

1 **Title**

2 **Positional cloning and comprehensive mutation analysis of a Japanese family with**
3 **lithium-responsive bipolar disorder identifies a novel *DOCK5* mutation**

4

5 Hiromi Umehara^a, Masayuki Nakamura^{a,*}, Mio Nagai^b, Yuko Kato^c, Shu-ichi Ueno^d,
6 Akira Sano^a

7

8 ^aDepartment of Psychiatry, Kagoshima University Graduate School of Medical and
9 Dental Sciences

10 8-35-1 Sakuragaoka, Kagoshima 890-8520, Japan, Phone: +81-99-2755346

11 ^bMatsuyama Red Cross Hospital, Division of Psychiatry, Matsuyama, Japan

12 ^cJiundo Hospital, Division of Psychiatry, Tokyo, Japan

13 ^dDepartment of Neuropsychiatry, Ehime University Graduate School of Medicine

14 ,Toon, Japan

15

16 *Corresponding author: Masayuki Nakamura

17 Department of Psychiatry, Kagoshima University Graduate School of Medical and Dental

18 Sciences, 8-35-1 Sakuragaoka, Kagoshima 890-8520, Japan

19 Email address: nakamu36@m.kufm.kagoshima-u.ac.jp

20 Telephone number: +81-99-2755346

21 Fax number: +81-99-265-7089

22 ORCID ID (Masayuki Nakamura): [0000-0001-5558-0418](https://orcid.org/0000-0001-5558-0418)

23

24 **Conflict of Interest Statement**

25 The authors have nothing to disclose.

26

27

28 **Abstract**

29 Bipolar disorder (BD) is a severe psychiatric disorder characterized by the recurrence of
30 depressive and manic episodes. Its heritability is high, and many linkage and association
31 studies have been performed. Although various linkage regions and candidate genes
32 have been reported, few have shown sufficient reproducibility, and none have identified
33 the pathogenic genes based on the results of the linkage analysis. To find functional
34 variants that are expected to be rare and have strong genetic effects, we recruited 10
35 healthy individuals, two individuals with unknown status, and six patients with BD or
36 recurrent major depressive disorder (MDD) from a Japanese family consisting of 21
37 members. We performed a genome-wide linkage analysis using a 100K single
38 nucleotide polymorphism (SNP) array and microsatellite markers to narrow linkage
39 regions within this family. Subsequently, we performed whole exome sequencing for
40 two patients with BD to identify genetic mutations in the narrowed linkage regions.
41 Then, we performed co-segregation analysis for DNA variants obtained from the results
42 of the exome sequencing. Finally, we identified a rare heterozygous mutation in exon 31
43 of *DOCK5* (c.3170A>G, p.E1057G). Convergent functional genomics analysis revealed
44 that *DOCK5* was listed as one of the biomarkers for mood state and suicidality.
45 Although *DOCK5* is still a functionally unknown gene, our findings highlight the

46 possibility of a pathological relationship between BD and *DOCK5*.

47

48 **Keywords**

49 *DOCK5*; lithium responder; whole-exome sequencing; linkage study

50

51 **1. Introduction**

52 Bipolar disorder (BD) is a severe psychiatric disorder characterized by the recurrence of
53 depressive and manic episodes. In the general population, the lifetime prevalence of BD
54 is around 1%, and those of subthreshold BD and bipolar spectrum disorder are 1.4% and
55 2.4%, respectively (1). In Japan, the prevalence of bipolar disorder is much smaller, at
56 0.2% (2). Although molecular genetic research indicates that BD is a multifactorial
57 disorder in which both genetic and environmental factors play a role (3), the etiology of
58 BD is not clearly understood. The overall heritability of bipolar spectrum disorders is
59 estimated to be 0.71 (4). Many linkage and association studies have been performed so
60 far, and various linkage sites have been reported (5). However, many of these results do
61 not have sufficient reproducibility, and the linkage analyses have not yet identified the
62 pathogenic gene. Furthermore, with advances in gene technology, several genes have
63 also been identified by genome-wide association studies (GWAS). However, their
64 functional correlations remain largely unknown. A study by Lee et al. showed that the
65 cumulative impact of small general alleles can only account for about 40% of BD
66 phenotypic differences (6). Therefore, it can be hypothesized that rare and highly
67 penetrant mutations may contribute to BD susceptibility. To evaluate this hypothesis, a
68 family with a high frequency of BD can be studied to extract and evaluate genetic

69 mutations inherited from a common ancestor (7). Therefore, conducted this study with
70 the aim of finding functional variants that are expected to be rare and have strong
71 genetic effects in a family in which BD and related diseases frequently occur. In this
72 study, we recruited 10 unaffected individuals, two individuals with unknown status,
73 three patients with BD, and three patients with recurrent major depressive
74 disorder (MDD) from a Japanese family consisting of 21 members. To narrow linkage
75 regions, we used the family's genetic material to perform a genome-wide linkage
76 analysis using a 100K single nucleotide polymorphism (SNP) array and microsatellite
77 markers. Subsequently, we performed whole exome sequencing on two of the patients
78 with BD to identify genetic mutations on the narrowed linkage regions. Then, we
79 performed co-segregation analysis for DNA variants obtained from the results of exome
80 sequencing.

