Supplementary materials for

Rapid gut dysbiosis induced by stroke exacerbates brain infarction in turn

Authors: Kaiyu Xu[1]*, Xuxuan Gao[1]*, Genghong Xia[2], Muxuan Chen[1], Nianyi Zeng[1], Shan Wang[1], Chao You[2], Xiaolin Tian[2], Huiling Di[1], Wenli Tang[1], Pan Li[1], Huidi Wang[2], Xiuli Zeng[2], Chuhong Tan[2], Fanguo Meng[3], Hailong Li[4], Yan He[1†], Hongwei Zhou[1,5†], Jia Yin[1,2†]

Affiliations:

[1]Microbiome Medicine Center, Department of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, 510282, China.

[2]Department of Neurology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, 510515, China.

[3]Redox Medical Center for Public Health, Soochow University, Suzhou, Jiangsu, 215123, China.

[4]Institute of Molecular Enzymology, Medical college of Soochow University, Suzhou, Jiangsu, 215123, China.

[5]State Key Laboratory of Organ Failure Research, Southern Medical University, Guangzhou, Guangdong, 510515, China.

* These authors contributed equally to this work.

†Correspondence to:

Professor Jia Yin, Microbiome Medicine Center, Department of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, 510282, China, and Department of Neurology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, 510515, China; Tel: +86(20)61641965; Fax: +86(20)62787664; Email: jiajiayin@139.com.

Professor Hongwei Zhou, Microbiome Medicine Center, Department of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, 510282, China and State Key Laboratory of Organ Failure Research, Southern Medical University, Guangzhou, Guangdong, 510515, China; Tel: +8618688489622; Fax: +86(20)62783662; Email: biodegradation@gmail.com.

Professor Yan He, Microbiome Medicine Center, Department of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, 510282, China; Tel: +8613824420986; Fax: +86(20)62783662; Email: bioyanhe@gmail.com.

This file includes:

Supplementary materials and methods

Supplementary results

Supplementary figures S1-S7

Supplementary tables S1-S4

Supplementary materials and methods

Human observational study

The human observational study was approved by the Ethics Committee of Southern Medical University (NFEC-2016-148, NFEC-2020-095). Written informed consent was obtained from all subjects.

Cohort 1 included a total of 28 acute ischaemic stroke (AIS) patients from Yanling Hospital (Southern Medical University, Guangzhou, China) and 28 healthy controls. Faecal samples were collected at the indicated time phases after stroke, namely, T1 (0-4 d post-stroke), T2 (5-7 d post-stroke), T3 (8-30 d post-stroke), and T4 (1-4 months post-stroke). Fresh faecal samples and accompanying demographic information were collected for the control group. Patients diagnosed with AIS were consecutively recruited if they met the following criteria: (1) presented with ischaemic stroke due to cerebral infarction (according to standard clinical criteria with supporting cerebral imaging evidence consisting of either computed tomography or magnetic resonance imaging and magnetic resonance angiography), (2) were >18 years of age, (3) were admitted within 48 h after stroke onset, and (4) had no intravenous thrombolysis or endovascular treatment. Exclusion criteria included the following: (1) evidence of hemorrhagic infarction; (2) severe comorbidities, including neoplastic diseases, massive gastrointestinal hemorrhage, and cirrhosis; (3) unstable medical condition; (4) chronic inflammatory disease; and (5) antibiotic, prebiotic or probiotic use during the preceding 1 month. The control group included 28 healthy people who resided in Guangzhou City for more than 5 years. They had no cardiovascular or cerebrovascular diseases and were matched according to age and gender to the patient group. The exclusion criteria for the control group were the same as those described above.

Cohort 2 included a total of 124 AIS patients from Nanfang Hospital (Southern Medical University, Guangzhou, China). Faecal samples were collected within 48 h after admission. Neurological deficits of AIS patients were assessed by the National Institutes of Health Stroke Scale (NIHSS) on admission and 7 d after admission. Patients were classified into the good primary outcome group (NIHSS score improved ≥40% after 7 d of standard treatments) or the poor primary outcome group (NIHSS score improved <40% after 7 days of standard treatments). Inclusion criteria were as follows: (1) presented with ischaemic stroke due to cerebral infarction, (2) were >18 years of age, (3) were admitted within 7 d after stroke onset, (4) were admitted with an acute moderate-to-severe ischaemic stroke (NIHSS score ≥4) and (5) had no intravenous thrombolysis or endovascular treatment. The exclusion criteria were the same as those described above.

All samples were frozen at -80°C until analysis.

Mice and treatments

All experiments involving mice were approved by the Experimental Animal Ethical Committee of Southern Medical University (No.L2020020). Studies involving animals were performed in compliance with all relevant ethical regulations. Male and female C57BL/6 mice were obtained from Guangdong Medical Laboratory Animal Center and bred at the Experimental Animal Research Center at Southern Medical University. At the beginning of the experiment, all mice were 8 weeks old. The temperature and humidity of the animal room were controlled at 23°C and 40%, respectively, along with a 12-h light/12-h dark cycle. For antibiotic treatment, mice were given a combination of vancomycin (0.5 mg/ml), gentamicin (1 mg/ml), metronidazole (1 mg/ml), and ampicillin (1 mg/ml) in drinking water for 1 week, according to a previous study.[1] For *Escherichia coli* and *E. coli waaL* mutant colonization, mice were treated with antibiotics for 1 week, followed by a 1-d washout period, and were gavaged with 200 μl of PBS-suspended bacteria (OD=1) each day for another week until middle cerebral artery occlusion (MCAO) was conducted.

For aminoguanidine (AG) and tungstate administration, AG (Sigma-Aldrich) and sodium tungstate (W) (Sigma-Aldrich) were used. Superoxide dismutase (SOD) and SOD^{H84A} were obtained from Hangzhou Redox Pharmatech Co., Ltd. (Hangzhou, China).[2] SOD is a pH-stable, temperature-stable and protease-stable enzyme. For the AG, SOD, SOD^{H84A} and W groups, mice were treated with 200 μl of 100 mg/kg body weight AG, 3000 U/ml SOD, 3000 U/ml SOD^{H84A}, and 0.8% w/v W, respectively, at 1 h after MCAO. For the pre-AG, pre-SOD, and pre-W groups, mice were given drinking water treated with 1 mg/ml AG for 5 d, 375 U/d SOD for 3 d, and 0.2% w/v W for 7 d, respectively, before MCAO.

For the dose-response relationship experiment, mice were treated with 200 μl of AG at concentrations of 1 mg/kg, 10 mg/kg, 25 mg/kg, 50 mg/kg, 100 mg/kg, or 200 mg/kg body weight at 1 h after MCAO. For the therapeutic time window experiment, mice were treated with 200 μl of 25 mg/kg body weight AG at different time points after MCAO, including 1 h, 3 h, 6 h, and 12 h post-MCAO.

Middle cerebral artery occlusion

Mice were anesthetized with 1.25% tribromoethanol by intraperitoneal injection (0.02 ml/g body weight). An incision in the midline neck region was made to isolate and ligate the common carotid artery and left external carotid artery. Then, we inserted a 2-mm silicon-coated filament into the right internal carotid artery. After 60 minutes of occlusion, the filament was removed to restore blood flow. During this procedure, a feedback-controlled heating pad was used to maintain the body temperature of the mice. This surgical procedure was completed by an investigator who was blinded to the experimental groups. The exclusion criteria were as follows: (1) the mouse died during the procedure or (2) brain ischemia induction was not observed during histological analysis. At the indicated time points post-MCAO, mice were anesthetized, and samples were collected and stored at -80°C for further analysis.

