

Published in final edited form as:

Science. 2010 November 5; 330(6005): 845–851. doi:10.1126/science.1194908.

Functional Compartmentalization and Viewpoint Generalization Within the Macaque Face-Processing System

 Winrich A. Freiwald^{1,*†} and Doris Y. Tsao^{2,*†}
¹The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

²Division of Biology, California Institute of Technology, MC 114-96, Pasadena, CA 91125, USA

Abstract

Primates can recognize faces across a range of viewing conditions. Representations of individual identity should thus exist that are invariant to accidental image transformations like view direction. We targeted the recently discovered face-processing network of the macaque monkey that consists of six interconnected face-selective regions and recorded from the two middle patches (ML, middle lateral, and MF, middle fundus) and two anterior patches (AL, anterior lateral, and AM, anterior medial). We found that the anatomical position of a face patch was associated with a unique functional identity: Face patches differed qualitatively in how they represented identity across head orientations. Neurons in ML and MF were view-specific; neurons in AL were tuned to identity mirror-symmetrically across views, thus achieving partial view invariance; and neurons in AM, the most anterior face patch, achieved almost full view invariance.

Primates can recognize faces accurately despite a plethora of transformations in size, position, makeup, illumination, and, perhaps the most drastic in terms of low-level feature characteristics, head orientation (1). A biological substrate for primate face recognition is likely provided by face-selective cells (2-8) and by face-selective brain regions, which can be identified by functional magnetic resonance imaging (fMRI) experiments (9-12). In macaques, fMRI reveals six discrete face-selective regions, consisting of one posterior face patch [posterior lateral (PL)], two middle face patches [middle lateral (ML) and middle fundus (MF)], and three anterior face patches [anterior fundus (AF), anterior lateral (AL), and anterior medial (AM)], spanning the entire extent of the temporal lobe (12). Why are there multiple face patches? Answering this question requires understanding the representation of faces in each patch. The six patches form strong, specific connections to each other (13). This suggests that the representations in each distinct patch are not independent but constitute transformations of each other. In particular, electrical microstimulation in the middle face patches activates both AL and AM. Determining how ML, MF, AL, and AM represent faces was the goal of the current study.

We first used fMRI to localize face patches in two monkeys (M1 and M2). Both animals exhibited the typical arrangement of six face patches along the temporal lobe (Fig. 1A and fig. S1A). We then targeted ML, MF, AL, and AM (Fig. 1B, fig. S1B, and table S1) for electrophysiological recordings. These face areas are defined solely by their anatomical locations. We found that response properties of cells within a given anatomically defined

 Copyright 2010 by the American Association for the Advancement of Science; all rights reserved.

[†]To whom correspondence should be addressed. wfreiwald@rockefeller.edu (W.A.F.); dortsao@caltech.edu (D.Y.T.).

^{*}These authors contributed equally to this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6005/845/DC1

face patch, for example, AL, were highly similar across animals. Due to similarity of results across animals, we present combined results from the two animals; due to similarity of results from ML and MF, we group them together as ML/MF.

To compare the face selectivity of ML/MF, AL, and AM, we first recorded neural responses to the 128-image set used in the fMRI localizer experiments, consisting of 16 pictures each of eight object categories (human faces, human bodies, fruits and vegetables, gadgets, human hands, scrambled patterns, monkey body parts, and monkey whole bodies), which we will refer to as the FOB (faces, objects, and bodies) stimulus set. We recorded from every cell encountered (table S1), and the analyses presented below include all cells for which we were able to obtain complete data from FOB as well as a second image set described below. In ML/MF, most cells responded more strongly to faces than to non-face objects: 97% of visually responsive cells were face selective (Fig. 1C and fig. S2A), with 90% selectively enhanced by faces (average face response at least twice as high as average non-face response) [see supporting online material (SOM)] and 7% selectively suppressed by faces.

Both AL and AM contained large fractions of face-selective cells as well (86% in AL, 89% in AM) (Fig. 1, D and E). However, the selectivity patterns in AL and AM (fig. S2, B and C) differed from that in ML/MF: in AL, a much larger fraction of cells than in ML/MF was selectively suppressed by faces (24% versus 7%) or appeared unselective (14% versus 3%) (Fig. 1D). Similarly, in AM, relatively more cells were selectively face-suppressed (10%) or appeared nonselective (11%) (Fig. 1E). These results are puzzling because (i) anatomy suggests that AL and AM receive their major inputs from ML/MF (13, 14) and (ii) blood flow changes in AM and AL are as face selective as those in ML/MF (12). Therefore, one would expect AL and AM to inherit or even enhance the strong face selectivity of ML/MF.

