Supplemental Material for

Evaluation of denoising strategies for task-based functional connectivity: equalizing residual motion artifacts between rest and cognitively demanding tasks

Daniele Mascali[a,b*], Marta Moraschi[a,b], Mauro DiNuzzo[a,b], Silvia Tommasin[c], Michela Fratini[b,d], Tommaso Gili[e], Richard G. Wise[f,g], Silvia Mangia[h], Emiliano Macaluso[i], Federico Giove[a,b]

[a]MARBILab, CREF - Centro Ricerche Enrico Fermi, Roma, Italy.

[b]Fondazione Santa Lucia IRCCS, Roma, Italy.

[c]Dipartimento di Neuroscienze umane, Sapienza Università di Roma, Roma, Italy.

[d]Istituto di Nanotecnologia, Consiglio Nazionale delle Ricerche, Roma, Italy.

[e]IMT Scuola Alti Studi Lucca, Lucca, Italy.

[f]Cardiff University Brain Research Imaging Centre, School of Psychology, Cardiff University, Cardiff, UK.

[g]Institute for Advanced Biomedical Technologies, Department of Neuroscience, Imaging and Clinical Sciences, "G. D'Annunzio University" of Chieti-Pescara, 66100, Chieti, Italy.

[h]Center for Magnetic Resonance Research, Dept. Of Radiology, University of Minnesota, Minneapolis, USA.

[i]ImpAct Team, Lyon Neuroscience Research Center, Lyon, France.

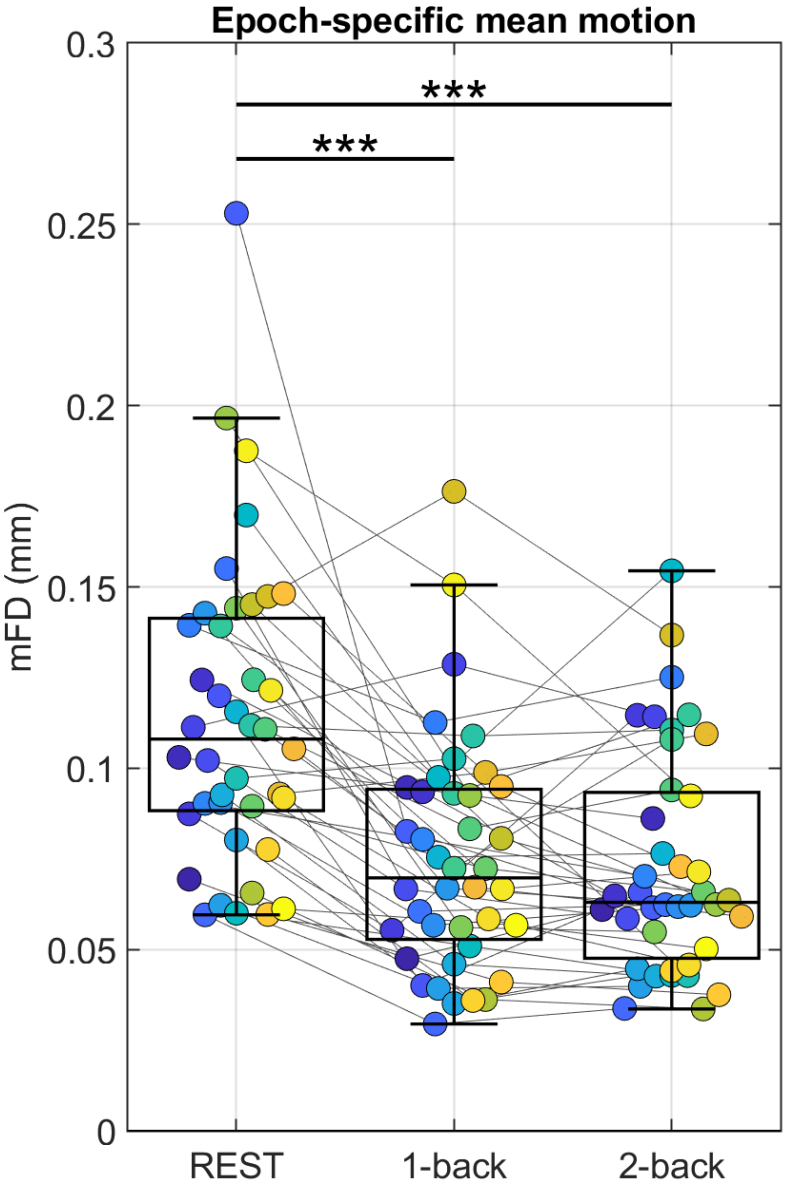

Fig S1. Box plots of mFD for the three functional conditions of the CF dataset. The two task conditions (1-back and 2-back) showed no significant difference in the amount of motion (paired t-test: t = 0.69, df = 19, p = 0.50). Rest was statistically different from both 1-back (paired t-test: t = 5.05, df = 19 , p = $7.1*10^{-5}$) and 2-back (paired t-test: t = 4.79, df = 19 , p = $1.3*10^{-4}$). Statistical tests were performed after averaging the two runs. ***, p < 0.001.

