Preprint

The genetics of ME: a commentary on Hajdarevic et al.

Chris P. Ponting[1] and Simon J. McGrath[2]
[1] MRC Human Genetics Unit, University of Edinburgh, Edinburgh EH4 2XU, UK
[2] c/o DecodeME, MRC Human Genetics Unit, University of Edinburgh, Edinburgh EH4 2XU, UK

Chris.Ponting@ed.ac.uk; psimcgrath@gmail.com

Published version: Brain, Behavior, and Immunity, Volume 104,2022,Pages 181-182,ISSN 0889-1591, https://doi.org/10.1016/j.bbi.2022.06.008. (Paywalled until 16 June 2023.)
Available free until 5 August 2022 with this link: https://authors.elsevier.com/c/1fFlT_KDvPN8-W

People with myalgic encephalomyelitis (ME; also known as chronic fatigue syndrome [CFS]) experience a poor quality of life and high levels of disability. ME is common, affecting approximately 0.3% of the population and, of these, up to one-in-four is either bed- or house-bound. Its core symptoms includes post-exertional malaise (a prolonged, disabling flare of symptoms after minimal physical or mental effort) and profound physical and/or cognitive fatigue that is unassuaged by rest (Institute of Medicine, 2015). Most people experience ME symptoms after viral infection leading to a fear that the SARS-CoV-2 pandemic will cause a steep rise in ME population prevalence. People with ME have no effective diagnostic test or therapy, they report being stigmatised, and they are underserved by health systems worldwide. Furthermore, they are not benefitting from ME research studies as these are mostly underpowered and have not formed a clear consensus on disease aetiology. In a recent paper in *Brain Behavior and Immunity*, Hajdarevic et al. (2022) perform a genome-wide association study (GWAS; Uffelmann, et al., 2021) on ME, seeking DNA variants that explain ME risk as a first step towards development of effective therapy.

Importantly, they found no support for many genetic variants that had previously been implicated in ME. They concluded that all genetic studies to date have not had sufficient participants to be well-powered to identify true associations. It was critical for Hajdarevic et al. (2022) to report such negative findings because this should now help ME research refocus on more productive lines of enquiry (Dibble et al., 2020). The Hajdarevic et al. (2022) study did not identify genetic associations with ME reaching genome-wide significance ($p < 5x10^{-8}$). They attribute this to two factors. First, too few ME cases were analysed: 460 in their Norwegian discovery cohort, and 2105 in the UK Biobank replication cohort. A larger study, such as DecodeME (25,000 intended participants; Devereux-Cooke et al., in press), and future meta-analyses across many studies, will be needed to pinpoint and tease apart ME's various causal pathways. Larger participant numbers are necessary because ME's many causal genetic variants are each expected to make only a small contribution to risk. Second, they stated a concern that the UK Biobank individuals were not phenotyped to the same standard as their Norwegian cohort who nearly all passed the Canadian Consensus Criteria. Rather than applying these criteria, UK Biobank participants were asked whether they had received a doctor's diagnosis of chronic fatigue syndrome. Participants who were unsure were assisted by a nurse who, if necessary, then sought further guidance from a doctor.

Inevitably, the lack of a diagnostic test for ME introduces misdiagnosed individuals into ME cohorts. There is interesting evidence of this in the UK Biobank. A common genetic variant strongly associated with diagnosed Lyme disease (Strausz et al., 2022) is also modestly associated with CFS in the UK Biobank cohort ($p=2.4x10^{-3}$; Dönertaş et al., 2021). This could be explained by some UK Biobank participants with Lyme disease being diagnosed with CFS instead. Future studies will need to be inclusive – resulting in larger cohort sizes and greater statistical power – and also exclusive – removing from consideration any individual lacking core symptoms of ME, especially postexertional malaise.

Hajdarevic et al. (2022) say that they have identified “several potential risk loci”. Among these, they highlight the *TPPP* gene, encoding the tubulin polymerisation promoting protein. This gene contains a rare DNA variant (~1% allele frequency in Northern Europeans) whose statistical association with ME among their Norwegian cases is $p = 8.5 \times 10^{-7}$. This does not reach the de facto standard for achieving robust statistical association of $p < 5 \times 10^{-8}$ and there is only very weak association among UK Biobank CFS cases ($p = 2.5 \times 10^{-2}$). Despite remaining unproven this locus deserves further attention when better-powered studies are performed.

In the meantime, what is known about *TPPP*? Hajdarevic et al. (2022) highlighted this gene’s high expression in brain, in particular in oligodendrocytes. This gene is also expressed more widely, including in ciliated lung and adipose cells, and only non-neurological diseases have systematically been linked to *TPPP*, namely lung disease, ulcerative colitis, Barrett's oesophagus, and haemorrhoidal disease (Mountjoy et al., 2021).

