Context: Dopamine receptor–mediated pathways play critical roles in the mechanism of addiction. However, associations of the D2 dopamine receptor gene (DRD2) with substance abuse are controversial.

Objective: To determine whether susceptibility sites resided at DRD2.

Design: Haplotype-based case-control analysis of 2 distinct populations using 10 single nucleotide polymorphisms (SNPs) with heroin dependence.

Setting: Universities of Mainz and Bonn, Germany, and 3 local hospitals in southwestern China.

Patients: Cases and control subjects recruited from China (486 cases, 313 controls) and Germany (471 cases, 192 controls).

Interventions: Genotyping for 10 SNPs by 5'-exonuclease fluorescence assays. The D' value of linkage disequilibrium and haplotypes were generated by the expectation-maximization algorithm.

Main Outcome Measures: Genotype, allele, and haplotype frequencies were compared between cases and controls by χ² tests constructed for each population. An additional 32 SNPs randomly distributed in the genome were genotyped for detecting population admixture in the 2 populations.

Results: A haplotype block of 25.8 kilobases (kb) was defined by 8 SNPs extending from SNP3 (TaqIB) at the 5' end to SNP10 site (TaqIA) located 10 kb distal to the 3' end of the gene. Within this block, specific haplotype cluster A (carrying TaqIB1 allele) was associated with a high risk of heroin dependence in Chinese patients (P = 1.425 × 10⁻²²; odds ratio, 52.80; 95% confidence interval, 7.290-382.5 for 8-SNP analysis). A putative recombination “hot spot” was found near SNP6 (intron 6 ins/del G), creating 2 new daughter haplotypes that were associated with a lower risk of heroin dependence in Germans (P = 1.94 × 10⁻¹¹ for 8-SNP analysis). There was no evidence of population stratification in either population.

Conclusions: These results strongly support a role of DRD2 as a susceptibility gene with heroin dependence in Chinese patients and was associated with low risk of heroin dependence in Germans.

Arch Gen Psychiatry. 2004;61:597-606

Although epidemiologic studies have shown that heroin dependence is strongly influenced by genetic factors (h²=0.54), the number and identity of susceptibility genes remain unknown. Animal and human studies of addiction indicate that the D2 dopamine receptor (DRD2) plays a critical role in the mechanism of reward and reinforcement behavior. Opiate rewarding effects were absent in mice lacking D2 receptors, while DRD2 overexpression in transgenic mice led to reduced self-administration of alcohol. A positron emission tomography study of human brain showed that D2 receptor density in the brain decreased significantly in alcoholic compared with control subjects. These findings suggest that genetically determined variation in DRD2 expression and function can alter reward responses to a variety of substances and may contribute to vulnerability to heroin dependence in humans.

DRD2 is located on 11q22-23 and is composed of 8 exons spanning 65.8 kilobases (kb) of genomic DNA. The first DRD2 genetic marker characterized was a single nucleotide polymorphism (SNP) originally detected as a restriction fragment length polymorphism (TaqIA) located 10 kb distal to the 3' end of the gene. This marker was extensively used in ge-
netic association and linkage investigations of addiction, with controversial results. Linkage of the \textit{TaqIA} restriction fragment length polymorphism was also evaluated in many other psychiatric disorders, also with varying results. Studies using known functional DRD2 SNPs (−141ins/delC and 311 Ser>Cys) in alcohol dependence and a mixture of other substance dependencies, detected no association with risk.

Among the causes of controversial findings in population-based studies are small sample size with reduced power to detect effect, linkage disequilibrium (LD) of associated markers with other unknown functional loci, and population structure (admixture). To detect association with moderately abundant alleles, the LD (also known as allele-based linkage) paradigm with functional alleles or highly informative haplotypes offers substantially greater power for mapping complex disease or trait genes than does the locus-based linkage approach. Linkage disequilibrium detects the physical correlation between the genetic markers that define a group of alleles or haplotype. A haplotype block defines a region of the genome showing little historical recombination. Thus, a panel of 5 to 6 moderately informative SNPs contained within a haplotype block captures the effect of any relatively abundant, but unknown, functional allele within the haplotype block. Haplotype association also has the advantage of narrowing the location of disease loci and reducing or clarifying discrepancies in results between studies using different populations, allowing disparate data to be reconciled or at least better understood. Thus, haplotype-based association becomes an important approach to investigate the relationship of DRD2 and addictive behavior, now that a detailed SNP map from public and private databases (ie, Celera Discovery System, Rockville, Md) is available for this gene.

