Network dysfunction of emotional and cognitive processes in those at genetic risk of bipolar disorder

Michael Breakspear,1,2,3,* Gloria Roberts,3,4,* Melissa J. Green,3,4,5,6 Vinh T. Nguyen,1 Andrew Frankland,3,4 Florence Levy,3 Rhoshel Lenroot3,6 and Philip B. Mitchell3,4

*These authors contributed equally to this work.

The emotional and cognitive vulnerabilities that precede the development of bipolar disorder are poorly understood. The inferior frontal gyrus—a key cortical hub for the integration of cognitive and emotional processes—exhibits both structural and functional changes in bipolar disorder, and is also functionally impaired in unaffected first-degree relatives, showing diminished engagement during inhibition of threat-related emotional stimuli. We hypothesized that this functional impairment of the inferior frontal gyrus in those at genetic risk of bipolar disorder reflects the dysfunction of broader network dynamics underlying the coordination of emotion perception and cognitive control. To test this, we studied effective connectivity in functional magnetic resonance imaging data acquired from 41 first-degree relatives of patients with bipolar disorder, 45 matched healthy controls and 55 participants with established bipolar disorder. Dynamic causal modelling was used to model the neuronal interaction between key regions associated with fear perception (the anterior cingulate), inhibition (the left dorsolateral prefrontal cortex) and the region upon which these influences converge, namely the inferior frontal gyrus. Network models that embodied non-linear, hierarchical relationships were the most strongly supported by data from our healthy control and bipolar participants. We observed a marked difference in the hierarchical influence of the anterior cingulate on the effective connectivity from the dorsolateral prefrontal cortex to the inferior frontal gyrus that is unique to the at-risk cohort. Non-specific, non-hierarchical mechanisms appear to compensate for this network disturbance. We thus establish a specific network disturbance suggesting dysfunction in the processes that support hierarchical relationships between emotion and cognitive control in those at high genetic risk for bipolar disorder.

1 QIMR Berghofer, Brisbane, Queensland, Australia
2 Metro North Mental Health Service, Brisbane, Queensland, Australia
3 School of Psychiatry, University of New South Wales, Randwick, NSW, Australia
4 Black Dog Institute, Prince of Wales Hospital, Randwick, NSW, Australia
5 Schizophrenia Research Institute, Sydney, NSW, Australia
6 Neuroscience Research Australia, Randwick, NSW, Australia

Correspondence to: Professor Michael Breakspear,
QIMR Berghofer Institute of Medical Research,
Herston Rd, Herston,
Queensland, Australia
E-mail: michael.breakspear@qimrberghofer.edu.au

Keywords: bipolar disorder; cognitive control; computational psychiatry; imaging; neuropsychiatry

Abbreviations: DCM = dynamic causal modelling; DLPFC = dorsolateral prefrontal cortex
Introduction

Bipolar disorder is characterized by episodic fluctuations in mood and cognition that disrupt function, identity and well-being. Converging structural and functional neuroimaging research has highlighted a number of cortical substrates associated with bipolar disorder (Phillips and Swartz, 2014). Key amongst these is the inferior frontal gyrus (Mazzola-Pomietto et al., 2009; Foland-Ross et al., 2012). Dysfunction of this region is consistent with the illness phenotype because of its crucial integrative role in emotion, perception and cognitive control (Liakakis et al., 2011; Cai et al., 2014). Understanding the processes that lead to this disturbance may hence elucidate neurobiological pathways in bipolar disorder.

Family studies have established a clear genetic contribution to the development of bipolar disorder, with heritability estimates between 59 and 85% (McGuffin et al., 2003; International Schizophrenia Consortium et al., 2009) and an odds ratio of ~7–14 in first-degree relatives (Mortensen et al., 2003; International Schizophrenia Consortium et al., 2009). Unaffected first-degree relatives also show subtle but distressing subclinical disturbances in mood and anxiety, and higher rates of major depressive disorder (DelBello and Geller, 2001; Perich et al., 2015). The late teenage years and twenties comprise the peak age of risk for illness-onset for bipolar disorder and hence include the period during which critical illness-related biological changes are presumed to occur. These years can therefore be conceptualized as a ‘high risk’ state, particularly in those also at increased genetic risk of the disorder (Loranger and Levine, 1978; Goodwin and Jamison, 2007). Thus studying brain network disturbances in unaffected young offspring and siblings of patients with bipolar disorder may shed light on biological risk factors preceding the development of bipolar disorder (Whalley et al., 2011) and identify pathological processes occurring in those who will later convert to the disorder (Goodwin and Jamison, 2007). Also of note, unaffected first-degree relatives of patients with bipolar disorder are largely free from psychotropic medication and illness burden. Such cohorts therefore eschew the problems associated with such confounds when using healthy participants as comparison groups.

We recently reported that young unaffected first-degree relatives of bipolar disorder show a functional impairment in the left inferior frontal gyrus during a task involving active inhibition of motor responses to emotionally salient stimuli (Roberts et al., 2013). Specifically, we acquired functional MRI data while first-degree relatives and matched control participants performed a task that has been well established to elicit interference between inhibitory and emotional processes, namely an emotional ‘Go/No-Go’ (Hare et al., 2008). Responses to fearful faces in the presence of motor inhibition robustly activated a cluster in the left inferior frontal gyrus. The strength of this activation, however, was significantly diminished in the cohort of unaffected first-degree relatives. By revealing a functional impairment in those at genetic risk for bipolar disorder, these findings—in conjunction with similar findings in high risk individuals (Brotman et al., 2014) and those with established bipolar disorder (Foland-Ross et al., 2012; Hajek et al., 2013; Hafeman et al., 2014)—support diminished integrity of the inferior frontal gyrus as a candidate endophenotype in bipolar disorder.

