# Supporting Information

## Douaud et al. 10.1073/pnas.1301816110

### SI Methods

**Ethics.** The study was carried out according to the principles expressed in the Declaration of Helsinki and was approved by the Oxfordshire National Health Service research ethics committee A on January 6, 2006 (Central Office for Research Ethics Committees, COREC 04/Q1604/100). Each subject gave written consent for their participation.

**Subjects.** Details of recruitment into the VITACOG trial, the cognitive testing, biochemical measurements, randomization, B-vitamin treatment, and ethical approval have been described previously (1, 2). Briefly, inclusion criteria included age ≥70 y, study partner available as informant, and diagnosis of amnestic or nonamnestic mild cognitive impairment (MCI) according to Petersen criteria (3). The diagnosis of MCI included a subjective concern about memory and a cognitive screening test using word-recall (TICS-M) with a score >1.5 SD below the norm that did not interfere with activities of daily living. The cutoff scores were based on locally established norms from subjects of the same age in the same city. Exclusion criteria included a diagnosis of dementia or being treated with antidementia drugs; active cancer; major stroke within the past 3 mo; treatment with methotrexate, anticancer, or antiepileptic drugs; taking folic acid >300 μg/d, pyridoxine >3 mg/d or vitamin B12 >1.5 μg/d by mouth or any dose by injection. Participants received a tablet containing 0.8 mg folic acid, 0.5 mg cyanocobalamin and 20 mg pyridoxine HCl, or a placebo tablet.

Baseline demographics characteristics, plasma biochemistry and cognitive test scores of these subjects are given in Table 1. In particular, the plasma total homocysteine (tHcy) concentrations at baseline were 11.4 ± 3.1 μmol/L in the placebo group and 11.8 ± 3.6 μmol/L in the B-vitamin group and, after 2 y, 12.6 ± 4.1 μmol/L in the placebo group and 8.9 ± 2.5 μmol/L in the B-vitamin group ($P < 0.001$). Thus, tHcy level was 29.4% lower in the B-vitamin group than in the placebo group at the end of the study.

**Image Acquisition.** From the 223 subjects who had been followed up for the 2 y, 187 had volunteered (84%) for two MRI scans and were eligible to be scanned (exclusion criteria included claustrophobia, severe back ache, presence of metallic implants, pacemaker and intraocular metallic foreign bodies). Seven subjects withdrew before the second scan and, of the 180 remaining, 12 T1-weighted scans were found to be technically unsatisfactory, leaving a total of 168 subjects who were included in our previous, whole-brain SIENA analysis (2). All subjects underwent the same imaging protocol with whole-brain T1-weighted scans acquired at baseline and after 2 y using the same 1.5-T Sonata MR imager (Siemens Medical Systems) with a standard quadrature head coil and maximum 40 $mT.m^{-1}$ gradient capability. Subjects were scanned with a 3D T1-weighted fast low-angle shot (FLASH) sequence using the following parameters: coronal orientation, matrix 256 × 256, 208 slices, 1 × 1 $mm^2$ in-plane resolution, slice thickness 1 mm, echo/repetition times 5.65/12 ms, flip angle 19°. To increase signal-to-noise ratio, scanning was repeated three times and averaged after acquisition and cross-repeat alignment.

**Image Preprocessing.** We wanted to investigate voxelwise, localized gray matter (GM) volume changes over the 2-y period between placebo and vitamin B-treated groups across the whole brain. As it was important that the use of nonlinear deformations to register native scans into a common space was carried out with appropriate accuracy, we conducted an "optimized" voxel-based morphometry (VBM) protocol (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLVBM) (4, 5) using FMRIB Software Library (FSL) tools (6), in which a symmetric study-specific GM template was built from the MCI participants' images. For this template, brain-extracted GM-segmented images from all subjects at both baseline time point (t0) and second time point 2 y later (t1) were nonlinearly registered to the Montreal Neurological Institute MNI152 GM template and averaged together with their respective mirror images to create the left/right symmetric study-specific GM template. All the native GM volume images were then nonlinearly normalized onto this template. Then, the optimized FSL-VBM protocol introduces a compensation (or "modulation") for the contraction/enlargement caused by the nonlinear component of the transformation: each voxel of each registered GM image was multiplied by the Jacobian of the warp field. Finally, every modulated normalized GM volume image at time point t1 was subtracted from its respective modulated normalized GM volume image at baseline t0. All "t0-t1" GM volume images were then concatenated and smoothed with an isotropic Gaussian kernel with a sigma of 3 mm (~7 mm full width at half maximum). The percentage change in total GM volume was derived from the average GM partial volume values for each subject inside the entire GM mask at t0 and t1.

