communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-09241-y

Mapping functional homologies between human and marmoset brain networks using movie-driven ultra-high field fMRI

Check for updates

Alessandro Zanini [1] ✉, Audrey Dureux[1], Ravi S. Menon [1,2] & Stefan Everling [1,3]

Naturalistic stimuli, such as movies, offer a powerful tool for probing functional brain organization across species. Using movie-driven functional magnetic resonance imaging (md-fMRI), we recorded brain activity in humans and awake marmosets exposed to the same dynamic audiovisual stimulus. We applied tensor independent component analysis (tICA) to identify functional networks in each species, hierarchically cluster them, and examine their within- and between-species temporal correlations to assess functional homologies. We found strong interspecies correspondence in core sensory networks, particularly those involved in visual and auditory processing, suggesting conserved mechanisms for sensory integration. In contrast, networks associated with higher-order cognition, including prefrontal and temporoparietal areas, were observed primarily in humans, highlighting species-specific specializations. These findings demonstrate the value of naturalistic paradigms and data-driven approaches in revealing both shared and divergent brain architectures. By openly sharing our data and pipelines, we aim to advance the marmoset as a model for investigating the evolutionary foundations of brain function.

The common marmoset (*Callithrix jacchus*) is gaining traction as a primate model in neuroscience due to its small size, rapid maturation, and rich cognitive and social behaviors[1–3]. The New World marmoset shares key features with Old World monkeys and humans, including a granular prefrontal cortex[4], a complex visual system[5], and comparable visuomotor behaviors[6,7]. Recent studies have also uncovered an extensive network dedicated to processing conspecific vocalizations, consistent with the species' reliance on social communication[8–10].

As interest in the marmoset grows, understanding the functional organization of its brain—and how it compares to the human brain—becomes increasingly important. Functional magnetic resonance imaging (fMRI) provides a powerful tool for this purpose, enabling whole-brain recordings during task-based or resting-state paradigms. Prior studies have successfully used both approaches to compare human and marmoset brain function, revealing homologous resting-state networks[11–14] and activations in response to shared stimuli[15,16].

However, both fMRI approaches have limitations. Resting-state fMRI can reveal functional connectivity across brain regions[17,18] but lacks the specificity to identify networks driven by behaviorally relevant stimuli. Moreover, although resting-state activity is spontaneous, it is also highly dependent on the subject's internal state—which is often unknown or uncontrolled—making interspecies comparisons of functional networks difficult. Task-based fMRI, by contrast, allows for direct comparison of stimulus-driven responses and has been used extensively in humans[16,19], macaques[20,21], and marmosets[9,22–26]. Yet this method is limited by the impracticality of designing and repeating numerous functional localizers across the entire brain, especially in non-human primates.

Movie-driven fMRI (md-fMRI) has emerged as a promising alternative. Naturalistic movie stimuli offer an ecologically valid alternative to conventional task paradigms by providing continuous, multisensory input that more closely resembles real-life experience. Prior work has shown that movies elicit robust and reliable neural responses across visual, auditory, and higher-order cognitive systems both in humans[27–33] and non-human primates[14,34,35], making them particularly suitable for comparative and translational research. Importantly, movie stimuli allow for the simultaneous engagement of multiple brain systems while preserving experimental control and repeatability. Crucially, such stimuli also maintain the engagement of non-human primates during long fMRI sessions. Finally, recent literature suggests that the brain may be more strongly "tuned" to naturalistic than to artificial stimuli, such that movies and other real-world narratives can evoke more reliable and representative patterns of neural activity across sensory and cognitive systems (for review, see ref. 36).

[1]Centre for Functional and Metabolic Mapping, Robarts Research Institute, University of Western Ontario, London, ON, Canada. [2]Department of Medical Biophysics, University of Western Ontario, London, ON, Canada. [3]Department of Physiology and Pharmacology, University of Western Ontario, London, ON, Canada. ✉e-mail: azanini4@uwo.ca

In this study, we presented a naturalistic movie featuring diverse visual and auditory stimuli to eight marmosets and nineteen humans. Using tensor independent component analysis (tICA), a data-driven approach well suited to naturalistic paradigms, we identified 20 functional components in each species. We then examined their temporal dynamics within and across species to identify potential functional homologs. Our findings reveal both conserved and divergent patterns of brain activity, shedding light on the evolutionary organization of sensory and cognitive networks in primates.

Results

Marmoset tICA components

Of the 20 components produced by tICA decomposition of the marmoset data, 12 were classified as functional (non-noise) based on spatial distribution and temporal power spectra (Fig. 1). These components encompassed a broad range of sensory, motor, and associative brain areas, and were grouped into networks based on anatomical distribution and functional relevance.

Three components (MA, MC, and MD) were categorized as *higher-order visual networks (HOV)* given their anatomical distribution and alignment with prior resting-state studies[13,37,38]. These exhibited strong bilateral activation in occipitotemporal visual areas as well as prefrontal cortex regions, including areas 8Av, 8C, 45, and 47. Subcortical activation in these networks included the pulvinar and superior colliculus, indicating integration of visual and attentional processing. Two additional components (MB and MF) were identified as *subcortical-visual networks (SUV)*. These shared overlap with the HOV components in both visual and prefrontal cortex but extended further into subcortical regions, including the caudate, putamen, and thalamus, and recruited parietal areas such as LIP, AIP, PE, PG, and PGM.

A single component (ME) was classified as a *somatomotor network (SOM)*, with strong activation in primary somatosensory cortex (areas 1/2, 3a, and 3b), motor cortex (area 4ab), and ventral premotor cortex (areas 6Va and 6Vb). Its spatial pattern closely matches the somatomotor ventral component previously described in the literature[12,13,37,38], likely reflecting widespread engagement of sensorimotor regions during the observation of behaviorally relevant stimuli in the movie. Component MG displayed a pattern consistent with a *default mode network (DMN)*[11,13,37–40], with activation in posterior cingulate, medial parietal areas (MIP, LIP, OPt), and dorsomedial prefrontal cortex (areas 6DC, 6DR, 8aD, and 8C), along with subcortical regions such as the hippocampus and caudate.

Component MH extended across medial occipital cortex, orbitofrontal cortex, and the putamen and was labeled an *occipitoparietal network (OP)*, likely involved in integrating visual and contextual information. Component MI was more spatially restricted to early visual areas, including V1 and MT, and was identified as a *primary visual network (PVIS)*[13,37,38]. In contrast to these cortical-dominant components, MJ exclusively encompassed subcortical structures—putamen, caudate, globus pallidus, and ventral thalamus—without cortical involvement, defining a *basal ganglia network (BG)*[12,13,37]. Component MK, the *motor-cerebellar network (MCE)*, revealed co-activation of cerebellar regions, motor and premotor cortex (4ab, 6DC, 6DR, and 6 M), and the periaqueductal gray, possibly reflecting coordinated sensorimotor integration. Lastly, component ML was localized to the core, belt, and parabelt auditory cortex and was labeled the *auditory network (AUD)*, consistent with prior studies using resting-state data in marmosets[12,13,37,38].

Figure. 2A illustrates that frequently recruited voxels were most prominent in prefrontal, auditory, and visual cortices, as well as in subcortical structures like the caudate and superior colliculus. These patterns were not simply a byproduct of data quality, as comparison with the temporal signal-to-noise ratio (tSNR) map (Fig. 2B) confirmed that component distributions did not merely reflect regions of high SNR. Quantitative comparison of mean tSNR between cortical and subcortical regions revealed significantly higher signal quality in the cortex (paired *t*-test: $t = 3.01$, df = 39, $p = 0.0046$)

Marmoset component relationships

To assess the interrelationships among the functional networks, we computed temporal correlations between component timecourses and performed hierarchical clustering (Fig. 3A). The resulting dendrogram revealed four distinct network groupings, which were mirrored in the within-species correlation matrix (Fig. 3B).

The first cluster consisted of *multisensory networks*, grouping together the HOV, SUV, and AUD components. HOV and SUV components were strongly inter-correlated ($r = 0.83$–0.96), suggesting coordinated engagement of visual systems during naturalistic movie viewing. The auditory component (AUD) showed moderate correlations with both HOV and SUV components ($r = 0.36$–0.50).

A second, largely independent cluster contained the SOM network, which showed weak correlation with the motor-cerebellar component (MCE; $r = 0.14$) and moderate to strong anticorrelations with all other functional components ($r = -0.32$ to -0.96). A third group—comprising the visual-default cluster—included the PVIS, DMN, and OP components, which were internally well correlated ($r = 0.70$–0.72), showed weak to moderate correlations with subcortical-visual components, and anticorrelated with the auditory network. Their correlations with the higher-order visual networks were more heterogeneous: PVIS and DMN showed generally weak to moderate correlations, while OP exhibited anticorrelations. Finally, a subcortical-motor cluster, composed of the BG and MCE networks, showed moderate correlation with each other ($r = 0.48$) and weak to moderate associations with the visual-default cluster. Interestingly, their correlation patterns with other components were diametrically opposed: the BG network showed fair-to-moderate positive correlations with the multisensory cluster and negative correlation with SOM, while the MCE network displayed the reverse pattern.

Human tICA

In the human dataset tICA decomposition, 14 of the 20 components were categorized as functionally relevant based on their anatomical distribution and frequency spectra (Fig. 4). These networks spanned primary sensory, association, and higher cognitive domains, enabling a direct comparison with the networks identified in marmosets.

Three components—HA, HB, and HE—were grouped as *HOV*. These included bilateral activation in occipital and parietal cortices, consistent with dorsal and ventral visual streams. Notably, component HB engaged the frontal eye fields (FEF), middle temporal areas and lateral intraparietal cortex, resembling an attentional control network or the dorsal attention network.

Two components, HC and HG, were classified as *audiovisual networks (AV)*. These showed strong activation in the superior temporal sulcus (STS) and auditory cortex, and in the case of HC, also included prefrontal regions such as the inferior frontal gyrus (IFG), area 8Av, and PEF. Component HG extended into the retrosplenial complex (RSC), suggesting involvement in integrating auditory, visual, and contextual information.

Component HF was identified as a *pure auditory network (AUD)*, with strong activation in the auditory cortex, STS, and insula. Like HG, it overlapped with the RSC, reflecting potential convergence of sensory and default-related processes.

Two *executive networks (EXE)* were detected: HD, a bilateral network spanning dorsolateral prefrontal cortex, anterior cingulate, and parietal regions, also engaging the basal ganglia; and HM, a left-lateralized network including the posterior temporal region, large portions of the medial and lateral prefrontal cortices, and parietal cortex, strongly resembling the frontoparietal executive network often reported in literature[41–44].

The *insular network (INS)*, component HH, encompassed the anterior insula, anterior cingulate cortex, and PF complex—regions consistent with the salience network described in humans[45].

Somatomotor (SOM) functions were represented by component HI, which showed strong bilateral activation in the precentral and postcentral gyri and aligned with typical somatomotor maps[41–44].

