Queensland University of Technology
School of Biomedical Sciences, 2 George St. Brisbane City QLD 4000

LQB684 - Advances in Medical Biotechnology

Submitted to:

Professor Dale Nyholt

Investigating the genetic relationship between Parkinson's disease and red hair colour

Submitted by:

Cæcilia Lind Skov-Jensen (n11368217)

October 30 2022

Abstract

Parkinson's disease is one of the most common neurodegenerative disorders and affects 1/1000 of the population at any time. Several studies have suggested a reciprocal association between melanoma and Parkinson's disease, however little is known about the underlying reason for this reciprocally increase in risk for both diseases. Studies suggest that individuals with red hair have a greater risk of melanoma than those with black hair, and it has likewise been shown that red hair colour is associated with a higher risk of Parkinson's disease. The genetic correlation and causal relationship between Parkinson's disease and red hair colour were investigated using linkage disequilibrium score regression (LDSC) and generalised summary-data-based Mendelian randomisation (GSMR), respectively. The LDSC analysis did not find evidence of a significant genetic correlation between Parkinson's disease and red hair. However, GSMR analysis suggested that Parkinson's disease had a negative causal effect on red hair. These findings seem to contradict the findings from the previous studies, and therefore it has been suggested what future studies should correct for when examining this.

Table of Contents

Introduction

Parkinson's disease is a brain disorder that results from damage to the nerve cells that produce the neurotransmitter dopamine, which is essential for the control of muscles and movements causing uncontrollable movements, such as shaking, stiffness, and difficulty with balance and coordination [1, 2]. Parkinson's disease is the second most common neurodegenerative disorder [3] that affects 1–2 individuals per 1000 of the population at any time. The prevalence of Parkinson's disease is increasing with age and affects approximately 1% of the population above 60 years [4].

A large number of studies have reported an association between Parkinson's disease and melanoma, where the occurrence of melanoma is higher than expected among subjects with Parkinson's disease, and the occurrence of Parkinson's disease is reciprocally higher than expected among patients with melanoma [5]. However, little is known about the underlying reason for this reciprocally increased risk for the two diseases [6].

It is currently known that melanoma is strongly tied to pigmentation, especially fair skin and red hair, which are phenotypes resulting from loss-of-function of the melanocortin-1 receptor gene (MC1R), the key pigmentation gene [7]. Studies suggest that individuals with red hair have an approximately threefold greater risk of melanoma than those with black hair [8, 9]. It is therefore interesting to further investigate whether the increased risk of melanoma among individuals with Parkinson's disease is related to an increased risk of Parkinson's disease depending on hair colour.

Several studies have been carried out to determine the relationship between Parkinson's disease and hair colour, and many have suggested that red hair colour and red hair-associated MC1R variants are associated with a higher risk of Parkinson's disease [10, 11], while other studies did not support this MC1R-Parkinson's disease link [12-14]. A meta-analysis from 2017 based on eight study cohorts from 2009 to 2016 that all focused on hair colour or red hair-associated MC1R variants, found evidence that red hair colour was associated with a significantly higher risk of Parkinson's disease, relative to black hair [6]. As these results were based on a meta-analysis of cohorts and no study yet suggests a genetic correlation between Parkinson's disease and red hair colour, this study will investigate precisely this.

Methods

Parkinson's disease GWAS

Parkinson's disease GWAS summary statistics were obtained from the Nalls et al. 2019 meta-analysis of Parkinson's disease risk [15] through the International Parkinson Disease Genomics Consortium (IPDGC) [16]. The meta-analysis comprised 56,306 Parkinson's disease cases (37,688 clinically diagnosed and 18,618 "Parkinson's disease-by-proxy") and 1.4 million controls of European ancestry, and a total of 7.8 million SNPs. The "Parkinson's disease-by-proxy" phenotype was defined by UK Biobank participants reporting whether they had a first-degree relative with Parkinson's disease. The analysed data set excluded the 23AndMe data, which gave a total sample size of 900,238. The downloading process for the Parkinson's disease GWAS data is further described in Appendix A, section 2.1.

