

GENETIC ANALYSIS OF STRUCTURAL BRAIN CONNECTIVITY USING DICCCOL MODELS OF DIFFUSION MRI IN 522 TWINS

Dajiang Zhu¹, Liang Zhan¹, Joshua Faskowitz¹, Madelaine Daianu¹, Neda Jahanshad¹, Greig I. de Zubicaray², Katie L. McMahon³, Nicholas G. Martin⁴, Margaret J. Wright⁴, Paul M. Thompson¹

¹Imaging Genetics Center, University of Southern California, CA, USA

²School of Psychology, University of Queensland, Brisbane, Australia

³Center for Advanced Imaging, University of Queensland, Brisbane, Australia

⁴QIMR Berghofer Medical Research Institute, Brisbane, Australia

ABSTRACT

Genetic and environmental factors affect white matter connectivity in the normal brain, and they also influence diseases in which brain connectivity is altered. Little is known about genetic influences on brain connectivity, despite wide variations in the brain's neural pathways. Here we applied the "DICCCOL" framework to analyze structural connectivity, in 261 twin pairs (522 participants, mean age: 21.8 y \pm 2.7SD). We encoded connectivity patterns by projecting the white matter (WM) bundles of all "DICCCOLs" as a tracemap (TM). Next we fitted an A/C/E structural equation model to estimate additive genetic (A), common environmental (C), and unique environmental/error (E) components of the observed variations in brain connectivity. We found 44 "heritable DICCCOLs" whose connectivity was genetically influenced ($a^2 > 1\%$); half of them showed significant heritability ($a^2 > 20\%$). Our analysis of genetic influences on WM structural connectivity suggests high heritability for some WM projection patterns, yielding new targets for genome-wide association studies.

Index Terms—DICCCOL, heritability

1. INTRODUCTION

Genetic and environmental factors contribute to individual differences in brain connectivity, and understanding their contributions is of fundamental importance in neuroscience. For two decades, a large number of studies [2-3, 5-10] have shown moderately high heritability for many brain measures derived from magnetic resonance imaging (MRI), diffusion tensor imaging (DTI) [1], and other neuroimaging methods. Early genetic studies of brain MRI examined gray matter (GM) similarities in pairs of monozygotic (MZ) and dizygotic (DZ) twins [2], and white matter (WM) integrity in the corpus callosum [3]. More recently, genome-wide association studies of over 30,000 individuals [16] have revealed common genetic variants that consistently affect

brain structure worldwide. Similar large-scale genetic analyses are also now focusing on DTI [5], relating imaging differences to clinical diagnosis for a range of neuropsychiatric diseases [17-19]. Even so it is crucial to discover which brain measures show significant heritability, to prioritize them for in-depth genetic screens. Some studies have shown moderate to strong genetic effects on DTI derived measures such as FA [5, 10], but we still know little about genetic influences on the structural connectivity of the brain's major neural pathways, which we examine here.

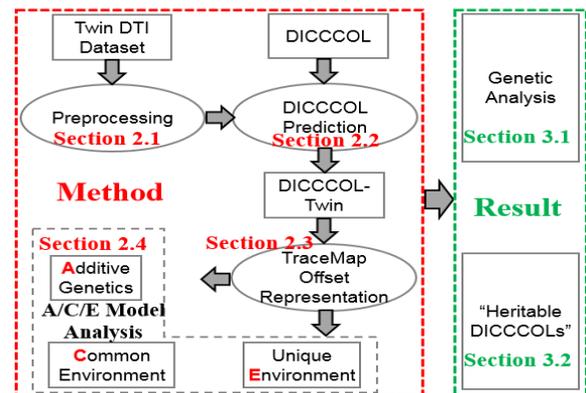


Fig.1. Flowchart of the proposed study.

