Genetic Etiology of Comorbid Reading Difficulties and ADHD

Erik G. Willcutt, John C. DeFries, Bruce F. Pennington, Shelley D. Smith, Lon R. Cardon, and Richard K. Olson

Plomin and McClearn's (1993) precursor to the present volume introduced psychologists to recent advances in behavioral genetics and suggested directions for future research. In the concluding chapter, Plomin (1993) highlighted several new research tools that were "fueling the current momentum of the field" (p. 467). These methodological advances included the use of multiple regression to understand the etiology of extreme phenotypic scores, the application of multivariate genetic analyses to test for pleiotropic effects of genes on two or more traits, and a paradigm shift toward the use of molecular genetic techniques to identify specific genes that influence behavior.

In a chapter in that previous volume, DeFries and Gillis (1993) summarized existing knowledge regarding the etiology of reading disability (RD). At that time, reading deficits had been shown to be familial, but studies that could disentangle genetic and environmental influences in families were still in their infancy. Results of several small twin studies were somewhat inconsistent but suggested tentatively that reading deficits were due at least in part to genetic influences. In a brief section at the conclusion of the chapter, DeFries and Gillis described preliminary evidence indicating that reading disability might be linked to quantitative trait loci (QTL) on chromosomes 6 and 15.

Subsequently, twin studies of substantially larger samples have provided conclusive evidence that reading disability is heritable, and linkage studies have localized possible genes for reading disability to at least four

This work was supported in part by program project and center grants from the National Institute of Child Health and Human Development (NICHD; HD-11681 and HD-27802) to John C. DeFries, grants from the National Institute of Mental Health (NIMH; MH-38820) and NICHD (HD-04024) to Bruce F. Pennington, and NIMH Individual National Research Service Award (MH-12100) to Erik G. Willcutt. We thank the school personnel and families who participated in the study, and Helen E. Datta, Terry Grupp, and Deborah Porter for database management.
chromosomal regions. The rapid accumulation of new knowledge about the genetic etiologies of reading disability and other related disorders illustrates the impact of both behavioral and molecular genetic methods on the development of knowledge about factors that influence both normal and abnormal development. However, results that have emerged in the past decade also reveal the complexity of the etiological pathways to reading difficulties and other multifactorial behaviors and underscore how much is yet to be learned.

We have four primary objectives in this chapter: (a) to review briefly previous family and twin studies of reading disability and to summarize the most recent results from the Colorado Learning Disabilities Research Center (CLDRC) twin study; (b) to summarize the results of genetic linkage studies of reading disability from our laboratory and others; (c) to describe results of studies that have applied behavioral and molecular genetic methods to understand the etiology of comorbidity between reading disability and attention deficit hyperactivity disorder (ADHD); and (d) to discuss future research directions for both behavioral and molecular genetic studies in this area.

**Behavioral Genetic Studies of Reading Disability**

Although the specific etiological mechanisms that lead to reading disability are still unknown, significant advances have been made in understanding the extent to which reading deficits are attributable to genetic or environmental influences. Previous studies demonstrate clearly that reading difficulties are familial (e.g., DeFries & Decker, 1982; DeFries, Vogler, & LaBuda, 1986; Finucci & Childs, 1983; Pennington et al., 1991). Specifically, the relative risk for reading disability is four to eight times higher in first-degree relatives of probands with a reading disability than in relatives of individuals without one (e.g., Gilger, Pennington, & DeFries, 1991; Pennington & Lefly, 2001). However, because members of intact nuclear families share both genetic and environmental influences, familial resemblance alone does not provide conclusive evidence for heritable variation.

By comparing the similarity of monozygotic (MZ) twins, who share all of their genes, with dizygotic (DZ) twins, who share half of their segregating genes on average, twin studies estimate the extent to which familiality is due to familial environmental versus genetic influences. Several early twin studies compared the probandwise concordance rates for reading disability in relatively small samples of MZ and DZ twin pairs (Bakwin, 1973; Stevenson, Graham, Fredman, & McLoughlin, 1984, 1987; Zerbin-Rüdin, 1967). Although results varied somewhat across these studies, all found that MZ twins were more likely than DZ twins to be concordant for reading disability, providing tentative evidence that reading deficits are due at least in part to genetic influences.
CLDRC Twin Study

Because of the paucity of well-designed twin studies of reading disability, a twin study was initiated in 1982 as part of the Colorado Reading Project (Decker & Vandenberg, 1985; DeFries, 1985). In 1991, this ongoing twin study was incorporated as one component of the CLDRC, a six-site research center devoted to the study of the etiology, assessment, and treatment of reading and other learning disabilities.

