

Exonic DNA Sequencing of *ERBB4* in Bipolar Disorder

Fernando S. Goes^{1*}, Michael Rongione¹, Yun-Ching Chen², Rachel Karchin², Eran Elhaik³, the Bipolar Genome Study[†], James B. Potash¹

1 Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, Maryland, United States of America, **2** Department of Biomedical Engineering and the Institute for Computational Medicine, Johns Hopkins University, Baltimore, Maryland, United States of America, **3** McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, United States of America

Abstract

The Neuregulin-ErbB4 pathway plays a crucial role in brain development and constitutes one of the most biologically plausible signaling pathways implicated in schizophrenia and, to a lesser extent, in bipolar disorder (BP). However, recent genome-wide association analyses have not provided evidence for common variation in *NRG1* or *ERBB4* influencing schizophrenia or bipolar disorder susceptibility. In this study, we investigate the role of rare coding variants in *ERBB4* in BP cases with mood-incongruent psychotic features, a form of BP with arguably the greatest phenotypic overlap with schizophrenia. We performed Sanger sequencing of all 28 exons in *ERBB4*, as well as part of the promoter and part of the 3'UTR sequence, hypothesizing that rare deleterious variants would be found in 188 cases with mood-incongruent psychosis from the GAIN BP study. We found 42 variants, of which 16 were novel, although none were non-synonymous or clearly deleterious. One of the novel variants, present in 11.2% of cases, is located next to an alternative stop codon, which is associated with a shortened transcript of *ERBB4* that is not translated. We genotyped this variant in the GAIN BP case-control samples and found a marginally significant association with mood-incongruent psychotic BP compared with controls (additive model: OR = 1.64, *P*-value = 0.055; dominant model: OR = 1.73, *P*-value = 0.039). In conclusion, we found no rare variants of clear deleterious effect, but did uncover a modestly associated novel variant that could affect alternative splicing of *ERBB4*. However, the modest sample size in this study cannot definitively rule out a role for rare variants in bipolar disorder and studies with larger sample sizes are needed to confirm the observed association.

Citation: Goes FS, Rongione M, Chen Y-C, Karchin R, Elhaik E, et al. (2011) Exonic DNA Sequencing of *ERBB4* in Bipolar Disorder. PLoS ONE 6(5): e20242. doi:10.1371/journal.pone.0020242

Editor: Katharina Domschke, University of Muenster, Germany

Received: December 14, 2010; **Accepted:** April 28, 2011; **Published:** May 26, 2011

Copyright: © 2011 Goes et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The study was funded by my NIMH grant, (K99 MH086049). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fgoes1@jhmi.edu

† Membership of the Bipolar Genome Study is provided in the Acknowledgments.

Introduction

Mounting evidence suggests that the major psychoses share, to some degree, a common genetic susceptibility [1–3]. In particular, we and others have proposed that within the broad bipolar disorder spectrum, the subtype of bipolar disorder with mood-incongruent psychotic features is likely to be the subphenotype most closely aligned with schizophrenia [4,5]. Among candidate genes implicated in the pathogenesis of both psychotic disorders, those in the Neuregulin1-ErbB4 signaling pathway have been frequently supported by association, copy number variation, and expression studies [6], although these findings have not always been consistent.

In the CNS, *NRG1* functions primarily as a signaling molecule that binds to *ERBB4*, a tyrosine kinase receptor predominantly expressed in inhibitory neurons [7]. The *ERBB4* gene spans over 1.16 Mb and consists of 28 exons. Alternative splicing of exons 15/16 and exon 26 results in the formation of at least four protein isoforms that differ in their susceptibility to extra-cellular and intracellular cleavage [8]. Bound *ERBB4* auto-phosphorylates several intracellular tyrosine residues, leading to the activation of key intracellular messengers such as phosphatidylinositol 3-kinase (PI3K) and AKT [9]. These proteins, among many functions,

inhibit glycogen synthase kinase-3 (GSK-3), which is arguably the protein most strongly implicated in lithium's mechanism of action [10]. Post-mortem studies have also provided initial evidence for a functional interaction of *ERBB4* with the NMDA receptor [11], also a promising therapeutic target for both mood disorders and schizophrenia.