Next, we probed cells with a second image set consisting of 200 pictures of 25 individuals each at eight different head orientations (left full profile, left half profile, straight, right half profile, right full profile, up, down, and back), a stimulus set that we will refer to as the FV (face views) set. Data obtained using the FV set from each of the three patches is presented side by side in Figs. 1, 2, and 4 to facilitate comparison. We first describe results from ML/MF and AL before proceeding to AM. Fig. 1F shows the responses of three typical cells in ML/MF to both the FOB (top) and FV (bottom) image sets. Whereas Cell 1 responded strongly to the faces in the FOB set, Cell 2 responded selectively, but weakly, to faces, and Cell 3 was unresponsive. In response to the FV set, Cell 1 responded to left profiles, straight, and upward views; Cell 2 responded to left full profiles, left half profiles, straight, and upward views, and more weakly to right half profiles, downward views, and the back of the head; and Cell 3 responded only to left half and full profiles. Fig. 1G shows the responses of three typical cells from face patch AL to FOB and FV image sets. Whereas Cell 1 responded strongly to the faces in the FOB set, Cells 2 and 3 were unresponsive and suppressed, respectively. The response properties of these three AL cells thus appeared rather similar to the three ML/MF cells of Fig. 1F, when tested by FOB stimuli. In response to the FV set, Cell 1 was selective for straight, up, and downward views, Cell 2 for left and right half profiles, and Cell 3 for left and right full profiles. Thus, selectivity for head orientation of the three AL cells differed from that of the three ML/MF cells.

This difference in head orientation tuning of ML/MF and AL cells was typical for the entire population of ML/MF and AL cells. Fig. 2, A and B show the response profiles of all cells recorded in ML/MF and AL, respectively, to both the FOB (left) and FV (right) stimuli, with identical cell ordering for both plots. In AL (Fig. 2B), two distinct face view-selective populations are evident. One was excited by straight, up, and downward faces, and often suppressed by left and right profiles (Fig. 2B, right plot, top part). When assessed by the FOB set, most of these cells proved face selective (Fig. 2B, left plot, top part). A second

population responded preferentially to left and right profiles (Fig. 2B, right plot, bottom part). When assessed by the FOB set, which contained only frontal faces, many of these cells appeared not face selective (Fig. 2B, left plot, bottom part). Of these profile-selective cells, 20% were suppressed by frontal faces. In contrast to AL, in ML/MF profile-selective cells responded to only one profile view, and only 11% were suppressed by frontal-view faces. Thus, the lower incidence of face-enhanced cells in AL compared with ML/MF (Fig. 1, C and D, and fig. S2, A and B) can be explained by the specifics of head-orientation tuning in AL. Corroborating this conclusion, the majority of AL cells that appeared visually unresponsive to FOB stimuli (Fig. 2B, left plot, bottom matrix) were also profile selective (Fig. 2B, right plot, bottom matrix).

The transformation of face representation from ML/MF to AL yields a novel property in AL, not found in ML/MF: mirror symmetry of head-orientation selectivity. Not only was the *de novo* appearance of mirror symmetry in AL surprising, also surprising was the fact that such a large fraction of cells exhibited this property: 92 of 215 AL cells responded at least twice as strongly to one of the two full profiles as to frontal faces (fig. S3). These profile-selective cells responded very similarly to both profiles (the response to the non-preferred profile view was, on average, 92% that of the preferred profile view). We next assessed the full tuning of AL cells to all head orientations. We probed 57 cells with faces randomly sampled from a three-dimensional (3D) head-orientation manifold, which was parameterized by up-down angle, left-right angle, and picture-plane angle (Fig. 3A). The faces were rendered using the face modeling software FACEGEN and refreshed at 6 Hz. To assess headorientation selectivity, we plotted tuning along the three possible pairs of rotation dimensions while averaging over the third dimension. Pairwise head-orientation tuning profiles are shown in Fig. 3B for four example cells (the first two cells are the same as the first two in Fig. 1G). Each of the four cells showed tuning along all three head-orientation axes, and this tuning was always mirror symmetric. Of 57 cells, 43 (75%) in the population tested had view tuning maps with two discrete peaks at mirror-symmetric positions like example cells 1 to 4.

A consequence of the emergence of mirror symmetry of head-orientation selectivity in AL is the development of partial invariance to head orientation. This can be seen in a multidimensional scaling (MDS) analysis of population response vectors (Fig. 4, A and B, and fig. S4). Although in ML/MF, each of the eight head orientations forms a discrete cluster, whose neighborhood relations reflect physical proximity of head orientation, in AL, this topological relationship is broken and responses to multiple head orientations are collapsed into joint clusters (Fig. 4B and fig. S4): Left and right full profiles are grouped into one cluster, left and right half profiles into another, and up, down, and straight into a third one.

