Identifying genetic markers that modify risk of carbamazepine-induced Stevens-Johnson Syndrome

Robert Zeibich[1], Terence O'Brien[1,2,3,4], Piero Perucca[1,2,3,4,5,6], Patrick Kwan[*,1,2,3,4], Alison Anderson[*,1]

Introduction

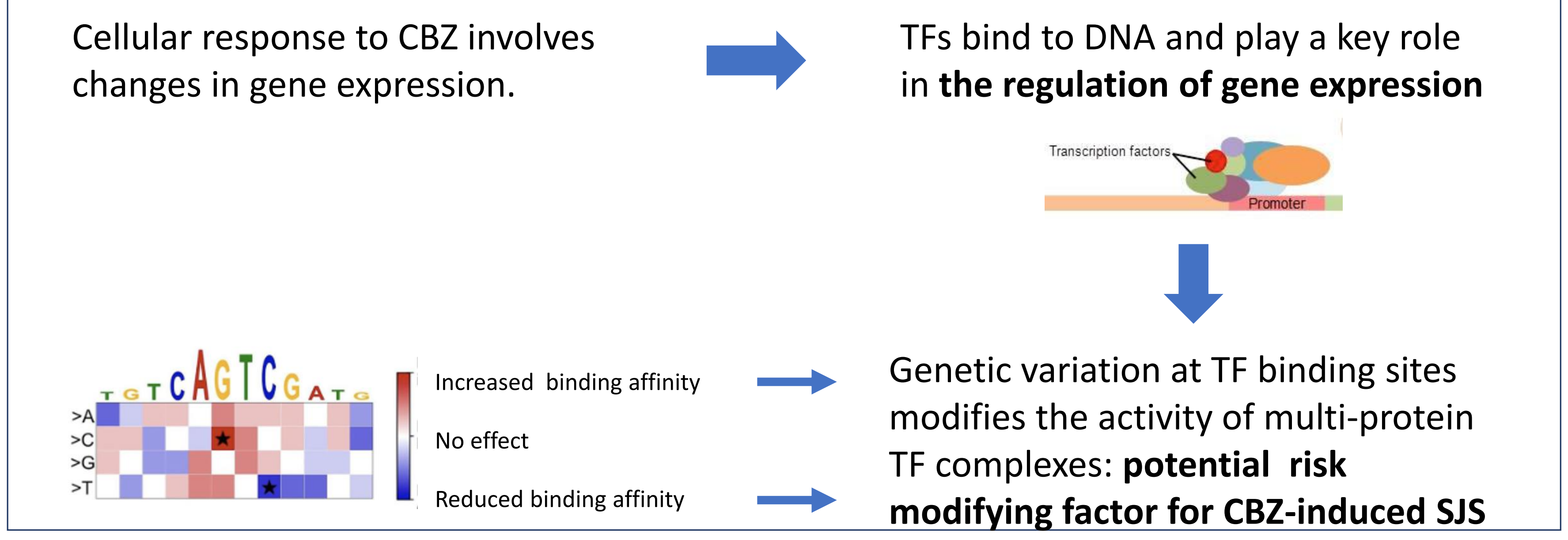
- Carbamazepine (CBZ) is a highly effective medication for people with epilepsy.
- Steven Johnson Syndrome (SJS,) a severe disorder of the skin, is a rare but life-threatening adverse reaction to CBZ treatment.

- In Han Chinese HLA-B*15:02 is associated with risk but is neither necessary nor sufficient so other genetic variations likely modify risk.

Objective: Identify additional genetic risk markers

Hypothesis

Hypothesis: Genetic variations in regions of the DNA where transcription factors (TF) bind modify risk of CBZ-induced SJS



Cellular response to CBZ involves changes in gene expression. → TFs bind to DNA and play a key role in **the regulation of gene expression**

Increased binding affinity → Genetic variation at TF binding sites modifies the activity of multi-protein TF complexes: **potential risk modifying factor for CBZ-induced SJS**

