

Genome-wide association study of social isolation in 63,497 Japanese individuals from the general population

Received: 29 September 2024

Revised: 21 December 2025

Accepted: 1 February 2026

Cite this article as: Ohseto, H., Inoue, K., Takahashi, I. *et al.* Genome-wide association study of social isolation in 63,497 Japanese individuals from the general population. *Transl Psychiatry* (2026). <https://doi.org/10.1038/s41398-026-03896-9>

Hisashi Ohseto, Kosuke Inoue, Ippei Takahashi, Taku Obara, Akira Narita, Mami Ishikuro, Masatsugu Orui, Keiko Murakami, Aoi Noda, Genki Shinoda, Masato Takase, Naoki Nakaya, Mana Kogure, Rieko Hatanaka, Kumi Nakaya, Ippei Chiba, Sayuri Tokioka, Yuka Kotozaki, Atsushi Shimizu, Kozo Tanno, Atsushi Hozawa, Gen Tamiya, Naoki Kondo & Shinichi Kuriyama

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Genome-wide Association Study of Social Isolation in 63 497 Japanese Individuals from the General Population

Running title: Genome-wide Association Study of Social Isolation

Hisashi Ohseto, MD,¹ Kosuke Inoue, MD, PhD,² Ippei Takahashi, MMSc,³ Taku Obara, PhD,^{3,4,5} Akira Narita, PhD,^{3,4} Mami Ishikuro, PhD,^{3,4} Masatsugu Orui, MD, PhD,^{1,3,4} Keiko Murakami, MPH, PhD,^{4,6} Aoi Noda, PhD,^{3,4,5} Genki Shinoda, MS,^{3,4} Masato Takase, MPH, PhD,^{3,4} Naoki Nakaya, PhD,^{3,4} Mana Kogure, PhD,^{3,4} Rieko Hatanaka, PhD,^{3,4} Kumi Nakaya, PhD,^{3,4} Ippei Chiba, PhD,^{3,4} Sayuri Tokioka, MD,^{3,4} Yuka Kotozaki, PhD,^{7,8} Atsushi Shimizu, PhD,⁸ Kozo Tanno, MD, PhD,^{7,8} Atsushi Hozawa, MD, PhD,^{3,4} Gen Tamiya, PhD,^{3,4,9} Naoki Kondo, MD, PhD,¹⁰ and Shinichi Kuriyama, MD, PhD^{1,3,4}

¹ International Research Institute of Disaster Science, Tohoku University, Sendai, Japan.

² Department of Health Promotion and Behavioral Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan.

³ Graduate School of Medicine, Tohoku University, Sendai, Japan.

⁴ Tohoku Medical Megabank Organization, Tohoku University, Sendai, Japan.

⁵ Tohoku University Hospital, Tohoku University, Sendai, Japan.

⁶ Department of Health and Social Behavior, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

⁷ Department of Hygiene and Preventive Medicine, Iwate Medical University, Iwate, Japan.

⁸ Iwate Tohoku Medical Megabank Organization, Iwate Medical University, Iwate, Japan.

⁹ RIKEN Center for Advanced Intelligence Project, Tokyo, Japan.

¹⁰ Department of Social Epidemiology, Graduate School of Medicine and School of Public Health, Kyoto University, Kyoto, Japan.

Corresponding author

Shinichi Kuriyama, MD, PhD

2-1, Seiryomachi, Aoba-ku, Sendai-shi, Miyagi-ken, 980-8573, Japan

Phone: +81-22-717-8104

Fax: +81-22-717-8106

E-mail: shinichi.kuriyama.e6@tohoku.ac.jp

ARTICLE IN PRESS

Abstract (216 words)

Social isolation, characterized by a lack of social connections with family, friends, and others, is associated with adverse health outcomes. However, the genetic contribution to the susceptibility to social isolation remains unclear. This study aimed to identify genetic loci associated with social isolation using the Lubben Social Network Scale (LSNS-6) in a Japanese population. The Tohoku Medical Megabank Community-Based Cohort Study was conducted between 2013 and 2016. The participants were genotyped using the Affymetrix Axiom Japonica Array. The LSNS-6 was used to assess familial and friend ties through six questions and social isolation statuses were defined using the total scale, family subscale, and friend subscale. Genome-wide association studies (GWASs) were conducted using a generalized linear mixed model, adjusting for age, sex, 10 genetic principal components and batch effects. In total, 63 497 participants who completed genotyping and the LSNS-6 were included. The mean age was 59.4 ± 11.9 years, and 41 126 (64.8%) were female. Significant genetic loci were identified in GWASs for the total scale (rs10736933 near ACADSB and HMX3) and friend subscale of LSNS-6 (rs1778366 near LINC02315 and LRFN5). This study provides the first genome-wide evidence of social isolation in the Japanese population, suggesting associations with ACADSB, HMX3, LINC02315, and LRFN5. These findings could enable personalized prevention and intervention for social isolation and related psychiatric disorders.

Introduction

The epidemic of social isolation, a concept defined as a lack of quality and quantity in social connections with family members, friends, and others,¹ has been an underappreciated public health crisis that has harmed individual and societal health.² This condition is linked to mental health,^{3,4} cardiovascular disease,⁵ and mortality.⁶ Social factors play a significant role in social isolation, but it is important to recognize that it is not just a result of external influences. Social isolation also has a biological basis, and relevant genetic and pathological factors need to be explored.⁷⁻⁹ From a psychiatric perspective, a deficit or reduction in social interaction, which is closely related to social isolation, is often observed in patients with autism spectrum disorders,¹⁰ major depressive disorder,¹¹ schizophrenia,¹² and Alzheimer's disease.¹³ A previous study revealed that these psychiatric disorders share common genetic variant risks,¹⁴ which may partly be explained by a common genetic predisposition for social isolation. Therefore, establishing the genetic architecture of social isolation is imperative to understand its role in psychiatric conditions and other health outcomes.