To minimize the possibility of referral bias, the researchers contact parents of all twins between ages 8 and 18 in 27 local school districts and obtain permission from parents to review the school records of both members of each pair of twins for evidence of reading problems. If either member of a twin pair manifests a positive history of reading problems (e.g., low reading achievement test scores, referral to a reading therapist because of poor reading performance, or reports by classroom teachers or school psychologists), both members of the pair are invited to complete an extensive battery of tests in our laboratories at the University of Colorado and the University of Denver.

The test battery includes the Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974) or Wechsler Adult Intelligence Scale—Revised (Wechsler, 1981) and the Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970). A reading achievement composite score is calculated on the basis of a discriminant function analysis of the PIAT Reading Recognition, Reading Comprehension, and Spelling subtests in a separate sample of nontwin individuals with and without a history of significant reading problems (DeFries, 1985). In addition, each participant completes several measures of language processes that are related to reading ability, and parents complete a battery of interviews regarding each twin’s emotional and behavioral functioning, including measures of ADHD.

The comparison of concordance rates in MZ and DZ twin pairs provides an initial test for genetic etiology of a discrete disorder. As of May 2001, 245 pairs of MZ twins and 195 pairs of same-sex DZ twins had been selected because at least one twin in the pair met criteria for reading disability based on a positive school history of reading difficulties and a score at least 1.5 standard deviations (SD) below the estimated population mean on the reading composite score. The probandwise concordance rate for reading disability is significantly higher in MZ twin pairs (65%) than in DZ twin pairs (35%), providing additional evidence that reading disability is at least partially attributable to genetic influences.

Multiple Regression Analysis of Twin Data

Although the simplicity of a comparison of concordance rates is appealing, increasing evidence suggests that reading disability, ADHD, and most other disorders are defined on the basis of arbitrary diagnostic cutoff points on a quantitative measure (e.g., Barkley, 1998; Stevenson et al.,
Transformation of a continuous measure such as reading performance into a categorical variable (e.g., reading disabled vs. unaffected) results in the loss of important information pertaining to the continuum of variation in reading performance. In contrast, multiple regression analysis of twin data provides a versatile and powerful test of the etiology of extreme scores on a continuous trait (DeFries & Fulker, 1985, 1988).

The Basic Regression Model

The DeFries–Fulker (DF) model is based on the regression of MZ and DZ cotwin scores toward the population mean when probands are selected because of extreme scores on a phenotype of interest (e.g., reading deficits). Although scores of both MZ and DZ cotwins would be expected to regress toward the mean of the unselected population (μ), scores of DZ cotwins should regress further than scores of MZ cotwins to the extent that extreme scores are influenced by genes (Figure 13.1). The magnitude of this differential regression by zygoisty provides an estimate of the heritability of the proband group deficit in reading (h²).

To illustrate the DF method, we fitted a multiple regression model to reading composite data from the current sample of probands and co-twins in the CLDRC twin study. All twins who scored more than 1.5 SD below the mean of the population on the reading discriminant function score were selected as probands. For pairs that were concordant for reading deficits, each twin was entered once as the proband and once as the cotwin. This “double-entry” procedure facilitates the most valid estimates of h² when a sample has been ascertained using truncate selection (e.g., DeFries & Alarcón, 1996; McGue, 1992). Because the double entry of con-

![Figure 13.1](attachment:image.png)

**Figure 13.1.** The distribution of reading achievement in the population: differential regression of monozygotic (MZ) and dizygotic (DZ) cotwin means toward the population mean (μ). RD = reading disability.
cordant pairs artificially inflates the sample size, the standard errors of
the regression coefficients and resulting \( t \) values were corrected for double
entry to obtain unbiased tests of significance (e.g., Stevenson, Pennington,
Gilger, DeFries, & Gillis, 1993).

Before the multiple regression analysis was conducted, standardized
scores were created based on the mean and standard deviation of the non-
reading disability sample to facilitate the estimation of \( h^2 \). The standard-
ized scores for the selected sample of MZ probands and cotwins were then
divided by the MZ proband mean, and the DZ proband and cotwin scores
were divided by the DZ proband mean. This procedure ensures that the
MZ and DZ probands are equally divergent from the mean of the popu-
lation prior to the regression analysis.

