



**HAL**  
open science

## Single neuron responses underlying face recognition in the human midfusiform face-selective cortex

Rodrigo Quian Quiroga, Marta Boscaglia, Jacques Jonas, Hernan G Rey, Xiaoqian Yan, Louis Maillard, Sophie Colnat-Coulbois, Laurent Koessler, Bruno Rossion

### ► To cite this version:

Rodrigo Quian Quiroga, Marta Boscaglia, Jacques Jonas, Hernan G Rey, Xiaoqian Yan, et al.. Single neuron responses underlying face recognition in the human midfusiform face-selective cortex. *Nature Communications*, 2023, 14 (1), pp.5661. 10.1038/s41467-023-41323-5. hal-04276234

**HAL Id: hal-04276234**

**<https://hal.science/hal-04276234>**

Submitted on 8 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Single neuron responses underlying face recognition in the human midfusiform face-selective cortex

Received: 2 November 2022

Accepted: 28 August 2023

Published online: 13 September 2023

 Check for updates

Rodrigo Quian Quiroga<sup>1,2,3,4,8</sup>✉, Marta Boscaglia<sup>3,8</sup>, Jacques Jonas<sup>5,6</sup>, Hernan G. Rey<sup>3</sup>, Xiaoqian Yan<sup>5</sup>, Louis Maillard<sup>5,6</sup>, Sophie Colnat-Coulbois<sup>5,7</sup>, Laurent Koessler<sup>5,6</sup> & Bruno Rossion<sup>5,6</sup>✉

Faces are critical for social interactions and their recognition constitutes one of the most important and challenging functions of the human brain. While neurons responding selectively to faces have been recorded for decades in the monkey brain, face-selective neural activations have been reported with neuroimaging primarily in the human midfusiform gyrus. Yet, the cellular mechanisms producing selective responses to faces in this hominoid neuroanatomical structure remain unknown. Here we report single neuron recordings performed in 5 human subjects (1 male, 4 females) implanted with intracerebral microelectrodes in the face-selective midfusiform gyrus, while they viewed pictures of familiar and unknown faces and places. We observed similar responses to faces and places at the single cell level, but a significantly higher number of neurons responding to faces, thus offering a mechanistic account for the face-selective activations observed in this region. Although individual neurons did not respond preferentially to familiar faces, a population level analysis could consistently determine whether or not the faces (but not the places) were familiar, only about 50 ms after the initial recognition of the stimuli as faces. These results provide insights into the neural mechanisms of face processing in the human brain.

Converging evidence has shown that visual object recognition is processed along the ventral visual pathway<sup>1–3</sup>. In macaque monkeys, face-selective neurons have long been identified in the highest stages of this pathway, i.e., the infero-temporal cortex<sup>4,5</sup>. More recent studies, combining functional magnetic resonance imaging (fMRI) with neurophysiological recordings, have described a number of face-selective patches containing almost exclusively neurons that respond selectively to faces<sup>6,7</sup>. In humans, for nearly three decades, neuroimaging studies have consistently shown larger responses to faces than to other

visual categories (i.e., face-selective responses) in the ventral occipito-temporal cortex (VOTC), with a peak of activation in the lateral portion of the midfusiform gyrus (the “Fusiform Face Area”, “FFA”<sup>8–11</sup>). The key role of this region in face recognition is further supported by intracranial electrophysiology studies<sup>12,13</sup>, face-selective identity recognition impairments elicited by focal lesions<sup>14</sup> and face-related perceptual effects following direct electrical stimulation<sup>15,16</sup>. There are basically two different mechanisms that can produce the observed fMRI responses: i) as in the monkey face patches<sup>6,11</sup>, the human midfusiform

<sup>1</sup>Hospital del Mar Research Institute (IMIM), Barcelona, Spain. <sup>2</sup>Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain. <sup>3</sup>Centre for Systems Neuroscience, University of Leicester, Leicester, UK. <sup>4</sup>Ruijin hospital, Shanghai Jiao Tong university school of medicine, Shanghai, China. <sup>5</sup>Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France. <sup>6</sup>Université de Lorraine, CHRU-Nancy, Service de Neurologie, F-54000 Nancy, France. <sup>7</sup>Université de Lorraine, CHRU-Nancy, Service de Neurochirurgie, F-54000 Nancy, France. <sup>8</sup>These authors contributed equally: Rodrigo Quian Quiroga, Marta Boscaglia.

✉ e-mail: [rqqg1@le.ac.uk](mailto:rqqg1@le.ac.uk); [bruno.rossion@univ-lorraine.fr](mailto:bruno.rossion@univ-lorraine.fr)

region may have a majority of neurons that respond selectively to faces (with other neurons being selective for other category of stimuli), and ii) an alternative hypothesis is that the human midfusiform face activations observed with fMRI are generated by broadly tuned neurons that respond strongly (but not exclusively) to faces and have also weaker responses to other category of stimuli.

In addition to the lack of a mechanistic understanding of the neuronal responses giving rise to face-selective neural responses in humans, the role of the midfusiform gyrus in the recognition of familiar faces is controversial, with neuroimaging studies reporting mixed results<sup>17–19</sup>. Moreover, there is very little work comparing neural responses to familiar versus unknown faces in monkeys<sup>20,21</sup> and, considering the thousands of familiar individuals that humans know by their faces<sup>22</sup>, testing face familiarity in humans may reveal unique outcomes. According to a prevalent view, the human face-selective midfusiform region holds similar representations of familiar and unknown faces, which are distinguished at later stages of processing in the temporal lobe<sup>18,23</sup>. Alternatively, the representation of familiar and unfamiliar faces may already differ in this region, shortly after, or even concurrently to face-selectivity<sup>17</sup>. To address these issues, and using the unique opportunity to record from multiple individual neurons in humans, in the present study we investigated: first, the neuronal responses underlying the human midfusiform face-selective activations, and second, whether the midfusiform neuronal responses hold information about the long-term familiarity of the face identities.

## Results

We recorded from 78 units (33 single-units and 45 multi-units) in the midfusiform gyrus of 5 patients implanted with intracerebral electrodes for refractory epilepsy pre-surgical exploration. During the experiment, the subjects viewed pictures corresponding to four categories: familiar faces, unknown faces, familiar places, and unknown places (Fig S1). Figure 1 shows the face-selective fMRI activations for one of the patients, together with the placement of the intracerebral electrode, ending in this region. For all patients, we verified that the electrodes for single neuron recordings ended in the midfusiform gyrus, although they had variable locations with respect to the peak of the face activations (Fig. S3). For this reason, we refer to responses in the face-selective midfusiform gyrus and not in the FFA (or even more specific functional subregions such as pFus-faces/FFA1, mFus-faces/FFA2;<sup>24</sup>). Yet, importantly, for all five patients, the regions in which we performed the recordings were face-selective, as determined from two independent fMRI and Local Field Potential (LFP) face localizer experiments (Fig. S4). Moreover, we found units responding to faces

and places in all patients (Fig. S5), and the proportion of both face and place responses was statistically the same across patients ( $\chi^2$  test,  $p = 0.2$  and  $p = 0.4$ , respectively).