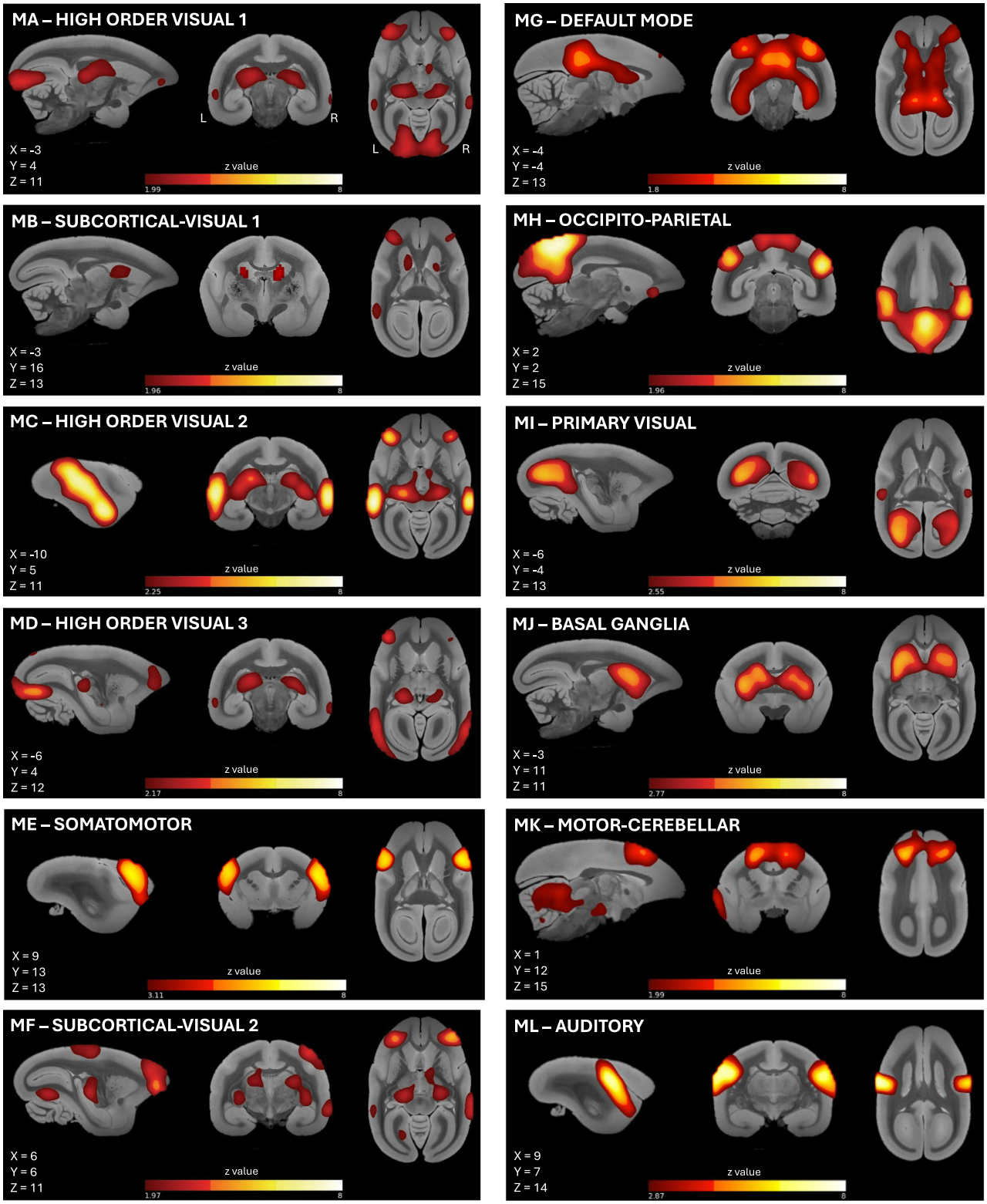


Fig. 1 | Spatial maps of the 12 functional marmoset tICA components. Components are represented on sagittal, coronal, and transversal slices of a high-resolution template of the marmoset brain[85]. Brain coordinates of the presented perspective are reported in the left bottom corner of each panel. Components include higher-order visual (HOV), subcortical-visual (SUV), somatomotor (SOM), default mode (DMN), occipitoparietal (OP), primary visual (PVIS), basal ganglia (BG), motor-cerebellar (MCE), and auditory (AUD) networks. Components were selected based on spatiotemporal features and visual inspection. Color indicates voxel-wise z-scores, thresholded at $p < 0.5$ (posterior probability that a voxel belongs to the active distribution rather than noise).

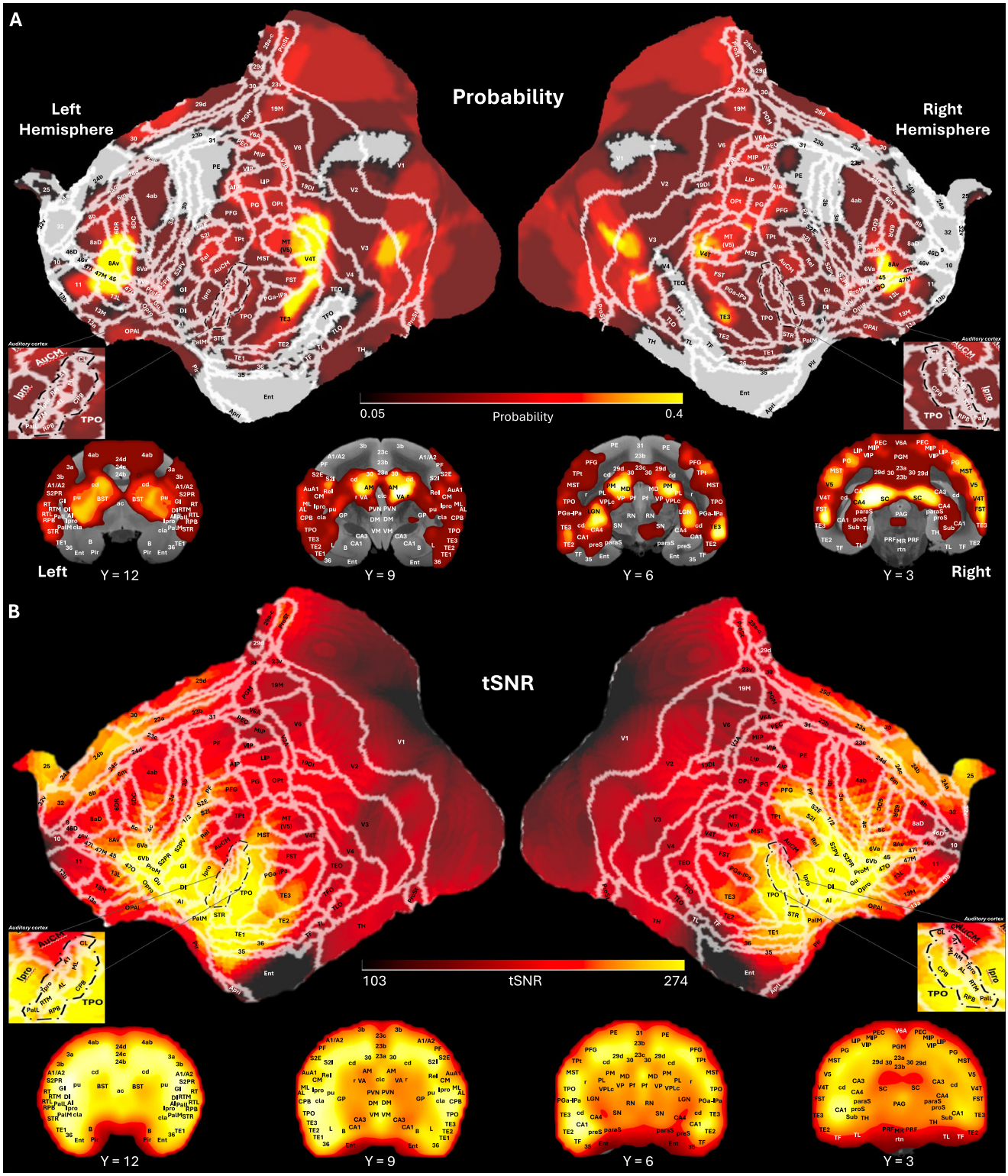


Fig. 2 | Spatial distribution and tSNR of marmoset components. A Probability map showing the number of functional components in which each voxel was active (thresholded as indicated in Fig. 1). Voxels most frequently recruited across components are concentrated in prefrontal, auditory, visual, and subcortical regions. **B** Mean tSNR map across marmosets showing the average quality of the BOLD signal across the brain, with higher tSNR values represented in yellow and lower values in black. The similarity between frequently recruited regions and high-tSNR areas confirms that component patterns were not driven solely by signal quality. In both panels, the maps are shown on flat surface representations of the marmoset brain, with white lines delineating the Paxinos parcellation[93] of the NIH marmoset brain atlas[85], and on coronal slices of a marmoset brain anatomical image at different inter-aural levels.