Red hair GWAS

Red hair colour GWAS summary statistics were obtained from Morgan et al. 2018 [17], which performed a GWAS comparing individuals with self-reported red hair to a combined group of black- and brown-haired individuals. The study comprised 15,731 cases (red hair) and 283,920 controls (non-red hair), which gave a total sample size of 299,651. The downloading process for the red hair GWAS data is further described in Appendix A, section 2.2.

Data preprocessing

The GWAS summary statistics data were prepared and processed to have the correct information and layout for upload to the Complex-Traits Genetics Virtual Lab (CTG-VL) [18], as the analyses were performed on this site. For the upload, the files should be tab-separated and contain columns for chromosome (CHR), base pair position (BP), variant (SNP), effect allele (A1), non-effect allele (A2), effect allele frequency (FREQ), beta regression coefficient (BETA), standard error (SE), and P-value (P).

The wanted columns were extracted for the Parkinson's disease GWAS, and missing rsIDs were added using the 1000G reference file [19]. Duplicates were removed, and only the SNPs with rsIDs were kept. After this, a total of 13,038,881 SNPs were left in the Parkinson's disease GWAS and uploaded to the CTG-VL. Further details are described in Appendix A, section 3.1.

The red hair GWAS summary statistics were missing the BETA, SE, FREQ, and A2 columns that were necessary for upload and further analyses in the CTG-VL. The BETA column was added using the odds ratio (OR), as BETA = ln(OR). The standard error was calculated from the Z-statistic and BETA as SE = BETA / Z. The effect (alternate) allele frequency and the missing alleles were obtained from the variant manifest file at the Pan-UK Biobank [20]. Hereafter, the correct columns were extracted to upload to the CTG-VL. The file was successfully uploaded to the CTG-VL, however, the LDSC did not produce a significant heritability (h^2). It was, therefore, necessary to remove the genome-wide significant SNPs before the LDSC analysis as they had done in Morgan et al. (2018) [17]. All SNPs with a p < 5e-08 were removed from the file, and the file was thereafter successfully uploaded to CTG-VL. Further details are described in Appendix A, section 3.2.

Linkage disequilibrium score regression (LDSC)

LDSC was carried out using the LDSC tool in the CTG-VL to estimate SNP-based heritability (h^2_{SNP}) [21] in the GWAS dataset summary statistics. Hereafter, cross-trait LDSC was used to estimate the genetic correlation (r_G) [22] between Parkinson's disease and red hair and to estimate sample overlap. R-studio software [23] was then used to generate Z-scores, 95% confidence intervals, and P-values for the data sets.

Generalised Summary-data-based Mendelian Randomisation (GSMR)

GSMR was carried out using the GSMR tool in the CTG-VL to test for a putative causal association between Parkinson's disease and red hair [24]. The HEIDI-outlier filtering method was included in the test to detect and eliminate SNPs with a pleiotropic effect as these could bias the estimate and result in an inflated test statistic [24].

Results

SNP-based heritability

Manhattan plots of Parkinson's disease GWAS and red hair GWAS both showed significant SNPs. The Manhattan plot for the Parkinson's disease GWAS showed strong genome-wide significant (P < 5e-08) associations across the chromosomes (**Figure 1**) indicating that genetics play a significant role in Parkinson's disease. The strongest signal of association with Parkinson's disease was located at chromosome 4 (rs356203, P =

