RESEARCH PROJECT PROPOSAL FORM - BIOL5352M

1. Details of Principal Applicant

Surname	Letton
Forename	Christopher
Title	Mr
Position Held	Postgraduate Student
Telephone	+44 7505032963
E-mail	Jsxm6270@leeds.ac.uk
Main contact address if different from company address	4 Glen Road, Leeds LS16 5NJ

2. Project title (not exceeding 120 characters)*

Gene Burden Analysis of Rare Germline Variants in Autoinflammatory and Immunodeficiency Genes Implicated in Giant Cell Arteritis

3. Abstract of proposed research project (not exceeding 750 characters)*

This project investigates the role of rare germline variants in the pathogenesis of giant cell arteritis (GCA), a large-vessel vasculitis of ageing associated with significant morbidity. Using gene burden analysis, the study will assess the enrichment of rare protein-altering variants in GCA patients compared to UK Biobank controls. Particular focus will be placed on genes from the Genomics England R413.1 autoinflammatory and R15 immunodeficiency panels, which catalogue immune-related monogenic disorders. Functional annotation, pathway and enrichment analysis, and therapeutic relevance assessment will be performed. Findings will be integrated with existing GWAS knowledge to uncover novel genetic mechanisms, identify potential therapeutic targets, and advance understanding of GCA pathogenesis.

4. Project length

Duration of Project (months)	**3.5**	Proposed Start Date	**5/05/2025**

5. Description of research project (in **no more than 6 A4 pages** describe the proposed research project providing a case for support. Font size Ariel 12pt or higher, 2 cm margins)

Aims and objectives

This project aims to investigate the contribution of rare germline variants to the risk and pathogenesis of GCA, focusing particularly on genes implicated in autoinflammatory processes. Through the use of gene burden analysis, this study will identify genes in which rare protein-altering variants are enriched in GCA patients compared to controls from the UK Biobank, excluding individuals with GCA or other vascular and rheumatological diseases. There will be a focus on genes included in the Genomics England PanelApp for autoinflammatory disorders (clinical indication R413.1) and the Primary Immunodeficiency or Monogenic Inflammatory

Bowel Disease panel (clinical indication R15), both of which contain genes involved in innate immunity and systemic inflammation.

The objectives are:

1) To conduct a gene-level burden analysis of rare germline variants in GCA cases compared to population controls, identifying genes with significant enrichment in cases.

2) To determine whether genes showing significant rare variant burden overlap with those listed in the Genomics England R413.1 and R15 panels.

3) To interpret the biological relevance of implicated genes, with a focus on immune, inflammatory, and vascular pathways involved in GCA pathogenesis.

4) To explore whether rare variants provide additional associations with GCA risk and pathogenesis beyond those identified by common variant studies.

5) To evaluate whether implicated genes highlight potential therapeutic targets or pathways for future clinical research.

This study addresses a critical gap in our understanding of GCA genetics. While common variant studies have highlighted key risk loci, they leave much of the heritable risk unexplained. Rare variants can exert strong effects and may help explain this missing heritability. This investigation is particularly relevant given increasing evidence that GCA shares immunological features with monogenic autoinflammatory syndromes such as Familial Mediterranean Fever (MEFV mutations). By focusing on biologically relevant gene sets and applying robust analytic methods, this project has the potential to uncover new disease mechanisms, refine our understanding of GCA pathogenesis, and support future efforts toward personalised medicine.

The use of UK Biobank data offers a robust control cohort, ensuring adequate power for statistical comparison. This project is positioned to possibly help guide future functional studies or personalised treatment approaches for GCA.

Background

Giant cell arteritis (GCA) is a granulomatous large-vessel vasculitis that predominantly affects older adults, with an incidence of approximately 20 per 100,000 persons over age 50 (Salvarani et al., 2008). Patients commonly present with symptoms such as new-onset headache, scalp tenderness, jaw claudication, visual disturbances, and constitutional symptoms including fatigue and weight loss. If not promptly diagnosed and treated, GCA can result in life-altering complications such as permanent vision loss, stroke, and aortic aneurysm (Kermani et al., 2013) (Gonzalez-Gay et al., 2005). GCA's clinical heterogeneity and immunopathogenesis indicate a complex interaction of genetic predisposition and immune dysregulation. The disease is characterised by infiltration of activated T cells and macrophages into

the vessel wall, resulting in intimal hyperplasia and luminal occlusion (Weyand & Goronzy, 2013).

