Review

Arguments for the genetic basis of the bipolar spectrum

John R. Kelsoe

Abstract

Family members of bipolar probands have been repeatedly shown to have an increased risk for mood disorders. However, a range of different syndromes in the bipolar spectrum are commonly observed in these relatives. This suggests the hypothesis that these different syndromes may be genetically related. It further suggests that bipolar disorder may be better conceptualized from a genetic standpoint as a quantitative trait. In such a model, the same genes may predispose to a variety of phenotypes ranging from schizoaffective disorder to cyclothymic temperament. Previous attempts to test such a multifactorial model have provided some limited support. However, other studies argue that some forms of bipolar disorder such as bipolar II may be genetically distinct. In this review, various quantitative and categorical models of illness are considered and the data supporting them reviewed. It is proposed that the existing data may best fit a model in which different sets of genes predispose to overlapping phenotypes that are in part both quantitative and distinct in nature.

Keywords: Bipolar disorder; Bipolar spectrum; Hypomania; Genetics

1. Introduction

The familial nature of bipolar disorder is apparent to the observant clinician in that most patients have family histories of mood disorders. Such observation suggests several features of this familial phenomenon. (1) Mood disorders occur at a higher rate in the families of bipolar patients, and (2) it is not only bipolar disorder that occurs in these families but a variety of mood syndromes and symptoms that differ both qualitatively and quantitatively from bipolar disorder itself (Gershon et al., 1982). These observations of the variety of mood symptoms in bipolar families have played a prominent role in inspiring theories about the relatedness of these “spectrum” phenomena (Akiskal, 1983) and the possible familial genetic basis for the spectrum (Akiskal, 2002). In fact, such spectrum phenotypes are common in genetics and suggest specific models of the nature of the genetic transmission of bipolar disorder. We are now on the verge of dramatic advances in the genetics of bipolar disorder. This specific knowledge of susceptibility genes will allow empirical testing of these ideas regarding the bipolar spectrum, as well
as, the genetic relationships among different forms of illness. In this review, the data regarding the familial and genetic nature of bipolar disorder and its spectrum phenotypes will be reviewed. Genetic models consistent with a spectrum or quantitative phenotype will be considered and compared to the existing data. Lastly, the scientific and clinical implications of identifying the specific genes involved in the bipolar spectrum will be considered.

2. Family epidemiology of bipolar disorder

Numerous family studies of bipolar disorder have been conducted over the past several decades. These studies consist of systematically identifying bipolar probands and determining what portion of first degree relatives are also ill. These studies, summarized in Table 1, include a wide range of samples and methodologies, and generally employ a narrow syndromic definition of bipolar disorder. The actual population prevalence of bipolar disorder is a topic of some debate and is discussed in more detail below. Given these qualifications, these data together argue that approximately 7% of the first degree relatives of bipolar disorder also have bipolar disorder. Given a population prevalence of 1% for the core bipolar phenotype, this indicates a sevenfold increase in risk. These data also indicate an approximate rate for unipolar disorder of 10%. Compared to a population prevalence of approximately 5%, this suggests a twofold increase in risk for this mood disorder and a possible genetic and etiological relationship between these two forms of mood disorder.

However, such family data do not prove that bipolar disorder is genetic, merely that it is familial. Families share both environment and genes, and many forms of non-genetic transmission can be imagined ranging from behavioral or learned experiences, to shared exposures to infectious agents or toxins. It is twin studies that have primarily been used to distinguish nature and nurture. Selected twin studies of bipolar disorder are summarized in Table 2. In these studies, monozygotic (MZ) and same sex dizygotic (DZ) twin pairs are identified in which one twin has bipolar disorder. These studies have found that approximately 70% of monozygotic pairs are concordant for illness, while only about 30% of dizygotic twin pairs are concordant. As both types of twins in these studies are raised together and environment is shared equally, these data argue that sharing all genes as opposed to half of genes increases the risk for illness over twofold.

Adoption studies represent another approach to separating nature and nurture. Typically in such studies probands are ascertained with bipolar disorder and the rate of illness in their biological and adoptive parents is compared. These studies are more

Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Relatives at risk</th>
<th>Morbid risk (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bipolar</td>
<td>Unipolar</td>
</tr>
<tr>
<td>Fieve et al. (1984)</td>
<td>2171</td>
<td>6.6</td>
</tr>
<tr>
<td>Gershon et al. (1982)</td>
<td>598</td>
<td>8.0</td>
</tr>
<tr>
<td>Rice et al. (1987)</td>
<td>557</td>
<td>5.7</td>
</tr>
<tr>
<td>Sadovnick et al. (1994)</td>
<td>1102</td>
<td>3.5</td>
</tr>
</tbody>
</table>

* Observed rates rather than morbid risk.

Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Monozygotic twins</th>
<th>Dizygotic twins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. twin pairs</td>
<td>Concordance (%)</td>
</tr>
<tr>
<td>Rosanoff et al. (1935)</td>
<td>23</td>
<td>69.6</td>
</tr>
<tr>
<td>Kallman (1954)</td>
<td>27</td>
<td>92.6</td>
</tr>
<tr>
<td>Bertelsen (1979)</td>
<td>55</td>
<td>58.3</td>
</tr>
<tr>
<td>Kendler et al. (1993)</td>
<td>154</td>
<td>69.7</td>
</tr>
<tr>
<td>Total</td>
<td>259</td>
<td>69.6*</td>
</tr>
</tbody>
</table>

* Weighted means.
difficult to conduct, and are fewer in number than the
twin studies. However, the existing adoption studies
are in general supportive of the twin studies, and
have found an elevated rate of illness in the bio-
logical parents, but only a population rate in the
adoptive parents (Mendlewicz and Rainer, 1977;
Wender et al., 1986).

3. Family epidemiology of the bipolar spectrum

Some of the first and most compelling data for the
genetic relationship of bipolar spectrum disorders to
the core phenotype come from these early twin and
family studies. Bertlesen et al. found not only strictly
defined bipolar disorder, but also a variety of other
psychiatric disorders and traits in the co-twins of
bipolar identical twins (Bertelsen et al., 1977; Ber-
telsen, 1979). Among the co-twins of 69 MZ bipolar
probands, they found 46 with bipolar disorder, but
also another 14 with other psychoses, affective
personality disorders or suicide. Using a narrow
“bipolar only” definition of illness yielded a pair-
wise direct concordance rate of 58%, however, using
a broader definition of illnesses raised this con-
cordance rate to 84%. Corresponding concordance
rates for the matched same sex DZ twins were 17
and 35%, respectively. MZ and DZ rates were
significantly different for both definitions of illness.
Similar results come from the twin studies of Kal-
lman (1954) and Da Fonseca (1963), who both
reported the occurrence of affective personality in
the identical co-twins of bipolar probands, and a
higher concordance rate for the broad definition of
phenotype. These and similar data argue that these
related traits are alternative phenotypes resulting
from the same genetic diathesis that produces bipolar
disorder in the MZ proband.

Several conclusions can be drawn regarding the
genetics of bipolar disorder and related phenotypes
from these early data. First, genes explain only a
portion of the etiology. Environmental factors likely
also play a substantial role. In genetic terms, the
penetrance, or probability of manifesting illness
given that one has the disease gene, is less than
100%. Secondly, the trait displays variable expres-
sivity. That is, a variety of related clinical pre-
sentations can result from the same genes. Together,
these characteristics indicate a lack of one-to-one
 correspondence between genotype and phenotype
 that is characteristic of complex, non-Mendelian
 disorders. Consistent with this and even more com-
plicating is the likely presence of genetic hetero-
genicity, or multiple genes for the disorder. These
features are characteristic of a variety of other non-
Mendelian disorders such as hypertension, Alzheim-
er’s disease, or cancer. The non-Mendelian nature
of bipolar disorder is also supported by several segrega-
tion analyses in which the patterns of transmission
are analyzed in families of bipolar probands (Rice et
al., 1987; Spence et al., 1995). These studies fail to
find a strict Mendelian form of transmission, though
they do find some support for a modest dominant
major locus effect. As discussed in more detail
below, a major locus describes a gene that conveys a
strong susceptibility to illness.

It should not be surprising that the genetics of a
disorder as common and multi-faceted as bipolar
disorder is complex. However, that complexity has
both qualitative and quantitative aspects which make
it difficult to model and suggest several layers of
complexity. The term bipolar spectrum derives from
the seeming gradation in symptom severity and
impairment from more to less as follows: bipolar
I—bipolar II—cyclothymia—affective personalities
(Akiskal, 1983). It is this last category, affective
personalities or temperaments (Akiskal et al., 1979,
1998), that may be controversial from a diagnostic
standpoint as it implies a genetic connection to
normal personality variation. Though, if bipolar
disorder is genetically transmitted as a quantitative
trait as is suggested by the spectrum concept and the
twin and family data, then such an idea is a logical
result of that genetic model. However, the problem
with this simple quantitative spectrum model is that
it does not include a number of diagnostic and
syndromic entities that also occur in the families and
co-twins of bipolar patients. Specifically, this in-
cludes unipolar depression, dysthymia, as well as,
psychotic mood disorder, schizoaffective disorder
and possibly schizophrenia. One can construct a
spectrum to include these syndromes such as: schizo-
phrenia—schizoaffective disorder—bipolar I—bipo-
lar II—recurrent unipolar depression—single episode
unipolar depression—cyclothymia—dysthymia—affect-
ive personalities. Such an attempt to be all
inclusive with the variety of clinical syndromes seen in the families of bipolar probands becomes increasingly unsatisfactory. (The inclusion of schizophrenia is admittedly even more controversial, and the evidence for and against it will be dealt with below). One has to ask exactly what is the quantity that is gradated across such a range of presentations. It is not “bipolar-ness” as not all include mania-like states. Similarly, it is not severity, as some forms of recurrent unipolar disorder are more impairing and cause greater morbidity than some forms of bipolar disorder.