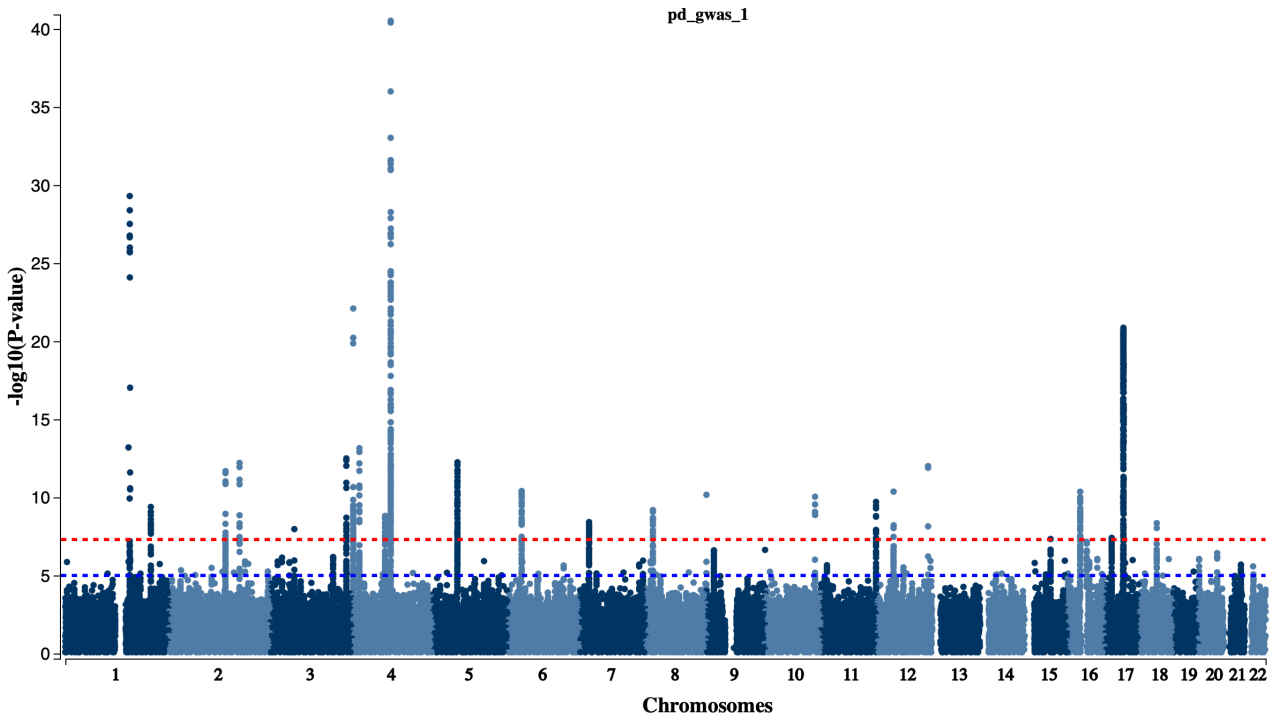


Figure 1: *Manhattan plot showing associations between SNPs and Parkinson's disease. The SNPs are plotted on the x-axis according to their positions on each chromosome against association with Parkinson's disease on the y-axis (-log$_{10}$ P-value). The red line indicates the genome-wide significant threshold of P = 5e-08. The blue line indicates the suggestive significant threshold of P = 1e-06.*