As mentioned earlier, population structure has been thought of as one of the reasons to explain unreplicated results from population-based association studies. When case and control samples are collected from different subpopulations, allele frequencies will tend to differ for most randomly chosen loci. Admixed populations can be detected by genotyping a number of markers and detecting systematic differences in allele frequency within the study population. Simulated analyses have suggested that 30 SNP markers should have reasonable power to detect stratification in subpopulations. With this approach, population admixture in African American subpopulations has been detected. To better understand whether DRD2 is associated with substance abuse, our strategy was to use a combined haplotype–functional locus approach in 2 large, ethnically well-defined heroin-dependent case-control samples derived from Chinese Han and German populations. To control for sample stratification, we genotyped 32 SNPs in our case and control groups for the Chinese and German populations. In addition, we genotyped an admixed African American population with the use of the same SNP panel as a positive reference sample set. To our knowledge, this is the first large-scale haplotype analysis of DRD2 in heroin dependence that controls for population admixture.

\section*{METHODS}

\subsection*{Chinese Data Set}

A total of 799 subjects, composed of 486 heroin-dependent cases and 313 unrelated and unaffected controls, were recruited in 3 waves during 1996, 1997, and 1999 from southwestern China, including Sichuan Province and Chongqing City, a federal district that is geographically adjacent to Sichuan Province. The Chinese Han population and data collection were described in more detail elsewhere. Patients were interviewed with the Structured Clinical Interview for DSM-III-R Axis I disorders, and diagnosed as opiate dependent by 2 psychiatrists using DSM-IV criteria. Other substance abuse, such as cocaine and cannabis, was uncommon in this area. Control subjects were recruited from students and staff at a local medical university. Control subjects were asked only if they had had a mental disorder, had been prescribed medication for a mental illness, or used a drug for a nonmedical purpose. The mean ± SD age of heroin-dependent subjects was 27.3 ± 5.80 years, and that of controls was 28.0 ± 10.0 years. Informed consent was obtained under a human research protocol approved by ethics committees at the 3 local hospitals and 1 local medical school.

\subsection*{German Data Set}

A total of 663 individuals were recruited, including 471 heroin-dependent subjects from 2 western German cities, Mainz and Bonn, and 192 unrelated controls from Bonn. Both cities are situated along the Rhine River, where they are separated by 150 km and have similar population structure. In Germany, citizens are obliged to register births and relocations with local authorities. Sample collection took place between 1993 and 2001 as part of a study on genetic and psychosocial risk factors in alcohol and heroin dependence. Cases were consecutive inpatients of the university hospital detoxification units at Mainz (1993-1999) and Bonn (1996-2001). Subjects were interviewed by senior psychiatrists using the Semi-structured Assessment for the Genetics of Alcoholism for psychiatric disorders and diagnosed as opiate dependent by DSM-III-R. Unrelated controls were randomly ascertained from the population registries of Bonn and represent the local population. The mean ± SD age of cases was 30.2 ± 6.8 years and that of controls was 31.8 ± 7.0 years. Control subjects were contacted by mail or telephone by the same research staff who recruited case subjects. Evaluation was the same as for the case group, including Semi-structured Assessment for the Genetics of Alcoholism diagnostic interview for psychiatric disorders. Informed consent was obtained under a human research protocol approved by the ethics committees at the University of Mainz and the University of Bonn.

\subsection*{SNP GENOTYPING BY 5’-EXONUCLEASE FLUORESCENCE ASSAY}

Genotyping for 10 SNPs of DRD2

Genotyping was performed by 5’-exonuclease fluorescence assay. We developed 10 SNP assays for DRD2 genotyping. From 5’ end to 3’ end, these 10 SNPs were as follows: −241 A>G, −141ins/delC, TaqIB A>G, TagID G>A, intron 4 T>C, intron 6 ins/del G, 311 Ser>Cys, 20236 C>T, exon 8 822640 C>G, and TaqIA G>A. Their corresponding National Center for Biotechnology Information SNP identification, Celera Discovery System identification, and primer-probe sequences are available from the corresponding author on request. The SNP locations are shown in Figure 1.

\textit{Arch Gen Psychiatry} 2004;61:598-605

\textbf{©2004 American Medical Association. All rights reserved.}
Figure 1. Human D2 dopamine receptor gene structure and single nucleotide polymorphism (SNP) sites for the haplotype SNP sets studied. The size of the gene is 65.8 kilobases (kb). The SNP10 (TaqIA) is located 10 kb downstream of 3′ with a total coverage for this study of 75.8 kb. A 50-kb intron separates exon 1 from exon 2, previously described as 250 kb. Exons are shown in black boxes. Each SNP was assigned a site number, and the SNPs are arranged from 5′ to 3′. The SNP2 and SNP7 have previously been shown to alter the function of the D2 dopamine receptor gene. Three SNP sets and their physical coverage for haplotype-based association are indicated by brackets.

