Regular Article

Enrichment of deleterious variants of mitochondrial DNA polymerase gene (POLG1) in bipolar disorder

Takaoki Kasahara, PhD,1 Mizuho Ishiwata, BSc,1 Chihiro Kakiuchi, MD, PhD,1† Satoshi Fuke, PhD,1‡ Nakao Iwata, MD, PhD,2 Norio Ozaki, MD, PhD,3 Hiroshi Kunugi, MD, PhD,4 Yoshiro Minabe, MD, PhD,5§ Kazuhiko Nakamura, MD, PhD,5*** Yasuhide Iwata, MD, PhD,5 Kumiko Fujii, MD, PhD,1,6 Shigenobu Kanba, MD, PhD,7 Hiroshi Ujike, MD, PhD,5†† Ichiro Kusumi, MD, PhD,9 Muneko Kataoka, MD, PhD,1§§ Nana Matoba, PhD,1** Atsushi Takata, MD, PhD,1*** Kazuya Iwamoto, PhD,10 Takeo Yoshikawa, MD, PhD11 and Tadafumi Kato, MD, PhD1*  

1Laboratory for Molecular Dynamics of Mental Disorders, RIKEN Brain Science Institute, Wako, 2Faculty of Medicine, Department of Psychiatry, Fujita Health University, 3Faculty of Medicine, Department of Psychiatry, Nagoya University, Nagoya, 4Department of Mental Disorder Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Kodaira, 5Department of Psychiatry and Neurology, Hamamatsu University School of Medicine, Hamamatsu, 6Department of Psychiatry, Dokkyo Medical University School of Medicine, Mibu, 7Department of Psychiatry, Kyushu University School of Medicine, Fukuoka, 8Department of Psychiatry, Okayama University School of Medicine, Okayama, 9Department of Psychiatry, Hokkaido University School of Medicine, Sapporo, 10Department of Molecular Brain Science, Graduate School of Medicine, Kumamoto University, Kumamoto, 11Laboratory for Molecular Psychiatry, RIKEN Brain Science Institute, Wako, Japan

Aim: Rare missense variants, which likely account for a substantial portion of the genetic ‘dark matter’ for a common complex disease, are challenging because the impacts of variants on disease development are difficult to substantiate. This study aimed to examine the impacts of amino acid substitution variants in the POLG1 found in bipolar disorder, as an example and proof of concept, in three different modalities of assessment: in silico predictions, in vitro biochemical assays, and clinical evaluation. We then tested whether deleterious variants in POLG1 contributed to the genetics of bipolar disorder.

Methods: We searched for variants in the POLG1 gene in 796 Japanese patients with bipolar disorder and 767 controls and comprehensively investigated all 23 identified variants in the three modalities of assessment. POLG1 encodes mitochondrial DNA polymerase and is one of the causative genes for a Mendelian-inheritance mitochondrial disease, which is occasionally accompanied by mood disorders. The healthy control data from the Tohoku Medical Megabank Organization were also employed.

Results: Although the frequency of carriers of deleterious variants varied from one method to another, every assessment achieved the same conclusion that deleterious POLG1 variants were significantly enriched in the variants identified in patients with bipolar disorder compared to those in controls.

Conclusion: Together with mitochondrial dysfunction in bipolar disorder, the present results
suggested deleterious POLG1 variants as a credible risk for the multifactorial disease.

**Key words:** bipolar disorder, mitochondrial dysfunction, POLG, POLG1, rare variants.

Bipolar disorder (BD) is a chronic and severe mental illness with a lifetime prevalence of 2.1% in the USA and 0.4% in Japan. Despite the strong heritability of BD, the underlying genetic architecture remains largely elusive. Genetic studies exploring common variants have identified several loci associated with BD; however, the odds ratios are not high enough to account for the entire architecture of BD. Targeted resequencing of specific genes and, recently, whole genome and exome sequencing have been exploited to identify rare variants, which may constitute a part of the architecture. In this study, on the strength of the hypothesis of mitochondrial dysfunction in BD and the high comorbidity of mood disorder in patients with mitochondrial diseases, we have focused on the POLG1 gene, which is one of the causative genes for chronic progressive external ophthalmoplegia (CPEO), a Mendelian-inheritance mitochondrial disease (OMIM 157640).

Although CPEO is a very rare disease (e.g., one in approximately 600,000 of the general population in Japan), dozens of mutations have been found mostly in the POLG1 gene in patients with autosomal dominant, autosomal recessive, or sporadic CPEO. POLG1 encodes the catalytic subunit of mitochondrial DNA (mtDNA) polymerase, which is the sole DNA polymerase of mtDNA and is expressed ubiquitously even in postmitotic cells. Thus, mutations in the POLG1 can have pleiotropic effects; CPEO-related mutations cause not only myopathy but also other somatic symptoms and a wide range of neuropsychiatric symptoms, including epilepsy, sensorimotor neuropathy, ataxia, Parkinsonism, and psychiatric diseases, such as major depressive disorder, BD, and panic disorder, but not schizophrenia or autism spectrum disorder.