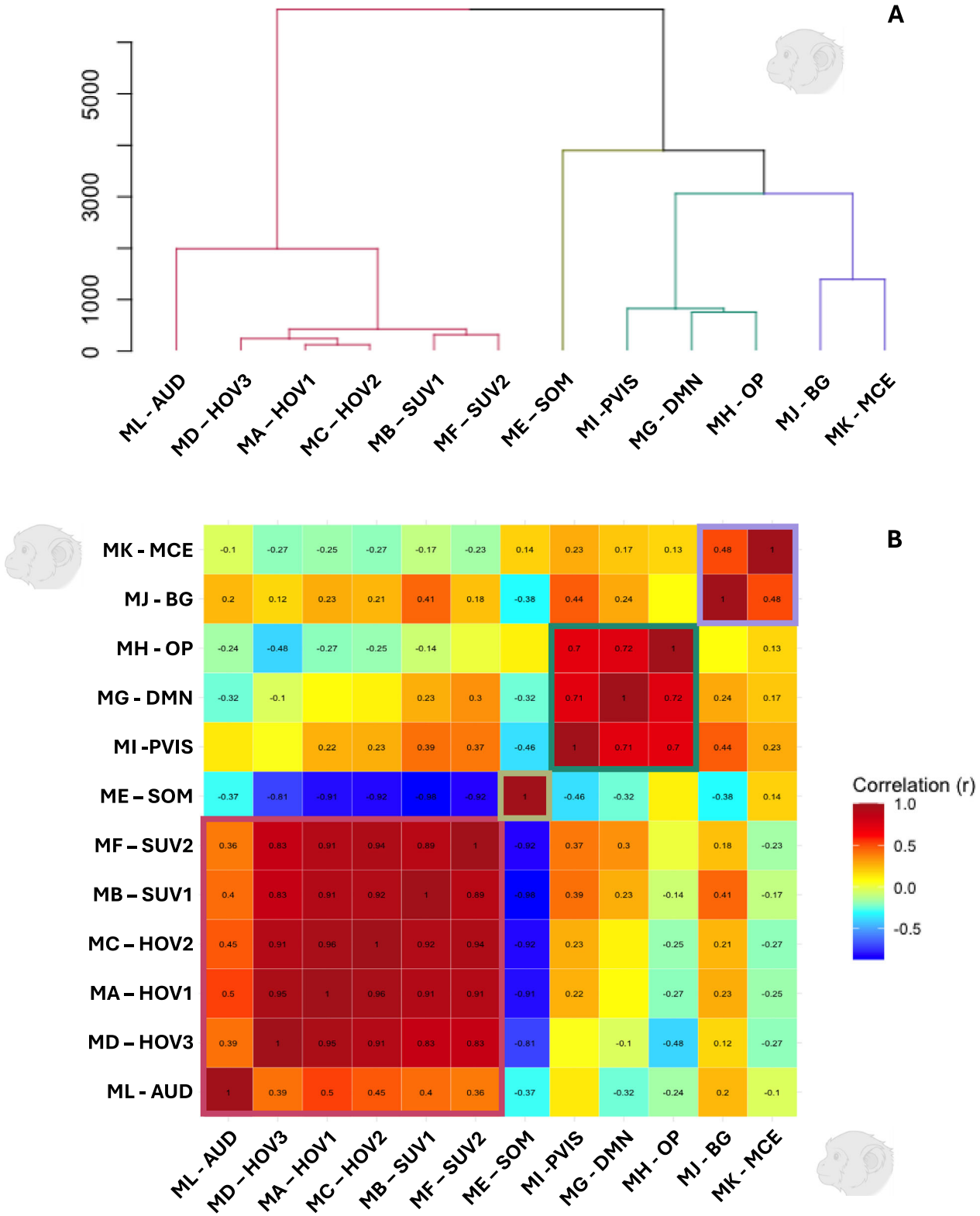


Fig. 3 | Marmoset component clustering and temporal relationships. **A** Hierarchical clustering dendrogram of the 12 functional marmoset components using squared Euclidean distance and Ward's linkage. Colors indicate functional clusters: multisensory (red), somatomotor (green), visual-default (turquoise), and subcortical-motor (purple). **B** Within-species correlation matrix showing pairwise temporal correlations (Pearson's r) between component timecourses. Components are ordered based on the sequence specified by the hierarchical clustering tree. Networks within the same functional group show strong positive correlations, while cross-cluster relationships are more variable, including anticorrelations between somatomotor and multisensory networks. Clusters identified through hierarchical clustering are highlighted on the matrices with colored squares, following the same color scheme presented in panel A. Red shades indicate positive correlations, while blue shades represent negative correlations. Only correlations with an absolute *r* value greater than 0.10 are displayed.

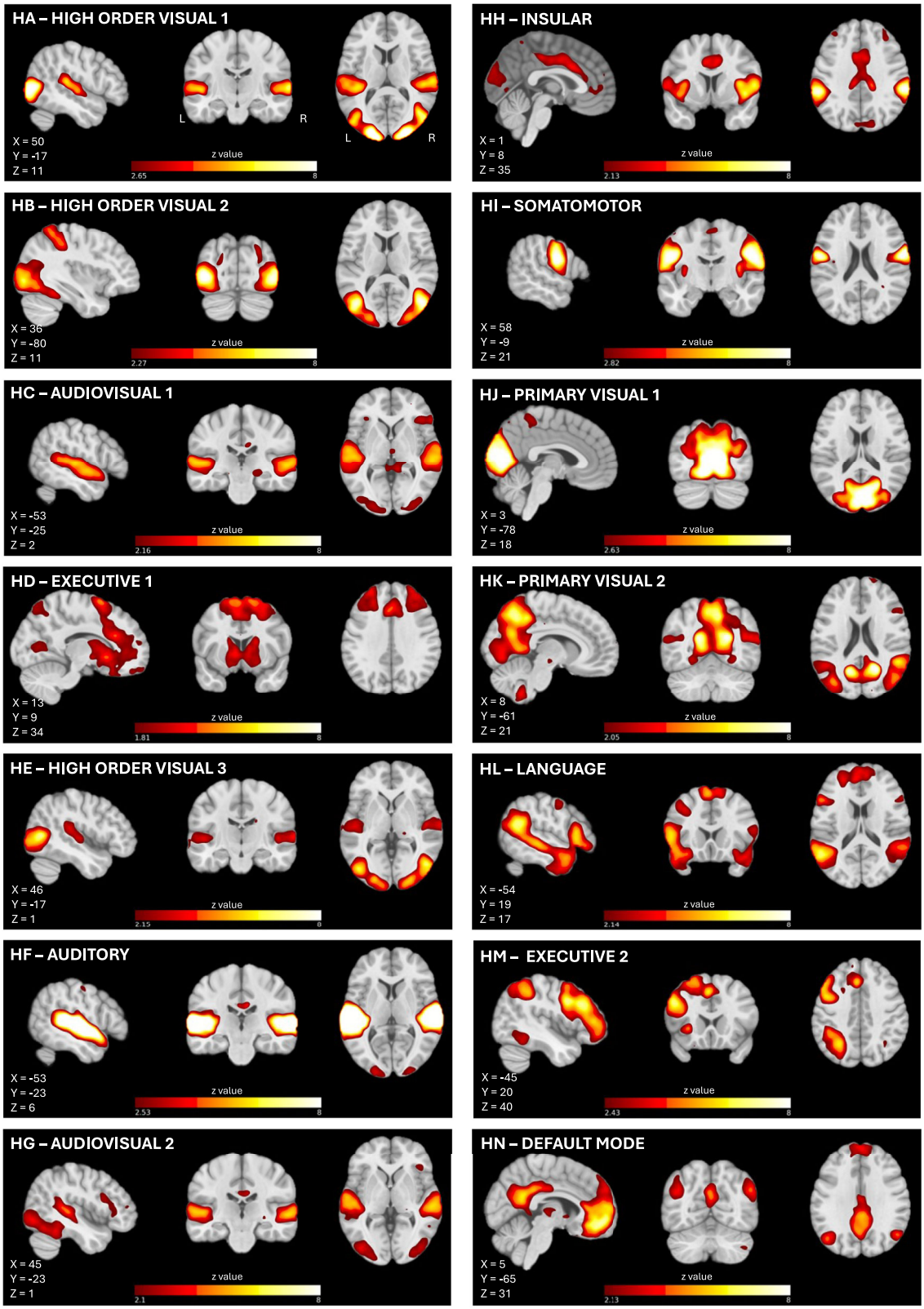


Fig. 4 | Spatial maps of the 14 functional human tICA components. Z-scored spatial maps of each human component are overlaid on the MNI152 template. MNI coordinates of the presented perspective are reported in the left bottom corner of each panel. Networks include higher-order visual (HOV), audiovisual (AV), auditory (AUD), executive (EXE), insular (INS), somatomotor (SOM), primary visual (PVIS), language (LAN), and default mode (DMN). Color indicates voxel-wise z-scores, thresholded at $p < 0.5$ (posterior probability that a voxel belongs to the active distribution rather than noise). Components were labeled based on spatial overlap with known functional networks and their correspondence to task content.

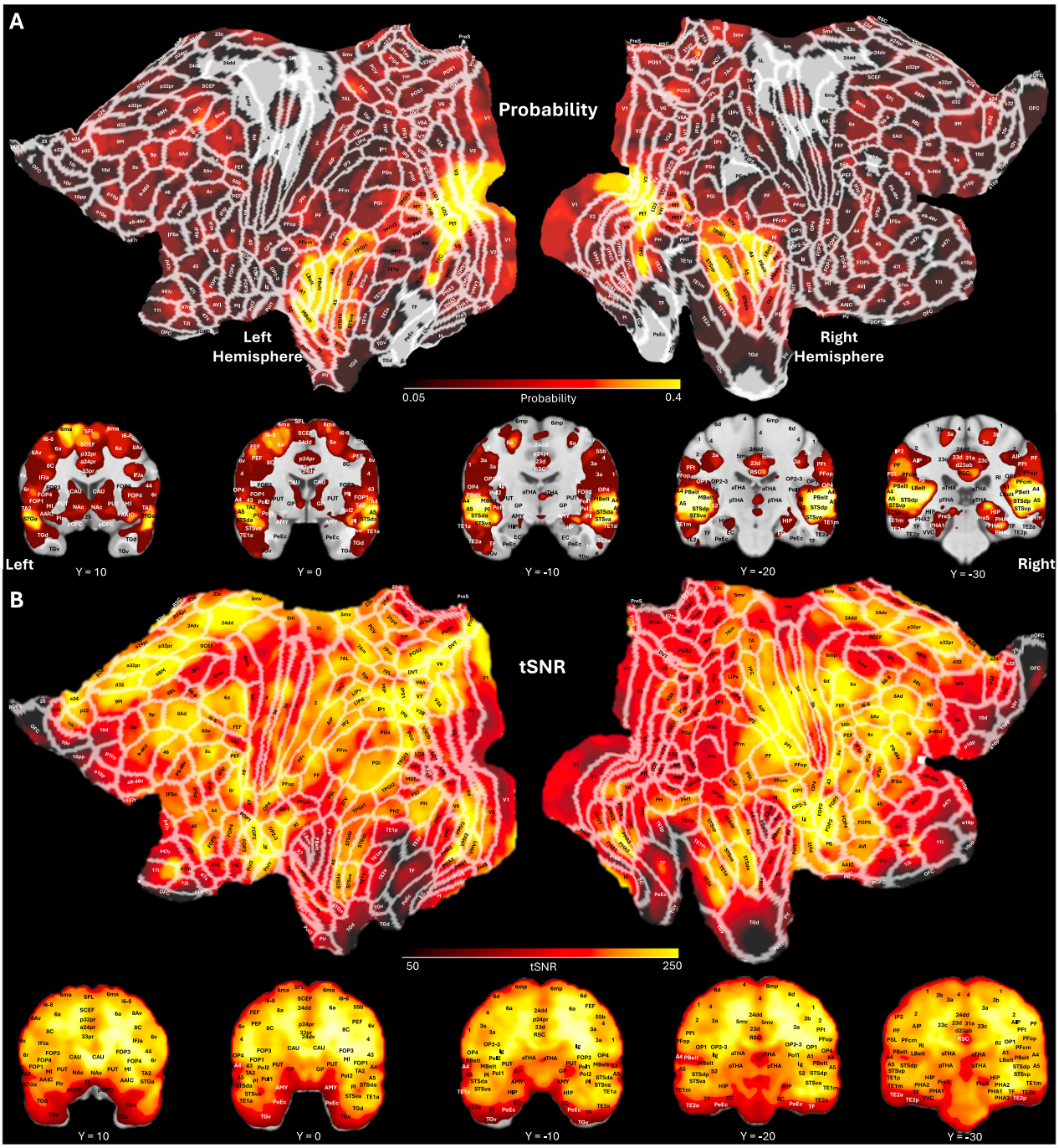


Fig. 5 | Spatial distribution and tSNR of human components. A Probability map showing the number of components in which each voxel was active across the 14 functional networks (thresholded as indicated in Fig. 4). High-frequency recruitment was observed in auditory and visual areas, with sparse representation in somatosensory and ventral temporal cortex. **B** Mean tSNR map across all human participants. Visual inspection confirms that functional component distributions were not solely determined by tSNR patterns. Both maps are projected onto inflated and flat surface representations of the human brain, as well as onto coronal slices from an anatomical template in MNI space. White outlines indicate cortical areas based on the Human Connectome Project's multi-modal parcellation atlas[94].

