

A linkage and exome study of multiplex families with bipolar disorder implicates rare coding variants of *ANK3* and additional rare alleles at 10q11-q21

Claudio Toma, PhD; Alex D. Shaw, PhD; Anna Heath BSc; Kerrie D. Pierce, BSc; Philip B. Mitchell, MBBS, MD; Peter R. Schofield, PhD, DSc; Janice M. Fullerton, PhD

Background: Bipolar disorder is a highly heritable psychiatric condition for which specific genetic factors remain largely unknown. In the present study, we used combined whole-exome sequencing and linkage analysis to identify risk loci and dissect the contribution of common and rare variants in families with a high density of illness. **Methods:** Overall, 117 participants from 15 Australian extended families with bipolar disorder (72 with affective disorder, including 50 with bipolar disorder type I or II, 13 with schizoaffective disorder–manic type and 9 with recurrent unipolar disorder) underwent whole-exome sequencing. We performed genome-wide linkage analysis using MERLIN and conditional linkage analysis using LAMP. We assessed the contribution of potentially functional rare variants using a gene-based segregation test. **Results:** We identified a significant linkage peak on chromosome 10q11-q21 (maximal single nucleotide polymorphism = rs10761725; exponential logarithm of the odds [LOD_{exp}] = 3.03; empirical p = 0.046). The linkage interval spanned 36 protein-coding genes, including a gene associated with bipolar disorder, ankyrin 3 (*ANK3*). Conditional linkage analysis showed that common *ANK3* risk variants previously identified in genome-wide association studies — or variants in linkage disequilibrium with those variants — did not explain the linkage signal (rs10994397 LOD = 0.63; rs9804190 LOD = 0.04). A family-based segregation test with 34 rare variants from 14 genes under the linkage interval suggested rare variant contributions of 3 brain-expressed genes: *NRBF2* (p = 0.005), *PCDH15* (p = 0.002) and *ANK3* (p = 0.014). **Limitations:** We did not examine non-coding variants, but they may explain the remaining linkage signal. **Conclusion:** Combining family-based linkage analysis with next-generation sequencing data is effective for identifying putative disease genes and specific risk variants in complex disorders. We identified rare missense variants in *ANK3*, *PCDH15* and *NRBF2* that could confer disease risk, providing valuable targets for functional characterization.

Introduction

Bipolar disorder is a complex psychiatric disorder with a population prevalence of approximately 1%.¹ It is typically characterized by cyclic manic or hypomanic and depressive affect.² Genetic factors are strongly implicated in the disease etiology, with heritability estimates of approximately 70%.³ Genome-wide association studies (GWAS) indicate that 20% to 30% of the genetic liability for bipolar disorder is due to common genetic variants,⁴ but a significant proportion of the genetic risk factors remain unknown. Overall, GWAS have been an effective approach for identifying susceptibility genes in bipolar disorder, and common susceptibility variants in *CACNA1C*, *ANK3*, *ODZ4* and other genes have been implicated.^{5–9} Recent sequencing studies also suggest a role for rare variants in disease pathogenesis;^{10–12} thus, the genetic

model in bipolar disorder is likely to comprise both common and rare variants, the latter with potential for higher penetrance effects.¹² Despite recent advances in high-throughput technologies, many individual risk alleles remain unknown, but they are potentially tractable in large multiplex families with multiple affected relatives.¹³

Linkage studies have been systematically applied to map disease genes for Mendelian diseases, and they represent a powerful and successful method when a single gene carries etiologic mutations. However, linkage approaches for complex disorders have not achieved similar success: replication of linkage signals among independent studies has typically failed, and follow-up association studies have been largely unable to reveal susceptibility variants in linkage regions, because of substantial heterogeneity both at the allelic and locus levels and across cohorts.

Correspondence to: J. M. Fullerton, Neuroscience Research Australia, Margarete Ainsworth Building, Barker Street, Randwick, Sydney, NSW 2031, Australia, j.fullerton@neura.edu.au; C. Toma, Centro de Biología Molecular “Severo Ochoa,” Universidad Autónoma de Madrid/CSIC, C/Nicolás Cabrera, 1, Campus UAM, 28049 Madrid, Spain, claudio.toma@cbm.csic.es

Submitted Apr. 29, 2020; Revised Sept. 18, 2020; Accepted Oct. 12, 2020

DOI: 10.1503/jpn.200083

More recently, whole-exome sequencing (WES) or whole-genome sequencing (WGS) have become essential tools for the identification of rare pathogenic variants in multiplex families with psychiatric diseases. In bipolar disorder, WGS or WES in multiplex or extended families have implicated pathogenic single nucleotide variants and rare truncating mutations.^{10–12,14–20} However, because of diverse family structures and the relatively small sample sizes thus far employed, results from these individual studies are not currently converging onto the same pool of genes.¹³

Although large family studies have been mostly abandoned in favour of GWAS over the past decade, they have re-emerged as a powerful alternative approach for gene discovery, particularly for mapping rare variants with relatively large effect sizes.¹³ Recently, studies in recessive or dominant Mendelian diseases have adopted a linkage analysis using WES- or WGS-derived genotypes, followed by rare variant analysis and the identification of the etiologic variant.^{21,22} This approach has only recently been applied to some psychiatric disorders, in which mostly single families have been examined for pedigree-specific linkage^{15,19,23–28} (although a study in attention-deficit/hyperactivity disorder also employed a linkage approach with WES in 3 families comprising 10 relatives²⁹); thus, the strength of this strategy for unravelling missing heritability is yet to be clarified. While a number of studies have performed rare variant analysis across the genome, rare variant analyses in specific linkage regions has been performed in only 3 independent studies of bipolar disorder, each examining a single extended family: (1) a large Old Order Amish pedigree with single nucleotide polymorphism (SNP) genotypes for linkage analysis of 388 family members and WGS of 50 people with bipolar disorder;¹⁵ (2) a family of 6 relatives with bipolar disorder with WES and SNP data;¹⁹ and (3) a 4-generation family with 11 patients with bipolar disorder examined using linkage and WES.²⁶ These studies identified significant linkage regions that bore potential rare etiologic variants, although they did not identify highly penetrant rare variants that would entirely explain the linkage peaks,¹² suggesting that polygenic inheritance also plays a role in these extended families.

In the present study, we aimed to extend our previously published work, which focused on rare-variant sharing in families with bipolar disorder.¹² We performed a genome-wide linkage analysis across 15 extended families (with an average of 4.2 relatives affected with bipolar disorder type I or schizoaffective disorder–manic type) to identify major susceptibility loci and simultaneously map common and rare pathogenic variants from WES that contribute to susceptibility to bipolar disorder.

Methods

Ethics approval

All experiments were carried out in accordance with regulations and guidelines approved by the University of New South Wales Human Research Ethics Committee (initial approval HREC04144; extensions HREC10078, HC15503,

HC16347). We obtained written informed consent from all participants.