More than 60 CPEO-related mutations are currently recorded in the Human DNA Polymerase gamma Mutation Database, however, some of them seem to be ambiguous. For example, Y831C (rs41549716, global minor allele frequency [MAF] = 0.0063) is listed not only as a CPEO-related mutation but also as a single nucleotide polymorphism (SNP), raising the question as to whether Y831C is deleterious or not. The Japanese Reference Panel Project (1070 individuals) of the Tohoku Medical Megabank (ToMMo) found that R964C (rs201477273) was the most frequent SNP in the POLG1 (MAF = 0.012), though it has been labeled as susceptible to a nucleoside analog reverse transcriptase inhibitor. In the era of whole genome and exome sequencing, it is becoming crucial to determine whether identified variants have deleterious impacts or are benign for complex disease studies. There are three different types of approach: a computational prediction, a biochemical method, and in vivo observation of knock-in mice or human subjects carrying certain variants. Each of them has both drawbacks and advantages. Computational prediction can quickly assess a number of variants, but the result is no more than prediction. In addition, it is problematic in practice to select one program out of a dozen in silico prediction programs. An in vitro biochemical method, which is laborious, apparently seems to produce correct results, but it is impossible to mimic the actual physiological condition. In contrast, in vivo observation has high reliability; however, generation of knock-in mice is more lengthy and costly, even if genome-editing technology is available. Human data are difficult to interpret due to individual variability, including genetic and environmental differences. This must lead to the difficulty in assessing the above-mentioned Y831C and R964C variants.

In this study, we examined the impacts of POLG1 variants identified in Japanese subjects on the POLG1 protein function and ideally on vulnerability for BD by the three different modalities of assessment, in silico prediction, biochemical assays, and clinical evaluation, which complementally support each other. We also present here a way to select most appropriate prediction programs for particular genes of interest.
METHODS

All experiments were approved by the Ethics Committees of RIKEN and all participating institutes. All participants gave informed consent and their anonymity was preserved. A detailed description of human subjects and methods, including mutation screening, computational prediction, preparation of recombinant POLG1 and POLG2 proteins, polymerase processivity, exonuclease assays, and statistics can be found in Appendix S1 in Supporting Information.

RESULTS

Identification of amino acid substitution variants in POLG1

We re-sequenced the entire coding sequence of POLG1 in 796 Japanese patients with BD (645 BD-I, 147 BD-II, and four schizoaffective disorder, bipolar type) and 767 controls. We identified 12 amino acid substitution variants (mutations or polymorphisms) in 23 patients and 15 variants in 26 controls (Table 1 and Fig. 1). All were heterozygotes. One control subject carried two variants (S731N and H734R). There was no substantial difference in variant frequency in patients and controls (2.9% vs 3.4%). A similar frequency of POLG1 variants was reported in the ToMMo data (40 variants in 1070 healthy Japanese subjects) (Table 1). Apparent loss-of-function variants, that is, nonsense or frame-shift variants, were found in one patient (E1046X) and one control (H927fs), and two CPEO-related mutations were found in patients (R562Q in one patient and R574W in two patients). We next investigated the impacts of the POLG1 variants that were identified in patients and controls using in silico prediction programs and in vitro biochemical methods.

In silico prediction: POLG1 variants found in BD are more deleterious

POLG1 shares conserved catalytic domains with bacterial DNA polymerase I and bacteriophage T7 DNA polymerase22 (Fig. 1), and the crystal structure of human POLG1 is available.23 It was highly likely that some prediction algorithms would give more correct predictions for POLG1 variants than others, depending on how the rich information about the POLG1 was dealt with. Thus, we first screened eight prediction algorithms and SNP (Fig. 2a). The accuracy (Matthews correlation coefficient [MCC]) of the binary classification was 0.395. The PolyPhen-2 program yielded a more accurate result (MCC = 0.554) in the classification test, when the threshold between ‘benign’ and ‘possibly damaging’ was used (Fig. 2b). It was predicted using the program that deleterious (or possibly damaging) variants of POLG1 were more frequent in the variants found in patients with BD than those found in controls plus ToMMo data (P = 0.026) (Fig. 2c).