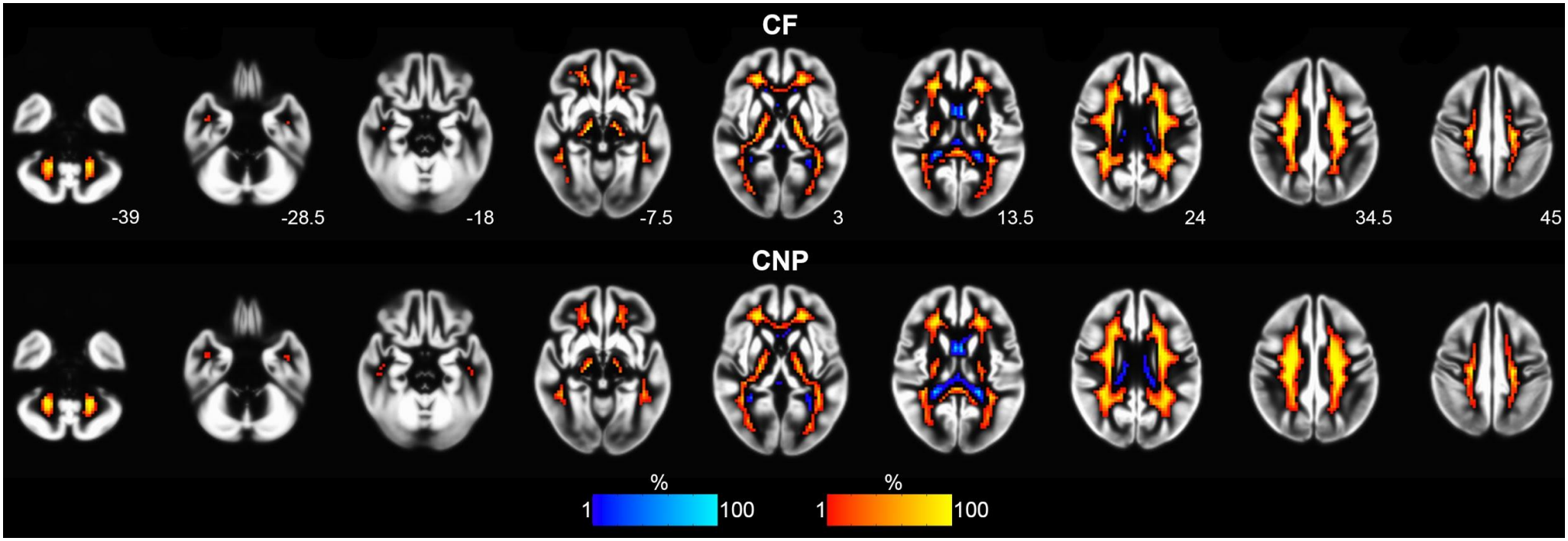

Fig S2. Frequency maps of the WM and CSF masks used for extracting confounding signals. The maps represent the within-dataset probability for voxels to be included in either the WM mask (red to orange) or the CSF mask (blue to cyan). The masks used for generating the probability maps underwent the processing described in the main text. The underlay image is the SPM12 grey matter tissue probability map; the numbers next to each slice refer to z coordinates in MNI space.

		|WM:GM correlation|		WM Voxel Count		|CSF:GM correlation|		CSF Voxel Count	
Dataset	**N**	**mean**	**std**	**mean**	**std**	**mean**	**std**	**mean**	**std**
CF	20	0.54	0.30	1408	296	0.26	0.28	40	24
CNP	120	0.39	0.24	1482	331	0.23	0.17	59	57

Table S1. Characteristics of the WM and CSF masks used for extracting confounding signals. For each tissue compartment, the table reports the mean and the standard deviation within each dataset of 1) the absolute Pearson's correlation between the signal extracted from the mask and the signal extracted from the gray matter (GM), and 2) the number of voxels composing the mask (voxel size = 27 mm^3). Before averaging, correlations were z-to-Fisher transformed and then back converted to Pearson's value. The mean GM signal was extracted with a mask obtained by thresholding at 95% the subject-specific T1-derived grey-matter probability map.

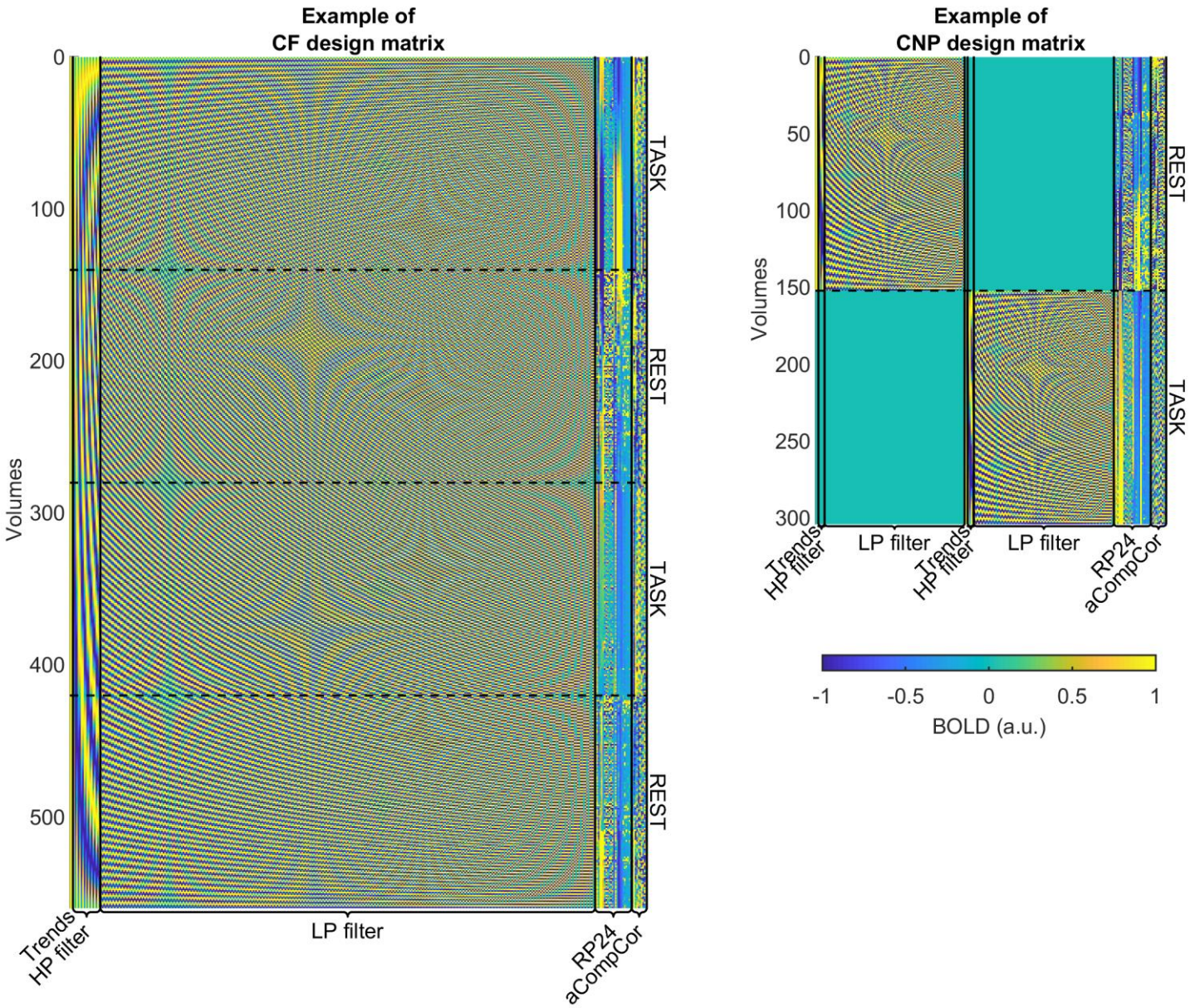

Fig S3. Representative denoising matrices for CF and CNP datasets. The plots show the explanatory variables (i.e., the denoising matrices) used for linear regression of nuisance signals in the case of the 24RP+aCompCor pipeline for an example subject. Columns in the matrix contain different explanatory variables grouped in blocks: trends, high-pass (HP) filter, low-pass (LP) filter, RP24 and aCompCor; while rows run on volumes. Note the different scheme used for processing the two datasets. CNP data were processed with two orthogonal sets of regressors for trends, HP and LP filtering, so to account for phase shifts between the two concatenated runs. For CF data such scheme was not required since the different epochs were acquired within the same functional run. The matrices are plotted preserving their aspect ratio so that the reader can appreciate the dissimilarities between datasets in terms of number of volumes and number of explanatory variables.

