European collaborative study of early-onset bipolar disorder: Evidence for genetic heterogeneity on 2q14 according to age at onset.

Flavie Mathieu, Marie-Hélène Dizier, Bruno Etain, Stéphane Jamain, Marcella Rietschel, Wolfgang Maier, Margot Albus, Patrick Mckeon, Siobhan Roche, Douglas Blackwood, et al.

European Collaborative Study of Early-Onset Bipolar Disorder: Evidence for genetic heterogeneity

on 2q14 according to age at onset.

Flavie Mathieu,[1,2*] Marie-Hélène Dizier,[3,4,5] Bruno Etain,[1,2,6] Stéphane Jamain,[1,2] Marcella Rietschel,[7] Wolfgang Maier,[8] Margot Albus,[9] Patrick McKeon,[10] Siobhan Roche,[11] Douglas Blackwood,[12] Walter J. Muir,[12] Chantal Henry,[1,2,6,13] Alain Malafosse,[14] Martin Preisig,[15] François Ferrero,[16] Sven Cichon,[17] Johannes Schumacher,[18] Stephanie Ohlraun,[8] Peter Propping,[18] Rami Abou Jamra,[18] Thomas G. Schulze,[7,21] Diana Zelenica,[19] Céline Charon,[19] Andrej Marusic,[20] Mojca C. Dernovsek,[20] Hugh Gurling,[22] Markus Nöthen,[18] Mark Lathrop,[19] Marion Leboyer,[1,2,6,13] and Frank Bellivier,[1,2,6,13]

[§] Mathieu F and Dizier MH contributed to this work equally.

1) INSERM, U 955, IMRB, Dept of Medical Genetics, Psychiatry Genetics, Creteil, F-94010, France

2) Fondation FondaMental, Creteil, France

3) INSERM, U946, Paris, France

4) Université Paris Diderot, Paris 7, Institut Universitaire d'Hématologie, Paris, France

5) Fondation Jean Dausset, Centre d'Etude du Polymorphisme Humain (CEPH), Paris, France

6) University Paris-Est, Faculty of Medicine, IFR10, Creteil, F-94010, France

7) AP-HP, Henri Mondor-Albert Chenevier Group, Department of Psychiatry, Creteil, F-94010, France

8) Central Institute of Mental Health - Division Genetic Epidemiology in Psychiatry - Mannheim - Germany

9) Department of Psychiatry - University of Bonn – Bonn - Germany

10) District Hospital Haar - Haar - Germany

11) Department of Psychiatry - St. Patrick's Hospital - Dublin - Ireland

12) Smurfit Institute of Genetics - Trinity College - Dublin 2 - Ireland

13) University of Edinburgh - Division of Psychiatry - Royal Edinburgh Hospital - Edinburgh - UK

14) Department of Neuropsychiatry - Belle-Idée Hospital - Geneva – Switzerland

15) Department of psychiatry - University Hospital - Lausanne - Switzerland

16) Department of psychiatry - University Hospital - Geneva - Switzerland

17) Life & Brain Center - Department of Genomics - University of Bonn and Institute of Human Genetics - Bonn - Germany

18) Institute of Human Genetics - University of Bonn - Bonn - Germany

19) Centre National de Génotypage - Evry - France

20) Institute of Public Health of the Republic of Slovenia - Ljubjana – Slovenia

21) Department of Psychiatry and Psychotherapy, Georg-August-Universität Göttingen, Göttingen, Germany

22) Molecular Psychiatry Laboratory - Department of Psychiatry and Behavioural Sciences - Windeyer Institute for Medical Sciences - Royal Free and University College London Medical School - London - UK

Running title: Genetic heterogeneity of bipolar disorder

Corresponding author :

Flavie MATHIEU, PhD.

INSERM Unité 955. EQ 15

Hôpital Chenevier. Bâtiment Hartmann.

40 rue de Mesly 94000 Créteil. France

Tel : (+33) 1 49 81 32 79 / Fax: (+ 33) 1 49 81 30 99

e-Mail : Flavie.Mathieu@inserm.fr

ABSTRACT

- 3 -

Bipolar disorder has a genetic component, but the mode of inheritance remains unclear. A previous genome scan conducted in 70 European families led to detect 8 regions linked to bipolar disease. Here, we present an investigation of whether the phenotypic heterogeneity of the disorder corresponds to genetic heterogeneity in these regions using additional markers and an extended sample of families. The MLS statistic was used for linkage analyses. The Predivided Sample Test and the Maximum Likelihood Binomial methods were used to test genetic homogeneity between early onset bipolar type I (cut-off of 22 years) and other types of the disorder (later onset of bipolar type I and early onset bipolar type II), using a total of 138 independent bipolar-affected sib-pairs. Analysis of the extended sample of families supports linkage in four regions (2q14, 3p14, 16p23 and 20p12) of the eight regions of linkage suggested by our previous genome scan. Heterogeneity testing revealed genetic heterogeneity between early and late onset bipolar type I in the 2q14 region (p=0.0001). Only the early form of the bipolar disorder but not the late form appeared to be linked to this region. This region may therefore include a genetic factor either specifically involved in the early onset bipolar type I or only influencing the age at onset. Our findings illustrate that stratification according to age at onset may be valuable for the identification of genetic vulnerability polymorphisms.

Keywords: Bipolar disorder, genetic heterogeneity, age at onset.