Currently, four genome-wide association studies (GWASs) of social isolation¹⁵⁻¹⁸ have been registered in the GWAS catalog¹⁹ (trait ID: EFO_0009592; access date: Oct 21st, 2023), all of which were conducted using the UK Biobank database. Although these studies identified more than 30 associations under various definitions of social isolation, challenges persist in this research area. First, the existing studies lack validated questionnaires for social isolation assessment. Typically, these studies utilized either singular or a combination of self-report questionnaires, which lack thorough validation in terms of consistency, structural stability, and clinical correlation. The need for validated measures is critical for reliable replication,

considering the variability in the reported loci across different GWASs, which is likely due to the varying definitions of social isolation. Second, no GWAS of social isolation has been conducted among East Asian populations. The GWAS of social isolation in East Asian populations is important, especially given the distinct allele frequencies and linkage disequilibrium between East Asian and European populations. In addition, the prevalence of social isolation is higher in Japan than in the UK,²⁰ accentuating the need for genetic studies on social isolation in East Asian populations.

Therefore, we conducted GWASs of social isolation, defined by a validated questionnaire—the Lubben Social Network Scale (LSNS-6)²¹—in East Asian populations using population-based cohorts of more than 60 000 people in Miyagi and Iwate Prefectures in Japan.

Method

Study Population

The Tohoku Medical Megabank (TMM) Project,²² which was initiated after the Great East Japan Earthquake, was conducted by the Tohoku University Tohoku Medical Megabank Organization (ToMMo) and the Iwate Medical University Iwate Tohoku Medical Megabank Organization (IMM). ToMMo and IMM conducted the TMM Community-Based Cohort Study (TMM CommCohort Study),²³ a population-based adult cohort study using two different recruitment methods. Type 1 surveys were conducted among those who underwent a specific municipal health checkup, which was provided by insurers and covered all persons living in Japan. Type 2 surveys were conducted among those who voluntarily visited the assessment center through mass media, distribution of advertising flyers to public institutions, and advertisements at specific health checkup sites. Therefore, our study included four cohorts: type 1 survey conducted by the

ToMMo, type 2 survey conducted by the ToMMo, type 1 survey conducted by the IMM, and type 2 survey conducted by the IMM. For the type 1 surveys, 97 419 people were invited, of whom 67 355 participated, and for the type 2 surveys, 19 846 participants were recruited. A total of 87 201 participants were included in this study. Participants who withdrew their consent (n = 4 704), those who did not have a genotype (n = 15 800) or outcome information (n = 4 017), and those who had >4 standard deviations for the first and second genetic principal components (n = 195) were excluded (Supplementary Figure 1).

All study participants provided written informed consent, and ethical approval was obtained from the Ethics Committee of the TMM Organization (2023-4-110). All methods involving human participants were performed in accordance with the ethical standards of the institutional and/or national research committee and with the Declaration of Helsinki, as well as relevant guidelines and regulations.

Genotyping, imputation, and quality control

The participants were genotyped using the Affymetrix Axiom Japonica Array (v2) in 19 batches, each consisting of 50 plates. Details about the genotyping conducted in TMM have been previously described.²⁴ After batch genotyping, samples with call rates < 0.95 or exhibiting unusually high identical by descent values compared to others were excluded. Additionally, variants with Hardy–Weinberg equilibrium (HWE) P-values < 1.00×10^{-5} , minor allele frequency < 0.01, or missing fraction > 0.01 were excluded from each batch. The genotype datasets from the 19 batches were merged to obtain a direct genotype dataset in the PLINK BED format. Subsequently, principal component analysis was performed using PLINK (version 2.0)²⁵ on the direct genotype dataset after excluding 3rd degree relatives and projecting them onto the

remaining participants. For the imputed genotype dataset, pre-phasing was conducted using SHAPEIT2.²⁶ The phased genotypes were then imputed with a cross-imputed panel of 3.5KJPNv2²⁷ and 1KGP3²⁸ using IMPUTE4.²⁹ The cross-imputation panel for 3.5KJPNv2²⁷ and 1KGP3²⁸ was created by the -merge_ref_panels tool in IMPUTE2.³⁰ After imputation, variants with minor allele frequencies < 0.01 , HWE P-value $< 1.00 \times 10^{-5}$, and imputation information scores < 0.4 were excluded. Finally, 9 040 166 variants were included in the GWASs.

Outcomes

At the baseline survey, participants answered the LSNS-6, which evaluates familial and non-familial ties, using identical sets of three questions.²¹ The familial items include “How many relatives do you see or hear from at least once a month?,” “How many relatives do you feel at ease with that you can talk about private matters?,” and “How many relatives do you feel close to such that you can call on them for help?” Similarly, the non-familial items evaluate the same aspects but with friends instead of relatives. Each LSNS-6 item is scored from 0 to 5, contributing to a total score ranging from 0 to 30, with higher scores representing broader social networks. Social isolation is defined as a score of < 12 , indicating an average of two or fewer connections across the six social network dimensions assessed by the LSNS-6.²¹ The three familial LSNS-6 items form the family subscale, where limited familial ties are determined by a score of less than 6. Likewise, the three non-familial LSNS-6 items form the friend subscale, where limited non-familial ties are determined by a score of < 6 . Although different response patterns may yield identical total scores, the LSNS-6 is designed to capture a common latent construct of social connectedness. Its reliability and construct validity have been consistently demonstrated across multiple populations and cultural contexts^{21,31}. This approach is consistent

with previous behavioral and psychiatric GWAS that have used validated composite scales, such as the Center for Epidemiologic Studies Depression Scale and the Generalized Anxiety Disorder 7-item scale, to quantify complex social or psychological traits.

