Correspondences between retinotopic areas and myelin maps in human visual cortex

Rouollah O. Abdollahi, Hauke Kolster, Matthew F. Glasser, Emma C. Robinson, Timothy S. Coalson, Donna Dierker, Mark Jenkinson, David C. Van Essen, Guy A. Orban

Abstract

We generated probabilistic area maps and maximum probability maps (MPMs) for a set of 18 retinotopic areas previously mapped in individual subjects (Georgieva et al., 2009 and Kolster et al., 2010) using four different inter-subject registration methods. The best results were obtained using a recently developed multimodal surface matching method. The best set of MPMs had relatively smooth borders between visual areas and group average area sizes that matched the typical size in individual subjects. Comparisons between retinotopic areas and maps of estimated cortical myelin content revealed the following correspondences: (i) areas V1, V2, and V3 are heavily myelinated; (ii) the MT cluster is heavily myelinated, with a peak near the MT/pMSTv border; (iii) a dorsal myelin density peak corresponds to area V3D; (iv) the phPIT cluster is lightly myelinated; and (v) myelin density differs across the four areas of the V3A complex. Comparison of the retinotopic MPM with cytoarchitectonic areas, including those previously mapped to the fs_LR cortical surface atlas, revealed a correspondence between retinotopy and myelin density map landmarks in human visual cortex.

© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

Introduction

Cortical areas are one level in the hierarchical organization of the primate cerebral cortex. Human neocortex has been estimated to contain 150–200 areas, perhaps 20–30% more than in the intensively studied macaque monkey (Van Essen et al., 2012a, 2012b). In monkeys, cortical areas have been defined using four general approaches: architectonic organization, anatomical connections, topographic organization, and functional properties, but congruence across all approaches has been demonstrated for only a few visual areas such as V1 and MT (Felleman and Van Essen, 1991). For human cortex, informative approaches to parcellation include post-mortem cytoarchitectonics (Amunts et al., 2000), post-mortem myeloarchitectonics (Nieuwenhuys, 2013), in-vivo myelin mapping (Glasser and Van Essen, 2011), task-activation fMRI (Tootell et al., 1995) resting-state fMRI (Mantini et al., 2012), and tractography based on diffusion imaging (Johansen-Berg et al., 2004). Finally, mapping of topographic organization, especially retinotopy, has proven highly successful in fMRI studies of human and nonhuman primates (Brewer et al., 2002; Fize et al., 2003; Kolster et al., 2009; Sereno et al., 1995; Wandell et al., 2007). In nonhuman primates, appropriate fMRI methods can map even areas with coarse retinotopic organization because fMRI averages over a large number of neurons (Kolster et al., 2010). The present study capitalizes on maps of 18 retinotopic areas in twelve healthy adults obtained in recent human fMRI studies (Ferrera et al., 2012; Georgieva et al., 2009; Kolster et al., 2010). By using novel surface-based methods to generate probabilistic area maps and maximum probability maps, we obtained valuable information about the most likely location of each area as well as its variability in location and extent. We compared these retinotopic maps to population-average cortical myelin maps derived from the Human Connectome Project (HCP: http://www.humanconnectome.org). The myelin maps were based on the ratio of T1w to T2w intensity, which reveals patterns that are closely correlated with cytoarchitectonic areas, particularly in

http://dx.doi.org/10.1016/j.neuroimage.2014.06.042
1053-8119/© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).
early sensory and motor areas (Glasser and Van Essen, 2011; Glasser et al., 2014). A previous study used quantitative T1 to identify high myelination content in eight retinotopic areas studied in six subjects (Sereno et al., 2013). The present study provides a more extensive analysis involving more retinotopic areas (18) and a larger number of subjects for each modality. We did not examine the laminar pattern of myelination, which has been analyzed in other post-mortem studies, but to date has proven diagnostic in vivo only for the stria of Gennari in V1 when examined at high spatial resolution and with less than full brain coverage (Clare and Bridge, 2005; Geyer et al., 2011; Hinds et al., 2008, 2009; Sanchez-Panchuelo et al., 2010). We did compare the retinotopic areas to a V1 atlas based on the stria of Gennari (Hinds et al., 2008) and to post-mortem cytoarchitectonic maps (Caspers et al., 2013), thereby extending the approach of others (Wilms et al., 2010; Wohlschläger et al., 2005) to a larger number of cortical areas.

Critical issues in inter-subject comparisons include the choice of an atlas target and the method(s) used to align individuals to the atlas (Yeo et al., 2010). We used the fs_LR atlas template to bring the two hemispheres into geographic correspondence (Van Essen et al., 2012b), then compared four independent methods for registration data to the fs_LR atlas. The PALS landmark-based registration (Van Essen, 2005) and FreeSurfer’s energy-based registration are constrained by geometric features related to cortical folding. Folding-constrained registration was also performed using a multimodal surface matching (MSM) algorithm that provides greater configuration flexibility and allows registration to be driven by different modalities and their combinations (Robinson et al., 2013). Finally, MSM was also used to directly align the retinotopic areas themselves, an example of areaal-feature-based registration (using features that identify or characterize cortical areas).

Our results provide valuable insights regarding the relationship between retinotopic areas, cytoarchitectonic areas, and patterns of cortical myelination. In addition, this freely available dataset provides a general reference framework for comparison with other human neuroimaging studies, past and present. Retinotopic mapping in individual subjects is often not feasible even when the location of visual areas is of high interest. Having a set of 18 retinotopic areas accurately mapped to the cortical sheet in a standard atlas space will likely be useful for a variety of analyses.

Material and methods

Single subject data

We used retinotopic-mapping data acquired using phase encoding of eccentricity and polar angle obtained from 12 healthy adults of either sex, as shown in Fig. 1 for an exemplar subject. An additional example is shown in figure S1 of Kolster et al. (2010). All 18 areas were identified, as described by Georgieva et al. (2009) and Kolster et al. (2010), in the 12 subjects. Ten subjects participated in the Kolster et al. study (and a subset of them in Georgieva et al. (2009)); the two remaining are taken from the study by Ferri et al. (2012). The visual stimuli used for retinotopic mapping extended from 0.25 to 7.75 degrees in eccentricity along both the horizontal and vertical axes (Kolster et al., 2010). Thus, the retinotopic maps for each area were incomplete and should span about half of the total extent of each area, at least for early visual areas (see Hinds et al., 2008). The functional data were mapped to the native mesh of each subject’s cortical surface using FreeSurfer. The estimated extent of each visual area was drawn manually on cortical flat maps based on independent mapping of polar angle and eccentricity. This resulted in a set of FreeSurfer (version 5.1) surface patches, with 18 retinotopic areal labels for each hemisphere of each of the 12 subjects, including: V1, V2, V3, hV4, VO, LO1, LO2, V3A, V3B, V3C, V3D, V7, MT, pMSTv, pFST, pV4t, phPITd, phPITv, where ‘p’ stands for putative and ‘h’ for human (Georgieva et al. (2009) and Kolster et al. (2010)). Vertices with overlapping labels (worst case 7% +/-3%) were reassigned to the most common label of neighboring vertices or zero in the case of ties. We refer to the four areas V3A-D as the V3A complex. Areas V3C/D of our scheme corresponds largely to V3A/B (Larsson and Heeger, 2006), although our V3C and V3D are comparable in size whereas their V3A was generally larger than V3B. V3A/B in our scheme largely corresponds to a gap between their V3A/B and LO1 (which frequently included a central representation) and to the dorsal part of their LO1, which

Fig. 1. The 18 retinotopic areas defined in the polar angle (A) and eccentricity (B) maps by Georgieva et al., 2009 and Kolster et al., 2010; right hemisphere of subject 1. Stars: central visual field, purple: eccentricity ridge, white dotted lines: horizontal meridian (HM), black full and dashed lines: lower and upper vertical meridian (VM). In case of small regions devoted to central vision, eg the center of the V3C–D cluster the central representations are affected by large prF sizes and do not reach the smallest values (hence remain yellow or green). Hence, the central representations of the visual field were identified as positions where a local minimum in eccentricity and representations of the upper VM, lower VM, and HM coincide.
included a proportionally larger central part than their LO2 (cf. Figure 2 of Larsson and Heeger, 2006).