A second difference between face representations in ML/MF and AL became apparent when we analyzed selectivity for facial identity, the second dimension of the FV stimulus set. Almost half the cells in AL were significantly modulated by facial identity (45%, analysis of variance) (fig. S5A)—significantly more than in ML/MF (19%). This is reflected in sharper identity tuning in AL compared with ML/MF (Fig. 4G and fig. S6). Because in AL partial view invariance by virtue of mirror-symmetric tuning is established, the question arises whether selectivity for facial identity generalizes across view directions in AL. We measured the similarity (correlation) between population responses to all FV images, that is, to all combinations of identity-head-orientation pairs to construct population response similarity matrices (Fig. 4, D and E) and corresponding significance matrices (fig. S7, A and B). The main feature of the ML/MF similarity matrix (Fig. 4D) is high-similarity 25 by 25 squares along the main diagonal, reflecting a view-specific representation; furthermore, there are no visible paradiagonal stripes ($y = x + n \times 25$), which would emerge if population

response vectors to specific individuals were similar across head orientations. The AL similarity matrix (Fig. 4E) is characterized by a different pattern of high-similarity 25 by 25 squares reflecting mirror-symmetric head-orientation tuning and by paradiagonal stripes indicating view-invariant individual selectivity. These paradiagonals are not continuous but are confined to specific combinations of views (e.g., left and right full profiles). Quantification of view-invariant identity selectivity in the populations of ML/MF and AL cells is shown in fig. S5, B and C. View-invariant identity tuning was significantly stronger in AL than in ML/MF (fig. S5D).

AM, like AL, showed a different pattern of face selectivity than ML/MF when probed with FOB stimuli (Fig. 1E and fig. S2C). Again, the FV stimuli revealed the cause. Fig. 1H shows the responses of three AM example cells to the FOB and FV stimuli. Cell 1 responded strongly to the FOB faces and to all non-backward head orientations in the FV set. Cells 2 and 3 were unresponsive to all images in the FOB set, and each responded strongly to only two individuals in the FV set, across almost all non-backward views. These example cells span the range of responses we observed across the population (Fig. 2C). From the population response, two new characteristics of face representations in AM, found neither in ML/MF nor in AL, are apparent. First, cells in AM were far less sharply and strongly tuned to head orientation than those in AL or ML/MF (Fig. 4H). This is reflected in the MDS analysis (Fig. 4C), which shows that in AM, unlike ML/MF and AL, population responses were only weakly organized by head orientation. Second, AM cells spanned a wide range of identity selectivity from complete lack of identity tuning to sharp identity tuning; in particular, we observed cells like Cells 2 and 3 in Fig. 1H with sparse, individual-specific, view-invariant response patterns only in AM. The specificity of some AM cells was so extreme that they did not respond to any of the faces in the FOB set, explaining the lower apparent face selectivity of AM compared with ML/MF. Even sparser cells may exist in AM, cells that do not respond to any of the FOB or FV stimuli.

Overall, an even larger fraction of cells in AM than in AL were tuned to facial identity (73% versus 45%) (fig. S5A), and tuning to facial identity was sharper in AM than in AL (Fig. 4G). How does identity selectivity depend on head orientation in AM? The population similarity matrix for AM (Fig. 4F and fig. S7C), unlike those for ML/MF and AL, exhibited only a weak dependence on head orientation; rather, its main feature was robust paradiagonal stripes across all head orientations, except the back of the head. For example, the population response to individual 7 at left full profile was more similar to the response to the same individual at upward head orientation than to the response to individual 12 at left full profile. Quantification of view-invariant identity selectivity (figs. S5, B and C) shows a further increase in AM, compared with ML/MF and with AL; both differences are highly significant (fig. S5D). Thus, response similarity in AM is abstracted from pixel-wise picture similarity: The population of AM cells approaches a view-invariant representation of facial identity. It may be surmised that view-invariant identity in AM is carried primarily by sparsely firing cells. Yet we found that both sparsely and non-sparsely responding cells contained view-invariant identity information (fig. S8).

Although we have emphasized feed-forward transformations between face patches as the major change in face representation, processing within face patches and recurrent processing between patches at different levels of the processing hierarchy are likely further mechanisms that, over time, bring about more elaborate representations (15). Indeed, view-invariant face selectivity as assessed by the similarity matrices changed substantially over the course of time in a manner not directly reconcilable with a simple, one-time feed-forward mechanism of transformation (Fig. 4I). Rather, the buildup over time suggests additional recurrent mechanisms. Only representations in AM and AL, but not ML/MF, profited from the passage of time to increase identity selectivity in a view-invariant manner (Fig. 4I). Thus,

whatever mechanism brings about view-invariant identity selectivity does not indiscriminately involve all face patches but only specific subsets.