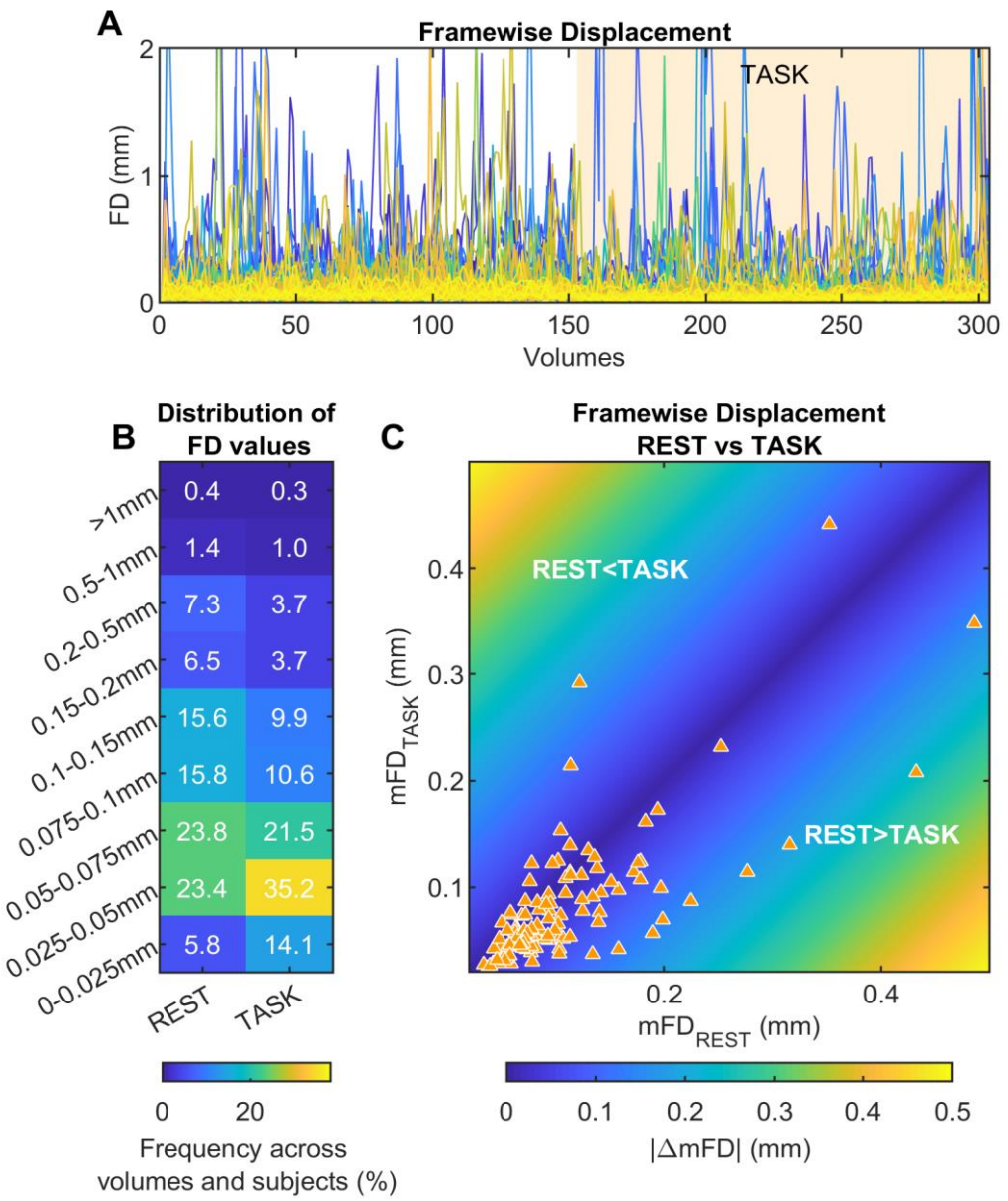

Fig S4. Evaluation of in-scanner head movement for the CNP dataset. A) FD series for each subjects. B) Distribution of FD values in 9 bins of different width, showing the marked difference in the distribution of FD values between rest and task epochs. C) Task-averaged FD vs rest-averaged FD