In vitro assessment: Deleterious POLG1 variants are enriched in BD variants

We analyzed the biochemical consequences of all POLG1 variants identified in the Japanese patients and controls as well as several other variants, such as A467T (the most frequent CPEO mutation),33 Y955C (the best-studied CPEO mutation),34–36 and
Y831C (see Introduction section). Mutant POLG1 proteins, including the nonsense and frame-shift variants, were expressed in Sf9 cells and were purified almost to homogeneity using the C-terminal histidine tag. DNA polymerase activity (processivity) and 3’ → 5’ exonuclease activity were examined using a

| Variant | Base number | Exon substitution | BD (n = 796) | Control (n = 767) | ToMMo Control (n = 1070) | Functional domain | Variant ID of the dbSNP (with EAS_MAF) | Position (GRCh37/hg19) | Position (GRCh38/hg38) |
|---------|-------------|-------------------|-------------|-----------------|----------------------|------------------|---------------------------------|-----------------------|------------------------|
| Q43R    | 128         | 2 A/G             | 0           | 1               | 0                     |                  | rs28567406 (EAS_MAF = 0.005) | 89 876 858          | 89 333 627            |
| L112F   | 334         | 2 C/T             | 0           | 0               | 1                     |                  | rs777228398            | 89 876 652          | 89 333 421            |
| W113G   | 337         | 2 T/G             | 1           | 0               | 2                     |                  | UCV                | 89 876 649          | 89 333 418            |
| N134S   | 401         | 2 A/G             | 0           | 0               | 1                     |                  | UCV                | 89 876 585          | 89 333 354            |
| R177Q   | 530         | 2 G/A             | 0           | 0               | 0                     |                  | UCV                | 89 876 456          | 89 333 225            |
| A219D   | 656         | 2 C/A             | 1           | 0               | 0                     |                  | UCV                | 89 876 330          | 89 333 099            |
| Q238R   | 713         | 3 A/G             | 1           | 0               | 2                     |                  | rs56410699           | 89 873 454          | 89 330 223            |
| P241L   | 722         | 3 C/T             | 0           | 0               | 1                     |                  | UCV                | 89 873 445          | 89 330 214            |
| S332F   | 995         | 4 C/T             | 0           | 0               | 1                     |                  | UCV                | 89 872 202          | 89 328 971            |
| V353I   | 1057        | 5 G/A             | 0           | 0               | 1                     |                  | UCV                | 89 872 029          | 89 328 798            |
| D384Y   | 1150        | 5 G/T             | 0           | 0               | 1                     |                  | UCV                | 89 871 936          | 89 328 705            |
| P412S   | 1234        | 6 C/T             | 1           | 0               | 0                     |                  | UCV                | 89 871 703          | 89 328 472            |
| P412L   | 1235        | 6 C/T             | 1           | 0               | 1                     |                  | rs587780420         | 89 871 702          | 89 328 471            |
| S462L   | 1385        | 7 C/T             | 0           | 0               | 1                     |                  | rs762878459          | 89 870 446          | 89 327 215            |
| C471Y   | 1412        | 7 G/A             | 0           | 0               | 1                     |                  | UCV                | 89 870 419          | 89 327 188            |
| R562Q   | 1685        | 9 G/A             | 0           | 0               | 1                     |                  | rs781168350          | 89 869 870          | 89 326 639            |
| R574W   | 1720        | 10 C/T            | 2           | 0               | 0                     |                  | rs774474723          | 89 868 910          | 89 325 679            |
| R628W   | 1882        | 10 C/T            | 0           | 0               | 1                     |                  | UCV                | 89 868 748          | 89 325 517            |
| S731N   | 2192        | 13 G/A            | 0           | 0               | 1                     |                  | UCV                | 89 866 708          | 89 323 477            |
| S998L   | 2993        | 19 C/T            | 0           | 0               | 1                     |                  | UCV                | 89 866 120          | 89 322 889            |
| A824V   | 2471        | 15 C/T            | 0           | 0               | 3                     |                  | UCV                | 89 865 202          | 89 321 971            |
| R866W   | 2596        | 16 C/T            | 2           | 1               | 0                     |                  | rs748777396          | 89 864 969          | 89 321 738            |
| H927fs  | 2779        | 18 A/−            | 0           | 0               | 1                     |                  | UCV                | 89 864 199          | 89 320 968            |
| A1217V  | 3650        | 23 C/T            | 0           | 2               | 0                     |                  | rs199751339          | 89 860 052          | 89 316 821            |

1The functional domains of human POLG1 are defined by the crystal structure.23
2EAS_MAF values are retrieved from the 1000 Genomes Project data.
3One of control subjects carrying the H734R variant had S731N as well.
BD, bipolar disorder; CDS, coding sequence; EAS_MAF, East Asian minor allele frequency; ToMMo, Tohoku Medical Megabank Organization; UCV, uncharacterized variant (novel variant).
As stated, Human POLG1 protein harbors a polyglutamine tract, which is critical for its enzymatic activities, respectively. The domains share considerable sequence and 3-D structure homologies with bacterial DNA polymerase vs processivity, Spearman’s rank correlation between the three activities of each variant (exonuclease activity and polymerase activity) interacts with and binds to POLG2 dimer. Human POLG1 protein harbors a polyglutamine tract (polyQ) of 13 glutamines in the N-terminal region. The locations of all variants identified in patients with bipolar disorder (BD) and controls, in which the Tohoku Medical Megabank data are not included, are identified. Two chronic progressive external ophthalmoplegia-related mutations (R562Q and R574W) are indicated in bold, and two apparent loss-of-function variants (H927fs and E1046X) are in italics.