Different schemes for processing facial information have been proposed. In particular, it has been suggested that after an early encoding stage two parallel streams emerge (16), a dorsal one in the superior temporal sulcus (STS) related to coding changeable aspects of faces, and a more ventral one coding structural aspects of faces (17). Applied to the macaque brain, this scheme implies a parallel arrangement of AL and AM. Alternatively, the three face patches could be arranged in a hierarchical manner. Thus, the transformation that gives rise to an almost fully viewinvariant representation in AM may originate directly from ML/MF (in parallel to the transformation occurring between ML/MF and AL) or mostly from AL (fig. S9A). The increase in invariance to head orientation from ML/MF to AL and further to AM (Fig. 4, A to F, and figs. S5 and S7) suggests a hierarchical arrangement. Also, consistent with AM receiving inputs from AL, the population similarity matrix for AM showed residual traces of mirror symmetry (Fig. 4F). A separate experiment in which we tested for invariance to spatial position showed a similar increase of invariance from ML/MF to AL, and further to AM (fig. S10). Independent and more direct evidence for parallel or hierarchical arrangements can be derived from the relative timing of neural responses across face patches. First, we consider local field potential (LFP) responses (fig. S11). In all three patches, the LFP evoked by faces was clearly distinct from that evoked by nonface objects: The latency of the first peak to faces was at least 12 ms shorter than that to nonface objects. The latencies of the first faceevoked response peaks increased from ML/MF (126 ms), to AL (133 ms), and further to AM (145 ms). This sequential increase is most easily reconciled with a hierarchical arrangement of the face patches. If a major direct projection from ML/MF to AM existed without additional synaptic relays, one would expect that a first activity wave would be triggered in AM at approximately the same time as that in AL, not at more than twice the ML/MF-AL delay. Second, response latencies of neurons in the three face patches (fig. S12) systematically increased from ML/MF (average 88 ms) to AL (104 ms) and further to AM (124 ms). Again, this is difficult to explain without assuming additional steps of synaptic transmission between ML/MF and AM compared with ML/MF and AL, and the most parsimonious explanation would be relay through AL.

The finding that view dependence of identity tuning was fundamentally different in the three face patches (Fig. 4, A to F) addresses a fundamental problem for computational neuroscience (18): How can increasingly shape-selective but otherwise invariant representations be generated? Our results suggest that pooling across mirrorsymmetric views may constitute a crucial step in this process. Fig. S9B presents a sequential model for establishing a view-invariant representation of faces, motivated by the response properties of neurons recorded in ML/MF, AL, and AM. At the first level (view-specific, I), each neuron responds most strongly to a specific individual at a specific view. At the next level (partially viewinvariant, II), each neuron pools together inputs representing mirror-symmetric views of the same individual (left and right full profiles, green axons; left and right half profiles, blue axons), or straight, upward, and downward views of the same individual (red axons), with lateral inhibition between profile- and frontal-selective neurons, and between neurons representing different identities. Finally, in a view-invariant stage (level III), each neuron pools together inputs representing all views of the same individual, with lateral inhibition to generate sparse responses. The model predicts similarity matrices (fig. S9C), which resemble the ones found in the three face patches (Fig. 4, D to F), capturing their main features of view tuning (square structures) and view-invariant identity selectivity (paradiagonal stripes).

The greatest obstacle to object recognition is the huge amount of variation that can occur in the retinal images cast by a 3D object. Our finding of individual-selective responses with a

high degree of invariance across head orientations in AM was obtained with an image set containing faces never encountered in real life. Thus, whatever learning has occurred before the experiments, it has resulted in a face representation that allows generalization of selectivity for new faces. The face system may already incorporate all a priori knowable invariances of a bilaterally symmetric 3D object, a face, into a canonical face space (19-21), such that for an a posteriorly encountered novel face, a largely invariant response can be generated without necessity for further learning. Although experience with an actual individual is not a necessary condition for representations in AM, such experience may yield an even more invariant representation (22-24).