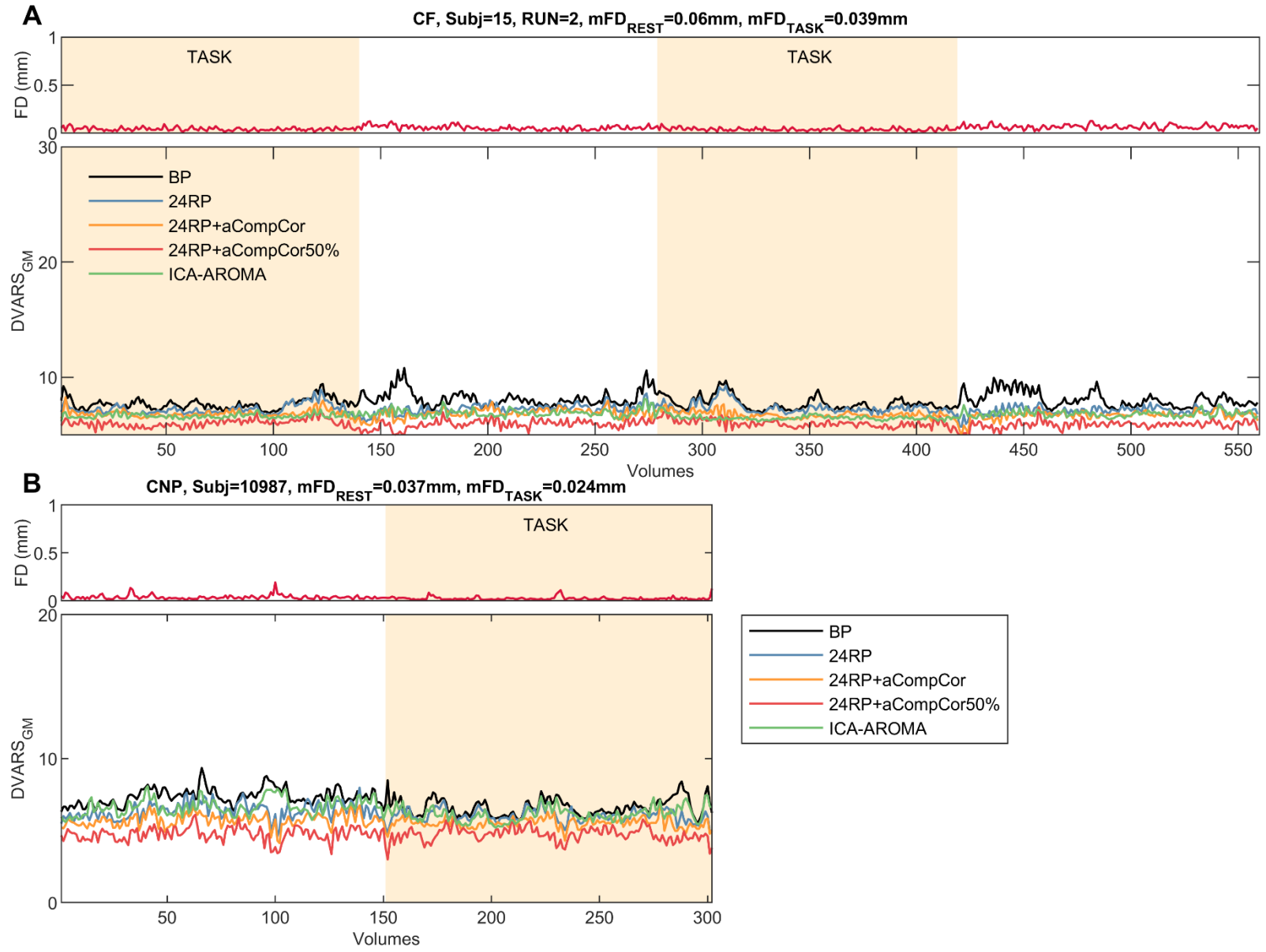

Figure S5. QC series for the stillest subject of the A) CF and B) CNP dataset. In each main panel, the first row shows the FD series, while the second row shows $DVARS_{GM}$ series calculated after applying 5 different denoising models. BP (band-pass, black line) is a denoising model containing only trends and band-pass regressors

Figure S6. QC series for an average-moving subject within the A) CF and B) CNP dataset.In each main panel, the first row shows the FD series, while the second row shows $DVARS_{GM}$ series calculated after applying 5 different denoising models. BP (band-pass, black line) is a denoising model containing only trends and band-pass regressors.