GCA is known to have a genetic component, with familial clustering and strong associations with the HLA class II region (Carmona et al., 2015). GWAS have further identified associations with non-HLA genes, including PTPN22 (involved in T cell signalling), PLG (plasminogen, angiogenesis), IL6, REL, and LRRC32 (Serrano et al., 2013). However, these common variants collectively explain only a portion of disease susceptibility, estimated to account for around 13–15% of GCA's heritability. A recent large-scale GWAS corroborated this view, estimating the total SNP-based heritability of GCA at just 15.1%, with HLA-specific contributions accounting for only 1.7% (Borrego-Yaniz et al., 2024). This leaves a large proportion of heritability unexplained, reinforcing the need to investigate other types of genetic variation, particularly rare variants with larger individual effect sizes.

In addition, the study identified three novel non-HLA loci: MFGE8, VTN, and CCDC25, involved in apoptotic cell clearance, vascular homeostasis, and neutrophil extracellular trap (NET) signalling, respectively. These findings broaden the range of biological processes implicated in GCA beyond traditional adaptive immunity, highlighting the importance of innate immune responses, vascular remodelling, and neutrophil-mediated inflammation (Borrego-Yaniz et al., 2024).

In GCA's closest clinical relative, polymyalgia rheumatica (PMR), a recent study found a significantly higher burden of rare damaging variants in inflammasome-related genes such as NLRP12, supporting an autoinflammatory component (Higuchi et al., 2024). Given the high clinical overlap between PMR and GCA (up to 50% of GCA cases present with PMR symptoms), similar rare variant patterns in GCA are plausible but remain largely unexplored.

The 2024 GWAS also found that many associated variants exerted regulatory effects in immune cells such as $CD4^+$/$CD8^+$ T cells, B cells, monocytes, and especially NK cells, which may act as key initiators of vessel wall inflammation in GCA (Borrego-Yaniz et al., 2024). This reinforces the idea that rare variants affecting these pathways, though not detected by GWAS, could still exert large effects through key immunological nodes.

Given the growing evidence that rare variants contribute significantly to inflammatory diseases, a targeted approach focusing on key immunological pathways is warranted.

The Genomics England PanelApp provides curated panels of genes with established relevance to specific clinical disorders. Of particular interest to this project are the "Autoinflammatory disorders" (R413.1) and the "Primary immunodeficiency or monogenic inflammatory bowel disease" (R15) panels. These panels collectively catalogue genes implicated in monogenic diseases of innate immune dysregulation, many of which overlap mechanistically with processes active in GCA, such as

cytokine signalling, inflammasome activation, and neutrophil function. Investigating rare variants within these genes offers a targeted approach to identify pathways that may contribute to GCA susceptibility, particularly given the increasing recognition of autoinflammatory components within its pathogenesis. By focusing on these panels, this study leverages existing clinical and genomic knowledge to prioritise biologically meaningful candidate genes for rare variant burden analysis.

No large-scale studies have been conducted that have systematically assessed the role of rare germline variants in GCA, particularly in genes implicated in innate immunity and autoinflammatory disease. This project fills that gap by applying gene burden analysis to identify genes in which rare functional variants are significantly enriched in GCA cases compared to controls. By integrating high-quality GCA case data with UK Biobank controls, and by employing functional annotation and pathway analysis, this project will provide novel insights into the genetic and biological basis of GCA. If successful, it could lead to improved understanding of disease mechanisms, genetic risk stratification, and ultimately, more personalised approaches to diagnosis and therapy.

Research Plan

This project will use gene burden analysis to investigate the role of rare germline variants in GCA, focusing on autoinflammatory genes. This will be completed in sections:

Gene-Level Burden Analysis:

Rare, protein-altering variants (nonsynonymous, splice-site, frameshift, or stop-gain) with a minor allele frequency (MAF) <1% will be extracted from whole-exome sequencing data of GCA cases and UK Biobank controls. Variants will be annotated using the Ensembl Variant Effect Predictor (VEP) and filtered based on functional prediction scores (e.g., REVEL, CADD, SIFT, PolyPhen).