**Templates and registration methods**

The Freesurfer data were converted to GIFTI format, and the cortical mid-thickness surface (average of pial and white matter surfaces) was generated for the native mesh of each subject. To enable inter-subject analyses, we transformed all single subject data to a common surface template, the fs_LR atlas surface previously derived by registering the two hemispheres of FreeSurfer's fsaverage template (Van Essen et al., 2012b). This process also entailed nonlinear registration of the individual brain volume to the MNI152 volume space using FreeSurfer, and application of this transformation to the cortical surfaces so that they were also in nonlinear volume space (Glasser et al., 2013b). We compared four methods to register the retinotopic data to the common surface-based template. Two recently introduced methods are: (1) folding-based surface registration to fs_LR performed using the Multimodal Surface Matching algorithm (MSM, Robinson et al., 2014), termed MSM-sulc throughout, and 2) retinotopic feature-driven registration performed using MSM (MSM-retinotopy). Two older methods are; (3) Landmark-based registration to the PALS template and then landmark-based inter-atlas registration of PALS to fs_LR (PALS) (Van Essen, 2005) and (4) FreeSurfer folding-based registration to fsaverage (Fischl et al., 1999) followed by landmark-based inter-atlas registration of fsaverage to fs_LR (Van Essen et al., 2012b). These methods used multiple stages of registration to align the native spherical mesh to the fs_LR template sphere and also to compensate for a “drift” that, if uncorrected, can create undesirable biases in group average results. To avoid erosion of spatial fidelity that can occur during multiple resampling steps, we concatenated successive spherical registrations (by a process of projecting to one sphere and unprojecting from a deformed version of that sphere) into a single registered sphere that was used for one-step mapping of native-mesh data to the fs_LR atlas. We first describe the two MSM methods followed by the PALS and Freesurfer methods to demonstrate how the results from all four methods were made directly comparable to one another.

**MSM sulc: single subject → fs_LR**

The multi-modal surface matching (MSM) algorithm is a fast, highly tunable generic registration algorithm that uses a discrete optimization strategy (Robinson et al., 2013, 2014) to align cortical surfaces. MSM was used to perform folding-based registration directly from the native mesh to the fs_LR template prior to running MSM, the affine (rigid rotational) component of the FreeSurfer registration was extracted using a linear regression. This compensates for the significant rotation between the native and fs_LR mesh. The linear regression minimizes the sum of squares of the coordinate differences between the original native spherical mesh and the concatenated native spherical mesh that was registered in the FreeSurfer methods. This affine matrix was used to rotate the original native sphere and initialize the MSM registration.

For MSM folding-based registration, we used average convexity (“sulc”, analogous to sulcal depth), as does Freesurfer's registration. Each subject's sulc map was registered to the group average sulc map using a normalized cross correlation cost function. For the MSM-sulc folding-based registration, the configuration parameters were optimized to produce an accurate geographic registration of major sulci (those that are consistently present across subjects) while minimizing areal distortion across the surface. This was parameterized by increasing the contribution of the regularization to the registration (by increasing the weighting parameter lambda Robinson et al., 2013, 2014); it prevents the registration from over-fitting incompatible folding patterns between the individual subjects and the template. Indeed, outside of primary areas such as V1, folding patterns are highly variable across subjects, and they are often poorly correlated with cortical areal boundaries (Fischl et al., 2008; Van Essen et al., 2012b), though there are some exceptions (Weiner et al., 2014; Witthoft et al., in press; see Discussion). Thus, the MSM-sulc method is designed to achieve approximate folding based alignment of cortical geography without the increases in distortion seen with FreeSurfer. The ‘native-to-fs_LR-via-MSM-sulc’ sphere was used for one-step resampling of the 18 retinotopic areas of each hemisphere to the fs_LR template mesh. This registration was also used to initialize Retinotopic alignment.

**MSM retinotopy: single subject → fs_LR**

The MSM algorithm was also used to perform an areal-feature-based registration to further improve the inter-subject alignment of retinotopically defined cortical areas, using the retinotopic areas as features. Retinotopic area labels were converted to a set of binary surface maps (one map per label) with values of one inside the area and zero outside. This generated an 18 dimensional feature vector for each vertex. These were registered to analogous surface maps from the group-defined MPM (see description below) by using the normalized mutual information (NMI) cost function (Robinson et al., 2014). The normalized mutual information cost function was chosen in this case because it performed better than the normalized correlation coefficient cost function on this kind of unsmoothed binary data.3

Because there were significant individual differences in the location and even the topological arrangement of retinotopic areas (see below), MSM retinotopic registration was tuned to tolerate greater distortion (low regularization lambda) and achieve better cross-subject areal alignment as a result. The group-defined MPM was initially generated from the MSM-sulc method and again after each of three iterations of MSM-retinotopy registration and template refinement. Three iterations were judged to be sufficient because there was little further improvement in template sharpness in the third iteration relative to the second. The multiple template generation and registration steps that follow the MSM-sulc folding registration generally resulted in some “drift” in the group average maps, creating a mismatch in areal size, shape and position between the group averages and the typical individual. We calculated this effect using the component of registration occurring after MSM-sulc (represented by a deformed sphere) and averaging the coordinate positions of the resampled deformed spheres across subjects. The result was a group average deformed sphere that reflects the group average drift in registration. The inverse of the drift was calculated and concatenated onto the end of each subject’s MSM-retinotopy registration, removing the drift from all subjects and yielding unbiased group average maps. Finally, the 18 retinotopic areas of each hemisphere were transformed from their native mesh to the fs_LR template in a single resampling step that combined MSM-sulc, MSM-retinotopy, and DeDrifting.

**PALS: single subject → PALS → fs_LR**

The PALS registration process was applied as described previously (Freesurfer_to_PALS-B12_Pipeline_Distribution, Anticevic et al., 2012). The process automatically generated the Core 6 landmarks, which were manually edited when necessary. After projection to the individual’s spherical mesh, these landmarks were used to register the individual to the PALS atlas target sphere (Anticevic et al., 2012; Dierker et al., 2013; Van Essen, 2005). The registered (deformed) native mesh sphere was concatenated with the ‘PALS-to-fs_LR’ registered sphere that represents a previously generated landmark-based registration between the PALS and fs_LR templates (Van Essen et al., 2012b).

---

3 We would not necessarily expect other kinds of continuous data (e.g. myelin maps) to perform well with normalized mutual information and do not recommend NMI for data having fewer features than used here. The default choice for most continuous data sets should be the normalized correlation coefficient cost function.
The group average drift between the PALS registration and MSM-sulc registration was calculated and removed. The resultant ‘native-to-fs_LR-via-PALS-and-DeDrifted’ registered sphere was used for one-step resampling of the 18 retinotopic areas of each hemisphere of each subject from the native mesh to the fs_LR template mesh in order to generate unbiased group average maps.

**FreeSurfer: single subject – >fsaverage – >fs_LR**

FreeSurfer’s folding-based registration also uses the ‘sulc’ folding map to align individual subjects to the fsaverage template, by creating an output registered sphere that is geographically aligned with fsaverage (Fischl et al., 1999). This ‘native-to-fsaverage’ sphere was concatenated to the ‘fsaverage-to-fs_LR’ registered sphere representing the landmark-based registration between the fsaverage and fs_LR templates (Van Essen et al., 2012b). The group average drift between the FreeSurfer registration and MSMSulc registration was calculated and removed. The resultant ‘native-to-fs_LR-via-fsaverage-and-DeDrifted’ registered sphere was used for one-step resampling of the 18 retinotopic areas of each hemisphere from the native mesh to the fs_LR template mesh in order to generate unbiased group average maps.

**Probabilistic area maps [PAMs] and maximum probability maps (MPMs)**

For each surface vertex and for a given surface registration method, we calculated the probability of belonging to each of the 18 visual areas based on the number of subjects having that area’s label at that vertex, making a probabilistic area map (PAM). While the probability area map provides the probability that each vertex belongs to a given area, it does not consider the probabilities of neighboring areas. Therefore, we calculated a maximum probability map (MPM), which for each vertex indicates the area to which it most likely belongs, given the registration method used. In addition, we created a “no area assigned” label in each hemisphere of each individual for all parts of cortex that did not have an areal assignment, and created also a PAM for this label. In the MPM, each vertex was assigned to the label that is most probable for that vertex (including the no area assigned option). If the maximum probability was ambiguous (two or more areas with equal maximum probabilities) for any given vertex, we used two rules. If one of the areas was the ‘no area assigned’ area this one always lost. For the neighboring retinotopic areas, we used neighboring vertices to choose the most probable area. To remove any remaining ambiguities (fewer than 1% of the vertices), smoothed versions of the probability atlases (Gaussian geodesic smoothing with sigma = 2 in Caret) were used. Since we included a no area assigned label, there was no need to apply a probability threshold as previously done for cytoarchitectonic MPMs (Amunts et al., 2000; Eickhoff et al., 2005).

