.51 [2.09–5.89]	4.99E-08	COL14A1	[COL14A1]	NA

Table 2

Sentinel hits from EA-female genome-wide association study. Sentinel variant reported for each locus with a genome-wide significant hit, *p*-value < 5 x 10^{-8}. The functional gene represents the variant-to-gene predicted by Open Targets V2G pipeline, which integrates evidence from molecular quantitative trait loci, chromatin interactions, in silico functional predictions from Ensembl, and distance between the variant and gene canonical transcription start site. NA (not available) reported for variants where Open Targets did not identify a gene. Base-pair positions listed according to human genome reference build 19. Dashes in location column indicate distance, where [] is contained within the transcripts of the specified gene; ″ = <1kb; '-' = < 10kb; '–' = < 100kb; '—' = < 1000kb to either upstream or downstream of the gene. Replicating Study indicates the study where replication was observed or 'NA' if replication was not observed in any of the seven tested independent studies.

rsid	chr	pos_b37	EA	NEA	EAF	OR [95% CI]	*P*-value	Functional Gene(s)	Location	Replicating Study
rs13107325	4	103188709	T	C	0.0813	1.11 [1.09–1.15]	3.81E-16	*SLC39A8*	[*SLC39A8*]	AdA-F
rs572319557	18	50846441	I	D	0.551	1.04 [1.03–1.06]	2.95E-10	*DCC*	[*DCC*]	AA-M
rs3801279	7	104904868	T	C	0.52	0.96 [0.94–0.97]	3.03E-09	*SRPK2*	[*SRPK2*]	NA
15:29934686	15	29934686	T	C	7.37E-04	0.23 [0.13–0.42]	7.53E-09	NA	*FAM189A1*–[]–*TJP1*	NA
rs535503154	5	151965756	I	D	0.271	0.95 [0.94–0.97]	2.78E-08	*NMUR2*	*NMUR2*–[]—*GRIA1*	NA
rs968163	20	51037935	G	A	0.258	0.95 [0.94–0.97]	3.81E-08	*TSHZ2*	*ZFP64*–[]—*TSHZ2*	NA
rs529593131	17	68255397	T	C	2.81E-04	0.043 [0.008–0.225]	3.84E-08	NA	*KCNJ2*–[]	NA
rs779897701	4	12449797	G	C	3.63E-04	0.096 [0.026–0.35]	4.05E-08	NA	[]—*RAB28*	NA
rs62252182	3	69881433	G	A	0.206	1.05 [1.03–1.07]	4.51E-08	*MITF*	[*MITF*]	NA

Table 3

Sentinel hits from non-EA genome-wide association studies. Sentinel variant reported for each locus with a genome-wide significant hit, *p*-value < 5 x 10^{-8}. The functional gene(s) represents the variant-to-gene predicted by Open Targets V2G pipeline, which integrates evidence from molecular quantitative trait loci, chromatin interactions, in silico functional predictions from Ensembl, and distance between the variant and gene canonical transcription start site. NA (not available) reported for variants where Open Targets did not identify a gene. Base-pair positions listed according to human genome reference build 19. Dashes in location column indicate distance, where [] is contained within the transcripts of the specified gene; ″ = <1kb; '-' = < 10kb; '–' = < 100kb; '—' = < 1000kb to either upstream or downstream of the gene. Replication was not observed in any of the seven tested independent genome-wide association studies.

Study	rsid	chr	pos_b37	EA	NEA	EAF	OR [95% CI]	*P*-value	Functional Gene(s)	Location	Avg RSQ	Min RSQ	Neff
AA-M	rs541395135	12	80825417	I	D	0.945	0.654 [0.564–0.758]	4.36E-08	*PTPRQ*	[*PTPRQ*]	0.686	0.596	179
AA-M	rs7333000	13	26535079	G	A	0.915	1.40 [1.23–1.59]	4.29E-08	*SHISA2*	[*ATP8A2*]	0.899	0.856	347
AA-M	rs192857772	22	37824152	G	A	0.998	5.83 [3.18–10.71]	2.24E-08	*CYTH4*	*ELFN2*[]–*MFNG*	0.872	0.868	8
AdA-F	rs35713684	10	109112494	G	A	0.993	2.22 [1.62–3.06]	4.58E-08	*SORCS1*	*SORCS1*—[]	0.725	0.708	77
AdA-M	rs556601931	13	41980338	T	C	6.76E-04	6.84 [3.65–12.84]	2.24E-08	*RGCC*	*NAA16*–[]-*RGCC*	0.682	0.401	6

Sentinel genome-wide significant hits from the EA-male GWAS identified signals implicating *VRK2, CAMTA1, MYO16, MMAB, CTNND2, SEMA6D, IRS2,* and *COL14A1* as the most likely impacted functional genes (Open Targets V2G pipeline;[41] Table 1). Sentinel genome-wide significant hits from the EA-female GWAS identified signals implicating *SLC39A8, DCC, SRPK2, NMUR2, TSHZ2,* and *MITF* as the most likely impacted functional

genes (Table 2). Sentinel genome-wide significant hits from the AA-male GWAS implicated *PTPRQ*, *SHISA2*, and *CYTH4* as likely functional genes; AdA-female GWAS implicated *SORCS1*; and AdA-male GWAS implicated *RGCC* (Table 3).

We aggregated association summary statistics across ancestry via multi-ancestry meta-regression, implemented in MR-MEGA,[35] identifying 15 loci associated with stuttering in our sex-combined meta-analysis (Extended Data Fig. 8, Table S3 and Figs. S25-39), five loci (all loci *p*-value < 5.0 x 10^{-8} in the sex-combined meta-analysis) in our female-specific meta-analysis (Extended Data Fig. 9, Table S4 and Figs. S40-44), and three loci (all loci *p*-value < 1.5 x 10^{-8} in the sex-combined meta-analysis) in our male-specific meta-analysis (Extended Data Fig. 9, Table S5 and Figs. S45-47). However, our concordance analysis, which compared summary statistics from each genetic ancestries- and sex-specific GWAS revealed genetic dissimilarity across the studies (Table S6). Only the EA-male and EA-female association studies had strong concordance (0.953, with other cross ancestries concordance rates ranging from 0.339 and 0.800) suggesting genetic heterogeneity by sex and ancestries. Therefore, we focused subsequent analyses of stuttering, which is known to be sexually dimorphic, on the eight genetic ancestries- and sex-specific GWAS.

36 unique signals for self-reported stuttering, across ancestries- and sex-specific, and trans-ancestries meta- and mega GWASs, were identified after establishing credible sets.

Genetic heritability

We calculated SNP-based liability scaled heritability for our male and female EA, AA, AdA, and EAA studies using LD Score regression.[36,37] The explained variance estimates were transformed from the observed scale to the underlying liability scale, accounting for an expected case prevalence of 0.1 on the basis of the observed frequency of stuttering cases (Table S1). Liability scaled heritability was 0.0906 (SE = 0.0051) for EA-females and 0.0947 (SE = 0.0054) for EA-males. For other sex- and genetic ancestries, the liability scaled heritability ranged from 0.0161 to 0.1531 (Table S7).

Partitioned SNP-based heritability of stuttering by broad functional annotation showed significant enrichments of conserved regions in EA-male, EA-female, and EA-sex-combined stuttering (Extended Fig. 10, Tables S8-10). EA-male and EA-sex-combined stuttering was enriched for weak enhancers, repressed marks, and chromatin marks of H3K4me1, a marker for enhancers, and H3K9ac, a marker for activate chromatin (Extended Fig. 10, Tables S8 and S10, *p*-value < 9.6 x 10^{-4}). EA-sex-combined stuttering was enriched for fetal and adult hypersensitive sites and chromatin mark H3K27ac, a marker of activate chromatin sites (Extended Fig. 10, Table S8, *p*-value < 9.6 x 10^{-4}). Furthermore, we used LDSC specifically enriched genes[38] to determine whether genes expressed in specific cell or tissue types are enriched for stuttering associated variants. For brain cell types, EA-female and EA-sex-combined stuttering was enriched for neurons (Extended Fig. 11, Tables S11-13, *p*-value < 0.017). After, we identified enrichment of brain tissues previously associated with stuttering in imaging studies.[38–45] For genes expressed within specific brain tissues, EA-female stuttering was enriched for genes expressed in one brain tissue, EA-male stuttering was enriched for genes expressed in four brain tissues, and EA-sex-combined stuttering was enriched for genes expressed in eight brain tissues (Extended Fig. 12, Tables S11-13, *p*-value < 6.25 x 10^{-3}). Enrichment was further investigated by examining tissue-specific annotations for activate chromatin and enhancers. EA-female stuttering was enriched for one brain tissue, identified by the chromatin mark of H3K27ac; EA-male stuttering was enriched for two brain tissues, identified by the presence of H3K9ac and H3K27ac chromatin marks; EA-sex-combined stuttering was enriched for four brain tissues, identified by the presence of H3K27ac, H3Kme1, H3K9ac, H3K27ac chromatin marks (Extended Fig. 13, Tables S11-13, *p*-value < 2.5 x 10^{-3})