Participants

Fifteen families with a high density of affective illness over 2 to 4 generations were selected for WES (Appendix 1 Figure S1, available at jpn.ca/200083-a1).¹² The sequenced sample comprised 117 participants: 44 were diagnosed with bipolar disorder I (31 female), 6 with bipolar disorder II (5 female), 13 with schizoaffective disorder–manic type (7 female) and 9 with recurrent unipolar depression (8 female); 45 were unaffected or their status was unknown (25 female; Appendix 1, Table S1). All diagnoses were ascertained using the Family Interview for Genetic Studies³⁰ and the Diagnostic Interview for Genetic Studies.³¹ Additional clinical and demographic information is as previously described.¹²

WES-derived genotypes and linkage analysis procedures

Exome enrichment, template sequencing and variant calling were performed at the Lotterywest State Biomedical Facility (Perth, Australia). AmpliSeq exome enrichment and Ion Proton sequencing were performed as previously described:¹² 92% of bases captured were covered by more than 10 reads, and the mean read depth was 112×. Genotypes were called from aligned WES reads using SAMtools pileup (<https://github.com/samtools/samtools>). The WES-derived genotypes were filtered to identify independent haplotype-informative markers (HapMap CEU population) using LINKDATAGEN³² for input into linkage analysis, which employed 4044 WES-derived SNPs from 22 autosomes and the X chromosome across all 15 families. Family relationships were confirmed with identity-by-descent analysis using independent WES-derived genotypes in PLINK.³³ Nonparametric linkage analyses were performed using the “all” statistic in MERLIN, under the Kong and Cox linear and exponential models.³⁴

Because of the high frequency of depression in the general population and to reduce genetic heterogeneity, for linkage analysis we differentiated between transmitting and nontransmitting parents with recurrent unipolar depression. People with transmitting recurrent unipolar depression were those who reached the threshold criteria for a diagnosis of recurrent unipolar depression and who had 1 offspring or more with bipolar disorder I or schizoaffective disorder–manic type. Those who reached the threshold criteria for a diagnosis of recurrent unipolar depression but who had no descendants, or whose descendants were not diagnosed with bipolar disorder I or schizoaffective disorder–manic type, were considered “nontransmitting” and were conservatively coded as “unknown” rather than “affected.” Thus, we coded 60 participants as “affected”: those with a diagnosis of severe bipolar disorder (bipolar disorder I or schizoaffective disorder–manic type; $n = 57$), or parents with recurrent unipolar depression who had affected offspring ($n = 3$; Appendix 1, Table S1). Participants with all other diagnoses were coded as “unknown,” and those who

did not reach any diagnostic criteria were coded as “unaffected” (Appendix 1, Figure S1).

Refining intervals and empirical significance of linkage

To reduce the inter-marker interval to less than 1 cM across the maximal linkage peak at chromosome 10q11-q21, we selected 5 additional SNPs with high heterozygosity in Caucasian Europeans between rs2271128 and rs2273773 (85.4–90.8 cM): rs7073837 (86.217 cM), rs1061259 (87.263 cM), rs2105702 (89.187 cM), rs4548513 (89.5 cM) and rs7896005 (90.717 cM). These markers were not in linkage disequilibrium (LD; $r^2 < 0.03$), based on the genotypes of 405 non-Finnish Europeans (www.internationalgenome.org).

After linkage interval refinement, we computed the relative family contribution to the overall linkage using the *-perFamily* option in MERLIN.

We calculated the empirical significance of linkage using the *-simulate* option in MERLIN, in which input genotype data were replaced with simulated chromosomes conditional on family structure and phenotypes, generating 10 000 simulated data sets. The maximum logarithm of the odds (LOD) for linkage from each simulation was identified, and the empirical significance was defined by the number of times the observed maximum LOD exceeded the simulated maximum LOD, divided by the total number of simulations. A significance threshold was set at empirical $p < 0.05$.

Conditional linkage analysis and conditional SNP selection

Because significant associations have been reported between bipolar disorder and a number of common SNP variants in the ankyrin 3 (*ANK3*) gene,^{5–9,35,36} we performed conditional linkage analysis to assess whether variants commonly associated with bipolar disorder were responsible for the 10q11-q21 linkage peak. We performed an assessment of LD between previously reported GWAS-associated variants using genotype data from Caucasian Europeans (www.internationalgenome.org) to identify independent risk variants (Appendix 1, Figure S2). We selected 2 independent risk variants of *ANK3* for conditional linkage analysis: rs10994397 (intron 2) and rs9804190 (intron 36).^{6,8,36} These non-coding variants were not covered by WES amplicons, so we imputed genotypes from Illumina PsychChip data obtained for all 117 relatives. The imputation procedure followed ENIGMA2 genetics protocols (<http://enigma.ini.usc.edu/>) and used the Michigan Imputation Server (<https://imputationserver.sph.umich.edu/index.html>). Imputed genotypes had high-quality metrics (rs9804190 $R_{sq} = 0.811$; rs10994397 $R_{sq} = 0.999$), and resulting allele frequencies were as expected for non-Finnish Europeans (rs10994397 minor allele frequency [MAF] = 0.061, non-Finnish Europeans = 0.054; rs9804190 MAF = 0.235, non-Finnish Europeans = 0.225); genotypes were in Hardy–Weinberg equilibrium. Imputed genotypes were 100% concordant with direct TaqMan SNP genotyping (Applied Biosystems; rs10994397 assay ID:C_31344967_10) for all 117 participants, confirming imputation accuracy.

We performed conditional linkage using LAMP software³⁷ to evaluate whether the observed linkage signal was ex-

plained by GWAS risk alleles. The LAMP analysis framework is based on a parametric single-locus penetrance model for the trait; likelihood is maximized over a range of penetrance models and family structures, so it is not dependent on a specific model,³⁷ unlike the Kong and Cox single parameter framework, which is based on identity-by-descent sharing statistics. Further, LAMP quantifies the degree of LD between the candidate SNP and putative disease locus through joint modelling of linkage and association to determine whether the SNP and the disease locus are (1) in linkage equilibrium (so that the SNP plays no causal role in the linkage signal) or (2) in complete LD (so that the SNP or a marker in complete LD with it may account fully for the linkage signal). LAMP provides 3 likelihood ratio tests of the null hypothesis, and LOD scores are generated as follows: (1) a test for linkage; (2) a test for association in presence of linkage (i.e., that a candidate variant is in complete LD with the disease variant); and (3) a test for other linked variants. We performed conditional linkage analysis considering 17 SNPs, which spanned a 1-LOD drop interval after fine-mapping (listed in Appendix 1, Table S2).

Family-based association analysis for rare variants across 10q11-q21

We extracted rare variants for all coding genes within the 95% confidence interval of the chromosome 10q11-q21 linkage peak for all individuals across the 15 pedigrees. Only variants with a MAF less than 5% were included, after examination of allele frequencies from the Exome Variant Server (<http://evs.gs.washington.edu/EVS/>), ExAC (<http://exac.broadinstitute.org>), dbSNP138 (www.ncbi.nlm.nih.gov/projects/SNP) and the 1000 Genomes Phase I integrated call set (www.internationalgenome.org). All variants were checked for adequate read coverage ($> 10\times$) and consistency with phased haplotypes in each family to exclude potential artifacts. We predicted variant pathogenicity using Variant Effect Predictor annotation tool software (www.ensembl.org/Tools/VEP), which incorporates both PolyPhen and SIFT.^{38,39} However, because the detected exonic variants could be in LD with a causal variant but be nonpathogenic themselves, we examined all rare variants regardless of predicted pathogenicity. This selection identified 34 rare variants, including all missense, indels or stop mutations.

Next, we performed a family-based association test, using the gene-based segregation test (GESE) package implemented in R (<https://cran.r-project.org/web/packages/GESE/>) and further described in Appendix 1. The parameters for GESE analysis were as follows: segregation of rare variants was assessed in people diagnosed with bipolar disorder I and schizoaffective disorder–manic type (bipolar disorder II and recurrent unipolar depression were considered unknown, as per nonparametric linkage parameters); the population frequency of each of the 34 rare variants was included in an allele-frequency weighting; p values were calculated after 100 000 simulations; and per-family weights were included based on a combination of the ratio of severely

affected patients (bipolar disorder I and schizoaffective disorder–manic type) per family and diagnosis age at onset, to include a family burden of illness variable.