**Group average surfaces**

Surface areas measured on native meshes in a subject’s native volume space represent our best estimate of the actual surface areas in that subject’s brain, and the average of these surface areas represents the surface area in a typical subject. Group-average mid-thickness surfaces were created separately for each registration method in order to calculate surface areas of retinotopic parcels. Prior to averaging across subjects, surfaces were transformed nonlinearly to FSL’s MNI152 template space using the warp field generated by an independent nonlinear volume registration step using FSL’s FNIRT (Glasser et al., 2014). The surface and volume registrations were independent insofar as the non-rigid volume transform was applied after the anatomical surfaces, spheres, and ‘sulc’ folding maps were generated in native volume space. Applying nonlinear volume registration to the surfaces likely reduces the variance of surface coordinate positions in 3D space when creating group average surfaces. Compensating for two types of distortion when computing group average surface area of identified cortical areas

When computing surface areas for retinotopic parcels represented on the group average midthickness, we corrected for two kinds of distortion in order to compute surface areas that are comparable to those measured on subjects’ native meshes in their undistorted native volume space (i.e., with nothing more than rigid-body alignment of the original data). The first source of distortion is the registration to MNI space, which has substantial drift owing to the way the MNI standard space was originally created (it did not include a compensatory DeDrifting process as described above for surface registration). MNI space brain volumes are ~40% larger than native space brain volumes (based on comparison of 196 HCP subjects’ native volume space brain masks to their MNI space brain masks; see also Van Essen et al., 2012b). A second source of distortion is the spatially nonuniform shrinkage of surface area when creating an average midthickness surface. This occurs because the group average midthickness has a surface area that is substantially less than the typical individual’s midthickness surfaces, especially in regions of high intersubject variability (because topologically incompatible individual differences in folding patterns are averaged out).

To correct for these two sources of distortion the following procedure was carried out after completion of the preceding analyses: 1) The MNI volume registration portion of the correction was calculated as follows (for each vertex); $VD = (\text{Individual Native Mesh Midthickness Native Volume Space Vertex Area})/(\text{Individual Native Mesh Midthickness MNI Volume Space Vertex Area})$. 2) The MNI registered sphere was used for one-step resampling of the 18 retinotopic areas of each hemisphere to the MNI space (i.e., with nothing more than rigid-body alignment of the original data). The $VD$ map was resampled from the native mesh onto the 164k Mesh. 3) The midthickness averaging portion of the correction was calculated as follows (for each vertex): $SA = (\text{Individual 164k Mesh Midthickness MNI Volume Space Vertex Area})/ (\text{Group Average 164 k Mesh Midthickness MNI Volume Space Vertex Area})$. 4) For each subject (on the 164k mesh), the vertex-wise $VD$ and SA measures were multiplied: $VD * SA = DL$. 5) DL was averaged across subjects to produce a vertex-wise group average distortion measure (DG) that corrects for both sources of distortion. 6) For each parcel (area), the DG map was set to zero outside the parcel. 7) The masked DG maps were integrated across vertices to calculate a corrected surface area for each parcel. This resulted in 18 values, each containing the surface area for a retinotopic parcel corrected for the various distortions. These areas are comparable to the native mesh, native volume space areas, having been corrected in a vertex-wise fashion for both sources of distortion/areal change. These distortion corrections were performed separately for each of the four registration methods (the $VD$ term stays the same but the SA term depends on the surface registration method).

**Comparing with myelin maps**

We compared the retinotopy data with group-average myelin maps from 196 Human Connectome Project subjects (Glasser et al., 2013b; Smith et al., 2013) having complete myelin maps and fMRI data from the data release for the first three quarters (Q1 + Q2 + Q3) after registration using MSM that were generated as a part of a separate study. The MSM registration used was MSM Sulc + Myelin + Resting State Network + DeDrift (Glasser et al., 2013b; Robinson et al., 2013, 2014), i.e., using myelin and resting state maps as areal features. While myelin maps provide excellent contrast for cortical areas in some regions, resting state network maps span the entire hemisphere (Glasser et al., 2013a) and hence improve the overall alignment quality related to cortical function. Resting state networks were not available for the registration of retinotopic areas. Because different modalities were available for aligning the retinotopy and myelin-map datasets, the cross-study alignment (alignment of group averages) relied on the MSM-sulc folding-based component of the registration and the removal of group average registration drift. In contrast, the within-study inter-subject alignment (sharpness of group averages) benefited from (different) areal features.
(retinotopy or myelin content + resting state networks, respectively) in the two approaches.

Comparison of the retinotopic MPM with architectonically defined areas

We also made comparisons to nine published occipital cytoarchitectonic areas (hOc1, 2, 3d,v, 4d, 4v, 5, FG1–2) aligned across subjects using volumetric registration. To compute the MPM of these areas, the PAMS in Colin Space provided by Juelich (Eickhoff et al., 2005) were mapped onto the surface of fs_LR as follows. The Colin brain was registered nonlinearly to the MNI–152 template, using FNIRT in fsl_anat. This transformation was applied to the probability volumes, bringing them into nonlinear MNI space. The 3D volumes of the 12 retinotopic subjects were each nonlinearly registered to MNI space again using FNIRT in fsl_anat. We applied the reverse transformations of the MNI-to-single-subject to the probability volumes in MNI space, thereby transferring each of the probability volumes from the Colin brain to each of the single retinotopic subjects' brains. We then mapped the probability maps from the 3D volume of each single subject to the cortical surface of that subject. This entailed attributing to each vertex of the native midthickness mesh the average of the probabilities in the voxels along surface normal through that vertex, using the volume to surface fraction projection average Freesurfer command (0.1 steps). Hence the cytoarchitectonic data were at that stage in the same space as the original retinotopic patches. They were registered to the fs_LR template using FreeSurfer registration (see above) between the subjects to ensure consistency with the pre-aligned regions (hOc1, hOc2 and hOc5, see below). This method aims to reduce the biases inherent in volume to surface mapping of group average volume data to group average surface data. However, it does not account for drifts in the volume-based registration, nor does it compensate for the effects of volume-based intersubject averaging. The probability maps of the nine cytoarchitectonic areas were averaged over the 12 subjects and the MPM was generated as described above for the retinotopic areas.

To evaluate the effect of volume intersubject averaging, we also compared our retinotopic MPM to probabilistic surface-based maps of three cyto-architectonic areas (hOc1, hOc2, and hOc5) previously mapped from surface reconstructions of individual post-mortem brains to the fsaverage surface (Fischl et al., 2008) and then to the fs_LR mesh (Van Essen et al., 2012b). These probability maps could be compared to our retinotopic MPM, even though surface-based group average registration drift from FreeSurfer could not be removed, because lack of access to individual data.

Finally, V1 has a prominent anatomical feature, the stria of Gennari, which can be identified using high resolution ex vivo MRI. Hinds et al. (2008) developed an algorithm to match the cortical folding pattern of individual subjects to their probabilistic V1 defined ex vivo. We used FreeSurfer’s automated version of this algorithm (Hinds et al., 2009) to estimate the extent of V1 in each hemisphere. These maps were registered to fs_LR to generate a probability map of V1 in each hemisphere for the group of 12 subjects, using two different thresholds to define V1: 8%, and 50% (see supplementary information).

Results

Size and variability of retinotopic areas

The surface patches corresponding to the 18 different retinotopic areas differed greatly in size. Table 1 (columns 2, 3) reports the average size (in mm² on the native mid-thickness surface of each individual) of the mapped portions of each visual area for both hemispheres separately, as well as the standard deviation across subjects. Average surface areas spans more than a factor of 10 from the largest area (V2, 1567 mm² average for the two hemispheres) to the smallest (pV4t, 140 mm²). These data serve as a reference for the typical sizes of the central visual field representations of retinotopic areas in individual subjects. The maximal eccentricity mapped in our subjects (7.75°) lies in between the 12° used by Dougherty et al. (2003) and the 5.5° of Schira et al. (2009). This is important when evaluating the relative sizes of V1–V3, since the dependence of magnification factor on eccentricity differs between the 3 areas (Schira et al., 2009).

