

## Genome-wide association studies identify polygenic effects for completed suicide in the Japanese population

**Cite this article as:** Ikuo Otsuka, Masato Akiyama, Osamu Shirakawa, Satoshi Okazaki, Yukihide Momozawa, Yoichiro Kamatani, Takeshi Izumi, Shusuke Numata, Motonori Takahashi, Shuken Boku, Ichiro Sora, Ken Yamamoto, Yasuhiro Ueno, Tatsushi Toda, Michiaki Kubo and Akitoyo Hishimoto, Genome-wide association studies identify polygenic effects for completed suicide in the Japanese population, *Neuropsychopharmacology* doi:10.1038/s41386-019-0506-5

This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Terms of use and reuse: academic research for non-commercial purposes, see here for full terms. <https://www.nature.com/authors/policies/license.html#AAMtermsV1>

Author accepted manuscript

## Genome-wide association studies identify polygenic effects for completed suicide in the Japanese population

Ikuo Otsuka<sup>1,2</sup>, MD, PhD; Masato Akiyama<sup>2,3</sup>, MD, PhD; Osamu Shirakawa<sup>4</sup>, MD, PhD; Satoshi Okazaki<sup>1</sup>, MD, PhD; Yukihide Momozawa<sup>5</sup>, DVM, PhD; Yoichiro Kamatani<sup>2,6</sup>, MD, PhD; Takeshi Izumi<sup>7</sup>, MD, PhD; Shusuke Numata<sup>8</sup>, MD, PhD; Motonori Takahashi<sup>9</sup>, MD, PhD; Shuken Boku<sup>1</sup>, MD, PhD; Ichiro Sora<sup>1</sup>, MD, PhD; Ken Yamamoto<sup>10</sup>, MD, PhD; Yasuhiro Ueno<sup>9</sup>, MD, PhD; Tatsushi Toda<sup>11</sup>, MD, PhD; Michiaki Kubo<sup>12</sup>, MD, PhD; Akitoyo Hishimoto<sup>1,\*</sup>, MD, PhD.

<sup>1</sup> Department of Psychiatry, Kobe University Graduate School of Medicine, Kobe, Japan.

<sup>2</sup> Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

<sup>3</sup> Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan.

<sup>4</sup> Department of Neuropsychiatry, Kindai University Faculty of Medicine, Osaka, Japan.

<sup>5</sup> Laboratory for Genotyping Development, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

<sup>6</sup> Laboratory of Complex Trait Genomics, Department of Computational biology and medical Sciences, Graduate school of Frontier Sciences, The University of Tokyo, Tokyo, Japan.

<sup>7</sup> Department of Pharmacology, Pharmaceutical Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan.

<sup>8</sup> Department of Psychiatry, Graduate School of Biomedical Sciences, Tokushima University, Tokushima, Japan.

<sup>9</sup> Division of Legal Medicine, Kobe University Graduate School of Medicine, Kobe, Japan.

<sup>10</sup> Department of Medical Biochemistry, Kurume University School of Medicine, Kurume, Japan.

<sup>11</sup> Department of Neurology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan.

<sup>12</sup> RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.

**\* Corresponding author:**

Akitoyo Hishimoto, MD, PhD

Associate Professor, Department of Psychiatry, Kobe University Graduate School of Medicine, 7-5-1, Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan

Tel: +81-78-382-6065, Fax: +81-78-382-6079, E-mail: [hishipon@med.kobe-u.ac.jp](mailto:hishipon@med.kobe-u.ac.jp)

**Running title:** GWAS of completed suicide

**ABSTRACT**

Suicide is a significant public health problem worldwide, and several Asian countries including Japan have relatively high suicide rates on a world scale. Twin, family, and adoption studies have suggested high heritability for suicide, but genetics lags behind due to difficulty in obtaining samples from individuals who died by suicide, especially in non-European populations. In this study, we carried out genome-wide association studies combining two independent datasets totaling 746 suicides and 14,049 non-suicide controls in the Japanese population. Though we identified no genome-wide significant single nucleotide polymorphisms (SNPs), we demonstrated significant SNP-based heritability (35–48%;  $P < 0.001$ ) for completed suicide by genomic restricted maximum likelihood analysis and a shared genetic risk between two datasets ( $P_{\text{best}} = 2.7 \times 10^{-13}$ ) by polygenic risk score analysis. This study is the first GWAS for suicidal behavior in an East Asian population, and our results provided the evidence of polygenic architecture underlying completed suicide.

(151 words)

## INTRODUCTION

Suicide is a significant public health problem that causes nearly one million deaths worldwide each year [1]. In addition, several Asian countries including Japan have relatively high suicide rates on a world scale [2,3].

There is evidence for high heritability in suicidal behavior, with an estimated  $h^2$  of 30–50% based on twin, family, and adoption studies in European populations [4]. To investigate the genetic components of suicidal behavior, large-scale genetic studies focusing on “completed suicide” should be crucial because other suicidal behaviors (suicide ideation or attempt) vary in terms of degree of lethality and suicidal intent, which may cause heterogeneity [5]. However, due to the difficulty in obtaining samples from individuals who died by suicide, few genome-wide association studies (GWASs) have been reported (only European populations) [6–9]. The largest case-control GWAS to date included 317 suicides, and revealed no significant findings [7]. In addition, no study has investigated single nucleotide polymorphism (SNP)-based heritability or polygenic effects for completed suicide. Consequently, genetic insights into suicide lag behind those of other mental problems, and no useful genetic biomarker of suicide risk has been found.

On the other hand, previous studies have associated different characteristics with suicidal behavior in different age groups; for instance, more impulsive, aggressive, and novelty seeking personality traits in the young, while heightened harm avoidance and higher comorbidity rate of depression and physical diseases were associated with increasing age of

suicidal behavior [10–13]. Family studies have indicated strong familial transmission for early onset suicidal behavior [14–16]. These studies encourage research focusing on the biological factors associated with age at suicide.