Although *ERBB4* has been far less studied than *NRG1*, preliminary evidence suggests a possible association with schizophrenia. Among the initial candidate gene studies of *ERBB4*, Silberberg et al. found an association between schizophrenia and three highly linked markers surrounding exon 3 of the gene (best allelic *P*-value = 0.0049) [12]. However, the sample size was small (total *N* = 199) and these findings have not been replicated. Subsequent studies have focused on the interactions between *NRG1* and *ERBB4*; while significant interactions between various markers were reported by each study, there was little overlap among the actual interacting markers across studies [13–16]. Moreover, in genome-wide association studies (GWAS) of Caucasians with schizophrenia or bipolar disorder, neither *NRG1* nor *ERBB4* have featured among the top hits in the original studies or in subsequent meta-analyses [17,18]. By contrast, the most highly associated SNP in the only GWAS of schizophrenia in African-Americans was found in *ERBB4* (rs1851196, *P*-value = 2.14×10^{-6}), though the

sample size was relatively modest by GWAS standards and the findings fell short of genome-wide significance.

The above studies of *ERBB4* have focused almost exclusively on common variants. However, increasing evidence indicates that rare variants might also play a role in the etiology of complex diseases such as bipolar disorder and schizophrenia [19]. Indeed, a large deletion (~400 kb) of the 3' region of *ERBB4* has been reported in a subject with schizophrenia [20], but, to our knowledge, no study has performed comprehensive sequencing of *ERBB4*.

While most of the evidence for association in the NRG1-ERBB4 pathway comes from studies of schizophrenia, we hypothesized that variation in the pathway might also be involved in susceptibility to bipolar disorder with mood-incongruent psychosis, where symptoms can often be indistinguishable from those of schizophrenia. In this study we have performed comprehensive sequencing of all 28 exons of *ERBB4* in cases with mood-incongruent psychosis from the GAIN BP sample [21]. While we did not find an excess of functional rare variants, we discovered a novel, potentially functional, common variant, which was additionally genotyped in a case-control association experiment. We demonstrate a modest excess of this variant that appears to be specific to the mood-incongruent form of BP.

Results

We analyzed 5.9 kb of DNA sequence representing all the 28 exons and surrounding sequences, as well as approximately 600 bp of the promoter sequence, all the 5' UTR, and 400 bp of the 3' UTR sequence in 188 BP subjects with mood-incongruent psychotic features. We found 42 variants; of these, 26 were single nucleotide variants (SNVs) present in dbSNP 132, including 17 common polymorphisms present in the HapMap CEU sample (Fig. 1). We discovered 16 novel variants across *ERBB4* (Table 1), and while no novel variant was non-synonymous, several were found to have bioinformatic evidence of a potential functional effect. There were two synonymous SNVs predicted to change splice site enhancers and silencers, as well as three variants in the 3' UTR sequence.

Among the 16 novel variants, our power analyses indicated that only two variants (SNVs 7 and 8) had allele frequencies sufficiently high enough (MAF>0.03) to warrant additional genotyping in our

available 999 independent controls. Of these, the potentially most interesting finding was SNV 7 (G>A), which was absent in dbSNP 132, but was present in 21 out of 188 cases (11.2% prevalence, MAF of 5.6%). As shown in Fig. 2, this SNV is 40 bp downstream of exon 20 and is located next to an alternate "bleeding" form of exon 20 that is associated with a prematurely truncated transcript of *ERBB4*, with no evidence of being translated (<http://genome.ucsc.edu/>).

