Predicting White Matter Integrity from Multiple Common Genetic Variants

Omid Kohannim[1], Neda Jahanshad[1,2], Meredith N Braskie[1], Jason L Stein[1], Ming-Chang Chiang[1,3], April H Reese[1], Derrek P Hibar[1], Arthur W Toga[1], Katie L McMahon[4], Greig I de Zubicaray[5], Sarah E Medland[6], Grant W Montgomery[6], Nicholas G Martin[6], Margaret J Wright[6] and Paul M Thompson*[,1]

[1]Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, CA, USA; [2]Medical Imaging Informatics Group, Department of Radiology, UCLA School of Medicine, Los Angeles, CA, USA; [3]Department of Biomedical Engineering, National Yang-Ming University, Taipei, Taiwan; [4]Center for Advanced Imaging, University of Queensland, Brisbane, QLD, Australia; [5]School of Psychology, University of Queensland, Brisbane, QLD, Australia; [6]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, QLD, Australia

Several common genetic variants have recently been discovered that appear to influence white matter microstructure, as measured by diffusion tensor imaging (DTI). Each genetic variant explains only a small proportion of the variance in brain microstructure, so we set out to explore their combined effect on the white matter integrity of the corpus callosum. We measured six common candidate single-nucleotide polymorphisms (SNPs) in the *COMT, NTRK1, BDNF, ErbB4, CLU,* and *HFE* genes, and investigated their individual and aggregate effects on white matter structure in 395 healthy adult twins and siblings (age: 20–30 years). All subjects were scanned with 4-tesla 94-direction high angular resolution diffusion imaging. When combined using mixed-effects linear regression, a joint model based on five of the candidate SNPs (*COMT, NTRK1, ErbB4, CLU,* and *HFE*) explained ∼6% of the variance in the average fractional anisotropy (FA) of the corpus callosum. This predictive model had detectable effects on FA at 82% of the corpus callosum voxels, including the genu, body, and splenium. Predicting the brain's fiber microstructure from genotypes may ultimately help in early risk assessment, and eventually, in personalized treatment for neuropsychiatric disorders in which brain integrity and connectivity are affected.
Neuropsychopharmacology (2012) **37,** 2012–2019; doi:10.1038/npp.2012.49; published online 18 April 2012

Keywords: neuroimaging; brain structure; DTI; genetics; genetic profiles; prediction

INTRODUCTION

Diffusion tensor imaging (DTI) is widely acknowledged as a useful tool for studying the white matter microstructure of the living brain. By mapping the diffusion of water through the brain's fibers, DTI can recover major fiber pathways in the brain, and patterns of anatomical connectivity, with broad applications in psychiatry, neurology, and brain mapping (Thomason and Thompson, 2011). DTI-based white matter abnormalities are widely reported in developmental and degenerative brain diseases including Alzheimer's disease and mild cognitive impairment (Fellgiebel *et al*, 2004; Naggara *et al*, 2006; Oishi *et al*, 2011), schizophrenia (White *et al*, 2008; Ellison-Wright and Bullmore, 2009; Patel *et al*, 2011), bipolar disorder (Sussmann *et al*, 2009; Heng *et al*, 2010), attention-deficit/hyperactivity disorder (Konrad and Eickhoff, 2010), and autism (Alexander *et al*, 2007; Ke *et al*, 2009). These studies show the utility of

DTI in neuropsychiatric research. In several studies, treatment of neuropsychiatric patients has also been associated with changes in DTI measures (Versace *et al*, 2008; Yoo *et al*, 2007). This also shows the promise of DTI for understanding therapeutic effects.

Measures of white matter integrity derived from DTI, such as fractional anisotropy (FA), are highly heritable (Lee *et al*, 2008; Chiang *et al*, 2009; Kochunov *et al*, 2010; Lee *et al*, 2010; Patel *et al*, 2010; Chiang *et al*, 2011b). As such, they may be useful as intermediate measures or 'endophenotypes' (Meyer-Lindenberg and Weinberger, 2006; de Geus *et al*, 2008; Hall and Smoller, 2010; Marenco and Radulescu, 2010) for assessing genetic influences on the brain. Several commonly carried genetic variants have already been identified that exert small effects on the brain's white matter as detected by DTI. These include highly prevalent polymorphisms in genes coding for brain-derived neurotrophic factor (*BDNF*; Chiang *et al*, 2011a), clusterin (*CLU*; Braskie *et al*, 2011), the neuregulin 1 receptor (*ErbB4*; Konrad *et al*, 2009), neurotrophic tyrosine kinase receptor-type 1 (*NTRK1*; Braskie *et al*, 2012), catechol-*O*-methyl transferase (*COMT*; Thomason *et al*, 2010), and the hemochromatosis (*HFE)* gene (Jahanshad *et al*, 2012a). We therefore considered these genes as candidates in this study.

The molecular and cellular effects of these genes and their protein products have been extensively investigated. *COMT* is a well-studied gene and codes for one of the group of enzymes that degrade catecholamines. Catecholamine levels are altered in many neuropsychiatric disorders, thereby making this molecule an ideal target for medications. Several of the genes above are also well known for their role in brain development. *BDNF*'s protein product is a neural growth factor or neurotrophin, vital for the healthy development and maintenance of the nervous system (Binder and Scharfman, 2004). Similarly, *NTRK1* codes for TrkA, which belongs to a tyrosine kinase receptor family, to which neurotrophin growth factors bind. Neurotrophins and their receptors, not surprisingly, are also important in neuropsychiatric disease and may offer new therapeutic targets in the form of small-molecule antagonists or mimickers (Allen and Dawbarn, 2006). *ErbB4* encodes another tyrosine kinase receptor, which by binding to its ligand, neuregulin-1 (coded by *NRG1*), participates in neural modulation and development and is thought to contribute to the pathophysiology of schizophrenia (Hahn *et al*, 2006). Lastly, *HFE* and *CLU* contain polymorphisms that increase the risk for neurodegenerative disease. Their protein products regulate iron metabolism—important in brain aging (Bartzokis *et al*, 2011)—and beta-amyloid metabolism (DeMattos *et al*, 2002), respectively.