The coefficient of variation (SD/mean) is around 0.2 for large areas, such as V1, V2, V3, but reaches values of 0.5 or larger for some of the smaller areas (Table 1, column 4). It is likely that methodological ‘noise’ is greater for areas that are smaller and less well organized retinotopically, but it is possible that intersubject variability is greater for small than large areas and also contributes to the observed difference.

The average surface areas for the original area patches were also expressed as a percentage of the cortical surface area in individual hemispheres (on the fs_LR mesh). Even the V1 patch represents only 1.3 % of the total surface, but of course only a limited range of eccentricities (0–7.75°) was mapped; the actual extent of V1 and perhaps other early visual areas is likely to be about 2-fold greater (see below and Discussion). In total, all mapped retinotopic areas represent just less than 8% of the cortical surface in each hemisphere.

Besides the variability in area size documented in Table 1, there is also variability in neighborhood relationships (topology) for some areas that are part of distinct retinotopic clusters (Kolster et al., 2009, 2010; see Discussion). This is particularly evident for the two clusters in the V3A complex: V3A–B, a p-cluster (shared periphery), and V3C–D, a c-cluster (shared center). Fig. 2 illustrates variable topology in the V3A complex, with the number of areas adjoining dorsal V3 ranging from 2 (Fig. 2B) to 4 (Fig. 2C), with two variants for 3 neighbors (Figs. 2A, D). This may reflect independent rotation of each cluster in individual subjects and even hemispheres (Kolster et al., 2009, 2010).

### Table 1

| Area      | LH Indv Mean ± SD | RH Indv Mean ± SD | Av CV (Mean ± SD) | LH MPM Mean ± SD | RH MPM Mean ± SD |
|-----------|------------------|-------------------|------------------|------------------|------------------|
| V1        | 1329 ± 310       | 1305 ± 263        | 0.22             | 1398 ± 133       | 1333 ± 244       |
| V2        | 1524 ± 258       | 1567 ± 214        | 0.15             | 1509 ± 1576      | 1572 ± 1458      |
| V3        | 1306 ± 181       | 1394 ± 313        | 0.18             | 1418 ± 1403      | 1403 ± 1403      |
| VO1       | 381 ± 115        | 399 ± 123         | 0.31             | 336 ± 274        | 374 ± 274        |
| hV4       | 494 ± 166        | 573 ± 118         | 0.27             | 489 ± 595        | 595 ± 595        |
| pHPTv     | 202 ± 101        | 204 ± 139         | 0.59             | 205 ± 177        | 177 ± 177        |
| pHPTd     | 189 ± 110        | 142 ± 63          | 0.51             | 202 ± 154        | 154 ± 154        |
| LIO2      | 339 ± 149        | 267 ± 86          | 0.38             | 365 ± 315        | 315 ± 315        |
| LIO1      | 254 ± 121        | 291 ± 110         | 0.43             | 267 ± 331        | 331 ± 331        |
| V3A       | 201 ± 51         | 183 ± 64          | 0.35             | 208 ± 181        | 181 ± 181        |
| V3B       | 183 ± 44         | 172 ± 76          | 0.34             | 182 ± 162        | 162 ± 162        |
| V3C       | 259 ± 64         | 251 ± 73          | 0.27             | 246 ± 267        | 267 ± 267        |
| V3D       | 315 ± 95         | 360 ± 115         | 0.31             | 323 ± 360        | 360 ± 360        |
| V7        | 231 ± 61         | 227 ± 79          | 0.31             | 235 ± 198        | 198 ± 198        |
| pV4t      | 143 ± 48         | 140 ± 66          | 0.33             | 153 ± 153        | 153 ± 153        |
| MT        | 224 ± 84         | 339 ± 135         | 0.39             | 214 ± 340        | 340 ± 340        |
| pMST      | 192 ± 53         | 181 ± 99          | 0.41             | 177 ± 176        | 176 ± 176        |
| pFST      | 151 ± 80         | 107 ± 50          | 0.55             | 85 ± 56          | 56 ± 56          |

Probabilistic area maps

The probabilistic area map (PAM) of a given area reflects the degree of consistency in size, shape and location across subjects after registration to an atlas by a particular method. As shown in Fig. 3 for three exemplar areas (V1, MT, and pFST) the distribution of the PAM on the surface represents a two-dimensional step function with a central maximum or plateau surrounded by a progressive step-wise decline whose steepness depends on the residual variation in size, shape, and location of the area (Fig. 3, leftmost column for FreeSurfer registration and third column for MSM-retino registration). These PAMs were further characterized using the contours for maximum probability, 50% probability, and the minimum probability threshold (8%), which corresponds to
the maximum possible extent of the PAM. These contours are shown in the second and fourth columns of Fig. 3 in increasingly dark shades of blue (V1, top row), orange (MT, middle row) and green (pFST, bottom row).

When subjects were registered with the standard FreeSurfer folding-based method (left columns), the probability reached a maximum of 100% for V1, but not for MT and pFST. Thus the maximum overlap includes only about half the subjects for both MT and pFST. While the 8% probability size is relatively similar in MT and pFST, the 50% probability size is larger in MT than pFST, and the 50%/8% ratio differs between the two areas. These differences are presumably due in part to differences in areal size (it is harder to achieve high overlap for small areas), and also to differences in alignment quality (MT and pFST are in regions of high folding variability across subjects and in relation to areal boundaries that are difficult to align based only on folding patterns).

Comparing the two right columns to the two left columns of Fig. 3 shows improved registration by MSM-retino (i.e., using the retinotopic areas themselves as features) relative to FreeSurfer-based registration: the area size at 50% probability is increased for all three areas; the 50%/8% ratio also increases, indicating tighter alignment across all subjects. This increase reflects not just an increased extent of the 50% contour, but also a reduced extent of the 8% contour. The maximum probability reaches 100% in MT and V1, but not pFST.

Fig. 4 extends this type of comparison systematically and quantitatively, by plotting three different characteristics of probability area maps as a function of the original average area size for each of the four registration methods. The top row shows that the four methods differ markedly in the curves relating maximum probability of an area to its size. Of the three methods based purely on shape/folding, PALS performs the worst, not even achieving perfect overlap for the largest areas (V1, V2, and V3), whereas FreeSurfer and MSM-sulc perform comparably well. MSM-retino performs easily the best, achieving 100% probability for medium-sized as well as larger areas, and having larger overlap for even the smallest areas (phPITd, phPITv, pV4t, and pFST in both hemispheres, pMST on the right and LO2 on the left). These results demonstrate that MSM-retino performed as expected, because the retinotopic areas themselves were used to drive the registration. For the 50%/8% threshold ratio (middle row) MSM-retino again performs the best and PALS most poorly: the ratios are highest for MSM-retino.
The plots of PAM size at 50% threshold as a function of the original size (bottom row) are linear for all four methods. The PALS registration is the only one in which 50% area size markedly undershoots the original size. The slope of the linear function for the PALS registration is significantly smaller than that for FreeSurfer ($t = 3.1, p < 0.05$). Moreover, the linear fit intercepts the x-axis well to the right of the origin for all but MSM-retino (see legend for values). The intercept of the 50% function for the MSM-retino was significantly different compared to that for the PALS ($t = 6.3, p < 0.01$), FreeSurfer ($t = 7.3, p < 0.01$) or MSM-sulc registration ($t = 5.13, p < 0.02$). Thus, MSM-retino is the only method that preserves a consistently positive 50% size for the smallest retinotopic areas. These results demonstrate that MSM-retino is the optimal method among those tested, PALS being worst and the FreeSurfer and MSM-sulc being comparable and in between. Notably, the non-linear functions differentiate amongst the four methods more readily than the linear one relating 50% size to the original size.

