VTT-Net: Learning Viral Tissue Tropism Using Graph Neural Networks

Dr. Haley Stone
MRC University of Glasgow Centre for Virus Research
VIBE Research Lab

Medical Research Council
University of Glasgow
CVR Centre for Virus Research

vibelab.co.uk

Background

- Viral tissue tropism shapes disease severity, transmission routes, host range, and spillover risk
- Tissue tropism remains underrepresented in comparative virology and host-range prediction because data is difficult to collect, curate, and standardise
- Why this is difficult:
  - Sparse and biased tissue-level evidence
  - Limited experimental validation across hosts and viruses
  - Strong reporting bias toward clinically sampled tissues
  - Host, tissue, and virus effects are biologically coupled rather than independent

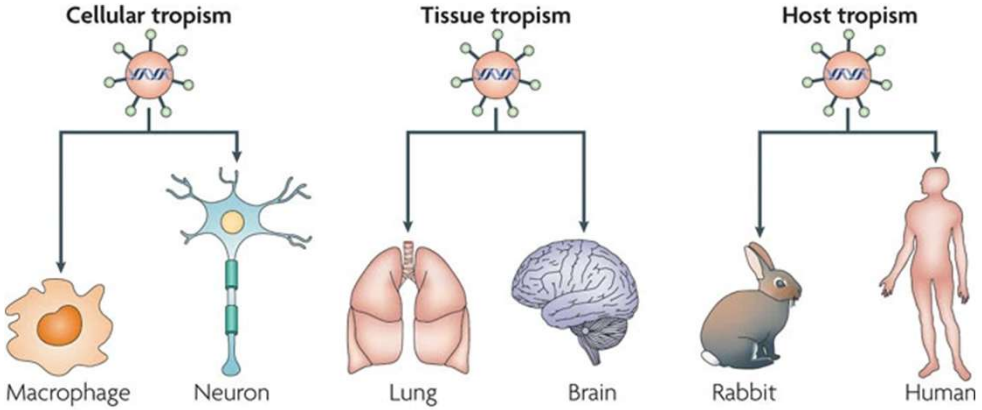
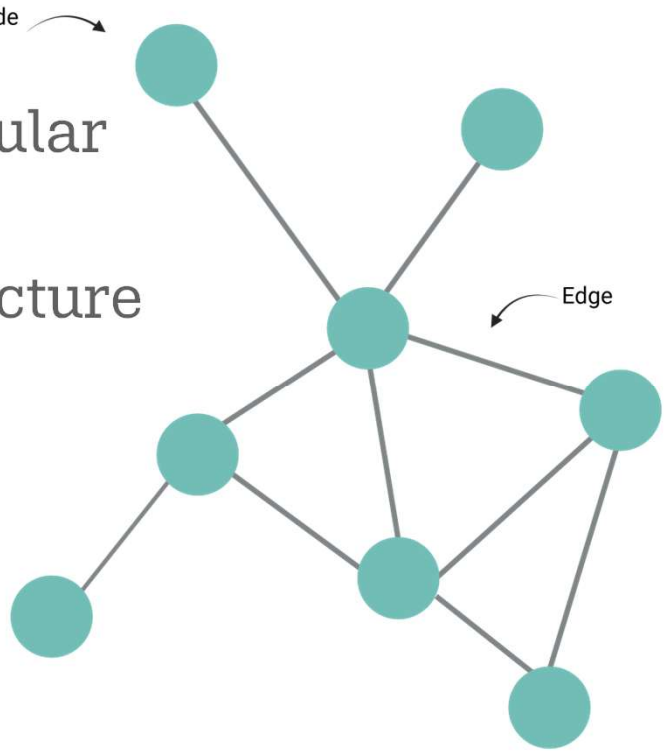