In prior work, we introduced the "tracemap" (TM) model [11] and DICCCOL [12] - novel methods to represent and discover WM connectivity. The tracemap is a computational model that transforms the directional information in the trajectory of an arbitrary WM bundle to a standard spherical surface, for quantitative comparison of structural connectivity patterns. The DICCCOL method identifies a group of cortical surface landmarks in each individual's coordinate space, and a correspondence across different individuals based on WM structural similarity, measured by TM distance. Here we combine the TM approach with the DICCCOL model, to investigate genetic influences on WM pathways in 522 twins. As illustrated in Fig. 1, after preprocessing (section 2.1), 358 DICCCOLs

were identified in the twin dataset (section 2.2). Then we calculated the TM offset (section 2.3), defined as the difference between individual TM and the model TM. Finally, we performed an ACE analysis [21] (section 2.4) to estimate the additive genetic (A), common environmental (C), and unique environmental (E) components of variance for the computed TM measures. We detected 44 “heritable DICCCOLs” displaying underlying genetic influences ($a^2 > 1\%$, $p < 0.037$ after FDR). These DICCCOLs were spread over the whole cortical surface, and included landmarks in the superior, middle and inferior frontal gyri, orbital gyri, sensorimotor areas, temporal lobe and visual cortex. Half of them showed significant heritability ($a^2 > 20\%$) and these were primarily located in the temporal lobes and visual cortices. WM tracts connecting to the “heritable DICCCOLs” with the largest a^2 came from the inferior fronto-occipital fasciculus (IFOF). This study offers a new approach to explore genetic influences on brain connectivity, by considering the connection patterns of WM pathways. It also provides promising targets for genome-wide association analyses.

2. METHODS

2.1. Data acquisition

We analyzed brain DTI data from 261 twin pairs (522 participants, M/F: 196/326, MZ/DZ: 214/308, mean age: $21.8 \text{ y} \pm 2.7\text{SD}$). Diffusion-weighted images were collected with a 4T Siemens Bruker Medspec MRI scanner, using single-shot echo planar imaging with a twice-refocused spin echo sequence to reduce eddy-current induced distortions. Scan parameters were: 23 cm FOV, TR/TE 6090/91.7 ms, 2-mm axial slice thickness and $1.79 \times 1.79 \text{ mm}^2$ in-plane resolution. More information on acquisition and preprocessing may be found in [6, 12].

2.2. DICCCOL prediction

DICCCOL is an acronym that stands for Dense Individualized and Common Connectivity-based Cortical Landmarks [12]. It includes 358 cortical landmarks defined in the individual space; each represents the locations that share the most consistent WM wiring patterns across different individuals. Several studies [13-15] have shown that this is an effective and robust ROI modeling framework, which may offer advantages compared to other registration methods [12]. DICCCOL prediction is very similar to the DICCCOL optimization procedure [12] as the cost function also aims to minimize the group-wise variation in structural connectivity, measured by TM [11]. The difference is that we only optimize the locations of DICCCOL candidates that need to be predicted in individual subject. This process may be summarized as:

$$E(S_M, S_N) = \min(\sum |TMD_{MN}|) \quad (1)$$

S_M represents the DICCCOL models, S_N is the new brain that needs to be predicted, TMD_{MN} is defined as the tracemap distance. We applied the above prediction to all of the 522 twin participants; each of the predicted DICCCOLs yields an automatically computed correspondence across all subjects.

2.3. TraceMap offset representation

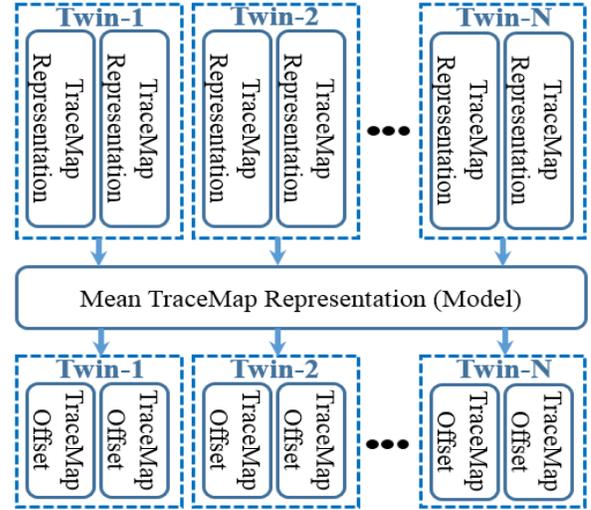


Fig.2. Illustration of tracemap offset for single DICCCOL.

After we computed DICCCOLs (locations) on every twin participant, we extracted the WM bundle connecting to each DICCCOL and transformed it as a TM, which may be represented as a vector describing a specific points-distribution. Then we calculated the mean TM (\overline{TM}_{model}) from the model and computed the offset (difference) between twin participants and \overline{TM}_{model} (Fig. 2). The TM offset (TMO) can be defined as:

$$TMO_{s,d} = ||TM_{s,d} - \overline{TM}_{model,d}||_2^2 \quad (2)$$

s and d are the indices of the participant and DICCCOLs. The calculation of TMO is conducted for each DICCCOL independently. Note that if two participants have similar WM wiring patterns for some DICCCOLs, their TMO will also be similar.

