Title: Pipeline for 4D whole-brain microscopy preprocessing and analysis.

Authors: Bonapersona, V.[1*], Schuler, H.[1,2], Damsteegt, R. D.[1], Adolfs, Y.[1], Pasterkamp, R.J.[1], van den Heuvel[3], M. P., Joëls, M[1,4^], Sarabdjitsingh[1^], R.A.

Affiliations:

[1] Dept. Translational Neuroscience, UMC Utrecht Brain Center, Utrecht University, Utrecht, The Netherlands.

[2] Integrated Program in Neuroscience, McGill University, Montréal, QC, Canada.

[3] Connectome Lab, Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands.

[4] University Medical Center Groningen, Groningen, The Netherlands.

[^] Contributed equally

[*] Corresponding author: v.bonapersona-2@umcutrecht.nl

Abstract

Whole-brain microscopy allows for high-resolution imaging of intact mouse brains. It is a promising technique to relate behaviour to underlying brain activity, for example when combined with staining of immediate early genes. Multiple tools have been developed to transform images to machine-readable numbers, but these do not address the statistical requirements for robust data analysis. We present a novel pre-processing and analytical pipeline to analyze whole-brain data in 4 dimensions, from macro- to micro-scale and over time. This practical pipeline guides researchers through the appropriate cleaning, normalization and transformation steps for each analysis type. We developed analyses at different spatial resolutions, from macroscale functional networks to the single cell. We validate and showcase the advantages of the method by analysing activity changes at various time-points after foot-shock in male mice. All data can be visualized in our interactive web-portal (https://utrecht-university.shinyapps.io/brain_after_footshock/). The pipeline is available as R-package, *abc4d*.

Introduction

Linking behavior to the underlying neuronal substrate is not trivial. Over the past decades, several experimental techniques have been developed to enable such link. Typically, these techniques are a trade-off between temporal and spatial resolution. In rodents for example, both functional magnetic resonance imaging (fMRI) and whole-brain microscopy can be used to delineate activity throughout the whole brain. Of the two, whole-brain microscopy has excellent spatial resolution (~5µm, (Renier et al., 2016)) yet very poor time resolution, usually confined to a single time-point. An appealing possibility would be to improve the temporal resolution of whole-brain microscopy by changes in the experimental design. This can be achieved by using immediate early genes (IEG) (Guzowski et al., 2005), such as c-fos (Supplementary Note 1). c-fos has been extensively used as a post-mortem marker of cellular activity (Kovács, 1998). Previous studies report that c-fos mRNA can peak at different times across brain areas after swim or restraint stress (Cullinan et al., 1995). This suggests that there may be multiple waves of c-fos activation throughout the brain, which could be used to map the temporal dynamics across all brain areas up to hours after the initial stimulation.

Whole-brain IEG microscopy has therefore the great potential of establishing itself as a translational technique to investigate the patterns of brain activation after an acute experience, throughout the whole-brain and over time. The experimental designs can be easily adapted to this purpose. The process of converting whole-brain images to analyzable data can be divided into three consecutive parts: 1) transforming images into machine-readable numbers (i.e., data), 2) cleaning the data, and 3) transforming the data specifically for distinct analyses. During the first step, active cells (e.g. c-fos+) are detected and annotated to brain areas via segmentation and alignment. During the second step, noisy data (e.g. unspecific antibody binding, halo around brains and ventricles) should be removed, and missing values should be imputed (e.g. damaged areas). The third step varies

depending on the type of analysis that will be performed. Here, batch effects should be corrected, and the data could be normalized (e.g. to control for the size of the brain area) and transformed (e.g. log transformation, scaling).

Over the past decade, several tools have been created to transform images into interpretable numbers (i.e., Step 1, for an excellent review of open source tools, see (Tyson & Margrie, 2021)). Although most of these tools offer built-in options for visualization, to date no study has formalized the entire pipeline for preprocessing (i.e., Step 2-3) and analysis of active cells across the whole brain.

Here, we present a preprocessing and analytical pipeline for whole-brain microscopy data. The pipeline can guide researchers through general data cleaning steps, as well as transformations related to specific analyses. This pipeline can be applied independently of the type of software used for alignment and annotation (Step 1), and it can analyze 3D (i.e. whole-brain) or 4D (over time) experimental designs. Furthermore, to get the best out of whole-brain data, we highlight analyses from the macro- (functional networks) to micro- (single cell resolution) scales. We showcase advantages of this approach on data related to foot-shock exposure in adult male mice using c-fos as an activity marker. Notably, we validated long-standing observations of the stress system and substantiated new hypotheses – all within one well-controlled experiment. All data can be interactively analyzed with an open-source user interface (https:// utrecht-university.shinyapps.io/brain_after_footshock/). We developed the R package *abc4d* (“Analysis Brain Cellular activation in 4 Dimensions”) to facilitate the implementation of the pipeline.

Results

Overview of the pipeline

Figure 1 summarizes the main features of the pipeline including experimental procedure, image processing (Step 1), data cleaning (Step 2), data pre-processing (Step 3) and analysis.

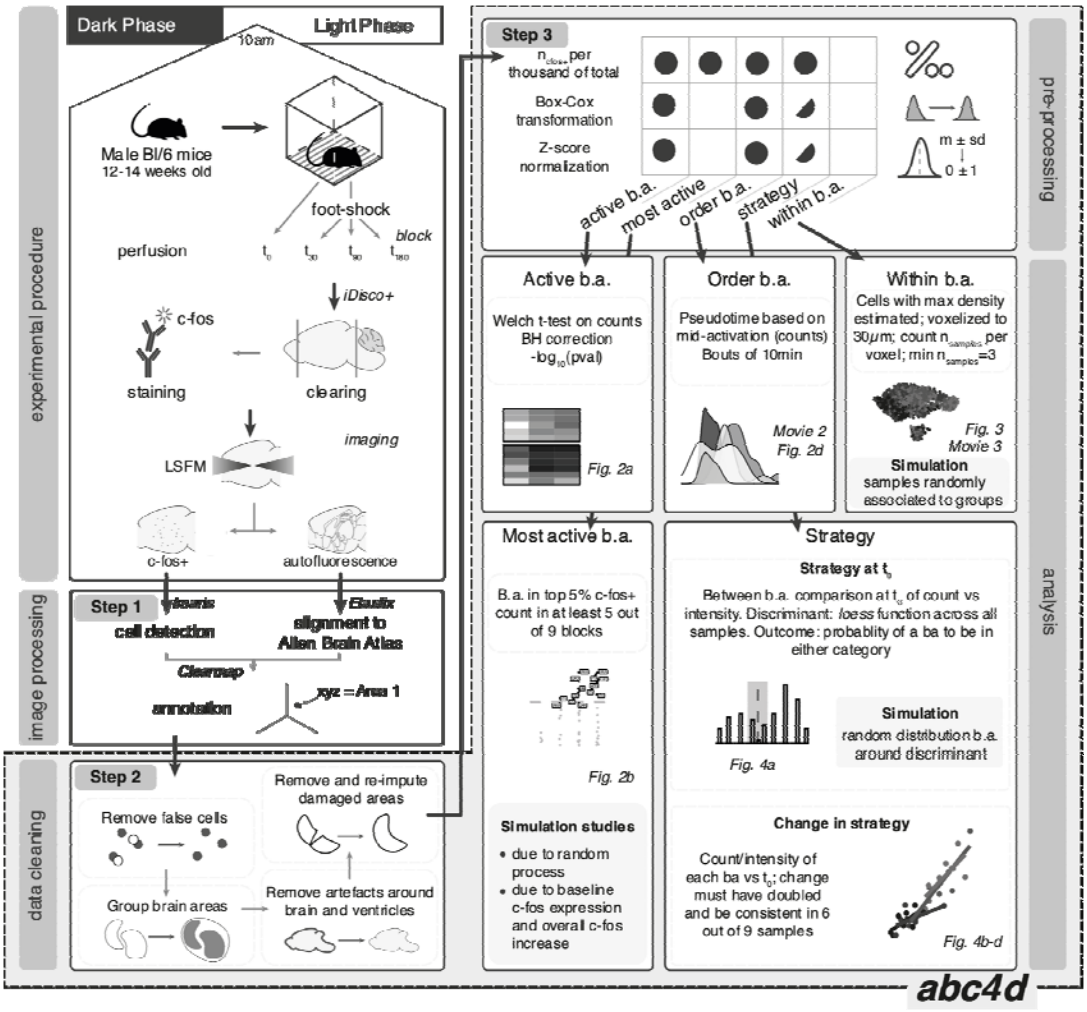

Figure 1. Schematic overview of the pipeline. Animals perfused at different time-points ($n_{animals} = n_{time\ point} * n_{blocks} = 4*9 = 36$) after foot-shock. Whole-brain samples processed with iDisco+ protocol; c-fos+ cells imaged with light-sheet fluorescent microscopy. Cells detected with Imaris and annotated to Allen Brain Atlas with Clearmap. Output is xyz coordinates per cell. Quality control (data cleaning) consists of removal of various artefacts and of grouping brain areas (b.a.) to the spatial resolution of interest. Data preprocessing of b.a. for each of the analyses. Circle: step required; half-circle: step recommended but not required. Strategy refers to t_0 strategy categorization, as well as change of strategy over time. Summary of the analyses conducted, and the main statistical decisions made. Data cleaning (Step 2), pre-processing (Step 3) steps as well as analyses are explained in detail in the Methods section and have been implemented in the *abc4d* package.

