1   **MorphoCellSorter: An Andrews plot-based sorting approach to rank microglia according to their**
2   **morphological features.**
3
4   **Running title: MorphoCellSorter**
5

6   Sarah Benkeder[1,2], Son-Michel Dinh[2,3], Paul Marchal[1,2], Priscille De Gea[1,2], Muriel Thoby-Brisson[4],
7   Violaine Hubert[1,5], Ines Hristovska[1,2], Gabriel Pitollat[4], Kassandre Combet[1,2], Laura Cardoit[4], Bruno
8   Pillot[1,5], Christelle Léon[1,5], Marlène Wiart[5,6], Serge Marty[7], Jérôme Honnorat[1,2,8], Olivier Pascual[1,2#],
9   Jean-Christophe Comte[1,9 #]
10

[1] Claude Bernard Lyon 1 University, Lyon, France.
[2] Institut MeLiS, INSERM U1314 CNRS UMR5284, Lyon, France.
[3] Institut National des Sciences Appliquées (INSA), Lyon, France.
[4] Institut de Neurosciences Cognitives et Intégratives d'Aquitaine, CNRS UMR 5287, Université de
Bordeaux, Bordeaux, France.
[5] CarMeN Laboratory, INRAE U1397, INSERM U1060, Claude Bernard Lyon 1 University, Lyon France.
[6] CNRS, Lyon, France.
[7] Sorbonne Université, Institut du Cerveau – Paris Brain Institute – ICM, Inserm, CNRS, APHP, Hôpital
de la Pitié Salpêtrière, Paris, France.
[8] French Reference Center on Paraneoplastic Neurological Syndromes and Autoimmune Encephalitis,
Hospices Civils de Lyon.
[9] Centre de Recherche en Neurosciences de Lyon (CNRL), INSERM 1028 CNRS UMR5292, Bron,
France.
[#] Contributed equally

Key words Microglia; Morphology; Andrews plots; Principal component analysis; Ranking; Ischemia;
Alzheimer's disease; Development; Neurodevelopmental disease.
28

29 **Word count:** 10620
30

Abstract

Microglia exhibit diverse morphologies reflecting environmental conditions, maturity or functional states. Thus, morphological characterization provides important information to understand microglial roles and functions. Most recent morphological analysis relies on classifying cells based on morphological parameters. However, this classification may lack biological relevance, as microglial morphologies represent a continuum rather than distinct, separate groups, and do not correspond to mathematically defined, clusters irrelevant of microglial cells function. Instead, we propose a new open-source tool, MorphoCellSorter, which assesses microglial morphology by automatically computing morphological criteria, using principal component analysis and Andrews plots to rank cells. MorphoCellSorter properly ranked cells from various microglia datasets in mice and rats of different age, from *in vivo*, *in vitro* and *ex vivo* models, that were acquired using diverse imaging techniques. This approach allowed for the discrimination of cell populations in various pathophysiological conditions. Finally, MorphoCellSorter offers a versatile, easy and ready-to-use method to evaluate microglial morphological diversity that could easily be generalized to standardize practices across laboratories.

Introduction

In the healthy brain, microglial cells are characterized by a complex arborization of highly ramified and fine processes that are distributed around the cell body. Their processes present a strong and constant dynamic of protractions and retractions in the adult (Hristovska et al., 2022). For many years, this dynamic has been associated with their monitoring role, as they are constantly looking for danger signals in their surroundings (Davalos et al., 2005; Nimmerjahn, Kirchhoff, & Helmchen, 2005). Microglia continuously adapt to their local environment, sensing even small changes via membrane receptors and transporters (Hanisch & Kettenmann, 2007). They are able to react quickly and adjust their functions in response to the detected signals. These functional changes are supported by morphological modifications. As an example, the formation of filopodia enables directed migration to lesion sites, while the formation of phagocytic cups and phagolysosome increases are linked to their phagocytic activity (Kettenmann, Hanisch, Noda, & Verkhratsky, 2011; Sierra, Abiega, Shahraz, & Neumann, 2013).

The traditional view of microglia being either ramified (homeostatic microglia) or amoeboid (reactive microglia) has now been proven to be oversimplified and inaccurate, as microglia display a whole range of various morphologies (Paolicelli et al., 2022). Indeed, many studies have shown a wide diversity of microglial morphotypes in pathological conditions, and microglia can even exhibit an increase in ramification within the first hours of injury (Vidal-Itriago et al., 2022).

In stroke models, morphological transformation of reactive microglia is correlated with the severity and duration of ischemia, as well as an enhanced expression of markers such as CD11B and CD68 (Zhang, 2019). In the rodent brain, different morphotypes can be observed depending on how the area has been impacted by the ischemic challenge (V. Hubert et al., 2021). Thus, a gradient of microglial morphologies can be observed from amoeboid microglia in the ischemic core to ramified complex cells in the contralateral cortex (Anttila, Whitaker, Wires, Harvey, & Airavaara, 2017; Fumagalli, Perego, Ortolano, & De Simoni, 2013). Similarly, in Alzheimer's disease (AD) mouse

models, microglial cells associated with Aβ plaques present strong morphological transformations in

opposition to plaque-distant cells mildly morphologically impacted (Plescher et al., 2018).

As a matter of fact, morphology is an indicator commonly accessible through live imaging or

immunohistochemistry, and it has been used in numerous studies to describe a broad range of

pathological contexts (Adeluyi et al., 2019; Ali et al., 2019; Morrison, Young, Qureshi, Rowe, &

Lifshitz, 2017; Plescher et al., 2018). To date, the literature contains a wide variety of criteria to

quantitatively describe microglial morphology, ranging from descriptive measures such as cell body

surface area, perimeter, and process length to indices calculating different parameters such as

circularity, roundness, branching index, and clustering (Adaikkan et al., 2019; Heindl et al., 2018;

Kongsui, Beynon, Johnson, & Walker, 2014; Morrison et al., 2017; Young & Morrison, 2018). Artificial

intelligence (AI) approaches such as machine learning have also been used to categorize

morphologies (Leyh et al., 2021). This variety of metrics and approaches to measure microglia

morphology makes it difficult to compare results between studies. Moreover, categorization

methods seem subjective with the number, name and composition of categories being non-

consistent between studies. In a biological and functional perspective mathematical of manual

clustering can be arbitrary as microglia morphology exists along a continuum. In addition, as

microglial morphology is highly diverse, it is not rare to be confronted with cells belonging to

categories that are not defined, or that could belong to several of them.

Here, we propose a fully automated program that quickly and unbiasedly computes most parameters

used in the literature to characterize microglial morphology, and overcomes the limitations of

current methods. Our approach consists in: 1) performing principal component analyses (PCA) to

select the parameters best fitted to describe the dataset of interest and 2) generating a sorting of the

cells based on their morphology. The originality of our study is the use of Andrews Plots to rank

microglia according to their morphology and discriminate distinct populations. In addition, our

working flow has been developed so that it can be used for a wide variety of models and image

acquisitions (confocal, wide field, two-photon, etc.). We developed this method on an exploratory

dataset from a rat brain with focal cerebral ischemia, and validated it on other independent datasets

from other models (other animal, age, pathology) to confirm its wide range of application.

Methods

Models

All experimental procedures were carried out in accordance with the French institutional guidelines

and ethical committee, and authorized by the local Ethics Committees of the University of Bordeaux

and "Comité d'éthique pour l'Expérimentation Animale Neurosciences Lyon": CELYNE; CNREEA no.

42, and the Ministry of National Education and Research (APAFIS#30666-2021032518063894).

Ischemic stroke rat model – fixed tissue

The experiment was conducted on 6-8-week-old male Sprague Dawley rat with transient (90

minutes) middle cerebral artery occlusion (tMCAo) followed by reperfusion. Twenty-four hours after

reperfusion, the animal was deeply anesthetized (xylazine at 12 mg/kg and ketamine at 90 mg/kg in

0.9% NaCl) and died by the collapse of the lungs when the thoracic cage was open. Infusion of

phosphate-buffered saline (PBS) was performed directly in the left ventricle. Next, 100 mL of

paraformaldehyde (PFA) 4% was infused at 10 mL/min before the brain collection. The brain was

then post-fixed in PFA 4% for 2-4 days, rinsed, and soaked in 30% sucrose for 72 hours before being

frozen in -40°C 2-methylbutane. It was then stored at -20°C until being cut. Floating 30 μm thick

sections were obtained using an NX50 cryostat and conserved in PB-Thimerosal until

immunostaining.

