

RESEARCH

Open Access



Variants in *CCL16* are associated with blood plasma and cerebrospinal fluid *CCL16* protein levels

Mark T. W. Ebbert^{1†}, Lyndsay A. Staley^{1†}, Joshua Parker¹, Sheradyn Parker¹, Matthew Bailey², for the Alzheimer's Disease Neuroimaging Initiative, Perry G. Ridge¹, Alison M. Goate³ and John S. K. Kauwe^{1*}

From 12th Annual Biotechnology and Bioinformatics Symposium (BIOT-2015)
Provo, UT, USA. 10-11 December 2015

Abstract

Background: *CCL16* is a chemokine predominantly expressed in the liver, but is also found in the blood and brain, and is known to play important roles in immune response and angiogenesis. Little is known about the gene's regulation.

Methods: Here, we test for potential causal SNPs that affect *CCL16* protein levels in both blood plasma and cerebrospinal fluid in a genome-wide association study across two datasets. We then use METAL to performed meta-analyses with a significance threshold of $p < 5 \times 10^{-8}$. We removed SNPs where the direction of the effect was different between the two datasets.

Results: We identify 10 SNPs associated with increased *CCL16* protein levels in both biological fluids.

Conclusions: Our results will help understand *CCL16*'s regulation, allowing researchers to better understand the gene's effects on human health.

Keywords: Blood, Brain, *CCL16*, Plasma, Cerebrospinal fluid, Genetics, Association

Background

Chemokines are specialized cytokines (cell-signaling proteins) that induce chemotaxis in proximal cells. The CC subfamily is primarily a chemoattractant to monocytes and lymphocytes [1], demonstrating their role in immune response. Chemokines are highly conserved across species, suggesting their roles are essential to viable offspring [2]. They have been implicated specifically in brain health, including neuronal migration during development and even neuronal death [2, 3], which has important implication on brain health throughout life and could have important implications in neurodegenerative diseases. *CCL16* is part of the CC chemokine subfamily and is predominantly expressed in the liver [4]. The *CCL16* protein is also known to be present and active in the brain and blood [3, 5].

CCL16 is located on the q arm of chromosome 17 amongst a cluster of other chemokines and is known to play a role in angiogenesis [5]. While the *CCL16* protein is known to be heavily active in the liver, little is known about its overall regulation in the brain and blood.

To date, most research regarding *CCL16* is limited to its expression in the liver and its role in chemotaxis, generally. Little is known about *CCL16* gene regulation and the protein's role across all tissues. It is clearly involved in immune response, as one of its primary functions is to attract lymphocytes and monocytes [1], making *CCL16* a potentially critical protein throughout the body. It is also unclear how *CCL16* protein levels across blood plasma and CSF are regulated.

Here, we measure *CCL16* protein levels in cerebrospinal fluid (CSF) and blood plasma, and perform a genome-wide association analysis to identify SNPs that are associated with *CCL16* levels in both CSF and blood plasma. These findings will help clarify *CCL16* regulatory mechanisms and their effects on human development and health.

* Correspondence: Kauwe@byu.edu

†Equal contributors

¹Department of Biology, Brigham Young University, Provo, UT 84602, USA
Full list of author information is available at the end of the article

Methods

Subjects and data description

Exactly 246 and 240 CSF and blood plasma samples, respectively, were used in this study from participants in the Knight-Alzheimer's disease Research Center at Washington University School of Medicine (Knight ADRC) and 297 and 347 CSF and blood plasma samples, respectively, were used from the Alzheimer's Disease Neuroimaging Initiative (ADNI). From the Knight ADRC samples, approximately 93 % of the samples were controls, and 7 % were Alzheimer's disease cases, and from the ADNI samples, approximately 85 % were controls, and 15 % were AD cases. We measured levels for CCL16 in each sample using the Human DiscoveryMAP Panel v1.0 and a Luminox 100 platform [6]. All samples were genotyped using the Illumina 610 or the Omniexpress chip. The Knight ADRC samples and associated collection methods were previously described [7, 8]. We collected the ADNI samples from the ADNI database (adni.loni.usc.edu), which were part of the ADNI biomarker study [9]. All samples are of European descent.

SNP imputation

The SNPs were imputed as previously described [6]. Briefly, data from the 1000 Genomes Project (June 2012 release) were used to impute SNPs using Beagle. Imputed SNPs with the following criteria were removed: (1) an r^2 of 0.3 or lower, (2) a minor allele frequency (MAF) lower than 0.05 (3) out of Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$), (4) a call rate lower than 95 %, or (5) a Gprobs score lower than 0.90. Exactly 5,815,690 SNPs passed the QC process.