81 **2. Materials and methods**

82 **2.1. Samples**

83 Participants across three generations in a Japanese family including patients with BD
84 and recurrent MDD were enrolled in this study (Figure 1). All affected individuals were
85 diagnosed independently according to the criteria in the Diagnostic and Statistical
86 Manual of Mental Disorders, Fourth Edition (DSM-IV). Written informed consent was

87 obtained from all participants. The clinical diagnoses of patients II-2, II-3, and III-5
88 were BD-I and those of III-1, III-2, and III-6 were recurrent MDD. II-8, with alcoholic
89 psychosis, as well as I-1 and I-2, with no clinical information, were classified as
90 unknown for this linkage study. Interestingly, all affected six individuals were
91 successfully treated with lithium carbonate (Table 1). The pedigree structure of the
92 family recruited in this study is shown in Figure 1. Participants' genomic DNA was
93 extracted from peripheral leukocytes using standard methods after informed consent and
94 approval by the local ethics committee were obtained. Altogether, the 18 available
95 family members were studied, including the six affected patients, 10 unaffected
96 individuals, and two individuals with unknown status.

97 **2.2. Genetic analysis**

98 Written informed consent was obtained at the time of blood sampling from all study
99 participants, including these 18 family members (Figure. 1) and 50 unrelated BD-I
100 patients and 67 healthy controls. The research protocol and consent form were approved
101 by the Institutional Review Boards of Kagoshima University.

102 **2.3. Linkage and haplotype analysis**

103 Genome-wide two-points linkage analysis was performed on six affected family members,
104 10 unaffected family members, and two family members with unknown status (Figure 1)

105 using the Gene Chip Human Mapping 100K set (Affymetrix, Santa Clara, CA, USA).
106 After genotyping, a two-points linkage analysis was performed using the GeneSpring
107 GT2.0 (Agilent Technologies, Santa Clara, CA, USA) for narrowing and screening of
108 candidate linkage regions of BD in this family. In this family, the autosomal dominant
109 mode of inheritance was assumed because phenotypes appeared in all generations with
110 no gender differences. Data were analyzed for dominant mode of inheritance with 90%
111 penetrance values and the disease allele frequency was assumed to be 0.001. We extracted
112 genotypable 114,816 SNPs out of the 116,204 SNPs in the Human Mapping Array 100K
113 set. For the extracted SNPs, we removed 825 SNPs which were inconsistent with
114 autosomal dominant inheritance pattern in this family and performed parametric analysis
115 using the remaining 113,991 SNPs for the genome-wide two-points linkage analysis.
116 Following preliminary haplotype and sequencing analysis revealed a candidate disease
117 mutation, c.3170A>G in *DOCK5* on chromosome 8p, whose proximate region showed
118 the highest LOD score in the genome-wide two-points linkage analysis. For chromosome
119 8p, we performed a multipoint linkage analysis using microsatellite markers using
120 MERLIN (Multipoint Engine for Rapid Likelihood Inference) v.1.1.2. We set the
121 penetrance to 80% and the disease allele frequency to 0.002, taking into account the
122 penetrance rate of the c.3170A>G in this family and the incidence of bipolar disorder in

123 Japan (2), respectively. In addition, multipoint linkage analysis and haplotype analysis
124 were performed on chromosome 8p to show genetic evidence of recombination using 41
125 SNPs from the Gene Chip Human Mapping 100K set and the Genome-Wide Human SNP
126 Array 6.0 (Affymetrix), and seven microsatellite markers on chromosomes 8p23.1 to
127 8q12.2: D8S277, D8S549, D8S258, D8S136, D8S1771 D8S1443, and D8S260 in the 18
128 available family members (Figure 2).

129 The microsatellite markers that are highly polymorphic in Japanese populations were
130 selected from the map of the human genome (8). The primer sequences were taken from
131 the Genome Database. We genotyped the seven microsatellite markers (Figure 2). The
132 microsatellite sequence was amplified with PCR. The forward primer in each marker set
133 was labeled with a fluorescent dye at the 5' end. The 12.5 μ l PCR mixture contained 1
134 μ l of template DNA and 0.25 μ l of each forward and reverse primer (10 μ mol), and the
135 DNA region was amplified using TaKaRa Taq DNA polymerase under the following
136 conditions: denaturation at 94°C for 2 min, 35 cycles at 94°C for 30 s, 58–62°C for 30 s
137 and 72°C for 30 s, followed by a final extension step of 5 min at 72°C. The amplified
138 products were observed on 1% GelRed (Biothium Inc, Hayward, CA, USA) pre-stained
139 agarose gel to determine the efficacy of PCR amplification.

140 The PCR products were then denatured and subjected to capillary electrophoresis on an

141 ABI PRISM 3130 Avant Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA,
142 USA). DNA fragment sizes were analyzed relative to Gene Scan 500 RIZ Dye size-
143 standard (Applied Biosystems, Foster City, CA, and USA) using the Genemapper
144 software (Applied Biosystems).