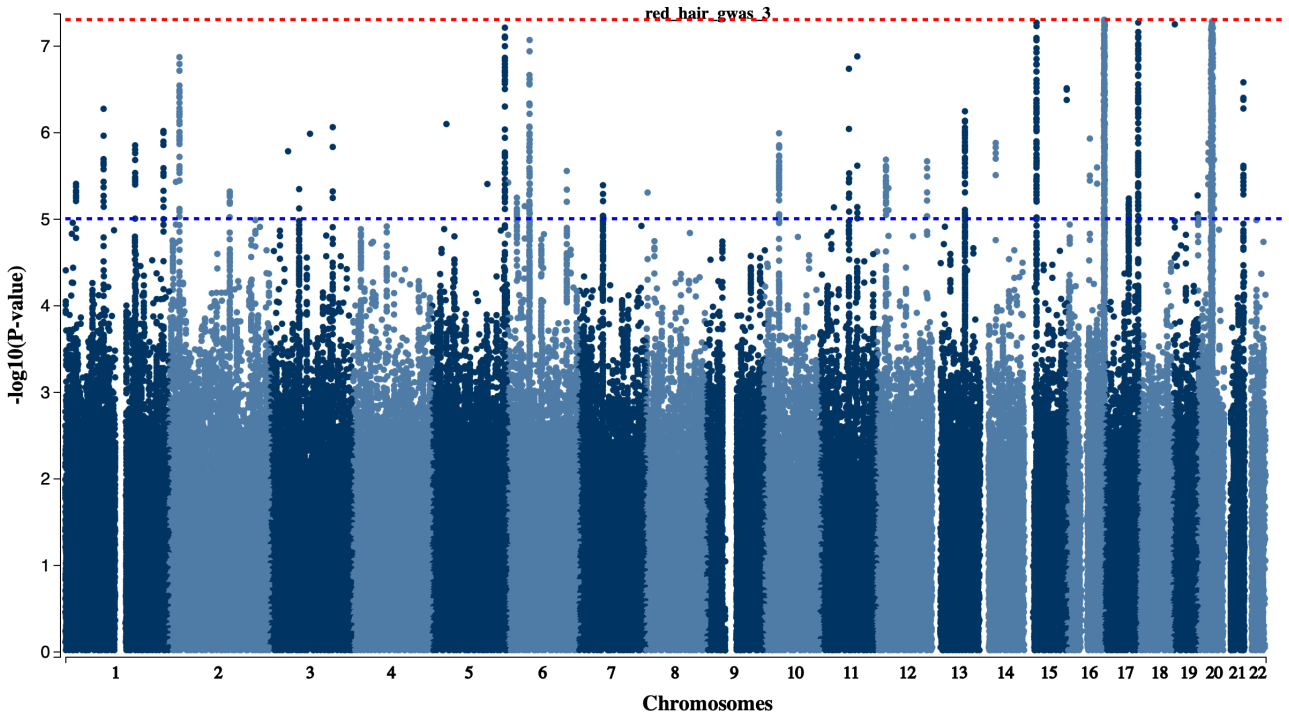


Figure 2: *Manhattan plot showing associations between SNPs and red hair. The SNPs are plotted on the x-axis according to their positions on each chromosome against association with red hair on the y-axis (-log$_{10}$ P-value). The red line indicates the genome-wide significant threshold of P = 5e-08. The blue line indicates the suggestive significant threshold of P = 1e-06.*

3.01e-41). Since all genome-wide significant SNPs were removed from the red hair GWAS data, the Manhattan plot did not show any genome-wide significant SNPs. There were however suggestive significant SNPs (P < 1e-05) across the chromosomes (**Figure 2**), which also suggested that genetics play a significant role in red hair colour.

The LD score regression analysis of SNP-based heritability gave highly significant observed SNP-based heritability (h^2) values for the Parkinson's disease and red hair GWAS datasets of 0.0085 (P = 2.28e-20) and 0.0108 (P = 2.66e-06), respectively (**Table 1**). When assuming a population prevalence of 5%, the converted liability scale heritability was estimated to be 0.0491 for Parkinson's disease and 0.0362 for red hair.

Phenotype	Observed h^2_{SNP}	SE	95% CI	Z-statistic	P-value
Parkinson's disease	0.0085	0.0008	0.0069 – 0.0101	10.625	2.28e-20
Red hair	0.0108	0.0023	0.0063 – 0.0153	4.6956	2.66e-06

Table 1: *Output on SNP-based heritability from LDSC analysis in CTG-VL.*

Genetic correlation

The cross-trait LD score regression analysis of genetic correlation showed no significant genetic correlation between Parkinson's disease and red hair, as the P-value was not below the threshold of 0.05 (r_G = 0.0040, P = 0.961) (**Table 2**). There was however no significant sample overlap between the two GWAS summary statistics (gcov_int = -0.0063, P = 0.226) (**Table 3**).

Phenotypes	r_G	SE	95% CI	Z-statistic	P-value
Parkinson's disease and red hair colour	0.0040	0.0814	-0.1555 – 0.1635	0.0491	9.61e-01

Table 2: *Output on genetic correlation from cross-trait LDSC analysis in CTG-VL.*

Phenotypes	gcov_int	SE	95% CI	Z-statistic	P-value
Parkinson's disease and red hair colour	-0.0063	0.0052	-0.0165 – 0.0039	1.2115	2.26e-01

Table 3: *Output on sample overlap from cross-trait LDSC analysis in CTG-VL.*

Causal relationship

GSMR could not be carried out as intended with red hair as exposure and Parkinson's disease as the outcome because of a lack of significant loci. However, the GSMR results for the reverse relationship indicated a significant negative effect value of -0.0486 (P = 0.0202) (**Table 4**), which indicated evidence of a negative causal effect of Parkinson's disease on red hair colour (OR = 0.8942, 95% CI -0.0895 – -0.0076). The HEIDI test for this analysis was not significant (P_{HEIDI} = 0.0715), which did not suggest the presence of SNPs with pleiotropic effects. The GSMR results are visualised in **Figure 3**.