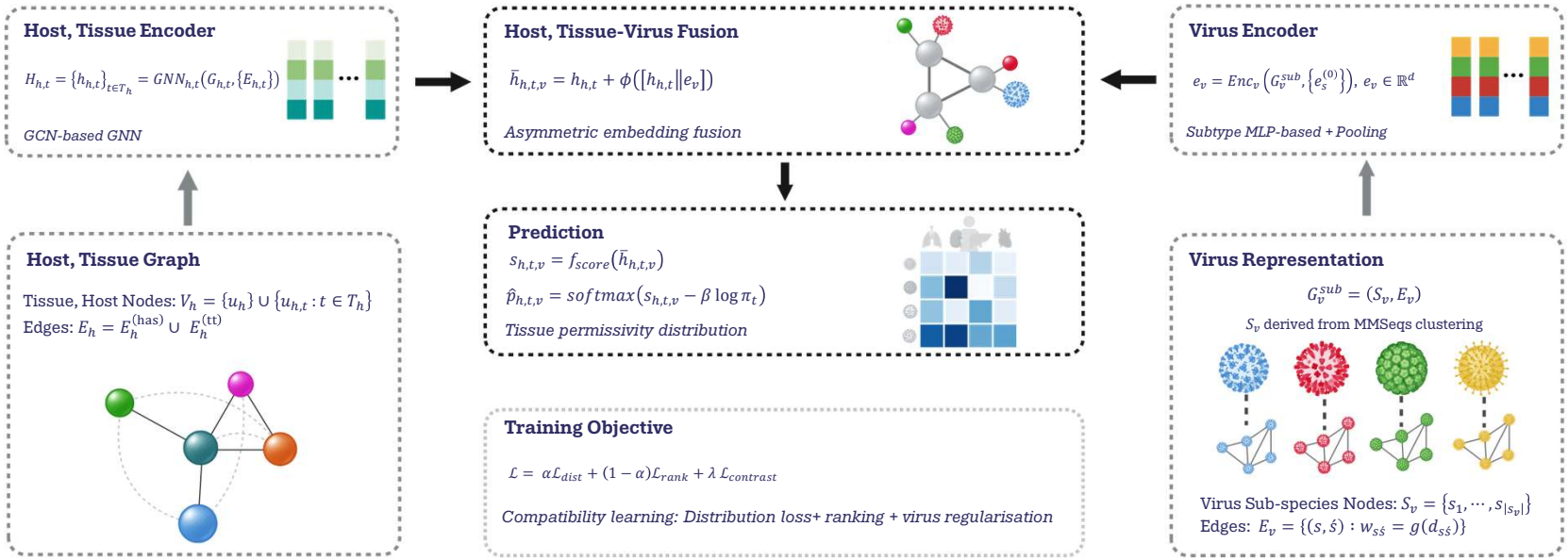
Figure adapted from: McFadden, G., Mohamed, M., Rahman, M. et al. Cytokine determinants of viral tropism. Nat Rev Immunol 9, 645–655 (2009). https://doi.org/10.1038/nri2623

Why Graph Neural Networks (GNNs)?

- Biological systems are relational
- Tissues are connected through shared molecular profiles
- Viruses are related through evolutionary structure
- GNNs allow:
  - Information sharing across related entities
  - Learning from indirect evidence
  - Modelling system level interactions
  - Integration of heterogeneous biological structure