145 **2.4. DNA sequence analysis using Whole-Exome sequencing**

146 Whole-exome sequencing was performed in II-2 and III-5 (both diagnosed with BD-I)
147 according to the manufacturers' protocol. Genomic DNA was captured with the
148 SureSelect Target Enrichment System Human All Exon V5 Kit (Agilent Technologies,
149 Santa Clara, CA, USA) and sequenced per lane on an Illumina HiSeq 2000 (Illumina,
150 San Diego, CA, USA) with paired-end 101 bp reads. Image analysis and base calling
151 were performed using sequencing control software with real-time analysis and
152 CASAVA ver.1.8.2 software (Illumina). Generated sequence data (FastQ files) were
153 processed using the pipeline with BWA (version 0.5.9), SAMtools, Picard (version
154 1.59, <http://picard.sourceforge.net/>) and GATK32 (version 1.6–5). Variant calls were
155 made using the GATK best practices recommendations. To prioritize variations, we
156 applied several filtering steps. First, we included non-synonymous variations. Second,
157 we included variations inherited in an autosomal dominant manner, considering the
158 inheritance pattern of this family. Third, we included variations that were co-segregated

159 in the two family members with BD-I who underwent whole exome sequencing. Fourth,
160 we included variations with mutant allele frequency less than 0.01 according to the
161 Japanese reference panel project of the Tohoku Medical Megabank (4.7KJPN) and
162 NCBI databases. Fifth, we included mutations that exist in the candidate region based
163 on the results of the linkage analysis described above (Figure 4). To validate prioritized
164 variations, Sanger sequencing was performed using a 3130 × 1 Genetic Analyzer
165 (Applied Biosystems).

166 **2.5. DNA sequence analysis using Sanger sequencing**

167 Sanger sequencing of the mutations on the candidate genes was performed for the 18
168 family members based on the results of linkage analysis and whole exome sequencing.
169 Then, we identified a mutation (c.3170A>G) in *DOCK5* gene. For candidate mutations,
170 sequences in *DOCK5* from 50 unrelated Japanese patients with sporadic BD-I and 67
171 unrelated Japanese healthy controls were searched.
172 For Sanger sequencing, polymerase chain reaction primers (sequences available upon
173 request) were designed to amplify all coding exons (non-coding RNA exons),
174 5'untranslated regions, and intronic regions surrounding each exon. All PCR products
175 were subjected to Sanger sequencing. PCR reactions were performed in a total volume
176 of 12.5 μ l containing 10–30 ng of each DNA sample, 0.2 μ M each of the forward and

177 reverse oligonucleotide primers, 0.2 mM each of dATP, dGTP, dTTP, and dCTP, and 0.5
178 units of TaKaRa Taq. The PCR program was 94°C for 2 min, 35 cycles of 94°C for 30 s,
179 58–64°C for 30 s and 72°C for 1 min, followed by 72°C for 5 min. The amplified PCR
180 products were separated on 1% agarose gels at 100 V for 30 min and were labeled using
181 a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with the same
182 primers used in the initial PCR and following the protocol described above. The
183 products were then directly sequenced on an ABI PRISM 3130 Avant Genetic Analyzer
184 (Applied Biosystems).

185 **2.6. Copy-number variants (CNV) analysis using Genome-Wide Human SNP**

186 **Array**

187 Whole genome CNV analysis was performed for the 17 (one data set from poor quality
188 DNA of patient III-1 was not usable) family members using the Genome-Wide Human
189 SNP Array 6.0 (Affymetrix), which features 1.8 million genetic markers, including
190 more than 906,600 SNPs and more than 946,000 comparative genomic hybridization
191 probes. Labeling and hybridization were performed according to the Affymetrix
192 protocol. Raw data were acquired from the scanner and transferred to a Genotyping
193 Console 3.0.2 (Affymetrix) to determine the genotypes and CNVs from fluorescence
194 intensity data. For all chromosomes, signal intensities for the copy-number predictions

195 of both duplications and deletions were investigated. CNV aberrations of more than 100
196 kbp were automatically detected by the genotyping console 3.0.2 software. In addition,
197 small CNV aberrations of a few hundred to several thousand bp were visually detected
198 in detail from the raw data. Because DNA from III-1 was unavailable for the Genome-
199 Wide Human SNP Array 6.0 analysis due to low quality, this array was not used in
200 genome-wide linkage analysis.

201

202 **3. Results**

203 **3.1. Linkage and haplotype analysis**

204 We performed a genome-wide scan using the DNA arrays to confirm the evidence for
205 linkage. In preliminary two-point linkage analysis, the highest LOD score of 2.3 ($\theta =$
206 0) was obtained for marker rs10503492 on chromosome 8p22. The second highest LOD
207 score of 2.10 ($\theta = 0$) was obtained for marker rs10491029 on chromosome 10q26.