Figure 1. Human POLG1 protein structure and amino acid substitution variants identified in Japanese cases and controls. Human POLG1 protein contains the exonuclease and polymerase domains that possess 3' → 5' exonuclease and DNA polymerase activities, respectively. The domains share considerable sequence and 3-D structure homologies with bacterial DNA polymerase I and bacteriophage T7 DNA polymerase. The linker region locating in the polymerase domain (or roughly between the two domains) interacts with and binds to POLG2 dimer. Human POLG1 protein harbors a polyglutamine tract (polyQ) of 13 glutamines in the N-terminal region. The locations of all variants identified in patients with patients compared to those in controls (P = 0.039, U-test) (Fig. 3c). Because several variants recently identified in the ToMMo controls were not examined in the in vitro assays, we excluded the ToMMo data from the analysis.

We compared the biochemical activities of POLG1 variants with the in silico predictions, PROVEAN score, and PolyPhen-2 binary classification. Both programs significantly predicted the biochemical impacts of POLG1 variants (PROVEAN, P = 0.0083; PolyPhen-2, P = 0.014, U-test) (Fig. S4), which supported the validity of the application of these prediction programs to the POLG1 variant analysis (Fig. 2c,d). Either the biochemical analysis or in silico predictions demonstrated that most of the POLG1 variants located in the linker domain and the polymerase domain were likely deleterious (Table 2 and Fig. S3). These POLG1 variants were severely deprived of both the exonuclease and polymerase activities, in contrast to variants in the exonuclease domain, which affected only the exonuclease activity (Table 2).

Evaluation based on reports in CPEO patients: CPEO-related POLG1 mutations are found only in BD

In addition to the in silico and in vitro assessments, we estimated the in vivo consequences of POLG1 variants based on the reported clinical data. As stated
Figure 2. Legend on next page.
in the Introduction, patients with CPEO frequently harbor mutations in the POLG1 gene; these CPEO-related mutations are apparently deleterious in vivo. We found two CPEO-related mutations, R562Q and R574W, in three patients with BD (Table 1). Both the mutations were classified as deleterious by the in silico and in vitro methods (Table 2). The R562Q mutation, which was found in a patient with BD-I, rapid cycling, had been identified in a CPEO patient with myopathy\(^{36}\) (Table S2). The R574W mutation, which was found in two unrelated patients with BD-I, had been identified in a CPEO patient with dysphagia and myopathy.\(^ {14}\) None of CPEO-related mutations that were listed in the POLG1 mutation database\(^ {18}\) were found in healthy controls or the ToMMo data. Although the frequency of patients carrying CPEO-related mutations was very low (0.38%), CPEO-related mutations were significantly more frequent in patients with BD compared to controls plus ToMMo data (\(P = 0.028,\) Fisher’s exact test).

**DISCUSSION**

In the present study, we thoroughly characterized the 23 POLG1 variants identified in Japanese patients with BD and controls by using three approaches of analysis. Every analysis achieved the same conclusion that deleterious variants were enriched in the variants identified in patients with BD (Figs 2c,d and 3c). Together with the established observations, namely, mitochondrial dysfunction in BD\(^ {9,10}\) and high frequent comorbidity of BD in mitochondrial diseases,\(^ {11,12}\) the present results suggest deleterious variants in the POLG1 gene as a risk for BD.

A great number of rare variants have recently been identified with the development of sequencing and have been assessed principally using in silico prediction programs. It is difficult to select one program out of several prediction algorithms with distinct characteristics. In this study, we selected the best algorithms for POLG1 variant prediction in an unbiased manner (Fig. 2a) and assessed variants we identified using the best algorithms. This procedure will be effective when variants of particular genes are assessed in silico. Computational algorithms predict deleteriousness of variants based on various kinds of information, such as phylogenetic similarity among homologs (orthologs and paralogs), 3-D structures, and epidemiological, clinical, and genetic characteristics of the variants. The width and depth of the information greatly vary among genes and how to handle the information also varies among algorithms, and as a result, variants in certain genes can be predicted more correctly by certain algorithms. In the case of POLG1 variants, although in silico predictions were significantly correlated with the in vitro assessment (Fig. S4), the prediction algorithms likely gave more deleterious scores than the biochemical assessment (Table 2 and Fig. S4). We cannot tell whether the in silico prediction algorithms overevaluated the impacts of POLG1 variants or the biochemical assessment underevaluated them, because the in silico method is still immature and also because in vitro experiments cannot mimic the physiological condition and cannot test as-yet-unrecognized activities or functions. Considering that both the methods robustly detected deleterious impacts of CPEO-related mutations (Table 2), however, we could argue that the combined assessment with three different modalities, in silico, in vitro, and in vivo (or clinical data survey), was favorable and fruitful.