To determine whether this novel SNV was associated with mood-incongruent psychotic BP, we performed an association study by genotyping additional controls (N=999) and the remaining (non mood-incongruent psychotic) BP cases (N=806) from the GAIN BP sample (Table 2). The novel marker was in Hardy-Weinberg equilibrium in both cases (P -value = 1.0) and controls (P -value = 0.36), and was present in 6.9% of controls, with a MAF of 3.6%. Using logistic regression with principal components as covariates we found an association of the novel SNV in cases with mood-incongruent psychotic BP compared with controls that was statistically significant in a dominant model (OR = 1.73, P -value = 0.039) and close to significance in an additive model (OR = 1.64, P -value = 0.055). This association appeared to be specific to cases with mood-incongruent psychotic BP, since a case-only analysis found an enrichment of the putative risk allele in cases with mood-incongruent psychosis compared with all other BP cases (Table 2). The second variant tested for association was a common A/- insertion-deletion polymorphism (SNV 8). This variant was genotyped in the 999 controls and 806 non-mood incongruent BP cases, but showed no evidence of association in either the case-control (P -value = 0.87) or the case-only analyses (P -value = 0.86).

Discussion

In this study we sequenced all coding regions of *ERBB4*, a candidate gene with strong biological plausibility, as well as suggestive evidence for genetic association with schizophrenia. We hypothesized that rare novel functional variants, potentially of large effect, would be over-represented in BP cases with mood-incongruent psychotic features, a subset of BP with phenotypic similarities to schizophrenia. Although sequencing of 188 cases revealed no evidence of a non-synonymous or loss of function variant, we identified an intriguing variant with a possible functional effect, and found an association of

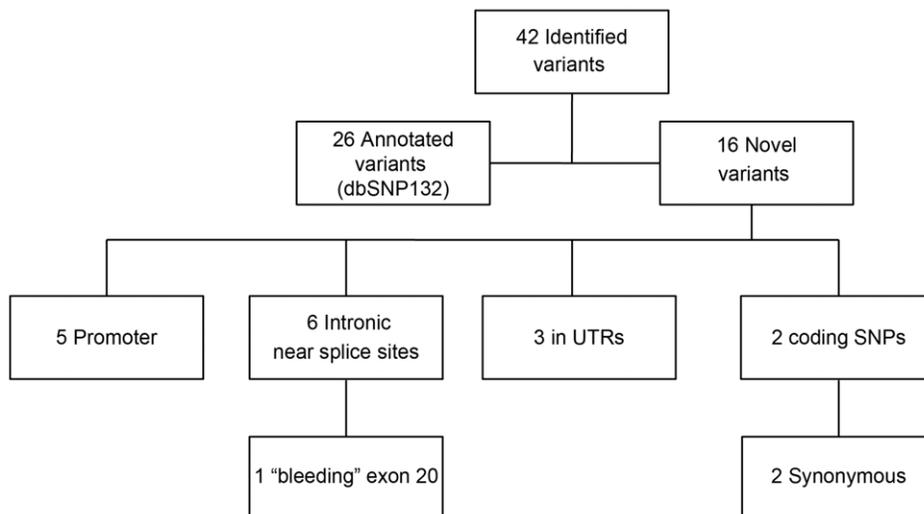


Figure 1. Variants identified by sequencing in *ERBB4*.

doi:10.1371/journal.pone.0020242.g001

Table 1. Description of novel variants identified by sequencing in *ERBB4*.

Novel variant number	Location (bp)	SNP	reference allele	new allele	minor allele frequency	location	Bioinformatic Annotation
1	213,403,857	W	T	A	0.003	5' upstream / in unspliced EST	
2	213,403,847	S	G	C	0.003	5' upstream / in unspliced EST	
3	212,590,021	M	A	C	0.003	intron 102 bp 5' of exon 7	
4	212,566,810	Y	C	T	0.003	exon 12	Synonymous; Loss of exon splice enhancer
5	212,543,982	K	G	T	0.003	intron 73 bp 5' of exon 13	
6	212,488,828	S	C	G	0.003	intron 59 bp 5' of exon 18	
7	212,426,588	R	G	A	0.056	intron 40 bp 3' of exon 20	Adjacent to alternative stop codon
8	212,295,591	A indel	-/-	A/-	0.247	intron 80 bp 3' of exon 21	
9	212,293,120-2	CTT indel	CTT/CTT	CTT/-	0.003	intron 85 bp 5' of exon 22	
10	212,286,804	R	A	G	0.003	exon 24	Synonymous; Gain of exon splice silencer
11	212,251,910	Y	C	T	0.008	intron 35 bp 5' of exon 27	
12	212,251,550	R	G	A	0.003	intron 28 bp 3' of exon 27	
13	212,251,537	M	C	A	0.003	intron 41 bp 3' of exon 27	
14	212,248,283	M	C	A	0.008	3' UTR (57 bp)	
15	212,248,281	S	C	G	0.005	3' UTR (59 bp)	
16	212,248,219	R	G	A	0.003	3' UTR (121 bp)	