Data cleaning and analysis

We excluded SNPs that exceeded thresholds for Hardy-Weinberg Equilibrium [10, 11] ($--hwe$ 0.00001), missing genotype rate ($--geno$ 0.05), and minor allele frequency ($--maf$ 0.01) for each data set, using PLINK version 1.07 [12], to perform genotype quality control. We then excluded individuals with a missing genotyping rate greater than 2 % ($--mind$ 0.02), leaving 246 individuals from Knight ADRC and 282 samples from ADNI after cleaning. Remaining Knight ADRC and ADNI samples consisted of 40 and 61 % males, respectively. The average age for ADNI samples was 76 years, ranging from 58 to 91 years, and the average age for Knight ADRC samples was 73, ranging from 49 to 91.

After data cleaning, we tested for an association between each remaining SNP and CCL16 CSF levels within each dataset, adjusting for age, gender, and the first two principal components generated by EigenSoft [13, 14]. We then performed a meta-analysis across both data sets, accounting for sample size, p-values, and direction of effect using the default METAL [15] settings. We

retained all SNPs that had a genome-wide significant meta-analysis p-value less than 5×10^{-8} and that had the same direction of effect in both datasets. We then tested associations between each SNP and CCL16 plasma levels following the same protocol and kept only those SNPs that were significantly associated with both CSF and plasma CCL16 levels.

To assess known functional effects and identify SNPs that are biologically likely to modify gene expression or function, we searched all significant SNPs in the NHGRI catalog of published genome-wide association studies [16] (downloaded July, 2015) for known disease associations, collected RegulomeDB annotations (accessed September, 2015) [17], and collected functional annotations from wANNOVAR [18, 19].

We performed a conditional analysis for all included SNPs that were in or near a given region to test whether there is one or multiple independent effects in the region [20]. Conditional analysis is a follow-up method used to test if there are secondary association signals within a region by retesting each SNP while including the top SNP as a covariate. We chose the most significant SNP in the region to use as a covariate in the conditional analysis.

Results

We identified 34 and 25 SNPs significantly associated with CCL16 protein levels in CSF and plasma, respectively (Additional file 1 and Additional file 2). Of these, 10 SNPs were significantly associated with increased CCL16 protein levels in both CSF and plasma, based on the meta-analyses, all of which are in or near the CCL16 gene (Table 1). The genomic inflation factor was 1.0 (q-q plots can be found in Additional file 3 and Additional file 4). None of the 10 SNPs have been reported to show association with disease in the NHGRI GWAS catalog. Two SNPs are located in the 3' untranslated region (UTR), 4 are intronic, 1 is downstream, and 3 are intergenic (Table 1). SNPs are identified as "downstream" if they are within 500 nucleotides of the 3' end of a gene, according to the National Center for Biotechnology Information's (NCBI) SNP FAQ Archive [21]. Minor allele frequencies (MAF) for the ten SNPs ranged from 0.06 to 0.14 and RegulomeDB scores ranged from '1f' to '6', with three SNPs having no known regulation data, according to RegulomeDB.

All ten SNPs associated with both CSF and plasma CCL16 protein levels are in high linkage disequilibrium, suggesting there is likely one association signal in the region (Figs. 1 and 2). The conditional analysis further supports a single signal as the p-values are no longer genome-wide significant when including the most significant SNP, rs80329614, as a covariate. SNP rs80329614 is the SNP identified as being "downstream"

Table 1 Significant SNPs in or near the *CCL16* gene on chromosome 17 that met our inclusion criteria with pertinent biological information implicating them in *CCL16* regulation

SNP	Base Pair position	Minor Allele	Major Allele	MAF	Predicted Function	RegulomeDB score	Meta-analysis p-value	
							CSF	Plasma
rs80329614	34303312	C	T	0.1406	downstream	3a	1.666E-19	5.853E-28
rs11080368	34305071	A	C	0.1040	intronic	No data	2.321E-19	1.777E-27
rs11080369	34305164	C	A	0.1042	intronic	1f	2.321E-19	1.777E-27
rs33995560	34303771	C	T	0.1122	UTR3	No data	2.941E-19	2.094E-27
rs7216969	34305048	A	G	0.1040	intronic	6	2.941E-19	2.094E-27
rs150951362	34304264	A	G	0.1008	UTR3	6	1.65E-17	2.05E-20
rs75236781	34306470	C	G	0.0613	intronic	6	2.958E-17	3.942E-20
rs149197550	34295254	T	C	0.0641	intergenic	No data	8.32E-16	9.20E-18
rs4795104	34287400	T	A	0.0639	intergenic	6	9.626E-16	1.015E-17
rs4796144	34293003	A	G	0.0641	intergenic	6	9.626E-16	1.015E-17