Here lies the chief difficulty in defining the bipolar spectrum phenotype. It is neither strictly quantitative or qualitative, but a combination of the two. Rather than discrete entities or a one dimensional spectrum, the actual phenotype seen in families is probably best described as a surface in a multidimensional space that allows for a quantitative relatedness among qualitatively different features of illness. From the standpoint of diagnostic or genetic models, this is quite complex and unwieldy. However, from the standpoint of the likely underlying biology and the relationship of multiple highly interconnected neural systems, such complexity in mood related brain “failure modes” seems highly probable. This difficulty in defining the most accurate or optimal phenotype not only makes clinical diagnosis complex but also is problematic for the molecular genetic mapping of disease genes.

4. A brief review of genetic models

A brief general review of genetic models may be useful at this point in considering the possible ways in which different gene effects might result in such a complex pattern of related phenotypes. Historically, the route from gene to phenotype has been conceptualized in two general ways. The first school of thought derives from the famous experiments of Gregor Mendel. In these models, phenotypes result from the action of single genes or major loci. Such major loci may be transmitted in a variety of ways including dominant, in which only one copy of the trait allele is necessary to cause the trait, or recessive, in which both chromosomes must carry the trait allele. X-Linked transmission occurs when the gene is located on the X chromosome, in this case the expression of the trait is gender dependent. Several variations on these simple principles can add levels of complexity. As mentioned above, reduced penetrance or variable expressivity can modify the effect of a major locus. Another complication may be if the trait sometimes occurs solely as the result of environmental causes, such cases are called phenocopies. Lastly, though in each individual and family, the trait results from a single gene, different genes may operate in different families. In this model genetic heterogeneity occur as the result of multiple major loci. Though these variations are not strictly Mendelian, they nonetheless, all fall within the general rubric of single major locus transmission as the trait results from one gene within each individual.

The historical alternative first proposed by Galton was quantitative or polygenic transmission. In this form of genetic transmission, illustrated in Fig. 1, the trait is a continuous variable whose value results from the combined effect of many genes. The more of these polygenes that are present, the numerically greater the value of the quantitative trait. In its purest form, quantitative traits derive from numerous genes each of small effect. The trait will display a normal distribution around the population mean which reflects the average number of polygene alleles carried in the population. In an additive quantitative model, the effect of the polygenes add together to produce a cumulative quantitative effect. Alternatively, the genes may combine in a greater than additive fashion termed epistasis. For example, gene A alone may lead to a 10% elevation in the trait value, and gene B alone may also lead to a 10% elevation. However, if one inherits the susceptibility alleles for both gene A and gene B, the trait value may be elevated not 20% but rather 40% above the mean. This makes great biological sense in considering systems of interacting genes and proteins. Each gene may be able to homeostatically compensate for a disturbance in the function of the other gene, however, when both genes are affected, the systems capacity to maintain homeostasis is gone and the effect is greatly amplified.

The quantitative trait model has also been extended to include qualitatively different phenotypes lying on a spectrum (Reich et al., 1972). In this multiple
threshold model, a latent “invisible” quantitative trait is hypothesized. As illustrated in Fig. 1, qualitatively different traits result when the value of this latent quantitative trait is above different critical threshold values. This requires the assumption that these different phenotypes are quantitatively ranked in relationship to each other.

These two models actually likely exist on a spectrum of their own. The key underlying variable in this genetic spectrum is the effect size of the gene or genes under consideration. Though there may be traits in which one gene explains all the variance, and other traits that are highly polygenic, it is more likely that for many traits, there is a mixture of genes and alleles of different effect sizes. It also seems increasingly likely that epistasis explains a substantial portion of the variance for many complex genetic traits. If this is the case, then it will be difficult to identify susceptibility genes for complex traits by examining only one gene at a time. Rather, progress may come from examining multiple genes simultaneously and considering the interaction effects.