Compared to other DTI-derived scalar measures, such as FA and MD (mean diffusivity), the TM representation can effectively capture the overall wiring patterns in the WM. At the same time, it is not over-sensitive to small variations or noise introduced by tractography errors (i.e., false positive fibers). Hence, it is a promising approach to characterize brain WM connection patterns for genetic analysis.

2.4. Genetic analysis

The classical A/C/E model assumes that the variance in a brain measure (F) can be partitioned into three components: variance due to additive genetic factors (A), shared

environment (C), and unshared environment (E) or measurement errors:

$$F = aA + cC + eE \quad (3)$$

In the above formula, a , c and e are the weights of each factor and the total variance of F ($a^2+c^2+e^2$) should be one. As we know, MZ twins share all the same genetic variants and the DZ twins, on average, share 50% of their genetic polymorphisms. Both MZ and DZ twins share the same environment (C). Because of this, we can estimate how much additive genetic factors (A) contribute to the variance in the observed variable (F).

The full A/C/E structural equation model, or its simpler sub models (i.e., AE or CE) can be preferred based on their relative goodness-of-fit. Here we modelled the TMO measures (section 2.3) using the full A/C/E model to estimate genetic effects.

Prior studies have shown genetic influences on the volumes of GM [2] and WM structures [5, 10] in the brain. Here, DICCCOL derived TMO measures provide a novel approach to study WM heritability in two respects: 1) the TMO reflects the similarity of WM wiring patterns in 3D space; and 2) DICCCOLs are defined on cortical surface, but they are closely related to the underlying WM structures. Therefore, DICCCOL derived measures (i.e., TMO) have the potential to link the GM and WM features together.

3. RESULTS

3.1. Genetic analysis

We predicted 358 DICCCOLs (Section 2.2) on all the 512 twin participants. For each pair of twins we calculated the tracemap offset (TMO) for each DICCCOL with Eq. 2 (Section 2.3). Then we performed the A/C/E variance components analysis on the TMO. The classical ACE model estimates the proportion of different factors contributing to the observed phenotype measure (here, TMO). We compared the full model (A/C/E) to the simpler AE model, as in [5, 10, 20]. At almost 1/3 of the DICCCOLs (100+), the ACE model showed a better fit than the AE model ($p < 0.05$). That means for those 100+ DICCCOLs, the shared environmental (C) factor was detectable. As such, we chose the full A/C/E model to report our results. To adjust for multiple hypothesis testing, all p-values used to screen DICCCOLs were corrected using the false discovery rate (FDR) [22] (0.037 after correction) at the 5% significance level. We also considered the goodness-of-fit of A/C/E to all DICCCOLs: in structural equation models (SEM), the χ^2 (goodness-of-fit) term determines a significance value, which indicates whether the model is a good fit to the data. Hence, we screened all DICCCOLs and only those with good fit ($P > 0.05$) were considered. In total, 44 DICCCOLs were found to be heritable ($a^2 > 1\%$, 95% CI (0-0.03, 0.33-0.72)), and these are subsequently referred to as “heritable DICCCOLs”.

Fig. 3 summarizes the results of the A/C/E analysis. The proportions of A/C/E, a^2 , c^2 and e^2 , for 44 “heritable DICCCOLs” are shown in **Fig. 3(a)** as red, orange and blue columns, respectively. Even though unshared environment (E) has significant contributions to many “heritable DICCCOLs”, the additive genetic factor (A) and shared environment (C) also play critical roles. We also showed the a^2 and the corresponding 95% confidence interval (CI) as the white triangle and shadow in **Fig. 3(b)**. No CI included zero, but for some of them the lower bounds were close to zero. The distribution of a^2 is shown in **Fig. 3(c)**. From the distribution, we can see that the a^2 for half of the “heritable DICCCOLs” is larger than 0.2, so they are substantially influenced by genetic factors.