Ischemic stroke mouse model – live tissue

An ischemic stroke was induced in 8 to 12 weeks CX3CR1$^{+/GFP}$ male mice (RRID:IMSR_JAX:005582) by

a permanent occlusion of their middle cerebral artery (pMCAo). Imaging was performed the next day

directly on living animals through a cranial window. Briefly, animals were anesthetized with

isoflurane (induction: 3-4%, surgery: 1.5-2%) and installed on a stereotaxic apparatus where their

temperature was monitored and maintained at 37°C. After cleaning and exposition of the skull, a

polyamide implant was glued on a zone at the periphery of the lesion in order to be able to image

both ischemic and extralesional regions. The skull was then cautiously thinned by drilling until a

thickness of 20 to 30 µm was reached. A glass coverslip was then glued to the thinned region. Mice

were then put under a heating lamp and individually housed for 24 hours until imaging (V. Hubert et

al., 2021).

Phox2b mutation and Embryonic dataset

The expansion of a 20-residue polyA tract in PHOX2B is used as a model for Central Hypoventilation

Syndrome (CCHS) (Amiel et al., 2003). Mutant pups were generated by mating conditional

Phox2b$^{27Ala/27Ala}$ females with Pgk::Cre males (Dubreuil et al., 2008). Since mutant newborns die

rapidly after birth, all experiments were performed blind on E18.5 embryos of either sex, and the

genotype of each embryo was determined a posteriori on tail DNA. Pregnant mice were killed by

cervical dislocation at E18.5. Embryos were rapidly excised from uterine horns and placed in artificial

oxygenated cerebrospinal fluid (aCSF) at 18-20°C until dissection. The aCSF solution composition was

(in mM): 120 NaCl, 8 KCl, 0.58 NaH2PO4, 1.15 MgCl2, 1.26 CaCl2, 21 NaHCO3, 30 Glucose, pH 7.4 and

equilibrated with 95% O_2-5% CO_2. Embryos were decerebrated and decapitated, and ventral tissues

were removed. Brainstems were then carefully disassociated from the surrounding tissues and

completely isolated by a rostral section made at the junction between the rhombencephalon and

mesencephalon and a caudal section at the spinal cord level. After dissection, brainstem

preparations were placed in 4% PFA for 2 - 3 h for tissue fixation. To obtain transverse frozen

sections, brainstems were placed in a 20% sucrose-PBS solution for cryoprotection overnight, then

they were embedded in a block of Tissue Tek (Leica Microsystems, France) and sectioned at 30 µm

using a cryostat (Leica Microsystems, France).

Alzheimer's Disease model

C57BL6 5-month-old female APPxPS1-KI and control (PS1-KI) mice were used(Casas et al., 2004). For

the brain collection, 0.5 mg xylazine was injected intraperitoneally followed by Euthasol. They were

then intracardially perfused first with 4% PFA in PB. After one hour, the brains were collected and

then post-fixed for 2 hours in 4% PFA and stored in PBS. They were rinsed in phosphate buffer (PB)

and soaked in 30% sucrose for 48 hours before being frozen at -40°C in 2-methylbutane for 2 minutes

and stored at -70°C until the cut. Free-floating sections (30 μm thick) were cut using a cryostat and

conserved in PB-Thimerosal until immunostaining.

In vitro microglia

We followed the protocol published by Bohen and collaborators using Magnetic Activated Cell

Sorting (MACS) (Bohlen, Bennett, & Bennett, 2019). Ten days old mice were anesthetized with

isoflurane, intracardially perfused with PBS enriched with magnesium and calcium (14040141, Gibco)

before decapitation and brain collection. The brains were then finely chopped in a dissociation buffer

and transferred to a tissue grinder to dissociate further the cells from the tissue. Myelin and debris

were then eliminated thanks to a Percoll® PLUS solution (E0414, Sigma-Aldrich) diluted with DPBS10X

(14200075, Gibco) and enriched in $MgCl_2$ and $CaCl_2$ (for 50 mL of myelin separation buffer: 90 mL of

Percoll PLUS, 10 mL of DPBS10X, 90 μL of 1 M $CaCl_2$ solution, and 50 μL of 1 M $MgCl_2$ solution). An

incubation with anti-myelin magnetic beads (130-096-433, Miltenyi Biotec): the cell suspension is

applied on a column fixed to a magnet so that the myelin will stay on the column while the cells will

exit. To isolate specifically microglia, the suspension was incubated with anti-CD11b magnetic beads

(130-097-142, Miltenyi Biotec). 20.000 cells were then seeded on each of the four compartments of

the 35 mm diameter dishes with a polymer bottom (80416, Ibidi) coated with poly-D-lysine (100

μg/mL, A-003-E, Sigma-Aldrich) and collagen IV (2μg/mL, 354233, Corning). The culture medium was

adapted from Bohlen et al., 2019 with DMEM/F12 (21041025, Gibco), 1% penicillin-streptomycin, 2

mM L-Glutamine (10378016, Gibco), 5 μg/mL N-acetylcysteine (A9165, Sigma-Aldrich), 100 μg/mL

apo-transferrin (T1147, Sigma-Aldrich), 100 ng/mL de sodium selenite (S5261, Sigma-Aldrich), 1.5

µg/mL cholesterol (700000P, Avanti), 0.1 µg/mL oleic acid (90260, Cayman Chemical), 1 ng/mL

gondoic acid (20606, Cayman Chemical), 1 µg/ml heparan sulfate (AMS.GAG-HS01, AMSBIO), 2

ng/mL TGFβ2 (100-35B, Preprotech) and 10 ng/mL M-CSF (315-02, Preprotech). Cells were kept at

37°C 5% CO_2 and the culture medium was half renewed every two days until imaging (4 to 10 days

after seeding).

Immunohistochemistry

The sections underwent 5-minute rinses three times with PBS and Triton 100X 0.3% (PBS-T). Sections

from the ischemic experiment were incubated for 20 minutes in a 1.5% hydrogen peroxide solution

diluted in PBS, and after 3*10 minutes of rinsing with PBS-T, they were incubated for 30 minutes in

0.5% sodium borohydride diluted in PBS and then thoroughly rinsed. One-hour saturation was then

performed on all sections with PBS-T, 3% bovine serum albumin and 5% normal goat serum (PBS-T-

BSA-NGS). Sections were then incubated overnight with anti-Iba1 antibodies (1/500, Wako, Germany,

#W1W019-19741) at 4°C and anti-Amyloid-β 4G8 (1/500, Bio-Legend, #800709) antibody. Sections

were rinsed 3*10 minutes the next day with PBS-T. They were then incubated in goat anti-rabbit

secondary antibodies coupled with Alexa 555 fluorochrome (1/300, Invitrogen, France, #A21429) and

goat anti-mouse coupled with Alexa 488 fluorochrome (1/300, Invitrogen, France, #A11029) for 2

hours at room temperature. They were then rinsed 3*10 minutes with PBS and mounted on

microscope slides using Fluoromount G (ThermoFisher Scientific, #00-4958-02) and a glass coverslip.

For the Phox2b mutant model and embryonic dataset, sections were incubated for 90 min in a

solution of PBS containing 0.3% Triton X-100 and 1% BSA to limit nonspecific labeling and favor

antibody tissue penetration. The primary rabbit anti-Iba1 antibody (1/500; Wako, Germany,

#W1W019-19741) was then applied overnight at room temperature and under slight agitation. To

amplify the primary antibody signal, after several washes, sections were incubated with an Alexa

Fluor 568 donkey anti-rabbit secondary antibody (1/500, Merck, France, #A10042) for 90 min at

room temperature. Immunostained sections were then mounted in Vectashield Hard Set medium

(Eurobio, France) cover-slipped and kept in the dark until imaging.

Image acquisitions

For the ischemic stroke model on fixed tissue, acquisitions were performed with a Zeiss confocal

microscope and a 40X objective (LSM 880, Zeiss). Z-stacks between 18 and 25 µm thickness were

executed with a 1 µm step. For the ischemic stroke model on live, microglia were observed thanks to

a two-photon microscope (Bruker Ultima) and a 20X objective, at a deepness of 50 to 150 µm. 10 to

15 µm thick acquisitions were acquired with a 1 µm step between plans. For the AD model, a Zeiss

Axio Scan 7 microscope slide scanner was used with a 20X objective. Z-stacks (30 µm thick) were

performed with a step of 1 µm. For the Phox2b mutant model, imaging was conducted using a Zeiss

confocal microscope (LSM 900, Zeiss). Z-stacks (30 µm thick) were performed with a step of 1 µm and

a 40X objective. *In vitro* microglia were imaged on a single plane using the BioStation (Nikon).

Image processing and segmentation

The general procedure is described in supplemental Fig. 7.

For the ischemic model on fixed tissue, all post-acquisition processing was executed with FIJI

software (2.9.0/1.54f). Individual microglial cells were cropped from the original acquisitions before

Z-maximum projections. The brightness of the projections was adjusted, and median (radius=2),

mean (radius=2) and unsharp mask (radius=5; mask=0.60) filters were applied via a macro. A

threshold was then applied to each crop to binarize the cells, and pixels not belonging to the

microglia of interest were manually deleted. The same procedure was performed for the Phox2b

mutant model, except the mean filter that was not applied.

For the ischemic model on live tissue all post-acquisition processing was also executed with FIJI

software. Z-maximum projections were performed, and individual crops of the cells were done.