For each SNP, genotyping error rates were determined by duplicate genotyping of an additional 10% of the samples randomly selected from each reaction plate.

Genotyping for Population Admixture

To detect population structure, an additional 32 SNPs were chosen from the National Center for Biotechnology Information public database. The SNPs were distributed on 17 chromosomes in the genome. The SNP rs identifications and physical locations are available from the corresponding author on request. Most markers showed large differences in allele frequencies across 8 different populations (K.X., unpublished data, 2003). Because DNA was available in limited amounts in Chinese samples, we were not able to genotype the entire sample set for all 32 markers. Therefore, we randomly selected 106 individuals from control (46) and case (60) groups for genotyping with this SNP set. We genotyped 194 control individuals and 286 case subjects in the Germans. In addition, we genotyped 174 African American individuals who were known to represent an admixed population for use as a reference by using the same 32-marker SNP set.

STATISTICAL ANALYSES

Association Analyses of DRD2 With Heroin Dependence

For individual SNP association analyses, genotype and allele frequencies in cases and controls were compared by χ² tests on 2×3 and 2×2 categorical tables constructed for each population. To exclude false-positive results due to multiple testing, Bonferroni correction was used. The P values were multiplied by the total number of loci genotyped (10). This was recognized to be a conservative correction because of extensive LD across DRD2.

For LD analysis, \( D = P_{ab} - (P_a \times P_b) \), where D is a parameter of LD, \( P_{ab} \) is the expected haplotype frequency, and \( P_a \) and \( P_b \) are observed frequencies for alleles at loci A and B, respectively. \( D^2 \) is \( D \) normalized against the maximum value of \( D \) possible, given allele frequencies \( P_a \) and \( P_b \). \( D^2 \) for each DRD2 SNP pair was computed with the help of PAIRWISE software (Jeffery C. Long, PhD, University of Michigan, Ann Arbor).

Ten-SNP DRD2 haplotype frequencies were inferred separately for cases and controls in each population by means of an expectation-maximization algorithm implemented in MLOCUS.27 A likelihood ratio test for global haplotype effects \((G)\) was performed with the following equation:

\[
\text{log}\left(\frac{G}{2\text{IxIn total}}\right) = (\text{ln} P_{\text{case}} + \text{ln} P_{\text{control}})
\]

Where \( P_{\text{case}} \) and \( P_{\text{control}} \) are the observed haplotype frequencies in cases and controls, respectively.

Population Admixture

We performed a contingency \( \chi^2 \) test for comparing allele frequencies for each marker and all markers between case and control groups in each population. Under the null hypothesis that the populations have the same allele frequencies, the sum of the statistics for all of the markers has a \( \chi^2 \) distribution with degrees of freedom equal to 1 less the total numbers of SNPs. We also compared overall the allele frequency among 3 control populations: Chinese, Germans, and African Americans.

We used the computer program Structure26 in an attempt to identify clusters of genetically similar individuals from multilocus genotype data.

RESULTS

ALLELE FREQUENCIES AND HARDY-WEINBERG EQUILIBRIUM FOR CONTROLS AND HEROIN-DEPENDENT SUBJECTS

Genotypes determined by 5′-exonuclease assay for the 10 DRD2 SNPs were highly accurate. Genotype discrepancy rates across the 10 loci were only 0.021±0.018 in...
### Genotype and Allele Frequencies for Individual SNPs in Cases and Controls in 2 Populations