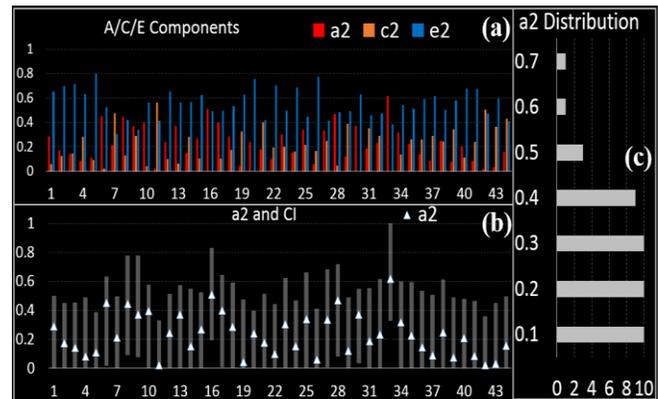


Fig.3. (a) A/C/E components of 44 “heritable DICCCOLs”. (b) The white triangles and the ranges shown represent a^2 and the corresponding 95% confidence interval (CI). (c) The distribution of a^2 for “heritable DICCCOLs”.

3.2. Visualization of “Heritable DICCCOLs”

As mentioned already, we identified 44 “heritable DICCCOLs” that were influenced by genetic factors ($a^2 > 0.1$). We projected them onto the cortical surface as shown in **Fig. 4(a)**. The bubbles represent “heritable DICCCOLs” and the colors show the value of a^2 . In general, these “heritable DICCCOLs” are spread over the whole cortical surface, including superior, middle and inferior frontal gyri, orbital gyri, sensorimotor areas, temporal lobes and the visual cortex. This result is consistent with prior reports that volumes of frontal, sensorimotor and temporal brain regions are significantly influenced by genetic factors [2]. From the color information, we can see that the “heritable DICCCOL” with the highest a^2 is located on the middle frontal gyrus (DICCCOL# 255, highlighted by the black arrow) and this DICCCOL is part of the emotional regulation network [12]. Other “heritable DICCCOLs” with relatively high a^2 are mainly located in the temporal lobe and visual cortices (*highlighted with yellow circles*).

To examine the underlying WM structures, we extracted the WM tracts connecting to these “heritable DICCCOLs”. In **Fig. 4(b)**, we show fiber tracts according to

the different a^2 for the “heritable DICCCOLs” that they are connected to. For example, the left subfigure shows WM tracts connecting to the “heritable DICCCOLs” with a^2 larger than 0.1. The middle WM tracts belong to those with a^2 larger than 0.3. The “heritable DICCCOLs” are also displayed accordingly. From this figure, we can clearly identify different WM tracts based on the hierarchy of “heritable DICCCOLs”. The final WM tracts are connected to the “heritable DICCCOLs” with the largest a^2 (>0.5). Most of the fibers are from inferior fronto-occipital fasciculus (IFOF). Interestingly, this result is surprisingly consistent with the previous finding that IFOF may be the most heritable tract when considering FA- derived measures [10].

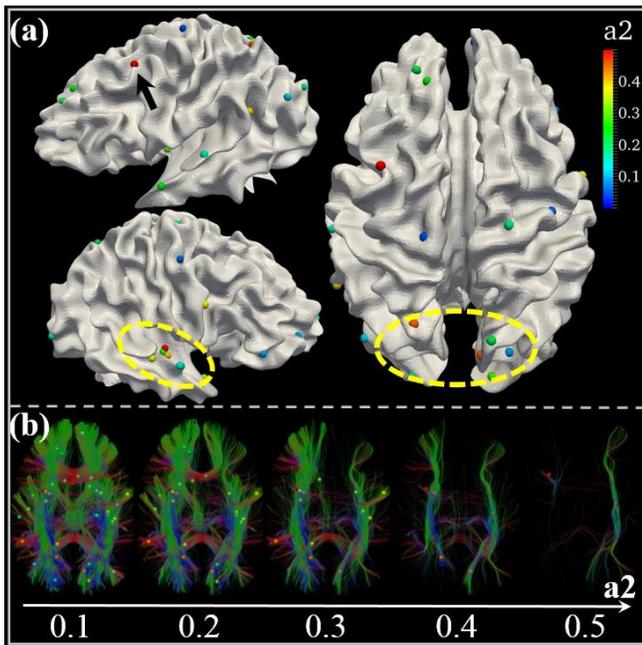


Fig.4. (a) The spatial distribution of “heritable DICCCOLs”. The bubbles represent “heritable DICCCOLs” and the colors reflect the value of a^2 . The “heritable DICCCOLs” with larger a^2 are highlighted with black arrows and yellow circles. (b) A visualization of WM tracts, according to different a^2 of “heritable DICCCOLs” they are connected to.