Genetic Correlation

We performed ancestries and sex-specific genetic correlation analysis comparing our EA-male stuttering GWAS results to summary statistics from independent GWAS of EA-males and our EA-female stuttering GWAS results to summary statistics from independent GWAS of EA-females for 17 traits previously reported in studies of stuttering[13, 46–50] [see Table S14 for trait details]. The 17 selected traits encompassed the following categories: behavioral, circadian rhythm, immune, metabolic, motor, neurological, and hearing traits (Fig. 2a). In addition, we explored the genetic correlation of stuttering with one trait where sex stratified summary statistics were not available: beat synchronization (Fig. 2b).[51] We observed a nominally significant positive genetic correlation with both our EA-male and EA-female studies for hearing loss (EA-male 95% confidence intervals (CI): 0.051−0.23, EA-male *P*: 1.90 x 10^{-3}, EA-female CI: 0.12−0.30, EA-female *P*: 6.50 x 10^{6}), daytime sleepiness (EA-male CI: 0.011−0.18, EA-male *P*: 2.75 x 10^{-2}, EA-female CI: 0.085−0.25, EA-female *P*: 6.96 x 10^{-5}), and depression (EA-male CI: 0.16−0.47, EA-male *P*: 6.82 x 10^{-5}, EA-female CI: 0.23−0.48, EA-female *P*: 4.53 x 10^{-5}), and a negative correlation with beat synchronization (EA-male CI: -0.17 - -0.048, EA-male *P*: 4.0 x 10^{-4}, EA-female CI: -0.20 - -0.089, EA-female *P* = 3.33 x 10^{-7}). We observed a significant positive genetic correlation in EA-females only for asthma (CI: 0.25−0.52, *P* = 2.91 x 10^{-8}), allergic rhinitis (CI: 0.047−0.31, *P* = 7.90 x 10^{-3}), suicide ideation (CI: 0.11−0.47, *P*: 1.7 x 10^{-3}), anxiety (CI: 0.046−0.31, *P* = 8.0 x 10^{-3}), ADHD (CI: 0.21−0.53, *P* = 4.66 x 10^{-6}), and BMI (CI: 0.17−0.28, *P* = 4.08 x 10^{-17}), as well as a significant negative genetic correlation for sleep duration (CI: -0.19 - -0.046, *P* = 1.4 x 10^{-3}), alcohol consumption frequency (CI: -0.25 - -0.071, *P* = 5.0 x 10^{-4}), and walking pace (CI: -0.30 - -0.16, *P* = 8.81 x 10^{-11}; Fig. 3a and 3b). EA-sex-combined was negatively correlated with beat synchronization (CI: -0.17 - -0.080, *P*: 1.13 x 10^{-7}). No traits were significantly associated with EA-males exclusively.

Mendelian Randomization

We performed a sex-specific Mendelian randomization analysis for each of the 17 traits, and sex-specific and sex-combined for beat synchronization, included in our genetic correlation analysis to determine if the genetic risk for any of these traits might harbor a causal (vertical) or horizontal relationship with the self-reported stuttering phenotype captured in our EA-male, EA-female, or EA-sex-combined GWAS. In particular, in EA-females, increased genetic risk of slower walking pace shows a significant causal effect on self-reported stuttering (Fig. 3a). We observed a causal relationship and evidence of pleiotropy between an increased genetic risk for higher body mass index, and evening predisposed chronotype on EA-female self-reported stuttering (Fig. 3a, Table S15). Additionally, in EA-males, we observed a causal relationship and evidence of pleiotropy between an increased genetic risk for higher testosterone levels on self-reported stuttering (Fig. 3a, Table S15). In EA-sex-combined, we found a causal relationship between increased genetic risk for poorer beat synchronization and self-reported stuttering (Fig. 3a). In EA-females, we observed a significant causal relationship of female stuttering on depression risk (Fig. 3b, Table S15). We observed significant causal effects across EA-males and EA-females stuttering on poorer beat synchronization (Fig. 3b, Table S15). In EA-males, we observed a pleiotropic relationship between self-reported stuttering on increased depression risk, and increased hearing loss (Table S15). Lastly, we observed a pleiotropic relationship between EA-female self-reported stuttering on increased risk of anxiety (Table S15). We did not observe any shared overlap in pleiotropy between EA-female and EA-male self-reported stuttering.

The genetic architecture of self-reported stuttering significantly predicts clinically validated stuttering

Due to power, stuttering polygenic scores (PGS) were derived from 23andMe association statistics in the EA-female and EA-male GWAS, and applied to EA participants within two independent cohorts of developmental stuttering, ISP (893 EA cases and 6,052 EA controls) and Add Health (588 EA cases and 6,621 EA controls). The final male derived PRS model included 1,024,432 variant predictors and the female PRS model included 1,024,431 variant predictors. Overall, male-specific PRS models out-performed female-specific PRS models (Fig. 4). The male-specific PRS model derived from the EA-male GWAS demonstrated good performance for both male and female EA in the ISP (AUC = 0.612 for male, and AUC = 0.607 for female; Fig. 4a, Extended Fig. 14), and Add Health (AUC = 0.537 for male, AUC = 0.553 for female; Fig. 4b, Extended Fig. 14). PRS scores of cases and controls within the ISP cohort and Add Health subjects can be found in Table S16. Cross-ancestries testing of the PRS models are presented in Extended Fig. 15.

Replication from prior studies

We sought replication for six genes previously implicated as causal genes in family studies by evaluating all variants that passed our QC metrics within each gene across all eight independent GWAS. Since previously described variants from family-based studies[20,21,23,25] were not directly genotyped or were too rare to impute, we sought replication of effects from variants in and around these genes. We uncovered variant signals reaching statistical significance after adjusting for multiple testing (see Methods) for the following genes: *GNPTAB, GNPTG, AP4E1*, and *CYP17A1* (Table S17). The variant in *GNPTAB*, rs76300806, represents a common indel (EAF = 0.484 in EA males) found in the 5' UTR region. The variant in *GNPTG*, rs111790048, represents a rare (EAF = 4.22×10^{-5} in EA males) intronic variant that is also in proximity to *TSR3* (~ 2Kb upstream). The variant in *AP4E1*, rs565776226, represents a rare (MAF = 0.001 in AdA-males) intronic variant. The variant in *CYP17A1*, rs777625933, represents a rare (EAF = 1.58×10^{-4} in EA-females) intronic variant.

We also sought SNP-based replication for stuttering associated variants reported in Shaw *et al.*[25] and Polikowsky *et. al.*[24] We did not replicate any signals in either study after applying a Bonferroni correction (8 tests, *p*-value > 6.25 x 10^{-3}); however, one variant, rs34919320, reported by Polikowsky *et al.* neared significance (Table S18).

Discussion

Our ancestries-specific genome-wide association studies of self-reported stuttering in men and women are the largest to date comprising ~ 100,000 cases and 1 million controls stratified by AA, AdA, EAA, and EA. Eight total loci from the trans-ancestries sex-stratified meta- and 15 loci from the trans-ancestries sex-combined mega-analyses and 24 total loci from sex- and ancestries-specific analyses, resulted in 36 unique signals from credible sets reaching genome wide significance in our GWAS of self-reported stuttering (Tables 1 − 3, Tables S3-5), none of which have been previously reported in stuttering literature. Eight variants replicated across our independent GWAS. Most prior studies of genetic risk factors for stuttering have explored rare variant effects in pedigrees, however we find effects that are consistent with high polygenicity, suggesting a genetic architecture similar to other common complex disease traits. Effects at significant loci ranged from β = 0.04 − 3.51. LD-score based estimates of heritability are often lower than those estimated from twin and family studies,[52,53] and our estimates from LD-score regression of the heritability of developmental stuttering ranged from 1.61% in AdA-female to 15.31% in AA-male[37]. These estimates are in line with other common, complex disease traits such as insomnia,[54] type 2 diabetes,[55] and beat synchronization.[56]

We developed PRS models from the sex-specific GWAS results from the AA and EA groups and applied them to the ISP and Add Health stuttering cohorts (which comprise AA and EA samples) for validation. The EA-male derived models showed significant differences in liability scores between stuttering cases and controls in both sexes in the ISP, a clinically ascertained cohort enriched with males and persistent cases of stuttering, while the EA-female model had significant predictive performance only in EA-females. In Add Health, both the male and female PRS models significantly

predict case/control status (here, stuttering cases are based on self-report like the 23andMe, Inc. analyses). The difference in the predictive performance in these two validation cohorts is notable with several possible explanations, including: the trait captured by 23andMe EA-females contains more false positives than the EA-male study (due to sex-differential participation bias[44] or poorer recall rate in females who more often recover as children), the genetic liability for developmental stuttering varies between males and females, and is perhaps confounded by differences in genetic susceptibility to persistent versus recovered stuttering, or genetic variation contributing to developmental stuttering risk may be confounded by horizontal pleiotropy modulated by sex. The first possibility, sex-differential participation bias, represents a documented phenomenon reported within 23andMe genetic data.[57] Sex-differential participation bias could also be confounded by adult recall, because adults who recovered from stuttering during childhood might not recall their childhood stuttering status. Since females more often recover from stuttering in very early childhood, accurate recall of stuttering in childhood may disproportionately impact female self-report of stuttering.[2] Future research is needed to deconvolute genetic risk factors that are specific to sex or persistence.