White matter structure is certainly influenced by nongenetic factors such as age (Chiang *et al*, 2011b), and sex differences (which are partly genetic and nongenetic), but we expect a moderate and significant proportion of an individual's white matter integrity to be predictable from their genetic profiles. This is corroborated by DTI findings of high heritability for white matter microstructure. As mentioned above, individual effects of single genetic variants on white matter structure have been explored, but a multilocus approach has not yet been taken. The utility of a multilocus candidate gene model in predicting an imaging-derived outcome was recently explored in the context of structural MRI (Biffi *et al*, 2010) and functional MRI (Nikolova *et al*, 2011), but its applications in DTI and detailed three-dimensional maps of brain structure appear novel. In this paper, we incorporate a subject's genetic signature, at key loci, into a multilocus model. We hypothesize that this would help predict brain integrity, as measured by DTI-derived FA, more powerfully than a single-locus genetic test. We focus on the corpus callosum, as it is the largest white matter structure in the brain, easy to examine at the brain's midline, highly heritable (Chiang *et al*, 2009; Brouwer *et al*, 2010; Kochunov *et al*, 2010), and well studied in neurology and psychiatry as the primary commissure connecting the two brain hemispheres (Foong *et al*, 2000; Alexander *et al*, 2007). We chose FA as the DTI measure of white matter structure, as it has been shown to have higher heritability than other DTI parameters, such as radial and axial diffusivity (Kochunov *et al*, 2010).

MATERIALS AND METHODS

Participants

A total of 395 subjects (23.7 ± 2.2 years of age; 143 men and 252 women; 47 siblings, 141 monozygotic twins (49 pairs and 43 singletons), and 207 dizygotic twins (1 triplet, 70 pairs, and 64 singletons) from the Brisbane young adult twins and siblings study (de Zubicaray *et al*, 2008) were included in our study, for whom both 105-gradient DTI scans and genome-wide genotype information were available. All twins in this study are Australians of European descent. Previously, principal component analysis was conducted in this cohort for population stratification analysis and correction (Medland *et al*, 2009). Subjects who were >6 SD from either of the top two average reference principal component scores—derived from non-Australian European populations—were identified as ancestry outliers and excluded from analysis. The first two principal components refer to differences between Africans and non-Africans and to differences between East Asians and others, respectively. Owing to migration patterns and the fact that this sample was originally recruited to study mole patterns, exclusions are usually because of Asian or Polynesian ancestry. All subjects were screened to exclude cases of pathology known to affect brain structure. Additionally, no subjects had a first-degree relative with a psychiatric disorder or reported a history of significant head injury, a neurological or psychiatric illness, substance abuse or dependence.

Diffusion Tensor Imaging

Whole-brain diffusion tensor MRI scans were collected with a 4-tesla Bruker Medspec MRI scanner. Images were acquired using single-shot echo planar imaging with a twice-refocused spin echo sequence to reduce eddy current-induced distortions. Acquisition parameters were optimized to yield the best signal-to-noise ratio for estimation of diffusion tensors (Jones *et al*, 1999). Imaging parameters were: 23 cm field-of-view, TR/TE 6090/91.7 ms, with a 128×128 acquisition matrix. 105 images were acquired for each subject: 11 with no diffusion sensitization and 94 diffusion-weighted images with gradient directions evenly distributed on the hemisphere. Standard protocols for skull-stripping and eddy current distortion correction were performed using FSL (http://www.fmrib.ox.ac.uk/fsl) and we adjusted for echo planar imaging distortions as detailed in prior studies (Leow *et al*, 2005; Jahanshad *et al*, 2010). FSL was also used to calculate tensors and scalar maps of FA from the corrected images. The LONI pipeline (http://pipeline. loni.ucla.edu) was used to parallelize the preprocessing steps.

A mean deformation template (MDT) was created for the DTI scans, to which subjects' FA maps (obtained from DWI elastically aligned to their high resolution T1-weighted anatomical scan) were registered as in Jahanshad *et al* (2010), using a 3D elastic warping technique with a mutual information cost function (Leow *et al*, 2005). The MDT and the registered FA maps were then thresholded at 0.25, as FA measures below this threshold may reflect contributions from non-white matter in healthy-appearing white matter. After registering the FA maps across subjects, all FA images were smoothed with a Gaussian filter with a 7-mm isotropic full-width at half-max (FWHM). The structure of the corpus callosum was identified automatically by using the Johns Hopkins University (JHU) white matter atlas (ICBM DTI 81; Mori *et al*, 2008), which tracks its 3D extent, extending laterally from the midline (Figure 1). The atlas FA image was linearly and then elastically registered to our study-specific

FA-MDT; the transformation matrix and deformation map were then applied to the JHU set of labels using nearest-neighbor interpolation to avoid intermixing of labels. The full label (composed of three regions: splenium, body, and genu) of the corpus callosum was then accurately extracted. This avoided subjectivity and rater dependency in defining the limits of the corpus callosum.

Genotyping and Selection of Candidate Single-Nucleotide Polymorphisms (SNPs)

We considered six candidate SNPs listed in Table 1 owing to the recent imaging genetics discoveries outlined in the introduction. These particular genetic variants are located in six different genes. All have been linked to structural differences detectable with DTI. Several of these polymorphisms (ie, rs6265 in *BDNF*, rs6336 in *NTRK1*, rs4680 in *COMT,* and rs1799945 in *HFE*) are exonic variants and lead to amino-acid changes in the protein products of these genes (val→met, his→tyr, val→met, and his→asp, respectively). These have been well studied in the neuropsychiatric literature (Egan *et al*, 2003; Zecca *et al*, 2004; Tunbridge *et al*, 2006; van Schijndel *et al*, 2011). The remaining candidate SNPs do not cause missense mutations, but have been discovered in genome-wide association and genetic risk studies of neuropsychiatric disease (Lambert *et al*, 2009; Silberberg *et al*, 2006; Konrad *et al*, 2009). To obtain genotype information, genomic DNA samples were analyzed on the Human610-Quad BeadChip (Illumina, San Diego, California, USA) according to the manufacturer's protocols (Infinium HD Assay; Super Protocol Guide; Rev. A, May 2008). Additionally, imputation was performed by mapping the genotyped information to HapMap (Release 22 Build 36) using the Mach software (http://www.sph.umich.edu/csg/abecasis/MACH/index.html). All candidate SNPs passed a platform-specific quality control score (>0.7) and genotype call rate (>0.95).