The PAMs of the different cortical areas clearly indicate that MSM-retino is the optimal registration method. This was further supported by two additional analyses. First, in a 'leave-one-out' procedure probabilistic area maps were generated using 11 subjects and the remaining subject was used as a test of the PAMs (see Supplementary information). This yielded much higher average (across areas) overlaps for MSM retina (see legend for values). The intercept of the 50% function for the MSM-retino was significantly different compared to that for the PALS ($t = 6.3, p < 0.01$), FreeSurfer ($t = 7.3, p < 0.01$) or MSM-sulc registration ($t = 5.13, p < 0.02$). Thus, MSM-retino is the only method that preserves a consistently positive 50% size for the smallest retinotopic areas. These results demonstrate that MSM-retino is the optimal method among those tested, PALS being worst and the FreeSurfer and MSM-sulc being comparable and in between. Notably, the non-linear functions differentiate amongst the four methods more readily than the linear one relating 50% size to the original size.

The plots of PAM size at 50% threshold as a function of the original size (bottom row) are linear for all four methods. The PALS registration is the only one in which 50% area size markedly undershoots the original size. The slope of the linear function for the PALS registration is significantly smaller than that for FreeSurfer ($t = 3.1, p < 0.05$). Moreover, the linear fit intercepts the x-axis well to the right of the origin for all but MSM-retino (see legend for values). The intercept of the 50% function for the MSM-retino was significantly different compared to that for the PALS ($t = 6.3, p < 0.01$), FreeSurfer ($t = 7.3, p < 0.01$) or MSM-sulc registration ($t = 5.13, p < 0.02$). Thus, MSM-retino is the only method that preserves a consistently positive 50% size for the smallest retinotopic areas. These results demonstrate that MSM-retino is the optimal method among those tested, PALS being worst and the FreeSurfer and MSM-sulc being comparable and in between. Notably, the non-linear functions differentiate amongst the four methods more readily than the linear one relating 50% size to the original size.

The PAMs of the different cortical areas clearly indicate that MSM-retino is the optimal registration method. This was further supported by two additional analyses. First, in a 'leave-one-out' procedure probabilistic area maps were generated using 11 subjects and the remaining subject was used as a test of the PAMs (see Supplementary information). This yielded much higher average (across areas) overlaps for MSM retina (see legend for values). The intercept of the 50% function for the MSM-retino was significantly different compared to that for the PALS ($t = 6.3, p < 0.01$), FreeSurfer ($t = 7.3, p < 0.01$) or MSM-sulc registration ($t = 5.13, p < 0.02$). Thus, MSM-retino is the only method that preserves a consistently positive 50% size for the smallest retinotopic areas. These results demonstrate that MSM-retino is the optimal method among those tested, PALS being worst and the FreeSurfer and MSM-sulc being comparable and in between. Notably, the non-linear functions differentiate amongst the four methods more readily than the linear one relating 50% size to the original size.

The plots of PAM size at 50% threshold as a function of the original size (bottom row) are linear for all four methods. The PALS registration is the only one in which 50% area size markedly undershoots the original size. The slope of the linear function for the PALS registration is significantly smaller than that for FreeSurfer ($t = 3.1, p < 0.05$). Moreover, the linear fit intercepts the x-axis well to the right of the origin for all but MSM-retino (see legend for values). The intercept of the 50% function for the MSM-retino was significantly different compared to that for the PALS ($t = 6.3, p < 0.01$), FreeSurfer ($t = 7.3, p < 0.01$) or MSM-sulc registration ($t = 5.13, p < 0.02$). Thus, MSM-retino is the only method that preserves a consistently positive 50% size for the smallest retinotopic areas. These results demonstrate that MSM-retino is the optimal method among those tested, PALS being worst and the FreeSurfer and MSM-sulc being comparable and in between. Notably, the non-linear functions differentiate amongst the four methods more readily than the linear one relating 50% size to the original size.

Fig. 3. Probability atlases (color code see inset) and 8% (dark color), 50% (medium color) and 100% (light color) probability contours obtained with FreeSurfer registration (left columns) and MSM-retino registration (right columns) for three cortical areas: V1 (blue–D), MT (brown–yellow, E–H) and pFST (green, I–L). Note that the 100% probability plateau is not continuous in V1, because of the cut in the flatmap along the calcarine sulcus used to generate the original patches in Freesurfer. Note also the presence of an outlier in the pFST maps.
methods. The boundaries between neighboring areas are very jagged for the PALS method, are smoother but with some local irregularities for the FreeSurfer and MSM-sulc methods, and are smoothest for the MSM-retino method. This has two consequences: (i), some areas in some maps are polytopic, consisting of small disjoint pieces as, e.g., left pFST in PALS and FreeSurfer, or right hV4 in FreeSurfer (Fig. 5); (ii), two neighboring areas can be intertwined as, e.g., left V3A and LO1 in PALS, right LO1 and LO2 in FreeSurfer or MSM-sulc, and right phPITd and phPITv in MSM-sulc. Another difference between methods relates to area size, as some smaller areas are reduced in size or missing altogether for some registration methods, e.g., right pFST in MSM-sulc and right hV4 in FreeSurfer (Fig. 5); (iii), two neighboring areas can be intertwined as, e.g., left V3A and LO1 in PALS, right LO1 and LO2 in FreeSurfer or MSM-sulc, and right phPITd and phPITv in MSM-sulc. Another difference between methods relates to area size, as some smaller areas are reduced in size or missing altogether for some registration methods, e.g., right pFST in MSM-sulc. For the MSM-retino registration method, all 18 areas are present and unitopic in the MPM. This method also yields the smoothest boundaries between areas, and the empty space between areas is reduced, except for the space between V3 and LO2 in the right hemisphere, which was, however, present in most of the individual subjects.

The neighborhood relationships among retinotopic areas obtained with the MSM-retino method generally matches that of individual subjects. However, there are mismatches in regions where the topology differs across individuals, such as in the V3A complex. In this region, the topology of the left MSM-retino MPM resembles that of subject 10’s left hemisphere (Fig. 2C), whereas the right MSM-retino MPM resembles subject 3’s right hemisphere (Fig. 2D).

To further evaluate the robustness of the MPMs we used the ‘leave-one-out’ procedure described above, but applied to the MPM obtained using MSM-retino. Figs. 7A, B illustrates the overlap between that MPM and the individual areas of two subjects. The overlap is excellent for large areas such as V1 and V2, but is reduced in smaller areas (compare pFST with pMSTv). In order to quantify these differences, Fig. 7C plots the overlap, averaged across subjects, for the 18 areas (dark blue bars). In most (13/18) areas the overlap exceeds 70%, reaching 80% in five areas. In another 4 areas the overlap exceeds 50% and only in one area, pFST, is it smaller than 50%. Fig. 7C also shows that the overlap is substantially larger for the MSM-retino method (70% average overlap, dark blue bars) than any other registration strategy (average overlap between 40 and 50%), all differences being very significant (see legend).

Furthermore using MSM-retino for the atlas, even when the left-out subject is registered with MSM-sulc or FreeSurfer, yields significantly more overlap than using only MSM-sulc or FreeSurfer (see legend).

Columns 5 and 6 of Table 1 show that the MPM sizes for the areas aligned using MSM-retino (after compensation for distortions – see Methods) are comparable to the average native-mesh surface areas for all but the smallest retinotopic areas. The MPM-based surface areas were lower using the other three registration methods, especially for the smaller areas (Fig. S3). Area size for the different MPMs depended linearly on the original size, with the slope the smallest for the PALS registration and the intercept the smallest for MSM-retino. The equation relating MPM area size for MSM-retino and the average native midthickness values has a slope of 1.03, close to unity, and an intercept of 9, close to zero (Fig. S3). These comparisons confirm that MSM-retino
is indeed the best of the four alignment methods tested, and therefore it is used to compare retinotopic data to anatomical parcellations in the remainder of this study.

Fig. 8 shows a labeled map of retinotopic areas (defined by the MPM for MSM-retino) on the inflated cortical surface of the HCP group-average left hemisphere (panels A, B) alongside a map of retinotopic organization for one subject on a flatmap (panel C; stars indicate the foveal representation). The visual hemifield is split for areas V1–3 but is contiguous for the other 15 areas. Several areas have their foveal representation in a central confluence shared with V1–V3. As noted by Kolster et al. (2010), three groups of areas share their foveal representation, disjoint from the central confluence, to form a center-based cluster (C-cluster): V3D/C, phPITd/v and MT/pMSTv/pFST/pV4t. Two pairs of areas share their peripheral representation to form a periphery-based cluster (p-cluster): V3A/B and hV4/VO1.