We observed replication-specific significance at four genes that have been previously reported in prior family-based studies of stuttering.[20,23,25] Although the previously reported variants were not directly genotyped and were too rare for accurate imputation in the 23andMe data, our replication analysis identified two extremely rare variants in *GNPTG* and *CYP17A1*, and one variant in *GNPTAB* and *AP4E1* that passed multiple test correction (see Methods). *AP4E1* interacts with previously reported gene *NAGPA*, and together these results provide modest additional support for the role of rare variants in genes that control intracellular trafficking in stuttering.

All variants reaching genome-wide significance in our study represent novel findings for developmental stuttering. We found one locus, which spans our two top hits in EA-male, rs35609938 and rs1040225, that the Open Targets variant-to-gene (V2G) prediction algorithm assigned *VRK2* as the most likely functional gene (R^2 between these variants is 0.31 in CEU).[58] Specifically, rs35609938 occurs downstream of *VRK2* and upstream of *FANCL*, and rs1040225 occurs within either an intronic or genic upstream region of *VRK2* (Table 1, Figs. S1-2). Interestingly, *FANCL* and *VRK2* were recently implicated in musical beat synchronization.[56] Rhythm perception impairments have been linked to a number of speech and language conditions, including stuttering.[59,60] Complex rhythm discrimination is below average in adults[59] and children[51] who stutter, consistent with the Atypical Rhythm Risk Hypothesis,[60] which posits that those with atypical rhythm may be at risk for developmental/speech language disorders. Clinically, synchronizing speech with external pacing cues, such as a metronome, can temporarily decrease stuttering disfluencies.[51, 61–63] Therefore, impairments in rhythm processing may be causal for stuttering, and our GWAS findings offer further support for this hypothesis.

Of the 30 unique significant genes identified across sex-specific and meta- and mega- analyses, 20 have been previously associated with traits (Mapped Genes) in the GWAS Catalog. Fifteen genes were previously associated with traits grouped into Obesity/Endocrine/Metabolic and Lifestyle/Behaviors categories, 13 of which overlapped the two categories. Additionally, 12 genes have been previously associated with traits grouped into Mental Disorders and Neurological categories, with nine genes overlapping between the two categories. While the etiology of stuttering remains largely obscure, the proportion of genes impacting stuttering risk that are also associated with neurological traits provides additional evidence for a neurological basis of stuttering.[39,44,64] Furthermore, 11 genes have been previously associated with Educational Attainment traits. The overlap of stuttering associated genes with those identified in educational attainment may be the result of social factors, including bias and stigma, that potentially hinder classroom performance of those who stutter.[11,65,66] For unique genes found within other categorized GWAS traits, please see Table S19 and Fig. 5.

Imaging studies have demonstrated that people who stutter exhibit differences in a variety of brain areas,[39,40] including the frontal cortex,[41] cingulate cortex,[38,41] basal ganglia [caudate, substantia nigra],[42–45] inferior temporal lobe, and cerebellum.[38] The enrichment of genes expressed and tissue-specific regulatory annotations, as well as our gene module enrichment analyses provide additional evidence for the association of these brain areas with stuttering risk. Findings from our gene module enrichment analysis also revealed enrichments in the frontal cortex, cerebellum, cortex, nucleus accumbens of the basal ganglia, and anterior cingulate cortex. Together, these findings provide new genetic evidence for previously described relationships between brain areas and stuttering.[38–41, 43–45,67]

Genetic correlation analysis showed significant correlations of increased stuttering risk in both EA-male and female with increased hearing loss, increased daytime sleepiness, decreased beat synchronization, and increased risk of depression. In addition, genetic correlation analysis showed significant correlations of increased stuttering risk in EA-female with increased alcohol consumption, increased BMI, decreased walking pace, decreased sleep duration, and increased risk of asthma, suicide ideation, anxiety, ADHD, and allergic rhinitis. Increased risk of EA-sex-combined stuttering showed significant correlations with poorer beat synchronization. These genetic correlations, and their respective directions, are largely consistent with previous literature identifying traits comorbid with stuttering.[13, 46–51]

We also performed Mendelian randomization analyses to explore causal relationships between stuttering and traits that have been previously reported as co-occuring with stuttering in the literature. The genetic liability of BMI, chronotype, walking pace, suicide ideation, testosterone showed significant causal effects on stuttering risk. We also observed significant causal effects of the genetic liability of stuttering on depression. Our results are consistent with several studies suggesting that males and females who stutter report elevated symptoms of depression compared to their fluent counterparts.[13,68,69] Communication difficulties due to stuttering can result in feelings of frustration and hopelessness and, paired with

broader societal stigma toward stuttering, can negatively impact psychological health.[70-72] We also observed significant bi-directional effects of stuttering on the ability to clap to a beat.

We did not observe any significant causal or pleiotropic relationships of self-reported stuttering across sexes, highlighting differences in genetic risk between sexes. The distinct causal pathways in males and females relating stuttering to genetically correlated traits are notable; however, since females are more likely to recover from stuttering than males, one limitation of this study is an inability to fully differentiate the factors of sex and stuttering persistence. Improved granularity of self-report with information regarding stuttering persistence will be necessary to resolve the confound of sex and persistence.

Overall, this study represents the largest GWAS of stuttering to date. We leveraged 99,776 cases and over 1 million controls to identify 36 unique genome-wide significant loci associated with sex- and ancestries-specific self-reported stuttering. This genetic architecture was validated in two independent stuttering datasets. These data provide insight into the genetic contributions to stuttering at the population level, demonstrating that genetic risk is complex and polygenic and dominated by modest to low genetic effects. After decades of progress examining the behavioral, neural, and physiological contributions of language, articulation, speech-motor coordination, and temperament and emotion to stuttering, the addition of genetics may provide a mechanistic framework for integrating findings across these domains. For the first time, we demonstrate shared molecular underpinnings between stuttering and other associated clinical features including depression and beat synchronization. An unresolved question in the field of stuttering, with lengthy historical speculation, is whether persistent stuttering and recovery from stuttering represent distinct subtypes.[73,74] Thus far, studies have yielded conflicting results with no clear relationship between pattern of recovery and genetic model.[75-77] These analyses motivate continued research into causal differences between females and males as well as between persistent and recovered stuttering. These findings represent a critical step toward the next era of research for this common, complex, costly, and heritable condition.

Methods

Studies

23andMe: Genome-wide association studies included participants from 23andMe, Inc. who self-reported stuttering status through a questionnaire. Cases included participants who answered "yes" (99,776 individuals) to the question: "Have you ever had a stammer or stutter?" Controls (1,023,243 individuals) included participants who answered "no" to this same question (see Demographic Table 1). As is a common standard in population-based studies investigating stuttering, our study relies on self-report (see Bloodstein and Ratner (2008) in which all but two of the reviewed papers were based on retrospective questionnaire or interview-style surveys).[2,78] All individuals included in the analyses provided informed consent and answered surveys online according to 23andMe human subject protocol, which was reviewed and approved by Ethical & Independent Review Services, a private institutional review board (http://www.eandireview.com).

ISP: Polygenic model testing was performed using individuals with developmental stuttering acquired through the International Stuttering Project. Stuttering status in the ISP cohort was confirmed by speech-language pathologists with expertise in stuttering and fluency disorders. See Polikowsky et al. 2022 *HGG Advances*[31] for a detailed description of this cohort, and genotyping information.

The National Longitudinal Study of Adolescent to Adult Health (Add Health): Polygenic model testing was performed using individuals who self-reported stuttering via an Add Health questionnaire. Add Health represents an ongoing, nationally representative, longitudinal study of the social, behavioral, and biological factors influencing health and developmental trajectories from early adolescence into adulthood.[33] Add Health collected demographic and health survey data as well as in-home physical and biological data from participants. See Harris et al. 2019 *Int J Epidemiol* for genotyping information. For our study, self-reported stuttering cases were defined as participants who at one point answered "yes" to the following survey question: "Do you have a problem with stuttering or stammering?" All control individuals answered "no" to the above question. Self-reported race/ethnicity was used to group participants.

Statistical Analysis

Eight ancestries- and sex-specific genome-wide association analyses were performed to determine variant association with stuttering risk (Table S1). Each performed GWAS used a logistic regression that assumed an additive model for allelic effects:

$$Stuttering\ status \sim age + pc.0 + pc.1 + pc.2 + pc.3 + pc.4 + v2\_platform + v3\_0\_platform + v3\_1\_platform + v4\_platform + genotype$$

Reported p-values were calculated using a computed likelihood ratio test. Principal components for each logistic regression model were derived independently for each ancestries, using ~ 65,000 high quality genotyped variants present across all five genotyping platforms. Principal components were computed on a subset of participants randomly sampled across all the genotyping platforms (137K, 102K, 1000K, and 360K participants were used for AA, EAA, EA and AdA, respectively). PC scores for participants not included in the analysis were obtained by projection, combining the eigenvectors of the analysis and the SNP weights. Summary statistics were reported for imputed autosomal variants that were successfully imputed across all platforms (v2v3v4v5) and reached the following quality control thresholds: average rsq > 0.5, minimum rsq > 0.3, and batch check p-value > 1x10^{-50}.

We aggregated association summary statistics across ancestries-specific association studies using multi-ancestries meta-regression, as implemented in MR-MEGA.[35] Analyses were performed for female-only GWAS, male-only GWAS, and a sex-combined meta-analysis. We included three axes of genetic variation as covariates in the sex-combined meta-analysis, and, due to the lower number of contributing analyses in the female- and male-only meta-analyses limiting the number of possible axes of genetic variation, included one axis as a covariate in the sex-specific analyses.