Statistical analysis

Age at recruitment, sex, and social isolation status were summarized for each cohort. For each social isolation status defined by the total scale, family subscale, and friend subscale, GWASs were conducted separately in the four cohorts. GCTA (1.94.0)³² was used with a generalized linear mixed model adjusted for age, sex, 10 genetic principal components, and batch effects. The results from the four cohorts were meta-analyzed with METAL³³ using a sample size-weighted method. Genome-wide statistical significance was set at a conventional P-value $< 5 \times 10^{-8}$. Clumping was applied using PRSice-2 (version 2.3.5)³⁴ to determine independent variants, with only significant or top-hit variants reported. These variants were annotated using ANNOVAR.³⁵ The genomic inflation factors were calculated for each GWAS meta-analysis by dividing the median of the observed chi-squared test statistics by the expected median of the corresponding chi-squared distribution. Heritability was estimated for each GWAS meta-analysis by LDSC (1.0.1), and the LDSC intercepts were examined to distinguish polygenicity from potential population stratification.³⁶ For the significant or top-hit variants, the relationship between genotype and social isolation status was assessed for each cohort by logistic regression analyses adjusted for age, sex, and 10 genetic principal components, and then meta-analyzed. In addition, cross-tabulations were performed according to the genotypes of participants' characteristics (Supplementary Methods). We additionally examined our lead variants in two previously published UK Biobank GWAS of social isolation.^{15,16}

Two sensitivity analyses were conducted. First, given that LSNS-6 responses can be misclassified depending on the participants' cognitive and mental status, the associations between significant or top-hit variants and social isolation status were analyzed in a population excluding participants who self-reported dementia or major depressive disorder. Second, considering that the impact of the Great East Japan Earthquake varied geographically and affected social ties, the analysis was conducted in a population excluding participants who reported that their houses were partially or completely damaged by the earthquake.

Results

A total of 63 497 participants were included in the ToMMo type 1 survey (n = 28 835), ToMMo type 2 survey (n = 9 510), IMM type 1 survey (n = 17 608), and IMM type 2 survey (n = 7 544) (Supplementary Figure 1). The mean age was 59.4 ± 11.9 years, and 41 126 (64.8%) participants were female. A total of 16 353 (25.8%), 9 418 (14.8%), and 21 638 (34.1%) participants showed social isolation on the total scale, family subscale, and friend subscale, respectively (Table 1). In each of the four cohorts, the number and percentage of socially isolated participants in the total scale were 7 339 (25.5%), 2 007 (21.1%), 4 940 (28.1%), and 2 067 (27.4%), respectively.

Two loci were significant in the GWAS meta-analyses (Figure 1). The significant variant in the total scale GWAS was rs10736933, located at 10q26.13 (Table 2). The rs10736933 variant has been annotated as an intergenic variant between Short/branched chain acyl-CoA dehydrogenase (ACADSB) and H6 family homeobox 3 (HMX3). ACADSB is a member of the acyl-CoA dehydrogenase family of enzymes with activity toward the short/branched chain acyl-CoA derivatives.³⁷ HMX3 is expected to have DNA-binding transcription factor activity, especially in nerve cell types.^{38,39} The significant variant in the friend subscale GWAS was

rs1778366, located at 14q21.1 (Table 2). The rs1778366 was annotated as an intergenic variant between LINC02315, leucine-rich repeats, and fibronectin type III domain-containing protein 5 (LRFN5). LINC02315 is a long intergenic non-protein-coding RNA. LRFN5, also known as synaptic adhesion-like molecule 5 (SALM5), is a small transmembrane protein gene that belongs to a family of postsynaptic organizers involved in synaptic development, organization, and plasticity.^{40,41} A top-hit but not significant variant in the family subscale GWAS was rs780848113, located at 1q32.1 (Table 2), annotated as an intergenic variant between leiomodulin 1 (LMOD1) and the translocase of the inner mitochondrial membrane 17A (TIMM17A). Each of the three variants was associated to a different degree with the three scales of social isolation, suggesting a scale-specific association (Figure 2). Those with the alternative allele of rs780848113, which was the top hit in the family subscale GWAS, had lower hemoglobin A1c measurements, while those with the alternative allele of rs1778366, which was significant in the friend subscale GWAS, had a higher prevalence of depressive symptoms (Supplementary Tables 1-3). Other participant characteristics were not associated with the genotypes of individual variants. Both significant variants were not replicated in the UK Biobank GWAS (Supplementary Table 4). Because rs780848113 was not available in either dataset, its high-LD proxy rs7542670 ($r^2 = 0.91$) was examined instead and showed significance ($P < 0.05$) in both GWAS.

Both variants were positively associated with social isolation across all four cohorts (Table 3). The genomic inflation factors, lambda, were 1.073, 1.046, and 1.085 in the GWASs of the total scale, family subscale, and friend subscale, respectively (Figure 1). The LDSC intercepts were 1.034, 1.035, and 1.046, suggesting that the slight inflation was mainly attributable to polygenicity rather than population stratification. Heritability was estimated at

4.0%, 2.4%, and 3.8% for the total, family, and friend subscales, respectively. The sensitivity analyses showed results comparable to those of the main analysis (Supplementary Table 5). To complement the binary analyses, we performed an additional analysis treating the LSNS-6 scores as continuous variables. The results were largely consistent with those of the main analyses (Supplementary Table 6).

Discussion

Using population-based cohorts of more than 60 000 people in Japan, we identified two loci associated with social isolation: rs10736933 near ACADSB and HMX3, and rs1778366 near LINC02315 and LRFN5. These variants showed consistent associations across the four cohorts, and the heritability estimates for social isolation ranged from 2.4% to 4.0%.