doi:10.1371/journal.pone.0020242.t001

this variant with the mood-incongruent psychotic form of BP (dominant *P*-value = 0.039).

There are at least five alternatively spliced transcripts of *ERBB4* documented in the UCSC Genome Browser, including one transcript that ends just after exon 20 and has no associated protein product. This transcript includes a retained 140 bp sequence of intron 20 (hg 19 coordinates 212,426,487–212,426,626), which, as shown in Fig. 2, leads to the transcription of a “bleeding” exon with an alternative stop codon. The newly identified single nucleotide

variant is one base downstream of the alternative stop codon and, if functional, may have an impact on transcription of the shorter and non-functional *ERBB4* isoform. Although this hypothesis remains to be experimentally validated, it raises the possibility that risk alleles may, among other mechanisms, disrupt the normal isoform “balance” of alternatively spliced genes [28,29].

The major psychotic syndromes are likely to be heterogeneous categories with many underlying etiologies. Clinical subphenotypes like mood-incongruent psychosis may help mitigate this

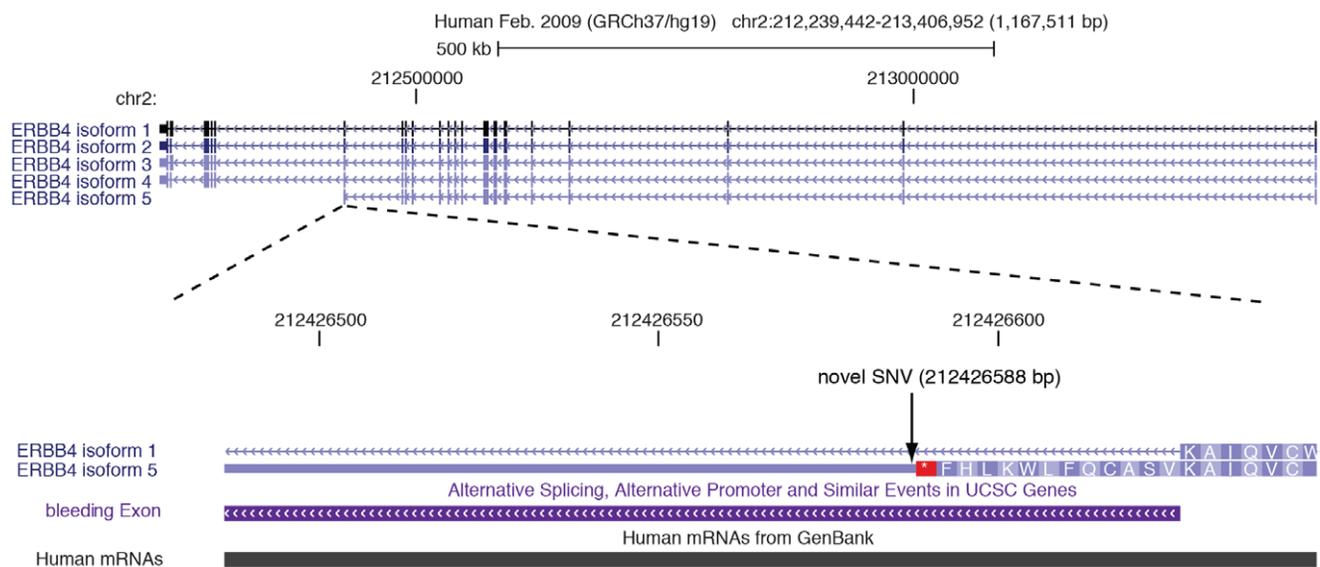


Figure 2. *ERBB4* gene structure with a focus on a novel variant within a “bleeding” exon 20.

doi:10.1371/journal.pone.0020242.g002

Table 2. Association analysis of a novel SNV (chr2:212426588).

