A Genome-Wide Search for Risk Genes Using Homozygosity Mapping and Microarrays With 1,494 Single-Nucleotide Polymorphisms in 22 Eastern Cuban Families With Bipolar Disorder


1Centre for Basic Psychiatric Research, Psychiatric Hospital in Aarhus, Risskov, Denmark
2Department of Clinical Biochemistry, Skejby Hospital, Brendstrupgårdsvej, Aarhus N, Denmark
3Molecular Genetics Laboratory, National Center of Medical Genetics, Reparto Cubanacín, Playa, Havana, Cuba
4Psychiatric Hospital, Holguín, Cuba
5Department of Clinical Biochemistry and Genetics, Odense University Hospital, Odense, Denmark

Homozygosity mapping is a very powerful method for finding rare recessive disease genes in monogenic disorders and may also be useful for locating risk genes in complex disorders, late onset disorders where parents often are not available, and for rare phenotypic subgroups. In the present study, homozygosity mapping was applied to 24 persons with bipolar disorder from 22 inbred families. The families were selected irrespective of whether other affected family members were present or not. A genome wide screen using genotypes from only a single affected person in each family was performed using the AFFYMETRIX GeneChip HuSNP Mapping Assay, which contains 1,494 single nucleotide polymorphisms. At chromosome 17q24-q25 a parametric multipoint LOD score of 1.96 was found at WIAF-2407 and WIAF-2405. When analyzing 19 additional microsatellite markers on chromosome 17q the maximum parametric multipoint LOD score was 2.08, 1.5 cM proximal to D17S868. The present study replicates a recent significant linkage finding.© 2005 Wiley-Liss, Inc.

KEY WORDS: homozygosity mapping; microarrays; genome scan; bipolar disorder; chromosome 17

INTRODUCTION

Bipolar disorder is a severe psychiatric disorder with a lifetime prevalence of 0.3–1.5%. Though most individuals with bipolar disorder do not have affected first-degree relatives, the high concordance ratio for mono- and dizygotic twins suggest that genetic factors are of a very strong influence. Bipolar disorder is a complex disorder in which rare, large families with multiple affected individuals do occur. It is possible that the inheritance in such families is monogenic or at least oligogenic. Most studies of bipolar disorder have applied parametric or non-parametric linkage analyses to larger families and a number of chromosome regions potentially containing risk genes have been identified [Baron, 2002].

Homozygosity mapping as a method for identifying chromosome regions harboring disease genes for rare recessive diseases was originally suggested by Lander and Botstein [1987]. The basic principle of homozygosity mapping is that an affected person has inherited two identical copies of the disease gene from a common ancestor. In consanguineous families, patients with rare recessive diseases are likely to be homozygous by descent (HBD) for the disease allele and nearby markers. This method has almost exclusively been applied to rare recessive diseases in patients whose parents are first or second cousins [e.g., Wang et al., 2001]. The average size of a segment, which is HBD surrounding a rare recessive disease gene, is on average 33.3, 25, and 20 cM, respectively, for a single affected offspring of first, second, or third cousins [Wright et al., 1997]. Offspring of first cousins are on average HBD on 1/16 of their genome, while offspring of second or third cousins are on average HBD on 1/64 and 1/256 of their genome, respectively. The disease gene for a rare recessive disease has been identified by initially searching for homozygous segments among more distantly related cases from isolated populations, with incomplete knowledge about their genealogy [Houwen et al., 1994; Bull et al., 1998, 1999]. It has been suggested to apply homozygosity mapping on complex disorders [Miano et al., 2000]. However, to our knowledge, this method has only been applied to an apparently recessive subgroup of early-onset Parkinsonism [van Duijn et al., 2001], and a single inbred pedigree with several siblings with bipolar disorder [Ewald et al., 2003].

Homozygosity mapping is a very powerful method when applying parametric, recessive LOD score analyses as one affected offspring of parents who are first cousins may yield a LOD score around 1.2 for a rare, fully penetrant, recessive disease (Table I).

