PNAS
www.pnas.org

Supplementary Information for

NTRK2 Methylation is Related to PTSD Risk and Symptoms in Two Independent African Cohorts of Trauma Survivors

Vanja Vukojevic[a,b,c,d,1], David Coynel[b,e,1], Navid Reza Ghaffari[b,e,f,1], Virginie Freytag[a,b,c], Thomas Elbert[g], Iris-Tatjana Kolassa[h], Sarah Wilker[i], James L. McGaugh[f,2], Andreas Papassotiropoulos[a,b,c,d,2], Dominique J.-F.de Quervain[b,c,d,e,2]

[a] *Department of Psychology, Division of Molecular Neuroscience, University of Basel, Birmannsgasse 8, CH-4055 Basel, Switzerland*
[b] *Transfaculty Research Platform, University of Basel, Birmannsgasse 8, CH-4055 Basel, Switzerland*
[c] *University Psychiatric Clinics, University of Basel, Wilhelm Klein-Strasse 27, CH-4055 Basel, Switzerland*
[d] *Department Biozentrum, Life Sciences Training Facility, University of Basel, Klingelbergstrasse 50-70, CH-4056 Basel, Switzerland*
[e] *Department of Psychology, Division of Cognitive Neuroscience, University of Basel, Birmannsgasse 8, 4055 Basel, Switzerland*
[f] *Center for the Neurobiology of Learning and Memory, Department of Neurobiology and Behavior, University of California, Irvine, CA 92697-3800*
[g] *Clinical Psychology, University of Konstanz, D-78457 Konstanz, Germany*
[h] *Clinical & Biological Psychology, Institute for Psychology & Education, Ulm University, Albert-Einstein-Allee 47, D-89069 Ulm, Germany*
[i] *Department for Psychology and Sports Science University Bielefeld, P.O. Box 100131, D-33501 Bielefeld, Germany*

[1] *these authors contributed equally to this work*
[2] *these authors jointly supervised this work*

Correspondence:

www.pnas.org/cgi/doi/10.1073/pnas.2008415117

Prof. Dr. James L. McGaugh
University of California, Irvine, Center for the Neurobiology of Learning and Memory, Department of Neurobiology and Behavior
CA 92697-3800, USA
e-mail: james.mcgaugh@uci.edu

Prof. Dr. Dominique J.-F. de Quervain
University of Basel, Department of Psychology, Division of Cognitive Neuroscience, Birmannsgasse 8, CH-4055 Basel, Switzerland
Phone/Fax: +41 61 207 02 37
e-mail: dominique.dequervain@unibas.ch
&
Dr. Vanja Vukojevic
University of Basel, Department of Psychology, Division of Molecular Neuroscience, Birmannsgasse 8, CH-4055 Basel, Switzerland
Phone/Fax: +41 61 207 02 58
e-mail: vanja.vukojevic@unibas.ch

This PDF file includes:

Supplementary Information Text

African Samples, conflict survivors

PTSD risk and symptomatology were assessed in two independent African samples of conflict zone survivors. For the African sample 1 we included 463 survivors of the rebel war in Northern Uganda. Data collection took place in the former internally displaced people (IDP) camps of Anaka, Pabbo (Amuru District) and Koch Goma (Nwoya District), and in resettled communities and villages of Gulu District, Northern Uganda (mean age 29 y, 18-55 y; 44.1 % females; 68.6 % with PTSD lifetime diagnosis; 17.4 % subjects with current PTSD). For the African sample 2 we included N = 350 survivors from the 1994 Rwandan genocide who lived as refugees in the Nakivale settlement in Uganda (mean age 34.8 y, 18-68 y; 49.1 % females; 78.8 % with PTSD lifetime diagnosis; 40.3 % subjects with current PTSD). This is a substantially enlarged sample of the African sample of our previously published findings on the relation of the epigenetic modification of the glucocorticoid receptor gene with traumatic memory and PTSD risk (1-4).

All subjects had experienced traumatic situations and were examined according to DSM-IV criteria. The Post-Traumatic Diagnostic Scale (PDS; (5-8)) was administered as a structured interview by expert psychologists from the Universities of Konstanz and Ulm, Germany, as well as by trained local interviewers. The PDS was used to assess current and lifetime symptoms of intrusions, avoidance, and hyperarousal (Table S1) as well as the current and lifetime diagnosis of PTSD according to DSM-IV. Since spontaneous remission might have occurred during the time between the traumatic experiences and the interview (5, 8, 9), the lifetime variables were chosen as the main outcomes for the statistical analyses. In order to determine lifetime symptoms, participants were asked to

report if they experienced the respective symptom or not, according to the worst period of 4 weeks since the traumatic event. Traumatic load was assessed with the number of traumatic event types experienced, which has been previously shown to be a reliable and efficient way of assessment (7, 8). Accordingly, the resulting lifetime scores represent the number of experienced symptoms, but not symptom severity. This approach has been validated and used in several previous studies (1-4, 7-9, 11).

Sum of lifetime traumatic events (both war-related and war-unrelated) was assessed with a 36-item checklist in African Sample 2 (Rwanda) (1-4). For African Sample 1 (Uganda), we used a 62-item checklist that included additional questions regarding atrocities specific to the LRA - Lord's Resistance Army (e.g., forced to eat human flesh) (5-8). The resulting difference in the mean traumatic load score between the two African samples (25.7 and 11.9, for African Sample 1 and 2, respectively), can be understood as a consequence of the different event lists used (5, 8, 9). The Traumatic event check lists did not differentiate between trauma in childhood and trauma in adulthood. Depressive symptoms were ascertained with the depression section of the Hopkins Symptom Checklist (HSCL-D) (7, 8, 10).

Traumatic load was estimated by assessing the number of different traumatic event types experienced or witnessed, which is a reliable measurement of traumatic event exposure and showed the strongest relationship with lifetime PTSD (1-4, 7-9, 11). To avoid known ceiling effects of trauma load on PTSD risk, subjects were selected to have experienced up to 59 or 19 traumatic event types, for the African Sample 1 and African Sample 2, respectively. For a trauma load higher than these cut-offs, the probability of lifetime PTSD approaches 100% (9). To exclude genetic relatives in the samples, only one

person per household was interviewed. Additionally, given the genetic data was also available, a fairly conservative IBD (identity by descent) threshold > 0.08 was applied (12). Individuals with current alcohol abuse, acute psychotic symptoms and individuals who were under psychiatric medication were excluded. Saliva samples were collected at the time-point of the main investigation for the DNA isolation. For details see (1, 5).