Here, we first conducted genome-wide association analyses using two independent datasets totaling ~746 suicides and 14,049 non-suicide controls in the Japanese population in order to identify genetic variants affecting suicide risk (case-control study). We then utilized these genome-wide SNP data to investigate SNP-based heritability and polygenic effects for completed suicide. Additionally, we also investigated individual variants and polygenicity affecting age at suicide (targeting only suicidal cohorts).

## **MATERIALS AND METHODS**

### **Subjects**

The entire study design and procedures were performed in accordance with the Declaration of Helsinki. This study was approved by the Ethics Committee for Genetic Studies of Kobe University and RIKEN.

#### ***Individuals who died by suicide***

Autopsies on suicide victims were conducted at the Division of Legal Medicine in the Department of Community Medicine and Social Health Science at the Kobe University Graduate School of Medicine. The verdict of “completed suicide” was made through discussion with the Medical Examiner’s Office of the Hyogo Prefecture and the Division of

Legal Medicine in the Kobe University Graduate School of Medicine [17]. In order to gather background information on completed suicides, psychological autopsy through their medical records and bereaved family interviews were conducted by professional staff from the Medical Examiner's Office of the Hyogo Prefecture and the Division of Legal Medicine in the Kobe University, where available.

### *Non-suicide controls*

As non-suicide controls, we used genome-wide genotype data from subjects in the Biobank Japan project who had been genotyped as case subjects for non-psychiatric disorders and from healthy volunteers of the Osaka-Midosuji Rotary Club and the Pharma SNP consortium. The controls were not psychiatrically evaluated [18–20].

### **Genotyping, QC and imputation**

We genotyped 434 individuals who died by suicide between June 1996 and July 2012 and 405 individuals who died by suicide between August 2012 and February 2017 using Illumina HumanOmniExpress and HumanOmniExpressExome BeadChips for the 1st and 2nd set, respectively. We obtained control data genotyped with the same arrays (N = 7,993 and 7,136 for the 1st and 2nd set, respectively) (**Table S1**).

We performed quality control (QC) using PLINK 1.9 [21]. Firstly, we excluded SNPs with a call rate  $< 0.98$  and minor allele frequency (MAF)  $< 0.01$ , and those with  $P < 1.0 \times 10^{-6}$  for Hardy-Weinberg equilibrium (HWE) in controls. Related individuals were excluded

(PI\_HAT  $\geq$  0.175). We then performed principal component analysis (PCA), and excluded samples outside the Japanese main islands cluster [22,23]. The results of PCA are shown in **Figure S1**. The final datasets included 386 suicides and 7,458 controls as the 1st set, and 360 suicides and 6,591 controls as the 2nd set. After estimating haplotypes using SHAPEIT2 (v2.r778) [24], we performed genotype imputation by Minimac3 (1.0.13) [25] using ALL samples in the 1000 Genomes Project phase 3v5 [26] as a reference.

To investigate the X chromosome, we called genotypes using GenomeStudio. First, we generated genoplots using only female samples. After that, we added male samples to genoplots and called genotypes. Genotypes called as heterozygotes were treated as missing in male samples. For the genotyping QC, we excluded SNPs with MAF  $<$  0.01 and SNP call rate  $<$  0.98 in either male or female samples. We also excluded SNPs with  $P$  value for HWE  $<$   $1.0 \times 10^{-6}$  in female samples. Haplotype phasing and imputation were performed separately for males and females. Allelic dosages were imputed from 0 to 2 in male samples under the assumption of full dosage compensation. The pseudo-autosomal region was excluded from the reference before imputation.

After imputation of autosomal and X chromosomes, we included only SNPs with high imputation quality ( $r^2 \geq 0.7$ ) and MAF  $\geq 0.01$  (**Figure S2**).

### **Association analysis and meta-analysis**



For all GWASs identifying genetic variants affecting suicide risk (case-control study; 746 suicides and 14,049 controls) and age at suicide (only suicidal cohorts; 719 suicides), single variant association tests were performed for common variants ( $MAF \geq 0.01$ ) using logistic regression and linear regression based on Wald test implemented in Rvtests software [27] with a correction for the top 10 PCs as covariates. Meta-analysis was performed with METAL software [28] using a fixed effects model with inverse-variance weighted approach. The significance level was set at  $P < 2.5 \times 10^{-8}$  due to correction for multiple comparison for two GWASs (case-control GWAS and GWAS for age at suicide).  $P$  for heterogeneity between two analyses (1st set and 2nd set) were calculated by Cochran's Q test. Regional association plots were generated using LocusZoom [29]. A QQ plot for each GWAS is shown in **Figure S3**.

### **Evaluation of previously implicated SNPs in prior GWAS of suicidal behavior**

Among variants that have previously been associated with suicidal behavior (suicide ideation, suicide attempt and suicide completion) in the published literature [7,9,30–36], we identified 26 variants with  $P < 1.0 \times 10^{-6}$  in previous GWASs and  $MAF > 0.01$  in JPT population (the 1000 Genomes Project phase 3). We looked up the association of each of these SNPs with the results from our case-control GWAS. Only the top SNP in the same region in each reference was selected. For a candidate SNP that was not directly genotyped or imputed in our GWAS, we identified a proxy SNP ( $r^2 > 0.8$  in East Asian samples of the 1000 Genomes Project phase

3) whenever possible. For SNPs that reached  $P < 0.05$ , we determined whether the direction of the association was consistent between the prior and current studies.

