Corresponding author(s): Gonçalo Castelo-Branco, Anna Williams, Charles ffrench-Constant

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒ Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Zeiss Zen 2.3 Blue Edition Leica Application SuiteX 3.5.2.18963 Vectra Polaris 1.0.7 Information regarding data collection is described in the methods section of the manuscript and data can be browsed at https://castelobranco.shinyapps.io/MSCtrl_CCA_18/ and https://ki.se/en/mbb/oligointernode.
Data analysis	QuPath v0.1.2 Fiji-ImageJ v1.52h Graph Pad Prism 7. Cytoscape version 3.7.0 python packages : python 3.6 Cellranger version 2.1.1 Velocyto.py 0.17.7 Own developed R code: https://github.com/Castelo-Branco-lab/GeneFocus

R packages: R version 3.4.3 and R version 3.5:
Velocyto.R_0.5
Matrix_1.2-14
Seurat_2.3.4 / Seurat.2.1.0
ggplot2_3.0.0
https://github.com/maggiecrow/MetaNeighbor.git , 2017-08-28-runMN-US.R
MAST_1.4.1
SummarizedExperiment_1.8.1
DelayedArray_0.4.1
matrixStats_0.53.1
GenomicRanges_1.30.3
GenomeInfoDb_1.14.0
S4Vectors_0.16.0
Biobase_2.38.0
BiocGenerics_0.24.0
cowplot_0.9.2
scater_1.8.1
SingleCellExperiment_1.2.0

All tools described in the methods section of the manuscript, source code and notebooks are available at https://github.com/Castelo-Branco-lab under the repository Jaekel_Agirre_et_al_2018

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

No restrictions apply on data availability. All figures have associated source code and/or raw data, these are: Fig 1-4 and Extended Data Figures 1-8. Count and annotation matrices are available at https://ki.se/en/mbb/oligointernode. Sequence data has been deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number EGAS00001003412 . UMI expression and cell type annotation tables have been deposited in GEO, accession number GSE118257. The data can be explored and visualized in https://castelobranco.shinyapps.io/MSCtrl_CCA_18/ and https://ki.se/en/mbb/oligointernode.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

In the following section (also as displayed in the manuscript in the Figure legends and in the Materials and Methods section), n represents the individual number of different tissue donors (=different individuals), in case not stated differently. No sample size calculation was performed to pre-determine sample sizes. Rationale for sample size choice is provided for each of the experiments described below:

For the snRNA-seq experiment we used a sample size of n=5 Ctr and n=4 MS donors. We further sub-sampled the MS tissue into 3 normal appearing white matter (NAWM) samples from 3 different patients, 4 chronic active lesion samples from 4 different patients, 3 active lesion samples from 2 different patients, 3 chronic inactive lesion samples from 3 different patients and 2 remyelinating lesion samples from 2 different patients. These number of samples was sufficient to obtain the number of cells necessary to perform a confident data analysis. It was also possible to perform confident comparisons between controls and MS patients, between different lesion types and between this and already published datasets from human brain snRNA-seq experiments.

The chosen sample size for validations of findings arising from snRNA-Seq is in the same range of other publications in the research field using human post mortem donor tissue. The number of samples is limited by tissue availability from the brain banks. For the qualitative validation of mRNA markers in combination with immunohistochemistry, the sample size was a minimum of n=2. For the qualitative validation of double immunohistochemistry, the sample size was n=3.

In Fig.1c, we used n=4 for LURAP1L.AS1+CDH20+, n=3 for the other combinations of OL subclass markers.
In Fig.3a we used n=5 Ctr and n=9 MS patients. In Fig.3b (SOX6) n represents the number of different donors. We used n=4 Ctr and n=5

NAWM and lesions.
In Fig. 3c (OPALIN) we used n=3 for Ctr and n=5 for NAWM and Lesions.
In Fig.4e (CDH20 BaseScope) we used n=5 for active lesion, n=7 for chronic inactive lesions and n=3 for chronic active lesions in a total of 7 MS patients.
For Extended Fig.4b-e we used n=3 different donors to validate the co-labeling of each marker.
For Extended Fig.4g (OPALIN bins) we used n=3 different donors for each group.
In Extended Fig. 8a (BCAN) we used n=4 Ctr, n=6 NAWM and n=5 Lesions. I
In Fig. 8b (KLK6) we used n=4 Ctr and n=5 NAWM and lesions.
In Extended Fig.8d (MYRF) we used n=6 Ctr and NAWM and n=7 lesions.
For Extended Fig.9a, the individual number of quantified mRNA molecules per field per patient (n=7) are shown. We used the following number of fields: MS235: n=10 for A and CI lesions, MS200: n=4 for A, CI and CA lesions, MS249: n=4 for A and n=8 for CI lesions, MS361: n=7 for A and n=10 for CI lesions, MS106: n=11 for CA and CI lesions, MS161: n=6 for CA and n=10 for CI lesions, MS300: n=7 for A and n=10 for CI lesions.
For Extended Fig.9b (WWOX) we used n=2 active lesions, n=5 chronic inactive lesions and n=4 chronic active lesions in a total of 5 MS patients. For Extended Fig.9c, the individual number of quantified mRNA molecules per field per patient (n=5) are shown. We used the following number of fields: MS245: n=8 for A, n=10 for CI and n=9 for CA lesions, MS361: n=6 for A and n=10 for CI lesions, MS101: n=6 for CI and n=11 for CA lesions, MS161: n=10 for CI and n=7 for CA lesions, MS296: n=11 for CA and n=6 for CI lesions.