Genes emerging from the association test were assessed for evidence of brain expression, defined from RNA sequencing data from the developmental transcriptomics project (www.brainspan.org/rnaseq/search/index.html), where genes with reads per kilobase per million (RPKM) greater than 1 at any age and in any brain region were defined as brain-expressed (35% of all genes).

Finally, we identified all rare variants in *ANK3* from WES regardless of variant type or pathogenicity (intronic, indels, synonymous or missense with MAF < 5%) and performed a segregation study for each variant in the corresponding family. All rare variants in *ANK3* reported here were validated by Sanger sequencing.

Results

Linkage analysis

We performed a genome-wide nonparametric linkage analysis using the WES-derived genotype data of 117 people from 15 extended bipolar disorder pedigrees, including 60 severely affected individuals with diagnoses of bipolar disorder I or schizoaffective disorder–manic type, plus their transmitting parents with recurrent unipolar depression (Appendix 1, Figure S1). The highest peak LOD score was on chromosome 10q11–q21, with an exponential LOD of 2.863 ($p = 1.347 \times E^{-4}$) and a linear LOD of 2.424 ($p = 4.163 \times E^{-4}$) at rs10761725 (87.25 cM; Figure 1). With the addition of 5 independent WES-derived markers to reduce inter-marker intervals, the evidence for linkage at 10q11–q21 increased (exponential LOD =

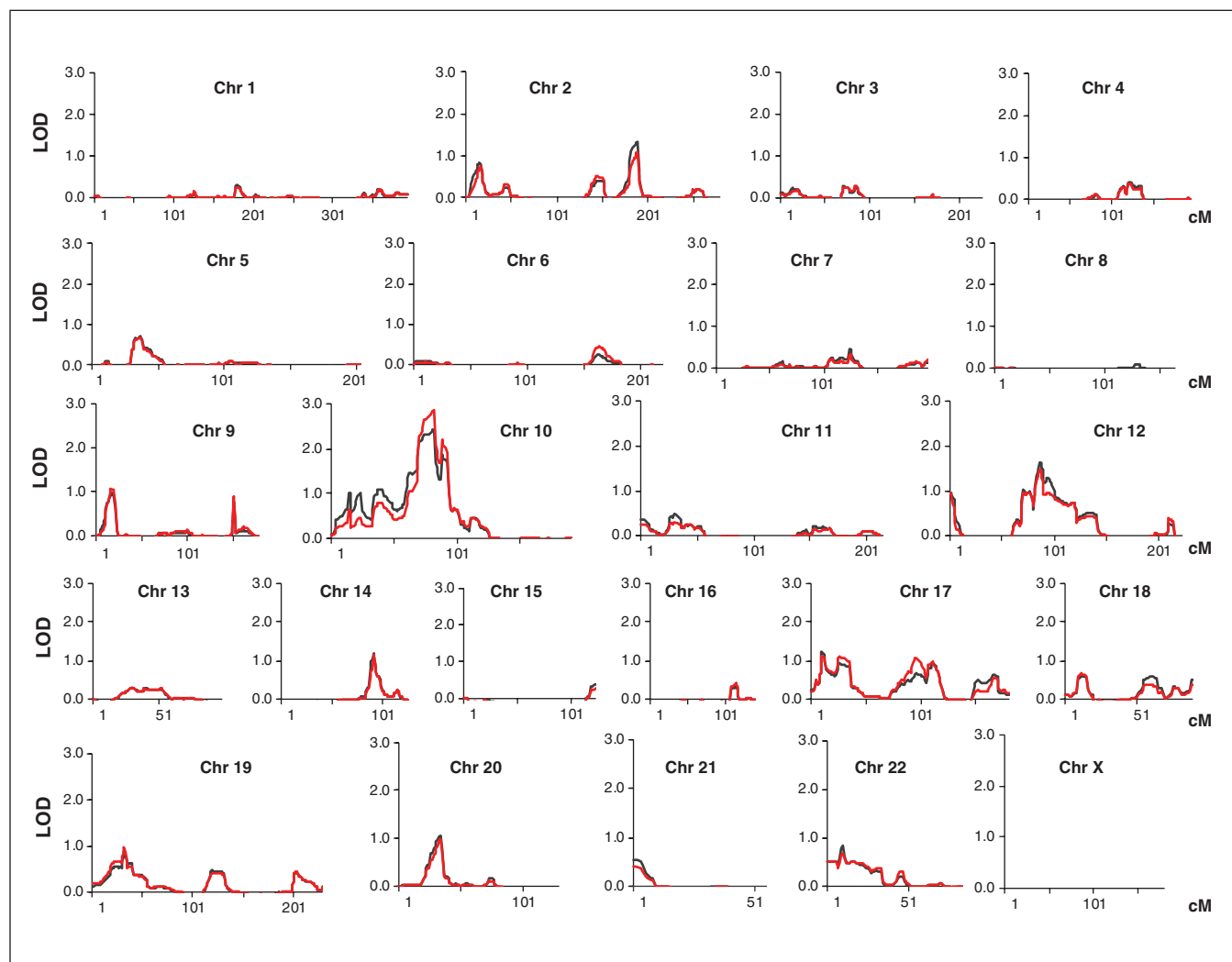


Fig. 1: Results of the genome-wide linkage analysis in 15 extended families with bipolar disorder. The results are shown for nonparametric linkage under linear (black line) and exponential (red line) models. Each chromosome is represented in a plot, including the X chromosome. A linkage signal with LOD greater than 2 was observed at chromosome 10q11.23–q21, with a maximum linear LOD score of 2.424 at marker rs10761725 (uncorrected $p = 0.00042$) and a maximum exponential LOD score of 2.863 (uncorrected $p = 0.00014$). Additional markers were subsequently added to refine the 10q11.23–q21 locus. Chr = chromosome; LOD = logarithm of the odds (for linkage).

3.103, $p = 7.84 \times E^{-5}$; linear LOD = 2.591, $p = 2.76 \times E^{-4}$) and rs10761725 remained the peak marker (Figure 2A; Appendix 1, Table S2). To assess the empirical genome-wide significance of the 10q11-q21 linkage, we permuted phenotypes across 10000 simulations, obtaining significance levels of $p = 0.045$ and $p = 0.046$ for the observed linkage under linear and exponential models, respectively.

The linkage interval, as defined by a 1-LOD drop interval, spanned a genetic distance of 15.88 cM between markers rs12571253 and rs2105702 (73.3 cM to 89.19 cM) and a physical interval of 16.1 Mb (hg19/chr10: 51589149 to 67726514). Per-family linkage analysis indicated that 11 families were contributing positively to the LOD score at 10q11-q21 (Appendix 1, Table S3). This 95% confidence interval contained