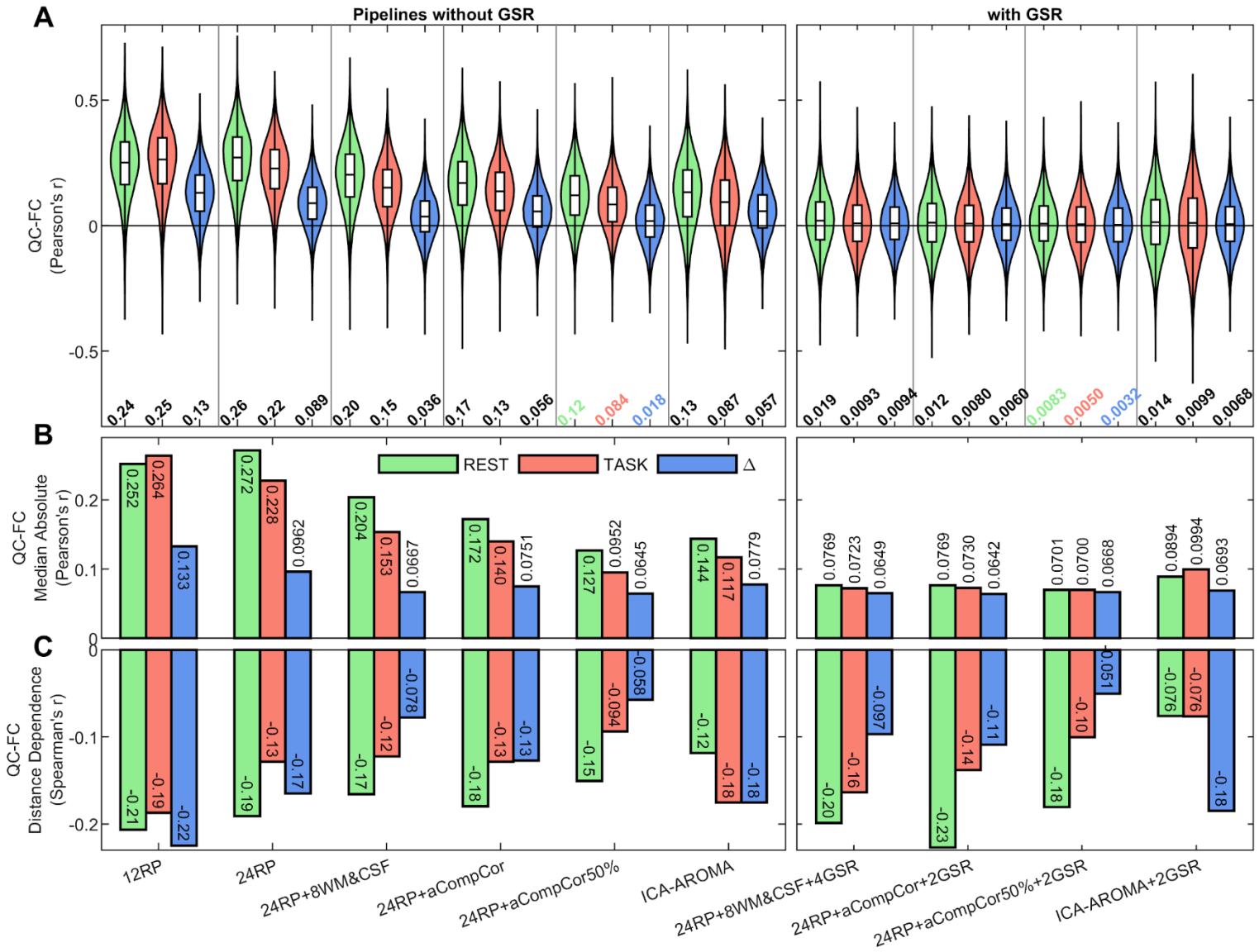

Figure S7. QC-FC plots for evaluating the across-subject relationship between motion (mFD) and connectivity estimates under different denoising strategies for the CNP dataset. The top panels (A) show the distribution of QC-FC correlations along with the absolute mean value of the correlations, aiming at quantifying the centering of the distributions. The middle panels (B) show the median value of the absolute QC-FC correlations, which takes into account both the centering and the spread of the distribution. The bottom panels (C) show the Spearman's correlation between QC-FC correlations and the Euclidean distance between pairs of nodes, indexing distance-dependent artifacts. QC-FC results are displayed for REST and TASK separately. For the task-based change in FC ($\Delta FC = FC_{task} - FC_{rest}$), the residual relationship with motion was evaluated with respect to the change in mFD ($mFD = mFD_{task} - mFD_{rest}$).

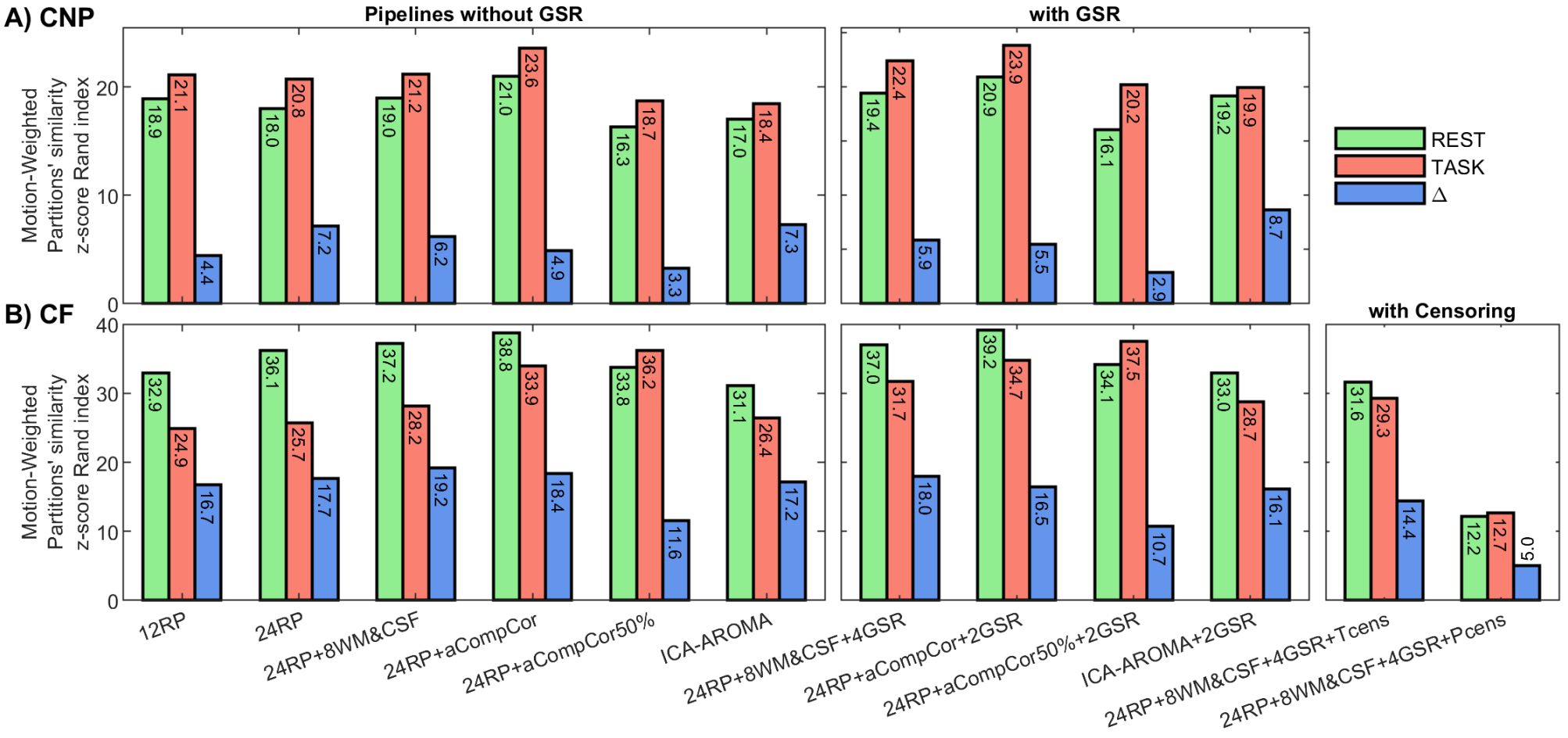

Figure S8. Motion-weighted partition's similarity for the (A) CNP and (B) CF dataset. Before computing the average similarity, the pairwise similarity matrix was weighted by 1/<mFD>, where the mean is calculated over the selected pair of subjects. The weights penalize partitions from high-moving subjects.