The basic regression model for the univariate case is as follows:

\[
C = B_1P + B_2R + K,
\]

where \( C \) is the expected cotwin score, \( P \) is the proband score, \( R \) is the
coefficient of relationship (1 for MZ pairs, 0.5 for DZ pairs), and \( K \) is the
regression constant. The \( B_1 \) coefficient represents the partial regression of
the cotwin’s score on the proband’s score and provides a measure of twin
resemblance irrespective of zygosity. The \( B_2 \) parameter represents the par-
tial regression of the cotwin’s score on the coefficient of relationship, and
after appropriate transformation of the data provides a direct estimate of
\( h^2 \). After adjustment of the standard errors of the regression coefficients
to correct for the double entry of concordant pairs, the significance of the
\( B_2 \) parameter provides a statistical test of the extent to which extreme
scores are attributable to genetic influences.

The means of MZ and DZ probands on the reading composite score
were almost identical, and both means were over 2.5 \( SD \) below the mean
of the comparison sample of twins without reading disability (see Table
13.1). The mean reading score of the MZ co-twins regressed less toward
the population mean than the reading score of the DZ co-twins, consistent
with what would be expected if genetic influences contribute to deficits in
reading. Indeed, when Equation 1 was fitted to the reading data, the re-
sulting estimate of extreme group heritability was highly significant: \( h^2 =
0.55 (\ .08), t = 6.88, p = 2.3 \times 10^{-11} \).

<p>| Table 13.1. Mean Reading Discriminant Scores of 245 Pairs of Monozygotic Twins and 195 Pairs of Dizygotic Twins |
|----------------------------------------------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Twin pair</th>
<th>Proband</th>
<th>Cotwin</th>
<th>Proband</th>
<th>Cotwin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monozygotic</td>
<td>( -2.61 )</td>
<td>( 0.85 )</td>
<td>( -2.39 )</td>
<td>( 1.03 )</td>
</tr>
<tr>
<td>Dizygotic</td>
<td>( -2.60 )</td>
<td>( 0.84 )</td>
<td>( -1.67 )</td>
<td>( 1.28 )</td>
</tr>
</tbody>
</table>

*Expressed as standardized deviations from the mean of 996 control twins.
Extensions of the Basic DF Model

The DF model is versatile and can be elaborated to test for the influence of other independent variables on the etiology of the trait under consideration. For example, the model to test whether $h^2_x$ for reading difficulties differs as a function of IQ is as follows:

$$C_{RD} = B_1P_{RD} + B_2P_{IQ} + B_3P_{IQ} \times P_{RD} + B_6P_{IQ} \times R + K,$$

where $C_{RD}$ is the expected cotwin reading score, $P_{RD}$ is the proband reading score, $R$ is the coefficient of relationship, $P_{IQ}$ is the IQ of the proband, $P_{IQ} \times P_{RD}$ is the product of the proband's IQ score and the proband's reading score, and $P_{IQ} \times R$ is the product of the proband's IQ and the coefficient of relationship. The $B_3$ term represents a test for the main effect of IQ on reading across zygoty, the $B_4$ coefficient tests for differential twin resemblance as a function of IQ, and the $B_5$ coefficient tests for differential $h^2_x$ of reading as a function of IQ.

The CLDRC dataset has recently been used to test whether the heritability of reading ability varies as a function of gender or IQ (Wadsworth, Knopik, & DeFries, 2000; Wadsworth, Olson, Pennington, & DeFries, 2000). Wadsworth, Knopik, and DeFries (2000) obtained almost identical estimates of the heritability of reading deficits in girls ($h^2_x = .59$) and boys ($h^2_x = .58$), providing little support for the hypothesis that genetic influences play a larger role in reading deficits in girls than in boys (Geschwind, 1981).

In contrast to the results for gender, Wadsworth, Olson, et al. (2000) found that the heritability of the group deficit in reading varied significantly as a function of IQ. Twin pairs were divided into pairs with an average IQ greater than 100 and pairs with an average IQ below 100, and Equation 2 was then fitted to data from both groups simultaneously. Results revealed that the estimate of $h^2_x$ was significantly lower in the group with IQ scores below 100 ($h^2_x = .43$) than the group with IQ scores of 100 or higher ($h^2_x = .72$). Similarly, the heritability of the reading deficit also increased significantly as a linear function of IQ ($p < .01$), suggesting that environmental influences may be a more salient cause of reading difficulties in children with lower general cognitive ability. These results demonstrate the potential utility of the DF method for identifying and understanding etiologically meaningful subtypes within a disorder.