---

**Figure 2.** Computational predictions of deleteriousness of POLG1 variants. (a) Performance of eight prediction programs. Deleteriousness of 60 known chronic progressive external ophthalmoplegia (CPEO)-related mutations and 19 single nucleotide polymorphisms (SNPs; Table S1) were predicted and shown in violin plots with a marker for the median and a gray box indicating the interquartile range. Dashed lines indicate the cut-off thresholds, which prediction algorithms set to classify deleterious and benign variants. (b) Statistical result of the performance of prediction programs. Values in bold face indicate the highest performance. Rank–sum value (U-test method) or accuracy (Matthews correlation coefficient [MCC]) is used to select the best program in a continuous score-based test or a binary classification test, respectively. PolyPhen-2 has two cut-off thresholds of classification: between ‘benign’ and ‘possibly damaging’ and between ‘possibly damaging’ and ‘probably damaging.’ The MCC value of 0.554 was obtained when the cut-off threshold between ‘benign’ and ‘possibly damaging’ was used. (c) PROVEAN prediction. POLG1 variants found in patients are significantly more deleterious than controls (\( \because P < 0.05, U\)-test). The area of each dot graphically represents a number of subjects who carry the corresponding variant. (d) PolyPhen-2 prediction. Deleterious variants are more enriched in the POLG1 variants identified in patients than those in controls (\( \because P < 0.05, \# P = 0.1,\) one-tailed Fisher’s exact test). ToMMo, Tohoku Medical Megabank.
We found two CPEO-related mutations, R562Q and R574W, in three patients with BD. The R562Q had been identified as a heterozygous mutation in a sporadic patient with CPEO\textsuperscript{38} and the R574W was found as a compound heterozygote mutation.\textsuperscript{14} Although CPEO is a Mendelian genetic disease, the inheritance pattern of POLG1 mutations is not simple; some mutations represent a dominant trait or a recessive trait of CPEO, and sometimes also lead to other diseases, such as Alpers syndrome, encephalopathy, epilepsy (and probably BD). The complex genetics involving the POLG1 gene may occur for not only CPEO and the other POLG1-related diseases but also for BD. The R562Q and
Table 2. Functional analyses of POLG1 variants

| Variant | Functional domain w/ POLG2 | w/o POLG2 | POLYmerase domain | w/ POLG2 | w/o POLG2 | Overall evaluation | PROVEAN score | PolyPhen-2 | BD prediction (n = 796) | Control (n = 767) | Note |
|---------|-----------------------------|-----------|-------------------|-----------|-----------|-------------------|---------------|-------------|------------------------|------------------|------|
| Q43R    | +                           | ++        | ++                | ++        | ++        | +                 | 0.19          | Benign      | 0                      | 1                | rs28567406 (EAS_MAF = 0.005) |
| L112F   | +++                         | ++        | ++                | ++        | ++        | −2.83             | Benign        | 0            | 0                      | UCV              | (unpublished) |
| W113G   | +                           | +         | +                 | +         | +         | −5.12             | Benign        | 1            | 0                      | UCV              |          |
| N134K   | +++                         | ++        | ++                | ++        | ++        | −3.47             | Benign        | −            | −                      | UCV in a non-Japanese patient (unpublished) |
| R177Q   | +                           | +         | +                 | +         | +         | −2.32             | Benign        | 0            | 1                      | UCV              |          |
| D198A   | −                           | −         | −                 | −         | −         | −7.19             | Deleterious   | −            | −                      | −                | Exonuclease-deficient mutation (artificial) |
| A219D   | NE                          | NE        | NE                | NE        | ++        | −2.51             | Benign        | 1            | 0                      | UCV              |          |
| Q238R   | +++                         | +++       | +++               | +++       | +         | −1.02             | Benign        | 1            | 2                      | ns56410699       |          |
| P241L   | +++                         | +++       | +++               | +++       | +         | 2.45              | Benign        | 0            | 1                      | Found in patients with Parkinsonism |
| G268A   | ++                          | +         | +                 | +         | +         | −5.19             | Deleterious   | −            | −                      | CPEO-related mutation in patients with Alpers' disease or others |
| H277L   | +                           | +         | +                 | +         | +         | −2.18             | Deleterious   | −            | −                      | −                |          |
| S332F   | +++                         | +++       | +++               | +++       | +         | −1.64             | Benign        | 0            | 1                      | UCV              |          |
| P412S   | +                           | +         | +                 | +         | +         | −6.99             | Benign        | 1            | 0                      | UCV              |          |
| P412L   | +                           | +         | +                 | +         | +         | −8.85             | Benign        | 1            | 0                      | ns587780420     |          |
| S462L   | +++                         | +++       | +++               | +++       | +         | −0.75             | Benign        | 0            | 1                      | ns762878459      |          |
| A467T   | −                           | −         | −                 | −         | −         | −3.27             | Benign        | −            | −                      | −                | ns113994095; Most frequent CPEO-related mutation (gMAF = 0.00051) |
| R562Q   | −                           | ±         | −                 | −         | −         | −2.01             | Deleterious   | 1            | 0                      | ns781168350; CPEO-related mutation (unpublished) |
| R574W   | NE                          | NE        | NE                | NE        | −         | −6.39             | Deleterious   | 2            | 0                      | ns74474723; CPEO-related mutation (unpublished) |
| S731N   | ±                           | ±         | +                 | +         | +         | 0.45              | Benign        | 0            | 1                      | UCV              |          |
| H734R   | ±                           | ±         | +                 | +         | ±         | −0.84             | Benign        | 7            | 8                      | ns56119329       |          |
| K731R   | ±                           | ±         | ±                 | +         | ±         | −3.15             | Benign        | −            | −                      | UCV in a non-Japanese patient (unpublished) |
| Y831C   | +                           | +         | +                 | +         | +         | −2.82             | Deleterious   | −            | −                      | −                | ns1549716; CPEO-related, but SNP (gMAF = 0.0063) |
| R866W   | −                           | −         | −                 | ±         | +         | −4.31             | Deleterious   | 2            | 1                      | ns748777396     |          |
| R927fs  | NE                          | NE        | NE                | NE        | NE        | −                | −             | 0            | 1                      | UCV              |          |
| H932Q   | +++                         | +++       | +++               | +++       | +         | −7.91             | Deleterious   | 1            | 0                      | UCV              |          |
| Y955C   | +++                         | +++       | +++               | +++       | +         | −8.77             | Deleterious   | −            | −                      | −                | ns113994099; Best-studied CPEO mutation (unpublished) |
| R964C   | ±                           | ±         | ±                 | ±         | ±         | −5.83             | Deleterious   | 4            | 3                      | −                | ns201477273 (EAS_MAF = 0.013) |
| S998L   | +                           | +         | +                 | +         | +         | −2.21             | Deleterious   | 0            | 1                      | ns79840247      |          |
| E1046X  | NE                          | NE        | NE                | NE        | NE        | −                | −             | 1            | 0                      | UCV              |          |
| T1119R  | −                           | −         | −                 | −         | −         | −5.52             | Deleterious   | 0            | 1                      | −                | UCV |
| A1217V  | +                           | +         | +                 | +         | +         | −2.85             | Deleterious   | 0            | 2                      | ns199751339 (EAS_MAF = 0.00058) |          |