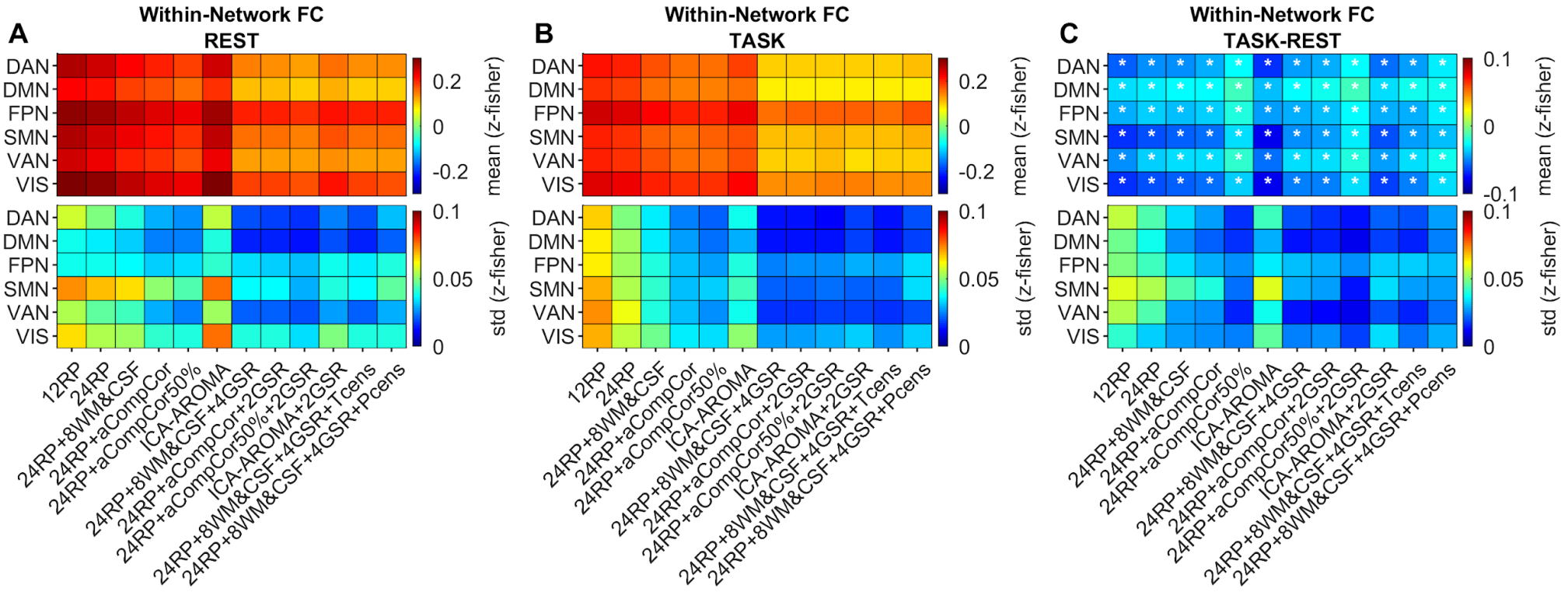

Figure S9. Within-network FC calculated under different processing pipelines using the CF dataset. Across-subject mean (top row) and standard deviation (middle) of within-network FC are shown for A) rest, B) task, and for the relative change C). For the latter, a white symbol marks the statistical significance of a second-level one-sample t-test, with "*" indicating $p < 0.001$. The magnitude of the task-associated change in FC is significantly negative for all pipelines, but tends to decrease in magnitude with more efficient pipelines, with the exception of ICA-AROMA, which shows the greatest magnitude. The across-subject standard deviation tends to decrease with more efficient pipelines for all considered networks, with the exception of ICA-AROMA that shows mixed results depending on the networks.

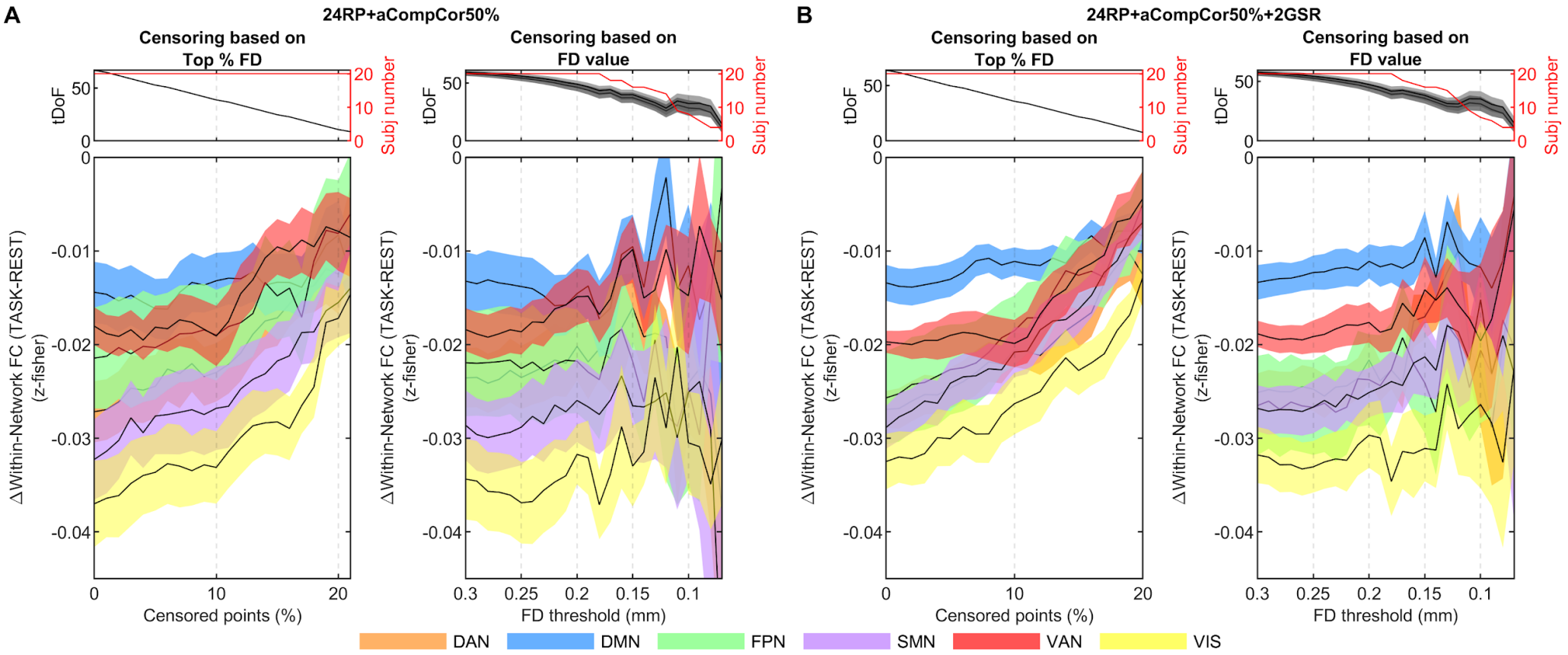

Figure S10. Incremental censoring analysis for within-network FC considering models (A) RP24+aCompCor50% and (B) RP24+aCompCor50%+2GSR. In each panel, the left plot shows a P-censoring analysis while the right plot a T-censoring analysis. In P-censoring an equal number of volumes were excised from rest and task conditions ensuring conditions comparability in terms of tDoF, yet at the expense of removing potentially good volumes in the task condition. In T-censoring a more efficient data cleaning comes at the expense of variable tDoF among conditions and, in case of severe thresholds, at the progressive elimination of subjects with the highest motion. For each network, the mean across subjects of FC is shown as a black line along with shades representing the standard error of the mean (SEM), color-coded based on the network. Likewise, tDoF for the T-censoring variant are represented with the mean and SEM (light-gray shade for task, darker for rest). DAN: Dorsal Attention, DMN: Default Mode, FPN: Frontoparietal, SMN: Somatomotor, VAN: Ventral Attention and VIS: Visual network.