Genetic Linkage Analysis

If a trait is found to be significantly heritable, genetic linkage analyses can be conducted to localize the genes that are associated with the trait. Linkage analysis is based on the extent to which alleles at two genetic loci are inherited together (for additional reading on linkage analysis, see Faroone, Tsuang, & Tsuang, 1999; Plomin, DeFries, McClearn, & McGuffin, 2001). If two genes are close together on the same chromosome, alleles at
the two loci are likely to be transmitted together from parent to child. In contrast, if the two genes are on different chromosomes or are far apart on the same chromosome, alleles at the two loci will be transmitted independently. Classical linkage analysis (CLA) tests whether a marker allele recombines with a putative gene for a categorical phenotype significantly less frequently than expected by chance across several generations of a family.

Although CLA is a powerful analytic technique in some situations, this method has several disadvantages for complex, heterogeneous disorders such as reading disability and ADHD. First, CLA is optimized for use with categorical definitions of disease and is best able to identify a "major gene" that accounts for a large amount of variance in the phenotype. As noted previously, a growing consensus suggests that disorders such as reading disability and ADHD may be conceptualized more appropriately as extreme scores on continuous symptom dimensions. Second, parametric linkage analyses require that the mode of inheritance and the penetrance of the locus be specified in the model, which may be difficult for a complex trait such as reading disability. Finally, CLA requires a reliable and valid measure of the phenotype in two or more consecutive generations of individuals in a large familial pedigree. If the nature of the phenotype changes significantly across development, this may further complicate the analysis.

**Multiple Regression Linkage Analysis**

As an alternative to CLA, sibling-pair linkage methods have been developed to localize QTL. QTLs are genes that influence susceptibility to a disorder but are not typically a necessary or sufficient cause of the disorder (see Plomin & Crabbe, 2000).

Haseman and Elston (1972) first advocated that regression analysis with pairs of siblings could be used to test for the effect of a QTL on individual differences in a dimensional trait. Fulker et al. (1991) then showed that when a sample is selected because at least one sibling exhibits an extreme score on a trait, an adaptation of the multiple regression model described by DeFries and Fulker (1985, 1988) provides a versatile and powerful test for linkage. If a QTL close to a chromosomal marker influences the selected trait, the scores of osibs of selected probands should regress differentially toward the population mean as a function of the number of marker alleles shared identical by descent (IBD; i.e., the same allele inherited from the same parent) with the proband. To adapt the basic DF multiple regression model for linkage analysis, the coefficient of relationship in Equation 1 is replaced by the estimated proportion of alleles shared IBD (π) by the pair of siblings.

The regression model for univariate QTL analysis is

\[ C = B_1 P + B_2 \pi + K, \]

where \( C \) is the expected cosib score, \( P \) is the proband score, and \( \pi \) is the
estimated proportion of alleles shared IBD at the marker. After appropriate adjustment of the standard errors of the regression coefficients to correct for the double entry of concordant sibling pairs (Gayán et al., 1999), the significance of the $B_2$ parameter provides a statistical test of the extent to which extreme scores are influenced by a QTL linked to the marker. Cardon and Fulker (1994) then developed this approach further to map the interval between markers. By analyzing multiple genetic markers simultaneously, this multipoint interval mapping procedure estimates the proportion of alleles shared IBD between siblings at points on the chromosome between available DNA markers. On the basis of these estimates, a continuous line is obtained indicating the significance of linkage at all points along a chromosomal region of interest.

**Linkage Studies of Reading Disability**

Linkage studies have identified several chromosomal regions that may contain genes that increase susceptibility to reading disability. There have been single reports of linkage to genes on chromosome 1 (Grigorenko et al., 1998), 2 (Fagerheim et al., 1999), 6q (Petryshen, Kaplan, & Field, 1999), and 18 (Fisher et al., 2002); QTLs on chromosome 6p21.3 and 15q21 have now been found in several independent samples.

An initial study of nine nuclear families suggested that reading disability may be associated with a gene near the centromere of chromosome 15 (Smith, Kimberling, Pennington, & Lubs, 1983), although this result was weakened when the sample was expanded to 21 families (Smith, Kimberling, & Pennington, 1991). Subsequently, several studies have reported significant linkage to markers in a different region of the chromosome (15q21) for reading disability (Fulker et al., 1991; Grigorenko et al., 1997) and spelling disability (Nöthen et al., 1999). Moreover, using a family-based association approach, Morris et al. (1999) reported a significant association between reading disability and a three-marker haplotype within the linkage region identified in the previous studies.