Numerous studies have examined the relevance of each of these two basic genetic models to the bipolar spectrum. What conclusions can be drawn from these data about the underlying genetic structure of bipolar disorder traits?

5. The bipolar spectrum as a quantitative phenotype

Several basic approaches have been taken to testing the hypothesis that bipolar disorder is a quantitative genetic trait. The first is to attempt to order the different affective diagnoses and examine families and patterns of transmission. As outlined above, this ranking in terms of severity has typically been from schizoaffective disorder—bipolar I—bipolar II—unipolar depression recurrent—unipolar depression single episode—dysthymia/cyclothymia. This model has several specific testable predictions.

First of all, within the population, the prevalence of these entities should roughly mirror that ranking. A number of studies provide data to test this hypothesis. In particular, two large recent epidemiological surveys have confirmed the longstanding estimate of approximately 1% for the lifetime prevalence of the core bipolar I phenotype. The Epidemiological Catchment Area study (ECA) found a rate of 0.8% (Weissman et al., 1990), while the National Comorbidity Study reported 1.6% (Kessler et al., 1994). However, there is more disagreement about the lifetime prevalence of hypomania and other forms of the bipolar spectrum. Nonetheless, higher lifetime prevalences of bipolar II and hypomania ranging up to 5% (Szadoczky et al., 1998; Akiskal et
al., 2000), and rates for subsyndromal manic symptoms as high as 11% (Angst, 1998) have been reported. This variation may be related to differences in populations and ascertainment, but at least in part is related to variation in severity as defined by the number of symptoms and episode duration required in these different studies. This later source of variation in prevalence estimates supports the quantitative genetic hypothesis. Further support comes from family studies that indicate that in families of bipolar I probands, bipolar II is the most common form of illness (Simpson et al., 1993); and from studies of the general population reporting a 6% prevalence of cyclothymia (Placidi et al., 1998). These data, therefore at least in part, are consistent with the population prevalence predictions of a quantitative genetic trait model.

Another prediction is that the degree of genetic loading or family history will be stronger for probands at the more severe end of this spectrum than for probands with less severe illness. Support for this has come from some (Gershon et al., 1982), but not other studies (Rice et al., 1987; Andreasen et al., 1986). The most consistent finding of the numerous family studies that have been conducted is increased rates of depression in the family members of probands of a variety of diagnoses. This is only partly consistent with the model. Another approach taken in many family studies is to examine the specific patterns of transmission within families and to use segregation analysis to test for fit of these data to single major locus versus polygenic models. As mentioned above, results from such studies have been largely inconclusive arguing for the presence of some dominant single major locus genes operating on a background of polygenes (Rice et al., 1987; Spence et al., 1995). Overall, these approaches have yielded only mixed support for the quantitative multiple threshold model.

Yet another approach to the bipolar spectrum as a quantitative genetic trait is the examination of affective temperament in the family members of bipolar probands. If one assumes that the genes involved in mental illness also play a role in temperament within the normal range, then one would expect to see an increase in affective temperaments in the families of bipolar probands. This predication derives directly from quantitative genetic models. Alternatively, it has been argued that it is this quantitative difference in affective temperament that is primary, and that this variable level of predisposition in turn predisposes one to episodes of mania, hypomania or depression (Akiskal, 1995). Ultimately, distinguishing these possibilities may be difficult, but has emerging backing from psychometric validation in population studies (Akiskal et al., 1998). If the underlying biological systems for the mild and severe forms of phenotype are the same, then likely these are simply different degrees or types of pathologies within them. The premorbid or interepisode presence of affective temperaments in patients with mood disorders has been repeatedly demonstrated (Hirschfeld et al., 1989; Akiskal, 1995). Affective temperaments have been shown to predict course and switching in depression (Akiskal et al., 1995). Similarly, the relationship of personality disorders such as borderline to mood disorders has been the subject of much study (Akiskal, 1981). However, fewer studies have addressed the presence of affective temperaments in the families of patients with bipolar disorder (Lauer et al., 1997; Maier et al., 1995). Such studies have, however, supported the idea that relatives of bipolar probands show an increased rate of affective temperaments in comparison to control groups. These studies have also indicated not only increases in affective traits such as depressive cognition and neuroticism, but also less mood related traits such as rigidity and compulsivity.