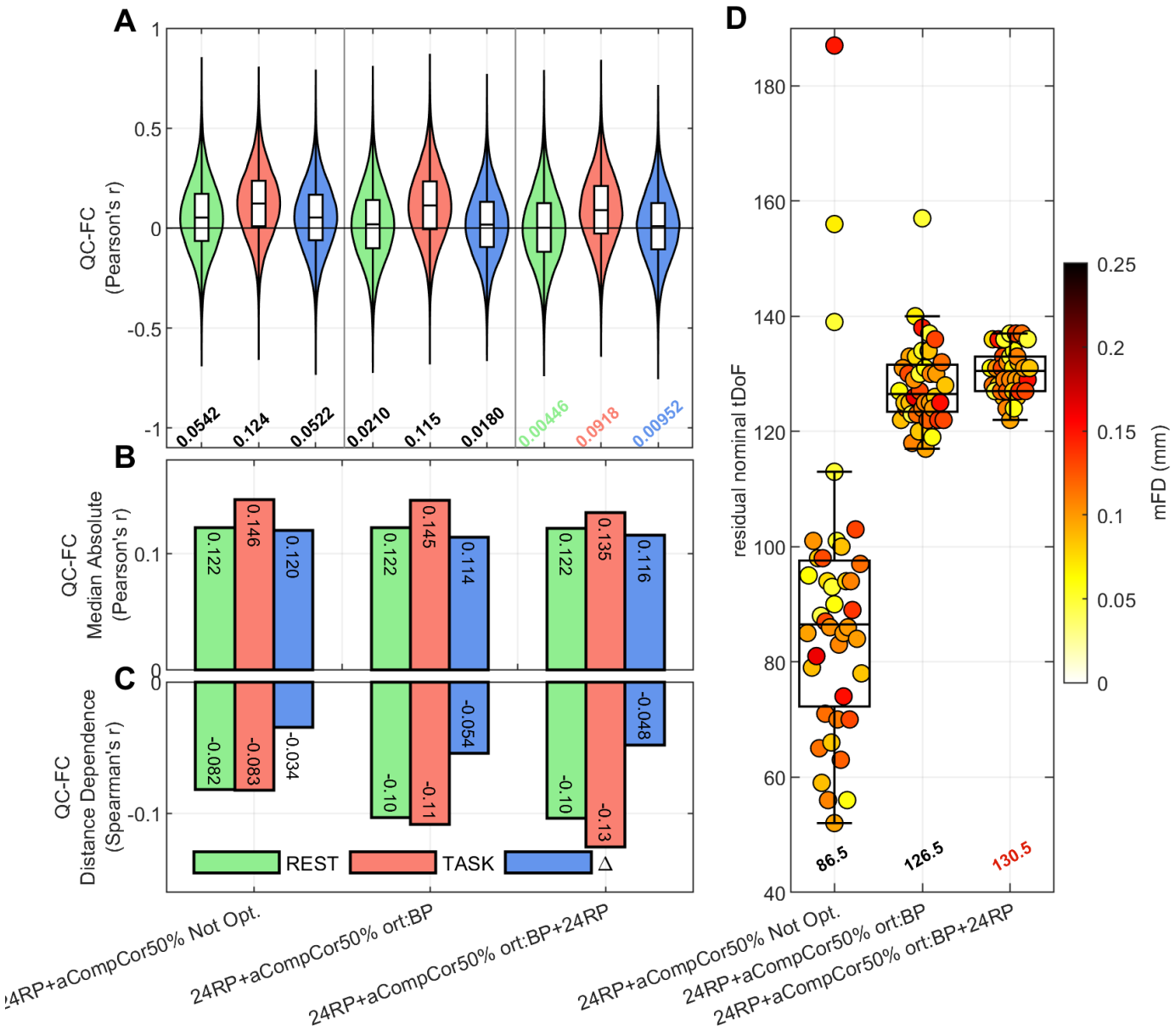

Figure S11. Illustration of the benefits produced by the optimization scheme adopted in the aCompCor pipelines. Panels A-C show the QC-FC benchmarks comparing the aCompCor50% model of the CF dataset with and without pre orthogonalization with respect to different subsets of the confounding variables that composed the model. The first model is without any optimization. In the second model data are orthogonalized with respect to the sine and cosine functions used for band-pass filtering (ort:BP), the third model with respect to both BP and the 24RP set (ort:BP+24RP). The metrics are described in Fig. 4 of the main text. Panel D shows the residual tDoF after the application of each of the above-defined aCompCor50% variants. The tDoF of each run are color-coded based on the mFD of the whole series.