Brightness and contrast of the images were modified before applying a manual threshold to binarize

the cells. Reminiscent noise and neighboring cells were manually removed.

For the AD model, crops of individual microglial cells located in the secondary visual cortex were

extracted from images using the *Zen* software (v3.5; Zeiss) and exported to the Tif image format. The

next operations were executed with FIJI software. Z stack cleaning was performed by background

subtraction (rolling=50), followed by a median filter (radius=3) and an unsharp mask (radius=5;

mask=0.60). Segmentation was performed by the FIJI Trainable Weka Segmentation 3D plugin

(Arganda-Carreras et al., 2017). The selected training features were based on the Gaussian blur,

Laplacian, Maximum, Mean, Minimum, Median and Variance (Minimum sigma: 1.0 and Maximum

sigma: 8.0). Binarized z-stacks were then manually cleaned to remove pixels from neighboring cells or

remaining noise. Finally, a maximum Z-projection was performed.

For the Phox2b mutant model, image processing was performed via a custom MATLAB R2018B

(Mathworks, Massachusetts, USA) script. First, a median filter followed by an unsharp mask filter was

applied to all images. An automated threshold was then applied. Manual tuning with a thresholding

tool (R, 2023) was applied if the automated threshold was not satisfying. The binarized images were

then automatically cropped around microglia. Briefly, the centroid of each detected object (i.e.

microglia), except the ones on the borders, were detected, and a crop of 300x300 pixels around the

objects were generated. Then, the pixels belonging to neighboring cells were manually removed on

each generated crop.

For the *in vitro* microglia, acquisitions were cropped around individual microglia, and the cells were

then manually traced on Procreate (5.2.9 version).

When the binarization process generated one or several holes inside the cell bodies, they were

manually filled to not interfere with the automated detection of the cell bodies. All binarized and

individualized microglial cells were then rescaled through a FIJI macro command to obtain a pixel

resolution equal to 0.5 µm. Finally, the images were resized to obtain a field size equal to 300 x 300

pixels.

MorphoCellSorter algorithm pipeline

Automated measure of 20 dimensionless morphological descriptors

All parameters are illustrated in Supp. Fig. 8.

First, we calculate the morphological parameters most commonly used in the literature, followed by

other parameters that we have developed and proposed here (Fig. 2B). To make our approach scale-

independent, we reformulated some morphological descriptors to obtain dimensionless parameters.

$$Circularity = \frac{2 \times \sqrt{\pi \times CellArea}}{CellPerimeter}$$

$$RamificationIndex = \frac{\dfrac{CellPerimeter}{CellArea}}{2 \times \sqrt{\pi \times CellArea} \times Image\ length \times Image\ width}$$

$$RoundnessFactor = \frac{4 \times SomaArea}{(Longest\ axis\ length)^2 \times \pi}$$

$$PerimeterAreaRatio = \frac{Perimeter^2}{Cell\ Area}$$

$$Density = \frac{CellArea}{Length \times width\ of\ the\ image}$$

$$Inertia = \frac{Longest\ axis\ length}{Smallest\ axis\ length}$$

Cell Body Recognition

The cell body identification is based on successive erosions until the center of the cell body is the

only pixels remaining using a 3*3 pixels square as a structuring element. The last erosion is the one

preceding the elimination of all pixels belonging to the cell and corresponds to the center of the cell

body. Then, successive dilations are performed to reconstruct the cell body according to the

morphology of the cells, using a cross (+) 3 pixels height and width (Leyh et al., 2021). Reconstruction

of the cell body relies on (1) the identification of the index *i* on which measured surfaces after

erosions are inferior to the surfaces measured after dilation, (2) the average of two copies of the

images: one on which $i + 1$ erosions have been performed, and the other one on which i dilatations

have been performed from the center of the cell body. The resulting mean image was then subjected

to a threshold fixed at 0.25 to obtain the cell body. From this, we can calculate the following

parameters.

$$Processes\ Soma\ Areas\ Ratio = \frac{Processes\ Area}{Soma\ Area}$$

$$Processes\ Cell\ Areas\ Ratio = \frac{Processes\ Area}{Cell\ Area}$$

$$Polarization\ Index = \frac{Soma\ Centroid\ |\ Cell\ Centroid}{\sqrt{CellArea}}$$

Skeleton

The cell skeleton is obtained via the MATLAB *bwskel* function. This function reduces all 2-D binary

objects to 1-pixel width curves while taking care of not changing the essential structure of the image.

The aim is to extract the image skeleton while keeping the topology properties such as the Euler

number and to extract branchpoints and endpoints. The skeleton length is defined by the number of

pixels forming the skeleton minus the pixels of the cell body.

Thus, we calculate the skeleton related parameters:

$$Skeleton\ Process\ Ratio = \frac{(Skeleton\ Length)^2}{Processes\ Area}$$

$$Branchpoints\ Endpoints\ Ratio = \frac{Number\ of\ Branchpoints}{Number\ of\ Endpoints}$$

275 *Sholl analysis*

276 Sholl analysis was performed on the microglial skeleton. Concentric circles with a step size of two

277 pixels were generated on the skeletons minus the soma, and the soma centroid was the center point

278 of the circles. Intersections between the growing circles and the skeleton were counted.

$$Branching\ Index = \sum \frac{(Intersections\ circle_n - Intersections\ circle_{n-1}) \times r_n}{\sqrt{Length\ \times width\ of\ the\ image}}$$

Fractal/Hausdorff dimension

The fractal dimension measures the complexity of a cell shape, that is, how an object fills the space

at different scales. The Hausdorff fractal dimension (A, 2023) corresponds to the slope of the

logarithm of the number of tiles of size ε needed to cover all the surfaces of the object of interest vs

the logarithm of ε.

Lacunarity

Lacunarity corresponds to the inhomogeneity of a cell shape. It measures the variance of how the

object fills the space at different scales. Patterns with larger gaps generally have higher lacunarity.

We used the gliding box algorithm (T, 2023; Tolle, McJunkin, & Gorsich, 2008), and the value

corresponds to the slope of the log-log representation.

Convex Hull

The convex hull is the smallest convex polygon (with angles inferior to 180°) enclosing the whole cell.

$$Solidity = \frac{Cell\ Area}{Convex\ hull\ Area}$$

$$Convexity = \frac{Cell\ Perimeter}{Convex\ hull\ Perimeter}$$

$$Convex\ Hull\ Circularity = \frac{2 \times \sqrt{\pi \times Convex\ hull\ Area}}{Convex\ hull\ Perimeter}$$

293   We also measure the length of the major and minor axes of the convex hull.

$$Convex\ Hull\ Span\ Ratio = \frac{Length\ Major\ axis}{Length\ Minus\ axis}$$

295   Finally, we measure the largest and smallest radii of the circles from the centroid of the convex hull

296   as the center to external points.

$$Convex\ Hull\ Radii\ Ratio = \frac{Maximum\ Convex\ hull\ radius}{Minimum\ Convex\ hull\ radius}$$

298   *Linearity*

299   A Principal Component Analysis is generated on all the points forming the microglial cell. Then, the

300   variance of the two first principal components (PCs) is calculated; the highest corresponds to the

301   Variance Max and the lowest to the Variance Min.

$$Linearity = \frac{Variance\ Max}{Variance\ Min}$$

Establishment of the cell ranking according to morphological criteria.

The second step is to determine which of the previously calculated parameters best discriminate the

cells in a given dataset. Principal component analysis (PCA) is a statistical method allowing dataset

dimensionality reduction (Jolliffe & Cadima, 2016). A PCA is performed on the 20 morphological

indexes calculated to determine the parameters that best explain the dispersion of the data.

Lacunarity, roundness factor, convex hull radii ratio, processes cell areas ratio and skeleton processes

ratio were subjected to an inversion operation in order to homogenize the parameters before

conducting the PCA: indeed, some parameters rank cells from the least to the most ramified, while

others rank them in the opposite order. By inverting certain parameters, we can standardize the

ranking direction across all parameters, thus simplifying data interpretation. To identify the main

parameters responsible for the data spreading, we perform a PCA

The principal component 1 (PC$_1$) is the direction in which data is most dispersed. The two first

components represent most of the information (about 70%), hence we can consider the plan PC_1, PC_2

as the principal plan reducing the dataset to a two dimensional space. Next, we calculate the

projection of the parameters F_n onto the two first principal components respectively PC_1 and $PC2$ (Fig.

2C) and rank those parameters according to the intensity (absolute value) of their projection onto PC_1

(Fig. 2D). Each parameter is then weighted by a factor $W_n = w_n./w_0$, where w_n is the absolute value of

the nth parameter projection onto PC_1, and w_0 is the absolute value of the strongest projection among

all the parameters. As shown in Fig. 2C, each parameter F_n is represented by a vector in the PC_1, PC_2.