Quality control

Concerning image processing (Step 1), cell detection was performed with Imaris's spot object (Supplementary Figure 1 a), after which it was aligned to the Allen Brain Reference Atlas (ABA) with Clearmap. Precision of alignment was assessed by comparing how sample images and template images would distort landmarks which were previously manually placed (Supplementary Figure 1 b). The average absolute difference was 8.39 ± 5.88 µm (mean ± SD) in the horizontal plane and 10.92 ± 12.44 µm (mean ± SD, maximal displacement = 23.36 µm) in the sagittal plane, with more laterally placed landmarks being less precise. Alignment did not differ per condition, suggesting that alignment error should not affect our results.

During data cleaning (Step 2), unspecific binding was mitigated by removing background signal and applying a mask of 3 voxels (~75 µm) around the brain and ventricles (Movie 1), and by removing cells with abnormally high intensity ($n_{\text{cells removed}}$ = 12). The background and mask step accounted for ~97% of the removed cells. Across all samples, ~5% of the brain areas showed some form of damage; these were removed from the analyses and re-imputed. Ultimately, the number of cells removed during the quality control procedure did not differ across groups (Supplementary Figure 2).

The next step in the pipeline is data pre-processing (Step 3). As summarized in the figure (Figure 1 upper left panel), data-preprocessing is specific for each analysis type. Of note, we used a block design, meaning that a 'block' (i.e., from the same cage, one animal for each time-point processed simultaneously) was the experimental unit of randomization and processing of samples. This type of design is essential for effective batch effect correction. Next, we illustrate the potential of analyses in the pipeline that can be carried out with thus obtained datasets for the situation of activation following foot-shock exposure of adult male mice.

c-fos+ expression is increased throughout the brain, but with spatial and temporal specificity

The total number of c-fos+ cells (n_{cfos+}) across the brain was increased 30 minutes (t_{30}) after foot-shock induction and remained elevated at t_{90} and t_{180}. It returned to t_0 levels 300 minutes after foot-shock (Supplementary Figure 3). Across batches, n_{cfos+} was comparable to that of previous literature (Kim et al., 2016; Renier et al., 2016). Of note, t_0 animals were placed in a foot-shock chamber but did not receive a foot-shock. As a consequence, they should be considered as a “mildly stressed” (novelty stressor) group rather than true baseline controls.

To test the extent of c-fos+ expression throughout the brain, we performed pairwise comparisons (Welch t-test, one sided) between each foot-shock time point (t_{30}, t_{90}, t_{180}) and t_0 (Figure 2A). 86 out of the 89 brain areas had a significant increase in c-fos+ cells in at least one of the time points. Only three brain areas were not significant, i.e. medial preoptic nucleus, ventral anterior-lateral complex of the thalamus, and ventro-posterior complex of the thalamus. The time point t_{180} had the highest number of significant brain areas ($n_{sig\ brain\ areas}$ = 85), followed by t_{90} ($n_{sig\ brain\ areas}$ = 79) and t_{30} ($n_{sig\ brain\ areas}$ = 40). The effect sizes (g±SD) ranged between -0.32±0.23 (Midbrain raphe nuclei, t_{30} vs t_0) and 5.17±0.96 (Subiculum, t_{90} vs t_0) with a mean 1.62±0.57.

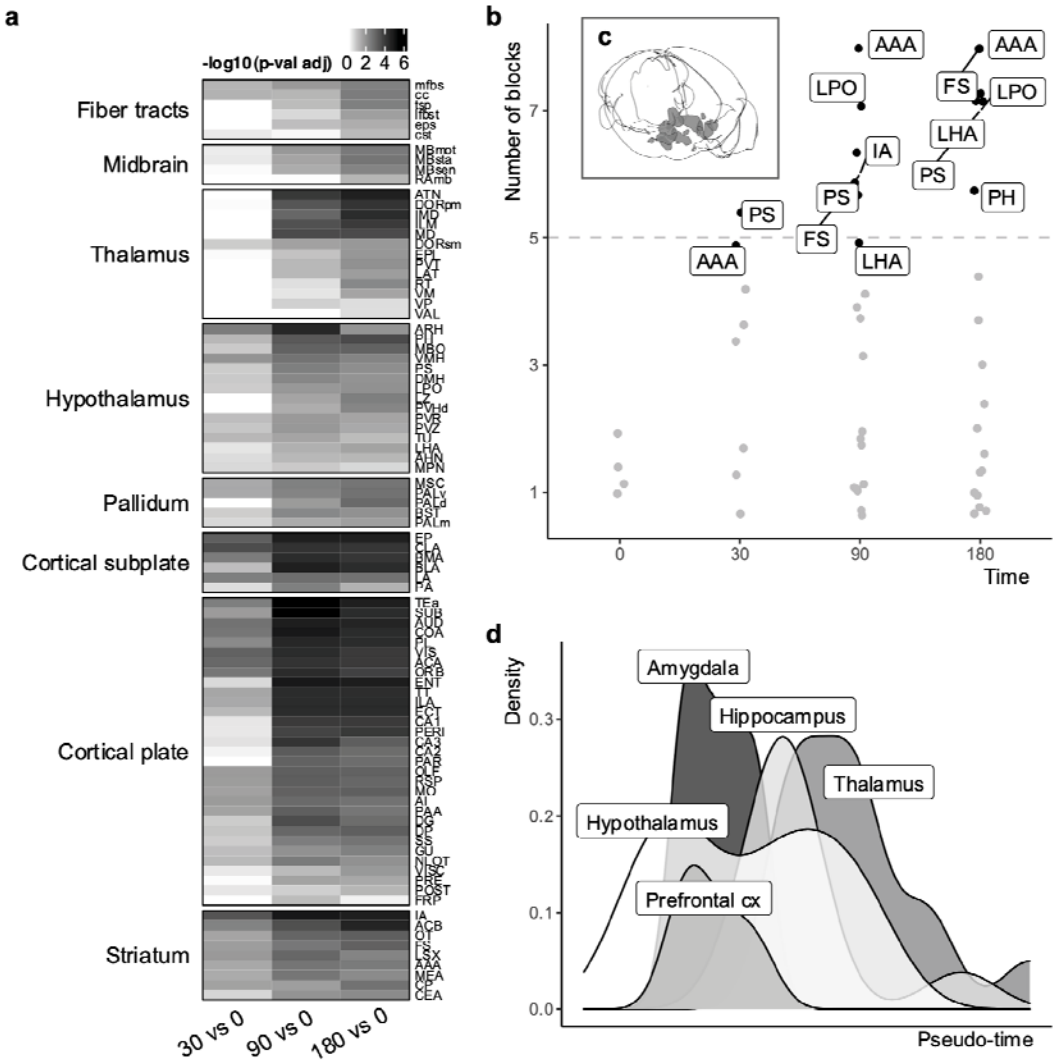

Figure 2. Most brain areas are activated by foot-shock, with spatial and temporal specificity. a) Heatmap of -$\log_{10}$ p-values derived from pairwise comparisons of each foot-shock time point (t_{30}, t_{90}, t_{180}) against t_0 for each brain area. White: p-val_{adj} >=0.05; $grey_{shades}$: p-val_{adj} < 0.05. The legend numbers correspond to the log values. b) A set of hypothalamic areas is consistently found as having the highest (top 5% of the distribution) number of c-fos+ cells (per thousand of total). The criterion for consistency was 5 out of 9 samples. Abbreviations are explained in Suppl. Table 2. c) Cartoon of the brain areas identified in b; created with *brainrender* (Claudi et al., 2020) d). Functional order of brain areas' c-fos activity following foot-shock. Brain areas were ordered based on a pseudo-time depending on c-fos+ activation across the time points, and grouped based on functional categorization important for the stress response (Henckens et al., 2015). Hypothalamic areas are the first to reach the mid-point of their activation, followed by amygdalar, prefrontal, hippocampal and lastly thalamic areas. For an interactive visualization of the single brain areas rather than the categorization, see Movie 2.

Since most brain areas (86 out of 89) were active in at least one time point, we aimed to identify whether some brain areas were more active than others. To answer this question, we identified for each block (i.e., a unique set of each time-point) the brain areas that had the highest (i.e., top 5% of the distribution) c-fos+ cell count (per thousand of total, $n_{cfos+/tot}$). Under random circumstances, the same brain area would be in the top 5% in at least 5 out of 9 samples in about 1% of the cases, as illustrated by a simulation study (Methods). Being selected by at least 5 samples was therefore used as a criterion to define consistency of highly active brain areas. In our experimental data, the criterion was met by 8 brain areas (Figure 2 b), which belonged mostly to the hypothalamus (Figure 2 c). This number ($n_{highly\ active} = 8$) was much higher than the 1% expected ($n_{randomly\ active} = 1$) by sheer randomness. With a simulation study, we confirmed that this activation could also not be attributed to the spatial localization of c-fos throughout the brain, as reported by the Allen Brain Atlas (Supplementary Figure 4).

Next, we hypothesized that although in most brain areas the number of c-fos+ cells is increased after foot shock, the peak of activation would not occur at the same time for every brain area. Brain areas are expected to be involved at different stages (and therefore at different times) of the stress response, as earlier proposed based on human fMRI studies (Hermans et al., 2014). We organized brain areas on a pseudo-time scale, based on the point in which a brain area (median across blocks) would reach the middle of its activation. This pseudo-time should only be interpreted relatively. We visualized the order of activation (Figure 2 d) of the brain areas by following a functional categorization valuable to the stress response (Henckens et al., 2015). Based on this classification, hypothalamic areas were found to be activated first, followed by amygdalar and prefrontal, hippocampal, and finally thalamic areas. Movie 2 shows a visualization of each brain area over time.