Four results provide new insights into the functional organization of the macaque facepatch system. First, all face regions recorded from contained a high degree of face-selective neurons (lower estimates using only the FOB stimulus set: 97% in ML/MF, 86% in AL, and 89% in AM). Extrapolating from these findings, it seems likely that the entire network of temporal lobe face patches constitutes a dedicated brain system for the processing of one high-level object-category, faces. Second, structure and function are highly correlated. A face patch at a particular anatomical location harbors a specific face representation that is qualitatively different from the face representation in a face patch at a different location, and further differences are likely to exist between face representations in the different face patches. Thus, attaching a name to a given face patch is now shown to be meaningful because it signifies functional identity across individuals. Together with earlier results (12, 13, 25), this shows that the face-processing system is a network composed of multiple, functionally specialized nodes. Third, the system contains one region (AM) that provides for population coding of identity across view conditions, using a hybrid representation of both coarse and sparse elements (26-28). Fourth, the finding of an entire face patch, yet only one, containing neurons with mirror-symmetric tuning is worth particular emphasis because it raises the possibility that such a representation constitutes a critical computational step for object recognition. Even though inferotemporal cells with mirror-symmetric tuning have been reported before (29, 30), this is the first time that such cells have been shown to be agglomerated within a single intermediate node within a form-processing network. Why would this step be useful? One possibility, following earlier work on view-tuned face-selective neurons in the STS (7, 31-33), is that this stage serves to extract socially important information from head orientation, because face (and gaze) aversion carries similar social meaning whether for leftward or rightward orientation. Alternatively, such a stage may provide computational advantages for efficient coding (34). Why are the three view stages of the face-processing system located in separated regions and not next to each other? One possibility is that the faceprocessing system interdigitates with representations for other objects implementing the same three view stages. The pattern of viewpoint generalization revealed here for faces may thus turn out to be a general organization principle of the entire inferotemporal cortex.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to M. Livingstone, S. Moeller, C. Cadieu, and A. Tsao for discussions; N. Schweers for outstanding technical assistance; S. Chang and the late D. Freeman for stimulus programming; the Humboldt Foundation Sofia Kovalevskaya Award, German Science Foundation grant FR 1437/3-1, the German Ministry of Science (Grant 01GO0506, Bremen Center for Advanced Imaging), and NIH 1R01EY019702 for financial support; and the Otto Loewi Minerva Center for logistic support. D.Y.T. is supported by the Alfred Sloan Foundation, the John Merck Foundation, the Searle Foundation, the Klingenstein Foundation, and NSF CAREER Award BCS-0847798. W.A.F. is supported by a Feodor Lynen Research Fellowship from the Alexander von Humboldt Foundation, an Irma T.

Hirschl and Monique Weill-Caulier Trusts Award, a Klingenstein Fellowship Award, a Sinsheimer Scholar Award, and a Pew Scholar Award.

