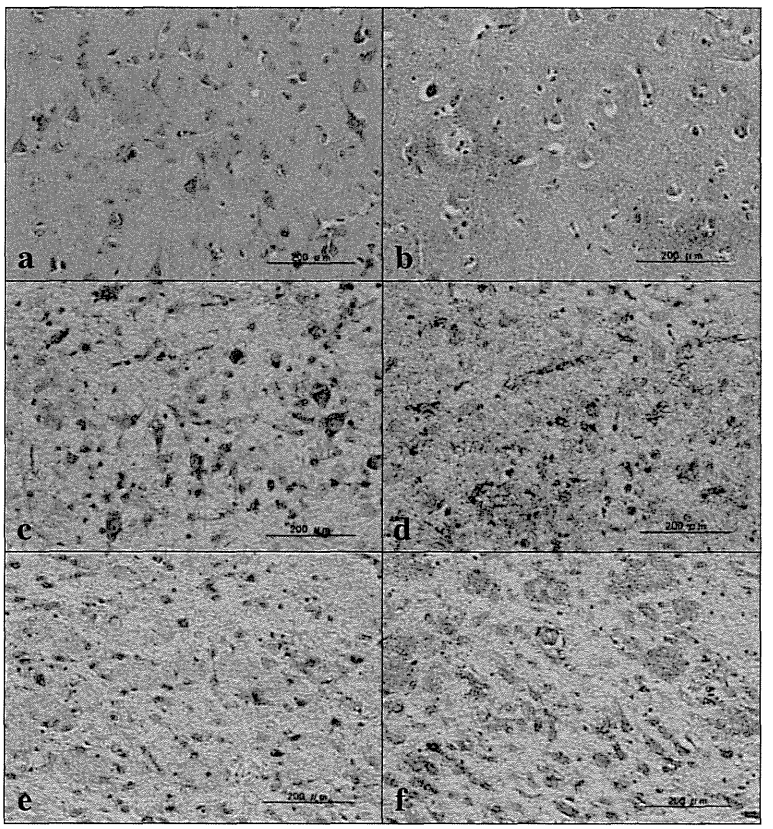

Figure 8 TMEM106B and PGRN immunoreactivities in Alzheimer's disease brains. Expression of TMEM106 and progranulin (PGRN) immunoreactivities was studied in six Alzheimer's disease brains presented in Table 1 by immunohistochemistry using the A303-439A antibody. **(a)** TMEM106B, the frontal cortex, moderate neuronal cytoplasmic staining and faint senile plaque staining; **(b)** PGRN, same region as **(a)**, moderate senile plaque staining and diffuse neuropil staining; **(c)** TMEM106B, the hippocampal CA1 region, intense neuronal and astroglial cytoplasmic staining; **(d)** PGRN, same region as **(c)**, intense perivascular neuropil staining; **(e)** TMEM106B, the hippocampal CA1 region, no staining of senile plaques and neurofibrillary tangles; **(f)** PGRN, same region as **(e)**, moderate staining of numerous senile plaques and neurofibrillary tangles.

TMEM106B in AD brains. Further studies are required to evaluate this possibility.

In AD brains, granulovaculoar degeneration bodies, a kind of autophagosome, express immunoreactivity for charged multivesicular body protein 2B (CHMP2B), whose genetic mutations definitely cause FTLD [33]. We found that granulovacuolar degeneration vacuoles (GVD) located in hippocampal CA1 pyramidal neurons of AD brains are devoid of TMEM106B immunoreacitivity. At present, the mechanisms responsible for reduced expression of TMEM106B in AD brains remain unknown. If downregulation of TMEM106B is directly or indirectly involved in neurodegeneration, we could propose the hypothesis that TMEM106B plays a protective role against the neurodegenerative processes in AD. Worthy of note is that by analyzing the promoter region of the *TMEM106B* gene with bioinformatics tools for the Database of Transcriptional Start Site [34] and the Matrix Search for Transcription Factor Binding Sites [35], we identified three potential binding sites for POU class 2 homeobox 1 (POU2F1; OCT1), a transcription factor of the POU transcription factor family (unpublished observations) whose SNP is closely associated with the genetic risk of AD [36].

Conclusions

TMEM106B mRNA and protein expression levels were reduced, while PGRN mRNA levels were elevated, in AD brains compared with the levels in non-AD brains. TMEM106B was expressed in the cytoplasm of cortical neurons, hippocampal neurons, and subpopulations of oligodendrocytes, reactive astrocytes, and microglia. In AD brains, surviving neurons expressed moderate/intense TMEM106B immunoreactivity, while senile plaques, neurofibrillary tangles, and the perivascular neuropil intensely expressed PGRN. These observations suggest an active role of TMEM106B in the pathological processes of AD.

Additional files

Additional file 1: Figure S1. Showing p.T185S genotyping analysis. The rs3173615 SNP composed of p.T185S (C760G) in exon 6 of the human *TMEM106B* gene was studied by direct sequencing of PCR product

amplified from brain cDNA. (a) T185/T185 homozygote, (b) T185/S185 heterozygote, and (c) S185/S185 homozygote.

Additional file 2: Figure S2. Showing elevated expression of TMEM106A and TMEM106C mRNA in AD brains. The TMEM106A and TMEM106C mRNA expression levels were studied by qPCR in human brain tissues derived from a REF, four NC cases, six ALS cases, four PD cases, and seven AD cases. The expression levels were standardized against those of G3PDH. (a) TMEM106A. (b) TMEM106C. (c) Difference in TMEM106A levels between AD and non-AD cases. $*P = 0.0002$ by Student's *t* test. (d) Difference in TMEM106C levels between AD and non-AD cases. $**P = 0.0005$ by Student's *t* test.

Additional file 3: Figure S3. Showing pTDP-43 immunoreactivity in AD and non-AD brains. The expression of phosphorylated TDP-43 (pTDP-43) immunoreactivity was studied in six AD brains and 13 non-AD brains presented in Table 1 by immunohistochemistry using anti-pS409/410 TDP-43 antibody. (a) AD, the hippocampal granule cell layer, neuronal cytoplasmic staining; (b) ALS, the hippocampal granule cell layer, neuronal cytoplasmic staining; (c) AD, the frontal cortex, microglial cytoplasmic staining; (d) ALS, the frontal cortex, neuronal cytoplasmic staining.

Additional file 4: Figure S4. Showing TMEM106B and PGRN immunoreactivities in AD and non-AD brains. The expression of TMEM106 and PGRN immunoreactivities was studied in six AD brains and 13 non-AD brains presented in Table 1 by immunohistochemistry using the A303-439A antibody. (a) TMEM106B, AD, the hippocampal CA1 region, vacuoles of granulovacuolar degeneration (GVD) devoid of staining; (b) TMEM106B, AD, the hippocampal molecular layer, intense astroglial cytoplasmic staining; (c) TMEM106B, AD, the periventricular white matter, intense oligodendroglial cytoplasmic staining; (d) PGRN, AD, the frontal white matter, intense microglial cytoplasmic staining; (e) TMEM106B, PD, the frontal cortex, moderate/intense neuronal cytoplasmic staining; (f) TMEM106B after absorption of the antibody, same region as (e), diminished neuronal cytoplasmic staining.

Additional file 5: Figure S5. Showing overexpression of TMEM106B or PGRN did not alter PGRN or TMEM106B mRNA expression levels in SK-N-SH neuroblastoma cells. SK-N-SH neuroblastoma cells expressing Xpress-tagged recombinant proteins were processed for western blot and qPCR. Immunoblot of (a) Xpress and (b) HSP60, an internal control for protein loading. Lanes represent the protein of (1) untransfected cells and the cells expressing (2) TMEM106B, (3) PGRN, and (4) LacZ tagged with Xpress. mRNA expression levels of (c) TMEM106B and (d) PGRN in SK-N-SH cells exposed to Lipofectamine 2000 alone (CNT) and following expression of TMEM106B, PGRN, and LacZ proteins tagged with Xpress. (Single star indicates P = 0.1204 by one-way ANOVA, while double star indicates P = 0.4726 by one-way ANOVA).

Abbreviations

AD: Alzheimer's disease; ALS: amyotrophic lateral sclerosis; bp: base pair; FTLD: frontotemporal lobar dementia; GFAP: glial fibrillary acidic protein; GRN: granulin; NC: non-neurological causes; NEUN: RNA binding protein, fox-1 homolog (*Caenorhabditis elegans*)-3 (RBFOX3); NFH: neurofilament, heavy polypeptide; PCR: polymerase chain reaction; PD: Parkinson's disease; PGRN: progranulin; qPCR: quantitative reverse transcriptase-polymerase chain reaction; SNP: single nucleotide polymorphism; TDP-43: TAR DNA-binding protein-43; TMEM106B: transmembrane protein 106B.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JS and KA designed the study. JS, YK, NK, and YY carried out qPCR, western blot, immunohistochemistry, and genetic analysis. TI, YS, and KA validated the pathological diagnosis of autopsied brains. JS, TI, YS, and KA cooperatively analyzed immunohistochemical data. JS drafted the manuscript. YK, NK, YY, TI, YS, and KA read the draft, critically revised the entire contents, and approved the final manuscript. All authors read and approved the final manuscript.

Acknowledgements

All autopsied brain samples were obtained from the Research Resource Network, Japan. This work was supported by JSPS KAKENHI (C22500322 and C25430054), by the Intractable Disease Research Center project, by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and by a grant from the Biobank of the National Center for Geriatrics and Gerontology (NCGC 26–20).

Author details

[1]Department of Bioinformatics and Molecular Neuropathology, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588, Japan. [2]Department of Pathology and Laboratory Medicine, Kohnodai Hospital, National Center for Global Health and Medicine, 1-7-1 Kohnodai, Ichikawa, Chiba 272-8516, Japan. [3]Department of Laboratory Medicine, National Center Hospital, National Center of Neurology and Psychiatry, 4-1-1 Ogawahigashi, Kodaira, Tokyo 187-8502, Japan. [4]Department of Psychiatry, National Center Hospital, National Center of Neurology and Psychiatry, 4-1-1 Ogawahigashi, Kodaira, Tokyo 187-8502, Japan.