In inbred families with more than one affected offspring, parametric as well as non-parametric linkage analyses can be applied [Ewald et al., 2003; Modin et al., 2003]. It has not yet been attempted to apply homozygosity mapping on families with a possibly more complex inheritance pattern with only a single affected offspring. Recessive risk alleles may be present in complex disorders such as bipolar disorder. Furthermore,
Monogenic hospital case notes, a clinical narrative was made for each of the SCAN [Wing et al., 1990]. Based on the interview and viewed by an experienced psychiatrist Dr. Mario Torralba, of cases in families F2V1, F3V9, F3V30, F7V1, F22V1, and F3V32 were first cousins. At present, no first-degree relatives of F3V1, and more F3V1 and F3V2 were related to F3V9 (second degree and the siblings F19V10–V16 were third degree relatives. Further- and the siblings F19V10–V16 were third degree relatives. Some of the other cases were more distantly related. Further-

The other family members have been coded as being phenotype unknown. The LOD score at zero recombination frequency is shown.

“sporadic” cases without known affected close relatives are compatible with recessive risk alleles.

The present study searched for homozygous chromosome regions among 24 cases with bipolar disorder whose parents were third cousins or more closely related. A genome wide scan was performed, using the AFFYMETRIX GeneChip HuSNP Mapping Assay with 1,494 single nucleotide polymorphisms (SNPs).

**TABLE I. Parametric LOD Scores Assuming a Recessive Mode of Inheritance Modeled as a Monogenic Disease (Penetrances 0, 0, and 1, Disease Allele Frequency 0.001) or a Complex Disease (Penetrances 0.005, 0.005, and 0.65, Disease Allele Frequency 0.10)**

<table>
<thead>
<tr>
<th>Parents</th>
<th>Monogenic</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st cousins</td>
<td>0.80 (1.29)</td>
<td>0.36 (0.85)</td>
</tr>
<tr>
<td>2nd cousins</td>
<td>0.92</td>
<td>0.20</td>
</tr>
<tr>
<td>3rd cousins</td>
<td>0.90</td>
<td>0.07</td>
</tr>
</tbody>
</table>

A single affected offspring is genotyped and homozygous for a single marker with frequency of the relevant marker allele of 0.1, 0.01, 0.001, or 0.0001 as shown above. For two affected siblings that are homozygous for the same marker with 1st cousin parents the LOD score is shown in parentheses. The other family members have been coded as being phenotype unknown. The LOD score at zero recombination frequency is shown.

**MATERIALS AND METHODS**

**Collection of Cases, Controls, and Diagnostic Assessment**

Well-documented cases with severe bipolar disorder were sought among in- or out-patients, who originated from a specific subregion around the city of Velasco, in the Holguin region in Eastern Cuba. The majority of the inhabitants in the region are of Caucasian descent, while a smaller minority is of Afro-American descent. The population size of the Velasco region is today around 44,000, and there has been a tradition of marriages between cousins and previously an average of nine children per marriage. In order to include the cases had to be Caucasian, of European descent and their parents had to be related, preferably at first or second cousins (Fig. 1). Cases were collected irrespective of their age of onset of bipolar disorder and the presence or absence of other affected family members. Some of the cases were close relatives. The affected persons in F3V1 and F3V2 were parent and child who were inbred through partly different loops (not shown in Fig. 1). F4V2 and F4V3 were siblings and F19V10 and F19V16 were siblings. F19V9 and the siblings F19V10–V16 were third degree relatives. Some of the other cases were more distantly related. Further-

The average distance between the SNP is approximately 3 cM, which should be sufficient to detect most HBD regions in affected offspring from first, second, or third cousins. The SNP heterozygosity of the 133 control persons tested by AFFYMETRIX has a median of 36% (25th and 75th percentile of 22 and 47%, respectively). The informativity of two to three SNPs corresponds roughly to one microsatellite marker [Kruglyak, 1997].

