Performing a principal component analysis and genome-wide association study on the Canadian Longitudinal Study on Aging self-reported hearing loss cohort

MacKenzie Wilke [1], Dr. Britt Drögemöller [1-3]
1. Department of Biochemistry and Medical Genetics, Winnipeg, Manitoba, 2. Children's Hospital of Manitoba, Winnipeg, Manitoba, 3. CancerCare Manitoba, Winnipeg, Manitoba

INTRODUCTION

•Age-related hearing loss has a high burden worldwide and has been shown to be genetically heterogenous with its genetic mechanisms largely unknown.

•The UKBiobank database has found promising genetic results for hearing loss etiology through genome-wide association studies (GWAS). The Canadian Longitudinal Study on Aging (CLSA)[1] data provides a Candian cohort to perform these analyses on.

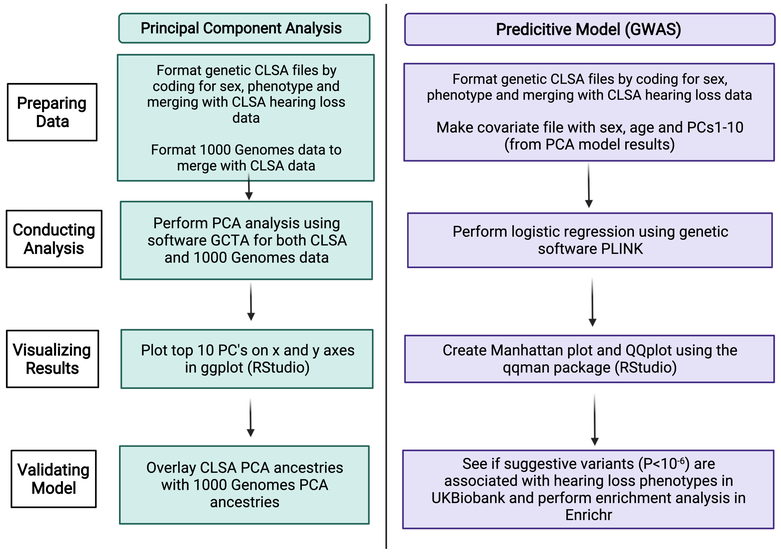
•Principal component analysis is a common quality control step in GWAS and reveals the data's intrinsic population structure[2]. 1000 Genomes data[3] has genetic ancestry data that can cluster with CLSA so validation of genetic substrata can occur.

OBJECTIVES

1. Perform a principal component analysis on CLSA participants with self-reported hearing loss
•Genetic ancestries intrinsic to participants will be identified

2. Conduct a GWAS on self-reported hearing loss CLSA participants
• Genetic variants associated with self-reported hearing loss will be discovered

METHODS



CLSA: Canadian Longitudinal Study on Aging, GWAS: Genome-wide association study, PCA: Principal Component Analysis, PCs: Principal Components, QQ: Quantile-quantile.

REFERENCES

1. Raina, P. S. *et al.* The Canadian longitudinal study on aging (CLSA). *Can. J. Aging Rev. Can. Vieil.* **28**, 221–229 (2009).
2. Uffelmann, E. *et al.* Genome-wide association studies. *Nat. Rev. Methods Primer* **1**, 1–21 (2021).
3. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

RESULTS

Figure 1. Top two principal components for the Candian Longitudinal Study on Aging self-reported hearing loss participants (case and control) overlaid with the top two principal components for 1000 Genomes. Ethnicity also refers to genetic ancestry.

