Imaging genomics
Paul M. Thompson\textsuperscript{a}, Nicholas G. Martin\textsuperscript{b} and Margaret J. Wright\textsuperscript{b}

\textsuperscript{a}Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, California, USA and \textsuperscript{b}Queensland Institute of Medical Research, Brisbane, Australia

Correspondence to Dr Paul Thompson, Laboratory of Neuro Imaging, Dept. of Neurology, UCLA School of Medicine, 635 Charles E. Young Drive South, Suite 225E, Los Angeles, CA 90095-7332, USA
Tel: +1 310 206 2101; fax: +1 310 206 5518; e-mail: thompson@loni.ucla.edu

Current Opinion in Neurology 2010, 23:368–373

Purpose of review
Imaging genomics is an emerging field that is rapidly identifying genes that influence the brain, cognition, and risk for disease. Worldwide, thousands of individuals are being scanned with high-throughput genotyping (genome-wide scans), and new imaging techniques [high angular resolution diffusion imaging and resting state functional magnetic resonance imaging (MRI)] that provide fine-grained measures of the brain’s structural and functional connectivity. Along with clinical diagnosis and cognitive testing, brain imaging offers highly reproducible measures that can be subjected to genetic analysis.

Recent findings
Recent studies of twin, pedigree, and population-based datasets have discovered several candidate genes that consistently show small to moderate effects on brain measures. Many studies measure single phenotypes from the images, such as hippocampal volume, but voxel-wise genomic methods can plot the profile of genetic association at each 3D point in the brain. This exploits the full arsenal of imaging statistics to discover and replicate gene effects.

Summary
Imaging genomics efforts worldwide are now working together to discover and replicate many promising leads. By studying brain phenotypes closer to causative gene action, larger gene effects are detectable with realistic sample sizes obtainable from meta-analysis of smaller studies. Imaging genomics has broad applications to dementia, mental illness, and public health.

Keywords
diffusion imaging, genome-wide association study, heritability, MRI, twins

Introduction
Imaging genetics is a rapidly emerging field that is opening up a new landscape of discovery in medicine and neuroscience. The field is a hybrid effort that merges methods and discoveries in both imaging and genetics; its power has recently taken a quantum leap for a number of reasons. First, many groups worldwide are scanning thousands of individuals with structural and functional magnetic resonance imaging (MRI). Samples are now large enough to discover and verify effects of specific genes on the brain [1–3]. Second, voxel-wise genomic methods are emerging that search every location in a brain image for statistical effects of genes [4\textsuperscript{*}]. These approaches identify coherent anatomical patterns of gene effects in 3D. Replication efforts can then focus on selected brain measures that show promise in preliminary analyses.

Finally, several imaging measures that may seem too unstable to be measured reproducibly have been shown to be highly heritable and feasible to collect in large numbers of individuals. Recent functional MRI studies in healthy adults found that patterns of task-related brain activation [5,6] and resting state functional connectivity [7,8\textsuperscript{*}] are heritable and highly reproducible, making it more likely to discover genes that influence these traits. MRI variants such as diffusion tensor imaging (DTI) – which is sensitive to the directional diffusion of water along neural pathways – can be used to identify entirely new measures of brain integrity and connectivity [9]. By scanning large cohorts of twins with high angular resolution diffusion imaging (HARDI), Chiang \textit{et al.} [10\textsuperscript{*}] revealed that the fiber integrity of the brain is under strong genetic control, and is highly correlated with cognitive measures such as intelligence quotient. Bivariate genetic analysis can also be used to show that common sets of underlying genes affect cognition and fiber integrity [10\textsuperscript{*}] or cognition and regional brain volumes [11,12].