1Activities of POLG1 variants are graded in five levels (−, ±, +, ++, and +++), where activity of wild-type POLG1 is ‘++’. 
2Recombinant triple mutant proteins containing D198A and E200A as well as each respective variant were prepared and assayed. 
3Overall evaluation of the biochemical activity is graded in four levels (−, ±, +, and ++). 

EAS_MAF values are retrieved from the 1000 Genomes Project data. 
gMAF values are retrieved from the ExAC data aggregated from multiple populations. 
BD, bipolar disorder; EAS_MAF, East Asian minor allele frequency; gMAF, global minor allele frequency; NE, not fully expressed; UCV, uncharacterized variant (novel variant).
R574W found in patients with BD were not tightly segregated in each family (Table S2). In addition, the frequency of carriers of deleterious variants of POLG1 in BD was low: CPEO-related mutations, 0.38%; deleterious variants biochemically defined, 2.3%; variants predicted as deleterious, 2.4%. However, CPEO-related mutations of POLG1 were significantly enriched in patients with BD (Table 2), suggesting that they confer a risk for BD. In contrast, genome-wide association studies on BD have not detected any significant signals in the POLG1 region, which indicates that there is no common functional polymorphism that is associated with BD. We assume that BD has genetic architectures constructed with common SNPs, copy number variants, and rare deleterious variants.

Given that deleterious variants of POLG1 constitute a part of the genetic architecture of BD, the brain circuit architecture is then of much interest. We have recently reported a transgenic mouse line, in which an exonuclease-deficient mutant of POLG1 (D198A in human POLG1) is expressed exclusively in the forebrain neurons, spontaneously exhibiting recurrent mood episodes. The mutant mice and post-mortem brains of CPEO patients with mood symptoms highlighted the paraventricular nucleus of thalamus (PVT), and PVT-specific inhibition of neural transmission triggered depression-like behavioral change in mice. The PVT plays a prominent role in stress and fear responses and participates in circadian behaviors that are relevant to BD. It seems reasonable to conjecture a converging path from deleterious variants in POLG1 into the final common pathway of BD as follows: deleterious POLG1 variants, then resultant mtDNA deletions and mitochondrial dysfunction in the PVT, and then impaired neural circuit(s) involving PVT, which would be close to the final common pathway. Further studies will be needed to elucidate the pathophysiology of BD.