Framework



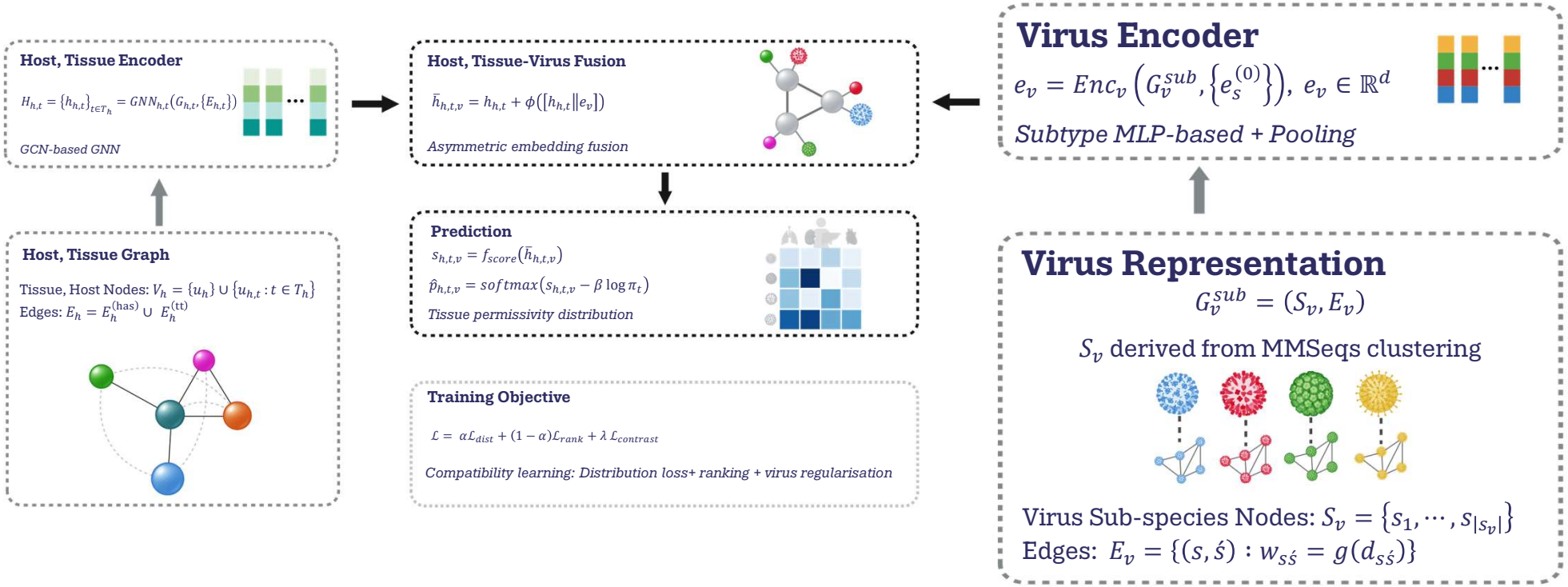
Medical Research Council

University of Glasgow

CVR Centre for Virus Research

vibelab.co.uk

Virus Representation



Medical Research Council

University of Glasgow

CVR Centre for Virus Research

vibelab.co.uk

In the model: what is the unit of the virus?

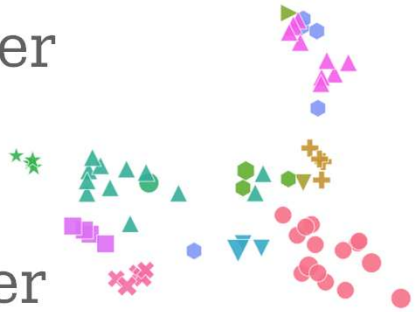
The model does not assume that each taxonomic species is homogeneous