Received: 3 December 2013 Accepted: 20 March 2014
Published: 31 March 2014

References

1. Van Deerlin VM, Sleiman PM, Martinez-Lage M, Chen-Plotkin A, Wang LS, Graff-Radford NR, Dickson DW, Rademakers R, Boeve BF, Grossman M, Arnold SE, Mann DM, Pickering-Brown SM, Seelaar H, Heutink P, van Swieten JC, Murrell JR, Ghetti B, Spina S, Grafman J, Hodges J, Spillantini MG, Gilman S, Lieberman AP, Kaye JA, Woltjer RL, Bigio EH, Mesulam M, Al-Sarraj S, Troakes C, *et al*: **Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions.** *Nat Genet* 2010, **42**:234–239.
2. van der Zee J, Van Langenhove T, Kleinberger G, Sleegers K, Engelborghs S, Vandenberghe R, Santens P, Van den Broeck M, Joris G, Brys J, Mattheijssens M, Peeters K, Cras P, De Deyn PP, Cruts M, Van Broeckhoven C: **TMEM106B is associated with frontotemporal lobar degeneration in a clinically diagnosed patient cohort.** *Brain* 2011, **134**:808–815.
3. Nicholson AM, Finch NA, Wojtas A, Baker MC, Perkerson RB, Castanedes-Casey M, Rousseau L, Benussi L, Binetti G, Ghidoni R, Hsiung GY, Mackenzie IR, Finger E, Boeve BF, Ertekin-Taner N, Graff-Radford NR, Dickson DW, Rademakers R: **TMEM106B p.T185S regulates TMEM106B protein levels: implications for frontotemporal dementia.** *J Neurochem* 2013, **126**:781–791.
4. Baker M, Mackenzie IR, Pickering-Brown SM, Gass J, Rademakers R, Lindholm C, Snowden J, Adamson J, Sadovnick AD, Rollinson S, Cannon A, Dwosh E, Neary D, Melquist S, Richardson A, Dickson D, Berger Z, Eriksen J, Robinson T, Zehr C, Dickey CA, Crook R, McGowan E, Mann D, Boeve B, Feldman H, Hutton M: **Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17.** *Nature* 2006, **442**:916–919.
5. Cruts M, Gijselinck I, van der Zee J, Engelborghs S, Wils H, Pirici D, Rademakers R, Vandenberghe R, Dermaut B, Martin JJ, van Duijn C, Peeters K, Sciot R, Santens P, De Pooter T, Mattheijssens M, Van den Broeck M, Cuijt I, Vennekens K, De Deyn PP, Kumar-Singh S, Van Broeckhoven C: **Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21.** *Nature* 2006, **442**:920–924.
6. Chen-Plotkin AS, Unger TL, Gallagher MD, Bill E, Kwong LK, Volpicelli-Daley L, Busch JI, Akle S, Grossman M, Van Deerlin V, Trojanowski JQ, Lee VM: **TMEM106B, the risk gene for frontotemporal dementia, is regulated by the microRNA-132/212 cluster and affects progranulin pathways.** *J Neurosci* 2012, **32**:11213–11227.
7. Lang CM, Fellerer K, Schwenk BM, Kuhn PH, Kremmer E, Edbauer D, Capell A, Haass C: **Membrane orientation and subcellular localization of transmembrane protein 106B (TMEM106B), a major risk factor for frontotemporal lobar degeneration.** *J Biol Chem* 2012, **287**:19355–19365.
8. Brady OA, Zheng Y, Murphy K, Huang M, Hu F: **The frontotemporal lobar degeneration risk factor, TMEM106B, regulates lysosomal morphology and function.** *Hum Mol Genet* 2013, **22**:685–695.
9. Schwenk BM, Lang CM, Hogl S, Tahirovic S, Orozco D, Rentzsch K, Lichtenthaler SF, Hoogenraad CC, Capell A, Haass C, Edbauer D: **The FTLD risk factor TMEM106B and MAP6 control dendritic trafficking of lysosomes.** *EMBO J* 2014, **33**:450-467.
10. Ahmed Z, Mackenzie IR, Hutton ML, Dickson DW: **Progranulin in frontotemporal lobar degeneration and neuroinflammation.** *J Neuroinflammation* 2007, **4**:7.

11. Cenik B, Sephton CF, Kutluk Cenik B, Herz J, Yu G: **Progranulin: a proteolytically processed protein at the crossroads of inflammation and neurodegeneration.** *J Biol Chem* 2012, **287**:32298–32306.
12. Hu F, Padukkavidana T, Vægter CB, Brady OA, Zheng Y, Mackenzie IR, Feldman HH, Nykjaer A, Strittmatter SM: **Sortilin-mediated endocytosis determines levels of the frontotemporal dementia protein, progranulin.** *Neuron* 2010, **68**:654–667.
13. Cruchaga C, Graff C, Chiang HH, Wang J, Hinrichs AL, Spiegel N, Bertelsen S, Mayo K, Norton JB, Morris JC, Goate A: **Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels.** *Arch Neurol* 2011, **68**:581–586.
14. Finch N, Carrasquillo MM, Baker M, Rutherford NJ, Coppola G, Dejesus-Hernandez M, Crook R, Hunter T, Ghidoni R, Benussi L, Crook J, Finger E, Hantanpaa KJ, Karydas AM, Sengdy P, Gonzalez J, Seeley WW, Johnson N, Beach TG, Mesulam M, Forloni G, Kertesz A, Knopman DS, Uitti R, White CL 3rd, Caselli R, Lippa C, Bigio EH, Wszolek ZK, Binetti G, *et al*: **TMEM106B regulates progranulin levels and the penetrance of FTLD in GRN mutation carriers.** *Neurology* 2011, **76**:467–474.
15. van Blitterswijk M, Mullen B, Nicholson AM, Bieniek KF, Heckman MG, Baker MC, Dejesus-Hernandez M, Finch NA, Brown PH, Murray ME, Hsiung GY, Stewart H, Karydas AM, Finger E, Kertesz A, Bigio EH, Weintraub S, Mesulam M, Hatanpaa KJ, White Iii CL, Strong MJ, Beach TG, Wszolek ZK, Lippa C, Caselli R, Petrucelli L, Josephs KA, Parisi JE, Knopman DS, Petersen RC, *et al*: **TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia.** *Acta Neuropathol* 2014, **127**:397-406.
16. Gallagher MD, Suh E, Grossman M, Elman L, McCluskey L, Van Swieten JC, Al-Sarraj S, Neumann M, Gelpi E, Ghetti B, Rohrer JD, Halliday G, Van Broeckhoven C, Seilhean D, Shaw PJ, Frosch MP, Alafuzoff I, Antonell A, Bogdanovic N, Brooks W, Cairns NJ, Cooper-Knock J, Cotman C, Cras P, Cruts M, De Deyn PP, Decarli C, Dobson-Stone C, Engelborghs S, Fox N, *et al*: **TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions.** *Acta Neuropathol* 2014, **127**:407-418.
17. Vass R, Ashbridge E, Geser F, Hu WT, Grossman M, Clay-Falcone D, Elman L, McCluskey L, Lee VM, Van Deerlin VM, Trojanowski JQ, Chen-Plotkin AS: **Risk genotypes at TMEM106B are associated with cognitive impairment in amyotrophic lateral sclerosis.** *Acta Neuropathol* 2011, **121**:373–380.
18. Rutherford NJ, Carrasquillo MM, Li M, Bisceglio G, Menke J, Josephs KA, Parisi JE, Petersen RC, Graff-Radford NR, Younkin SG, Dickson DW, Rademakers R: **TMEM106B risk variant is implicated in the pathologic presentation of Alzheimer disease.** *Neurology* 2012, **79**:717–718.
19. Lu RC, Wang H, Tan MS, Yu JT, Tan L: **TMEM106B and APOE polymorphisms interact to confer risk for late-onset Alzheimer's disease in Han Chinese.** *J Neural Transm* 2014, **121**:283-287.
20. Satoh J, Tabunoki H, Ishida T, Saito Y, Arima K: **Dystrophic neurites express C9orf72 in Alzheimer's disease brains.** *Alzheimers Res Ther* 2012, **4**:33.
21. Mirra SS, Gearing M, McKeel DW Jr, Crain BJ, Hughes JP, van Belle G, Heyman A: **The Consortium to Establish a Registry for Alzheimer's Disease (CERAD), Part II. Standardization of the neuropathologic assessment of Alzheimer's disease.** *Neurology* 1991, **41**:479–486.
22. Braak H, Alafuzoff I, Arzberger T, Kretzschmar H, Del Tredici K: **Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry.** *Acta Neuropathol* 2006, **112**:389–404.
23. Satoh J, Tabunoki H, Ishida T, Saito Y, Arima K: **Accumulation of a repulsive axonal guidance molecule RGMa in amyloid plaques: a possible hallmark of regenerative failure in Alzheimer's disease brains.** *Neuropathol Appl Neurobiol* 2013, **39**:109–120.
24. **dbSNP Short Genetic Variations.** [www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=3173615]
25. van der Zee J, Van Broeckhoven C: **TMEM106B a novel risk factor for frontotemporal lobar degeneration.** *J Mol Neurosci* 2011, **45**:516–521.
26. Busch JI, Martinez-Lage M, Ashbridge E, Grossman M, Van Deerlin VM, Hu F, Lee VM, Trojanowski JQ, Chen-Plotkin AS: **Expression of TMEM106B, the frontotemporal lobar degeneration-associated protein, in normal and diseased human brain.** *Acta Neuropathol Commun* 2013, **1**:36.
27. Pereson S, Wils H, Kleinberger G, McGowan E, Vandewoestyne M, Van Broeck B, Joris G, Cuijt I, Deforce D, Hutton M, Van Broeckhoven C, Kumar-Singh S: **Progranulin expression correlates with dense-core amyloid plaque burden in Alzheimer disease mouse models.** *J Pathol* 2009, **219**:173–181.
28. Suh HS, Choi N, Tarassishin L, Lee SC: **Regulation of progranulin expression in human microglia and proteolysis of progranulin by matrix metalloproteinase-12 (MMP-12).** *PLoS One* 2012, **7**:e35115.
29. Vercellino M, Grifoni S, Romagnolo A, Masera S, Mattioda A, Trebini C, Chiavazza C, Caligiana L, Capello E, Mancardi GL, Giobbe D, Mutani R, Giordana MT, Cavalla P: **Progranulin expression in brain tissue and cerebrospinal fluid levels in multiple sclerosis.** *Mult Scler* 2011, **17**:1194–1201.
30. Pickford F, Marcus J, Camargo LM, Xiao Q, Graham D, Mo JR, Burkhardt M, Kulkarni V, Crispino J, Hering H, Hutton M: **Progranulin is a chemoattractant for microglia and stimulates their endocytic activity.** *Am J Pathol* 2011, **178**:284–295.
31. Baralle M, Buratti E, Baralle FE: **The role of TDP-43 in the pathogenesis of ALS and FTLD. The role of TDP-43 in the pathogenesis of ALS and FTLD.** *Biochem Soc Trans* 2013, **41**:1536–1540.
32. Sephton CF, Cenik C, Kucukural A, Dammer EB, Cenik B, Han Y, Dewey CM, Roth FP, Herz J, Peng J, Moore MJ, Yu G: **Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes.** *J Biol Chem* 2011, **286**:1204–1215.
33. Funk KE, Mrak RE, Kuret J: **Granulovacuolar degeneration (GVD) bodies of Alzheimer's disease (AD) resemble late-stage autophagic organelles.** *Neuropathol Appl Neurobiol* 2011, **37**:295–306.
34. **Database of Transcriptional Start Site.** [http://dbtss.hgc.jp]
35. **Matrix Search for Transcription Factor Binding Sites.** [www.gene-regulation.com]
36. Taguchi K, Yamagata HD, Zhong W, Kamino K, Akatsu H, Hata R, Yamamoto T, Kosaka K, Takeda M, Kondo I, Miki T: **Identification of hippocampus-related candidate genes for Alzheimer's disease.** *Ann Neurol* 2005, **57**:585–588.

doi:10.1186/alzrt247
Cite this article as: Satoh *et al.*: TMEM106B expression is reduced in Alzheimer's disease brains. *Alzheimer's Research & Therapy* 2014 6:17.

Libertas Academica
FREEDOM TO RESEARCH

Open Access: Full open access to this and thousands of other papers at http://www.la-press.com.

Gene Regulation and Systems Biology

A Comprehensive Profile of ChIP-Seq-Based PU.1/Spi1 Target Genes in Microglia

Jun-ichi Satoh, Naohiro Asahina, Shouta Kitano and Yoshihiro Kino

Department of Bioinformatics and Molecular Neuropathology, Meiji Pharmaceutical University, Kiyose, Tokyo, Japan.

ABSTRACT: Microglia are resident mononuclear phagocytes that play a principal role in the maintenance of normal tissue homeostasis in the central nervous system (CNS). Microglia, rapidly activated in response to proinflammatory stimuli, are accumulated in brain lesions of neurodegenerative diseases, such as Alzheimer's disease and Parkinson's disease. The E26 transformation-specific (ETS) family transcription factor PU.1/Spi1 acts as a master regulator of myeloid and lymphoid development. PU.1-deficient mice show a complete loss of microglia, indicating that PU.1 plays a pivotal role in microgliogenesis. However, the comprehensive profile of PU.1/Spi1 target genes in microglia remains unknown. By analyzing a chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq) dataset numbered SRP036026 with the Strand NGS program, we identified 5,264 Spi1 target protein-coding genes in BV2 mouse microglial cells. They included *Spi1, Irf8, Runx1, Csf1r, Csf1, Il34, Aif1* (*Iba1*), *Cx3cr1, Trem2*, and *Tyrobp*. By motif analysis, we found that the PU-box consensus sequences were accumulated in the genomic regions surrounding ChIP-Seq peaks. By using pathway analysis tools of bioinformatics, we found that ChIP-Seq-based Spi1 target genes show a significant relationship with diverse pathways essential for normal function of monocytes/macrophages, such as endocytosis, Fc receptor-mediated phagocytosis, and lysosomal degradation. These results suggest that PU.1/Spi1 plays a crucial role in regulation of the genes relevant to specialized functions of microglia. Therefore, aberrant regulation of PU.1 target genes might contribute to the development of neurodegenerative diseases with accumulation of activated microglia.