Two previous GWASs have reported significant genetic variants associated with social isolation. The first study¹⁵ using the UK Biobank in 2018 assessed information on social isolation status based on three questions (loneliness, rarely interacting with others, and ability to confide) and identified 15 distinct genome-wide significant loci, including BPTF, GPX1, and MTCH2. The second study¹⁶ using the UK Biobank in 2021, assessed sociability scores based on four social functioning-related self-report questions and identified 18 independent genome-wide significant loci, including ARNTL, DRD2, and ELAVL2. Although both studies were conducted in the UK Biobank, the genome-wide significant loci were quite different, possibly because of differences in social isolation definitions. Our GWAS, conducted outside the UK Biobank, provides novel insights by identifying genes not previously associated with social isolation, although our lead variants were not replicated in the previous UK Biobank GWASs. This could be attributed to several factors. First, our study was the first to use a validated questionnaire, the

LSNS-6, to evaluate social isolation. Second, the current study was a GWAS conducted outside the UK Biobank, which may have led to different results. Social isolation, which is a complex interplay of genetic and environmental factors, suggests that gene-environment interactions could account for the variations observed in the GWAS results. Third, differences in allele frequencies and linkage disequilibrium may alter the strength of associations for each variant.

Our study highlighted two variants, rs10736933 and rs1778366, around three protein-coding genes: ACADSB, HMX3, and LRFN5. The pattern of association between genotypes and the three scales of social isolation differed, suggesting a scale-specific etiology and genetic basis. In the GWAS catalog, intergenic variants between ACADSB and HMX3 were reported to be associated with six traits, including Lewy body dementia⁴² and response to antidepressant.⁴³ As a mendelian disease gene, ACADSB was reported to cause short/branched-chain acyl-CoA dehydrogenase deficiency, a rare autosomal recessive inborn error of metabolism. This causes neurometabolic symptoms such as epilepsy and developmental delay.⁴⁴ In GWAS catalog, variants around LRFN5 were reported in 57 traits, nearly half of which are neurobehavioral or psychiatric phenotypes, including well-being,⁴⁵ neuroticism,⁴⁶ educational attainment,⁴⁷ major depressive disorder,⁴⁸ and generalized anxiety disorder.⁴⁹ Moreover, structural changes in LRFN5 were reported to be associated with developmental delay and autism in several cases.⁵⁰ In this study, rs1778366 was associated with depressive symptoms, indicating a possible involvement of psychiatric etiology in social isolation in the friend subscale. In summary, our GWAS findings suggest the genetic relevance of social isolation to neurobehavioral or psychiatric phenotypes.

Two recent post-GWASs of social isolation revealed a relationship between social isolation and neurobehavioral or psychiatric phenotypes. The first study¹⁵ revealed a significant

enrichment in the expression of social isolation-associated genes in the central nervous system and strong positive genetic overlaps with neuroticism and depressive symptoms. The second study¹⁶ showed that the sociability score (the higher the score, the more sociable) had negative genetic correlations with autism spectrum disorders, depression, and schizophrenia. These two studies highlighted that social isolation has its own genetic background and is genetically related to other neurobehavioral and psychiatric phenotypes. Consistent with these findings, our study highlighted the genes implicated in these conditions. When considering heritability, previous studies have estimated the heritability of social isolation at 4.2%¹⁵ and 6.0%,¹⁶ which were higher than our results: 4.0%, 2.4%, and 3.8%. This may be because the LSNS-6 focused on objective facts, such as “How many relatives do you see or hear from at least once a month?,” whereas previous studies have also used subjective loneliness to measure social isolation; thus, the LSNS-6 was less affected by the psychological state of the participants, which generally exhibits high heritability.⁵¹ The distinction between objective and subjective measures of social isolation is critical and warrants careful consideration in future research. In addition, differences in social norms and community structures between Japan and Western societies such as the strength of family networks and patterns of daily communication may influence the variability and genetic architecture of social isolation. Slight differences in age distribution between cohorts (mean 59.4 ± 11.9 years in our study vs. 56.6 ± 8.0 years in the UK Biobank¹⁶) may also have played a minor role.

Implications

This study has two implications. First, it reveals two novel susceptibility loci that provide insights into the biological mechanisms involved in social isolation. Because reduced social

interaction is a common symptom of psychiatric disorders, these loci may provide clues to the pathogenesis of psychiatric disorders. Second, our GWAS results can be used for further post-GWAS research in East Asian populations. Methodologies such as genetic correlation, cell-type enrichment, and Mendelian randomization are potential approaches for uncovering significant insights into the impact of social isolation on health but require ethnic concordance with the target population. Given that the prevalence of social isolation is higher in Japan than in the UK,²⁰ we hope that future post-GWAS will help us understand its pathology and guide effective interventions for East Asian populations.

Strength and limitations

Our study is the first to conduct a GWAS of social isolation in East Asian populations, highlighting two novel loci. Social isolation was measured using a validated questionnaire (the LSNS-6). However, this study has several limitations. First, owing to the nature of the TMM project, which was initiated after the Great East Japan Earthquake, it is possible that some of the TMM project participants were affected by the disaster,⁵²⁻⁵⁴ which limits the generalizability of the present results. Sensitivity analyses focusing on participants whose houses were less damaged by the earthquake yielded results similar to those of the main analysis. However, the impact of the earthquake on social connections could be significant, and the results should be interpreted with caution. Second, our GWAS showed slightly high genomic inflation factors, which may indicate population stratification. However, since TMM project participants were recruited from only two prefectures in Japan, they are expected to be relatively homogeneous. In addition, the LDSC intercepts were close to 1.0, suggesting that the inflation was mainly attributable to polygenicity. Given social isolation is a neurobehavioral phenotype, which tends

to be polygenic,^{55,56} it is reasonable to attribute part of the high genomic inflation factor to polygenicity. Third, our study population included over 60 000 participants, which was relatively small for a GWAS, since previous studies in the UK biobank included nearly 500 000 participants. This may have limited the detection power of our study.

Conclusions

Our GWASs revealed the first genome-wide evidence for a genetic predisposition to social isolation in the East Asian population, highlighting the possible involvement of ACADSB, HMX3, LINC02315, and LRFN5.

Acknowledgments

The authors would like to thank all the participants who consented to participate in this study and all the staff at the Tohoku Medical Megabank Organization, Tohoku University, Iwate Tohoku Medical Megabank Organization, and Iwate Medical University.