Thirteen additional SNPs were genotyped in regions where the distance between neighboring SNPs was greater than 15 Mb. In addition ten randomly chosen SNPs, which were in Hardy–Weinberg equilibrium and not associated with the disorder, were regenotyped in order to evaluate genotyping error frequencies. These 23 SNPs were amplified using touchdown PCR. Excess primers and dNTPs were degraded using shrimp alkaline phosphatase (SAP) (USB) and exonuclease I (USB). The SNP was analyzed by a primer extension reaction according to the SNaPshot protocol (Applied Biosystems, Foster City, CA). Excess labeled nucleotides were degraded and reaction products were analyzed on an ABI 3100 automated sequencer (Applied Biosystems).

Microsatellite markers were amplified using a Perkin-Elmer 9700 thermocycler and were subsequently analyzed on an ABI Prism 3100 or 310 Genetic Analyzer (Applied Biosystems). Specific PCR conditions, when available, from http://www-gdb.cmbi.kun.nl/gdb/ were further optimized for multiplex PCR.

**Genotyping**

The AFFYMETRIX GeneChip HuSNP Mapping Assay, with 1,494 SNPs distributed on the 22 autosomes and the X chromosome was applied for the 24 cases and 54 controls. DNA was assayed and scanned according to the HuSNP protocol supplied by AFFYMETRIX, Inc., Santa Clara, CA, as described previously [Primdahl et al., 2002].

SNPs were positioned for genetic mapping using the SNP database at NCBI, build 101, (based on genome assembly build 27 (http://www.ncbi.nlm.nih.gov/SNP/)). SNPs with ambiguous positions were excluded, except where the differences in position did not influence the order of the markers. Additional sixteen positioned on the X chromosome, were excluded from the analyses because all families except families F12, F19, and F27, had transmission from father to son. Based on this strategy a total of 1,168 SNP markers were analyzed.

The 54 controls were used for estimating allele frequencies. However, if any specific SNP was unambiguously genotyped on less than 60 out of the 108 control chromosomes, the frequency obtained from AFFYMETRIX was used. These allele frequencies are generated from 133 unrelated individuals, 113 of Western European descent, ten African-Americans and ten of Asian descent. If an allele was not observed among any of the 108 control chromosomes its frequency was set to 1/108 (0.0093).

The average distance between the SNP is approximately 3 cM, which should be sufficient to detect most HBD regions in affected offspring from first, second, or third cousins. The SNP heterozygosity of the 133 control persons tested by AFFYMETRIX has a median of 36% (25th and 75th percentile of 22 and 47%, respectively). The informativity of two to three SNPs corresponds roughly to one microsatellite marker [Kruglyak, 1997].

Thirteen additional SNPs were genotyped in regions where the distance between neighboring SNPs was greater than 15 Mb. In addition ten randomly chosen SNPs, which were in Hardy–Weinberg equilibrium and not associated with the disorder, were regenotyped in order to evaluate genotyping error frequencies. These 23 SNPs were amplified using touchdown PCR. Excess primers and dNTPs were degraded using shrimp alkaline phosphatase (SAP) (USB) and exonuclease I (USB). The SNP was analyzed by a primer extension reaction according to the SNaPshot protocol (Applied Biosystems, Foster City, CA). Excess labeled nucleotides were degraded and reaction products were analyzed on an ABI 3100 automated sequencer (Applied Biosystems).

Microsatellite markers were amplified using a Perkin-Elmer 9700 thermocycler and were subsequently analyzed on an ABI Prism 3100 or 310 Genetic Analyzer (Applied Biosystems). Specific PCR conditions, when available, from http://www-gdb.cmbi.kun.nl/gdb/ were further optimized for multiplex PCR.
Statistical Analysis