Thus, we calculate the projection y_n of each parameter onto PC_2 and calculate the following phase

factor

$\phi_n = arctan(y_n/w_n)$.

The following stage of our method consists in introducing the previously calculated factors (Weight:

W_n, and phase: ϕ_n) and their corresponding parameter into a special Fourier synthesis called

"Andrews Plot" or "Andrews curve" (Andrews, 1972; García-Osorio & Fyfe, 2005) allowing to visualize

the data structure in high-dimensional. The general form of the spectral Fourier synthesis of an

individual j writes:

$$S_j(t) = \sum_{n=1}^{N} A_n \, cos(2\pi \, n \, t) + B_n \, sin(2\pi \, n \, t) \, , \qquad (4)$$

331   that we rewrite as,

$$S_j(t) = \sum_{n=1}^{N} C_n \, cos(2\pi \, n \, t \, + \phi_n), \qquad (5)$$

where n, is the index of the ranked parameters, N is the full number of parameters F_n, t is a virtual

time $0 \leq t \leq 1$, and $C_n = F_n \times W_n$. Finally, we derive this method to represent data in high dimensional

by a simple one-dimension signal composed of orthogonal trigonometric functions in researching the

time $t = t^*$ for which the curves present the maximum of variance. At this specific time point, we rank

individuals according to the magnitude of their own function $S_j(t^*)$ value as represented Fig. 2E. We

now call $S(t^*)$ the "Andrews-Score".

Statistics

Rankings comparisons, performed with Spearman's correlation, were also used to measure the

monotonic correlation between the rankings. They were executed with MATLAB.

We used mixed effect models on RStudio, that are adapted to complex experimental designs with

dependent and non-dependent data. This allows the addition of random factors to fixed effects.

Model fittings and analysis were performed with the lme4 package (Amiel et al., 2003). Condition

(Control/APP-PS1xKI or WT/CCHS) was considered a fixed effect, and the factor mice was considered

a random intercept. When the mixed models were not applicable, we used Wilcoxon tests. Statistical

significance is shown on the graphs (*p < 0.05; **p < 0.01; ***p < 0.001).

Number of bins of Andrews scores histograms was automatically determined by the Freedman-

Diaconis method based on the following formula:

$$k = (x_{max} - x_{min})/(2 \times IQR \times N^{(-1/3)})$$

With x_{max} and x_{min} being respectively the maximum and minimum Andrews scores, IQR the

interquartile range and N the number of cells.

Results

Evaluation of classification methods

Morphological categorization has been at the center of recently developed approaches to

characterize microglial morphology. It enables the evaluation of the morphological diversity existing

in a cell population. Moreover, in the context of morphological comparison between populations, the

study of the morphotypes' proportions is relevant as it can highlight meaningful differences that the

sole global comparison of morphological criteria cannot. One way to categorize the cells based on

their morphology is to perform clustering to group the cells having similar characteristics. We tested

a commonly used approach, k-means, to categorize microglia from a rat brain cortex following an

occlusion of the middle cerebral artery (Fig. 1A). In this model, microglia contralateral to the lesion

are mildly to not impaired whereas ipsilateral microglia undergo light to drastic morphological

changes (Anttila et al., 2017; Fumagalli et al., 2013) (Fig. 1A). We selected 20 non dimensional

parameters from the literature and developed by the team, on which we conducted a Principal

Component Analysis (PCA). The k-means clustering algorithm was applied on the first two principal

components (PCs), and we fixed at 4 the number of clusters as commonly chosen in the literature

(Verdonk et al., 2016) (Fig. 1B).

The obtained microglial groups could be named as ramified microglia (cluster 1), with cells having a

morphology consistent with a physiological state. Cluster 2 could be referred to as "reactive

microglia", as it was composed of microglia displaying various morphologies; they had an

intermediary ramification level with large cell bodies. Cluster 3 was composed of cells having few

short or no processes and corresponded to amoeboid microglia. Lastly, cluster 4 was rod-like

microglia, with two processes extending on either side of the cell body or a single long polarized

extension. A sample of the cells of the dataset (ipsilateral and contralateral) and their associated

cluster is represented in Fig. 1C. These 4 clusters were not equally represented across each condition

as no amoeboid nor rod-like microglia were found in the contralateral side (Fig. 1D). The large

majority of the cells were ramified microglia (92%), and the rest were reactive microglia (8%). The

proportion of ramified cells drastically decreased ipsilaterally by a reduction factor of 4 (23%), while

reactive microglia increased (39%). 25% of the cells were amoeboid microglia, while rod-like cells

represented 13% of the ipsilateral population (Fig. 1D). These results show that the microglial

populations in each region are different, with microglia having a "non-pathological" morphology

mainly found contralaterally, with only few cells being potentially impacted by the

ischemia/reperfusion challenge as they showed morphological signs of reactivity with a less ramified

morphology. Ipsilateral microglia were morphologically highly altered, with a decrease in ramified

cells, replaced by reactive and amoeboid microglia, and rod-like microglia.

However, the clusters themselves were highly morphologically heterogeneous, making these

quantifications problematic (Fig. 1C). For instance, Cluster 1 (ramified microglia), was composed of

cells with branched processes but also included cells that were visibly affected by the ischemic lesion,

exhibiting sometimes short, few, or less branched processes, as well as elongated cell bodies. These

microglia might have been more appropriate in Cluster 2 (reactive microglia). The composition of the

latter is also questionable, with the presence of branched processes that would fit better in the first

cluster, or cells with few very short extensions that resemble some cells found in Cluster 3. Several

non-bipolar or non-unipolar cells are also found in Cluster 2. We tested to decrease or increase the

number of clusters with no improvement (data not shown).

Other classification methods exist, relying on supervised artificial intelligence (AI) algorithms to

identify specific morphotypes. These approaches can have high rates of success (96% for the method

developed by Leyh and collaborators (Leyh et al., 2021)), meaning that they are in theory efficiently

classifying cells. However, the problems of cells which do not fit in a defined category, or to several

categories can be raised. For example, in our dataset we found that about 60% of cells would not

belong to any category or belong to several categories. In this context, the way the networks are

trained might be very subjective to the experimenter and could influence the results. Moreover,

depending on the papers, semantic problems are also met, with similar morphologies having

different names or on the contrary same names having different definitions (Choi et al., 2022; Leyh et

al., 2021), making the comparisons between studies complicated.

We know now that microglial morphology is rather a continuum than closed categories (Paolicelli et

al., 2022). That is why we propose a continuous ranking rather than a classification of the cells.

Development of a ranking approach based on Andrews plots to order rat fixed microglia in a

context of ischemic stroke lesion.

We developed MorphoCellSorter, a novel ranking method, on the same dataset of rat microglia in

the context of an ischemic stroke lesion, on fixed tissue acquired *via* confocal imaging. This model is

known to induce major morphological modifications in microglia (Violaine Hubert et al., 2021),

enabling us to constitute a heterogeneous dataset, with varied morphologies ranging from amoeboid

to hyper-ramified.

The method is summed up on Fig. 2A, and takes as inputs binarized and individualized cells form

projected Z stacks. To generate the ranking, we first gathered 20 morphological indices, that we also

call parameters, either from the literature (Clarke, Crombag, & Hall, 2021; Fernández-Arjona,

Grondona, Granados-Durán, Fernández-Llebrez, & López-Ávalos, 2017; Madry et al., 2018) or

developed by the team (see Methods). These parameters, describing morphological complexity

(circularity, ramification index, roundness factor, perimeter area ratio, density, processes soma area

ratio, processes cell area ration, skeleton processes ratio, branchpoints endpoints ratio, fractal

dimension, lacunarity, solidity, convexity, convex hull circularity, convex hull radii ratio and branching

index), linearity (linearity, convex hull span ratio, inertia) and directionality (polarization index) of the

cells, were normalized to suppress any dimension in order to be exempted from any scale effect. The

first step of our method involves PCA, which helps us determine the main parameters responsible for

the data dispersion. The first principal component (PC1) captures the direction of maximum variance

in the data, and subsequent component (PC2) captures the direction of the next highest variance

(Fig. 2B). Typically, the first two principal components account for a significant proportion (often over

70%) of the data's total variance, allowing us to reduce the dataset's complexity to a two-

dimensional space, such as a two-dimensional plane defined by PC1 and PC2 (Fig. 2C).

Next, each parameter was projected onto these principal components, which helps us rank

parameters based on their influence on the overall data structure (Fig. 2D). To determine the number

of parameters used to elaborate the ranking, we selected the median number of the normalized

cumulative projection (threshold = 0.5). We assign weights to the selected parameters based on their

projections onto PC1, which allows us to emphasize the most influential parameters. This weighted

projection, along with phase factors calculated for each parameter (referred as $\square_n$ in the Methods

section), forms the basis of an Andrews Plot — a graphical representation that simplifies high-

dimensional data into a two-dimensional signal using trigonometric functions.