While incipient pathology within the inferior frontal gyrus might underlie this disturbance, emerging research has highlighted the role of wider brain network disturbances underlying psychiatric disorders (Menon, 2011; Fornito and Bullmore, 2012). Although the notion of a ‘disconnection syndrome’ was originally proposed in relation to schizophrenia (Friston and Frith, 1995; Stephan et al., 2006) brain network disturbances have been reported in other disorders, such as bipolar disorder (Leow et al., 2013), Alzheimer’s disease (Seeley et al., 2007) and major depressive disorder (Lord et al., 2012; Hyett et al., 2015). Failure to activate the inferior frontal gyrus during inhibition to emotional stimuli in those at risk of bipolar disorder might reflect suboptimal convergence between the circuits underlying cognitive control and those subserving emotion perception.

Models of emotion dysregulation in bipolar disorder highlight circuits that underlie cognitive control and subserve emotion perception (Phillips et al., 2008; Phillips and Swartz, 2014). For example, aberrant functional connectivity between fronto-limbic regions are evident during emotional perception, cognitive control and emotion regulation in bipolar disorder (Rich et al., 2008; Etkin et al., 2011; Foland-Ross et al., 2011; Morris et al., 2012; Passarotti et al., 2012; Townsend et al., 2013; Bertocci et al., 2014). Functional dysconnectivity of prefronto-limbic systems has also been observed in offspring at familial risk of developing bipolar disorder during an emotional working memory paradigm (Ladouceur et al., 2013), and between the prefrontal cortex and pregenual cingulate during anticipation of reward and loss in the same population (Singh et al., 2014).

We proposed that hypoactivation of the inferior frontal gyrus in the at-risk cohort may arise through a failure in hierarchical network dynamics in these key circuits. In particular, as fronto-temporal circuitry has been implicated during both normative cognitive control of emotion (Ochsner et al., 2012) and in the aetiology of bipolar disorder, we hypothesized that impairment in the inferior frontal gyrus would reflect wider fronto-limbic network disturbances. We tested this hypothesis using dynamic causal modelling (DCM). We first identified cortical substrates in our data for the perception of fear and the execution of motor control. We then constructed network models of the convergence of these effects in the inferior frontal gyrus, considering a constellation of serial, parallel, hierarchical and non-linear models. We report differences in these models and their connectivity parameters between a young, high genetic risk group in comparison to matched healthy controls and those with established bipolar disorder.
Materials and methods

Participants

Our study comprised 41 participants ‘at-risk’ for bipolar disorder, 45 controls and 55 participants with bipolar disorder (Roberts et al., 2013; Perich et al., 2015). In brief, at-risk participants were first-degree relatives of an individual (parent and/or sibling) with an established diagnosis of bipolar disorder, but who did not meet diagnostic criteria for bipolar disorder themselves. The bipolar disorder group were sex- and age-matched participants with established bipolar disorder. The control group was matched for age and sex. Details of sample ascertainment, clinical assessment, plus summary demographic and clinical data are provided in Table 1 and the online Supplementary material, and are described in more detail elsewhere (Perich et al., 2015). We excluded 10 participants from our previous study of this cohort (Roberts et al., 2013) who were either depressed at the time of scanning, on psychotropic medication, or did not complete all functional imaging tasks.

Emotional face Go/No-Go task

Functional MRI data were acquired during two event-related Go/No-Go tasks. Participants were requested to respond to ‘Target’ facial expressions with a button press and ignore any other ‘Distractor’ faces (Fig. 1A). The experiment incorporated an overt emotional task (‘Fear Target/Calm Distractor’, ‘Happy Target/Calm Distractor’, ‘Calm Target/Fear Distractor’, and ‘Calm Target/Happy Distractor’) and a non-emotional (sex-differentiation) task (‘Male Target/Female Distractor’, ‘Female Target/Male Distractor’). Further details of the task, image acquisition and preprocessing are provided in the Supplementary material.

Dynamic causal modelling

Model specification

As discussed above, our previously reported group × condition effect (Roberts et al., 2013) reflected hypoactivation of the left inferior frontal gyrus in the at-risk group during fearful distractor trials. Here, we used DCM to model this effect and, in particular, to infer patterns of effective connectivity underlying the interaction of motor inhibition and fear perception. DCM is a computational approach that allows construction and comparison of network models of functional imaging data (Friston et al., 2003). Specifying dynamic causal models requires two steps. First, regions (network ‘nodes’) that express the specific effects of interest are identified using a standard general linear model. Second, a space of models that embody a variety of possible interactions amongst these regions is specified. Each of these models should offer a plausible explanation of the effects observed in the network nodes.

For the first of these steps, we identified regions showing responses to (i) stimulus inputs; (ii) the effect of fear; (iii) the effect of motor inhibition; and (iv) the interaction of fear and inhibition. These regions were identified using the relevant contrasts in second level group general linear models (GLMs), using all subjects in our cohort and using stringent statistical thresholds (P < 0.05, corrected for family wise error; Fig. 1B and Table 2). Further statistical details on identification of network nodes are provided in the Supplementary material.