208 Multipoint linkage analysis using SNPs and microsatellite markers revealed a
209 consistently high LOD score at around D8S136 at 8p (Figure 3). Haplotype analysis
210 using SNPs and additional microsatellite markers analysis revealed a common
211 haplotype delimited by rs10503492 and rs10504053 on chromosome 8p23.1 to 8p11.1,
212 which was segregated in all the available patients (Figure 2). The proximal
213 recombination site was between rs10503452 and rs10503492, and the distal site was

214 between rs10504053 and rs1474386. These results demonstrate that the region linked to
215 BD and its related diseases was possibly within the approximately 30 Mb section
216 spanned by rs10503492 and rs10504053 on chromosomes 8p23.1 to 8p11.1 (Figure 2).
217 Chromosome 10 was analyzed in the same way, but no mutation that was co-segregated
218 in the patients was found (described in next section).

219 **3.2. DNA sequence analysis**

220 A region of approximately 69.35 Mb was captured from each of the II-2 and III-5
221 exomes (Supplementary Table). The average reading depth was 82, and about 93.1% of
222 the target area was covered by more than 10 reads. In II-2, 74,386 mutations including
223 65,858 SNPs and 5,089 indels were detected in the coding region, and in III-5, 70,947
224 mutations including 68,615 SNPs and 5,771 indels were detected in the coding region.
225 Filtering was performed on the detected mutations, as described in the Materials and
226 Methods section. Only three mutations were found under the linkage peak region on
227 chromosome 8p. In addition, for the three variants, Sanger sequencing revealed that
228 only a g.25224432A>G (E1057G) mutation on *DOCK5* was co-segregated in the
229 affected individuals; it was also found in two healthy family members.

230 In 67 healthy controls and 50 patients with BD-I, Sanger sequencing was performed for
231 the *DOCK5* mutation (E1057G). However, no similar mutation was found. According to

232 4.7KJPN, this is an extremely rare mutation with a frequency of 0.0002. This mutation
233 was not found in the gnomAD (The Genome Aggregation Database). Additionally, in
234 *DOCK5*, Sanger sequencing for all 52 coding exons and exon-intron boundaries was
235 performed for 50 patients with BD-I. However, there were no pathological mutations,
236 except for high minor allele frequency (MAF) polymorphisms (>0.01).

237 We evaluated the E1057G mutation in *DOCK5* using the following five *in silico* tools:
238 poyphen2, SHIFT, PANTHER, Mutation Taster, and PROVEAN. In poyphen2 and
239 SHIFT, the results were “Benign” and “tolerated”, respectively. The results of
240 PANTHER, PROVEAN, and Mutation Taster were “potentially damaging”,
241 “deleterious”, and “disease causing”, respectively.

242 **3.6. Copy-number variants (CNV) analysis using Genome-Wide Human SNP**

243 **Array**

244 In all chromosomes, the signal intensity of both the replication and deletion copy-
245 number predictions was examined in detail from the raw data. However, no specific
246 CNVs were found in affected individuals.

247 **4. Discussion**

248 The aim of this study was to identify potential susceptibility variants that contribute to
249 the risk of BD. In this study, we identified an E1057G mutation in *DOCK5*, which was

250 carried by all family members affected by BD-I and recurrent major depression.

251 However, two healthy individuals also carried this mutation. Thus, we considered two

252 possibilities. The first was penetrance rate. If an odds ratio of three, frequency of

253 0.0002, a dominant model and a one-tailed test for $p = 0.05$ were assumed, then power

254 is $> 80\%$ for a sample of 15,000 BPD patients and 15,000 controls. With an odds ratio

255 of five, 5,700 case-control pairs are needed. Such odds ratios would be compatible with

256 two unaffected carriers found in this family. The second was that individuals with this

257 mutation that were healthy at the time of analysis may be affected later. II-9 was a 50-

258 year-old woman without history of psychiatric treatment and detailed clinical

259 information although she was considered to be depressive state with under threshold of

260 diagnose at the interview of the informed consent for the DNA analysis. II-10 was a 47-

261 year-old woman without history of psychiatric treatment and detailed clinical

262 information although she showed hyperthymia with under threshold of diagnose of

263 mood disorder at the interview of the informed consent for the DNA analysis. *In silico*

264 analysis showed inconsistent results. The inheritance of this family suggests an

265 autosomal dominant form, which is generally explained by a gain-of-function mutation.

266 It has been reported that polyphen2 and SHIFT are better at predicting loss-of-function

267 mutations than gain of function mutations (9), suggesting that these *in silico* programs

268 may not be ideal for predicting the effects of the E1057G mutation in *DOCK5*. In
269 general, in order to avoid missing causative mutations and to take the best advantage of
270 in silico analysis, different computational platforms should be used (10). Although
271 further functional analysis is required, one of the rationales of the c.3170A>G mutation
272 in *DOCK5* as risk for BD was that the results of other three platforms, PANTHER,
273 PROVEAN, and Mutation Taster showed “potentially damaging”, “deleterious”, and
274 “disease causing”, respectively.