Statistical Analysis

Linear mixed-effects models were used to study the joint and individual associations of genotypes with imaging measures, to take into account the relatedness between the subjects. For n subjects and p independent predictors (SNPs or other covariates), regression coefficients (β) were obtained, using the efficient mixed-model association (EMMA) software with restricted maximum likelihood estimation (Kang *et al*,

2008), according to the formula:

$$y = X\beta + Zb + \varepsilon$$

Here, y represents an n-component vector of voxelwise or mean FA measures, X is a matrix of SNP genotypes (coded additively as 0, 1, or 2 for the number of minor alleles) and/or covariates (eg, sex and age), Z is the identity matrix, and b is a vector of random effects with a variance of $\sigma_g^2 K$, where K is the n by n kinship matrix for the twins and siblings (here, monozygotic twins are coded as 1, dizygotic twins and siblings as 0.5, and unrelated subjects as 0, corresponding to the expected proportion of their shared genetic polymorphisms, respectively). ε is a matrix of residual effects with a variance of $\sigma_e^2 I$, and I is an identity matrix. P-values for the significance of individual and joint SNP associations with FA were assessed using an F-test, according to the formula,

$$F = \frac{(RSS_{\text{covariates}} - RSS_{\text{full}})/(p_{\text{full}} - p_{\text{covariates}})}{RSS_{\text{full}}/(n - p_{\text{full}})}$$

where RSS represents the residual sum-of-squares, a reduced model includes only covariates, and a full model contains both

Table 1 Associations of Single SNPs with Mean Callosal FA

SNP	Gene	Chromosome	MAF	HWE (*p*-value)	B (*p*-value)
rs6336	*NTRK1*	1	0.04	0.062	−0.015 (**0.0066**)
rs11136000	*CLU*	8	0.45	0.072	0.0051 (**0.037**)
rs839523	*ErbB4*	2	0.30	0.42	0.0053 (**0.036**)
rs4680	*COMT*	22	0.48	0.50	0.0039 (0.11)
rs1799945	*HFE*	6	0.16	0.10	0.0032 (0.29)
rs6265	*BDNF*	11	0.21	0.085	0.00031 (0.91)

Abbreviations: FA, fractional anisotropy; HWE, Hardy–Weinberg equilibrium; MAF, minor allele frequencies; SNP, single-nucleotide polymorphism. Candidate SNPs are listed with corresponding gene names and chromosome numbers. MAFs were estimated from unrelated subjects in our data set, for each genetic variant; the resulting MAF estimates were, in all cases, comparable to corresponding MAF estimates reported in the literature. *P*-values arising from tests of deviation from HWE are also reported for each SNP, none of which were significant. Linear regression coefficients, β, are also shown (with directions of effect corresponding to the number of minor alleles). *P*-values report the significance of association of each SNP with the average FA across the corpus callosum (CC). The top three genes have SNPs that are each predictive of CC integrity on their own (nominally significant *p*-values are shown in bold).
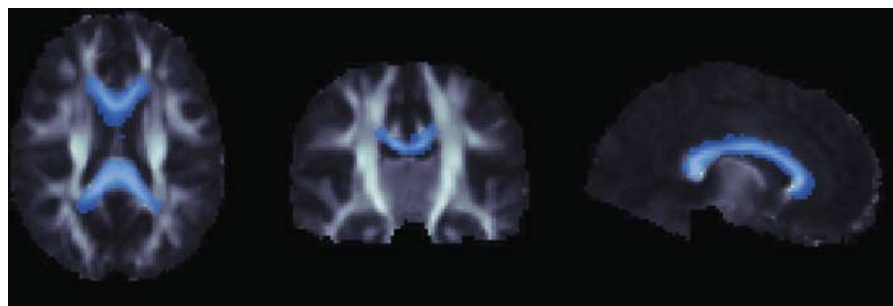

Figure 1 The three-dimensional structure of the corpus callosum, as defined by the JHU white matter atlas, is displayed in axial, coronal, and sagittal views in *blue*, overlaid on the study template.

SNPs and covariates. For all statistical analyses, the LONI pipeline (http://pipeline.loni.ucla.edu/) was used for parallelization on a multi-CPU grid computer. The standard false discovery rate (FDR) method (Benjamini and Hochberg, 1995) was used for multiple comparison correction across voxels in the corpus callosum.

RESULTS

We assessed six candidate SNPs that have recently been implicated, to varying degrees, in affecting the brain, at the gross anatomical or microstructural level. We first used linear mixed-effects models to regress each subject's genotype at each candidate SNP against average FA measures across the corpus callosum (all callosal voxels with an FA above 0.25), to study their individual effects on white matter integrity. The regression β- and p-values for each SNP on its own, treated as an independent predictor, are shown in Table 1.

We then assessed the joint effect of our set of candidate SNPs on the corpus callosum, using a partial *F*-test and linear mixed-effects model to compute p-values. When regressed against average FA across the corpus callosum, a 5-SNP model with *NTRK1, CLU, COMT, ErbB4,* and *HFE* containing top SNPs explained 5.6% of the variance in FA ($p = 0.0001$; model included sex and age; a model including only sex and age explained 0.42% of the variance). Prediction of mean callosal FA was improved by adding candidate SNPs in a stepwise fashion (Table 2). Addition of the *BDNF* SNP, however, did not improve the model. To ensure multicollinearity was not present among the genotypes, we assessed the correlation structure between the candidate SNPs, and none was correlated with any of the others (Supplementary Table S1).