Time-dependent wave of activation within the Basolateral Amygdala

The resolution of 3D microscopy stretches as far as the single cell level, which is a great advantage over e.g. fMRI studies. To illustrate this advantage, we focused on the basolateral amygdala (BLA) since this area is key for the cognitive processing of a foot-shock (Silva et al., 2016) and stressful conditions in general (Sharp, 2017). We hypothesized that the increase of c-fos+ cells was not uniform across the BLA, but that it may be restricted to different sub-parts of the brain area. For each sample independently, we identified the most densely activated part of the brain area, i.e., the part with the highest number of c-fos+ cells. When all samples are considered, there are obvious regional distributions across time points (Figure 3 a, Movie 3), which are not evident when samples are randomly associated to the experimental groups (Supplementary Figure 5 a).

We voxelized the xyz coordinates (voxel size: approx. 30µm x 30µm x 30µm) and visualized per time point which voxels have at least one cell from three different samples. As shown by Figure 3 b, at t_{90} and t_{180}, the highest density of c-fos+ cells in the BLA was found to be more posterior (difference of ~96µm from anterior to posterior, 23% of BLA) than at t_0 and t_{30}.

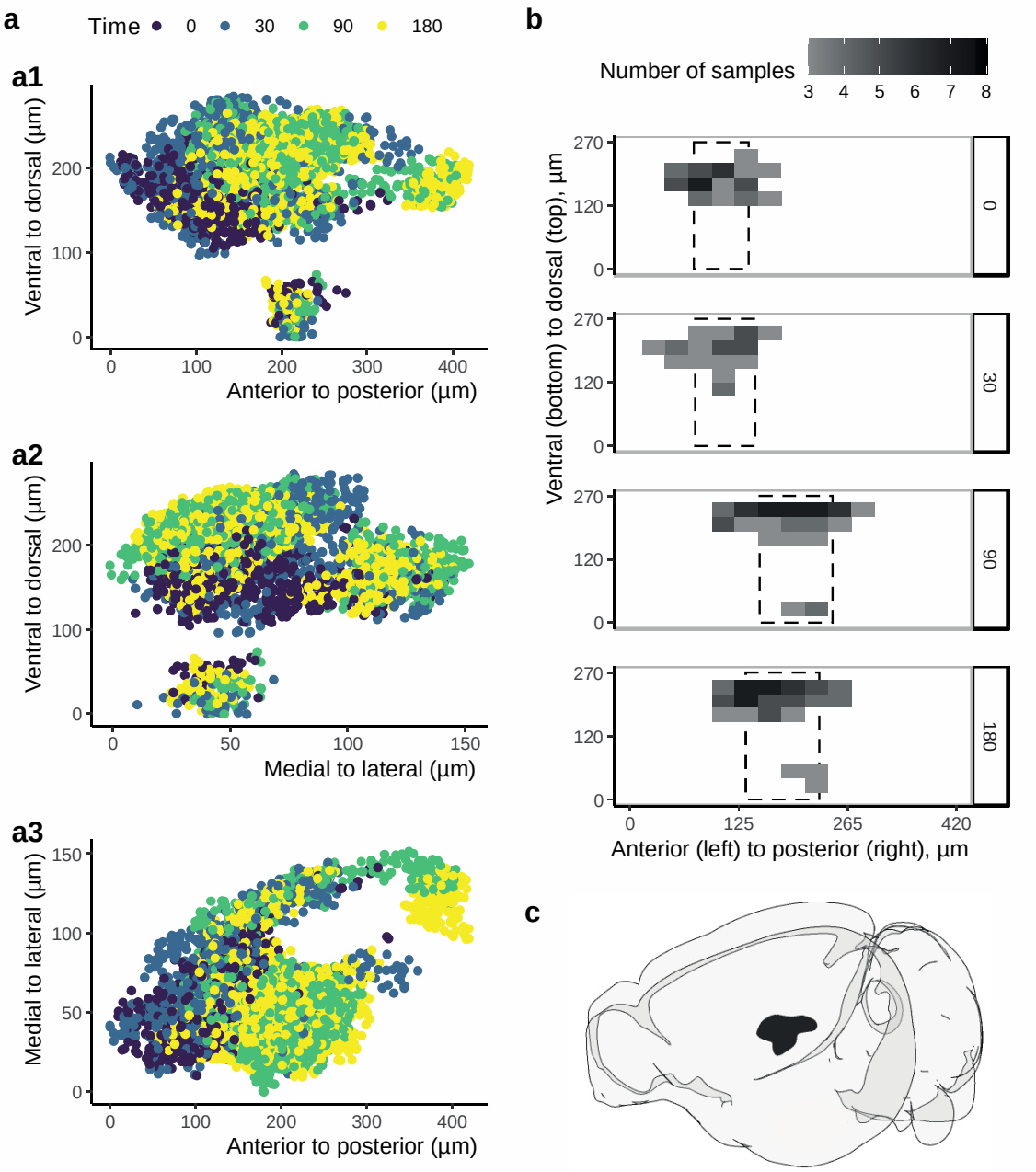

Figure 3. Changes in c-fos+ cell density within the Basolateral Amygdala. a) Cells in high density regions of the right BLA. The 3D cell coordinates are represented as a set of three 2D graphs, one for each couple of coordinates (xy, yz, xz). Each dot is a cell of a sample in a region with highest density. The colors refer to the different time points. b) The densest c-fos+ sub-part of the BLA moved from more anterior (t_0 and t_{30}) to more posterior (t_{90} and t_{180}). The BLA has been voxelized (voxel size 30x30x30µm), and the fill color refers to the number of samples with at least one cell in that voxel. The dashed box indicates the mean and SD per group along the posterior-anterior axis. c) Cartoon visualization of the right BLA in the same orientation of a1 and b, created with *brainrender* (Claudi et al., 2020).

Brain areas use different strategies of activation, which can change after stress

With 3D microscopy, one can not only count the $n_{c\text{-}fos+}$, but also quantify the intensity of c-fos staining of each cell. Among other parameters, a cell is considered c-fos+ if the intensity of c-fos is higher than the background (i.e., signal-to-noise ratio). As a consequence, one would expect the $n_{c\text{-}fos+}$ per brain area to strongly correlate with the average c-fos intensity. However, this was not the case (Supplementary Figure 5 b). We hypothesized that different brain areas have a preferential strategy of activation: a few very active (i.e. low count, high intensity) cells, versus many lowly active cells (i.e. low intensity, high count). Intensity and count are therefore expected to be related within brain areas, rather than across the whole brain.

We categorized the brain areas of each t_0 sample based on their strategy, i.e. their preference for increasing in count or intensity. Across samples, we then calculated the probability of each brain area to belong to either categorizations (Figure 4 a). The results showed a bimodal distribution (Figure 4 a), which is clearly different than the normal distribution expected under the null hypothesis (Supplementary Figure 5 c). Therefore, whether a brain area is activated by increasing n_{cfos+} or by increasing the average intensity of c-fos is unlikely to be the result of a random process.

Next, we examined whether the strategy of a brain area changes after foot-shock, relatively to t_0. For each foot-shock time-point (t_{30}, t_{90}, t_{180}), we selected brain areas with a consistent (at least 6 out of 9 samples) increase or decrease in both count and intensity, and calculated to what extent count and intensity were increasing compared to each other. We categorized brain areas as “changing strategy” if they at least doubled the increase in one category. 43 brain areas met these criteria (Figure 4b). Of these, 30 increased activation by means of intensity rather than c-fos+ cell count, especially in the amygdala, hypothalamus and thalamus. Of note, the increase in intensity for the amygdalar nuclei was present only for the

time points t_{90} and t_{180}. Figure 4 c displays two brain areas (BLA and Subiculum) as representative examples of activation towards intensity and count, respectively. We also added a visualization of the dentate gyrus (Figure 4 c, part 3) as a validation from literature.