How can gene effects be studied?
Current efforts to relate genetic variation to imaging data typically proceed in one of several directions. The most direct method is to study an illness with a genetic basis,
such as fragile X syndrome, Turner syndrome, Williams syndrome, or 22q11.2 deletion syndrome [13]. Neuroimaging of groups of patients with neurodevelopmental disorders has been extremely useful in discovering the brain systems affected, when and where in the brain abnormalities emerge, and where identifiable abnormalities of brain structure and function fit on the pathway from molecular dysfunction to behavior [14]. Fragile X syndrome, for example, involves a genetic abnormality in the expression of a protein, FMRP, which is involved in dendritic pruning [15]. As such, blood levels of the protein have been related to identifiable patterns of brain hypertrophy in the caudate nucleus, when the normal pruning process is derailed [16]. Brain mapping studies of Williams syndrome have also revealed regionally altered cortical complexity [17], cortical folding [18], and diffusion anisotropy [19] in characteristic patterns that may be related to the expression of genes in the deleted region of the genome. Some of these genes, such as elastin, are involved in cortical folding. In many of these studies, imaging provides a quantifiable phenotype in the brain that can be monitored over time, leading to a more mechanistic understanding of behavioral abnormalities in neurogenetic disorders, and a means to assess their onset and progression.

A second approach to imaging genetics is to fit quantitative genetic models to data from twins or other related individuals [20]. This can identify features or traits that are genetically influenced. Even the most associated common genetic polymorphisms are individually expected to explain less than 1–5% of the variation in the form of a 3D map. This and subsequent analyses showed that cortical gray matter volumes, especially in the frontal lobes, were highly heritable, consistent with reports of gray matter deficits in relatives of unaffected patients with schizophrenia [33]. Hulshoff Pol et al. [34] used an image warping method to deform over 100 twins’ MRI scans to a common brain template to estimate the volumes of different parts of the brain. Most structure volumes were highly heritable; other structures, such as the cerebellum, were more environmentally influenced. Statistical mapping methods may also be used to transfer other types of imaging data to a common template, and perform voxel-by-voxel genetic analysis. In one morphometric study of twins, the degree of genetic control was found to be higher for earlier-maturing regions of the brain, such as the visual processing areas of the occipital lobes. In line with expectation, frontal lobe white matter structures with a more protracted maturational time-course were more environmentally influenced [35]. Paradoxically, however, there is some evidence for an increasing genetic effect on cortical thickness with age [36], in agreement, perhaps, with the rising heritability of intelligence quotient with age [37,38]. Importantly, these genetic mapping studies find anatomical and time-dependent gradients in the pattern of genetic influences. Mapping methods can then home in on selected regions of the brain in which genetic effects were most likely to be detected. This type of voxel-selection approach – when candidate brain regions are prescreened for future study – is the mainstay of functional brain imaging, and avoids the need for heavy corrections for multiple comparison.
statistical comparisons that occur when surveying the entire brain.

**Candidate genes**

Given the enormous number of possible genes that could influence the brain, candidate genes offer a more principled method to focus on promising regions of the genome. Perhaps the most studied polymorphism in brain imaging is the apolipoprotein E4 (ApoE4) allele, which is carried by approximately 25% of normal individuals; each allele confers a three-fold increased risk for developing late-onset Alzheimer’s disease by 75 years of age. ApoE4 is now widely recognized as leading to reduced gray matter [39,40] and white matter [41] volumes in the elderly, perhaps due to the toxic effects of beta-amyloid burden, which is also higher in ApoE4 carriers [42,43]. Somewhat disturbingly, Shaw et al. [44] found that children carrying this common gene also showed a pattern of cortical thinning and an altered developmental trajectory; resting-state brain activity is also altered in young ApoE4 carriers (age 20–35; [45]). In one of a series of studies associating BDNF variants with alterations in brain structure, activation, and cognition (e.g. [46]), Chiang et al. [47] discovered that a common polymorphism in the BDNF gene influences the fiber integrity of the brain, as seen with DTI. The BDNF polymorphism accounted for around 15% of the variance in diffusion anisotropy in the posterior outflow of the corpus callosum [47], but even so, scans from 455 twins were required to detect and replicate this association, suggesting the need for large imaging databases.