Analysis	M.A.F.	Additive Model		Dominant Model	
	(N subjects)	OR	P-value	OR	P-value
Mood-incongruent psychotic BP	5.6% (189)				
vs. Controls	vs. 3.6% (999)	1.64	0.055	1.73	0.039
Mood-incongruent psychotic BP	5.6% (189)				
vs. All other BP cases	vs. 3.6% (806)	1.65	0.063	1.65	0.063

doi:10.1371/journal.pone.0020242.t002

complexity. If the BP phenotype is composed of multiple subphenotypes with partially distinct genetic causes, then any particular genetic variant might contribute to causation of one or several subphenotypes, but not all. Subphenotypes of the disorder might be more informative for purposes like gene mapping if the increased effect size of a risk allele due to genetic homogeneity within the subgroup outweighs the reduction in power from a smaller sample size [30]. In this study, the association of a novel SNP with mood-incongruent BP yielded a larger effect size (OR~1.6–1.7) than those typically seen in GWAS of psychiatric disorders.

This study has several limitations, perhaps the most important being the relatively small sample size. Assuming a binomial distribution, our sample of 188 sequenced cases had an 80% probability to find a mutation with a case frequency >0.4%, which is likely to miss many rare variants, particularly if extensive allelic heterogeneity is present. In our genotyped case-control sample we had 80% power to detect OR \geq 1.9, but only ~55% power to detect ORs in the range found in this study (ORs \geq 1.6). Modest sample sizes such as these are also more likely to produce inflated effect size estimates and are vulnerable to chance findings (type I error). A further limitation which may increase type I error stems from our strategy to sequence cases, while only genotyping potentially functional variants in controls [31], although this is likely lessened by our tested variant being uncommon rather than rare, and by the control sample size being over five times larger than the case sample size. Additional limitations include the limited coverage of the promoter and the 3' UTR sequences and the very restricted coverage of the introns. Given the length of *ERBB4* (1.16 MB) full sequencing of the gene is likely to be feasible only in the context of whole genome sequencing. Finally, we note that although the discovered SNV may potentially affect splicing, experimental validation is necessary to test this hypothesis.

In conclusion, we find no evidence of unambiguous loss of function mutations in 188 cases with mood-incongruent psychotic BP. We discovered a novel variant present in 11% of cases and 6% of controls that may have functional importance, but additional studies are necessary to replicate this association and to study the impact of the variant on splicing of *ERBB4*.

Methods

Ethics Statement

All samples were collected from study participants after obtaining written informed consent under clinical research protocols approved by the Johns Hopkins University School of Medicine institutional review board.

Subjects

Cases were selected from the 1,001 BP cases and 1,033 controls of European-American descent genotyped through the GAIN consortium by the Bipolar Genome Study (BiGS) [21]. All cases were interviewed with the Diagnostic Interview for Genetic Studies (DIGS) and best-estimate diagnoses were made by two research psychiatrists or PhD psychologists. Among the BP cases, we initially selected for sequencing the 189 subjects from the GAIN BP sample who had a lifetime history of mood-incongruent psychosis as previously defined [5]. Briefly, subjects were classified as cases with mood-incongruent psychotic bipolar disorder if they had a lifetime history of running commentary auditory hallucinations, or passivity delusions such as delusions of being controlled, or delusions of thought insertion, withdrawal, or broadcasting. Subjects were also included if their psychotic symptoms during their most severe depression or mania were judged by the interviewer to be “inconsistent” with typical depressive or manic themes. Of the 189 subjects, one subject was sequenced in duplicate, and DNA for one subject was unavailable, leading to a final count of 188 subjects sequenced across the *ERBB4* gene.