The Andrews Plot represents each individual microglia, as a unique curve, allowing for visual

comparison based on their trajectories within this synthesized signal (Fig. 2E). This method enables

efficient visual comparisons and insights into complex datasets with potentially significant

implications for data analysis and interpretation. We took advantage of this representation to rank

the cells based on their morphology; the ranking is established at the time point displaying the

highest variance, i.e. the time point where the curves are the most dispersed, also named ⍺ (Fig. 2E).

The obtained ranking is displayed on Fig. 3A. To evaluate the quality of the ranking method, we

compared the automated ranking to a subjective manual ranking by experts based on visual

morphological criteria (cell size; number, length and branching of the processes). To assess the

similarity degree between the rankings, we looked at the distribution of rank differences for each

microglia between rankings (Fig. 3B) and observed that the majority of the cells' rank

differences did not exceed an absolute value of 50 rank difference with a peak around 0. We

also calculated Spearman's rank correlation coefficient (r_s=0.96; Fig. 3E). This test revealed that the

automated ranking is close to the manual ranking. To ensure the reliability of the manual ranking, we

compared the automated ranking with another expert and obtained the same results (r_s=0.97; Fig.

3C,F). Interestingly, the comparison between the two manual rankings shows similar results to those

obtained when compared to the automated ranking (r_s=0.93; Fig. 3D,G), indicating that variability is

the same between automated or manual rankings. The distribution of the rank differences

between the same microglial cells in the two rankings elaborated by the experts even show a

greater dispersion of the rank differences, with more cells having rank differences higher

than 50 and around 100, showing a part of subjectivity in the cell ranking based on their

morphologies. However, automated ranking provided an objective and repeatable result, validating

our method. Thus, MorphoCellSorter is able to generate an accurate ranking of the cells based on

morphological criteria.

Evaluation of MorphoCellSorter on other datasets of microglia from different animals, age, models

and imaging techniques

One limitation to the use of a general and common method to characterize microglial morphology is

the high diversity of study material: the models (*in vivo, ex vivo, in vitro,* fixed or live tissue), the type

of collected signal (DAB or fluorescent staining, endogenous fluorescence…), the imaging technique

(influencing the resolution and the background noise). We thus wanted to evaluate the capacity of

MorphoCellSorter to accurately rank microglia from various datasets having different morphological

characteristics (Table 1). As previously, we compared the results to manual rankings performed by

two independent experts to evaluate the quality of the automated rankings (Table 1). We tested the

accuracy of the rankings using several thresholds impacting the number of parameters selected as

entries for the Andrews plots (Table 1). To maximize the accuracy of the automated ranking for all

datasets, we determined which threshold provided the best rankings overall (Supp. Table 1). We

observed that a threshold of 0.8 (Supp. Table 1), leading to 11 to 12 selected parameters (Supp. Fig.

1A-E), constituted a good compromise to obtain good rankings for all datasets. Thus, we chose to fix

this threshold for the following analyses. The Andrews plots and rankings obtained at the 0.8

threshold are respectively displayed in Supplemental Fig. 2 and 3.

Thus, our Andrews plot-based ranking method, relying on an automatic selection of relevant

parameters thanks to the PCA, allows a reliable automated ranking of the cells based on their

morphologies. The diversity of the tested datasets, with different models, types of signal and imaging

techniques shows that our approach offers a good sorting of the cells independently of their shape

and context.

486 **APPxPS1-KI microglia in the visual cortex have significant morphological alterations compared to**

487 **control cells.**

488 To determine the usability of our approach to compare microglial morphology in different conditions,

489 we used MorphoCellSorter to assess the presence of morphological alteration in the context of an

490 adult pathological murine model of Alzheimer's disease. This model is known to induce amyloid

491 plaques with dystrophic neurites, synaptic dysfunction and behavioral deficits (Casas et al., 2004).

492 We compared the morphology of PS1-KI (control) and APPxPS1-KI microglia in the visual cortex (Fig.

493 4A). Eleven parameters were automatically selected by the PCA: circularity, ramification index,

494 perimeter area ratio, fractal dimension, skeleton processes ratio, density, convexity, processes cell

495 areas ratio, lacunarity, branching index and processes soma areas ratio (Supp. Fig. 4A-C) and used to

496 generate the Andrews plots (Supp. Fig. 4D). The automated ranking obtained revealed that control

497 and APPxPS1-KI microglia were well segregated with APPxPS1-KI microglia at the beginning of the

498 ranking and control mice at the end (Fig. 4B). The distributions showed a 16.4% overlap between the

499 populations, meaning that a large number of APPxPS1-KI cells have a morphology that differs from

500 that of controls (Fig. 4C).

501 The morphological indices showed significant differences between APPxPS1-KI and control microglia.

502 The parameters measuring the morphological complexity of the cells highlighted a clear decrease in

503 the ramification level of APPxPS1-KI microglia, with for example a fractal dimension and convexity

504 decrease and a circularity increase. However, there were no significant differences in the cells'

505 linearity and polarity between groups (Fig. 4D-W).

506 Thus, this approach allowed the detection of morphological changes of microglia in the visual cortex

507 of APPxPS1-KI mice.

508 **MorphoCellSorter pinpoints no morphological alterations in a model of congenital central**

509 **hypoventilation syndrome.**

Congenital central congenital hypoventilation syndrome (CCHS) is a rare genetic disease is associated

with mutations of the PHOX2B gene and characterized by life-threatening breathing deficiencies

(Ceccherini, Kurek, & Weese-Mayer, 2022; Dubreuil et al., 2008). Phox2b is a transcription factor

required for the specification of the autonomous nervous system that contains respiratory control

centers. Because some of our preliminary observations strongly suggest a possible involvement of

microglia in the disease (unpublished data), we tested whether MorphoCellSorter highlighted

eventual morphological alterations of microglia in Phox2b mutants (CCHS embryos) compared to

wildtype embryos (WT), focusing on brainstem tissue that hosts respiratory networks involved in

breathing rhythmogenesis.

In this case using a threshold of 0.8, the PCA highlighted 12 parameters that were used to rank

microglia with Andrews plots: circularity, ramification index, perimeter area ratio, skeleton

processes ratio, fractal dimension, processes cell areas ratio, density, branching index, processes cell

areas ratio, convexity, lacunarity and solidity (Supp. Fig. 5A-C). These parameters were weighted to

generate the Andrews plots (Supp. Fig. 5D) used to rank the cells according to their morphology (Fig.

5B; complete ranking in Supp. Fig. 6).

Both populations displayed a large variability of morphologies from amoeboid to few ramified

processes and were well represented throughout the ranking (Fig. 5B). This is concordant with

previous studies on developing microglia: they can concomitantly be found in a large variety of

shapes and morphologies (Orłowski, Sołtys, & Janeczko, 2003). Indeed, the distributions of Andrews

scores showed no difference between WT and mutant mice. The overlap of distributions between

the populations found was 78.9% (Fig. 5C). Accordingly, when taking into consideration the various

morphological indices, there was no significant difference between WT and Phox2b mutant

microglial morphologies (Fig. 5D-W). Additional experiments are required to test whether this is

specific for embryonic stages and whether this is associated or not to different microglial functional

states.

Discussion

MorphoCellSorter enables the automated measurement of morphological descriptors, the

calculation of several non-dimensional indices, and a meaningful comparison of cell populations

based on their morphological characteristics. The method was initially applied to an ischemic stroke

model on confocal images with significant morphological changes. MorphoCellSorter was

subsequently tested on five different datasets with heterogeneous characteristics (various models,

age, sex, imaging technique and type of collected signal), showcasing its adaptability. We then

successfully used our approach to compare the cells' morphology in an Alzheimer's disease context,

and in embryonic microglia using a CCHS model. Our method can autonomously identify the most

effective parameters for discriminating the studied cell set, thereby providing valuable insights into

microglial morphology. Additionally, it offers a useful automated ordering of cells based on a

combination of selected parameters using Andrews plots. This feature allows for precise positioning

of cells along an axis, ranging from round cells to highly branched cells, facilitating the discrimination

of various cell populations based on their morphologies.

The use of the Andrews plots offers the possibility to combine all automatically selected most

discriminant parameters and to weight them according to their relevance, making the rankings more

accurate and less subjective than only sorting the cells based on one parameter. Even if we

recommend a threshold at 0.8 for the number of parameters selected, as it gave good rankings for all

datasets we tested, we included the possibility to change this threshold if the generated ranking is

not satisfying, knowing that all thresholds give accurate rankings when compared to expert rankings.

