UNIVERSITY OF LEEDS

FACULTY OF BIOLOGICAL SCIENCES

ELECTRONIC Coursework Coversheet

For use with *individual* assessed work

Student Identification Number:	2	0	1	9	0	8	1	0	9

Module Code: BIOL5352M

Module Title: Res Proj: Genomics and Analytics

Assessment title: Dissertation

Name of marker for this work:

Statement of Intent

By submitting this work I am agreeing to the University Statement on Academic Integrity and confirm that this is all my own work.

Word count / Page count	**5264 / 23**

Please note that failure to state word/page count, stating and incorrect word/page count or falsifying of word count will result in application of penalties as per the FBS Code of Practice on Assessment

UNIVERSITY OF LEEDS

Exome-Wide Rare Variant Burden Testing of Cranial Ischaemic Complications in Giant Cell Arteritis

Christopher Letton

The University of Leeds

Faculty of Biological Sciences

School of Molecular and Cellular Biology

August 2025

Acknowledgements

I would first like to acknowledge the patients who generously contributed their data, without whom this research would not have been possible.

I am deeply grateful to my supervisors, Dr Mark Iles, Professor Ann Morgan and Dr Zak Thornton, for their guidance, support, and encouragement throughout this project. Their expertise and constructive feedback have been invaluable in shaping both the research and my development as a scientist.

I would also like to thank the University of Leeds and the Leeds Institute for Data Analytics (LIDA) for providing the facilities, resources, and academic environment that enabled this work.

On a more personal note, I am thankful to my friends and family for their patience and encouragement during my degree. In particular, I wish to thank William Edwards for his advice and support.

Finally, I owe a special thanks to my family, for their support that made this journey possible.

Abstract

Giant cell arteritis (GCA) is a systemic vasculitis in which cranial ischaemic complications (CICs) such as vision loss and stroke represent the most severe outcomes. While common variants have been implicated in GCA susceptibility, the role of rare coding variants in CICs remains unclear. This study investigated whether rare germline variants contribute to the risk of CICs in patients with GCA.

Whole-exome sequencing data from 1,497 clinically characterised GCA patients was analysed using gene-based rare variant burden testing. Patients were stratified into those without CICs, those with transient events, and those with permanent complications. Analyses were conducted with the SKAT-O method under three functional masks (predicted loss-of-function (pLoF), pLoF+missense, and pLoF+missense+synonymous), adjusting for sex and the first ten genomic principal components.

In the primary comparison of patients with no CICs versus those with permanent CICs, 17,472 genes were tested under the pLoF+missense mask. No associations remained significant after multiple-testing correction. Several genes, including *XDH*, *HYAL2*, *EYA1*, and *VPS13D*, showed the lowest nominal p-values, mapping to processes such as oxidative stress, angiogenesis, and extracellular matrix turnover.

No rare variants of large effect were identified, suggesting that CICs in GCA are unlikely to be driven by single high-impact alleles. Instead, multiple small contributions across immune, vascular, and metabolic pathways are more probable. These results highlight the need for larger collaborative studies and integration with functional and clinical data to clarify the genetic basis of ischaemic complications in GCA.

Introduction

Giant Cell Arteritis: Clinical and Pathological Overview

Giant cell arteritis (GCA) is the most common systemic vasculitis in adults over 50, with an incidence of approximately 20 per 100,000 persons in this age group(Salvarani et al., 2008). It predominantly affects medium and large arteries, particularly the extracranial branches of the carotid artery, leading to clinical features such as headache, scalp tenderness, jaw claudication, and visual disturbances (Gonzalez-Gay et al., 2005; Kermani et al., 2013). Severe complications include cranial ischemic events (CICs), such as transient or permanent vision loss and stroke, which contribute significantly to morbidity (Lyons et al., 2020). While glucocorticoid therapy is highly effective for controlling systemic inflammation, cranial ischemic complications remain a substantial clinical challenge, often occurring despite treatment initiation (Hellmich et al., 2020). Understanding the mechanisms predisposing to these complications is critical for improving patient outcomes.

Cranial Ischemic Complications in GCA

Cranial ischemic complications represent some of the most severe outcomes of GCA, with vision loss being one of the most devastating consequences. Permanent visual impairment occurs in approximately 15–20% of patients at diagnosis, while transient visual symptoms are

reported in up to one-third of cases (Donaldson and Margolin, 2022). These complications are driven by inflammatory occlusion of the posterior ciliary arteries and other branches of the ophthalmic artery, resulting in anterior ischemic optic neuropathy, central retinal artery occlusion, or cerebral ischemia (Patil et al., 2013; Lazaar et al., 2025). Despite prompt glucocorticoid therapy, patients who develop permanent vision loss rarely regain visual function, highlighting the urgency of prevention (Hellmich et al., 2020).

Recent large-scale cohort analysis has identified specific clinical predictors of CICs. Older age and pre-existing hypertension were associated with an increased risk of ischemic complications at presentation, while anticoagulant therapy appeared to be protective. Positional mapping of polygenic risk scores for cardiovascular traits also implicated immune- and coagulation-related pathways, including signals at the TEK, CD96 and MROH9 loci (Chaddock et al., 2025). These findings highlight the heterogeneity underlying CIC development, but clinical predictors alone cannot fully explain inter-individual variation, raising the possibility that germline genetic variation contributes to susceptibility.

Genetic Contributions to GCA

GCA is known to have a strong genetic component, with familial clustering and a well-established association with the HLA class II region (Carmona et al., 2015). Genome-wide association studies (GWAS) have identified additional non-HLA loci, including *PTPN22*, *PLG*, and *MFGE8*, implicating adaptive and innate immune pathways in disease pathogenesis (Borrego-Yaniz et al., 2024). However, these common variants collectively explain only a small proportion of the heritable risk (estimated at ~15%), leaving much of the genetic architecture unresolved (Borrego-Yaniz et al., 2024). Importantly, existing GWAS have not addressed the genetic basis of phenotypic heterogeneity in GCA, such as why some patients develop severe cranial ischemic complications while others do not.

Rare Variants and Phenotypic Variation

Rare variants, particularly predicted loss-of-function (pLoF) alleles, often exert larger biological effects than common variants and can provide mechanistic insights into disease pathogenesis (Momozawa and Mizukami, 2021). In related inflammatory conditions, such as polymyalgia rheumatica, rare damaging variants in immune-related genes have been linked to disease susceptibility (Higuchi et al., 2024). However, little is known about the contribution of rare germline variants to GCA, and no previous studies have systematically investigated their role in the development of cranial ischemic complications. This represents a critical gap in our understanding of disease heterogeneity and may help explain why certain patients experience more severe vascular outcomes.

Study Aim

This study aimed to investigate whether rare germline variants are associated with the presence and severity of cranial ischemic complications in GCA patients.

1. Analyse rare variants across three functional categories: predicted loss-of-function (pLoF), missense, and synonymous variants.

2. Compare and evaluate the burden of these variants between GCA patients with no CICs, transient CICs, and permanent CICs.

3. Perform pathway enrichment analysis of any genes reaching statistical significance, to evaluate whether implicated variants have shared biological pathways relevant to immune or vascular processes.

Methods

Study Cohort and clinical stratification

Whole-exome sequencing (WES) data from 1,512 patients with biopsy-proven or clinically confirmed giant cell arteritis (GCA) were obtained via the UK GCA Consortium, with data processing and analysis conducted at the University of Leeds within the Leeds Institute of Cardiovascular and Metabolic Medicine (LICAMM) and the Leeds Institute for Data Analytics (LIDA). Sequencing was performed by Regeneron Pharmaceuticals, and project-level files containing all "QC pass" samples were provided. Patients were diagnosed with GCA primarily using the American College of Rheumatology (ACR) classification criteria, with diagnosis supported in some cases by confirmatory testing such as temporal artery biopsy (TAB) or vascular imaging. Twelve samples were excluded due to failing quality control, and three patients with unknown cranial ischemic complication (CIC) status were also excluded, leaving a final cohort of 1,497 patients for analysis.

The final dataset included 1,497 patients, of which 464 were male and 1,033 were female. Principal component analysis (PCA), performed using reference data from the 1000 Genomes Project, estimated that 99.2% of the cohort were of European ancestry (Chaddock et al., 2025). Cohort characteristics stratified by CIC status are summarised in Table 1.

Table 1. Cohort characteristics stratified by cranial ischemic complication (CIC) status. Values are counts, with percentages in parentheses.

CIC Group	*Number of Patients*	*Males n (%)*	*Females n (%)*
No CICs	568	181 (31.9%)	387 (68.1%)
Transient CICs	677	200 (29.5%)	477 (70.5%)
Permanent CICs	252	83 (33.0%)	169 (67.0%)

Patients were stratified into three groups based on the presence and severity of cranial ischemic complications at diagnosis:

- No CICs (n = 568): Patients with no reported ischemic complications.
- Transient CICs (n = 677): Patients with temporary visual or ischemic symptoms without permanent deficits.
- Permanent CICs (n = 252): Patients with irreversible vision loss or other permanent ischemic damage.

Variant Processing and Quality Control

All samples that passed quality control were used in the analysis. Variants were decomposed (one variant per line) and normalised prior to downstream processing. The Goldilocks filter pipeline was applied to the genotype-level data set, which excludes low-confidence calls in two stages. First, depth-based filtering was applied: SNP calls with a read depth (DP) <7 and INDEL calls with DP <10 were set to "no-call". Second, allele balance (AB) filters were applied: SNP sites required at least one alternate allele with AB ≥15%, and INDEL sites required at least one alternate allele with AB ≥20%. Sites failing these thresholds were removed from the data set.