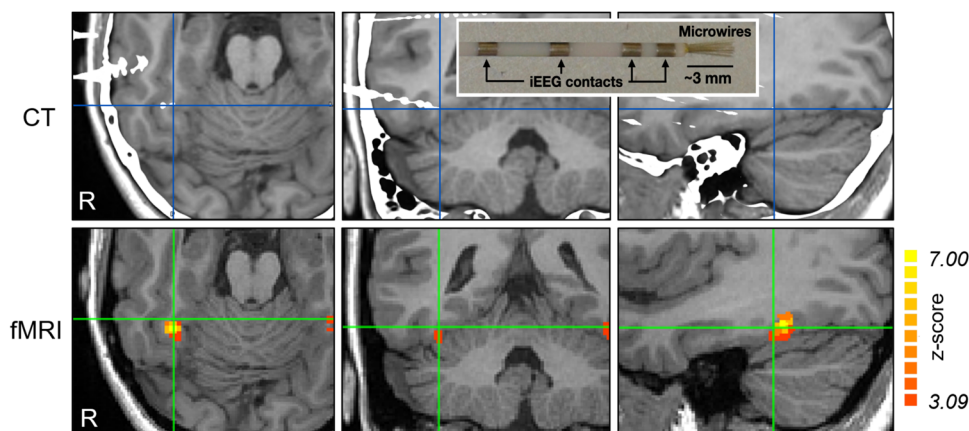
## Midfusiform neurons have a variety of responses

All patients showed units responding to faces and places. Four exemplary units are displayed in Fig. 2, showing a variety of responses to faces, to places and to both category of stimuli. Altogether, from the 78 detected units, 43 (55%) were responsive (paired *t*-test,  $p < 0.05$ ) and, of these: 19 responded only to faces (as in Fig. 2a, c), 7 responded only to places (as in Fig. 2d), and 17 responded to both faces and places (as in Fig. 2b)—i.e. 36 (46%) of the recorded units showed a response to faces and 24 (30%) to places (see Table S1). The number of units responding to faces and to places was significantly larger than chance ( $p < 10^{-25}$  and  $p < 10^{-12}$ , respectively, binomial test), and there were more units responding to faces than to places ( $p < 10^{-3}$ , binomial test). Of the 36 responses to faces, 25 were enhanced responses and 11 were suppressed, whereas of the 24 responses to places, 13 were enhanced, and 11 were suppressed. Eleven (30%; 9 enhanced and 2 suppressed) of the face-responsive units showed a significant familiarity modulation (*t*-test,  $p < 0.05$ ), whereas a familiarity modulation was found for 6 (25%; 5 enhanced and 1 suppressed) place-responsive units (*t*-test,  $p < 0.05$ ). Altogether, we could not identify any obvious visual feature in our stimulus set (Fig. S1B) that triggered the unit responses. As shown with the examples of Fig. 2, and as also shown in the next section with a wide variety of responses, in terms of their strength, latency, selectivity, and familiarity, different neurons had different tuning preferences given by non-trivial features of the stimuli.

## Similar average responses to faces and places

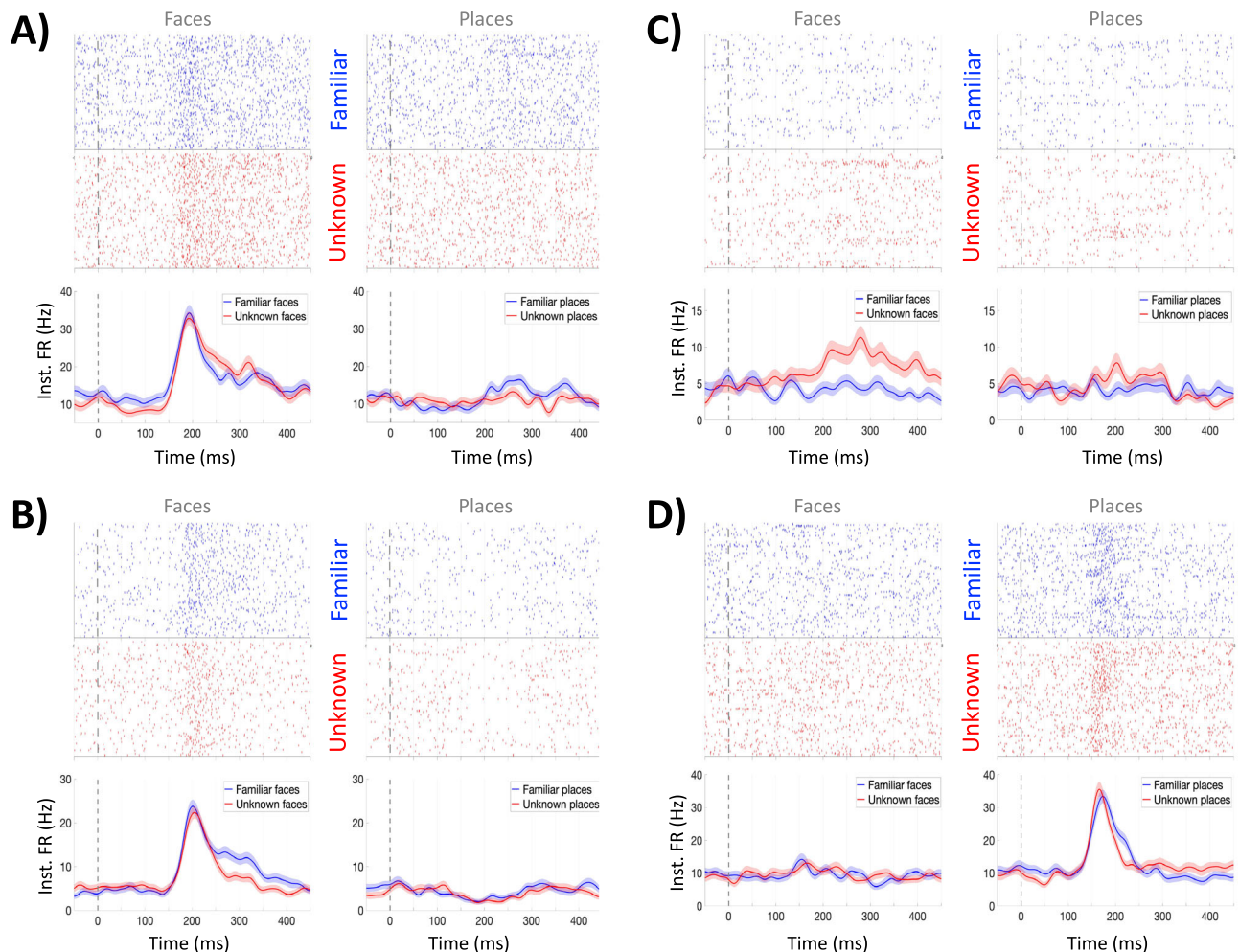
Figure 3a, b show the normalized grand averages for the enhanced responses to faces ( $N = 25$ ) and places ( $N = 13$ ). (Analogous results were obtained for the suppressed responses and are shown in Fig. S6). We should clarify that we did not find any systematic trend in the responses across trials, showing adaptation or facilitation effects, for any of the category of stimuli (familiar and unknown faces and places; see Fig. S7). Therefore, when referring to familiarity effects, we refer to the responses to the pictures initially classified as familiar (or unknown; see “Methods” section) and not to how pictures may have become more familiar after several presentations.

For both places and faces, in the grand averages we observe similar responses to the familiar and the unknown pictures. To quantify this observation, we calculated the strength and latency of the individual responses (see “Methods” section). For each responsive



**Fig. 1 | Electrode location in one patient.** Axial, coronal, and sagittal views of a CT-MRI coregistration, showing the estimated position of the microwires (at -3 mm from the tip of the intracerebral electrode) marked with a cross, and functional Magnetic Resonance Imaging (fMRI) showing the midfusiform gyrus activation.

The colorbar on the right denotes the normalized (*z*-score) activation compared to baseline (see “Methods” section). The inset on the top right shows the depth electrodes with microwires used for single neuron recordings. Source data are provided as a Source Data file.



**Fig. 2 | Example of 4 neuronal responses in the face-selective midfusiform cortex.** **A** A single neuron responding to pictures of faces (and not to pictures of places), without a significant difference between the responses to familiar and unknown faces (two-sided  $t$ -test). **B** A single neuron also showing face-responses but with a significantly larger response to the familiar faces ( $p < 0.05$ , two-sided  $t$ -test). For this neuron, there was also a less pronounced but significant decrease in firing in response to pictures of places ( $p < 0.05$ , two-sided  $t$ -test). **C** A face-

responsive single neuron, with a significantly larger response to unknown faces ( $p < 0.05$ , two-sided  $t$ -test). This modulation, however, is very different to that of the previous neurons, showing a more variable response pattern, responding selectively to a few unknown faces. **D** A multi-unit with a significant response to places, without modulation by familiarity. The shaded areas represent SEM. Source data are provided as a Source Data file.