The automated parameter selection relies on principal component analysis, which is used to

determine the parameters that best disperse the data. In contrast, other approaches use PCA for

dimensionality reduction to create new morphological parameters that are difficult to apprehend as

they do not refer to any physical measure (Clarke et al., 2021; Heindl et al., 2018; Salamanca et al.,

2019). The strength of MorphoCellSorter lies in its personalized approach, with automated

parameter selection adapted to the dataset, making it easily applicable to segmented microglia in

diverse contexts, including those with distinct morphologies, such as microglia in culture or *in vivo*

acquisitions. It does not require program adaptation or the need to train a neural network for

deep/machine learning approaches. We observed that some selected parameters were consistent

between all the tested datasets (ramification index and perimeter area ratio for example), showing

their robustness to differentiate microglial cells with various morphologies. On the contrary, some

parameters were never selected (inertia and linearity for example). The information contained in

these parameters were not discriminant enough in the datasets we used in this present work, but we

can imagine that they would have been selected if we would have chosen a model in which microglial

growth towards a specific site is observed, such as focal laser injury (Davalos et al., 2005; Haynes et

al., 2006).

We also introduce a ranking tool to further enhance morphological analysis. In contrast to recent

studies, we opt for ordering cells rather than classifying them. Microglial morphologies can span a

continuum with no clear boundaries between morphotypes. Automated clustering methods such as

k-means (Young & Morrison, 2018) may be more suggestive, as they require an *a priori*

determination of the number of clusters. Hierarchical clustering (Clarke et al., 2021) can avoid this

requirement but may result in non-reproducible clusters, or clusters without biological meaning.

Attempts to identify distinct morphotypes, as previously done with artificial intelligence (Leyh et al.,

2021), are also questionable due to varying definitions of morphotypes between laboratories and a

lack of consideration for subtle morphological differences and intermediate forms. This problem is

overcome by the ranking approach, as it does not lock the cells in defined categories but rather

position the cells in relation to one another. However, the ability to compare the evolution of

clusters of microglial morphologies in different conditions remains valuable, especially when

morphologies are heterogeneous in the same population, such as in developmental microglia or

pathological contexts. Comparing medians or means of certain parameters may not reveal

differences due to high variability, whereas comparing the evolution of the proportions of specific

morphotypes may provide more informative insights. The Andrews plots used to generate a ranking

of the cells could also be used to generate clusters of the cells presenting morphological

resemblances, as similar cells display similar Andrews curves' trajectories.

We chose not to integrate a segmentation tool into our program, as the segmentation method

should be tailored to the specific imaging conditions, materials, and microscopes used. Diverse

segmentation methods are available in different software, with varying levels of automation and

manual input requirements. Although, automated thresholding would be ideal. In our case, image

acquisitions were not entirely uniform, even within the same sample. For instance, in ischemic brain

samples, lipofuscin from cell death introduces background noise that can artificially impact threshold

levels. This effect is observed even when comparing contralateral and ipsilateral sides of the same

brain. In our experience, manually adjusting the threshold provides a more accurate, reliable, and

comparable selection of cellular elements, even though it introduces some subjectivity. To ensure

consistency in segmentation, we recommend that the same person performs the analysis across all

conditions. In this work, we present methods using FIJI. It is important to note that achieving perfect

cell segmentation is not a prerequisite for using MorphoCellSorter since the tool does not aim to

provide highly precise measurements of microglial morphology. For instance, discontinuous

processes are not problematic. Consequently, it is possible to work with lower-resolution

microscopes, such as epifluorescence wide-field microscopes, although the segmentation process

may be more labor-intensive if the source images are of lower quality. Obviously, the results of the

ranking will be highly affected by the quality and the homogeneity of segmentation in a given data

set. However, we did not assess how much the resolution of the microscope used could affects the

quality of the ranking as low resolution and poor signal to noise ratio would mainly affect the

segmentation step which is not the purpose of this paper.

One potential limitation is that we decided to work with 2D Z-stack projections, as accurate 3D

reconstructions are time-consuming and require high-resolution acquisitions. Our goal was to

develop a fast and robust method that could be broadly applicable to most microscopes and imaging

scenarios. Therefore, our approach may not yield precise measurements of parameters such as the

exact number and length of processes or their hierarchy. However, other research groups have

successfully developed tools capable of providing these types of measurements, such as 3DMorph

(York, LeDue, Bernier, & MacVicar, 2018) and MorphOMICS via Imaris (Colombo et al., 2022), Mic-

Mac (Salamanca et al., 2019). Instead, MorphoCellsorter is a rapid way to determine if a treatment,

genotype or location of cells affects morphology without going through the laborious work of 3D

reconstruction. Our tool can thus be complementary to methods that provide those precise metrics.

In summary, MorphoCellSorter is an efficient and user-friendly tool for morphological analysis that

operates with minimal computation time. The code is open-source on GitHub, accessible by the

community (https://github.com/Pascuallab/MorphCellSorter) and improved according to the needs.

It can be complemented with additional measures, such as the expression of specific markers, to

investigate the relationship between morphology and gene expression. The emergence of spatial

transcriptomics is a promising avenue for furthering our understanding of the link between

morphology and phenotype in the microglia field in which MorphoCellSorter will be of great interest.

Figure legends

Figure 1: k-means clustering identifies highly different populations contralaterally vs ipsilaterally in

an ischemic stroke model but offers a heterogeneous cluster composition.

A T2-weighted MRI acquisition of the tMCAO rat brain 24 hours after reperfusion (top). Z maximum

projections of IBA1 staining of microglia on the ipsilateral side to the lesion (bottom left) and

contralateral side (bottom right). Scale bar: 20 µm. **B** Scatter plot with each dot representing one

microglia from the complete dataset (ipsilateral and contralateral), plotted in function of their PC1

against PC2 values from the PCA conducted on all 20 initial morphological parameters. The colors

represent the clusters identified by the k-means method, and the 'x' symbols are the centers of each

cluster. **C** Sample of microglial cells and their affiliated cluster **D** Pie chart illustrating the

proportion of each cluster constituting the contralateral and ipsilateral sides.

Figure 2: **MorphoCellSorter approach to generate a ranking of the cells based on their morphology**

applied to the ischemic stroke model on fixed tissue.

A Summary of the fully automated MorphoCellSorter approach: morphological indexes of the

individualized and binarized microglia are computed on a first script (MorphoCellMeter), and the

resulting data table is used as an entry for MorphoCellSorter to generate the ranking of the cells

based on their morphological characteristics. **B** Cascade of normalized eigenvalues. The blue bars

represent the eigenvalues for each principal component (PCs), and the red curve is the cumulative

sum of the values. The first two eigenvalues corresponding to the first two PCs that account for over

70% of the total sum. **C** Correlation circle in the PC1 and PC2 planes. **D** Parameters ranked according

to their projection on PC1. Values are normalized by the highest projection value. To determine the

number of morphological parameters selected, here we consider the median number of the

normalized cumulative projection (pink curve), leading to a threshold at 0.5 (red line), and thus 7

kept parameters in this case. **E** Andrews plots generated with the 7 parameters weighted according

to their contribution onto PC1. Each curve represents one microglia in the high-dimensional space of

the 7 strongest parameters' projections onto PC1. The maximum variance between curves is

obtained at 359°.

Figure 3: MorphoCellSorter offers a reliable ranking of microglia in an ischemic stroke model based

on their morphologies.

A Automatic ranking generated by MorphoCellSorter of the 347 microglia constituting the ischemic

stroke model dataset. **B** Distribution of the rank difference between expert 1 manual and

MorphoCellSorter rankings. **C** Distribution of the rank difference between expert 2 manual and

MorphoCellSorter rankings. **D** Distribution of the rank difference between expert 1 and expert 2

manual rankings. **E** Correlation between expert 1 and MorphoCellSorter rankings. Spearman's

correlation coefficient Rs = 95968, p<0.0001. **F** Correlation between expert 2 and MorphoCellSorter

rankings. Spearman's correlation coefficient Rs = 97188, p<0.0001. **G** Correlation between expert 1

and expert 2 rankings. Spearman's correlation coefficient Rs = 93498, p<0.0001.

Table 1: Evaluation of MorphoCellSorter rankings on highly heterogeneous datasets of microglia.

The evaluation of the rankings is conducted by comparing the Spearman's correlation coefficients. w:

weeks, m: months, E: embryonic day, P: postnatal day, DIV: day *in vitro*, Ex: expert.

Figure 4: MorphoCellSorter identifies morphological alterations in the visual cortex of APPxPS1-KI

mice.

A Z maximum projections of IBA1 (microglia) and 4G8 (Aβ plaques) stainings of control and APPxPS1-

KI mice in the visual cortex. The white arrowheads point at Aβ plaques positive for 4G8, green

arrowheads designate ramified microglia far from Aβ plaques and red arrowheads point to highly

morphologically altered microglia close to Aβ plaques. Scale bar: 20µm. **B** MorphoCellSorter

automated ranking. Control microglia are represented in green and APPxPS1-KI microglia in orange. **C**

Distribution of the Andrews values at the maximum variance (Andrews scores) for control and

APPxPS1-KI microglia. **D-W** Morphological indexes computed by MorphoCellSorter. The data shown

are the mean ± standard deviation (std). Linear mixed models were applied when the application

conditions were respected and Wilcoxon tests were performed otherwise (for the Span Ratio) (N=3

mice for each condition, $n_{APP-PS1xKI}$ and $n_{Control}$ = 180). *p<0.05; **p<0.01; ***p<0.001.