Smith et al. (1991) first reported evidence suggesting that a gene on chromosome 6 might influence reading deficits. Subsequently, Cardon et al. (1994, 1995) used the interval mapping technique described previously to localize a QTL for reading disability to chromosome 6p21.3. Although one subsequent study did not find evidence of a QTL in this region (Petryshen, Kaplan, Liu, & Field, 2000), this localization has now been confirmed in three independent samples (Fisher et al., 1999; Gayán et al., 1999; Grigorenko et al., 1997). Moreover, these subsequent studies revealed significant linkage for both overall reading ability and several specific reading and language skills that influence reading ability. Thus, this QTL represents one of the most consistently replicated linkages in genetic studies of complex traits with respect to both phenotype definition and chromosomal refinement (e.g., Flint, 1999). The fact that the evidence for this QTL has replicated more consistently than linkages for other complex traits such as schizophrenia (see review by Riley & McGuffin, 2000) or
bipolar disorder (Craddock & Jones, 2001) may be attributable to its relatively large effect (Knopik et al., in press) or to the power gained through analyses of continuous phenotypic scores versus the categorical phenotype used in linkage studies of psychiatric illnesses.

In summary, substantial progress has been made in understanding the etiology of reading disability since the publication of the previous volume (Plomin & McClearn, 1993). Twin studies of larger samples have provided conclusive evidence that reading disability is heritable, and linkage studies have localized putative QTLs for reading disability to regions on chromosomes 1, 2, 6, 15, and 18. In the next section, we turn to studies that have applied these methods to understand the relation between reading disability and other disorders such as ADHD.

**Etiology of Comorbidity Between Reading Disability and ADHD**

The genetics of ADHD is reviewed in more detail elsewhere in this volume (see Thapar, chapter 22). Here we focus on the finding that reading disability and ADHD co-occur more frequently than expected by chance; 25%–40% of children with either reading disability or ADHD also meet criteria for the other disorder (e.g., August & Garfinkel, 1990; Semrud-Clikeman et al., 1992; Willcutt, Chhabildas, & Pennington, 2001; Willcutt & Pennington, 2000). However, the etiology of this association is not well understood. Reading disability and ADHD are significantly comorbid in both clinical and community samples, indicating that it is not a selection artifact. Similarly, because reading disability is assessed by cognitive tests, whereas ADHD is usually assessed only by behavioral ratings, the relation between reading disability and ADHD does not appear to be due to shared method variance (e.g., Willcutt, Pennington, Boada, et al., 2001). In the following two sections, we describe a series of analyses that we have conducted over the past several years to test if comorbidity of reading disability and ADHD is attributable to common genetic influences.

**Applying the DF Model to Bivariate Data**

A simple generalization allows the univariate multiple regression models described in Equations 1 and 2 to be applied to bivariate twin data. Rather than comparing the relative similarity of MZ and DZ twins for the same trait, bivariate analyses compare the relation between the proband’s score on one trait and the cotwin’s score on a second trait across zygodity. Therefore, if common genetic influences contribute to the association between reading disability and ADHD, the ADHD score of the cotwins of MZ probands with reading disability would be expected to regress less toward the population mean than the ADHD score of DZ cotwins.

The regression equation to apply the DF method to bivariate data is expressed as follows:
where \( C_{\text{ADHD}} \) is the expected cotwin score on the nonselected measure (ADHD), \( P_{\text{RD}} \) is the proband score on the selected measure (reading disability), \( R \) is the coefficient of the relationship, and \( K \) is the regression constant. The \( B_2 \) coefficient provides a direct estimate of the bivariate heritability of reading disability and ADHD (\( h_{\text{RD/ADHD}}^2 \)), the extent to which the proband reading deficit is attributable to genetic influences that are also associated with elevations of ADHD.

Our most recent article used a sample of 313 same-sex twin pairs (183 MZ, 130 DZ) from the CLDRC twin study in which at least one twin met criteria for reading disability (Willcutt, Pennington, & DeFries, 2000b). The DSM–III (Diagnostic and Statistical Manual of Mental Disorders, 3rd ed.; American Psychiatric Association, 1980) version of the Diagnostic Interview for Children and Adolescents was used to assess symptoms of ADHD because it had been completed by the largest proportion of participants, and inattention and hyperactivity–impulsivity factor scores were computed on the basis of results of a factor analysis conducted previously (Willcutt & Pennington, 2000). The overall ADHD composite and the inattention and hyperactivity–impulsivity factor scores were all highly heritable in this sample (\( h_r^2 = .78-.94 \); Willcutt, Pennington, & DeFries, 2000a).