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype/Allele</th>
<th>No. (Frequency)</th>
<th>(\chi^2)</th>
<th>(P) Value</th>
<th>Genotype/Allele</th>
<th>No. (Frequency)</th>
<th>(\chi^2)</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-214</td>
<td>1 = A; 2 = G</td>
<td>Control: 11 210 (0.695)</td>
<td>6.299</td>
<td>.042</td>
<td>Case: 12 258 (0.621)</td>
<td>493 (0.883)</td>
<td>1.85</td>
<td>.397</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1 = A; 2 = G</td>
<td>Control: 22 15 (0.050)</td>
<td>2.360</td>
<td>.124</td>
<td>Case: 22 192 (0.201)</td>
<td>0 (4.009)</td>
<td>23 (0.121)</td>
<td>50 (0.108)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-141</td>
<td>1 = G; 2 = DelC</td>
<td>Control: 12 397 (0.783)</td>
<td>0.042</td>
<td>.897</td>
<td>Case: 22 15 (0.003)</td>
<td>5 (0.010)</td>
<td>2 (0.010)</td>
<td>5 (0.010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaqIB</td>
<td>1 = G; 2 = A</td>
<td>Control: 11 567 (0.917)</td>
<td>5.480</td>
<td>.002 (.02)</td>
<td>Case: 2 112 (0.118)</td>
<td>41 (0.107)</td>
<td>98 (0.104)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 1</td>
<td>1 = T; 2 = C</td>
<td>Control: 22 101 (0.457)</td>
<td>1.477</td>
<td>.478</td>
<td>Case: 22 151 (0.357)</td>
<td>0.121</td>
<td>50 (0.108)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 2</td>
<td>1 = G; 2 = DelC</td>
<td>Control: 22 46 (0.202)</td>
<td>1.890</td>
<td>.124</td>
<td>Case: 22 115 (0.309)</td>
<td>0.187</td>
<td>145 (0.157)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>311 Ser&gt;Cys</td>
<td>2 = DelC</td>
<td>Control: 22 112 (0.539)</td>
<td>0.766</td>
<td>.382</td>
<td>Case: 22 112 (0.539)</td>
<td>0.766</td>
<td>.382</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 8</td>
<td>1 = G; 2 = G</td>
<td>Control: 22 168 (0.542)</td>
<td>0.900</td>
<td>.367</td>
<td>Case: 22 168 (0.542)</td>
<td>0.900</td>
<td>.367</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaqIA</td>
<td>1 = G; 2 = A</td>
<td>Control: 22 44 (0.141)</td>
<td>0.819</td>
<td>.366</td>
<td>Case: 22 78 (0.161)</td>
<td>0.17</td>
<td>33 (0.073)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Cys, cysteine; Ser, serine; SNP, single nucleotide polymorphism. *Significant \(P\) value after Bonferroni correction.

The Chinese and 0.010±0.017 in the Germans. No significant deviation from Hardy-Weinberg expectations occurred in Chinese controls, German controls, or German cases. In Chinese cases, SNPs8 showed a slight departure from Hardy-Weinberg equilibrium that remained marginally significant (\(P = .05\) after correction for multiple testing).

**INDIVIDUAL SNP ASSOCIATIONS WITH HEROIN DEPENDENCE**

Cases and controls were compared for genotype and allele frequencies across the 10 DRD2 markers (Table). In the Chinese population, genotype frequencies at 4 sites differed significantly between cases and controls. The sig-
significant sites were SNP1 (-214 A>G; P = .042); SNP3 (TaqIB A>G; P = 1.71 x 10^{-4}), SNP4 (TaqID G>A; P = .01), and SNP5 (intron 4 T>C; P = .024). Only the SNP3 genotype frequency remained significant after applying a conservative Bonferroni adjustment for multiple testing (P = 1.71 x 10^{-3}). Allele frequency comparisons between cases and controls were significant at 4 sites: SNP2 (-141ins/delC; P = .002), SNP3 (TaqIB A>G; P = 3.8 x 10^{-5}), SNP4 (TaqID G>A; P = .006), and SNP5 (intron 4 T>C; P = .005). Only 2 SNPs remained significant after Bonferroni correction: SNP2 (-141ins/delC; P = .02) and SNP3 (TaqIB; P = 3.8 x 10^{-5}). In the German population, genotype and allele frequency comparisons were not significantly different between case and control groups across 10 markers.

LINKAGE DISEQUILIBRIUM

Pairwise LD for the 10 SNPs at DRD2 is presented separately for cases and controls from each population in Figure 2. Values on the abscissa and ordinate are physical distances (logarithmic scale). Levels of D' differ between populations and between clinical diagnoses. In Chinese controls, 24% of SNP pairs were in complete LD (D'>0.99), while in Chinese cases only 4% of pairs were in complete LD. In German controls, 13% of SNP pairs were in complete LD, while in German cases 42% of pairs were in complete LD. In both populations, 2 SNPs within the promoter region (SNP1 and SNP2) presented weak LD with the other 8 SNPs (SNP3 to SNP10) in the rest of DRD2 region. The 8 SNPs (SNP3 to SNP10) spanned 25.8 kb, with high LD levels displayed in both Chinese and Germans (D' = 0.804 ± 0.196 in Chinese; D' = 0.801 ± 0.228 in Germans). A core conservative LD block included 6 SNPs (from SNP4 to SNP9) spanning 10.8 kb with strong D' (0.934 ± 0.069 in the Chinese; 0.897 ± 0.174 in the Germans). The strength of LD provided a justification to divide the entire region into discrete windows for subsequent haplotype-based association analyses (Figure 2).