INTRODUCTION

Although bipolar disorder (BP) has a genetic component, its mode of inheritance remains obscure and no vulnerability gene has been clearly identified. Promising loci for BP have been identified by candidate gene studies and numerous regions have been implicated by several genome wide linkage scans (Serretti and Mandelli, 2008). However, such findings have not always been replicated (Craddock and Forty, 2006), and meta-analysis of 18 genome scans on bipolar disorder failed to yield statistically significant linkage results (Segurado et al., 2003). Recently, three groups have performed independent whole genome association studies of bipolar disorder (WTCCC, 2007; Baum et al., 2008; Sklar et al., 2008), but the studies agreed only for a limited number of regions and for the ankyrin 3 (ANK3) (Ferreira et al., 2008; O'Donovan et al., 2009; Smith et al., 2009; Schulze et al., 2009; Scott et al., 2009; Moskvina et al., 2009) gene and alpha 1C subunit of the voltage-dependent calcium channel (CACNA1C) (Ferreira et al., 2008; O'Donovan et al., 2009) genes. This may reflect clinical and genetic heterogeneity of bipolar disorder and lack of a consensus definition for the affected phenotype (Leboyer et al., 1998; Leboyer, 2003). Three different phenotype definitions have commonly been used in genetic studies of BP: a narrow phenotype including BP type I (BPI) characterized by clear manic episodes; an intermediate phenotype including BPI and BP type II (BPII, involving milder forms of mania); and a broad phenotype including BPI, BPII and Major Depressive Episode (MDE).

A 'symptom candidate approach' has been proposed to disentangle the genetic and clinical heterogeneity of BP, and age at onset (AAO) has proved to be a relevant candidate symptom (Leboyer et al., 1998). Several admixture analyses of AAO demonstrated the existence of three sub-groups, showing a consistent cut-off of 22 years to define early onset forms of BP (Bellivier et al., 2001; Bellivier, 2003; Lin et al., 2006; Hamshere et al., 2008; Manchia et al., 2008). Clinical studies have shown that early-onset BP (EO-BP) is a more severe form of the disorder than later onset forms, and is characterized by more mixed episodes, greater overall psychiatric comorbidity, more lifetime psychotic symptoms and poorer response to prophylactic lithium (Schurhoff et al., 2000; Suominen K et al., 2007). A greater heritability of early-onset than late-onset forms of the disorder has been demonstrated (Weissman et al. 1984, Faraone et al. 2003). Moreover, the pattern of disease inheritance seems to differ between early- and late-onset BP families (Grigoroiu-Serbanescu et al. 2001), suggesting that different genetic mechanisms might be involved.

Through a European Collaborative Study (France, Germany, Ireland, Scotland, Switzerland, England, Slovenia) (Etain et al., 2006), we previously carried out a genome-wide linkage analysis. Two phenotypes were analyzed: a "narrow phenotype" (EO-BPI only) and a "broad phenotype" (including as affected, siblings suffering from either BPI or BPII regardless of AAO, or Major Depressive Episode (MDE) with an early onset). This study suggested linkage for five regions using the broad phenotype (2p21, 2q14.3, 7q36, 10q23 and 20p12); for two regions (5q33 and 16q23) using the narrow phenotype and for one region (3p14) using both narrow and broad phenotypes. These results suggest that the genetic factors involved in the EO-BPI differ from those involved in the other forms of the disease. Genetic heterogeneity according to AAO has already been considered but only in two studies (Faraone et al., 2004; Zandi et al., 2007) and has never been formally tested.

Our aims were to investigate whether phenotypic heterogeneity of the disorder reflects genetic heterogeneity in the eight regions described in our previous study (Etain et al., 2006) using additional markers and an extended sample of families. The MLS statistic was used for linkage analyses. We used two different statistical approaches to test genetic homogeneity: the Predivided Sample Test (PST) (24) and the Maximum Likelihood Binomial (MLB) method (25, 26). This study is the first formal analysis of the relationship between genetic heterogeneity in bipolar disorder and phenotypic heterogeneity, including AAO and severity (BPII *versus* BPI).

Materials and methods

Subjects

Families were recruited as part of the European Collaborative Study of Early-Onset Bipolar Disorder (see Table I). The ascertainment scheme is detailed elsewhere (Etain et al., 2006). Families were included on the basis of a BP type I proband who suffered one or more thymic episodes before the age of 22 years, and an affected sibling suffering from either BP type I or II (regardless of AAO) or single or recurrent early-onset major depressive episode. In addition to the previously reported sample of 69 families, we include 51 further families (one from Germany, 29 from Spain, four from France, ten from Scotland and seven from Ireland), leading to a total sample size of 120 families. Similar ascertainment criteria were

applied to all families. Informed consent was obtained from all participants. The study complied with all ethical guidelines of the institutions involved in each participating country.

Phenotypes analyzed

Age at-onset of BP was defined as the age of the first thymic episode (depressive, manic or hypomanic, determined by reviewing medical case notes and information from semi-structured interviews). Early AAO (EO) was defined as onset before age 22 years. Siblings were considered as affected for the "broad phenotype" if they had BP type I or II or early-onset major depressive episode (MDE). For the "narrow phenotype", only siblings suffering from EO-BPI were considered to affected, in the same way as for probands. As patients classified at the time of the study as having had a major depressive episode may subsequently suffer from single or recurrent early-onset major depressive episode, or become BPI or BPII, we chose to exclude siblings with this phenotype from the heterogeneity analysis. As only two siblings suffered from late onset BPII, they were also excluded. Thus, for the heterogeneity analyses, we used only three phenotypes: EO-BPI, LO-BPI and EO-BPII.

Genotyping

Genomic DNA was isolated from peripheral blood leukocytes using a phenol/chloroform method. The initial genome screen included 384 highly polymorphic microsatellite markers (STR for Short Tandem Repeat, see Etain et al., 2006 for details). All families were genotyped for 91 additional STRs in the eight linkage regions. New pedigrees were not genotyped for all genome-scan markers but only for STRs in the eight regions included in this analysis. Marker locations are reported in Kosambi centiMorgans (cM), according to the Marshfield map (http://research.marshfieldclinic.org/). The mean distance between markers was 1.89 (range 1.2 cM to 3 cM). Genotyping was performed at the Centre National de Génotypage (CNG - France).