We have recently presented data to empirically test this hypothesis (Evans and Kelsoe, 2001). Akiskal et al. have developed a self-rated version of their instrument to assess affective temperament (Akiskal et al., 1998) along the five dimensions of dysthymia, hyperthymia, cyclothymia, irritability and anxiety. As will be reported in greater detail separately (Evans et al., in preparation), we employed this instrument in the evaluation of 220 relatives of 86 bipolar probands, and 53 controls. For the dysthymia and cyclothymia scales, the highest values were obtained for the probands and lowest for controls. Values for affectively ill relatives were not different from probands for any of the scales. However, for the dysthymia and cyclothymia scales, the values for unaffected relatives were significantly elevated in comparison with controls. The results for these scales were consistent with the above cited twin data
and with a quantitative trait genetic model. Surprisingly, hyperthymia was not significantly different between groups. These results for hyperthymia were contrary to prediction and to results from an earlier observer rated version of the instrument. This might be, in part, an artifact of converting the instrument to a self-rated format.

6. The bipolar spectrum as overlapping phenotypes

Though the above data provide some support for a quantitative genetic model, the fit of the model is only partial. The alternative approach is that these diagnoses are discrete genetic entities. Many of the same family studies cited above have also addressed this question framed essentially as whether bipolar disorder “breeds true.” Again here, the answer is a mixed “yes” and “no”. Though bipolar disorder is found at substantially elevated rates among family members of bipolar probands, the rate of unipolar depression is also elevated. Studies of schizoaffective disorder have also yielded mixed results that suggest that the depressive type is more related to schizophrenia, while the bipolar type is more genetically related to bipolar disorder (Andreasen et al., 1987).

Similar results emerge from studies of the relationship between bipolar I and bipolar II. Coryell et al. examined a series of families of bipolar probands and compared rates of bipolar I and bipolar II illness in the family members of bipolar I vs. bipolar II probands (Coryell et al., 1984). In the first degree relatives of bipolar I probands bipolar I and bipolar II were of approximately similar prevalence (8.5 vs. 6.1%, respectively). However, among first degree relatives of bipolar II probands bipolar II was significantly more prevalent (3 vs. 30%, respectively). These data suggest that genes for bipolar I predispose equally to bipolar I and bipolar II. However, in addition to this set of genes, a separate and more common set of genes predisposes preferentially or exclusively to bipolar II illness.

At the more distant edges of the bipolar spectrum, a similar picture emerges. Many family studies have examined the genetic relationship between schizophrenia and bipolar disorder. Early studies argued that these two disorders were genetically separate. Bipolar disorder was found in the families of schizophrenia probands at no more than the population rate and vice versa (Tsuang et al., 1980). However, more recent studies question this result and suggest that the genetic diathesis may be less specific (Erlenmeyer-Kimling et al., 1997). Similar complexity emerges in considering “co-morbid” diagnoses. Evidence for a shared vulnerability with bipolar disorder has been reported for alcoholism, substance abuse, and panic disorder (Winokur et al., 1998; MacKinnon et al., 1998; Lauer et al., 1997; Maier et al., 1995). Together, these data suggest a model of partially overlapping phenotypes, and a mixture of genes some of which may be specific and some non-specific.

7. Further evidence for overlapping spectra of illness from molecular genetic studies

Further support for this idea has recently come from molecular genetic linkage studies of bipolar disorder and schizophrenia. The past 15 years has seen an accelerating number of studies and growing body of data attempting to employ the molecular genetic methods of positional cloning in order to identify the specific genes underlying affective disorders. Investigators in this field are challenged by the complexity of the phenotype, as the choice of phenotypic definition may be critically important to the odds of success. Ultimately, it is the success of this enterprise and the identification of genes that will lead to a fundamental understanding of the genetic relationship of these different phenotypes. However, studies to date have already suggested some unexpected results. The most surprising of these is the apparent overlap of regions of the genome implicated in bipolar disorder with those for schizophrenia (Kelsoe, 1999; Berrettini, 2000). We reported one of the first such overlaps on chromosome 22q where in a study of 20 bipolar families, the genome-wide maximum lod score for bipolar disorder was located at a marker previously implicated in several studies of schizophrenia (Kelsoe et al., 2001). Several other genomic regions also demonstrating such an overlap include: 18p, 13q, 10p, 10q, 6q, 5q. These data suggest the possibility that the
same gene in each of these regions predisposes to both bipolar disorder and schizophrenia. This will not be conclusively known until the genes are definitively identified. It is possible that each region contains separate genes for each disorder that are coincidentally near each other. However, if a substantial portion of the genes for these two disorders are common, it will raise the challenging question of how the same genes can lead to two different phenotypes. Several possible models will be discussed below. One result of these data may be an easier incorporation of schizoaffective disorder into a comprehensive genetic model. In addition to the overlap with schizophrenia, recently a similar overlap has been reported for both bipolar disorder and alcoholism to a locus on chromosome 1 (Nurnberger Jr. et al., 2001).