Exposure	Outcome	Effect	SE	95% CI	P-value	# SNPs	HEIDI P
Parkinson's disease	Red hair	-0.0486	0.0209	-0.0895 – -0.0076	2.02e-02	14	7.15e-02

Table 4: *Output from GSMR analysis in CTG-VL.*

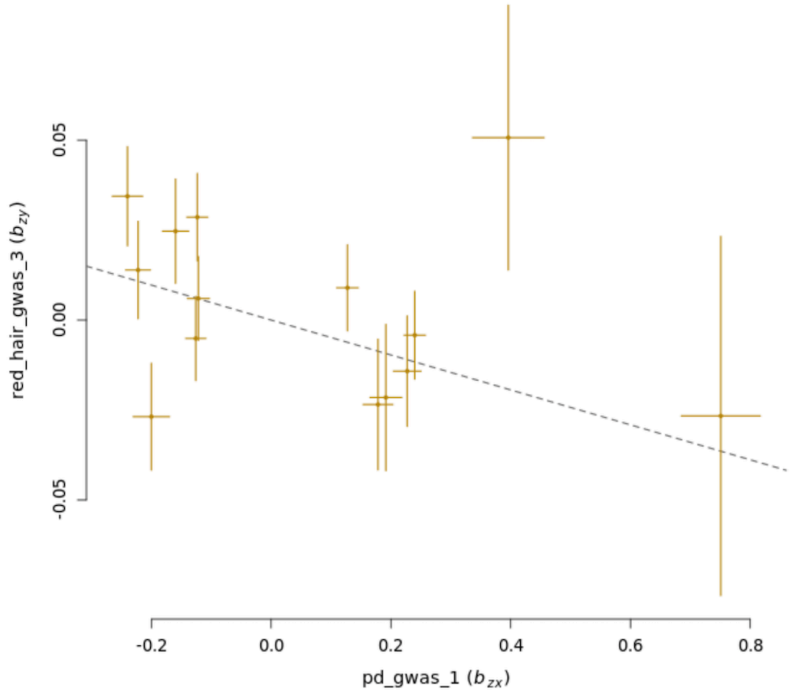


Figure 3: *Visualisation of GSMR results.*

Discussion

In this research project, the genetic relationship between Parkinson's disease and red hair colour was examined using statistical methods on GWAS datasets. No significant genetic correlation was found between Parkinson's disease and red hair colour. There was found significant evidence that Parkinson's disease had a negative causal effect on red hair colour (vs black and brown hair colour), but the causal relationship could not be tested in the other direction. The results from these analyses are not supported from any previous studies and therefore future studies with more powerful datasets are needed to provide robust evidence of this. The GSMR result that Parkinson's disease should have a negative causal effect on red hair does not intuitively make sense, as Parkinson's disease has a late onset whereas hair colour is determined at birth. Confounding could possibly be the explanation of this significant causal association. Furthermore, the removal of all the genome-wide significant SNPs from the red hair GWAS could influence the outcome from the genetic correlation. By correcting for this it might produce a significant result.

Acknowledgments

First of all, I would like to thank professor Dale Nyholt for his contribution to this research project. I would also like to acknowledge Queensland University of Technology and the School of Biomedical Science for organising this research project and enabling us students to challenge ourselves and expand our knowledge in a fun and independent way. I would like to send gratitude to the International Parkinson Disease Genomics Consortium as well as Ian J. Jackson for giving me access to the Parkinson's GWAS data and the red hair GWAS data, respectively. At last, I would like to thank the Complex-Traits Genetics Virtual Lab for their contribution to the analyses.

References

[1] H. Direct. "Parkinson's disease." https://www.healthdirect.gov.au/parkinsons-disease (accessed 26 Aug, 2022).

[2] N. I. o. A. (NIA). "Parkinson's Disease: Causes, Symptoms, and Treatments." https://www.nia.nih.gov/health/parkinsons-disease (accessed 26 Aug, 2022).