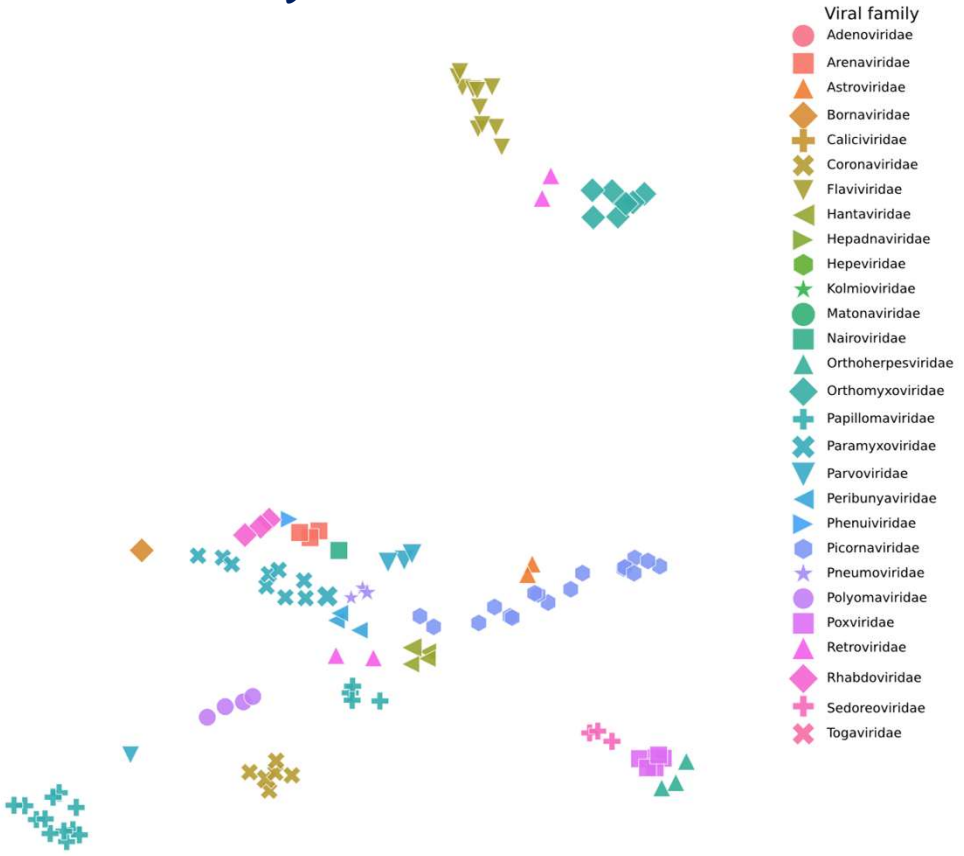
	Non-Segmented	Segmented
Accession	Complete Sequence	Individual segment belonging to the same assembly number
MMSeqs Clustering	Same taxonomic species	Separately for each segment within same taxonomic species
What unit do we consider?	Sequences in separate clusters are treated as distinct units in the model	Assemblies are separated into distinct units if every segment does not cluster

Virus Representation

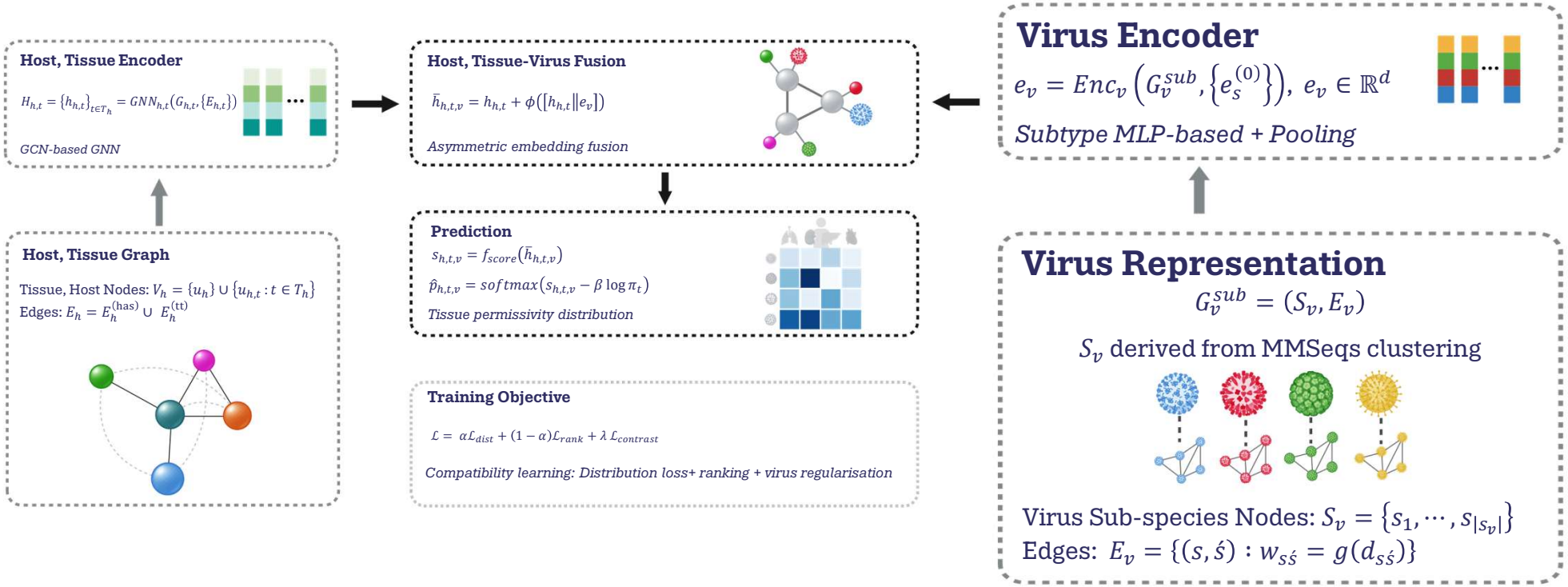
$$e_v = \frac{1}{|S_v|} \sum_{s \in S_v} e_s$$