We also investigated the combined influence of the candidate SNPs on more detailed, spatial maps of the corpus callosum. In a stepwise fashion, in order of the SNPs' individual effects (strongest to weakest, as shown in Table 1), we studied multilocus effects on voxel-by-voxel maps of callosal white matter structure that are shown in Table 2. The 5-SNP

model showed the most widespread, statistically significant influence on the corpus callosum, where 82% of voxels (encompassing the callosal body, genu, and splenium) survived the FDR correction for multiple comparisons across all callosal voxels, at a critical p-value threshold of 0.041 (Figure 2); we note that in FDR, a higher critical p-value denotes a stronger effect, as it is the highest threshold that controls the FDR; as such this effect is widespread and strong. Both the number of statistically significant
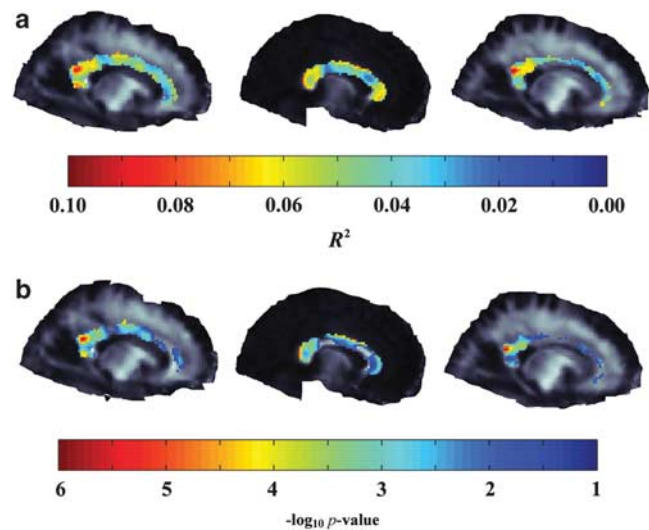

Figure 2 Voxelwise R^2 and p-values are shown in three representative sagittal slices for the joint effect of five SNPs in the *NTRK1, CLU, COMT, ErbB4,* and *HFE* genes on the corpus callosum microstructure, measured by DTI fractional anisotropy (FA). (a) The coefficient of determination (R^2) or predictability of the 5-SNP model at each voxel is shown in the selected slices. Warmer colors represent higher fractions of variance in FA explained by the multi-SNP model. (b) *P*-values are shown for the 5-SNP model at each voxel; maps are corrected for multiple comparisons across voxels by applying a critical p-value threshold to control the FDR. Warmer colors represent more significant associations (greater effect sizes). For associations at each voxel, we adjusted for any effects of sex and age, and accounted for kinship structure via mixed-effects models.

Table 2 Multilocus Effects on Average Callosal FA and 3-Dimensional Maps of the Corpus Callosum

No. of top SNPs	Mean callosal FA	3D maps of the corpus callosum			
	R^2	Critical FDR p	Voxels below critical p (%)	Voxels with $p < 0.05$ (%)	Minimum p
1	0.031	3.49×10^{-2}	4303 (70)	4709 (77)	1.49×10^{-5}
2	0.039	4.02×10^{-2}	4946 (80)	5147 (84)	8.54×10^{-7}
3	0.046	3.96×10^{-2}	4874 (79)	5080 (83)	1.23×10^{-6}
4	0.052	4.06×10^{-2}	5003 (81)	5164 (84)	1.54×10^{-6}
5	0.056	4.08×10^{-2}	5024 (82)	5184 (84)	1.25×10^{-6}
6	0.056	3.66×10^{-2}	4501 (73)	4830 (79)	4.25×10^{-6}

Abbreviations: FA, fractional anisotropy; FDR, false discovery rate; SNP, single-nucleotide polymorphism.
Groups of candidate SNPs are considered in joint association with mean callosal FA and voxelwise FA measures in the corpus callosum, in a stepwise fashion, adding SNPs in order of their individual effects (strongest to weakest). For mean FA, the fractions of variability explained by the top SNPs are shown, using the multiple regression R^2. For voxelwise results, we show the critical p-value thresholds after correcting for multiple comparisons with the FDR method, the number of voxels passing FDR at the critical thresholds (also as a percentage of all callosal voxels), the number of voxels with joint effect p-values less than $p = 0.05$ (also as a percentage of all callosal voxels) and the minimum joint effect p-values across the corpus callosum. Values are displayed for each progressively expanding set of SNP predictors. For all associations, we adjusted for sex and age at each voxel; we accounted for kinship structure via mixed-effects models.

voxels and the critical *p*-value threshold were strongest for the 5-SNP model. Figure 2 also shows the voxelwise distribution of the fraction of variance explained by the 5-SNP model in the corpus callosum.

In addition to the additive, linear model, we included two-way SNP–SNP interactions in the mean callosal FA mixed-effect model. No significant interactions were found (Supplementary Table S2). We also explored prediction of voxel-by-voxel FA from the five SNPs using two popular machine-learning models (support vector regression and artificial neural networks; see Supplementary Methods), within a cross-validation framework, which similar to the mixed-effect model, led to statistically significant predictions across the corpus callosum (Supplementary Figure S1). At each voxel, mean-squared errors of predictions of FA were obtained from the candidate genotypes. The artificial neural network and support vector regression models' predictive errors were then compared with those of null predictors (ie, where FA is randomly assigned to subjects) through permutations. The artificial neural network and support vector regression learning models were found to be statistically significant across 75% and 40% of the corpus callosum voxels, after correcting for multiple comparisons, with critical *p*-value thresholds of 0.037 and 0.019, respectively.

DISCUSSION

In this work, we aimed to predict neuroanatomical white matter micro-structure based on multiple genetic risk factors, while covarying for sex and age. Five of the six candidate polymorphisms that we considered in the study—*CLU* (Braskie *et al*, 2011), *ErbB4* (Konrad *et al*, 2009), *NTRK1* (Braskie *et al*, 2012), *COMT* (Thomason *et al*, 2010), and *HFE* (Jahanshad *et al*, 2012a)—explained close to 6% of the variability in mean callosal FA, using a linear mixed-effect model. This is a considerable fraction of the variance explained by only a few SNPs, taking into account the complexity of the structure and the non-genetic factors that influence it. It is also comparable to previous findings in the literature for multilocus models of a brain-imaging phenotype. Biffi *et al* (2010) found that 3% of the variance in MRI-derived volumes of several brain regions could be explained from a number of candidate genes for Alzheimer's disease. Nikolova *et al* (2011) showed 11% of the variance in ventral striatal reactivity could be explained from their panel of five polymorphisms. We also found that our candidate polymorphisms displayed extensive, significant effects on 82% of the volume of the corpus callosum, when cumulatively modeled at a voxelwise FA basis, which captures more spatial detail than an average measure of FA across the corpus callosum. We also confirmed significant predictions across the corpus callosum from the five SNPs using multilocus machine-learning models. These yielded similar predictions, but were less spatially widespread, as only a subset of subjects were considered who were unrelated to each other.