Early visual processing was captured by two *PVIS networks*—components HJ and HK—covering V1, V2, and adjacent areas. Notably, component HK has sometimes been considered part of the human default mode network due to its anatomical distribution, which includes the boundary between the medial occipital cortex and the posterior cingulate cortex[41,42].

Component HL, designated the *language network (LAN)*, activated frontal areas (left IFG, 44, 45, 47), superior frontal language area (SFL), posterior STS, and temporal pole areas (TPOJ, TG), with an additional cluster in the posterior cingulate[46,47].

Finally, component HN was identified as the *DMN*, encompassing the medial prefrontal cortex, PCC, angular gyrus, and hippocampus, consistent with canonical DMN architecture described in resting-state and naturalistic fMRI studies[41,42,48,49].

Voxel-wise probability maps (Fig. 5A) indicated the most consistent recruitment across participants in auditory and visual cortices. By contrast, somatosensory and ventral temporal areas showed relatively sparse recruitment. These spatial patterns were not explained by signal quality, as confirmed by comparison with the tSNR map (Fig. 5B), supporting the

functional relevance of the extracted components. Similarly, comparison of mean tSNR between cortical and subcortical regions did not reveal a significant difference (paired t-test: $t_{(18)} = -1.45$, $p = 0.164$), suggesting that the lack of a basal ganglia network in the human sample is unlikely to be explained only by signal quality differences.

Human component relationships

Hierarchical clustering of the 14 human functional components (Fig. 6A) revealed three major groupings, reflected in the corresponding within-species correlation matrix (Fig. 6B).

The first branch of the dendrogram encompassed two functionally distinct but moderately linked clusters. The first comprised the two PVIS networks (PVIS1 and PVIS2), which were strongly correlated with each other ($r = 0.64$), reflecting their shared involvement in early visual processing. The second cluster was a *sensorimotor-executive-insular* group, comprising the SOM, EXE1, and INS networks, which also showed strong mutual correlations ($r = 0.58$–0.75), suggesting co-activation related to action observation, executive demands, or salience-driven processes. These two clusters were grouped together based on moderate cross-cluster correlations, most notably between SOM and PVIS1 ($r = 0.40$), though SOM was weakly anticorrelated with PVIS2 ($r = -0.15$), indicating some degree of functional separation.

The second branch was a *multisensory network group*, containing the HOV, AV, and AUD components. These networks were moderately or highly intercorrelated ($r = 0.24$–0.93), reflecting the integrated audiovisual nature of the stimulus and the simultaneous activation of visual and auditory pathways.

Three components—LAN, EXE2, and DMN—were more weakly correlated with other networks, forming functionally distinct entities. The LAN component showed minimal temporal correlation with either sensory or attentional networks. EXE2 showed weak correlation with visual components and anticorrelation with EXE1, indicating a potentially distinct functional role. The DMN component exhibited little to no correlation with other networks.

Cross-species correlations

To assess the degree of temporal alignment between functional brain networks across species, we computed the cross-species correlation matrix based on the timecourses of the 12 marmoset and 14 human functional components (Fig. 7). This analysis revealed two major network groupings that differed in the strength and structure of their interspecies correspondence.

The first was a *conserved multisensory cluster*, comprising audiovisual and visual networks in both humans and marmosets. This group included the HOV, SUV, AUD, and AV components. Networks within this cluster demonstrated modest to strong interspecies correlations, with values ranging from r = 0.27 to 0.60. Notably, the marmoset HOV and SUV components showed robust temporal alignment with human HOV and AV networks. The marmoset AUD was most strongly correlated with the human AUD and AV networks, with cross-species correlations reaching $r = 0.56$–0.59.

The second major grouping represented a *divergent cognitive cluster*, encompassing components associated with somatomotor, executive, default mode, and subcortical-motor functions. This cluster showed only weak to modest cross-species correlations ($r = 0.12$–0.38) and differed in composition between species. In humans, it included the SOM, INS, and EXE1 components, along with the PVIS network. In marmosets, the corresponding cluster included PVIS, SOM, OP, DMN, BG, and MCE components.

Importantly, several human components—particularly those associated with higher cognition—exhibited minimal or no correlation with any marmoset network. Notably, the *LAN network*, characterized by strong left-lateralization and robust activation of frontal and temporal language areas, showed weak correlations across all marmoset components.

Similarly, the *left-lateralized EXE2 network* showed minimal alignment with any marmoset component and was weakly anticorrelated with the marmoset DMN and subcortical-motor clusters.

The *DMN* in humans also lacked a clear homolog in the marmoset data. Although a marmoset DMN component was identified (component MG), it did not correlate strongly with the human DMN and instead aligned more closely with marmoset visual and parietal networks and with the human primary visual, EXE1 and INS networks.

Discussion

In this study, we used md-fMRI to compare brain activity in humans and common marmosets as they watched the same naturalistic audiovisual stimulus. This paradigm provided a rich, ecologically valid context to assess cross-species functional network organization, expanding upon our earlier work identifying interspecies homologies in face-selective visual areas[14]. By introducing diverse auditory and visual content—including conspecific and heterospecific vocalizations, human speech, and dynamic scenes—we engaged a broader range of sensory and associative regions, allowing for a more comprehensive examination of shared and divergent brain networks.

To analyze these responses, we applied tICA, a fully data-driven method that decomposes brain activity into temporally and spatially independent components without reliance on predefined network templates. This approach is well-suited to complex, continuous stimuli like movies, enabling the extraction of distinct functional networks that reflect moment-to-moment neural dynamics. Our analysis extracted 20 components per species, of which 12 in marmosets and 14 in humans were identified as functionally relevant (i.e., non-noise). We then compared temporal dynamics of these networks within and across species.

We observed moderate to strong cross-species correlations among core sensory systems, particularly those supporting audiovisual processing. HOV, SUV, and AUD networks formed tightly correlated clusters in both species. These findings suggest a shared temporal structure in how primate brains process dynamic audiovisual input, despite differences in cortical architecture and specialization[32,50]. The high degree of within- and between-species temporal correlation among these networks reinforces the idea that sensory integration, especially in naturalistic contexts, is supported by evolutionarily conserved mechanisms.

In contrast, components linked to higher cognitive functions—including language, executive control, and default mode processes—were uniquely observed in humans and did not correlate with any marmoset networks. The LAN network, for instance, exhibited strong left-lateralized recruitment of frontal and temporal regions traditionally associated with speech comprehension and semantic processing. Its absence in marmosets aligns with the lack of linguistic capacity in non-human primates and underscores the human specificity of this network[51]. Similarly, a left-lateralized executive component resembling the frontoparietal executive network[41–44] emerged only in humans, suggesting a functional specialization not mirrored in the marmoset brain.

The internal organization of networks also differed markedly between species. Human components exhibited greater segregation, with distinct correlation clusters reflecting sensory, cognitive, and default mode systems. In contrast, correlations among marmoset components were generally higher, reflecting a tendency toward more broadly shared temporal fluctuations across networks. This pattern suggests a more distributed and overlapping functional organization in marmosets, where networks are less differentiated and more broadly engaged by naturalistic stimuli. Such differences may reflect evolutionary divergence in brain architecture and processing capacity, with the human brain exhibiting a higher degree of functional specialization.

Importantly, several findings confirmed and extended prior resting-state fMRI observations. In marmosets, components resembling default mode, somatomotor, visual, auditory, and basal ganglia networks closely aligned with those identified in earlier studies[12,13,37,38], lending support to the reproducibility and reliability of tICA under naturalistic stimulation.

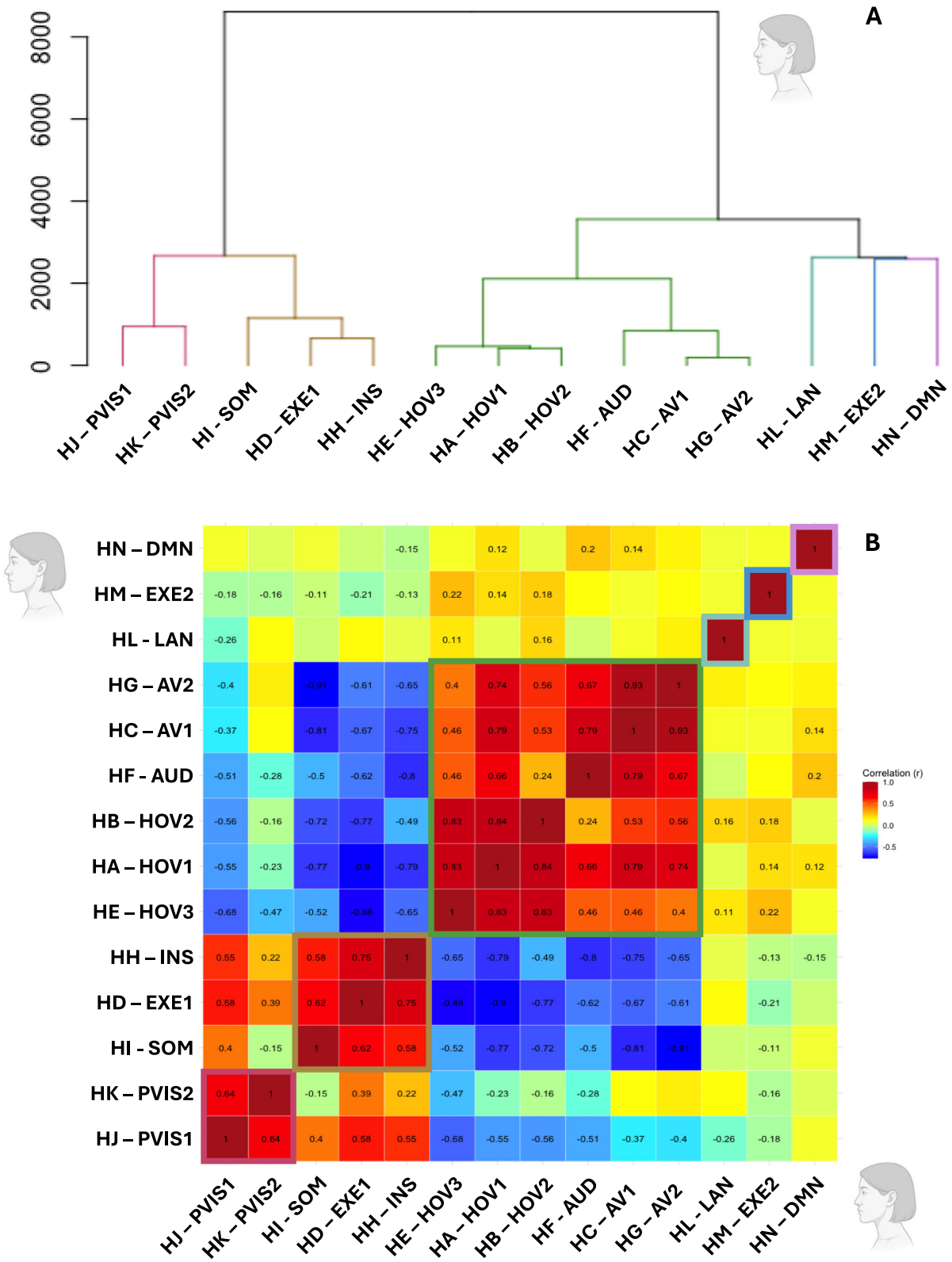


Fig. 6 | Human component clustering and temporal relationships. A Hierarchical clustering dendrogram of the 14 functional human components using squared Euclidean distance and Ward's linkage. Clusters include primary visual (red), sensorimotor-executive-insular (gold), multisensory (green), and functionally distinct components (light blue, dark blue, purple). **B** Within-species correlation matrix showing pairwise temporal correlations (Pearson's r) between component time-courses. Components are ordered based on the sequence specified by the hierarchical clustering tree. Sensorimotor and audiovisual networks show high internal correlation, while the language, executive-2, and DMN components are temporally distinct from other clusters. Clusters identified through hierarchical clustering are highlighted on the matrices with colored squares, following the same color scheme presented in panel A. Red shades indicate positive correlations, while blue shades represent negative correlations. Only correlations with an absolute *r* value greater than 0.10 are displayed.