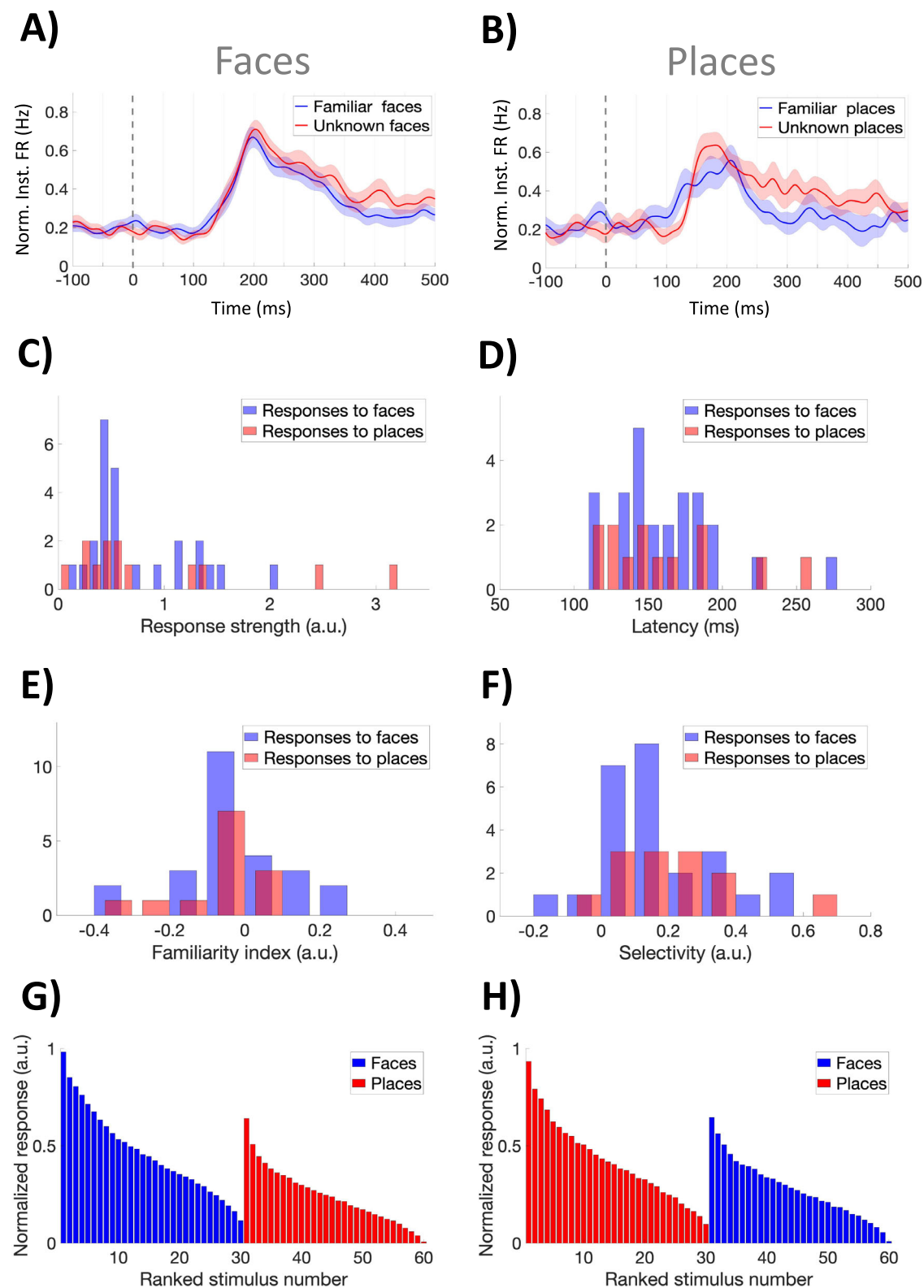
neuron, the strength was normalized to the baseline firing activity (see “Methods” section and Fig. S11). There were no significant differences in the strength of the face and place response distributions (Fig. 3c; Wilcoxon rank-sum test); in addition, for both faces and places there were no significant differences between the familiar and unknown stimuli (Wilcoxon rank-sum test). Figure 3d shows the response onset latencies, which varied mostly between 100 and 200 ms, and were not significantly different between the face and place responses (median for face responses: 156 ms; median for place responses: 150 ms; Wilcoxon rank-sum test), or between the familiar and unknown stimuli in each case (Wilcoxon rank-sum test). Altogether, given that the responses to faces and places were not significantly different, the larger responses to faces observed at the population level in the midfusiform gyrus can be attributed to a larger proportion of neurons responding to faces in this region.

In the above analyses, we do not observe major differences between the responses to familiar and unknown faces or places. To further quantify whether midfusiform units tend to respond preferentially to familiar stimuli, we defined a familiarity index ( $FI$ ), which gives positive (negative) values for units responding more strongly to familiar (unknown) stimuli, and values around zero for those

responding similarly to both types of stimuli (see “Methods” section). Figure 3e shows the distributions of the familiarity index values, which were not significantly different for the face and place responses (Wilcoxon rank-sum test) and had values centered around zero (for each distribution the median was not significantly different from zero; sign test), meaning that, overall, the familiar and unknown stimuli elicited similar activations.

As in previous work<sup>25</sup>, to study the selectivity of the responses we used a measure that is independent of any threshold for defining responsiveness, and gives values around 0 for non-selective neurons and close to 1 for the selective ones (see “Methods” section). In Fig. 3f we observe that the responses of midfusiform units had a wide range of selectivity values, which were not significantly different between the face and the place responses, or between the familiar and unknown stimuli in each case (Wilcoxon rank-sum test). Considering our responsiveness criterion (see Methods), the responsive neurons responded to 15% of the presented stimuli on average. To visualize the overall selectivity of these neurons, in Fig. 3g, h we show the ranked normalized responses for the face and place selective units, respectively. In these plots we observe a gradual decrease of the responses, showing a graded tuning, as has been described in object and





**Fig. 3 | Characterization of single neuron responses.** **A, B** Normalized grand average of enhanced responses to faces ( $N=25$ ) and places ( $N=13$ ). In both cases, the responses to the familiar and unknown stimuli were very similar. **C, D** Distribution of response strength and latency, with no significant differences between the responses to faces and places (**C**: bin size=0.1; **D**: bin size=10). **E** Familiarity index values for the place and the face responses. Values are centered around zero—i.e. not showing any clear tendency to respond to familiar or

unknown pictures—and, as before, there were no differences between the face and place responses (**E**: bin size=0.1). **F** Selectivity values, also showing no differences between the face and place responses- (**F**: bin size=0.1). **G, H** Average of the normalized ranked responses for the face- and place-responsive units, respectively. In both cases, a graded pattern of responses was observed, with face-responsive neurons also responding to places, and vice versa. The shaded areas represent SEM. Source data are provided as a Source Data file.

face-selective regions of the monkey visual cortex<sup>1,5,7</sup> and in contrast to the nearly binary coding by downstream neurons in the human hippocampal formation<sup>26</sup>. Moreover, although with a lower strength, the face-responsive units also responded to some of the places (and vice versa), with the maximum response to the non-preferred stimulus category reaching about 60–70% of the maximum value for the preferred stimulus category.

### Coding of face familiarity

Given the mixed results obtained with the analysis of individual unit responses, we carried out a population decoding analysis—predicting in each trial the stimulus being displayed based on the neurons' responses—to assess whether the midfusiform neurons provide a reliable discrimination between faces and places, on the one hand, and of familiar versus unknown faces, on the other. Considering all 43 responsive units (similar results were obtained when using all the recorded units), the mean decoding performance discriminating the presentation of faces *vs.* places was much larger than chance (91%,  $p < 10^{-87}$ ; permutation test; Fig. 4a). Likewise, taking only the 36 face-responsive units, the discrimination between familiar and unknown faces was also significantly larger than chance (70%,  $p < 10^{-11}$ ; permutation test; Fig. 4b), but the discrimination between familiar and unknown places, considering the 24 place-responsive neurons, was only marginally significant (54%,  $p = 0.085$ ; permutation test; Fig. 4c). Considering all responsive units and all picture categories (familiar or unknown faces or places), the decoder was able to distinguish whether the pictures corresponded to a face or a place (Fig. 4d). In fact, in only 9% of the cases a face was confused with a place—a much lower rate than that expected by chance ( $p < 10^{-92}$ ; permutation test). As before, the decoder could predict whether the stimulus corresponded to a familiar or unknown face, but it confused familiar *vs.* unknown places.