Statistical methods

Maximum Likelihood Score (MLS) method

Sib-pair linkage analysis assesses the sharing by affected sib-pairs of identical-by-descent (IBD) marker alleles. If the observed IBD distribution differs significantly from that expected under the null hypothesis

of independent segregation of the disease and markers, it provides evidence of linkage. Linkage analysis was carried out using the MLS statistic (Risch, 1990), assuming that the IBD proportions, z_2, z_1 and z_0, are constrained within a triangle defined by $2z_0 \leq z_1 \leq 0.5$ referred to as the triangle constraints (Suarez et al., 1978; Holmans, 1993). However, if the phenotype of each member of a sib-pair is determined by different genetic models, *i.e.* corresponding to different genotype relative risks, the triangle constraints may not be valid (Dizier et al., 2000). This can occur when the two sibs differ for a factor, such as a characteristic of the disease, *e.g.* severe *vs.* mild form, which modifies the genotype relative risks for the disease. We used a p-value of 0.01 as the cut-off and p-values between 0.0001 and 0.01 were considered to provide indication of linkage.

The Predivided Sample Test (PST)

Let us consider two sub-phenotypes, A and A'. Under the null hypothesis of genetic homogeneity (*i.e.* the same genotype relative risks for the two sub-phenotypes), the Z (z_2, z_1, z_0) vectors are expected to be equal for all types of pairs, *i.e.* concordant pairs (AA and A'A'), and discordant pairs (AA'). Therefore this null hypothesis can be tested by the PST statistic as proposed by Morton and Steinberg (Morton and Steinberg, 1956): $-2[Ln(L(Z_{AA}) \cdot L(Z_{AA'}) \cdot L(Z_{A'A'})/L(Z))]$ where $L(Z_{AA})$, $L(Z_{AA'})$, $L(Z_{A'A'})$ and $L(Z)$ are the likelihoods of the parameter vectors Z, estimated in each sample of pairs (AA, AA' and A'A') and the whole sample, respectively. The Z vectors were estimated here without constraints to allow departure from the triangle constraint among discordant sib-pairs in the case of heterogeneity. In the case of three phenotypic categories of sib-pairs, the PST statistic would follow a χ^2 with four degrees of freedom (df).

Maximum Likelihood Binomial for ordered categorical traits

The Maximum Likelihood Binomial (MLB) method is a linkage model-free approach which is likelihood-based and can be applied to the whole sibship of affected subjects (Majumder and Pal, 1987; Satsangi et al., 1996; Abel L and Muller-Myhsok, 1998). This method is based on the binomial distribution of the number of affected sibs receiving a given parental allele from heterozygous parents. The probability α for an affected sib to receive from his/her parent the marker allele transmitted with the disease allele is equal to 0.5 under the null hypothesis of no linkage; α is greater than 0.5 under the hypothesis of linkage. The

test for linkage is performed using a likelihood ratio test statistic, $\Lambda=2Ln[L(\alpha)/L(\alpha=0.5)]$ with the statistic Λ being distributed asymptotically as a mixture distribution of 0.5 χ^2_{0df} and 0.5 χ^2_{1df} or expressed as a one-sided standard normal deviate denoted $Z_{MLB}=\Lambda^{\frac{1}{2}}$. This method has been extended to categorical traits by introducing a latent binary variable ($Y=\{0;1\}$) which captures the linkage information between the observed categorical phenotype (Z) and the marker (M) (Alcais and Abel, 1999). The method requires the assignment of the probability of the latent variable (*i.e*. being affected or not) according to each observed category of the phenotype. MLBGH (Abel L and Muller-Myhsok, 1998) which incorporates the MLB method into the multipoint approach of Genehunter (Kruglyak et al., 1996) was used for multipoint linkage analyses of polychotomous phenotypes. Note that the MLB method for categorical traits can detect linkage accounting for the presence of heterogeneity, and that the PST method can detect both linkage and heterogeneity.

Analysis strategy

Linkage analyses were conducted with the MLS using the triangle constraints, in the whole sample of families and using all available genotypes in the eight regions. For multiple sibships, all independent pairs of affected sibs were considered. Families were ascertained through an EO-BPI proband, and therefore we focused on the detection of genetic heterogeneity between EO-BPI and other forms of the disorder (LO-BPI or EO-BPII).

To limit the number of tests, heterogeneity testing was only applied to markers giving a p-value lower than or equal to 0.05 with the MLS statistic, using either the broad or the narrow phenotypic definition. Two different tests were conducted with the PST, each test contrasting concordant pairs (both sibs being EO-BPI) to discordant-pairs (first sib being EO-BPI and second sib being either LO-BPI for the first test or EO-BPII for the second test). If one or both tests provided evidence of heterogeneity ($p\leq0.005$), we also conducted the PST contrasting the concordant pairs to the pooled discordant pairs. Because only two categories of sib-pairs were considered, the degree of freedom of the PST was equal to 2. Note that the PST was applied using the MLS without constraints for both concordant and discordant sib-pairs, although departure from the Holmans triangle was expected only among discordant sib-pairs. Thus, for markers showing a significant heterogeneity with the PST, we verified that the p-value ≤0.05 for concordant pairs for BP1-EO using MLS with the triangle constraints.

Using the MLB method, we considered three models of ordered category trait, each model including three categories:

Model 1: EO-BPI > LO-BPI > UNAFFECTED (UA)

Model 2: EO-BPI > EO-BPII> UA

Model 3: EO-BPI > (LO-BPI + EO-BPII)> UA

For the three models, the corresponding probabilities of being affected for the unobserved binary trait given the three ordered categories were fixed to 1, 0.5 and 0.

RESULTS

Description of the sample

A total of 120 nuclear families ascertained through an early onset BPI proband were analyzed. 46.7% of the parents were genotyped. The initial sample of 69 families included 33 independent affected sib-pairs with the narrow phenotype-definition and the additional sample of 51 families included 20 such sib-pairs. Using broad definition the numbers were 76 and 62 sib-pairs respectively. Thus, 53 narrow and 138 broad sib-pairs were used for analysis.