In reviewing the scans while still blind to the group membership of each participant, we had to exclude from the FSL-VBM analysis an additional 12 subjects for the following reasons: subcortical lacunar infarct, megadolicobasilar artery, meningioma, very focal atrophy (lateral fissure and intraparietal sulcus), insufficient gray/white matter contrast, and two instances for which "t0-t1" differences exceeded 3σ from the average of all subjects. Thus, scans from 156 subjects remained for analysis (87% of the 180 subjects scanned at the two time points).

Upon visual inspection, white matter hypointensities in T1-weighted images did not affect the nonlinear registration of the segmented images. Furthermore, they did not impact results, as they were outside the GM mask used for the statistical analysis (GM mask automatically generated following the FSL-VBM protocol and masked by the brain mask of the MNI152 template). On average, subjects with low and high tHcy levels had similar volume of WM "lesions" (~1.3 × 1.3 × 1.3 $cm^3$, as measured crudely by assessing the volume in the registered segmented images represented by the voxels above a partial volume value of 0.8 outside of the GM mask).

**Statistical Analysis.** Based on a priori criteria, a median split for tHcy (11.06 μmol/L) was used to ensure sufficient sample size in the two subgroups. We thus considered four groups in our statistical design: subjects in the placebo group with low level of tHcy at baseline (P_h1), subjects in the placebo group with high tHcy level (P_h2), subjects in the B-vitamin group with low tHcy level (B_h1), and subjects in the B-vitamin group with high tHcy level (B_h2).

We first looked for:

- The regions that showed significant changes over the 2 y separating the two scanning time points for the placebo group (*Results* and Fig. 1*A*) and the B-vitamin group (*Results* and Fig. 1*B*): (P_h1 + P_h2) and (B_h1 + B_h2),
- The regions where the "t0-t1" changes between the two time points in the placebo group were significantly different from the changes in the B-vitamin group (*Results*, Fig. 2, and Fig.

S1): (P_h1 + P_h2) − (B_h1 + B_h2) = (P_h1 − B_$h_1$) + (P_h2 − B_h2).

Then, examining the impact of baseline level of tHcy on the results, we looked for:

- The regions where, looking at the placebo group and at the B-vitamin group separately, the "t0-t1" changes in the subjects with high tHcy level were significantly different from the changes in the subjects with low tHcy (main text and Fig. S2): (P_h2 − P_h1) and (B_h2 − B_h1),
- The regions where, looking at the subjects with low tHcy level and with high tHcy separately, the t0 − t1 changes in the placebo group were significantly different from the changes in the B-vitamin group (Fig. 3, main text and Fig. S3): (P_h1 − B_h1) and (P_h2 − B_h2); and
- The regions where there was a formal interaction in the t0 − t1 changes between baseline tHcy level and treatment status (main text and Fig. S4): (P_h2 − P_h1) − (B_h2 − B_h1) = (P_h2 − B_h2) − (P_h1 − B_h1).

We also compared baseline GM volume between placebo and B-vitamin groups and found no significant difference [minimum familywise error (FWE), $P = 0.33$]. In a further test, we investigated the possible influence of GM volume at baseline on the second time point GM volume by regressing out the former from the latter voxelwise. We found extremely similar results between this approach and subtracting GM volume at baseline from the second timepoint as in our main analysis.

Baseline GM volume differences between low and high tHcy level groups were also investigated. There was no significant difference between low and high tHcy level groups after correction for multiple comparisons, although a trend ($P < 0.01$, uncorrected) could be seen with lower GM volume with higher tHcy level essentially in the left amygdalohippocampal complex.

Furthermore, similarly to our assessment of the impact of baseline tHcy levels on GM atrophy, we looked for the impact of carrying apolipoprotein E (ApoE) ε4 on GM atrophy. No results reached significance after correction for multiple comparisons. There was, however, a trend for higher atrophy in ApoE ε4 carriers in the placebo group, especially in the medial temporal lobe, fusiform and piriform cortex, which did not exist in the treatment group ($P < 0.01$, uncorrected), although a formal interaction between treatment and ApoE ε4 status was not significant.

Finally, for all subjects, we correlated "t0-t1" GM volume changes with the changes in global cognition and memory function using the clinical dementia rating–sum of boxes (CDR-SOB), Mini-Mental State Examination (MMSE), the Hopkins Verbal Learning Test (HVLT) delayed recall and the category fluency (animals).