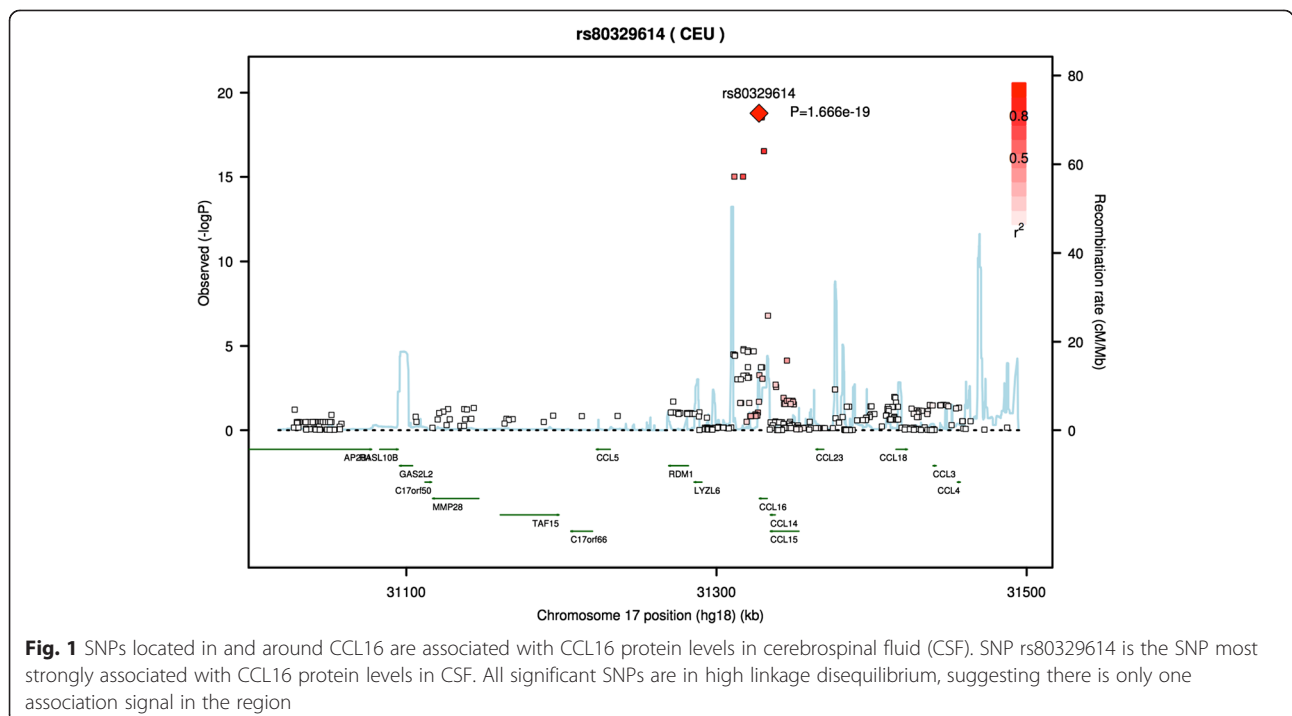
of *CCL16* (Table 1, Figs. 1 and 2) and has a RegulomeDB score of ‘3a’.