ACKNOWLEDGMENTS

We wish to thank all volunteers for participating in this study. We are grateful to staff members in the Research Resources Center, RIKEN Brain Science Institute for technical assistance, and members in our laboratory for valuable discussions. We are also grateful to Ms Mizue Kametani for technical assistance. We thank Professor Shin-ichiro Nanko and Dr Koji Matsuo for providing DNA samples and Professor Anu Suomalainen-Wartiovaara for valuable discussion. This work was supported by the Japan Agency for Medical Research and Development (T. Kato), Grants-in-Aid from the Japanese Ministry of Education, Culture, Sports, Science and Technology (T. Kato) and an Independent Investigator Award from NARSAD (T. Kato).

DISCLOSURE STATEMENT

There is no conflict of interest.

AUTHOR CONTRIBUTIONS

The study was designed by T. Kasahara and T. Kato. T. Kasahara performed bioinformatics analysis on the impact of variants. T. Kasahara and M.I. performed biochemical assays. C.K., S.F., and T. Kato performed resequencing. A.T. and N.M. analyzed exome data. H.K., N.O., N.I., Y.M., K.N., Y.I., K.F., S.K., H.U., I.K., M.K., T.Y., and T. Kato collected DNA samples. T.Y. organized the COSMO. K.I. performed SNP chip analysis. T. Kasahara and T. Kato wrote the manuscript.

REFERENCES

1. Merikangas KR, Jin R, He JP et al. Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. Arch. Gen. Psychiatry 2011; 68: 241–251.
2. Kawakami N. Large-scale epidemiological study on the prevalence of mental disorders: World Mental Health Japan Survey. A report of the Health Labour Sciences Research Grant from The Ministry of Health Labour and Welfare. 2014 (in Japanese).
3. Kato T. Whole genome/exome sequencing in mood and psychotic disorders. Psychiatry Clin. Neurosci. 2015; 69: 65–76.
4. Kerner B. Toward a deeper understanding of the genetics of bipolar disorder. Front. Psychiatry 2015; 6: 105.
5. Chen DT, Jiang X, Akula N et al. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. Mol. Psychiatry 2013; 18: 195–205.
6. Hou I, Bergen SE, Akula N et al. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. Hum. Mol. Genet. 2016. doi: 10.1093/hmg/ddw181
7. Ament SA, Szelering S, Glusman G et al. Rare variants in neuronal excitability genes influence risk for bipolar disorder. Proc. Natl. Acad. Sci. U. S. A. 2015; 112: 3576–3581.
8. Kataoka M, Matoba N, Sawada T et al. Exome sequencing for bipolar disorder points to roles of de novo loss-of-
function and protein-altering mutations. Mol. Psychiatry 2016; 21: 885–893.
9. Kato T. Mitochondrial dysfunction as the molecular basis of bipolar disorder: Therapeutic implications. CNS Drugs 2007; 21: 1–11.
10. Scaini G, Rezin GT, Carvalho AF, Streck EL, Berk M, Quevedo J. Mitochondrial dysfunction in bipolar disorder: Evidence, pathophysiology and translational implications. Neurosci. Biobehav. Rev. 2016; 68: 694–713.
11. Fattal O, Link J, Quinn K, Cohen BH, Franco K. Psychiatric comorbidity in 36 adults with mitochondrial cytopathies. CNS Spectr. 2007; 12: 429–438.
12. Smits BW, Fermont J, Delnooz CC, Kalkman JS, Bleijenberg G, van Engelen BG. Disease impact in chronic progressive external ophthalmoplegia: More than meets the eye. Neuromuscul. Disord. 2011; 21: 272–278.
13. Kurihara T. Mitochondrial neurogastrointestinal encephalomyopathy and its pathophysiology. Intern. Med. 2006; 45: 415–416.
14. Horvath R, Hudson G, Ferrari G et al. Phenotypic spectrum associated with mutations of the mitochondrial polymerase γ gene. Brain 2006; 129: 1674–1684.
15. Stumpf JD, Saneto RP, Copeland WC. Clinical and molecular features of POLG-related mitochondrial disease. Cold Spring Harb. Perspect. Biol. 2013; 5: a011395.
16. Krishnan KJ, Reeve AK, Samuels DC et al. What causes mitochondrial DNA deletions in human cells? Nat. Genet. 2008; 40: 275–279.
17. Longley MJ, Graziiewicz MA, Bienstock RJ, Copeland WC. Consequences of mutations in human DNA polymerase γ. Gene 2005; 354: 125–131.
18. Copeland B. Human DNA polymerase gamma mutation database. [Cited 15 November 2016.] Available from URL: http://tools.niehs.nih.gov/polg/
19. Nagasaki M, Yasuda J, Katsuoka F et al. Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. Nat. Commun. 2015; 6: 8018.
20. Yamanaka H, Gatanaga H, Kosalaraksa P et al. Novel mutation of human DNA polymerase γ associated with mitochondrial toxicity induced by anti-HIV treatment. J. Infect. Dis. 2007; 195: 1419–1425.
21. Peterson TA, Doughty E, Kann MG. Towards precision medicine: Advances in computational approaches for the analysis of human variants. J. Mol. Biol. 2013; 425: 4047–4063.
22. Kaganis LS. DNA polymerase γ, the mitochondrial replicase. Annu. Rev. Biochem. 2004; 73: 293–320.
23. Lee YS, Kennedy WD, Yin YW. Structural insight into processive human mitochondrial DNA synthesis and disease-related polymerase mutations. Cell 2009; 139: 312–324.
24. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 2014; 46: 310–315.
25. González-Pérez A, López-Bigas N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. Am. J. Hum. Genet. 2011; 88: 440–449.
26. Shihab HA, Gough J, Cooper DN et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum. Mutat. 2013; 34: 57–65.
27. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: Application to cancer genomics. Nucleic Acids Res. 2011; 39: e118.
28. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: Expanded protein families and functions, and analysis tools. Nucleic Acids Res. 2016; 44: D336–D342.
29. Azdzhubei IA, Schmidt S, Peshkin L et al. A method and server for predicting damaging missense mutations. Nat. Methods 2010; 7: 248–249.
30. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One 2012; 7: e46688.
31. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003; 31: 3812–3814.
32. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015; 526: 68–74.
33. Chan SS, Longley MJ, Copeland WC. The common A467T mutation in the human mitochondrial DNA polymerase (POLG) compromises catalytic efficiency and interaction with the accessory subunit. J. Biol. Chem. 2005; 280: 31341–31346.
34. Baruffini E, Lodi T, Dallabona C, Puglisi A, Zeviani M, Ferrero I. Genetic and chemical rescue of the Saccharomyces cerevisiae phenotype induced by mitochondrial DNA polymerase mutations associated with progressive external ophthalmoplegia in humans. Hum. Mol. Genet. 2006; 15: 2846–2855.
35. Lewis W, Day BJ, Kohler JJ et al. Decreased mtDNA, oxidative stress, cardiomyopathy, and death from transgenic cardiac targeted human mutant polymerase γ. Lab. Invest. 2009; 87: 326–335.
36. Estep PA, Johnson KA. Effect of the Y955C mutation on mitochondrial DNA polymerase nucleotide incorporation efficiency and fidelity. Biochemistry 2011; 50: 6376–6386.
37. Longley MJ, Ropp PA, Lim SE, Copeland WC. Characterization of the native and recombinant catalytic subunit of human DNA polymerase γ: Identification of residues critical for exonuclease activity and deoxyribonucleotide sensitivity. Biochemistry 1998; 37: 10529–10539.
38. Di Fonzo A, Bordoni A, Crimi M et al. POLG mutations in sporadic mitochondrial disorders with multiple mtDNA deletions. Hum. Mutat. 2003; 22: 498–499.
39. Kasahara T, Takata A, Kato TM et al. Depression-like episodes in mice harboring mtDNA deletions in paraventricular thalamus. Mol. Psychiatry 2016; 21: 39–48.
40. Bubser M, Deutch AY. Stress induces Fos expression in neurons of the thalamic paraventricular nucleus that innervate limbic forebrain sites. *Synapse* 1999; 32: 13–22.