The phenotypic distribution was similar in the initial and the additional samples (p=0.44): the proportions of EO-BPI (p=0.18), LO-BPI (p=0.66) and EO-BPII (p=0.85) siblings did not differ between the two samples (Table I). All the 138 independent affected sib-pairs were genotyped: 53 siblings were EO-BPI (corresponding to the narrow phenotype), 45 were LO-BPI, 20 were EO-BPII, 18 were EO-MDE and two were LO-BPII. In the whole sample, the mean AAO was 17.2 years for probands and 21.2 years for affected siblings. The sex-ratio (male/female) was 69% for probands and 67.1% for siblings. The distributions of AAO and the sex-ratio between probands and siblings did not differ between the initial and the additional family samples (Table I).

Non parametric linkage analysis in regions of linkage

Genotyping of an additional set of linkage markers led to a mean information content of 90% (*versus* 0.80 in our initial genome screen) in our eight regions of linkage. Linkage analysis of the initial sample with all markers gave results very similar to those reported previously (results not shown). Multipoint linkage analysis for the broad and narrow phenotypes in the extended sample of families is

described in Table 2. For the broad phenotype, linkage was indicated for regions 2q14 (p=0.001), 3p14 (p=0.005) and 20p12 (p=0.007) as in the initial scan; by contrast, no evidence of linkage was found for regions 2p21, 7q36 and 10q23. For the narrow phenotype, linkage was indicated for regions 3p14 (p=0.006) and 16p23 (p=0.004), as in the initial scan. By contrast, there were indications of linkage for the region 2q14 (p=0.004), which had not been identified in the initial genome scan. Unlike previously, no linkage was found for region 5q33 with the narrow phenotype in the whole sample.

Heterogeneity testing

The PST and MLB methods were used to search for genetic heterogeneity for markers in regions of linkage showing a p≤0.05 for narrow and/or broad phenotypes in the whole sample of families. Thirty-six markers fulfilled this criterion. Using the PST, two regions show heterogeneity with a p-value lower than or equal to 0.005 (Table 3). The strongest linkage heterogeneity was detected for the 2q14 region between the phenotypes EO-BPI and LO-BPI (p=0.0001). In contrast, no heterogeneity was found between EO-BPI and EO-BPII in this region. The allele sharing proportion was higher between concordant EO-BPI pairs (0.58) than between BPI pairs discordant for AAO (0.28). There was also some indication of heterogeneity in the region 16p23, between EO-BPI and other forms of the disease (LO-BPI + EO-BPII). The categorical MLB method (Table 4) gave results consistent with those obtained using the PST, in particular confirming the heterogeneity of the 2q14 region. Linkage was detected (p=0.005) for the categorical model EO-BPI > LO-BPI > UA, but not for the model EO-BPI > EO-BPII > UA.

DISCUSSION

Of the eight regions of linkage suggested by our previous genome scan, analysis of the extended sample of families provided indication of linkage to four regions (2q14, 3p14, 16p23 and 20p12). The 3p14 region, detected here, has been reported in several other genome scans (McInnes et al., 1996; Cichon et al., 2001; Kelsoe et al., 2001; Radhakrishna et al., 2001; Ewald et al., 2003; Fallin et al., 2004). However, the 5q33 region, showing linkage in numerous independent reports (Edenberg et al. 1997, Cichon et al., 2001; Kelsoe et al., 2001; McInnis et al., 2003; Herzberg et al. 2006; Jones et al., 2007) did not show linkage in our sample (p=0.04). Our results for the 20p12 region are consistent with several

studies (McInnis et al., 2003; Fallin et al., 2004; Jones et al., 2007; Etain et al., 2009). By contrast, the 7q36 region has been reported in only two scans (Jones et al., 2007; Cassidy et al., 2009). As far as we know, the 16p23 region has never previously been reported in bipolar genome scan.

Indication of linkage was shown here with the 2q14 region for both the narrow and the broad phenotypes in the extended sample of families although no evidence for linkage was found for the narrow phenotype in the initial sample of families. The initial sample included only 29 concordant EO-BPI sib-pairs, whereas 53 were included in this study. The region has been identified in three previous genome scans (Dick et al. 2003; Fallin et al., 2004; Pato et al., 2004) and by the meta-analysis reported by Segurado et al (Segurado et al., 2003); note that none of these studies considered AAO in their analyses. One of the four genome-wide association studies conducted for bipolar disorder (WTCCC, 2007; Baum et al., 2008, Ferreira et al., 2008, Sklar et al., 2008), the Wellcome Trust Case control consortium, reported moderate evidence for association of bipolar disorder in the 2q14 region ($p=10^{-6}$). This region however gave no or only weak signals for association in the three other studies. Similarly to previous linkage studies, AAO was not taken into account in these genome-wide association studies. Recently, Le Niculescu et al. proposed a convergent genetic approach, integrating the results from GWA studies, functional genomics and animal models, and suggested the Dpp10 gene as a relevant candidate gene in this region (Le-Niculescu et al., 2009). Note that, although not yet tested, the 2q14 region also contains other candidate genes, particularly CNTNAP5, GPR7 and HS6ST1.