A number of different analyses were performed in order to search for homozygous segments. Multipoint and single point parametric LOD scores were calculated for the autosomal markers with known physical order, using GENEHUNTER version 2.0 [Kruglyak et al., 1996]. This was done using a recessive model, and performing an affected-only analysis. Bipolar disorder is a complex disease and penetrances were therefore set to 0.005, 0.005, and 0.65 for carriers of none, one and two disease alleles, respectively. The disease allele frequency was set to 0.1 as in previous studies [Ewald et al., 2002]. A 1 Mb physical distance between SNPs was roughly transformed to a 1 cM genetic distance. SNPs closer than 10,000 base pairs were set to a distance of 0.0001. The maximum obtainable LOD score for a single affected offspring being homozygous for a marker allele depends on the relatedness of the parents, the penetrances, and marker and disease allele frequencies. In order to evaluate power, two-point LOD scores for a monogenic or a complex recessive disease were calculated using GENEHUNTER with the parameters specified (Table I). A total LOD score of above 3 may be obtained if five unrelated affected persons, whose parents are first cousins, are homozygous for an allele or haplotype with a frequency of 0.01 or less.

RESULTS

Single- and multipoint parametric LOD scores above 1.0 are shown in Table II. High single point LOD scores at WIAF-3227 on chromosome 14 and at WIAF-1706 on chromosome 22 were found to be errors upon regenotyping. These SNPs were excluded from the multipoint analysis.

The highest LOD score found was a multipoint LOD score of 1.96 distally on chromosome 17q25 for the two neighboring markers WIAF-2407 and WIAF-2405. On chromosome 10, a multipoint LOD score of 1.24 was found at WIAF-4519 in a very broad region with 25 neighboring markers yielding positive multipoint LOD scores. A multipoint LOD score of 1.19 was found at WIAF-1969 on chromosome 20q in a region with nine neighboring markers yielding positive multipoint LOD scores. On chromosome 11p, a peak multipoint LOD score of 1.13 at WIAF-3023 was found in a region with 15 neighboring markers all with positive multipoint LOD scores.

The highest single point LOD score was 1.26 at WIAF-1805 on chromosome 12p (Table II). In this region nine neighboring markers yielded positive multipoint LOD scores with a maximum of 0.67 at WIAF-3049. No positive multipoint LOD scores or single point LOD scores above 0.15 were found distally on chromosome 8q, suggesting that the single point
TABLE II. Single- and Multipoint Parametric LOD Scores Above 1.0 Are Shown

<table>
<thead>
<tr>
<th>LOD score</th>
<th>Marker</th>
<th>No</th>
<th>Chromosome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.96</td>
<td>WIAF-2405</td>
<td>50, alpha = 1, marker 45 (WIAF-2450) to 51 (WIAF-3305) &gt;1.50</td>
<td>17</td>
</tr>
<tr>
<td>1.24</td>
<td>WIAF-4519</td>
<td>24, alpha = 1, marker 21 (WIAF-1585) to 23 (WIAF-2438) &gt;0.90</td>
<td>10</td>
</tr>
<tr>
<td>1.19</td>
<td>WIAF-1969</td>
<td>22, alpha = 1, marker 17 (WIAF-1648) to 24 (WIAF-3937) &gt;0.81</td>
<td>20</td>
</tr>
<tr>
<td>1.13</td>
<td>WIAF-3492</td>
<td>15, alpha = 1, marker 13 (WIAF-4592) to 16 (WIAF-2313) &gt;0.81</td>
<td>11</td>
</tr>
<tr>
<td>1.26</td>
<td>WIAF-1805</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>1.26</td>
<td>WIAF-583</td>
<td>80</td>
<td>8</td>
</tr>
<tr>
<td>1.18</td>
<td>WIAF-2858</td>
<td>48</td>
<td>17</td>
</tr>
<tr>
<td>1.10</td>
<td>WIAF-3130</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>1.05</td>
<td>WIAF-1805</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>1.04</td>
<td>WIAF-2857</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1.04</td>
<td>WIAF-583</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>5.91</td>
<td>WIAF-2467</td>
<td>46, marker 44 (WIAF-364) to 47 (WIAF-2531) all &gt;5.86</td>
<td>12</td>
</tr>
<tr>
<td>5.66</td>
<td>WIAF-3336</td>
<td>16, marker 15 (WIAF-3797) to 17 (WIAF-3901) all &gt;5.53</td>
<td>10</td>
</tr>
<tr>
<td>5.42</td>
<td>WIAF-3136</td>
<td>3, marker 2 (WIAF-3130) to marker 4 (WIAF-2158) all &gt;5.40</td>
<td>10</td>
</tr>
<tr>
<td>5.33</td>
<td>WIAF-3684</td>
<td>23, marker 21 (WIAF-953) to 24 (WIAF-3937) all &gt;4.64</td>
<td>20</td>
</tr>
</tbody>
</table>

Multipoint LOD scores above 5.3 with assumed allele frequencies of 0.5 for both marker alleles are shown. Only the maximum multipoint LOD score for contiguous markers are shown. No is the marker number from pter. Alpha is the fraction of pedigrees linked.