A full list of the members of the Tohoku Medical Megabank Organization is available at <https://www.megabank.tohoku.ac.jp/english/a230901/>.

We used the artificial intelligence tool GPT-4 language model, version 4.0, developed by OpenAI, to assist in the initial drafting and editing of this manuscript. It provided suggestions for sentence structure and grammatical corrections. However, all intellectual contributions and final edits were made by the authors.

Funding

This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) KAKENHI [grant numbers 19H03894 and 22H03346] and the Japan Agency for Medical Research and Development (AMED) [grant numbers JP17km0105001, JP21tm0124005, and JP21tm0424601].

Authors' contributions

HO, KI, and IT conceptualized and designed the study.

HO and IT conducted the statistical analyses and generated the figures.

HO drafted the initial version of the manuscript.

KI and NK provided conceptual advice on social isolation and contributed to interpretation of the findings.

AN (Narita) and GT contributed to the processing, quality control, and interpretation of genomic data.

TO, MI, MO, KM, AN (Noda), GS, MT, NN, MK, RH, KN, IC, and ST contributed to data acquisition, cohort management, and data curation within ToMMo.

YK, AS, and KT contributed to data acquisition, cohort management, and data curation within IMM.

AH and GT provided methodological advice and contributed to interpretation of the findings. SK and NK supervised the study.

Information on previous presentation

This work has not been previously presented anywhere.

Conflict of interest

KM is an employee of the Ministry of Education, Culture, Sports, Science and Technology, Japan. The other authors have no conflict of interests.

Data Sharing Statement

Individual data are available upon request after the approval of the Ethical Committee and the Materials and Information Distribution Review Committee of Tohoku Medical Megabank Organization.

References

1. Wang J, Lloyd-Evans B, Giacco D, Forsyth R, Nebo C, Mann F et al. Social isolation in mental health: a conceptual and methodological review. *Soc Psychiatry Psychiatr Epidemiol* 2017; **52**(12): 1451–1461.
2. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. *Public Health* 2017; **152**: 157–171.
3. Santini ZI, Jose PE, York Cornwell E, Koyanagi A, Nielsen L, Hinrichsen C et al. Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *Lancet Public Health* 2020; **5**(1): e62–e70.
4. Noguchi T, Saito M, Aida J, Cable N, Tsuji T, Koyama S et al. Association between social isolation and depression onset among older adults: a cross-national longitudinal study in England and Japan. *BMJ Open* 2021; **11**(3): e045834.
5. Steptoe A, Kivimäki M. Stress and cardiovascular disease: an update on current knowledge. *Annu Rev Public Health* 2013; **34**: 337–354.
6. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. *PLOS Med* 2010; **7**(7): e1000316.
7. Cacioppo JT, Hawkley LC, Norman GJ, Berntson GG. Social isolation. *Ann N Y Acad Sci*. 2011; **1231**(1): 17–22.

8. Nicholson NR. A review of social isolation: an important but underassessed condition in older adults. *J Prim Prev* 2012; **33**(2-3): 137–152.
9. Jang Y, Park NS, Chiriboga DA, Yoon H, Ko J, Lee J et al. Risk factors for social isolation in older Korean Americans. *J Aging Health* 2016; **28**(1): 3–18.
10. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. *Lancet* 2018; **392**(10146): 508–520.
11. Kupferberg A, Bicks L, Hasler G. Social functioning in major depressive disorder. *Neurosci Biobehav Rev* 2016; **69**: 313–332.
12. Galderisi S, Mucci A, Buchanan RW, Arango C. Negative symptoms of schizophrenia: new developments and unanswered research questions. *Lancet Psychiatry* 2018; **5**(8): 664–677.
13. Reichman WE, Negron A. Negative symptoms in the elderly patient with dementia. *Int J Geriatr Psychiatry* 2001; **16**(suppl 1): S7–S11.
14. Anttila V, Bulik-Sullivan B, Finucane HK, Finucane HK, Walters RK, Bras J et al. Analysis of shared heritability in common disorders of the brain. *Science* 2018; **360**(6395): eaap8757.
15. Day FR, Ong KK, Perry JRB. Elucidating the genetic basis of social interaction and isolation. *Nat Commun* 2018; **9**(1): 2457.
16. Bralten J, Mota NR, Klemann CJHM, Witte WD, Laing E, Collie DA et al. Genetic underpinnings of sociability in the general population. *Neuropsychopharmacology* 2021;

- 46(9):** 1627–1634.
17. Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* 2021; **599**(7886): 628–634.
 18. Jiang L, Zheng Z, Fang H, Yang J. A generalized linear mixed model association tool for biobank-scale data. *Nat Genet* 2021; **53**(11): 1616–1621.
 19. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* 2023; **51**(D1): D977–D985.
 20. Ikeda T, Cable N, Saito M, Koyama S, Tsuji T, Noguchi T et al. Association between social isolation and smoking in Japan and England. *J Epidemiol* 2021; **31**(10): 523–529.
 21. Lubben J, Blozik E, Gillmann G, Iliffe S, von Renteln Kruse W, Beck JC et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European Community-dwelling older adult populations. *Gerontologist* 2006; **46**(4): 503–513.
 22. Kuriyama S, Yaegashi N, Nagami F, Arai T, Kawaguchi Y, Osumi N et al. The Tohoku Medical Megabank project: design and mission. *J Epidemiol* 2016; **26**(9): 493–511.
 23. Hozawa A, Tanno K, Nakaya N, Nakamura T, Tsuchiya N, Hirata T et al. Study profile of the Tohoku Medical Megabank community-based cohort study. *J Epidemiol* 2021; **31**(1): 65–76.
 24. Yamada M, Motoike IN, Kojima K, Fuse N, Hozawa A, Kuriyama S et al. Genetic loci for lung function in Japanese adults with adjustment for exhaled nitric oxide levels as airway

- inflammation indicator. *Commun Biol* 2021; **4**(1): 1288.
25. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 2015; **4**: 7.
26. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013; **10**(1): 5–6.
27. Tadaka S, Katsuoka F, Ueki M, Kojima K, Makino S, Saito S et al. 3.5KJPNv2: an allele frequency panel of 3552 Japanese individuals including the X chromosome. *Hum Genome Var* 2019; **6**: 28.
28. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM et al. A global reference for human genetic variation. *Nature* 2015; **526**(7571): 68–74.
29. Bycroft C, Freeman C, Petkova D, et al. The UK biobank resource with deep phenotyping and genomic data. *Nature* 2018; **562**(7726): 203–209.
30. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLOS Genet* 2009; **5**(6): e1000529.
31. Kurimoto A, Awata S, Ohkubo T, et al. Reliability and validity of the Japanese version of the abbreviated Lubben social network scale. *Nihon Ronen Igakkai Zasshi* 2011; **48**, 149–157
32. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**(1): 76–82.

33. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome wide association scans. *Bioinformatics* 2010; **26**(17): 2190–2191.
34. Choi SW, O'Reilly PF. PRSice-2: polygenic Risk Score software for biobank-scale data. *GigaScience* 2019; **8**(7): giz082.
35. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; **38**(16): e164.
36. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J; Schizophrenia Working Group of the Psychiatric Genomics Consortium et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47**(3): 291–295.
37. Arden KC, Viars CS, Fu K, Rozen R. Localization of short/branched chain acyl-CoA dehydrogenase (ACADSB) to human chromosome 10. *Genomics* 1995; **25**(3): 743–745.
38. Wang W, Lo P, Frasch M, Lufkin T. Hmx: an evolutionary conserved homeobox gene family expressed in the developing nervous system in mice and *Drosophila*. *Mech Dev* 2000; **99**(1-2): 123–137.
39. Wang W, Grimmer JF, Van De Water TR, Lufkin T. Hmx2 and Hmx3 homeobox genes direct development of the murine inner ear and hypothalamus and can be functionally replaced by *Drosophila* Hmx. *Dev Cell* 2004; **7**(3): 439–453.
40. Choi Y, Nam J, Whitcomb DJ, Song YS, Kim D, Jeo S et al. SALM5 trans-synaptically interacts with LAR-RPTPs in a splicing-dependent manner to regulate synapse

- development. *Sci Rep* 2016; **6**: 26676.
41. Goto-Ito S, Yamagata A, Sato Y, Uemura T, Shiroshima T, Maeda A et al. Structural basis of trans-synaptic interactions between PTPdelta and SALMs for inducing synapse formation. *Nat Commun* 2018; **9**(1): 269.
42. Gibson G, Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M et al. Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLOS Genet* 2014; **10**(9): e1004606.
43. Wigmore EM, Hafferty JD, Hall LS, Howard DM, Clarke T-K, Fabbri C et al. Genome-wide association study of antidepressant treatment resistance in a population-based cohort using health service prescription data and meta-analysis with GENDEP. *Pharmacogenomics J* 2020; **20**(2): 329–341.
44. Porta F, Chiesa N, Martinelli D, Spada M. Clinical, biochemical, and molecular spectrum of short/branched-chain acyl-CoA dehydrogenase deficiency: two new cases and review of literature. *J Pediatr Endocrinol Metab* 2019; **32**(2): 101–108.
45. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP et al. Multivariate genome-wide analyses of the well-being spectrum. *Nat Genet* 2019; **51**(3): 445–451.
46. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* 2018; **50**(7): 920–927.

47. Okbay A, Wu Y, Wang N, Jayashankar H, Bennett M, Nehzati SM et al. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nat Genet* 2022; **54**(4): 437–449.
48. Levey DF, Stein MB, Wendt FR, Pathak GA, Zhou H, Aslan M et al. Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat Neurosci* 2021; **24**(7): 954–963.
49. Wendt FR, Pathak GA, Deak JD, De Angelis F, Koller D, Cabrera-Mendoza B et al. Using phenotype risk scores to enhance gene discovery for generalized anxiety disorder and posttraumatic stress disorder. *Mol Psychiatry* 2022; **27**(4): 2206–2215.
50. Lybaek H, Robson M, de Leeuw N, Hehir-Kwa JY, Jeffries A, Haukanes BI et al. LRFN5 locus structure is associated with autism and influenced by the sex of the individual and locus conversions. *Autism Res* 2022; **15**(3): 421–433.
51. Baselmans BML, Yengo L, van Rheenen W, Wray NR. Risk in relatives, heritability, SNP-based heritability, and genetic correlations in psychiatric disorders: a review. *Biol Psychiatry* 2021; **89**(1): 11–19.
52. Tsuchiya N, Nakaya N, Nakamura T, Narita A, Kogure M, Aida J et al. Impact of social capital on psychological distress and interaction with house destruction and displacement after the Great East Japan Earthquake of 2011. *Psychiatry Clin Neurosci* 2017; **71**(1): 52–60.
53. Murakami K, Ishikuro M, Obara T, Ueno F, Noda A, Onuma T et al. Traumatic experiences

- of the Great East Japan Earthquake and postpartum depressive symptoms: the Tohoku Medical Megabank Project Birth and Three-Generation Cohort Study. *J Affect Disord* 2023; **320**: 461–467.
54. Ohseto H, Ishikuro M, Obara T, Murakami K, Onuma T, Noda A et al. Dietary calcium intake was related to the onset of pre-eclampsia: the TMM BirThree Cohort Study. *J Clin Hypertens (Greenwich)* 2023; **25**(1): 61–70.
55. Wendt FR, Pathak GA, Tylee DS, Goswami A, Polimanti R. Heterogeneity and polygenicity in psychiatric disorders: A genome-wide perspective. *Chronic Stress (Thousand Oaks)* 2020; **4**: 2470547020924844.
56. Abdellaoui A, Verweij KJH. Dissecting polygenic signals from genome-wide association studies on human behaviour. *Nat Hum Behav* 2021; **5**(6): 686–694.

