Univariate Approaches: Multiple Testing & Voxelwise Whole Genome Association

Jason L. Stein
Laboratory of Neuro Imaging
University of California, Los Angeles
steinja@ucla.edu
June 6, 2010

Outline

- Why use a genome-wide analysis?
- Why use a genome-wide analysis on imaging data?
- Why search voxelwise for genetic effects?
- How many tests are involved and how do we deal with them?
- Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function
- Application of vGWAS in ADNI dataset

Outline

- Why use a genome-wide analysis?
- Why use a genome-wide analysis on imaging data?
- Why search voxelwise for genetic effects?
- How many tests are involved and how do we deal with them?
- Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function
- Application of vGWAS in ADNI dataset

Brain structure is highly heritable

Must be specific genetic variants
explaining the high heritability
(most of which are unknown)

(Kremen et al., 2010)

Low prior probability of any candidate to have effects on brain structure

"Choosing candidate genes is generally on the basis of limited information and therefore exclude the vast majority of genes expressed in the central nervous system" (Glatt & Freimer, 2002)

Percentage of Genes Expressed in Human Cortex (Gene Chip)

(Myers et al., 2007)

Percentage of Genes Expressed in Mouse Brain (ISH)

(Lein et al., 2007)

We generally don't know the theoretical genetic underpinnings of a phenotype

(Freimer & Sabatti, 2004)

Structure of genome allows for high coverage with SNP chips

http://hapmap.ncbi.nlm.nih.gov

Table 1. Estimates of Genomic Coverage for Currently Available Genome-wide SNP Platforms Alone and after Imputation

	Percentage of Genomic Coverage at $r^2 \geq 0.8$	Percentage of Genomic Coverage at $r^2 = 1$
Affymetrix SNP Array 5.0	65	43
Affymetrix SNP Array 5.0 plus imputed SNPs	73	54
Affymetrix SNP Array 6.0	80	59
Illumina HumanHap 300	77	42
Illumina HumanHap 300 plus imputed SNPs	81	50
Illumina HumanHap 550	87	57
Illumina HumanHap 1M	91	68

Estimates evaluated with Phase II HapMap data from the CEU population. Coverage estimates for Illumina HumanHap 1M and Affymetrix SNP-array-6.0 are likely to be biased downward because the genotypes at approximately 10% of the SNPs on each platform are not currently publicly available for the CEU HapMap individuals. Where imputations are included, all SNPs passing imputation-filter thresholds and with an $r^2 \geq 0.8$ between known and imputed genotypes are included along with the SNPs on the genome-wide SNP chip.

(Anderson et al, 2008)

And SNP genotyping chips are cheap!

Outline

- Why use a genome-wide analysis?
  - Much heritability left to explain!
  - Candidate genes have given many insights, but low prior probability of selecting right one
  - The structure of the genome allows for genome-wide search
- <u>Why use a genome-wide analysis on imaging data?</u>
- Why search voxelwise for genetic effects?
- How many tests are involved and how do we deal with them?
- Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function
- Application of vGWAS in ADNI dataset

Two reasons to use genome-wide association on imaging data

(1)

Genetic Variation (SNPs) → Brain Structure and /or function

Interested in finding the genetic variants that influence the brain structures/functions of interest

(2)

Genetic Variation (SNPs) → Endophenotype (Brain Structure, expression levels, Insulin, etc.) → Diagnostic Status

Interested in the genetic variants that influence disease state and believe that brain traits are quantitative traits closer to the genetics

GWAS has only been moderately successful in psychiatric and neurological disorders

Table 1 | Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained	Heritability measure
Age-related macular degeneration[72]	5	50%	Sibling recurrence risk
Crohn's disease[21]	32	20%	Genetic risk (liability)
Systemic lupus erythematosus[73]	6	15%	Sibling recurrence risk
Type 2 diabetes[74]	18	6%	Sibling recurrence risk
HDL cholesterol[75]	7	5.2%	Residual* phenotypic variance
Height[15]	40	5%	Phenotypic variance
Early onset myocardial infarction[76]	9	2.8%	Phenotypic variance
Fasting glucose[77]	4	1.5%	Phenotypic variance

* Residual is after adjustment for age, gender, diabetes.

(Manolio et al, 2009)

<u>Why?</u>

- Disease caused by rare variants which are untested in GWAS?
- Epistatic interactions?
- Gene x Environment interactions?
- **Disease too complex/constructs derived from clinical criteria so unlikely biologically homogeneous**

Quantitative traits more powerful than clinical diagnosis

Figure 1. The graph shows the power distribution curves for QT analysis contrasted with a case-control design at $p < .01$ and 10^{-7} (OR $= 1.5$). The x axis portrays the sample sizes and the y axis the power at each value of the sample size for a 10% percentage variance explained for the QT, a QT MAF of 10% and a marker SNP MAF at 20%. The results for QT and case-control are displayed in black and grey respectively. Compared to QT, the case-control curves are shifted to the right indicating that much larger sample sizes are required to reach the same power.