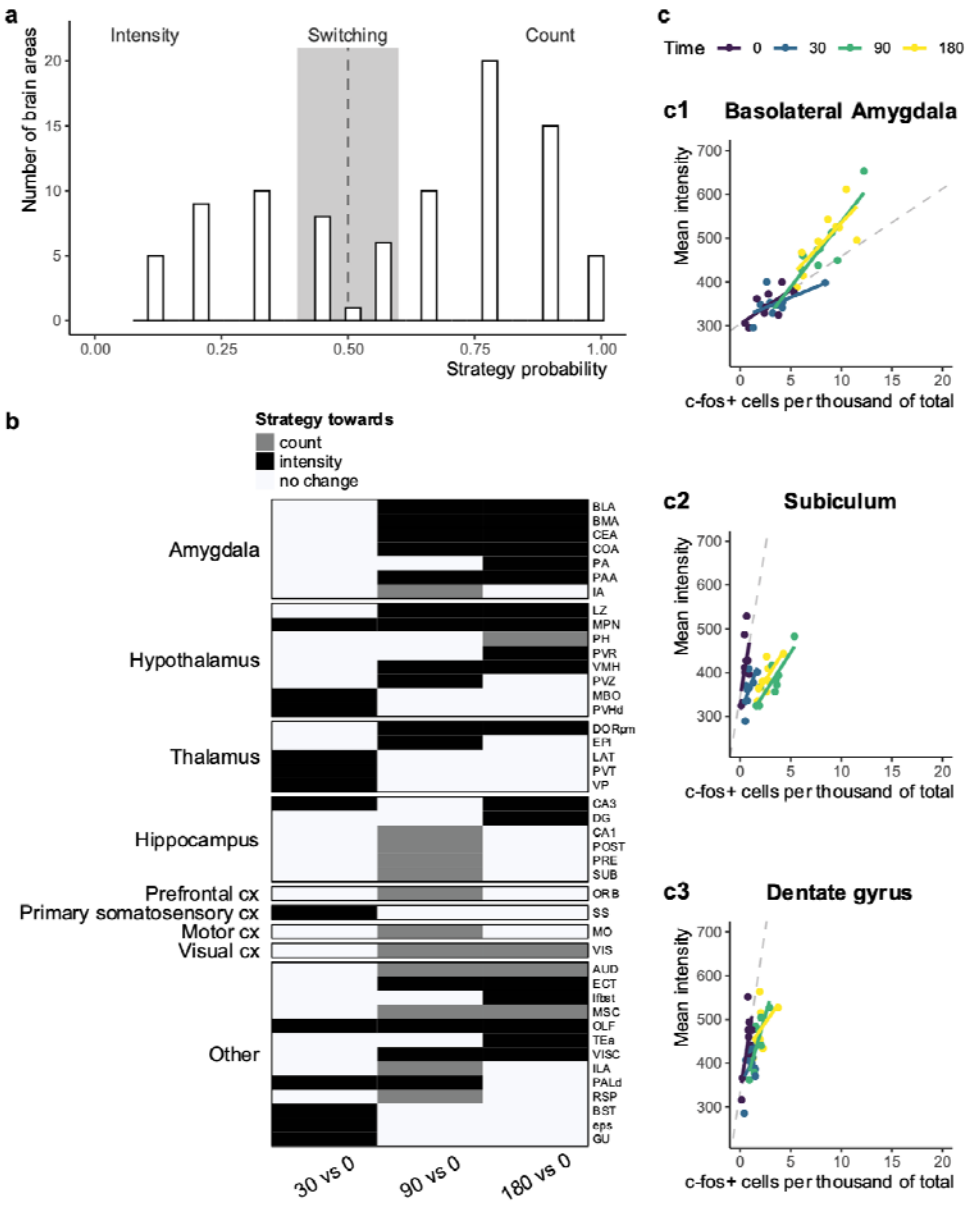

Figure 4. Brain areas have a preferential strategy of activation, which may change after foot-shock. a) Preferential strategy of brain areas based on t_0 data, i.e. the relationship between intensity and count per brain area. Histogram of the number of brain areas across the strategy probability. If a brain area would be activated by indiscriminately increasing c-fos+ cells and expression, the distribution would be Normal (Supplementary Figure 5c), with μ around 0.5. The bimodal distribution suggests that certain brain areas prefer increasing the number of c-fos+ cells (probability > 0.5, count), whereas other increase the mean c-fos expression (probability < 0.5, intensity). b) The strategy of brain areas can change after foot-shock. Binary heatmap of how strategy can change across brain areas for pairwise comparisons of time points. White corresponds to no change in strategy, black to a change towards intensity and grey to a change towards count. Of note, a change towards intensity does not necessarily mean that the brain area does not increase count; rather, it means that the increase in intensity cannot be explained by the increase in count alone. c) Representative examples of a brain area that after foot-shock changes strategy towards intensity (BLA, c1) or count (Subiculum, c2). The dentate gyrus (c3) was added for literature validation (see

Discussion). Each dot represents one sample; the line represents that correlation between count and intensity per group. Dashed line represents what one would expect based on t_0 activation.

Discussion

One of the biggest challenges in neuroscience is to understand how behavior emerges from the activation of brain circuits. The development of functional MRI (fMRI) in particular has greatly expanded the possibilities of connecting brain activity to behavior. However, spatial resolution of fMRI is rather limited (mm scale), especially when overall brain size is small, as in mice. Moreover, fMRI is less suitable for the study of circuits that are acutely activated in mice, such as stress or fear. The use of anesthetics hampers the activation of stress circuits, while results in awake animals are likely confounded by the stress of the condition or the extensive habituation to the experimental set-up. On the other end of the technical spectrum, whole-brain microscopy offers high spatial resolution, but it is usually confined to a single timepoint. To improve the latter approach, we investigated a functional outcome over time, by staining for the IEG c-fos and introducing a pseudo-time metric. With our developed analytical methodology, we could investigate how activity dynamically changes across the entire mouse brain, from the circuit to the single cell level. As an example, we presented the specific case of foot-shock. To dive deeper into any area of interest, an interactive interface on the data is available at our web portal (https://utrecht-university.shinyapps.io/brain_after_footshock/). We propose this approach as a translational technique to aid bridging the gap between conventional animal experiments and functional outcomes of human studies. This approach has the added value to confirm previous research and develop new hypotheses at different spatial resolutions, all within one single experiment. The major advance of the current report is that we provide researchers with an easy-to-use preprocessing and analytical pipeline to apply this technology. The pipeline is available in the R package *abc4d*, to which new improvements e.g. in terms of cell identification can be easily added in future. *Abc4d* also includes a framework of simulation studies, where the null hypotheses for different analyses can be investigated. An overview of the package is provided in the cheat-sheet (Supplementary Figure 6).

Data cleaning and preprocessing are essential for any data analysis. Several tools have been developed to transform images into machine-readable data (for example, (Fürth et al., 2018; Renier et al., 2016; Tyson et al., 2021)), but no publication has so far thoroughly explored the subsequent steps of data analysis, i.e. dealing with missing values, batch effect corrections, normalization and transformation. These share commonalities to other big-data analyses (e.g. omics). For example, whole-brain data is highly dimensional such as in genomics, although to a different order of magnitude. The number of brain areas is in the dozens, whereas the number of genes is in the thousands. There are also obvious differences. For example, for normalization, the size of brain areas is known, whereas gene expression is not related to the size of a transcript. Concerning function, several gene ontologies have been developed, whereas this resource is not available for brain areas. Therefore, preprocessing steps require to be specifically tailored to the characteristics of whole-brain data. Figure 1 summarizes which steps to take for each analysis, in order to comply with the statistical assumptions. Importantly, so far analyses have been limited to a region-based approach, where the number of active cells is calculated for each brain area separately. Voxel-based analyses similar to MRI have only recently been developed (Vandenberghe et al., 2018). Here, we take this one step further, and suggest analyses for *i)* the time dynamic and *ii)* the single cell level. The resolution up to the level of single cells is certainly one of the major advantages of whole-brain microscopy. In the future, c-fos+ cells could be characterized in more detail, being able to distinguish excitatory from inhibitory neurons, or neurons from glia cells, as the current findings confirm that c-fos staining is not confined to neurons only (e.g. (Condorelli et al., 1989)). After technical development, this could be achieved by multiple concurrent stainings, or by computationally categorizing cellular morphology (Tyson & Margrie, 2021). However, even in the absence of this follow-up characterization, the current methodology can be helpful in generating new hypotheses about cellular responses.

To illustrate the strength of a single cell approach, we refer to the in-depth analysis of the basolateral amygdala – an area highly relevant for the cognitive processing of acutely stressful situations like a foot-shock (Sharp, 2017; Silva et al., 2016) . We show a clear shift after foot-shock from activation of cells in the lateral-anterior part towards a more posterior-medial subset of cells. This type of insight would not be possible with fMRI, even using high-field (e.g. 9.4 T) magnetic resonance. Alongside a recent tracing study (O'Leary et al., 2020), one could speculate that this sub-regional difference may be linked to the involvement of different projections, specifically from the prefrontal cortex and nucleus accumbens. Such hypotheses could be investigated in the future either computationally or experimentally.

Another example of a new research avenue is the observation that areas use a distinct strategy in c-fos cell activation under mildly stressful conditions such as a novel environment (in this case the shock-box at t_0) and after exposure to a very stressful situation like an inescapable foot-shock. Interestingly, the dentate gyrus is one of the areas that stands out as an 'intensity-prone' area, meaning it shows a limited number of c-fos+ cells with high intensity. This confirms earlier findings after a swim stress, revealing DNA demethylation at specific CpG sites close to the c-fos transcriptional start site, in the gene promoter region of early growth response protein 1, specifically in a small subset of dentate neurons (Chandramohan et al., 2008; Saunderson et al., 2016). Earlier studies showed that the expression of c-fos is proportional to the rate of firing of the cell (Fields et al., 1997). If these observations are considered together, one could hypothesize that in certain brain areas many cells are slightly more activated (i.e. express c-fos above detection threshold), whereas other brain areas may have only a few cells that are strongly excited, which would lead to an increase in their c-fos intensity. Of note, especially in amygdalar areas the increase in intensity was observed only at 90 and 180 minutes, suggesting a glucocorticoid-dependent mechanism. The functional relevance of these findings clearly requires follow-up.

Next to these important technical advances illustrated for the case of brain activation after a single foot shock, there are some limitations to consider. The choice of c-fos as an activity marker is arguably appropriate in case of acute stress exposure (Kovács, 1998), but it is by no means the only IEG one could choose for the current approach. Other markers of cell activity may afford additional insights into the circuits being activated after stress. For example, several other markers of cell activity such as *Arc* and *Egr1* were transcriptionally activated following acute stress in a multi-omics approach (Ziegler et al., 2021). Ideally, a combination of markers should be applied to get a completer view. Another technical limitation is linked to the size restraints of imaging with light-sheet microscopy. In our study, we trimmed the most frontal and most caudal parts of the brains. Although this does not impact the methodology we present, brain areas involved in the acute stress response (e.g. locus coeruleus) are missing. A third consideration concerns the pseudo-time approach. On the one hand, it overcomes the absence of high frequency sampling (as possible with fMRI). On the other hand, it is a mere approximation of real-time processes. Interestingly, at the macro-scale, the order in which brain regions are activated following the acute stress of a foot-shock strongly resembles that seen with fMRI after acute stress in the human brain (Hermans et al., 2014). This lends credibility to our approach and also emphasizes its translational potential. Importantly though, while it might be an acceptable approximation of the activation phase, the method gives little insight in the gradual turning off of brain areas, since this also depends on the half-life time of the c-fos protein. The half-life time may differ across brain areas (Kovács, 1998), something that could be investigated with a meta-analytic approach.