We focused on DTI-derived FA of the corpus callosum as our imaging measure in this study. The corpus callosum is the largest white matter structure in the brain, containing over 300 million axons (Hofer and Frahm, 2006). This fiber bundle transfers motor, sensory, and cognitive information between the two cerebral hemispheres (Huang *et al*, 2005). With the advent of DTI, it has become increasingly clear that the structure of the corpus callosum is impaired in several brain disorders. In a recent meta-analysis, for instance, Patel *et al* (2011) concluded that the splenium of the corpus callosum has significantly lower FA in patients with schizophrenia *vs* controls. Recent DTI studies have also identified callosal abnormalities in patients with other brain disorders such as bipolar disorder (Benedetti *et al*, 2011), post-traumatic stress disorder (Jackowski *et al*, 2008), and autism (Alexander *et al*, 2007). It would therefore be beneficial, clinically, to know an individual's personalized genetic risk for a corpus callosum structural abnormality. In addition, the microstructure of the corpus callosum has been shown to be highly heritable in studies including those with the same Australian twins as in this paper. Chiang *et al* (2009) mapped out genetic contributions to white matter structure in the Australian twins and discovered significant voxelwise effects in the callosal genu and splenium. In that paper, a classical twin design was used to estimate the overall genetic contribution to the observed variance, but effects of specific SNPs were not assessed or modeled. Similarly, Kochunov *et al* (2010) found mean FA values from the corpus, body, and genu of the corpus callosum were highly heritable (all with $h^2 > 0.5$) in members of the San Antonio Family Study. Similar results have also been reported in studies of young children (Brouwer *et al*, 2010) and in older individuals (eg, Pfefferbaum *et al*, 2001). Recently, it was also shown in the same Australian population as ours that the heritability of callosal FA, particularly in the genu, is high regardless of imaging protocol differences (Jahanshad *et al*, 2012b). Here, we show that predictions of microstructural measures may be made based on a few common polymorphisms. We focused on the corpus callosum here, but our results may also have implications for other white matter tracts in the brain. The millions of axons in the corpus callosum connect numerous regions of the brain with each other. Genetic variants that affect this brain structure may also have roles in other white matter regions.

We selected six candidate SNPs for our study based on their reported individual effects on white matter structure on DTI and their importance in neuropsychiatric disease. The val158met missense mutation resulting from the candidate SNP in *COMT* causes reduced degradation and thus increased availability of dopamine, thereby leading to alterations in reward experience, executive function, and working memory, with implications on risk for neuropsychiatric disease and differential response to therapy (Tunbridge *et al*, 2006; Wichers *et al*, 2008). *BDNF*'s val66met polymorphism, which affects the neurotrophin's secretion and its function in long-term potentiation, has been investigated in many studies and shown to alter memory performance at a young age, among other associations with neuropsychiatric disease (Egan *et al*, 2003; Hariri *et al*, 2003). Similarly, although not as fully characterized, the candidate SNP in the neurotrophin receptor gene, *NTRK1*, leads to a his598tyr amino-acid change in the kinase domain of TrkA, and has been significantly associated with risk for schizophrenia (van Schijndel *et al*, 2009; van Schijndel *et al*, 2011). In our study, this SNP had the strongest effect of all candidates on white matter

micro-structure. We also found subjects with greater numbers of minor alleles of the *NTRK1* polymorphism had lower FA, which is consistent with data suggesting that the minor allele is over-represented in schizophrenia patients (van Schijndel *et al*, 2011). Another receptor gene we considered was the neuregulin receptor, *ErbB4*, with an intronic variant associated with schizophrenia risk in several studies (Konrad *et al*, 2009; Nicodemus *et al*, 2006, Silberberg *et al*, 2006). Another intronic variant, rs11136000, in *CLU* has been discovered and replicated in genome-wide association studies of Alzheimer's disease (Lambert *et al*, 2009). Similarly, the his63asp mutation in iron-related *HFE* gene has been linked to Alzheimer's disease, along with other neurodegenerative disorders (Connor and Lee, 2006). We found all variants except for the one in *BDNF* contributed additively to prediction of average callosal FA as well as three-dimensional maps of voxelwise FA across the corpus callosum.

Personalized prediction of individuals' disease-related measures is being advocated by some as a vital component of future diagnosis and treatment of brain disorders (Koslow *et al*, 2010). Some of the genetic variation may account for some of the broad heterogeneity in patients' disease status (Cummings, 2000; Folstein and Rosen-Sheidley, 2001) and the extent to which they respond to therapy (Gordon, 2007). Multilocus models are particularly appealing for personalized prediction of disease. Several groups have explored multilocus models of candidate risk variants in the context of brain disorders. Carayol *et al* (2010), for instance, reported on the cumulative effect of four candidate SNPs on the risk for autism, using a case–control approach. These models are beginning to be applied to brain imaging in the context of neuropsychiatric disorders (Biffi *et al*, 2010; Hibar *et al*, 2011a; Nikolova *et al*, 2011), and may provide more biologically meaningful predictions with implications for personalized diagnosis and therapy.