KEYWORDS: ChIP-Seq, GenomeJack, KeyMolnet, microglia, microgliopathy, Nasu–Hakola disease, PU.1, Spi1, Strand NGS

CITATION: Satoh et al. A Comprehensive Profile of ChIP-Seq-Based PU.1/Spi1 Target Genes in Microglia. *Gene Regulation and Systems Biology* 2014:8 127–139 doi: 10.4137/GRSB.S19711.

RECEIVED: September 1, 2014. **RESUBMITTED:** November 2, 2014. **ACCEPTED FOR PUBLICATION:** November 10, 2014.

ACADEMIC EDITOR: James Willey, Editor in Chief

TYPE: Original Research

FUNDING: This work was supported by the JSPS KAKENHI (C25430054), the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and the grant from the National Center for Geriatrics and Gerontology (NCGC26–20). The authors confirmed that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: The authors declare no competing interests.

CORRESPONDENCE: satoj@my-pharm.ac.jp

Paper subject to independent expert blind peer review by minimum of two reviewers. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Introduction

Microglia are resident mononuclear phagocytes that play a principal role in the maintenance of normal tissue homeostasis in the central nervous system (CNS).[1] They are derived from primitive c-kit$^+$ erythromyeloid precursors (EMPs) in the yolk sac emerging as early as on day 8 post-conception during embryogenesis.[2,3] EMPs develop into CD45$^+$ c-kit^{lo} CX3CR1$^-$ immature A1 cells that subsequently differentiate into CD45$^+$ c-kit$^-$ CX3CR1$^+$ A2 cells. Proliferating A2 cells enter into the developing CNS and are incorporated into the brain parenchyma as resident microglia. Microglia have a capacity to constantly scavenge invading pathogens, dying cells, and unwanted synapses by sensing them with a panel of pattern recognition receptors (PRRs).[1] Microglia show a ramified morphology under physiological conditions. When exposed to infectious and traumatic stimuli, they rapidly adopt an amoeboid morphology, followed by secretion of various cytokines, chemokines, and reactive oxygen and nitrogen species. Depending on their microenvironment, microglia are activated to acquire two distinct priming states. Stimulation with lipopolysaccharide (LPS) or interferon-gamma (IFNγ) induces the "classically" activated (M1; proinflammatory) state relevant to defense against bacterial and viral infection, whereas exposure to interleukin (IL)-4 or IL-13 promotes the conversion to the "alternatively" activated (M2; anti-inflammatory) state involved in tissue repair and remodeling.[1]

Microglia play a central role in the pathophysiology of human neurodegenerative diseases that are characterized by chronic inflammation associated with accumulation of activated microglia in affected areas, such as Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD).[4,5] In AD, amyloid-beta (Aβ) activates microglia by signaling through Toll-like receptors (TLRs) and NOD-like receptors (NLRs), leading to production of proinflammatory mediators potentially toxic to neurons.[6,7] In PD, alpha-synuclein (α-Syn), which serves as a danger-associated molecular pattern, directly activates microglia.[8] In HD, mutant huntingtin promotes transcriptional activation of numerous proinflammatory genes in microglia.[9] However, at present, the precise mechanism underlying gene regulation relevant to microglial activation in human neurodegenerative diseases remains largely unknown.

The E26 transformation-specific (ETS) family transcription factor PU.1, also named as Spi1 or Sfpi1 in mouse, acts as a master regulator of myeloid and lymphoid development, expressed chiefly in monocytes/macrophages, neutrophils, mast cells, B cells, and early erythroblasts.[10] PU.1 comprises an N-terminal transactivation domain, a C-terminal DNA-binding domain, and an intervening PEST domain for protein–protein interactions. It activates expression of hundreds of downstream genes by binding to a purine-rich DNA sequence named the PU-box located on the targets. The expression levels of PU.1 target genes are highly variable in different cell types, owing to the difference in cellular concentration of PU.1, chromatin accessibility, motif-binding affinity, and cooperation with neighboring transcription factors.[11] Importantly, PU.1-deficient mice show a complete loss of microglia, along with a lack of mature macrophages, monocytes, neutrophils, and B cells, indicating that PU.1 regulates key genes involved in differentiation and maturation of not only hematopoietic cells but also brain microglia.[12,13] However, at present, the comprehensive profile of PU.1 target genes involved in microgliogenesis remains uncharacterized. In the adult human microglia, MCSF (CSF1) treatment elevates the expression levels of PU.1 and stimulates phagocytosis of Aβ, while knockdown of PU.1 reduces their viability and phagocytic capability.[14,15] Interferon regulatory factor 8 (Irf8) serves as an essential regulator of development of A2 microglial progenitor cells.[3] Irf8-deficient microglia show fewer elaborated processes with decreased expression of Iba1 and reduced proliferative and phagocytic activities.[16] Runt-related transcription factor 1 (Runx1), whose expression levels are elevated in amoeboid microglia, promotes reverse transition from amoeboid to ramified microglia.[17]

Recently, the rapid progress in the next-generation sequencing (NGS) technology has revolutionized the field of genome research. Chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq) serves as one of NGS applications for genome-wide profiling of DNA-binding proteins, histone modifications, and nucleosomes.[18] ChIP-Seq, with advantages of higher resolution, less noise, and greater coverage of the genome, compared with microarray-based ChIP-Chip, provides an innovative tool for studying gene regulatory networks on the whole genome scale. Furthermore, recent advances in systems biology help us to investigate the cell-wide map of the complex molecular interactions by using the literature-based knowledgebase of molecular pathways.[19] Therefore, the integration of high dimensional ChIP-Seq NGS data with underlying molecular networks represents a rational approach to characterize the genome-wide network-based molecular mechanisms of gene regulation. To clarify the biological role of PU.1 in regulation of microglial functions, we attempted to characterize the comprehensive set of ChIP-Seq-based PU.1/Spi1 target genes in microglial cells by analyzing a dataset retrieved from public database.

Methods

ChIP-Seq dataset of microglial cells. A ChIP-Seq dataset of microglial cells was retrieved from DDBJ Sequence Read Archive (DRA) under the accession number SRP036026. The researchers in Dr. Christopher K. Glass's Laboratory, University of California, San Diego, performed the original experiment to study the role of reactive microglia in HD.[9] The raw data are open to public from March 2, 2014. Currently, no alternative datasets are publicly available for PU.1 ChIP-Seq of microglia. They cloned the N-terminus of wild-type (15Q) or mutant (128Q) human huntingtin in the pCDH-CMV-MCS-EF1-Puro vector (System Bioscience). Either the cloned vector or the empty vector was expressed in BV2 mouse microglial cells,[20] by using the Lentiviral expression system (System Bioscience). Then, they were processed for ChIP-Seq analysis. We studied ChIP-Seq data derived from the cells transduced with the empty vector (no exogenous huntingtin). Following fixation with formaldehyde, sonicated nuclear lysates were immunoprecipitated with a rabbit polyclonal anti-PU.1 (Spi1) antibody (sc-352; Santa Cruz Biotechnology) (SRX451619) or a rabbit polyclonal anti-CCAAT-enhancer-binding protein alpha (C/EBPα, Cebpa) antibody (sc-61; Santa Cruz Biotechnology) (SRX451622). NGS libraries constructed from adapter-ligated ChIP DNA fragments were processed for deep sequencing on Genome Analyzer IIx (Illumina).

First, we evaluated the quality of NGS short reads by searching them on the FastQC program (www.bioinformatics.babraham.ac.uk/projects/fastqc). Then, we removed the reads of insufficient quality by filtering them out with the FASTX-toolkit (hannonlab.cshl.edu/fastx_toolkit). After cleaning the data, we mapped them on the mouse genome reference sequence version mm9 by a mapping tool named COBWeb of the Strand NGS2.0 program, formerly named Avadis NGS (Strand Genomics), or by the Bowtie2 version 2.1.0 program (bowtie-bio.sourceforge.net/bowtie2/index.shtml). Then, we identified the peaks of binding sites with fold enrichment (FE) ≥5 by using the Model-based Analysis of ChIP-Seq (MACS) program or the Probabilistic Inference for ChIP-Seq (PICS)

program.[21,22] We determined the genes corresponding to the peaks by a neighboring gene analysis tool of Strand NGS in the setting within a distance of 5,000 bp from peaks to genes. We characterized the genomic location of binding peaks by a peak-finding tool of Strand NGS that classifies the locations into the upstream region, 5' untranslated region (5'UTR), exon, intron, and 3'UTR. We also imported the processed data into a genome viewer named GenomeJack v1.4 (Mitsubishi Space Software). We identified the consensus motif sequences in the genomic regions surrounding the peaks by using the GADEM program.[23]

Molecular network analysis. To identify molecular networks biologically relevant to ChIP-Seq-based Spi1 target genes, we imported the corresponding Entrez Gene IDs into the Functional Annotation tool of Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (david.abcc.ncifcrf.gov).[24] DAVID identifies relevant pathways constructed by Kyoto Encyclopedia of Genes and Genomes (KEGG), composed of the genes enriched in the given set, followed by statistical evaluation by a modified Fisher's exact test corrected by Bonferroni multiple comparison test. KEGG (www.kegg.jp) is a publicly accessible knowledgebase that covers a wide range of pathway maps on metabolic, genetic, environmental, and cellular processes, and human diseases, currently composed of 332,680 pathways generated from 466 reference pathways.[25]

We also imported Entrez Gene IDs into Ingenuity Pathways Analysis (IPA) (Ingenuity Systems; www.ingenuity.com). IPA is a commercial knowledgebase that contains approximately 3,000,000 biological and chemical interactions and functional annotations with definite scientific evidence. Upon uploading the list of Gene IDs, the network-generation algorithm identifies focused genes integrated in global molecular pathways and networks. IPA calculates the score *P*-value that reflects the statistical significance of association between the genes and the pathways or networks by the Fisher's exact test.

KeyMolnet (KM Data; www.km-data.jp), a different commercial knowledgebase, contains manually curated content on 164,000 relationships among human genes and proteins, small molecules, diseases, pathways, and drugs.[26] They include the core content collected from selected review articles with the highest reliability. Upon importing the list of Gene IDs, KeyMolnet automatically provides corresponding molecules as nodes on the network. The neighboring network-search algorithm selected one or more molecules as starting points to generate the network of all kinds of molecular interactions around starting molecules, including direct activation/inactivation, transcriptional activation/repression, and the complex formation within one path from starting points. The generated network was compared side by side with 501 human canonical pathways of the KeyMolnet library. The algorithm counting the number of overlapping molecular relations between the extracted network and the canonical pathway makes it possible to identify the canonical pathway showing the most significant contribution to the extracted network.

Results

Identification of 5,264 ChIP-Seq-based Spi1 target genes in mouse microglia. First, we evaluated the quality of ChIP-Seq NGS data examined in the present study. After cleaning, the quality scores mostly exceeded 30 across the bases on FastQC, indicating an acceptable quality for downstream analysis (Supplementary Fig. 1, panels a, b). After mapping them on mm9 by COBWeb, we identified 56,278 Spi1-ChIP peaks detected by MACS and 15,141 Spi1-ChIP peaks detected by PICS. From these, we selected peaks located within a distance of 5,000 bp from protein coding genes, and then extracted 5,264 genes overlapping between data derived from two distinct peak-finding algorithms MACS and PICS termed as the most reliable Spi1 targets (Supplementary Table 1). The peaks were accumulated in the upstream (18.7%) and intronic (69.5%) regions. The motif analysis by GADEM revealed an existence of the PU-box consensus sequences defined as 5'-GAGGAA-3' located within the genomic regions surrounding ChIP-Seq peaks (Fig. 1).