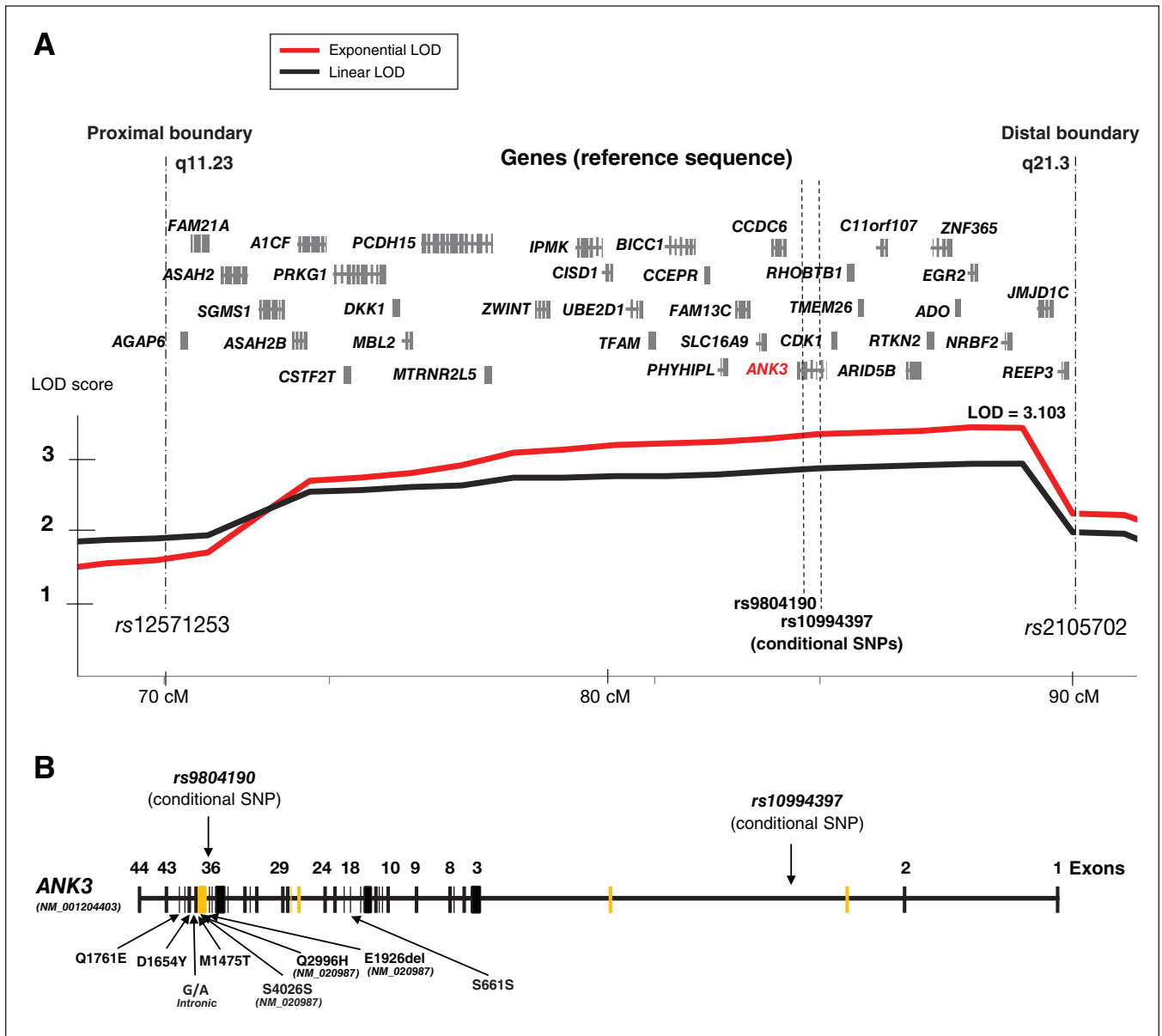


Fig. 2: Linkage on chromosome 10q11.23-q21.3 after refining the interval with additional markers. (A) A schematic of the linkage interval between the proximal and distal boundaries from rs12571253 to rs2105702 (73–89 cM) on chromosome 10q11.23-q21.3 after refining the linkage interval with an additional 5 SNPs. The maximum LOD score was found at rs10761725 under the exponential model ($LOD_{exp} = 3.103$, $p = 7.84 \times E^{-5}$) with a significant empirical p value ($p = 0.046$) after 10000 simulations of permuted phenotypes. The locations of known protein-coding genes in the linkage interval are provided, and the position of the conditional linkage SNPs rs10994397 and rs9804190 in *ANK3* are shown. (B) A schematic of the *ANK3* gene is depicted in the context of the position of the rare variants identified (reference sequence for *ANK3* isoform-3: NM_001204403, ENST00000373827; protein NP_001191332, unless indicated with reference to NM_020987 or NP_066267 for isoform-1). Alternative exons from the main isoform are depicted in yellow. The location of the bipolar-disorder-associated SNPs rs10994397 and rs9804190 used in the conditional linkage are shown in introns 2 and 36, respectively. LOD = logarithm of the odds (for linkage); SNP = single nucleotide polymorphism.

36 protein-coding genes (Figure 2A), including the *ANK3* gene, previously implicated in GWAS studies of bipolar disorder.

Conditional linkage analysis and *ANK3* SNP selection

Significant or suggestive association with a number of SNPs within the *ANK3* locus have been reported in GWAS studies of bipolar disorder: rs10994336,⁵ rs10994397,^{6,8} rs10994299,⁷ rs10994338,³⁵ rs4948418,⁴⁰ rs10994359,⁸ rs10994415⁹ and rs9804190.^{6,8,36} Therefore, we examined whether common variants associated with bipolar disorder were partly or entirely responsible for the observed linkage signal at 10q11-q21 by conditional linkage analysis.

First, we performed an LD study to assess the independence of these 8 risk alleles associated with bipolar disorder. Using 1000 Genomes genotype data from 405 non-Finnish Europeans, we observed high to moderate LD among the low-frequency SNPs associated with bipolar disorder in the 5' regulatory region of *ANK3* (Appendix 1, Figure S2; hg19/chr10: 62076628 to 62322034 bp; high LD among rs10994299, rs10994336, rs10994338, rs4948418, rs10994359 and rs10994397; moderate LD with rs10994415; $D' > 0.65$ and $R^2 > 0.38$; average MAF = 0.054). We observed an independent signal with the more common SNP variant rs9804190 (MAF = 0.225) within *ANK3* intron 36 (Figure 2B). Thus, conditional linkage with 1 SNP from the 5' LD block would index the same risk effect, and in conjunction with rs9804190, would capture most of the common variant association signal from *ANK3*. Therefore, we selected rs10994397 (with the highest imputation quality metric; $R_{sq} = 0.999$) as the tagging SNP for the 5' LD block, which maps within intron 2 of *ANK3* (Figure 2B).

Conditional linkage analysis using LAMP revealed that the risk variants rs10994397 and rs9804190 only partially explained the 10q11-q21 linkage signal (rs10994397 LOD = 0.63; rs9804190 LOD = 0.04), providing evidence that those variants contributed minimally to the overall LAMP linkage signal of a LOD greater than 3 (Appendix 1, Table S4). Furthermore, the test for other linked variants suggested that variants other than rs10994397 and rs9804190 (with "other linked variant" LODs of 3.4 and 2.91, respectively) were likely responsible for the observed linkage signal at 10q11-q21.

Family-based association analysis for variants across 10q11-q21

Next, we explored the possibility that rare alleles — in *ANK3* or in other genes in the region and not indexed by SNPs identified by GWAS (typically with MAF > 5%) — explained the linkage peak at 10q11-q21. We extracted all rare variants (MAF < 5% in Exome Variant Server) from WES data in protein-coding genes under the linkage peak interval (Figure 2A); after quality-control filters, we selected 34 variants from 14 genes for segregation analysis (Appendix 1, Table S5).