LOD score of 1.26 at WIAF-583 in this region was a chance finding. In the region on chromosome 17q25 suggested by multipoint LOD scores, WIAF-2858 yielded a single point LOD score of 1.18. In this region eight of the ten most distal markers yielded positive single point LOD scores. On chromosome 10p, WIAF-3130 yielded a LOD score of 1.10, however, the neighboring markers yielded single point LOD scores close to zero. WIAF-303 on chromosome 10 yielded a LOD score of 1.05. Nearby markers yielded negative single point LOD scores. WIAF-786 on chromosome 19 yielded a LOD score of 1.04, but was only weakly supported by LOD scores from other neighboring markers and all multipoint LOD scores in the region were negative. Finally, a single point LOD score of 1.04 at WIAF-2137 on chromosome 18p received support from a number of nearby markers.

In individual families the highest LOD scores were 1.25 at WIAF-2503 for the affected sib-pair F19V10–V16 at chromosome 10q23.31. 1.17 at WIAF-3130 for sib-pair F4V2–V3 at chromosome 10p and 1.02 at WIAF-3310 for sib-pair F19V10–V16 at chromosome 9p34.2.

Concerning the ten randomly chosen SNPs, that were regenotyped several times, eight markers produced consistent and identical results, for one marker a single genotyping was discrepant, while a single marker yielded very inconsistent results. Considering all ten markers a total of 56 out of 706 genotypings were different (7.9%).

For further evaluation 19 microsatellite markers on 17q (D17S1299-D17S928: 62–126.5 cM) were genotyped. Microsatellites in the region, yielding a maximum multipoint parametric LOD score of 2.08 1.5 cM proximal to D17S928. The region identified by the microsatellite markers was around 4 Mb and 20 cM more distal than the most significant SNP marker (WIAF-2407 and WIAF-2405) and around 3 Mb and 14 cM more distal than the most distal SNP marker (WIAF-3305) the distal neighbor of WIAF-2405. Additional information was obtained when genotyping seven more telomeric microsatellite markers as the LOD score increased and as the final region of most interest were found to be more distal.

A non-parametric significant LOD score of 3.63 between bipolar disorder and D17S928 has recently been reported in a sample of 250 pedigrees with bipolar disorder [Dick et al., 2003]. D17S928 is located around 200,000 bp more distal than our most interesting marker, D17S928. Our finding formally replicates their finding. Bennet et al. [2002] investigated 154 affected sib-pairs from Ireland and United Kingdom and found a LOD score (MLS) of 1.38 also peaking around D17S928. Curtis et al. [2003] investigated seven larger pedigrees with bipolar disorder and unipolar disorder from United Kingdom and Iceland. Assuming a recessive mode of inheritance and only including unipolar patients from bipolar families they found a four-point LOD score of 1.9 at D17S939 at 105.7 cM, i.e., proximal to the loci suggested by the other studies. This region did not receive much support in a recent meta-analysis of bipolar disorder [Segurado et al., 2003]. However, the meta-analysis ranked findings from individual investigations and the genome-wide scans included in the meta-analysis mainly investigated larger pedigrees, which may preferentially harbor dominant risk alleles. The finding of a common region on chromosome 17q in four studies using different linkage analysis methods and different family material, ranging from single affected individuals to large pedigrees from various mainly Caucasian populations, is encouraging. The 17qter is a gene rich region which contains a number of interesting neurogenes.