Variant Annotation and Categorisation

Variants were annotated using the Ensembl Variant Effect Predictor (VEP; version 114) with the GRCh38 reference genome build. VEP was configured with the annotation distance set to 0 kb in order to avoid the default behaviour of reporting variants located within 5 kb of a gene as upstream or downstream consequences. Restricting this distance prevented the inclusion of variants outside coding regions being misassigned to nearby genes.

To restrict analyses to protein-coding regions, the Ensembl release 114 GTF file was used to extract the Ensembl gene identifiers (ENSG IDs) corresponding to protein-coding genes. This approach was used because filtering by consequence terms alone was not a reliable proxy for identifying protein-coding loci. In the case of filtering using VEP annotation it can inadvertently include read-through or antisense transcripts, retained intron isoforms of protein-coding genes, or pseudogenes that may still receive annotations such as "missense_variant" or "synonymous_variant." By filtering directly against a curated list of protein-coding ENSG IDs, non-coding loci were excluded, and only genuine protein-coding variants were retained for downstream analysis.

Whilst consequence terms were not used to define protein-coding status, they were used to generate the functional masks for analysis. Sets of consequence terms were grouped according to Sequence Ontology definitions to classify variants into predicted loss-of-function, missense, or synonymous categories.

Variants were grouped into three functional categories ("masks"):

- Predicted loss-of-function (pLoF)**:** Stop-gain, stop-loss, start-loss, frameshift, splice acceptor, and splice donor variants.
- Missense: Non-synonymous single nucleotide variants.
- Synonymous: Variants not altering amino acid sequence.

Grouping variants into functional "masks" is a widely used approach in rare variant burden testing because it enriches gene-level tests for variants most likely to affect function. Individually rare variants have little power, but collapsing them by predicted consequence (e.g. pLoF, missense) increases sensitivity and reduces noise from neutral variants (Vali-Pour et al., 2022; Zhou et al., 2022). Using different masks also helps interpret associations: for example, a signal in the pLoF mask suggests loss-of-function mechanisms, while inclusion of missense variants can detect protein-altering effects. Synonymous variants were retained to confirm that associations were not driven by artefactual signals.

Variants with a minor allele frequency (MAF) <1%, calculated from the cohort using PLINK (version 1.9), were included in the burden analysis.

Three comparisons were performed:

- No CICs vs transient + permanent CICs
- No CICs vs permanent CICs
- No CICs + transient CICs vs permanent CICs

Burden Analysis

Gene-level burden analyses were conducted by collapsing rare variants within each functional mask per gene set, as defined by VEP annotation. Variants were assigned to genes using VEP, and individuals were then classified as carriers or non-carriers of at least one qualifying variant within each gene set. Analyses were performed in R (version 4.3) using the Sequence Kernel Association Test package (version 2.2.5), adjusting for sex and the first ten principal components to account for population stratification calculated using PLINK (version 1.9) (**Figure 1**). The SKATBinary function was run with default settings, including a missingness rate threshold of 15% and the default imputation method "bestguess," which replaces missing genotypes with the most likely value. This approach reduces false positives that can occur with rare variants when missing rates differ between cases and controls. For p-value computation, SKATBinary uses a hybrid strategy that adaptively selects between several available methods. These include quantile adjustment (QA), which calibrates test statistics against the null distribution; efficient resampling (ER), which estimates mid-p values accurately in the presence of very rare variants; and adaptive efficient resampling (ER.A), which applies ER selectively depending on allele counts and case–control balance. Other options such as moment adjustment (MA) are also available, but were not specifically applied here. The hybrid framework ensures robust type I error control across a different patterns of genetic variation (Lee et al., 2016).

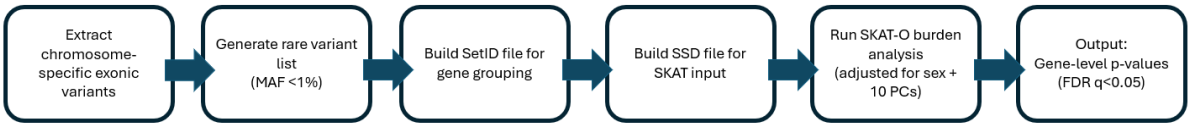


Figure 1. Workflow for rare variant burden analysis. Steps illustrate processing of exonic variants through to gene-level burden testing using SKAT-O.

Results

A total of 1,497 samples were analysed. Of these, 568 had no cranial ischaemic complications (CICs), 677 experienced transient CICs, and 252 had permanent CICs. The cohort comprised 464 males and 1,033 females.

A gene-based, whole-exome wide association analysis was performed using the SKAT-O method for rare variants (minor allele frequency < 1%), adjusting for sex and the first ten genomic principal components (PCs). Analyses were conducted across three clinical comparisons: patients with no CIC versus those with permanent CIC, patients with no CIC versus those with either permanent or transient CIC, and patients with no or transient CIC versus those with permanent CIC. For each comparison, three functional annotation masks were applied: predicted loss-of-function (pLoF) variants only, pLoF plus missense variants,

and pLoF plus missense plus synonymous variants. Across all analyses, no associations remained statistically significant after correction for multiple testing using both Bonferroni adjustment and the Benjamini–Hochberg false discovery rate (FDR) (**See appendix**). The analysis producing the strongest associations was the pLoF plus missense mask in the none versus permanent CIC comparison, in which a total of 17,472 genes were tested (**Table 2**).

Table 2. Top genes identified by SKAT-O under the pLoF plus missense mask across three clinical comparisons. (a) Patients with no CIC versus permanent CIC. (b) Patients with no CIC plus transient CIC versus permanent CIC. (c) Patients with no CIC versus transient plus permanent CIC.

For each comparison, the table shows the top genes ranked by lowest nominal p-values. Columns include the Ensembl gene ID (SetID), gene symbol (GeneID), raw p-value, number of variants in the gene (N.Marker.All), number of variants tested (N.Marker.Test), total minor allele count (*MAC*), number of variant carriers (*m*), test method (Method.bin), minimum single-variant p-value (*MAP*), Bonferroni-adjusted p-value (Bonf), and Benjamini–Hochberg false discovery rate (*FDR*). None of the associations remained significant after multiple testing correction.

(a)

SetID	GeneID	P.value	N.Marker.All	N.Marker.Test	MAC	m	Method.bin	MAP	Bonf	FDR
ENSG00000048707	*VPS13D*	3.17E-05	60	60	107	100	QA	-1	0.554	0.279
ENSG00000183648	*NDUFB1*	3.19E-05	6	6	14	14	ER	1.65E-08	0.557	0.279
ENSG00000104313	*EYA1*	0.000152	11	11	14	14	ER	3.17E-08	1	0.69
ENSG00000158125	*XDH*	0.000158	40	40	64	62	QA	-1	1	0.69
ENSG00000113790	*EHHADH*	0.000435	20	20	35	34	ER.A	-1	1	0.949
ENSG00000005238	*ATOSB*	0.000476	8	8	21	21	ER.A	-1	1	0.949
ENSG00000068001	*HYAL2*	0.000566	6	6	6	6	ER	0.000566	1	0.949
ENSG00000117616	*RSRP1*	0.000602	6	6	23	22	ER.A	-1	1	0.949
ENSG00000171827	*ZNF570*	0.000668	7	7	9	9	ER	6.75E-06	1	0.949
ENSG00000123415	*SMUG1*	0.000683	5	5	9	9	ER	8.26E-06	1	0.949

(b)

SetID	GeneID	P.value	N.Marker.All	N.Marker.Test	MAC	m	Method.bin	MAP	Bonf	FDR
ENSG00000104313	*EYA1*	5.64E-05	14	14	21	21	ER.A	-1	1	0.733
ENSG00000160789	*LMNA*	8.32E-05	20	20	70	67	QA	-1	1	0.733
ENSG00000100888	*CHD8*	0.000137	36	36	91	80	QA	-1	1	0.733
ENSG00000221826	*PSG3*	0.000189	31	31	62	59	QA	-1	1	0.733
ENSG00000186474	*KLK12*	0.000211	12	12	19	19	ER	3.35E-16	1	0.733
ENSG00000048707	*VPS13D*	0.000248	92	92	183	171	QA	-1	1	0.733
ENSG00000155393	*HEATR3*	0.000377	20	20	37	35	ER.A	-1	1	0.887
ENSG00000113593	*PPWD1*	0.000441	10	10	13	13	ER	9.35E-12	1	0.887
ENSG00000119715	*ESRRB*	0.000449	9	9	29	29	ER.A	-1	1	0.887
ENSG00000150045	*KLRF1*	0.000576	5	5	10	10	ER	1.53E-08	1	0.889

(c)

SetID	GeneID	P.value	N.Marker.All	N.Marker.Test	MAC	m	Method.bin	MAP	Bonf	FDR
ENSG00000124201	*ZNFX1*	6.23E-05	31	31	80	78	QA	-1	1	0.611
ENSG00000099194	*SCD*	6.88E-05	5	5	11	11	ER	1.22E-05	1	0.611
ENSG00000150456	*EEF1AKMT1*	0.000137	4	4	8	8	ER	0.000137	1	0.814
ENSG00000153015	*CWC27*	0.000287	9	9	39	23	ER.A	-1	1	0.895
ENSG00000172146	*OR1A1*	0.000308	7	7	9	9	ER	9.11E-05	1	0.895
ENSG00000145649	*GZMA*	0.000312	7	7	8	8	ER	0.000312	1	0.895
ENSG00000104936	*DMPK*	0.00044	22	22	39	37	ER.A	-1	1	0.895
ENSG00000162687	*KCNT2*	0.000543	11	11	16	16	ER	1.31E-07	1	0.895
ENSG00000239389	*PCDHA13*	0.000597	23	23	66	65	QA	-1	1	0.895
ENSG00000196946	*ZNF705A*	0.000607	8	8	45	43	QA	-1	1	0.895

To assess the overall behaviour of the association statistics, a histogram of SKAT-O p-values across all genes was generated (**Figure 2**). The near-uniform distribution indicates an absence of systematic inflation or global enrichment, consistent with a null distribution.