- Within-virus subgraphs:
  - Nodes: viral genome sequences
  - Edges: patristic distance
  - Aggregation: pooled into a unit-level virus embedding
- Encoder: Multilayer Perceptron (MLP)
- Units not connected together





Virus Representation



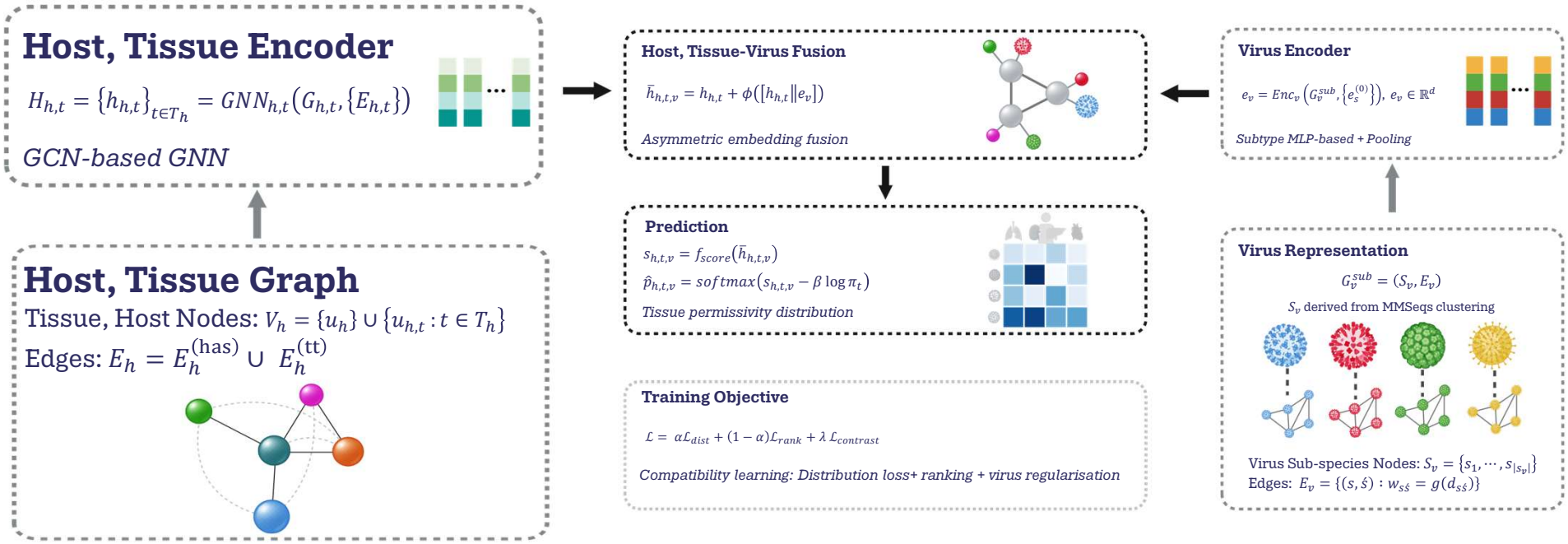