In our association study we genotyped all 189 cases as well as 810 non mood-incongruent BP cases and 999 healthy controls from the GAIN BP consortium sample. These controls were previously ascertained using an Internet based adaptation of the Composite International Diagnostic Interview-Short Form (CIDI-SF) [22]. Controls were selected to have no self-reported history of hallucinations, bipolar disorder, or schizophrenia, and no history of sufficient lifetime depressive symptoms to meet DSM-IV criteria for major depressive disorder.

Sequencing

Conventional PCR amplification and Sanger sequencing were performed by Polymorphic DNA technologies Inc. (Alameda, CA, USA). Primers were designed based on the NCBI36/hg18 reference sequence of the longest *ERBB4* transcript (RefSeq NM_005235; CCDS 2394). The sequenced regions included all 28 exons, 1 kb of the promoter region, the 5'UTR, and 400 bp of the 7.9 kb 3'UTR. Sequencing was performed on both strands and chromatograms were aligned and visualized using CodonCode Aligner (CodonCode Corporation, Dedham, MA, USA). One sample was sequenced in duplicate across all PCR amplifications and showed 100% concordance. Sequencing of the promoter region was divided into five PCR amplicons; however, the first three amplicons yielded poor sequence quality in all samples and were excluded from the analysis. The remaining two amplicons (closest to the promoter) yielded approximately 600 bp of high quality sequence.

All novel single nucleotide variants (SNVs) were confirmed either with bidirectional sequencing, or, if the complimentary strand was of poor sequencing quality, with additional genotyping (see below).

Genotyping

Genotyping was performed by pyrosequencing using the PyroMark MD system (QIAGEN). To confirm the novel variants that were seen only on one strand we genotyped nine novel SNVs, validating seven of these nine variants. In the case-control association, we genotyped 995 BP cases (189 with and 806 without mood-incongruent psychosis) and 999 controls from the BiGS study. Among this sample, 23 individuals were genotyped in duplicate and showed 100% concordance.

Association analysis

To account for potential population stratification between cases and controls, we used Eigensoft [23] to derive principal components from the available GWAS data for all samples. Based on a scree plot, we selected the top two principal components to include as covariates in our association analysis. We performed association analyses using additive and dominant models.

Bioinformatic Annotation

Novel variants were visualized in the UCSC genome browser (GRCh37/hg19) with all available annotation tracks. RESCUE-ESE was used to identify potential exonic splicing enhancers [24]. We queried UTR variants for disruption of miRNA binding sites with miRBase [25], and for changes in RNA secondary structure with RNAFold [26].

Probability and power calculations

For the sequenced sample size of 188, we calculated the smallest minor allele frequency (MAF) that could be detected with a probability of 80% using an integral over a cumulative binomial distribution. Power for the association analysis of 188 cases and 999 controls was calculated using the Genetic Power Calculator [27] with the following assumptions: (1) full linkage disequilibrium with the pathogenic mutation; (2) an additive model; and (3) an $\alpha = 0.05$. With the above parameters and a hypothesized effect size of OR = 2.0, our sample had 80% power to detect an association with variants of minor allele frequencies >0.03.

