

Genetic Linkage of Bipolar Disorder to Chromosome 6q22 Is a Consistent Finding in Portuguese Subpopulations and May Generalize to Broader Populations

Carlos N. Pato,^{1,2,3,4*} Frank A. Middleton,^{1,2,3,5} Karen L. Gentile,^{2,5} Christopher P. Morley,^{2,3} Helena Medeiros,^{1,2,3,4} Antonio Macedo,⁶ M. Helena Azevedo,⁶ and Michele T. Pato^{1,2,3,4}

¹Veterans Affairs Medical Center, Washington, District of Columbia

²Center for Neuropsychiatric Genetics, Upstate Medical University, Syracuse, New York

³Department of Psychiatry, Upstate Medical University, Syracuse, New York

⁴Department of Psychiatry, Georgetown University, Washington, District of Columbia

⁵Department of Neuroscience and Physiology, Upstate Medical University, Syracuse, New York

⁶Psicologia Medica, Universidade de Coimbra, Coimbra, Portugal

We recently reported genome-wide significant linkage to chromosome 6q for bipolar disorder, in a study of 25 Portuguese families, using the Human Mapping Assay Xba 131 (HMA10K). To explore the generalizability of this finding, we reanalyzed our SNP linkage data according to the families' geographic origin. Specifically, the 25 families included 20 families from the Portuguese island collection (PIC; 15 families from the Azores Islands and 5 from the Madeira Islands) and 5 families from continental Portugal. Non-parametric linkage analysis (NPL) was performed as previously described and indicated that each of these subpopulations showed evidence of linkage for the same region. To further address the potential generalizability of these findings to other populations, we have also examined allelic heterozygosity in our subpopulations and in three reference populations (Caucasian, East Asian, and African-American). This analysis indicated that the PIC population is highly correlated to the Caucasian reference population ($R=0.86$) for all of chromosome 6. In contrast allelic heterozygosity was more weakly correlated between PIC and both East Asian ($R=0.37$) and African-American ($R=0.32$) reference populations. Taken together these observations suggest a shared genetic liability among Portuguese populations for bipolar disorder on chromosome 6q, and that the PIC population is likely representative of Caucasians in general. © 2005 Wiley-Liss, Inc.

KEY WORDS: microarray; genechip; non-parametric; heterozygosity; SNP

We have recently reported genome-wide significant linkage non-parametric linkage analysis (NPL-Z=4.2) for bipolar

disorder on chromosome 6q22 [Middleton et al., 2004]. This work was primarily based on families from the Portuguese Island Collection (PIC) (encompassing the Azores and Madeira Islands) but also included a small number of families from continental Portugal ($n=5$). Our finding on 6q complements the findings of the NIMH Genetics Initiative [see Dick et al., 2003; and published Erratum, 2003]. Those authors reported in a study of 250 families with bipolar disorder a maximum LOD of 2.2 at marker D6S1021 (104.7 MB) on chromosome 6q, that was subsequently raised to a genome-wide significant level of 3.8 in a combined analysis of all 399 pedigrees with bipolar disorder from their collection. This same marker produced the highest linkage signal in our original report of 16 families [Pato et al., 2004]. However, in an expanded family set ($n=25$), using the HMA10K SNP assay, with far more complete coverage of this region, we identified a more telomeric peak NPL at 126 MB. This suggests that linkage to 6q may represent a shared genetic liability for bipolar disorder in both of these populations.

The PIC population is drawn from the Azores and Madeira Islands. These islands were settled more than 500 years ago almost exclusively by the Portuguese. The islands had no native population when they were first settled by Portugal in the early 1400s. The current population of Madeira is 300,000, and of the Azores is 249,000. To address whether families from these two archipelagoes share the same genetic forms of illness, we analyzed our recently reported linkage to chromosomes 6q for bipolar disorder in our Azorean families and in our Madeiran families, as well as, a small set of families from continental Portugal.

Our results indicate that families from all three Portuguese subpopulations appear to share this linkage to chromosome 6q. This evidence for disease homogeneity on the islands argues for treating this population as one. Moreover, this finding has proven generalizable to mainland Portugal and possibly other European and European-American populations.

We performed two primary analyses. In the first, we specifically sub grouped the 25 families from our original report based on their geographic origin. Their origin was defined through probands with both parents and all four grandparents being from the same geographic origin (thus, families had to have three consistent generations to be considered from that population base). All 25 families in our original report could be classified by these criteria into: Azorean ($n=15$), Madeiran ($n=5$), PIC (Azorean or Madeiran), or continental ($n=5$) origin.

In the second analysis, we compared allelic heterozygosity between populations. We included all subjects for which we had HMA10K data. For this analysis, members of 35 families were

Grant sponsor: VA Merit Award (to M.T.P.); Grant sponsor: NIMH (to C.N.P. and M.T.P.); Grant numbers: MH52618, MH058693.