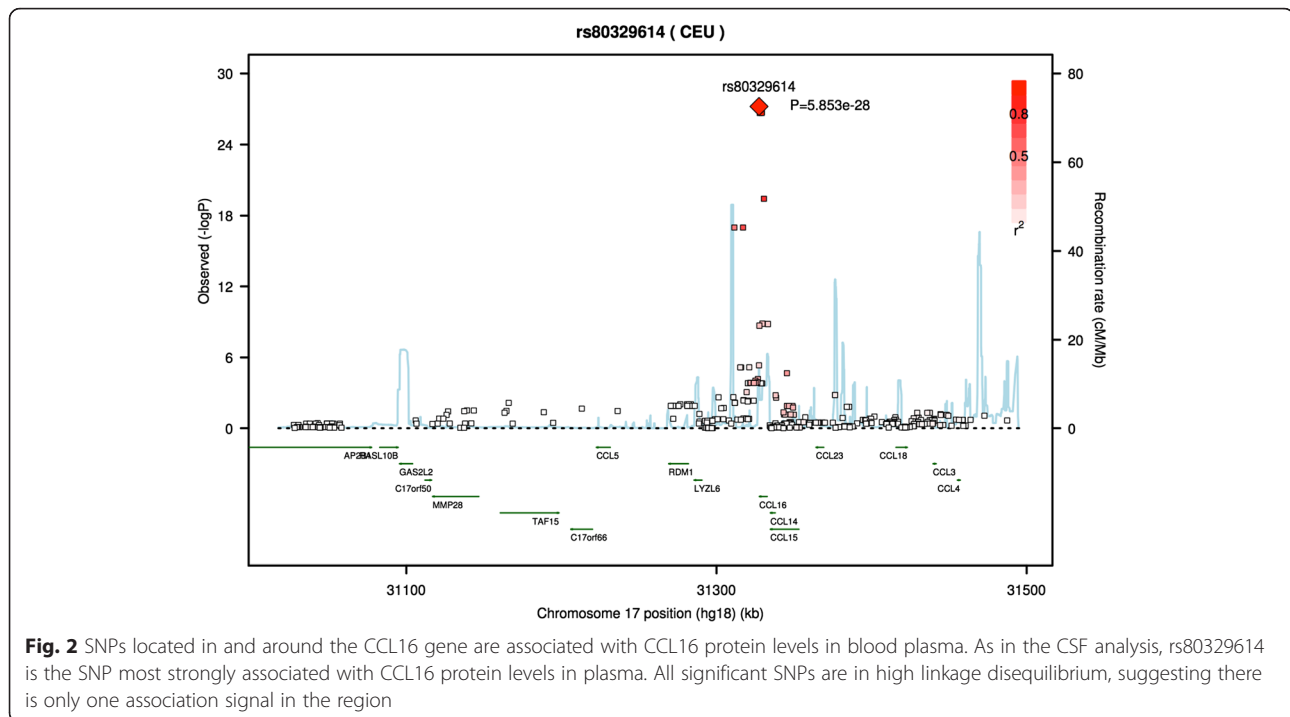
SNP rs11080369 is an intronic SNP with a RegulomeDB score of ‘1f’ and was previously demonstrated to be within an expression quantitative trait locus (eQTL) that specifically modifies *CCL16* gene expression [22]. The remaining 8 SNPs do not have known regulatory associations or functions, according to RegulomeDB, but 5 of the 8 are located in genic regions, such as UTRs and introns.

Discussion

We identified exactly 10 SNPs associated with increased *CCL16* protein levels in both CSF and blood plasma, all

of which were located in or around the *CCL16* gene and based on the conditional analysis, all representing on single signal. The SNP most strongly associated with *CCL16* protein levels was rs80329614, which was identified as being “downstream” (within 500 nucleotides downstream) of *CCL16* with a RegulomeDB score of “3a”. RegulomeDB scores range from “1a” to “6” where lower scores indicate stronger evidence that the SNP affects gene regulation based on both empirical data, such as ChIP-seq, and whether the SNP is within a known transcription factor binding motif. A score of “3a” indicates minimal evidence that a SNP is involved in gene regulations, but RegulomeDB can only represent what is





currently known based on experimental data. Thus, further investigation of this SNP may be warranted. SNP rs11080369 received a score of “1F”, indicating that it is known to be part of an eQTL. As such, the rest of the SNPs are likely part of the same eQTL since they are all in strong linkage disequilibrium.

Two of the remaining SNPs (rs33995560 and rs150951362) are located in the 3'UTR of *CCL16*, which can play an important role in gene transcription and translation [23–25], while three others (rs11080368, rs7216969, rs75236781) are located within *CCL16* introns, which can also affect gene regulation [25]. The remaining three SNPs (rs149197550, rs4795104, and rs4796144) are located between approximately 8000 and 16000 nucleotides downstream of the gene. Intergenic variants are generally less likely to affect transcription than variants within the promoter region or the gene itself, though it is possible if the variant affects transcription factor binding. Many genes have enhancers both upstream and downstream that can be active in specific tissues, depending on the transcription factors expressed in the tissue [26, 27]. However, given that rs149197550, rs4795104, and rs4796144 are in high linkage disequilibrium with the other significant SNPs, we believe they are less likely to be causal variants. The remaining three are intergenic. Identifying which SNP(s) directly affect *CCL16* regulation will require experimental data, but we believe the most suspect from these 10 is rs80329614 because it has the strongest association, its proximity to the gene, and the

fact that many genes have regulatory elements (e.g., enhancers) downstream [26, 27], though the VISTA enhancer database does not have data on *CCL16* enhancers [26].