References and Notes

1. Biederman I. *Spat. Vis.* 2000; 13:241. [PubMed: 11198235]
2. Bruce C, Desimone R, Gross CG. *J. Neurophysiol.* 1981; 46:369. [PubMed: 6267219]
3. Perrett DI, Rolls ET, Caan W. *Exp. Brain Res.* 1982; 47:329. [PubMed: 7128705]
4. Rolls ET. *Hum. Neurobiol.* 1984; 3:209. [PubMed: 6526707]
5. Desimone R, Albright TD, Gross CG, Bruce C. *J. Neurosci.* 1984; 4:2051. [PubMed: 6470767]
6. Yamane S, Kaji S, Kawano K. *Exp. Brain Res.* 1988; 73:209. [PubMed: 3208858]
7. Hasselmo ME, Rolls ET, Baylis GC, Nalwa V. *Exp. Brain Res.* 1989; 75:417. [PubMed: 2721619]
8. Tanaka K, Saito HA, Fukada Y, Moriya M. *J. Neurophysiol.* 1991; 66:170. [PubMed: 1919665]
9. Kanwisher N, McDermott J, Chun MM. *J. Neurosci.* 1997; 17:4302. [PubMed: 9151747]
10. Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB. *Nat. Neurosci.* 2003; 6:989. [PubMed: 12925854]
11. Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S. *Proc. Natl. Acad. Sci. U.S.A.* 2005; 102:6996. [PubMed: 15860578]
12. Tsao DY, Moeller S, Freiwald WA. *Proc. Natl. Acad. Sci. U.S.A.* 2008; 105:19514. [PubMed: 19033466]
13. Moeller S, Freiwald WA, Tsao DY. *Science.* 2008; 320:1355. [PubMed: 18535247]
14. Saleem KS, Tanaka K, Rockland KS. *Cereb. Cortex.* 1993; 3:454. [PubMed: 8260813]
15. Sugase Y, Yamane S, Ueno S, Kawano K. *Nature.* 1999; 400:869. [PubMed: 10476965]
16. Bruce V, Young A. *Br. J. Psychol.* 1986; 77:305. [PubMed: 3756376]
17. Haxby JV, Hoffman EA, Gobbini MI. *Trends Cogn. Sci.* 2000; 4:223. [PubMed: 10827445]
18. Riesenhuber M, Poggio T. *Nat. Neurosci.* 1999; 2:1019. [PubMed: 10526343]
19. Valentine T. *Q. J. Exp. Psychol. A.* 1991; 43:161. [PubMed: 1866456]
20. Leopold DA, O'Toole AJ, Vetter T, Blanz V. *Nat. Neurosci.* 2001; 4:89. [PubMed: 11135650]
21. McKone E, Aitkin A, Edwards M. *J. Exp. Psychol. Hum. Percept. Perform.* 2005; 31:1181. [PubMed: 16366783]
22. Miyashita Y. *Nature.* 1988; 335:817. [PubMed: 3185711]
23. Wallis G, Bühlhoff HH. *Proc. Natl. Acad. Sci. U.S.A.* 2001; 98:4800. [PubMed: 11287633]
24. Li N, DiCarlo JJ. *Science.* 2008; 321:1502. [PubMed: 18787171]
25. Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS. *Science.* 2006; 311:670. [PubMed: 16456083]
26. Barlow HB. *Perception.* 1972; 1:371. [PubMed: 4377168]
27. Young MP, Yamane S. *Science.* 1992; 256:1327. [PubMed: 1598577]
28. Rolls ET. *Neuropsychologia.* 2007; 45:124. [PubMed: 16797609]
29. Rollenhagen JE, Olson CR. *Science.* 2000; 287:1506. [PubMed: 10688803]
30. Logothetis NK, Pauls J, Poggio T. *Curr. Biol.* 1995; 5:552. [PubMed: 7583105]
31. Perrett DI, et al. *Proc. R. Soc. Lond. B Biol. Sci.* 1985; 223:293. [PubMed: 2858100]
32. Perrett DI, et al. *Exp. Brain Res.* 1991; 86:159. [PubMed: 1756786]
33. Eifuku S, De Souza WC, Tamura R, Nishijo H, Ono T. *J. Neurophysiol.* 2004; 91:358. [PubMed: 14715721]
34. Vetter T, Poggio T, Bühlhoff HH. *Curr. Biol.* 1994; 4:18. [PubMed: 7922306]