Figure 4d shows the performance obtained when decoding the 60 pictures presented using all responsive units. We observe a clear pattern along the diagonal, with an average hit rate of 22%, significantly higher than chance (1/60 = 1%;  $p < 0$ ; permutation test). In particular, the average number of correct predictions for faces (27%) was significantly higher than for places (18%;  $p < 0.05$ , *t*-test) and, for both faces and places, there were no significant differences in the predictions of the familiar versus the unknown individual pictures (*t*-test). The high decoding performance obtained for the individual pictures raises the concern that the familiarity information obtained for faces (Fig. 4b) could be a byproduct of the relatively high discriminability of the individual faces—i.e. midfusiform neurons could just discriminate individual faces without providing any specific information about their familiarity. To address this issue, we repeated the decoding analysis of familiar versus unknown faces but shuffling the familiar/unknown labels of each picture, and found that the performance obtained was significantly lower than when not shuffling the picture labels ( $p < 0.005$ , permutation test), thus confirming that these units contain information about both the familiarity of the faces and picture identity. Likewise, when shuffling the 'face'/'place' labels of the stimuli, we obtained a significantly lower performance decoding faces versus places, compared to the one obtained without shuffling the labels shown in Fig. 4a ( $p < 10^{-32}$ ). However, reshuffling the familiar/unknown labels of the place pictures did not significantly change the decoding performance ( $p = 0.6$ ) and therefore, the marginally significant performance when decoding familiar *vs.* unknown places was likely due to the discriminability of one or relatively few individual places. To further confirm these results, we also performed a K-fold decoding crossvalidation—excluding from the training set all the trials corresponding to the picture being tested—and obtained qualitatively similar results (Fig. S8).

### Three face-related discriminations at the population level

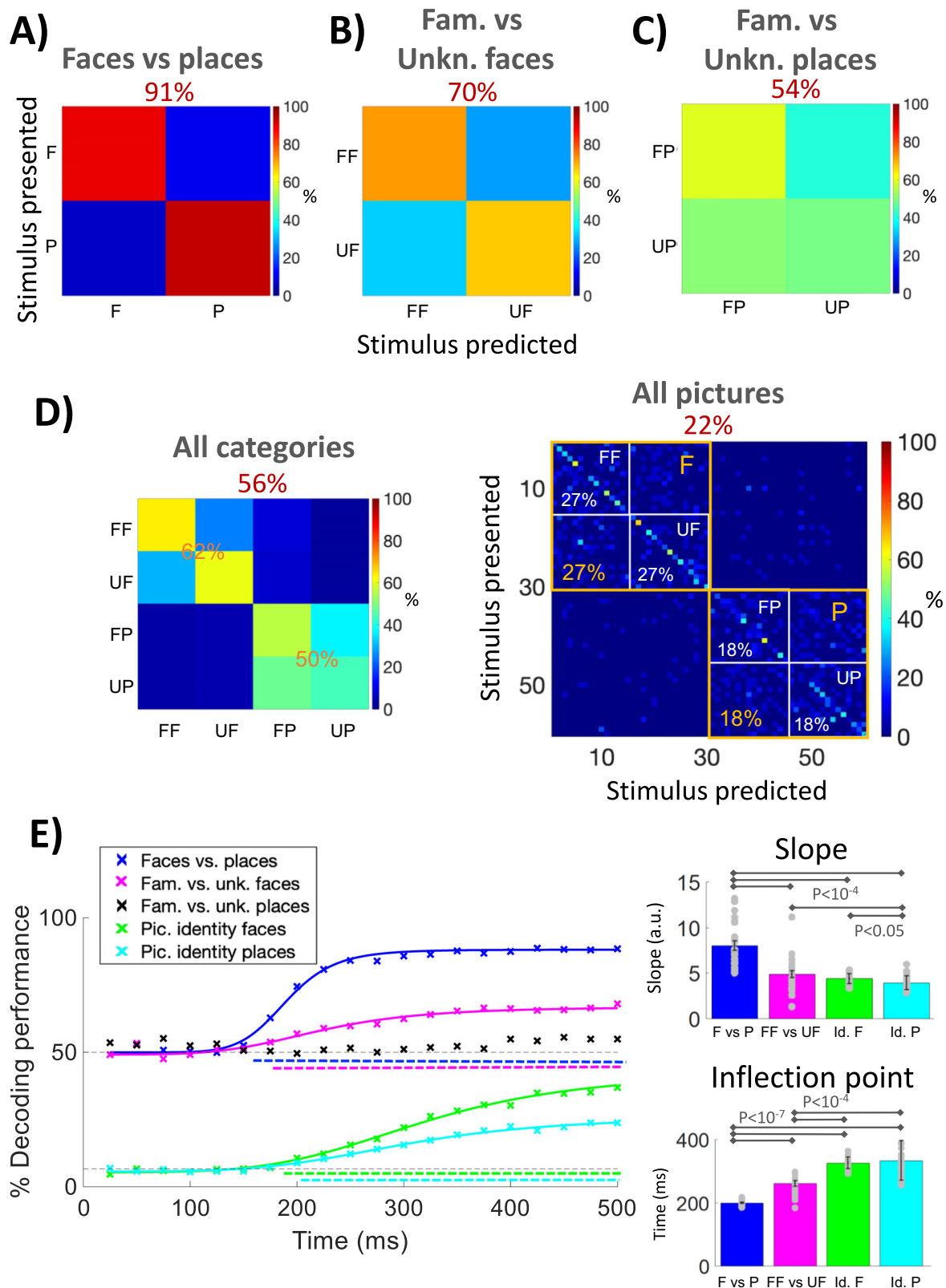
From the decoding results described above, we infer that there are 3 discriminations taking place at the population level in the face-

selective midfusiform gyrus: face detection, familiarity face recognition, and picture identification. To investigate the time profile of these discriminations, we carried out a decoding analysis in which we successively expanded the time window considered (i.e. the information that would in principle be accessible to downstream neurons integrating inputs from the midfusiform population), predicting whether the picture was a face or a place, a familiar or an unknown face, and the identity of the face (place) picture. Figure 4e shows that the decoding curves had different time profiles, with the face *vs.* place discrimination showing an earlier increase compared to the other curves, and an inflection point occurring significantly earlier and with a significantly higher slope (left panels of Fig. 4e). The inflection point was also significantly earlier in the familiar *vs.* unknown face curve compared to the face and the place picture identity decoding curves. Both a time-resolved decoding analysis using a sliding window, and a cross-time decoding analysis considering different time windows in the training and test sets, gave qualitatively similar results (shown in Figs. S9 and S10, respectively).

### Discussion

Contrasting with the vast neuroimaging and electrophysiological evidence characterizing different aspects of face-selective activity in the human midfusiform gyrus<sup>8–13,18,19</sup>, the neural mechanisms producing such activations have remained elusive, due to the lack of single neuron recordings in this region (for a few recent studies recording face cells in the human VOTC, see ref. 27 and the single cases reported in refs. 28,29). Neurons responding to faces have long been identified in the temporal cortex of macaque monkeys<sup>1,4,5</sup>, with fMRI-defined face-selective clusters containing nearly exclusively neurons responding to faces<sup>6,7</sup> (but see ref. 30) and suggesting that a similar type of neuronal responses may underlie the fMRI activations observed in humans<sup>11</sup>. Here, recording directly from individual neurons in the face-selective human midfusiform gyrus, we found that: First, in humans there is also a tendency to find face-responsive neurons. Second, midfusiform neurons did not fire only to faces, as 30% of the recorded units fired to places. Third, the responses to faces and places in face- and place-responsive units, respectively, were statistically the same in terms of their strength, latency and selectivity. These results suggest that the stronger responses to faces compared to non-face objects found with neuroimaging, scalp and intracranial EEG recordings in humans<sup>9,10,12</sup> are due to a higher proportion of face-responsive neurons in this region rather than broadly tuned neurons responding stronger to faces compared to other category of stimuli.

Although in our study there was a tendency to find face-responsive units in the fusiform gyrus (46%), the proportion of these units was not as high as reported in macaques (~90% or more)<sup>6,7</sup>. In this respect, our results are actually closer to those in non-human primate studies that sampled more broadly within face-selective clusters (ref. 30, see also ref. 31 for intermediary values). The lack of nearly exclusive face-selective responses reported here could be due to differences across species and regions (i.e., recordings in monkeys were performed in the lateral temporal cortex, whereas here we recorded ventrally in the fusiform gyrus—a region that is not present in macaques and is specific to hominoids/apes<sup>32</sup>). Another possibility is that this structure is indeed specialized for faces but also represents other non-face stimuli, given the richness of perceptual discriminations across a vast number of meaningful stimuli categories determined essentially by natural experience (and not potentially biased by extensive reward-driven training with particular sets of stimuli). However, it should also be noted that, due to clinical constraints, the placement of the electrodes varied from patient to patient (see Fig. S3) and was not necessarily at the peak of the face-selective activations (as defined from the independent fMRI and LFP face-localizer experiments), which is the area typically targeted in monkeys<sup>6,7</sup>. This might be particularly relevant, considering the changes in the tuning of single



neurons in a specific area reported at the submillimeter scale<sup>33</sup>. However, it is important to mention that, in spite of the variable localization of the electrodes within the midfusiform gyrus, a similar proportion of face and place responses was found for all patients in this study (Fig. S5). Further studies including more subjects with recording locations close to or at the peak of the face-selective activations will be

required to determine if the finding of nearly-exclusive face-selective neurons, as in the monkeys' face patches, is also present in the human brain.