We identified both *Spi1* (FE = 19.3) and *Irf8* (FE = 27.8) in the list of Spi1 target genes (Supplementary Table 1; Figs. 2 and 3). Both of them are known to serve as a crucial regulator of differentiation of microglia from EMPs during early embryogenesis.[3] We also found *Runx1* (FE = 26.7), a transcription factor acting to constitute a negative feedback loop of PU.1,[27] along with *Csf1r* (FE = 10.7), *Csf1* (FE = 17.8), and *Il34* (FE = 31.5), acting as a key growth factor for differentiation of microglia,[28] as Spi1 targets (Supplementary Table 1). Furthermore, we identified known cell type-specific markers for microglia, such as *Aif1* (Iba1, FE = 32.9), *Cx3cr1*

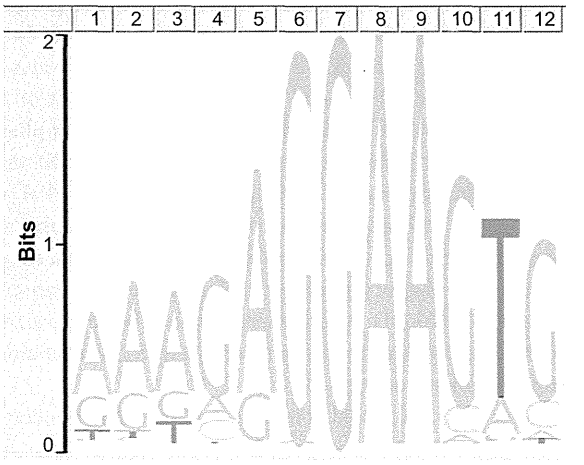

Figure 1. Spi1-binding consensus sequence motif. The consensus motif sequences surrounding Spi1 ChIP-Seq peaks were identified by the GADEM program. The PU-box consensus sequences defined as 5'-GAGGAA-3' were located on 80.3% of the peaks detected by MACS.

Figure 2. Genomic location of Spi1 ChIP-Seq peak on the *Spi1* gene. The genomic location of Spi1 ChIP-Seq peaks was determined by importing the processed data into GenomeJack. An example of transcription factor PU.1 (Spi1; Entrez Gene ID 20375) is shown, where a MACS peak numbered 66326 in the Spi1.bam Coverage track is located in the promoter region of the *Spi1* gene (panel a) with a Spi1-binding consensus sequence motif highlighted by orange square (panel b).

(FE = 17.8), *Cd68* (FE = 20.3), *Trem2* (FE = 12.8), and *Tyrobp* (Dap12) (FE = 14.6) in the list of Spi1 target genes (Supplementary Table 1; Supplementary Figs. 2 and 3). Importantly, loss of function of either TREM2 or DAP12, components of a receptor/adapter complex on human microglia, plays a causative role in Nasu–Hakola disease (NHD).[29] Furthermore, we found *Syk* (FE = 17.8), a downstream signal transducer of the Trem2/Dap12 pathway, as a Spi1 target gene.

Next, we studied ChIP-Seq-based Cebpa target genes in BV2 microglial cells. We identified 12,685 Cebpa-ChIP peaks detected by MACS and 10,311 Cebpa-ChIP peaks detected by PICS. From these, we selected peaks located within a distance of 5,000 bp from protein coding genes, and then extracted 3,106 genes overlapping between data derived from MACS and PICS termed as the most reliable Cebpa targets. We found that 1,844 genes are shared between Spi1 targets and Cebpa targets, suggesting the possibility that Cebpa coregulates a substantial proportion (35%) of Spi1 target genes in microglial cells (Supplementary Table 1, underline).

A recent study by direct RNA sequencing of flow cytometry-sorted mouse brain microglia has characterized a set of 100 transcripts exclusively expressed in microglia.[30] The study designated them as "the microglial sensome" (Supplementary Table 2). Importantly, we found that 63 out of 100 microglial sensome genes correspond to ChIP-Seq-based Spi1 target genes, indicating that Spi1 plays a pivotal role in regulation of the genes relevant to specialized functions of microglia (Table 1).

Molecular networks of ChIP-Seq-based Spi1 target genes in microglia. Next, we studied molecular networks of the set of 5,264 ChIP-Seq-based Spi1 target genes by using three distinct pathway analysis tools of bioinformatics. By using DAVID, we identified functionally associated gene ontology (GO) terms. The most significant GO terms included "phosphate metabolic process" (GO:0006796; P = 2.21E-16 corrected by Bonferroni multiple comparison test) for biological process, "plasma membrane" (GO:0005886; P = 1.26E-15) for cellular component, and "GTPase regulator activity" (GO:0030695; P = 1.27E-22) for molecular function.

By using KEGG, we found that the set of 5,264 Spi1 targets showed a significant relationship with the pathways defined as "Lysosome" (mmu04142; P = 5.08E-08 corrected by Bonferroni multiple comparison test), "Focal adhesion" (mmu04510; P = 1.27E-07), "Endocytosis" (mmu04144; P = 4.54E-07), "Fcγ receptor-mediated phagocytosis" (mmu04666; P = 5.03E-07) (Fig. 4), and "MAPK signaling pathway" (mmu04010; P = 6.23E-06) (Table 2). Furthermore, they also exhibited significant association with the pathways

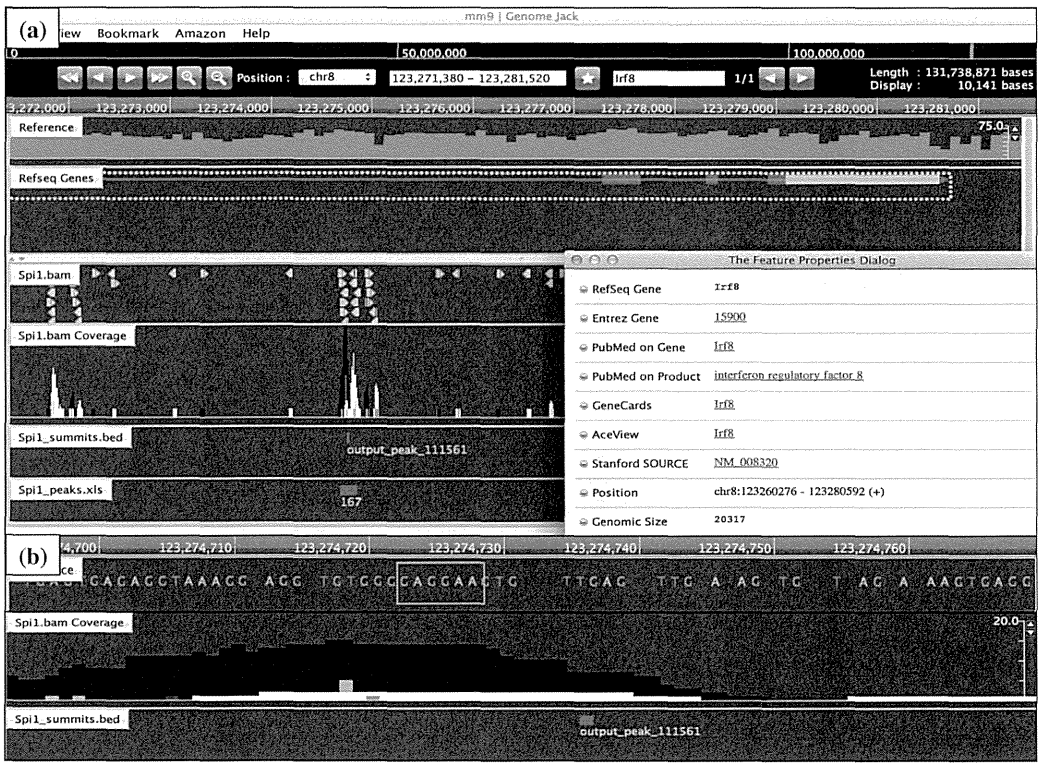

Figure 3. Genomic location of Spi1 ChIP-Seq peak on the *Irf8* gene. The genomic location of Spi1 ChIP-Seq peaks was determined by importing the processed data into GenomeJack. An example of interferon regulatory factor 8 (Irf8; Entrez Gene ID 15900) is shown, where a MACS peak numbered 111561 in the Spi1.bam Coverage track is located in the intronic region of the *Irf8* gene (panel a) with a Spi1-binding consensus sequence motif highlighted by orange square (panel b).

defined as "Pathways in cancer" (mmu05200; P=1.39E-04), "B cell receptor signaling pathway" (mmu04662; P = 2.18E-04), "Apoptosis" (mmu04210; P=4.23E-04), "Leukocyte transendothelial migration" (mmu04670; P=6.57E-04), "Chemokine signaling pathway" (mmu04062; P=8.52E-04), and "Chronic myeloid leukemia" (mmu05220; P = 1.16E-03) (Table 2). Importantly, the top-ranked "Lysosome" pathway included the set of 10 cathepsin genes, such as *Ctsa, Ctsb, Ctsc, Ctsd, Ctse, Ctsf, Ctsk, Ctsl, Ctss,* and *Ctsz*, essential for degradation of lysosomal proteins in microglia (Table 2).

Next, we studied molecular networks of 5,264 Spi1 target genes by using the core analysis tool of IPA. They showed a significant relationship with canonical pathways defined as "Fcγ receptor-mediated phagocytosis in macrophages and monocytes" (P=2.11E-15), "Molecular mechanisms of cancer" (P=6.92E-15), "B cell receptor signaling" (P=2.77E-14), "Role of NFAT in regulation of the immune response" (P=1.17E-12), and "PI3K signaling in B lymphocytes" (P = 2.15E-12). The results of KEGG and IPA combined together indicated that Spi1 regulates expression of not only the genes crucial for normal function of monocytes/macrophages and B cells but also those involved in oncogenesis, particularly in leukemogenesis. IPA also identified functional networks relevant to Spi1 target genes (Supplementary Table 3). The most significant network was defined as "Cell Morphology, Cellular Function and Maintenance, Cell Death and Survival" (P = 1.00E-53), where key components of autophagosomes, such as ATG3, ATG5, ATG7, and ATG10, are clustered (Fig. 5). The second rank network represented "RNA Post-Transcriptional Modification, Cellular Assembly and Organization, Infectious Disease" (P = 1.00E-53).

Finally, we studied molecular networks of 5,264 Spi1 target genes by using KeyMolnet. The neighboring network-search algorithm extracted the highly complex network composed of 5,788 molecules and 13,719 molecular relations (Supplementary Fig. 4). It showed the most significant relationship with "Transcriptional regulation by RB/E2F" (P = 7.27E-193). We identified *Rb1* (FE = 17.7), *Rbl1* (FE = 25.8), *Rbl2* (FE = 37.1), and *E2f1* (FE = 18.6) as a group of Spi1 target genes (Supplementary Table 1).

Discussion

Mice lacking PU.1/Spi1 are devoid of microglia, indicating that PU.1 acts as an indispensable transcription factor for development and differentiation of microglia.[13,31,32] By analyzing a ChIP-Seq dataset, we identified 5,264 Spi1 target protein-coding genes in BV2 mouse microglial cells. BV2 cells are derived from immortalized microglia of newborn

Table 1. The set of 63 microglial sensome genes corresponding to ChIP-Seq-based Spi1 target genes.