(Potkin et al., 2009)

Outline

- Why use a genome-wide analysis?
- Why use a genome-wide analysis on imaging data?
  - More powerful way to find genetics associated with diseases of the brain
- <u>Why search voxelwise for genetic effects?</u>
- How many tests are involved and how do we deal with them?
- Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function
- Application of vGWAS in ADNI dataset

ROI based approach are interesting and can be successful

ROI based phenotype

GWAS to ROI based phenotypes

(Stein, et al., 2010a)

Voxelwise vs. ROI approach

Dependent on geometry of the signal

★ = signal

Signal overlaps with ROI definition
ROI more powerful

Signal does not overlap with ROI definition
Voxelwise more powerful

In search for genetic effects on brain structure – we generally are not clear where they are

(Desikan et al., 2006)

Outline

- Why use a genome-wide analysis?
- Why use a genome-wide analysis on imaging data?
- Why search voxelwise for genetic effects?
  - When signal location is unknown, search entire space
- <u>How many tests are involved and how do we deal with them?</u>
- Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function
- Application of vGWAS in ADNI dataset

Multiple Testing Problem

~1000 subjects

~30,000 voxels in the brain

1.8×10^{10} tests!

~600,000 genetic markers (SNPs)

Multiple Comparisons: GWAS

One SNP

600,000 SNPs

Null *P*-values: Uniform Distribution

Independent null *P*-values: *Beta*(1,600000) distribution

Multiple Comparisons Example

Error: Not accounting for multiple comparisons

Null *P*-values: Uniform Distribution

600,000 draws from a uniform distribution (Like GWAS on one phenotype). Minimum P-value gives very low P-values (1.7057×10^{-6} , 1.1026×10^{-6})

Wow, I have such low P-values! But all of this is randomness (simulated from null distributions)

Accounting for multiple comparisons assuming independence: Beta Distribution

Beta(1,600000) distribution

Models the multiple comparisons by picking the minimum *P*-value after 600,000 draws from a uniform distribution.

Adjustment through CDF of *Beta*(1,600000) gives corrected P-values

Raw P-value	**Corrected P-value**
1.7057×10^{-6}	0.646
1.1026×10^{-6}	0.484

Multiple Comparisons: Correlation of Genetic Markers

Linkage disequilibrium (LD; correlation between genetic markers) means that all tests are **not** independent.

simpleM: a method to determine the effective number of tests conducted M_{eff} where $M_{eff} \leq M$

1. Create correlation matrix
2. Calculate eigenvalues through PCA
3. Number of principal components which jointly explain 99.5% of variance = M_{eff}

TABLE I. Derived significance thresholds for the Illumina 1M data

		Permutation		SimpleM	
Chromosome	Number of SNPs	M_{eff}	α_p	M_{eff}	α_p
1	63,560	30,881	1.66e−06	31,898	1.61e−06
2	64,367	31,760	1.62e−06	33,386	1.54e−06
3	53,622	26,328	1.95e−06	27,675	1.85e−06
4	47,096	24,292	2.11e−06	24,474	2.10e−06
5	48,758	24,187	2.12e−06	25,229	2.03e−06
6	52,870	25,667	2.00e−06	26,165	1.96e−06
7	43,084	23,040	2.23e−06	23,123	2.22e−06
8	41,932	23,254	2.21e−06	22,314	2.30e−06
9	35,591	20,078	2.55e−06	20,061	2.56e−06
10	40,195	21,186	2.42e−06	21,277	2.41e−06
11	39,420	21,101	2.43e−06	20,359	2.52e−06
12	38,787	20,875	2.46e−06	21,318	2.41e−06
13	28,656	14,936	3.43e−06	15,087	3.40e−06
14	26,026	14,414	3.56e−06	13,829	3.71e−06
15	23,896	12,277	4.18e−06	13,312	3.85e−06
16	24,517	14,346	3.58e−06	14,626	3.51e−06
17	23,054	11,558	4.44e−06	12,578	4.08e−06
18	22,885	11,629	4.41e−06	12,583	4.08e−06
19	16,918	9,152	5.60e−06	9,646	5.32e−06
20	19,995	10,177	5.04e−06	11,169	4.59e−06
21	11,012	6,150	8.34e−06	6,247	8.21e−06
22	12,388	6,708	7.65e−06	7,004	7.32e−06
Genome-wide	778,629	378,368	1.36e−07	413,360	1.24e−07