275 Although there was no significant association between genes on chromosome 8 and BD
276 in the data base of the second GWAS of the multinational Psychiatric Genomics
277 Consortium BD Working Group (11), in a study of approximately 400 families affected
278 by BD, Park et al. reported that the highest LOD score (higher than 3) from a linkage
279 analysis was obtained on 8p21 at the D8S382 region, whose physical position is from
280 25,952,418 to 25,952,386 (GRCh37) (12). The distance between D8S382 and the
281 E1057G mutation in *DOCK5* is approximately 700,000 bp, which is a remarkably close
282 genetic distance equivalent to 0.7 cM. Another study reported that a *DOCK5* mutation
283 was significantly associated with various psychiatric disorders, according to a whole
284 exome analysis of 32 patients from two families with various psychiatric disorders,
285 including bipolar disorder, and 33 controls (13). Furthermore, Le-Niculescu et al.

286 reported that variable expression levels of *DOCK5* were associated with mood
287 disorders, psychosis and suicidality, and that *DOCK5* may be a key genetic biomarker
288 reflecting those diseases (14).

289 *DOCK5* is a member of a family of genes encoding an atypical Rho guanine nucleotide
290 exchange factor (GEF) for Rac and/or Cdc42 GTPases. The Dock family is genetically
291 conserved from *Drosophila melanogaster* and *Caenorhabditis elegans* to mammals
292 (15). In addition, 11 members of the Dock family have been identified and divided into
293 four groups (Dock-A, B, C, and D) based on their domain structure differences. It has
294 been reported that the Dock family is associated with several neurodegenerative and
295 neuropsychiatric disorders (16). *Dock2*, belonging to the Dock-A group (the same group
296 as Dock5) is a regulator of microglia and is associated with Alzheimer's disease (17).
297 Recently, an association between microglial activation and mental disorders such as
298 schizophrenia and depression was identified (18), suggesting that *DOCK5* may be
299 involved in microglial activation. In addition, *DOCK5* is an indispensable molecule for
300 transmitting GSK3 β signals in mast cells during allergic reactions (19). A similar
301 mechanism may also exist in the central nervous system, because GSK3 β -related
302 proteins and *DOCK5* are expressed in the central nervous system. It is also interesting
303 to note that all affected individuals in the present study responded to lithium carbonate.
304 In a GWAS study using 2,000 BD lithium carbonate responders, only ch8 rs7833426
305 was associated with the BD when the study was restricted to subjects with Asian
306 ancestors, although at just a nominal significant level (20). rs7833426 is within the
307 linkage region in the present study. These results suggest that *DOCK5* is associated with
308 lithium-reactive BD. Although further genetic analysis of additional family members

309 and functional analysis are required, our results suggest that the DOCK5 protein is a
310 possible pathological molecule for BD and related disorders.

311

312 **5. Conclusions**

313 Positional cloning of the genes of a Japanese family affected by BD was performed
314 using a SNP array and whole exome sequencing. The analysis revealed an E1057G
315 mutation in the *DOCK5* gene, which was located in the region with the highest LOD
316 score. *DOCK5* may be involved in the pathology of BD and its related diseases.

317

318 **Acknowledgments**

319 The authors thank all involved patients and their families for their participation. The
320 authors also thank Ms. Hiwatashi, Ms. Yokoyama, and Ms. Meguro for their technical
321 assistance. We would like to thank Uni-edit (<https://uni-edit.net/>) for editing and
322 proofreading this manuscript.

323 **Source of funding**

324 This study has no financial support or funding.

325 **Declaration of Interest**

326 The authors declare that they have no known competing financial interests or personal
327 relationships that could have appeared to influence the work reported in this paper.

328

329 **References**

- 330 1. Merikangas KR, Jin R, He J-P, Kessler RC, Lee S, Sampson NA, et al. National
331 Institute of Mental Health, Genetic Epidemiology Research Branch, Building 35,
332 Room 1A201, 35 Convent Drive. 2011;3720(3):241–51. Available from:
333 www.hcp.med.harvard.edu/wmh/
- 334 2. Ishikawa H, Kawakami N, Kessler RC. Lifetime and 12-month prevalence,
335 severity and unmet need for treatment of common mental disorders in Japan:
336 Results from the final dataset of World Mental Health Japan Survey.
337 *Epidemiology and Psychiatric Sciences*. 2016;25(3):217–29.
- 338 3. Barnett JH, Smoller JW. The genetics of bipolar disorder. *Neuroscience*.
339 2009;164(1):331–43.
- 340 4. Edvardsen J, Torgersen S, Røysamb E, Lygren S, Skre I, Onstad S, et al.
341 Heritability of bipolar spectrum disorders. Unity or heterogeneity? *Journal of*
342 *Affective Disorders*. 2008;106(3):229–40.
- 343 5. Barnett JH, Smoller JW. The genetics of bipolar disorder. *Neuroscience*. 2009.
- 344 6. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability
345 for disease from genome-wide association studies. *American Journal of Human*