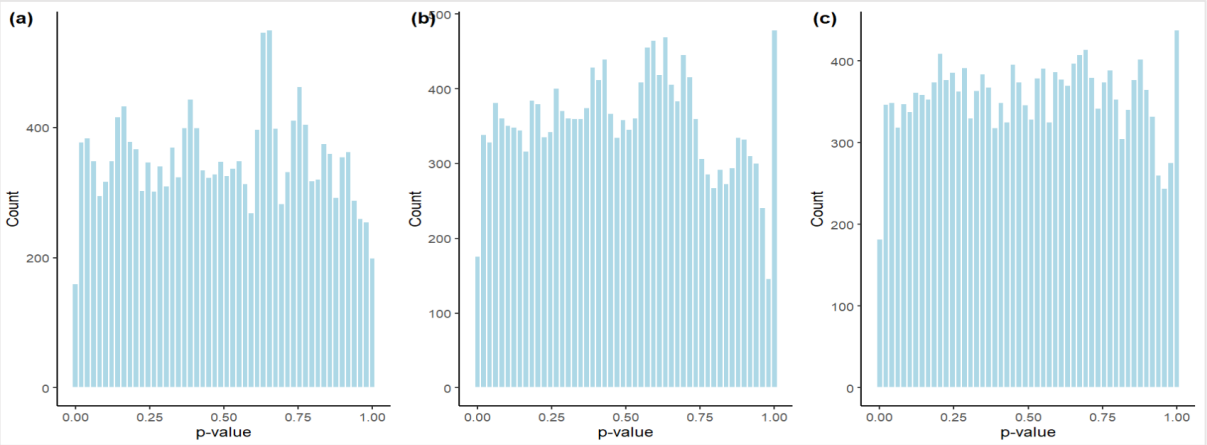


Figure 2. Histograms of SKAT-O gene-level p-values under the pLoF plus missense mask across three clinical comparisons. (a) Patients with no CIC versus permanent CIC. (b) Patients with no CIC plus transient CIC versus permanent CIC. (c) Patients with no CIC versus transient plus permanent CIC.

In all comparisons, the distribution of p-values was approximately uniform, consistent with the null expectation and indicating no evidence of systematic inflation.

This finding was further supported by the quantile–quantile (Q-Q) plot of p-values (**Figure 3**). The observed values closely followed the expected null distribution, demonstrating that the test was well-calibrated. A slight deviation at the extreme tail reflects the few genes with the lowest nominal p-values, although none remained significant after multiple-testing correction.

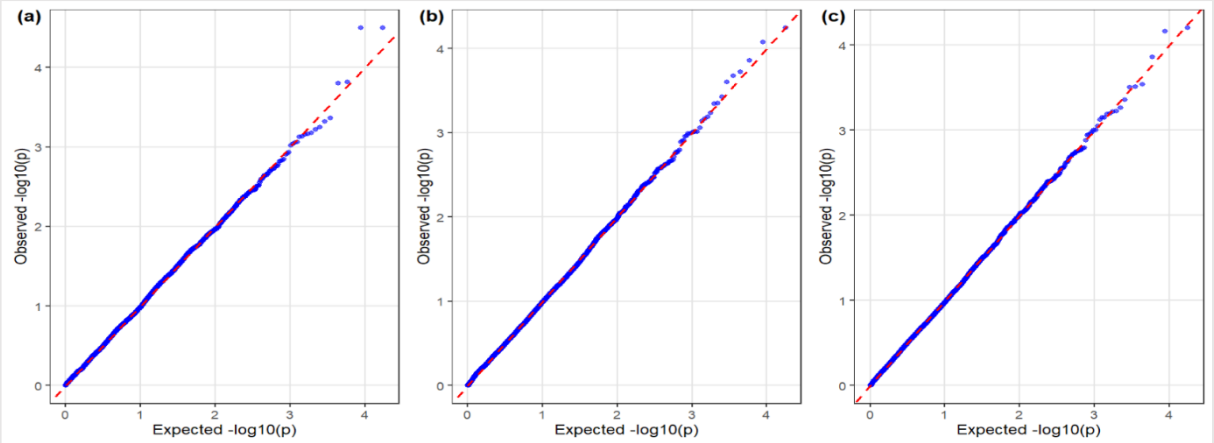


Figure 3. Q-Q plots of SKAT-O gene-level p-values under the pLOF plus missense mask across three clinical comparisons. (a) Patients with no CIC versus permanent CIC. (b) Patients with no CIC plus transient CIC versus permanent CIC. (c) Patients with no CIC versus transient plus permanent CIC.

In all comparisons, the observed p-values closely followed the expected null distribution, indicating well-calibrated test statistics. Minor upward deviations at the extreme tail represent genes with the lowest nominal p-values, although none remained significant after multiple testing correction.

Discussion

Study Aim and Context

This study investigated whether rare genetic variants contribute to the risk of cranial ischaemic complications in patients with giant cell arteritis. Cranial ischaemia represents one of the most severe and clinically important outcomes of GCA, leading to permanent vision loss or stroke in a subset of patients. Clarifying whether rare germline variants predispose to these complications could provide valuable insight into disease mechanisms and may ultimately support improved risk stratification or therapeutic development.

To explore this, a gene-based rare variant burden analysis was applied to whole-exome sequencing data from a clinical cohort of GCA patients stratified by cranial ischaemic status. Comparisons were made between patients without CICs, those with transient events, and those with permanent complications. This subgroup-based design, rarely applied in GCA genetics, was intended to shed light on the genetic contribution to the most severe vascular outcomes of the disease.

Summary of Main Results

Using the pLoF plus missense mask in the comparison of patients with no cranial ischaemic complications versus those with permanent CICs, 17,472 genes were tested by rare variant burden analysis with SKAT-O in 1,497 patients with GCA. No gene reached significance after multiple-testing correction by either Bonferroni adjustment or the Benjamini–Hochberg false discovery rate (FDR). The top-ranked genes by nominal p-value included *VPS13D*, *NDUFB1*, *EYA1*, *XDH*, *EHHADH*, *HYAL2*, *RSRP1*, *ATOSB*, and *ZNF570*, although none approached the corrected significance threshold. The distribution of p-values, as illustrated by the histogram and Q-Q plot, was consistent with the null expectation.

Interpretation

The absence of statistically significant associations does not imply that rare variants are irrelevant to cranial ischaemic complications in GCA. Rather, it reflects the inherent challenges of rare-variant association testing and the complexity of the phenotype under study.

A major factor is statistical power. Although the overall cohort of 1,497 patients is large for a rare disease, the number of permanent CIC cases (n = 252) is modest. Rare-variant analyses are constrained by allele frequency: very low-frequency variants may not be observed at all, or only in a handful of individuals. For example, an allele with a minor allele frequency of 0.01% would be expected in fewer than one individual in this cohort, making it extremely difficult to detect differences between groups. Literature consistently emphasises that thousands of cases are often required to detect rare-variant signals reliably (Zuk et al., 2014; Zhang et al., 2019). This study should therefore be viewed as exploratory rather than definitive, with replication in larger or multi-centre cohorts needed.

The use of SKAT-O provided a robust analytic framework under these constraints. SKAT-O adaptively combines a burden test, which assumes all variants influence risk in the same direction, with a variance-component test, which allows for heterogeneity of effects. This is appropriate for complex inflammatory diseases such as GCA, where some rare variants may increase susceptibility, others may be protective, and many may have no observable effect. SKAT-O also includes a small-sample adjustment, automatically applied when total sample size falls below 2,000 individuals, which corrects the conservative bias that can otherwise occur in binary traits with small case numbers (Lee et al., 2012). This ensures calibration of type I error rates and makes SKAT-O a widely accepted choice for studies of this scale.

An alternative approach for rare variant association testing in binary traits is SKATBinary.Robust, which was developed to address type I error inflation that can occur in imbalanced case–control designs or very small samples. This method calibrates single-variant score statistics using a saddlepoint approximation and efficient resampling before aggregation, providing accurate p-values even when case numbers are low or case:control ratios are extreme (Zhao et al., 2020). Simulation studies and applications in UK Biobank and other cohorts have shown that robust SKAT-O controls false positives under severe imbalance (e.g. 1:100 ratios) while maintaining power. In the present study, I chose to use the standard SKATBinary.SSD.All function with SKAT-O, which already includes small-sample adjustment by default, to ensure consistency across the three phenotype contrasts analysed, some of which were relatively well balanced. SKATBinary.Robust remains a valid alternative, and its use should be considered in future work, particularly for replication studies or when analysing more extremely unbalanced phenotypes.