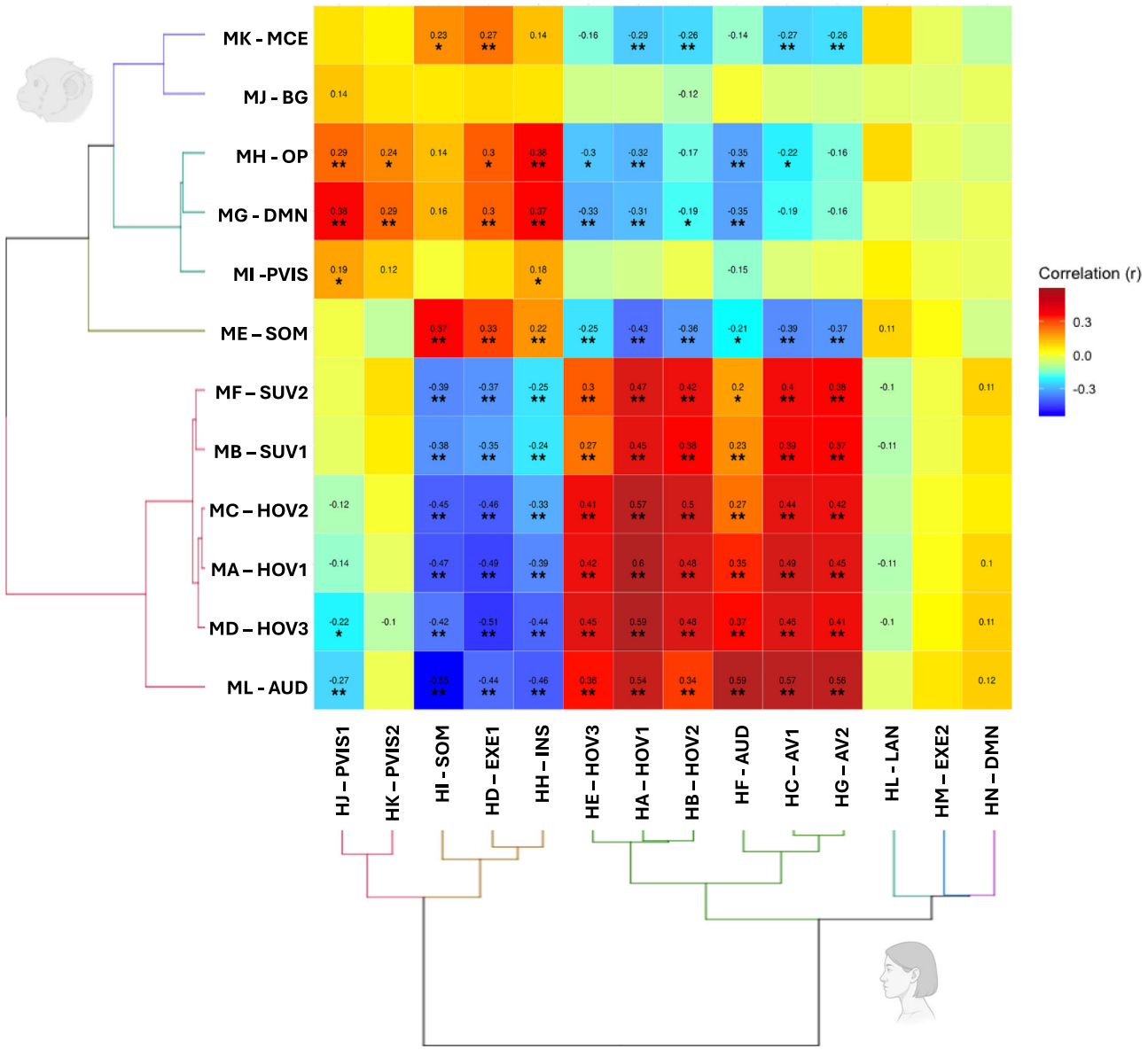


Fig. 7 | Cross-species correlation matrix. Matrix showing temporal correlations (Pearson's r) between the 12 marmoset and 14 human functional components. Warm colors indicate positive correlation; cool colors indicate anticorrelation. Two major interspecies clusters emerge: a conserved multisensory cluster (upper left), including visual, auditory, and audiovisual networks, and a divergent cognitive cluster (lower right) comprising somatomotor, executive, and default mode components. Language and lateralized executive networks in humans show little or no correspondence in the marmoset. Only correlations with an absolute *r* value greater than 0.10 are displayed. Asterisks indicate correlations that reached statistical significance after FDR correction based on a permutation test ($p < 0.05$, $p < 0.01^{**}$, $p < 0.001^{***}$).

Moreover, the networks extracted recruit all the major functional hubs previously described in the awake marmoset brain[12,52]. Although many of these hubs-identified in resting-state studies (e.g., caudate, putamen, thalamus, areas 8Av and TE3)-are also highly recruited during movie viewing, others appear less involved in the networks extracted by tICA. This discrepancy raises the possibility that certain hubs play a more prominent role in organizing intrinsic brain activity during rest, but are less dynamically engaged during complex, multisensory stimulation. In other words, these areas may act as central nodes within the brain's default communication architecture, but their functional influence may shift when the brain is engaged in externally driven, stimulus-bound processing. This distinction suggests that the role of functional hubs is not static, but rather dynamic and context-dependent, modulated by the cognitive and sensory demands of the environment.

In humans, canonical resting-state networks—including default mode, somatomotor, executive, and insular networks[12,41–44,48]—were robustly extracted, reinforcing their stability across task-free and naturalistic paradigms. However, the movie-driven context allowed for the emergence of additional components, particularly in sensory and multisensory domains, that are less frequently resolved in resting-state conditions. This highlights the utility of naturalistic stimuli for engaging diverse functional systems and enhancing component separability.

The dynamic nature of the movie facilitated the emergence of multisensory integration networks, especially in humans. Several audiovisual components centered on the STS, a region known for integrating visual and auditory input and implicated in speech and social perception[53]. In contrast, marmosets showed fewer distinct audiovisual components, and these often involved subcortical regions such as the superior colliculus and pulvinar—

structures known to contribute to multisensory processing in non-human primates[54,55]. The presence of multisensory networks in both species underscores the shared computational demands imposed by naturalistic stimuli[30,31,56]. Nevertheless, these differences suggests that cortical multisensory integration may be more anatomically and functionally elaborated in humans, possibly reflecting adaptations for processing complex communicative stimuli such as speech.

Despite the absence of direct somatosensory stimulation, both species exhibited well-defined somatomotor networks. These findings align with previous studies reporting strong intrinsic connectivity in these regions during rest[12,13,37,38,41,42,44] and support the notion that spontaneous fluctuations and action observation can activate sensorimotor systems. The content of the movie, which included scenes of movement, social interaction, and goal-directed behavior, may have engaged these networks via internal simulation mechanisms[57–59]. This is consistent with evidence for action observation networks in both humans[60–62] and marmosets[26], suggesting that cross-species similarities in somatomotor engagement may reflect not only structural homology but shared computational roles during passive observation. Otherwise, the co-activation of somatomotor, insular, and executive regions may also reflect engagement of broader integrative systems often referred to as the salience network[45,63] or, in other contexts, the so-called pain matrix[64]. These overlapping designations highlight both the importance and the difficulty of assigning unique functional labels to networks that flexibly participate in multiple processes, including salience detection, interoception, and cognitive control. In our study, the relations observed between somatomotor, executive, and insular components may therefore index a more general mechanism for orienting attention and behavior toward salient, socially and biologically relevant events in the movie, rather than reflecting a single "executive" function. This underscores the need for caution in applying rigid naming conventions and supports a more dynamic view of large-scale networks as multifunctional systems recruited according to contextual demands.

An intriguing aspect of our findings lies in the interpretation of the marmoset DMN. While anatomically similar to the human DMN, the marmoset component labeled as DMN clustered with visual and occipitoparietal networks and showed strong temporal coupling with sensory and subcortical components—patterns not observed in humans. The human DMN was largely functionally segregated, exhibiting weak correlations with sensory and executive systems. Furthermore, the marmoset DMN does not show temporal coupling with the human DMN, whereas it shows moderate correlations with the human executive and insular networks, which respectively resemble known cognitive control and salience networks in humans[41–43,48]. This discrepancy highlights a critical gap in our understanding of the marmoset DMN: unlike the human DMN, which is extensively characterized in the literature[41,42,44,48,49], resting-state studies in marmosets report inconsistent findings, with some identifying networks that resemble the human DMN anatomically and others reporting divergent patterns[11–13,37–40].

These observations raise the possibility that the so-called marmoset DMN may serve a different function, perhaps related to visual attention or environmental monitoring, rather than internally directed thought or mind-wandering as in humans[65]. The lack of interspecies correlation between DMN components, despite periods of reduced stimulation in the movie, underscores the need to reconsider the functional equivalence of these networks across species. However, it is important to note that during naturalistic stimulation, DMN subregions may not exhibit the strong anticorrelations with sensory and attentional networks typically reported in resting-state studies[66–69]. Instead, emerging evidence suggests that DMN activity in such contexts may be more dynamic and multifaceted, supporting processes such as narrative comprehension, social cognition, or integration of external and internal information[70,71]. This broader functional repertoire could account for the atypical pattern of within-species correlations observed in our human sample and may help explain why the human DMN showed limited correlations with the marmoset DMN: while the human network may flexibly adapt to complex cognitive demands, the marmoset counterpart might engage in more basic or distinct functions, leading to reduced interspecies correspondence during naturalistic stimulation. Given the rich, naturalistic design and cross-species comparability of our paradigm, this work may offer a valuable framework to further clarify the architecture and role of the DMN in marmosets and its potential evolutionary divergence from the human counterpart.