### **Estimation of the proportion of the variance in completed suicide explained by the genotyped SNPs (SNP-based heritability)**

To assess SNP-based heritability ( $h^2_{\text{SNP}}$ ), we used GCTA [37] to generate genetic relatedness matrices (GRMs) among 506,645 SNPs that passed QC in both genotyped datasets, and then performed genomic restricted maximum likelihood (GREML) analysis. This assumed prevalence rates of 0.1 and 0.5% for completed suicide, considering the reported incidence of completed suicide in Japan (Ministry of Health, Labour and Welfare of Japan) [3] and the estimates from previous papers [38–40]. We strictly controlled the cryptic relatedness of the analyzed samples using the `-grm-cutoff` option (threshold of 0.05) implemented by GCTA, including 385 cases and 7,409 controls for the 1st set, and 357 cases and 6,560 controls for the 2nd set, respectively. We then estimated SNP-based heritability using `reml` function implemented by GCTA with top 10 PCs as covariates.

### **PRS analysis**

Polygenic risk score (PRS) analyses were performed using PRSice v1.23 [41]. The  $P$  threshold ( $P_t$ ) for selecting ‘risk’ SNPs was sequentially set at 0.1, 0.2, 0.3, 0.4, and 0.5 without SNPs in the major histocompatibility complex (MHC) region. We then performed LD clumping (used

by the default setting of the software) to select the eligible SNPs for PRS. To calculate the PRS, we analyzed two discovery/target sets, using the 1st set (386 suicides and 7,458 controls) as the discovery set and the 2nd set (360 suicides and 6,591 controls) as the target set, and vice versa. We included the top 10 PCs derived from each genotyped dataset as covariates, respectively. The variance explained for the PRS was estimated based on Nagelkerke's  $R^2$  from a logistic regression model. For the analysis of polygenic effects on age at suicide (1st set, 366 suicides; 2nd set, 353 suicides), we also applied the same procedure as the above to the case-control GWAS datasets.

### Pathway enrichment analysis

Using the results of a meta-analysis of the case-control GWAS datasets, we ran PASCAL [42] for gene-based enrichment analysis using 1,077 gene sets including KEGG [43], REACTOME [44], and BIOCARTA ([http://cgap.nci.nih.gov/Pathways/BioCarta\\_Pathways](http://cgap.nci.nih.gov/Pathways/BioCarta_Pathways)). The significance level was set at  $\chi^2 P < 4.6 \times 10^{-5}$  after Bonferroni correction for 1,077 tests.

## RESULTS

### GWAS for risk of completed suicide

Sample characteristics are shown in **Table S1**. We performed a meta-analysis of the case-control GWAS datasets for both suicides and non-suicide controls (totaling 746 cases and 14,049 controls) using 8,625,325 SNPs ( $\lambda_{GC} = 1.07$ ; **Figure S3a**). This analysis

identified no genome-wide significant SNPs (**Figure S4**), although some loci, including SNPs on *GRM1* and *CTPS2*, were suggestive with  $P_{\text{meta}} < 1.0 \times 10^{-6}$  (**Table S2**). In addition, we looked up the 26 variants showing  $P < 1.0 \times 10^{-6}$  in prior European GWASs for suicidal behavior [7,9,30–36]. Among these, rs7989250, reported in the recent UK Biobank GWAS for ordinal suicidality, reached the threshold for our replication analysis ( $P < 0.05$ ) with the same direction of allelic effect (**Table S3**). We ran PASCAL [42] using our meta-GWAS to estimate the enrichments in 1,077 gene sets, but no pathways were enriched at the Bonferroni level of significance (Results for all tested 1,077 pathways are listed in an Excel spreadsheet that is included in the Supplementary Material).

### **GREML analysis reveals high SNP-based heritability for completed suicide**

To investigate SNP-based heritability ( $h^2_{\text{SNP}}$ ) for completed suicide, we used GCTA [37] to GRMs using 506,645 directly genotyped SNPs, and then estimated SNP-based heritability with GREML. These analyses revealed significant SNP-based heritability in both datasets ( $P < 0.001$ ; **Figure 1**). By combining the results, we estimated that 35.9 and 48.3% of phenotypic variance in the two datasets could be explained by SNPs with a prevalence of 0.1 and 0.5%, respectively; this implies the underlying polygenic architecture composed by numerous ‘risk’ SNPs for completed suicide.

### **PRS analysis reveals polygenic effects for completed suicide**

To investigate whether two independent datasets share genetic components conferring the risk of completed suicide, we performed PRS analyses by PRSice (v1.23) [41], setting the 1st set as the discovery set and the 2nd set as the target set, and vice versa. The analyses demonstrated significant polygenic effects on completed suicide ( $P < 0.001$ ) when the SNPs were stratified by  $P$  values, obtaining the most significant  $P$  values ( $P_{\text{best}} = 2.7 \times 10^{-13}$ ) with an inclusion threshold of  $P_t = 0.4$  in each set, with explained 1.3%–2.4% of the variance (**Figure 2** and **Table S4**). These results strongly suggest shared genetic components for completed suicide between the two datasets, providing additional evidence of polygenic architecture underlying completed suicide.