While these 10 SNPs are the most significant and biologically likely to affect *CCL16* regulation based on our criteria, there may be other SNPs in the individual CSF and plasma lists that regulate *CCL16* independently, including those found in other genes. More biological data will be necessary to identify causal SNPs. Additionally, our data are not whole exome or genome and there may be causal variants in LD with our top hits, associated with the single signal we've seen in our results that were not yet genotyped. Full sequencing data within the region may reveal other candidate causal variations.

Conclusions

Our results show that one or more SNPs in or around the *CCL16* gene are associated with increased *CCL16* protein levels in both CSF and plasma, but it is not clear whether that is due to regulation in multiple tissues such as the brain and liver, or whether the protein is crossing the blood–brain barrier. Our results are an important step to resolving this relationship, as they will allow researcher to explore experimentally how these SNPs affect *CCL16* gene regulation in different tissues. Understanding *CCL16* gene regulation in greater depth will be important to better understanding its role in human health.

Additional files

Additional file 1: File contains a table of SNPs significantly associated with CCL16 levels in blood plasma by meta-analysis. (DOCX 88 kb)

Additional file 2: File contains a table of SNPs significantly associated with CCL16 levels in CSF by meta-analysis. (DOCX 94 kb)

Additional file 3: File contains a Q-Q plot of the CSF data used in this study. (DOCX 82 kb)

Additional file 4: File contains a Q-Q plot of the plasma data used in this study. (DOCX 83 kb)

Abbreviations

ADNI, Alzheimer's Disease Neuroimaging Initiative; CSF, Cerebrospinal Fluid; eQTL, expression quantitative trait locus; Knight ADRC, Knight-Alzheimer's Disease Research Center at Washington University School of Medicine; LD, Linkage Disequilibrium; SNP, single nucleotide polymorphism; UTRs, untranslated regions

Acknowledgements

The authors acknowledge that many scientists contributed in developing the clinical and genetic resources necessary to collect these data and complete this project. The authors also gratefully acknowledge the efforts of hundreds of individuals who participated as subjects in these studies.

The NIH (R01 AG035053, R01 AG042611, P50 AG05681, P01 AG03991, P01 AG026276), the Alzheimer's Association (MNIRG-11-205368), and the Brigham Young University Gerontology Program supported this work. We also acknowledge the Alzheimer's Disease Genetics Consortium (ADGC) and Genetic and Environmental Risk for Alzheimer's Disease Consortium (GERAD) for providing genotype data used in this work.

GERAD was supported by the Medical Research Council (Grant nu 503480), Alzheimer's Research UK (Grant nu 503176), the Wellcome Trust (Grant nu 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant nu 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Some of the samples used in this study were genotyped by the ADGC and GERAD. ADGC is supported by grants from the NIH (#U01AG032984) and GERAD from the Wellcome Trust (GR082604/MA) and the Medical Research Council (G0300429).

Some data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

The ADNI Executive Committee consists of: Michael Weiner, MD UC San Francisco; Paul Aisen, MD UC San Diego; Ronald Petersen, MD, PhD Mayo Clinic, Rochester; Clifford R. Jack, Jr., MD Mayo Clinic, Rochester; William Jagut, MD UC Berkeley; John Q. Trojanowki, MD, PhD U Pennsylvania; Arthur W. Toga, PhD USC; Laurel Beckett, PhD UC Davis; Robert C. Green, MD, MPH Brigham and Women's Hospital/Harvard Medical School; Andrew J. Saykin, PsyD Indiana University; John Morris, MD Washington University St. Louis; Leslie M. Shaw University of Pennsylvania. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Declarations

Publication for this article was funded by Brigham Young University's Department of Biology.

This article has been published as part of *BMC Genomics* Volume 17 Supplement 3, 2016: Selected articles from the 12th Annual Biotechnology and Bioinformatics Symposium: genomics. The full contents of the supplement are available online at <https://bmcbgenomics.biomedcentral.com/articles/supplements/volume-17-supplement-3>.

Availability of data and materials

Data are available to researchers by applying to the respective organizations, ADNI and ADGC consortia. Application is required to protect participant confidentiality. The ADNI data are available at (<http://adni.loni.usc.edu/>), the Knight ADRC data are available through dbGAP (<http://www.ncbi.nlm.nih.gov/gap>).