No effect

Reduced binding affinity →

Methods

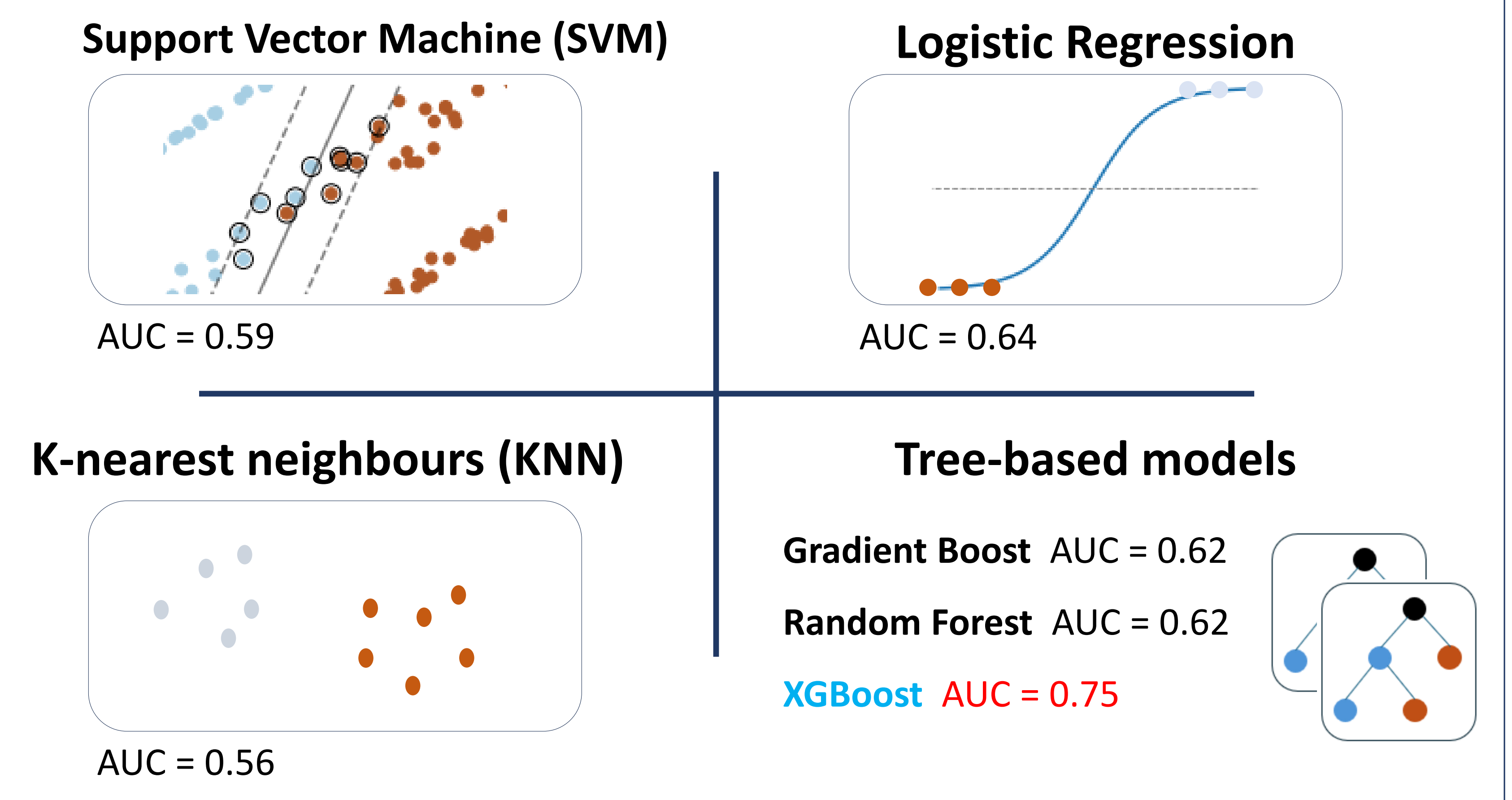
Step 1: Evaluate predictive performance of machine learning (ML) models

- Existing whole genome data obtained from 117 individuals with CBZ-induced SJS and 81 CBZ-tolerant.
- No one model is best in all situations. It depends on the underlying structure in the data
- Applied 6 ML models with 5-fold-cross validation to find the model with the best AUC for the knows risk region (HLA-B region in chromosome 6)

Step 2: Use the best model to systematically look for risk in regions of the DNA where TFs bind