Data exclusions

For single nuclei RNA-Seq data, we excluded data points through our quality control pipeline, as indicated in the methods section in the paper. In short, data was excluded based on thresholding of cells with transcript count less than 300, less than 200 expressing genes, more than 6000 UMI counts, a mitochondrial count ratio more than 0.20 and with 1 count above in at least 2 cells. Then, specific datasets from published studies were filtered specifically as explained in the methods section. After clustering we identify one nuclei cluster with high mitochodrial and ribosomal gene expression that was removed.

Replication

Verification of the experimental findings derived from single nuclei RNA-seq is performed through validation by mRNA (BaseScope) and protein (Immunohistochemistry) on a different set tissue sections derived from different donors. All validation data was reproducible and no patient datasets were excluded.

Randomization

For both the snRNA-seq and the in-situ validations, we used donor tissue from both sexes randomly distributed in each group (see manuscript Table 1).
The age of the donors was also randomly chosen and distributed between the groups and does not display significant differences (see Table 1 and Methods section).
For Immunohistochemistry, the fields for image quantification have been chosen randomly throughout the tissue section not to introduce any regional bias.

Blinding

The tissue for the snRNA-seq experiment could not be collected blinded, as specific areas had to be chosen. However, the experiment and subsequent data analysis has been performed fully automated and does hence not display any bias.
Where possible, analysis of Immunohistochemistry has been performed automated with a image analysis software to exclude any bias. In the other cases, quantification has been validated by 2 independent people.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Primary antibodies:
- anti-OLIG1 (Abcam, Cat. number ab68105, Lot number GR236765-4);
- anti-OLIG2 (rabbit, Atlas Antibodies, Cat. number HPA003254, Lot number CC81836);
- anti-OLIG2 (goat, R&D Systems, Cat. number AF2418, Lot number UPA0718031);
- anti-MYRF (Millipore, Cat. number ABN45, Lot number 2652011 and 2909641)
- anti OPALIN (Abcam, Cat number ab121425, Lot number GR264580-3)
- anti KLK6 (Life technologies, Cat number PA5-47239, TD2560533B)
- anti-CNP (Atlas Antibodies, Cat number AMAb91072, Lot number 02942)
-anti-SOX6 (Millipore, Cat number AB5805, Lot number 2921391)
-anti-NOGOA (R&D Systems, Cat number MAB3098, Lot number YYM0215021)

Secondary antibodies:
- rb-HRP IgG (Vector laboratories, Cat number MP-7401, Lot number ZD1019)
- rb-AP IgG (Vector laboratories, Cat number MP-5401, Lot number ZE0530)
- gt-HRP IgG (Vector laboratories, Cat number MP-7405, Lot number ZE0425)
- ms-HRP IgG (Vector laboratories, Cat number MP-7402, Lot number ZE0227)
- ms-AP IgG (Vector laboratories, Cat number MP-5402, Lot number ZC0831)

Colour/Fluorescent reaction kits:
- Fluorescein (Perkin Elmer, Cat number NEL741B001KT, Lot number 2779696)
- Cyanine 3 (Perkin Elmer, Cat number NEL744B001KT, Lot number 2328502)
- Cyanine 5 (Perkin Elmer, Cat number NEL745B001KT, Lot number 1859566)
- DAB (Vector laboratories, Cat number SK-4100, Lot number ZE0102)
- VectorBlue (Vector laboratories, Cat number SK-5300, Lot number ZE0402)

Validation

- anti-OLIG1: reactivity validated by the company for Rat, predicted to react with Human Mouse, Pig, Chimpanzee, Monkey, Baboon, Common marmoset. Has been used for human in Jakovcevski et al. 2005. Validated by the company for IHC on PFA-fixed frozen and fresh frozen tissue sections.
- anti-OLIG2: reactivity validated by the company for Human. Validated by the company for IHC. Has been validated by the Human Protein Atlas in 44 human control brain samples.
- anti-OLIG2: reactivity validated by the company for Human, Rat and Mouse. Validated by the company for IHC. Company has 30 citations for the use of this antibody.
- anti-MYRF: reactivity validated by the company for Human and Mouse. Validated by the company for IHC. MYRF antibody has further been validated by WB for specific binding as well as with a combination of mRNA and protein labeling of MYRF in the same cell (Extended Data Fig 6).
- anti OPALIN: reactivity validated by the company for Human. Validated by the company for IHC on paraffin-embedded tissue.
- anti KLK6: reactivity validated by the company for Human. Validated by the company for IHC on paraffin-embedded tissue.
- anti-CNP: reactivity validated by the company for Human, Rat and Mouse. Validated by the company for IHC and WB. Antibody has an entry at the Human Protein Atlas as validated antibody.
-anti-SOX6: reactivity validated by the company for Human and Mouse. Validated by the company for ICC.
-anti-NOGOA: reactivity validated by the company for Rat. This antibody has been used in human tissue in Kuhlmann et al. 2008. Validated by the company for IHC.

Human research participants

Policy information about studies involving human research participants

Population characteristics

For the snRNA-seq experiment:
The study included 5 Ctr (4 male, and 1 female) and 4 MS (3 male and 1 female) donors all in an age range between 35 and 82 years. Non of the Ctr donors had any known neurological disorders and no known medical treatment in this regard. The MS donors were all diagnosed with Chronic Multiple sclerosis (<14 years of disease duration) and non of them had a specific disease treatment.
For the in-situ validations:
The study included 11 Ctr (5 male and 6 female) and 15 MS (7 male and 8 female) donors in an age range between 36 and 77 years. Non of the Ctr donors had any known neurological disorders and no known medical treatment in this regard. The MS donors were all diagnosed with Chronic Secondary Progressive Multiple sclerosis (<18 years of disease duration for the cases with available information) and non of them had a specific disease treatment.

Detailed information is given in the Materials and Methods section of the manuscript and in Table1.

Recruitment

No donors were recruited, the tissue has been obtained from an accredited UK tissue bank.