Evaluation of brain infarct lesions

At the indicated time points post-MCAO, mice were anesthetized before they were perfused with 10 ml of ice-cold PBS. The perfused brains were removed and immediately frozen on powdered dry ice. Then, the perfused brain tissue was cut into coronal cryosections 1 mm thick. The cryosections were further stained with 2,3,5-triphenyltetrazolium chloride (TTC, Sigma-Aldrich), consistent with standard protocols, for 15 minutes and fixed with 4% paraformaldehyde (PFA) for an additional 30 minutes. The stained brain sections were imaged under the microscope for brain infarction quantification. The cerebral infarct volume was measured in each section using ImageJ software (National Institution of Health, Pro Plus 6.0). The total cerebral infarct volume was calculated by integrating the measured volumes of different sections in one mouse brain tissue. We used an edema correction formula for the infarct volume, as previously described.[3] The brain infarction ratio was calculated by the proportion of ischaemic volume to the total volume of brain tissue of one mouse.

16S RNA sequencing and analysis

For the human observational study, bacterial genomic DNA was extracted from faecal samples using a PowerSoil DNA Extraction Kit (Mo Bio) following the manufacturer's specifications. The barcoded primers 514F (GTGCCAGCMGCCGCGGTAA) and 805R (GGACTACHVGGGTWTCTAAT) were used to amplify the 16S rRNA gene V4 variable region. The PCR was performed as described previously,[4] and the PCR cycling conditions were as follows: 94°C for 2 minutes; followed by 30 cycles at 94°C for 30 s, 52°C for 30 s, and 72°C for 45 s; and a final extension at 72°C for 5 minutes. All PCR amplicons were mixed and sequenced by Illumina paired-end sequencing following the manufacturer's protocol. The raw sequences were preprocessed using the BIPES protocol. A QIIME 2 workflow script was used to construct the BIOM file (biological observation matrix format).[5] Operational taxonomic units (OTUs) were selected using the Usearch algorithm.[6] Based on sequence frequency, representative sequences for each OTU were determined. Taxonomic assignments for representative sequences were determined. PyNAST algorithms were used to align representative sequences, and FastTree was used to determine phylogenetic relationships based on representative sequences.[7, 8]

For animal experiments, bacterial genomic DNA was extracted using a MinkaGene Stool DNA Kit (Guangzhou mCHIP BioTech) following the manufacturer's specifications. The barcoded primers V4F (GTGYCAGCMGCCGCGGTAA) and V4R (GGACTACNVGGGTWTCTAAT) were used to amplify the 16S rRNA gene V4 variable region. PCR was performed as described previously.[4] All PCR amplicons were mixed and sequenced using an Illumina iSeq 100 following the manufacturer's protocol. The raw sequences were preprocessed and quality controlled using QIIME 2 with default parameters.[9] A QIIME workflow script, pick_closed_reference_otus.py,

was used to perform referenced-based OTU clustering, and USEARCH61 was used in the reference mode against the Greengenes database v13_8 and for BIOM generation.

The sequences were clustered into species-level OTUs with 97% similarity. Phylogenetic relationships and the taxonomy of representative OTU sequences were determined using the Greengenes database in default mode.

Both the Shannon index and PD whole tree index represent α diversity. The Shannon index indicates the number and distribution of microbial species in a sample. The PD whole tree index indicates the range of phylogenetic distances among microbial species. The UniFrac distances were used to analyse the β diversity by illustrating the phylogenetic dissimilarity among samples. A smaller UniFrac distance between two samples indicates a higher similarity. As a dimensionality reduction method, principal coordinate analysis (PCoA) was used to describe the relationship among samples based on the distance matrix and visualize the unsupervised grouping pattern of the complex data set, i.e., the microbiome. Linear discriminant analysis effect size (LEfSe) was used to compare the discriminative data between groups.[10] As an algorithm for high-dimensional biomarker discovery, LEfSe identifies genomic features that characterize differences between two or more biological conditions. By emphasizing statistical significance, biological consistency and effect relevance, LEfSe determines the abundant features most different between conditions in accordance with biologically meaningful categories. For each differential feature detected by LEfSe, we calculated a linear discriminant analysis value, representing the difference in the feature between groups.

E. coli abundance

The mouse jejunum, ileum, cecum, and colon contents in the sham group and MCAO group at the 1 d time point were extracted using a MinkaGene Stool DNA Kit (Guangzhou mCHIP BioTech) following the manufacturer's specifications. A 2-μl sample of the bacterial DNA was used as the template for SYBR Green-based real-time PCR. The gene copy number was determined according to a standard curve generated using a DNA fragment of the respective bacterial 16S rRNA gene. The *E. coli* fraction as part of the whole bacterial population was calculated by dividing the gene copy number of *E. coli* by the total gene copy number using 16S rRNA V1 primers. The primers used in this study are listed in online supplementary table S3.

Evaluation of cecum blood flow using laser speckle imaging

Mouse cecum blood flow was recorded before and after MCAO and sham operations at various time points using an RFLSI Pro+ laser speckle system and review software (RWD). Mice were anesthetized and placed on a warming plate maintained at 37°C. For cecum blood flow measurements, mouse hair was removed from the abdomen by using an electrical shaver, and a small incision was made to separate the cecum. Care was taken to not stretch or puncture blood vessels, and if excessive bleeding was observed, mice were terminated and excluded from the analysis. The exposed cecum

was kept moist with saline prewarmed to 37°C. Data were acquired from a 1.6 cm × 1.6 cm field of view using a 785 nm, 90 mW laser with a sampling rate of 60 Hz at a working distance of 11 cm. Flux over time was analysed by LSCI review software (RWD). In each visible cecum branching vessel, regions of interest (ROIs) were measured based on the intensification of blood flow. All assessments and analyses were performed in a blinded manner.

Blood samples

Mouse blood samples were collected from the inferior vena cava and centrifuged to obtain serum for biochemical analyses. Mouse serum lipopolysaccharide (LPS), LPS-binding protein (LBP), D-lactate (D-lac), tumor necrosis factor alpha (TNF-α), interleukin (IL)-6, IL-1β, malonaldehyde (MDA), and 8-hydroxy-2-deoxyguanosine (8-OHdG) levels were determined using ELISA kits (Elisalab) according to the manufacturer's protocols.

Measurement of nitrate concentrations in the mouse mucus layer

Cecum mucus layers were collected at 3 h, 6 h, and 1 d post-MCAO. Then, they were extracted with 0.2 ml of ultrapure water and centrifuged at 20,000×g for 2 minutes. The filter sterilized supernatant was stored at -80°C until analysis. The nitrate concentration in the mouse mucus layer was determined using a Nitrite/Nitrate Assay Kit (Sigma-Aldrich) according to the manufacturer's protocol.