Medical Research Council

University of Glasgow

CVR Centre for Virus Research

vibelab.co.uk

Host, Tissue Graph

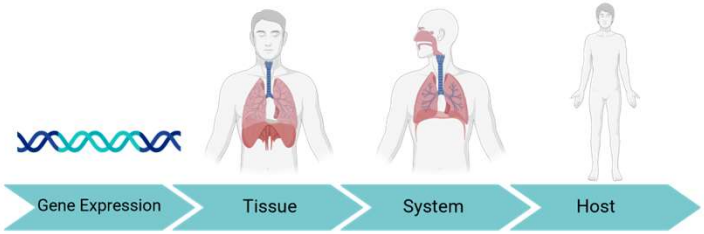


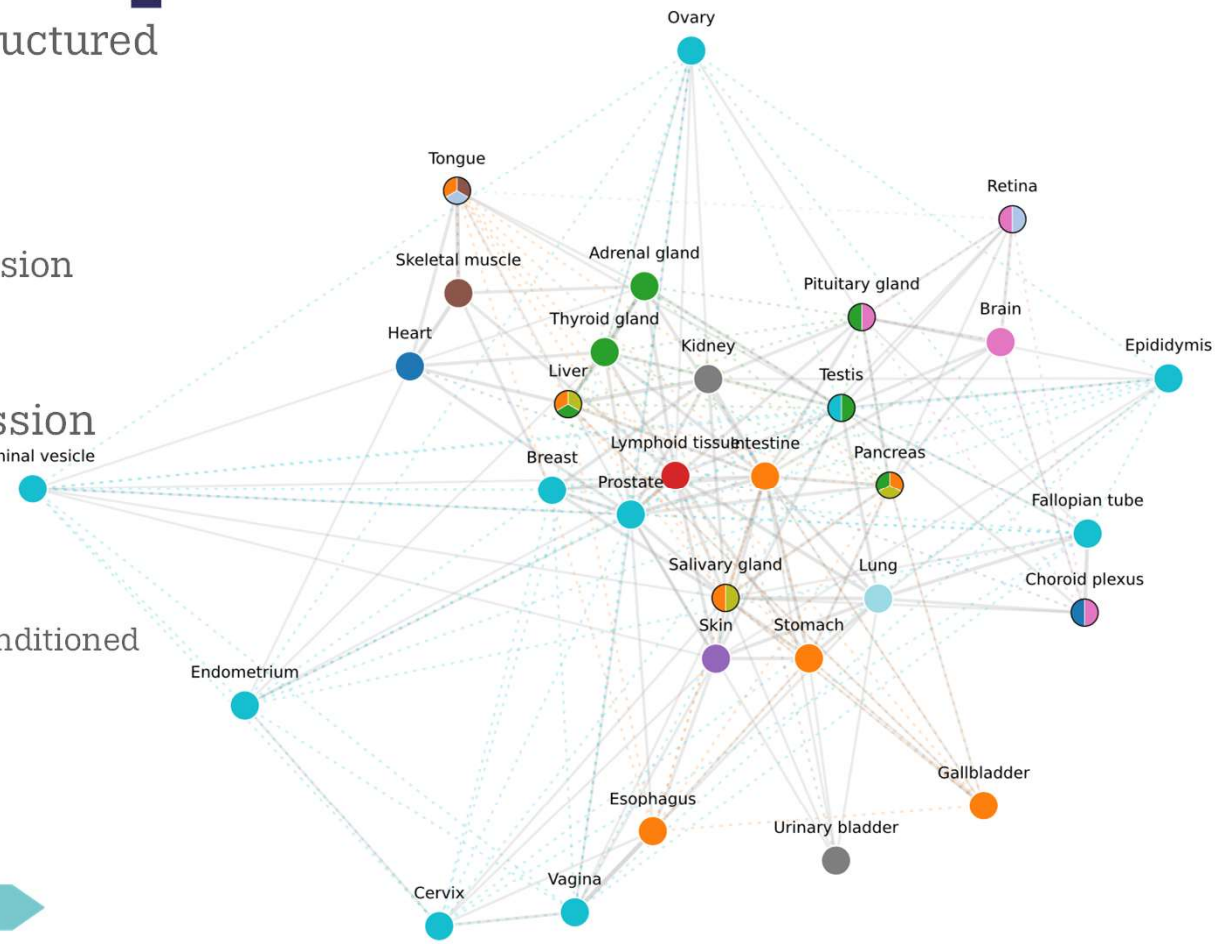
Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Host-Tissue Graph