Despite some limitations, future researchers interested in event-related brain activation, from the network to the single cell level, will be greatly served by the ready-to-use pipeline for 4D

immunohistochemical whole-brain analysis presented in this report, and supported by a new R package.

Acknowledgments:

We would like to acknowledge the MIND facility (UMC Utrecht Brain Center) for 3D imaging by lightsheet microscopy (mindresearchfacility.nl), and Roger Koot for the IT infrastructure. We would like to thank Kevin Kemna for suggesting the creation of the R package and the helpful feedback on the analyses. This work was supported by the Consortium of Individual Development (CID) and BRAINSCAPES, which are funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant numbers 024.001.003 and 024.004.012).

Author contributions:

V.B. contributed with the conceptualization, data curation, analysis, investigation, methodology, software, visualization and writing the manuscript; H.S. contributed with the methodology, performance of experiments, analysis and reviewing the manuscript; R.D. contributed with the performance of experiments; Y.A. contributed with support with experiments and reviewing the manuscript; R.J.P. provided infrastructure, contributed with funding acquisition and reviewing the manuscript; M.P.vdH. contributed with the conceptualization of the project and reviewing the manuscript; M.J. contributed with the conceptualization, funding acquisition, project administration, supervision, and writing the manuscript; R.A.S. contributed with the conceptualization, project administration, supervision and writing the manuscript.

Competing interests:

The authors have no competing interests as defined by Nature Research, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

References (up to 40 / 50 depending on type)

Chandramohan, Y., Droste, S. K., Arthur, J. S. C., & Reul, J. M. H. M. (2008). The forced swimming-induced behavioural immobility response involves histone H3 phospho-acetylation and c-Fos induction in dentate gyrus granule neurons via activation of the N-methyl-d-aspartate/extracellular signal-regulated kinase/mitogen- and stress-activated kinase signalling pathway. *European Journal of Neuroscience*, *27*(10), 2701–2713. https://doi.org/10.1111/j.1460-9568.2008.06230.x

Claudi, F., Tyson, A. L., Petrucco, L., Margrie, T. W., Portugues, R., & Branco, T. (2020). *Brainrender: A python-based software for visualizing anatomically registered data* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.02.23.961748

Condorelli, D. F., Kaczmarek, L., Nicoletti, F., Arcidiacono, A., Dell'Albani, P., Ingrao, F., Magrì, G., Malaguarnera, L., Avola, R., Messina, A., & Stella, A. M. G. (1989). Induction of protooncogene fos by extracellular signals in primary glial cell cultures. *Journal of Neuroscience Research*, *23*(2), 234–239. https://doi.org/10.1002/jnr.490230214

Cullinan, W. E., Herman, J. P., Battaglia, D. F., Akil, H., & Watson, S. J. (1995). Pattern and time course of immediate early gene expression in rat brain following acute stress. *Neuroscience*, *64*(2), 477–505. https://doi.org/10.1016/0306-4522(94)00355-9

Erö, C., Gewaltig, M.-O., Keller, D., & Markram, H. (2018). A Cell Atlas for the Mouse Brain. *Frontiers in Neuroinformatics*, *12*. https://doi.org/10.3389/fninf.2018.00084

Fields, R. D., Eshete, F., Stevens, B., & Itoh, K. (1997). Action Potential-Dependent Regulation of Gene Expression: Temporal Specificity in Ca2+, cAMP-Responsive Element Binding Proteins, and Mitogen-Activated Protein Kinase Signaling. *Journal of Neuroscience*, *17*(19), 7252–7266. https://doi.org/10.1523/JNEUROSCI.17-19-07252.1997

Fürth, D., Vaissière, T., Tzortzi, O., Xuan, Y., Märtin, A., Lazaridis, I., Spigolon, G., Fisone, G., Tomer, R., Deisseroth, K., Carlén, M., Miller, C. A., Rumbaugh, G., & Meletis, K. (2018). An interactive framework for whole-brain maps at cellular resolution. *Nature Neuroscience*, *21*(1), 139–149. https://doi.org/10.1038/s41593-017-0027-7

Guzowski, J. F., Timlin, J. A., Roysam, B., McNaughton, B. L., Worley, P. F., & Barnes, C. A. (2005). Mapping behaviorally relevant neural circuits with immediate-early gene expression. *Current Opinion in Neurobiology*, *15*(5), 599–606. https://doi.org/10.1016/j.conb.2005.08.018

Henckens, M. J. A. G., van der Marel, K., van der Toorn, A., Pillai, A. G., Fernández, G., Dijkhuizen, R. M., & Joëls, M. (2015). Stress-induced alterations in large-scale functional networks of the rodent brain. *NeuroImage*, *105*, 312–322. https://doi.org/10.1016/j.neuroimage.2014.10.037

Hermans, E. J., Henckens, M. J. A. G., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences*, *37*(6), 304–314. https://doi.org/10.1016/j.tins.2014.03.006

Kim, Y., Perova, Z., Mirrione, M. M., Pradhan, K., Henn, F. A., Shea, S., Osten, P., & Li, B. (2016). Whole-Brain Mapping of Neuronal Activity in the Learned Helplessness Model of Depression. *Frontiers in Neural Circuits*, *10*. https://doi.org/10.3389/fncir.2016.00003

Kimbrough, A., Lurie, D. J., Collazo, A., Kreifeldt, M., Sidhu, H., Macedo, G. C., D'Esposito, M., Contet, C., & George, O. (2020). Brain-wide functional architecture remodeling by alcohol dependence and abstinence. *Proceedings of the National Academy of Sciences*, *117*(4), 2149–2159. https://doi.org/10.1073/pnas.1909915117

Klein, S., Staring, M., Murphy, K., Viergever, M. A., & Pluim, J. P. W. (2010). elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging*, *29*(1), 196–205. https://doi.org/10.1109/TMI.2009.2035616

Kovács, K. J. (1998). Invited review c-Fos as a transcription factor: A stressful (re)view from a functional map. *Neurochemistry International*, *33*(4), 287–297. https://doi.org/10.1016/S0197-0186(98)00023-0

McReynolds, J. R., Christianson, J. P., Blacktop, J. M., & Mantsch, J. R. (2018). What does the Fos say? Using Fos-based approaches to understand the contribution of stress to substance use disorders. *Neurobiology of Stress*, *9*, 271–285. https://doi.org/10.1016/j.ynstr.2018.05.004

Morimoto, M., Morita, N., Ozawa, H., Yokoyama, K., & Kawata, M. (1996). Distribution of glucocorticoid receptor immunoreactivity and mRNA in the rat brain: An immunohistochemical and in situ hybridization study. *Neuroscience Research*, *26*(3), 235–269. https://doi.org/10.1016/S0168-0102(96)01105-4

O'Leary, T. P., Sullivan, K. E., Wang, L., Clements, J., Lemire, A. L., & Cembrowski, M. S. (2020). Extensive and spatially variable within-cell-type heterogeneity across the basolateral amygdala. *ELife*, *9*, e59003. https://doi.org/10.7554/eLife.59003

Qian, X., Droste, S. K., Gutièrrez-Mecinas, M., Collins, A., Kersanté, F., Reul, J. M. H. M., & Linthorst, A. C. E. (2011). A Rapid Release of Corticosteroid-Binding Globulin from the Liver Restrains the Glucocorticoid Hormone Response to Acute Stress. *Endocrinology*, *152*(10), 3738–3748. https://doi.org/10.1210/en.2011-1008

Renier, N., Adams, E. L., Kirst, C., Wu, Z., Azevedo, R., Kohl, J., Autry, A. E., Kadiri, L., Umadevi Venkataraju, K., Zhou, Y., Wang, V. X., Tang, C. Y., Olsen, O., Dulac, C., Osten, P., & Tessier-Lavigne, M. (2016). Mapping of Brain Activity by Automated Volume Analysis of Immediate Early Genes. *Cell*, *165*(7), 1789–1802. https://doi.org/10.1016/j.cell.2016.05.007

Saunderson, E. A., Spiers, H., Mifsud, K. R., Gutierrez-Mecinas, M., Trollope, A. F., Shaikh, A., Mill, J., & Reul, J. M. H. M. (2016). Stress-induced gene expression and behavior are controlled by DNA methylation and methyl donor availability in the dentate gyrus. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(17), 4830–4835. https://doi.org/10.1073/pnas.1524857113

Sharp, B. M. (2017). Basolateral amygdala and stress-induced hyperexcitability affect motivated behaviors and addiction. *Translational Psychiatry*, *7*(8), e1194–e1194. https://doi.org/10.1038/tp.2017.161

Silva, B. A., Gross, C. T., & Gräff, J. (2016). The neural circuits of innate fear: Detection, integration, action, and memorization. *Learning & Memory*, *23*(10), 544–555. https://doi.org/10.1101/lm.042812.116

Tyson, A. L., & Margrie, T. W. (2021). Mesoscale microscopy for micromammals: Image analysis tools for understanding the rodent brain. *ArXiv:2102.11812 [q-Bio]*. http://arxiv.org/abs/2102.11812

Tyson, A. L., Rousseau, C. V., Niedworok, C. J., Keshavarzi, S., Tsitoura, C., Cossell, L., Strom, M., & Margrie, T. W. (2021). A deep learning algorithm for 3D cell detection in whole mouse brain image datasets. *BioRxiv*, 2020.10.21.348771. https://doi.org/10.1101/2020.10.21.348771