Quantitative determination of mRNA levels

The relative transcription levels of mRNAs for genes encoding NADPH oxidase 1 (NOX1), dual oxidase 2 (DUOX2), nitric oxide synthase 2 (NOS2), zonula occludens-1 (ZO-1), occludin, claudin-2, Toll-like receptor-4 (TLR4), TNF receptor-associated factor 6 (TRAF6), Toll receptor-associated molecule (TRAM), myeloid differentiation primary response 88 (Myd88), nuclear factor kappa light chain enhancer of activated B cells (NF-κB), TNF-α, IL-6, IL-17, interferon-γ (IFN-γ), IL-1β, chemokine (C-X-C motif) ligand 1, and chemokine (C-X-C motif) ligand 2 encoded by the *Nox1, Duox2, Nos2, Tjp1, Ocln, Cldn2, Tlr4, Traf6, Tram, Myd88, Nfkb, Tnf, Il6, Il17, Ifng, Il1b, Kc,* and *Cxcl2* genes, respectively, in mice colon tissue were determined by qRT-PCR. Thawed tissues in TRIzol reagent (Invitrogen) were immediately homogenized at 70 Hz for 2 minutes in a tissue homogenizer (JXFSTPRP-32) with sterile steel balls. RNA was extracted following the standard protocol from Thermo Fisher Scientific. cDNA was generated with Takara reverse transcription reagents. Real-time PCR was performed using SYBR Premix Ex Taq™ II (Takara), and data were acquired with a ViiA 7 real-time PCR system (Takara). The comparative C_t method was used to analyse the data. Target gene transcription of each sample was normalized to the respective levels of *Gapdh* mRNA. The primers used in this study are listed in online supplementary table S3.

RNA isolation and sequencing

Mouse cecum tissue samples were snap-frozen in liquid nitrogen. Thawed tissues were immediately homogenized using a tissue homogenizer (OMNI-TH) for 10 s in TRIzol reagent (Takara). Total RNA was isolated following the standard protocol from Takara. The total RNA was further purified with VAHTS RNA Clean Beads (Vazyme). RNA quality was assessed on a Bioanalyzer 2100 (Agilent). RNA libraries were prepared with a VAHTS^{TM} mRNA-seq V2 Library Prep Kit for Illumina (Vazyme). RNA sequencing was performed with an Illumina NovaSeq 6000 using the 150 bp paired-end mode. Read quality control was performed with FastQC. Each read was mapped onto the mouse genome using Hisat2. Counts per gene were assessed using Stringtie software. Differential gene expression analysis was performed using edgeR. Adjusted log2 fold change and adjusted p-values (corrected for multiple testing) were used to determine the significant differences in gene expression. Pathway analysis was performed using the WEB-based GEne SeT AnaLysis Toolkit (http://www.webgestalt.org/).

Nissl staining

For the sham group, MCAO group, AG group, SOD group, and W group, mice were anesthetized before cardiac perfusion with saline and fixation with PFA. Brain tissue was removed carefully, postfixed in 4% PFA for 24 h, and then cryoprotected with 30% sucrose for 48 h. Serial frozen coronal sections (5 μm thick) were cut using a cryostat (Leica, CM1950) at -20°C, and all sections containing the hippocampus were stored at -20°C. Then, the sections were dried in air, washed twice with distilled water (2 minutes/time), and stained with 1% toluidine blue for 5 minutes. The sections were washed 3 times with distilled water (3 minutes/time), placed in 70% ethanol for 2 minutes, washed twice with 95% ethanol (2 minutes/time), followed by washing with xylene for 5 minutes. Sections were observed by a microscope (Leica DM2500).

In vitro anaerobic growth assays

Brain heart infusion (BHI) broth was prepared by dissolving BHI broth (Sigma-Aldrich) in sterile water. Sodium nitrate was added at final concentrations of 0 μM, 100 μM, or 200 μM. Next, 2.2 ml of BHI broth was inoculated with 200 μL of each mouse faecal mixture and incubated anaerobically for 48 h at 37°C. The abundance of Enterobacteriaceae was analysed by using the 16S rRNA sequence as described previously.

Strain isolation and identification

Ten-fold serial dilutions of the faecal contents of C57BL/6 MCAO mice or faecal samples of an AIS patient were plated on MacConkey medium (10 g/l pancreatic digest of gelatin, 3 g/l peptone, 10 g/l lactose, 1.5 g/l bile salts, 5 g/l sodium chloride, 13.5 g/l agar, 30 mg/l neutral red, 1 mg/l crystal violet) and incubated aerobically at 37°C overnight. To isolate murine *E. coli* strains, single colonies were isolated and subjected to species identification using MALDI-TOF and 16S rDNA sequencing.

Plasmids and construction of mutants

All primers are listed in online supplementary table S3. The bacterial strains and plasmids are listed in online supplementary table S4. The *E. coli* strains used in this study were grown on LB plates, and 100 mg/l ampicillin or 30 mg/l chloramphenicol was added if needed. Gene mutagenesis was performed with Lambda Red recombinase.[11] PCR products (FRT-ChlR cassettes) were generated with primers with 50-nt extensions that are homologous to regions adjacent to the gene to be inactivated and the template plasmid pKD3 carrying a chloramphenicol resistance gene flanked by FLP recognition target (FRT) sites. Mutants were isolated as chloramphenicol-resistant colonies after the respective FRT-ChlR cassettes were introduced into bacteria carrying the Lambda Red expression plasmid pKD46. Deletion of the target gene was confirmed by PCR and sequencing. The plasmid pKD46 was then easily cured when the mutants were subcultured at 37 °C.

Data availability

The sequencing data have been deposited at the European Nucleotide Archive (ENA) database under the accession nos. PRJEB38502, PRJEB38503, and PRJEB38504.

Statistical analysis

Univariate and multivariate logistic regression analyses were performed, and odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to identify risk factors for stroke outcome. Confounding factors used in the multivariate logistic regression analysis include age, gender, smoking history, diabetes, hypertension, coronary artery disease, white blood cell, hemoglobin A1c, creatinine, uric acid, glucose, triglycerides, total cholesterol, high-density lipoprotein, low-density lipoprotein, homocysteine, NIHSS, and Enterobacteriaceae.