CHROMOSOME	START OF PEAKS	END OF PEAKS	FOLD ENRICHMENT	ENTREZ GENE ID	GENE SYMBOL	GENE NAME
chr7	132757191	132757492	38.659794	60504	Il21r	interleukin 21 receptor
chr3	89709374	89709685	37.668518	16194	Il6ra	interleukin 6 receptor, alpha
chr6	122902489	122902814	37.181004	73149	Clec4a3	C-type lectin domain family 4, member a3
chr11	60955751	60956142	35.720745	57916	Tnfrsf13b	tumor necrosis factor receptor superfamily, member 13b
chrX	13213539	13213888	33.61721	23890	Gpr34	G protein-coupled receptor 34
chr4	66504343	66504608	33.29038	21898	Tlr4	toll-like receptor 4
chr7	108112592	108113189	33.136967	233571	P2ry6	pyrimidinergic receptor P2Y, G-protein coupled, 6
chr18	35879188	35879366	32.216496	68545	Ecscr	endothelial cell-specific chemotaxis regulator
chr4	149501529	149501774	30.927835	56485	Slc2a5	solute carrier family 2 (facilitated glucose transporter), member 5
chr3	30763607	30763924	30.746971	71862	Gpr160	G protein-coupled receptor 160
chr7	16829398	16829780	29.842648	319430	C5ar2	complement component 5a receptor 2
chr11	120818969	120819371	29.78236	80879	Slc16a3	solute carrier family 16 (monocarboxylic acid transporters), member 3
chr1	121972010	121972524	28.747026	170706	Tmem37	transmembrane protein 37
chr9	110948641	110949068	27.061855	17002	Ltf	lactotransferrin
chr9	116075399	116075615	26.661926	21813	Tgfbr2	transforming growth factor, beta receptor II
chr18	60966814	60967338	26.09944	16149	Cd74	CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen-associated)
chr5	114094521	114094878	26.044493	14747	Cmklr1	chemokine-like receptor 1
chr7	135258809	135259228	25.773195	16409	Itgam	integrin alpha M
chr13	37488117	37488288	25.773195	17084	Ly86	lymphocyte antigen 86
chr3	106592873	106593332	25.472342	12508	Cd53	CD53 antigen
chr16	38789019	38789447	25.28691	22268	Upk1b	uroplakin 1B
chr11	46280548	46280823	24.811836	171285	Havcr2	hepatitis A virus cellular receptor 2
chr6	40535749	40536059	24.257126	23845	Clec5a	C-type lectin domain family 5, member a
chr11	78799024	78799325	23.375689	16859	Lgals9	lectin, galactose binding, soluble 9
chr13	103496918	103497108	22.680412	17079	Cd180	CD180 antigen
chr3	105735367	105735926	22.315817	433638	I830077J02Rik	RIKEN cDNA I830077J02 gene
chr5	65352028	65352431	21.753523	21899	Tlr6	toll-like receptor 6
chr16	36642687	36643132	21.577559	12524	Cd86	CD86 antigen
chr1	140026784	140027054	21.477663	19264	Ptprc	protein tyrosine phosphatase, receptor type, C
chr7	106859678	106859981	21.273113	101488	Slco2b1	solute carrier organic anion transporter family, member 2b1
chr6	122805926	122806252	20.940722	12267	C3ar1	complement component 3a receptor 1
chr7	16849252	16849697	20.837904	12273	C5ar1	complement component 5a receptor 1

chr4	132139773	132139985	20.618557	19204	Ptafr	platelet-activating factor receptor
chr19	40782595	40783023	20.34726	12495	Entpd1	ectonucleoside triphosphate diphosphohydrolase 1
chr11	69479440	69480081	20.306154	12514	Cd68	CD68 antigen
chr1	172903138	172903598	20.071161	14130	Fcgr2b	Fc receptor, IgG, low affinity IIb
chr10	19317003	19317223	19.63672	15979	Ifngr1	interferon gamma receptor 1
chr3	87180536	87180918	19.435524	229499	Fcrl1	Fc receptor-like 1
chr6	125030472	125030733	19.091257	381810	Lpar5	lysophosphatidic acid receptor 5
chr3	59064135	59064423	19.032515	70839	P2ry12	purinergic receptor P2Y, G-protein coupled 12
chr16	33947248	33947454	18.990776	16419	Itgb5	integrin beta 5
chr1	172989424	172989659	18.93541	14131	Fcgr3	Fc receptor, IgG, low affinity III
chr4	133654539	133654895	18.744143	23833	Cd52	CD52 antigen
chr7	4018702	4019029	18.126204	52855	Lair1	leukocyte-associated Ig-like receptor 1
chrX	103346689	103346910	17.842981	279572	Tlr13	toll-like receptor 13
chr9	119977884	119978196	17.774618	13051	Cx3cr1	chemokine (C-X3-C) receptor 1
chr1	174408269	174408572	17.634293	98365	Slamf9	SLAM family member 9
chr3	100835299	100835620	16.568483	630146	Cd101	CD101 antigen
chr2	93289154	93289478	16.108248	12521	Cd82	CD82 antigen
chr4	47408473	47408654	16.108248	21812	Tgfbr1	transforming growth factor, beta receptor I
chr3	105716542	105717356	15.942183	11542	Adora3	adenosine A3 receptor
chr15	103140748	103141572	14.66406	80910	Gpr84	G protein-coupled receptor 84
chr7	31198532	31198988	14.567458	22177	Tyrobp	TYRO protein tyrosine kinase binding protein
chr5	138278454	138278742	14.318442	231805	Pilra	paired immunoglobin-like type 2 receptor alpha
chr3	96097068	96098277	14.087213	14129	Fcgr1	Fc receptor, IgG, high affinity I
chr9	114661993	114662369	14.058106	67213	Cmtm6	CKLF-like MARVEL transmembrane domain containing 6
chr17	48491514	48491850	12.772557	83433	Trem2	triggering receptor expressed on myeloid cells 2
chr15	78135354	78135537	12.326311	12984	Csf2rb2	colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)
chr18	61280165	61280583	10.698308	12978	Csf1r	colony stimulating factor 1 receptor
chr1	172948811	172949254	10.376222	246256	Fcgr4	Fc receptor, IgG, low affinity IV
chr4	144830530	144830848	9.163803	21938	Tnfrsf1b	tumor necrosis factor receptor superfamily, member 1b
chr10	59834857	59835131	8.8232565	74048	4632428N05Rik	RIKEN cDNA 4632428N05 gene
chr9	20820541	20820889	7.9916887	15894	Icam1	intercellular adhesion molecule 1

Notes: From the ChIP-Seq dataset, we identified 5,264 Spi1-target genes in BV2 mopuse microglia showing fold enrichment (FE) ≥5. Among them, those corresponding to microglial sensome genes (Supplementary Table 2) are listed with the chromosome, the position of the peak (start, end), FE, Entrez Gene ID, Gene Symbol, and Gene Name. The Cebpa-target genes are underlined.

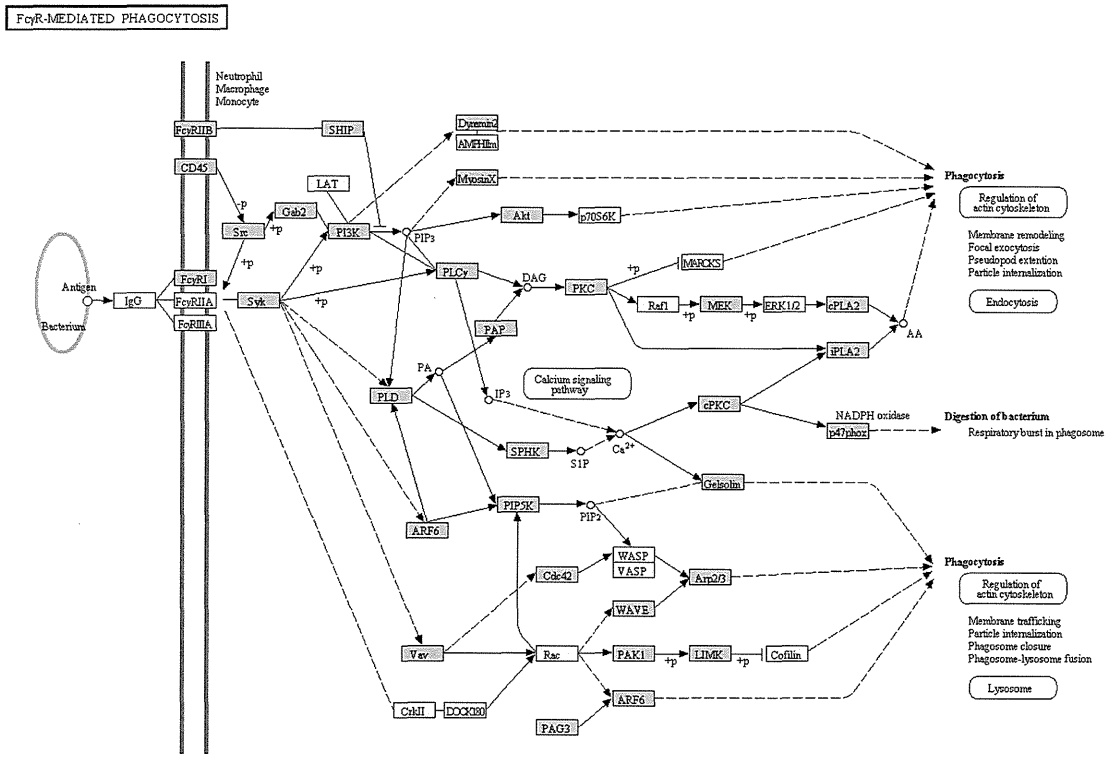

Figure 4. KEGG "Fcγ receptor-mediated phagocytosis" pathway relevant to Spi1 target genes. Entrez Gene IDs of 5,264 ChIP-Seq-based Spi1 target genes were imported into the Functional Annotation tool of DAVID. It extracted the KEGG "Fcγ receptor-mediated phagocytosis" pathway (mmu04666) as the fourth rank significant pathway as listed in Table 2. Spi1 target genes are colored by orange.

mouse origin that express morphological, phenotypical, and functional characteristics of primary microglia.[33] We extracted the genes overlapping between data derived from two distinct peak-finding algorithms MACS and PICS termed as the set of most reliable Spi1 targets. Enrichment of the PU-box consensus sequences within the genomic regions surrounding ChIP-Seq peaks validated the reliability of our analysis. By using pathway analysis tools named KEGG and IPA, we found that ChIP-Seq-based Spi1 target genes show a significant relationship with diverse pathways essential for normal function of monocytes/macrophages, such as endocytosis, Fcγ receptor-mediated phagocytosis, and lysosomal degradation, along with the pathway closely related to leukemogenesis. Relevantly, mice with reduced expression of PU.1 develop acute myeloid leukemia.[34] Approximately two-thirds (63%) of "microglial sensome" genes that reflect the microglia-specific gene signature[30] corresponded to Spi1 targets. These observations suggest that Spi1 plays a pivotal role in regulation of the genes relevant to specialized functions of microglia. KeyMolnet constructed the complex network of Spi1 target genes showing the most significant relationship with transcriptional regulation by RB/E2F. PU.1, capable of interacting physically with the C pocket of phosphorylated retinoblastoma (Rb) protein, blocks erythroid differentiation by repressing GATA-1 in mouse erythroleukemia cells.[35]

The set of 5,264 Spi1 target genes include Spi1 itself, supporting the previous observations that PU.1 activates its own promoter elements via an autoregulatory loop.[36] We found that approximately one-third of Spi1 target genes in microglia are potentially coregulated by Cebpa, a transcription factor essential for the development of monocytic and granulocytic lineage cells.[37] Importantly, Cebpa directly activates PU.1 gene transcription by binding to its promoter and distal enhancer.[38] A previous study showed that the Cebpa–Spi1 pathway plays a central role in regulation of microglial proliferation in a mouse model of prion diseases.[39]

NHD is a rare autosomal recessive disorder characterized by progressive dementia and multifocal bone cysts, caused by genetic mutations of either *DAP12* or *TREM2*.[29] Pathologically, NHD brains exhibit extensive demyelination and gliosis distributed predominantly in the frontal and temporal lobes and the basal ganglia, accompanied by marked accumulation of axonal spheroids and microglia.[40] TREM2 acts as a phagocytic receptor expressed on osteoclasts, dendritic cells, macrophages, and microglia, where it constitutes a signaling complex with an adaptor molecule DAP12, leading to phosphorylation and activation of the downstream kinase Syk. TREM2 expressed on microglia plays a key role in the clearance of damaged neural tissues to resolve damage-induced inflammation.[41] We identified *Trem2*, *Tyrobp* (*Dap12*), and

Table 2. KEGG pathways relevant to ChIP-Seq-based Spi1 target genes in microglia.