### **GWAS for age at suicide**

Age-divided sample characteristics focusing on psychological background information, comorbid severe physical diseases, and distribution of age at suicide in both the 1st and 2nd set of our suicidal cohorts are shown in **Table S5** and **Figure S5**. We performed GWAS for the age at completed suicide in each dataset (only suicides with accurate age information available;  $N = 366$  and  $353$  for the 1st and 2nd set, respectively), and integrated these results by fixed-effects meta-analysis. The meta-GWAS included 8,810,873 SNPs, and the genomic inflation factor suggested low possibility of bias from population stratification and cryptic relatedness ( $\lambda_{GC} = 1.02$ ; a QQ plot is shown in **Figure S3b**). We identified a novel suggestive locus on 7q11.23 in GTF2I repeat domain containing 1 (*GTF2IRD1*) (top SNP

rs73135307 G>C;  $P = 3.3 \times 10^{-8}$ , beta = -12.3 years; **Figure 3, Table 1, Table S6 and Figure S6**) without significant heterogeneity between the two data sets ( $P_{\text{het}} = 0.20$ ). This means that the effect allele of rs73135307 (G allele) could lead to 12.3 years younger of age at suicide in our suicidal cohort compared to C allele. G allele frequency variation of rs73135307 in divided groups based on age at suicide in the 1st and 2nd set are shown in **Figure S7**. According to HaploReg (v4.1) [45], rs73135307 and the four variants in LD ( $r^2 > 0.6$  in East Asian samples of the 1000 Genomes Project [26]) with rs73135307 affect various histone modifications (H3K4me1, H3K4me3, H3K27ac, and H3K9ac) in various brain regions, including the hippocampus and dorsolateral prefrontal cortex. However, neither nonsynonymous SNP nor significant eQTL was found. Our PRS analysis showed no significant polygenic effects on age at suicide (**Table S7**).

## DISCUSSION

Through the GWAS including 746 suicides in the Japanese population, we demonstrated SNP-based heritability (35–48%) and polygenic effects, indicating polygenic inheritance of completed suicide. We also showed that the *GTF2IRD1* locus is suggestively associated with age at completed suicide.

Using the two case-control GWAS datasets, we estimated SNP-based heritability with GREML and conducted PRS analyses. Our GREML analysis revealed that completed suicide has significant SNP-based heritability, with estimates of 35–48% in both individual

GWAS sets. We estimated SNP-based heritability using the GREML method with two different prevalence rates (0.1% and 0.5%) assumed from the reported incidence of completed suicide in Japan [3] and the previous studies [38–40]. These estimates are qualified by the difficulty of estimating the lifetime prevalence of completed suicide, compared to those of “diseases”, presumably due to the influence of historical and cultural contexts on the epidemiology of suicide, along with other national/regional differences [46]. Moreover, our PRS analyses demonstrated significant polygenic effects on completed suicide, indicating that two independent cohorts of suicides shared genetic components conferring suicide risk. Consequently, to our knowledge, we provide the first evidence of SNP-based heritability in completed suicide derived from case-control GWAS dataset, other than from epidemiological data such as twin, family, and adoption studies.

Due to the lack of well-designed epidemiological studies and GWAS for suicidal behavior in Asia, this is the first study demonstrating SNP-based heritability of suicide in an Asian population. This is a major strength of the current study, since the dearth of GWASs in non-European populations compared with the abundance of European-descent studies is causing the disparity and poor generalizability of genetic studies across populations [47]. The SNP-based heritability of completed suicide found in our analysis dramatically exceeded the previous estimates from past GWASs of suicide attempt in European populations, and further comparable to twin heritability estimated by prior studies of European populations despite SNP-based heritability estimates by GWASs usually explaining roughly half of those based on

twin modeling [48]. In order to elucidate the reasons for the different levels of SNP-based heritability estimated by previous GWASs for suicide attempt and our GWAS, we should focus on the following points regarding SNP-based heritability differences; 1) between suicide attempt and suicide completion, 2) between Europeans and Asians. On the other hand, significant high SNP-based heritability for completed suicide shown here might be overestimated to some extent due to residual population stratification (e.g., regional differences in Japan) or the genetic effects of comorbid psychiatric disorders. Indeed, genetic influences on suicidal behavior seem to be confounded by genetic vulnerability for psychiatric disorders; for instance, a previous twin study reported lower heritability (approximately 17%) for suicidal behavior after controlling for the inheritance of psychiatric disorders [49]. Publicly available GWAS data focusing on psychiatric problems in East Asian population are limited, which makes difficult to perform replication of GWAS for suicidal behavior and in-depth analyses (e.g., suicide GWAS divided into cases with and without psychiatric disorders, or genetic correlation analysis with various psychiatric disorders/conditions) using GWAS datasets from the same Asian population. Since we have made our summary statistics of GWAS publicly available, our data will enable the above analyses in the future.

Our meta-analysis of two GWASs of age at completed suicide identified a suggestive locus, intronic SNPs in the *GTF2IRD1* gene (top hit rs73135307,  $P = 3.3 \times 10^{-8}$ ) at 7q11.23. This SNP and variants in LD with the lead SNP are associated with various histone modifications across brain regions related to suicide pathophysiology [5], although none of



those SNPs were nonsynonymous or overlapped with eQTLs. *GTF2IRD1* has been reported as one of the promising genes for Williams Syndrome (WS), particularly responsible for neurodevelopmental abnormalities [50], which is known as a risk factor for suicidal behavior [51]. While the typical social phenotype of WS is characterized by optimistic personality, there is also evidence of diametric characteristics in individuals with WS such as heightened anxiety, social relationship difficulties, and higher prevalence of autism spectrum disorder (ASD) [52–54]. The associations of *GTF2IRD1* with anxiety and social impairment in WS have been previously reported [53,55]. In addition, various copy number variations in 7q11.23 region have been strikingly implicated in the genetic etiology of schizophrenia and ASD [56,57], which are associated with high risk of youth suicide [58,59]. Focusing on this chromosomal region by more detailed genome sequencing is necessary to yield further insight into the genetic feature of suicide in young people. Indeed, the aberrant personality traits, lower comorbidity rate of depression and physical diseases, and familial transmission previously linked with younger age at suicidal behavior [10–16] indicated the possibility that future genetic studies for suicide in young people may identify genetic clues.