We performed a family gene-based segregation test of the 34 rare variants using the GESE package,⁴¹ with no selection for the predicted pathogenicity of variants or the total count of rare alleles among affected or unaffected relatives in the 15 pedigrees. Allele-frequency weighted segregation analysis

revealed significant rare variant segregation with bipolar disorder in 4 genes: *NRBF2* ($p = 0.005$), *PCDH15* ($p = 0.002$), *ASAH2* ($p = 0.004$) and *ANK3* ($p = 0.014$; Table 1). The only truncating variant (nonsense) was found in the *ASAH2* gene, and although sphingolipid metabolism is important for neuronal differentiation and signalling, this gene appears to show extremely low or nondetectable expression in the human brain over development, so its potential contribution to bipolar disorder is unlikely. Interestingly, the 4 rare variants that segregated with bipolar disorder in the genes *NRBF2*, *PCDH15* and *ANK3* were all missense variants predicted to be pathogenic either in PolyPhen or SIFT, or both, and were found in 3 families that contributed positively to the overall LOD score (Table 2). Furthermore, annotation using the combined annotation-dependent depletion (CADD)⁴² and deleterious annotation of genetic variants using neural networks (DANN)⁴³ algorithms predicted protein-disrupting or -altering pathogenicity of variants in *ASAH2*, *PCDH15*, *ANK3* and *NRBF2*, with scores of greater than 20 and greater than 0.99, respectively, placing them in the top 1% to 0.01% of deleterious variants in the human genome.

Further exploration of variation in *ANK3*

Given the established role of *ANK3* common variants in risk for bipolar disorder, we explored the potential involvement of rare variants in this gene by considering coding and non-coding alleles in all bipolar disorder subphenotypes (bipolar disorder I, bipolar disorder II, schizoaffective disorder–manic type and recurrent unipolar depression) with any segregation pattern among affected or unaffected participants (Table 3). We also included rare intronic or rare synonymous variants, given their potential effect on splicing or exon splicing enhancers, which are typically discarded during the WES or WGS pathogenic variant selection process.

We observed a single missense variant (rs780899852) in 1 unaffected parent with segregation to all 4 affected offspring and none of the 2 unaffected offspring of pedigree 18 (Table 2; Table 3). This was predicted to be damaging in 3 *ANK3* transcripts, including the longest canonical

Table 1: Top 5 genes from a gene-based segregation test (GESE) of 34 rare variants in 14 genes under the linkage peak

Gene	Rare variants/ segregating rare variants, <i>n</i>	Observed probability	<i>p</i> value	Weighted <i>p</i> value
<i>NRBF2</i>	1/1	0.00015	0.00051	0.00538
<i>PCDH15</i>	10/2	0.00017	0.00774	0.00220
<i>ASAH2*</i>	1/1	0.00141	0.01275	0.00408
<i>ANK3</i>	5/1	0.00454	0.04415	0.01403
<i>JMJD1C</i>	7/2	0.02067	0.24728	0.09346

The number of rare variants per gene relative to the number of rare variants segregating with bipolar disorder status across the 15 families with bipolar disorder ($n = 117$ individuals). The observed segregating probability for each gene is provided, as is the basic gene-based *p* value for association after simulations and the gene-based *p* value after per-family weight corrections. Suggestive *p* values (< 0.1) are reported.

*Genes not expressed in the brain (defined by reads per kilobase per million [RPKM] < 1 in developmental transcriptomics RNA sequence data, representing 65% of mapped genes).

Table 2: Most significantly associated rare variants in the family-based association test and their predicted pathogenicity*

Chr:BP (hg19)	dbSNP identifier	Gene	Reference allele/ alternative allele	Variant type	PolyPhen	CADD Phred score	DANN score	Alternative allele count†	Pedigree (per-family LOD)
10:52005095	rs116049719	<i>ASAH2</i> ‡	G/A	Nonsense	—	35	0.9974	9	131 (0.31) 18 (0.42)§ 1 (0.68) 74 (–0.03)
10:55566719	rs570828018	<i>PCDH15</i>	G/A	Missense	Damaging	19.18	0.9984	5	138 (0.43)¶
10:55581787	Novel	<i>PCDH15</i>	T/A	Missense	Benign	23	0.9664	5	18 (0.42)§
10:64967951	rs139722368	<i>JMJD1C</i>	CTAAAC/–	Indel	—	19.05	—	15	129 (0.24) 138 (0.43)¶ 19 (–0.124)
10:64974380	rs41274074	<i>JMJD1C</i>	G/C	Missense	Benign	15.64	0.8880	15	129 (0.24) 138 (0.43)¶ 19 (–0.124)
10:61815652	rs780899852	<i>ANK3</i>	G/C	Missense	Damaging	24.3	0.9927	5	18 (0.42)§
10:64913602	rs62623680	<i>NRBF2</i>	A/G	Missense	Damaging	28.3	0.9980	6	17 (0.57)

CADD = Combined Annotation Dependent Depletion; Chr:BP (hg19) = the chromosome (chr) and base pair (BP) position of each variant, based on Genome Reference Consortium Human Build 37 (GRCh37/hg19); DANN = Deleterious Annotation of genetic variants using Neural Networks; dbSNP ID = SNP identifier; LOD = logarithm of the odds (for linkage); PolyPhen = polymorphism phenotyping, functional prediction of SNP impact on protein product; SNP = single nucleotide polymorphism.

*Pathogenicity prediction is presented for missense variants using PolyPhen, the CADD Phred scaled score and DANN scores. A CADD Phred scaled score⁴² of 20 means that the variant is among the top 1% of deleterious variants in the human genome, and a score of 30 means that the variant is in the top 0.1%. For DANN,⁴³ a score of 0.96 identifies 92.1% of pathogenic variations from ClinVar (www.ncbi.nlm.nih.gov/clinvar) and 18.1% of false-positive benign variations; values greater than 0.98 predict protein-disrupting or -altering variants. Positive values for the per-family LOD score indicate a positive contribution to the linkage signal from each family carrying that variant.

†Total number of alternative alleles observed across 117 sequenced individuals.

‡Genes not expressed in the brain (defined by reads per kilobase per million [RPKM] < 1 in developmental transcriptomics RNA sequencing data, representing 65% of mapped genes) are indicated.

§Variants present on the same haplotype in pedigree 18.

¶Variants present on the same haplotype in pedigree 138.

Table 3: Inventory of coding and non-coding rare variants identified in ANK3, which were identified in 8 of the 15 families

Chr:BP (hg19)	dbSNP identifier	Reference allele/ alternative allele	Alternative allele count*	Pedigree†	Affected, any diagnosis (alternative/total, n)‡	Unaffected (alternative/total, n)§	Protein impact	PolyPhen	CADD Phred score	DANN score
10:61815652	rs780899852	G/C	5	18	4/4	1/4	Q4277E Q1761E¶	Damaging	24.3	0.9927
10:61819543	rs139092048	C/A	7	19	3/5	0/2	D1654Y¶	Damaging	20.1	—
10:61824133	rs71495624	G/A	6	121	1/8	2**/5	D1654Y¶	—	—	—
10:61827731	rs748863853	A/G	1	17	3/4	3/4	Intronic	—	0.294	—
10:61828561	rs200605397	G/A	2	131	1/5	0/2	M4094T M1475T¶	Damaging	26.3	0.9839
10:61831651	rs41274672	C/G	3	74	2/6	0/2	S4026	—	9.294	—
10:61834861–3	rs764632652	CTC/–	5	1	1/5	0/3	Q2996H	Benign	0.49	0.9712
10:61946557	rs146438080	G/A	4	131	84	0/2	Q2996H	—	0.49	—
					3/8	2/5	E1926/–	—	22	—
					4	1/2	S661¶	—	13.29	—

CADD = Combined Annotation Dependent Depletion; Chr:BP (hg19) = the chromosome (chr) and base pair (BP) position of each variant, based on Genome Reference Consortium Human Build 37 (GRCh37/hg19); DANN = Deleterious Annotation of genetic variants using Neural Networks; dbSNP ID = SNP identifier; PolyPhen = polymorphism phenotyping, functional prediction of SNP impact on protein product; SNP = single nucleotide polymorphism.