RANK	CATEGORY	FOCUSED GENES IN THE PATHWAY	*P*-VALUE CORRECTED BY BONFERRONI	FDR
1	mmu04142: Lysosome	Abcb9, Ap1b1, Ap1g1, Ap1s3, Ap3b1, Ap3b2, Ap3m2, Ap3s2, Ap4e1, Ap4s1, Arsa, Arsb, Arsg, Atp6v0a1, Atp6v0c, Atp6v0d1, Atp6v0d2, Cd68, Clta, Cltb, Cltc, Ctsa, Ctsb, Ctsc, Ctsd, Ctse, Ctsf, Ctsk, Ctsl, Ctss, Ctsz, Ctns, Fuca1, Galc, Gga1, Gla, Gnptab, Gns, Gusb, Hexa, Hexb, Hyal1, Igf2r, Lamp1, Lamp2, Laptm4a, Laptm4b, Laptm5, Lgmn, LipA, Man2b1, NagA, Neu1, Npc1, Pla2g15, Ppt1, Psap, Scarb2, Sgsh, Slc11a1, Slc11a2, Slc17a5, Sort1, Tcirg1	5.08E-08	3.30E-07
2	mmu04510: Focal adhesion	Actb, Actn1, Akt1, Akt3, Arhgap5, Bcar1, Bcl2, Birc2, Birc3, Capn2, Cav2, Ccnd2, Ccnd3, Cdc42, Col1a1, Col2a1, Col4a1, Col4a2, Col4a6, Col5a1, Col5a3, Diap1, Dock1, Egf, Flna, Flnb, Flnc, Fyn, Grb2, Grlf1, Gsk3b, Igf1, Igf1r, Itga2, Itga4, Itga5, Itga6, Itga7, Itga8, Itga9, Itgav, Itgb3, Itgb5, Itgb7, Kdr, Lama5, Lamb1, Lamc1, Map2k1, Mapk9, Met, Myl10, Myl12a, Mylk, Pak1, Pak2, Parvb, Parvg, Pdgfc, Pgf, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Pip5k1c, Ppp1r12a, Prkca, Prkcb, Pten, Ptk2, Pxn, Rap1a, Rapgef1, Rasgrf1, Rock2, Shc1, Shc2, Shc4, Spp1, Src, Tln1, Tln2, Tnr, Vav1, Vav2, Vav3, Vcl, Vegfa, Vwf, Zyx	1.27E-07	8.24E-07
3	mmu04144: Endocytosis	Acap2, Acvr1b7, Adrb2, Adrbk1, Agap1, Ap2m1, Arap1, Arap3, Arf6, Arrb1, Asap1, Asap3, Cbl, Cblb, Cdc42, Chmp3, Chmp4b, Chmp6, Clta, Cltb, Cltc, Csf1r, Cxcr2, Cxcr4, Dab2, Dnm1, Dnm2, Dnm3, Eea1, Egf, Ehd1, Ehd2, Ehd4, Epn1, Epn2, Eps15, Fgfr2, Fgfr4, Git1, Glit2, Grk5, Grk6, H2-D1, H2-K1, H2-Q6, H2-Q9, H2-T3, Hspa1a, Igf1r, Il2ra, Iqsec1, Itch, Kdr, Met, Mvb12b, Nedd4l, Ntrk1, Pard3, Pard6b, Pip4k2b, Pip5k1a, Pip5k1b, Pip5k1c, Pld1, Prkcz, Psd, Psd3, Psd4, Rab11a, Rab11b, Rab11fip1, Rab11fip2, Rab11fip5, Rab22a, Rab31, Rab5b, Rab5c, Rabep1, Sh3gl1, Sh3gl3, Sh3glb1, Sh3kbp1, Smap2, Smurf1, Src, Tfrc, Tgfbr1, Tgfbr2, Tsg101, Vps37b, Vps37C, Vps45, Vps4b	4.54E-07	2.94E-06
4	mmu04666: Fc gamma R-mediated phagocytosis	Akt1, Akt3, Arf6, Arpc1b, Arpc2, Arpc4, Arpc5l, Asap1, Asap3, Cdc42, Dnm1, Dnm2, Dnm3, Fcgr1, Fcgr2b, Gab2, HCck, Igh, Inpp5d, Limk1, Limk2, Lyn, Map2k1, Myo10, Ncf1, Pak1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Pip4k2b, Pip5k1a, Pip5k1b, Pip5k1c, Pla2g4a, Pla2g6, Plcg2, Pld1, Ppap2a, Ppap2b, Prkca, Prkcb, Prkcd, Prkce, Ptprc, Scin, Sphk2, Syk, Vav1, Vav1, Vav3, Wasf2	5.03E-07	3.26E-06
5	mmu04010: MAPK signaling pathway	Acvr1b, Akt1, Akt3, Arrb1, Atf2, B230120H23Rik, Cacna1a, Cacna1b, Cacna1d, Cacna1f, Cacna1g, Cacna2d3, Cacnb2, Cacnb4, Cacng2, Cacng4, Casp3, Cdc42, Chuk, Dusp16, Dusp2, Dusp3, Dusp4, Dusp5, Dusp6, Egf, Fas, Fgf14, Fgf18, Fgfr1, Fgfr2, Fgfr4, Flna, Flnb, Flnc, Gadd45a, Gna12, Gng12, Grb2, Hspa1a, Ikbkb, Il1a, Il1r1, Il1r2, Map2k1, Map2k3, Map2k6, Map3k1, Map3k11, Map3k12, Map3k13, Map3k14, Map3k2, Map3k3, Map3k5, Map3k7, Map4k2, Map4k3, Map4k4, Mapk14, Mapk9, Mapkapk2, Mapt, Max, Mef2c, Mknk1, Mras, Myc, Nf1, Nfatc2, Nfatc4, Nfkb1, Nr4a1, Ntrk1, Pak1, Pak2, Pla2g12a, Pla2g2e, Pla2g4a, Pla2g5, Pla2g6, Ppm1a, Ppm1b, Ppp3ca, Ppp5c, Prkaca, Prkca, Prkcb, Ptpn5, Ptpn7, Rap1a, Rapgef2, Rasgrf1, Rasgrf2, Rasgrp1, Rasgrp3, Rps6ka2, Rps6ka4, Rps6ka5, Rras2, Srf, Stk3, Tab2, Taok3, Tgfb1, Tgfb2, Tgfbr1, Tgfbr2, Tm4sf19, Tnfrsf1a	6.23E-06	4.04E-05
6	mmu04810: Regulation of actin cytoskeleton	Abi2, Actb, Actn1, Arhgef12, Arhgef6, Arhgef7, Arpc1b, Arpc2, Arpc4, Arpc5l, Baiap2, Bcar1, Cdc42, Chrm3, Csk, Cyfip1, Cyfip2, Diap1, Diap2, Diap3, Dock1, Egf, Fgd1, Fgd3, Fgf14, Fgf18, Fgfr1, Fgfr2, Fgfr4, Git1, Gna12, Gng12, Iqgap1, Iqgap3, Itga2, Itga4, Itga5, Itga6, Itga7, Itga8, Itga9, Itgad, Itgae, Itgam, Itgav, Itgax, Itgb3, Itgb5, Itgb7, Limk1, Limk2, Map2k1, Mras, Msn, Myh10, Myh14, Myh9, Myl10, Myl12a, Mylk, Nckap1, Nckap1l, Pak1, Pak2, Pdgfc, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Pip4k2b, Pip5k1a, Pip5k1b, Pip5k1c, Ppp1r12a, Ptk2, Pxn, Rock2, Rras2, Scin, Ssh1, Ssh2, Tiam1, Tmsb4x, Vav1, Vav2, Vav3, Vcl, Wasf2	1.38E-04	8.96E-04
7	mmu05200: Pathways in cancer	Abl1, Acvr1b, Akt1, Akt3, Arnt, Arnt2, Axin1, Bcl2, Bcl2l1, Bcr, Bid, Birc2, Birc3, Casp3, Casp8, Casp9, Cbl, Cblb, Ccdc6, Ccne1, Cdc42, Cdk6, Cdkn1a, Cdkn1b, Chuk, Col4a1, Col4a2, Col4a6, Crebbp, Csf1r, Csf3r, Ctbp2, Ctnna2, Dapk1, Dapk2, E2f1, Egf, Egln3, Epas1, Fas, Fgf14, Fgf18, Fgfr1, Fgfr2, Flt3, Fzd3, Gli2, Grb2, Gsk3b, Hsp90ab1, Igf1, Igf1r, Ikbkb, Itga2, Itga6, Itgav, Jak1, Lama5, Lamb1, Lamc1, Lef1, Map2k1, Mapk9, Max, Met, Mitf, Msh3, Msh6, Myc, Nfkb1, Ntrk1, Pax8, Pgf, Pias1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Plcg2, Pml, Ppard, Pparg, Prkca, Prkcb, Pten, Ptk2, Ralb, Ralgds, Rara, Rassf1, Rassf5, Rb1, Runx1, Rxra, Rxrb, Spi1, Skp2, Slc2a1, Smad3, Smo, Stat1, Stat3, Stat5a, Stat5b, SufU, Tceb1, Tcf7l2, Tfg, Tgfb1, Tgfb2, Tgfbr1, Tgfbr2, Tpm3, Traf1, Traf3, Traf5, Vegefa, Wnt1, Wnt2b, Wnt5b, Wnt7b, Zbtb16	1.39E-04	8.99E-04

(*Continued*)

Table 2. (*Continued*)