Annotation

The sentinel variant for each genome-wide significant locus was reported for each ancestries and sex-specific study. The genome-wide significance was defined as $P < 5 \times 10^{-8}$ [79] this threshold applies a Bonferroni correction where $\alpha = 0.05$ and assumes there are approximately 1 million independent (i.e. not in linkage disequilibrium) common signals across the human genome. Annotated gene(s) for each locus included the predicted functional gene(s) for each loci (when available) according to Open Targets "Variant-to-gene (V2G) pipeline", which integrates evidence from molecular quantitative trait loci, chromatin interactions, in silico functional predictions from Ensembl, and distance between the variant and gene canonical transcription start site.[80,81] Loci were defined according to independent linkage disequilibrium blocks identified in 1000 Genomes reference using the matched ancestries reference. Reported sentinel variants represent the variant with the smallest p-value within each associated region. All reported positional coordinates (chromosome and base-pair locations) refer to human genome reference build 37. We also looked for replication of any genome-wide significant signal among the other independent ancestries- and sex-specific GWAS.

Credible Sets

95% credible sets were established to determine if genome-wide significant hits were unique across all sex- and ancestries- specific GWAS and trans-ancestries meta- and mega- GWAS. The potential causal variants for SNPs within significant regions was based on approximate Bayes factor[82] assuming a prior variance of .1, and using the method from Maller et al.[83] to define these sets. A hit was determined to be unique if there were no overlaps in SNPs with other credible sets from the same chromosome.

Variant-effect size concordance analysis

We compared summary statistics from each ancestries- and sex-specific to one another to determine whether the concordance rate between the two summary statistics was high (Table S6). The concordance rate was calculated by the proportion of overlapping LD blocks that had the same direction of effect over the total variants present in both GWAS analyses with p-values below 0.005 threshold. See Table S6 for details regarding the number of variants used in each concordance combination.

SNP Heritability and Partitioned Heritability

Genome-wide SNP based heritability (h^2) was calculated using summary statistics resulting from the EA, AA, and AdA GWAS results using the LD Score regression software.[37] We used LDSC to estimate liability scaled h^2 assuming a 10% population prevalence. LD maps were estimated from the 1000 Genomes phase 3 European, Admixed American, and African populations for the respective ancestries. Sample size prevented h^2 calculations in the East Asian cohort since estimates are likely to be unreliable (sample size below range of $5,000-10,000$).[84]

To better understand the types of variation that contribute most to stuttering, we partitioned SNP heritability of EA-male, EA-female and EA-sex-combined (See Supplemental Methods for sex-combined meta-analysis details) stuttering using stratified LDSC.[85] LD scores, regression weights, and allele frequencies from European ancestries were obtained from: https://alkesgroup.broadinstitute.org/LDSCORE. We performed 80 different tests, resulting in a p-value = 6.25×10^{-4} Bonferroni-corrected significance level globally. Partitioning was performed for 52 baseline annotations as described by Finucane et al.[85]. Enrichment was considered significant if p-value < 9.6×10^{-4}, derived by Bonferroni correction (52 gene-sets).

Next, we estimated enrichments for cell-type-specific and tissue-specific heritability[86] on EA-male, EA-female and EA-sex-combined stuttering, while controlling for the baseline models. Brain cell types used to estimate enrichment of heritability consisted of neurons, astrocytes, and oligodendrocytes using data from Cahoy et al.[87]. Enrichments were considered significant if p-value < 0.017, derived by Bonferroni correction (3 gene-sets). Gene expression data (computed from GTEx[88] database) used to estimate enrichment of heritability consisted of eight brain regions with empirical evidence relating to stuttering.[38–41, 43–45] Enrichments were considered significant if p-value < 6.25×10^3, derived by Bonferroni correction (8 gene-sets). Lastly, 20 chromatin annotations, derived from Roadmap Epigenomics consortium[89] and EN-TEx[86,90], with epigenetic marks of me1, me2, me3, and ac, from four brain regions previously associated with stuttering,[38,40,41, 43–45,67] were used to estimate enrichment of heritability with stuttering. These marks were considered significant if p-value < 2.5×10^{-3}, Bonferroni adjusted (20 gene-sets).

Literature Replications

Gene replication analysis performed using methods detailed in Polikowsky et. al.[31]; however, previously calculated effective number of tests was then multiplied by eight, since we looked for replication of signals across our eight independent sex- and ancestries- specific GWAS. As such, the

effective number of tests used for our Bonferroni correction represented the number of independent tag SNPs in each gene with pairwise $r^2 < 0.4$ multiplied by eight. Gene replication results were Bonferroni corrected for the effective number of tests in each gene and the variant with the minimum p-value within each gene was reported.

SNP-based replications looked for replication of the top hits reported in Shaw *et al.*[32] and Polikowsky *et. al.*[31] across all eight sex- and ancestries-specific studies. Bonferroni multiple-test correction applied (p-value = 0.05/8 tests or significance: p-value $< 6.25 \times 10^{-3}$).

Stuttering polygenic risk-score model development

Polygenic risk-score models were trained using GWAS results from each sex- and ancestries-stratified in PRScs,[91] using a continuous shrinkage prior to adjust individual SNP weight for LD and variant significance. Default auto-phi parameters were used in both the male and female derived models and were not optimized to prevent overfitting, with LD reference panels constructed using 1KG phase 3 EUR reference for EA, and AFR for AA. Each model was applied to both the international stuttering project (ISP) sample[31] and Add Health sample,[33] matched according to ancestries and stratified by sex. The ISP testing set included 651 EA male stuttering cases and 4,264 sex-matched controls were included, as well as 242 EA female cases and 1,788 female-controls; 48 AA male stuttering cases and 308 sex-matched controls were included as well as 16 AA female stuttering cases and 90 sex-matched controls. The Add Health testing set included 352 EA male stuttering cases and 3,104 sex-matched controls; 236 EA female cases and 3,517 female-controls; 117 AA male stuttering cases and 847 sex-matched controls were included as well as 107 AA female stuttering cases and 1,101 sex-matched controls.

Genetic datasets were scored using PLINK 1v9.[92] Genetic liability scores were z-score normalized. Liability score distributions between cases and controls were compared via student's two-sample t-test.

Genetic Correlation

To assess common underlying genetic architecture between stuttering and various other comorbid traits, we performed a genetic correlation analysis, comparing the EA-male and EA-female GWAS results to sex-specific summary statistics from 17 traits with available sex-specific GWAS results obtained from http://www.nealelab.is/uk-biobank/ (see Table S14). Genetic correlations were also performed with EA-male, EA-female and EA-sex-combined with non-sex-stratified summary statistics of beat synchronization (see Table S14). These traits were selected to include phenotypes previously identified as comorbidities with stuttering.[13, 46–51] Each binary trait needed a case size > 1000, due to these constraints genetic correlations were only performed in the European-specific GWAS. All genetic correlation estimates were calculated between the European stuttering GWAS results (male and female-specific for 17 sex-stratified traits, and male, female, and sex-combined for beat synchronization) through LDSC.[36,37]

Mendelian Randomization

We performed both Egger and weighted median Mendelian randomization analysis using the MendelianRandomization R package.[93] We performed a sex-specific Mendelian randomization analysis for each of 18 traits with prior evidence of association with stuttering in the literature[13, 46–51] (hearing loss, asthma, dermatitis/eczema, hayfever/allergic rhinitis, sleep duration, daytime sleepiness, chronotype, recent thoughts of suicide or self-harm/suicide ideation, depression, anxiety, epilepsy, Attention-Deficit/Hyperactivity Disorder, alcohol dependency, alcohol consumption frequency, walking pace, body mass index, testosterone, and beat synchronization obtained from http://www.nealelab.is/uk-biobank/, PGC + iPSYCH data,[94] or 23andMe,Inc.[56]) included in our genetic correlation analysis to determine if the genetic risk for any trait was either causally related or pleotropic to stuttering in our EA-male and EA-female GWAS. In addition to EA-male and EA-female, we performed Mendelian randomization with our EA-sex-combined GWAS with our sex-combined trait: beat synchronization. We filtered input variants for each included trait to only include variants that 1) were included in both the trait and stuttering GWAS and 2) variant p-value in each GWAS $< 5 \times 10^{-6}$, and the independent instrumental SNPs were selected based on linkage disequilibrium from 1000 Genome European (LD pruning with 1000 kb windows, 1 SNPs each step, and LD < 0.2 by PLINK). Analysis details and results annotated in Table S15.

Gene module Enrichment

We performed an enrichment test for gene modules using our top identified genes associated with variant signals in either the EA-male or EA-female GWAS to determine if any sets of highly correlated genes (gene modules) were associated with stuttering risk (See Supplemental Methods). Top associated genes were determined for all variants with a p-value $< 5 \times 10^{-6}$ using Open Targets Genetics V2G pipeline.[80,81] Gene co-expression networks comprised groups of functionally related genes or "modules" Gerring *et al.*[95] identified from GTEx v.7 tissue gene expression data. Module enrichments were reported for any gene tissue-specific analysis with an FDR adjusted p-value < 0.05 among any of the 49 available GTEx tissues. We performed a competitive gene pathway analysis for reported module enrichments using g:Profiler and subsequently annotated the outputted biological pathways (Tables S20-24).