Even with these methodological safeguards, SKAT-O cannot overcome the fundamental limitation of few variant carriers in small cohorts. In the smallest stratum (252 permanent CIC cases versus 568 with no CIC), the number of carriers of any given rare variant was inevitably low, limiting power even after aggregation. This reflects a common challenge in rare-variant studies: reliable discovery typically requires larger collaborative datasets or meta-analyses.

A further limitation concerns the clinical definition of cases. Current GCA diagnostic guidelines recommend at least one confirmatory test, such as temporal artery biopsy (TAB) or vascular imaging, to distinguish arteritic anterior ischaemic optic neuropathy (AION) from non-arteritic AION, the latter accounting for ~95% of all AION cases (Mackie et al., 2020). In the wider cohort of 1,946 patients, from which the 1,497 analysed here were drawn, approximately 30% did not undergo confirmatory diagnostic testing (Chaddock et al., 2025). Consequently, some patients included in this analysis may have lacked biopsy or imaging-confirmed GCA, introducing the possibility of misclassification. Such misclassification would dilute true genetic

signals by including non-GCA cases. This highlights the importance of precise phenotyping in rare variant studies, where even small numbers of misclassified cases can disproportionately reduce power.

Finally, the phenotype itself is inherently heterogeneous. Cranial ischaemic complications arise from a complex interplay of genetic predisposition, vascular pathology, and inflammatory processes, with contributions distributed across many variants of small effect. Even within the strict comparison of permanent CIC versus no CIC, heterogeneity in clinical manifestations (e.g. visual loss versus stroke) may have diluted gene-level signals.

Overall, the lack of significant findings should not be interpreted as evidence against a genetic contribution to CICs in GCA. Rather, it reflects the multifactorial nature of the outcome and the modest effect sizes expected for rare variants. This study demonstrates the feasibility of applying gene-based burden tests such as SKAT-O in GCA and underscores the importance of expanding sample sizes, refining phenotypic definitions, and ensuring robust diagnostic criteria in future research.

Comparison with Existing Literature

Genome-wide and immunochip studies of GCA have consistently highlighted the strong role of HLA class II alleles, particularly *HLA-DRB1*∗04, in disease susceptibility, which remains the most robust common genetic determinant across populations (Carmona et al., 2015). Beyond HLA, several non-HLA loci have been implicated. Early studies identified immune-regulatory genes such as *PTPN22*, *LRRC32*, and *REL*, pointing toward dysregulation of T-cell and NF-κB pathways, while subsequent GWAS added loci involved in vascular biology, including *PLG* and *P4HA2* (Carmona et al., 2017). These findings introduced fibrinolysis, collagen synthesis, and extracellular matrix remodelling as key processes in GCA pathogenesis. The largest GWAS to date confirmed the central role of HLA region and *PLG* and further identified three novel loci: *MFGE8* and *VTN*, linked to angiogenic pathways, and *CCDC25*, a regulator of neutrophil extracellular trap formation (Borrego-Yaniz et al., 2024). Collectively, GWAS findings suggest that GCA susceptibility reflects an interplay between immune dysregulation and vascular injury/repair mechanisms.

Most GWAS have not specifically interrogated cranial ischaemic complications as a phenotype, though smaller studies provide insights. Carriage of *HLA-DRB1*∗04 has been associated not only with overall GCA risk but also with increased likelihood of visual ischaemic complications (Amoli et al., 2001). Likewise, polymorphisms in *VEGF* have been linked to outcomes such as vision loss and stroke, consistent with the idea that impaired angiogenic compensation predisposes to ischemia (Prieto-Peña et al., 2022). These findings suggest that while GCA susceptibility is strongly immune-driven, the risk of complications may additionally depend on vascular repair capacity and angiogenesis.

In contrast, rare variant analyses in GCA remain sparse. To date, no exome-wide rare variant study has reported genome-wide significant findings, and the contribution of uncommon alleles remains largely unexplored. In related conditions, however, rare variants have been implicated: for example, damaging variants in *NLRP12* and other inflammasome genes in polymyalgia rheumatica (Higuchi et al., 2024), or rare alleles in *NFKB1* and *IL15* in Takayasu arteritis. More broadly, rare variant association studies in autoimmune disease have successfully applied SKAT-O and burden tests, identifying protective *IL23R* alleles in inflammatory bowel disease and *IL36RN* mutations underlying pustular psoriasis (Onoufriadis et al., 2011; Delgado-Vega et al., 2018; Venkataraman and Rivas, 2019). These studies

highlight the value of gene-based rare variant approaches, but also the need for very large cohorts, often tens of thousands of cases, to detect exome-wide significant signals.

With regard to cranial ischemia in GCA, both clinical and genetic factors are recognised. Clinically, advanced age and hypertension are consistently associated with higher risk of ischemic complications, while anticoagulant therapy at diagnosis appears protective (Chaddock et al., 2025). Genetically, in addition to *HLA-DRB1*04* and *VEGF* polymorphisms, polygenic risk analyses suggest variants near *TEK*, which regulates angiogenesis, and *CD96*, an immunoregulatory receptor, have been associated with ischemic outcomes in GCA cohorts (Chaddock et al., 2025). These findings further reinforce the dual contribution of immune responses and vascular repair capacity to ischemic risk.

Considering the literature, the present study adds complementary evidence. Although no associations reached statistical significance, several genes with the lowest nominal p-values map to biologically plausible processes. *XDH* encodes xanthine dehydrogenase, a source of reactive oxygen species in vascular tissue, linking to oxidative stress and endothelial dysfunction (Noda et al., 2023). *HYAL2* regulates extracellular matrix turnover through hyaluronan degradation, central to vascular remodelling (de la Motte et al., 2009). *EYA1* promotes angiogenesis, consistent with prior associations of *VEGF* and *MFGE8* in GCA (Tadjuidje et al., 2012). *VPS13D* and *NDUFB1* are involved in mitochondrial function and oxidative stress regulation, processes increasingly recognised as modulators of inflammatory responses (Anding et al., 2018). Notably, none of these overlap with known GCA susceptibility loci, suggesting that the genetic architecture of ischaemic complications may diverge from that of disease onset.

Taken together, these results align with prior evidence that no single rare variant of large effect drives GCA or its complications. Instead, they suggest that multiple small contributions across immune, metabolic, and vascular pathways are likely involved. This underscores both the complexity of GCA genetics and the likelihood that ischaemic complications reflect a convergence of immune dysregulation and impaired vascular resilience rather than a single genetic driver.

Limitations

Several limitations of this study should be acknowledged. First, although the overall cohort of 1,497 patients with GCA is large for a rare disease, the number of individuals with permanent cranial ischaemic complications (n = 252) remained modest. Rare-variant tests rely on sufficient carriers of low-frequency alleles, and in this context statistical power was limited. It is therefore possible that true associations exist but could not be detected in a study of this size.

Second, phenotype definition introduces uncertainty. Approximately 30% of patients in the wider dataset did not undergo confirmatory temporal artery biopsy or vascular imaging, raising the possibility of misclassification. Inclusion of misdiagnosed cases would attenuate genuine signals. In addition, the “permanent cranial ischaemia” category is heterogeneous, encompassing visual loss, stroke, and other manifestations that may not share identical biology, further diluting potential associations.

Third, the scope of the analysis was restricted to germline rare coding variants. Non-coding regions, regulatory elements, and somatic mutations were not assessed, and integration with transcriptomic or epigenomic data was beyond the scope of this project.

Fourth, allele frequency thresholds were defined using the case-only cohort of 1,497 patients. This introduces two important sources of bias. In a dataset of this size, many truly rare variants present in the wider population may not appear at all simply by chance, meaning their absence here does not exclude a role in cranial ischaemic complications. Conversely, variants enriched among GCA patients may appear "common" internally and therefore be excluded, even though they remain rare in the general population. For example, a variant with a population frequency of 0.1% could be observed in ~15 individuals here (≈1% of the cohort) and wrongly discarded as non-rare. Large-scale population resources such as gnomAD provide more reliable frequency estimates across diverse ancestries, and are now recommended for filtering rare variants to avoid stochastic losses and case-enrichment bias (Pedersen et al., 2021; Gudmundsson et al., 2022). Incorporating such reference data in future analyses would help ensure that variants are filtered according to true population rarity.

Finally, there were analytical considerations. In the pLoF plus missense mask for the none versus permanent comparison, 19 genes returned *N.Marker.Test = 0*, indicating that no variants passed quality control for those genes. These were inadvertently retained in the false discovery rate adjustment, although this did not alter conclusions as even the closest genes to significance were far from the corrected threshold. Nonetheless, functions such as Get_EffectiveNumberTest within the SKAT package could be used in future to refine multiple-testing adjustment more formally. In addition, many genes, particularly in the pLoF-only mask, contained very few qualifying variants (**See appendix**). Around 2,000 genes had only one or two, and the majority fewer than four. Gene-based tests with such small sets, especially when minor allele counts are extremely low, are known to yield unstable or spurious results (Lee et al., 2017). Including them in multiple-testing correction also inflates the apparent number of tests, making thresholds overly stringent. Standard practice in rare-variant studies is to impose minimum thresholds, such as requiring ≥3 variants or a total MAC ≥5 for a gene to be tested (Lee et al., 2017; Gao et al., 2022). Applying such filters in future analyses would reduce noise, improve calibration of test statistics, and ensure that multiple-testing correction reflects only genuinely testable genes.