Due to lack of direct single neuron recordings, the role of the midfusiform region in coding long-term face familiarity has been controversial<sup>11,18,19,34</sup>, with some neuroimaging studies finding different

**Fig. 4 | Population decoding results.** **A** Decoding of whether the picture was a face (F) or a place (P), using all responsive units ( $N = 43$ ); **B** decoding of whether the picture was a familiar face (FF) or unknown face (UF), using face-responsive units ( $N = 36$ ); **C** decoding of whether the picture was a familiar place (FP) or unknown place (UP), using place-responsive units ( $N = 24$ ). In all cases, the overall decoding performance is shown on top. **D** decoding of all 4 categories (familiar face, unknown face, familiar place, and unknown place) using all responsive units (left), and decoding of the pictures presented using all responsive units (right). Decoding performance was similar for familiar and unknown stimuli (for both faces and places) and was higher for faces (27%) compared to places (18%). **E** Cumulative decoder using all responsive units and discriminating: faces vs. places, familiar vs. unknown faces, familiar vs. unknown places (not significant), and the picture

identity of the face and place stimuli. The solid lines mark the corresponding logistic regressions (see “Methods” section). The top and bottom dashed horizontal lines mark the chance level for the binary decisions (face vs. place; familiar vs. unknown face; familiar vs. unknown place) and discrimination of face/place identify, respectively. The colored dashed lines below the chance levels show the time periods where each decoder is significantly larger than chance (permutation test,  $p < 0.05$ ). We observe different time profiles of the decoders, with the decoding of faces vs. places having an earlier inflection point and higher slope compared to the other ones (right panels; two-sided  $t$ -test;  $N = 24$ ). The inflection point was also earlier for the familiar vs. unknown face decoder compared to the face and place picture identification. Source data are provided as a Source Data file.

activations for familiar compared to unfamiliar faces<sup>8,17,35–38</sup> but the majority reporting no familiarity effects in the fusiform gyrus<sup>34,39–45</sup> (among others). Humans are familiar with thousands of faces<sup>22</sup>, and familiar faces have a special status, being associated with rich semantic networks<sup>22,23,46</sup>. Our results show that neurons in the midfusiform region have an equal tendency to respond to familiar and unknown faces, and no significant differences were observed in the grand averages, or in the average strength, latency, and selectivity of the responses. However, given that about 30% of the face responsive units showed a familiarity modulation (with a larger response either to the familiar or the novel faces but, as mentioned above, without showing a difference in the averages) when performing a single-trial decoding analysis of the midfusiform population activity – exploiting the single neuron resolution of our recordings and extracting information beyond the one provided by grand averages or non-invasive recordings – the decoder gave reliable information about whether the pictures corresponded to a familiar or unknown face. This result indicates that neurons in this relatively posterior region within the cortical face network respond beyond mere physical features of the faces and are already shaped by the subjects’ previous experience, i.e., making them familiar with certain faces identities – an information that could be read by downstream anterior temporal lobe areas to differentially process familiar faces and their meaning<sup>23,47</sup>.

Overall, the single-trial population analysis led to identifying three face-related discriminations taking place in the midfusiform gyrus: the categorization of the stimulus as a face, familiarity recognition, and (face and place) picture identification. While the discrimination of the stimulus as a face saturated early, followed by the discrimination of its familiarity, the discrimination of the individual (face and place) pictures kept increasing in the time window considered, reaching a performance more than 10 times larger than what would be expected by chance (Fig. 4d). This contrasts to relatively small or no increases of performance above chance reported with fMRI multivoxel pattern analyses in face-selective cortex<sup>11,48,49</sup>, which led some authors to argue that the midfusiform cortex may merely detect faces, with the identity of the faces being recognized in more anterior cortical regions<sup>48</sup>. Instead, our results at the level of individual neurons agree with the view derived from fMRI and iEEG adaptation paradigms, showing that the face-selective midfusiform cortex is involved in both face detection and face individuation<sup>50,51</sup>, although it should be noted that we did not test different views of the same faces to evaluate generalization in individual face discrimination. The finding of three different time profiles of face-related activations (face detection, followed by familiarity recognition, and then followed by face image identification) is in line with a coarse-to-fine processing of information, with e.g., an earlier representation of the age and gender of the faces before identity information<sup>52</sup>.

The human medial temporal lobe (MTL), and within it, the hippocampus, receives direct projections from high-level visual areas<sup>53</sup> and has neurons with very selective and invariant responses to specific persons or places<sup>54,55</sup>. Compared to the upstream midfusiform neuron responses reported here, hippocampal neuron responses occur much

later (i.e. with an onset of -300 ms compared to -150 ms in the midfusiform). This latency gap suggests the existence of much further neocortical processing – possibly to derive stimulus meaning<sup>54,56</sup> – before the neural signals reach the hippocampal memory system. Moreover, human hippocampal neurons have a lower baseline activity (2.2 Hz<sup>57</sup> vs. 4.8 in the fusiform gyrus; see Fig. S11), nearly binary tuning curves (rather than graded responses, as shown here in the midfusiform), and respond mainly to familiar stimuli<sup>54</sup>. Another notable difference is the much lower selectivity we found in the fusiform gyrus, with responsive neurons being on average activated by 15% of the stimuli, compared to the selectivity of hippocampal neurons, responding on average to 1.7% of the stimuli<sup>57,58</sup>. Future research should provide more insights into how the representation of faces, involved in face recognition in the midfusiform gyrus, leads to a selective and invariant representation of familiar persons, for memory functions in the MTL<sup>54</sup>.

## Methods

Subjects performed a main experiment in which they were shown pictures of familiar and unknown faces and places, while we recorded the activity of multiple single neurons in the midfusiform gyrus. Moreover, in order to determine the location of the electrodes with respect to the face-selective midfusiform region, we performed two additional, independent experiments with the same frequency-tagging approach to isolate face-selective voxels in fMRI<sup>59</sup> (see section “fMRI face localizer” below) and face-selective local field potentials<sup>13</sup> (see section “LFP face localizer” below). Except explicitly mentioned otherwise, all quantifications described in this study (e.g., responsiveness, selectivity, etc.) correspond to the unit responses of the main experiment, as described in the following sections.

## Main experimental design

The data comes from 7 sessions in 5 patients with pharmacologically intractable epilepsy (all right-handed, 4 females, 24–42 years old), who were implanted with depth electrodes for about a week to determine the seizure onset zone for possible subsequent surgical resection<sup>54</sup>, at the epilepsy unit of the University Hospital of Nancy, France. All patients gave written informed consent and the study, including number of patients, selection criteria, and electrode implantation locations (REUNIE, trial No. 2015-A01951-48, ClinicalTrials.gov identifier NCT02877576) was approved by the ethical committee CPP Est III, No 16.02.01. The electrode locations were based exclusively on clinical criteria and were verified by CT fused with preoperative MRI. Here we report data from sites in the right (2 patients) and left (3 patients) face-selective midfusiform region. Each electrode probe had a total of 9 micro-wires at its end, 8 active recording channels and one reference (AdTech), protruding -3 mm from the tip of the depth electrodes<sup>54</sup>. The differential signal from the micro-wires was amplified using a Blackrock system (NeuroPort Central Suite version 6.5.4), filtered between 0.3 and 7500 Hz and sampled at 30 kHz.