Swiss Sample, healthy young adults

Memory was assessed in a sample of healthy, young adults in the city of Basel, Switzerland who participated in a behavioral and imaging genetics study (Swiss Sample: N = 568; mean age 23.8 y, 18.3-36.8 y; 59% females). Participants were healthy, free of any neurological or psychiatric illness, and did not take any medication at the time of the experiment (except hormonal contraceptives).

Subjects performed different consecutive tasks as described in detail previously (13). As in our previous study (1), we also focused in the present study on a picture recognition memory task (N = 537). Incidental encoding of the pictures was achieved through an event-related design. It consisted of 100 trials, including two primacy and two recency trials depicting neutral information, 24 scrambled pictures, and 24 pictures per valence category (positive, negative, neutral). The pictures were presented for 2.5 s in a quasi-randomized order so that a maximum of four pictures of the same category were shown consecutively. A fixation-cross appeared on the screen for 500 ms before each picture presentation. Trials were separated by a variable intertrial period (period between appearance of a picture and the next fixation cross) of 9–12 s (jitter). During the intertrial period, participants subjectively rated the meaningful pictures according to valence (positive, neutral, negative)

and arousal (high, medium, low) on a three-point scale (Self Assessment Manikin). For scrambled pictures, participants rated form (vertical, symmetric, horizontal) and size (small, medium, large) of the geometrical object in the foreground.

After encoding, participants performed a free recall task of the pictures in a separate room (no time limit was set for this task). Approximately 80 min. after the presentation of the last picture in the encoding task, participants performed a recognition task for 20 min. We used an event-related design consisting of 144 trials. Pictures from two different sets were presented. Each set contained 72 pictures (24 pictures for each stimulus category), one of the sets of stimuli was new (i.e., not presented before), the other old (i.e., presented during the encoding task). The pictures were presented for 1 s in a quasi-randomized order so that at most four pictures of the same category (i.e., negative new, negative old, neutral new, neutral old, positive new, positive old) were shown consecutively. A fixation-cross appeared on the screen for 500 ms before each picture presentation. Trials were separated by a variable intertrial period of 6–12 s (jitter) that was equally distributed for each stimulus category. During the intertrial period, participants subjectively rated the picture as remembered, familiar or new on a three-point scale by pressing a button with the fingers of their dominant or nondominant hand. Correct recognition performance was measured as the number of old pictures correctly remembered minus the number of old ones incorrectly identified as new or familiar.

For DNA isolation, saliva samples were collected at the time-point of the main investigation. Additionally, subjects were re-invited for an additional saliva and blood sampling, which took place on average 360 days (median 341 days) after the main investigation. Samples were collected between midday and evening (mean time 2:30 p.m.,

range 1:00 p.m. - 8:00 p.m.). Hematological analysis, including blood cell counts, was performed with Sysmex pocH-100iTM Automated Hematology Analyzer (Sysmex Co, Kobe, Japan).

MRI data acquisition

Measurements were performed on a Siemens Magnetom Verio 3 T whole-body MR unit equipped with a twelve-channel head coil. During the course of the study there were 2 changes in gradient coils and one change in MR software. All group-level analyses included those factors as covariates.

A high-resolution T1-weighted anatomical image was acquired using a magnetization prepared gradient echo sequence (MPRAGE) with the following parameters: TR = 2000 ms; TE = 3.37 ms; TI = 1000 ms; flip angle = 8°; 176 slices; FOV = 256 mm; voxel size = 1 x 1 x 1 mm3.

During the encoding and recognition tasks blood oxygen level-dependent fMRI was acquired, using a single-shot echo-planar sequence (EPI) using parallel imaging (GRAPPA). The following acquisition parameters were used: TE (echo time) = 35 ms, FOV (field of view) = 22 cm, acquisition matrix = 80 × 80 (interpolated to 128 × 128, voxel size: 2.75 × 2.75 × 4 mm3), GRAPPA acceleration factor R = 2.0. Using a midsagittal scout image, 32 contiguous axial slices placed along the anterior–posterior commissure (AC–PC) plane covering the entire brain with a TR = 3000 ms (α = 82°) were acquired using an ascending interleaved sequence.

From the N=537 subjects with methylation data, N=498 were included in the fMRI data analysis. N=39 subjects were excluded because of missing fMRI data, technical issues with MR acquisition, or data quality issues (T1 quality or extreme movement, see below).

fMRI QC steps and filters

T1 data were visually checked for movement artefacts by 3 raters. Additionally, Qoala v1.1 was used on freesurfer v6.0 measures (https://qoala-t.shinyapps.io/qoala-t_app/): subjects labelled "exclude without manual QC" were excluded, while subjects labelled "exclude with manual QC" and "include with manual QC" underwent a second visual QC. A matlab-based plot of standard deviation of signal outside of the brain (SPM12-based brain mask; in-house code) was used to detect potential noise in the acquisition, due to technical artefacts. Subjects with more than 1 noisy slice were excluded. SPM12 brain segmentations were visually checked. Subjects for which it failed were excluded from further fMRI analyses. Because of the need of a good quality T1 image for EPI normalization, subjects who had an excluded T1 image were excluded from all subsequent fMRI analyses.

The framewise displacement was computed for every subject's functional scan (14). Subjects where more than 5% of the volumes had a >1mm framewise displacement were excluded. The "Power plot" was also visually inspected to detect scanner-related artifacts. EPI-to-T1 coregistration as well as EPI-to-MNI normalization were visually checked for every subject.

fMRI pre-processing

fMRI data were preprocessed using SPM12 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging; http://www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB R2016b (MathWorks).

Volumes were slice-time corrected to the first slice, realigned using the 'register to mean' option, and coregistred to the anatomical image by applying a normalized mutual

information 3-D rigid-body transformation. Subject-to-template normalization was done using DARTEL (15), which allows registration to both cortical and subcortical regions and has been shown to perform well in volume-based alignment (16). Normalization incorporated the following steps: 1) Structural images of each subject were segmented using the "Segment" procedure in SPM12; 2) The resulting gray and white matter images were used to normalize the subject's image to the template space; 3) Subject-to-template and template-to-MNI transformations were combined to map the functional images to MNI space. The functional images were smoothed with an isotropic 8mm full-width at half-maximum (FWHM) Gaussian filter.

Normalized functional images were masked using information from their respective T1 anatomical file as follows. At first, the three-tissue classification probability maps of the "Segment" procedure (grey matter, white matter, and CSF) were summed to define a brain mask. The mask was binarized, dilated and eroded with a 3 × 3 × 3 voxels kernel using fslmaths (FSL) to fill in potential small holes. The previously computed DARTEL flowfield was used to normalize the brain mask to MNI space, at the spatial resolution of the functional images. The resulting non-binary mask was thresholded at 50% and applied to the normalized functional images. Consequently, the implicit intensity-based masking threshold usually employed to compute a brain mask from the functional data during the first level specification (spm_get_defaults('mask.thresh'), by default fixed at 0.8) was not needed any longer and set to a lower value of 0.05.