Future Work

Future studies will need to address the limitations of this analysis in order to clarify the genetic contribution to cranial ischaemic complications in GCA. Most important among these is the requirement for larger sample sizes. Rare variant association tests demand large numbers of variant carriers, and single-centre studies are unlikely to achieve this. Multi-centre collaborations and meta-analyses across cohorts will therefore be essential to achieve sufficient power.

Methodologically, future analyses should also validate findings by applying alternative gene-based rare variant methods alongside SKAT-O. Classical burden tests are most powerful when all causal variants act in the same direction, variance-component tests such as SKAT are better suited to mixed effects, and p-value combination methods such as the Aggregated Cauchy Association Test (ACAT) and the Harmonic Mean p-value (HMP) offer advantages when only a subset of variants drive the association or when only summary statistics are available (Lee et al., 2014; Wilson, 2019; Liu et al., 2019). Because no single test is uniformly optimal, employing a range of complementary approaches would provide an important validation step: consistent results across methods would strengthen confidence in any gene-level signals, while divergence could provide insight into the underlying genetic architecture.

Replication of the nominal gene signals identified here should also be pursued in expanded datasets. Genes such as *XDH*, *HYAL2*, *EYA1*, and *VPS13D* map to processes central to vascular inflammation and ischemia, including oxidative stress, angiogenesis, and extracellular matrix turnover. Independent validation will be essential to distinguish true biological associations from chance findings and to prioritise targets for further investigation.

Finally, there is potential for clinical translation. A clearer understanding of genetic risk factors for cranial ischaemia could ultimately support stratification of GCA patients. Individuals with a genetic predisposition to complications might benefit from intensified monitoring, earlier intervention, or adjunctive therapies. Achieving this will require not only genetic discovery but also functional validation of candidate genes and integration of genetic data with established clinical risk models.

In summary, while this study provides an initial framework for rare variant analysis of GCA complications, substantial progress will depend on larger collaborative efforts, replication of preliminary signals, and integration with functional and clinical data.

Conclusion

This study represents the first application of gene-based rare variant analysis to investigate cranial ischaemic complications in GCA. Although no associations reached statistical significance, several nominal signals implicated genes involved in biologically plausible processes, including oxidative stress, extracellular matrix turnover, angiogenesis, and mitochondrial function. These findings highlight that rare variants of large effect are unlikely to underlie GCA complications and instead suggest a model in which multiple small contributions converge across immune, vascular, and metabolic pathways.

The absence of significant associations should be interpreted in the context of limited statistical power and heterogeneous clinical phenotypes. Nonetheless, the analyses demonstrate the feasibility of applying rare variant burden testing in clinically stratified GCA cohorts and provide a framework for future work. Larger collaborative datasets, replication of nominal signals, and integration with functional and clinical data will be required to clarify the genetic architecture of ischaemic complications and to realise their potential for clinical risk stratification.

Code availability

All analyses were conducted using open-source software. Rare variant association testing was performed with the SKAT R package v2.2.5 (https://cran.r-project.org/web/packages/SKAT), using the SKATBinary.SSD.All function. Quality control of genotype data was performed with PLINK v1.9 (https://www.cog-genomics.org/plink/1.9/) and supporting scripts. Variant annotation was carried out using the Ensembl Variant Effect Predictor (VEP) v101.0 (https://www.ensembl.org/info/docs/tools/vep/). Data visualisation, including histograms, Q–Q plots, and gene-level tables, was performed in R v4.3.0.

Codes are available at: https://github.com/ChristopherLetton/Research_Project

References

Amoli, M.M., Shelley, E., Mattey, D.L., Garcia-Porrua, C., Thomson, W., Hajeer, A.H., Ollier, W.E. and Gonzalez-Gay, M.A. 2001. Lack of association between intercellular adhesion molecule-1 gene polymorphisms and giant cell arteritis. *The Journal of Rheumatology*. **28**(7), p.1600.

Anding, A.L., Wang, C., Chang, T.-K., Sliter, D.A., Powers, C.M., Hofmann, K., Youle, R.J. and Baehrecke, E.H. 2018. Vps13D Encodes a Ubiquitin-Binding Protein that Is Required for the Regulation of Mitochondrial Size and Clearance. *Current Biology*. **28**(2), pp.287-295.e6.

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., Holle, J., Schönau, V., Pugnet, G., Papo, T., Haroche, J., Mahr, A., Mouthon, L., Molberg, Ø., Diamantopoulos, A.P., Voskuyl, A., Daikeler, T., Berger, C.T., Molloy, E.S., Blockmans, D., van Sleen, Y., Iles, M., Sorensen, L., Luqmani, R., Reynolds, G., Bukhari, M., Bhagat, S., Ortego-Centeno, N., Brouwer, E., Lamprecht, P., Klapa, S., Salvarani, C., Merkel, P.A., Cid, M.C., González-Gay, M.A., Morgan, A.W., Martin, J., Márquez, A., Callejas, J.L., Caminal-Montero, L., Corbera-Bellalta, M., de Miguel, E., Díaz-López, J.B., García-Villanueva, M.J., Gómez-Vaquero, C., Guijarro-Rojas, M., Hidalgo-Conde, A., Marí-Alfonso, B., Martínez-Berriochoa, A., Morado, I.C., Narváez, J., Ramentol-Sintas, M., Martínez-Zapico, A., Martínez-Taboada, V.M., Miranda-Filloy, J.A., Monfort, J., Pérez-Conesa, M., Prieto-González, S., Raya, E., Ríos-Fenández, R., Sánchez-Martín, J., Sopeña, B., Tío, L., Unzurrunzaga, A., Wordsworth, O., Whitwell, I., Brock, J., Douglas, V., Hettiarachchi, C., Bartholomew, J., Jarrett, S., Smithson, G., Green, Michael, Brown, P.C., Lawson, C., Gordon, E., Lane, S., Francis, R., Dasgupta, B., Masunda, B., Calver, J., Patel, Y., Thompson, C., Gregory, L., Levy, S., Menon, A., Thompson, A., Dyche, L., Martin, M., Li, C., Laxminarayan, R., Wilcox, L., de Guzman, R., Isaacs, J., Lorenzi, A., Farley, R., Hinchcliffe-Hume, H., Bejarano, V., Hope, S., Nandi, P., Stockham, L., Wilde, C., Durrant, D., Lloyd, M., Ye, C.-S., Stevens, R., Jilani, A., Collins, D., Pegler, S., Rivett, A., Price, L., McHugh, N., Skeoch, S., O'Kane, D., Kirkwood, S., Vadivelu, S., Pugmire, S., Sultan, S., Dooks, E., Armstrong, L., Sadik, H., Nandagudi, A., Abioye, T., Ramos, A., Gumus, S., Sofat, N., Harrison, A., Seward, A., Mollan, S., Rahan, R., Hawkins, H., Emsley, H., Bhargava, A., Fleming, V., Hare, M., Raj, S., George, E., Allen, N., Hunter, K., O'Sullivan, E., Bird, G., Magliano, M., Manzo, K., Sanghera, B., Hutchinson, D., Hammonds, F., Sharma, P., Cooper, R., McLintock, G., Al-Saffar, Z.S., Green, Mike, Elliott, K., Neale, T., Mallinson, J., Lanyon, P., Pradere, M.-J., Jordan, N., Htut, E.P., Mushapaidzi, T., Abercrombie, D., Wright, S., Rowlands, J., Mukhtyar, C., Kennedy, J., Makkuni, D., Wilhelmsen, E., Kouroupis, M., John, L., Hughes, R., Walsh, M., Buckley, M., Mackay, K., Camden-Woodley, T., Redome, J., Pearce, K., Marianayagam, T., Cruz, C., Warner, E., Atchia, I., Walker, C., Black, K., Duffy, S., Fothergill, L., Jefferey, R., Toomey, J., Rhys-Dillon, C., Pothecary, C., Green, L., Toms, T., Maher, L., Davis, D., Sayan, A., Thankachen, M., Abusalameh, M., Record, J., Khan, A., Stafford, S., Hussein, A., Williams, C., Fletcher, A., Johson, L., Burnett, R., Moots, R., Frankland, H., Dale, J., Black, K., Moar, K., Hollas, C., Parker, B., Ridings, D., Eapen, S., John, S., Robson, J., Guthrie, L.B., Fyfe, R., Tait, M., Marks, J., Gunter, E., Hernandez, R., Bhat, S., Johnston, P., Khurshid, M., Barclay, C., Kapur, D., Jeffrey, H., Hughes, A., Slack, L., Thomas, E., Royon, A., Hall, A., King, J., Nyathi, S., Morris, V., Castelino, M., Hawkins, E., Tomson, L., Singh, A., Nunag, A., O'Connor, S., Rushby, N., Hewitson, N., O'Sunmboye, K., Lewszuk, A., Boyles, L., Perry, M., Williams, E., Graver, C., Defever, E., Kamanth, S., Kay, D., Ogor, J., Winter, L., Horton, S., Welch, G., Hollinshead, K., Peters, J., Labao, J., Dmello, A., Dawson, J., Graham, D., De Lord, D., Deery, J., Hazelton, T., Carette, S., Chung, S., Cuthbertson, D., Forbess, L.J., Gewurz-Singer, O., Hoffman, G.S., Koening, C.L., Maksimowicz-McKinnon, K.M., McAlear, C.A., Moreland, L.W., Pagnoux, C., Seo, P., Specks, U., Spiera, R.F., Sreih, A., Warrington, K.J., Monach, P.A. and Weisman, M. 2024. Risk loci involved in giant cell arteritis susceptibility: a genome-wide association study. *The Lancet Rheumatology*. **6**(6), pp.e374–e383.