Subcortical networks also revealed interesting cross-species differences. In marmosets, we identified a distinct BG component encompassing the caudate, putamen, and thalamus—structures frequently observed as independent networks in marmoset resting-state studies[12,13,37,38] and major functional hubs of the marmoset brain[52]. This BG network showed strong correlations with subcortical-visual components and the primary visual network, suggesting integrated processing during movie viewing. In contrast, no basal ganglia component emerged from the human tICA decomposition. This discrepancy is unlikely to reflect differences in signal quality, as the mean tSNR in human subcortical structures was comparable to that of cortical regions, whereas the marmoset exhibited lower subcortical tSNR yet still yielded a clear basal ganglia network. Across all human components—both those retained for analysis and those identified as noise—no component showed distinctly higher weights in basal ganglia regions, suggesting that these areas are represented more diffusely within the human functional architecture. While some human studies do extract basal ganglia components[12,44,72,73], others do not[41–43], pointing to variability in detectability and functional coupling depending on species, analysis methods, and stimulus context. It is therefore plausible that the observed interspecies difference arises from other factors, such as differences in the intrinsic functional organization of subcortical networks or the specific parameters used for temporal ICA decomposition, including the number of components, initialization, and intersubject variability, all of which can affect component reliability.

A similar pattern was observed with cerebellar networks. The motor-cerebellar component in marmosets included motor cortex, cerebellum, and periaqueductal gray—regions seldom observed as a unified network in human resting-state studies. Its positive correlation with human somatomotor and cognitive control networks, and negative correlation with audiovisual systems, suggests a broader role in sensorimotor integration, potentially extending into cognitive domains. In humans, no cerebellar network was extracted, consistent with known difficulties in resolving posterior fossa activity in whole-brain ICA (53–56, but see ref. 74). However, targeted ICA approaches have demonstrated the cerebellum's rich functional architecture[75]. We note that although a few noise components included parts of the cerebellum, these were not reliable due to artificial spatial distributions or noise-like frequency spectra. Other factors, including tICA parameters, intersubject variability, motion, and limited image coverage, may have further limited the recovery of cerebellar networks. Our findings suggest that naturalistic paradigms may facilitate the recovery of cerebellar activity, particularly in species like the marmoset where the cerebellum may play a more integrated role in behavior.

Finally, in addition to the previously discussed default mode network, our analysis revealed two other human networks—the language and EXE2 components—that lacked clear analogs in the marmoset dataset. Their absence could reflect several factors: species differences in anatomical expansion and specialization[76]; divergence in cognitive capacity; and differences in the relevance of the movie content. For instance, the LAN network[77] likely emerged due to the presence of continuous speech, which holds semantic value for human viewers but likely lacks meaningful content for marmosets[51]. The EXE2 component, strongly lateralized and resembling the frontoparietal executive network described in resting-state studies[41–44,78], may reflect uniquely human mechanisms for sustained attention, memory orgoal-directed processing consistent with evolutionary expansions in lateral prefrontal cortex. The emergence of such networks highlights the sensitivity of md-fMRI and tICA to cognitive specialization and underscores the value of cross-species comparison in understanding the evolution of brain function.

In summary, our findings demonstrate that movie-driven fMRI, when combined with data-driven analytical techniques like tICA, provides a

powerful platform for mapping functional brain organization across species. This approach enables the identification of both conserved and divergent neural dynamics, offering insight into the evolutionary underpinnings of sensory integration, cognitive specialization, and large-scale brain network architecture. Importantly, emerging evidence suggests that the brain is more strongly tuned to naturalistic than to artificial stimuli (for review, see ref. 36), which supports and is consistent with our observation of robust interspecies convergence in sensory networks using movie-driven fMRI.

To facilitate transparency and reproducibility, we have made all datasets and analysis scripts used in this study openly available on Zenodo (see "Data Availability" section). The repository contains the individual- and sample-level fMRI data for both humans and marmosets, as well as the full preprocessing and analysis pipelines used in the present work. A detailed text file is provided to guide users through the structure of the repository and the steps necessary to reproduce our results. By sharing both the data and the computational workflow, we aim to enable other researchers to replicate our analyses, adapt our pipeline to their own data, and extend cross-species comparisons in future studies.

Limitations

While our current movie-driven fMRI approach offers a rich and ecologically valid stimulus, it may nevertheless be too complex for precise functional localization. However, more targeted movies could be designed to emphasize specific features, such as increased frequency of conspecific calls or varying levels of motion and contrast. It is also possible to focus analyses on selected segments of the movie to isolate responses to particular stimulus categories (e.g., faces, bodies, or vocalizations).

Another limitation is that brain regions with closely related functions —such as those along the marmoset's occipito-temporal axis—may be grouped together due to high temporal overlap in their activity patterns. In cases where finer functional distinctions are required, more targeted approaches such as functional localizers, electrophysiological recordings, or the use of more specifically designed movie stimuli may be necessary.

Finally, our functional interpretations were primarily based on the spatial distribution of components. While this strategy is common in the literature, it involves a degree of reverse inference. For example, while we labeled component HM as an executive network, and component HB as a higher-order visual/attentional network, their respective inclusion and exclusion of regions such as the frontal eye fields and superior temporal cortex illustrate how different plausible interpretations can arise depending on the spatial emphasis. Analyses linking component timecourses to stimulus features would provide stronger evidence for functional roles[33,79,80].

Methods

Common marmosets

All procedures followed Canadian Council on Animal Care guidelines and were approved by the University of Western Ontario Animal Care Committee, and we have complied with all relevant ethical regulations for animal use. Eight healthy adult common marmosets (4 females; age: 28–45 months, mean: 39.8 months) participated in the study. Each animal underwent surgery to implant a PEEK head post[81], following the protocol detailed in our recent methods article[82]. Briefly, under gas anesthesia (0.5–3% isoflurane), the skull was exposed, prepared with adhesive resin (All-Bond Universal, Bisco), and affixed with a resin composite (Core-Flo DC Lite, Bisco). Vital signs were continuously monitored throughout surgery. After 2 weeks of recovery, animals were acclimated to the head-fixation system over 3 weeks in a mock MRI environment.

Human participants

Nineteen healthy, self-reported right-handed participants (11 females; age: 25–45 years, mean: 32.7 years), with normal or corrected-to-normal vision and no neurological or psychiatric history, were recruited. Fourteen had previous fMRI experience. Demographic information including sex, age, prior experience with MRI scanning, and languages spoken are available in the Zenodo repository (https://doi.org/10.5281/zenodo.12746414). All participants provided written informed consent, and the study was approved by the University of Western Ontario Human Research Ethics Board. All ethical regulations relevant to human research participants were followed.

Stimuli

Marmosets and humans viewed a 33-min naturalistic movie composed of alternating baseline periods (5:50 min total) and excerpts from two nature documentaries: *Monkey Kingdom* (Disneynature, Spanish narration) and *Hidden Kingdoms – Urban Jungles* (BBC, English narration). The movie contained a wide variety of visual (e.g., marmosets, humans, animals, cityscapes, landscapes) and auditory stimuli (e.g., speech, vocalizations, music, environmental sounds), including dynamic social and ecological content. These features were selected to ensure engagement of multiple sensory modalities and higher-order networks in both humans and marmosets, while maintaining relevance and interest across species. Moreover, the documentaries were selected based on their frequent inclusion of scenes depicting marmoset or macaque monkeys, able to trigger the engagement of our marmoset participants. Baseline periods displayed a fixation target (black circle, 0.36° visual angle) on a gray screen without audio. The structure of the naturalistic movie is represented in detail in Supplementary Fig. 1.

Experimental setup

Marmosets. During scanning, awake animals were seated in a custom 3D-printed sphinx-style chair with head fixation using the headpost[81,82]. Visual stimuli were rear-projected onto a screen 119 cm from the eyes using a Sony VLP-FE40 LCSD projector reflected off a front-surface mirror and were presented using PowerPoint, synchronized with the MRI TTL pulse via a Raspberry Pi (model 3B+) running custom Python software. Auditory stimuli were delivered via Sensimetrics S14 tubes, secured with earplugs and veterinary bandage. An MRI-compatible camera (MRC Systems) monitored the animals, though eye tracking was unreliable due to partially closed eyelids. A drop of marshmallow-flavored liquid reward was delivered every 4.5 s via a tube to maintain alertness. Each marmoset watched the same 33-min naturalistic movie across five separate sessions. To minimize possible reductions in attention caused by repeated exposure, sessions were spaced at least 2 weeks apart. The vigilance state of the animals was monitored throughout using the MRI-compatible camera, ensuring that the eyes remained open during stimulus presentation.

Humans

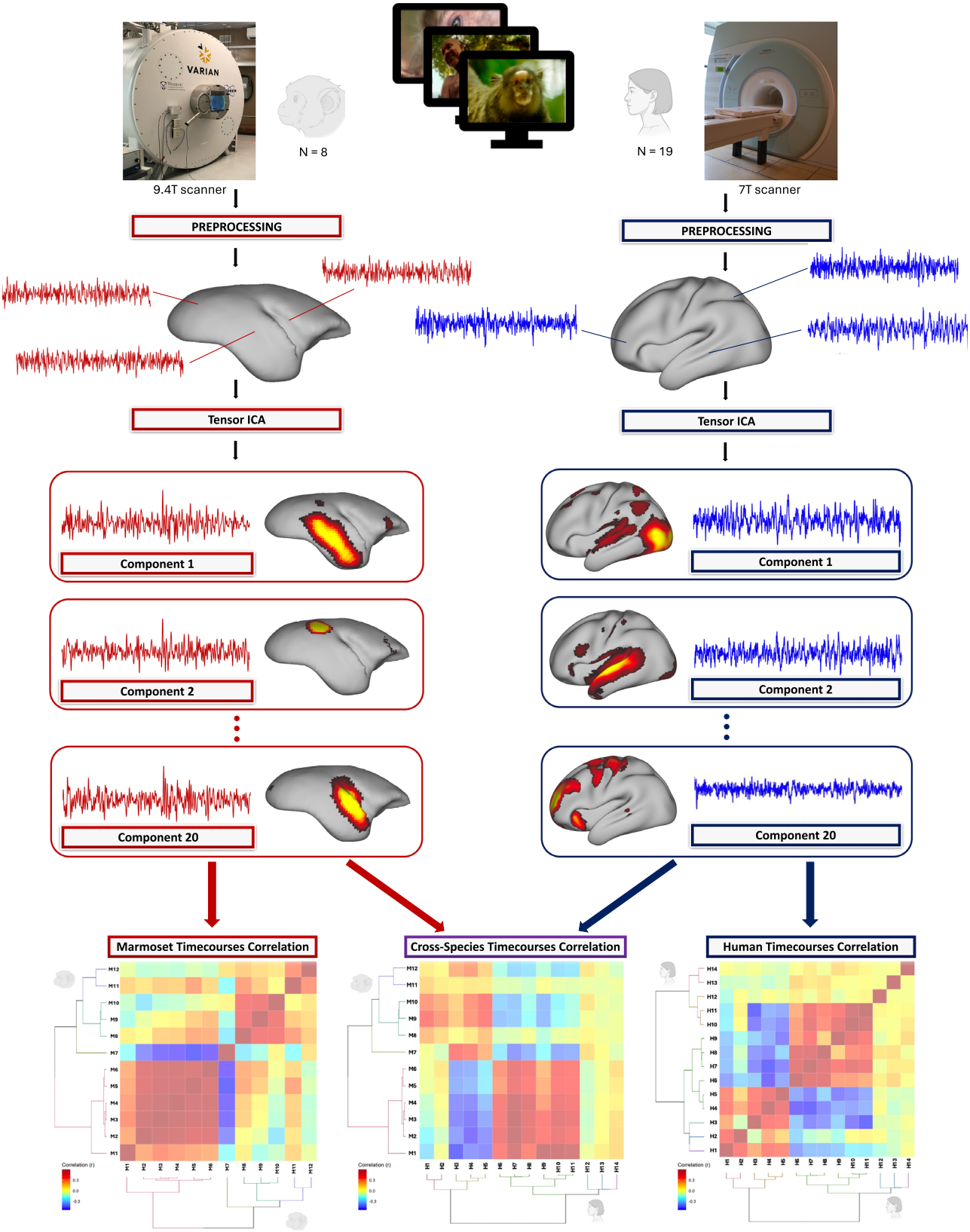
Participants lay supine in the scanner and viewed the stimulus through a mirror mounted on the head coil. Visual stimuli, projected using an Avotech SV-6011 projection system, were presented using PowerPoint, synchronized with the MRI TTL pulse via a Raspberry Pi (model 3B+) running custom Python software. Audio was delivered through Sensimetrics T14 tubes, with participants confirming acceptable volume prior to scanning. Each subject viewed the movie once.