The second step in DCM specification involves construction of a space of models that embody various hypotheses about the manner in which these nodes interact—that is, the (effective) connectivity, or network ‘edges’, between the nodes. Restricting the space of models to a relatively small family that test specific hypotheses is an important way to constrain the number (and utility) of models to be tested (Stephan et al., 2010). As the present objective was to use DCM to explore the integration of cognition control and emotion perception (hence, not focused on basic visual processing per se), we restricted our analyses to a small number of models that shared a common input base, and added candidate integrative mechanisms on top of this base. We introduced eight separate models (four bilinear and four non-linear) on top of the common base that modelled serial or parallel integrative mechanisms (see Fig. 2 and ‘Results’ section, for a representation of the specified models). Non-linear models specify hierarchical relationships between the network nodes, i.e. where the neuronal activity in one region gates the flow of activity between other regions (Stephan et al., 2010); bilinear models mirror their more complex non-linear counterparts, except they lack this hierarchical relationship between regions: non-specific modulatory inputs instead fulfil this gating (interaction) function.

Model selection and parameter estimation

Following model specification, DCM uses Bayesian model selection to identify which model is the most likely to have generated the observed data. The process of adjudicating between models essentially balances their goodness of fit against a factor that penalises models for their relative complexity (for review, see Marreiros et al., 2010). Bayesian model selection yields the evidence for each model—the (posterior) probability of the model given the data—as well as the estimated (posterior) parameter values that reflect the strength of interactions between regions. Relative evidence for all models can be used to identify the most likely model, or the best family of models. We performed Bayesian model selection within each group using random effects analysis (Stephan et al., 2009a), then compared models and parameters between groups. To study the role of hierarchical, state-dependent influences, we grouped our models into two families (bilinear versus non-linear) and inferred the most likely family for each group (Penny et al., 2010). To characterize group differences in connectivity parameters, we used a multivariate algorithm, namely canonical variates analysis (CVA). This test accommodates parametric random effects between subjects using classical inference and yields a single P-value that tests for the significance of group effects on sets of (non-linear, bilinear, intrinsic) parameters (Friston et al., 1995).

Results

Demographic and clinical data

The control and at-risk groups did not differ significantly on age, education, IQ, gender distribution, or symptom severity.
Table 1 Demographic and clinical data for ‘at-risk’, control, and bipolar groups