Similar to permutation derived values

Correct P-values through a Beta(1, M_{eff}) distribution

(Gao et al., 2008; Gao et al., 2010)

Multiple Comparisons: Correction Across Voxels Through False Discovery Rate (FDR)

Signal + Noise

Control of False Discovery Rate at 10%

6.7% 10.4% 14.9% 9.3% 16.2% 13.8% 14.0% 10.5% 12.2% 8.7%

Percentage of Activated Pixels that are False Positives

(Tom Nichols website: http://www.sph.umich.edu/~nichols/FDR/; Genovese et al., 2002)

Histogram visualization

FDR significant – over-representation of low *P*-values

Null *P*-values – no violations of assumptions

Violation of assumptions – bimodal histogram

Violation of assumptions – discrete *P*-value distribution

Figure 1: Examples of *P*-value histograms.

(Pounds, 2006; Dabney & Storey, 2006)

Outline

- Why use a genome-wide analysis?
- Why use a genome-wide analysis on imaging data?
- Why search voxelwise for genetic effects?
- How many tests are involved and how do we deal with them?
  - ~ 1.8 x 10^{10} tests
  - Can use Beta(1, M_{eff}) to correct across genetics followed by FDR to correct across voxels
- <u>Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function</u>
- Application of vGWAS in ADNI dataset

J. Craig Venter says it's important!

human genetic variation and biological outcomes such as physiology and disease, will require the complete genomes of tens of thousands of humans together with comprehensive digitized phenotype data. A simplistic query could be easily scored, for example, "do you have diabetes: yes or no". A more comprehensive view would include such things as age of onset and a scoring of the range of clinical manifestations associated with the disease, including extent of nerve damage, vascular issues, doses of medications used and family history. The scoring system would subcategorize characteristics such as disease type, progression and severity.

Even if we had all this information today, we wouldn't be able to make use of it because we don't have the computational infrastructure to compare even thousands of genotypes and phenotypes with each other. The need for such an analysis could be the best justification for building a proposed 'exascale' supercomputer, which would run 1,000 times faster than today's fastest computers. Scientists need to work

EUROPEAN NUCLEOTIDE ARCHIVE AND ENSEMBL GENOMES (EBI, EMBL)

(Venter, 2010)

vGWAS

Figure 1: Summary of vGWAS method for conducting association and assessing statistical significance.

(Stein et al., 2010b)

Outline

- Why use a genome-wide analysis?
- Why use a genome-wide analysis on imaging data?
- Why search voxelwise for genetic effects?
- How many tests are involved and how do we deal with them?
- Voxelwise Genome Wide Association (vGWAS): a method for finding genes affecting brain structure/function
- <u>Application of vGWAS in ADNI dataset</u>

vGWAS

Figure 1: Summary of vGWAS method for conducting association and assessing statistical significance.

(Stein et al., 2010b)

ADNI Dataset

Subjects

740 Caucasian subjects to avoid population stratification

Diagnosis:
- 173 Alzheimer's disease patients
- 361 Mild Cognitive Impairment
- 206 healthy elderly

Demographics:
- 75.52 +/- 6.82 years
- 438 male

Genetics

Illumina 610-Quad BeadChip

Exclusions:
- genotype call rate < 95%,
- deviation from Hardy-Weinberg equilibrium $P<5.7x10^{-7}$
- minor allele frequency < 0.10

448,293 SNPs in analysis

Imaging Phenotype

Tensor Based Morphometry

Each voxel encodes volume change relative to a study-specific template

31,622 voxels in the brain when downsampled to 4x4x4 mm^3 voxels

(Stein et al., 2010b)

vGWAS

Figure 1: Summary of vGWAS method for conducting association and assessing statistical significance.

(Stein et al., 2010b)

Computationally Intensive: GWAS on each voxel

Genome-wide association on each phenotype takes ~9 minutes / phenotype.
31,622 voxels * 9 minutes = 198 days of computation!
Across 300 nodes total computation time is 27h.

http://pipeline.loni.ucla.edu/

(Image courtesy of D. Hibar)

Raw minimum *P*-value at each voxel

Figure 3: The significance of the most associated SNP at each voxel. Each voxel is colored by the $-\log_{10}$ of the *P*-value of the genetic association at each point (warmer colors are more strongly associated).