Vandenberghe, M. E., Souedet, N., Hérard, A.-S., Ayral, A.-M., Letronne, F., Balbastre, Y., Sadouni, E., Hantraye, P., Dhenain, M., Frouin, F., Lambert, J.-C., & Delzescaux, T. (2018). Voxel-Based Statistical Analysis of 3D Immunostained Tissue Imaging. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00754

Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., Dolbeare, T., Blanchard, B., Dee, N., Wakeman, W., Hirokawa, K. E., Szafer, A., Sunkin,

S. M., Oh, S. W., Bernard, A., … Ng, L. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*, *181*(4), 936-953.e20. https://doi.org/10.1016/j.cell.2020.04.007

Wheeler, A. L., Teixeira, C. M., Wang, A. H., Xiong, X., Kovacevic, N., Lerch, J. P., McIntosh, A. R., Parkinson, J., & Frankland, P. W. (2013). Identification of a Functional Connectome for Long-Term Fear Memory in Mice. *PLOS Computational Biology*, *9*(1), e1002853. https://doi.org/10.1371/journal.pcbi.1002853

Ziegler, L. M. von, Floriou-Servou, A., Waag, R., Gupta, R. R. D., Sturman, O., Gapp, K., Maat, C. A., Kockmann, T., Lin, H.-Y., Duss, S. N., Privitera, M., Hinte, L., Meyer, F. von, Zeilhofer, H. U., Germain, P.-L., & Bohacek, J. (2021). Molecular roadmap of the healthy stress response in the mouse hippocampus. *BioRxiv*, 2021.03.26.436970. https://doi.org/10.1101/2021.03.26.436970

Methods

Animals and husbandry

8 to 10 weeks old male C57Bl/6JOlaHsd mice were purchased from Envigo (Harlan, France) in 5 separate batches. The animals were habituated to the animal facility for at least two weeks, then tested at the age of 10 to 14 weeks. Until the experimental day, the animals were housed in groups of five in type II L cages (365x207x140mm, Tecniplast®) on a 12h dark/light cycle (light phase: 9.00AM–9.00PM), 22±2 °C, humidity at ± 64%, with *ad libitum* access to standard chow (Special Diet Services, UK, sdsdiets.com) and tap water. Experimental cages were placed on racks without a specific order and left undisturbed in the same experimental room, except for weekly cleaning by animal caretakers unfamiliar with the study design. A copy of the work-protocol (license: 527/16/644/01/06, 527/18/4806/01/01) as well as a step-by-step protocol can be found on the Open Science Framework page of the project (osf.io/8muvw). All animal procedures were approved by the Animal Ethical Committee at Utrecht University (license: AVD1150020184806), the Netherlands. Every effort was taken to minimize animal suffering in accordance with the FELASA guidelines and the Dutch regulation for housing and care of laboratory animals.

Experimental design

The primary aim of this study was to develop a methodology to be able to analyze brain-wide activation over time, specifically of the stress system. The experiments were designed to address batch effects, missing values and normalization. An overview of the experimental procedure, data cleaning, preprocessing, and analysis can be found in Figure 1.

The main experiment used a uniform block design, where each block corresponded to a separate cage of animals. Each cage contained an animal of each experimental group (being 'time after foot-shock' and control; $n_{\text{time point}} = 4$). The animals were killed at three different time points after receiving a foot-shock ($n_{\text{time point foot-shock}} = 3$). A control group underwent the same identical procedure but did not receive the foot-shock. This group is considered time point 0 min (t_0). Of note, this control group did experience a novel

environment, and its activation should therefore be considered not as 'baseline' but as mildly stressed. The time points refer to the moment of transcardial perfusion, and were chosen to comprehensively model all phases of the stress response. They cover the initiation (t = 30 min, t_{30}), maintenance (t = 90 min, t_{90}) and termination (t = 180 min, t_{180}) of the HPA axis response, while considering the required time to synthesize c-fos (Kovács, 1998) as well as the delay of 30 minutes of increase of corticosteroid in brain tissue compared to blood (Qian et al., 2011). Randomization of the experimental groups occurred within the cage, and samples of the same block were processed together both pre- and post- mortem. In other words, each experimental group was represented in every four animals (block), but the order of experimental groups differed across blocks. This type of design is important for batch effects' correction.

The following control experiments were performed. A 15 min time point ($n_{animals}$ = 3) to validate that indeed 30 min is the earliest time point to detect an increase in c-fos protein expression. A 300 min time point ($n_{animals}$ = 2) to validate that c-fos expression eventually decreases. A no primary antibody staining ($n_{animals}$ = 3) to estimate the extent of unspecific secondary antibody binding. A home-cage control group ($n_{animals}$ = 3) was used to validate that the c-fos expression of the t_0 group was due to the experimental procedure (i.e., placement into a novel cage). To minimize the sample size of animals used, the animals of these control experiments were processed together with those of the main experiments, so that the comparison groups could be reused. As a consequence, these control experiments are only qualitatively assessed, and no formal statistical analysis is performed to avoid unnecessary multiple testing. Data of the control experiments is presented paired to the t_0 group of the same batch.

Initially, we aimed to include males as well as females (see work protocol at osf.io/8muvw) in the experiment. However, the percentage of animals lost due to unforeseen circumstances (due to the antibody) was much larger than anticipated (30% instead of 10%). That is, the

secondary antibody was less stable than expected; as a consequence, the quality of the scans was often insufficient for a brain-wide quantification. We therefore simplified the experimental design and used only 9 male mice per experimental group (n_{block} = 9, $n_{time\ point}$ = 4, $n_{animals}$ = 36). These were processed in three different batches of 3, 4, and 2 blocks, respectively. We chose males instead of females because more male samples had sufficient quality after the first two batches. This qualitative assessment took place before c-fos quantification. The sample size of 9 animals per group was in line with or larger than manuscripts previously published using whole-brain microscopy (Kimbrough et al., 2020; Renier et al., 2016; Wheeler et al., 2013). Furthermore, it was sufficiently large to identify, study and mitigate batch effects. No experimental animal was excluded from the study for reasons other than insufficient staining quality. For a full description of the missing animals, see Supplementary Table 1.

Foot-shock induction

At any given experimental day, a cage was brought to the experimental room. Each animal of the cage was placed in a separate foot-shock box (floor area: 250x300mm) at the same time. Since the animals were not earmarked, selection bias was limited by randomizing the order of the shock boxes when placing the animals. In this way, we aimed to limit a "shock box" specific effect. After 60 seconds, the experimental animals received a single foot-shock (0.8 mA, 2 sec). The t_0 animal of each block was always placed in a "sham shock box". This box was identical to the others, but it did not give a shock. After another 30 seconds, all animals were removed from the shock boxes. The t_0 animals were euthanized immediately, whereas animals of the foot-shock groups were single-housed in new cages waiting for transcardial perfusion according to their specified time point. The new cages were enriched with *ad libitum* chow and water, as well as bedding material of the home cage to limit arousal due to a novel environment.

To acquire meaningful, yet above threshold c-fos expression, the optimal t_0 condition (home-cage vs novel environment) and foot-shock intensity (0.4 mA vs 0.8 mA, assessed at t = 90 min) were established in a pilot study. The results of the pilot study were only qualitatively assessed.

The investigator (RD) performing the foot-shock procedure and the perfusion was not blinded to condition, since she needed to confirm that the foot-shock was successfully applied and that animals were perfused at the correct time point.

Perfusion and tissue preparation

Euthanesia was performed with an intraperitoneal injection of 0.1 mL pentobarbital (Euthanimal 20%, 200 mg/mL) ~10 minutes prior to transcardial perfusion. The animals were perfused with ice-cold 1x PBS until blood clearance, followed by perfusion with ice-cold 4% PFA/1x PBS. Brains were extracted from the skull of perfused animals and stored in 4% PFA/1x PBS overnight for post-fixation.

Brains were cleared and stained following the iDisco+ protocol (Renier et al., 2016). Briefly, 24h post-fixation, samples were washed with 1x PBS. The olfactory bulbs and cerebellum were trimmed. A methanol/H_2O gradient was applied to dehydrate the tissue, followed by decolorization with 5% hydrogen-peroxide. After rehydration with a methanol/H_2O gradient, brains were permeabilized and remained in a blocking buffer for two days. Samples were then incubated in the primary antibody for seven days (rabbit anti c-fos, Synaptic Systems, cat. 226003, 1/1000 in 92% PTwH / 3% Donkey Serum / 5% DMSO). Following washing with 1x PBS / 0.2%Tween-20 / 0.1%Heparin (PTwH), brains were incubated for another seven days in the secondary antibody (Donkey anti-Rabbit IgG (H+L) Alexa 647™, Thermo Fisher Scientific, cat. A31573, 1/1000 in 97% PTwH / 3% Donkey Serum). Lastly, samples were washed with PTwH, dehydrated with a methanol/H_2O gradient, incubated in 66% DCM / 33% Methanol, washed with 100% DCM, and then cleared and stored in the clearing agent

dibenzyl ether (DBE), protected from the light. The total time required to complete the protocol is 25 days. The investigators processing the tissue samples were blind to the experimental groups.