To achieve accurate inference, including FWE correction for multiple comparisons across space, we used permutation-based nonparametric inference within the framework of the general linear model (5,000 permutations) (7). Results were considered significant for $P < 0.05$, FWE-corrected for multiple comparisons across space using a Threshold-Free Cluster Enhancement (TFCE) approach, which avoids the use of an arbitrary threshold for the initial cluster formation (8).

All results were identified using a combination of three complementary atlases: the Harvard–Oxford structural cortical probability maps based on MRI T1-weighted images, Jülich cytoarchitectonic probabilistic maps based on postmortem brains, and Talairach Daemon labels corresponding approximately to Brodmann areas (9, 10).

**Directed Acyclic Graph Analysis in Subjects with High tHcy Level.** We modeled nine different variables as a directed acyclic graph: (*i*) treatment category and changes ("delta") over the 2-y period in: (*ii*) plasma folate, (*iii*) plasma vitamin B12 and (*iv*) plasma tHcy, (*v*) GM volume, (*vi*) CDR-SOB, (*vii*) MMSE, (*viii*) HVLT delayed recall, and (*ix*) category fluency. Indeed, as changes in vitamin B6 were shown to have no impact on tHcy over time ($P = 0.19$), we considered changes only in folate and vitamin B12 concentrations in this analysis. To ensure that there was no bias in the GM measure, we considered the regions where all participants with high tHcy lost GM volume, regardless of their treatment status. We looked for the optimal Bayesian network explaining these variables using Greedy Equivalence Search (GES) (11) (www.phil.cmu.edu/projects/tetrad/), which allowed us to establish statistical conditional dependencies (edges) between these variables—primarily to estimate direct causal connections between variables. The Linear, Non-Gaussian, Acyclic causal Models (LiNGAM) (12) approach was then used to determine further the direction of the edges, i.e. the causality between the variables. We estimated the goodness of fit of the model using a $\chi^2$ test, whereby the null hypothesis was that our model of interest is equivalent to a saturated model with all possible connections between all variables. For this $\chi^2$ test, a large, nonsignificant $P$ value is evidence of good fit, in that the fitted model is indistinguishable from a saturated model with many more parameters. Moreover, we used a subsampling procedure to evaluate the stability of the direction of the edges in our optimized network. Using five random splits of the data (into two groups of 38 and 39 subjects), we created 10 reduced versions of the data. Holding fixed the undirected edges found in the complete dataset by GES, we found highly consistent directionality determined by PC-LiNGAM. Only one edge had less than 80% stability (directionality the same in eight out of 10 subsamples): "delta_b12" → "delta_tHcy" was identified five out of 10 times as "delta_tHcy" → "delta_b12".


1. de Jager CA, Oulhaj A, Jacoby R, Refsum H, Smith AD (2012) Cognitive and clinical outcomes of homocysteine-lowering B-vitamin treatment in mild cognitive impairment: A randomized controlled trial. *Int J Geriatr Psychiatry* 27(6):592–600.
2. Smith AD, et al. (2010) Homocysteine-lowering by B vitamins slows the rate of accelerated brain atrophy in mild cognitive impairment: A randomized controlled trial. *PLoS ONE* 5(9):e12244.
3. Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *J Intern Med* 256(3):183–194.
4. Good CD, et al. (2001) A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14(1 pt 1):21–36.
5. Douaud G, et al. (2007) Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain* 130(pt 9):2375–2386.
6. Smith SM, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(suppl 1):S208–S219.
7. Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Hum Brain Mapp* 15(1):1–25.
8. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1):83–98.
9. Eickhoff S, et al. (2005) High-resolution MRI reflects myeloarchitecture and cytoarchitecture of human cerebral cortex. *Hum Brain Mapp* 24(3):206–215.
10. Lancaster JL, et al. (2007) Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp* 28(11):1194–1205.
11. Ramsey J, Zhang J, Spirtes P (2006) Adjacency-faithfulness and conservative causal inference. *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence (UAI-2006)* (Association for Uncertainty in Artificial Intelligence), pp 401–408.
12. Shimizu S, Hoyer PO, Hyvarinen A, Kerminen A (2006) A linear non-Gaussian acyclic model for causal discovery. *J Mach Learn Res* 7:2003–2030.