This is not the first time a rare DNA variant has been associated with ME or CFS (Dibble et al., 2020). A variant on chromosome 10 (rs148723539) reached genome-wide significance ($p = 2.3 \times 10^{-9}$) in a 2018 analysis by Neale and co-workers (Mountjoy et al., 2021). This variant lies within the *EBF3* gene yet also could be regulating one or more of three neighbouring genes, namely *MGMT*, *C10orf143*, and *GLRX3* (Mountjoy et al., 2021). *EBF3* encodes a transcription factor involved in B-cell differentiation, bone development and neurogenesis, and is linked to variation in age at menopause; *MGMT* encodes a DNA repair protein involved in cellular defence against mutagenesis and toxicity which has been linked to variation in height and weight; *C10orf143* encodes a protein of unknown function; and, *GLRX3* encodes an oxidoreductase required for haemoglobin maturation.

It needs to be stressed that such statistical associations should be treated with an abundance of caution. Many variants with non-significant associations ($p > 5 \times 10^{-8}$) will never be independently replicated. Others, even those with $p < 5 \times 10^{-8}$, may be technical artefacts, especially variants with low population frequency ($p \leq 1\%$ approximately). Future genetic associations involving more common variants with even smaller *p*-values will be needed. For this, as Hajdarevic et al. (2022) say, ME genetics cohorts will need to exceed current sizes. As a recent review states (Uffelmann et al., 2021) “GWAS for most traits require large (>10,000) sample sizes to yield reproducible results. Such sample sizes can only be generated through collaboration and data sharing agreements.”

People who live with ME should always be at the heart of ME studies because their experience will always improve scientific quality and delivery. They can also consent for their own data to be shared ethically in a spirit of genuine collaboration so that reproducible genetic results can be attained sooner.

985 words.

References

Devereux-Cooke, A., Leary, S., McGrath, S.J., Northwood, E., Redshaw, A., Shepherd, C., Stacey, P., Tripp, C., Wilson, J., Mar, M., Boobyer, D., Bromiley, S., Chowdhury, S., Dransfield, C., Almas, M., Almelid, Ø., Buchanan, D., Garcia, D., Ireland, J., Kerr, S.M., Lewis, I., McDowall, E., Migdal, M., Murray, P., Perry, D., Ponting, C.P. Vitart, V., Wolfe, J.C., in press. BMC Neurol.

Dibble, J.J., McGrath, S.J., Ponting, C.P., 2020. Genetic risk factors of ME/CFS: a critical review. Hum Mol Genet. 29:R117-R124. https://doi.org/10.1093/hmg/ddaa169.

Dönertaş, H.M., Fabian, D.K., Valenzuela, M.F., Partridge, L., Thornton, J.M., 2021. Common genetic associations between age-related diseases. Nat Aging. 1, 400-412. https://doi.org/10.1038/s43587-021-00051-5.

Hajdarevic, R., Lande, A., Mehlsen, J., Rydland, A., Sosa, D.D., Strand, E.B., Mella, O., Pociot, F., Fluge, Ø., Lie, B.A., Viken, M.K., 2022. Genetic association study in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) identifies several potential risk loci. Brain Behav Immun. 102, 362-369. https://doi.org/10.1016/j.bbi.2022.03.010.

Institute of Medicine, 2015. Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness. Washington, DC: The National Academies Press, https://doi.org/10.17226/19012.

Mountjoy, E., Schmidt, E.M., Carmona, M., Schwartzentruber, J., Peat, G., Miranda, A., Fumis, L., Hayhurst, J., Buniello, A., Karim, M.A., Wright, D., Hercules, A., Papa, E., Fauman, E.B., Barrett, J.C., Todd, J.A., Ochoa, D., Dunham, I., Ghoussaini, M. 2021. An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. Nat Genet. 53, 1527-1533. https://doi.org/10.1038/s41588-021-00945-5.

Strausz, S., Blacker, G., Galloway, S., Hansen, P., Jones, S.E., Sanders, E., Sinnott-Armstrong, N., FinnGen, Weissman, I.L., Daly, M., Aivelo, T., Caspi Tal, M., Ollila, H.M. 2022. Secretoglobin family 1D member 2 (*SCGB1D2*) protein inhibits growth of *Borrelia burgdorferi* and affects susceptibility to Lyme disease. bioRxiv. https://doi.org/10.1101/2022.05.27.493784

Uffelmann, E., Huang, Q.Q., Munung, N.S., de Vries, J., Okada, Y., Martin, A.R., Martin, H.C., Lappalainen, T., Posthuma, D. 2021. Genome-wide association studies. Nat Rev Methods Primers 1, 59. https://doi.org/10.1038/s43586-021-00056-9.