For better understanding the “heritable DICCCOLs” and the corresponding WM bundles, we randomly picked one “heritable DICCCOL” as an example to illustrate its structural connectivity profile, in **Fig. 5**. This DICCCOL is located at the junction of the temporal and parietal lobes. Each column shows two members of the same twin pair. Based on the figure, we have two observations: first, since we are showing the WM fibers extracted from the same DICCCOL, their overall shape and patterns are very similar. Second, even though there are some differences due to individual structural variability and the tractography method, the WM structural connectivity patterns within the

same twin pair (same column) are more consistent compared to the members of other twin pairs. Of course, this cross-twin structural similarity may also be influenced by non-genetic factors. However, for these identified “heritable DICCCOLs”, additive genetic effects clearly play a critical role in influencing WM structural similarity.

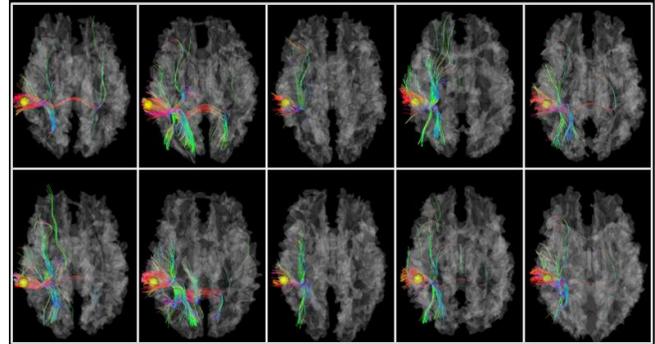


Fig. 5. WM structural connectivity profiles for 10 twin participants. The yellow bubbles represent the selected “heritable DICCCOL”, which is located at the junction of the temporal and parietal lobes. Each column shows two subjects who belong to the same twin pair; there are 5 twin pairs in total.

4. CONCLUSION

In this paper, we used a novel method to investigate white matter connectivity and its heritability. By applying DICCCOL as the structural foundation, we proposed a new DTI derived measure, TMO, and assessed the extent that genetic and environmental factors that influence it using structural equation modeling (the classical A/C/E twin model). We identified 44 “heritable DICCCOLs” exhibiting significant heritability. Both the spatial distribution patterns of these “heritable DICCCOLs”, and the WM tracts that they connect, are consistent with prior reports of moderately high heritability for GM and WM volumes. Therefore, the proposed method has the potential to link GM and WM features together and further provide promising targets for genome-wide association studies.

ACKNOWLEDGMENTS

This work was supported in part by NIH *Big Data to Knowledge* (BD2K) Center of Excellence grant U54 EB020403, funded by a cross-NIH consortium including NIBIB and NCI. Data collection was supported by R01 HD050735 and the National Health and Medical Research Council (NHMRC 486682, 1009064).

5. REFERENCES

- [1] P.J. Basser, et al., “MR Diffusion Tensor Spectroscopy and Imaging,” *Biophys. J.*, 66:259–267, 1994.
- [2] P.M. Thompson, T.D. Cannon, K.L. Narr, T. van Erp, V.P. Poutanen, M. Huttunen, J. Lonnqvist, C.G. Standertskjold-Nordenstam, J. Kaprio, M. Khaledy, R. Dail,