HAPLOTYPE-BASED ASSOCIATION

Chinese Haplotype Structure and Association With Heroin Dependence

We used 3 SNP sets for haplotype-based association analyses, grouping SNPs on the basis of the level of LD strength. We used 6-SNP (SNP set 4-9), 8-SNP (SNP set 3-10), and 10-SNP (SNP set 1-10) sets to create the windows for performing each of the 3 separate analyses.

In 6-SNP core haplotype block, there were 2 configurations (A: 111121; and B: 112112) that accounted for 89% of all chromosomes in Chinese subjects (highlighted in yellow and blue, respectively in Figure 3). The global haplotype pattern differed significantly between controls and heroin-dependent subjects (G_{10} = 129.7, P < 1.771 x 10^{-10} after multiple test correction). However, specific haplotypes in the 6-SNP block did not differ significantly (Figure 3).

Adding 2 flanking loci, SNP3 (TaqIB A>G) and SNP10 (TaqIA G>A), that were in strong LD with the markers in the 6-SNP block increased the information content within the 25.8-kb region defined by this window. With the use of 8 SNPs, 6 haplotypes were generated and grouped as 4 major haplotype clusters: A, B, C, and D (Figure 3). Each was defined by the core 6-SNP haplotype and by 1 allele of each of the 2 flanking SNPs: TaqIB, at the 5' end, and TaqIA, at the 3' end of the haplotype block. Tests for global haplotype association with heroin dependence were significant for the 8-SNP haplotype (G_{9} = 322.3, P < 4.720 x 10^{-10}, after multiple test correction). Among 4 haplotype clusters, 2 haplotype clusters, 8S-A and 8S-C, were observed in cases but not in controls (frequency, 0.149 vs 0.000; Fisher P = 1.425 x 10^{-3}; odds ratio [OR], 52.80; 95% confidence interval [CI], 7.290-382.5 for cluster A; 0.063 vs 0.000, Fisher P = 3.471 x 10^{-4}; OR, 40.19; 95% CI, 5.550-291.1 for cluster C). In contrast, haplotype cluster 8S-B, corresponding to 10S-B (see below), was at higher frequency in controls than cases (0.460 vs 0.347; P = 1.140 x 10^{-5}; OR, 0.667; 95% CI, 0.456-0.857). With this approach, it became apparent that both adjacent SNPs (TaqIB and TaqIA) at opposite ends of the block added critical information, thus localizing the effective locus to this 25.8-kb region. For example, allele 1 (shown in red in Figure 3) of the TaqIB locus combined with 111121 (coded yellow in Figure 3) defined a high-risk haplotype for heroin dependence. Without the added information from TaqIA allele 1, haplotype 2111211 appeared to be low-risk (coded as green and yellow in Figure 3). On the basis of frequencies of alleles and haplotypes, the TaqIB appeared to add more predictive information than TaqIA.

Finally, using all 10 available SNPs, we simultaneously evaluated the entire 75.8-kb region. The global 10-SNP haplotype test for association was significantly different between control subjects and heroin addicts (G_{10} = 237.2, P < 1.916 x 10^{-10} after multiple test correction). Two of the clusters, 10S-A and 10S-C, corresponding to 8S-A and 8S-C, were observed only in cases (cluster 10S-A: frequency, 0.119 in cases vs 0.000 in controls, Fisher P = 2.499 x 10^{-4}; OR, 77.79; 95% CI, 4.793-1268; and cluster 10S-C: 0.036 in cases vs 0.000 in controls, Fisher P = .001; OR, 22.43; 95% CI, 1.340-375.3), while 10S-B was significantly more abundant in controls (0.422 in controls vs 0.326 in cases; P = .013; OR, 0.678; 95% CI, 0.499-0.922). These data suggested that haplotype clusters 10S-A and 10S-C represented high-risk copies of DRD2, while haplotype cluster 10S-B may represent low-risk copies of DRD2 with heroin dependence.

Haplotype Structure and Association With Heroin Dependence in the German Population

Applying the same strategy used with the Chinese dataset, 3 SNP haplotype sets were analyzed for association with heroin dependence in Germans (Figure 4). Overall haplotype tests showed that DRD2 was significantly associated with heroin dependence in 3 SNP haplotype sets analyses (G_{9} = 105.0, P < 1.617 x 10^{-10} for 6-locus after Bonferroni correction; G_{12} = 134.0 P<1.000 x 10^{-10} for 8-lo-
cus after Bonferroni correction; \( G_{10} = 138.4, \ P < 4.570 \times 10^{-10} \) for 10-locus after Bonferroni correction.