Several other molecular genetic observations support the role of genetics in personality variation relevant to affective temperaments. These data provide some support for the possible genetic relationship of these traits to illness. A repeat variant in the promoter of the serotonin transporter has been reported to be associated with neuroticism in normal volunteers (Lesch et al., 1996). This variant has also been associated with unipolar disorder, bipolar disorder, response to SSRI antidepressants and SSRI induced hypomania (Collier et al., 1996; Smeraldi et al., 1998; Mundo et al., 2001). Similarly, a repeat variant in the D4 dopamine receptor has been associated with novelty seeking in normal volunteers and with attention deficit hyperactivity disorder (Ebstein et al., 1996; LaHoste et al., 1996).

8. Models of bipolar genetics

Given the family, twin and molecular mapping data, and the possible genetic models reviewed above, what inferences can be drawn about the underlying genetic structure of affective disorders? Which genetic models can explain the observed data regarding the bipolar spectrum and overlapping phenotypes? To reframe the above discussion, these issues really break down into two basic questions. Why are milder versions of the phenotype seen in the relatives of probands with the more severe “core” phenotype? And, what is the relationship between qualitatively different phenotypes that seem to associate together in families?

Both the major locus model and the polygenic model provide possible explanations for the first question. Under a major locus model, variable expressivity and incomplete penetrance may explain such a phenomenon. Relatives of the proband may inherit the bipolar disorder susceptibility major locus, but required environmental factors may be absent such that the phenotype is only partially or incompletely manifest. Such a model could explain the quantitative spectrum of traits seen in families ranging from bipolar I to bipolar II to cyclothymic temperament. The polygenic model similarly can explain such transmission. In this case, the proband has the most polygenic susceptibility alleles, while the relatives have a fewer number of alleles resulting in a quantitatively reduced phenotype.

However, it is easy to demonstrate that these two models predict different patterns of transmission in families. A major locus retains its larger effect across many generations. The degree of penetrance might vary in different individuals depending on their environmental exposures, but one would still expect to see even quite distantly removed relatives still manifesting the full phenotype at the rate predicted by the penetrance probability. The polygenic model, however, predicts a more rapid drop off of the value of the quantitative trait as the polygenes are effectively diluted through subsequent breeding. These two models may in principle be distinguished in studies of the patterns of segregation in families. However, in practice the numerous unknown variables limit the power of such segregation analyses. As described above, many such studies exclude all models, however, a few have supported a modest major locus effect. These results lead to a third, more complicated and more likely scenario, a mixed model. The phenotype may derive from the action of a few major loci acting on a background of polygenes.

As if this were not complex enough, another complicating and likely factor, is the prominent role of epistasis or gene interaction. Here, as described above the interaction of specific combinations of genes may lead to a greater than additive increase in illness risk or severity. Genes of small effect might work together or interact with a major locus in order
to influence the phenotypic expression. Such a model is very difficult to distinguish by segregation analysis alone. Yet, linkage studies have already revealed some evidence for epistatic interactions in alcoholism and anxiety (Cloninger et al., 1998; Reich et al., 1998).

The second question posed above is no less complex in terms of the range of underlying genetic models that might result in the observed pattern of observations. How could the involved susceptibility genes result in the range of different phenotypes seen segregating together. As described above, some of the involved phenotypes are easily conceptualized as on a spectrum, however, for others it is difficult to place the clinical syndromes on a single spectrum. Further, several fairly robust observations are simply incompatible with a single continuum or the multiple threshold model described above. Specifically, as mentioned above such a model would predict that schizophrenia should be more rare than bipolar disorder, while they are actually seen at approximately equal prevalence. Similarly, the above cited data suggesting that bipolar II breeds true argues that it is not simply on a continuum with bipolar I, but may derive from separate genes.

Given the limited knowledge at this time, one can only speculate, but the observed pattern of overlapping but distinct phenotypes suggests a model of partially overlapping genes as illustrated in Fig. 2. In this model, some of the genes involved are specific for each of the phenotypes considered, bipolar I, bipolar II and unipolar disorder. Others are less specific and may predispose to either form of bipolar disorder or unipolar disorder. These genes in lesser number or in different environmental context may result in the softer spectrum phenotypes in the normal range such as temperaments. Furthermore, epistasis may play a role in the qualitative outcome of gene interactions as well as the quantitative mentioned above. Genes may result in qualitatively different phenotypes when they interact with different other genes.