Only two previous genome-wide linkage studies of BP considered AAO (Faraone et al., 2004; Zandi et al., 2007). Faraone et al. carried out linkage analysis using AAO as a quantitative trait in BPI families and concluded that genes localized in 12p, 14q and 15q regions influenced AAO of mania and were distinct from those influencing the liability to develop BP. Zandi et al (Zandi et al., 2007) reported the 3q28 region in a genome-wide linkage analysis of BP adjusted for AAO. Thus, there are discrepancies between the findings of these two earlier studies that took AAO into account and our study. There are several methodological issues that may explain these discrepancies. There are differences in the statistical methods used in the three studies: the variance-component method for Faraone's study, the conditional logistic linkage approach in Zandi's study, and the MLS, PST and MLB statistics used here. In addition,

in Faraone's study, AAO was defined as the age at the first manic episode and was analyzed as a quantitative trait, whereas Zandy and colleagues, like us, defined AAO as the age at the first thymic episode, whether manic, hypomanic or depressive, and analyzed it as a binary trait (using a threshold of 21 years old). The modes of ascertainment used by these studies also differ: our families were ascertained through an EO-BPI proband and an affected sibling suffering either from BPI or BPII (regardless of AAO) or EO-MDE, whereas the other studies included families with two BPI patients (Faraone et al., 2004) or with a BPI proband and at least two first-degree relatives affected by BP (Zandi et al., 2007).

Heterogeneity testing allowed the identification of genetic heterogeneity between EO-BPI and LO-BPI in the 2q14 region, suggesting linkage heterogeneity according to the AAO. The excess sharing of parental alleles between pairs concordant for EO-BPI but not between discordant sib-pairs (one sib with EO BPI, the other LO), indicated linkage of this region to only the early form of the disorder. The failure to detect heterogeneity between EO-BPI and EO-BPII could be a consequence of a lack of power due to the small number of discordant pairs, or of a dependence of the two forms of the disease (EO-BPI and EO-BPII) on the same genetic factor in 2q14 region. The MLB results were consistent with heterogeneity in the 2q14 region: this approach supports linkage to the 2q14 region allowing for heterogeneity under the model EO-BPI > LO-BPI > UA. The significance level was lower than with the PST, the power of the two methods being dependent on the underlying model. In our sample, the PST method detected linkage in the 2q14 region with greater significance than found with classical linkage analysis (p=0.0001 *versus* p=0.001, without correction for multiple testing). We used a relatively stringent threshold (p ≤ 0.005) to select regions because of the increase of type one error due to multiple testing. We conducted tests for 36 markers, for two independent models (EO-BPI *versus* LO-BPI and EO-BPI *versus* EO-BPII), and with both PST and MLB methods. A total of 144 tests were conducted, but not all the tests were independent. Although Bonferroni's correction is known to be over-conservative for non-independent tests, the results of the PST analysis in the 2q14 region using this correction remained significant (p=0.01). By contrast, the result for the 16p23 region was not significant after Bonferroni's correction.

Our results clearly illustrated the value of such heterogeneity testing for detecting linkage, and particularly its greater power to detect linkage accounting for possible genetic heterogeneity when it

exists. Moreover, even without large samples, this approach allowed the detection of significant genetic heterogeneity in the 2q14 region according to AAO of BPI.

These various results indicate that a genetic factor located in the 2q14 region is either specifically involved in the early onset BP, or is involved in only the AAO acting like a modifier gene (a gene which modifies the effect of another gene). As BP is a multi-factorial disease, it is difficult to prove that the gene in 2q14 influences the AAO and thereby modifies the effect of another gene; numerous genes are involved in BP and they are likely to interact. We could have conducted heterogeneity tests, controlling for the other linked regions (by stratifying between families linked or not linked to another region), but our sample size is not large enough, the resulting groups would be too small for satisfactory statistical analysis. For the other linkage regions detected in this study, no heterogeneity was detected according AAO. This suggests that there are also some genetic factor(s) shared by early and late onset BP. The existence of such shared genetic component(s) is supported by the presence of late onset BP individuals among sibships of early onset BP probands.

A major issue in the identification of susceptibility genes for BP disorder is the choice of the heritable phenotype to study. BPI, BPII and MDE are defined as distinct entities in DSM-IV, which allows a high inter-rater concordance but which was not constructed to distinguish between different clinical entities depending on distinct genetic factors. Most genome-wide linkage scans have analyzed broad, intermediate and narrow phenotypes using a gradient from MDE to BPI through BPII (Garner et al., 2001; McInnis et al., 2003; Fallin et al., 2004; Schumacher et al., 2005; Zandi et al., 2007, Vazza et al., 2007; Nwulia et al., 2007; Venken et al., 2007). This phenotypic classification may not be the most appropriate for defining phenotypes for the identification of linkage regions. Indeed, our results for the 2q14 region show that stratification according to AAO may be more relevant than DSM-IV classifications of mood disorders. Further genome-wide association studies including candidate genes in this region will help shed light on the various etiological pathways involved in BP according to the AAO.

REFERENCES

Abel L, Alcais A, Mallet A. Comparison of four sib-pair linkage methods for analyzing sibships with more than two affecteds: interest of the binomial maximum likelihood approach. Genet Epidemiol 1998; 15(4): 371-390.

Abel L, Muller-Myhsok B. Robustness and power of the maximum-likelihood-binomial and maximum-likelihood-score methods, in multipoint linkage analysis of affected-sibship data. Am J Hum Genet 1998 Aug; 63(2): 638-647.

Alcais A, Abel L. Maximum-Likelihood-Binomial method for genetic model-free linkage analysis of quantitative traits in sibships. Genet Epidemiol 1999; 17(2): 102-117.

Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry 2008 Feb; 13(2): 197-207.

Bellivier F, Golmard JL, Henry C, Leboyer M, Schurhoff F. Admixture analysis of age at onset in bipolar I affective disorder. Arch Gen Psychiatry 2001 May; 58(5): 510-512.

Bellivier F. Genetic association studies. Definition of cases and controls. Methods Mol Med 2003; 77: 127-141.

Cassidy F, Zhao C, Badger J, Claffey E, Dobrin S, Roche S et al. Genome-wide scan of bipolar disorder and investigation of population stratification effects on linkage: support for susceptibility loci at 4q21, 7q36, 9p21, 12q24, 14q24, and 16p13. Am J Med Genet B Neuropsychiatr Genet 2007 Sep 5; 144B(6): 791-801.