As seen previously in the Chinese, a 6-SNP core haplotype block was observed in both German cases and controls. Within the core haplotype block, 3 major 6-SNP haplotypes (6S-H1, 6S-H2, and 6S-H3) accounted for 79% of the chromosomes in controls and 91% in cases. Two major haplotypes (6S-H2: 111121; and 6L-H3: 112112) were identical to haplotypes in the Chinese (yellow and blue blocks, Figure 3) but the most frequent core haplotype (6S-H1: 221112; Figure 4) in Germans differed from that in the Chinese, accounting for 47% of Germans but representing only 7% of Chinese. We observed a possible recombination event in the German population between SNP5 and SNP6 produced from the 2 abundant haplotypes, 6S-H1 and 6S-H3, resulting in 2

**Figure 2.** Pairwise single nucleotide polymorphism (SNP) linkage disequilibrium of the D2 dopamine receptor gene across control and heroin-dependent groups in Chinese and German populations. Linkage disequilibrium (LD) was measured by \( D' \) with the MLOCUS program. \( D' \) lies in range from 0 to 1 and is shown in different colors (highest \( D' \) is in red, while lowest \( D' \) is in blue). Numbers on the x-axis show log values of the actual physical distance for pairwise \( D' \) for SNPs 1 through 10 (5' to 3'). The SNP order is repeated top to bottom in each panel. There are 4 LD panels: A, Chinese control; B, Chinese heroin dependent; C, German control; and D, German heroin dependent. An LD block contained 8 SNPs from SNP3 to SNP10. A core, more conservative LD block contained 6 SNPs from SNP4 to SNP9. The percentages of complete LD (\( D' > 0.99 \)) SNP pairs for each panel were 24%, 4%, 13%, and 42% for panels A, B, C, and D, respectively. The strongest LD was panel D, with mean ± SD \( D' = 0.796 ± 0.211 \), while the weakest LD was panel B, with mean ± SD \( D' = 0.573 ± 0.338 \) across 10 SNPs.
daughter haplotypes (6S-H5 and 6S-H6) that were not seen in Chinese subjects. These 2 daughter haplotypes, which accounted for 10.2% of all haplotypes, were represented only in the control group (Fisher $P = 1.614 \times 10^{-11}$). This difference in frequency strongly suggested that haplotypes 6S-H5 and 6S-H6 were associated with lower risk of heroin dependence in the German population. In fact, the 6-SNP region covering 10.8 kb even more narrowly defined the affected region than in the Chinese population.

The 8-SNP analysis showed a different pattern and predictive outcome among haplotypes in Germans as compared with Chinese (Figure 4). Similar to the 6-SNP analysis, we also observed 2 common haplotypes (8S-H1 and 8S-H3) whose recombination near SNP 6 resulted in 2 daughter haplotypes (8S-H5 and 8S-H6) that predicted low risk of heroin dependence in German populations (Fisher $P = 1.940 \times 10^{-11}$). The SNPs TaqIA and TaqIB were in the LD block but did not contribute additional information here, a result that differed from the Chinese population.

With the 10-SNP window used for analysis, 1 haplotype (10S-H2) was more frequent in the cases than in the controls in Germans (0.100 in the controls, 0.148 in the cases), and was modestly significant ($P = .020$; OR, 1.595; 95% CI, 1.089-2.338) (Figure 4). We also found evidence of recombination by means of the 10S-locus haplotype set. However, only haplotype 10S-H5 was represented at significantly higher frequency in controls (0.050 vs 0.000; Fisher $P = 1.100 \times 10^{-5}$).

Testing for Admixture in the 2 Populations

Although 32 SNPs were initially selected for analysis, 2 SNPs from Chinese and 4 SNPs from Germans were removed from the test because of high genotyping failure rate or deviation from Hardy-Weinberg equilibrium or for being monomorphic in a population. Thus, a total of 30 SNPs for the Chinese and 28 SNPs for the Germans were used for these analyses. Within each population, a comparison of allele and genotype frequencies between case and control groups for each marker failed to show any significant difference (data not shown). In addition, overall allele frequencies for SNP loci did not differ between case and control groups in either the Chinese ($P = .744$) or the Germans ($P = .183$), as expected. Between the populations, allele frequencies for all markers showed significant differences among Chinese, Germans, and African Americans ($P < .001$ for each comparison). By this approach, there was no evidence