Supplementary results

MCAO causes significant *E. coli* overgrowth in female mice

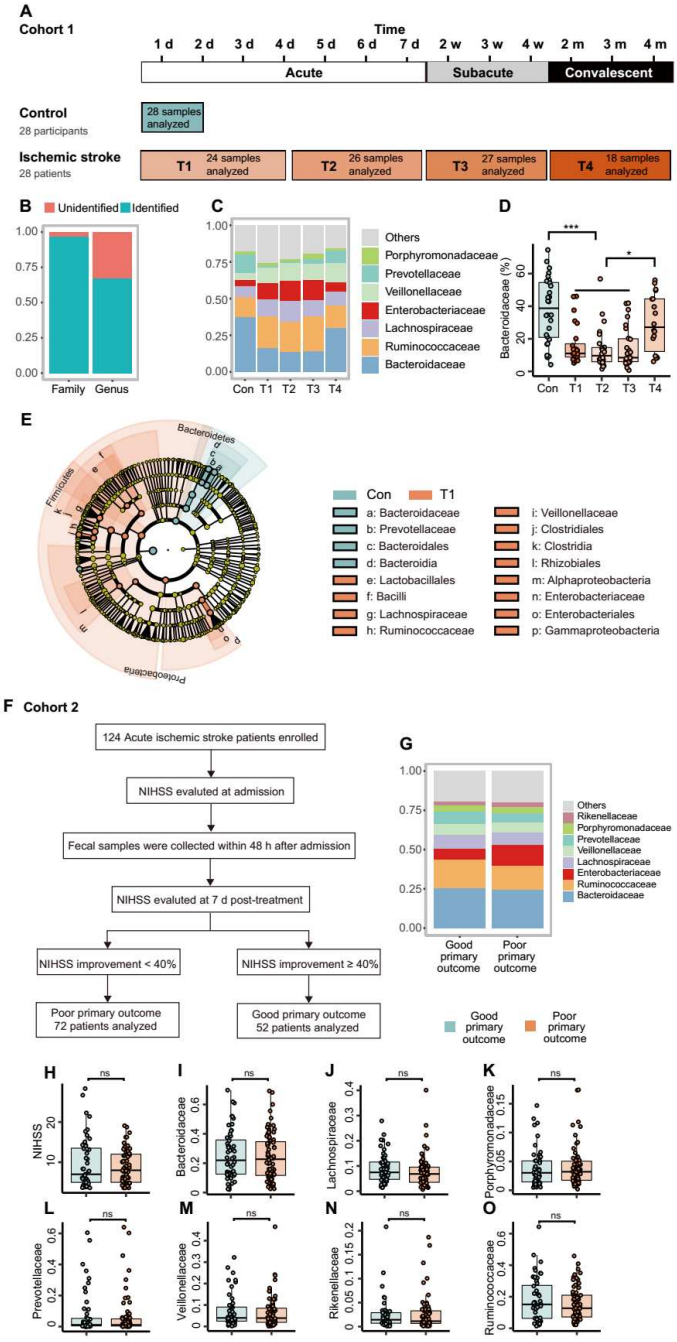
Because gender differences may interfere with ischaemic stroke sensitivity[12] and gut microbiota composition,[13] we applied the MCAO model to age-matched female C57BL/6 mice. We found that the female mice showed lower cerebral infarction ratios compared with the male mice at 1 d post-stroke (online supplemental figure S3A, B). According to the qPCR results with *E. coli* specific primers, *E. coli* showed significant overgrowth from hundreds to thousands of folds in the four GI sections of MCAO group at 1 d post-stroke compared with sham group at 1 d time point in female mice (online supplemental figure S3C-F). These results indicated that the stroke induced *E. coli* overgrowth is gender independent.

Preventive roles of AG, SOD and tungstate in ischaemic stroke

From another perspective, we wondered whether the preventive use of AG, SOD and tungstate would also ameliorate brain infarction. Here, we treated mice with AG, SOD or tungstate for 3 to 7 days before MCAO, and they were referred to as the pre-AG, pre-SOD and pre-W groups, respectively (online supplemental figure S5B, S6A, and S7A). We found that pretreatment with AG, SOD or tungstate successfully suppressed Enterobacteriaceae overgrowth as well as gut microbiome dysbiosis to mimic that of the sham group (online supplemental figure S5D-I, S6C-H, and S7B-G). Accordingly, the brain infarction in the pre-AG, pre-SOD and pre-W groups was also significantly less than that of the MCAO group (online supplemental figure S5C, S6B, and S7H).

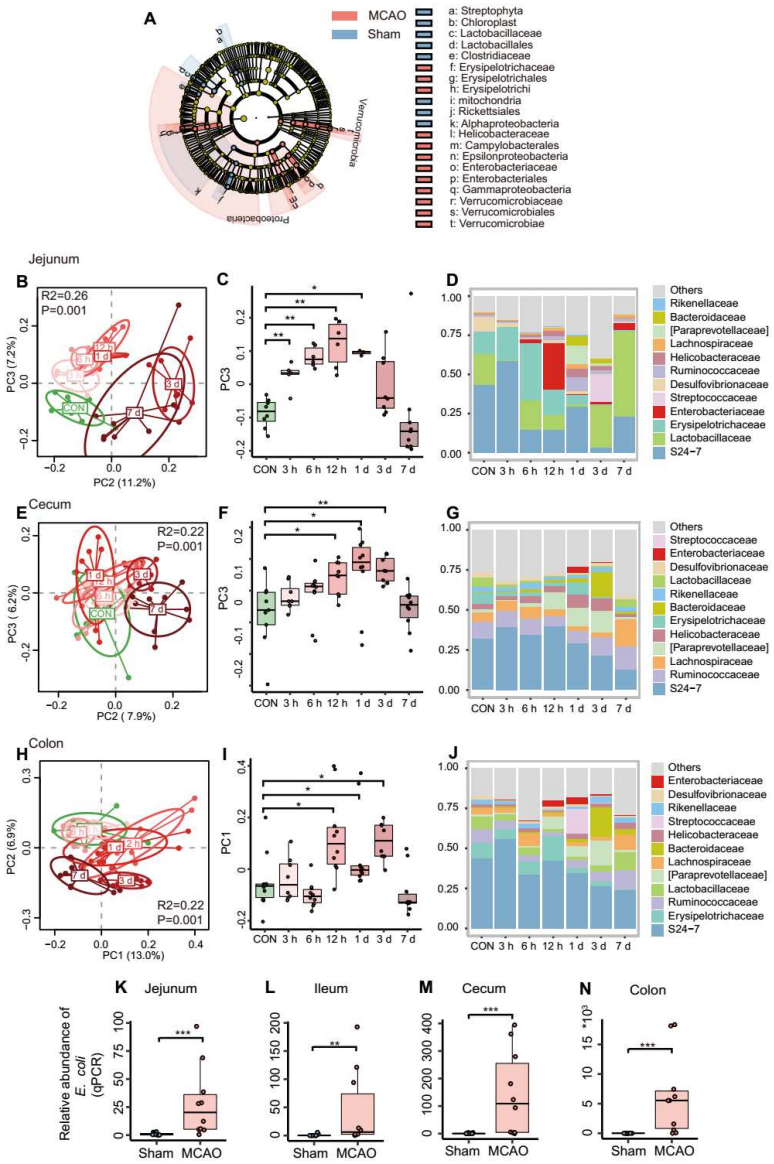
Preventive administration of AG, SOD or W restored gut barrier dysfunction as demonstrated by the LPS, LBP, D-lac and MDA results and the relative expression of tight junction protein genes, including *Cldn2*, *Tjp1* and *Ocln* (online supplemental figure S5J-L, O-Q, S6I-K, N-P, and S7I, J, M, N). These approaches alleviated systematic inflammation as demonstrated by the significantly decreased serum TNF-α and IL-6 concentrations and expression of proinflammatory cytokine genes, such as *Tnf*, *Il6*, *Il1b*, *Kc*, *Cxcl2*, *Ifng*, and *Il17* (online supplemental figure S5M, N, R-X, S6L, M, Q-W, and S7K, L, O-R). However, whether the preventive effects of the three inhibitors were due to the prevention of stroke induced gut dysbiosis or improvement of the baseline gut microbiome and immune system remains unclear. Future studies monitoring the effects of AG, SOD and tungstate on baseline gut microbiome and immune system may unravel the mechanisms of their preventive effects on brain infarction post-stroke.