Figure legends

Figure 1. GWAS meta-analyses for social isolation status

The Manhattan and QQ plots for social isolation status are presented. Three social isolation statuses—the total scale, family subscale, and friend subscale—were measured using the LSNS-6. One locus in the total scale GWAS and one in the friend subscale GWAS were significant. GWAS: Genome-wide association study; QQ plot: Quantile-quantile plot; LSNS-6: Lubben Social Network Scale.

Figure 2. The relationship between genotype and social isolation status

For significant or top-hit variants, the relationships between genotype and social isolation status were assessed for each cohort by logistic regression analyses adjusted for age, sex, and 10 genetic principal components, and then meta-analyzed.

For rs1778366, because the alternative allele frequency was low, the Ref/Alt and Alt/Alt genotypes were combined and presented together to ensure statistical stability in visualization. When analyzed separately, results (n/N %, OR [95% CI]) for Ref/Alt and Alt/Alt vs Ref/Ref were as follows—Total: 494/1,660 (29.8%), 1.23 (1.10–1.37); 6/16 (37.5%), 2.46 (0.82–7.36); Family: 251/1,660 (15.1%), 1.02 (0.89–1.17); 2/16 (12.5%), 1.30 (0.25–6.83); Friend: 673/1,660 (40.5%), 1.33 (1.20–1.46); 9/16 (56.2%), 2.45 (0.90–6.64).

Ref: Reference allele; Alt: Alternative allele; OR: Odds ratio; CI: Confidence interval.

Table 1. Characteristics of the four cohorts in the TMM CommCohort Study

	ToMMo		IMM	
	Type 1 survey	Type 2 survey	Type 1 survey	Type 2 survey
n	28 835	9 510	17 608	7 544
Female, %	17 903 (62.1)	7 249 (76.2)	10 947 (62.2)	5 027 (66.6)
Age at recruitment, years	59.62 ± 11.55	57.71 ± 13.36	61.32 ± 10.61	56.38 ± 13.26
Social isolation status				
Total scale, %	7 339 (25.5)	2 007 (21.1)	4 940 (28.1)	2 067 (27.4)
Family subscale, %	4 212 (14.6)	1 396 (14.7)	2 678 (15.2)	1 132 (15.0)
Friend subscale, %	9 754 (33.8)	2 619 (27.5)	6 514 (37.0)	2 751 (36.5)

TMM CommCohort Study: Tohoku Medical Megabank Community-Based Cohort Study;

ToMMo: Tohoku University Tohoku Medical Megabank Organization; IMM: Iwate Medical University; Iwate Tohoku Medical Megabank Organization.

Data are shown as the mean ± standard deviation for continuous variables and n (%) for categorical variables.

Social isolation was evaluated using the Lubben Social Network Scale. The total score ranged from 0 to 30. Scores < 12 were considered socially isolated. The family and friend subscales were calculated by summing up each of the three items, and the total ranged from 0 to 15. Scores < 6 were considered socially isolated.

Table 2. Significant variants in GWAS meta-analyses of social isolation status

Outcome	dbSNP ID	Locus	Position	Ref	Alt	EAF	Function	Gene	P-value	Directions
Total scale	rs107369	10q26.13	124832178	C	T	T	intergenic	ACADS	4.9*10 ⁻⁸	++++
								(distance = 14347)		
Family subscale	rs780848	1q32.1	201919817	AAA	C	A	intergenic	HMX3	4.6*10 ⁻⁷	----
								(distance = 4141)		
Friend subscale	rs177836	14q21.1	41766603	C	T	T	intergenic	TIMM17	1.1*10 ⁻⁸	++++
								(distance = 315)		
								LINC02		

= 156

352)

LRFN5

(distance

= 309

456)

GWAS, Genome-wide association study; Ref: Reference allele; Alt: Alternative allele; EA: Effect allele.

Social isolation was defined using the total scale, family subscale, and friend subscale of the Lubben Social Network Scale. Variants were annotated by ANNOVAR. The direction column indicates whether the loci were positively or negatively associated in each of the four GWASs prior to the meta-analysis. Only significant or top-hit loci were reported.

Table 3. GWAS results of the four cohorts in the TMM CommCohort Study

	dbSNP ID	EA freq.	Beta	SE	P-value
ToMMo	Type 1				
	rs10736933	0.680	0.07	0.02	5.4×10^{-4}
	rs780848113	0.486	-0.07	0.02	2.4×10^{-3}
	rs1778366	0.014	0.22	0.08	4.5×10^{-3}
	Type 2				
	rs10736933	0.679	0.12	0.04	2.7×10^{-3}
	rs780848113	0.492	-0.08	0.04	5.7×10^{-2}
	rs1778366	0.014	0.54	0.14	1.8×10^{-4}
IMM	Type 1				
	rs10736933	0.670	0.08	0.03	3.0×10^{-3}
	rs780848113	0.473	-0.12	0.03	7.5×10^{-5}
	rs1778366	0.014	0.32	0.10	1.0×10^{-3}
	Type 2				
	rs10736933	0.673	0.05	0.04	2.5×10^{-1}
	rs780848113	0.483	-0.02	0.05	6.3×10^{-1}
	rs1778366	0.013	0.28	0.15	7.0×10^{-2}

TMM CommCohort Study: Tohoku Medical Megabank Community-Based Cohort Study;

ToMMo: Tohoku University Tohoku Medical Megabank Organization; IMM: Iwate Medical