Subjects sat in bed, facing a laptop computer on which 60 images were presented for 500 ms, in pseudorandom order and without



blanks in between, with a software written in Matlab (R2018a) that used the Psychophysics Toolbox. This sequence of picture presentations was triggered with a keypress by the patient, it lasted 30 s and was repeated 15 times. To ensure that they paid attention to the pictures, the patients had to indicate with a key press whether one of two color bars (either red or green), on the top and bottom borders of the pictures, changed color (Fig. S1a). The experiment lasted about 10 min in total.

The laptop was placed in a hospital tray, -50–70 cm away from the patient, with each picture sustaining about 2 degrees visual angle from their view. In order to assess whether neurons respond preferentially to faces and how they encode familiarity, the pictures corresponded to 4 categories: 15 known faces (famous actors and politicians, in principle known to the patients), 15 unknown faces (local actors and politicians from Argentina, in principle unknown to the patients), 15 known places (Eiffel Tower, Arc de Triomphe, etc.), and 15 unknown places (local and mostly unknown scenes from UK towns; Fig. S1b). We contrasted faces to places because: (1) places are commonly contrasted to faces in human neuroimaging and electrophysiological studies<sup>10</sup>, and (2) familiar and unknown items can be easily selected for both types of stimuli to evaluate specific effects of face familiarity. To ensure that the patients were familiar with the people and places of the ‘known’ category compared to the ‘unknown’ one, they were asked to rank how much they knew the person/place of the pictures using a scale from 1 (completely unknown) to 7 (very familiar). Four of the five patients filled this questionnaire and for each of these patients we found a significantly higher score both for the faces and places of the ‘known’ category (in all cases  $p < 0.05$ ; T-test). Rather than transforming the images in order to match their low-level visual features, which is prone to make the images look less natural compared to the original versions<sup>60</sup>, we chose a variety of images with overlapping distributions of low-level features in the 4 categories, with which we verified that the FFA neuron responses cannot be trivially attributed to low-level visual features (see Fig. S2).

### Statistics analysis

The significance of the single neuron responses was established using paired t-tests, comparing baseline with response time windows (see below), and for the main statistical analysis, the differences of the responses to the four categories were assessed using non-parametric Wilcoxon rank-sum tests (see below). No data were excluded from the analyses and no statistical method was used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### Spike sorting and responsiveness criteria

From the continuous wide-band data, spike detection and sorting was carried out using ‘Wave\_Clus’, an unsupervised clustering algorithm<sup>61,62</sup>. Neurons were classified into single- or multi-units based on: (1) the spike shape and its variance; (2) the ratio between the spike peak value and the noise level; (3) the ISI distribution of each cluster; and (4) the presence of a refractory period for single-units; i.e., <1% spikes within <3 ms ISI<sup>63</sup>. Altogether we recorded a total of 78 units (33 single-units and 45 multi-units). The mean number of detected units per micro-wire was 1.39 (s.d.: 0.82).

For each stimulus we defined a “response window” between 100 and 400 ms from stimulus onset, and a corresponding “baseline window” between -100 and 100 ms. A unit was considered to be responsive if, either for the face or for the place stimuli, the average number of spikes in the response window was significantly different than the one in the baseline window, according to a paired t-test with  $p < 0.05$ . Considering the probability of observing a response to the face/place stimuli with a chance level of 0.05, we estimated the probability of

obtaining  $N_f$  and  $N_p$  responses by chance (with  $N_f$  and  $N_p$  the number of face and place responses from the 78 recorded units, respectively) using a binomial test. For the face-responsive units, we assessed whether they showed a significant familiarity modulation, by comparing the responses to unknown versus known faces, and likewise, we compared the responses to unknown versus known places for the units responding to places (t-test,  $p < 0.05$ ).

### Familiarity index

In order to quantify a tendency of the neurons to respond more to familiar or unknown pictures (faces or places), we defined a familiarity index as:

$$FI = \frac{\langle \text{Response Familiar} \rangle - \langle \text{Response Unknown} \rangle}{\langle \text{Response Familiar} \rangle + \langle \text{Response Unknown} \rangle} \quad (1)$$

where  $\langle \text{Response Familiar} \rangle$  and  $\langle \text{Response Unknown} \rangle$  correspond to the mean response to the familiar and unknown stimuli (faces or places). Units with no familiarity modulation will have values around zero, whereas units with a larger response to familiar (unknown) pictures will tend to have positive (negative) values. The familiarity index values obtained for the face and the place responses were compared using a non-parametric Wilcoxon rank-sum test, and we further used a sign test to evaluate if the median of each of the distributions was significantly different from zero (thus showing a preference for familiar or novel pictures).

### Instantaneous firing rate and latency estimation

For each significant response, the response latency was estimated from the instantaneous firing rate, which was obtained by convolving the spike train with a Gaussian of 10 ms s.d. and then averaging across trials. The latency was then computed as the time where the instantaneous firing crossed the baseline plus 2 s.d. value (or minus 2 s.d. for suppressed responses) for at least 50 ms. For some responsive units, a response latency could not be calculated according to this criterion (i.e., the instantaneous firing rate did not cross the baseline plus 2 s.d. for more than 50 ms) and therefore these cases were not included in the latency distribution. Latency distributions were compared using a non-parametric Wilcoxon rank-sum test. For the grand average curves (Fig. 3a, b and S6a, b), each face/place response was normalized to have a maximum range (either for familiar or unknown faces/places) between 0 and 1 for enhanced responses, and 0 and -1 for the suppressed responses.

### Strength and selectivity

For each face- (place-) responsive unit, we calculated the normalized response strength as:  $RS = \frac{\langle R - B \rangle}{B}$ , where  $R$  are the responses to faces (places),  $B$  the baseline activity and  $\langle \dots \rangle$  denotes the mean across face (place) stimuli. We also estimated the selectivity of each responsive unit using a measure that it is not dependent on a responsiveness threshold, as used in previous works<sup>25,64</sup>. For this, we calculated the relative number of stimuli  $R(T)$  eliciting a response larger than a threshold  $T$  as:

$$R(T) = \frac{1}{N} \sum_{i=1}^N \theta(f_i - T) \quad (2)$$

where  $f_i$  denotes the mean response to each face/place stimuli,  $N (= 30)$  denotes the number of face/place stimuli, and  $\theta(x) = 1$  for  $x > 0$  and  $\theta(x) = 0$  for  $x \leq 0$ . Then, we defined the selectivity index as:

$$S = 1 - \frac{2}{M} \sum_{j=1}^M R(T_j) \quad (3)$$

where  $T_j = f_{\min} + j(f_{\max} - f_{\min})/M$  are equidistant thresholds between the minimum and maximum response, respectively, and  $M$  denotes the

number of thresholds used ( $M = 1000$ ). This measure is independent of a particular threshold used to define responses, is close to 0 for uniformly distributed random responses, and approaches 1 the more selective the unit is. For more details and comparisons with other selectivity measures, see<sup>25</sup>. Statistical comparisons for both the response strength and the selectivity distributions were done using non-parametric Wilcoxon rank-sum tests.

### Decoding analysis

A linear Bayesian classifier was used to decode the pictures presented to the subjects based on the firing of the responsive units in the response window (from 100 ms to 400 ms after stimulus onset<sup>65</sup>). One at a time, the spike count of each responsive unit in each trial was used to test the classifier, which was trained with the remaining trials (leave-one-out cross-validation). The decoding performance was estimated as the percentage of trials correctly predicted, and its statistical significance was assessed in comparison to the performances obtained with a population of 1000 surrogates created by randomly shuffling the trial labels.

The cumulative decoding performance was calculated from time zero until 500 ms after stimulus onset, in steps of 25 ms. Cumulative decoding performances, showing a sigmoidal increase of information until reaching saturation, were fitted with a 4-parameter logistic regression:

$$f(x) = d + \frac{a - d}{1 + \left(\frac{x}{c}\right)^b} \quad (4)$$

where  $a$  and  $d$  are the upper and lower asymptotes, determined by the maximum and chance level decoding performance, respectively,  $c$  is the inflection point, and  $b$  the slope of the curve at the inflection point. The four coefficients were estimated with an iterative least squares estimation using the MATLAB function *nlinfit*. To compare the different cumulative decoding curves, we created 30 realizations with a random selection of half the number of units each, discarding 6 realizations in which the root mean square error of the fit of at least one of the cumulative curves was  $>0.1$ . The slope and inflection points of the curves were compared using an ANOVA test and  $t$ -test post-hoc comparisons. The statistical significance of decoding performance was established using a permutation test (comparing with the values of the 1000 surrogates shuffling labels, using a rank test).