Future studies are needed to replicate our findings in independent cohorts of subjects, even though we found significant predictions using cross-validation in the support vector regression and artificial neural network analyses. In addition, as new candidate gene studies and genome-wide searches using DTI measures (eg, Kochunov *et al*, 2011) identify effects of new variants, candidate genes may be added or removed from this panel, to better predict white matter integrity. We did not find evidence for two-way interactions between the SNPs in our study, which is probably reasonable, as the SNPs are likely contributing independently and additively to white matter integrity, and interactions are second-order effects (modulations of the main effect of a gene) that may require large samples to identify, if present at all. Such interactions, however, may be identified in follow-up studies particularly with SNPs that directly share the same pathway, like *NTRK2* and *BDNF* (Perroud *et al*, 2009), *NRG1* and *ErbB4* (Nicodemus *et al*, 2010), or *COMT* and *5-HTTLPR* (Borroni *et al*, 2006). In this paper, we took a voxelwise approach to study genetic associations with FA. In addition to voxelwise maps of FA, tract- and fiber-based measures from diffusion imaging may also be considered as predictive outputs. Such measures, along with multivariate methods that simultaneously consider not only multiple genes, but also multiple voxels (Vounou *et al*, 2010; Hibar *et al*, 2011b; Wan *et al*, 2011) may help provide more statistical power. For instance, our voxelwise, multilocus model improved only slightly beyond the 2-SNP model with polymorphisms in *NTRK1* and *CLU*. This may be because of the strong effects of *NTRK1* and *CLU* SNPs on their own, but it may also be because multiple variants do not necessarily affect the same exact voxels. This may make it difficult to obtain substantially more expansive voxelwise effects by adding more variants to the model. Here, we considered genetic polymorphisms as predictors, and these explained a small but significant proportion of the heritable variation in white matter structure across young, healthy individuals. Although it remains to be determined, it is plausible that a measure of white matter integrity, such as DTI-derived FA, relates to a person's lifetime risk for developing mental and neurodegenerative disorders, especially for disorders in which FA is abnormally low.

DISCLOSURE

The authors declare that, except for income received from their primary employer, no financial support or compensation has been received from any individual or corporate entity over the past 3 years of research or professional service and there are no personal financial holdings that could be perceived as constituting a potential conflict of interest. One co-author, Dr April Ho Reese, contributed to the paper when she was a PhD student at UCLA; Dr Reese now is employed at Forest Pharmaceuticals but has not worked on this project while at Forest.

REFERENCES

Alexander AL, Lee JE, Lazar M, Boudos R, DuBray MB, Oakes TR et al (2007). Diffusion tensor imaging of the corpus callosum in Autism. *Neuroimage* **34**: 61–73.

Allen SJ, Dawbarn D (2006). Clinical relevance of the neurotrophins and their receptors. *Clin Sci* **110**: 175–191.

Bartzokis G, Lu PH, Tingus K, Peters DG, Amar CP, Tishler TA et al (2011). Gender and iron genes may modify associations between brain iron and memory in healthy aging. *Neuropsychopharmacology* **36**: 1375–1384.

Benedetti F, Yeh PH, Bellani M, Radaelli D, Nicoletti MA, Poletti S et al (2011). Disruption of white matter integrity in bipolar depression as a possible structural marker of illness. *Biol Psychiatry* **69**: 309–317.

Benjamini Y, Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* **57**: 289–300.

Biffi A, Anderson CD, Desikan RS, Sabuncu M, Cortellini L, Schmansky N *et al* (2010). Genetic variation and neuroimaging measures in Alzheimer disease. *Arch Neurol* **67**: 677–685.

Binder DK, Scharfman HE (2004). Brain-derived neutrotrophic factor. *Growth Factors* **22**: 123–131.

Borroni B, Grassi M, Agosti C, Archetti S, Costanzi C, Cornali C *et al* (2006). Cumulative effect of COMT and 5-HTTLPR polymorphisms and their interaction with disease severity and comorbidities on the risk of psychosis in Alzheimer disease. *Am J Geriatr Psychiatry* **14**: 343–351.

Braskie MN, Jahanshad N, Stein JL, Barysheva M, McMahon KL, de Zubicaray GI *et al* (2011). Common Alzheimer's disease risk variant within the *CLU* gene affects white matter microstructure in young adults. *J Neurosci* **31**: 6764–6770.

Braskie MN, Jahanshad N, Stein JL, Barysheva M, Johnson K, McMahon KL *et al* (2012). Relationship of a variant in the *NTRK1* gene to white matter microstructure in young adults. *J Neurosci* (in press).

Brouwer RM, Mandl RC, Peper JS, van Baal GC, Kahn RS, Boomsma DI *et al* (2010). Heritability of DTI and MTR in nine-year-old children. *Neuroimage* **53**: 1085–1092.

Carayol J, Schellenberg GD, Tores F, Hager J, Ziegler A, Dawson G (2010). Assessing the impact of a combined analysis of four common low-risk genetic variants on autism risk. *Mol Autism* **1**: 4.

Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C *et al* (2009). Genetics of brain fiber architecture and intellectual performance. *J Neurosci* **29**: 2212–2224.

Chiang MC, Barysheva M, Toga AW, Medland SE, Hansell NK, James MR *et al* (2011a). BDNF gene effects on brain circuitry replicated in 455 twins. *Neuroimage* **55**: 448–454.

Chiang MC, McMahon KL, de Zubicaray GI, Martin NG, Hickie I, Toga AW *et al* (2011b). Genetics of white matter development: a DTI study of 705 twins and their siblings aged 12 to 29. *Neuroimage* **54**: 2308–2317.

Connor JR, Lee SY (2006). HFE mutations and Alzheimer's disease. *J Alzheimer Dis* **10**: 267–276.

Cummings JL (2000). Cognitive and behavioral heterogeneity in Alzheimer's disease: seeking the neurobiological basis. *Neurobiol Aging* **21**: 845–861.

de Geus E, Goldberg T, Boomsma DI, Posthuma D (2008). Imaging the genetics of brain structure and function. *Biol Psychol* **79**: 1–8.

de Zubicaray GI, Chiang MC, McMahon KL, Shattuck DW, Toga AW, Martin NG *et al* (2008). Meeting the challenges of neuroimaging genetics. *Brain Imaging Behav* **2**: 258–263.

DeMattos RB, O'Dell MA, Parsadanian M, Taylor JW, Harmony JA, Bales KR (2002). Clusterin promotes amyloid plaque formation and is critical for neuritic toxicity in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci USA* **99**: 10843–10848.