It must be emphasized that our findings should be interpreted in the context of several other limitations. First, although our sample size was one of the largest ever for research on completed suicide, larger sample sizes would be preferable in order to draw robust conclusions regarding SNP-based heritability and polygenicity for completed suicide, and to detect reliable genetic markers. In particular, although we reported that the *GTF2IRD1* locus is

suggestively associated with age at completed suicide, GWAS for age at suicide here contained too small sample size ( $N = 719$ ) and replication studies are indeed required. Second, our study cohort was restricted to only the Japanese population; thus, findings from the present study might not be generalizable to other populations. Third, the subjects we used non-suicide controls who had not been psychiatrically screened. In addition, most of the controls had various non-psychiatric disorders [18–20]. However, this approach has been already applied in the previous GWAS [20] of which results was replicable in other recent GWAS [60], supporting the reliability of our results. Fourth, the biological interpretation of the significant polygenic effects for completed suicide and the variants reported here remains largely unknown due to lack of functional evidence yielded by our pathway analysis and database searches, respectively.

In conclusion, we provided the first evidence of SNP-based heritability and polygenic effects in completed suicide, and polygenic effects in age at suicide.

#### **FUNDING AND DISCLOSURE**

This work was supported in part by JSPS KAKENHI Grant Number JP17H04249, BBJ, and The Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan. The authors declare no conflict of interest.

#### **ACKNOWLEDGMENTS**

We would like to acknowledge the staff of the Biobank Japan (BBJ) project for collecting samples and clinical information. We are grateful to Dr Tatsuhiro So, Dr Woraphat Ratta-Apha, Yasuko Nagashima, and the staff of the RIKEN Center for Integrative Medical Sciences for genotyping and data management.

#### **SUPPLEMENTARY INFORMATION**

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (<http://www.nature.com/npp>).

Author accepted manuscript

## REFERENCES

1. World Health Organization. Suicide Fact Sheet. 2017; <http://www.who.int/mediacentre/factsheets/fs398/en/>. Accessed August 29, 2017.
2. Värnik P. Suicide in the world. *Int J Environ Res Public Health*. 2012; **9**: 760–71.
3. Statistics and Information Department, Minister's Secretariat, Ministry of Health, Labour and Welfare. *Vital Statistics of Japan*. 2017.
4. Brent DA, Turecki G. Suicide and suicidal behavior. *Lancet*. 2017; **387**: 1227–39.
5. Oquendo MA, Sullivan GM, Sudol K, Baca-Garcia E, Stanley BH, Sublette E, et al. Toward a biosignature for suicide. *Am J Psychiatry*. 2014; **171**: 1259–77.
6. Coon H, Darlington TM, DiBlasi E, Callor WB, Ferris E, Fraser A, et al. Genome-wide significant regions in 43 Utah high-risk families implicate multiple genes involved in risk for completed suicide. *Mol Psychiatry*. 2018 [Epub ahead of print].
7. Galfalvy H, Haghghi F, Hodgkinson C, Goldman D, Oquendo MA, Burke A, et al. A genome-wide association study of suicidal behavior. *Am J Med Genet B Neuropsychiatr Genet*. 2015; **168**: 557–63.
8. Mirkovic B, Laurent C, Podlipski MA, Frebourg T, Cohen D, Gerardin P. Genetic Association Studies of Suicidal Behavior: A Review of the Past 10 Years, Progress, Limitations, and Future Directions. *Front Psychiatry*. 2016; **7**: 158.
9. Strawbridge RJ, Ward J, Ferguson A, Graham N, Shaw RJ, Cullen B, et al. Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation

- with psychiatric disorders and polygenic association with completed suicide. *EBioMedicine*. 2019; pii: S2352-3964: 30077–5.
10. Bozzay ML, Liu RT, Kleiman EM. Gender and age differences in suicide mortality in the context of violent death: findings from a multi-state population-based surveillance system. *Compr Psychiatry*. 2014; **55**: 1077–84.
  11. Fässberg MM, Cheung G, Canetto SS, Erlangsen A, Lapierre S, Lindner R, et al. A systematic review of physical illness, functional disability, and suicidal behaviour among older adults. *Aging Ment Health*. 2016; **20**: 166–94.
  12. McGirr A, Renaud J, Bureau A, Seguin M, Lesage A, Turecki G. Impulsive-aggressive behaviors and completed suicide across the life cycle: a predisposition for younger age of suicide. *Psychol Med*. 2008; **38**: 407–17.
  13. Rich CL, Young D, Fowler RC. San Diego suicide study. I. Young vs old subjects. *Arch Gen Psychiatry*. 1986; **43**: 577–82.
  14. Brent DA, Melhem NM, Oquendo M, Burke A, Birmaher B, Stanley B, et al. Familial pathways to early-onset suicide attempt: a 5.6-year prospective study. *JAMA Psychiatry*. 2015; **72**: 160–8.
  15. Brent DA, Oquendo M, Birmaher B, Greenhill L, Kolko D, Stanley B, et al. Familial pathways to early-onset suicide attempt: risk for suicidal behavior in offspring of mood-disordered suicide attempters. *Arch Gen Psychiatry*. 2002; **59**: 801–7.

16. Melhem N, Brent DA, Ziegler M, Iyengar S, Kolko D, Oquendo M, et al. Familial pathways to early-onset suicidal behavior: familial and individual antecedents of suicidal behavior. *Am J Psychiatry*. 2007; **164**: 1364–70.
17. Otsuka I, Izumi T, Boku S, Kimura A, Zhang Y, Mouri K, et al. Aberrant telomere length and mitochondrial DNA copy number in suicide completers. *Sci Rep*. 2017; **7**: 3176.
18. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: study design and profile. *J Epidemiol*. 2017; **27**: S2–8.
19. Hirata M, Kamatani Y, Nagai A, Kiyohara Y, Ninomiya T, Tamakoshi A, et al. Cross-sectional analysis of BioBank Japan clinical data: a large cohort of 200,000 patients with 47 common diseases. *J Epidemiol*. 2017; **27**: S9–21.
20. Ikeda M, Takahashi A, Kamatani Y, Okahisa Y, Kunugi H, Mori N, et al. A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Mol Psychiatry*. 2018; **23**: 639–47.
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; **81**: 559–75.
22. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; **38**: 904–9.

23. Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M, et al. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am J Hum Genet.* 2008; **83**: 445–56.
24. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods.* 2011; **9**: 179–81.
25. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol.* 2010; **34**: 816–34.
26. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010; **467**: 1061–73.
27. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics.* 2016; **32**: 1423–6.
28. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010; **26**: 2190–1.
29. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics.* 2010; **26**: 2336–7.

30. Willour VL, Seifuddin F, Mahon PB, Jancic D, Pirooznia M, Steele J, et al. A genome-wide association study of attempted suicide. *Mol Psychiatry*. 2012; **17**: 433–44.
31. Erlangsen A, Appadurai V, Wang Y, Turecki G, Mors O, Werge T, et al. Genetics of suicide attempts in individuals with and without mental disorders: a population-based genome-wide association study. *Mol Psychiatry*. 2018 [Epub ahead of print].
32. Kimbrel NA, Garrett ME, Dennis MF; VA Mid-Atlantic Mental Illness Research, Education, and Clinical Center Workgroup, Hauser MA, Ashley-Koch AE, et al. A genome-wide association study of suicide attempts and suicidal ideation in U.S. military veterans. *Psychiatry Res*. 2018; **269**: 64–9.
33. Levey DF, Polimanti R, Cheng Z, Zhou H, Nuñez YZ, Jain S, et al. Genetic associations with suicide attempt severity and genetic overlap with major depression. *Transl Psychiatry*. 2019; **9**: 22.
34. Mullins N, Perroud N, Uher R, Butler AW, Cohen-Woods S, Rivera M, et al. Genetic relationships between suicide attempts, suicidal ideation and major psychiatric disorders: a genome-wide association and polygenic scoring study. *Am J Med Genet B Neuropsychiatr Genet*. 2014; **165B**: 428–37.
35. Perroud N, Uher R, Ng MY, Guipponi M, Hauser J, Henigsberg N, et al. Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J*. 2012; **12**: 68–77.



36. Stein MB, Ware EB, Mitchell C, Chen CY, Borja S, Cai T, et al. Genomewide association studies of suicide attempts in US soldiers. *Am J Med Genet B Neuropsychiatr Genet.* 2017; **174**: 786–97.
37. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011; **88**: 76–82.
38. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry.* 1999; **56**: 617–26.
39. Bostwick JM, Pabbati C, Geske JR, McKean AJ. Suicide Attempt as a Risk Factor for Completed Suicide: Even More Lethal Than We Knew. *Am J Psychiatry.* 2016; **173**: 1094–100.
40. Nock MK, Green JG, Hwang I, McLaughlin KA, Sampson NA, Zaslavsky AM, et al. Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: results from the National Comorbidity Survey Replication Adolescent Supplement. *JAMA Psychiatry.* 2013; **70**: 300–10.
41. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics.* 2015; **31**: 1466–8.
42. Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol.* 2016; **12**: e1004714.

43. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 2012; **40**: D109–14.
44. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 2011; **39**: D691–7.
45. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012; **40**: D930–4.
46. Russell R, Metraux D, Tohen M. Cultural influences on suicide in Japan. *Psychiatry Clin Neurosci.* 2017; **71**: 2–5.
47. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet.* 2019; **51**: 584–91.
48. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. *Nat Genet.* 2017; **49**: 1304–10.
49. Fu Q, Heath AC, Bucholz KK, Nelson EC, Glowinski AL, Goldberg J, et al. A twin study of genetic and environmental influences on suicidality in men. *Psychol Med.* 2002; **32**: 11–24.
50. Fan CC, Schork AJ, Brown TT, Spencer BE, Akshoomoff N, Chen CH, et al. Williams Syndrome neuroanatomical score associates with GTF2IRD1 in large-scale magnetic resonance imaging cohorts: a proof of concept for multivariate endophenotypes. *Transl Psychiatry.* 2018; **8**: 114.

51. Sokolowski M, Wasserman J, Wasserman D. Polygenic associations of neurodevelopmental genes in suicide attempt. *Mol Psychiatry*. 2016; **21**: 1381–90.
52. Klein-Tasman BP, Li-Barber KT, Magargee ET. Honing in on the social phenotype in Williams syndrome using multiple measures and multiple raters. *J Autism Dev Disord*. 2010; **41**: 341–51.
53. Kopp ND, Parrish PCR, Lugo M, Dougherty JD, Kozel BA. Exome sequencing of 85 Williams-Beuren syndrome cases rules out coding variation as a major contributor to remaining variance in social behavior. *Mol Genet Genomic Med*. 2018; **6**: 749–65.
54. Richards C, Jones C, Groves L, Moss J, Oliver C. Prevalence of autism spectrum disorder phenomenology in genetic disorders: A systematic review and meta-analysis. *Lancet Psychiatry*. 2015; **2**: 909–16.
55. Schneider T, Skitt Z, Liu Y, Deacon RM, Flint J, Karmiloff-Smith A, et al. Anxious, hypoactive phenotype combined with motor deficits in *Gtf2ird1* null mouse model relevant to Williams syndrome. *Behav Brain Res*. 2012; **233**: 458–73.
56. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet*. 2017; **49**: 27–35.
57. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*. 2015; **87**: 1215–33.