Primary analysis will apply a gene-level burden test by collapsing rare variants within each gene: individuals will be classified as carriers or non-carriers of at least one qualifying variant. Association between gene burden and GCA status will be tested using SKAT-O (Sequence Kernel Association Test-Optimal), adjusting for sex and the first 10 ancestry principal components. Firth's penalised logistic regression will be used for sparse data.

Secondary analyses will use ACAT-V to combine p-values across variant categories or filtering strategies. Statistical significance will be adjusted for multiple testing using FDR and/or Bonferroni correction. Sensitivity analyses will explore alternative allele frequency thresholds (e.g., MAF <0.5%) and stricter variant filters (e.g. restricting to predicted loss-of-function variants).

Targeted Testing of Autoinflammatory and Immunodeficiency Genes:

Genes from the Genomics England Autoinflammatory Disorders panel (R413.1) and the Primary Immunodeficiency or Monogenic Inflammatory Bowel Disease panel (R15) will undergo focused burden testing. Multiple testing correction will be applied appropriately, and a cumulative burden across each panel will be evaluated to identify enrichment patterns in GCA cases compared to controls.

Functional Interpretation and Pathway Analysis:

Significant genes and variants will be annotated using tools such as SIFT, PolyPhen, and ClinVar. Pathway analysis will assess biological relevance, with a focus on cytokine signalling, innate immunity, and inflammasome regulation.

Integration with Existing Genetic Knowledge:

Genes identified through rare variant burden analysis will be compared to genes highlighted in previous GCA genome-wide association studies (GWAS) to assess overlap and novelty. This comparison will help determine whether rare variants contribute to disease risk through the same genes implicated by common variants, or identify novel genes not detected by GWAS. Where possible, findings will be interpreted in relation to clinical subtypes and therapeutic targets.

Evaluation of Therapeutic Potential:

Genes identified through rare variant burden analysis will be assessed for therapeutic relevance by reviewing their biological functions and consulting drug-target databases such as DrugBank and Open Targets. Genes already targeted by existing therapies, or involved in druggable pathways (e.g., cytokine signalling, inflammasome activation), will be highlighted. Findings will be discussed in the context of current GCA treatments and potential opportunities for therapeutic repurposing.

Research Impact

Who Will Benefit:

Researchers in Genetics and Immunology:
This study will provide novel insights into the genetic architecture of GCA, particularly the underexplored contribution of rare germline variants. It will also inform the broader fields of vasculitis, autoimmunity, and autoinflammation.

Clinicians:
Findings may help explain differences in disease onset, severity, or treatment response in GCA patients, potentially informing risk stratification and personalised medicine approaches.

Patients with GCA:
Improved understanding of the genetic underpinnings of GCA may eventually lead to earlier diagnosis, better-targeted treatments, and inclusion of genetic screening in clinical care.

How They Will Benefit:

New Genetic Insights: Identification of rare variant burden in key inflammatory genes could offer mechanistic explanations for disease development not captured by common variant studies.

Translational Potential: Genes or pathways identified may highlight new therapeutic targets or suggest existing treatments could be repurposed for specific patient subsets.

Framework for future research: For rare variant burden analysis in other inflammatory diseases

Engagement and Communication Methods:

Internal Engagement:
Regular updates and discussions with academic supervisors and collaborators.

Academic Dissemination:
Submission of results to peer-reviewed journals.

Ethical Considerations

This project involves the analysis of existing human genetic and clinical data from two separate sources.

GCA Case Data:

Genetic data for GCA cases was generated through a collaboration with Regeneron and has been shared under specific agreements. Use of this data is subject to the ethical standards, data protection regulations, and access conditions set by Regeneron and the collaborating institutions. All analyses will be conducted in accordance with these agreements and in compliance with applicable data governance frameworks.

UK Biobank Control Data:

Control data will be obtained from the UK Biobank, under a registered institutional application and within the scope of participant consent for health-related research. Data handling will comply with the General Data Protection Regulation (GDPR) and UK Biobank's Access and Ethics Policies.