Egan MF, Weinberger DR, Lu B (2003). Schizophrenia, III: brain-derived neurotropic factor and genetic risk. *Am J Psychiatry* **160**: 1242.

Ellison-Wright I, Bullmore E (2009). Meta-analysis of diffusion tensor imaging studies in schizophrenia. *Schizophr Res* **108**: 3–10.

Fellgiebel A, Wille P, Müller MJ, Winterer G, Scheurich A, Vucurevic G *et al* (2004). Ultrastructural hippocampal and white matter alterations in mild cognitive impairment: a diffusion tensor imaging study. *Dement Geriatr Cogn Disord* **18**: 101–108.

Folstein SE, Rosen-Sheidley B (2001). Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet* **2**: 943–955.

Foong J, Maier M, Clark CA, Barker GJ, Miller DH, Ron MA (2000). Neuropathological abnormalities of the corpus callosum in schizophrenia: a diffusion tensor imaging study. *J Neurol Neurosurg Psychiatry* **68**: 242–244.

Gordon E (2007). Integrating genomics and neuromarkers for the era of brain-related personalized medicine. *Pers Med* **4**: 201–215.

Hahn CG, Wang HY, Cho DS, Talbot K, Gur RE, Berrettini WH *et al* (2006). Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nat Med* **12**: 824–828.

Hall MH, Smoller JW (2010). A new role for endophenotypes in the GWAS era: functional characterization of risk variants. *Harv Rev Psychiatry* **18**: 67–74.

Hariri AR, Goldberg TE, Mattay VS, Kolachana BS, Callicott JH, Egan MF *et al* (2003). Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance. *J Neurosci* **23**: 6690–6694.

Heng S, Song AW, Sim K (2010). White matter abnormalities in bipolar disorder: insights from diffusion tensor imaging studies. *J Neural Transm* **117**: 639–654.

Hibar DP, Kohannim O, Stein JL, Chiang MC, Thompson PM (2011b). Multilocus genetic analysis of brain images. *Front Genet* **2**. doi:10.3389/fgene.2011.00073.

Hibar DP, Stein JL, Kohannim O, Jahanshad N, Saykin AJ, Shen L *et al* (2011a). Voxelwise gene-wide association study (vGene-WAS): multivariate gene-based association testing in 731 elderly subjects. *Neuroimage* **56**: 1875–1891.

Hofer S, Frahm J (2006). Topography of the human corpus callosum revisited–comprehensive fiber tractography using diffusion tensor magnetic resonance imaging. *Neuroimage* **32**: 989–994.

Huang H, Zhang J, Jiang H, Wakana S, Poetscher L, Miller MI *et al* (2005). DTI tractography based parcellation of white matter: application to the mid-sagittal morphology of corpus callosum. *Neuroimage* **26**: 195–205.

Jackowski AP, Douglas-Palumberi H, Jackowski M, Win L, Schultz RT, Staib LW *et al* (2008). Corpus callosum in maltreated children with posttraumatic stress disorder: a diffusion tensor imaging study. *Psychiatry Res* **162**: 256–261.

Jahanshad N, Kohannim O, Hibar DP, Stein JL, McMahon KL, de Zubicaray GI *et al* (2012a). Brain structure in healthy adults is related to serum transferrin and the H63D polymorphism in the HFE gene. *Proc Natl Acad Sci USA* **109**: E851–E859.

Jahanshad N, Kohannim O, Toga AW, McMahon KL, de Zubicaray GI, Martin NG *et al* (2012b). Diffusion imaging protocol effects on genetic associations. *ISBI ' 12. IEEE International Symposium on Biomedical Imaging*. Barcelona.

Jahanshad N, Lee AD, Barysheva M, McMahon KL, de Zubicaray GI, Martin NG *et al* (2010). Genetic influences on brain asymmetry: a DTI study of 374 twins and siblings. *Neuroimage* **52**: 455–469.

Jones DK, Horsfield MA, Simmons A (1999). Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn Reson Med* **42**: 515–525.

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ *et al* (2008). Efficient control of population structure in model organism association mapping. *Genetics* **178**: 1709–1723.

Ke X, Tang T, Hong S, Hang Y, Zou B, Li H *et al* (2009). White matter impairments in autism, evidence from voxel-based morphometry and diffusion tensor imaging. *Brain Res* **1265**: 171–177.

Kochunov P, Glahn DC, Lancaster JL, Winkler AM, Smith S, Thompson PM *et al* (2010). Genetics of microstructure of cerebral white matter using diffusion tensor imaging. *Neuroimage* **53**: 1109–1116.

Kochunov P, Glahn DC, Nichols TE, Winkler AM, Hong EL, Holcomb HH *et al* (2011). Genetic analysis of cortical thickness and fractional anisotropy of water diffusion in the brain. *Front Neurosci* **5**: 120.

Konrad A, Vucurevic G, Musso F, Stoeter P, Dahmen N, Winterer G (2009). ErbB4 genotype predicts left frontotemporal structural

connectivity in human brain. *Neuropsychopharmacology* **34**: 641–650.

Konrad K, Eickhoff SB (2010). Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. *Hum Brain Mapp* **31**: 904–916.

Koslow SH, Williams LM, Gordon E (2010). Personalized medicine for the brain: a call for action. *Mol Psychiatry* **15**: 229–230.

Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M et al (2009). Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* **41**: 1094–1099.

Lee AD, Lepore N, Barysheva M, Chou Y-Y, Brun CC, Madsen SK et al (2008). Gene effects mapped using fractional and geodesic anisotropy in diffusion tensor images of 92 monozygotic and dizygotic twins. In: Alexander D, Gee J, Whitaker R (eds). *MICCAI Workshop on Computational Diffusion MRI*. Springer: New York, pp 31–40.

Lee AD, Lepore N, de Leeuw J, Brun CC, Barysheva M, McMahon KL et al (2010). Multivariate variance-components analysis in DTI. In: *ISBI '10. IEEE International Symposium on Biomedical Imaging*. IEEE Press: Rotterdam, pp 157–1160.