<table>
<thead>
<tr>
<th>Demographic data</th>
<th>Control (n = 45)</th>
<th>At-risk (n = 41)</th>
<th>Post hoc at-risk versus control differences</th>
<th>Bipolar disorder (n = 55)</th>
<th>Difference statistic</th>
<th>Post hoc Bipolar disorder differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD)</td>
<td>23.3 (3.5)</td>
<td>24.4 (3.8)</td>
<td></td>
<td>25.0 (3.5)</td>
<td>F = 2.86</td>
<td>-</td>
</tr>
<tr>
<td>Years of education, mean (SD)</td>
<td>16.1 (2.1)</td>
<td>15.6 (2.3)</td>
<td></td>
<td>15.7 (2.2)</td>
<td>F = 0.57</td>
<td>-</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>119.5 (12.1)</td>
<td>119.7 (9.6)</td>
<td></td>
<td>115.5 (12.1)</td>
<td>F = 1.97</td>
<td>-</td>
</tr>
<tr>
<td>Females, n (%)</td>
<td>31 (68.9)</td>
<td>20 (48.8)</td>
<td></td>
<td>37 (67.3)</td>
<td>$\chi^2 = 4.61$</td>
<td>-</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>14 (31.1)</td>
<td>21 (51.2)</td>
<td></td>
<td>18 (32.7)</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Global functioning at time of interview</td>
<td>92.5 (4.1)</td>
<td>87.2 (8.9)</td>
<td>CON &gt; AR**</td>
<td>78.4 (11.4)</td>
<td>F = 32.2***</td>
<td>CON &gt; BD*** AR &gt; BD***</td>
</tr>
<tr>
<td>Symptom severity scales</td>
<td>22–30 years</td>
<td>22–30 years</td>
<td>22–30 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DASS stress, mean (SD)</td>
<td>6.8 (7.7)</td>
<td>8.5 (6.6)</td>
<td></td>
<td>16.1 (11.2)</td>
<td>F = 10.67***</td>
<td>BD &gt; AR** BD &gt; CON***</td>
</tr>
<tr>
<td>DASS anxiety, mean (SD)</td>
<td>2.6 (4.1)</td>
<td>3.5 (5.0)</td>
<td></td>
<td>10.1 (9.7)</td>
<td>F = 11.98***</td>
<td>BD &gt; AR** BD &gt; CON***</td>
</tr>
<tr>
<td>DASS depression, mean (SD)</td>
<td>3.6 (7.3)</td>
<td>2.9 (4.4)</td>
<td></td>
<td>13.23 (11.6)</td>
<td>F = 15.4***</td>
<td>BD &gt; AR** BD &gt; CON***</td>
</tr>
<tr>
<td>ASRM, mean (SD)</td>
<td>3.4 (3.5)</td>
<td>3.0 (2.8)</td>
<td></td>
<td>3.85 (4.4)</td>
<td>F = 0.4</td>
<td>-</td>
</tr>
<tr>
<td>ISS Perceived conflict, mean (SD)</td>
<td>34.2 (46.6)</td>
<td>35.4 (48.9)</td>
<td></td>
<td>94.87</td>
<td>F = 9.48**</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>ISS Wellbeing, mean (SD)</td>
<td>178.7 (43.5)</td>
<td>175.4 (53.9)</td>
<td></td>
<td>149.49</td>
<td>F = 2.70</td>
<td>-</td>
</tr>
<tr>
<td>ISS Activation, mean (SD)</td>
<td>83.9 (96.5)</td>
<td>69.3 (73.9)</td>
<td></td>
<td>117.69</td>
<td>F = 2.18</td>
<td>-</td>
</tr>
<tr>
<td>ISS Depression, mean (SD)</td>
<td>12.9 (24.0)</td>
<td>12.9 (18.6)</td>
<td></td>
<td>52.05</td>
<td>F = 12.29***</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>MADRS, mean (SD)</td>
<td>1.3 (1.8)</td>
<td>2.1 (2.7)</td>
<td></td>
<td>11.65 (11.1)</td>
<td>F = 21.42**</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>BDRS, mean (SD)</td>
<td>1.4 (1.3)</td>
<td>2.1 (2.7)</td>
<td></td>
<td>10.87 (5.5)</td>
<td>F = 22.4**</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>YMRS, mean (SD)</td>
<td>1.0 (1.2)</td>
<td>0.5 (1.0)</td>
<td></td>
<td>4.92 (4.7)</td>
<td>F = 21.3**</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>CDI, mean (SD)</td>
<td>7.3 (3.6)</td>
<td>8.2 (4.0)</td>
<td></td>
<td>23.4 (7.9)</td>
<td>F = 24.73***</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>CDI Ineffectiveness, mean (SD)</td>
<td>2.1 (1.0)</td>
<td>2.1 (0.8)</td>
<td></td>
<td>4.12 (1.8)</td>
<td>F = 10.36**</td>
<td>BD &gt; AR* BD &gt; CON**</td>
</tr>
<tr>
<td>CDI Negative mood, mean (SD)</td>
<td>1.8 (1.1)</td>
<td>1.5 (0.8)</td>
<td></td>
<td>6.37 (2.6)</td>
<td>F = 25.77**</td>
<td>BD &gt; AR** BD &gt; CON**</td>
</tr>
<tr>
<td>CDI Anhedonia, mean (SD)</td>
<td>3.0 (1.8)</td>
<td>3.1 (1.8)</td>
<td></td>
<td>6.87 (2.6)</td>
<td>F = 10.5**</td>
<td>BD &gt; AR* BD &gt; CON**</td>
</tr>
<tr>
<td>CDI Interpersonal problems, mean (SD)</td>
<td>0.5 (0.6)</td>
<td>0.3 (0.3)</td>
<td></td>
<td>1.50 (1.1)</td>
<td>F = 8.50**</td>
<td>BD &gt; AR** BD &gt; CON**</td>
</tr>
<tr>
<td>CDI Negative self-esteem, mean (SD)</td>
<td>1.3 (0.5)</td>
<td>1.3 (0.3)</td>
<td></td>
<td>4.50 (2.6)</td>
<td>F = 19.47**</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>Consensus DSM-IV diagnosis</td>
<td>(n = 45)</td>
<td>(n = 41)</td>
<td>(n = 55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any lifetime psychiatric diagnosis, n (%)</td>
<td>9 (20.0)</td>
<td>25 (61.0)</td>
<td>AR &gt; CON**</td>
<td>55 (100.0)</td>
<td>$\chi^2 = 68.16$***</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>Lifetime major depressive episode, n (%)</td>
<td>2 (4.4)</td>
<td>14 (34.1)</td>
<td>AR &gt; CON**</td>
<td>50 (90.9)</td>
<td>$\chi^2 = 78.04$***</td>
<td>BD &gt; AR*** BD &gt; CON***</td>
</tr>
<tr>
<td>Lifetime recurrent MDD, n (%)</td>
<td>0 (0)</td>
<td>4 (9.8)</td>
<td>AR &gt; CON**</td>
<td>0 (0)</td>
<td>$\chi^2 = 10.04$**</td>
<td>BD &gt; AR* BD &gt; CON**</td>
</tr>
<tr>
<td>Lifetime anxiety disorder, n (%)</td>
<td>3 (6.7)</td>
<td>10 (24.4)</td>
<td>AR &gt; CON**</td>
<td>26 (49.1)</td>
<td>$\chi^2 = 23.61$**</td>
<td>BD &gt; AR* BD &gt; CON**</td>
</tr>
<tr>
<td>Lifetime behavioural disorder, n (%)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td>8 (15.4)</td>
<td>$\chi^2 = 14.04$**</td>
<td>-</td>
</tr>
<tr>
<td>Lifetime substance disorder, n (%)</td>
<td>1 (2.2)</td>
<td>4 (9.8)</td>
<td></td>
<td>14 (25.5)</td>
<td>$\chi^2 = 12.14$**</td>
<td>-</td>
</tr>
</tbody>
</table>
measures of depressive, anxiety, or hypo/manic symptoms. Current and lifetime consensus best-estimate DSM-IV diagnoses are detailed in Table 1. At the time of testing, no control or at-risk participant met DSM-IV criteria for a diagnosis of a current episode of major depression mania/hypomania and no control or at-risk participant was currently prescribed psychotropics. Pearson’s chi-square tests revealed that the lifetime occurrence of at least one DSM-IV psychiatric diagnosis was significantly higher in the at-risk compared to control group (\(P < 0.05\)), consistent with prior reports of at-risk populations (Birmaher et al., 2009; Nurnberger et al., 2011). In particular, a greater proportion of the at-risk group reported either a lifetime single major depressive episode (\(P < 0.05\)) or recurrent major depressive disorder (\(P < 0.05\)) compared to the control group. Rates of anxiety disorders (\(P < 0.05\)) were also significantly higher among the at-risk group compared to the control group. As expected, a number of clinical and lifetime psychiatric measures differed between the bipolar disorder and both control and at-risk groups (Table 1).