(Stein et al., 2010b)

Most associated SNPs

Chr	Base Pair	SNP	MAF	Number of subjects in genotype groups			Volume (mm^3)	Minimum *P*-value	mean *P*-value	Gene or EST (±50 kb)
				Maj	Het	Min				
6q16.2	99778735	rs2132683	0.3257	340	318	82	4224	2.56×10^{-10}	1.01×10^{-6}	
6q15	91474473	rs713155	0.3966	274	345	121	7296	3.11×10^{-10}	5.08×10^{-7}	
1p35.1	34020651	rs476463	0.1203	567	168	5	1472	3.18×10^{-10}	1.27×10^{-6}	***CSMD2***
7q31.32	121989829	rs2429582	0.3417	319	331	86	2496	4.23×10^{-10}	6.46×10^{-7}	***CADPS2***
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2048	5.34×10^{-10}	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	14528	[illegible]	[illegible]	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.41×10^{-9}	6.54×10^{-7}	*WFDC2, SPINT3*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	64	[illegible]	6.57×10^{-7}	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	34	2.27×10^{-9}	1.00×10^{-6}	**BG436399**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	24	2.29×10^{-9}	6.21×10^{-7}	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	68	2.30×10^{-9}	1.21×10^{-6}	***SHB***
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	56	2.65×10^{-9}	1.10×10^{-7}	***KIAA0090***, *MRT04, AKR7L*
9p13.1	38031142	rs7873102	0.3821	283	340	109	832	2.96×10^{-9}	6.42×10^{-7}	***SHB***
20p12.1	12822585	rs2073233	0.4291	234	369	131	2560	3.17×10^{-9}	1.42×10^{-6}	**BC036700**
2q37.3	242151629	rs12479254	0.4049	255	341	119	1408	3.88×10^{-9}	5.70×10^{-7}	***BOK***, *THAP4*
16p12.1	24439219	rs11643520	0.1160	574	146	11	1920	4.39×10^{-9}	6.06×10^{-7}	*RBBP6*
5p12	44222425	rs4296809	0.1448	539	177	17	12736	4.41×10^{-9}	8.75×10^{-7}	**BG334794**
13q32.2	97764318	rs688872	0.3804	283	345	106	3392	4.68×10^{-9}	1.06×10^{-6}	***FARP1***
14q22.1	51080549	rs7140150	0.4566	219	353	160	1856	5.78×10^{-9}	4.77×10^{-7}	***FRMD6***
6p12.3	49596867	rs9473582	0.3973	274	339	121	4416	5.98×10^{-9}	8.25×10^{-7}	***GLYATL3***

Highest expression in the brain

Regulates synaptic and large dense core vesicle priming in neurons, associations to autism

(Stein et al., 2010b)

Most associated voxels for most associated SNPs

(Stein et al., 2010b)

vGWAS

Figure 1: Summary of vGWAS method for conducting association and assessing statistical significance.

(Stein et al., 2010b)

M_{eff} Estimation

Table 1
Number of SNPs measured and the effective number of tests (M_{eff}) on each chromosome. The effective number of tests was not estimated in sex chromosomes or for mitochondrial DNA, where the effective number of tests was set to be equal to the number of SNPs measured in those regions, rather than a smaller number. This is the most conservative estimate. Chromosome XY refers to SNPs on both the X and Y chromosomes.

Chr	Number of SNPs	M_{eff}
1	33,850	20,205
2	36,384	21,747
3	30,765	18,275
4	27,072	15,972
5	27,396	16,245
6	30,054	16,669
7	24,446	14,680
8	24,768	14,710
9	21,283	13,231
10	23,089	13,806
11	21,796	13,066
12	21,461	12,787
13	16,605	9809
14	14,501	8916
15	13,355	8557
16	13,460	9111
17	11,721	7764
18	13,198	8232
19	7,895	5425
20	11,169	7148
21	6,582	4193
22	6,757	4341
X	10,637	10,637
Y	12	12
XY	25	25
Mitochondrial	12	12
Total	448,293	275,575

$$M_{eff} << M$$

(Stein et al., 2010b)

vGWAS

Figure 1: Summary of vGWAS method for conducting association and assessing statistical significance.

(Stein et al., 2010b)

How well do results fit distributions?

Raw *P*-value distribution

Corrected *P*-value distribution

(Stein et al., 2010b)

vGWAS

Figure 1: Summary of vGWAS method for conducting association and assessing statistical significance.

(Stein et al., 2010b)

Significance Testing through FDR and pFDR

Fig. 3. The cumulative distribution function of corrected *P*-values. The cumulative distribution function of P_c-values is shown (red) with two lines representing thresholds of $q = 0.50$ (blue), and $q = 0.05$ (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

q-value = 0.25 for SNP rs2132683

(Stein et al., 2010b)

Sample size needed for replication

N=312 for rs2132683**; N=263** for rs713155**; N=291** for rs476463**; N=299** for rs2429582**; N=319** for rs9990343

(Stein et al., 2010b)

Replication through collaboration

http://enigma.loni.ucla.edu

Useful web resources

UCSC genome browser: http://genome.ucsc.edu/cgi-bin/hgGateway
Genome visualization magic.