---

**Figure 3.** Haplotype clusters and frequencies of 3 single nucleotide polymorphism (SNP) sets at the D2 dopamine receptor gene in Chinese case and control samples. High-risk haplotype and low-risk haplotype clusters for heroin dependence were determined with 3 sets of haplotype analyses: 6-SNP, 8-SNP, and 10-SNP for case and control groups, performed separately, using the MLOCUS program. Four clusters, A, B, C, and D, were generated in 8- and 10-loci analyses: core haplotypes A and B were obtained in the 6-SNP analysis. The combination between the block and surrounding SNPs, SNP3 and SNP10, showed significant differences between case and control groups (8-SNP and 10-SNP). The block shown in yellow with allele 2 of SNP3 (cluster B) was more abundant in the control group than the case group; while the block shown in yellow with allele 2 of SNP3 (cluster B) was more abundant in the control group than the case group; the combination of the block shown in yellow containing allele 1 of SNP10 (TaqIB) (cluster C) was only represented in the case group. OR indicates odds ratio; CI, confidence interval; asterisk, Fisher exact test; and dagger, comparison of haplotype cluster between control and heroin-dependent groups.
of population admixture between case and control groups in each of the study populations.

Using the Structure 2.0 program (available at: http://pritch.bsd.uchicago.edu; Jonathan Pritchard, PhD, The University of Chicago, Chicago, Ill) for detecting population admixture in either Chinese or Germans produced only 1 cluster when applied to the combined population or to separate case and control groups (K=1, postprobability=0.999 for each test). However, in African Americans, there was evidence of population admixture (K=2, postprobability=0.999). These data indicated that the markers selected were able to detect population structure in an admixed African American population and provided support that the Chinese and German populations used in the present study were homogeneous.

In this study, we found that specific DRD2 haplotypes were highly associated with heroin dependence in both Chinese and German populations. In addition, singleton association with heroin dependence in Chinese was significant. Global tests of haplotype association were significant at the level of 3 SNP sets in both populations. A 25.8-kb region defined by 8 SNPs was implicated more strongly over any individual SNP analyzed in the Chinese, while a 10.8-kb region containing 6 SNPs supported a low-risk region for heroin dependence in Germans. Moreover, our data showed that there was no evidence of population admixture in either Chinese or Germans by testing additional genetic markers.

Previous studies using the known functional alleles have been contradictory or nonsupportive of DRD2 association with alcoholism and other addictions. Therefore, it would be advantageous to use markers spanning the entire DRD2 region and incorporate into the analysis any new in vitro functional variants available. In this study, we found that the −141delC allele at SNP2 was slightly more abundant in Chinese heroin addicts, yet genotype-based comparison between cases and controls did not share this difference. Also, the significance level for −141ins/delC was less than for the TaqIB at SNP3, even in Chinese (P=0.021 for −141ins/delC vs P=3.8×10⁻³ for TaqIB). The −141ins/delC is outside the implicated haplotype block, suggesting that −141delC plays a minor role in heroin dependence in Chinese. Our study also showed that TaqIB was strongly associated with heroin dependence in Chinese, which was consistent with previous studies. Therefore, TaqIB is located within intron 1, but it may be in LD with an unknown functional SNP within the LD block. It should be noted that this SNP was also in strong LD with other SNPs within the LD block. Furthermore, haplotype data substantially increased the significance level of the association. As discussed in the introduction, the TaqIA marker was previously implicated in alcoholism and substance dependence, but not in heroin dependence. Although our data did not support a particular role for the TaqIA polymorphism in heroin de-
pendsence, this SNP did add information to the 8-SNP haplotype, increasing the strength of linkage in Chinese but not in Germans. These results supported the idea that association of haplotypes rather than any individual SNP points to an unknown effective variant or variants within the 25.8-kb region.