41. Hsu DT, Price JL. Paraventricular thalamic nucleus: Subcortical connections and innervation by serotonin, orexin, and corticotropin-releasing hormone in macaque monkeys. *J. Comp. Neurol.* 2009; 512: 825–848.

42. Penzo MA, Robert V, Tucciarone J et al. The paraventricular thalamus controls a central amygdala fear circuit. *Nature* 2015; 519: 455–459.

43. Salazar-Juárez A, Escobar C, Aguilar-Roblero R. Anterior paraventricular thalamus modulates light-induced phase shifts in circadian rhythmicity in rats. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 2002; 283: R897–R904.

44. Murray G, Harvey A. Circadian rhythms and sleep in bipolar disorder. *Bipolar Disord.* 2010; 12: 459–472.

**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Figure S1.** Processivity assay of POLG1 mutants.

**Figure S2.** Exonuclease activity assay of POLG1 mutants.

**Figure S3.** Polymerase activity assay of POLG1 mutants.

**Figure S4.** Comparison of *in vitro* assessment of POLG1 variants with *in silico* predictions.

**Figure S5.** Deleterious scores of all possible single nucleotide polymorphisms (SNPs) of POLG1.

**Table S1.** POLG1 variants used for screening of prediction algorithms.

**Table S2.** Bipolar patients carrying chronic progressive external ophthalmoplegia (CPEO)-related POLG1 mutations

**Appendix S1.** Supporting methods.

**Appendix S2.** Supporting references for Appendix S1 and Table S1.