Phenotypic analyses indicated that individuals with reading disability exhibited significantly more symptoms of both inattention and hyperactivity–impulsivity than individuals without reading disability but that this effect was significantly larger for inattention scores. When Equation 4 was fitted to bivariate twin data, results revealed significant bivariate heritability for reading disability and the overall ADHD symptom count (\( h_{\text{RD/ADHD}}^2 = .23, p < .05 \)). However, additional analyses indicated that the etiology of the relation between reading disability and ADHD varied as a function of the ADHD symptom dimension. Specifically, the bivariate heritability of reading disability and inattention symptoms was significant (\( h_{\text{RD/ADHD}}^2 = .39, p < .01 \)), whereas reading disability and hyperactivity–impulsivity symptoms were not significantly coheritable (\( h_{\text{RD/ADHD}}^2 = .05 \)).

Interpretation of our study and other previous studies of the relation between reading disability and ADHD is complicated by recent changes in the diagnostic criteria for ADHD. Because our initial results were based on a measure of DSM–III attention deficit disorder (American Psychiatric Association, 1980), additional analyses were conducted for this chapter to test if these results replicated when a measure of DSM–IV ADHD (American Psychiatric Association, 1994) was used (see Table 13.2). Results revealed significant bivariate heritability of reading disability and parent ratings of total DSM–IV ADHD symptoms (\( h_{\text{RD/ADHD}}^2 = .42 \)). The estimate of bivariate heritability was similar for teacher ratings of ADHD (\( h_{\text{RD/ADHD}}^2 = .37 \)), but this result was not significant because teacher ratings were only available for a smaller subset of the sample.

When symptoms of inattention and hyperactivity–impulsivity were analyzed separately, estimates of bivariate heritability were again higher
Table 13.2. Bivariate Heritability of RD and DSM-IV ADHD in Twin Pairs Selected for RD

<table>
<thead>
<tr>
<th>Cotwin measure</th>
<th>MZ cotwins</th>
<th>DZ cotwins</th>
<th>Bivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Parent ratings(^a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total ADHD</td>
<td>1.02</td>
<td>1.63</td>
<td>0.49</td>
</tr>
<tr>
<td>Hyperactivity–impulsivity</td>
<td>0.59</td>
<td>1.45</td>
<td>0.25</td>
</tr>
<tr>
<td>Inattention</td>
<td>1.19</td>
<td>1.54</td>
<td>0.57</td>
</tr>
<tr>
<td>Teacher ratings(^b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total ADHD</td>
<td>1.58</td>
<td>1.96</td>
<td>1.10</td>
</tr>
<tr>
<td>Hyperactivity–impulsivity</td>
<td>0.96</td>
<td>2.16</td>
<td>0.69</td>
</tr>
<tr>
<td>Inattention</td>
<td>2.03</td>
<td>1.60</td>
<td>1.49</td>
</tr>
</tbody>
</table>

Note. RD = reading disability; DSM-IV = Diagnostic and Statistical Manual of Mental Disorders (4th ed.; American Psychiatric Association, 1994); ADHD = attention deficit hyperactivity disorder.

\(^a\)Expressed as standardized deviations from the mean of 367 control twins. N = 66 MZ pairs, 69 DZ pairs. \(^b\)Expressed as standardized deviations from the mean of 202 control twins. N = 34 MZ pairs, 36 DZ pairs. Degrees of freedom are 133 for parent ratings and 67 for teacher ratings.

*p < .05. †p < .10.

for reading disability and inattention \(h^2_{RD\,Inatt} = .42 \text{ and } .49\) than reading disability and hyperactivity–impulsivity \(h^2_{RD\,Hyp} = .20 \text{ and } .25\), although those differences were not significant. Moreover, a similar pattern of results was obtained when probands were selected for DSM-IV ADHD \(h^2_{Inatt\,RD} = .36, h^2_{Hyp\,RD} = .14\), although the sample of probands with ADHD is still relatively small.

In summary, our results suggest that common genetic influences contribute to comorbidity of reading disability and ADHD and that these common genes may influence symptoms of inattention more strongly than symptoms of hyperactivity–impulsivity. In the next section, we describe results of linkage analyses that we conducted recently to attempt to localize a QTL that influences both disorders.

**Bivariate Linkage Analyses**

The influence of genes on behavior is likely to be pleiotropic, such that the same genes affect more than one phenotype (e.g., Falconer & MacKay, 1996). For example, the gene for albinism in mice is also associated with significantly higher levels of emotionality (e.g., DeFries, Hegmann, & Weir, 1966; Turri, Datta, DeFries, Henderson, & Flint, 2001). However, genes that influence both reading disability and ADHD have not been previously localized. On the basis of the finding that common genetic influences contribute to comorbidity of reading disability and ADHD, we recently tested whether the well-replicated QTL for reading disability on chromosome 6p is also associated with increased susceptibility to ADHD (Willcutt et al., 2002).