Carmona, F.D., Martín, J. and González-Gay, M.A. 2015. Genetics of vasculitis. *Current Opinion in Rheumatology*. **27**(1).

Carmona, F.D., Vaglio, A., Mackie, S.L., Hernández-Rodríguez, J., Monach, P.A., Castañeda, S., Solans, R., Morado, I.C., Narváez, J., Ramentol-Sintas, M., Pease, C.T., Dasgupta, B., Watts, R., Khalidi, N., Langford, C.A., Ytterberg, S., Boiardi, L., Beretta, L., Govoni, M., Emmi, G., Bonatti, F., Cimmino, M.A., Witte, T., Neumann, T., Holle, J., Schönau, V., Sailler, L., Papo, T., Haroche, J., Mahr, A., Mouthon, L., Molberg, Ø., Diamantopoulos, A.P., Voskuyl, A., Brouwer, E., Daikeler, T., Berger, C.T., Molloy, E.S., O'Neill, L., Blockmans, D., Lie, B.A., Mclaren, P., Vyse, T.J., Wijmenga, C., Allanore, Y., Koeleman, B.P.C., Callejas, J.L., Caminal-Montero, L., Corbera-Bellalta, M., de Miguel, E., López, J.B.D., García-Villanueva, M.J., Gómez-Vaquero, C., Guijarro-Rojas, M., Hidalgo-Conde, A., Marí-Alfonso, B., Berriochoa, A.M., Zapico, A.M., Martínez-Taboada, V.M., Miranda-Filloy, J.A., Monfort, J., Ortego-Centeno, N., Pérez-Conesa, M., Prieto-González, S., Raya, E., Fernández, R.R., Sánchez-Martín, J., Sopeña, B., Tío, L., Unzurrunzaga, A., Gough, A., Isaacs, J.D., Green, M., McHugh, N., Hordon, L., Kamath, S., Nisar, M., Patel, Y., Yee, C.-S., Stevens, R., Nandi, P., Nandagudi, A., Jarrett, S., Li, C., Levy, S., Mollan, S., Salih, A., Wordsworth, O., Sanders, E., Roads, E., Gill, A., Carr, L., Routledge, C., Culfear, K., Nugaliyadde, A., James, L., Spimpolo, J., Kempa, A., Mackenzie, F., Fong, R., Peters, G., Rowbotham, B., Masqood, Z., Hollywood, J., Gondo, P., Wood, R., Martin, S., Rashid, L.H., Robinson, J.I., Morgan, M., Sorensen, L., Taylor, J., Carette, S., Chung, S., Cuthbertson, D., Forbess, L.J., Gewurz-Singer, O., Hoffman, G.S., Koening, C.L., Maksimowicz-McKinnon, K.M., McAlear, C.A., Moreland, L.W., Pagnoux, C., Seo, P., Specks, U., Spiera, R.F., Sreih, A., Warrington, K.J., Weisman, M., Barrett, J.H., Cid, M.C., Salvarani, C., Merkel, P.A., Morgan, A.W., González-Gay, M.A. and Martín, J. 2017. A Genome-wide Association Study Identifies Risk Alleles in Plasminogen and P4HA2 Associated with Giant Cell Arteritis. *The American Journal of Human Genetics*. **100**(1), pp.64–74.

Chaddock, N.J.M., Harden, C.J., Sorensen, L., Mathieson, H.R., Zulcinski, M., Lawson, C.A., O'Sullivan, E., Mollan, S.P., Martin, J., Mackie, S.L., Iles, M.M. and Morgan, A.W. 2025. Age, anticoagulants, hypertension and cardiovascular genetic traits predict cranial ischaemic complications in patients with giant cell arteritis. *Annals of the Rheumatic Diseases*. **84**(2), pp.329–340.

Delgado-Vega, A.M., Martínez-Bueno, M., Oparina, N.Y., López Herráez, D., Kristjansdottir, H., Steinsson, K., Kozyrev, S. V and Alarcón-Riquelme, M.E. 2018. Whole Exome Sequencing of Patients from Multicase Families with Systemic Lupus Erythematosus Identifies Multiple Rare Variants. *Scientific Reports*. **8**(1), p.8775.

Donaldson, L. and Margolin, E. 2022. Vision loss in giant cell arteritis. *Practical Neurology*. **22**(2), p.138.

Gao, X.R., Chiariglione, M. and Arch, A.J. 2022. Whole-exome sequencing study identifies rare variants and genes associated with intraocular pressure and glaucoma. *Nature Communications*. **13**(1), p.7376.

Gonzalez-Gay, M.A., Barros, S., Lopez-Diaz, M.J., Garcia-Porrua, C., Sanchez-Andrade, A. and Llorca, J. 2005. Giant Cell Arteritis: Disease Patterns of Clinical Presentation in a Series of 240 Patients. *Medicine*. **84**(5).

Gudmundsson, S., Singer-Berk, M., Watts, N.A., Phu, W., Goodrich, J.K., Solomonson, M., Consortium, G.A.D., Rehm, H.L., MacArthur, D.G. and O'Donnell-Luria, A. 2022. Variant interpretation using population databases: Lessons from gnomAD. *Human Mutation*. **43**(8), pp.1012–1030.

Hellmich, B., Agueda, A., Monti, S., Buttgereit, F., de Boysson, H., Brouwer, E., Cassie, R., Cid, M.C., Dasgupta, B., Dejaco, C., Hatemi, G., Hollinger, N., Mahr, A., Mollan, S.P., Mukhtyar, C., Ponte, C., Salvarani, C., Sivakumar, R., Tian, X., Tomasson, G., Turesson, C., Schmidt, W., Villiger, P.M., Watts, R., Young, C. and Luqmani, R.A. 2020. 2018 Update of the EULAR recommendations for the management of large vessel vasculitis. *Annals of the Rheumatic Diseases*. **79**(1), pp.19–30.

Higuchi, T., Oka, S., Furukawa, H. and Tohma, S. 2024. The contributions of deleterious rare alleles in NLRP12 and inflammasome-related genes to polymyalgia rheumatica. *Scientific Reports*. **14**(1), p.490.

Kermani, T.A., Warrington, K.J., Crowson, C.S., Ytterberg, S.R., Hunder, G.G., Gabriel, S.E. and 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), pp.1989–1994.

de la Motte, C., Nigro, J., Vasanji, A., Rho, H., Kessler, S., Bandyopadhyay, S., Danese, S., Fiocchi, C. and Stern, R. 2009. Platelet-Derived Hyaluronidase 2 Cleaves Hyaluronan into Fragments that Trigger Monocyte-Mediated Production of Proinflammatory Cytokines. *The American Journal of Pathology*. **174**(6), pp.2254–2264.

Lazaar, H., Moumni, A., Aziz, A., Boutaj, T., Amazouzi, A., Boutimzine, N., El Ouazzani, H. and Cherkaoui, L.O. 2025. Unilateral Central Retinal Artery Occlusion Revealing Giant Cell Arteritis: A Case Report. *Cureus*. **17**(3), p.e80452.

Lee, S., Abecasis, G.R., Boehnke, M. and Lin, X. 2014. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *The American Journal of Human Genetics*. **95**(1), pp.5–23.

Lee, S., Emond, M.J., Bamshad, M.J., Barnes, K.C., Rieder, M.J., Nickerson, D.A., Christiani, D.C., Wurfel, M.M. and Lin, X. 2012. Optimal Unified Approach for Rare-Variant Association Testing with Application to Small-Sample Case-Control Whole-Exome Sequencing Studies. *The American Journal of Human Genetics*. **91**(2), pp.224–237.

Lee, S., Fuchsberger, C., Kim, S. and Scott, L. 2016. An efficient resampling method for calibrating single and gene-based rare variant association analysis in case–control studies. *Biostatistics*. **17**(1), pp.1–15.

Lee, S., Kim, S. and Fuchsberger, C. 2017. Improving power for rare-variant tests by integrating external controls. *Genetic Epidemiology*. **41**(7), pp.610–619.

Liu, Y., Chen, S., Li, Z., Morrison, A.C., Boerwinkle, E. and Lin, X. 2019. ACAT: A Fast and Powerful p Value Combination Method for Rare-Variant Analysis in Sequencing Studies. *The American Journal of Human Genetics*. **104**(3), pp.410–421.

Lyons, H.S., Quick, V., Sinclair, A.J., Nagaraju, S. and Mollan, S.P. 2020. A new era for giant cell arteritis. *Eye*. **34**(6), pp.1013–1026.

Mackie, S.L., Dejaco, C., Appenzeller, S., Camellino, D., Duftner, C., Gonzalez-Chiappe, S., Mahr, A., Mukhtyar, C., Reynolds, G., de Souza, A.W.S., Brouwer, E., Bukhari, M., Buttgereit, F., Byrne, D., Cid, M.C., Cimmino, M., Direskeneli, H., Gilbert, K., Kermani, T.A., Khan, A., Lanyon, P., Luqmani, R., Mallen, C., Mason, J.C., Matteson, E.L., Merkel, P.A., Mollan, S., Neill, L., Sullivan, E.O., Sandovici, M., Schmidt, W.A., Watts, R., Whitlock, M., Yacyshyn, E., Ytterberg, S. and Dasgupta, B. 2020. British Society for Rheumatology guideline on diagnosis and treatment of giant cell arteritis. *Rheumatology*. **59**(3), pp.e1–e23.