References

- Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, et al. (2009) Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: A population-based study. *Lancet* 373(9659): 234–239.
- Potash JB, Bienvenu OJ (2009) Neuropsychiatric disorders: Shared genetics of bipolar disorder and schizophrenia. *Nat Rev Neurol* 5(6): 299–300.
- Williams HJ, Craddock N, Russo G, Hamshere M, Moskvina V, et al. (2011) Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross traditional diagnostic boundaries. *Hum Mol Genet* 20(2): 387–391.
- Green EK, Raybould R, Macgregor S, Gordon-Smith K, Heron J, et al. (2005) Operation of the schizophrenia susceptibility gene, neuregulin 1, across traditional diagnostic boundaries to increase risk for bipolar disorder. *Arch Gen Psychiatry* 62(6): 642–648.
- Goes FS, Zandi PP, Miao K, McMahon FJ, Steele J, et al. (2007) Mood-incongruent psychotic features in bipolar disorder: Familial aggregation and suggestive linkage to 2p11-q14 and 13q21-33. *Am J Psychiatry* 164(2): 236–247.
- Buonanno A (2010) The neuregulin signaling pathway and schizophrenia: From genes to synapses and neural circuits. *Brain Res Bull* 83(3–4): 122–131.
- Fazzari P, Paternain AV, Valiente M, Pla R, Lujan R, et al. (2010) Control of cortical GABA circuitry development by Nrg1 and ErbB4 signaling. *Nature* 464(7293): 1376–1380.
- Birchmeier C (2009) ErbB receptors and the development of the nervous system. *Exp Cell Res* 315(4): 611–618.
- Mei L, Xiong WC (2008) Neuregulin 1 in neural development, synaptic plasticity and schizophrenia. *Nat Rev Neurosci* 9(6): 437–452.
- Beaulieu JM, Sotnikova TD, Yao WD, Kockeritz L, Woodgett JR, et al. (2004) Lithium antagonizes dopamine-dependent behaviors mediated by an AKT/glycogen synthase kinase 3 signaling cascade. *Proc Natl Acad Sci U S A* 101(14): 5099–5104.
- Hahn CG, Wang HY, Cho DS, Talbot K, Gur RE, et al. (2006) Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nat Med* 12(7): 824–828.
- Silberberg G, Darvasi A, Pinkas-Kramarski R, Navon R (2006) The involvement of ErbB4 with schizophrenia: Association and expression studies. *Am J Med Genet B Neuropsychiatr Genet* 141B(2): 142–148.
- Norton N, Moskvina V, Morris DW, Bray NJ, Zammit S, et al. (2006) Evidence that interaction between neuregulin 1 and its receptor erbB4 increases susceptibility to schizophrenia. *Am J Med Genet B Neuropsychiatr Genet* 141B(1): 96–101.
- Benzel I, Bansal A, Browning BL, Galwey NW, Maycox PR, et al. (2007) Interactions among genes in the ErbB-neuregulin signalling network are associated with increased susceptibility to schizophrenia. *Behav Brain Funct* 3: 31.
- Shiota S, Tochigi M, Shimada H, Ohashi J, Kasai K, et al. (2008) Association and interaction analyses of NRG1 and ERBB4 genes with schizophrenia in a Japanese population. *J Hum Genet* 53(10): 929–935.

Acknowledgments

The authors express their appreciation to the families who participated in this project, and to the many clinicians who facilitated the referral of participants to the study. Genome-wide SNP genotyping of the NIMH samples was performed through the Genetic Association Information Network under the direction of the Bipolar Genome Study (BiGS). The principal investigators and co-investigators were: John R. Kelsoe, Tiffany A. Greenwood, Caroline M. Nievergelt, Rebecca McKinney, Paul D. Shilling, Nicholas Schork, Erin N. Smith, Cinnamon Bloss, John Nurnberger, Howard J. Edenberg, Tatiana Foroud, Daniel L. Koller, Elliot Gershon, Chunyu Liu, Judith A. Badner, William A. Scheftner, William B. Lawson, Evaristus A. Nwulia, Maria Hipolito, William Coryell, John Rice, William Byerley, Francis McMahon, Thomas G. Schulze, Wade Berrettini, James B. Potash, Peter P. Zandi, Pamela B. Mahon, Melvin G. McInnis, Sebastian Zöllner, Peng Zhang, David W. Craig, Szabolcs Szélinger, Thomas B. Barrett.