### Behavioural performance

All participants scored over 65% accuracy on the Go/No-Go task (range 65–100%). There were no significant group differences in overall accuracy for responses to targets (\(P > 0.16\)), distractors (\(P > 0.52\)), or for reaction time during target trials (\(P > 0.12\)) across the three groups. The same behavioural results held for emotional and non-emotional conditions when analysed separately (Supplementary material).

### Network nodes

The regions showing responses to faces, the effect of motor inhibition, the effect of fear, and the interaction of fear and inhibition are shown in Fig. 1B and Table 2. The main effect of face stimuli elicited a robust effect in fusiform visual cortex, a region well documented for processing of faces (Kanwisher and Yovel, 2006). Motor inhibition was associated with a strong response in the left dorsolateral prefrontal cortex, a canonical hub in the cognitive control circuitry (Miller, 2000; Rushworth et al., 2007) evidencing a particularly strong association with response inhibition (Menon et al., 2001; Rubia et al., 2003; Buchsbaum et al., 2005). The fear contrast revealed a robust activation in the left dorsal anterior cingulate consistent with prior reports of the role of this region in perception of fear expression (Milad et al., 2007; Etkin et al., 2011). Finally, the interaction contrast (fear \(\times\) inhibition) revealed a robust and specific effect in the left inferior frontal gyrus. Notably, this cluster revealed here in the whole group (at-risk + control) contrast corresponds to the region where we previously observed a group difference (Roberts et al., 2013) (at-risk cohort showing diminished activation). Figure 1C summarizes the relative location of these four nodes and the effects each embody.

### Dynamic causal modelling

We next specified dynamic causal models of these data. All models shared a common input base, beginning with stimulus inputs (i.e. faces) directed to the facial fusiform area (Fig. 1D). An effective connection from the facial fusiform area to the anterior cingulate, modulated by the presence of fear faces, modelled the effect of fear observed in the anterior cingulate. Likewise, the effect of inhibition expressed in the DLPFC was modelled by an effective connection from facial fusiform area to DLPFC, modulated by motor inhibition.

We specified eight separate models (four bilinear and four non-linear) on top of this common base that represent serial, parallel or hierarchical processes (Fig. 2). As the name suggests, in serial models (both bilinear and non-linear), information passes in a serial manner from the facial fusiform area via the anterior cingulate or the DLPFC (or both) \textit{en route} to the inferior frontal gyrus. In parallel models, there is a direct effective connection from the facial fusiform area to the inferior frontal gyrus, hence in parallel to the anterior cingulate and DLPFC connections. Additional modulatory influences are introduced on top of these architectures to explain the interaction effect in the inferior frontal gyrus. In the non-linear models (\(m5–8\)) the modulation of inputs to...
inferior frontal gyrus is mediated by modulation of connections from one area by another (namely anterior cingulate or DLPFC). This activity-dependent modulation can be considered hierarchical. In contrast, in bilinear models (m1–4), this modulation is attributed directly to experimental inputs (namely, inhibition and fear). In short, both bilinear and non-linear models allow for context or state-dependent changes in afferents to the inferior frontal gyrus. However, non-linear models consider this state-dependent modulation to be dynamic and activity-dependent. These eight models encompass all possible such serial, parallel and hierarchical arrangements.

Table 2 Network nodes

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Cluster region</th>
<th>Peak coordinates (x,y,z)</th>
<th>Peak -Level t</th>
<th>Peak Z-score</th>
<th>Peak-level P (FWE corrected)</th>
<th>Cluster-level P (FWE corrected)</th>
<th>Cluster size (voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1-Faces</td>
<td>L. fusiform/primary visual cortex</td>
<td>−42, −49, −26</td>
<td>13.10</td>
<td>Inf</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>460</td>
</tr>
<tr>
<td>X2-Inhibition</td>
<td>L. dorsolateral prefrontal cortex</td>
<td>−39, 17, 25</td>
<td>4.99</td>
<td>4.91</td>
<td>0.010</td>
<td>&lt;0.001</td>
<td>60</td>
</tr>
<tr>
<td>X3-Fear</td>
<td>L. anterior cingulate</td>
<td>−3, 47, −2</td>
<td>5.40</td>
<td>5.31</td>
<td>0.002</td>
<td>0.001</td>
<td>28</td>
</tr>
<tr>
<td>X4-Fear × inhibition</td>
<td>L. inferior frontal gyrus</td>
<td>−21, 11, −17</td>
<td>6.39</td>
<td>6.24</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>48</td>
</tr>
</tbody>
</table>
considered separately. Because we sought a parsimonious and non-redundant model space, we did not consider models that combine these basic features (for example both serial and parallel connections). A more detailed explanation of these models is provided in the Supplementary material.

Model selection (Bayesian model selection) was performed within each of our three cohorts. We first inverted this family of eight models from the data of the matched control cohort (controls), grouping the non-linear and bilinear models into two sub-families. The non-linear, hierarchical models had far higher exceedance probability than their bilinear counterparts (Fig. 3A). That is, the family of non-linear models provided a more accurate and parsimonious description of the effects in the control data than the bilinear models. This implies that models that embody the interaction of fear and motor inhibition through implicit hierarchical interactions perform better than those models where these effects are introduced by external, bilinear terms.