Another layer of complexity is the likely role of allelic heterogeneity, or the presence of different mutations within the same susceptibility gene. In this way, the same gene could have different molecular defects that might impact the phenotypic outcome in either a quantitative or qualitative fashion. In this way, Fig. 2 is probably best considered not just in terms of genes, but all mutations in the involved genes that exist in the population. This possibility is most apparent in the molecular data that indicates possible linkage of bipolar disorder and schizophrenia to the same genomic regions. This could be explained by the presence of different mutations in the same genes which differentially predispose to different phenotypes. There are abundant examples of this sort of variability in human, rodent and invertebrate genetics.

As complex as these possibilities seem, we have so far primarily considered the role of genes alone. It is highly likely that genes or combinations of genes interact with environment in very specific ways. That is, a susceptibility gene may only cause illness in a specific environmental context. An example might be a gene that predisposes to a central nervous system (CNS) virus, might only cause psychiatric illness if the mother is exposed to the virus while pregnant. Fig. 3 illustrates just a few of the myriad ways in which genes could operate in combination with other genes or environment to produce the different phenotypes. The top of the figure illustrates a simple model of environmental interaction in which a single
9. Evolutionary speculations

Bipolar spectrum phenotypes also raise some intriguing questions about the population genetics of susceptibility genes for bipolar disorder and their possible role in evolution. These traits are clearly deleterious for those individuals most seriously affected. Yet, bipolar disorder seems to have existed in human populations for a long time. Descriptions of cycling mood appear throughout our history in literature and art, dating even to antiquity. Furthermore, the illness occurs around the world at a remarkably consistent prevalence. This suggests that these alleles have been present for a long time in order to become so evenly distributed around the world. What perpetuates the presence of bipolar susceptibility alleles in the human population over the millennia? There are many examples of genetic traits that are deleterious in one situation, but advantageous in another. Homozygotes for the sickle cell trait, suffer from sickle cell disease, a terrible illness with early mortality. However, for every one homozygote with this recessive disease, there are two heterozygotes who not only have minimal symptoms...
of the illness, but also have a resistance to malaria. It has been hypothesized in the cold calculus of evolution, that in regions in which malaria is endemic, these mutations to the hemoglobin gene are advantageous to the overall population despite those who are severely affected. Is it possible that a similar equation plays out in a more complex way for bipolar disorder? This question has been considered by several authors (Gardner Jr., 1982; Brody, 2001).

It is easy to imagine how in “low doses,” the symptoms of hypo-mania could work to one’s advantage. To have just a very slight amount of euphoria, increase in energy, faster thoughts and need less sleep could be advantageous. Unfortunately, patients with bipolar I or bipolar II rarely stay in these milder stages of mood acceleration for any length of time. For many, clinical depression or more extreme states of mania predominate with no advantage. Yet those fleeting periods of hypomania can be associated with genuine accomplishment. The association between creativity and bipolar disorder has been repeatedly documented. Rates of mood disorders have been shown to be elevated among creative individuals (Richards and Kinney, 1989; Jamison, 1989). Similarly, creativity has been shown to occur at an increased frequency among not only bipolar patients, but also their first degree relatives (Andreasen, 1987; Andreasen and Glick, 1988; George et al., 1988; Richards et al., 1988). These data suggest that not only those with bipolar disorder itself, but also their relatives who may have milder manifestations of the trait, may benefit from enhanced achievement. Hence, one can speculate that the same susceptibility alleles that cause illness in some, may be related to affective temperaments, such as hyperthymia or cyclothymia temperament, and in turn to creativity among some non-ill individuals.

Is there a selective advantage conveyed to those individuals? In evolutionary terms, selective advantage refers strictly to the ability to survive and procreate and thereby propagate one’s genes. The simplistic explanation is that the hypersexuality associated with bipolarity leads to more offspring. However, it is easy to imagine how other aspects of the trait such as creativity and energy might also enhance individual selective fitness. Yet, other more complex mechanisms also might play a role such as the value of individuals with the trait to the survival and expansion of the social group. It has been argued that individuals with bipolar traits may be more likely to serve as leaders within a social group (Gardner Jr., 1982). To speculate even further, as these traits profoundly affect social function, a diversity of such genes may contribute to a diversity of personality and social styles that in turn lead to more effective function of the social group. Such issues as group selection and evolutionary psychology have been argued at length in the literature of evolution and psychology. These are questions that are extremely difficult to test empirically, but the evidence for the quantitative nature of the bipolar spectrum suggest that such mechanisms may play a role in the population genetics of the trait.