58. Radomsky ED, Haas GL, Mann JJ, Sweeney JA. Suicidal behavior in patients with schizophrenia and other psychotic disorders. *Am J Psychiatry*. 1999; **156**: 1590–5.
59. Chen MH, Pan TL, Lan WH, Hsu JW, Huang KL, Su TP, et al. Risk of Suicide Attempts Among Adolescents and Young Adults With Autism Spectrum Disorder: A Nationwide Longitudinal Follow-Up Study. *J Clin Psychiatry*. 2017; **78**: e1174–9.
60. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*. 2019; **51**: 793–803.

Author accepted manuscript

**FIGURE LEGENDS**

**Figure 1.** Forest plot showing estimated SNP-based heritability in completed suicide.

SNP-based heritability ( $h^2_{\text{SNP}}$ ) was estimated with assumed prevalence rates of 0.1 and 0.5% using the GREML method. Averages and 95% confidence intervals are shown.

**Figure 2.** Polygenic risk score (PRS) analysis of completed suicide.

We analyzed two discovery/target sets, using the 1st set as the discovery set and the 2nd set as the target set, and vice versa. The  $P$  threshold ( $P_t$ ) for selecting ‘risk’ SNPs was sequentially set at 0.1, 0.2, 0.3, 0.4, and 0.5 without SNPs in the major histocompatibility complex (MHC) region. The y-axis indicates the explained variation (Nagelkerke’s  $R^2$ ).

**Figure 3.** Regional plots of the top hit in the association results based on the meta-analysis of the 1st and 2nd sets. Blue lines indicate the recombination rate for East Asian samples of the 1000 Genomes Project phase 3. The Y axis is  $-\log_{10}(P \text{ value})$  of the SNPs and the X axis is chromosomal position (hg19). The linkage disequilibrium ( $r^2$ ) between the top and the remaining SNPs is indicated by color.