Momozawa, Y. and Mizukami, K. 2021. Unique roles of rare variants in the genetics of complex diseases in humans. *Journal of Human Genetics*. **66**(1), pp.11–23.

Noda, M., Kikuchi, C., Tarui, R., Nakamura, T., Murase, T. and Matsunaga, T. 2023. Effects of Xanthine Oxidoreductase Inhibitors on Reactive Oxygen Species Produced *In Vitro* from Xanthine Oxidase. *BPB Reports*. **6**(6), pp.189–192.

Onoufriadis, A., Simpson, M.A., Pink, A.E., Di Meglio, P., Smith, C.H., Pullabhatla, V., Knight, J., Spain, S.L., Nestle, F.O., Burden, A.D., Capon, F., Trembath, R.C. and Barker, J.N. 2011. Mutations in IL36RN/IL1F5 Are Associated with the Severe Episodic Inflammatory Skin Disease Known as Generalized Pustular Psoriasis. *The American Journal of Human Genetics*. **89**(3), pp.432–437.

Patil, P., Karia, N., Jain, S. and Dasgupta, B. 2013. Giant cell arteritis: a review. *Eye and Brain*. **5**(null), pp.23–33.

Pedersen, B.S., Brown, J.M., Dashnow, H., Wallace, A.D., Velinder, M., Tristani-Firouzi, M., Schiffman, J.D., Tvrdik, T., Mao, R., Best, D.H., Bayrak-Toydemir, P. and Quinlan, A.R. 2021. Effective variant filtering and expected candidate variant yield in studies of rare human disease. *npj Genomic Medicine*. **6**(1), p.60.

Prieto-Peña, D., Remuzgo-Martínez, S., Genre, F., Ocejo-Vinyals, J.G., Atienza-Mateo, B., Muñoz-Jimenez, A., Ortiz-Sanjuán, F., Romero-Yuste, S., Moriano, C., Galíndez-Agirregoikoa, E., Calvo, I., Ortego-Centeno, N., Álvarez-Rivas, N., Miranda-Filloy, J.A., Llorente, I., Blanco, R., Gualillo, O., Martín, J., Márquez, A., Castañeda, S., Ferraz-Amaro, I., López-Mejías, R. and González-Gay, M.A. 2022. Vascular endothelial

growth factor haplotypes are associated with severe ischaemic complications in giant cell arteritis regardless of the disease phenotype. *Clinical and experimental rheumatology*. **40**(4), pp.727–733.

Salvarani, C., Cantini, F. and Hunder, G.G. 2008. Polymyalgia rheumatica and giant-cell arteritis. *The Lancet*. **372**(9634), pp.234–245.

Tadjuidje, E., Wang, T. Sen, Pandey, R.N., Sumanas, S., Lang, R.A. and Hegde, R.S. 2012. The EYA Tyrosine Phosphatase Activity Is Pro-Angiogenic and Is Inhibited by Benzbromarone. *PLOS ONE*. **7**(4), pp.e34806-.

Vali-Pour, M., Park, S., Espinosa-Carrasco, J., Ortiz-Martínez, D., Lehner, B. and Supek, F. 2022. The impact of rare germline variants on human somatic mutation processes. *Nature Communications*. **13**(1), p.3724.

Venkataraman, G.R. and Rivas, M.A. 2019. Rare and common variant discovery in complex disease: the IBD case study. *Human Molecular Genetics*. **28**(R2), pp.R162–R169.

Wilson, D.J. 2019. The harmonic mean p-value for combining dependent tests. *Proceedings of the National Academy of Sciences*. **116**(4), pp.1195–1200.

Zhang, X., Basile, A.O., Pendergrass, S.A. and Ritchie, M.D. 2019. Real world scenarios in rare variant association analysis: the impact of imbalance and sample size on the power in silico. *BMC Bioinformatics*. **20**(1), p.46.

Zhao, Z., Bi, W., Zhou, W., VandeHaar, P., Fritsche, L.G. and Lee, S. 2020. UK Biobank Whole-Exome Sequence Binary Phenome Analysis with Robust Region-Based Rare-Variant Test. *The American Journal of Human Genetics*. **106**(1), pp.3–12.

Zhou, W., Bi, W., Zhao, Z., Dey, K.K., Jagadeesh, K.A., Karczewski, K.J., Daly, M.J., Neale, B.M. and Lee, S. 2022. SAIGE-GENE+ improves the efficiency and accuracy of set-based rare variant association tests. *Nature Genetics*. **54**(10), pp.1466–1469.

Zuk, O., Schaffner, S.F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., Daly, M.J., Neale, B.M., Sunyaev, S.R. and Lander, E.S. 2014. Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences*. **111**(4), pp.E455–E464.

Appendix

Appendix A: pLoF Mask Analyses

Appendix Table A1.

Top results from gene-based burden tests using the predicted loss-of-function (pLoF) only mask across the three clinical phenotype comparisons: (a) patients with no cranial ischaemic complications (CICs) vs permanent CICs, (b) patients with no CICs vs transient + permanent CICs, (c) patients with none + transient CICs vs permanent CICs.

(a)

SetID	P.value	N.Marker.A	N.Marker.T	MAC	m	Method.bi	MAP	Bonf	FDR
ENSG00000108395	0.00357	4	4	4	4	ER	0.00357	1	0.80911
ENSG00000119912	0.003942	3	3	4	4	ER	0.003942	1	0.80911
ENSG00000163728	0.004035	3	3	4	4	ER	0.004035	1	0.80911
ENSG00000104805	0.004279	1	1	4	4	ER	0.004279	1	0.80911
ENSG00000204161	0.00455	1	1	8	8	ER	1.97E-05	1	0.80911
ENSG00000130270	0.004811	4	4	6	6	ER	1.76E-04	1	0.80911
ENSG00000138036	0.005511	2	2	4	4	ER	0.005511	1	0.80911
ENSG00000185238	0.005889	2	2	4	4	ER	0.005889	1	0.80911
ENSG00000198920	0.006035	4	4	4	4	ER	0.006035	1	0.80911

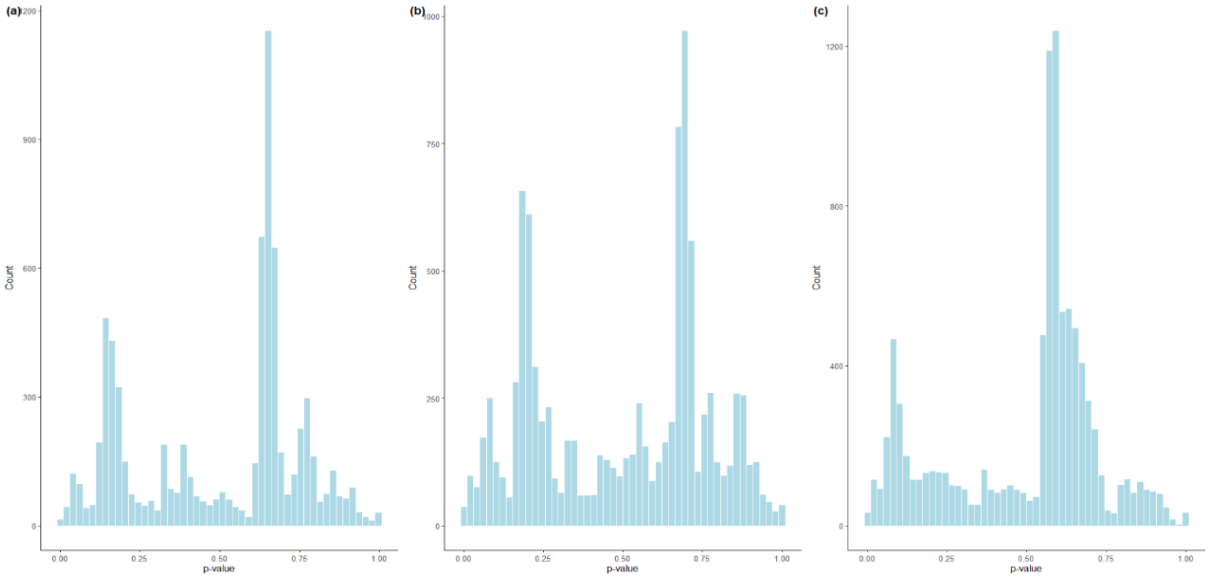
(b)

SetID	P.value	N.Marker.A	N.Marker.T	MAC	m	Method.bi	MAP	Bonf	FDR
ENSG00000183091	1.77E-04	10	10	12	12	ER	1.21E-06	1	0.820682
ENSG00000175894	3.27E-04	7	7	12	12	ER	5.10E-06	1	0.820682
ENSG00000143520	5.07E-04	8	8	24	24	ER.A	-1	1	0.820682
ENSG00000153015	6.34E-04	3	3	18	18	ER	8.01E-09	1	0.820682
ENSG00000070031	0.001064	3	3	13	12	ER	4.16E-06	1	0.820682
ENSG00000104321	0.001244	6	6	6	6	ER	0.001244	1	0.820682
ENSG00000172456	0.00203	6	6	17	17	ER	6.13E-08	1	0.820682
ENSG00000225781	0.002118	2	2	11	11	ER	1.87E-05	1	0.820682
ENSG00000182022	0.002574	5	5	19	19	ER	5.96E-09	1	0.820682