We then inverted models from both the at-risk and bipolar disorder groups and compared the parameters of these non-linear models between our two groups. Canonical variates analysis revealed a significant and specific between group effect for the values of the four non-linear (hierarchical) parameters ($P < 0.0156$). Post hoc analyses of each of these parameters using individual $t$-tests showed that this group effect was largely driven by a significant between-group effect ($P < 0.0249$) of the (non-linear) parameter representing the hierarchical influence of anterior cingulate on the DLPFC to inferior frontal gyrus connection (Fig. 3B). There is thus a group difference between the control and at-risk groups in the hierarchical gating of fear perception—expressed in the anterior cingulate, on the inhibition of motor response—represented by the influence of the DLPFC on the inferior frontal gyrus. The other non-linear parameters pooled from these non-linear models did not significantly differ ($P > 0.29$; Supplementary Fig. 2).

There were no significant differences for the four non-linear parameters between the control and bipolar disorder groups ($P > 0.65$).

We next compared bilinear to non-linear models in the at-risk cohort (Fig. 3C). In marked contrast to the controls, the bilinear models showed a far higher exceedance probability than their non-linear counterparts. That is, the models in which the interaction between fear and motor inhibition were introduced as external modulators (and

---

**Figure 2 Model space.** (A) Bilinear models. Asymmetric ($m_1$, $m_2$) and symmetric ($m_3$) serial models. Parallel model ($m_4$). (B) Non-linear models. These mirror the bilinear models with the exception that the interaction effects into inferior frontal gyrus are directly mediated by state dependent (non-linear) effects in anterior cingulate and/or DLPFC.
not through intrinsic hierarchical connections) provided the most accurate and parsimonious account of the at-risk functional MRI data. Taking the ratio of the posterior probability of models in both the control and at-risk groups yields an odds ratio of non-linear versus bilinear models of 2.6, corresponding to a robust effect (Stephan et al., 2009a).

Inversion of all models in the bipolar disorder group showed that the non-linear family was the most likely (Fig. 3D). That is, the bipolar disorder group mirrored the control group, both in terms of the individual non-linear parameters, and for the overall most likely family of models.

Inspecting the posterior probability of individual models (Supplementary Fig. 1) echoes the results when grouping models into families (non-linear versus bilinear). The three most likely models in the control group (m5, m7, and m8) and two most likely in the bipolar disorder group (m7, m8) are all non-linear, while the bilinear models have a fairly uniform (and low) model-likelihood. The asymmetric bilinear models (m1, m2) have substantially higher likelihood in the at-risk group than the control group, whereas all non-linear models have lower likelihood in the at-risk group than their corresponding bilinear counterparts. Taken together, however, the model space in the control and at-risk groups is reasonably flat, hence we draw our main conclusions from the family-level inference that tends to average out variability within groups.

**Post hoc analyses**

Rates of non-bipolar psychopathology (such as anxiety) were significantly higher in the at-risk and bipolar disorder groups compared to the control group (Table 1). To determine whether non-bipolar psychopathology influenced network connectivity abnormalities, we carried out secondary correlational analyses with symptom severity measures (total Children’s Depression Inventory scores of 12–21 year olds and total Montgomery Åsberg Depression Rating Scale scores of 22–30 year olds), and examined group differences in network connectivity based on the presence of any lifetime psychiatric diagnoses. No significant correlations were found between the hierarchical influence of the anterior cingulate on the DLPFC to inferior frontal gyrus connection and symptom severity (when confined to the at-risk group or considered across the whole cohort), and overall lifetime psychiatric diagnosis had no
effect on network connectivity differences among groups (see Supplementary material for further details).

**Discussion**

The inferior frontal gyrus is a key cortical hub that integrates emotion and executive processes (Ochsner and Gross, 2005; Liakakis et al., 2011; Cai et al., 2014) and exhibits a functional disturbance in those at high genetic risk of bipolar disorder during inhibition of responses to fearful faces (Roberts et al., 2013). We here show that this disturbance is associated with a failure of hierarchical relationships between emotion processing in the anterior cingulate and cognitive control in the DLPFC. In particular, we observed a strong difference between the control and at-risk groups in the parameter corresponding to the hierarchical gating of effective connectivity from the DLPFC to the inferior frontal gyrus by state-dependent activity in the anterior cingulate. This difference in non-linear, gating parameter values was accompanied by a failure of non-linear, hierarchical models of effective connectivity that was not present in the bipolar disorder group, and hence unique to the at-risk group. Our findings thus suggest a disruption in the influence of fearful emotion processing on the neuronal mechanisms of cognitive control specific to the at-risk group, in which the usual gating function of the anterior cingulate appears to be modulated instead by non-specific dynamic mechanisms.

Of note, these different effective network mechanisms underlying inhibition of fearful stimuli were not associated with impaired performance on this task. We credit this maintenance of performance to the bilinear network effects in the at-risk group (Fig. 3C). That is, given the disturbance in the hierarchical gating by the anterior cingulate, alternative modulatory influences on the DLPFC to inferior frontal gyrus connection appear to have been instituted in the at-risk group. These alternative influences may represent early markers, or may serve as compensatory mechanisms, allowing for preservation of adaptive emotion regulation in those with genetic vulnerabilities for bipolar disorder. Based on prior research, we speculate that the medial prefrontal cortex and amygdala may serve as candidate alternative regions: activity in the medial prefrontal cortex was previously found to be greater in bipolar disorder and at-risk compared to control participants in response to fearful stimuli (Surguladze et al., 2010). In an emotion regulation task, at-risk participants showed increased amygdala reactivity during the presentation of negative stimuli (Heissler et al., 2014). Reduced ventrolateral prefrontal modulation of the amygdala was found during the presentation of emotional distractors in an emotional working memory paradigm (Ladouceur et al., 2013). There is, nonetheless, diminished activation in the inferior frontal gyrus during response inhibition to fearful stimuli in our data, suggesting suboptimal compensation in at least some of our at-risk participants. Partial failure of these compensatory processes during emotionally salient experiences of greater intensity might, in principle, underlie the subtle, subclinical disturbances in mood and anxiety seen in this at-risk cohort (Perich et al., 2015).