General linear models (GLMs) were specified for each subject to identify voxels activated by the recognition task. Regressors modeling the onsets and duration of stimulus events were convolved with a canonical hemodynamic response function (HRF). More

precisely, the model comprised regressors for button presses modeled as stick/delta functions, picture presentations modeled with an epoch/boxcar function (duration: 1s), and rating scales modeled with an epoch/boxcar function of variable duration (depending on when the subsequent button press occurred). Pictures were modeled separately depending on whether they were correctly identified as old (previously seen) or new (never seen). Previously seen pictures were further divided into pictures correctly rated as old or familiar. Serial correlations were removed using a first-order autoregressive model, and a high-pass filter (128s) was applied to remove low-frequency noise. The six movement parameters from movement correction were also entered as nuisance covariates. Subject-level contrasts were computed to estimate brain activity of [correctly remembered - correctly identified as new] pictures.

ICA was used as an unbiased, data-driven method to reduce the dimensionality of the first-level data to a lower number of statistically independent components (ICs) (17, 18). We applied ICA to a matrix X ([correctly remembered - correctly identified as new] contrast), comprised of m observations (participants) and n variables (voxels). ICA estimates a matrix of k x n latent sources S that underlie the variables, while holding the source estimates (voxel loadings) as independent from each other as possible. Therefore, in the ICA decomposition voxel loadings describe statistically independent latent sources that underlie the contrast estimates. Additionally, ICA provides a matrix of m x k mixing coefficients A (participants scores) for each IC. The mixing coefficients of each component represent the component's activity strength, per participant (19).

The optimal number of components for the ICA decomposition was chosen through a stability analysis of several decompositions. In total 20 decompositions were computed,

with N = 3:20, 22, 24 ICs respectively. For each decomposition, we used a resampling method with 100 repetitions and 90% of randomly selected participants from the whole sample, producing 100 similar, but non-identical subsamples, in order to prevent overfitting. For each of the 20 ICA solutions: a) an ICA was performed on each of the 100 subsamples using the fastICA algorithm (R-package "fastICA"); b) the stability of each IC was estimated by computing the absolute value of Pearson's correlation to its voxel loadings across subsamples; c) ICA solution stability was calculated as the mean stability across its ICs. The final number of components was chosen from the solution with the highest stability (N=13), and the final ICA decomposition was run with ICASSO approach (20) implemented in the "clusterFastICARuns" function from the MineICA R-package (21) (https://www.bioconductor.org/packages/release/bioc/html/MineICA.html). This approach runs the fastICA algorithm 100 times with random initializations. The obtained components are clustered and the medoids of these clusters are used as the final estimates. This decomposition was performed on a larger cohort of 1576 subjects, comprising the subjects from this study.

fMRI statistical analyses

Following the ICA decomposition, the IC's participants scores relationship with correct recognition memory performance and *NTRK2* methylation were examined by means of linear models. Correction for multiple comparisons across the 13 ICs and two independent variables was applied using the Bonferroni procedure, with a corrected threshold of p_{Bonf}=0.0019. Age, sex and two MR-related technical batches were included as covariates in all models.

DNA isolation

Saliva DNA was collected in the African Samples using an Oragene DNA Kit (DNA Genotek, Ottawa, ONT) and initially extracted using the precipitation protocol recommended by the manufacturer. High-purity DNA was obtained by additional re-purification. For this purpose, 2µg of DNA isolated via the Oragene procedure was incubated overnight at 50°C with proteinase K (lysis buffer: 30 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, 150 ng/l proteinase K), agitated by gentle orbital shaking. Next, the DNA was purified using a Genomic DNA Clean & Concentrator Kit (Zymo Research, Irvine, CA).

Blood samples were collected in the Swiss Samples from all subjects using the BD Vaccutainer Push Button blood collection set and 10.0 mL BD Vacutainer® Plus plastic whole blood tube, BD Hemogard™ closure with spray-coated K2EDTA (Becton Dickinson, Franklin Lakes, NJ). Sampling was performed in the period between June 2011 and May 2013. Immediately after collection, blood samples were processed in a hematological lab. Upon plasma removal, the remaining fraction was first stored overnight on –20°C, and then moved to –80°C biobank storage. DNA isolation was then performed from the plasma-removed blood fraction on average 2.3 weeks later (range: 1-31 d). Isolation was performed with QIAmp Blood Maxi Kit (Qiagen AG, Hilden, Germany), using the recommended spin protocol. DNA quality and concentration were assessed using spectrophotometry (Nanodrop 2000; ThermoScientific, Waltham, MA) and fluorometry (Qubit dsDNA BR Assay Kit, Invitrogen, Carlsbad, CA). DNA was then aliquoted and long-term stored in the biobank at –80°C. DNA aliquots not affected by the freeze-thaw

cycles were used for the Infinium BeadChip methylation analyses performed in August 2013.

Infinium 450K and EPIC BeadChip methylation analyses

DNA isolated from saliva (African samples) or peripheral blood (Swiss Sample) was investigated with the 450K (African Sample 2, Swiss Sample) or EPIC (African Sample 1) array (Illumina, Inc., San Diego, CA; see Supplementary Material). For the African Sample 1 (6 plates, N = 463) and African Sample 2 (4 plates, N = 350) all subjects were processed in a single batch. The subjects of Swiss sample (N = 568) were processed in two batches (2 plates and 4 plates). Within a batch, samples were processed with a randomized plate assignment and with a single bisulfite conversion.

Preprocessing of data was done separately for each batch. Data were extracted and analysed from the generated idat files using the R package RnBeads version 0.99.9 (22). CpG annotation was based on the manufacturer's annotation file (Human-Methylation450_15017482_v.1.2). During preprocessing, the background was subtracted using the "noob" method in the methylumi package (23), and the signal was further normalized using the SWAN algorithm (24). The following probe categories were excluded from the final data sets, based on the annotation provided within the RnBeads package: non-CpG context probes (due to underrepresentation on the 450 K array, 0.6%, (25); functional differences when compared to the CpG context as well as very low abundance of non-CpG methylation in somatic tissues (26) ; N = 3091); probes with a SNP mapping directly to the target CpG site, as well as probes with three and more SNPs mapping within the 50mer probe (see Supplementary Fig. 2; MAF threshold was set to 0.01; N = 18,998

CpGs); gonosomal probes (N = 11,473 CpGs); non-specific probes. Using the Greedycut algorithm, we iteratively removed the probes and data sets of the highest impurity (rows and columns in the detection p-value table that contain the largest fraction of unreliable measurements; $p < 0.05$; for each sample (22)).