Supplementary figures S1-S7:



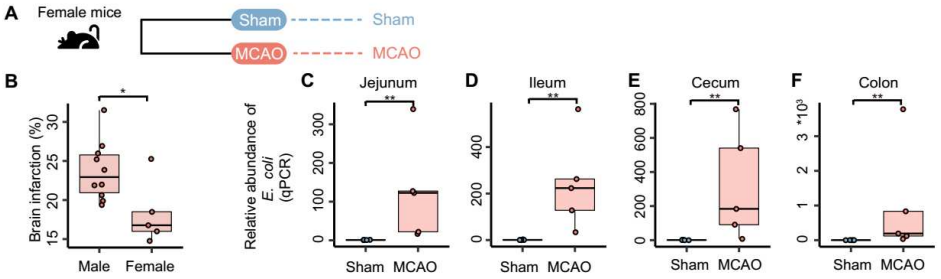
Supplementary figure S1. Trial profile and gut microbiota composition of cohort 1 and cohort 2. A, Trial profile of cohort 1. **B,** The relative abundances of identified and unidentified bacteria at the family and genus levels. **C,** Gut microbial composition at the family level of cohort 1. **D,** Relative Bacteroidaceae abundance in cohort 1. **E,** Cladogram based on LEfSe revealing differential gut microbiota between healthy controls and the T1 group (cohort 1). **F,** Trial profile of cohort 2. (**G** to **O**) Gut microbial

composition at the family level (G), NIHSS scores (H), and relative abundances of bacterial families, including Bacteroidaceae (I), Lachnospiraceae (J), Porphyromonadaceae (K), Prevotellaceae (L), Veillonellaceae (M), Rikenellaceae (N), and Ruminococcaceae (O), between the good and poor primary outcome groups of cohort 2. Medians ± quartiles are plotted. * $p<0.05$ and *** $p<0.001$ based on the Wilcoxon rank sum test.

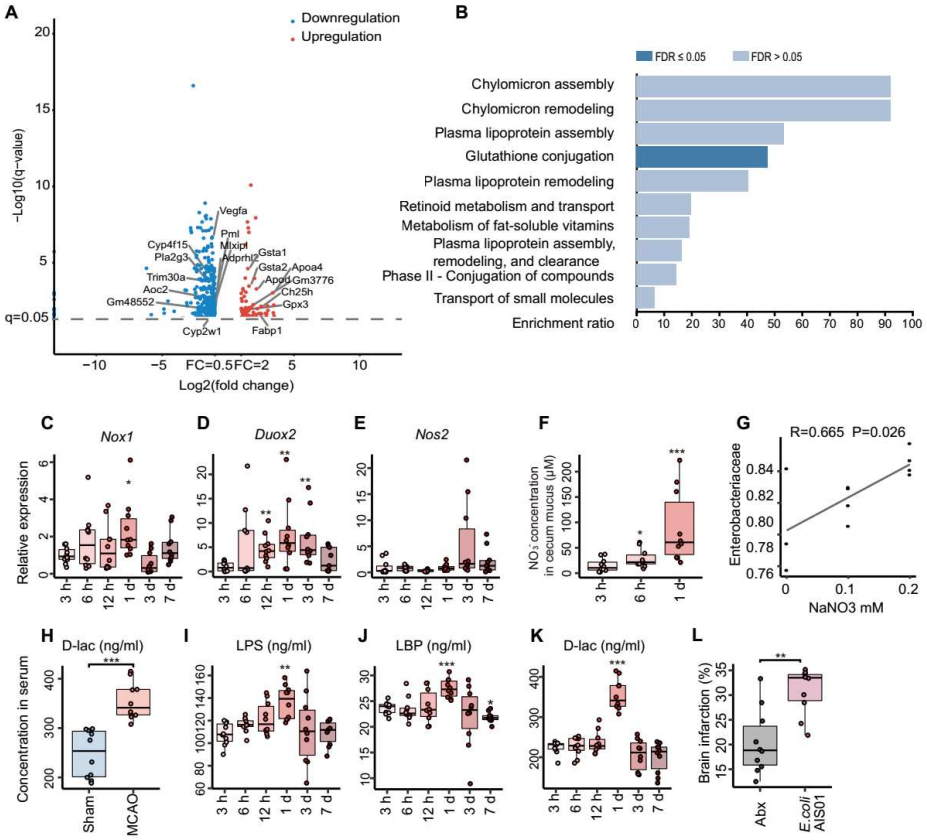


Supplementary figure S2. Temporal shifts in the gut microbiota in the MCAO groups at all time points. A, Cladogram based on LEfSe between the sham group and MCAO group at the 1 d time point. (**B** to **D**) PCoA based on unweighted UniFrac distances (B), time series analysis of the community structure on principal coordinate 3 (PC3) (C), and microbiota composition (D) at the jejunum section among the control group and all MCAO groups. (**E** to **G**) PCoA based on unweighted UniFrac distances (E), time series analysis of the community structure on PC3 (F), and microbiota composition (G) at the cecum section among the control group and all MCAO groups. (**H** to **J**) PCoA based on unweighted UniFrac distances (H), time series analysis of the community structure on PC1 (I), and microbiota composition (J) at the colon section among the control group and all MCAO groups. (**K** to **N**) Relative *E.*

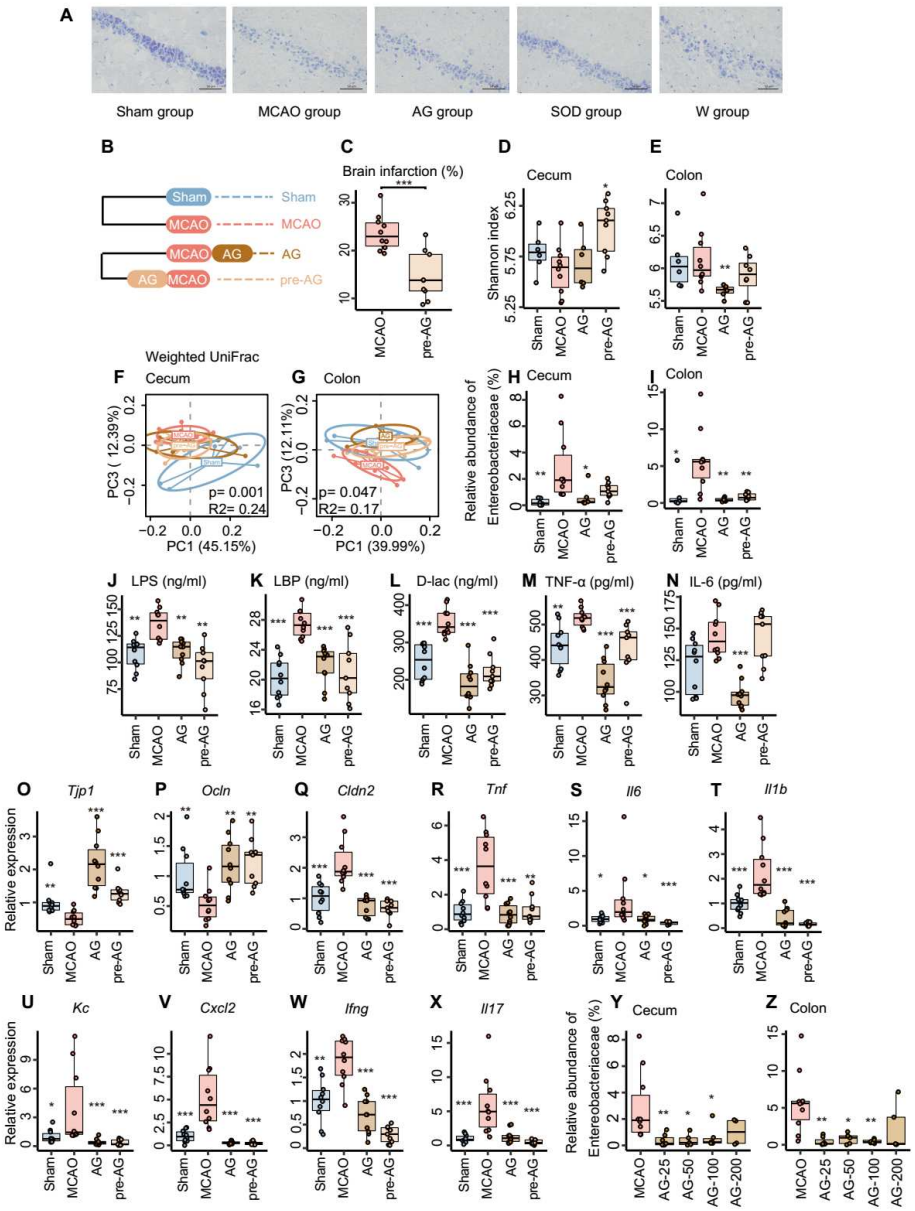
coli abundance using real-time PCR in the jejunum (K), ileum (L), cecum (M), and colon (N) sections between the sham group and MCAO group at the 1 d time point (n=10 mice/group). Medians ± quartiles are plotted. * $p<0.05$, ** $p<0.01$ and *** $p<0.001$ based on the Wilcoxon rank sum test.