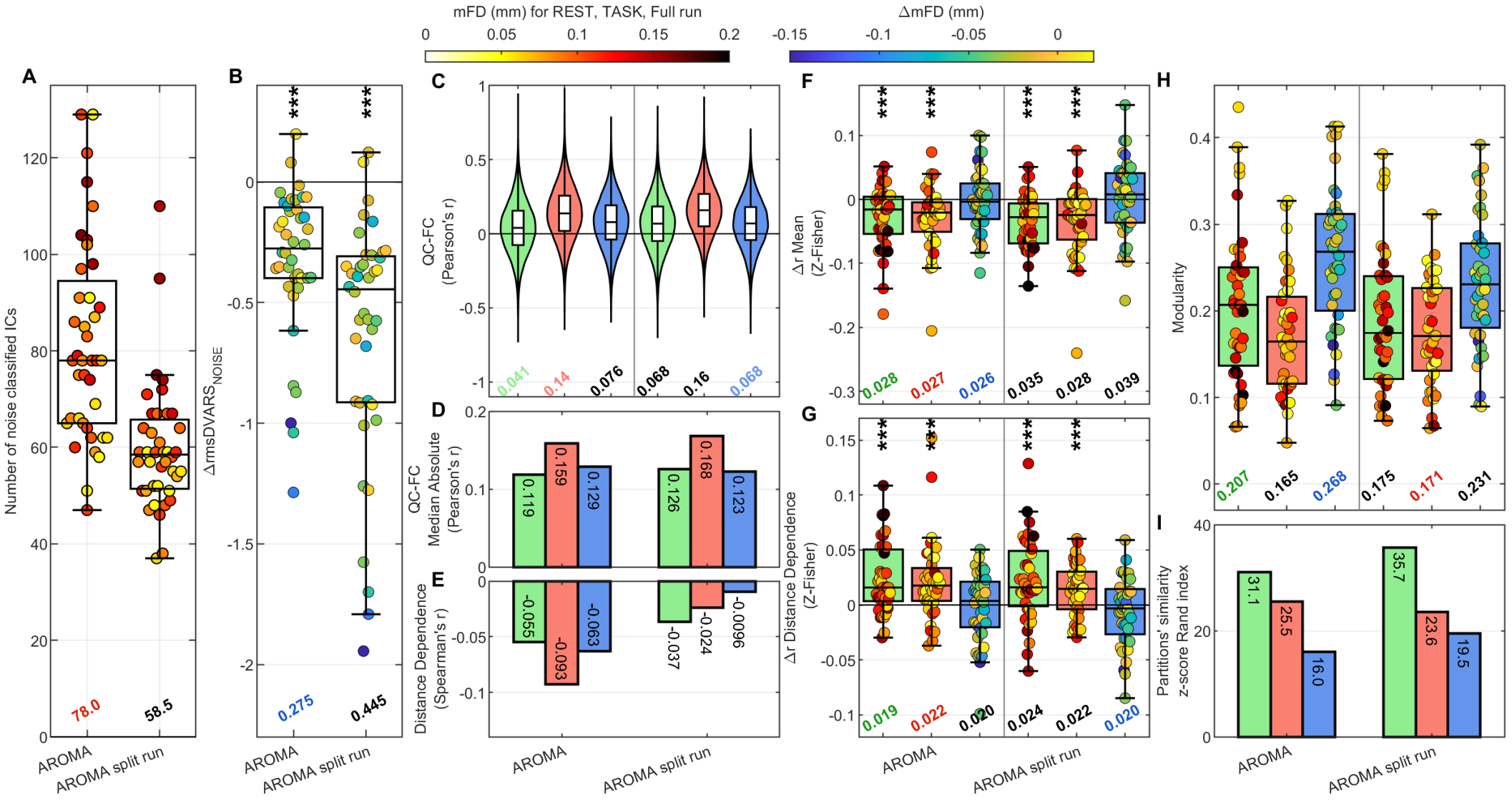

Figure S12. Comparison of ICA-AROMA performed on the whole run (AROMA) vs performed on each epoch separately (AROMA split run) for the CF dataset. Splitting the run before applying ICA-AROMA yields a lower number of total (i.e., summing the 4 epochs) noise-classified components (panel A).

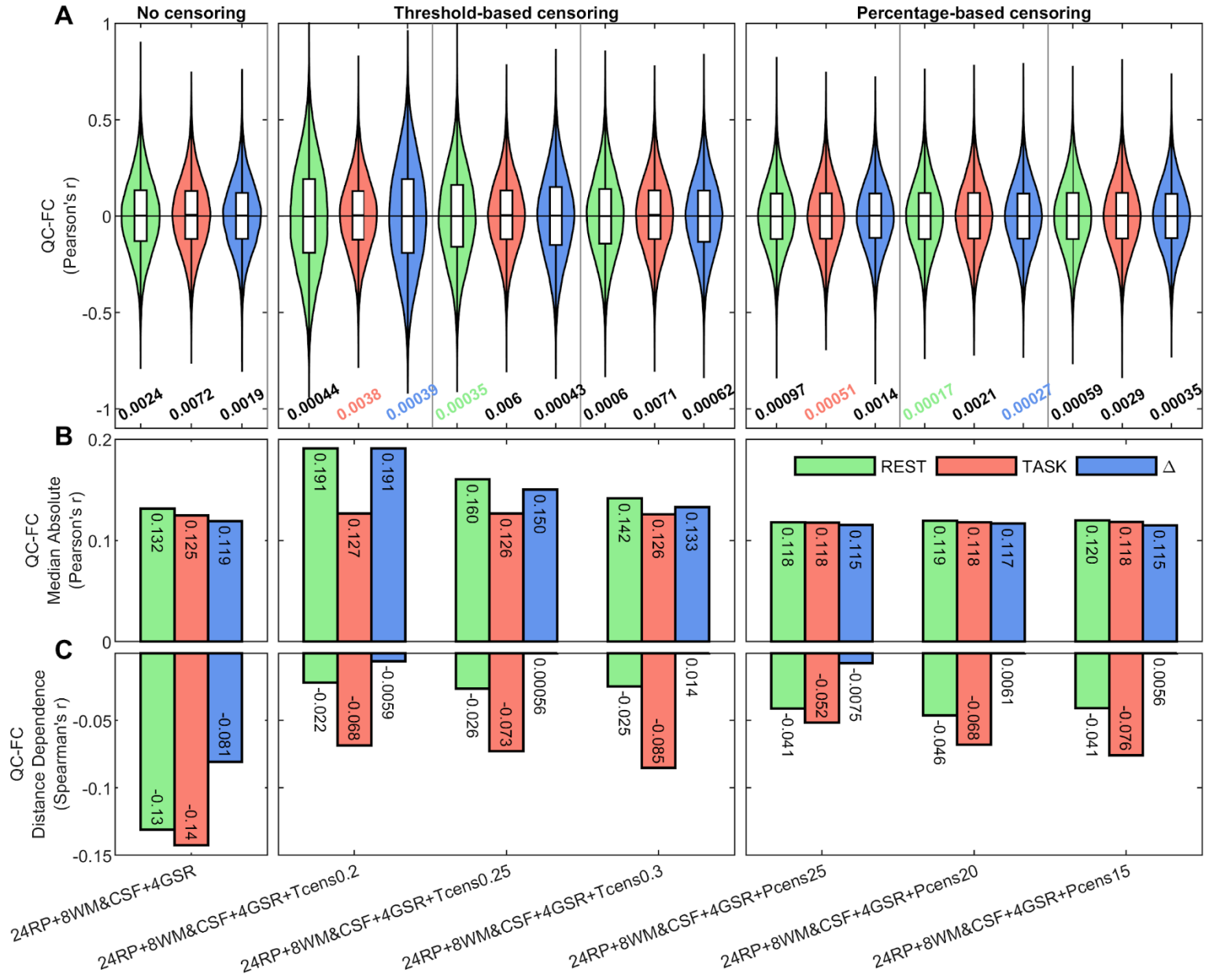

Figure S13. QC-FC plots as a function of different censoring thresholds using CF data. The metrics are described in Fig. 4 of the main text.