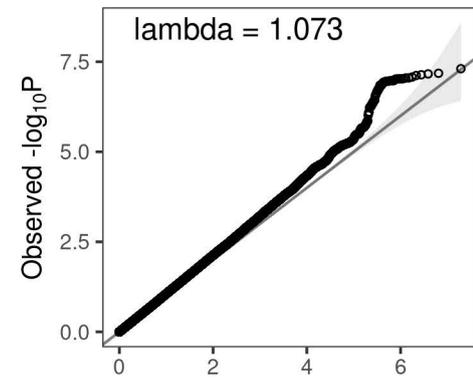
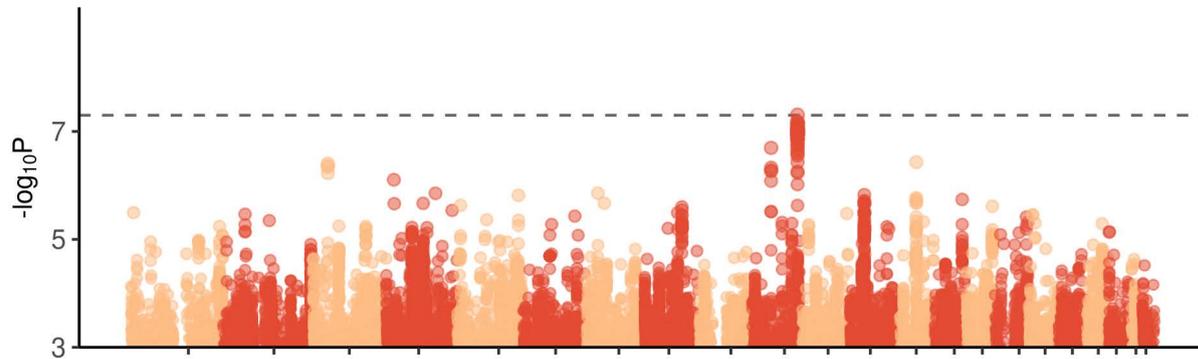
University Iwate Tohoku Medical Megabank Organization; EA Freq: Effect allele frequency; SE:

Standard error.

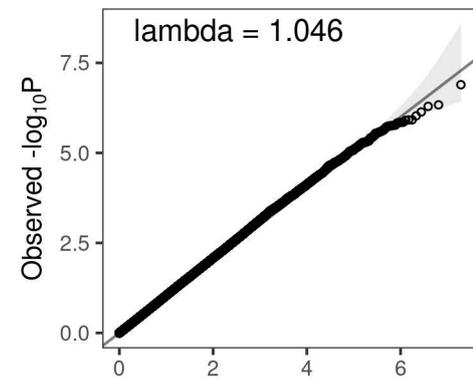
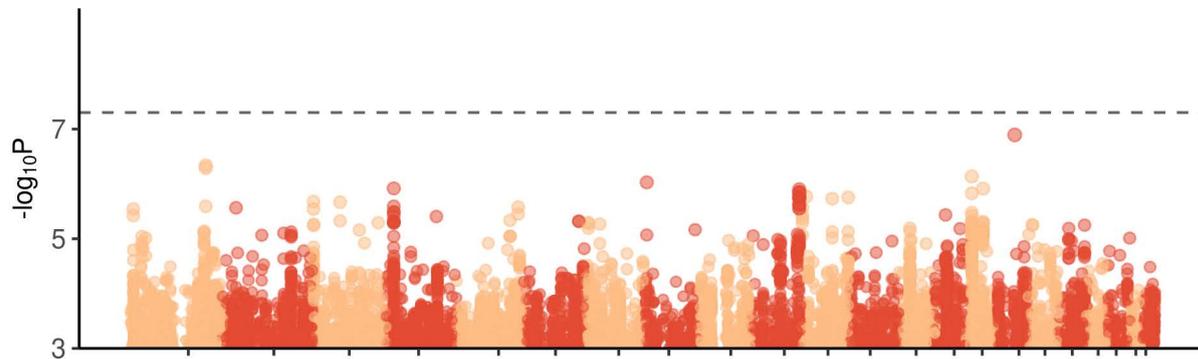
Only significant or top-hit loci are shown in this table.

ARTICLE IN PRESS

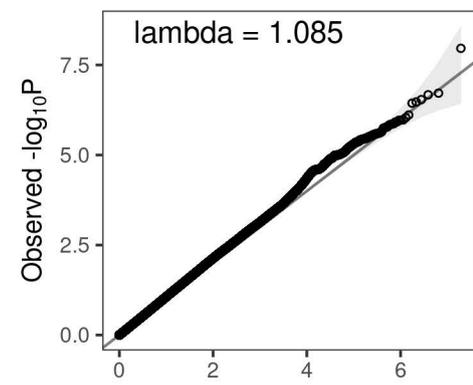
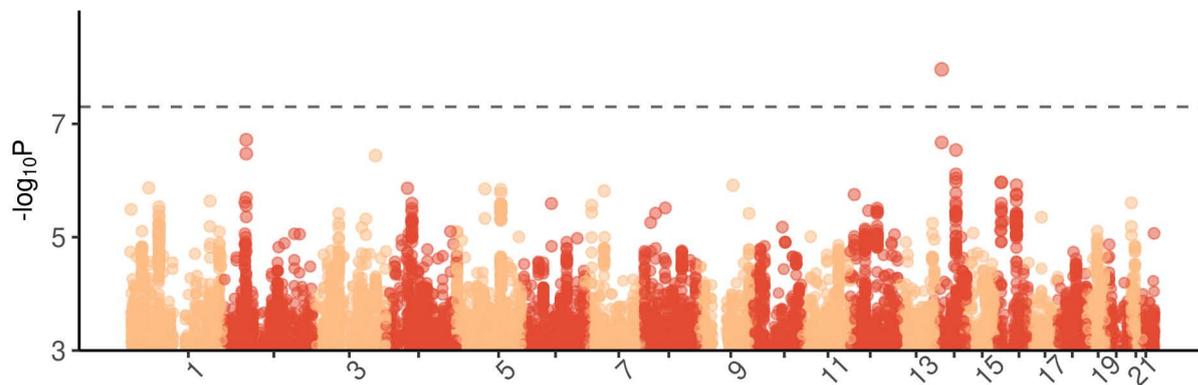
LSNS-6: Total scale



LSNS-6: Family subscale



LSNS-6: Friend subscale



Chromosome

Expected $-\log_{10}P$