Cichon S, Schumacher J, Muller DJ, Hurter M, Windemuth C, Strauch K et al. A genome screen for genes predisposing to bipolar affective disorder detects a new susceptibility locus on 8q. Hum Mol Genet 2001 Dec 1; 10(25): 2933-2944.

Craddock N, Forty L. Genetics of affective (mood) disorders. Eur J Hum Genet 2006 Jun; 14(6): 660-668.

Dick DM, Foroud T, Flury L, Bowman ES, Miller MJ, Rau NL et al. Genomewide linkage analyses of bipolar disorder: a new sample of 250 pedigrees from the National Institute of Mental Health Genetics Initiative. Am J Hum Genet 2003 Jul; 73(1): 107-114.

Dizier MH, Quesneville H, Prum B, Selinger-Leneman H, Clerget-Darpoux F. The triangle test statistic (TTS): a test of genetic homogeneity using departure from the triangle constraints in IBD distribution among affected sib-pairs. Ann Hum Genet 2000 Sep; 64(Pt 5): 433-442.

Dmitrzak-Weglarz M, Rybakowski JK, Slopien A, Czerski PM, Leszczynska-Rodziewicz A, Kapelski P et al. Dopamine receptor D1 gene -48A/G polymorphism is associated with bipolar illness but not with schizophrenia in a Polish population. Neuropsychobiology 2006; 53(1): 46-50.

Edenberg HJ, Foroud T, Conneally PM, Sorbel JJ, Carr K, Crose C et al. Initial genomic scan of the NIMH genetics initiative bipolar pedigrees: chromosomes 3, 5, 15, 16, 17, and 22. Am J Med Genet 1997 May 31; 74(3): 238-246.

Etain B, Dumaine A, Mathieu F, Chevalier F, Henry C, Kahn JP et al. A SNAP25 promoter variant is associated with early-onset bipolar disorder and a high expression level in brain. Mol Psychiatry 2009 Jan 6.

Etain B, Mathieu F, Rietschel M, Maier W, Albus M, McKeon P et al. Genome-wide scan for genes involved in bipolar affective disorder in 70 European families ascertained through a bipolar type I early-onset proband: supportive evidence for linkage at 3p14. Mol Psychiatry 2006 Jul; 11(7): 685-694.

Ewald H, Kruse TA, Mors O. Genome wide scan using homozygosity mapping and linkage analyses of a single pedigree with affective disorder suggests oligogenic inheritance. Am J Med Genet B Neuropsychiatr Genet 2003 Jul 1; 120B(1): 63-71.

Fallin MD, Lasseter VK, Wolyniec PS, McGrath JA, Nestadt G, Valle D et al. Genomewide linkage scan for bipolar-disorder susceptibility loci among Ashkenazi Jewish families. Am J Hum Genet 2004 Aug; 75(2): 204-219.

Faraone SV, Glatt SJ, Su J, Tsuang MT. Three potential susceptibility loci shown by a genome-wide scan for regions influencing the age at onset of mania. Am J Psychiatry 2004 Apr; 161(4): 625-630.

Faraone SV, Glatt SJ, Tsuang MT. The genetics of pediatric-onset bipolar disorder. Biol Psychiatry 2003 Jun 1; 53(11): 970-977.

Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 2008 Sep; 40(9): 1056-1058.

Garner C, McInnes LA, Service SK, Spesny M, Fournier E, Leon P et al. Linkage analysis of a complex pedigree with severe bipolar disorder, using a Markov chain Monte Carlo method. Am J Hum Genet 2001 Apr; 68(4): 1061-1064.

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007 Jun 7; 447(7145): 661-678.

Grigoroiu-Serbanescu M, Martinez M, Nothen MM, Grinberg M, Sima D, Propping P et al. Different familial transmission patterns in bipolar I disorder with onset before and after age 25. Am J Med Genet 2001 Dec 8; 105(8): 765-773.

Hamshere ML, Gordon-Smith K, Forty L, Jones L, Caesar S, Fraser C et al. Age-at-onset in bipolar-I disorder: Mixture analysis of 1369 cases identifies three distinct clinical sub-groups. J Affect Disord 2008 Dec 3.

Herzberg I, Jasinska A, Garcia J, Jawaheer D, Service S, Kremeyer B et al. Convergent linkage evidence from two Latin-American population isolates supports the presence of a susceptibility locus for bipolar disorder in 5q31-34. Hum Mol Genet 2006 Nov 1; 15(21): 3146-3153.

Horiuchi Y, Nakayama J, Ishiguro H, Ohtsuki T, Detera-Wadleigh SD, Toyota T et al. Possible association between a haplotype of the GABA-A receptor alpha 1 subunit gene (GABRA1) and mood disorders. Biol Psychiatry 2004 Jan 1; 55(1): 40-45.

Jones I, Hamshere M, Nangle JM, Bennett P, Green E, Heron J et al. Bipolar affective puerperal psychosis: genome-wide significant evidence for linkage to chromosome 16. Am J Psychiatry 2007 Jul; 164(7): 1099-1104.

Kelsoe JR, Spence MA, Loetscher E, Foguet M, Sadovnick AD, Remick RA et al. A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22. Proc Natl Acad Sci U S A 2001 Jan 16; 98(2): 585-590.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996 Jun; 58(6): 1347-1363.

Leboyer M, Bellivier F, McKeon P, Albus M, Borrman M, Perez-Diaz F et al. Age at onset and gender resemblance in bipolar siblings. Psychiatry Res 1998 Nov 16; 81(2): 125-131.

Leboyer M, Bellivier F, Nosten-Bertrand M, Jouvent R, Pauls D, Mallet J. Psychiatric genetics: search for phenotypes. Trends Neurosci 1998 Mar; 21(3): 102-105.