- 346 Genetics. 2011;88(3):294–305.
- 347 7. Collins AL, Kim Y, Szatkiewicz JP, Bloom RJ, Hilliard CE, Quackenbush CR, et
348 al. Identifying bipolar disorder susceptibility loci in a densely affected pedigree.
349 Molecular Psychiatry. 2013;18(12):1245–6.
- 350 8. Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardson
351 B, et al. A high-resolution recombination map of the human genome. Nature
352 Genetics. 2002;31(3):241–7.
- 353 9. Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-
354 function and gain-of-function mutations. Genetic Testing and Molecular
355 Biomarkers. 2010;14(4):533–7.
- 356 10. Duarte AJ, Ribeiro D, Moreira L, Amaral O. In silico analysis of missense
357 mutations as a first step in functional studies: Examples from two
358 sphingolipidoses. International Journal of Molecular Sciences. 2018;
- 359 11. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al.
360 Genome-wide association study identifies 30 loci associated with bipolar disorder.
361 Nature Genetics. 2019; May ; 51(5): 793–803.
- 362 12. Park N, Juo SH, Cheng R, Liu J, Loth JE, Lilliston B, et al. Linkage analysis of

- 363 psychosis in bipolar pedigrees suggests novel putative loci for bipolar disorder
364 and shared susceptibility with schizophrenia. *Molecular Psychiatry*.
365 2004;9(12):1091–9.
- 366 13. Ganesh S, Ahmed P. H, Nadella RK, More RP, Seshadri M, Viswanath B, et al.
367 Exome sequencing in families with severe mental illness identifies novel and rare
368 variants in genes implicated in Mendelian neuropsychiatric syndromes.
369 *Psychiatry and Clinical Neurosciences*. 2019;73(1):11–9.
- 370 14. Le-Niculescu H, Kurian SM, Yehyawi N, Dike C, Patel SD, Edenberg HJ, et al.
371 Identifying blood biomarkers for mood disorders using convergent functional
372 genomics. *Molecular Psychiatry*. 2009;14(2):156–74.
- 373 15. Côté JF, Vuori K. Identification of an evolutionary conserved superfamily of
374 DOCK180-related proteins with guanine nucleotide exchange activity. *Journal of*
375 *Cell Science*. 2002;115(24):4901–13.
- 376 16. Shi L. Dock protein family in brain development and neurological disease.
377 *Communicative and Integrative Biology*. 2013;6(6):1–9.
- 378 17. Cimino PJ, Yang Y, Li X, Hemingway JF, Cherne MK, Khademi SB, et al.
379 Ablation of the microglial protein DOCK2 reduces amyloid burden in a mouse
380 model of Alzheimer’s disease. *Experimental and Molecular Pathology*.

- 381 2013;94(2):366–71.
- 382 18. Kato TA, Watabe M, Kanba S. Neuron-glia interaction as a possible glue to
383 translate the mind-brain gap: A novel multi-dimensional approach toward
384 psychology and psychiatry. *Frontiers in Psychiatry*. 2013;4(OCT):1–7.
- 385 19. Ogawa K, Tanaka Y, Uruno T, Duan X, Harada Y, Sanematsu F, et al. DOCK5
386 functions as a key signaling adaptor that links FcεRI signals to microtubule
387 dynamics during mast cell degranulation. *Journal of Experimental Medicine*.
388 2014;211(7):1407–19.
- 389 20. Hou L, Heilbronner U, Degenhardt F, Adli M, Akiyama K, Akula N, et al.
390 Genetic Variants associated with response to lithium treatment. Vol. 387, *The*
391 *Lancet*. 2016. p. 1085–93.
- 392

393 **Figure legends**

394

395 **Figure 1**

396 Pedigree of a Japanese family having a high frequency of bipolar disorder and recurrent
397 major depressive disorder. Affected status is indicated with a filled square (male) or
398 circle (female). Gray squares and circles show unknown disease status. Black spots
399 indicate individuals who were sampled and analyzed.

400 **Figure 2**

401 Haplotypes of each individual, with the disease haplotype boxed. The absence of a
402 haplotype indicates that the individual was not genotyped at that microsatellite marker
403 or single-nucleotide polymorphism locus.