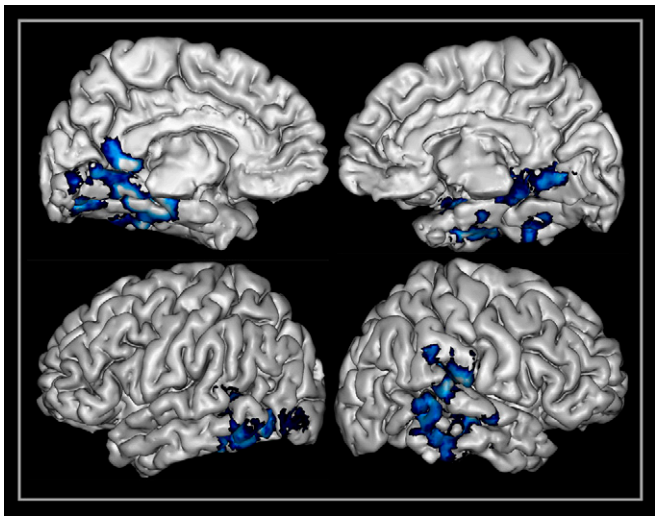

**Fig. S1.** B-vitamin treatment significantly reduces regional loss of GM. 3D rendering of brain regions in blue–white where B-vitamin treatment significantly reduces GM loss over the 2-y period ($P < 0.05$ FWE-corrected). All blue/white areas correspond to regions of significant loss in placebo and known to be vulnerable in Alzheimer's disease (Fig. 1).

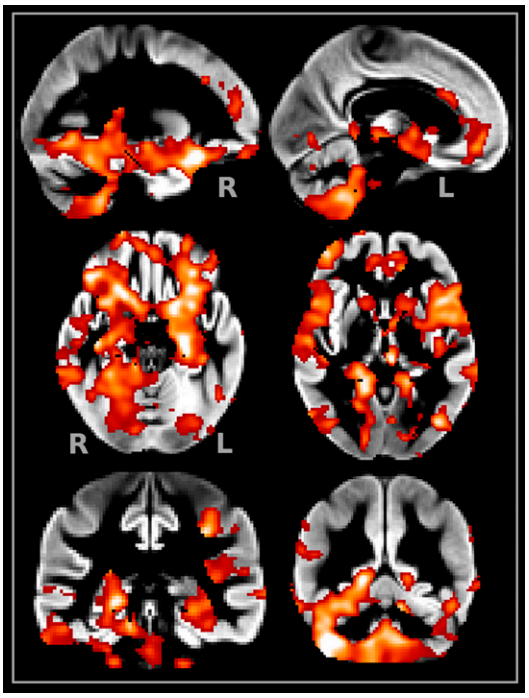


**Fig. S2.** Increased GM loss over the 2-y period in participants in the placebo group with higher tHcy compared with lower tHcy. The 76 participants in the placebo group were split into two categories depending on whether they had higher or lower tHcy levels than the overall median value (11.06 μmol/L). Participants with higher baseline tHcy levels ($n = 35$) had greater GM atrophy (red, $P < 0.05$ FWE-corrected) over the 2-y period compared with those with lower tHcy levels. In contrast, there was no difference in atrophy between participants with high and low baseline tHcy levels in the B-vitamin group.

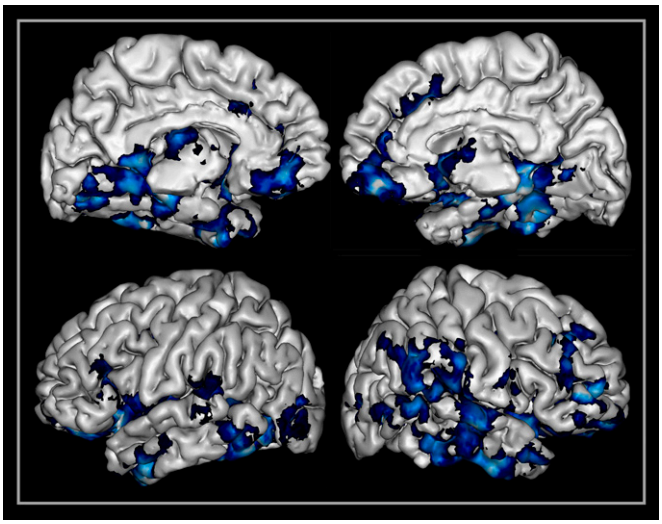

**Fig. S3.** B-vitamin treatment is effective only in participants with higher tHcy levels. 3D rendering of brain regions in blue–white where B-vitamin treatment significantly reduces GM loss in participants with high tHcy levels (>11.06 μmol/L) at baseline ($P < 0.05$ FWE-corrected). No significant effect of B-vitamin treatment was found in the participants with low baseline tHcy levels.