Colocalization

We performed a Bayesian colocalization analysis between our EA-male and EA-female genome-wide significant hits and tissue-specific eQTL signals from GTEx v.8 data[88] using fast enrichment aided colocalization analysis (See Supplemental Methods).[96,97] We looked for colocalization solely within regions with a variant identified as a top hit (Tables 1–3). Evaluated regions included all sentinel variants in either the EA-male, EA-female, AA-male, AdA-male, or AdA-female GWAS as well as any other variant found in the same LD block. LD blocks were defined according to European-based LD calculated from 1000 Genomes reference.[98] Colocalization analyses were tissue-specific and included all tissues available in GTEx v.8. We reported the results of any colocalization signal with a regional colocalization probability (RCP) (i.e., the probability that one of two SNPs in an LD block is responsible for a genuine association) > 0.05 (Table S25).

GWAS Catalog

After filtering the GWAS Catalog[99] (Release Date: 2022-21-12) to contain only genome-wide significant loci (p-value < 5.0×10^{-8}), 20 of 30 unique genes from sex-specific and meta- and mega- analyses were successfully found in the GWAS Catalog by querying the listed mapped genes. From these queried findings, GWAS Catalog associated traits were binned into 20 trait categories (Table S19a). Our genome-wide significant hits, and the associated GWAS Catalog findings can be found in (Table S19b). Number of unique genes per category can be seen within Fig. 5.

Declarations

Author Contributions

J.E.B. oversaw the entire study. J.E.B. and S.K. conceived the study. H.G.P, A.C.S., D.M.S., and J.E.B. drafted the manuscript, with sections contributed by D.G.P. and H.C. H.G.P., A.C.S., D.M.S., H.C., L.E.P., and H.M.H. performed analyses. A.S.P. provided computational support to all analyses. H.M.H., C.L.A., and K.M.H. managed the Add Health Study. J.M.B., E.L., S.K., and K.Z.V., managed the ISP cohort. R.L.G. contributed the Beat Synchronization GWAS summary statistics. 23andMe, Inc. performed all sex- and ancestries-specific GWAS analyses. All authors critically reviewed the manuscript.

Ethics Declarations

Competing Interests

Members of the 23andMe Research team are employed by 23andMe, Inc. H.G.P. is currently employed by Verily Life Sciences, an independent subsidiary of Alphabet Inc.

Ethical Standards

We have complied with all ethical guidelines. All participants provided informed consent for participating in research. This study has been approved by Vanderbilt IRB #181575 and #180583.

Code Availability

The code for the concordance analysis can be found on GitHub (https://github.com/belowlab/Concordance-Analysis).

Data Availability

Sex- and ancestries specific summary statistics of self-reported stuttering will be made available through the 23andMe website (https://research.23andme.com/dataset-access/) to qualified researchers under agreement with 23andMe that protects the privacy of the 23andMe participants.

References

1. Brady, N. C., Thiemann-Bourque, K., Fleming, K. & Matthews, K. Predicting Language Outcomes for Children Learning Augmentative and Alternative Communication: Child and Environmental Factors. *J. Speech Lang. Hear. Res.* **56**, 1595–1612 (2013).
2. Yairi, E. & Ambrose, N. Epidemiology of stuttering: 21st century advances. *J. Fluen. Disord.* **38**, 66–87 (2013).
3. Craig, A., Hancock, K., Tran, Y., Craig, M. & Peters, K. Epidemiology of Stuttering in the Community Across the Entire Life Span. *J. Speech Lang. Hear. Res.* **45**, 1097–1105 (2002).
4. Månsson, H. Childhood stuttering. *J. Fluen. Disord.* **25**, 47–57 (2000).
5. Singer, C. M., Hessling, A., Kelly, E. M., Singer, L. & Jones, R. M. Clinical Characteristics Associated With Stuttering Persistence: A Meta-Analysis. *J. Speech Lang. Hear. Res.* **63**, 2995–3018 (2020).
6. Singer, C. M., Otieno, S., Chang, S.-E. & Jones, R. M. Predicting Persistent Developmental Stuttering Using a Cumulative Risk Approach. *J. Speech Lang. Hear. Res.* **65**, 70–95 (2022).

7. Riley, J., Riley, G. & Maguire, G. Subjective Screening of Stuttering severity, locus of control and avoidance: research edition. *J. Fluen. Disord.* **29**, 51–62 (2004).

8. Baxter, S. *et al.* The state of the art in non-pharmacological interventions for developmental stuttering. Part 1: a systematic review of effectiveness. *Int. J. Lang. Commun. Disord.* **50**, 676–718 (2015).

9. Costa, D. & Kroll, R. Stuttering: an update for physicians. *CMAJ Can. Med. Assoc. J. J. Assoc. Medicale Can.* **162**, 1849–1855 (2000).

10. Daniels, D. E. & Gabel, R. M. The Impact of Stuttering on Identity Construction: *Top. Lang. Disord.* **24**, 200–215 (2004).

11. Daniels, D. E., Gabel, R. M. & Hughes, S. Recounting the K-12 school experiences of adults who stutter: A qualitative analysis. *J. Fluen. Disord.* **37**, 71–82 (2012).

12. McAllister, J., Collier, J. & Shepstone, L. The impact of adolescent stuttering on educational and employment outcomes: Evidence from a birth cohort study. *J. Fluen. Disord.* **37**, 106–121 (2012).

13. Briley, P. M., Gerlach, H. & Jacobs, M. M. Relationships between stuttering, depression, and suicidal ideation in young adults: Accounting for gender differences. *J. Fluen. Disord.* **67**, 105820 (2021).

14. Klein, J. F. & Hood, S. B. The impact of stuttering on employment opportunities and job performance. *J. Fluen. Disord.* **29**, 255–273 (2004).

15. Craig, A., Blumgart, E. & Tran, Y. The impact of stuttering on the quality of life in adults who stutter. *J. Fluen. Disord.* **34**, 61–71 (2009).

16. Shugart, Y. Y. *et al.* Results of a genome-wide linkage scan for stuttering. *Am. J. Med. Genet.* **124A**, 133–135 (2004).

17. Riaz, N. *et al.* Genomewide Significant Linkage to Stuttering on Chromosome 12. *Am. J. Hum. Genet.* **76**, 647–651 (2005).

18. Suresh, R. *et al.* New Complexities in the Genetics of Stuttering: Significant Sex-Specific Linkage Signals. *Am. J. Hum. Genet.* **78**, 554–563 (2006).

19. Wittke-Thompson, J. K. *et al.* Genetic studies of stuttering in a founder population. *J. Fluen. Disord.* **32**, 33–50 (2007).

20. Kang, C. *et al.* Mutations in the Lysosomal Enzyme–Targeting Pathway and Persistent Stuttering. *N. Engl. J. Med.* **362**, 677–685 (2010).

21. Lan, J. *et al.* Association between dopaminergic genes (SLC6A3 and DRD2) and stuttering among Han Chinese. *J. Hum. Genet.* **54**, 457–460 (2009).

22. Domingues, C. E. F. *et al.* A genetic linkage study in Brazil identifies a new locus for persistent developmental stuttering on chromosome 10. *Genet. Mol. Res.* **13**, 2094–2101 (2014).

23. Mohammadi, H. *et al.* Sex steroid hormones and sex hormone binding globulin levels, CYP17 MSP AI (−34 T:C) and CYP19 codon 39 (Trp:Arg) variants in children with developmental stuttering. *Brain Lang.* **175**, 47–56 (2017).

24. Raza, M. H., Amjad, R., Riazuddin, S. & Drayna, D. Studies in a consanguineous family reveal a novel locus for stuttering on chromosome 16q. *Hum. Genet.* **131**, 311–313 (2012).

25. Raza, M. H. *et al.* Association between Rare Variants in AP4E1, a Component of Intracellular Trafficking, and Persistent Stuttering. *Am. J. Hum. Genet.* **97**, 715–725 (2015).

26. van Beijsterveldt, C. E. M., Felsenfeld, S. & Boomsma, D. I. Bivariate Genetic Analyses of Stuttering and Nonfluency in a Large Sample of 5-Year-Old Twins. *J. Speech Lang. Hear. Res.* **53**, 609–619 (2010).

27. Fagnani, C., Fibiger, S., Skytthe, A. & Hjelmborg, J. V. B. Heritability and environmental effects for self-reported periods with stuttering: A twin study from Denmark. *Logoped. Phoniatr. Vocol.* **36**, 114–120 (2011).

28. Kazemi, N., Estiar, M. A., Fazilaty, H. & Sakhinia, E. Variants in GNPTAB, GNPTG and NAGPA genes are associated with stutterers. *Gene* **647**, 93–100 (2018).

29. Kang, C. *et al.* Evaluation of the association between polymorphisms at the DRD2 locus and stuttering. *J. Hum. Genet.* **56**, 472–473 (2011).

30. Frigerio Domingues, C. E. *et al.* Are variants in sex hormone metabolizing genes associated with stuttering? *Brain Lang.* **191**, 28–30 (2019).

31. Polikowsky, H. G. *et al.* Population-based genetic effects for developmental stuttering. *Hum. Genet. Genomics Adv.* **3**, 100073 (2022).

32. Shaw, D. M. *et al.* Phenome risk classification enables phenotypic imputation and gene discovery in developmental stuttering. *Am. J. Hum. Genet.* **108**, 2271–2283 (2021).