404

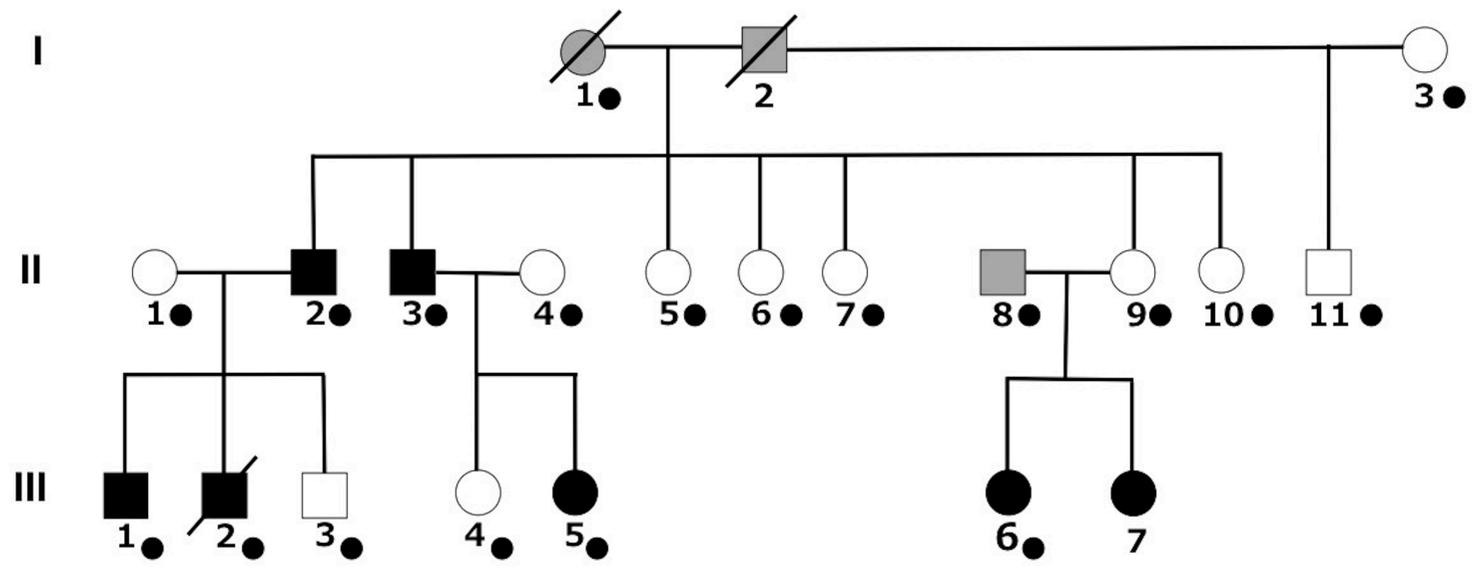
405 **Figure 3**

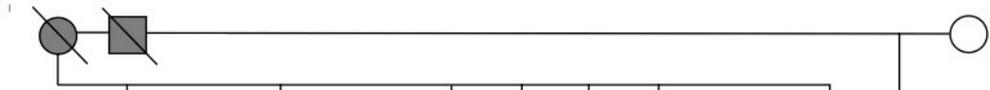
406 Multipoint logarithm of odds (LOD) scores for genetic locations on chromosome
407 8q21.2. The maximum LOD score was 2.1 for marker D8S136.

408

409 **Figure 4**

410 Filtering steps applied to variations called from the whole exome sequencing of II-2 and
411 III-5.





II	8P23.1	8p22	8p21.3	8p21.2	8p12	8p11.21	8q11.23
5 1	3 1	1 4	3 1	4 4	5 5	5 1	3 1
8487614 rs10503379	B A	A A	B B	A A	B B	B A	A A
10703353 rs2278335	A A	A A	A A	A A	B A	B A	B A
11102432 rs6601577	A A	A A	A A	A A	B A	B A	B A
12607134 rs6991754	B B	B B	B B	B B	B B	B B	B B
13187265 rs1671378	B A	B A	B B	B B	B A	B A	B A
13391847 rs10503452	B B	B B	A A	A B	B B	B B	B B
14142896 rs10503492			B B	B B	B B	B B	B B
15549577 D85549	3 3	2 3	3 3	2 3	3 3	3 3	3 3
17550523 rs4921804	B B	A B	B B	A B	B B	B B	B B
18547448 rs10503627	A B	A B	A A	A B	A B	A B	A B
19017835 rs05901	A A	B A	A A	B A	A A	A A	A A
20267167 D85258	2 4	2 4	1 1	2 4	2 4	2 4	2 4
20569147 rs2616231	B A	B A	B B	B A	A A	A A	A A
20571519 rs1868664	A B	A B	A A	A B	B B	B B	B B
21820763 rs7843479	A B	A B	B A	A B	A B	A B	A B
22333395 D85136	3 1	1 1	1 1	1 1	7 1	7 1	7 1
22752698 rs1509815	B B	B B	A B	B B	B B	B B	B B
23298735 rs2280832	A A	B A	A B	A B	A A	A A	A A
24623861 rs10503750	B B	A B	A B	A B	B B	B B	B B
25441113 D851771	1 3	1 3	2 3	1 3	3 3	3 3	3 3
25880150 rs7008914	A A	B A	A A	B A	B A	B A	B A
27019574 D851443	3 1	2 3	3 2	2 1	3 2	1 1	1 1
28137161 rs10503825	A A	B A	B A	B A	B A	A A	A A
29255484 rs10503841	B B	A B	A B	A B	B B	B B	B B
30275242 rs952774	B B	A B	A A	A B	A B	A B	A B
30959270 rs10503882	B B	A B	B B	A B	B B	B B	B B
31960276 rs9297186	B B	A B	B B	A B	B B	B B	B B
33272321 rs10503946	A B	A A	A A	A B	A A	B A	B A
35151470 rs1112204	B B	A B	A A	A B	B B	A B	A B
36509115 rs7006271	B B	A A	A A	A B	A B	A B	A B
37132195 rs422457	A A	B A	B A	B A	B A	B A	B A
38664074 rs7818605	A A	B A	A B	B A	A A	A A	A A
39791066 rs10504013	A A	A A	A A	A A	B A	B A	B A
40111754 rs1378415	A A	A A	A A	A A	B A	B A	B A
41595513 rs516606	B B	A A	B B	B B	B B	B B	B B
41895213 rs2277132	A A	A B	B A	A A	B A	B A	B A
43343418 rs10504053	B B	B B	B B	B B	B B	B B	B B
47558387 rs1474386	B A	B B	B B	B A	B B	B B	B A
47907403 rs10504055	A A	B B	A B	B A	B A	B A	B A
50106759 rs4873328	A A	B B	B B	B A	B A	B A	B A
50236445 rs1519833	A A	B B	B B	B A	B A	B A	B A
50244700 rs10504078	A A	B B	B B	B A	B A	B A	B A
50263471 rs10504082	A A	B B	B B	B A	B A	B A	B A
51937811 rs966216	A A	B A	A A	A B	A A	A A	A A
53324319 rs3758125	A A	B A	A A	A A	A A	A A	A A
55471517 rs10504167	A A	A B	A A	A A	B A	B A	B A
61721781 D85260	2 4	5 2	5 2	5 2	2 3	7 2	7 2