- C.I. Zoumalan, A.W. Toga, "Genetic influences on brain structure," **Nat Neurosci.**, 4:1253–1258, 2001.
- [3] A. Pfefferbaum, E.V. Sullivan, D. Carmelli, "Genetic regulation of regional microstructure of the corpus callosum in late life," **Neuroreport**, 12: 1677-1681, 2001.
- [4] J.L. Stein, et al., "Identification of common variants associated with human hippocampal and intracranial volumes," **Nat Genet.**, 44:552–561, 2012.
- [5] N. Jahanshad, et al., "Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the ENIGMA-DTI working group," **NeuroImage**, 81:455-69, 2013.
- [6] M. Daianu, et al., "Left versus right hemisphere differences in brain connectivity: 4-Tesla HARDI tractography in 569 twins," IEEE International Symposium on Biomedical Imaging (ISBI), 526-529, 2012.
- [7] M.C. Chiang, M. Barysheva, D.W. Shattuck, A.D. Lee, S.K. Madsen, C. Avedissian, A.D. Klunder, A.W. Toga, K.L. McMahon, G.I. de Zubicaray, M.J. Wright, A. Srivastava, N. Balov, P.M. Thompson, "Genetics of brain fiber architecture and intellectual performance," **J. Neurosci.**, 29 (7), 2212–2224, 2009.
- [8] M.M. Bohlken, R.C. Mandl, R.M. Brouwer, M.P. van den Heuvel, A.M. Hedman, R.S. Kahn, H.E. Hulshoff Pol, "Heritability of structural brain network topology: a DTI study of 156 twins," **Hum Brain Mapping**, 35(10):5295-305, 2014.
- [9] M.C. Chiang, K.L. McMahon, G.I. de Zubicaray, N.G. Martin, I. Hickie, A.W. Toga, M.J. Wright, P.M. Thompson, "Genetics of White Matter Development: A DTI Study of 705 Twins and Their Siblings Aged 12 to 29," **NeuroImage**, 54:2308–2317, 2011.
- [10] N. Jahanshad, A.D. Lee, M. Barysheva, K.L. McMahon, G.I. de Zubicaray, N.G. Martin, M.J. Wright, A.W. Toga, P.M. Thompson, "Genetic influences on brain asymmetry: a DTI study of 374 twins and siblings," **Neuroimage**, 52:455–469, 2010.
- [11] D. Zhu, K. Li, C. Faraco, F. Deng, D. Zhang, L. Guo, L.S. Miller, T. Liu, "Optimization of functional brain ROIs via maximization of consistency of structural connectivity profiles," **NeuroImage**, 59(2):1382-1393, 2011.
- [12] D. Zhu*, K. Li*, L. Guo, X. Jiang, T. Zhang, D. Zhang, H. Chen, F. Deng, C. Faraco, C. Jin, C. Wee, Y. Yuan, P. Lv, Y. Yin, X. Hu, L. Duan, X. Hu, J. Han, L. Wang, D. Shen, L. S. Miller, L. Li, T. Liu, "DICCCOL: Dense Individualized and Common Connectivity-based Cortical Landmarks," **Cerebral Cortex**, 23(4):786-800, 2012.
- [13] D. Zhu, K. Li, D. P. Terry, A. N. Puente, L. Wang, D. Shen, L. S. Miller, T. Liu, "Connectome-scale Assessments of Structural and Functional Connectivity in MCI," **Human Brain Mapping**, 35(7):2911-23, 2013.
- [14] K. Li*, D. Zhu*, L. Guo, Z. Li, M. E. Lynch, C. Coles, X. Hu**, T. Liu**, "Connectomics Signatures of Prenatal Cocaine Exposure Affected Adolescent Brains," **Human Brain Mapping**, 34(10):2494-510, 2012.
- [15] D. Zhu, T. Zhang, X. Jiang, X. Hu, N. Yang, J. Lv, J. Han, L. Guo, T. Liu, "Fusing DTI and FMRI Data: A Survey of Methods and Applications," **NeuroImage**, 102(1): 184-191, 2013.
- [16] D. Hibar, et al., "Common genetic variants influence human subcortical brain structures". **Nature**, in press, 2014.
- [17] P.M. Thompson, et al., "The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data," **Brain Imaging and Behavior**, 8(2): 153-182, 2014.
- [18] S.E. Medland, N. Jahanshad, B.M. Neale, P.M. Thompson, "Whole-genome analyses of whole-brain data: working within an expanded search space," **Nature Neuroscience**, 17:791-800, 2014.
- [19] D. Hibar, et al., "Robust subcortical volumetric abnormalities in bipolar disorder: findings from the ENIGMA Bipolar Disorder Working Group." Submitted to **Molecular Psychiatry**, 2014.
- [20] G. Prasad, et al., "Automatic clustering and population analysis of white matter tracts using maximum density paths," **NeuroImage**, 97:284-95, 2014.
- [21] M. Neale, S. Baker, G. Xie, H. Maes, "Mx: Statistical modeling," 6th ed., 2002.
- [22] Y. Benjamini, Y. Hochberg, "Controlling the false discovery rate - a practical and powerful approach to multiple testing," **J Roy Stat Soc B Met**, 57:289–300, 1995.