Imaging with Light-Sheet microscopy

Starting three days after clearing, samples were imaged with a light-sheet microscope (UltraMicroscopeII, LaVision BioTec), equipped with an MVPLAPO 2x/0.5NA objective (Olympus), an MVX-10 Zoom Body (Olympus) and a 10mm working distance dipping cap. The images were recorded with an sCMOS camera (Neo 5.5 sCMOS, Andor Technology Ltd; image size: 2560x2160 pixels; pixel size: 6.5x6.5µm^2) using the software ImspectorPro (v.5.0285.0; LaVision BioTec). The samples were scanned in horizontal slices (step size: 3µm; effective magnification: $zoom_{body\ objective}$ + dipping lens 2x0.63x=1.26x; sheet width: 60%) with two-sided illumination using the built-in blending algorithm. Two image stacks per sample were taken consecutively, without moving the sample in between recordings. This is essential for the later correct annotation of c-fos positive (c-fos+) cells to brain areas. To record the fluorescence of c-fos+ cells, we used a Coherent OBIS 647-120 LX laser (emission filter: 676/29). The images were recorded at 70% laser power, sheet NA of 0.076 (results in a 10µm thick sheet), and an exposure time of 100 msec, as well as with horizontal focus to reduce z-slice distortion (steps: 20). To highlight the intrinsic fluorescence of the tissue for registration of the sample to a template, we imaged with a Coherent OBIS 488-50 LX Laser (filter: 525/50nm). The images were recorded at 80% laser power, sheet NA at 0.109 (results in a 7µm thick sheet), and exposure time of 100 msec, without horizontal focus.

The investigators (VB, HS) conducting imaging and image processing were blind to condition. Samples were imaged in ascending order, with sample numbers being randomly assigned during perfusion.

Image processing: cell detection.

c-fos+ cells were detected from the 647nm image stack using Imaris (v9.2.0, Bitplane). A spots object was created and parameterized to detect cells with estimated xy-diameter of 10µm and an estimated z-diameter of 35µm (to avoid overcounting cells due to z-plane distortion). Thereafter, detected cells were filtered by quality (>18) and required to cross a threshold of minimum intensity (>225). The quality filter verifies that the shape of the detected cells aligns to the spots object within a threshold specified automatically by the algorithm. These parameters were optimized on pilot imagined brains to achieve optimal signal-to-noise ratio while avoiding ceiling effects of the cell intensity parameter. The xyz-position of each cell was exported for later annotation to a reference atlas.

Image processing: alignment and annotation.

Each cleared brain was registered to the Allen Mouse Brain 25 µm reference Atlas (Wang et al., 2020) by using ClearMap (Renier et al., 2016) as an interface to the open-source software Elastix v5.0 (Klein et al., 2010). Registration was performed using the autofluorescent images (488nm image stack). As per our setting, the images were first rotated, sheared and scaled with affline transformation, then translated onto the reference atlas with b-spline transformation.

The transformation matrix as calculated by Elastix was applied to the xyz coordinates of the Imaris detected cells using Transformix. In this way, detected cells were transformed into a template space, and each c-fos+ cell was assigned to a brain region. This was possible because each sample remained in the same position for both image stacks.

To estimate the error in the approach, we drew in ImageJ (Fiji) artificial spots (n_{spots} = 72) in both hemispheres ($n_{animals}$ = 12) in three different locations (i.e., within dentate gyrus, mammillothalamic tract, amygdalar capsule). To estimate the error of the alignment, we calculated the distance between the expected position of the artificial spots and the position resulting from the alignment procedure.

The output for further data analysis is the xyz-position of every cell with their corresponding area code. The code was then translated to the respective brain area according to the brain region table provided by Renier and colleagues (Renier et al., 2016), which follows the hierarchical organization of the Allen Brain Atlas (ABA).

At this point, images have been transformed in machine-readable numbers. Other tools can be used until here. To continue with the following steps, one is only required to have files with xyz coordinates for each cell, together with their annotated brain area.

Quality control and data pre-processing

Quality control, data pre-processing and analysis were planned on a subset of the data (n_{blocks} = 3), and then later extended to the full dataset. The experimenter coding the analyses (VB) was blinded to the experimental condition.

To mitigate technical noise, a series of quality control steps were performed on the xyz annotated coordinates of c-fos+ cells. We removed (false positive) cells from brain areas in which no counts were expected, either because these areas contain no brain tissue (background, ventricular system) or because they were trimmed from the sample (olfactory bulb, cerebellum, hindbrain). Next, we grouped the highest resolution areas of the ABA in line with the ABA hierarchical organization. The aim was to preserve as much spatial specificity as allowed by alignment inaccuracies in areas likely to be stress or c-fos sensitive, while minimizing the total number of brain areas to ease interpretation and to avoid unnecessary subsequent multiple testing. Accordingly, the categorization considered the region-specific distribution of glucocorticoid receptors as well as the region-specific c-fos expression after acute stress (Cullinan et al., 1995; Morimoto et al., 1996). The hierarchical relationship of ABA areas is not complete, meaning not all larger brain areas can be fully

subdivided into smaller brain areas. These “left over” spaces were removed from the analysis since they were deemed not interpretable.

An illumination artifact was present in all samples on the outside borders of the brain and the ventricles, presumably due to unspecific antibody binding. Initially, we aimed to use the no primary antibody control group to correct for this artifact; however, this was not possible as the number of c-fos+ cells in the no primary antibody control group was minimal. As an alternative solution, a mask of 75µm thickness was modeled along the inside border of the brain and the ventricles of the aligned samples (Supplementary Figure 3 a), and cells that fell within the mask coordinates were removed from further analysis. The size of the mask (25 through 175µm) was piloted in 3 samples. Ultimately, 89 brain areas were included in the analysis (Supplementary Table 2).

Lastly, we removed xyz coordinates with extremely high c-fos intensity. We qualitatively assessed histograms of the maximum intensity of c-fos+ xyz coordinates per brain area, and compared them across samples to identify potential unspecific binding of the protein (“spots”). The potential candidates had 2- to 10-fold higher intensity than others within the same brain area. These were checked against the raw scans and removed if they did not appear as “cells” during a qualitative evaluation.

Outlying values

The selection of parameters for cell identification, the removal of the illumination artefact, the managing of areas with small volumes, and the removal of mis-labelled spots are procedures to limit as much as possible the presence of outliers. Despite these efforts, residual biological / technical outliers could be expected, either at the single cell level (e.g. unspecific antibody binding) or at the brain area level (e.g. disproportionate activation). Due to the limited sample size and the batch effects, the identification of outlying values was not trivial. We therefore chose to not use any rule (e.g. 3SD away from the mean) or statistical

test to detect / exclude / replace outliers. Rather, we assumed that they may occur uniformly across samples, thereby giving rise to increased variation. To mitigate their effects, we used medians and quantiles rather than means to summarize the data.

Missing values

The main source of missing value was the loss of animals due to insufficient staining quality (see ‘Experimental Design’ in Methods).

A second source of missing values was due to damaged brain areas during the experimental procedure. c-fos+ cells were counted per brain area across the whole brain. Damaged areas were manually detected in the 488nm image stack independently by at least two of three researchers (VB, HS, RD). The researchers were blinded to the experimental condition, and discrepancies were resolved with discussion. c-fos+ cells of damaged areas were removed, and then imputed by mirroring the xyz cells' coordinates of the same brain area of the opposite hemisphere. Although this approach inherently assumes no differences between hemispheres, we preferred it to a multiple imputation approach because it did not require batch effects’ mitigation and it could be performed at a single cell rather than at brain area level.

A third source of missing values was linked to cell detection. The cell detection algorithm requires the definition of a minimum c-fos+ intensity. This parameter was optimized in pilot experiments, and it was kept identical throughout all experimental brains. In principle this is not a problem, since by rigorous standardization it is possible to mitigate batch effects and obtain comparable relative statistics. However, when a brain area has no active cells at t_0, it needs to be further evaluated to conduct a proper standardization. In our experiment, two brain areas (FRP and AHN) had no c-fos+ cells in one t_0 sample. Since this occurred only in one sample, we considered these brain areas as missing, not as zeros for analyses that required standardization with ratios to baseline.

Preprocessing for region-based statistics

Additional pre-processing is required for region-based analyses. In Figure 1, we summarize which pre-processing steps were required for which type of analysis.

In region-based analyses, the total number of c-fos+ cells (i.e. absolute counts) was calculated per brain area. However, cell counts are by definition not normally distributed; rather, they follow a Poisson or (negative) binomial distribution. We therefore applied a Box-Cox transformation per block (i.e. each set for four different timepoints), so that our data would resemble a normal distribution.

Different brain areas have different sizes; therefore, absolute counts of c-fos+ cells are not indicative of how active a certain brain area is. In analysis where different brain areas are compared, absolute counts need to be normalized to the size of the brain area. We therefore calculated the number of c-fos+ cells per thousand of the total cells in each brain area, by adapting the atlas by Erö and colleagues (Erö et al., 2018). We used the total cell count estimation rather than that of only neurons because several publications have reported c-fos+ glia and astrocytes (for a review, see (McReynolds et al., 2018)), and it is in agreement with the presence of c-fos+ cells in the fiber tracts of our own data.

The number of c-fos+ cells differed across batches, although the relationship across time points was consistent within batches. Therefore, a normalization step was required to make the data more comparable across batches. Z-score normalization was performed per block, i.e. a unit of one sample per experimental group. With z-score normalization, the data is scaled with a mean of 0 and a standard deviation of 1, according to the formula $(x - \mu)/\sigma$, where x is the observed value, μ is the mean of the sample, and σ corresponds to the standard deviation of the sample.

Region-based analyses: active brain regions

With the exception of the single-cell strategy analysis, the analyses were planned on a subset of data (n=3), and later extended to the full dataset. The experimenter coding the analyses (VB) was blinded to the experimental groups.