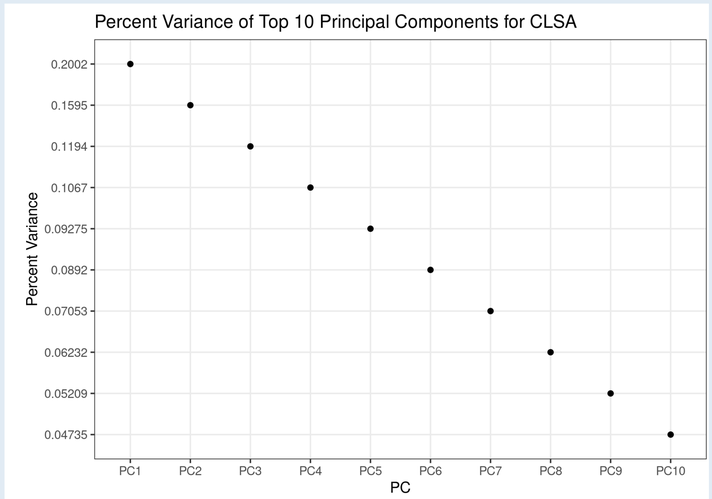


Figure 2. Percent variance of the top 10 principal components (PCs) in the Canadian Longitudinal Study on Aging (CLSA).

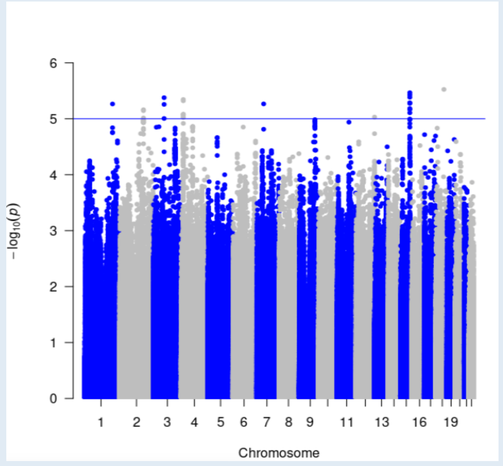


Figure 3. Manhattan plot of genome-wide association study results. Suggestive significance threshold is P<10^{-6}.

Gene	Associated genetic variants	P-value for hearing aid user phenotype in the UKBiobank	P-value for hearing difficulties phenotype in the UKBiobank
ISG20	rs12441297	$3.0x10^{-7}$	$6.5x10^{-6}$
	rs12441287	$7.6x10^{-7}$	$7.5x10^{-5}$
	rs6496519	$2.7x10^{-7}$	$7.4x10^{-6}$
	rs4128832	$2.1x10^{-7}$	$1.2x10^{-5}$
	rs11073804	$3.8x10^{-7}$	$1.2x10^{-5}$
	rs4293349	$1.3x10^{-6}$	$2.6x10^{-5}$
	rs10852113	$6.7x10^{-7}$	$2.1x10^{-5}$
	rs4293350	$6.4x10^{-7}$	$2.3x10^{-5}$

Table 1. Detailed information on genetic variants associated with *ISG20* in UKBiobank.

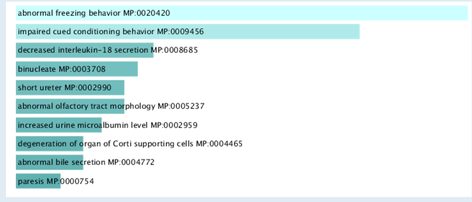


Figure 4. Enrichr enrichment analyses for genes that reached suggestive significance (*n*= 8).

CONCLUSIONS

•Ancestries in the CLSA were diverse and clustering with the correct 1000 Genomes ancestries. Europeans clustered the best and the top ten principal components explained all variance in the data.

•The gene, *ISG20*, was found to be signficant in the CLSA GWAS and has been previously associated with hearing aid users (average p-value= $5.7x10^{-7}$) and hearing diffculty (average p-value= $2.3x10^{-5}$) in the UKBiobank. Enrichment analyses found a phenotype relating to inner ear pathophysiology.

ACKNOWLEDGEMENTS

The VADA Program
Visual and Automated Disease Analytics Graduate Training Program
University of Manitoba
CIHR IRSC
Canadian Institutes of Health Research / Instituts de recherche en santé du Canada