MRI data acquisition

Marmosets. Imaging was performed at 9.4T (31-cm bore Varian magnet interfaced to a Bruker Avance NEO console) with a custom 15-cm gradient coil and eight-channel receive coil inside a quadrature birdcage transmit coil. Functional images were acquired in five sessions per animal using gradient-echo EPI (TR = 1.5 s, TE = 15 ms, flip angle = 40°, FOV = 64 × 48 mm, matrix = 96 × 128, resolution = 0.5 mm isotropic, 42 axial slices, bandwidth = 400 kHz, GRAPPA = 2). Additional EPI runs with reversed phase encoding were collected for distortion correction. A T2-weighted structural scan was acquired in one session (TR = 7 s, TE = 52 ms, FOV = 51.2 ×51.2 mm, bandwidth 50 kHz, resolution = 0.133 × 0.133 × 0.5 mm).

Humans

Data were acquired at 7 T (Siemens Magnetom MRI Plus) with a 32-channel receive and 8-channel parallel transmit coil. Functional images were acquired using multi-band EPI (TR = 1.5 s, TE = 20 ms, flip angle = 30°,



FOV = 208 × 208 mm, matrix = 104 × 104, resolution = 2 mm isotropic, 62 slices, GRAPPA = 3, multi-band factor = 2). Field maps were derived from magnitude and dual-phase images. MP2RAGE structural images were collected (TR = 6 s, TE = 2.13 ms, TI1/TI2 = 800/2700 ms, resolution = 0.75 mm isotropic).

fMRI data preprocessing

Marmoset data. Preprocessing was performed using AFNI[83] and FSL[84]. DICOMs were converted with dcm2niix, reoriented (fslswapdim, fslorient), and distortion-corrected (topup, applytopup). Volumes were despiked, slice-time corrected, motion-corrected, and spatially smoothed

Fig. 8 | Experimental design and stimulus structure. Subjects viewed a 33-min movie composed of alternating baseline and movie blocks. Baseline periods (totaling 5:50 min) consisted of a fixation point on a gray background with no audio. Movie segments included clips from two nature documentaries, featuring a variety of species and environments. The film contained naturalistic visual and auditory stimuli, including conspecific and heterospecific vocalizations, human speech, music, and environmental sounds. At the top, the 7 Tesla and 9.4 Tesla scanners used for fMRI sessions in humans ($n = 19$) and common marmosets ($n = 8$, with 5 functional runs per marmoset) are shown. On the left and right sides of the figure is represented the statistical approach used for both marmosets (left) and human (right). After preprocessing, individual fMRI maps were analyzed using tICA, extracting 20 independent components per species. The timecourses of non-noise components ($n = 12$ for marmosets and $n = 14$ for humans) were then correlated within and across species, generating correlation matrices (example shown at the bottom). Additionally, these functional components were clustered using hierarchical clustering to further explore functional network relationships in both species.

(1.5 mm FWHM). Nuisance regression included motion parameters. Temporal filtering (0.01–0.08 Hz) was applied using 3dBandpass. Functional data were registered to individual anatomy (FLIRT), then to the NIH marmoset template[85] using ANTs[86].

Human data

Preprocessing was done with SPM12 (Wellcome Center for Human Neuroimaging, London, UK) and AFNI. After conversion with dcm2niix, images were realigned, slice-time corrected, field map–corrected, coregistered to structural MP2RAGE, normalized to MNI space, and smoothed (6 mm FWHM). Temporal filtering (0.01–0.08 Hz) matched the marmoset pipeline.

Statistics and reproducibility

We applied tICA using MELODIC[87] with a 20-component cutoff for each species, using the entire 33-min movie. This method decomposes data into spatially and temporally independent components, enabling identification of stimulus-driven functional networks without relying on predefined templates. For the marmosets, the data from the five sessions were averaged within each animal prior to performing tICA. Twelve components in marmosets and fourteen in humans were classified as non-noise based on spatial and spectral criteria.

We computed within- and between-species correlation matrices of component timecourses (following conventional benchmarks, we classified correlations as weak ($r < 0.2$), moderate ($0.2 < r < 0.5$), or strong ($r > 0.5$)). To assess the statistical significance of cross-species temporal correlations, we implemented a non-parametric permutation test. For each pair of human–marmoset components, we kept the human timecourses fixed and applied random circular shifts to the marmoset timecourses ($N = 1000$ permutations), recomputing the correlations at each iteration. This procedure generated a null distribution of correlations for each component pair, against which the observed correlation was compared to obtain a p value. Resulting p values were corrected for multiple comparisons across the matrix using false discovery rate (FDR) adjustment (Benjamini–Hochberg). Then, we performed hierarchical clustering (squared Euclidean distance, Ward's method) in R[88] using the cluster[89], dendextend[90], and ggdendro[91] packages. These analyses allowed us to explore the organization of functional networks and to investigate interspecies relationships. This approach is displayed in Fig. 8.

To assess spatial coverage, non-noise components were converted to z-statistic images by MELODIC and thresholded using the component-specific values automatically estimated by the mixture-model fit. These thresholds emphasize voxels most strongly associated with each component, and should be understood as a heuristic for separating structured signal from background noise rather than as a formal voxelwise statistical test. Then, thresholded maps were binarized and averaged to generate probability maps. These were compared with tSNR maps to evaluate recruitment consistency and signal quality. Specifically, for both cortical and subcortical regions, mean tSNR values were extracted using binary masks from the corresponding atlases. Paired t-tests were then performed to compare cortical and subcortical signal quality, to determine whether differences in tSNR could confound the detection of subcortical networks in either species.

Reporting summary

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

Data availability

The individual maps for the human and marmoset samples are fully available online on Zenodo: https://doi.org/10.5281/zenodo.12746414[92].

Code availability

The code underpinning this study is fully available online on Zenodo: https://doi.org/10.5281/zenodo.12746414[92].

Received: 5 June 2025; Accepted: 12 November 2025;
Published online: 06 January 2026

References

1. Okano, H. Current status of and perspectives on the application of marmosets in neurobiology. *Annu. Rev. Neurosci.* **44**, 27–48 (2021).
2. Miller, C. T. et al. Marmosets: a neuroscientific model of human social behavior. *Neuron* **90**, 219–233 (2016).
3. Burkart, J. M. & Finkenwirth, C. Marmosets as model species in neuroscience and evolutionary anthropology. *Neurosci. Res.* **93**, 8–19 (2015).
4. Burman, K. J. & Rosa, M. G. P. Architectural subdivisions of medial and orbital frontal cortices in the marmoset monkey (Callithrix jacchus). *J. Comp. Neurol.* **514**, 11–29 (2009).
5. Solomon, S. G. & Rosa, M. G. P. A simpler primate brain: the visual system of the marmoset monkey. *Front Neural Circuits* **8**, 8 (2014).
6. Chen, C. Y. et al. Properties of visually guided saccadic behavior and bottom-up attention in marmoset, macaque, and human. *J. Neurophysiol.* **125**, 437–457 (2021).
7. Mitchell, J. F. & Leopold, D. A. The marmoset monkey as a model for visual neuroscience. *Neurosci. Res.* **93**, 20–46 (2015).
8. Jafari, A. et al. A vocalization-processing network in marmosets. *Cell Rep.* **42**, 112526 (2023).
9. Dureux, A., Zanini, A. & Everling, S. Mapping of facial and vocal processing in common marmosets with ultra-high field fMRI. *Commun. Biol.* **7**, 1–15 (2024).
10. Dureux A., Zanini A., Jafari A. & Everling S. Ultra-high field fMRI reveals effect of ketamine on vocal processing in common marmosets. *J. Neurosci*. **45**, e0651242025 (2025).
11. Garin, C. M. et al. An evolutionary gap in primate default mode network organization. *Cell Rep.* **39**, 110669 (2022).
12. Ghahremani, M., Hutchison, R. M., Menon, R. S. & Everling, S. Frontoparietal functional connectivity in the common marmoset. *Cereb. Cortex* **27**, 3890–3905 (2017).
13. Hori, Y. et al. Cortico-subcortical functional connectivity profiles of resting-state networks in marmosets and humans. *J. Neurosci.* **40**, 9236–9249 (2020).
14. Hori, Y. et al. Interspecies activation correlations reveal functional correspondences between marmoset and human brain areas. *Proc. Natl Acad. Sci.* **118**, e2110980118 (2021).
15. Schaeffer, D. J. et al. Task-based fMRI of a free-viewing visuo-saccadic network in the marmoset monkey. *NeuroImage* **202**, 116147 (2019).
16. Dureux, A., Zanini, A., Selvanayagam, J., Menon, R. S. & Everling, S. Gaze patterns and brain activations in humans and marmosets in the Frith-Happé theory-of-mind animation task. *Elife* **12**, e86327 (2023).