The DLPFC and anterior cingulate are among the key regions that have been implicated in emotion dysregulation in bipolar disorder (Phillips et al., 2008; Phillips and Swartz, 2014). Meta-analysis of mixed emotional and cognitive activities showed increased activity in regions such as the DLPFC in at-risk compared to control and bipolar disorder participants (Lee et al., 2014). Dysconnectivity between the prefrontal cortex and pregenual cingulate has also been observed during a reward and loss task in at-risk participants (Singh et al., 2014). Moreover, at-risk individuals who subsequently converted to bipolar disorder show increased activation of the anterior cingulate cortex during an emotional memory task (Whalley et al., 2015). These findings in bipolar disorder are also consistent with previous findings of increased activation in the subgenual anterior cingulate cortex in bipolar disorder but not at-risk participants, relative to controls, during performance of a (non-emotional) cognitive flexibility task (Kim et al., 2012). In line with the current proposal of a disruption in the influence of fearful emotion processing in our at-risk cohort, a prior meta-analysis provides evidence for decreased left anterior cingulate cortex activation in patients with bipolar disorder compared to controls during processing of fearful stimuli (Delvecchio et al., 2012).

The present findings shed new light on a number of issues: first, they show that at-risk participants demonstrate distinct brain network mechanisms for the integration of cognitive control and emotion perception. Second, this finding adds further weight to the utility of studying the human brain ‘connectomics’ in psychiatric research (Fornito and Bullmore, 2012; Fornito and Harrison, 2012; Fornito et al., 2015). While most connectivity research on bipolar disorder has focused on structural and functional connectivity, our work uses DCM to study effective (dys)-connectivity—that is, disturbed influences inferred at the neuronal level using model inversion to move beyond signal correlations. The previous papers that have used this approach in psychiatry include the study of young people at risk of schizophrenia (Diwadkar et al., 2012) as well as the analysis of network dysfunction in melancholia (Hyett et al., 2015). This study is the first application of DCM to those at risk to bipolar disorder. Our findings derive from a second order (non-linear) dynamic network effect—that is, the influence of the anterior cingulate on the connection from the DLPFC to the inferior frontal gyrus. Such a ‘state-dependent effect’ requires the use of computational models (Freyer et al., 2012) and cannot be captured by the application of functional connectivity, which is confined to the analysis of first-order correlations.

A third important finding derives from model evidence within the control group (Supplementary Fig. 1). Here we see that the two most likely models that can account for the data are both symmetric non-linear models (m7, m8), such
that the gating influences of the anterior cingulate and DLPFC on the inferior frontal gyrus are balanced. Put differently, neither the neuronal correlates of emotion, nor those of cognitive inhibition, enjoy a position of privilege. Rather there exist dynamic mechanisms allowing each to gate the influence of the other. As noted above, in the at-risk group, there is a substantial between-difference in the gating influence of one half of this balance (the anterior cingulate on DLPFC) under the conditions posed by this task. Interestingly, the most likely model in the at-risk group is the asymmetric bilinear model (m1), whereby the modulatory effect of fearful stimuli of the DLPFC to inferior frontal gyrus connection is subordinate to the earlier effect of inhibition. Further work is required to understand how these network effects reflect the subtle changes in emotion regulation in those at high genetic risk of bipolar disorder and to test whether these changes pre-empt full illness expression in those who convert to the disorder.

Traditionally, findings of illness-associated 'hypo-activation' in a region such as the inferior frontal gyrus triggered molecular biological studies to elucidate signs of incipient pathology in the affected region. Contemporary research is increasingly focused on brain network disturbances, highlighting system level disturbances that cannot be simplified to molecular or cellular level disturbances in one specific region. However, the two approaches are not mutually exclusive. The integrity of cortical tissue in any brain region is dependent upon activity-dependent trophic influences sustained by inputs arising from other regions (for review, see Menon et al., 2001). Hence, even if the hypoactivation in the inferior frontal gyrus did arise through a network disturbance, local structural changes might consequently arise (Fornito and Harrison, 2012). Changes in cortical thickness as well as structural and functional (dys)connectivity in this cohort are yet to be determined. Future attempts to identify longitudinal trajectories in each of these domains may assist in disambiguating between local and network-level events in the causal chain of effects underlying progression to bipolar disorder.

While the focus of the present study is on unaffected individuals at increased genetic risk, we also included comparison to an age- and gender-matched bipolar disorder cohort to aid in interpretation. Not surprisingly, the bipolar disorder group differed markedly across a broad range of clinical indices, suggesting caution when interpreting contrasts involving this group. Notwithstanding this limitation, effective connectivity amongst this network of regions did not appear to be impaired in the bipolar disorder group, reflecting the same non-linear mechanisms as the control cohort. Hence, failure of hierarchical mechanisms appears to be unique to the at-risk group. We offer several possible interpretations of the consistency between the control and bipolar disorder groups. First, we note that the inferior frontal gyrus region was specifically identified by a prior contrast between the control and at-risk groups for task-related activity. That is, we have biased our search for network mechanisms towards those likely showing dysfunction in the at-risk group. Future work is required to identify candidate regions and networks that may differ in bipolar disorder, considered alone or in a three-way group contrast. Second, patients with bipolar disorder are by necessity often on psychotropic medication, particularly mood stabilizers such as lithium; it is possible that state-dependent differences in neuronal dynamics are the target of such compounds which may thus show a normalizing effect. Third, there may actually be a unique endogenous marker unique to the high risk group as they transition through the highest developmental period for the development of the disorder. Future, longitudinal work is required to disambiguate these possibilities.