Supplementary figure S3. MCAO causes significant *E. coli* overgrowth in female mice. (**A**) Animal experiment design for the sham group (n=5 mice) and MCAO group (n=5 mice) in female mice. (**B**) Brain infarction ratios in the MCAO mice between the male group and the female group. (**C** to **F**) Relative *E. coli* abundance in the jejunum (C), ileum (D), cecum (E), and colon (F) sections using real-time PCR between the sham group and MCAO group at the 1 d time point in female mice (n=5 mice/group). Medians ± quartiles are plotted. * $p<0.05$ and ** $p<0.01$ based on the Wilcoxon rank sum test.

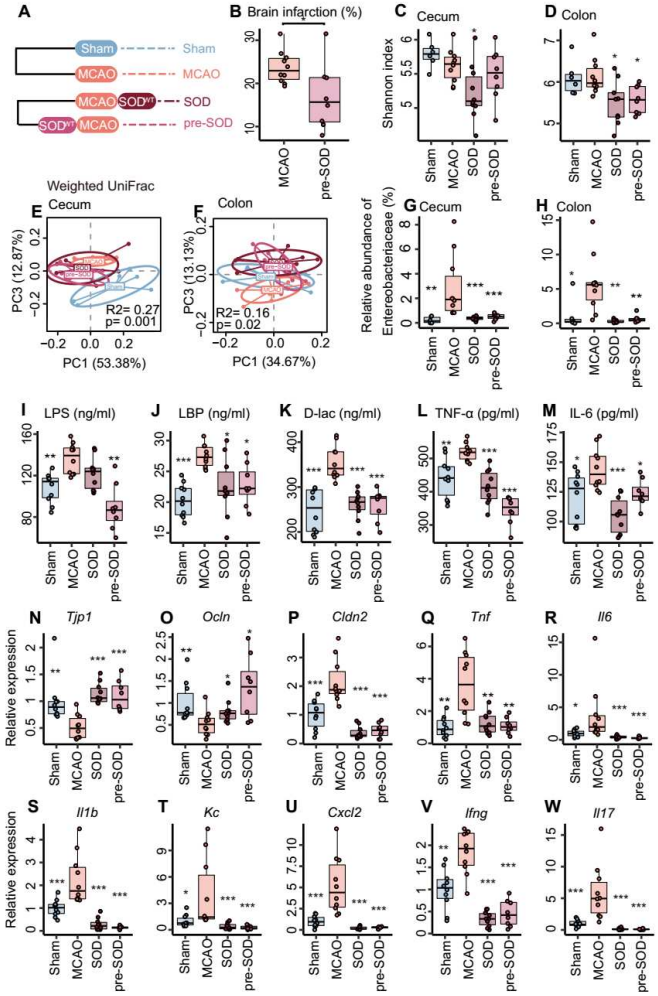


Supplementary figure S4. Gut dysbiosis is caused by stroke-induced intestinal ischemia and reperfusion through nitrate respiration. (**A** and **B**) RNA sequencing analysis of cecum tissue from the MCAO group compared with that from the sham group. Volcano plot showing significantly altered gene expression between the two groups (n=5 mice/group; A). Oxidative stress-related genes are annotated. Pathway analysis revealing that the genes involved in the glutathione conjugation pathway were significantly enriched in the MCAO group (B). (**C** to **K**) Temporal shifts in intestinal ischemia- and reperfusion-related genes, nitrate concentrations and gut barrier biomarkers in the MCAO groups. Relative *Nox1* (C), *Duox2* (D), and *Nos2* (E) expression in the MCAO groups at all time points (n=10 mice/group). Nitrate concentrations in mouse cecum mucus layers in the MCAO groups at the 3 h, 6 h, and 1 d time points (n=10 mice/group; F). Correlation analysis of the relative Enterobacteriaceae abundance and sodium nitrate concentrations in vitro (G). Serum D-lac levels between the sham group and MCAO group (H). Serum LPS (I), LBP (J) and D-lac (K) levels among all the MCAO groups. (**L**) *E. coli* AIS01 significantly increased brain infarctions compared with the Abx group. Medians ± quartiles are plotted. ** p<0.01 and *** p<0.001 based on the Wilcoxon rank sum test.



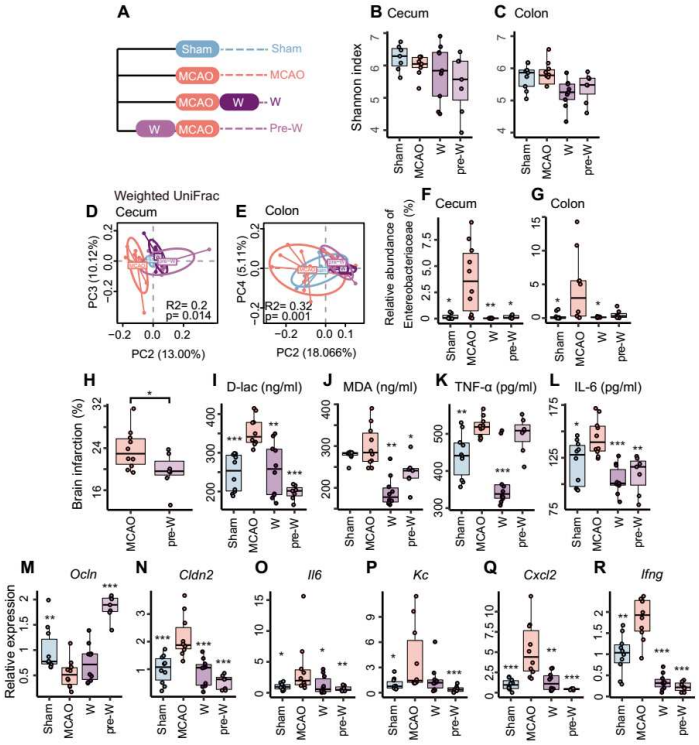
Supplementary figure S5. Aminoguanidine improves the gut microbiota composition, reduces Enterobacteriaceae abundance, suppresses gut barrier dysfunction and systemic inflammation, and alleviates brain infarction. A, Representative images of hippocampal neurons for the experiments shown in Fig. 5. Nissl-stained coronal sections of the hippocampus (CA1) area in the sham group, MCAO group, AG group, SOD group and W group. Scale bar, 50 μm. **B**, Animal experiment design for the sham group (n=10 mice), MCAO group (n=10 mice), aminoguanidine (AG) group (n=10 mice), and pre-AG group (n=9 mice). **C**, Brain infarction ratios between the MCAO group and the pre-AG group. (**D** and **E**) Shannon index of