Comparisons between retinotopic MPM and myelin maps

Fig. 9 illustrates two additional associations between retinotopy and myelin density. (i) The phPIT cluster is located in a region of moderate myelin density wedged in between hV4 and LO2. (ii) Area V3D, the most dorsal in the V3A complex, shows a distinct peak in myelin density in both hemispheres, suggesting an association between myelin density and a single area, rather than clusters. Interestingly, the portions of MT and pMST having the highest average myelin content (red and yellow in Fig. 9) involve the upper-field representation of each area (cf. Figs. 1 and 8C), whereas the regions with moderate average myelin content (green in Fig. 9) involve lower-field representations (Kolster et al., 2010). However, it is unclear from the group average maps to what degree this reflects a genuine
difference in myelin content for upper vs lower field representations in MT and pMST and to what degree it reflects residual intersubject areal misalignment, perhaps combined with intersubject differences in the orientation of the polar angle map in the cluster. To address this issue, we examined myelin maps in all 196 individual HCP subjects. Fig. 10 illustrates exemplar results for two individual subjects. In one subject (Fig. 10A), a patch of heavy myelination (red) occupied nearly all of MPM area MSTv and extended into area MT (orange and yellow). In the other (Fig. 10B), an even larger patch of heavy myelination lay mainly outside areas MT/MSTv (above on the inflated surface). In many other cases, the heavily myelinated patch was similar in size or smaller than the illustrated examples, and usually was overlapping with both MT and MSTv. From this, we infer that imperfect registration definitely contributed to the profile of the group average myelin map, but that nonuniformity within each area in some subjects likely was a contributing factor as well.

The estimated myelin content appears to vary within individual areas besides the MT complex already discussed. In V2, myelin content is high in the central representation and appears to be lower in the peripheral representation of both dorsal and ventral V2. Similarly the estimated myelin content is a bit higher in dorsal V1 than in ventral V1. However, this may be related more to the folding patterns in V1 and V2 than to retinotopy per se. Myelin maps tend to have lower values in sulcal fundi, because cortex is thinner overall than in sulcal banks and gyral crowns; the thickness reduction is more pronounced in the heavily myelinated deep layers. In area V3, a variation with eccentricity is suggestive only for the dorsal part.
More strikingly, for many of the ‘fourth-tier’ areas adjoining V3, including LO2, LO1, V3A, V3B, and V3C, there is a gradient in estimated myelin density, which is highest adjoining V3 (mainly posterior) and lowest on the side away from V3 (mainly anterior). This corresponds to a higher average myelin content in the lower field representation of dorsal occipital areas V3A, V3B, and V3C and ventral occipital cortex hV4 and VO1 and in the central visual field representation of LO1/LO2. In the individual-subject map in Fig. 10A, a heavily myelinated patch occupies nearly all of MPM area V3D; neighboring areas V3A/B/C are more lightly myelinated on average, but the pattern is irregular and is consistent with part of heavily myelinated area V3 encroaching into the V3A complex. In Fig. 10B, the myelin map is patchy and irregular in all areas of the V3A complex, but also suggestive of imperfect alignment of areal boundaries. Hence, we infer that, as with the MT complex, residual areal misalignment is likely a major contributor to the group-average gradient in average myelin content.

The myelin density in a given area of the MPM, averaged across the 196 HBP subjects, is plotted for the 18 areas of the two hemispheres in Fig. 11. Confirming the results illustrated in Fig. 9, the myelin density (in percentiles) varies considerably across areas. A two-way repeated-measures ANOVA revealed significant main effects for the factor area ($F_{17, 1} = 524, p < 10^{-10}$) and hemisphere ($F_{17, 1} = 11.15, p < 10^{-3}$).
However, the interaction was also significant ($F_{17, 1} = 18.8, p < 10^{-4}$). Of the 8 areas which showed significant inter-hemispheric differences (post hoc t-tests), three (phPITd, V3C, pV4t) favored the left and five (VO1, LO1, V3A, V7, pFST) the right hemisphere. For both hemispheres, t-tests (Table 2) confirmed statistical significance of the differences in myelin density between several areas in the MT cluster: pMSTv vs pFST; pMSTv vs MT; and MT vs pV4t. Also significant are the differences between the phPITs and their neighbors and between V3 and several of its neighbors (Table 2). Interestingly, t-tests also revealed significant differences between the neighboring areas of the V3A complex, between V3B and V3C as well as between V3C and V3D, and between V3D and V7. Finally, density differed significantly between LO1 and LO2 and between hV4 and VO1.

Comparison of the retinotopic MPM with architectonically defined areas

We compared the retinotopic MPM, obtained with MSM-retino, with the MPM of 9 cytoarchitectonic areas (cytoMPM) derived from the volumetric probability maps of these areas (Fig. 12). There is closer agreement between retinotopic V1 and hOc1 in the cytoMPM, than between V2 and hOc2. The computed overlap was 93% for V1 vs cytoMPM hOc1, compared to 62% for V2 and cytoMPM hOc2 (see in line

![Fig. 10. A. Myelin map (top) and sulc map (bottom) for HCP subject 118730, displayed on the group average inflated surface (postero-lateral view) after MSM-All registration (see Methods), with superimposed MSM-retino area outlines. B. Myelin map and sulc map for HCP subject 119833, with MSM-retino area outlines superimposed.](image-url)

![Fig. 11. Myelin density, averaged across the 196 subjects, in percentiles, for the 18 areas of the left (dark bars) and right (light bars) hemispheres. Vertical lines: standard error of the mean.](image-url)
supplementary information Table S1). Between V3 and hOc3d and hOc3v combined, the overlap is strongest for hOc3v in the left hemisphere and for hOc3d in the right. The overlap of V3 with hOc3d and hOc3v combined was 61% (Table S1). The effect of volume based inter-subject averaging could be evaluated to some degree by comparing our MPM with the probabilistic surface-based maps of cyto-architectonic areas hOc1 and hOc2, previously mapped from surface reconstructions of individual post-mortem brains (Fischl et al., 2008). The overlap between V1 and hOc1 changed little, but that between V2 and hOc2 increased to 81% (Fig. S4, Table S2B). This indicates that the volume based averaging does not solve the issue completely. hOc5, previously mapped from surface reconstructions of individual post-mortem brains (Fischl et al., 2008) overlapped most with pMST and MT in both hemispheres, but hOc5 extended posteriorly into pV4t and pFST (Fig. S4A). This suggests that hOc5, like the heavily myelinated patch to which it seems linked (Fig. S4B), is not restricted to MT proper, but may also include pMST. Given the uncertainty in the alignment of hOc5 across subjects (Van Essen et al., 2012b) and the potential drift between FreeSurfer and MSM-Retino registrations, it remains possible that hOc5 is restricted to MT and/or pMST.

Despite the methodological limitations, these results suggest that beyond V3, cytoarchitectonic maps do not match the retinotopic parcellation, simply because most retinotopic maps are much smaller than the cytoarchitectonic areas in this part of extrastriate cortex.

Comparison of our retinotopic MPM to a probabilistic surface-based map of V1 based on MRI mapping of the stria of Gennari in post-mortem brains, adapted to our 12 subjects (see supplementary information: Methods). The V1/V2 border in the retinotopic MPM was close to the probabilistic V1 boundary, either the 8% or 50% border, depending on the region (see in line supplementary information Fig. S4A). Calculated overlap values (see in line supplementary information Table S2A) showed that the 8% threshold favored the overlap of retinotopic V1 with Hinds’ V1 (95% on average), whereas the 50% threshold minimized the overlap of retinotopic V2 with Hinds’ V1 (5% on average). Thus the V1 in our retinotopic MPM matches V1 defined by the stria of Gennari reasonably well, despite the remaining methodological differences in registration and the difficulties in mapping central visual field with fMRI.

Discussion

Our results demonstrate striking correlations between the retinotopic organization of human occipital cortex as reflected in maximum probability maps and the maps of estimated myelin density. High myelin density occurred in early visual areas V1–3, the MT cluster and area V3D, whereas the phPITv cluster is only moderately myelinated.