Post-processing was further done for each sample separately, combining the B-values of the preprocessed data of all batches per sample. The B-values were further processed step-by-step in order to correct for further influential and putative confounding factors: 1) using logit- transformation (M-value, (27), done with the R-package car (28)); 2) z-transformation per plate (correcting for plate and batch effects); 3) regressing out the first 10 (African Samples) or 8 (Swiss Sample) axes of a principal component analysis (PCA, done with the R-package pcaMethods (29)). The PCA was based on CpGs with no missing values (>95% of the included CpGs). The PCA-based approach corrected for technical biases as well as for part of the variability induced by blood cell composition (European descent Samples); 4) regressing out the effects of sex and age; 5) regressing out the effects of variants in the 50mer probe sequence, if the total variance explained by these variants exceeded 0.1%.

The accepted missing rate per CpG was set to <1%. We further excluded cross-hybridizing probes and polymorphic CpG sites (30, 31) (Nmax = 63,974). Only CpGs surviving all filtering steps in all samples were used for the downstream analyses (N = 394,043 common CpGs across 450K und EPIC arrays). For details and validation of Infinium 450K array processing see (32).

Finally, we used the genome-wide regional segmentation analysis and clustered the individual CpGs to regional elements, as previously described (22, 33). In short: 1) Genome

(GRCh37/hg19) was segmented in 5 kb sliding-window regional elements; 2) Regional methylation was then calculated as mean value of DNA-methylation for CpGs clustered to every individual regional element; 3) Instead of performing epigenome wide association analysis (EWAS), we further focused on the targeted analysis inside the RGRS gene set (GO:2000322; defined based on the mammalian gene ontology repository (http://amigo.geneontology.org/amigo)). Such region-of-interest–based DMR analyses provide an effective way of increasing the statistical power to detect differential DNA methylation, and furthermore it also increases the interpretability of identified DMRs (22); 4) The RGRS gene set contained 12 genes (*CLOCK, PPP5C, BDNF, PHB, PER1, CRY1, NCOA2, LMO3, NR3C1, CRY2, ARNTL, NTRK2*) and predefined regional methylation elements (step 1 & 2) were then annotated to these genes by using the BioMart repository ((34); GRCh37/hg19, Ensembl Version 75). In order to have 5' and 3' regulatory regions included, extended gene borders were applied, covering 5 kb upstream and downstream from gene beginning and end, respectively. Given the distribution of the CpG sites examined by the *Infinium* (Illumina) arrays, regional methylation elements covering 5'UTR and TSS were overrepresented compared to intragenic, 3'UTR and intergenic regions, increasing the interpretability of identified regions; 5) Finally, the association of the epigenetic regulation of the RGRS pathway with PTSD and cognitive phenotypes was then performed by targeted analysis of the DNA methylation patterns of the regional methylation elements annotated to the RGRS genes only. Given our sample sizes, in all three examined cohorts, we were well powered. For the extensive list of the examined regional methylation elements of the RGRS pathway genes please refer to the

Supplemental Table S1 (total number of regional methylation elements within RGRS gene set shared across *450K* and *EPIC* (Illumina) platforms was 94).

Statistical analyses

In the African samples, the association between lifetime PTSD risk and symptom scores as dependent variables and DNA methylation of predefined regional methylation elements was assessed. To account for trauma load as a principal factor in the development of PTSD (2) sum of lifetime traumatic event types was used as a covariate in the regression models.

The association of the epigenetic regulation of the RGRS pathway with the PTSD was assessed by first modelling PTSD lifetime risk against every regional methylation element annotated to the RGRS gene set (Supplemental Table S1), by logistic regression. Following, the corrected significance was estimated by applying a Bonferroni threshold for the total number (n=175) of analyzed regional methylation elements (the discovery African sample 1 had 175 regional methylation elements annotated to RGRS pathway, out of which 94 were shared across all 3 examined samples – due to differences across *450K* and *EPIC* (Illumina) platforms). Finally, for the *NTRK2* regional element (chr9:87285001-87290000) the relationship of DNA methylation with PDS sum- and sub-scores in both African samples was calculated with linear regression using *NTRK2* regional element methylation as a quantitative predictor.

In the Swiss sample, recognition performance as dependent variable was modelled against the DNA methylation of predefined *NTRK2* regional element (chr9:87285001-87290000) in a linear model. The interaction of *NTRK2* DNA methylation with the valence

of pictures used in the emotional picture-encoding task was additionally tested in a linear mixed model.

A comparison of methylation levels of *NTRK2* regional element between the Swiss and the African population was done by Kruskal–Wallis one-way ANOVA. Furthermore, we assesse*d the equality of distributions and variability between the populations with Kolmogorov–Smirnov two-sample and Siegel–Tukey tests, respectively*. Effect sizes were calculated by Cohen's f^2 (35) for *NTRK2* regional DNA methylation within multiple regression models, by using the effectsizes R package (36). These indices represent an estimate of how much variance in the response variables is accounted for by the specific explanatory variable (NTRK2 regional DNA methylation). All laboratory procedures were conducted in a blinded, randomized order, including DNA isolations, bisulfite conversion and DNA methylation analysis. Only after performing all procedures and excluding samples with low quality controls and outliers, further analysis with phenotypic data was performed.