*Correspondence to: Carlos N. Pato, M.D., ACOS for Research and Development, Veterans Affairs Medical Center, Washington, DC 20422. E-mail: CMPATO@MAC.COM

Received 5 October 2004; Accepted 21 October 2004

DOI 10.1002/ajmg.b.30159

included. Family origins were: Azorean ($n=20$), Madeiran ($n=9$), and continental ($n=6$). The average number of subjects genotyped per SNP from these populations was: Azorean ($n=103$), Madeiran ($n=47$), and continental ($n=40$). These genotypes were further compared to genotype data obtained from three Affymetrix reference populations (Caucasian ($n=42$), East Asian ($n=20$), and African-American ($n=37$)) [Matsuzaki et al., 2004; reference data are available for free download at http://www.affymetrix.com/Auth/analysis/downloads/taf/Mapping10K_Xba131_annot_csv.zip].

Methods for subject ascertainment and classification are the same as previously described [Pato et al., 2004]. Families with two or more affected individuals were ascertained from systematic screening of all treating clinicians, treatment facilities, social services, and extensive family interviews. Informed consent was obtained in writing from all subjects for participation in the genetic and family studies. Best estimate diagnoses were made according to DSM-IV.

DNA was isolated and processed as previously described [Middleton et al., 2004]. Genotype data were obtained on each subject for approximately 11,195 SNPs with known cytogenetic coordinates based on the HMA10K SNP assay. We used MERLIN (Multipoint Engine for Rapid Likelihood Inference) software [Abecasis et al., 2002] to perform multipoint NPL of the subsets of families used in this report, and also to calculate heterozygosity for each SNP. The NPL-Z scores from the MERLIN runs were used to generate graphical plots. All data presented are plotted using physical positions specified in the July 2003 freeze of the human genome assembly (<http://genome.ucsc.edu/>).

We determined the heterozygosity of all SNPs in the HMA10K SNP assay. These data were used to assess correlations between each of the Portuguese cohorts for the entire genome. Further, we compared chromosome 6 heterozygosity correlations among and between the Portuguese cohorts and the three Affymetrix reference populations [Matsuzaki et al., 2004] to determine whether there were significant differences between them using the Fisher's R to Z test (with 675 dof).

The results of the genome-wide linkage analysis for the Portuguese 25 family data set are presented in Figure 1 (as the plot in black labeled "All Portugal;" Middleton et al., 2004]. This analysis provided further validation of the finding on chromosome 6q22 [initially reported in Pato et al., 2004], achieving a maximum NPL of 4.20 ($P=0.00001$) and maximum LOD of 3.56 ($P=0.00003$) at position 126 MB. This peak and the surrounding region of chromosome 6 that exceeded the threshold for suggestive genome-wide significance (NPLs >2.89) were interrogated by a total of 38 SNPs, with an average inter-SNP distance of 0.20 MB [Middleton et al., 2004].

The population was then analyzed for each specific sub-population. We have defined the PIC as a population of subjects from the Azores and Madeira islands. This combined sub-population includes 20 families. These families yield strong evidence for linkage with a NPL of 3.71 in the exact same region (Fig. 1, PIC). Further dividing the PIC population into the two archipelagoes revealed consistent evidence for linkage to this region. The Azorean families ($n=15$) again supported linkage with a NPL of 3.18 (Fig. 1, Azores). The Madeiran families ($n=5$) also showed evidence for linkage with a NPL of 1.94 (Fig. 1, Madeira). Even the continental families that had been part of our pilot studies proved to provide evidence for linkage in this region with a NPL of 2.0 (not shown). As is clearly demonstrated all of the population subgroups in our sample contribute to the overall significant linkage finding on chromosome 6q (Fig. 1, all Portugal).

To further characterize the population of the Portuguese islands (PIC) that we have used for these and other studies of complex disorders, we examined genome-wide genotype data

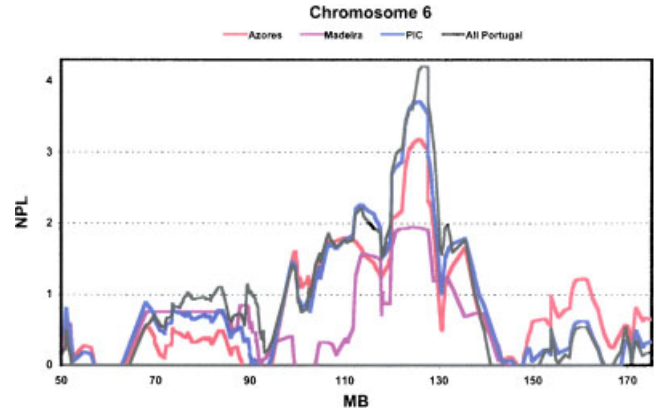


Fig. 1. Common linkage to chromosome 6q22 in Portuguese subpopulations. Non-parametric linkage analysis (NPL) of multiple families from the Azores ($n=15$), Madeira ($n=5$), and combined (PIC) families produces linkage peaks on 6q22 that completely overlap with that of the complete 25 family set (all Portugal), which included five families from the continental Portuguese population. These data argue in favor of a shared genetic liability for bipolar disorder on chromosome 6q22 in these subpopulations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