Figure 5: Accurate sorting of embryonic microglia reveals no morphological difference between

CCHS and WT embryos' microglia.

A IBA1 stainings in WT and CCHS embryos' brainstems. **B** MorphoCellSorter automated ranking. WT

microglia are displayed in light purple and mutated microglia are represented in pink. Only every 3

microglia composing the dataset are displayed for readability purposes. **C** Distribution of the

Andrews scores for CCHS and WT microglia. **D-W** Morphological indexes computed by

MorphoCellSorter. The data's distribution is represented as well as the mean ± std. Linear mixed

models were applied (N_{WT}=5, N_{mutant}=4; n_{WT}=498 and n_{mutant}=444).

Supplementary information

Supplemental figure 1: Parameters ranked according to their normalized projection onto PC1 for

the 5 tested datasets.

Parameters ranked according to their projection onto PC1 for the following datasets: temporary

mean cerebral artery occlusion (tMCAo) on fixed tissue (**A**), permanent mean cerebral artery

occlusion (pMCAo) on live tissue (**B**), Alzheimer's disease (AD) model (**C**), embryonic microglia (**D**) and

microglia in culture (**E**). Values are normalized by the highest projection value. The number of

morphological parameters selected is determined by the threshold corresponding value of the

normalized cumulative projection selected (pink curve).

Supplemental table 1: Identification of the optimal threshold.

EM: Embryonic microglia, MC: Microglia in culture.

Supplemental figure 2: Andrews plots at a 0.8 threshold.

Andrews plots for the following datasets: temporary mean cerebral artery occlusion (tMCAo) on fixed

tissue (**A**), permanent mean cerebral artery occlusion (pMCAo) on live tissue (**B**), Alzheimer's disease

(AD) model (**C**), embryonic microglia (**D**) and microglia in culture (**E**).

Supplemental figure 3: MorphoCellSorter rankings generated for the 5 datasets at a 0.8 threshold.

A Ranking for the tMCAo fixed tissue data set, n=347 cells**. B** Ranking for the pMCAo live tissue data

set, n= 168 cells**. C** Ranking for the AD model data set, n=180 cells. **D** Ranking for the embryonic data

set, n=407 cells. **E** Ranking for microglia in culture data set, n=96

Supplemental figure 4: Elaboration of MorphoCellSorter ranking for the AD dataset.

A Cascade of normalized eigenvalues. **B** Correlation circle in the PC1 and PC2 planes. **C** Parameters

ranked according to their projection on PC1. Values are normalized by the highest projection value.

To determine the number of morphological parameters selected, we selected a threshold of 0.8 (red

line), and thus 11 kept parameters. **D** Andrews plots generated with the 11 parameters weighted

according to their contribution onto PC1. Each curve represents one microglia in the high-

dimensional space of the selected parameters' projections onto PC1. The maximum variance

between curves is obtained at 1°.

Supplemental figure 5: Elaboration of MorphoCellSorter ranking for the CCHS dataset.

A Cascade of normalized eigenvalues. **B** Correlation circle in the PC1 and PC2 planes. **C** Parameters

ranked according to their projection on PC1. Values are normalized by the highest projection value.

To determine the number of morphological parameters selected, here we used a threshold of 0.8

(red line), and thus 12 kept parameters. **D** Andrews plots generated with the 12 parameters

weighted according to their contribution onto PC1. Each curve represents one microglia in the high-

dimensional space of the selected parameters. The maximum variance between curves is obtained at

359°.

Supplemental figure 6: MorphoCellSorter automated ranking of the CCHS dataset.

WT microglia are displayed in light purple and mutated microglia are represented in pink.

Supplemental figure 7: Pre-processing and segmentation procedure for all used datasets.

Steps in grey were performed on FIJI, blue on MATLAB, yellow on ZEISS ZEN and purple on Procreate.

Supplemental figure 8: Illustration of the 20 morphological indexes automatically computed.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The original code of MorphoCellSorter is available on GitHub (https://github.com/PascualLab/MorphocellSorter).

The datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study is supported by research grants from Fondation pour la recherche médicale (reference DQ20170336751) and has been developed within the BETPSY project, which is supported by a public grant overseen by the French national research agency (Agence nationale de la recherche, ANR), as part of the second "Investissements d'Avenir" program (reference ANR-18-RHUS-0012). This work was also supported by the French National Research Agency within the framework of the LABEX CORTEX ANR-11-LABX-0042.

Authors' contributions

SB, SMD, PM, JC and OP conceptualized and designed the study. JC and OP were in charge of overall direction. The experiments were conducted by SB, MTB, LC, BP, GP CL, IH, VH, MW and SM. JH and OP secured the fundings. Coding was executed by SB, SMD, PM and JC. SB, PDG and KC participated in the analysis. All authors provided critical feedback and helped shape the research, analysis and

manuscript. SB and OP drafted the manuscript, and all authors provided critical review, commentary

or revision.

Acknowledgments

We acknowledge the contribution of SFR Santé Lyon-Est (UAR3453 CNRS, US7 Inserm, UCBL) facility:

CIQLE (a LyMIC member), especially Bruno Chapuis and Denis Ressnikoff for their help.

References

A, C. (2023). Hausdorff (Box-Counting) Fractal Dimension.

Adaikkan, C., Middleton, S. J., Marco, A., Pao, P. C., Mathys, H., Kim, D. N., . . . Tsai, L. H. (2019).

Gamma Entrainment Binds Higher-Order Brain Regions and Offers Neuroprotection. *Neuron,*

102(5), 929-943.e928. doi:10.1016/j.neuron.2019.04.011

Adeluyi, A., Guerin, L., Fisher, M. L., Galloway, A., Cole, R. D., Chan, S. S. L., . . . Turner, J. R. (2019).

Microglia morphology and proinflammatory signaling in the nucleus accumbens during

nicotine withdrawal. *Science Advances, 5*(10), eaax7031. doi:10.1126/sciadv.aax7031

Ali, S., Liu, X., Queen, N. J., Patel, R. S., Wilkins, R. K., Mo, X., & Cao, L. (2019). Long-term

environmental enrichment affects microglial morphology in middle age mice. *Aging, 11*(8),

2388-2402. doi:10.18632/aging.101923

Amiel, J., Laudier, B., Attié-Bitach, T., Trang, H., de Pontual, L., Gener, B., . . . Lyonnet, S. (2003).

Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B

in congenital central hypoventilation syndrome. *Nature Genetics, 33*(4), 459-461.

doi:10.1038/ng1130

Andrews, D. F. (1972). Plots of High-Dimensional Data. *Biometrics, 28*(1), 125-136.

doi:10.2307/2528964

Anttila, J. E., Whitaker, K. W., Wires, E. S., Harvey, B. K., & Airavaara, M. (2017). Role of microglia in

ischemic focal stroke and recovery: focus on Toll-like receptors. *Progress in neuro-*

psychopharmacology & biological psychiatry, 79(Pt A), 3-14.

doi:10.1016/j.pnpbp.2016.07.003

Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K. W., Schindelin, J., Cardona, A., & Sebastian

Seung, H. (2017). Trainable Weka Segmentation: a machine learning tool for microscopy pixel

classification. *Bioinformatics, 33*(15), 2424-2426. doi:10.1093/bioinformatics/btx180

Bohlen, C. J., Bennett, F. C., & Bennett, M. L. (2019). Isolation and culture of microglia. *Current*

protocols in immunology, 125(1), e70. doi:10.1002/cpim.70

Casas, C., Sergeant, N., Itier, J.-M., Blanchard, V., Wirths, O., van der Kolk, N., . . . Pradier, L. (2004).

Massive CA1/2 neuronal loss with intraneuronal and N-terminal truncated Abeta42

accumulation in a novel Alzheimer transgenic model. *The American Journal of Pathology,*

165(4), 1289-1300. doi:10.1016/s0002-9440(10)63388-3

Ceccherini, I., Kurek, K. C., & Weese-Mayer, D. E. (2022). Developmental disorders affecting the

respiratory system: CCHS and ROHHAD. *Handb Clin Neurol, 189*, 53-91. doi:10.1016/B978-0-

323-91532-8.00005-7

Choi, S., Hill, D., Guo, L., Nicholas, R., Papadopoulos, D., & Cordeiro, M. F. (2022). Automated

characterisation of microglia in ageing mice using image processing and supervised machine

learning algorithms. *Scientific Reports, 12*(1), 1806. doi:10.1038/s41598-022-05815-6

Clarke, D., Crombag, H. S., & Hall, C. N. (2021). An open-source pipeline for analysing changes in

microglial morphology. *Open Biology, 11*(8), 210045. doi:10.1098/rsob.210045

Colombo, G., Cubero, R. J. A., Kanari, L., Venturino, A., Schulz, R., Scolamiero, M., . . . Siegert, S.