17. Ji, J. L. et al. Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage* **185**, 35–57 (2019).
18. Yeo, T. B. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
19. Chao, L., Martin, A. & Haxby, J. Are face-responsive regions selective only for faces? *NeuroReport* **10**, 2945–2950 (1999).
20. Pinsk, M. A. et al. Neural representations of faces and body parts in macaque and human cortex: a comparative fMRI study. *J. Neurophysiol.* **101**, 2581–2600 (2009).
21. Laurent, M. A. et al. Towards an optimization of functional localizers in non-human primate neuroimaging with (fMRI) frequency-tagging. *NeuroImage* **270**, 119959 (2023).
22. Dureux, A., Zanini, A. & Everling, S. Face-selective patches in marmosets are involved in dynamic and static facial expression processing. *J. Neurosci.* **43**, 3477–3494 (2023).
23. Schaeffer, D. J. et al. Face selective patches in marmoset frontal cortex. *Nat. Commun.* **11**, 4856 (2020).
24. Hung, C. C. et al. Functional mapping of face-selective regions in the extrastriate visual cortex of the marmoset. *J. Neurosci.* **35**, 1160–1172 (2015).
25. Hung, C. C. et al. Functional MRI of visual responses in the awake, behaving marmoset. *NeuroImage* **120**, 1–11 (2015).
26. Zanini, A., Dureux, A., Selvanayagam, J. & Everling, S. Ultra-high field fMRI identifies an action-observation network in the common marmoset. *Commun. Biol.* **6**, 1–11 (2023).
27. Hasson, U., Furman, O., Clark, D., Dudai, Y. & Davachi, L. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* **57**, 452–462 (2008).
28. Hasson, U., Malach, R. & Heeger, D. J. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* **14**, 40–48 (2010).
29. Lerner, Y., Honey, C. J., Silbert, L. J. & Hasson, U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* **31**, 2906–2915 (2011).
30. Naci, L., Cusack, R., Anello, M. & Owen, A. M. A common neural code for similar conscious experiences in different individuals. *Proc. Natl Acad. Sci.* **111**, 14277–14282 (2014).
31. Rajimehr, R. et al. Functional architecture of cerebral cortex during naturalistic movie watching. *Neuron* **112**, 4130–4146 (2024).
32. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
33. Bartels, A. & Zeki, S. Brain dynamics during natural viewing conditions —a new guide for mapping connectivity in vivo. *NeuroImage* **24**, 339–349 (2005).
34. Russ, B. E. & Leopold, D. A. Functional MRI mapping of dynamic visual features during natural viewing in the macaque. *NeuroImage* **109**, 84–94 (2015).
35. Park, S. H. et al. Parallel functional subnetworks embedded in the macaque face patch system. *Sci. Adv.* **8**, eabm2054 (2022).
36. Sonkusare, S., Breakspear, M. & Guo, C. Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* **23**, 699–714 (2019).
37. Belcher, A. M. et al. Large-scale brain networks in the awake, truly resting marmoset monkey. *J. Neurosci.* **33**, 16796–16804 (2013).
38. Muta, K. et al. Commonality and variance of resting-state networks in common marmoset brains. *Sci. Rep.* **14**, 8316 (2024).
39. Ngo, G. N., Hori, Y., Everling, S. & Menon, R. S. Joint-embeddings reveal functional differences in default-mode network architecture between marmosets and humans. *NeuroImage* **272**, 120035 (2023).
40. Liu, C. et al. Anatomical and functional investigation of the marmoset default mode network. *Nat. Commun.* **10**, 1975 (2019).
41. Damoiseaux, J. S. et al. Consistent resting-state networks across healthy subjects. *Proc. Natl Acad. Sci.* **103**, 13848–13853 (2006).
42. Damoiseaux, J. S. et al. Reduced resting-state brain activity in the "default network" in normal aging. *Cereb. Cortex* **18**, 1856–1864 (2008).
43. Beckmann, C. F., DeLuca, M., Devlin, J. T. & Smith, S. M. Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. B: Biol. Sci.* **360**, 1001–1013 (2005).
44. Allen, E. A. et al. A baseline for the multivariate comparison of resting-state networks. *Front. Syst. Neurosci*. **5**, 2 (2011).
45. Seeley, W. W. The salience network: a neural system for perceiving and responding to homeostatic demands. *J. Neurosci.* **39**, 9878–9882 (2019).
46. Binder J. R. fMRI of language systems. In *fMRI Techniques and Protocols* (ed Filippi, M.) 323–351 (Humana Press, 2009).
47. Turker, S., Kuhnke, P., Eickhoff, S. B., Caspers, S. & Hartwigsen, G. Cortical, subcortical, and cerebellar contributions to language processing: A meta-analytic review of 403 neuroimaging experiments. *Psychol. Bull.* **149**, 699–723 (2023).
48. Seitzman, B. A., Snyder, A. Z., Leuthardt, E. C. & Shimony, J. S. The state of resting state networks. *Top. Magn. Reson Imaging* **28**, 189–196 (2019).
49. Greicius, M. D., Krasnow, B., Reiss, A. L. & Menon, V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl Acad. Sci. USA* **100**, 253–258 (2003).
50. Mantini, D. et al. Interspecies activity correlations reveal functional correspondence between monkey and human brain areas. *Nat. Methods* **9**, 277–282 (2012).
51. Friederici, A. D. *Language in Our Brain: The Origins of a Uniquely Human Capacity* 300 (MIT Press, 2017).
52. Belcher, A. M. et al. Functional connectivity hubs and networks in the awake marmoset brain. *Front. Integr. Neurosci.* **10**, 9 (2016).
53. Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H. & Martin, A. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat. Neurosci.* **7**, 1190–1192 (2004).
54. Froesel, M., Cappe, C. & Ben Hamed, S. A multisensory perspective onto primate pulvinar functions. *Neurosci. Biobehav. Rev.* **125**, 231–243 (2021).
55. Stein, B. E., Meredith, M. A. & Wallace, M. T. The visually responsive neuron and beyond: multisensory integration in cat and monkey. In *Progress in Brain Research* (eds Hicks, T. P., Molotchnikoff, S., Ono, T.) Ch. 8, 79–90 (Elsevier, 1993).
56. Calhoun, V. D. et al. Different activation dynamics in multiple neural systems during simulated driving. *Hum. Brain Mapp.* **16**, 158–167 (2002).
57. Valchev, N., Gazzola, V., Avenanti, A. & Keysers, C. Primary somatosensory contribution to action observation brain activity—combining fMRI and cTBS. *Soc. Cogn. Affect. Neurosci.* **11**, 1205–1217 (2016).
58. Gazzola, V. & Keysers, C. The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb. Cortex* **19**, 1239–1255 (2009).
59. Keysers, C., Kaas, J. H. & Gazzola, V. Somatosensation in social perception. *Nat. Rev. Neurosci.* **11**, 417–428 (2010).
60. Caspers, S., Zilles, K., Laird, A. R. & Eickhoff, S. B. ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage* **50**, 1148–1167 (2010).
61. Cross, E. S., Hamilton, A. F. deC., Kraemer, D. J. M., Kelley, W. M. & Grafton, S. T. Dissociable substrates for body motion and physical experience in the human action observation network. *Eur. J. Neurosci.* **30**, 1383–1392 (2009).
62. Gazzola, V., Rizzolatti, G., Wicker, B. & Keysers, C. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *NeuroImage* **35**, 1674–1684 (2007).

63. Uddin, L. Q. Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* **16**, 55–61 (2015).
64. Iannetti, G. D. & Mouraux, A. From the neuromatrix to the pain matrix (and back). *Exp. Brain Res.* **205**, 1–12 (2010).
65. Menon, V. 20 years of the default mode network: a review and synthesis. *Neuron* **111**, 2469–2487 (2023).
66. Fox, M. D. et al. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl Acad. Sci.* **102**, 9673–9678 (2005).
67. Fox, K. C. R., Spreng, R. N., Ellamil, M., Andrews-Hanna, J. R. & Christoff, K. The wandering brain: meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *NeuroImage* **111**, 611–621 (2015).
68. Fransson, P. Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* **26**, 15–29 (2005).
69. Long, X. Y. et al. Default mode network as revealed with multiple methods for resting-state functional MRI analysis. *J. Neurosci. Methods* **171**, 349–355 (2008).
70. Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl Acad. Sci.* **113**, 12574–12579 (2016).
71. Simony, E. et al. Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* **7**, 12141 (2016).
72. Robinson, S. et al. A resting state network in the motor control circuit of the basal ganglia. *BMC Neurosci.* **10**, 137 (2009).
73. Kim, D. J., Park, B. & Park, H. J. Functional connectivity-based identification of subdivisions of the basal ganglia and thalamus using multilevel independent component analysis of resting state fMRI. *Hum. Brain Mapp.* **34**, 1371–1385 (2013).
74. Song, L. et al. Mapping the time-varying functional brain networks in response to naturalistic movie stimuli. *Front. Neurosci*. **17**, 1199150 (2023).
75. Dobromyslin, V. I. et al. Distinct functional networks within the cerebellum and their relation to cortical systems assessed with independent component analysis. *Neuroimage* **60**, 2073–2085 (2012).
76. Mars, R. B. et al. Whole brain comparative anatomy using connectivity blueprints. *eLife* **7**, e35237 (2018).
77. Ferstl, E. C., Neumann, J., Bogler, C. & von Cramon, D. Y. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum. Brain Mapp.* **29**, 581–593 (2008).
78. Galletti, C. & Fattori, P. The dorsal visual stream revisited: stable circuits or dynamic pathways? *Cortex* **98**, 203–217 (2018).
79. Ylipaavalniemi, J. et al. Dependencies between stimuli and spatially independent fMRI sources: towards brain correlates of natural stimuli. *NeuroImage* **48**, 176–185 (2009).
80. Lahnakoski, J. M. et al. Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Front. Hum. Neurosci*. **6**, 233 (2012).
81. Gilbert, K. M. et al. A radiofrequency coil to facilitate task-based fMRI of awake marmosets. *J. Neurosci. Methods* **383**, 109737 (2023).
82. Zanini, A. et al. In vivo functional brain mapping using ultra-high-field fMRI in awake common marmosets. *STAR Protoc.* **4**, 102586 (2023).
83. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
84. Smith, S. M. et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).
85. Liu, C. et al. A digital 3D atlas of the marmoset brain based on multi-modal MRI. *NeuroImage* **169**, 106–116 (2018).
86. Avants, B. B. et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* **54**, 2033–2044 (2011).
87. Smith, S. M. et al. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl Acad. Sci.* **106**, 13040–13045 (2009).
88. R. Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; Available from: https://www.R-project.org/ (2023).
89. Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. cluster: cluster analysis basics and extensions [Internet]. Available from: https://CRAN.R-project.org/package=cluster (2023).
90. Galili, T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31**, 3718–3720 (2015).
91. de Vries, A. & Ripley, B. D. ggdendro: create dendrograms and tree diagrams using 'ggplot2' [Internet]. Available from: https://CRAN.R-project.org/package=ggdendro (2024).
92. Zanini, A. Identifying Functional Homologues in Human and Marmoset Brain Networks via Movie-Driven Ultra-High Field fMRI (sample and human data) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.12746414 (2024).
93. Paxinos, G., Watson, C., Petrides, M., Rosa, M. G. P. & Tokuno, H. *The Marmoset Brain in Stereotaxic Coordinates* (Elsevier Academic Press, 2012).
94. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).

Acknowledgements

Support was provided by a Discovery grant by the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research (FRN 183973). We also acknowledge the support of the Government of Canada's New Frontiers in Research Fund (NFRF), [NFRF-T-2022-00051]. We wish to thank Cheryl Vander Tuin, Whitney Froese, Hannah Pettypiece, and Miranda Bellyou for animal preparation and care, Dr. Alex Li and Trevor Szekeres for scanning assistance, Dr. Kyle Gilbert, and Peter Zeman for coil designs.

Author contributions

A.Z. and A.D. acquired the human and marmoset data. A.Z. performed the analyses and wrote the manuscript. S.E. developed the study concept and provided resources. R.M. and S.E. supervised the project. All authors contributed to manuscript revision and approved the final version.

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/s42003-025-09241-y.

Correspondence and requests for materials should be addressed to Alessandro Zanini.

Peer review information *Communications Biology* thanks Fumiko Seki and Juha Lahnakoski for their contribution to the peer review of this work. Primary Handling Editors: Shenbing Kuang and Jasmine Pan. 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025