the cecum (D) and colon (E) contents in AG-treated mice. (**F** and **G**) PCoA based on weighted UniFrac distances showing the cecum (F) and colon (G) content distribution among the sham group, MCAO group, AG group and pre-AG group. (**H** and **I**) Relative Enterobacteriaceae abundance in the cecum (H) and colon (I) contents of AG-treated mice. (**J** to **N**), Serum LPS (J), LBP (K), D-lac (L), TNF-α (M), and IL-6 (N) levels among the sham group, MCAO group, AG group, and pre-AG group. (**O** to **X**) Relative *Tjp1* (O), *Ocln* (P), *Cldn2* (Q), *Tnf* (R), *Il6* (S), *Il1b* (T), *Kc* (U), *Cxcl2* (V), *Ifng* (W), and *Il17* (X) expression among the sham group, MCAO group, AG group, and pre-AG group. (**Y** to **Z**) Relative Enterobacteriaceae abundance in the cecum (Y) and colon (Z) contents for the AG dose-response relationship. Medians ± quartiles are plotted. * $p<0.05$, ** $p<0.01$ and *** $p<0.001$ based on the Wilcoxon rank sum test.



Supplementary figure S6. Superoxide dismutase improves the gut microbiota composition, reduces Enterobacteriaceae abundance, suppresses gut barrier dysfunction and systemic inflammation, and alleviates brain infarction. A, Animal experiment design for the sham group (n=10 mice), MCAO group (n=10 mice), SOD group (n=10 mice), and pre-SOD group (n=8 mice). **B**, Brain infarction ratios between the MCAO group and the pre-SOD group. (**C** and **D**) Shannon index of the cecum (C) and colon (D) contents in SOD-treated mice. (**E** and **F**) PCoA based on weighted UniFrac distances showing the cecum (E) and colon (F) content distribution among the sham group, MCAO group, SOD group and pre-SOD group. (**G** and **H**) Relative Enterobacteriaceae abundance in the cecum (G) and colon (H) contents of SOD-treated mice. (**I** to **W**) Serum LPS (I), LBP (J), D-lac (K), TNF-α (L), and IL-6 (M) levels and relative *Tjp1* (N), *Ocln* (O), *Cldn2* (P), *Tnf* (Q), *Il6* (R), *Il1b* (S), *Kc* (T), *Cxcl2* (U), *Ifng* (V), and *Il17* (W) expression among the sham group, MCAO group, SOD group, and pre-SOD group. Medians ±

quartiles are plotted. ** $p<0.01$ and *** $p<0.001$ based on the Wilcoxon rank sum test.



Supplementary figure S7. Tungstate improves gut microbiota composition, reduces Enterobacteriaceae abundance, suppresses gut barrier dysfunction and systemic inflammation, and alleviates brain infarction. A, Animal experiment design for the sham group (n=10 mice), MCAO group (n=10 mice), W group (n=10 mice), and pre-W group (n=7 mice). (**B** and **C**) Shannon index of the cecum (B) and colon (C) contents in W-treated mice. (**D** and **E**) PCoA based on weighted UniFrac distances showing the cecum (D) and colon (E) content distribution among the sham group, MCAO group, W group and pre-W group. (**F** and **G**) Relative Enterobacteriaceae abundance in the cecum (F) and colon (G) contents of tungstate-treated mice. **H**, Brain infarction ratios between the MCAO group and the pre-W group. (**I** to **R**) Serum D-lac (I), MDA (J), TNF-α (K), and IL-6 (L) levels and relative *Ocln* (M), *Cldn2* (N), *Il6* (O), *Kc* (P), *Cxcl2* (Q), and *Ifng* (R) expression among the sham group, MCAO group, W group, and pre-W group. Medians ± quartiles are plotted. ** p<0.01 and *** p<0.001 based on the Wilcoxon rank sum test.

Supplementary tables S1-S4

Supplementary table S1. Baseline characteristics of ischaemic stroke patients in cohort 1.

	Control group	IS group (T1)	P value
Number	28	28	1
Age	56 (3.75)	54 (15.75)	0.786
Sex, M/F	22/6	22/6	1
WBC, 10^9/L	6.86 (2.02)	7.56 (3.47)	0.039
RBC, 10^9/L	5.19 (0.66)	4.83 (0.88)	0.009
HGB, g/L	151 (11)	141 (23)	0.006
PLT, 10^9/L	259 (72)	228 (70)	0.011
TP, g/L	72.50 (4.13)	62.90 (9.37)	0.001
ALB, g/L	43 (3.57)	40.55 (6.75)	0.019
Cr, μmol/L	74.5 (29.50)	58.5 (24.75)	0.010
UA, μmol/L	372.50 (74.5)	326.50 (190.75)	0.248
Glu, mmol/L	4.86 (1.13)	4.56 (1.90)	0.694
TC, mmol/L	5.74 (1.71)	4.37 (1.91)	0.003
TG, mmol/L	1.53 (0.89)	1.56 (1.47)	0.471
HDL, mmol/L	1.23 (0.33)	1.17 (0.43)	0.238
LDL, mmol/L	3.56 (1.32)	2.48 (1.46)	0.003

Data represent the median (interquartile range). WBC, white blood cell count; RBC, red blood cell count; HGB, hemoglobin; PLT, platelets; TP, total protein; ALB, albumin; Cr, creatinine; UA, uric acid; Glu, glucose; TC, total cholesterol; TG, triglycerides; HDL, high-density lipoprotein; and LDL, low-density lipoprotein.

Supplementary table S2. Baseline characteristics of ischaemic stroke patients in cohort 2.

	Good primary outcome group	Poor primary outcome group	P value
Number	52	72	
Age	55.50 (18.75)	62 (18)	0.164
Sex, M/F	30/22	50/22	0.177
NIHSS on admission	12 (12.75)	9 (7)	0.976
WBC, 10^9/L	8.56 (4.05)	9.28 (3.37)	0.419
RBC, 10^9/L	4.67 (0.90)	4.90 (0.92)	0.534
HGB, g/L	134.50 (28.75)	133 (20)	0.710
PLT, 10^9/L	249 (61.25)	224 (91.50)	0.189
TP, g/L	72.55 (10.05)	71 (7.65)	0.153
ALB, g/L	43.35 (4.87)	42.90 (4.60)	0.347
Cr, μmol/L	78 (33.75)	82 (20)	0.172
UA, μmol/L	396.50 (157.25)	361 (177.50)	0.603
Glu, mmol/L	7.41 (2.88)	6.97 (2.47)	0.226
TC, mmol/L	4.81 (1.36)	5.14 (1.47)	0.262
TG, mmol/L	1.57 (1)	1.30 (1.19)	0.683
HDL, mmol/L	0.97 (0.35)	0.98 (0.27)	0.064
LDL, mmol/L	2.84 (1.06)	3.16 (1.22)	0.219
HbA1c, %	6.10 (1.50)	6 (2)	0.860
HCY, μmol/L	12 (4)	16 (6)	0.004

Data represent the median (interquartile range). NIHSS, national institutes of health stroke scale; WBC, white blood cell count; RBC, red blood cell count; HGB, hemoglobin; PLT, platelets; TP, total protein; ALB, albumin; Cr, creatinine; UA, uric acid; Glu, glucose; TC, total cholesterol; TG, triglycerides; HDL, high-density lipoprotein; LDL, low-density lipoprotein; HbA1c, hemoglobin A1c; and HCY, homocysteine.