Genes coding for common variants in monoamine neurotransmitter receptors and transporters are a common target of study [e.g. [48–51] on catechol-O-methyl transferase (COMT); [52] on the dopamine transporter, DAT1; and [53] relating variations in the dopamine D2 receptor to functional connectivity]. Hariri et al. [54] and others have reported that people carrying one or two copies of the short allele of the serotonin transporter (5-HTT) promoter polymorphism, which has been associated with reduced 5-HTTT expression and function and increased fear and anxiety-related behaviors, exhibit greater neuronal activity in the amygdala, as assessed by BOLD functional MRI, in response to fearful stimuli. Other candidate gene discoveries are relevant to public health. Ho et al. [55] found that FTO, an obesity-related gene variant carried by nearly half of all Western Europeans, may also be associated with brain degeneration. In 3D maps based on over 200 healthy elderly Caucasian individuals’ MRI scans, carriers’ brains demonstrated 8% lower volume than noncarriers in the frontal lobes and 12% lower volume in the occipital lobes. FTO may exert an additive effect on brain degeneration beyond the influence of an individual’s body mass index, which is also associated with atrophy and differences in brain structure [56].

**Genome-wide association studies**

Candidate gene studies can provide a better understanding of the pathways involved in several brain diseases, but the advent of high-throughput genotyping has offered still greater potential to search the entire genome for causative variants. Genome-wide association (GWA) scanning, in particular, searches the entire genome for single-nucleotide polymorphisms (SNPs) that may be associated with certain behaviors, diseases, or, in the case of imaging genomics, with imaging measures. Figure 1 shows how such an approach works (adapted from [57]).

### Figure 1 Genome-wide association study of temporal lobe structure

The GWA study identified a common glutamate receptor variant (GRIN2b) with suggestive evidence of association with temporal lobe volume and increased risk for Alzheimer’s disease in 740 brain MRI scans. The plot on the left is a standard output from the PLINK software [58] — it shows the significance of association between the image-derived measure of interest (here temporal lobe volume) and variants at each of 600,000 locations on the genome. By performing genetic association to assess the effects of variants in this SNP at millions of points in a brain image, associations were detected and mapped in 3D (right). Adapted from [57].
In this preliminary GWA study of neurodegeneration, Stein *et al.* [57] searched 546,314 genomic markers using the PLINK software [58]. Two SNPs were suggestively associated with temporal lobe volumes ($P < 5 \times 10^{-7}$), and with increased atrophy in all three diagnostic categories [Alzheimer’s disease, mild cognitive impairment (MCI), and controls]. One SNP, with genome-wide evidence, was in the GRIN2B gene that encodes the N-methyl-D-aspartate (NMDA) glutamate receptor NR2B subunit. It was also over-represented in Alzheimer’s disease and MCI patients versus controls (odds ratio (OR) = 1.273; $P = 0.039$). This protein, involved in learning and memory and excitotoxic cell death, has age-dependent prevalence in the synapse, and is already a therapeutic target in Alzheimer’s disease [59]. Voxel-by-voxel, 3D maps of genetic association with regional brain volumes revealed intense temporal lobe reductions of around 1.5% per risk allele (Fig. 1).

**Rare variants**

Dickson *et al.* [60*] have proposed, somewhat provocatively, that uncommon or rare genetic variants can easily create synthetic associations that are credited to common variants, and this should be considered when interpreting GWA scanning signals. Conventional GWA scans may also miss a large component of causal variance due to rare functional variation. With this in mind, Choi *et al.* [61] and others have advocated finer-scale exome sequencing of pedigrees, to identify all polymorphisms that might be relevant to a brain phenotype. As sequencing becomes less expensive, the relative contribution of rare versus common variation to brain phenotypes is likely to be better understood. Imputation to the 1000 Genomes Project (up to 8 million SNPs) and the advent of much denser SNP chips will provide cheaper options in the short term.

**Voxel-wise genomics: statistical and computational challenges**

Most genome-wide studies of brain images reduce the dimension of one or both datasets to avoid performing statistical tests on an astronomical scale. Meda *et al.* [62], for example, advocated independent component analysis to identify the association between sets of genomic variations and phenotypic traits, identifying interactions between patterns of brain function and genetic information. Other multivariate methods, such as canonical correlation analysis, have been used to find optimal ‘projections’ to best correlate the multivariate measures of genotypes and phenotypes [63].