Data and biomaterials for the NIMH samples were collected as part of 10 projects that participated in the NIMH Bipolar Disorder Genetics Initiative. From 1991 to 1998, the principal investigators and co-investigators were: John Nurnberger, Marvin Miller, Elizabeth Bowman, Theodore Reich, Allison Goate, John Rice, J. Raymond DePaulo Jr., Sylvia Simpson, Colin Stine, Elliot Gershon, Diane Kazuba, Elizabeth Maxwell. From 1999 to 2003, the principal investigators and co-investigators were: John Nurnberger, Marvin J. Miller, Elizabeth S. Bowman, N. Leela Rau, P. Ryan Moe, Nalini Samavedy, Rif El-Mallakh, Hussein Manji, A. Glitz, Eric T. Meyer, Carrie Smiley, Tatiana Foroud, Leah Flury, Danielle M. Dick, Howard Edenberg, John Rice, Theodore Reich, Allison Goate, Laura Bierut, Melvin McInnis, J. Raymond DePaulo Jr., Dean F. MacKinnon, Francis M. Mondimore, James B. Potash, Peter P. Zandi, Dimitrios Avramopoulos, Jennifer Payne, Wade Berrettini, William Byerley, Mark Vawter, William Coryell, Raymond Crowe, Elliot Gershon, Judith Badner, Francis McMahon, Chunyu Liu, Alan Sanders, Maria Caserta, Steven Dinwiddie, Tu Nguyen, Donna Harakal, John R. Kelsoe, Rebecca McKinney, William Scheftner, Howard M. Kravitz, Diana Marta, Annette Vaughn-Brown, Laurie Bederow, Layla Kassem, Sevilla Detera-Wadleigh, Lisa Austin, Dennis L. Murphy. Control subjects from the National Institute of Mental Health Schizophrenia Genetics Initiative (NIMH-GI), data and biomaterials were collected by the “Molecular Genetics of Schizophrenia II” (MGS-2) collaboration. The investigators and co-investigators are: Pablo V. Gejman, Alan R. Sanders, Farooq Amin, Nancy Buccola, William Byerley, C. Robert Cloninger, Raymond Crowe, Donald Black, Robert Freedman, Douglas Levinson, Bryan Mowry, Jeremy Silverman.

Author Contributions

Conceived and designed the experiments: FSG JBP. Performed the experiments: FSG MR. Analyzed the data: FSG Y-CC RK EE. Wrote the paper: FSG JBP.

16. Nicodemus KK, Law AJ, Radulescu E, Luna A, Kolachana B, et al. (2010) Biological validation of increased schizophrenia risk with *NRG1*, *ERBB4*, and *AKT1* epistasis via functional neuroimaging in healthy controls. *Arch Gen Psychiatry* 67(10): 991–1001.
17. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, et al. (2009) Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 460(7256): 753–757.
18. Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, et al. (2008) Collaborative genome-wide association analysis supports a role for *ANKK3* and *CACNA1C* in bipolar disorder. *Nat Genet* 40(9): 1056–1058.
19. Cirulli ET, Goldstein DB (2010) Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* 11(6): 415–425.
20. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, et al. (2008) Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 320(5875): 539–543.
21. Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, et al. (2009) Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* 14(8): 755–763.
22. Sanders AR, Levinson DF, Duan J, Dennis JM, Li R, et al. (2010) The internet-based MGS2 control sample: Self report of mental illness. *Am J Psychiatry* 167(7): 854–865.
23. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2(12): e190.
24. Fairbrother WG, Yeh RF, Sharp PA, Burge CB (2002) Predictive identification of exonic splicing enhancers in human genes. *Science* 297(5583): 1007–1013.
25. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: Tools for microRNA genomics. *Nucleic Acids Res* 36(Database issue): D154–8.
26. Bindewald E, Shapiro BA (2006) RNA secondary structure prediction from sequence alignments using a network of k-nearest neighbor classifiers. *RNA* 12(3): 342–352.
27. Purcell S, Cherny SS, Sham PC (2003) Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19(1): 149–150.
28. Johnson MB, Kawasawa YI, Mason CE, Krsnik Z, Coppola G, et al. (2009) Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* 62(4): 494–509.
29. Wang GS, Cooper TA (2007) Splicing in disease: Disruption of the splicing code and the decoding machinery. *Nat Rev Genet* 8(10): 749–761.
30. Potash JB, Toolan J, Steele J, Miller EB, Pearl J, et al. (2007) The bipolar disorder phenome database: A resource for genetic studies. *Am J Psychiatry* 164(8): 1229–1237.
31. Li B, Leal SM (2009) Discovery of rare variants via sequencing: Implications for the design of complex trait association studies. *PLoS Genet* 5(5): e1000481.