The number of individuals in each pedigree carrying alternative alleles for *ANK3* variants is listed, employing a broader definition of "affected" to include any mood disorder diagnosis. Two families (pedigrees 1 and 131) carried 2 different *ANK3* variants in separate arms of the family, and 2 variants were identified in 2 families (rs139092048 and rs41274672).

Pathogenicity prediction is presented for missense variants using PolyPhen, the CADD Phred scale score and DANN scores. The reference sequence for *ANK3* protein is NP_066267 (ENST00000280772, isoform-1) unless otherwise indicated.

*Total number of alternative alleles observed across 117 sequenced individuals.

†The pedigree or family in which the variant was detected.

‡Number of alternative allele carriers in family members affected with any mood disorder diagnosis (i.e., bipolar disorder I, bipolar disorder II, schizoaffective disorder manic type or recurrent unipolar depression; see Appendix 1, Table S1) out of the total number of family members with any diagnosis.

§Number of alternative allele carriers in unaffected family members out of the total number of unaffected family members.

¶Reference sequence for *ANK3* protein: NP_001191332 (ENST00000373827, isoform-3).

**This variant was identified in homozygous form in the unaffected spouse of a person with bipolar disorder I; the 2 people were the parents of a person with bipolar disorder I.

isoform-1 (ENST00000280772, Q4277E; NP_066267), isoform-3 (ENST00000373827, Q1761E; NP_001191332) and the short most brain-abundant transcript (ENST00000373820,

Q359E; H0Y3A4). We found no evidence for full segregation of any of the observed synonymous or intronic variants of *ANK3* (rs71495624, rs200605397, rs146438080) with

bipolar disorder in these families. Other missense or indel variants (rs139092048, rs748863853, rs41274672, rs764632652) were observed in a few affected relatives or were predicted to be benign (Table 3).

Discussion

The combination of next-generation sequencing with linkage analysis has exponentially increased the identification of genes implicated in rare Mendelian diseases.⁴⁴ In complex disorders, linkage studies can be promising when combined with whole-exome or genome sequence data, where genetic loci can be simultaneously inspected for rare variant segregation with putative higher penetrance effects in large families.^{13,45} This combined strategy has been employed in several large individual families with a range of psychiatric conditions,^{12,15,19,23–28} and also in a joint analysis of 3 families affected by attention-deficit/hyperactivity disorder (including a total of 10 cases with WES).²⁹ Here, we expanded this approach by performing a linkage study across 15 unrelated Australian-European families comprising 4 or more relatives per family with an affective disorder (including 117 participants with WES), with the aim of identifying genes that carried rare pathogenic variants through family-based segregation tests. Linkage analysis across the 15 extended families revealed a significant linkage peak at chromosome 10q11-q21, encompassing 16.13 Mb and 36 protein-coding genes, including *ANK3*.

Several large GWAS studies have reported significant associations with bipolar disorder at variants in *ANK3*,^{5–9,35,36,40} and the majority of these SNPs are in high LD, suggesting that they are largely indexing a similar genetic effect. A conditional linkage analysis based on independent GWAS-significant SNPs rs10994397 and rs9804190 showed that common bipolar disorder risk alleles in *ANK3* contributed only marginally to the observed linkage signal at 10q11-q21, although this did not exclude the involvement of other common risk variants not indexed by rs10994397 and rs9804190.

To uncover the possible effects of rare variants with higher penetrance, we conducted a gene-based segregation analysis that revealed significant segregation in individual families for rare alleles in the protein-coding regions of the brain-expressed genes *ANK3*, *PCDH15* and *NRBF2*. The conditional analysis and family-based segregation results suggest that common variants play a minor role in explaining the linkage to 10q11-q21, which may be more adequately explained by rare genetic variants. However, segregating rare variants in *ANK3*, *PCDH15* and *NRBF2* were observed only in some individual families, and not all families who contributed positively to the linkage signal had segregating rare variants in any of the 36 protein-coding genes examined, suggesting that other variants not captured by WES (such as those in regulatory regions or non-protein-coding genes) also contribute to the genetic risk of bipolar disorder and remain to be elucidated.

A compelling body of evidence makes the *ANK3* gene the most plausible candidate gene to emerge from our study. *ANK3* belongs to the family of postsynaptic density genes and encodes the ankyrin-G protein, which acts as scaffolding

adaptor between ion channels and the β -spectrin-actin cytoskeleton. There are 3 main ankyrin-G isoforms: the canonical 190 kDa isoform is expressed in most tissues and acts as a perisynaptic scaffold in dendritic spine heads and necks;⁴⁶ the 270 and 480 kDa isoforms are found predominantly in the brain and nervous system, acting as master organizers of ion channels. These latter isoforms are involved in signal propagation at the axon initial segment and formation of the nodes of Ranvier.⁴⁷

Mutations in *ANK3* are reported in several psychiatric diseases, suggesting that this gene has a pleiotropic effect. Several studies also suggest a correlation between the severity of *ANK3* mutations and the severity of the phenotype: homozygous truncating mutations in an *ANK3* isoform were associated with a neurodevelopmental phenotype characterized by intellectual disability, speech delay and behavioural abnormalities;⁴⁸ a balanced translocation in *ANK3* affecting all isoforms was reported in a patient with autism spectrum disorder, intellectual disability and muscular hypotonia;⁴⁸ heterozygous missense mutations were found in autism spectrum disorder with mild developmental delay;⁴⁹ and common variant associations have been reported in bipolar disorder.^{5–9,35,40} These studies suggest that highly disrupting loss-of-function coding mutations that segregate in bipolar disorder are unlikely, although it is plausible that not only missense variants but also non-coding variants in *ANK3* may explain the observed linkage. Thus, common and rare non-coding variants of *ANK3* may be implicated in bipolar disorder and act in multiple processes. For instance, the complex alternative splicing of this gene, which includes up to 51 exons and 6 different start exons, is regulated in a temporal and tissue-specific manner.⁵⁰ A rare variant leading to a novel alternative-spliced exon in a specific *ANK3* isoform has been recently associated with bipolar disorder and schizophrenia,⁵¹ and common intronic variants may also influence *ANK3* splicing.⁵⁰ Second, genetic variation in *ANK3* may also modulate epigenetic mechanisms through abolishing or creating CpG islands. Methylation occurs primarily in the first intron and the 5' untranslated region (UTR) of *ANK3*, with a specific temporal profile that has been shown to affect gene expression.⁵² Third, post-transcriptional regulation through specific sites at the 3'-UTR may mediate binding with miRNAs. Rare variants that modify consensus regions at the 3'-UTR recognized by miRNA seed regions may affect gene function,⁵³ as is the case for miR-34a that regulate *ANK3* transcripts.⁵⁴ The rare variants reported here may provide valuable targets to take forward into functional analyses examining the molecular mechanisms through which *ANK3* increases risk to psychiatric disorders.

However, other genes also appear to contribute to genetic risk for bipolar disorder in the 10q11-q21 locus. From our study, *PCDH15* and *NRBF2* carried rare variants that segregated with bipolar disorder. Protocadherin-related 15 (*PCDH15*) is part of the large family of cadherin genes involved in neuronal differentiation and synapse formation.⁵⁵ Several members of this superfamily have been associated with psychiatric disorders.⁵⁵ *PCDH15* is responsible for Usher syndrome type 1F, a recessive Mendelian condition characterized by congenital deafness, retinitis pigmentosa and

vestibular dysfunction; of those with this syndrome, approximately 20% manifest mental or behavioural disorders, including schizophrenia-like psychosis, bipolar disorder and atypical autism.⁵⁶ Structural variants across *PCDH15* have also been reported in patients with bipolar disorder,⁵⁷ and common variants influence the personality trait extraversion,⁵⁸ which is associated with bipolar disorder. A recent functional study of neurons derived from induced pluripotent stem cells of 2 patients with bipolar disorder carrying *PCDH15* deletions indicated dendrite shortening and decreasing synapse numbers.⁵⁹ However, the contribution of rare variants in *PCDH15* to psychiatric disorder is unclear, and although evidence of pleiotropy with respect to other psychiatric risk genes is compelling,^{60–62} studies on *PCDH15* have been too small to draw definitive conclusions.