RANK	CATEGORY	FOCUSED GENES IN THE PATHWAY	*P*-VALUE CORRECTED BY BONFERRONI	FDR
8	mmu04662: B cell receptor signaling pathway	Akt1, Akt3, Blnk, Btk, Card11, Cd72, Cd81, Chuk, Dapp1, Fcgr2b, Grb2, Gsk3b, Igh, Ikbkb, Inpp5d, Lyn, Map2k1, Nfat5, Nfatc1, Nfatc2, Nfatc3, Nfatc4, Nfkb1, Nfkbie, Pik3ap1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Pirb, Plcg2, Ppp3ca, Prkcb, Ptpn6, Rasgrp3, Syk, Tm4sf19, Vav1, Vav2, Vav3	2.18E-04	0.001413938
9	mmu04210: Apoptosis	Akt1, Akt3, Apaf1, Atm, Bcl2, Bcl2l1, Bid, Birc2, Birc3, Capn2, Casp3, Casp7, Casp8, Casp9, Cflar, Chuk, Csf2rb, Csf2rb2, Endod1, Fas, Ikbkb, Il1a, Il1r1, Il1rap, Irak2, Irak3, Irak4, Map3k14, Myd88, Nfkb1, Ntrk1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Ppp3ca, Prkaca, Prkar1b, Prkar2a, Prkar2b, Tm4sf19, Tnfrsf1A	4.23E-04	0.002744042
10	mmu05222: Small cell lung cancer	Akt1, Akt3, Apaf1, Bcl2, Bcl2l1, Birc2, Birc3, Casp9, Ccne1, Cdk6, Cdkn1b, Chuk, Col4a1, Col4a2, Col4a6, E2f1, Fhit, Ikbkb, Itga2, Itga6, Itgav, Lama5, Lamb1, Lamc1, Max, Myc, Nfkb1, Pias1, Pik3cb, Pik3cd, Plik3cg, Pik3r1, Pik3r3, Pik3r5, Pten, Ptk2, Rb1, Rxra, Rxrb, Skp2, Traf1, Traf3, Traf5	5.60E-04	0.003631259
11	mmu04670: Leukocyte transendothelial migration	Actb, Actn1, Arhgap5, Bcar1, Cdc42, Cldn14, Cldn23, Ctnna2, Ctnnd1, Cxcr4, Cyba, Cybb, Esam, F11r, Gnai2, Gnai3, Grlf1, Icam1, Itga4, Itgam, Jam3, Mapk14, Mllt4, Msn, Myl10, Myl12a, Ncf1, Ncf2, Ncf4, Nox1, Pecam1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Plcg2, Prkca, Prkcb, Ptk2, Ptk2b, Ptpn11, Pxn, Rap1a, Rapgef4, Rassf5, Rock2, Sipa1, Txk, Vav1, Vav2, Vav3, Vcam1, Vcl	6.57E-04	0.004260774
12	mmu04062: Chemokine signaling pathway	Adcy3, Adcy7, Adrbk1, Akt1, Akt3, Arrb1, Bcar1, Ccl1, Ccl2, Ccl3, Ccl4, Ccl5, Ccl9, Ccr1, Ccr6, Cdc42, Chuk, Csk, Cx3cl1, Cx3cr1, Cxcr2, Cxcr3, Cxcr4, Elmo1, Fgr, Foxo3, Gnai2, Gnai3, Gnb1, Gng12, Gng2, Gng4, Gngt2, Grb2, Grk5, Grk6, Gsk3a, Gsk3b, Hck, Ikbkb, Lyn, Map2k1, Ncf1, Nfkb1, Pak1, Pard3, Pf4, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Plcb2, Plcb4, Prex1, Prkaca, Prkcb, Prkcd, Prkcz, Ptk2, Ptk2b, Pxn, Rap1a, Rock2, Shc1, Shc2, Shc4, Stat1, Stat3, Stat5b, Tiam1, Vav1, Vav2, Vav3, Xcr1	8.52E-04	0.005526401
13	mmu05220: Chronic myeloid leukemia	Abl1, Acvr1b, Akt1, Akt3, Bcl2l1, Bcr, Cbl, Cblb, Cdk6, Cdkn1a, Cdkn1b, Chhuk, Ctbp2, E2f1, Gab2, GRrb2, Ikbkb, Map2k1, Myc, Nfkb1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Ptpn11, Rb1, Runx1, Shc1, Shc2, Shc4, Smad3, Stat5a, Stat5b, Tgfb1, Tgfb2, Tgfbr1, Tgfbr2	0.001156477	0.007503174
14	mmu04070: Phosphatidylinositol signaling system	Calm1, Calm3, Cds2, Dgkd, Dgkg, Dgki, Dgkz, Inpp4a, Inpp5a, Inpp5d, Inpp5k, Itpk1, Itpkb, Itpr1, Itpr2, Itpr3, Ocrl, Pl4kb, Pik3c2a, Pik3c2b, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Pip4k2b, Pip5k1a, Pip5k1b, Pip5k1c, Plcb2, Plcb4, Plcd3, Plce1, Plcg2, Prkca, Prkcb, Pten, Synj1	0.002271981	0.014748224
15	mmu04660: T cell receptor signaling pathway	Akt1, Akt3, Card11, Cbl, Cblb, Cd247, Cd28, Cd4, Cdc42, Chuk, Ctla4, Fyn, Grap2, Grb2, Gsk3b, Icos, Ikbkb, Il10, Il4, Lcp2, Map2Kk1, Map3k14, Map3k7, Mapk14, Mapk9, Nck1, Nck2, Nfat5, Nfatc1, Nfatc2, Nfatc3, Nfatc4, Nfkb1, Nfkbie, Pak1, Pak2, Pdk1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Ppp3ca, Prkcq, PtpnN6, Ptprc, Rasgrp1, Tec, Tm4sf19, Vav1, Vav2, Vav3	0.002792449	0.018131185
16	mmu05212: Pancreatic cancer	Acvr1b, Akt1, Akt3, Arhgef6, Bcl2l1, Casp9, Cdc42, Cdk6, Chuk, E2f1, Egf, Ikbkb, Jak1, Map2k1, Mapk9, Nfkb1, Pgf, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Pld1, Ralb, Ralgds, Rb1, Smad3, Stat1, Stat3, Tgfb1, Tgfb2, Tgfbr1, Tgfbr2, Vegfa	0.015811557	0.103295105
17	mmu04650: Natural killer cell mediated cytotoxicity	Bid, Casp3, Cd244, Cd247, Fas, Fcgr3, Fcgr4, Fyn, Grb2, H2-D1, H2-K1, Hcst, Icam1, Icam2, Ifnab, Ifngr1, Ifngr2, Igh, Klrb1c, Lcp2, Map2k1, Nfat5, Nfatc1, Nfatc2, Nfatc3, Nfatc4, Pak1, Pik3cb, Pik3cd, Pik3cg, Pik3r1, Pik3r3, Pik3r5, Plcg2, Ppp3ca, Prkca, Prkcb, Ptk2b, Ptpn11, Ptpn6, Raet1d, Raet1e, Sh3bp2, Shc1, Shc2, Shc4, Syk, Tm4sf19, Tyrobp, Vav1, Vav2, Vav3	0.038233058	0.252464917

Notes: By importing Entrez Gene IDs of 5,264 ChIP-Seq-based Spi1 target genes into the Functional Annotation tool of DAVID, KEGG pathways showing significant relevance to the set of imported genes were identified. They are listed with pathways, focused genes, p-value corrected by Bonferroni multiple comparison test, and false discovery rate (FDR).

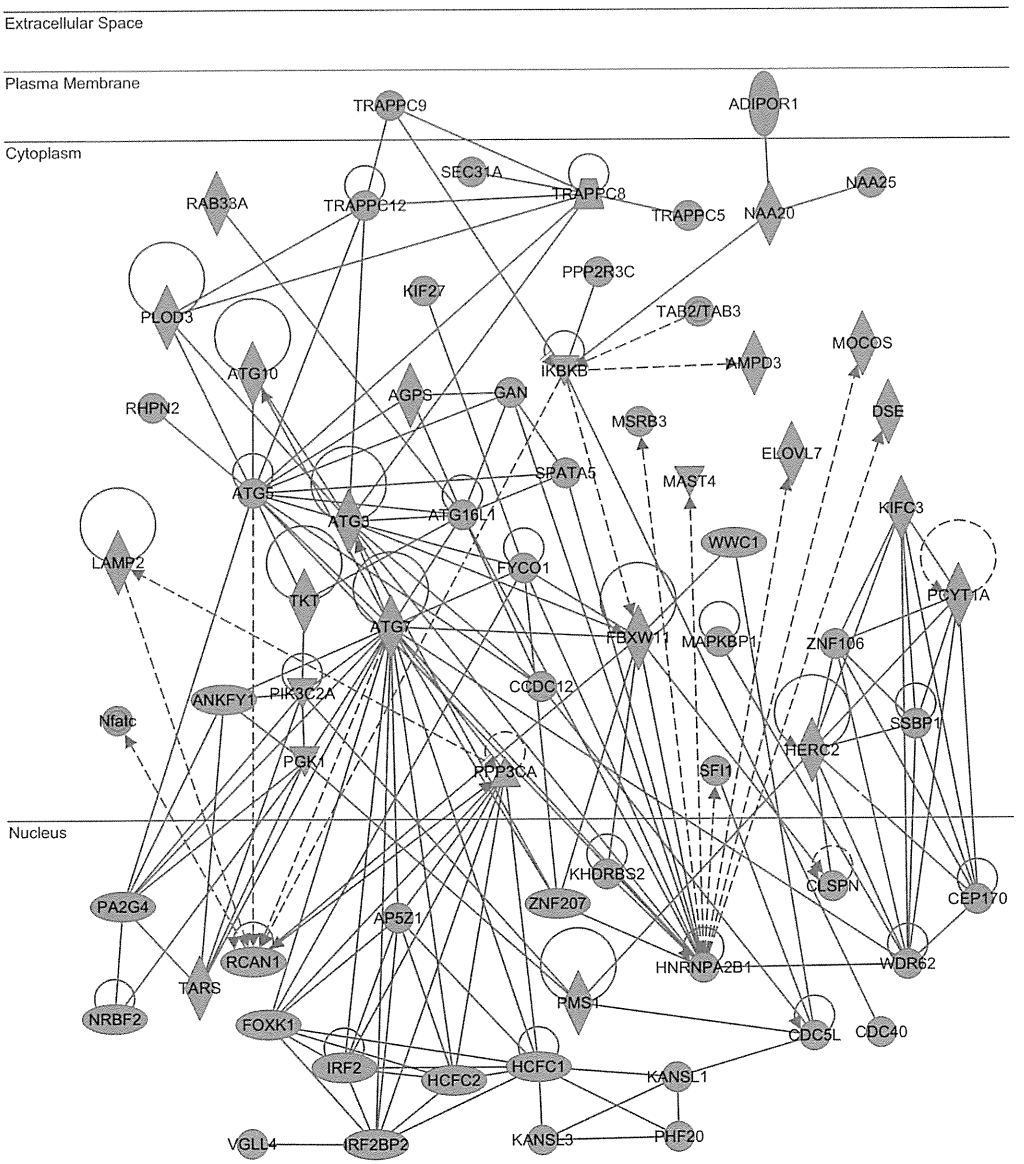

Figure 5. IPA "Cell Morphology, Cellular Function and Maintenance, Cell Death and Survival" network relevant to Spi1 target genes. Entrez Gene IDs of 5,264 ChIP-Seq-based Spi1 target genes were imported into the Core Analysis tool of IPA. It extracted the "Cell Morphology, Cellular Function and Maintenance, Cell Death and Survival" network as the first rank significant functional network as listed in Supplementary Table 3. Spi1 target genes are colored by red.

Syk as a group of Spi1 target genes, consistent partly with previous observations.[42] Importantly, Dap12 serves as a hub of the "microglial sensome" network, on which major molecular connections are concentrated.[30] These observations indicate that aberrant function of microglia plays a central role in the pathogenesis of NHD.

A recent study by combining genome-wide linkage analysis and exome sequencing identified several mutations in the *CSF1R* gene in patients with hereditary diffuse leukoencephalopathy with spheroids (HDLS), a rare autosomal dominant disease that affects predominantly the CNS white matter.[43] Clinically, HDLS exhibits early-onset personality and behavioral disturbances, dementia, and parkinsonism. HDLS shows striking similarities to the pathology of NHD, in view of diffuse demyelination and gliosis with morphologically abnormal microglia and marked accumulation of axonal spheroids, although HDLS never exhibits bone cysts and basal ganglia calcification, both of which are characteristic features of NHD. We identified *Csf1r*, *Csf1*, and *Il34* as another group of Spi1 target genes, indicating that HDLS represents a disease entity designated as "microgliopathy" caused by microglial dysfunction. Based on these observations, we could

propose a hypothesis that microglial dysfunction caused by aberrant regulation of PU.1 target genes contributes to the pathogenesis of various neurodegenerative and neuroinflammatory diseases. Importantly, a recent study indicates that DAP12 acts as a central regulator in gene networks of the late-onset AD.[44]

Although ChIP-Seq serves as a highly efficient method for genome-wide profiling of transcription factor-binding sites, the method intrinsically requires several technical considerations to achieve reproducibility of the results.[45] The specificity of antibodies, the sequencing depth and coverage, the source of target cell types and relevant controls, developmental stages, and culture conditions constitute critical factors capable of affecting both genetic and epigenetic features. Motif analysis of a defined set of high-quality peaks makes it possible to evaluate the antibody specificity and to predict the specificity of DNA–protein interaction to some extent.[45] In general, DNA-binding by transcription factors is a highly dynamic process following recruitment of the complex of auxiliary factors, such as coactivators and corepressors. However, in most occasions, ChIP-Seq data generally reflect a snapshot of binding actions of limited DNA-binding factors onto responsive elements, not always corresponding to their biological activities. Because of these limitations, it is highly important to validate main results by examining technical and biological replicates of samples with different sources of ChIP-quality antibodies, along with transcriptome analysis.