DISCUSSION

The present study is to our knowledge one of the first studies which applies homoyzogosity mapping to a complex disorder and one of the first linkage studies using microarrays. Using GENEHUNTER it was possible to calculate multipoint parametric LOD scores including all markers in one analysis on every autosomal except chromosomes 1, 2, 6, and 8 on which markers were subdivided into two segments. Combining all families the highest parametric multipoint LOD scores were found at chromosome 17q25, peaking at a LOD score of 1.96. This finding was further supported by analyzing 19 microsatellites in the region, yielding a maximum multipoint parametric LOD score of 2.08 1.5 cM proximal to D17S928. The region identified by the microsatellite markers was around 4 Mb and 20 cM more distal than the most significant SNP marker (WIAF-2407 and WIAF-2405) and around 3 Mb and 14 cM more distal than the most distal SNP marker (WIAF-3305) the distal neighbor of WIAF-2405. Additional information was obtained when genotyping seven more telomeric microsatellite markers as the LOD score increased and as the final region of most interest were found to be more distal.

A non-parametric significant LOD score of 3.63 between bipolar disorder and D17S928 has recently been reported in a sample of 250 pedigrees with bipolar disorder [Dick et al., 2003]. D17S928 is located around 200,000 bp more distal than our most interesting marker, D17S928. Our finding formally replicates their finding. Bennet et al. [2002] investigated 154 affected sib-pairs from Ireland and United Kingdom and found a LOD score (MLS) of 1.38 also peaking around D17S928. Curtis et al. [2003] investigated seven larger pedigrees with bipolar disorder and unipolar disorder from United Kingdom and Iceland. Assuming a recessive mode of inheritance and only including unipolar patients from bipolar families they found a four-point LOD score of 1.9 at D17S939 at 105.7 cM, i.e., proximal to the loci suggested by the other studies. This region did not receive much support in a recent meta-analysis of bipolar disorder [Segurado et al., 2003]. However, the meta-analysis ranked findings from individual investigations and the genome-wide scans included in the meta-analysis mainly investigated larger pedigrees, which may preferentially harbor dominant risk alleles. The finding of a common region on chromosome 17q in four studies using different linkage analysis methods and different family material, ranging from single affected individuals to large pedigrees from various mainly Caucasian populations, is encouraging. The 17qter is a gene rich region which contains a number of interesting neurogenes.

Though the families were obtained from a specific subregion in Cuba it is uncertain whether this is a founding population. Around 700 inhabitants lived in the Holguin region in 1726 later from other regions of Spain. In 1862, the population had
Homozygosity Mapping in Bipolar Disorder

increased to 53,128. Today the size of the population is above 1 million. In a heterogenous population many different haplo-
types may harbor a risk locus in a specific region. This will also
decrease the power when applying homozygosity mapping as many
affected individuals will be heterozygous at neighboring marker
loci though they may be homozygous at the risk locus.

In each family multipoint LOD scores close to the maximum
obtainable (Table I) were identified in one or more chromosome
regions. This indicate that the density and informativity of
SNPs in general is sufficient to detect segments which are
inherited HBD, whether containing a disease allele or not,
through the specified genealogy. The affected sib-pairs in
families 4 and 19 yielded LOD scores on chromosome 10 close to
the maximum obtainable value (Table I). All offspring of
parents who are first cousins yielded from one to five regions
with LOD scores above 0.5, most of which were above 0.6. So
even when genotyping only one person, the affected offspring,
in each family the LOD scores were close to the maximum
obtainable value given the genetic parameters chosen (Table I).
Concerning LOD scores in families where the parents of cases
were more distantly related, more variable maximum LOD
scores were found ranging from a high LOD score of 0.66 in
family F7V1 (first and second cousins, Fig. 1) to no LOD scores
above 0.3 in family F9V2 (second cousins, Fig. 1). However, all
offspring of second cousins except F9V2 yielded LOD scores
above 0.4. The highest LOD score in F3V1, whose parents were
second and third cousins, was 0.34.