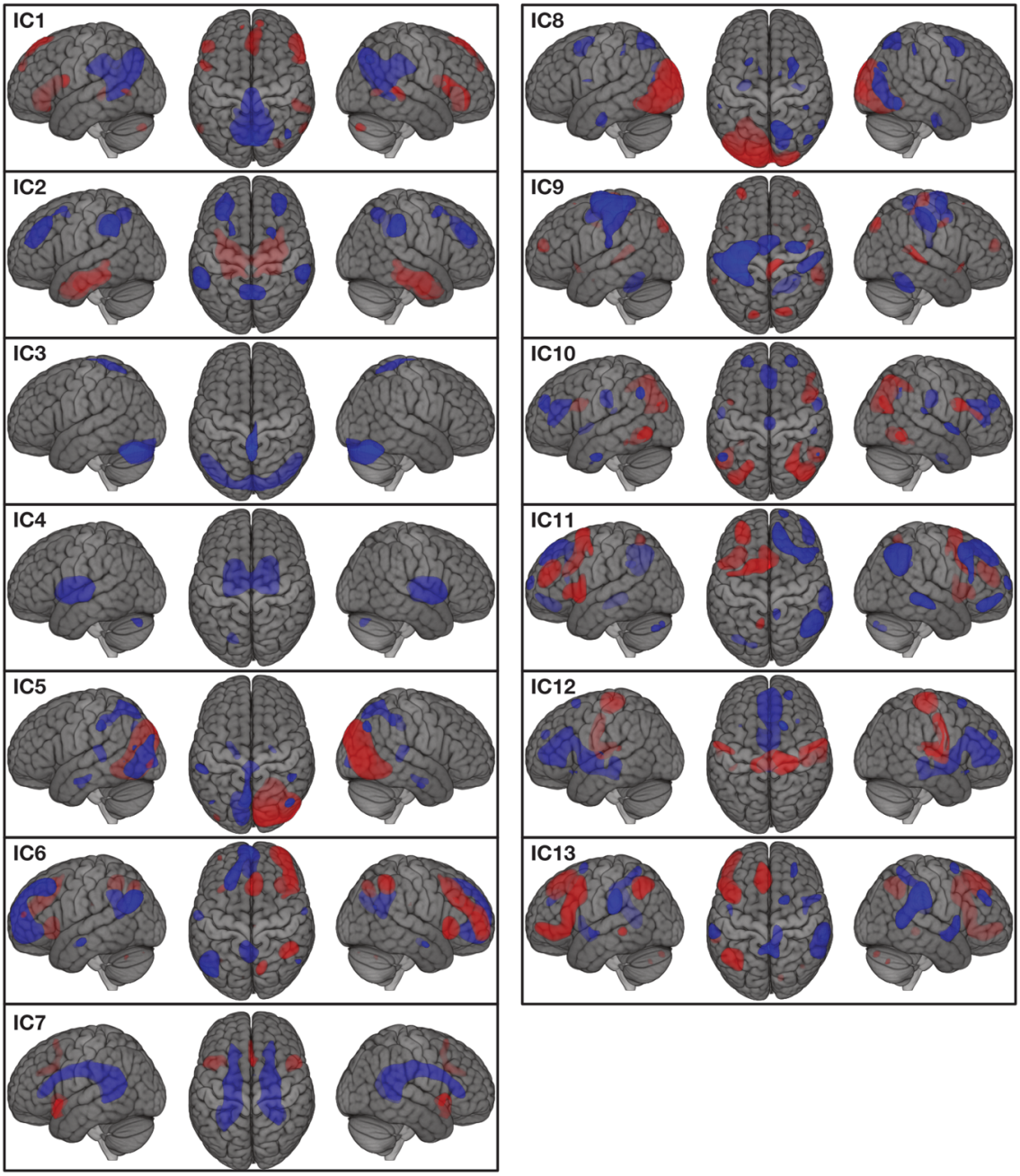

Fig. S1. Brain co-activation networks.

Independent Component Analysis (ICA) was applied to decompose brain activation into co-activation networks across participants during the recognition memory task. Thirteen components were identified. Red areas represent positive loadings and blue areas negative loadings. Only loadings below the 10^{th} percentile and above the 90^{th} percentile per component are represented.

Table S1. The examined regional DNA methylation elements annotated to the RGRS pathway genes. Genome was segmented in 5 kb sliding-window regional elements. Regional methylation was calculated as mean value of DNA-methylation for CpGs annotated to every individual element, as previously described (22, 33). Regional elements with DNA methylation signatures nominally associated with PTSD lifetime risk are marked in bold red font.

Chromosome	Start	End	Gene
4	56295001	56300000	CLOCK
4	56300001	56305000	CLOCK
4	56310001	56315000	CLOCK
4	56315001	56320000	CLOCK
4	**56325001**	**56330000**	**CLOCK**
4	56330001	56335000	CLOCK
4	56340001	56345000	CLOCK
4	56365001	56370000	CLOCK
4	56375001	56380000	CLOCK
4	56380001	56385000	CLOCK
4	56385001	56390000	CLOCK
4	56390001	56395000	CLOCK
4	56405001	56410000	CLOCK
4	56410001	56415000	CLOCK
5	142655001	142660000	NR3C1
5	142675001	142680000	NR3C1
5	142685001	142690000	NR3C1
5	142690001	142695000	NR3C1
5	**142695001**	**142700000**	**NR3C1**
5	142700001	142705000	NR3C1
5	142715001	142720000	NR3C1
5	**142720001**	**142725000**	**NR3C1**
5	142725001	142730000	NR3C1
5	142730001	142735000	NR3C1
5	142735001	142740000	NR3C1
5	142740001	142745000	NR3C1
5	142755001	142760000	NR3C1
5	142760001	142765000	NR3C1
5	142765001	142770000	NR3C1

Chromosome	Start	End	Gene
5	142770001	142775000	NR3C1
5	142775001	142780000	NR3C1
5	142780001	142785000	NR3C1
5	142785001	142790000	NR3C1
5	142790001	142795000	NR3C1
5	142795001	142800000	NR3C1
5	142800001	142805000	NR3C1
5	142805001	142810000	NR3C1
5	142810001	142815000	NR3C1
5	**142815001**	**142820000**	**NR3C1**
8	71025001	71030000	NCOA2
8	71035001	71040000	NCOA2
8	71040001	71045000	NCOA2
8	71045001	71050000	NCOA2
8	71050001	71055000	NCOA2
8	71060001	71065000	NCOA2
8	71065001	71070000	NCOA2
8	71070001	71075000	NCOA2
8	71085001	71090000	NCOA2
8	71090001	71095000	NCOA2
8	71095001	71100000	NCOA2
8	71100001	71105000	NCOA2
8	71105001	71110000	NCOA2
8	71120001	71125000	NCOA2
8	71125001	71130000	NCOA2
8	71130001	71135000	NCOA2
8	71135001	71140000	NCOA2
8	71145001	71150000	NCOA2
8	71150001	71155000	NCOA2
8	71155001	71160000	NCOA2
8	71170001	71175000	NCOA2
8	**71180001**	**71185000**	**NCOA2**
8	71185001	71190000	NCOA2
8	71195001	71200000	NCOA2
8	71225001	71230000	NCOA2
8	71235001	71240000	NCOA2
8	71240001	71245000	NCOA2
8	71245001	71250000	NCOA2
8	71255001	71260000	NCOA2
8	71265001	71270000	NCOA2
8	71275001	71280000	NCOA2

(continued)