### Table 2

The statistical comparison (t-test) of average myelin density between neighboring areas (see Fig. 11).

| Areas             | LH t-value | P   | RH t-value | P   |
|-------------------|------------|-----|------------|-----|
| MT–pMST           | 7          | *** | 10.2       | *** |
| pMST–pFST         | 20.8       | *** | 16.4       | *** |
| MT–pV4t           | 8.6        | *** | 12.2       | *** |
| phPITd–LO2        | 8.5        | *** | 14         |     |
| phPITv–pV4t       | 25.7       | *** | 23.4       | *** |
| V3–LO2            | 34.2       | *** | 31.3       | *** |
| V3–LO1            | 16         |     | 10.6       | *** |
| V3–V3A            | 21.6       | *** | 6.4        | *** |
| V3–V3B            | 20.2       | *** | 24.5       | *** |
| V3B–V3C           | 11.3       |     | 5.9        | *** |
| V3C–V3D           | 4.7        |     | 16         | *** |
| V3D–V7            | 14.3       | *** | 12.4       | *** |
| V01–hV4           | 14.3       | *** | 8.5        | *** |
| L01–L02           | 9.2        | *** | 14.8       | *** |

*: p < 10^-3; **: p < 10^-6; ***: p < 10^-9, all corrected for 28 comparisons.

### Fig. 12

Outlines (white) of retinotopic MPM superimposed on the cytoarchitectonic MPM of left (L) and right (R) hemispheres. Black lines: 50% contours of PAsof hOc1, hOc2, hOc5 from Fischl et al. (2008). Inset color code of cytoarchitectonic areas. In evaluating the overlap (see Table S1) one needs in addition to take into account that the MT cluster as a whole can differ in overall orientation by as much as 60 degrees in different individuals (Kolster et al., 2010, p.201).
We also demonstrated a good match between retinotopic V1–2–3 and cytoarchitectonic areas hOc1–2–3.

Different registration methods

Surface-based registration has a fundamental advantage in principle over conventional volumetric registration for cerebral cortex, because it respects the topology of the cortical surface (Fischl et al., 1999; Van Essen et al., 2012a, 2012b). In practice, however, the magnitude of any observed advantage depends upon the exact registration methods and criteria used for comparison. Here, we used four distinct surface registration methods to align retinotopic areas to an atlas template and used unbiased methods to compare the quality of surface-based alignment. We did not make explicit comparisons with volumetric registration because that has previously been done for both PALS (Anticevic et al., 2008, 2012) and FreeSurfer (Fischl et al., 2008; Van Essen et al., 2012b). One new method (MSM-sulc) achieves comparably good areal alignment the current standard method (FreeSurfer) but achieved this with much lower spatial distortion. Another method (MSM-retino) achieved superior alignment of retinotopic areas because it utilized the areas themselves as registration constraints. This enabled us to generate maximum probability maps of 18 retinotopic areas with regular boundaries and unitopic (non-fragmented) extent. Our method of alignment and MPM generation also avoided confounds of applying thresholds or spatial smoothing. In calculating group average surface areas, we corrected for multiple sources of distortion, resulting in estimated sizes that were very similar to those of a typical individual subject.

Further progress might be obtained by using more detailed functional information to drive the MSM registration, such as the signed distance transform of areas (Yeo et al., 2010) or by using more detailed retinotopic information (e.g., horizontal and vertical meridia, or even continuously varying eccentricity and polar angle maps).

We also addressed an underappreciated problem of systematic drift that can occur during iterative methods of registration to a template. By applying a DeDrifting process, the resultant MPM maps of retinotopic areas can be accurately compared with similarly processed datasets, including the population average myelin maps that we generated from Human Connectome Project data. Failure to include compensatory steps of this type has caused conventional volumetric brain atlas spaces such as MN152 to be substantially larger than the dimensions of the typical contributing individual (see Van Essen et al., 2012b). In the current study it was not possible to compare myelin maps and retinotopic maps directly in individual subjects (as done by Sereno et al., 2013; see below), because the retinotopic data were acquired without accompanying T2w scans needed for generating myelin maps.

Comparison with earlier studies

At a methodological level, the MSM-retino registration used here compares favorably with the results of Frost and Goebel (2013), who also used functional information to improve surface based registration. The average probabilities we obtained (Fig. S2), are a measure similar to the overlap calculated by Frost and Goebel (2013). Most (14/18) of our average probabilities exceeded a value of 0.3, whereas the largest overlap obtained by Frost and Goebel (2013) reached only 0.24. Our average probabilities for the areas of the MT cluster all exceeded the 0.11 overlap obtained for their MT+.

Our findings are in reasonable agreement with those of Sereno et al. (2013), who also compared retinotopic and myelin maps, although their retinotopic mapping of occipital cortex identified fewer areas and made comparisons only with regions of heavy myelination. They described high myelin density in V1 (but not V2–3), MT, FST, V3A and V8/hV4. We found that, all three early areas V1–3 were heavily myelinated, especially in the central visual field representation. Sereno et al. (2013) reported a heavily myelinated patch in the vicinity of human MT and concluded that MT occupied its posterior part, consistent with our interpretation, but they designated the more anterior part of this patch as area FST. Our analysis parcellated the MT cluster into four areas (Kolster et al., 2010), as in the monkey (Kolster et al., 2009) and indicates that the high myelin density area adjoining MT is in area pMSTv; pFST in our parcellation has one of the lowest myelin densities. The maximum in dorsal occipital cortex that Sereno et al. (2013) labeled V3A, corresponds to V3d in our parcellation of the V3A complex. Finally, they reported a relative increase in myelin density in V8/hV4, which corresponds to the higher density in hV4/V01 in our study, but they did not describe the local minimum in the neighboring phPIT cluster.

Our results are also consistent with earlier studies suggesting that retinotopic V1–3 correspond to hOc1, hOc2, and hOc3d&v combined respectively (Wilm et al., 2010; Wohlschlager et al., 2005), and that retinotopic MT does not correspond to all of hOc5 (Wilm et al., 2005).

The early visual areas: V1–3

The three early visual areas V1–3 are characterized by a high myelin density and distinct cytoarchitecture, plus split-hemifield representations with segregated dorsal and ventral quadrants for V2 and V3 (Schira et al., 2009). These evolutionarily conserved areas are common to old world monkeys and humans (Lyon and Kaas, 2002). In the macaque, V2 and V3 lack distinguishing cytoarchitectural differences, but dorsal V3 is more heavily myelinated (Burghalke et al., 1986). These early areas may share a columnar organization for orientation in both species. In monkeys these areas have orientation columns (Vanduffel et al., 2002), while in humans there is evidence for columnar organization of orientation selectivity from high-field fMRI for V1 (Yacoub et al., 2008) and from multivoxel pattern analysis for V2/V3 (Kamitani and Tong, 2005).

The myelin maps suggest possible regional inhomogeneity in the 3 early areas. V1 has the highest, relatively homogeneous density, perhaps increasing with eccentricity. In V2d, V2v and V3d myelin density appears lower in the peripheral representation, but it is difficult to exclude systematic biases related to cortical folding and/or cortical thickness in these regions. Cytoarchitectonic differences have been reported between hOc3d and hOc3v (Kujovic et al., 2013). In the macaque, functional and architectonic differences between V3d and V3v have been reported, though of opposite polarity in terms of myeloarchitecture (Burghalke et al., 1986), but V3d and V3v are now widely regarded as subdivisions of a single retinotopic area (Lyon and Kaas, 2002).