The final potential candidate gene emerging from our study was the nuclear receptor binding factor 2 (*NRBF2*) gene that positively regulates autophagy.⁶³ Autophagic processes are essential for neuronal survival and function, preserving the balance between synthesis, degradation and recycling of cellular components in response to stress. Cellular autophagy has been implicated in the risk of both schizophrenia and bipolar disorder.⁶⁴ Lithium, a first-line medicinal treatment for bipolar disorder, has been shown to induce autophagy,⁶⁵ providing neuroprotection against mitochondrial dysfunction and the accumulation of reactive oxygen species.⁶⁶ Thus, *NRBF2* warrants further attention as a potential contributor to risk of bipolar disorder.

Limitations

Although the conditional SNPs chosen to represent *ANK3* common risk variants appeared to index the majority of associated SNPs reported thus far, we could not exclude the involvement of other common risk variants not indexed by rs10994397 or rs9804190. As well, it is plausible that other non-coding variants in *ANK3* or other genes in the linkage region may explain the remaining linkage we observed. Non-coding variants in UTRs and cis-regulatory elements may play a significant role in psychiatric disorders.^{10,67} However, intergenic, promoter and intronic sequences were not captured by the WES employed in this study, and UTRs had limited coverage by the amplicons used for sequence library preparation.

Conclusion

Linkage to chromosome 10q11.23-q21.3 in 15 extended families was not explained by common variations in *ANK3* associated with bipolar disorder, but it was consistent with the contribution of multiple coding rare variants from multiple genes, including *ANK3*, *PCDH15* and *NRBF2*. Rare coding variants at the 10q11.23-q21.3 locus explain part of the genetic risk for bipolar disorder, and further studies are warranted to identify additional functional alleles, with particular focus on rare non-coding variants that may affect the splicing, methylation and post-translational regulation of *ANK3* and additional candidate genes at this locus.

Acknowledgments: We are grateful to all participants and their families, as well as clinical collaborators who were originally involved in collecting and phenotyping these families. We would also thank Richard J. N. Allcock from the University of Western Australia (Australia) for WES library preparation and analytic support. This study was funded by the Australian National Medical and Health Research Council (NHMRC) project grants 1063960 and 1066177, and program grant 1037196. We gratefully acknowledge the Janette Mary O'Neil Research Fellowship (to JMF) and Mrs. Betty Lynch (dec.) for supporting this work and our team. DNA was extracted by Genetic Repositories Australia, an Enabling Facility that was supported by NHMRC Enabling Grant 401184. Samples were sequenced at the Lottery State Biomedical Genomics Facility, University of Western Australia, and genotyped on Illumina PsychChip at the Australian Genome Research Facility. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

Affiliations: From Neuroscience Research Australia, Sydney, Australia (Toma, Shaw, Heath, Pierce, Schofield); the School of Medical Sciences, University of New South Wales, Sydney, Australia (Toma, Shaw, Schofield, Fullerton); the Centro de Biología Molecular 'Severo Ochoa', Universidad Autónoma de Madrid/CSIC, Madrid, Spain (Toma); the School of Psychiatry, University of New South Wales, Sydney, Australia (Mitchell); and the Black Dog Institute, Prince of Wales Hospital, Sydney, Australia (Mitchell).

Competing interests: P. Mitchell reports personal fees from Sanofi (Hangzhou) and Janssen-Cilag, outside the submitted work.

Contributors: C. Toma, A. Shaw, P. Mitchell, P. Schofield and J. Fullerton designed the study. All authors acquired the data, which C. Toma, A. Shaw, P. Mitchell, P. Schofield and J. Fullerton analyzed. C. Toma and J. Fullerton wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors or in the acknowledgments have made substantial contributions to the paper.

Content licence: This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

References

- 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;68:241–51.
- Mitchell PB, Johnston AK, Frankland A, et al. Bipolar disorder in a national survey using the World Mental Health Version of the Composite International Diagnostic Interview: the impact of differing diagnostic algorithms. *Acta Psychiatr Scand* 2013;127:381–93.
- McGuffin P, Rijsdijk F, Andrew M, et al. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* 2003;60:497–502.
- Stahl EA, Breen G, Forstner AJ, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 2019;51:793–803.
- Ferreira MA, O'Donovan MC, Meng YA, et al. Collaborative genome-wide association analysis supports a role for *ANK3* and *CACNA1C* in bipolar disorder. *Nat Genet* 2008;40:1056–8.
- Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. *Nat Genet* 2011;43:977–83.
- Charney AW, Ruderfer DM, Stahl EA, et al. Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. *Transl Psychiatry* 2017;7:e993.
- Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major