Hapmap: http://hapmap.ncbi.nlm.nih.gov/
Allele frequencies in multiple populations.

BioGPS: http://biogps.gnf.org/#goto=welcome
See what tissue the gene is expressed in.

Entrez Gene: http://www.ncbi.nlm.nih.gov/gene/
See the gene ontology (what it does).

dbSNP: http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp
The database of every documented genetic variation.

Plink: http://pngu.mgh.harvard.edu/~purcell/plink/
Incredibly useful tool for genome-wide analysis, organization, etc. Excellent documentation.

dbGaP: http://www.ncbi.nlm.nih.gov/gap/
Database of genotypes and phenotypes.

Acknowledgements

LONI UCLA

Paul Thompson

Derrek P. Hibar

Suh Lee

Xue Hua

Alex Leow

ADNI Genetics Core (Indiana University)

Andrew Saykin

Li Shen

Tatiana Foroud

Nathan Pankratz

NeuroImaging Training Program Training Grant
NIH/NIDA 1-T90-DA022768:02

UCLA Affiliates Scholarship
Dr. Ursula Mandel Scholarship

ARCS Scholar

Pre-doctoral NRSA 1F31MH087061-01

References

Anderson, C.A., Pettersson, F.H., Barrett, J.C., Zhuang, J.J., Ragoussis, J., Cardon, L.R., Morris, A.P., 2008. Evaluating the effects of imputation on the power, coverage, and cost efficiency of genome-wide SNP platforms. *Am J Hum Genet* 83(1), 112-119.

Dabney, A.R., Storey, J.D., 2006. A reanalysis of a published Affymetrix GeneChip control dataset. *Genome Biol* 7(3), 401.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31(3), 968-980.

Freimer, N., Sabatti, C., 2004. The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. *Nat Genet* 36(10), 1045-1051.

Gao, X., Becker, L.C., Becker, D.M., Starmer, J.D., Province, M.A., 2010. Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet Epidemiol* 34(1), 100-105.

Gao, X., Starmer, J., Martin, E.R., 2008. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol* 32(4), 361-369.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15(4), 870-878.

Glatt, C.E., Freimer, N.B., 2002. Association analysis of candidate genes for neuropsychiatric disease: the perpetual campaign. *Trends Genet* 18(6), 307-312.

Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., et al., 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445(7124), 168-176.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., et al., 2009. Finding the missing heritability of complex diseases. *Nature* 461(7265), 747-753.

Munafo, M.R., Brown, S.M., Hariri, A.R., 2008. Serotonin transporter (5-HTTLPR) genotype and amygdala activation: a meta-analysis. *Biol Psychiatry* 63(9), 852-857.

Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., et al., 2007. A survey of genetic human cortical gene expression. *Nat Genet* 39(12), 1494-1499.

Potkin, S.G., Turner, J.A., Guffanti, G., Lakatos, A., Torri, F., Keator, D.B., Macciardi, F., 2009. Genome-wide strategies for discovering genetic influences on cognition and cognitive disorders: methodological considerations. *Cogn Neuropsychiatry* 14(4-5), 391-418.

Pounds, S.B., 2006. Estimation and control of multiple testing error rates for microarray studies. *Brief Bioinform* 7(1), 25-36.

Stein, J.L., Hua, X., Morra, J.H., Lee, S., Hibar, D.P., Ho, A.J., Leow, A.D., Toga, A.W., Sul, J.H., Kang, H.M., Eskin, E., Saykin, A.J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M.J., Craig, D.W., Gerber, J.D., Allen, A.N., Corneveaux, J.J., Stephan, D.A., Webster, J., DeChairo, B.M., Potkin, S.G., Jack, C.R., Jr., Weiner, M.W., Thompson, P.M., 2010a. Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. *Neuroimage* 51(2), 542-554.

Stein, J.L., Hua, X., Lee, S., Ho, A.J., Leow, A.D., Toga, A.W., Saykin, A.J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M.J., Craig, D.W., Gerber, J.D., Allen, A.N., Corneveaux, J.J., Dechairo, B.M., Potkin, S.G., Weiner, M.W., Thompson, P.M., 2010b. Voxelwise genome-wide association study (vGWAS). *Neuroimage*. In press.

Venter, J.C., Multiple personal genomes await. *Nature* 464(7289), 676-677.