Leow A, Huang SC, Geng A, Becker J, Davis S, Toga A et al (2005). Inverse consistent mapping in 3D deformable image registration: its construction and statistical properties. *Inf Process Med Imaging* **19**: 493–503.

Marenco S, Radulescu E (2010). Imaging genetics of structural brain connectivity and neural integrity markers. *Neuroimage* **53**: 848–856.

Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G et al (2009). Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* **85**: 750–755.

Meyer-Lindenberg A, Weinberger DR (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci* **7**: 818–827.

Mori S, Oishi K, Jiang H, Jiang L, Li X, Akhter K et al (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage* **40**: 570–582.

Naggara O, Oppenheim C, Rieu D, Raoux N, Rodrigo S, Dalla Barba G et al (2006). Diffusion tensor imaging in early Alzheimer's disease. *Psychiatry Res* **146**: 243–249.

Nicodemus KK, Law AJ, Radulescu E, Luna A, Kolachana B, Vakkalanka R et al (2010). Biological validation of increased schizophrenia risk with NRG1, ERBB4, and AKT1 epistasis via functional neuroimaging in healthy controls. *Arch Gen Psychiatry* **67**: 991–1001.

Nicodemus KK, Luna A, Vakkalanka R, Goldberg T, Egan M, Straub RE et al (2006). Further evidence for association between ErbB4 and schizophrenia and influence on cognitive intermediate phenotypes in healthy controls. *Mol Psychiatry* **11**: 1062–1065.

Nikolova YS, Ferrell RE, Manuck SB, Hariri AR (2011). Multilocus genetic profile for dopamine signaling predicts ventral striatum reactivity. *Neuropsychopharmacology* **36**: 1940–1947.

Oishi K, Mielke MM, Albert M, Kyketsos CG (2011). DTI analyses and clinical applications in Alzheimer's disease. *J Alzheimers Dis* **26**(Suppl): 287–296.

Patel S, Mahon K, Wellington R, Zhang J, Chaplin W, Szeszko PR (2011). A meta-analysis of diffusion tensor imaging studies of the corpus callosum in schizophrenia. *Schizophr Res* **129**: 149–155.

Patel V, Chiang M-C, Thompson PM, McMahon KL, de Zubicaray GI, Martin NG et al (2010). Scalar connectivity measures from fast marching tractography reveal heritability of white matter architecture. In: *ISBI '10. IEEE International Symposium on Biomedical Imaging*. IEEE Press: Rotterdam, pp 1109–1112.

Perroud N, Aitchison KJ, Uher R, Smith R, Huezo-Diaz P, Marusic A et al (2009). Genetic predictors of increase in suicidal ideation during antidepressant treatment in the GENDEP project. *Neuropsychopharmacology* **34**: 2517–2528.

Pfefferbaum A, Sullivan EV, Carmelli D (2001). Genetic regulation of regional microstructure of the corpus callosum in late life. *Neuroreport* **12**: 1677–1681.

Silberberg G, Darvasi A, Pinkas-Kramarski R, Navon R (2006). The involvement of ErbB4 with schizophrenia: association and expression studies. *Am J Med Genet B Neuropsychiatr Genet* **141**: 142–148.

Sussmann JE, Lymer GK, McKirdy J, Moorhead TW, Muñoz Maniega S, Job D et al (2009). White matter abnormalities in bipolar disorder and schizophrenia detected using diffusion tensor magnetic resonance imaging. *Bipolar Disord* **11**: 11–18.

Thomason ME, Dougherty RF, Colich NL, Perry LM, Rykhlevskaia EI, Louro HM et al (2010). COMT genotype affects prefrontal white matter pathways in children and adolescents. *Neuroimage* **53**: 926–934.

Thomason ME, Thompson PM (2011). Diffusion imaging, white matter, and psychopathology. *Annu Rev Clin Psychol* **7**: 63–85.

Tunbridge EM, Harrison PJ, Weinberger DR (2006). Catechol-o-methyltransferase, cognition, and psychosis: Val158Met and beyond. *Biol Psychiatry* **60**: 141–151.

van Schijndel JE, van Loo KM, van Zweeden M, Djurovic S, Andreassen OA, Hansen T et al (2009). Three-cohort targeted gene screening reveals a non-synonymous TRKA polymorphism associated with schizophrenia. *J Psychiatr Res* **43**: 1195–1199.

van Schijndel JE, Van Zweeden M, Van Loo KM, Djurovic S, Andreassen OA, Hansen T et al (2011). Dual association of a TRKA polymorphism with schizophrenia. *Psychiatr Genet* **21**: 125–131.

Versace A, Almeida JR, Hassel S, Walsh ND, Novelli M (2008). Elevated left and reduced right orbitomedial prefrontal fractional anisotropy in adults with bipolar disorder revealed by tract-based spatial statistics. *Arch Gen Psychiatry* **65**: 1041–1052.

Vounou M, Nichols TE, Montana G, For the Alzheimer's disease neuroimaging initiative (2010). Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. *Neuroimage* **53**: 1147–1159.

Wan J, Kim S, Inlow M, Nho K, Swaminathan S, Risacheri SL et al (2011). Hippocampal surface mapping of genetic risk factors in AD via sparse learning models. *Med Image Comput Comput Assist Interv* **14**: 376–383.

White T, Nelson M, Lim KO (2008). Diffusion tensor imaging in psychiatric disorders. *Top Magn Reson Imaging* **19**: 97–109.

Wichers M, Aguilera M, Kenis G, Krabbendam L, Myin-Germeys I, Jacobs N et al (2008). The catechol-O-methyl transferase Val158Met polymorphism and experience of reward in the flow of daily life. *Neuropsychopharmacology* **33**: 3030–3036.

Yoo SY, Jang JH, Shin YW, Kim DJ, Park HJ (2007). White matter abnormalities in drug-naive patients with obsessive-compulsive disorder: a diffusion tensor study before and after citalopram treatment. *Acta Psychiatr Scand* **116**: 211–219.

Zecca L, Youdim MB, Riederer P, Connor JR, Crichton RR (2004). Iron, brain ageing and neurodegenerative disorders. *Nat Rev Neurosci* **5**: 863–873.