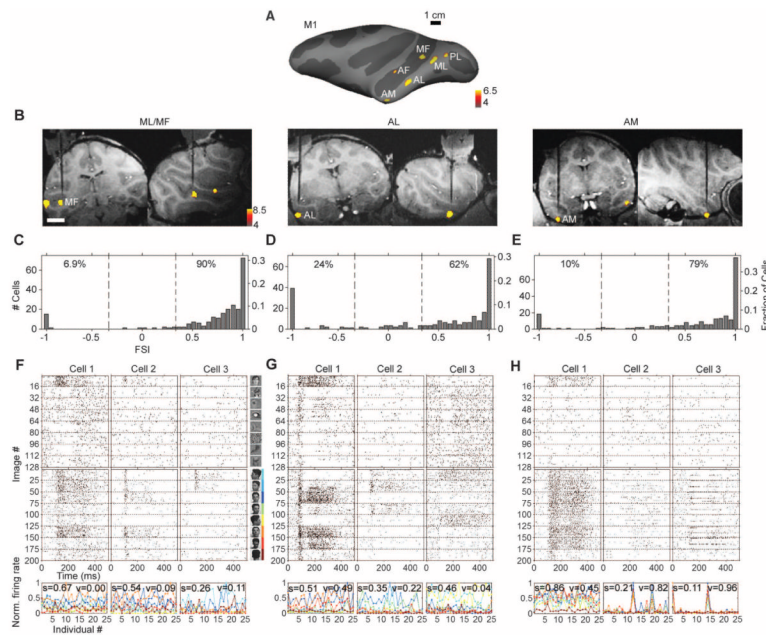


Fig. 1. Face selectivity in different parts of the macaque temporal lobe. **(A)** Inflated macaque left hemisphere (dark gray areas mark sulci, light gray-dark gray boundaries mark the middle of the bank within a sulcus) showing six regions in the temporal lobe of monkey M1 that responded significantly more to faces than to objects in fMRI experiments. Color scale indicates negative common logarithm of the P value. **(B)** Coronal (left) and sagittal (right) anatomical fMRI images showing the electrode descending into MF (located 3 mm anterior to the interaural line, AP (anterior-posterior) + 3 mm), AL (at AP + 12 mm), and AM (at AP + 19 mm), respectively, in monkey M1. Coregistered face-selective functional activation is overlaid on the fMRI images. **(C to E)** Face selectivity of neural population responses in ML/MF, AL, and AM, respectively. Shown are distributions of face selectivity indices (FSIs) (see SOM) for visually responsive cells in ML/MF, AL, and AM; dotted lines indicate FSI of ± 0.33 , corresponding to 1:2 and 2:1 response ratios to faces versus nonface objects. **(F to H)** Mean response time courses of three typical cells to the 128-image FOB set (top) and the 200-image FV set (middle) in ML/MF (F), AL (G), and AM (H), respectively. For clarity, responses are shown using a binary color scale. For the FV data, the first 25 rows are responses to 25 individuals looking to the left at full profile, the next 25 rows are responses to the same 25 individuals looking to the left at half profile, and so on; the eight different views of one example individual are shown on the right of (F). [Colored traces at the bottom of (F) to (H)]: Mean response levels to the 25 individuals at each head orientation, with the color corresponding to each view indicated in (F); *s* and *v* denote sparseness and view-invariant identity correlation coefficients, respectively (see SOM).

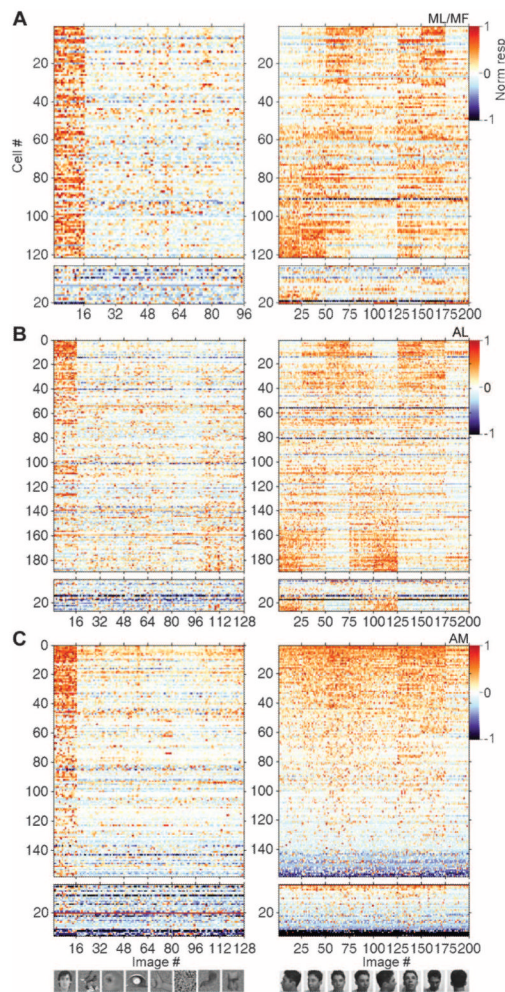


Fig. 2. Selectivity of neural populations in ML/MF, AL, and AM to faces varied in view and identity. (A to C) Population response matrices to the FOB image set (left) and to the FV set (right), for cells visually responsive (top) and nonresponsive (bottom) in ML/MF (A), AL (B), and AM (C), respectively. Responses are sorted from top to bottom by the first (C) or second (A and B) principal component of the FV responses. Data combined from monkeys M1 (recordings in left hemisphere) and M3 (recordings in right hemisphere) for ML/MF, and from monkeys M1 and M2 (recordings in right hemisphere) for AL and AM. The FOB matrix for ML/MF contains only 96 elements because only the first 96 images were presented during recordings in monkey M3.