Authors' contributions

LS, SP, JP, MB, PGR and JSKK performed the data analysis. LS and ME annotated and analyzed the SNPs for significance and drafted the manuscript. All other authors participated in the conception of the project and obtaining data. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Data and analyses in this manuscript were approved by the Brigham Young University Institutional Review Board.

Author details

¹Department of Biology, Brigham Young University, Provo, UT 84602, USA.

²Biology and Biomedical Sciences, Washington University, St. Louis, MO 63110, USA. ³Department of Neuroscience Icahn School of Medicine, New York, NY 10029, USA.

Published: 29 June 2016

References

1. Youn BS, Zhang S, Broxmeyer HE, Antol K, Fraser MJ, Hangoc G, et al. Isolation and characterization of LMC, a novel lymphocyte and monocyte chemoattractant human CC chemokine, with myelosuppressive activity. *Biochem Biophys Res Commun*. 1998;247:217–22.
2. Florio T, Schettini G. Chemokines, their Receptors and Significance in Brain Function. In: *Biology B-N*, editor. Elsevier; 2008 [cited 2015 Oct 13]. p. 242–73. Available from: <http://www.sciencedirect.com/science/article/pii/S1567744307100120>. Accessed Oct 2015.
3. Cartier L, Hartley O, Dubois-Dauphin M, Krause K-H. Chemokine receptors in the central nervous system: role in brain inflammation and neurodegenerative diseases. *Brain Res Rev*. 2005;48:16–42.
4. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. *Science*. 2015;347:1260419.
5. Strasly M, Doronzo G, Cappello P, Capello P, Valdembrì D, Arese M, et al. CCL16 activates an angiogenic program in vascular endothelial cells. *Blood*. 2004;103:40–9.
6. Kauwe JSK, Bailey MH, Ridge PG, Perry R, Wadsworth ME, Hoyt KL, et al. Genome-Wide Association Study of CSF Levels of 59 Alzheimer's Disease

- Candidate Proteins: Significant Associations with Proteins Involved in Amyloid Processing and Inflammation. *PLoS Genet.* 2014;10, e1004758.
7. Cruchaga C, Kauwe JSK, Harari O, Jin SC, Cai Y, Karch CM, et al. GWAS of cerebrospinal fluid tau levels identifies risk variants for Alzheimer's disease. *Neuron.* 2013;78:256–68.
 8. Anne M, Fagan MAM. Inverse relation between in vivo amyloid imaging load and CSF A β 42 in humans. *Ann Neurol.* 2006;59:512–9.
 9. Trojanowski JQ, Vandevertichele H, Korecka M, Clark CM, Aisen PS, Petersen RC, et al. Update on the biomarker core of the Alzheimer's Disease Neuroimaging Initiative subjects. *Alzheimers Dement J Alzheimers Assoc.* 2010;6:230–8.
 10. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet.* 2005;76:887–93.
 11. Jan Graffelman VM. The mid p-value in exact tests for Hardy-Weinberg equilibrium. *Stat Appl Genet Mol Biol.* 2013;12:433–48.
 12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool Set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81:559–75.
 13. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet.* 2006;2, e190.
 14. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38:904–9.
 15. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinforma Oxf Engl.* 2010;26:2190–1.
 16. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014;42:D1001–6.
 17. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 2012;22:1790–7.
 18. Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. *J Med Genet.* 2012;49:433–6.
 19. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38, e164.
 20. Yang J, Ferreira T, Morris AP, Medland SE, Consortium GI of AnT (GIANT), Consortium DiaGRAM (DIAGRAM), et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet.* 2012;44:369–75.
 21. SNP FAQ Archive [Internet] [Internet]. Bethesda, MD: National Center for Biotechnology Information; 2005. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK3848/>. Accessed Oct 2015.
 22. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* 2008;6, e107.
 23. Delay C, Calon F, Mathews P, Hébert SS. Alzheimer-specific variants in the 3'UTR of Amyloid precursor protein affect microRNA function. *Mol Neurodegener.* 2011;6:70.
 24. Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibé B, et al. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet.* 2006;38:813–8.
 25. Greenwood TA, Kelsoe JR. Promoter and intronic variants affect the transcriptional regulation of the human dopamine transporter gene. *Genomics.* 2003;82:511–20.
 26. Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA enhancer browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* 2007;35:D88–92.
 27. Klefogiannis D, Kalnis P, Bajic VB. DEEP: a general computational framework for predicting enhancers. *Nucleic Acids Res.* 2015;43:e6–e6.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