(2022). A tool for mapping microglial morphology, morphOMICs, reveals brain-region and

sex-dependent phenotypes. *Nature Neuroscience, 25*(10), 1379-1393. doi:10.1038/s41593-

022-01167-6

Davalos, D., Grutzendler, J., Yang, G., Kim, J. V., Zuo, Y., Jung, S., . . . Gan, W.-B. (2005). ATP mediates

rapid microglial response to local brain injury in vivo. *Nature Neuroscience, 8*(6), 752-758.

doi:10.1038/nn1472

Dubreuil, V., Ramanantsoa, N., Trochet, D., Vaubourg, V., Amiel, J., Gallego, J., . . . Goridis, C. (2008).

A human mutation in Phox2b causes lack of CO_2 chemosensitivity, fatal central apnea, and

specific loss of parafacial neurons. *Proc Natl Acad Sci U S A, 105*(3), 1067-1072.

doi:10.1073/pnas.0709115105

Fernández-Arjona, M. d. M., Grondona, J. M., Granados-Durán, P., Fernández-Llebrez, P., & López-

Ávalos, M. D. (2017). Microglia Morphological Categorization in a Rat Model of

Neuroinflammation by Hierarchical Cluster and Principal Components Analysis. *Front Cell*

Neurosci, 11, 235. doi:10.3389/fncel.2017.00235

Fumagalli, S., Perego, C., Ortolano, F., & De Simoni, M.-G. (2013). CX3CR1 deficiency induces an early

protective inflammatory environment in ischemic mice. *Glia, 61*(6), 827-842.

doi:10.1002/glia.22474

García-Osorio, C., & Fyfe, C. (2005). Visualization of High-dimensional Data via Orthogonal Curves.

JUCS - Journal of Universal Computer Science, 11((11)), 1806-1819. doi:10.3217/jucs-011-11-

1806

Hanisch, U.-K., & Kettenmann, H. (2007). Microglia: active sensor and versatile effector cells in the

normal and pathologic brain. *Nat Neurosci, 10*(11), 1387-1394.

Haynes, S. E., Hollopeter, G., Yang, G., Kurpius, D., Dailey, M. E., Gan, W. B., & Julius, D. (2006). The

P2Y12 receptor regulates microglial activation by extracellular nucleotides. *Nat Neurosci,*

9(12), 1512-1519.

Heindl, S., Gesierich, B., Benakis, C., Llovera, G., Duering, M., & Liesz, A. (2018). Automated

Morphological Analysis of Microglia After Stroke. *Front Cell Neurosci, 12*.

doi:10.3389/fncel.2018.00106

Hristovska, I., Robert, M., Combet, K., Honnorat, J., Comte, J. C., & Pascual, O. (2022). Sleep

830   decreases neuronal activity control of microglial dynamics in mice. *Nat Commun, 13*(1), 6273.

831   doi:10.1038/s41467-022-34035-9

832   Hubert, V., Hristovska, I., Karpati, S., Benkeder, S., Dey, A., Dumot, C., . . . Wiart, M. (2021).

833   Multimodal Imaging with NanoGd Reveals Spatiotemporal Features of Neuroinflammation

834   after Experimental Stroke. *Adv Sci (Weinh), 8*(17), e2101433. doi:10.1002/advs.202101433

835   Hubert, V., Hristovska, I., Karpati, S., Benkeder, S., Dey, A., Dumot, C., . . . Wiart, M. (2021).

836   Multimodal Imaging with NanoGd Reveals Spatiotemporal Features of Neuroinflammation

837   after Experimental Stroke. *Adv Sci (Weinh), 8*(17), e2101433. doi:10.1002/advs.202101433

838   Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments.

839   *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences,*

840   *374*(2065), 20150202. doi:10.1098/rsta.2015.0202

841   Kettenmann, H., Hanisch, U.-K., Noda, M., & Verkhratsky, A. (2011). *Physiology of Microglia* (Vol. 91).

842   Kongsui, R., Beynon, S. B., Johnson, S. J., & Walker, F. R. (2014). Quantitative assessment of microglial

843   morphology and density reveals remarkable consistency in the distribution and morphology

844   of cells within the healthy prefrontal cortex of the rat. *J Neuroinflammation, 11*, 182.

845   doi:10.1186/s12974-014-0182-7

846   Leyh, J., Paeschke, S., Mages, B., Michalski, D., Nowicki, M., Bechmann, I., & Winter, K. (2021).

847   Classification of Microglial Morphological Phenotypes Using Machine Learning. *Front Cell*

848   *Neurosci, 15*, 701673. doi:10.3389/fncel.2021.701673

849   Madry, C., Kyrargyri, V., Arancibia-Cárcamo, I. L., Jolivet, R., Kohsaka, S., Bryan, R. M., & Attwell, D.

850   (2018). Microglial Ramification, Surveillance, and Interleukin-1β Release Are Regulated by the

851   Two-Pore Domain K+ Channel THIK-1. *Neuron, 97*(2), 299-312.e296.

852   doi:10.1016/j.neuron.2017.12.002

Morrison, H., Young, K., Qureshi, M., Rowe, R. K., & Lifshitz, J. (2017). Quantitative microglia analyses

reveal diverse morphologic responses in the rat cortex after diffuse brain injury. *Scientific*

Reports, 7(1), 13211. doi:10.1038/s41598-017-13581-z

Nimmerjahn, A., Kirchhoff, F., & Helmchen, F. (2005). Resting microglial cells are highly dynamic

surveillants of brain parenchyma in vivo. *Science, 308*(5726), 1314-1318.

Orłowski, D., Sołtys, Z., & Janeczko, K. (2003). Morphological development of microglia in the

postnatal rat brain. A quantitative study. *International Journal of Developmental*

Neuroscience: The Official Journal of the International Society for Developmental

Neuroscience, 21(8), 445-450. doi:10.1016/j.ijdevneu.2003.09.001

Paolicelli, R. C., Sierra, A., Stevens, B., Tremblay, M.-E., Aguzzi, A., Ajami, B., . . . Wyss-Coray, T.

(2022). Microglia states and nomenclature: A field at its crossroads. *Neuron, 110*(21), 3458-

3483. doi:10.1016/j.neuron.2022.10.020

Plescher, M., Seifert, G., Hansen, J. N., Bedner, P., Steinhäuser, C., & Halle, A. (2018). Plaque-

dependent morphological and electrophysiological heterogeneity of microglia in an

Alzheimer's disease mouse model. *Glia, 66*(7), 1464-1480. doi:10.1002/glia.23318

R, B. (2023). Thresholding Tool.

Salamanca, L., Mechawar, N., Murai, K. K., Balling, R., Bouvier, D. S., & Skupin, A. (2019). MIC-MAC:

An automated pipeline for high-throughput characterization and classification of three-

dimensional microglia morphologies in mouse and human postmortem brain samples. *Glia,*

67(8), 1496-1509. doi:10.1002/glia.23623

Sierra, A., Abiega, O., Shahraz, A., & Neumann, H. (2013). Janus-faced microglia: beneficial and

detrimental consequences of microglial phagocytosis. *Front Cell Neurosci, 7*, 6.

doi:10.3389/fncel.2013.00006

T, V. (2023). Lacunarity of a binary image.

Tolle, C. R., McJunkin, T. R., & Gorsich, D. J. (2008). An efficient implementation of the gliding box

lacunarity algorithm. *Physica D: Nonlinear Phenomena, 237*(3), 306-315.

doi:10.1016/j.physd.2007.09.017

Verdonk, F., Roux, P., Flamant, P., Fiette, L., Bozza, F. A., Simard, S., . . . Danckaert, A. (2016).

Phenotypic clustering: a novel method for microglial morphology analysis. *J*

Neuroinflammation, 13, 153. doi:10.1186/s12974-016-0614-7

Vidal-Itriago, A., Radford, R. A. W., Aramideh, J. A., Maurel, C., Scherer, N. M., Don, E. K., . . . Morsch,

M. (2022). Microglia morphophysiological diversity and its implications for the CNS. *Frontiers*

in Immunology, 13, 997786. doi:10.3389/fimmu.2022.997786

York, E. M., LeDue, J. M., Bernier, L.-P., & MacVicar, B. A. (2018). 3DMorph Automatic Analysis of

Microglial Morphology in Three Dimensions from Ex Vivo and In Vivo Imaging. *eNeuro, 5*(6),

ENEURO.0266-0218.2018. doi:10.1523/ENEURO.0266-18.2018

Young, K., & Morrison, H. (2018). Quantifying Microglia Morphology from Photomicrographs of

Immunohistochemistry Prepared Tissue Using ImageJ. *Journal of Visualized Experiments:*

JoVE(136), 57648. doi:10.3791/57648

Zhang, S. (2019). Microglial activation after ischaemic stroke. *Stroke and Vascular Neurology, 4*(2),

71-74. doi:10.1136/svn-2018-000196

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5