The MT cluster

Our maps of the MT cluster include four retinotopic areas: MT, pMSTv, pFST and pV4t (Arcaro et al., 2011; Kolster et al., 2010). This organization differs from the two areas proposed in previous studies (Amano et al., 2009; Henriksson et al., 2012; Huk et al., 2002) and is similar to that observed recently in the monkey (Kolster et al., 2009). The myelin density peak in lateral occipital cortex is centered on the MT/pMSTv border; pFST is much less myelinated. In the macaque monkey MT is heavily myelinated and M5T (or its subdivisions M5St, MSTv, and MSTdp) tends to be moderately to heavily myelinated (Janssens et al., 2013; Lewis and Van Essen, 2000; Mansell and Van Essen, 1983; Van Essen et al., 1981), whereas FST is moderately myelinated (Lewis and Van Essen, 2000). This is similar but not identical to the pattern in humans. There may be genuine species differences or methodological differences relating to how myelin content is estimated and how it is averaged across subjects. The species difference view is supported by comparing in vivo myelin maps in macaque and human (Glasser et al., 2012, 2014). The MT cluster overlaps with cytoarchitectonic hOc5, but retinotopic MT is only a part of this cytoarchitectonic area, consistent with the results of Malikovic et al. (2007). Further work is needed to better understand which parts of the MT cluster correspond to hOc5 and its subdivisions hOc5v and hOc5d (Malikovic et al., 2007).
Functional characteristics of the MT cluster include responsiveness to moving visual stimuli (Tootell et al., 1995; Zeki et al., 1991) and also to images of body parts, insofar as the extrastriate body area (EBA) overlaps considerably with the motion localizer activation (Peelen and Downing, 2005; Peelen et al., 2006). However, MT is only one of many cortical areas associated with these functional specializations (Ferri et al., 2012; Kolster et al., 2010). While the four areas of the cluster are all motion sensitive (Ferri et al., 2012; Kolster et al., 2010), the myelin density varies largely between pMSTv and pFST, suggesting that there is no simple match between the motion localizer and myelination. On the other hand, the two areas of the cluster with low myelin density, pV4t and pFST, are shape sensitive (Ferri et al., 2012; Kolster et al., 2010), a property they share with the pFPT cluster.

The association of body and motion sensitivity suggests the MT cluster does more than just analyze motion and optic flow in retinal images (Born and Bradley, 2005), and it is also implicated in the analysis of observed actions (Jastorff and Orban, 2009). Indeed, all four areas of the cluster respond when subjects observe others’ actions (Ferri et al., 2012; Kolster et al., 2010). Other functions attributed to this cluster in humans or monkeys involve the control of pursuit of moving targets, including pursuit eye movements (Komatsu and Wurtz, 1988a, 1988b) pursuit by directed attention (Lebrancho et al., 2010) and possibly the arm (Ilg and Schumann, 2007). Macaque MT and MSTv have been implicated in the sensory control of pursuit (Komatsu and Wurtz, 1988a, 1988b) and their human counterparts are densely myelinated, as are human V1 and FEF (Glasser and Van Essen, 2011). Thus, the regions controlling pursuit might transmit visual signals rapidly, explaining the short latency of visual responses in FEF, as short as 50 ms (Schmolesky et al., 1998) and of pursuit, as short as 83 ms in the monkey (Lisberger and Westbrook, 1985). Further work is needed to understand the nonuniform distribution of average myelin content in human MT and pMSTv to determine the relative importance of imperfect registration vs myelination being genuinely denser in the upper visual field representation.

The V3A complex

Several parcellations of cortex in the vicinity of the V3A complex have been proposed, including one with two areas: V3A and 3B (Larsson and Heeger, 2006; Wandell et al., 2007), one with four areas V3A–D (Georgieva et al., 2009) and one with only a single area V3Aβ (Henriksson et al., 2012). The first two descriptions are based in part on properties of monkey V3A: a central representation slightly offset dorsally from the V1–3 confluence (4 areas scheme) and a peripheral representation close to V3d peripheral and V6 (two areas scheme). This suggests that the V3A complex might be expanded in humans compared to old world monkeys. Since V3A in the two-area scheme corresponds largely to V3D as defined in the four-area scheme used here, the present data agree with those of Sereno et al. (2013), as the areas share a high myelin density. However, the association of V3D with a myelin peak was more prominent in the present data than that of V3A in Sereno et al. (2013). In that study V3A was just the lower part of a large myelinated region. Furthermore, the statistical differences in myelin density between V3D, V3C and V3B support our parcellation of the complex into more than two areas. Clearly further work is needed to understand the exact organization of this region, which is of particular interest given its involvement in stereo processing (Georgieva et al., 2009; Taso et al., 2003).

Utilizing retinotopic maps as reference datasets

The MPM and probabilistic retinotopic datasets as well as the HCP myelin maps are accessible at https://db.humanconnectome.org/app/template/Abdollahi_Neuroimage2014.vmr/ and can be used in a variety of ways as reference data for other neuroimaging studies. The choice depends on the type of analysis being done and the reference space in which the analysis is carried out. Suppose, for example, a task-fMRI study has revealed a pattern of activation in extrastriate visual cortex, and the investigator wants to estimate the relative contributions of different retinotopic areas despite the lack of retinotopic maps or other areal features (myelin maps, resting-state fMRI) in the individual subjects. One option would be to register each individual surface to the fs_LR template using MSM-sulc, and then to compare task-fMRI activations to the corresponding MPM map (MSM-sulc). This would match the registration methods as closely as is feasible, but the MPM maps would not accurately reflect the average size of individual retinotopic areas. We favor the use of the MSM-retino version of the MPM map, which provides a much more accurate representation of average areal size, and the DeFrosting process avoids major misalignment even when comparing across different registration methods. Indeed this strategy yields significantly more overlap between areas of a test subject and those of the atlas (Fig. 7C).

If a group analysis for task-fMRI has already been done on a different mesh (specifically, FreeSurfer or PALS), another option is to register the data from one atlas mesh to another using bi-directional inter-atlas transformations available at http://sumsdb.wustl.edu/sums/directory.do?id=82917575&dir_name=inter-atlas_deformation_maps. This would enable mapping of the group-average task-fMRI data to the fs_LR mesh or, conversely, mapping any of the MPM or probabilistic retinotopic maps from fs_LR to the isovalue or PALS atlas mesh.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.06.042.

Acknowledgments

This work was supported by a grant from Telenet NV to GAO, by ERC Paralateralization (PI GAO), by NIH grants MH60974 and F30-MH97312, and by NIH grant 1U54MH091657, funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research. The authors are grateful to M Hohberg for providing the probability maps of the cytoarchitectonic areas.

References

Amano, K., Wandell, B.A., Dumoulin, S.O., 2009. Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. J. Neurophysiol. 102, 2704–2718.

Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., Zilles, K., 2000. Brodmann’s areas 17 and 18 brought into stereotaxic space – where and How variable? NeuroImage 11, 65–84.

Anticevic, A., Dierker, D.L., Gillespie, S.K., Repovs, G., Csernansky, J.G., Van Essen, D.C., Barch, D.M., 2008. Comparing surface-based and volume-based analyses of functional neuroimaging data in patients with schizophrenia. NeuroImage 41, 835–848.

Anticevic, A., Repovs, G., Dierker, D.L., Harwell, J.W.,Coalson, T.S., Barch, D.M., Van Essen, D.C., 2012. Automated landmark identification for human cortical surface-based registration. NeuroImage 59, 2539–2547.

Arcaroli, J.M., Pinsk, M.A., Li, X., Kastner, S., 2011. Visuotopic organization of macaque posterior parietal cortex: a functional magnetic resonance imaging study. J. Neurosci. 31, 2064–2078.

Born, R.T., Bradley, D.C., 2005. Structure and function of visual area MT. Annu. Rev. Neurosci. 28, 157–189.

Brewer, A.A., Press, W.A., Logothetis, N.K., Wandell, B.A., 2002. Visual areas in macaque cortex measured using functional magnetic resonance imaging. J. Neurosci. 22, 10416–10426.

Burkhalter, A., Felleman, D.J., Newsome, W.T., Van Essen, D.C., 1986. Anatomical and physiological asymmetries related to visual areas V3 and VP in macaque extrastriate cortex. Vis. Res. 26, 63–80.

Caspers, J., Zilles, K., Eichhoff, S.B., Schleicher, A., Mohlberg, H., Amunts, K., 2013. Cytoarchitectonic analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. Brain Struct. Funct. 218, 511–526.

Clare, S., Bridge, H., 2005. Methodological issues relating to in vivo cortical myeloigraphy using MRI. Hum. Brain Mapp. 26, 240–250.

Dierker, D.L., Feistch, E., Pruetz Jr, J.E., Petersen, S.E., Schlaggar, B.L., Constantino, J.N., Harwell, J.W.,Coalson, T.S., Van Essen, D.C., 2013. Analysis of cortical shape in children with simpleton autism. Cereb. Cortex Oct 27. [Epub ahead of print].

Dougherty, R.F., Koch, V.M., Brewer, A.A., Fischer, B., Modersitzki, J., Wandell, B.A., 2003. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. J. Vis. 3, 586–598.

Eichhoff, S.B., Stephan, K.E., Mohlberg, H., Greffes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and function-imaging data. NeuroImage 25, 1323–1335.