Leboyer M. Searching for alternative phenotypes in psychiatric genetics. Methods Mol Med 2003; 77: 145-161.

Le-Niculescu H, Patel SD, Bhat M, Kuczenski R, Faraone SV, Tsuang MT et al. Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms. Am J Med Genet B Neuropsychiatr Genet 2009 Mar 5; 150B(2): 155-181.

Lin PI, McInnis MG, Potash JB, Willour V, MacKinnon DF, DePaulo JR et al. Clinical correlates and familial aggregation of age at onset in bipolar disorder. Am J Psychiatry 2006 Feb; 163(2): 240-246.

Majumder PP, Pal N. Nonrandom segregation: uniformly most powerful test and related considerations. Genet Epidemiol 1987; 4(4): 277-287.

Manchia M, Lampus S, Chillotti C, Sardu C, Ardau R, Severino G et al. Age at onset in Sardinian bipolar I patients: evidence for three subgroups. Bipolar Disord 2008 May; 10(3): 443-446.

McInnes LA, Escamilla MA, Service SK, Reus VI, Leon P, Silva S et al. A complete genome screen for genes predisposing to severe bipolar disorder in two Costa Rican pedigrees. Proc Natl Acad Sci U S A 1996 Nov 12; 93(23): 13060-13065.

McInnis MG, Dick DM, Willour VL, Avramopoulos D, MacKinnon DF, Simpson SG et al. Genome-wide scan and conditional analysis in bipolar disorder: evidence for genomic interaction in the National Institute of Mental Health genetics initiative bipolar pedigrees. Biol Psychiatry 2003 Dec 1; 54(11): 1265-1273.

Morton NE, Steinberg AG. Sequential test for linkage between cystic fibrosis of the pancreas and the MNS locus. Am J Hum Genet 1956 Sep; 8(3): 177-189.

Moskvina V, Craddock N, Holmans P, Nikolov I, Pahwa JS, Green E; Wellcome Trust Case Control Consortium, Owen MJ, O'Donovan MC. Gene-wide analyses of genome-wide association data sets: evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk. Mol Psychiatry. 2009 Mar;14(3):252-60.

Ni X, Trakalo JM, Mundo E, Macciardi FM, Parikh S, Lee L et al. Linkage disequilibrium between dopamine D1 receptor gene (DRD1) and bipolar disorder. Biol Psychiatry 2002 Dec 15; 52(12): 1144-1150.

Nwulia EA, Miao K, Zandi PP, Mackinnon DF, DePaulo JR, Jr., McInnis MG. Genome-wide scan of bipolar II disorder. Bipolar Disord 2007 Sep; 9(6): 580-588.

O'Donovan MC, Craddock NJ, Owen MJ. Genetics of psychosis; insights from views across the genome. Hum Genet. 2009 Jul;126(1):3-12

Ohtsuki T, Ishiguro H, Detera-Wadleigh SD, Toyota T, Shimizu H, Yamada K et al. Association between serotonin 4 receptor gene polymorphisms and bipolar disorder in Japanese case-control samples and the NIMH Genetics Initiative Bipolar Pedigrees. Mol Psychiatry 2002; 7(9): 954-961.

Pato CN, Pato MT, Kirby A, Petryshen TL, Medeiros H, Carvalho C et al. Genome-wide scan in Portuguese Island families implicates multiple loci in bipolar disorder: fine mapping adds support on chromosomes 6 and 11. Am J Med Genet B Neuropsychiatr Genet 2004 May 15; 127B(1): 30-34.

Radhakrishna U, Senol S, Herken H, Gucuyener K, Gehrig C, Blouin JL et al. An apparently dominant bipolar affective disorder (BPAD) locus on chromosome 20p11.2-q11.2 in a large Turkish pedigree. Eur J Hum Genet 2001 Jan; 9(1): 39-44.

Risch N. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 1990 Feb; 46(2): 229-241.

Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K et al. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Genet 1996 Oct; 14(2): 199-202.

Schulze TG, Detera-Wadleigh SD, Akula N, Gupta A, Kassem L, Steele J, Pearl J, Strohmaier J, Breuer R, Schwarz M, Propping P, Nöthen MM, Cichon S, Schumacher J; NIMH Genetics Initiative Bipolar Disorder Consortium, Rietschel M, McMahon FJ. Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder.Mol Psychiatry. 2009 May;14(5):487-91.

Schumacher J, Kaneva R, Jamra RA, Diaz GO, Ohlraun S, Milanova V et al. Genomewide scan and fine-mapping linkage studies in four European samples with bipolar affective disorder suggest a new susceptibility locus on chromosome 1p35-p36 and provides further evidence of loci on chromosome 4q31 and 6q24. Am J Hum Genet 2005 Dec; 77(6): 1102-1111.

Schurhoff F, Bellivier F, Jouvent R, Mouren-Simeoni MC, Bouvard M, Allilaire JF et al. Early and late onset bipolar disorders: two different forms of manic-depressive illness? J Affect Disord 2000 Jun; 58(3): 215-221.

Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, Tozzi F, Li JZ, Burmeister M, Absher D, Thompson RC, Francks C, Meng F, Antoniades A, Southwick AM, Schatzberg AF, Bunney WE, Barchas JD, Jones EG, Day R, Matthews K, McGuffin P, Strauss JS, Kennedy JL, Middleton L, Roses AD, Watson SJ, Vincent JB, Myers RM, Farmer AE, Akil H, Burns DK, Boehnke M. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. Proc Natl Acad Sci U S A. 2009 May 5;106(18):7501-6.

Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, Gill M, Nurnberger JI, Jr. et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. Am J Hum Genet 2003 Jul; 73(1): 49-62.

Serretti A, Mandelli L. The genetics of bipolar disorder: genome 'hot regions,' genes, new potential candidates and future directions. Mol Psychiatry 2008 Aug; 13(8): 742-771.