### FMRI face localizer

To verify the position of the electrodes with respect to the face-selective midfusiform gyrus, patients performed an additional fMRI face-localizer experiment (independently from the experiments described above) within 4 weeks after the intracerebral electrode implantation. We used a frequency-tagging fMRI design, which provides high signal-to-noise ratio (SNR) and reliability to define brain regions selective to human faces versus non-face objects, independently of low-level visual cues<sup>59</sup>. In this paradigm, inspired from EEG frequency-tagging studies, images are presented at a fast 6 Hz rate, using different images of nonface objects ( $N = 188$ ) displayed for 7 sec, interleaved with mini-blocks of 7 (out of 28) face images shown every 9 sec (i.e.,  $1/9$  s = 0.11 Hz). Within a mini-block, the 7 face images are interleaved with six object images to avoid adaptation effects and increase SNR<sup>59</sup>. Since all face images are presented at the same 6 Hz base frequency rate as the nonface object images, neural activity that is selective to faces is expressed at 0.11 Hz in the frequency spectrum of the fMRI signal<sup>59</sup>. Images subtended a viewing angle of 8° and were back-projected onto a projection screen by an MRI-compatible LCD projector. Each fMRI sequence (run) lasted 333 sec. One to four sequences (runs, mean  $2.5 \pm 1$  runs) were tested for each patient.

**MR image acquisition.** All patients were scanned at the University Hospital of Nancy, using a 3 T Siemens Magnetom Prisma system with a

64-channel head-neck coil. Anatomical images were collected using a high-resolution T1-weighted magnetization-prepared gradient-echo image (MP-RAGE) sequence (192 sagittal slices, voxel size = 1 mm isotropic; flip angle (FA) = 9°, field of view (FOV) =  $256 \times 256$  mm<sup>2</sup>, matrix size =  $256 \times 256$ ). The acquisition time for the anatomical scan was about five min. Functional images were collected with a T2\*-weighted simultaneous multi-slice echo planar imaging (SMS EPI) sequence (TR = 1500 ms, TE = 34 ms, FA = 72°, FOV =  $240 \times 240$  mm<sup>2</sup>, voxel size = 2.5 mm isotropic, matrix size =  $96 \times 96$ , interleaved), which acquired 44 oblique-axial slices covering the entire temporal and occipital lobes.

**fMRI Analysis.** The functional runs were motion-corrected with reference to the average image of the first functional run using a 6-degree rigid body translation and rotation using a motion correction software (MCFLIRT; [www.fmrib.ox.ac.uk/fsl](http://www.fmrib.ox.ac.uk/fsl)), and spatially smoothed with a Gaussian kernel of 3 mm (FWHM; i.e. full width at half maximum). Linear trends from the preprocessed time series of each voxel were removed. The SNR of the face-selective response at 0.11 Hz was calculated for each voxel independently, as in previous studies<sup>59</sup>. The activation and deactivation of the neural responses at the stimulation frequency was defined by the phase of the BOLD responses<sup>59</sup>. To define face-selective activations, we averaged each voxel's time series across all face runs (Fig. S3). A response was considered as face-selective if the  $Z$ -score at the frequency bin of face stimulation exceeded 3.09 (i.e.,  $p < 0.001$  one-tailed: signal  $>$  noise).

**Visualization.** The functional activation map across runs for each patient was registered with their high-resolution T1-weighted image. To assess the spatial relationship between face-selective activations at the microwire locations, the T1-weighted images were fused with the post-operative CT-scan, and the tip of the macroelectrodes, as they appear on the CT-scan, was manually detected, and we estimated the coordinate of the microwires to be at about 3 mm from the macroelectrode tip (marked with a cross in Fig. 1 and Fig. S3). Finally, we rendered the estimated microwire coordinates into the T1-weighted MRI co-registered with the functional activation maps.

### LFP face localizer

To further validate the location of the electrodes with respect to the face activation areas, a second independent experiment was performed, in which the Local field potential (LFP) activity was recorded from the microwires with a frequency-tagging face localizer that has been used and validated in previous studies<sup>13</sup>. Participants viewed continuous sequences of natural grayscale images of objects ( $N = 200$ ) presented at 6 Hz, with images of faces ( $N = 50$ ) shown every 5th stimulus (i.e., 1.2 Hz). A sequence lasted 70 s. Four participants viewed 8 sequences, and one viewed 6 sequences. Half of the sequences were typically run before the main experiment. In this paradigm (analogous to the frequency tagging paradigms described in the previous sections), common neural activity to faces and non-face objects is expressed at the 6 Hz base frequency rate (and harmonics, i.e., 12 Hz, 18 Hz, etc.), while neural activity that is selective to faces is expressed at 1.2 Hz and harmonics in the LFP frequency spectrum<sup>13,66</sup>.

**LFP analysis.** A fourth-order zero-phase low-pass filter at 300 Hz was applied to each sequence. Data segments were cropped to contain an integer number of 1.2 Hz cycles beginning 2 s after the onset of the sequence, until approximately 68 s, averaged in the time-domain separately for each participant and each active recording channel (8 channels) and Fast Fourier Transformed (FFT). The FFT spectrum was cut into 4 segments centered at the fundamental face frequency and 3 harmonics (1.2 Hz to 4.8 Hz). Amplitude values in these 4 segments were summed and averaged across the 8 active channels in each participant.  $Z$ -scores were computed using the mean and standard

deviation of 20 surrounding frequency bins. A response was considered as face-selective if the Z-score at the frequency bin of face stimulation exceeded 3.09 (i.e.,  $p < 0.001$  one-tailed: signal > noise; Fig. S4).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The data used in this study is available at: <https://doi.org/10.25392/leicester.data.23459153.v1>. Source data are provided with this paper.

### Code availability

Codes for data processing (spike sorting) are available at: [https://github.com/csn-le/wave\\_clus](https://github.com/csn-le/wave_clus).