Chromosome	Start	End	Gene
8	71280001	71285000	NCOA2
8	71285001	71290000	NCOA2
8	71290001	71295000	NCOA2
8	71295001	71300000	NCOA2
8	71300001	71305000	NCOA2
8	71305001	71310000	NCOA2
8	71310001	71315000	NCOA2
8	71315001	71320000	NCOA2
9	**87280001**	**87285000**	**NTRK2**
9	**87285001**	**87290000**	**NTRK2**
9	**87305001**	**87310000**	**NTRK2**
9	87310001	87315000	NTRK2
9	87315001	87320000	NTRK2
9	87320001	87325000	NTRK2
9	87325001	87330000	NTRK2
9	87345001	87350000	NTRK2
9	87355001	87360000	NTRK2
9	87360001	87365000	NTRK2
9	87365001	87370000	NTRK2
9	87370001	87375000	NTRK2
9	87395001	87400000	NTRK2
9	87400001	87405000	NTRK2
9	87430001	87435000	NTRK2
9	87435001	87440000	NTRK2
9	87445001	87450000	NTRK2
9	**87465001**	**87470000**	**NTRK2**
9	87475001	87480000	NTRK2
9	87485001	87490000	NTRK2
9	87560001	87565000	NTRK2
9	87565001	87570000	NTRK2
9	87580001	87585000	NTRK2
9	87585001	87590000	NTRK2
9	87620001	87625000	NTRK2
9	87635001	87640000	NTRK2
11	13295001	13300000	ARNTL
11	13300001	13305000	ARNTL
11	13305001	13310000	ARNTL
11	13310001	13315000	ARNTL
11	13315001	13320000	ARNTL
11	13320001	13325000	ARNTL
11	13325001	13330000	ARNTL

(continued)

Chromosome	Start	End	Gene
11	13335001	13340000	ARNTL
11	13340001	13345000	ARNTL
11	13345001	13350000	ARNTL
11	13350001	13355000	ARNTL
11	13355001	13360000	ARNTL
11	13360001	13365000	ARNTL
11	13365001	13370000	ARNTL
11	13370001	13375000	ARNTL
11	13375001	13380000	ARNTL
11	13380001	13385000	ARNTL
11	13385001	13390000	ARNTL
11	13390001	13395000	ARNTL
11	13395001	13400000	ARNTL
11	27675001	27680000	BDNF
11	27680001	27685000	BDNF
11	27685001	27690000	BDNF
11	27690001	27695000	BDNF
11	27695001	27700000	BDNF
11	27705001	27710000	BDNF
11	27715001	27720000	BDNF
11	27720001	27725000	BDNF
11	27725001	27730000	BDNF
11	27730001	27735000	BDNF
11	27735001	27740000	BDNF
11	27740001	27745000	BDNF
11	45865001	45870000	CRY2
11	45870001	45875000	CRY2
11	45875001	45880000	CRY2
11	45880001	45885000	CRY2
11	45885001	45890000	CRY2
11	45890001	45895000	CRY2
11	45895001	45900000	CRY2
11	45900001	45905000	CRY2
12	16700001	16705000	LMO3
12	16715001	16720000	LMO3
12	16720001	16725000	LMO3
12	16725001	16730000	LMO3
12	16755001	16760000	LMO3
12	16760001	16765000	LMO3
12	107385001	107390000	CRY1
12	107390001	107395000	CRY1

(continued)

Chromosome	Start	End	Gene
12	107395001	107400000	CRY1
12	107405001	107410000	CRY1
12	107420001	107425000	CRY1
12	107440001	107445000	CRY1
12	107450001	107455000	CRY1
12	107470001	107475000	CRY1
12	107480001	107485000	CRY1
12	107485001	107490000	CRY1
17	8040001	8045000	PER1
17	8045001	8050000	PER1
17	8050001	8055000	PER1
17	8055001	8060000	PER1
17	47480001	47485000	PHB
17	47485001	47490000	PHB
17	47490001	47495000	PHB
19	46850001	46855000	PPP5C
19	46855001	46860000	PPP5C
19	46870001	46875000	PPP5C
19	46875001	46880000	PPP5C
19	46880001	46885000	PPP5C
19	46885001	46890000	PPP5C
19	46890001	46895000	PPP5C
19	46895001	46900000	PPP5C

(continued)

Table S2. The associations of the examined regional DNA methylation elements annotated to the *NTRK2* gene with lifetime PTSD risk. Regional methylation was calculated as mean value of DNA-methylation for CpGs annotated to every individual element at *NTRK2* gene. Associations of *NTRK2* gene regional methylation with lifetime PTSD risk was assessed in African Sample 1 ($N = 463$) and African Sample 2 ($N = 350$) with logistic regression. Fixed effects analysis *t*-statistics and nominal *p*-values are shown. The regional DNA methylation element surviving the multiple comparisons correction is marked with bold red font (chr9:87285001-87290000). Elements present only on the EPIC array are marked with blue font. Effect sizes of *NTRK2* regional DNA methylation are provided as Cohen's f^2 (90% CI).

African_Sample_1						
Chromosome	Start	End	Gene	*T.statistic*	*P.value*	Effect Size f^2
9	87280001	87285000	NTRK2	-2.582	1.01E-02	0.12 (0.04, 0.2)
9	**87285001**	**87290000**	**NTRK2**	**-3.758**	**1.93E-04**	0.17 (0.1, 0.25)
9	87305001	87310000	NTRK2	2.403	1.67E-02	0.11 (0.03, 0.19)
9	87310001	87315000	NTRK2	-0.849	3.96E-01	0.03 (0, 0.11)
9	87315001	87320000	NTRK2	-0.050	9.60E-01	0 (0, 0)
9	87320001	87325000	NTRK2	0.245	8.06E-01	0.01 (0, 0.02)
9	87325001	87330000	NTRK2	-0.259	7.96E-01	0.02 (0, 0.09)
9	87345001	87350000	NTRK2	-0.051	9.59E-01	0 (0, 0.04)
9	87355001	87360000	NTRK2	0.786	4.33E-01	0.04 (0, 0.11)
9	87360001	87365000	NTRK2	-0.547	5.85E-01	0.01 (0, 0.07)
9	87365001	87370000	NTRK2	0.153	8.78E-01	0 (0, 0.04)
9	87370001	87375000	NTRK2	0.911	3.63E-01	0.04 (0, 0.11)
9	87395001	87400000	NTRK2	0.530	5.96E-01	0.03 (0, 0.04)

African_Sample_1		*(continued)*				
Chromosome	Start	End	Gene	*T.statistic*	*P.value*	Effect Size f^2
9	87400001	87405000	NTRK2	-0.385	7.00E-01	0 (0, 0)
9	87430001	87435000	NTRK2	1.158	2.48E-01	0.05 (0, 0.13)
9	87435001	87440000	NTRK2	-0.215	8.30E-01	0 (0, 0.02)
9	87445001	87450000	NTRK2	-1.423	1.55E-01	0.06 (0, 0.14)
9	87465001	87470000	NTRK2	2.432	1.54E-02	0.11 (0.03, 0.19)
9	87475001	87480000	NTRK2	0.988	3.24E-01	0.05 (0, 0.12)
9	87485001	87490000	NTRK2	-0.268	7.88E-01	0.02 (0, 0.09)
9	87560001	87565000	NTRK2	-0.171	8.64E-01	0.01 (0, 0.07)
9	87565001	87570000	NTRK2	-0.841	4.01E-01	0.05 (0, 0.12)
9	87580001	87585000	NTRK2	-0.105	9.16E-01	0 (0, 0.03)
9	87585001	87590000	NTRK2	-1.141	2.54E-01	0.06 (0, 0.13)
9	87620001	87625000	NTRK2	-0.416	6.78E-01	0.02 (0, 0.09)
9	87635001	87640000	NTRK2	0.091	9.28E-01	0 (0, 0)