[3] O. Corti, S. Lesage, and A. Brice, "What genetics tells us about the causes and mechanisms of Parkinson's disease," (in eng), *Physiol Rev,* vol. 91, no. 4, pp. 1161-218, Oct 2011, doi: 10.1152/physrev.00022.2010.

[4] O. B. Tysnes and A. Storstein, "Epidemiology of Parkinson's disease," (in eng), *J Neural Transm (Vienna),* vol. 124, no. 8, pp. 901-905, Aug 2017, doi: 10.1007/s00702-017-1686-y.

[5] Q. Ye, Y. Wen, N. Al-Kuwari, and X. Chen, "Association Between Parkinson's Disease and Melanoma: Putting the Pieces Together," (in eng), *Front Aging Neurosci,* vol. 12, p. 60, 2020, doi: 10.3389/fnagi.2020.00060.

[6] X. Chen, D. Feng, M. A. Schwarzschild, and X. Gao, "Red hair, MC1R variants, and risk for Parkinson's disease – a meta-analysis," *Annals of Clinical and Translational Neurology,* vol. 4, no. 3, pp. 212-216, 2017, doi: https://doi.org/10.1002/acn3.381.

[7] E. M. Roider and D. E. Fisher, "Red Hair, Light Skin, and UV-Independent Risk for Melanoma Development in Humans," (in eng), *JAMA Dermatol,* vol. 152, no. 7, pp. 751-3, Jul 1 2016, doi: 10.1001/jamadermatol.2016.0524.

[8] S. Gandini *et al.*, "Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors," (in eng), *Eur J Cancer,* vol. 41, no. 14, pp. 2040-59, Sep 2005, doi: 10.1016/j.ejca.2005.03.034.

[9] J. Han, P. Kraft, G. A. Colditz, J. Wong, and D. J. Hunter, "Melanocortin 1 receptor variants and skin cancer risk," (in eng), *Int J Cancer,* vol. 119, no. 8, pp. 1976-84, Oct 15 2006, doi: 10.1002/ijc.22074.

[10] X. Gao, K. C. Simon, J. Han, M. A. Schwarzschild, and A. Ascherio, "Genetic determinants of hair color and Parkinson's disease risk," (in eng), *Ann Neurol,* vol. 65, no. 1, pp. 76-82, Jan 2009, doi: 10.1002/ana.21535.

[11] G. Tell-Marti *et al.*, "The MC1R melanoma risk variant p.R160W is associated with Parkinson disease," (in eng), *Ann Neurol,* vol. 77, no. 5, pp. 889-94, May 2015, doi: 10.1002/ana.24373.

[12] Z. Gan-Or *et al.*, "The role of the melanoma gene MC1R in Parkinson disease and REM sleep behavior disorder," (in eng), *Neurobiol Aging,* vol. 43, pp. 180.e7-180.e13, Jul 2016, doi: 10.1016/j.neurobiolaging.2016.03.029.

[13] O. Lorenzo-Betancor, Z. K. Wszolek, and O. A. Ross, "Rare variants in MC1R/TUBB3 exon 1 are not associated with Parkinson's disease," (in eng), *Ann Neurol,* vol. 79, no. 2, p. 331, Feb 2016, doi: 10.1002/ana.24581.

[14] S. J. Lubbe *et al.*, "Is the MC1R variant p.R160W associated with Parkinson's?," (in eng), *Ann Neurol,* vol. 79, no. 1, pp. 159-61, Jan 2016, doi: 10.1002/ana.24527.

[15] M. A. Nalls *et al.*, "Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies," (in eng), *Lancet Neurol,* vol. 18, no. 12, pp. 1091-1102, Dec 2019, doi: 10.1016/s1474-4422(19)30320-5.

[16] IPDGC. "IPDGC Resources." International Parkinson Disease Genomics Consortium. https://pdgenetics.org/resources (accessed 28 Aug, 2022).

[17] M. D. Morgan *et al.*, "Genome-wide study of hair colour in UK Biobank explains most of the SNP heritability," (in eng), *Nat Commun,* vol. 9, no. 1, p. 5271, Dec 10 2018, doi: 10.1038/s41467-018-07691-z.