### References

- Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* **19**, 577–621 (1996).
- Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19**, 109–139 (1996).
- Kravitz, D., Saleem, K., Baker, C., Ungerleider, L. & Mishkin, M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**, 26–49 (2013).
- Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* **4**, 2051–2062 (1984).
- Gross, C. Single neuron studies of inferior temporal cortex. *Neuropsychologia* **46**, 841–852 (2008).
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. & Livingstone, M. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
- Freiwald, F. A. & Tsao, D. Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**, 845–851 (2010).
- Sergent, J., Ohta, S. & MacDonald, B. Functional neuroanatomy of face and object processing: a positron emission tomography study. *Brain* **115**, 15–36 (1992).
- Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).
- Grill-Spector, K., Weiner, K. S., Kay, K. & Gomez, J. The functional neuroanatomy of human face perception. *Annu. Rev. Vis. Sci.* **3**, 167–196 (2017).
- Kanwisher, N. The quest for the FFA and where it led. *J. Neurosci.* **37**, 1056–1061 (2017).
- Allison, T., Puce, A., Spencer, D. & McCarthy, G. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* **9**, 415–430 (1999).
- Jonas, J. et al. A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proc. Natl Acad. Sci. USA* **113**, E4088–E4097 (2016).
- Barton, J. Structure and function in acquired prosopagnosia: lessons from a series of 10 patients with brain damage. *J. Neuropsychol.* **2**, 197–225 (2008).
- Parvizi, J. et al. Electrical stimulation of human fusiform face-selective regions distorts face perception. *J. Neurosci.* **32**, 14915–14920 (2012).
- Jonas, J. et al. A face identity hallucination (palinopsia) generated by intracerebral stimulation of the face-selective right lateral fusiform cortex. *Cortex* **99**, 296–310 (2018).
- Rossion, B., Schiltz, C. & Crommelinck, M. The functionally defined right occipital and fusiform “face areas” discriminate novel from visually familiar faces. *Neuroimage* **19**, 877–883 (2003).
- Gobbini, M. I. & Haxby, J. V. Neural systems for recognition of familiar faces. *Neuropsychologia* **45**, 32–41 (2007).
- Natu, V. & O’Toole, A. The neural processing of familiar and unfamiliar faces: a review and synopsis. *Br. J. Psychol.* **102**, 726–747 (2011).
- Landi, S. & Freiwald, W. A. Two areas for familiar face recognition in the primate brain. *Science* **357**, 591–595 (2017).
- Landi, S., Vishwanathan, P., Serene, S. & Freiwald, W. A. A fast link between face perception and memory in the temporal pole. *Science* **373**, 581–585 (2021).
- Jenkins, R., Dowsett, A. & Burton, A. How many faces do people know? *Proc. R. Soc. B* **285**, 20181319 (2018).
- Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. The distributed human neural system for face perception. *Trends Cogn. Sci.* **4**, 223–233 (2000).
- Weiner, K. S. & Grill-Spector, K. Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. *Psychol. Res.* **77**, 74–97 (2013).
- Quiaro, R., Reddy, L., Koch, C. & Fried, I. Decoding visual inputs from multiple neurons in the human temporal lobe. *J. Neurophysiol.* **98**, 1997–2007 (2007).
- Rey, H. et al. Encoding of long-term associations through neural unitization in the human medial temporal lobe. *Nat. Commun.* **9**, 4372 (2018).
- Khuvis, S. et al. Face-selective units in human ventral temporal cortex reactivate recall. *J. Neurosci.* **41**, 3386–3399 (2020).
- Axelrod, V. et al. Face-selective neurons in the vicinity of the human fusiform face area. *Neurology* **92**, 197–198 (2019).
- Decramer, T. et al. Single-unit recordings reveal the selectivity of a human face area. *J. Neurosci.* **41**, 9340–9349 (2021).
- Bell, A. et al. Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J. Neurosci.* **31**, 12229–12240 (2011).
- Issa, E. & Di Carlo, J. J. Precedence of the eye region in neural processing of faces. *J. Neurosci.* **32**, 16666–16682 (2012).
- Bryant, K. et al. Organization of extrastriate and temporal cortex in chimpanzees compared to humans and macaques. *Cortex* **118**, 223–243 (2019).
- Issa, E., Papanastassiou, A. & Di Carlo, J. J. Large-scale, high-resolution neurophysiological maps underlying fMRI of macaque temporal lobe. *J. Neurosci.* **33**, 15207–15219 (2013).
- Ramon, M., Vizioli, L., Liu-Shuang, J. & Rossion, B. Neural microgenesis of personal familiar face recognition. *Proc. Natl Acad. Sci. USA* **112**, E4835–E4844 (2015).
- Henson, R., Shallice, T. & Dolan, R. J. Neuroimaging evidence for dissociable forms of repetition priming. *Science* **287**, 1269–1272 (2000).
- Gobbini, M. I., Leibenluft, E., Santiago, N. & Haxby, J. V. Social and emotional attachment in the neural representation of faces. *Neuroimage* **22**, 1628–1635 (2004).
- Ewbank, M. P. & Andrews, T. J. Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *Neuroimage* **40**, 1857–1870 (2008).
- Weibert, K. & Andrews, T. J. Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. *Neuropsychologia* **75**, 588–596 (2015).
- Gorno-Tempini, M. & Price, C. Identification of famous faces and buildings: a functional neuroimaging study of semantically unique items. *Brain* **124**, 2087–2097 (2001).
- Gorno-Tempini, M. et al. The neural systems sustaining face and proper-name processing. *Brain* **121**, 2103–2118 (1998).



41. Dubois, S. et al. Effect of familiarity on the processing of human faces. *Neuroimage* **9**, 278–289 (1999).
42. Shah, N. et al. The neural correlates of person familiarity. A functional magnetic resonance imaging study with clinical implications. *Brain* **124**, 804–815 (2001).
43. Leveroni, C. et al. Neural systems underlying the recognition of familiar and newly learned faces. *J. Neurosci.* **20**, 878–886 (2000).
44. Davies-Thompson, J., Newling, K. & Andrews, T. J. Image-invariant responses in face-selective regions do not explain the perceptual advantage for familiar face recognition. *Cereb. Cortex* **23**, 370–377 (2013).
45. Weibert, K. et al. An image-invariant neural response to familiar faces in the human medial temporal lobe. *Cortex* **84**, 34–42 (2016).
46. Young, A. & Burton, M. Are we face experts? *Trends Cogn. Sci.* **22**, 100–110 (2018).
47. Olson, I., McCoy, D., Klobusicky, E. & Ross, L. Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc. Cogn. Affect Neurosci.* **8**, 123–133 (2013).
48. Kriegeskorte, N., Formisano, E., Sorger, B. & Goebel, R. Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc. Natl Acad. Sci. USA* **104**, 20600–20605 (2007).
49. Nestor, A., Plaut, D. & Behrmann, M. Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proc. Natl Acad. Sci. USA* **108**, 9998–10003 (2011).
50. Gauthier, I. et al. The fusiform “face area” is part of a network that processes faces at the individual level. *J. Cogn. Neurosci.* **12**, 495–504 (2000).
51. Jacques, C. et al. The neural basis of rapid unfamiliar face individuation with human intracerebral recordings. *Neuroimage* **221**, 117174 (2020).
52. Dobs, K., Isik, L., Pantazis, D. & Kanwisher, N. How face perception unfolds over time. *Nat. Commun.* **10**, 1258 (2019).
53. Lavenex, P. & Amaral, D. G. Hippocampal-Neocortical interaction: a hierarchy of associativity. *Hippocampus* **10**, 420–430 (2000).
54. Quian Quiroga, R. Plugging in to human memory: advantages, challenges and insights from human single neuron recordings. *Cell* **179**, 1015–1032 (2019).
55. Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005).
56. Quian Quiroga, R. No pattern separation in the human hippocampus. *Trends Cogn. Sci.* **24**, 994–1007 (2020).
57. Mormann, F. et al. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J. Neurosci.* **28**, 8865–8872 (2008).
58. Quian Quiroga, R. Concept cells: the building blocks of declarative memory functions. *Nat. Rev. Neurosci.* **13**, 587–597 (2012).
59. Gao, X., Gentile, F. & Rössion, B. Fast periodic stimulation (FPS): a highly effective approach in fMRI brain mapping. *Brain Struct. Funct.* **223**, 2433–2454 (2018).
60. Rössion, B. Understanding face perception by means of human electrophysiology. *Trends Cogn. Sci.* **18**, 310–318 (2014).
61. Chaure, F., Rey, H. & Quian Quiroga, R. A novel and fully automatic spike-sorting implementation with variable number of features. *J. Neurophysiol.* **120**, 1859–1871 (2018).
62. Quian Quiroga, R., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* **16**, 1661–1687 (2004).
63. Quian Quiroga, R., Kraskov, A., Koch, C. & Fried, I. Explicit encoding of multimodal percepts by single neurons in the human brain. *Curr. Biol.* **19**, 1308–1313 (2009).
64. Ison, M. et al. Selectivity of pyramidal cells and interneurons in the Human Medial Temporal Lobe. *J. Neurophysiol.* **106**, 1713–1721 (2011).
65. Quian Quiroga, R. & Panzeri, S. Extracting information from neural populations: Information theory and decoding approaches. *Nat. Rev. Neurosci.* **10**, 173–185 (2009).
66. Hagen, S. et al. Spatially dissociated intracerebral maps for face- and house-selective activity in the human ventral occipito-temporal cortex. *Cereb. Cortex* **30**, 4026–4043 (2020).

## Acknowledgements

The authors thank the patients for their collaboration and funding from the Biotechnology and Biological Sciences Research Council (BBSRC), UK (BB/T001291/1), ERC Adg HUMANFACE 101055175 and Région Grand est and Rotary-FRC-Espoir en tête.

## Author contributions

R.Q.Q. and B.R. designed the experiments and did the project supervision; R.Q.Q. and M.B. analyzed the electrophysiology single-cell data; R.Q.Q., M.B., and H.G.R. wrote software for presentation of the stimuli and analysis of the electrophysiology single-cell data; J.J., X.Y., and B.R. designed and analyzed the LFP and fMRI face localizer experiments; R.Q.Q., M.B., H.R., J.J., and L.K. did the electrophysiology single-cell data collection; R.Q.Q., B.R., and L.M. obtained funding to support the experiments; S.C.-C. performed the surgeries; S.C.-C., L.M., J.J., and B.R. planned the electrode implantations; R.Q.Q. and B.R. wrote the paper; R.Q.Q. and M.B. prepared the figures; all coauthors contributed to review and editing of the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-41323-5>.

**Correspondence** and requests for materials should be addressed to Rodrigo Quian Quiroga or Bruno Rössion.

**Peer review information** *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023