Our study indicates that it may be possible to identify regions
HBD by genotyping only a single offspring from first or second
cousins using microarrays with 1,100 or more SNPs. Cases
without affected first degree relatives also yielded regions with
high LOD scores, i.e., three, four, three, five, two, and zero
regions with a LOD score above 0.5 in families F2V1, F3V9,
F3V30, F7V1, F2V21, and F26V1, respectively. Therefore,
it seems reasonable also to apply homozygosity mapping to
“sporadic” cases with bipolar disorder and probably also in
other complex disorders with high twin concordance ratios.

The genetic intermarker distances used in the calculations
are derived directly from the physical distances by translating
1 Mb to 1 cM. Even if these estimates are not correct, we feel
confident that the LOD scores are rather robust towards mis-
specification of genetic distances as the high density of SNPs
ensures detection of HBD regions.

False low LOD scores may be obtained by multipoint anal-
yses if the specified marker order is incorrect. We chose also to
perform single point analyses in order to diminish the risk of
missing interesting chromosome regions due to wrongly order-
ated markers.

By regenotyping ten SNPs, we found that a single of these
had a high error rate while the others were quite accurate. A
high error rate in individual markers may lead to false nega-
tive findings. However, as the average size of the homozygous
segments surrounding a recessive disease gene should be
around 25–30 cM, corresponding to eight to ten markers, such
segments might be identified by applying single point analyses
in addition to multipoint analyses. In addition, if a single
heterozygous marker is surrounded by a long homozygous
segment this may be a genotyping error which could be de-
tected by regenotyping. Concerning the direction of the errors
if the biallelic marker is in Hardy–Weinberg equilibrium, most
alleles will be truly homozygous (except when both alleles
are of equal frequency). If the percentages of misclassified
genotyping are equal for homozygous and heterozygous ge-
notypings, more homozygous genotypes will be wrongly classified
as heterozygous, than vice versa.

We suggest that homozygosity mapping might be applied for
finding recessive risk alleles in complex disorders. It may be
possible among offspring from first or second cousins and even
more distant related parents. It may also be useful for finding
risk genes in late onset disorders where parents are often not
available, for finding rare phenotypic subgroups, endopheno-
types or modifying genes in monogenic or complex diseases. If
a number of affected siblings could be investigated it may even be
applied in very heterogeneous disorders [Wright et al., 1997]. It
seems to be the only possible method for doing linkage analyses
on families with only one affected person. Sporadic cases are
more common than familial cases, and thus clinically very
important, in most complex disorders. In the future homo-
zygosity mapping of even distantly related cases may be faci-
litated by using microarrays with a very high density of SNPs.
Finally, the use of SNP microarray genotypings is a flexible
method to apply in genome scans as additional cases may be
included during the study.

ACKNOWLEDGMENTS

The authors thank the following persons and institutions: Jorge Martínez for his contributions to the Velasco history.
Director and senior psychiatrist Aksel Bertelsen, WHO Collaborating Centre for Research and Training in Mental Health,
Psychiatric Hospital in Aarhus, for thorough reading of all reports, as well as performing the diagnostic evaluations.
The Center of Neuroscience of Havana, Cuba, for their support in training the Cuban psychiatrists. The health and local authori-
ties from the Holguín province and Velasco region. The ABI 3100 was supported by grants from the Fonden Psykiatriske Forskninger
and Psykiatrisk Forskningsfond. Torben F. Orntoft was supported by The Danish Medical Research Council, Institute of Experimental
Clinical Research, Faculty of Health Sciences, University of Aarhus, Konsul Meyers Fund and Aarhus County.

REFERENCES

Association.
pedigrees multiply affected with bipolar disorder provides further support for the presence of a susceptibility locus on chromosome 12q23-24, and
suggests the presence of additional loci on 1p and 1q. Psychiatr Genet 13:77–84.
Ewald H, Flint T, Kruse TA, Mors O. 2002. A genome-wide scan shows significant linkage between bipolar disorder and chromosome 12q24.3.
and suggestive linkage to chromosomes 1p22-21, 4p16, 6q14-22, 10q26, and 16p13.3. Mol Psychiatry 7:734–744.