African_Sample_2						
Chromosome	Start	End	Gene	*T.statistic*	*P.value*	Effect Size f^2
9	87280001	87285000	NTRK2	-1.324	1.86E-01	0.08 (0, 0.17)
9	**87285001**	**87290000**	**NTRK2**	**-3.072**	**2.32E-03**	0.18 (0.08, 0.27)
9	87305001	87310000	NTRK2	-1.103	2.71E-01	0.06 (0, 0.15)
9	87430001	87435000	NTRK2	-0.758	4.49E-01	0.05 (0, 0.08)
9	87445001	87450000	NTRK2	-1.724	8.57E-02	0.1 (0, 0.2)
9	87485001	87490000	NTRK2	-0.974	3.31E-01	0.06 (0, 0.16)
9	87635001	87640000	NTRK2	-2.505	1.28E-02	0.15 (0.06, 0.25)

Table S3. *NTRK2* regional DNA methylation element (chr9:87285001-87290000) associations with picture arousal and valence ratings, as well as depression and anxiety scores in the Swiss Sample.

Subjects from the healthy Swiss Sample performed several different consecutive tasks as described in detail previously, including encoding task during which also emotional valence and arousal ratings of presented pictures was determined (3). Questionnaires were also used to determine Depression and Anxiety Scores (Montgomery-Asberg Depression Rating Scale - MADRS & State-Trait Anxiety Inventory - STAI Trait; respectively). The association *NTRK2* regional DNA methylation element (chr9:87285001-87290000) with picture arousal & valence ratings, as well as MADRS and STAI was assessed. Fixed effects analysis *t*-statistics and nominal *p*-values are shown. Effect sizes of *NTRK2* regional DNA methylation are provided as Cohen's f^2 (90% CI).

Swiss Sample (*N=568*) ‡	**Phenotype**	***t-statistic***	***p-value***	**Effect Size f^2**
Arousal Ratings	Negative Pictures	-0.04	0.997	0 (0, 0)
	Positive Pictures	-0.33	0.741	0.01 (0, 0.08)
	Neutral Pictures	-1.67	0.092	0.07 (0, 0.14)
Valence Ratings	Negative Pictures	-1.58	0.114	0.07 (0, 0.14)
	Positive Pictures	1.57	0.117	0.07 (0, 0.14)
	Neutral Pictures	-0.94	0.348	0.04 (0, 0.11)
Depression	MADRS Scale	-0.56	0.577	0.02 (0, 0.09)
Anxiety	STAI Trait	0.27	0.789	0.01 (0, 0.07)

Table S4. Significant associations between ICs' subject scores and recognition memory performance in the Healthy Swiss Sample (*N=498*).

IC	t-statistic	p-value
IC4	4.47	9.63E-06
IC7	-7.84	2.89E-14
IC12	8.42	4.07E-16
IC13	-9.58	4.80E-20

Table S5. Anatomical distribution of the highest-loading voxels in IC13. Only clusters of more than 20 voxels are represented. The coverage is expressed as a percentage of the total number of gray matter voxels. Labels follow FreeSurfer's Desikan atlas (http://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/AnatomicalROI/FreeSurferColorLUT).

cluster	coverage extreme negative loadings
ctx_rh_supramarginal	15.46%
ctx_lh_supramarginal	9.11%
ctx_rh_superiorfrontal	6.49%
ctx_rh_rostralmiddlefrontal	6.40%
ctx_rh_precuneus	6.09%
ctx_rh_insula	6.00%
ctx_lh_superiortemporal	5.55%
ctx_rh_inferiorparietal	5.33%
ctx_lh_precuneus	3.95%
ctx_rh_middletemporal	3.29%
ctx_lh_insula	3.15%
ctx_rh_bankssts	2.93%
ctx_rh_parsopercularis	2.93%
ctx_rh_superiortemporal	2.89%
ctx_rh_posteriorcingulate	2.44%
ctx_rh_paracentral	2.09%
ctx_rh_precentral	2.00%
ctx_lh_precentral	1.69%
ctx_rh_superiorparietal	1.42%
ctx_rh_rostralanteriorcingulate	1.38%
ctx_rh_medialorbitofrontal	1.33%
ctx_lh_postcentral	1.07%
ctx_lh_paracentral	0.98%
ctx_lh_middletemporal	0.93%

cluster	coverage extreme positive loadings
ctx_lh_superiorfrontal	19.35%
ctx_lh_rostralmiddlefrontal	12.11%
ctx_lh_caudalmiddlefrontal	11.92%
ctx_lh_inferiorparietal	11.36%
ctx_lh_parsopercularis	7.99%
Right_Cerebellum_Cortex	7.99%
ctx_lh_parstriangularis	7.95%
ctx_lh_middletemporal	3.74%
ctx_lh_parsorbitalis	3.27%
ctx_lh_supramarginal	3.27%
ctx_lh_superiorparietal	2.90%
ctx_lh_precentral	2.06%
ctx_rh_superiorfrontal	2.06%
ctx_lh_lateralorbitofrontal	1.82%
ctx_lh_bankssts	1.03%

(continued)

Table S6. Association of *NTRK2* 5' regulatory region DNA methylation with the expression of the *NTRK2* gene in a TCGA dataset.

NTRK2 5' regulatory region DNA methylation correlations with the expression of the *NTRK2* gene across The Cancer Genome Atlas (TCGA) datasets. Correlation coefficients and nominal p-values are shown (http://www.unimd.org/dnmivd/, (37)).