In both cases, all data will be pseudonymised or anonymised at source, securely stored, and used exclusively for research purposes. No new data collection or

participant contact is planned. Individual-level results will not be made publicly available, and findings will be presented in aggregate form only.

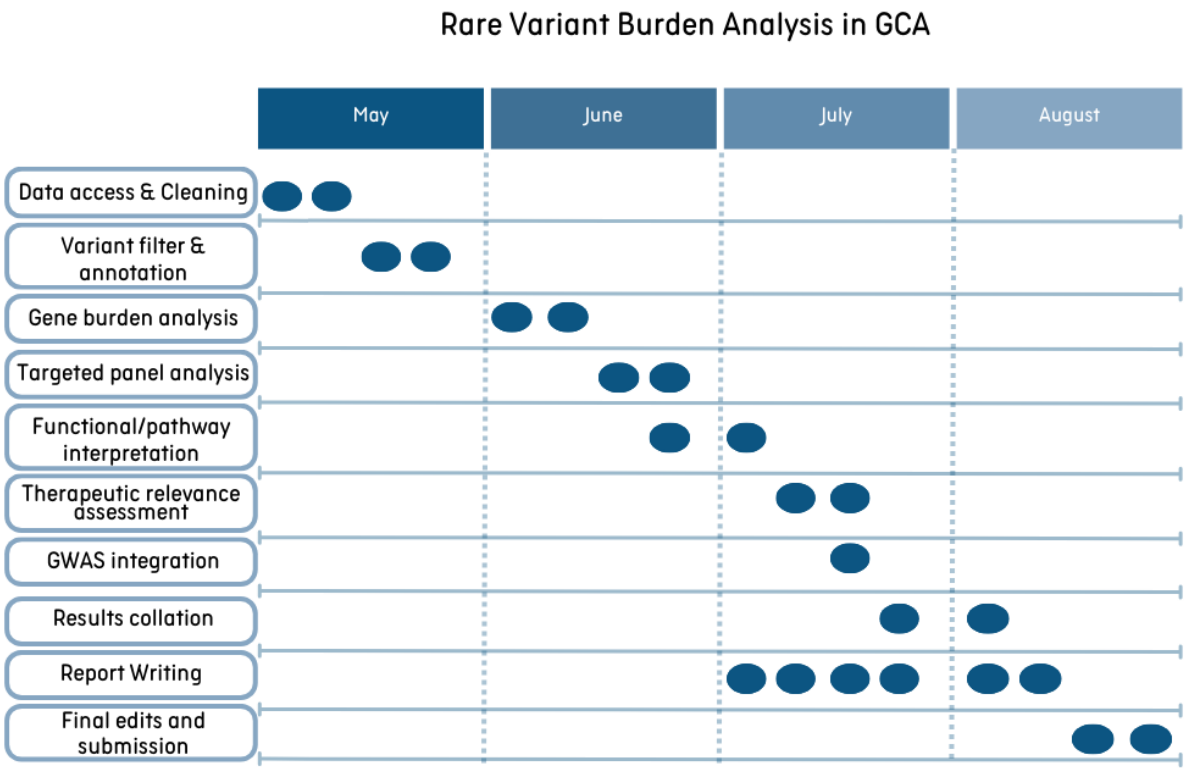
Data Sources and Consent:

GCA case data will be obtained from primary patient samples recruited at Leeds Teaching Hospitals NHS Trust (LTHT), collected under appropriate ethical approvals and with informed consent for genetic research and data sharing. Control data will be drawn from the UK Biobank, where participants have provided broad consent for the use of anonymised health and genetic data in approved research studies.

Ethical Approvals:

As no new data will be collected and no participant contact is required, existing ethics approvals and data access agreements are sufficient.

Project Plan



The key parts of the project plan are briefly explained covering the activity and actions:

- Data access and Cleaning:
  - Gain access to GCA case data and UK biobank controls, standardize datasets.