Figure 1

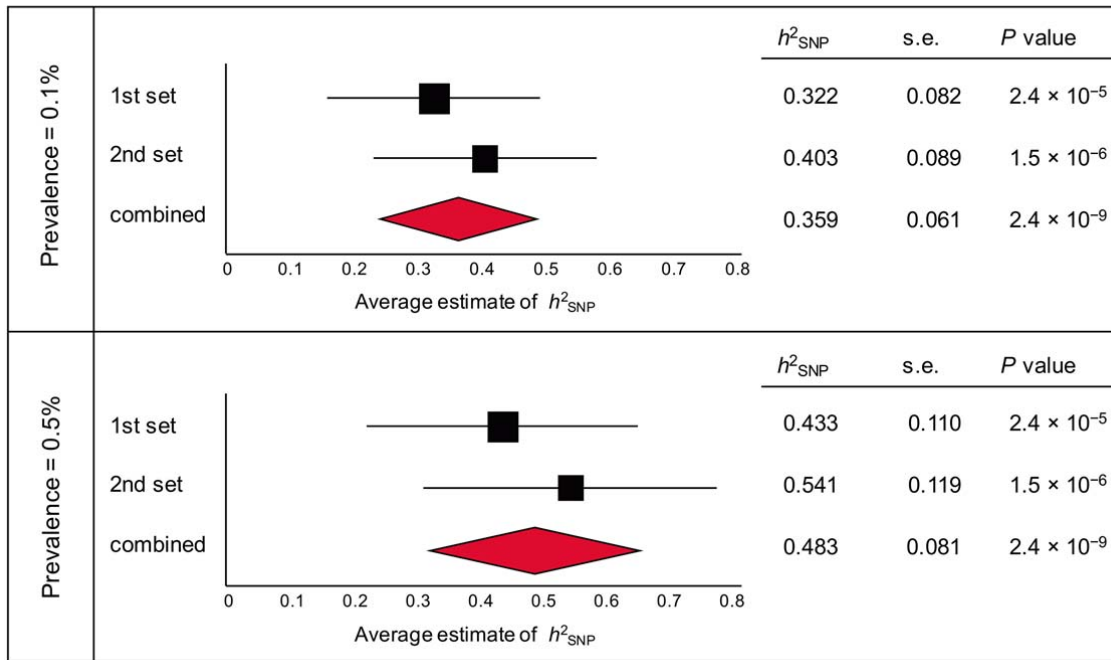


Figure 2

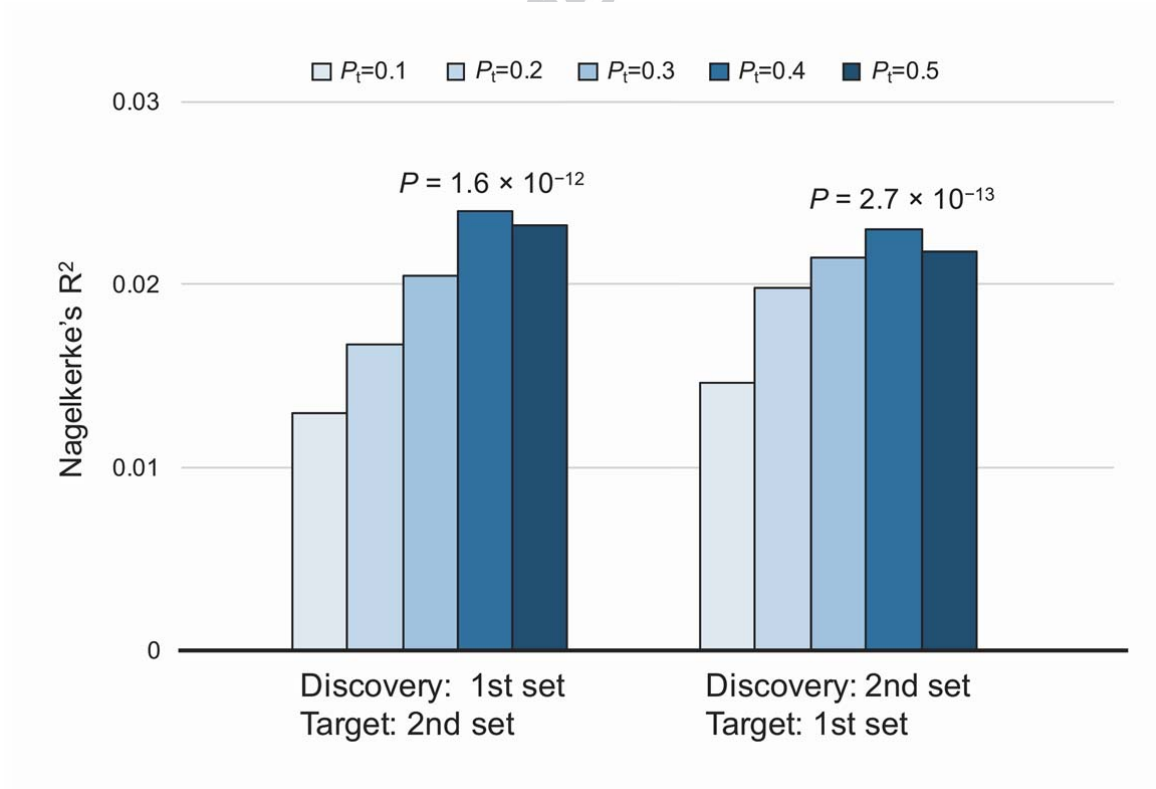
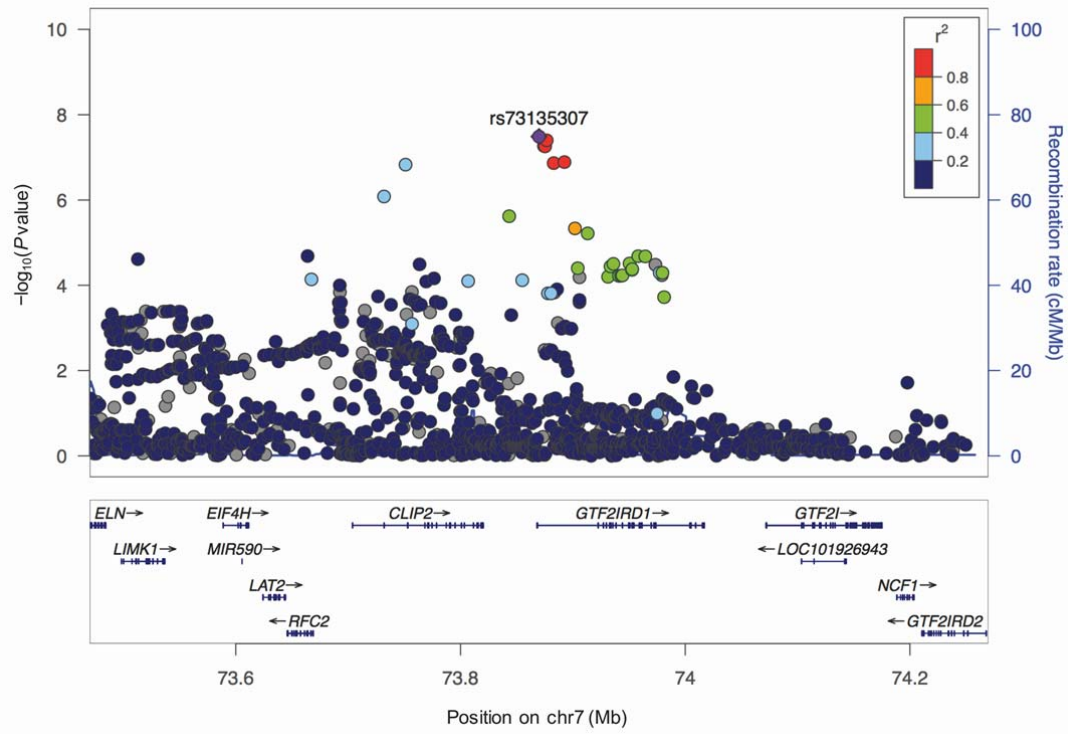


Figure 3



Author accepted

**Table 1. Top variant identified by meta-analysis of genome-wide association analyses of age at completed suicide**

SNP	Chr <sup>a</sup>	Position <sup>a</sup> (bp)	Gene <sup>a</sup> (location)	Effect allele (REF allele)	Stage	beta <sup>b</sup>	s.e.	<i>P</i> value	<i>P</i> <sub>het</sub> <sup>c</sup>
rs73135307	7	73869857	<i>GTF2IRD1</i> (intronic)	G (C)	1st set	-9.84	2.97	$9.17 \times 10^{-4}$	
					2nd set	-15.6	3.39	$4.15 \times 10^{-6}$	
					Meta-analysis	-12.3	2.23	<b><math>3.25 \times 10^{-8}</math></b>	0.20

Suggestive significance ( $P < 5.0 \times 10^{-8}$ ) after meta-analysis of 1st and 2nd set is shown in bold. Only top SNP in the same region is shown here. Full results ( $P < 1.0 \times 10^{-6}$ ) can be seen in Table S6. Chr, chromosome; bp, base pair; REF, reference allele.

<sup>a</sup>Position and Gene are based on Human Genome version 19 (hg19), build 37.

<sup>b</sup>beta were calculated by linear regression based on Wald tests implemented in Rvtests software.

<sup>c</sup>*P* for heterogeneity between two analyses (1st set and 2nd set) were calculated by Cochran's Q test.