DNMIVD: DNA Methylation Interactive Visualization Database - TCGA datasets					
Dataset	**GeneSymbol**	**Pearson_r**	**Pearson_pvalue**	**Spearman_r**	**Spearman_pvalue**
BLCA	NTRK2	-0.195	5.37E-05	-0.206	1.92E-05
BRCA	NTRK2	-0.240	0.00E+00	-0.304	4.81E-20
CESC	NTRK2	-0.255	5.70E-06	-0.355	1.00E-10
CHOL	NTRK2	-0.077	6.17E-01	-0.220	1.46E-01
COAD	NTRK2	-0.290	1.54E-07	-0.437	3.77E-16
ESCA	NTRK2	-0.388	1.60E-07	-0.493	0.00E+00
GBM	NTRK2	-0.435	3.29E-04	-0.399	1.09E-03
HNSC	NTRK2	-0.220	4.01E-07	-0.216	6.19E-07
KIRC	NTRK2	-0.212	7.76E-05	-0.249	3.22E-06
KIRP	NTRK2	-0.022	7.07E-01	-0.143	1.35E-02
LIHC	NTRK2	-0.051	3.01E-01	0.104	3.52E-02
LUAD	NTRK2	-0.196	1.59E-05	-0.304	0.00E+00
LUSC	NTRK2	-0.321	2.00E-10	-0.314	5.00E-10
PAAD	NTRK2	-0.172	2.02E-02	-0.032	6.64E-01
PCPG	NTRK2	0.046	5.35E-01	-0.049	5.04E-01
PRAD	NTRK2	-0.023	6.01E-01	0.032	4.67E-01
READ	NTRK2	-0.252	1.10E-02	-0.444	3.40E-06
SARC	NTRK2	-0.159	1.00E-02	-0.226	2.22E-04
SKCM	NTRK2	-0.123	7.79E-03	-0.125	6.85E-03
STAD	NTRK2	-0.315	3.50E-09	-0.391	0.00E+00
THCA	NTRK2	-0.027	5.22E-01	-0.055	1.96E-01
THYM	NTRK2	-0.103	2.61E-01	-0.234	9.84E-03

References

1. Vukojevic V, et al. (2014) Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *Journal of Neuroscience* 34(31):10274–10284.

2. Kolassa I-T, et al. (2010) Association study of trauma load and SLC6A4 promoter polymorphism in posttraumatic stress disorder: evidence from survivors of the Rwandan genocide. *The Journal of clinical psychiatry* 71(5):543–547.

3. de Quervain DJ-F, et al. (2007) A deletion variant of the alpha2b-adrenoceptor is related to emotional memory in Europeans and Africans. *Nature Publishing Group* 10(9):1137–1139.

4. de Quervain DJ-F, et al. (2012) PKCα is genetically linked to memory capacity in healthy subjects and to risk for posttraumatic stress disorder in genocide survivors. *Proceedings of the National Academy of Sciences* 109(22):8746–8751.

5. Wilker S, et al. (2013) The role of memory-related gene WWC1 (KIBRA) in lifetime posttraumatic stress disorder: evidence from two independent samples from African conflict regions. *Biological Psychiatry* 74(9):664–671.

6. Foa EB, Cashman L, Jaycox L, Perry K (1997) The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological Assessment* 9(4):445–451.

7. Wilker S, et al. (2015) How to quantify exposure to traumatic stress? Reliability and predictive validity of measures for cumulative trauma exposure in a post-conflict population. *European Journal of Psychotraumatology* 6:28306.

8. Conrad D, et al. (2017) Does trauma event type matter in the assessment of traumatic load? *European Journal of Psychotraumatology* 8(1):1344079.

9. Kolassa I-T, et al. (2010) Spontaneous remission from PTSD depends on the number of traumatic event types experienced. *Psychological Trauma: Theory, Research, Practice, and Policy* 2(3):169–174.

10. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L (1974) The Hopkins Symptom Checklist (HSCL). A measure of primary symptom dimensions. *Mod Probl Pharmacopsychiatry* 7(0):79–110.

11. Wilker S, et al. (2018) Genetic variation is associated with PTSD risk and

aversive memory: Evidence from two trauma-Exposed African samples and one healthy European sample. *Translational Psychiatry* 8(1):75.

12. Stevens EL, Baugher JD, Shirley MD, Frelin LP, Pevsner J (2012) Unexpected Relationships and Inbreeding in HapMap Phase III Populations. *PLoS ONE* 7(11):e49575.

13. Heck A, et al. (2017) exome sequencing of healthy phenotypic extremes links TROVE2 to emotional memory and PTSD. *Nat hum behav* 1:0081.

14. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 59(3):2142–2154.

15. Ashburner J (2007) A fast diffeomorphic image registration algorithm. *NeuroImage* 38(1):95–113.

16. Klein A, et al. (2009) Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* 46(3):786–802.

17. Egli T, et al. (2018) Identification of Two Distinct Working Memory-Related Brain Networks in Healthy Young Adults. *eNeuro* 5(1). doi:10.1523/ENEURO.0222-17.2018.

18. Hyvärinen A, Oja E (2000) Independent component analysis: algorithms and applications. *Neural Netw* 13(4-5):411–430.

19. Chiappetta P, Roubaud MC, Torrésani B (2004) Blind Source Separation and the Analysis of Microarray Data. *Journal of Computational Biology* 11(6):1090–1109.

20. Himberg J, Hyvärinen A, Esposito F (2004) Validating the independent components of neuroimaging time series via clustering and visualization. *NeuroImage* 22(3):1214–1222.

21. Biton A (2019) *MineICA: Analysis of an ICA decomposition obtained on genomics data*.

22. Assenov Y, et al. (2014) Comprehensive analysis of DNA methylation data with RnBeads. *Nat Meth* 11(11):1138–1140.

23. Davis S, Du P, Bilke S, Triche JT, Bootwalla M (2014) methylumi: Handle Illumina methylation data. R package version 2.28.0. *bioconductororg*.

24. Maksimovic J, Gordon L, Oshlack A (2012) SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips.

Genome Biol 13(6):R44.

25. Bibikova M, et al. (2011) High density DNA methylation array with single CpG site resolution. *Genomics* 98(4):288–295.

26. Ziller MJ, et al. (2013) Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500(7463):477–481.

27. Du P, et al. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 11:587.

28. Fox J, Weisberg S (2011) *An R Companion to Applied Regression, Second Edition* (socserv.socsci.mcmaster.ca, Thousand Oaks (CA)). SAGE.

29. Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007) pcaMethods--a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* 23(9):1164–1167.

30. Chen Y-A, et al. (2013) Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *epigenetics* 8(2):203–209.

31. Price ME, et al. (2013) Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics & Chromatin* 13(1):R44.

32. Milnik A, et al. (2016) Common epigenetic variation in a European population of mentally healthy young adults. *Journal of Psychiatric Research* 83:260–268.

33. Müller F, et al. (2019) RnBeads 2.0: comprehensive analysis of DNA methylation data. *Genome Biol* 20(1):1–12.

34. Smedley D, et al. (2015) The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Research* 43(W1):W589–98.

35. Cohen J (1988) *Statistical power analysis for the behavioral sciences* (L. Erlbaum Associates, Hillsdale, N.J.).

36. Ben-Shachar MS, Makowski D, Lüdecke D (2020) Compute and interpret indices of effect size. *cranr-projectorg*.

37. Ding W, et al. (2020) DNMIVD: DNA methylation interactive visualization database. *Nucleic Acids Research* 48(D1):D856–D862.