- psychiatric disorders: a genome-wide analysis. *Lancet* 2013; 381:1371-9.
9. Muhleisen TW, Leber M, Schulze TG, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun* 2014;5:3339.
 10. Ament SA, Szelinger S, Glusman G, et al. Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proc Natl Acad Sci U S A* 2015;112:3576-81.
 11. Goes FS, Pirooznia M, Parla JS, et al. Exome sequencing of familial bipolar disorder. *JAMA Psychiatry* 2016;73:590-7.
 12. Toma C, Shaw AD, Allcock RJN, et al. An examination of multiple classes of rare variants in extended families with bipolar disorder. *Transl Psychiatry* 2018;8:65.
 13. Glahn DC, Nimgaonkar VL, Raventos H, et al. Rediscovering the value of families for psychiatric genetics research. *Mol Psychiatry* 2019;24:523-35.
 14. Cruceanu C, Ambalavanan A, Spiegelman D, et al. Family-based exome-sequencing approach identifies rare susceptibility variants for lithium-responsive bipolar disorder. *Genome* 2013;56:634-40.
 15. Georgi B, Craig D, Kember RL, et al. Genomic view of bipolar disorder revealed by whole genome sequencing in a genetic isolate. *PLoS Genet* 2014;10:e1004229.
 16. Strauss KA, Markx S, Georgi B, et al. A population-based study of KCNH7 p.Arg394His and bipolar spectrum disorder. *Hum Mol Genet* 2014;23:6395-406.
 17. Rao AR, Yourshaw M, Christensen B, et al. Rare deleterious mutations are associated with disease in bipolar disorder families. *Mol Psychiatry* 2017;22:1009-14.
 18. Salvoro C, Bortoluzzi S, Coppe A, et al. Rare risk variants identification by identity-by-descent mapping and whole-exome sequencing implicates neuronal development pathways in schizophrenia and bipolar disorder. *Mol Neurobiol* 2018;55:7366-76.
 19. Zhang T, Hou L, Chen DT, et al. Exome sequencing of a large family identifies potential candidate genes contributing risk to bipolar disorder. *Gene* 2018;645:119-23.
 20. Cruceanu C, Schmouh JF, Torres-Platas SG, et al. Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors. *Mol Psychiatry* 2018;23:2050-56.
 21. Guimier A, Gabriel GC, Bajolle F, et al. MMP21 is mutated in human heterotaxy and is required for normal left-right asymmetry in vertebrates. *Nat Genet* 2015;47:1260-3.
 22. Marsh AP, Heron D, Edwards TJ, et al. Mutations in DCC cause isolated agenesis of the corpus callosum with incomplete penetrance. *Nat Genet* 2017;49:511-4.
 23. Chapman NH, Nato AQ Jr, Bernier R, et al. Whole exome sequencing in extended families with autism spectrum disorder implicates four candidate genes. *Hum Genet* 2015;134:1055-68.
 24. Amin N, de Vrij FMS, Baghdadi M, et al. A rare missense variant in RCL1 segregates with depression in extended families. *Mol Psychiatry* 2018;23:1120-6.
 25. Ryan NM, Lihm J, Kramer M, et al. DNA sequence-level analyses reveal potential phenotypic modifiers in a large family with psychiatric disorders. *Mol Psychiatry* 2018;23:2254-65.
 26. Anjanappa RM, Nayak S, Moily NS, et al. A linkage and exome study implicates rare variants of KANK4 and CAP2 in bipolar disorder in a multiplex family. *Bipolar Disord* 2020;22:70-8.
 27. John J, Sharma A, Kukshal P, et al. Rare variants in tissue inhibitor of metalloproteinase 2 as a risk factor for schizophrenia: evidence from familial and cohort analysis. *Schizophr Bull* 2019;45:256-63.
 28. John J, Bhattacharyya U, Yadav N, et al. Multiple rare inherited variants in a four generation schizophrenia family offer leads for complex mode of disease inheritance. *Schizophr Res* 2020; 216:288-94.
 29. Corominas J, Klein M, Zayats T, et al. Identification of ADHD risk genes in extended pedigrees by combining linkage analysis and whole-exome sequencing. *Mol Psychiatry* 2020;25:2047-57.
 30. NIMH Genetics Initiative: Family Interview for Genetic Studies (FIGS). Rockville (MD): National Institute of Mental Health; 1992.
 31. Numberger JI Jr, Blehar MC, Kaufmann CA, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Arch Gen Psychiatry* 1994;51:849-59, discussion 63-4.
 32. Smith KR, Bromhead CJ, Hildebrand MS, et al. Reducing the exome search space for Mendelian diseases using genetic linkage analysis of exome genotypes. *Genome Biol* 2011;12:R85.
 33. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-75.
 34. Abecasis GR, Cherny SS, Cookson WO, et al. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;30:97-101.
 35. Liu Y, Blackwood DH, Caesar S, et al. Meta-analysis of genome-wide association data of bipolar disorder and major depressive disorder. *Mol Psychiatry* 2011;16:2-4.
 36. Schulze TG, Detera-Wadleigh SD, Akula N, et al. Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder. *Mol Psychiatry* 2009;14:487-91.
 37. Li M, Boehnke M, Abecasis GR. Joint modeling of linkage and association: identifying SNPs responsible for a linkage signal. *Am J Hum Genet* 2005;76:934-49.
 38. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-9.
 39. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003;31:3812-4.
 40. 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.
 41. Qiao D, Lange C, Laird NM, et al. Gene-based segregation method for identifying rare variants in family-based sequencing studies. *Genet Epidemiol* 2017;41:309-19.
 42. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310-5.
 43. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* 2015;31:761-3.
 44. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Rev Genet* 2015;16:275-84.
 45. Wijsman EM. The role of large pedigrees in an era of high-throughput sequencing. *Hum Genet* 2012;131:1555-63.
 46. Smith KR, Kopeikina KJ, Fawcett-Patel JM, et al. Psychiatric risk factor ANK3/ankyrin-G nanodomains regulate the structure and function of glutamatergic synapses. *Neuron* 2014;84:399-415.
 47. Chang KJ, Zollinger DR, Susuki K, et al. Glial ankyrins facilitate paranodal axoglial junction assembly. *Nat Neurosci* 2014;17: 1673-81.
 48. Iqbal Z, Vandeweyer G, van der Voet M, et al. Homozygous and heterozygous disruptions of ANK3: at the crossroads of neurodevelopmental and psychiatric disorders. *Hum Mol Genet* 2013;22: 1960-70.
 49. Codina-Sola M, Rodriguez-Santiago B, Homs A, et al. Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. *Mol Autism* 2015;6:21.
 50. Rueckert EH, Barker D, Ruderfer D, et al. Cis-acting regulation of brain-specific ANK3 gene expression by a genetic variant associated with bipolar disorder. *Mol Psychiatry* 2013;18:922-9.
 51. Hughes T, Sonderby IE, Polushina T, et al. Elevated expression of a minor isoform of ANK3 is a risk factor for bipolar disorder. *Transl Psychiatry* 2018;8:210.
 52. Luoni A, Massart R, Nieratschker V, et al. Ankyrin-3 as a molecular marker of early-life stress and vulnerability to psychiatric disorders. *Transl Psychiatry* 2016;6:e943.
 53. Toma C, Torrico B, Hervas A, et al. Common and rare variants of microRNA genes in autism spectrum disorders. *World J Biol Psychiatry* 2015;16:376-86.
 54. Bavamian S, Mellios N, Lalonde J, et al. Dysregulation of miR-34a links neuronal development to genetic risk factors for bipolar disorder. *Mol Psychiatry* 2015;20:573-84.
 55. Hawi Z, Tong J, Dark C, et al. The role of cadherin genes in five major psychiatric disorders: a literature update. *Am J Med Genet B Neuropsychiatr Genet* 2018;177:168-80.
 56. Domanico D, Fragiotta S, Cutini A, et al. Psychosis, mood and behavioral disorders in Usher syndrome: review of the literature. *Med Hypothesis Discov Innov Ophthalmol* 2015;4:50-5.
 57. Noor A, Lionel AC, Cohen-Woods S, et al. Copy number variant study of bipolar disorder in Canadian and UK populations implicates synaptic genes. *Am J Med Genet B Neuropsychiatr Genet* 2014; 165B:303-13.

58. Lo MT, Hinds DA, Tung JY, et al. Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat Genet* 2017;49:152-6.
59. Ishii T, Ishikawa M, Fujimori K, et al. In vitro modeling of the bipolar disorder and schizophrenia using patient-derived induced pluripotent stem cells with copy number variations of *PCDH15* and *RELN*. *eNeuro* 2019;6:ENEURO.0403-18.2019.
60. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* 2019;179:1469-82.e11.
61. Torrico B, Anton-Galindo E, Fernandez-Castillo N, et al. Involvement of the 14-3-3 gene family in autism spectrum disorder and schizophrenia: genetics, transcriptomics and functional analyses. *J Clin Med* 2020;9:1851.
62. Torrico B, Shaw AD, Mosca R, et al. Truncating variant burden in high-functioning autism and pleiotropic effects of LRP1 across psychiatric phenotypes. *J Psychiatry Neurosci* 2019;44:350-9.
63. Zhong Y, Morris DH, Jin L, et al. Nrbf2 protein suppresses autophagy by modulating Atg14L protein-containing Beclin 1-Vps34 complex architecture and reducing intracellular phosphatidylinositol-3 phosphate levels. *J Biol Chem* 2014;289:26021-37.
64. Kim Y, Santos R, Gage FH, et al. Molecular mechanisms of bipolar disorder: progress made and future challenges. *Front Cell Neurosci* 2017;11:30.
65. Sarkar S, Rubinsztein DC. Inositol and IP3 levels regulate autophagy: biology and therapeutic speculations. *Autophagy* 2006;2:132-4.
66. Hou L, Heilbronner U, Degenhardt F, et al. Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *Lancet* 2016;387:1085-93.
67. Brandler WM, Antaki D, Gujral M, et al. Paternally inherited cis-regulatory structural variants are associated with autism. *Science* 2018;360:327-31.