33. Harris, K. M. *et al.* Cohort Profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). *Int. J. Epidemiol.* **48**, 1415–1415k (2019).

34. Durand, E. Y., Do, C. B., Mountain, J. L. & Macpherson, J. M. *Ancestry Composition: A Novel, Efficient Pipeline for Ancestry Deconvolution*. http://biorxiv.org/lookup/doi/10.1101/010512 (2014) doi:10.1101/010512.

35. Mägi, R. *et al.* Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum. Mol. Genet.* **26**, 3639–3650 (2017).

36. ReproGen Consortium *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).

37. Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

38. Lu, C. *et al.* The neural substrates for atypical planning and execution of word production in stuttering. *Exp. Neurol.* **221**, 146–156 (2010).

39. Chang, S.-E., Garnett, E. O., Etchell, A. & Chow, H. M. Functional and Neuroanatomical Bases of Developmental Stuttering: Current Insights. *The Neuroscientist* **25**, 566–582 (2019).

40. Etchell, A. C., Civier, O., Ballard, K. J. & Sowman, P. F. A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *J. Fluen. Disord.* **55**, 6–45 (2018).

41. Liu, J. *et al.* A Functional Imaging Study of Self-Regulatory Capacities in Persons Who Stutter. *PLoS ONE* **9**, e89891 (2014).

42. Neef, N. E. *et al.* Altered morphology of the nucleus accumbens in persistent developmental stuttering. *J. Fluen. Disord.* **55**, 84–93 (2018).

43. Toyomura, A., Fujii, T. & Kuriki, S. Effect of an 8-week practice of externally triggered speech on basal ganglia activity of stuttering and fluent speakers. *NeuroImage* **109**, 458–468 (2015).

44. Chang, S.-E. & Zhu, D. C. Neural network connectivity differences in children who stutter. *Brain* **136**, 3709–3726 (2013).

45. Chang, S.-E., Horwitz, B., Ostuni, J., Reynolds, R. & Ludlow, C. L. Evidence of Left Inferior Frontal–Premotor Structural and Functional Connectivity Deficits in Adults Who Stutter. *Cereb. Cortex* **21**, 2507–2518 (2011).

46. Arenas, R. M., Walker, E. A. & Oleson, J. J. Developmental Stuttering in Children Who Are Hard of Hearing. *Lang. Speech Hear. Serv. Sch.* **48**, 234–248 (2017).

47. Pruett, D. G. *et al.* Identifying developmental stuttering and associated comorbidities in electronic health records and creating a phenome risk classifier. *J. Fluen. Disord.* **68**, 105847 (2021).

48. Briley, P. M. & Merlo, S. Presence of Allergies and Their Impact on Sleep in Children Who Stutter. *Perspect. ASHA Spec. Interest Groups* **5**, 1454–1466 (2020).

49. Iverach, L. *et al.* Prevalence of anxiety disorders among children who stutter. *J. Fluen. Disord.* **49**, 13–28 (2016).

50. Alm, P. A. & Risberg, J. Stuttering in adults: The acoustic startle response, temperamental traits, and biological factors. *J. Commun. Disord.* **40**, 1–41 (2007).

51. 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).

52. Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five Years of GWAS Discovery. *Am. J. Hum. Genet.* **90**, 7–24 (2012).

53. Gusev, A. *et al.* Quantifying Missing Heritability at Known GWAS Loci. *PLoS Genet.* **9**, e1003993 (2013).

54. The 23andMe Research Team *et al.* Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* **51**, 394–403 (2019).

55. Xue, A. *et al.* Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. *Nat. Commun.* **9**, 2941 (2018).

56. Niarchou, M. *et al.* Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nat. Hum. Behav.* **6**, 1292–1309 (2022).

57. Pirastu, N. *et al.* Genetic analyses identify widespread sex-differential participation bias. *Nat. Genet.* **53**, 663–671 (2021).

58. Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants: Fig. 1. *Bioinformatics* **31**, 3555–3557 (2015).

59. Garnett, E. O. *et al.* Auditory rhythm discrimination in adults who stutter: An fMRI study. *Brain Lang.* **236**, 105219 (2023).

60. 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).

61. Wingate, M. E. & Howell, P. *Foundations of Stuttering*. *J. Acoust. Soc. Am.* **112**, 1229–1231 (2002).

62. Brady, J. P. Metronome-conditioned speech retraining for stuttering. *Behav. Ther.* **2**, 129–150 (1971).

63. Brady, J. P. Studies on the metronome effect on stuttering. *Behav. Res. Ther.* **7**, 197–204 (1969).

64. Chow, H. M. *et al.* Linking Lysosomal Enzyme Targeting Genes and Energy Metabolism with Altered Gray Matter Volume in Children with Persistent Stuttering. *Neurobiol. Lang.* **1**, 365–380 (2020).

65. Davis, S., Howell, P. & Cooke, F. Sociodynamic relationships between children who stutter and their non-stuttering classmates. *J. Child Psychol. Psychiatry* **43**, 939–947 (2002).

66. Walden, T. A. & Lesner, T. A. Examining implicit and explicit attitudes toward stuttering. *J. Fluen. Disord.* **57**, 22–36 (2018).

67. Loucks, T., Kraft, S. J., Choo, A. L., Sharma, H. & Ambrose, N. G. Functional brain activation differences in stuttering identified with a rapid fMRI sequence. *J. Fluen. Disord.* **36**, 302–307 (2011).

68. Ardila, A. *et al.* An epidemiologic study of stuttering. *J. Commun. Disord.* **27**, 37–48 (1994).

69. Bernard, R., Hofslundsengen, H. & Frazier Norbury, C. Anxiety and Depression Symptoms in Children and Adolescents Who Stutter: A Systematic Review and Meta-Analysis. *J. Speech Lang. Hear. Res.* **65**, 624–644 (2022).

70. Corcoran, J. A. & Stewart, M. Stories of stuttering. *J. Fluen. Disord.* **23**, 247–264 (1998).

71. Boyle, M. P. Enacted stigma and felt stigma experienced by adults who stutter. *J. Commun. Disord.* **73**, 50–61 (2018).

72. Briley, P. M., Merlo, S. & Ellis, C. Sex Differences in Childhood Stuttering and Coexisting Developmental Disorders. *J. Dev. Phys. Disabil.* **34**, 505–527 (2022).

73. Seider, R. A., Kidd, K. K. & Gladstien, K. L. Recovery and Persistence of Stuttering among Relatives of Stutterers. *J. Speech Hear. Disord.* **48**, 402–409 (1983).

74. Cox, N. J. & Kidd, K. K. Can recovery from stuttering be considered a genetically milder subtype of stuttering? *Behav. Genet.* **13**, 129–139 (1983).

75. Ambrose, N. G., Yairi, E., Loucks, T. M., Seery, C. H. & Throneburg, R. Relation of motor, linguistic and temperament factors in epidemiologic subtypes of persistent and recovered stuttering: Initial findings. *J. Fluen. Disord.* **45**, 12–26 (2015).

76. Yairi, E., Ambrose, N. & Cox, N. Genetics of Stuttering: A Critical Review. *J. Speech Lang. Hear. Res.* **39**, 771–784 (1996).

77. Yairi, E. Subtyping stuttering I: A review. *J. Fluen. Disord.* **32**, 165–196 (2007).

78. Bloodstein, O. & N Bernstein Ratner. *A handbook on stuttering*. (2008).

79. Risch, N. & Merikangas, K. The Future of Genetic Studies of Complex Human Diseases. *Science* **273**, 1516–1517 (1996).

80. Ghoussaini, M. *et al.* Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res.* **49**, D1311–D1320 (2021).

81. Mountjoy, E. *et al.* An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nat. Genet.* **53**, 1527–1533 (2021).

82. Wakefield, J. A Bayesian Measure of the Probability of False Discovery in Genetic Epidemiology Studies. *Am. J. Hum. Genet.* **81**, 208–227 (2007).

83. The Wellcome Trust Case Control Consortium *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).

84. INSIGHTS FROM ESTIMATES OF SNP-HERITABILITY FOR >2,000 TRAITS AND DISORDERS IN UK BIOBANK. http://www.nealelab.is/blog/2017/9/20/insights-from-estimates-of-snp-heritability-for-2000-traits-and-disorders-in-uk-biobank.

85. ReproGen Consortium *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).

86. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

87. Cahoy, J. D. *et al.* A Transcriptome Database for Astrocytes, Neurons, and Oligodendrocytes: A New Resource for Understanding Brain Development and Function. *J. Neurosci.* **28**, 264–278 (2008).

88. The GTEx Consortium *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).

89. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).

90. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

91. 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).

92. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

93. Yavorska, O. O. & Burgess, S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.* **46**, 1734–1739 (2017).

94. Martin, J. *et al.* A Genetic Investigation of Sex Bias in the Prevalence of Attention-Deficit/Hyperactivity Disorder. *Biol. Psychiatry* **83**, 1044–1053 (2018).

95. Gerring, Z. F., Gamazon, E. R., Derks, E. M., & for the Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. A gene co-expression network-based analysis of multiple brain tissues reveals novel genes and molecular pathways underlying major depression. *PLOS Genet.* **15**, e1008245 (2019).

96. Pividori, M. *et al.* PhenomeXcan: Mapping the genome to the phenome through the transcriptome. *Sci. Adv.* **6**, eaba2083 (2020).