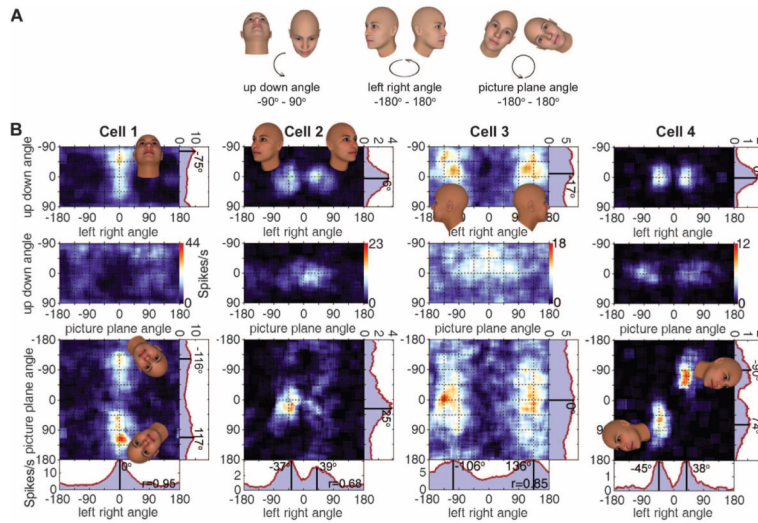
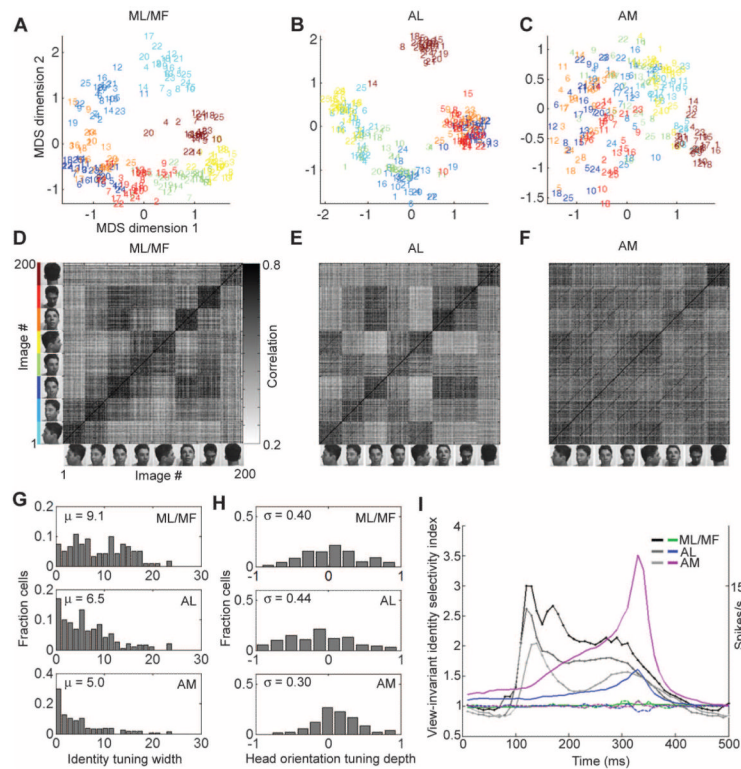


Fig. 3. Tuning of AL cells to a head randomly rotated in three dimensions. **(A)** Illustration of stimulus head and three axes of rotation. **(B)** View tuning in four typical cells. Cells 1 and 2 here are the same as Cells 1 and 2 in Fig. 1G. Cell 4 did not respond to any of the FV stimuli. (Top) Tuning to updown angle versus left-right angle (responses averaged across pictureplane angle). (Middle) Tuning to up-down angle versus picture-plane angle (responses averaged across left-right angle). (Bottom) Tuning to picture-plane angle versus left-right angle (responses averaged across up-down angle). Marginal tuning curves are also shown (vertical lines indicate tuning peak positions).

**Fig. 4.**

Population representations of face view and identity in ML/MF, AL, and AM. (**A** to **C**) Comparison of multidimensional scaling plots of responses to the FV image set in ML/MF, AL, and AM. Each plot shows the location of the 25 faces (indicated by numbers 1 to 25) at eight head orientations [indicated by eight colors, key in (**D**)] within the first two dimensions of the MDS space (corresponding Eigenvalues in fig. S3). (**D** to **F**) Population similarity matrices in the three face patches. A 200 by 200 matrix of correlation coefficients was computed between responses of all visually responsive cells to the 200 FV stimuli from ML/MF ($N = 121$ cells), AL ($N = 189$ cells), and AM ($N = 158$ cells). The correlation patterns do not change when only the first 121 cells of each patch are considered (fig. S13). (**G**) Sharpness of identity tuning in ML/MF, AL, and AM (top to bottom). Central tendencies of distributions of identity tuning half-widths (see SOM) were significantly different from each other ($P \ll 0.001$, Mann-Whitney U tests). (**H**) Distributions of head-orientation tuning depths (see SOM). Tuning depths close to 0 indicate broad tuning. These distributions are significantly different from each other ($P \ll 0.001$ for ML/MF versus AL and AL versus AM, Mann-Whitney U tests; $P < 0.002$ for ML/MF versus AM and AL versus AM, F test). (**I**) Evolution of view-invariant identity selectivity over time. View-invariant identity-selectivity index, computed over a 200-ms sliding response window beginning at the indicated time point, plotted for AM, AL, and ML/MF (solid curves). The dotted curves show the mean view-invariant identity-selectivity index over time computed from shuffled similarity matrices. The grayscale traces show the time course of the mean response to the FV stimuli across the population in each face patch. The substantial delay between the peak of the mean response to the FV stimuli and the peak of the view-invariant identity-selectivity index suggests that recurrent mechanisms are involved in the computation of viewinvariant identity.