The univariate DF model for linkage analysis can be applied to bivar-
iate data to test whether a QTL has pleiotropic effects on two traits. Rather than comparing the relative similarity of siblings for the same trait, bivariate analyses test if the relation between the proband score on the selected trait and the cosib score on a second, unselected trait varies as a function of $\hat{\pi}$. Therefore, if the QTL for reading disability on chromosome 6p is also a susceptibility locus for ADHD, the number of symptoms of ADHD exhibited by cosibs of reading disability probands would be expected to increase with allele sharing at DNA markers close to the susceptibility locus.

The regression model for the bivariate case is

$$C_{ADHD} = B_1P_{RD} + B_2\hat{\pi} + K,$$

where $C_{ADHD}$ is the predicted cosib score on the nonselected measure (ADHD) and $B_2$ tests for pleiotropic effects of the QTL on the two phenotypes.

Eight informative DNA markers spanning a region of 14.7 centimorgans (cM) on the short arm of chromosome 6 were used for these analyses (see Gayán et al., 1999). These markers span the region reported in previous 6p linkage studies of reading disability (Cardon et al., 1994, 1995; Fisher et al., 1999; Gayán et al., 1999; Grigorenko et al., 1997; Petryshen et al., 2000). Siblings and both parents from each family were genotyped for these markers from blood/cheek samples using published methods (Idury & Cardon, 1997). MAPMAKER/SIBS, a statistical package for genetic linkage analysis (Kruglyak & Lander, 1995), was used to estimate $\hat{\pi}$ or single-marker analyses. For multipoint analyses, MAPMAKER/SIBS uses available information at all markers to estimate the proportion of alleles-shared IBD at positions between markers ($\hat{\pi}_s$). Multipoint tests for linkage were conducted at 0.5 cM intervals across the region of interest.

Initial univariate analyses were conducted to test if the QTL on chromosome 6p21.3 is also a susceptibility locus for ADHD. Multipoint analyses revealed significant linkage for the overall ADHD composite and for both inattention and hyperactivity symptoms (see Figure 13.2). The area of peak linkage fell between markers D6S276 and D6S105 for all phenotypes. The maximum log of the odds (LOD) score for the ADHD measures (LOD = 1.8) was somewhat lower than the LOD scores obtained for the reading phenotypes (LOD = 1.9–3.2; Gayán et al., 1999), suggesting that this QTL may have a stronger effect on reading than symptoms of ADHD.

To test for pleiotropic effects of the QTL on reading disability and ADHD, we conducted bivariate linkage analyses for each reading measure based on the regression model in Equation 5. For each analysis, probands were selected for a score below the 1.5 SD cutoff on the relevant reading measure, and the cosib's ADHD score was regressed onto the proband's reading score and the proportion of alleles-shared IBD. Multipoint analyses revealed significant bivariate linkage for ADHD and each of the three reading phenotypes (Figures 13.3A–C), with the strongest results for the orthographic choice measure. Moreover, in contrast to the results of the twin analyses, the evidence for bivariate linkage with reading was some-
Figure 13.2. Results of multipoint linkage analyses of overall attention deficit hyperactivity disorder (ADHD) composite, inattention symptoms, and hyperactivity symptoms to eight DNA markers on chromosome 6p (N = 46-50 pairs). P values indicate one-tailed significance levels.

what higher for symptoms of hyperactivity, although the general pattern of results was similar for all three ADHD phenotypes. Because of power considerations (Fulker et al., 1991), however, much larger samples would be required to test for differential bivariate QTL heritability.

Future Directions

Our results indicate that both reading disability and ADHD are significantly heritable and that the frequent comorbidity of reading disability and ADHD is attributable substantially to common genetic influences. Linkage analyses indicate that the QTL for reading disability on chromosome 6p is also a susceptibility locus for ADHD and suggest that comorbidity between reading disability and ADHD may be due at least in part to pleiotropic effects of this QTL. In the remainder of this chapter, we discuss future directions for research on the etiology of reading disability and ADHD and their comorbidity.