Stein *et al.* [4*], however, used a massively parallel computer to associate approximately 600,000 SNPs at approximately 200,000 voxels in the brain, requiring approximately $1.2 \times 10^{11}$ tests on data from 742 individuals. To address the problem of multiple comparisons across both the genome and the image, only the ‘winning’ (most highly associated) SNP was retained at each voxel, and its $P$ value was plotted into the image. A beta function was fitted to model the distribution of minimum $P$ values that would be obtained under the null hypothesis. False discovery rate methods [64] were used to adjust for multiple tests across the image. Although the list of promising genes requires replication, clearly if several centers perform such an analysis, the resulting maps can be combined in several simple ways. Statistical conjunction maps [65], for example, could identify subsets of voxels and SNPs with weak or strong evidence of association in multiple datasets. Clearly, genes are more likely to operate in networks, and interactions are likely; machine learning methods, such as adaptive boosting have been proposed to identify sets of SNPs with large combined effects that would be hard to detect independently. Recognizing that individual variants are likely to have small effect, risk profiling has also been proposed [66] to identify sets of SNPs that jointly explain the observed variance. Genes may also interact [e.g. *BDNF* and *SLC6A4* (5-HTTLPRI)], or the dopamine receptor and transporter [67], and tools to detect these interactions are just beginning to be developed [68,69].

**Replication by joining datasets**

A major barrier in imaging genomics is replication; without it, GWA findings are met with skepticism [70]. Recently, Chiang *et al.* [65] discovered, and replicated, a finding that a nonsynonymous coding variant in the *BDNF* gene affects white matter microstructure and its relation to intelligence quotient. In parallel work, Kochunov *et al.* [3] performed GWA on a large diffusion imaging dataset from a Mexican-American pedigree, and revealed a list of promising genes that affect white matter circuitry and brain structure. Given the large samples needed to discover and verify promising hits, several imaging genomics groups are now working collaboratively to replicate findings [71*]. The ENIGMA network (Enhancing Neuroimaging Genetics through Meta-Analysis; http://enigma.loni.ucla.edu) seeks to accelerate replication by sharing information on promising findings, and identifying cohorts with sufficient power for meta-analysis. This approach has been fruitful in psychiatry and behavioral genetics [72*].

**Conclusion**

Imaging genetics is an emerging field evaluating imaging measures as quantitative traits. Many imaging phenotypes have high precision, and are readily standardized across centers [73]. They may also require smaller sample sizes to detect association [74] – and arguably, they may be closer to the underlying biology of disease than behavioral or standard diagnostic measures, making contributing genes easier to identify.
The flurry of recent discoveries in imaging genomics underscores the fact that imaging datasets are now large enough to discover single gene effects on brain structure and function. Several candidate genes show reproducible effects on the brain and risk for disease; long lists of new candidates have been discovered by mining genetic data. Replication is key and collaborations such as the ENIGMA network have been set up to replicate promising findings across multiple independently collected samples. In parallel, several groups have extended genomic analysis to new types of images: diffusion-based maps of fiber integrity and connectivity [9], and 3D maps of task-related brain activation and ‘resting-state’ functional connectivity [5,8*]. Together, these advances are spurring many new discoveries, providing valuable mechanistic information about the brain and behavior, and about factors that affect the expression of neuropsychiatric illnesses.

Acknowledgements

The authors are supported by NIH grant number R01 HD050735 from the National Institute of Child Health and Human Development, USA, and Project Grant 496882 from the National Health and Medical Research Council, Australia. P.T. is also supported, in part, by NIH grants EB008432, EB008281, EB007813, AG036535, RR021813 and RR013842. The authors report no conflicts of interest, financial or otherwise, regarding the work in this study.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest
•• of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (pp. 434–435).


59 Dickson SP, Wang K, Krantz I, et al. Rare variants create synthetic genome-wide associations. PLoS Biol 2010; 8:e1000294. This provocative study argues that rare variants may be responsible for some of the results seen in GWA studies. It advocates deeper sequencing of the genome than is customary in GWA scanning studies.