[18] G. Cuellar-Partida, M. Lundberg, P. F. Kho, S. D'Urso, L. F. Gutierrez-Mondragon, and L.-D. Hwang, "Complex-Traits Genetics Virtual Lab: A community-driven web platform for post-GWAS analyses," *bioRxiv,* p. 518027, 2019, doi: 10.1101/518027.

[19] A. Auton *et al.*, "A global reference for human genetic variation," *Nature,* vol. 526, no. 7571, pp. 68-74, 2015/10/01 2015, doi: 10.1038/nature15393.

[20] "Variant manifest file." Pan-UKB team. https://pan.ukbb.broadinstitute.org (accessed Oct 28, 2022).

[21] B. K. Bulik-Sullivan *et al.*, "LD Score regression distinguishes confounding from polygenicity in genome-wide association studies," *Nature Genetics,* vol. 47, no. 3, pp. 291-295, 2015/03/01 2015, doi: 10.1038/ng.3211.

[22] B. Bulik-Sullivan *et al.*, "An atlas of genetic correlations across human diseases and traits," *Nature Genetics,* vol. 47, no. 11, pp. 1236-1241, 2015/11/01 2015, doi: 10.1038/ng.3406.

[23] *RStudio: Integrated Development for R*. (2020). RStudio Team. [Online]. Available: http://www.rstudio.com/

[24] Z. Zhu *et al.*, "Causal associations between risk factors and common diseases inferred from GWAS summary data," *Nature Communications,* vol. 9, no. 1, p. 224, 2018/01/15 2018, doi: 10.1038/s41467-017-02317-2.

Appendix A: Lab Notebook

The lab notebook is an HTML file that can be opened with the following link: http://rpubs.com/caeciliaskov/lab_notebook_OMICS_2022.

Emne: We've processed your Assignment extension request FORM-AEX-278580
Dato: torsdag den 27. oktober 2022 kl. 13.44.33 Østaustralsk normaltid
Fra: no-reply@qut.edu.au
Til: Caecilia Lind Skov-Jensen

Hi Caecilia Lind,

Thank you for your assignment extension request **(FORM-AEX-278580)**.

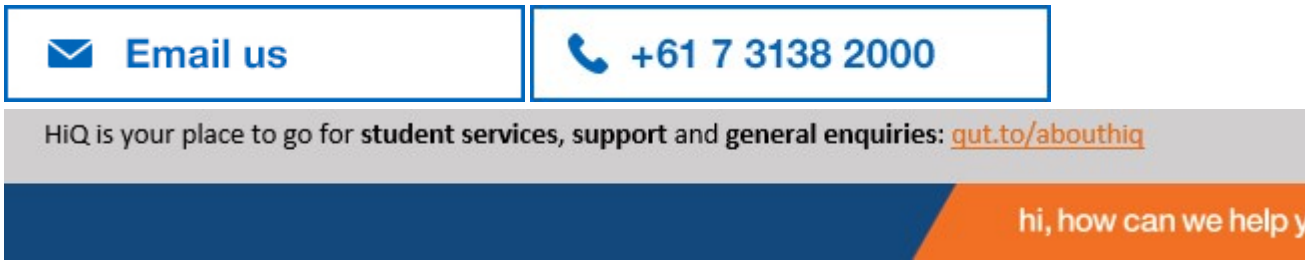
We have approved your request and the due date for your assignment **OMICS Research Report**, for unit LQB684 has been extended by 48 hours from the original due date. If your unit outline does not specify that your assignment is eligible for an extension, this confirmation email is not valid and unless you submit by the original due date, the late assessment policy will apply.

You are responsible for ensuring that this assignment is eligible for extension before submitting it after the original due date. Check your unit outline for eligibility.

Be aware that a copy of this email is kept on file. You should not alter this email in any way. Email notifications that have been altered or differ in any way from the original may result in an allegation of student misconduct as set out in the Student Code of Conduct.

Need extra support? You can access free, confidential counselling with qualified professionals. We also offer planning and support if you have a disability, injury or health condition that affects your ability to study. If you are struggling to meet your university commitments, academic support is also available.

Have a question? You can contact us by email or phone. We're also available to help you on campus or via online chat. Visit the HiQ website for our contact details and opening hours.



You have received this email because you have submitted an assignment extension request. View our Privacy and Security statement.

Ref ID: 11368217 FORM-AEX-278580