Localization of QTLs

Linkage analysis is an essential first step toward the localization of susceptibility loci for a complex trait. However, the region of significant linkage for a QTL is frequently large, often including hundreds of genes (e.g., Morris et al., 1999). In contrast to family-based linkage analysis, associ-
ation and linkage disequilibrium occur when alleles at a marker locus cosegregate with alleles at a susceptibility locus more frequently than expected by chance in an entire population (e.g., Cardon & Bell, 2001). Because linkage disequilibrium extends a relatively short distance in most regions of the genome, a DNA marker must be close to a QTL that influences a trait to detect significant association. Therefore, after linkage analysis has been used to identify chromosomal regions that are likely to contain a QTL for a trait, association methods can be used to refine the estimated chromosomal location of the QTL.

Evidence for linkage disequilibrium can be further enhanced by testing if a disorder is associated with a haplotype, a specific combination of alleles at a set of closely spaced markers. As noted previously, Morris et al. (1999) tested for an association between reading disability and eight markers on chromosome 15q, the site of a putative QTL for reading disability (e.g., Grigorenko et al., 1997). Their results revealed a significant association between reading disability and a three-marker haplotype in the region of significant linkage identified in previous studies, providing additional evidence that a QTL in this region increases susceptibility to reading difficulties. Moreover, because linkage disequilibrium rarely extends more than 0.5 cM from a marker (e.g., Abecasis et al., 2001), this result identifies a chromosomal region that is narrower than the 5–8 cM region identified in linkage studies (Grigorenko et al., 1997; Schulte-Korne et al., 1998).

As increasing numbers of DNA markers become available and genotyping techniques continue to become more efficient and less expensive, association methods will provide a powerful method to identify the location of QTLs for complex traits. We currently are genotyping additional markers in our sample across the region of significant linkage on chromosome 6p and flanking regions of the chromosome. Moreover, although no obvious candidate genes for reading disability or ADHD have been identified in this region to date, future molecular genetic studies will determine the function of additional genes in this region. On the basis of these findings, it is likely that plausible candidate genes will be identified that are expressed in the brain or influence other developmental processes related to reading disability or ADHD.

**Multivariate QTL Analyses**

In addition to the QTL on chromosome 6, possible QTLs for reading disability have been localized to regions on chromosomes 1 (e.g., Rabin et al.,

---

**Figure 13.3.** Results of bivariate multipoint linkage analyses of reading deficits and the overall attention deficit hyperactivity disorder (ADHD) composite, inattention symptoms, and hyperactivity symptoms. A: Proband selected for deficits in orthographic choice (N = 50 pairs); B: Probands selected for deficits in phonological decoding (N = 58 pairs); C: Probands selected for deficits on the reading composite score (N = 69 pairs). P values indicate one-tailed significance levels.
1993), 2 (Fagerheim et al., 1999), and 15 (e.g., Fulker et al., 1991; Grigorenko et al., 1997; Smith et al., 1983), and additional QTLs are likely to be identified in genomewide scans that are currently under way. Moreover, it is plausible that epistatic interactions among subsets of these genes may contribute to susceptibility to reading difficulties (e.g., Grigorenko, chapter 14, this volume). Similarly, candidate gene studies have demonstrated a significant association between ADHD and polymorphisms in at least five genes in the dopamine system (e.g., Barr et al., 2000; Cook et al., 1995; Daly, Hawi, Fitzgerald, & Gill, 1999; Eisenberg et al., 1999; Faraone et al., 1999), although these results have not been replicated in all samples (e.g., Asherson et al., 1998; Castellanos et al., 1998). These results underscore the complexity of the genetic etiology of reading disability and ADHD and suggest that additional genes are likely to contribute to comorbidity of these disorders. In future studies, it will be useful to obtain genotypes for markers in these chromosomal regions and for each of these candidate genes. These genotypes may then be included in multivariate analyses to test which genes contribute independently to reading disability, ADHD, or their comorbidity and whether epistatic interactions among any of the loci play a significant role in the etiology of these disorders or their overlap.

Conclusion

Our findings regarding reading disability and ADHD suggest that the boundaries between putatively distinct diagnoses may prove to be blurry, with the same genetic influences conferring risk for more than one disorder. In some cases, such findings may indicate that two disorders may be better conceptualized as alternate forms of the same disorder, whereas in other cases common risk factors may contribute to two or more distinct disorders. In either case, etiologically informative methods such as those described in this chapter provide an important tool that can be used in future studies to improve the validity of the diagnostic nosology of psychiatric disorders by revealing the etiology of comorbidity between complex syndromes.

References


Willcutt, E. G., & Pennington, B. F. (2000). Comorbidity of reading disability and attention-