Supplementary table S3. Primers used in this study.

Organism	Target	Sequence
Mus musculus	*Duox2*	5'-ACGCAGCTCTGTGTCAAAGGT-3'
		5'-TGATGAACGAGACTCGACAGC-3'
M. musculus	*Nox1*	5'-GTGATTACCAAGGTTGTCATGC-3'
		5'-AAGCCTCGCTTCCTCATCTG-3'
M. musculus	*Nos2*	5'-CCAGCCTTGCATCCTCATTGG-3'
		5'-CCAAACACCAAGCTCATGCGG-3'
M. musculus	*Tjp1*	5'-GGGGCCTACACTGATCAAGA-3'
		5'-TGGAGATGAGGCTTCTGCTT-3'
M. musculus	*Ocln*	5'-ACGGACCCTGACCACTATGA-3'
		5'-TCAGCAGCAGCCATGTACTC-3'
M. musculus	*Cldn2*	5'-TGGTTCCTGACAGCATGAAA-3'
		5'-CTTTGGGCTGTTGAGCAGAT-3'
M. musculus	*Tlr4*	5'-CTGGGTGAGAAAGCTGGTAA-3'
		5'-AGCCTTCCTGGATGATGTTGG-3'
M. musculus	*Traf6*	5'-CATCTTCAGTTACCGACAGCTCAG-3'
		5'-TGGTCGAGAATTGTAAGGCGTAT-3'
M. musculus	*Tram*	5'-GGCCTGGACCATCTTGTTAC-3'
		5'-CATGGGTATGACGGAGTTGT-3'
M. musculus	*Myd88*	5'-GTTGTGTGTGTCCGACCGT-3'
		5'-GTCAGAAACAACCACCACCATGC-3'
M. musculus	*Nfkb*	5'-CCAAAGAAGGACACGACAGAATC-3'
		5'-GGCAGGCTATTGCTCATCACA-3'
M. musculus	*Tnf*	5'-TTGGGTCTTGTTCACTCCACGG-3'
		5'-CCTCTTTCAGGTCACTTTGGTAGG-3'
M. musculus	*Il6*	5'-CGGAGGCTTGGTTACACATGTT-3'
		5'-CTGGCTTTGTCTTTCTTGTTATC-3'
M. musculus	*Il17*	5'-ACGTTTCTCAGCAAACTTAC-3'
		5'-CCCCTTTACACCTTCTTTTC-3'
M. musculus	*Ifng*	5'-GCTCTGAGACAATGAACGCTACAC-3'
		5'-TTCTTCCACATCTATGCCACTTGAG-3'
M. musculus	*Il1b*	5'-AAGGGCTGCTTCCAAACCTTTGAC-3'
		5'-TGCCTGAAGCTCTTGTTGATGTGC-3'
M. musculus	*Cxcl1*	5'-TGCACCCAAACCGAAGTCAT-3'
		5'-TTGTCAGAAGCCAGCGTTCAC-3'
M. musculus	*Cxcl2*	5'-TCCAGGTCAGTTAGCCTTGC-3'
		5'-CGGTCAAAAAGTTTGCCTTG-3'
M. musculus	*Gapdh*	5'-TGTAGACCATGTAGTTGAGGTCA-3'
		5'-AGGTCGGTGTGAACGGATTTG-3'
E. coli	16S rRNA	5'-GTGCCAGCMGCCGCGGTAA-3'
		5'-GCCTCAAGGGCACAACCTCCAAG-3'
V1	16S rRNA	5'-AGAGTTTGATCCTGGCTCAG-3'
		5'-TTACTCACCCGTNCGCCRCT-3'

E. coli	*dwaaL*	5'-TTATCTGGAGGATGGTAAGACTGGTAAATAGAAAAAAGAATAAACGTTAGagcgattgtgtaggctggag-3' 5'-GCAAATGCTGTTAAGCGGTTCATTGAAGAAAACGCTGCCATGATTTAAAttaacggctgacatgggaattag-3'
E. coli	*VwaaL*	5'-CAACATGAGACGGGAGGAG-3' 5'-ACAGCGGCACTGGATAGAC-3'

Supplementary table S4. Bacterial strains and plasmids used in this study.

Designation	Genotype or phenotype	Source or reference
pKD46	Lambda Red homologous recombination plasmid containing the exo, bet, and gam genes regulated by the arabinose promoter; its replicon is a low-copy temperature-sensitive replicon to facilitate plasmid elimination; Amp^R	Purchased from Beijing ZOMAN Co., Ltd.
pKD3	Lambda Red homologous recombination plasmid containing a FRT cassette, Amp^R, and Cm^R	Purchased from Beijing ZOMAN Co., Ltd.
E. coli	Strain isolated from intestinal contents from the MCAO group	This study
E. coli ΔwaaL:: Cm^R	*E. coli* carrying a mutation in *waaL*; Cm^R	This study
E. coli AIS01	Strain isolated from faecal samples from an AIS patient	This study

Reference

1 Benakis C, Brea D, Caballero S*, et al.* Commensal microbiota affects ischemic stroke outcome by regulating intestinal gammadelta T cells. *Nat Med* 2016;22:516-23.
2 Meng F. Novel recombinant high-stability superoxide dismutase and application thereof. Google Patents, 2018.
3 Singh V, Roth S, Llovera G*, et al.* Microbiota Dysbiosis Controls the Neuroinflammatory Response after Stroke. *J Neurosci* 2016;36:7428-40.
4 Zhou HW, Li DF, Tam NF*, et al.* BIPES, a cost-effective high-throughput method for assessing microbial diversity. *Isme j* 2011;5:741-9.
5 McDonald D, Clemente JC, Kuczynski J*, et al.* The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *GigaScience* 2012;1:7.
6 Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26:2460-1.
7 Caporaso JG, Bittinger K, Bushman FD*, et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010;26:266-7.
8 Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;26:1641-50.
9 Caporaso JG, Kuczynski J, Stombaugh J*, et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335-6.
10 Segata N, Izard J, Waldron L*, et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.
11 Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 2000;97:6640-5.
12 Manwani B, Bentivegna K, Benashski SE*, et al.* Sex differences in ischemic stroke sensitivity are influenced by gonadal hormones, not by sex chromosome complement. *J Cereb Blood Flow Metab* 2015;35:221-9.
13 Kaliannan K, Robertson RC, Murphy K*, et al.* Estrogen-mediated gut microbiome alterations influence sexual dimorphism in metabolic syndrome in mice. *Microbiome* 2018;6:205.