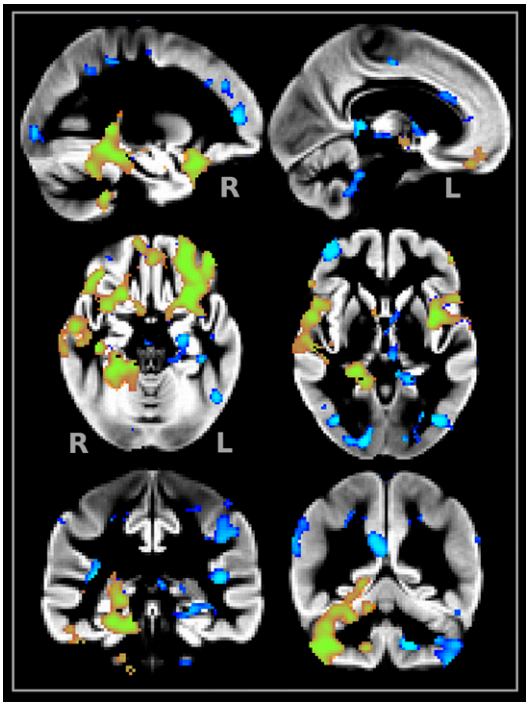


**Fig. S4.** Interaction between tHcy levels and treatment status in GM loss. The interaction between high and low tHcy levels and placebo and B-vitamin groups showed significant regions corresponding to where GM loss was reduced with treatment in participants with high tHcy levels (in red–yellow $P < 0.05$ FWE-corrected; in blue $P < 0.05$ uncorrected; overlap in green).

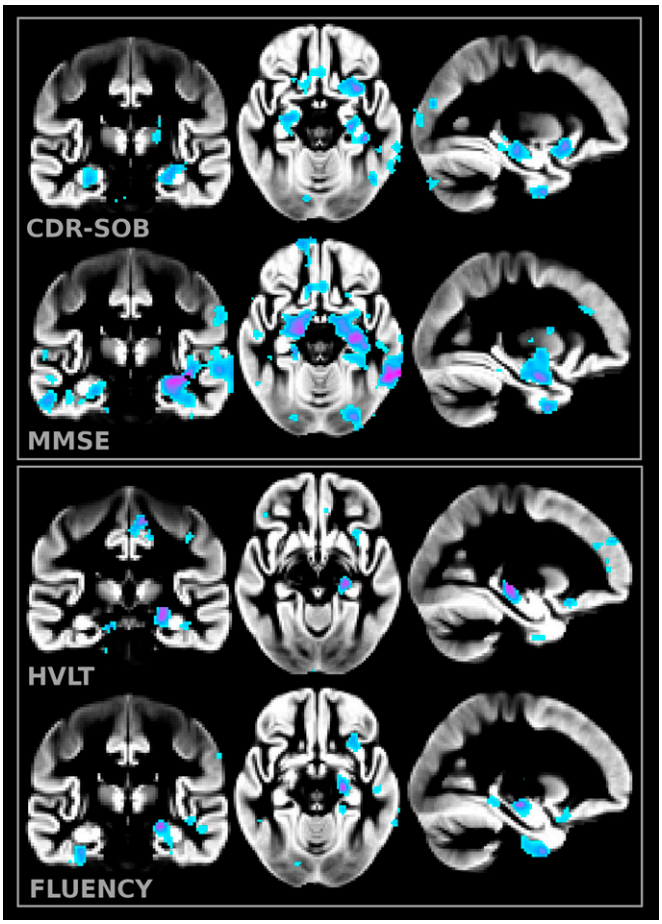


**Fig. S5.** Correlation between GM loss and cognitive decline. Significant correlation was found between GM loss over 2 y and global function mainly in the medial temporal areas, entorhinal and piriform cortex (CDR-SOB and MMSE, $P < 0.05$ FWE-corrected) and, at an uncorrected threshold, with memory function mainly in the left medial temporal areas (HVLT-R delayed recall and category fluency, $P < 0.001$).

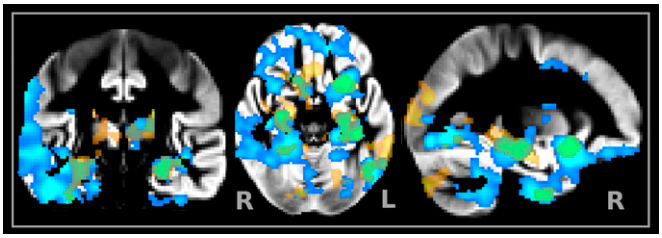


**Fig. S6.** Regions related to the global cognitive decline overlap with those benefiting from B-vitamin treatment. Correlation between GM loss and worsening of CDR-SOB scores (yellow) is overlaid onto the significant effect of B-vitamin treatment on atrophy in those with higher tHcy level (in blue, adapted from Fig. 3). Overlap is shown in green.