We tested which brain areas had a significant increase from baseline in c-fos+ cell count per thousand of total cells ($n_{cfos+ / tot}$). The dependent variable was scaled and normalized as explained in the data pre-processing section. We performed pairwise comparisons (Welch t-test, one-sided, alpha = 0.05) for each time point (t_{30}, t_{90}, t_{180}) against t_0. P-values were adjusted with the Benjamini-Hochberg (BH) procedure. For visualization only, we transformed the p-values with a $-\log_{10}$ transformation, and we grouped the brain areas according to the ABA embryological origin.

Next, we tested which brain areas were most active. The analysis was independently performed per block; therefore, no other batch-effect correction step was taken. For each block, we calculated the top 5% of $n_{cfos+ / tot}$ independently of time point, and thereby identified per block the most active brain areas. Next, we counted how often a specific brain area was categorized as most active. We considered a brain area to be consistent across samples if it was present in at least 5 out of 9 blocks in a particular time point. If we consider the process to be random under the null hypothesis, the probability of a brain area to be present in 5 out of 9 blocks would be 0.1%. This probability was estimated with a simulation study. We simulated 1000 independent experiments, and each experiment consisted of 4 time points with 9 independent iterations (i.e., $n_{blocks} = 9$). For each iteration, we selected 18 brain areas, meaning the 5% of 90 ($n_{brain\ areas}$) * 4 ($n_{time\ points}$). Each brain area had an equal probability of being selected (i.e., uniform distribution), and a brain area could be picked multiple times within each iteration (i.e., block), up to 4 ($n_{time\ point} = 4$). Then, we calculated across 1000 experiments the probability of a brain area to be in the top 5% of the

distribution. This simulation gave information about how likely it is that the representation in the top 5% was chance.

Since c-fos is not uniformly distributed across the brain, we performed a simulation study to assess whether the pattern obtained was due to the baseline spatial distribution of c-fos. We downloaded via the ABA's API the mRNA c-fos expression levels of 3 experiments that passed the ABA quality check (id: 80342219; 79912554; 68442895). We calculated the mean and standard deviation for each of the brain areas available ($n_{brain\ areas} = 8$). Since the resolution available for c-fos expression is lower than the resolution in our dataset, we assumed that the c-fos expression available corresponded to the location and scale parameters of a normal distribution defined by all the sub-areas. In other words, for each sub-area we sampled values from a normal distribution with location and scale parameters equal to those derived from the ABA atlas. This was performed for 9 independent samples ($n_{samples} = 4$) for each time point ($n_{time\ point} = 4$). Since the selection may not be linked only to baseline c-fos expression, but also to a natural increase due to the foot-shock, we multiplied the baseline expression levels with the overall increase in c-fos across time points in our experiment. For each brain area, we calculated the median of the ratio between each foot-shock time point (t_{30}, t_{90}, t_{180}) and t_0. This ratio was then multiplied by the estimated c-fos expression values. Of note, the ratio at t_0 was always 1, meaning that the expression levels were estimated only from the ABA. We performed this simulation 1000 times, thereby simulating 1000 independent experiments. For each simulated experiment, we considered 9 blocks and 4 time points, as in our actual experiment. Then, for each block in each experiment, we selected the brain areas whose expression was in the top 5% of the distribution. Across the 1000 experiments, we then calculated the mean and standard deviation.

Region-based analyses: order of activation

Since brain areas displayed a temporal dynamic pattern, we aimed to order the brain areas based on their c-fos+ expression. Ordering brain areas based on the time of their activation is not trivial, especially since in 3D microscopy time is discrete ($n_{\text{time point}} = 4$) rather than continuous (as, for example, in fMRI). Additionally, c-fos protein is not transient, but it peaks ~90 min and decays ~180 min after a stimulus. This dynamic may even be different depending on the brain area (Kovács, 1998). With these challenges in mind, we aimed to analytically create a pseudo-time to increase the temporal resolution, which would in turn allow to order brain areas.

Among the approaches considered (Supplementary Table 3), we ultimately ordered areas based on the estimated time of mid-activation across blocks. c-fos+ cell counts were Z-transformed. Then, for each brain area we calculated the median across blocks of each time point. We interpolated a linear model between each two consecutive timepoints: this line is the "continuum" of pseudo-time. To order the brain areas, we then considered at which pseudo-time point, c-fos activation reached its mid-activation level. Since the data was Z transformed, this corresponded to reaching the value 0. To limit the sensitivity of the pseudo-time, we binned the pseudo-time variable in bouts of 10 minutes. Each brain area was therefore grouped to the closest bout (binning). This approach has the advantage to create a criterion on which to categorize brain areas, but it does not consider the range (error) among which it could happen. The approach is ideal for areas that have one point of activation; it is biased for brain areas with a biphasic activation (e.g., at the beginning and at the end of our time curve).

For interpretation and visualization purposes, we classified the brain areas (Supplementary Table 4) according to functional networks relevant to stress. We followed Henckens and colleagues' (Henckens et al., 2015) results, to which the an amygdalar group was added.

Combining voxel based and single cell analysis: sub-brain areas

We questioned whether c-fos+ cells are uniform within a brain area, or whether there are locations in which c-fos+ cells are most dense, i.e., are in closer proximity to each other. For this, we selected a brain area important for the stress response, i.e. the basolateral amygdala (BLA). All other brain areas can be visualized on the *abc4d* app.

For each sample separately, we estimated the probability density of the BLA at each cell, by using a kernel density estimation with Gaussian function. Kernel densities are routinely used to smooth data from a finite sample to make inferences about a population. Here, they were used to estimate the cell density within a brain area. Next, we filtered the cells with highest density per sample.

The BLA was divided into voxels of 30µm per side. Considering that there could be an alignment error of (maximum) ~23µm (Supplementary Figure 2b), we considered 30µm the minimum, interpretable size. To look for consistency across samples, we calculated how many samples had at least one cell in each voxel, and considered 3 the minimum for consistency. In each xyz direction, we calculated per time point the median and interquartile of the voxels' position. Of note, due to the batch effects, calculating number of cells (or other measures of activation) across samples would have had little value, and would need to be standardized. We therefore opted for this more straightforward approach.
To determine whether our observations were due to a chance process, we randomly attributed each sample to a time point, and perform the exact same analysis.

Combining region-based & single cell analysis: Strategy

We hypothesized that brain areas can show activation with different strategies. With the "count strategy", a brain area increases the *number* of c-fos+ cells with a low c-fos expression (intensity); with an "intensity strategy", c-fos+ cells increase in *intensity* rather than number.

To test this hypothesis, we analyzed t_0 samples, where we calculated the n_{cfos+} of each brain area, as well as the mean intensity of the cells in that area (intensity refers to the maximum intensity as reported by Imaris). Here, the mean rather than the median was intentionally used to be able to observe the increase in intensity due to a subgroup of cells within a brain area. For this analysis, we did not perform a batch effects correction; rather we took advantage of the differences across blocks. In our cell detection methodology ('Image processing: cell detection' in Methods), cells are identified as c-fos+ depending on intensity. This relationship should always be the same, irrespective of batch effects or time points. To quantify the relationship analytically, we therefore interpolated a linear model between the raw c-fos+ cell count and median intensity, of all brain areas of all samples.

The linear model was used as a discriminant criterion to classify whether a brain area had a strategy more towards intensity (above the regression line) or towards count (below the regression line). This categorization was performed for each brain area of each t_0 sample ($n_{animal} = 9$) independently. We then calculated the probability of a brain area to belong to a certain categorization. The resulting variable was continuous between the values of 0 (i.e., all samples had an intensity strategy) and 1 (i.e., all samples had a count strategy).

If the categorization of brain areas would be a random process, the probability of brain areas to belong to a certain categorization would be normally distributed around $\mu = 0.5$ under the null hypothesis (i.e. brain areas do not have a strategy). To validate that the null hypothesis would indeed follow a normal distribution, we performed a simulation study. In this study, we used the exact same analysis, but the values for intensity and c-fos+ cell count were drawn independently from a Poisson distribution $P(\lambda = 3.5)$. The value for lambda λ was selected by qualitatively comparing the distribution of intensity and count in the data with computer generated Poisson distributions with different lambdas. However, the interpretation would not change if different values of lambda would be selected.

Next, we questioned whether brain areas may change strategy after stress, relatively to t_0. Our experiment was not powered sufficiently to answer this particular research question, and therefore results should be interpreted as exploratory. From the categorization probability, we selected those brain areas that were consistent across samples, i.e. that had a specific categorization in at least 6 out of 9 samples. Of these, we selected those with a consistent change (increase or decrease) in count and / or intensity in at least one of the foot-shock time points (t_{30}, t_{90}, t_{180}). We calculated the pairwise difference between t_0 and each foot-shock time point for count as well as intensity. From this, we calculated the rate of change (count over intensity) per block for each time point. To compare data across brain areas, we converted the rates across samples to standardized mean differences (Hedge's g). We then classified brain areas as having changed strategy after stress if the effect size was below 0.5 or higher than 2, meaning that the relative increase in activity must have doubled towards intensity or towards count after stress compared to baseline.

Software

We developed the R package *abc4d* ("Analysis Brain Cellular activation in 4 Dimensions") to ease the implementation of data pre-processing and analysis. Furthermore, we developed a web tool (https://vbonapersona.shinyapps.io/brain_after_footshock/) to interactively visualize the effects of acute stress on the brain area of choice.

All analyses were conducted with R (version 4.0.0) in the R studio environment on a macOS Mojave (version 10.14.6). The following R packages were core to this study: 1) *tidyverse* (version 1.3.0) for general data handling and visualization; *shiny* (v 1.6.0) for the generation of the web interface; *ComplexHeatmap* (v 2.4.3) for heatmap visualization.