- Expression quantitative trait loci (eQTL) data was used to find a signal specific to a tissue type
- TF binding sites were obtained from the Gene Transcription Regulation Database (GTRD)

Results step 1: Best model



Support Vector Machine (SVM)

AUC = 0.59

Logistic Regression

AUC = 0.64

K-nearest neighbours (KNN)

AUC = 0.56

Tree-based models

Gradient Boost AUC = 0.62

Random Forest AUC = 0.62

XGBoost AUC = 0.75

Result: XGBoost performed with an Area Under the Curve (AUC) 0.75 the best

Results step 2: Identifying risk modifying regions

ANALYSIS & RESULTS

Risk in additional chromosomes
Chr6 (AUC = 0.79)
Chr1 (AUC = 0.78)
Chr7 (AUC = 0.77)

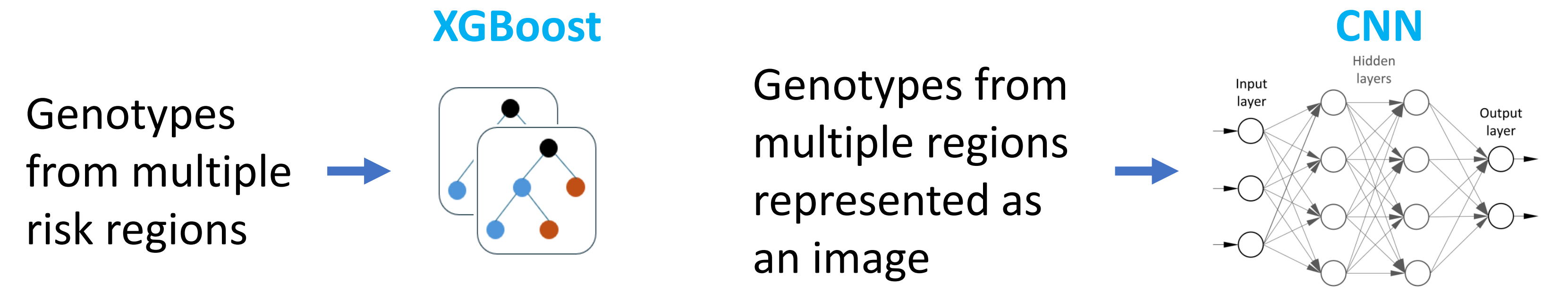
eQTL impacted tissues
Thyroid, Small Intestine and Colon Sigmoid (AUC = 0.81)
Colon Transverse (AUC = 0.80)

TF binding sites nearby HLA-B and HLA-A
HLA-B (AUC = 0.67)
HLA-A (AUC= 0.59)

TFs binding in chromosomes 1 to 22
H2AZ1 (AUC = 0.87)
ESR1 (AUC = 0.83)
CTCF (AUC = 0.83)

Next steps: Use deep learning model to predict risk

Hypothesis: An image-based model (e.g., Convolutional Neural Network [CNN]) will capture relationships between risk modifying regions and outperform the XGBoost model



Genotypes from multiple risk regions → **XGBoost**

Genotypes from multiple regions represented as an image → **CNN**

Conclusion

This study has identified regions in the DNA where genetic variations potentially modify risk of CBZ-induced SJS. This work will facilitate future targeted deep-learning approaches to identify robust risk biomarkers.

Acknowledgement

We thank the patients, clinicians and researchers for participating and making the study possible.

This study was supported by a National Health and Medical Research Council (NHMRC) Project Gant (APP1103979).

References

1. Alipanahi B, et al. Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. Nature biotechnology. 2015 Aug;33(8):831-8.
2. SJS image from https://dermnetnz.org/topics/sjs-ten-images licensed under Creative Commons Attribution-NonCommercial-NoDerivs 3.0 (New Zealand)

Affiliations

1. Department of Neuroscience, Central Clinical School, Monash University; 2. Department of Neurology, The Alfred Hospital; 3. Department of Neurology, The Royal Melbourne Hospital; 4. Department of Medicine, The Royal Melbourne Hospital, The University of Melbourne; 5. Department of Medicine, Austin Health, The University of Melbourne; 6. Comprehensive Epilepsy Program, Department of Neurology, Austin Health; *Co-senior authors