on each subject for all the mapped SNPs in the HMA10K assay. In Table I, we present the heterozygosity correlations for all of chromosome 6 between the PIC population, the Azores, and Madeira, and contrast them to correlations to continental Portugal, and the three Affymetrix reference populations.

Genome-wide, we observed a high heterozygosity correlation between the Azores and Madeira ($R=0.74$). More specifically, on chromosome 6, heterozygosity for the Azorean and Madeiran populations was similarly correlated ($R=0.71$). Furthermore, heterozygosity on chromosome 6 in the PIC population, as a whole, was highly correlated to the Affymetrix Caucasian reference population ($R=0.87$). In contrast, the PIC population was only weakly correlated with either of the other two Affymetrix reference populations; East Asians ($R=0.37$) and African-Americans ($R=0.32$) (Table I). We compared chromosome 6 correlations, using the Fisher's R to Z test (with 675 dof), and found significant differences between the high correlation with Caucasians and low correlations with East Asians ($Z=18.4$, $P<10^{-6}$) and African-Americans ($Z=19.5$, $P<10^{-6}$). Interestingly, both the Asian and African-American reference populations displayed similar weak correlations to the PIC population ($Z=1.2$, $P=0.25$). We present a plot of allelic heterozygosity in the PIC population and the Caucasian reference population across the region of maximal linkage on chromosome 6q (120–130 MB) (Fig. 2).

The choice of the Portuguese island population (PIC) for our studies is based on the hypothesis that the population is characterized by a higher degree of disease homogeneity than populations found in the United States. The geographic isolation may prove to have resulted in a higher percentage of families sharing the same genetic risks for bipolar disorder, possibly the same forms of illness, than in less isolated

TABLE I. Correlation in Allelic Heterozygosity for Chromosome 6

Population	Azores	Madeira	PIC	Continental
Azores		0.7109	0.9663	0.7165
Madeira	0.7109		0.8678	0.6065
PIC	0.9663	0.8678		0.7276
Continent	0.7165	0.6065	0.7276	
Caucasian	0.8565	0.7183	0.8676	0.7284
East Asian	0.3708	0.2985	0.3715	0.3252
African-American	0.3071	0.2831	0.3198	0.2333

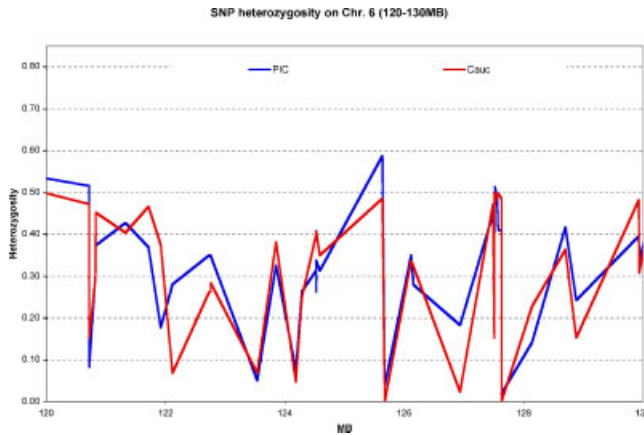


Fig. 2. High similarity of heterozygosity between PIC and Caucasian populations. The average heterozygosities of the SNPs on chromosome 6 were found to be highly similar between the PIC population and North American Caucasians ($R=0.87$), that was clearly evident in the region spanning the 6q linkage peak for bipolar disorder. This observation suggests that findings in the PIC population may generalize to other populations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

populations. This may, in fact, be the case for a number of complex disorders that may encompass a number of different subtypes. Furthermore, in our studies of schizophrenia, linkage signals on chromosomes 5q and 8p [Sklar et al., 2004] were detected in both the Azorean families and Madeiran families providing evidence for shared disease homogeneity in the two PIC subpopulations.

Our studies in the Portuguese Island population have been quite successful at identifying replicable genetic linkages in complex disorders. However, it is important to determine whether our findings can generalize to broader populations. Therefore, we carried out studies to determine whether the PIC population offers a specific genetic architecture that might facilitate the positional cloning of candidate genes. In the present study, we have explored these questions in two fundamental ways: genome wide genetic diversity and replicability of specific linkage findings.