97. Wen, X., Pique-Regi, R. & Luca, F. Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. *PLOS Genet.* **13**, e1006646 (2017).

98. Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* btv546 (2015) doi:10.1093/bioinformatics/btv546.

99. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
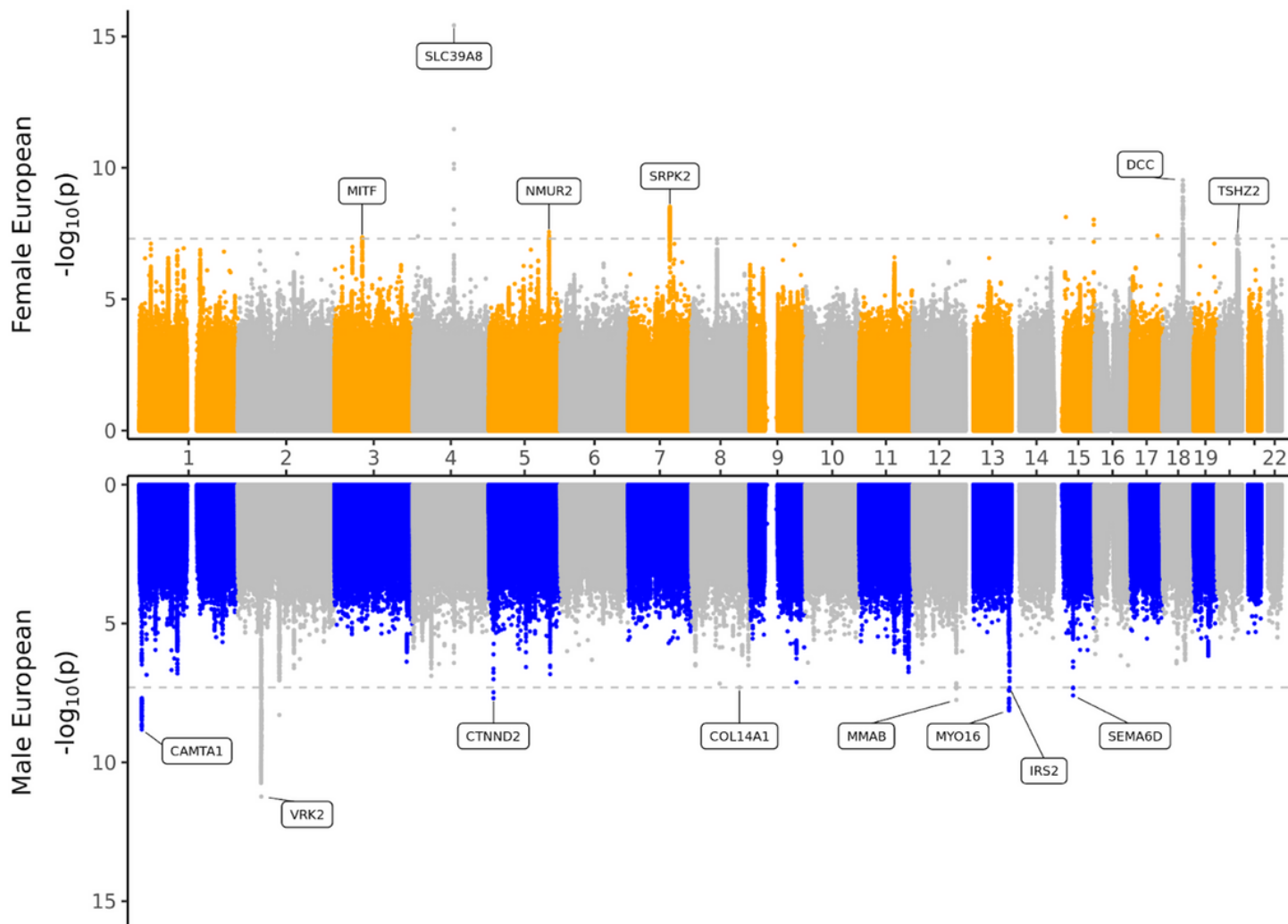
Figures

Figure 1

Miami plot for EA-female and EA-male genome-wide association studies. The EA-female association study (top panel in orange) included 570,071 total samples (40,137 self-reported stuttering cases) and 29,449,463 autosomal variants. Nine loci reached genome-wide significance (dotted line p-value < $5x10^{-8}$). The EA-male association study (bottom panel in blue) included 374,279 total samples (38,257 self-reported stuttering cases) and 29,409,446 autosomal variants. Ten loci reached genome-wide significance (dotted line p-value < $5x10^{-8}$). The x-axis represents chromosome base pair coordinates in human genome build 37 and the y-axis represents observed $-log_{10}$(p-value) for each GWAS. Annotated genes for each GWAS includes the predicted functional gene for each loci (when available) according to Open Targets V2G pipeline, which integrates evidence from molecular quantitative trait loci, chromatin interactions, in silico functional predictions from Ensembl, and distance between the variant and gene canonical transcription start site.
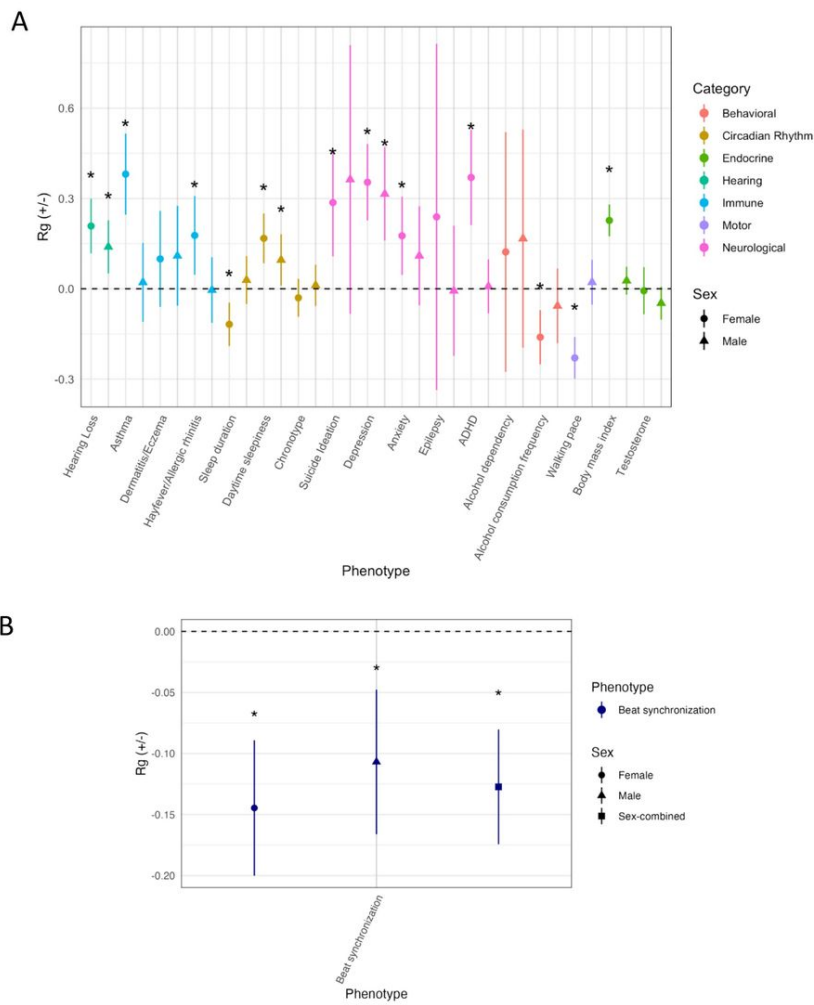
Figure 2