Because no single functional variant of DRD2 has previously been associated with heroin dependence, and because it is unknown whether the known variants that alter function in vitro also alter in vivo dopamine biology, LD analysis is an important step in detecting the action of an effective variant or variants somewhere in DRD2. For the pairwise LD matrix using DRD2 gene SNPs, we determined that a strong LD block extended to 25.8 kb in the DRD2 gene, across both populations. Similar to a report by Kidd et al,30 we observed similar LD patterns across the DRD2 gene in our study populations. Three SNPs (TaqIB, TaqID, and TaqIA) used in this study were the same as those used by Kidd and colleagues, but we applied these markers to much larger sample sizes in this study. The mean D' for these 3 SNPs was 0.885 ± 0.084 in our study compared with 1.000 ± 0.000, determined by Kidd and coworkers’ study30 for the Chinese Han population, while mean D’ was 0.903 ± 0.144 in our German population compared with 0.700 ± 0.111 in a Finnish population. In addition, the ancestral haplotype defined by the Kidd et al study corresponded to the same ancestral haplotype found in both Chinese and Caucasian populations. The most frequent haplotype in Chinese, B1D2A1, also the ancestral haplotype, had a frequency of 0.37 compared with 0.36 from Kidd and coworkers’ study.30 Another haplotype B2D1A2 was the most abundant (0.450) for the German population in this study and had a frequency that compared with 0.417 for a Finnish population.30 More interestingly, we also found that the strength of LD in Chinese was greater than in the German population, where approximately 10% recombination has occurred in this genomic region in the German population. This accounted for the different pattern of haplotype diversity between heroin addicts and controls in the 2 populations. This interpretation may explain why different haplotypes were associated with heroin addiction in the 2 populations.

It is well known that allele-based LD analysis is a powerful tool for identifying effective loci, assuming that a sufficiently large sample size is used and that stratification-produced results can be minimized or eliminated.31,32 In the German and Han Chinese case-control populations we studied, individuals were recruited from the same geographic areas and represented relatively well-defined populations. Neither Germans nor Han Chinese are isolated or semi-isolated populations. However, our results for detecting sample stratification indicated no evidence of subpopulation (admixture) in either case or control group in the 2 populations. Therefore, the association of DRD2 with heroin dependence was unlikely false positive owing to stratified samples.

By mapping haplotype blocks in different populations, the evolutionary history of genes can help us to put in order apparently disparate linkage and association data on disease-associated genes. Furthermore, differences in block structure between populations may be identified so that populations with smaller block sizes can be chosen for the purpose of homing in on effective SNPs or microsatellite polymorphisms.33,34 However, in both of the populations we studied, the implicated DRD2 haplotype block is large, so additional markers in the region increase information content and improve the strength of association. Use of other analytical strategies (for example, sequence variant detection and establishment of function) or linkage disequilibrium studies in other populations will therefore be required to identify the effective variants in the near future.

Submitted for publication September 9, 2003; final revision received January 5, 2004; accepted January 21, 2004.

From the Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, Rockville, Md (Drs Xu, Lipsky, and Goldman; Miss Ferro, Finch, Terry, and Taubman; and Mr Astor); Department of Psychiatry, University of Bonn, Bonn, Germany (Drs Lichtermann, Franke, Schwab, Wildenauer, and Maier); Department of Psychiatry, Sichuan University Medical School, Sichuan, China (Drs Liu and Cao); Laboratory of Population Genetics, National Cancer Institute, Bethesda, Md (Dr Hu); and Department of Genetica, Instituto de Biociencias, Universidad Federal do Rio Grande do Sul, Porto Alegre, Brazil (Dr Bau).

This study was supported by the German Federal Ministry for Education and Research (BMBF), Bonn, in part by grants 01EB9418/5 and 01EB9802/0, and by grant 01EB0133 within the framework of the Nordrhein-Westfälische Interdisziplinäre Netzwerk on Alcoholism Study Research (Dr Maier at the University of Bonn). This study was also partly funded by the Chinese National Science Foundation, Beijing (Dr Liu at the Sichuan University).

We thank D. J. Yuan, MD, and Z. H. Zhu, MD, of the Department of Psychiatry, Medical Center of Sichuan University, Sichuan, China, and J. C. Long, PhD, Department of Human Genetics, University of Michigan, Ann Arbor.

Corresponding author and reprints: Ke Xu, MD, PhD, Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, 12420 Parklawn Dr, Parklawn Building, Room 451, Rockville, MD 20852 (e-mail: ke@mail.nih.gov).

REFERENCES

Correction

Error in Figure. In the Original Article by Lotspeich et al titled “Investigation of Neuroanatomical Differences Between Autism and Asperger Syndrome,” published in the March issue of the Archives (2004;61:291-296), an error occurred in Figure 4 on page 296. In that figure, the lines representing the Asperger syndrome (ASP) and control groups were reversed. Figure 4 is reprinted correctly here.

Figure 4. Correlation between cerebral white tissue volume and performance IQ. Groups included 21 subjects with Asperger syndrome (ASP), 18 with high-functioning autism (HFA), and 21 control subjects. For between-group differences, ASP vs HFA, z = −2.13 (P = .03); ASP vs controls, z = −2.88 (P = .004).