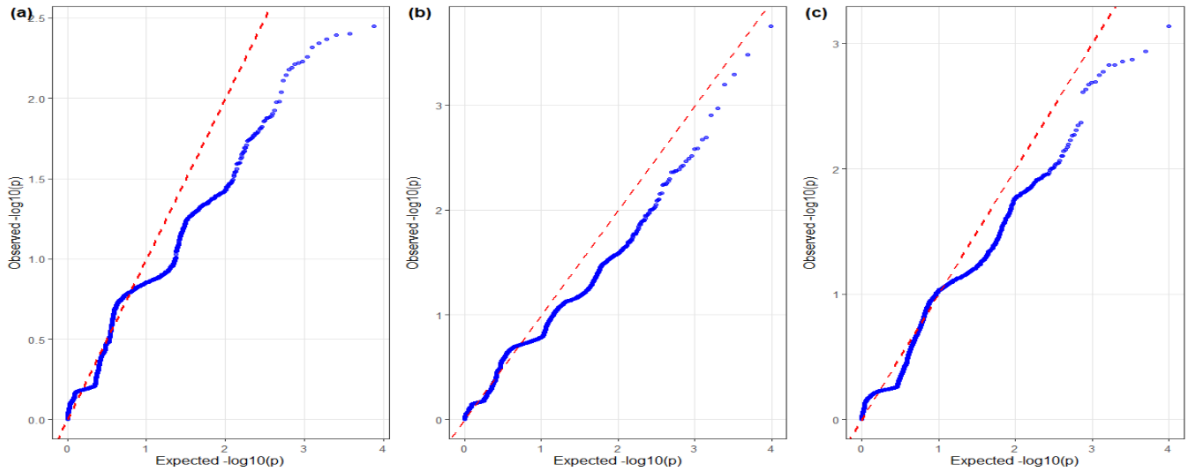
(c)

SetID	P.value	N.Marker.A	N.Marker.T	MAC	m	Method.bi	MAP	Bonf	FDR
ENSG00000163728	7.32E-04	4	4	5	5	ER	4.01E-05	1	0.800222
ENSG00000198920	0.001159	5	5	5	5	ER	8.94E-05	1	0.800222
ENSG00000131778	0.00135	3	3	3	3	ER	0.00135	1	0.800222
ENSG00000174358	0.001386	8	8	12	12	ER	1.74E-10	1	0.800222
ENSG00000197808	0.001486	3	3	4	4	ER	2.30E-04	1	0.800222
ENSG00000184530	0.001488	2	2	3	3	ER	0.001488	1	0.800222
ENSG00000108395	0.001676	5	5	5	5	ER	7.48E-05	1	0.800222
ENSG00000164649	0.001784	3	3	3	3	ER	0.001784	1	0.800222
ENSG00000260691	0.00202	3	3	18	18	ER	1.12E-15	1	0.800222

Appendix Figure A1.
Histograms of p-values from gene-based rare variant burden testing using the predicted loss-of-function (pLoF only) mask across the three clinical phenotype comparisons: (a) no cranial ischaemic complications (CICs) vs permanent CICs, (b) no CICs vs transient + permanent CICs, (c) none + transient CICs vs permanent CICs.



Appendix Figure A2.
Quantile–quantile (Q–Q) plots of observed versus expected $-\log_{10}(p)$ values from gene-based rare variant burden testing using the *predicted loss-of-function (pLF only)* mask across the three clinical phenotype comparisons: (a) no cranial ischaemic complications (CICs) vs permanent CICs, (b) no CICs vs transient + permanent CICs, (c) none + transient CICs vs permanent CICs.



Appendix B: pLOF + Missense + Synonymous Mask Analyses

Appendix Table B1.

Top results from gene-based burden tests using the pLoF + missense + synonymous mask across the three clinical phenotype comparisons: (a) no cranial ischaemic complications (CICs) vs permanent CICs, (b) no CICs vs transient + permanent CICs, (c) none + transient CICs vs permanent CICs.

(a)

SetID	P.value	N.Marker.	N.Marker. MAC	m		Method.b	MAP	Bonf	FDR
ENSG00000175664	4.24E-05	9	9	16	16	ER	2.11E-09	0.756049	0.444922
ENSG00000205030	5.91E-05	8	7	13	13	ER	8.02E-08	1	0.444922
ENSG00000183648	7.49E-05	7	7	15	15	ER	4.72E-09	1	0.444922
ENSG00000132164	2.12E-04	19	19	48	48	QA	-1	1	0.943273
ENSG00000117616	2.74E-04	9	9	26	25	ER.A	-1	1	0.961057
ENSG00000113790	4.50E-04	24	24	39	38	ER.A	-1	1	0.961057
ENSG00000188883	4.96E-04	14	14	23	23	QA	-1	1	0.961057
ENSG00000048707	5.06E-04	89	89	175	155	QA	-1	1	0.961057
ENSG00000158125	5.46E-04	56	56	101	95	QA	-1	1	0.961057

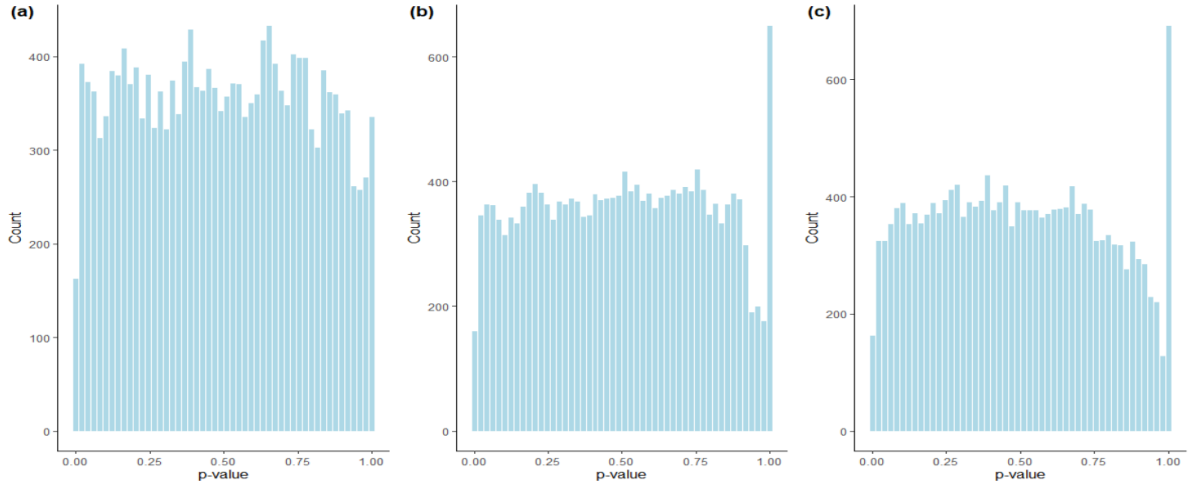
(b)

SetID	P.value	N.Marker.	N.Marker. MAC	m		Method.b	MAP	Bonf	FDR
ENSG00000197892	3.45E-05	78	78	160	149	QA	-1	0.617677	0.617677
ENSG00000153015	2.96E-04	10	10	43	27	QA	-1	1	0.923613
ENSG00000171346	3.75E-04	21	21	31	31	ER.A	-1	1	0.923613
ENSG00000051596	3.78E-04	12	11	38	38	ER.A	-1	1	0.923613
ENSG00000099194	4.85E-04	10	10	17	17	ER	1.25E-08	1	0.923613
ENSG00000188732	6.05E-04	13	13	73	71	QA	-1	1	0.923613
ENSG00000249158	6.17E-04	46	46	99	96	QA	-1	1	0.923613
ENSG00000164867	6.24E-04	70	70	183	168	QA	-1	1	0.923613
ENSG00000215343	7.30E-04	10	8	11	11	ER	1.44E-05	1	0.923613

(c)

SetID	P.value	N.Marker.	N.Marker. MAC	m		Method.b	MAP	Bonf	FDR
ENSG00000205030	2.78E-05	11	10	19	19	ER	1.33E-16	0.498604	0.498604
ENSG00000160789	7.19E-05	28	28	86	81	QA	-1	1	0.64489
ENSG00000092067	1.80E-04	16	16	43	41	QA	-1	1	0.802515
ENSG00000221826	1.81E-04	35	35	67	64	QA	-1	1	0.802515
ENSG00000150045	4.35E-04	6	6	11	11	ER	3.04E-09	1	0.802515
ENSG00000100888	4.56E-04	62	62	163	146	QA	-1	1	0.802515
ENSG00000275591	4.77E-04	48	48	109	103	QA	-1	1	0.802515
ENSG00000131142	5.05E-04	4	4	4	4	ER	5.05E-04	1	0.802515
ENSG00000119715	5.17E-04	11	11	31	31	ER.A	-1	1	0.802515

Appendix Figure B1.
Histograms of p-values from gene-based rare variant burden testing using the pLoF + missense + synonymous mask across the three clinical phenotype comparisons: (a) no CICs vs permanent CICs, (b) no CICs vs transient + permanent CICs, (c) none + transient CICs vs permanent CICs.



Appendix Figure B2.
Quantile–quantile (Q–Q) plots from gene-based rare variant burden testing using the pLoF + missense + synonymous mask across the three clinical phenotype comparisons: (a) no CICs vs permanent CICs, (b) no CICs vs transient + permanent CICs, (c) none + transient CICs vs permanent CICs.