- Variant filter and annotation:
  - Extract rare protein-altering variants, annotate using VEP and prediction tools.
- Gene burden analysis:
  - Conduct SKAT-O and Firth logistic regression tests to identify gene-level rare variant associations.
- Targeted panel analysis:
  - Focus burden analysis on genes from the R413.1 and R15 panels
- Functional/pathway interpretation:
  - Review implicated genes for biological relevance using pathway analysis tools
- Therapeutic relevance assessment:
  - Investigate possible drug interventions of implicated genes using DrugBank and Open Targets databases.
- GWAS integration:
  - Compare identified genes with published GWAS loci for GCA.
- Results collation:
  - Generate summary tables, figures and gather key findings.
- Report Writing:
  - Prepare written reports including methods, results, discussion and conclusion.
- Final edits and submission:
  - Complete final edits, proofreading and submission of report.

Signed:  **Date: 27/04/2025**

Bibliography

Borrego-Yaniz, G., Ortiz-Fernández, L., Madrid-Paredes, A., Kerick, M., Hernández-Rodríguez, J., Mackie, S. L., Vaglio, A., Castañeda, S., Solans, R., Mestre-Torres, J., Khalidi, N., Langford, C. A., Ytterberg, S., Beretta, L., Govoni, M., Emmi, G., Cimmino, M. A., Witte, T., Neumann, T., … Weisman, M. (2024). Risk loci involved in giant cell arteritis susceptibility: a genome-wide association study. *The Lancet Rheumatology*, *6*(6), e374–e383. https://doi.org/https://doi.org/10.1016/S2665-9913(24)00064-X

Carmona, F. D., Martín, J., & González-Gay, M. A. (2015). Genetics of vasculitis. *Current Opinion in Rheumatology*, *27*(1). https://journals.lww.com/co-rheumatology/fulltext/2015/01000/genetics_of_vasculitis.3.aspx

Gonzalez-Gay, M. A., Barros, S., Lopez-Diaz, M. J., Garcia-Porrua, C., Sanchez-Andrade, A., & Llorca, J. (2005). Giant Cell Arteritis: Disease Patterns of Clinical Presentation in a Series of 240 Patients. *Medicine*, *84*(5). https://journals.lww.com/md-journal/fulltext/2005/09000/giant_cell_arteritis__disease_patterns_of_clinical.2.aspx

Higuchi, T., Oka, S., Furukawa, H., & Tohma, S. (2024). The contributions of deleterious rare alleles in NLRP12 and inflammasome-related genes to polymyalgia rheumatica. *Scientific Reports*, *14*(1), 490. https://doi.org/10.1038/s41598-024-51320-3

Kermani, T. A., Warrington, K. J., Crowson, C. S., Ytterberg, S. R., Hunder, G. G., Gabriel, S. E., & Matteson, E. L. (2013). Large-vessel involvement in giant cell arteritis: a population-based cohort study of the incidence-trends and prognosis. *Annals of the Rheumatic Diseases*, *72*(12), 1989–1994. https://doi.org/10.1136/annrheumdis-2012-202408

Salvarani, C., Cantini, F., & Hunder, G. G. (2008). Polymyalgia rheumatica and giant-cell arteritis. *The Lancet*, *372*(9634), 234–245. https://doi.org/10.1016/S0140-6736(08)61077-6

Serrano, A., Márquez, A., Mackie, S. L., Carmona, F. D., Solans, R., Miranda-Filloy, J. A., Hernández-Rodríguez, J., Cid, M. C., Castañeda, S., Morado, I. C., Narváez, J., Blanco, R., Sopeña, B., García-Villanueva, M. J., Monfort, J., Ortego-Centeno, N., Unzurrunzaga, A., Marí-Alfonso, B., Sánchez-Martín, J., … Martín, J. (2013). Identification of the PTPN22 functional variant R620W as susceptibility genetic factor for giant cell arteritis. *Annals of the Rheumatic Diseases*, *72*(11), 1882–1886. https://doi.org/10.1136/annrheumdis-2013-203641

Weyand, C. M., & Goronzy, J. J. (2013). Immune mechanisms in medium and large-vessel vasculitis. *Nature Reviews Rheumatology*, *9*(12), 731–740. https://doi.org/10.1038/nrrheum.2013.161