III	8P23.1	8p22	8p21.3	8p21.2	8p12	8p11.21	8q11.23
3 1	3 4	1 4	3 4	3 4	3 4	5 2	5 2
8487614 rs10503379	B B	A B	A B	A B	B A	B A	B A
10703353 rs2278335	B B	B A	A A	B A	B A	B A	B A
11102432 rs6601577		A A	A A	A B	A A	A A	A A
12607134 rs6991754		B B	B B	B B	B B	B B	B B
13187265 rs1671378	B B	B B	A B	B A	B A	B A	B A
13391847 rs10503452		B A	B B	B B	B B	B B	B B
14142896 rs10503492		A B	B B	B B	B B	B B	B B
15549577 D85549	2 3	2 3	3 3	3 3	2 2	2 2	2 2
17550523 rs4921804	B B	A B	B B	B B	A A	A A	A A
18547448 rs10503627	B B	A A	B A	B A	A A	A A	A A
19017835 rs05901		B A	A A	A A	B A	B A	B A
20267167 D85258	2 1	2 1	1 4	4 4	2 2	2 2	2 2
20569147 rs2616231		B B	A B	B B	A B	A B	A B
20571519 rs1868664		B B	B A	B B	A B	A B	A B
21820763 rs7843479		A A	B A	B B	A A	A A	A A
22333395 D85136	1 1	1 1	1 1	1 2	1 1	1 1	1 1
22752698 rs1509815		B A	B A	B B	B A	B A	B A
23298735 rs2280832		B B	A B	A A	B B	B B	B B
24623861 rs10503750		A A	B A	B B	A A	A A	A A
25441113 D851771		B A	A A	A A	B A	B A	B A
25880150 rs7008914	1 3	1 3	3 3	3 1	1 3	1 3	1 3
27019574 D851443		B A	A A	A A	B A	B A	B A
28137161 rs10503825	2 2	2 3	3 2	1 3	2 2	2 2	2 2
29255484 rs10503841		B B	A A	A B	A B	A B	A B
30275242 rs952774		A A	B B	B B	A B	A B	A B
30959270 rs10503882		A A	B B	B B	A B	A B	A B
31960276 rs9297186		A B	B B	B B	A B	A B	A B
33272321 rs10503946	A A	A A	A A	B A	A A	A A	A A
35151470 rs1112204		A A	B A	B B	A A	A A	A A
36509115 rs7006271	A A	A A	B A	B A	A B	A B	A B
37132195 rs422457		A A	A B	B A	A B	A B	A B
38664074 rs7818605	A A		A B	A A	B A	B A	B A
39791066 rs10504013		A A	A A	A A	A A	A A	A A
40111754 rs1378415		A A	A A	A A	A A	A A	A A
41595513 rs516606		A B	B B	B B	A B	A B	A B
41895213 rs2277132	A A	A B	A B	A B	A B	A B	A B
43343418 rs10504053		B B	B B	B B	B B	B B	B B
47558387 rs1474386		B B	A B	A B	B B	B B	B B
47907403 rs10504055	A	B A	A A	A A	B A	B A	B A
50106759 rs4873328	A	B B	A B	A B	B B	B B	B B
50236445 rs1519833	B	B B	A B	A B	B B	B B	B B
50244700 rs10504078	B	B B	A B	A B	B B	B B	B B
50263471 rs10504082		B B	A B	A B	B B	B B	B B
51937811 rs966216		B A	A A	A A	B B	B B	B B
53324319 rs3758125		B A	A A	A A	B A	B A	B A
55471517 rs10504167	5	A A	A A	A A	A A	A A	A A
61721781 D85260	4	5 2	2 1	2 3	5 2	4 3	4 3