Forest plot showing genetic correlations for stuttering susceptibility with previously reported comorbid traits. A) Sex and ancestries-specific genetic correlations performed for each indicated trait against self-reported stuttering in EA-males and EA-females. Male-specific correlations designated by triangles and female-specific correlations designated by circles. Each trait color-coded according to its descriptive category (behavioral, circadian rhythm, immune, metabolic, motor, neurological, speech and language). Bars indicate standard error for each performed genetic correlation analysis. Significant correlations marked with an asterisk. B) Sex-specific and sex-combined ancestries-specific genetic correlations performed for beat synchronization. Male-specific correlations designated by triangles, female-specific correlations designated by circles, and sex-combined correlations designed by squares. Bars indicate standard error for each performed genetic correlation analysis. Significant correlations marked with an asterisk.
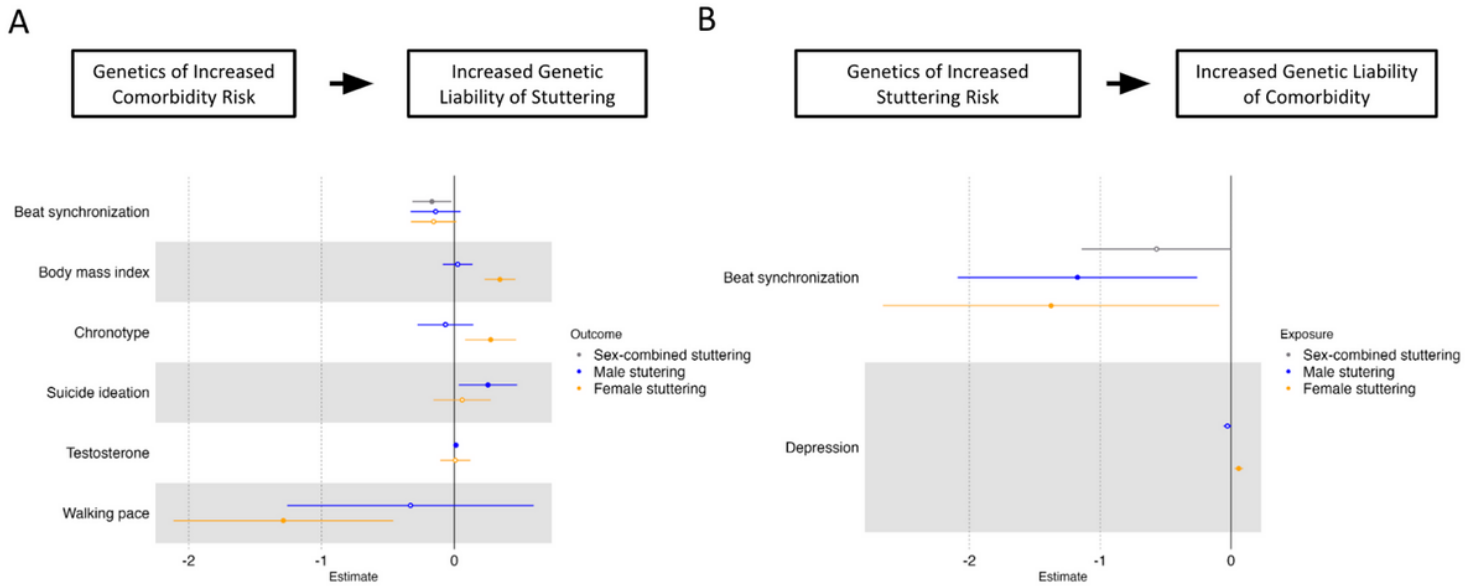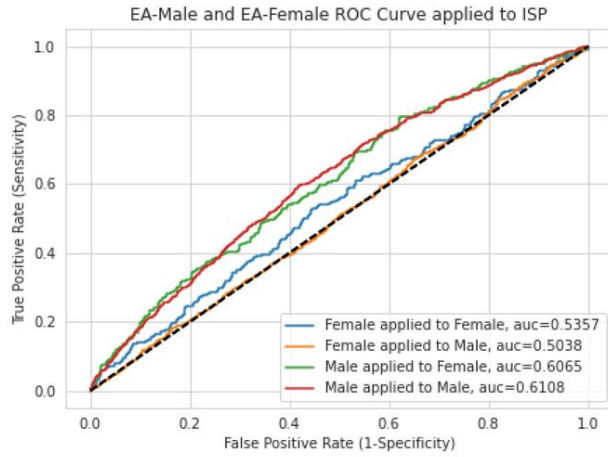
Figure 3

Forest plot showing Mendelian randomization of stuttering susceptibility with previously-reported comorbid traits. A) Mendelian randomization analysis (Egger approach) investigating vertical pleiotropy (comorbid trait → stuttering). B) Mendelian randomization analysis (Egger approach) investigating vertical pleiotropy (stuttering → comorbid trait). All results with closed circles are statistically significant (Egger *p*-value < 0.05)
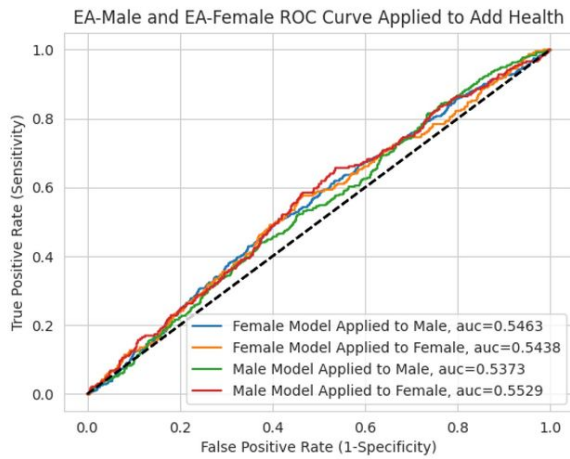
A

EA-Male and EA-Female ROC Curve applied to ISP

- Female applied to Female, auc=0.5357
- Female applied to Male, auc=0.5038
- Male applied to Female, auc=0.6065
- Male applied to Male, auc=0.6108

B

EA-Male and EA-Female ROC Curve Applied to Add Health

- Female Model Applied to Male, auc=0.5463
- Female Model Applied to Female, auc=0.5438
- Male Model Applied to Male, auc=0.5373
- Male Model Applied to Female, auc=0.5529

Figure 4

Self-report stuttering polygenic risk-score model performance in EA-stuttering test sets. Polygenic risk scores developed using EA-male or female GWAS results and applied to clinically validated International Stuttering Project and self-report Add Health subjects demonstrate increased stuttering liability within stuttering cases. Model was developed and trained using default auto-phi shrinkage parameter through PRScs. LD panels were constructed using 1KG phase 3 EUR reference. A) Represents AUC metric performance in the ISP cohort B) Represents AUC metric performance in the Add Health cohort
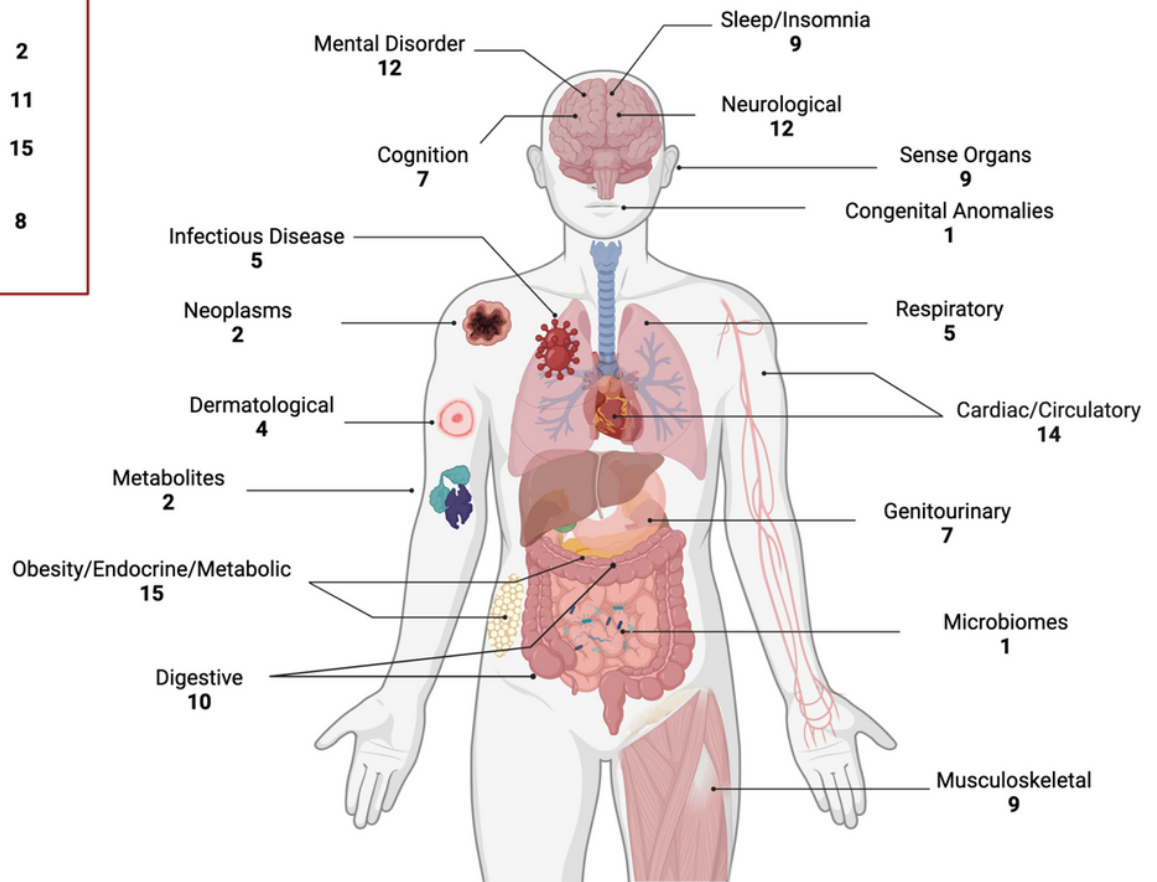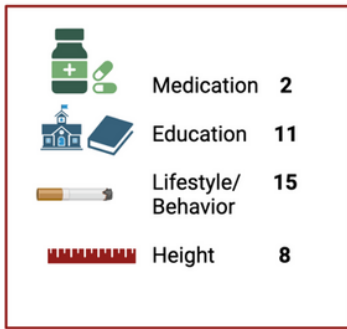
Figure 5

GWAS trait categories and sum of unique stuttering risk genes. Human body figure showing the number of unique self-reported stuttering genes associated with categorized traits from GWAS Catalog (Release Date: 21-12-2022). Twenty genes out of the 30 genes across sex-specific and meta- and mega- analyses were found in our GWAS Catalog search. All GWAS Catalog traits and categories associated with our genome-wide significant hits are listed in STable 19.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 23andMeStutteringSupplementalTables.xlsx
- 23andMeStutteringSupplementalText.docx
- 23andMeStutteringExtendedFigures.pdf
- 23andMeStutteringSupplementalFigures.pdf