This brings us to several important study limitations. Our participant group is likely to be heterogeneous; some at-risk participants may develop bipolar disorder while others will not. Longitudinal follow-up is required to identify early pathways from at-risk to bipolar disorder. In this vein, we predict that those at-risk participants who develop fully established bipolar disorder will be those in whom the neurobiological processes we have quantified are already furthest from the estimates in the control cohort. It is for this reason that the present cohort forms the basis of a longitudinal study. However, given their young age, the annual conversion is low, such that testing this prediction mandates a lag of 6 to 10 years from the present baseline assessment. Regarding the extent of networks tested, we deliberately restricted our analyses to a relatively small space of models, focusing on network mechanisms for the integration of emotion and inhibition. Additional model features, such as reciprocal connections amongst all nodes could, in theory have been added to expand the repertoire of models. However, two parsimonious principles guided us against this. First, although family-wise inference is well suited to large model spaces (Penny et al., 2010; Harding et al., 2014) restricting the number of causal models to the bare minimum of features that can explain the data avoids redundancies and allows simpler interpretations (Nguyen et al., 2014). Second, our focus on a particular dynamic mechanism, namely the presence or absence of hierarchical (non-linear) effects between emotion and executive function meant that model space could be restricted to match these specific core hypotheses, following ‘good practice’ guidelines for DCM (Stephan et al., 2010). Future work could also use other tasks that involve explicit instructions to regulate emotion (Ochsner and Gross, 2005) to quantify the network contributions of other relevant structures including subgenual anterior cingulate cortex, insula and amygdala.

We also note that a previous analysis of effective connectivity during emotion perception in a (schizophrenia) high risk cohort incorporated the amygdala in the basic ‘emotion route’ (Diwadkar et al., 2012). This accords with substantial prior literature implicating the amygdala in the processing of emotion (Phelps and LeDoux, 2005). However, despite strong effects in other regions, there was no significant effect in the amygdala for the contrast
between fearful and calm faces in our data. Thus we did not include the amygdala in DCM analyses owing to the lack of (amygdala) effects to model. While it may be argued that the amygdala could—in principle—have replaced the bilinear ‘fear’ modulation in all models (blue arrows), this could not be justified according to the effects in our data. We concede that this does not mean that the amygdala was not involved in routing (all) facial stimuli onwards from the facial fusiform area, but only that this did not substantially differ between processing of fearful versus calm faces. Despite these differences in model specification (as well as task instructions), Diwadkar et al. (2012) also implicated changes in effective connectivity of the DLPFC in their cohort of high risk for schizophrenia. Future work using DCM (Brodersen et al., 2011) might be helpful in disambiguating between the neurobiological correlates of risk for these two genetically related disorders (International Schizophrenia Consortium et al., 2009).

Recent commentaries have highlighted the limitations of imaging studies containing small sample sizes (often \( n = 20 \) or less) (Button et al., 2013). However, our present age- and gender-matched cohort (\( n = 144 \)) shows a strong, and specific (single cluster) group effect (Roberts et al., 2013) that indicates evidence for adequate power in this imaging study (David et al., 2013). Further, the core result of our study—a failure of hierarchical processes in the at-risk cohort—rests upon the ratio of (log) model probabilities which compliments reliance upon P-values alone with a quantitative estimate of the odds ratio (Wagenmakers et al., 2014). Our restricted model space embodies a small number of hypotheses about the effects in our data and, being hypothesis-driven, avoids the multiple comparison problem that plagues even large ‘discovery’ studies. Notwithstanding these issues, we note the known small size of genetic effects on imaging-based brain markers (Stein et al., 2012) and hence decided against testing for brain–gene associations.

These findings lay the platform for several future lines of enquiry. First, the presence of a group difference in the non-linear, gating parameter offers an intriguing mechanism that has been previously highlighted in regards to non-linear DCMs (Stephan et al., 2006, 2008). That is, state-dependent effects—as embodied in non-linear DCMs—correspond to dynamic, activity-dependent processes at the neuronal level. Whilst there are a myriad of possible mechanisms, the proximity of ligand-gated AMPA and voltage-dependent NMDA receptors has been previously invoked in this regard (Stephan et al., 2008). This putative mechanism could be tested through pharmacological manipulations using low doses of agents with action at NMDA receptors, such as ketamine (Moran et al., 2015). Second, as noted above, the present study is embedded in a longitudinal high risk project which will permit explicit testing of the predictive validity of the non-linear parameters highlighted here. Finally, it would be of interest to understand the possible role of structural dysconnectivity in the present cohort through use of a suitable multimodal framework (Stephan et al., 2009b).

Acknowledgements

We are grateful to the many individuals and families who have generously agreed to participate in this research study.

Funding

This study was funded by an Australian National Health and Medical Research Council Program Grant (1037196) and the Lansdowne Foundation. M.B. acknowledges a Health Research Fellowship from the Queensland Office of Health and Medical Research. M.J.G. was supported by a NHMRC R.D. Wright Biomedical Career Development Award (1061875) and a grant to the Netherlands Institute for Advance Studies.

Supplementary material

Supplementary material is available at Brain online.

References