Conclusions

By analyzing a ChIP-Seq dataset numbered SRP036026 with the Strand NGS program, we identified 5,264 Spi1 target protein-coding genes in BV2 mouse microglial cells. They included *Spi1, Irf8, Runx1, Csf1r, Csf1, Il34, Aif1* (*Iba1*), *Cx3cr1, Trem2*, and *Tyrobp*. Motif analysis identified the PU-box consensus sequences in the genomic regions surrounding ChIP-Seq peaks. By using pathway analysis tools of bioinformatics, we found that ChIP-Seq-based Spi1 target genes show a significant relationship with diverse pathways essential for normal function of monocytes/macrophages, such as endocytosis, Fcγ receptor-mediated phagocytosis, and lysosomal degradation. These results suggest that PU.1/Spi1 plays a pivotal role in regulation of the genes relevant to specialized functions of microglia. Therefore, aberrant regulation of PU.1 target genes might contribute to the development of neurodegenerative diseases with accumulation of activated microglia.

Acknowledgments

The authors thank Ms. Mutsumi Motouri for her invaluable help.

Author Contributions

JS designed the methods, analyzed the data, and drafted the manuscript. NA, SK, and YK helped with the data analysis. All authors have read and approved the final manuscript.

Supplementary Materials

Supplementary Figure 1. FastQC analysis of ChIP-Seq data. FASTQ format files composed of cleaned NGS data derived from Spi1 (panel a) or Cebpa (panel b) ChIP-Seq were imported into the FastQC program. The per base sequence quality score is shown with the median (red line), the mean (blue line), and the interquartile range (yellow box).

Supplementary Figure 2. Genomic locations of Spi1 ChIP-Seq peaks on the *Trem2* gene. The genomic locations of Spi1 ChIP-Seq peaks were determined by importing the processed data into GenomeJack. An example of triggering receptor expressed on myeloid cells 2 (Trem2; Entrez Gene ID 83433) is shown, where a MACS peak numbered 51970 in the Spi1.bam Coverage lane is located in the intronic region of the *Trem2* gene (panel a) with a Spi1-binding consensus sequence motif highlighted by orange square (panel b).

Supplementary Figure 3. Genomic locations of Spi1 ChIP-Seq peaks on the *Tyrobp* gene. The genomic locations of Spi1 ChIP-Seq peaks were determined by importing the processed data into GenomeJack. An example of TYRO protein kinase binding protein (Tyrobp, Dap12; Entrez Gene ID 22177) is shown, where a MACS peak numbered 100752 in the Spi1.bam Coverage lane is located in the promoter region of the *Tyrobp* gene (panel a) with a Spi1-binding consensus sequence motif (reverse complement) highlighted by orange square (panel b).

Supplementary Figure 4. KeyMolnet molecular network relevant to Spi1 target genes. Entrez Gene IDs of 5,264 Spi1 target genes were imported into KeyMolnet. The neighboring network-search algorithm extracted the extremely complex network composed of 5,788 molecules and 13,719 molecular relations, showing the most significant relationship with "Transcriptional regulation by RB/E2F". Red nodes indicate those closely related to imported genes. White nodes exhibit additional nodes extracted automatically from the core contents of KeyMolnet to establish molecular connections. The molecular relation is indicated by solid line with arrow (direct binding or activation), solid line with arrow and stop (direct inactivation), solid line without arrow (complex formation), dash line with arrow (transcriptional activation), and dash line with arrow and stop (transcriptional repression).

Supplementary Table 1. The set of 5,264 ChIP-Seq-based Spi1 target genes in microglia. From the ChIP-Seq dataset, we identified 5,264 Spi1-target genes in BV2 mouse microglia showing fold enrichment (FE) ≥5. They are listed with the chromosome, the position of the peak (start, end), FE, Entrez Gene ID, Gene Symbol, and Gene Name. The set of 1,844 Cebpa-target genes are underlined.

Supplementary Table 2. The list of 100 microglial sensome genes. The set of 100 microglial sensome genes (Ref. 30) are listed in order of their expression levels with Entrez Gene ID, Gene Symbol, Gene Name, and an existence of ChIP-Seq-based peaks for Spi1 or Cebpa.

Supplementary Table 3. IPA functional networks relevant to ChIP-Seq-based Spi1 target genes in microglia. By importing Entrez Gene IDs of 5,264 ChIP-Seq-based Spi1 target genes into the core analysis tool of IPA, functional networks showing significant relevance to the imported genes were identified. They are listed with rank, category of functional networks, focused molecules, and *P*-value by the Fisher's exact test.

REFERENCES

1. Saijo K, Glass CK. Microglial cell origin and phenotypes in health and disease. *Nat Rev Immunol.* 2011;11(11):775–87.
2. Ginhoux F, Greter M, Leboeuf M, et al. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science.* 2010;330(6005):841–5.
3. Schulz C, Perdiguero EG, Wieghofer P, et al. Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat Neurosci.* 2013;16(3):273–80.
4. Perry VH, Nicoll JA, Holmes C. Microglia in neurodegenerative disease. *Nat Rev Neurol.* 2010;6(4):193–201.
5. Heneka MT, Kummer MP, Latz E. Innate immune activation in neurodegenerative disease. *Nat Rev Immunol.* 2014;14(7):463–77.
6. Tahara K, Kim HD, Jin JJ, Maxwell JA, Li L, Fukuchi K. Role of toll-like receptor signalling in Ab uptake and clearance. *Brain.* 2006;129(pt 11):3006–19.
7. Halle A, Hornung V, Petzold GC, et al. The NALP3 inflammasome is involved in the innate immune response to amyloid-β. *Nat Immunol.* 2008;9(8):857–65.
8. Béraud D, Twomey M, Bloom B, et al. α-Synuclein alters toll-like receptor expression. *Front Neurosci.* 2011;5:80.
9. Crotti A, Benner C, Kerman BE, et al. Mutant Huntingtin promotes autonomous microglia activation via myeloid lineage-determining factors. *Nat Neurosci.* 2014;17(4):513–21.
10. Turkistany SA, DeKoter RP. The transcription factor PU.1 is a critical regulator of cellular communication in the immune system. *Arch Immunol Ther Exp (Warsz).* 2011;59(6):431–40.
11. Pham TH, Minderjahn J, Schmidl C, et al. Mechanisms of in vivo binding site selection of the hematopoietic master transcription factor PU.1. *Nucleic Acids Res.* 2013;41(13):6391–402.
12. McKercher SR, Torbett BE, Anderson KL, et al. Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J.* 1996;15(20):5647–58.
13. Beers DR, Henkel JS, Xiao Q, et al. Wild-type microglia extend survival in PU.1 knockout mice with familial amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A.* 2006;103(43):16021–6.
14. Smith AM, Gibbons HM, Oldfield RL, et al. M-CSF increases proliferation and phagocytosis while modulating receptor and transcription factor expression in adult human microglia. *J Neuroinflammation.* 2013;10:85.
15. Smith AM, Gibbons HM, Oldfield RL, et al. The transcription factor PU.1 is critical for viability and function of human brain microglia. *Glia.* 2013;61(6):929–42.
16. Horiuchi M, Wakayama K, Itoh A, et al. Interferon regulatory factor 8/interferon consensus sequence binding protein is a critical transcription factor for the physiological phenotype of microglia. *J Neuroinflammation.* 2012;9:227.
17. Zusso M, Methot L, Lo R, Greenhalgh AD, David S, Stifani S. Regulation of postnatal forebrain amoeboid microglial cell proliferation and development by the transcription factor Runx1. *J Neurosci.* 2012;32(33):11285–98.
18. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet.* 2009;10(10):669–80.
19. Satoh J, Kawana N, Yamamoto Y. Pathway analysis of ChIP-Seq-based NRF1 target genes suggests a logical hypothesis of their involvement in the pathogenesis of neurodegenerative diseases. *Gene Regul Syst Bio.* 2013;7:139–52.
20. Blasi E, Barluzzi R, Bocchini V, Mazzolla R, Bistoni F. Immortalization of murine microglial cells by a v-raf/v-myc carrying retrovirus. *J Neuroimmunol.* 1990;27(2–3):229–37.
21. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.
22. Zhang X, Robertson G, Krzywinski M, et al. PICS: probabilistic inference for ChIP-seq. *Biometrics.* 2011;67(1):151–63.
23. Li L. GADEM: a genetic algorithm guided formation of spaced dyads coupled with an EM algorithm for motif discovery. *J Comput Biol.* 2009;16(2):317–29.
24. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
25. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 2012;40(Database Issue):D109–14.
26. Satoh J, Tabunoki H. Comprehensive analysis of human microRNA target networks. *BioData Min.* 2011;4:17.
27. Jin H, Li L, Xu J, et al. Runx1 regulates embryonic myeloid fate choice in zebrafish through a negative feedback loop inhibiting Pu.1 expression. *Blood.* 2012;119(22):5239–49.
28. Wang Y, Szretter KJ, Vermi W, et al. IL-34 is a tissue-restricted ligand of CSF1R required for the development of Langerhans cells and microglia. *Nat Immunol.* 2012;13(8):753–60.
29. Klünemann HH, Ridha BH, Magy L, et al. The genetic causes of basal ganglia calcification, dementia, and bone cysts: DAP12 and TREM2. *Neurology.* 2005;64(9):1502–7.
30. Hickman SE, Kingery ND, Ohsumi TK, et al. The microglial sensome revealed by direct RNA sequencing. *Nat Neurosci.* 2013;16(12):1896–905.
31. Kierdorf K, Prinz M. Factors regulating microglia activation. *Front Cell Neurosci.* 2013;7:44.
32. Ginhoux F, Lim S, Hoeffel G, Low D, Huber T. Origin and differentiation of microglia. *Front Cell Neurosci.* 2013;7:45.
33. Henn A, Lund S, Hedtjärn M, Schrattenholz A, Pörzgen P, Leist M. The suitability of BV2 cells as alternative model system for primary microglia cultures or for animal experiments examining brain inflammation. *ALTEX.* 2009;26(2):83–94.
34. Rosenbauer F, Wagner K, Kutok JL, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet.* 2004;36(6):624–30.
35. Rekhtman N, Choe KS, Matushansky I, Murray S, Stopka T, Skoultchi AI. PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation. *Mol Cell Biol.* 2003;23(21):7460–74.
36. Chen H, Ray-Gallet D, Zhang P, et al. PU.1 (Spi-1) autoregulates its expression in myeloid cells. *Oncogene.* 1995;11(8):1549–60.
37. Wang D, D'Costa J, Civin CI, Friedman AD. C/EBPα directs monocytic commitment of primary myeloid progenitors. *Blood.* 2006;108(4):1223–9.
38. Friedman AD. C/EBPa induces PU.1 and interacts with AP-1 and NF-kB to regulate myeloid development. *Blood Cells Mol Dis.* 2007;39(3):340–3.
39. Gómez-Nicola D, Fransen NL, Suzzi S, Perry VH. Regulation of microglial proliferation during chronic neurodegeneration. *J Neurosci.* 2013;33(6):2481–93.
40. Satoh J, Tabunoki H, Ishida T, et al. Immunohistochemical characterization of microglia in Nasu-Hakola disease brains. *Neuropathology.* 2011;31(4):363–75.
41. Takahashi K, Rochford CD, Neumann H. Clearance of apoptotic neurons without inflammation by microglial triggering receptor expressed on myeloid cells–2. *J Exp Med.* 2005;201(4):647–57.
42. Weigelt K, Ernst W, Walczak Y, et al. Dap12 expression in activated microglia from retinoschisin-deficient retina and its PU.1-dependent promoter regulation. *J Leukoc Biol.* 2007;82(6):1564–74.
43. Rademakers R, Baker M, Nicholson AM, et al. Mutations in the colony stimulating factor 1 receptor (CSF1R) gene cause hereditary diffuse leukoencephalopathy with spheroids. *Nat Genet.* 2011;44(2):200–5.
44. Zhang B, Gaiteri C, Bodea LG, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell.* 2013;153(3):707–20.
45. Landt SG, Marinov GK, Kundaje A, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* 2012;22(9):1813–31.