Severino G, Congiu D, Serreli C, De Lisa R, Chillotti C, Del Zompo M et al. A48G polymorphism in the D1 receptor genes associated with bipolar I disorder. Am J Med Genet B Neuropsychiatr Genet 2005 Apr 5; 134B(1): 37-38.

Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K et al. Whole-genome association study of bipolar disorder. Mol Psychiatry 2008 Jun; 13(6): 558-569.

Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, Byerley W, Coryell W, Craig D, Edenberg HJ, Eskin E, Foroud T, Gershon E, Greenwood TA, Hipolito M, Koller DL, Lawson WB, Liu C, Lohoff F, McInnis MG, McMahon FJ, Mirel DB, Murray SS, Nievergelt C, Nurnberger J, Nwulia EA, Paschall J, Potash JB, Rice J, Schulze TG, Scheftner W, Panganiban C, Zaitlen N, Zandi PP, Zöllner S, Schork NJ, Kelsoe JR. Genome-wide association study of bipolar disorder in European American and African American individuals. Mol Psychiatry. 2009 Aug;14(8):755-63. Epub 2009 Jun 2.

Suominen K, Mantere O, Valtonen H, Arvilommi P, Leppamaki S, Paunio T et al. Early age at onset of bipolar disorder is associated with more severe clinical features but delayed treatment seeking. Bipolar Disord 2007 Nov; 9(7): 698-705.

Vazza G, Bertolin C, Scudellaro E, Vettori A, Boaretto F, Rampinelli S et al. Genome-wide scan supports the existence of a susceptibility locus for schizophrenia and bipolar disorder on chromosome 15q26. Mol Psychiatry 2007 Jan; 12(1): 87-93.

Venken T, Del-Favero J. Chasing genes for mood disorders and schizophrenia in genetically isolated populations. Hum Mutat 2007 Dec; 28(12): 1156-1170.

Weissman MM, Wickramaratne P, Merikangas KR, Leckman JF, Prusoff BA, Caruso KA et al. Onset of major depression in early adulthood. Increased familial loading and specificity. Arch Gen Psychiatry 1984 Dec; 41(12): 1136-1143.

Zandi PP, Badner JA, Steele J, Willour VL, Miao K, MacKinnon DF et al. Genome-wide linkage scan of 98 bipolar pedigrees and analysis of clinical covariates. Mol Psychiatry 2007 Jul; 12(7): 630-639.

Table 1: Description of the sample

	INITIAL SAMPLE		ADDITIONAL SAMPLE		INITIAL *vs.* ADDITIONAL		EXTENDED SAMPLE	
	Probands	Sibs	Probands	Sibs	Chi²(1df)	p-value	Probands	Sibs
Number of affected	69	76	51	62	-	-	120	138
Mean age at onset (Std dev)	17.41 (2.41)	21.66 (6.84)	16.94 (2.61)	20.65 (7.20)	-	-	17.21 (2.49)	21.21 (7.0)
Sex ratio (male/female)	72.5% (29/40)	62.5% (30/48)	64.5% (20/31)	73.0% (27/37)	0.05	0.82	69.0% (49/71)	67.1% (57/85)
Early BPI	69 (100%)	33 (43.4%)	51 (100%)	20 (32.3%)	1.80	0.18	120 (100%)	53 (38.4%)
Later BPI	-	26 (34.2%)	-	19 (30.6%)	0.20	0.66	-	45 (32.6%)
Early BPII	-	11 (14.5%)	-	9 (14.5%)	0.04	0.85	-	20 (14.5%)
Later BPII	-	0	-	2 (3.2%)	-	-	-	2 (1.5%)
Early MDE	-	6 (7.9%)	-	12 (19.4%)	3.95	0.05	-	18 (13.0%)

Table 2: Results of the MLS using the triangle constraints with the eight linkage regions in the extended family sample

Region of linkage	Phenotype	Genetic location Distance from p-ter*(cM)	MLS max	p-value
2p21	Broad	65.94	0.64	Ns
	Narrow	65.94	0.70	Ns
2q14	Broad	131.51	2.26	p =0.001
	Narrow	128.41	1.75	p = 0.004
3p14	Broad	85.97	1.66	p=0.005
	Narrow	85.97	1.60	p =0.006
5q33	Broad	153.17	0.32	Ns
	Narrow	150.34	0.86	*p =0.04*
7q36	Broad	165.18	1.17	*p =0.02*
	Narrow	162.33	0.54	Ns
10q23	Broad	116.34	0.43	Ns
	Narrow	116.34	0.11	Ns
16p23	Broad	100.39	0.04	Ns
	Narrow	100.39	1.71	p =0.004
20p12	Broad	32.30	1.50	p =0.007
	Narrow	37.65	0.91	*p =0.03*

*: Marker locations (distance from p-ter) were extracted from the Marshfield database

(http://research.marshfieldclinic.org).

Table 3

Results of the PST: regions showing a p ≤0.005 for at least one model

Region	Distance from p-ter	Physical location Mb	Phenotype tested	p-value
2q14	118	106.5	EO-BPI *vs.* LO-BPI	0.0001
			EO-BPI *vs.* EO-BPII	> 0.05
			EO-BPI *vs.* LO-BPI or EO-BPII	0.001
16p23	97	77.5	EO-BPI *vs.* LO-BPI	0.03
			EO-BPI *vs.* EO-BPII	0.01
			EO-BPI *vs.* LO-BPI or EO-BPII	0.005

Table 4

Results of the categorical MLB method: regions showing p ≤0.005 for at least one model

Region	Distance from p-ter	Phenotype tested	p-value
2q14	127 cM	EO-BPI > LO-BPI > UA	0.005
		EO-BPI > EO-BPII > UA	0.01
		EO-BPI > LO-BPI or EO-BPII > UA	0.005