Whether or not there may be a selective advantage to the bipolar trait, a polygenic model does have specific implications regarding the nature of the genetic transmission. As shown in Fig. 4, under a polygenic model, a much larger portion of the population has a few bipolar trait alleles, than the portion of the population that has many alleles, and hence has the disorder. This is consistent with the observation cited above that cyclothymic temperament occurs in about 6% of the normal population as compared to 1% who have bipolar disorder (Placidi et al., 1998). Therefore, under this model, the bipolar trait alleles are primarily transmitted through the population by individuals who do not have bipolar disorder, but rather only a few susceptibility alleles. At the population level, the impact of the phenotype of these much more mildly affected individuals may be much more important in terms of selection and genetic transmission, than the smaller number of individuals who have bipolar disorder. It may require only a small increase in selective fitness in this larger number of individuals, to offset the loss of fitness in those who have a more severe disorder.

In the absence of clear data to define the mode of transmission of bipolar disorder, a definitive theory about the role of bipolar traits in the selection of individuals or groups, and the difficulty in any empirical testing of such hypotheses, these ideas must be considered highly speculative. However, they are discussed in order to raise the issue that the view of bipolar disorder as stemming from genes that are defective and whose phenotype can only be deleterious, may be simplistic. As more data about
Fig. 4. An evolutionary and population genetics perspective on the bipolar spectrum. If bipolar disorder is to some extent a polygenic trait, then it is likely that the susceptibility alleles occur and are transmitted primarily in mildly affected individuals who may have affective temperaments, but not illness. Selective advantage in these individuals may greatly outweigh any selective disadvantage in severely ill individuals who are far fewer in number.

the relationship of bipolar genes to spectrum phenotypes becomes available, more specific hypotheses and theories may be constructed.

10. Conclusions, implications, and directions for future research

To distill all of the above, a complex phenotype is almost certainly the result of a complex genetics. And to go a step beyond that, a complex biology doubtless underlies the complex genetics. There has been a great deal of debate about the nature if not the existence of the bipolar spectrum. This has in particular been controversial in its overlap with other psychiatric diagnoses, and its extension into normal temperaments. An increasing amount of clinical data argue strongly that these are real phenomena. The intent of this paper has been to review the data supporting the idea that the bipolar spectrum has a genetic basis. A further intent was to review the possible genetic models that might underlie such a spectrum and argue that not only are such complex models plausible, but commonly seen in a variety of genetic systems in man and other species. From this genetic perspective, a spectrum of partly qualitative, partly quantitative phenotypes is not surprising at all.

Clearly, the identification of the underlying genes will lead to a much more detailed understanding of the complex genotype–phenotype relationships operating for bipolar traits. However, as is so often the case in science, the reverse is also true. A better understanding of the phenotype and its relationship to genes would greatly facilitate our ability to
identify the genes. As usual, these two must proceed hand in hand as we advance in understanding this illness. The exciting development in this enterprise is availability of powerful tools from the human genome project. The identification of all genes and all human sequence, as well as, tools for large scale high resolution genotyping make even the problems of complex polygenic illness begin to look tractable.

The fruits of progress in this difficult and long term venture will not only be scientifically fascinating, but have the potential to bring powerful new treatments to our patients. The identification of disease genes will lead to the identification of new targets for drug discovery, new methods of diagnosis and ways to predict drug response. It is likely that in the not too distant future, gene therapy could be applied to these disorders, thereby offering a much more specific therapeutic modality. Diagnosis will likely go through a dramatic change such that it may bear little resemblance to our current system. The numerous phenotypic overlaps and co-morbidities observed suggest that behavioral phenotype bears only a limited connection to the underlying pathophysiology. Ultimately, mania and depression, may go the way of jaundice. Jaundice was once considered a diagnosis, but now is known to be a syndrome which can have dozens of different etiologies. Jaundice is only the starting point towards identifying the cause and hence the appropriate pathophysiologically directed treatment. One day we may see mania and depression in a similar light, a starting point towards other testing that points to many different pathophysiologies and treatments that are obscure to us now.

Lastly, though depression appears to have analogs in other species, mania and psychosis do not. It is intriguing to speculate that these behaviors are uniquely human. If the genes for the bipolar spectrum do indeed extend into behaviors within the normal range, then their identification may shed light on some of the most uniquely human behaviors and in turn on what brain processes distinguishes man and his behavior from other species.

Acknowledgements

J.R.K. was supported by grants from the Department of Veterans Affairs, the NIMH (MH59567) and NIDA (DA13769). Support was also provided by the VA VISN 22 Mental Illness Research and Education Center, the UCSD Mental Health Clinical Research Center (MH30914) and the UCSD General Clinical Research Center (M01 RR00827). The author would also like to thank Dr. Hagop Akiskal for many stimulating discussions.

References


Andreasen, N.C., Rice, J., Endicott, J., Reich, T., Coryell, W.,


Mundo, E., Walker, M., Cate, T., Macciardi, F., Kennedy, J.L., 2001. The role of serotonin transporter protein gene in antide-