When we examined maximal linkage scores in each of the subpopulations a number of suggestive and even more nominal linkage peaks are identified. A few were comparable or even exceeded the chromosome 6q signal in each population. However, chromosomes 6 and 11 are the most consistent findings in the combined set [Middleton et al., 2004; Pato et al., 2004].

The present report clearly demonstrates the common finding of linkage for bipolar disorder to chromosome 6q in the three different Portuguese subpopulations, in the PIC population as a whole, as well as in the overall Portuguese population (Fig. 1). The complete overlap of the linkage signals from these populations causes this linkage signal to be the most significant in the combined 25 family set, and the only one that reaches genome-wide significance [Middleton et al., 2004].

Our analysis of heterozygosity correlations between populations provides a basis for exploring the potential generalizability of chromosome 6q linkage for bipolar disorder in other populations. We found highly significant correlations between the PIC populations, but more importantly the PIC population was found to be highly correlated with Caucasians. Furthermore, to date we have not detected any difference in the common haplotype frequencies between the PIC population and western European reference populations (i.e., CEPH data set). Collectively these observations suggest that linkage

findings from the PIC population may well be generalizable to European and North American Caucasian populations, and also support the value of this population for follow-up studies.

As an example of the generalizability of our findings in the PIC population, we point out the combined analysis of the NIMH waves 1–3 bipolar families (399 pedigrees) produces a genome-wide significant LOD score of 3.8 on chromosome 6q at marker D6S1021 (104.7 MB) (see Dick et al., 2003; and published Erratum, 2003). As we noted, D6S1021 was also the microsatellite that produced our maximal linkage signal on chromosome 6 in our 10 cM genome scan [Pato et al., 2004]. Additional microsatellite markers spanning the region began to demonstrate a telomeric shift in the position of this peak that was clearly evident with the much higher resolution SNP based assay [Middleton et al., 2004]. We believe that a higher resolution examination of this region in the NIMH families, as well as in other populations, may well further confirm and define the location of this linkage.

In conclusion, the genome wide significant linkage for bipolar disorder to chromosome 6q22 is a product of contributions from all of the Portuguese subpopulations. The striking overlap in the linkage signals from each of the component populations further strengthens this observation. Our examination of population architecture suggests that the Azores and Madeira (PIC) can be appropriately considered a single population. Moreover, findings obtained in this population may well generalize to western European and North American Caucasians.

ACKNOWLEDGMENTS

We thank the families and individuals who participated in these studies. We also thank Xin Zhao for assistance in software development, and Celia Carvalho, Ana Dourado, Isabel Coelho, M. J. Soares, Jose Valente, and Carlos Paz Ferreira for assistance in the clinical ascertainment of subjects. Support for this work was derived, in part, from a VA Merit Award (to M.T. Pato) and NIMH grants MH52618 and MH058693 (to C.N. Pato and M.T. Pato).

REFERENCES

- Abecasis GR, Chemy SS, Cookson WO, Cardon LR. 2002. Merlin: Rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101.
- Dick DM, Foroud T, Flury L, Bowman ES, Miller MJ, Rau NL, Moe PR, et al. 2003. Genomewide linkage analysis of bipolar disorder: A new sample of 250 pedigrees from the National Institute of Mental Health genetics initiative. *Am J Hum Genet* 73:107–114 (erratum 73:979).
- Matsuzaki H, Loi H, Dong S, Tsai YY, Fang J, Law J, Di X, Liu WM, Yang G, Liu G, Huang J, Kennedy GC, Ryder TB, Marcus GA, Walsh PS, Shriver MD, Puck JM, Jones KW, Mei R. 2004. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res* 14:786.
- Middleton FA, Pato MT, Gentile KL, Morley CP, Zhao X, Eisener AF, et al. 2004. Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: A comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. *Am J Human Genetics* 74:886–897.
- Pato CN, Pato MT, Kirby A, Petryshen TL, Medeiros H, Carvalho C, Macedo A, Dourado A, Coelho I, Valente J, Soares MJ, Ferreira CP, Lei M, Verner A, Hudson TJ, Morley CP, Kennedy JL, Azevedo MH, Daly MJ, Sklar P. 2004. 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 Genetics* 127B(1):30–34.
- Sklar P, Pato MT, Kirby A, Petryshen TL, Medeiros H, Carvalho C, Macedo A, Dourado A, Coelho I, Valente J, Soares MJ, Ferreira CP, Lei M, Verner A, Hudson TJ, Morley CP, Kennedy JL, Azevedo MH, Lander E, Daly MJ, Pato CN. 2004. Genome-wide scan in Portuguese Island families identifies 5q31-5q35 as a susceptibility locus for schizophrenia and psychosis. *Mol Psychiatry* 9:213–218.