- Tissues are modelled as a structured system within each host
- Nodes: host, tissues
- Edges:
  - kNN similarity -- gene expression from Human Tissue Atlas [1]
  - Host-tissue membership
- Features: compressed expression profiles
- Encoder:
  - Multi-layer Graph Convolutional Network (GCN)
  - Learns tissue representations conditioned on neighbouring structure

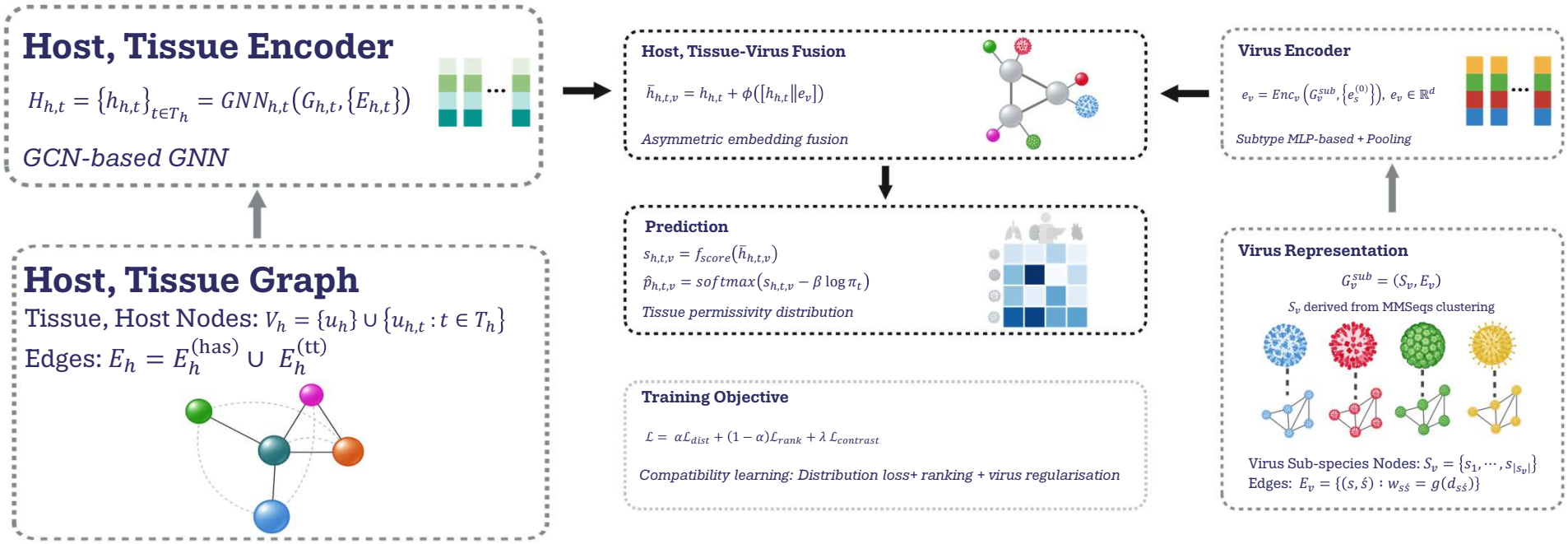




1. Mathias Uhlén et al. ,Tissue-based map of the human proteome. Science347,1260419(2015). DOI:10.1126/science.1260419

vibelab.co.uk

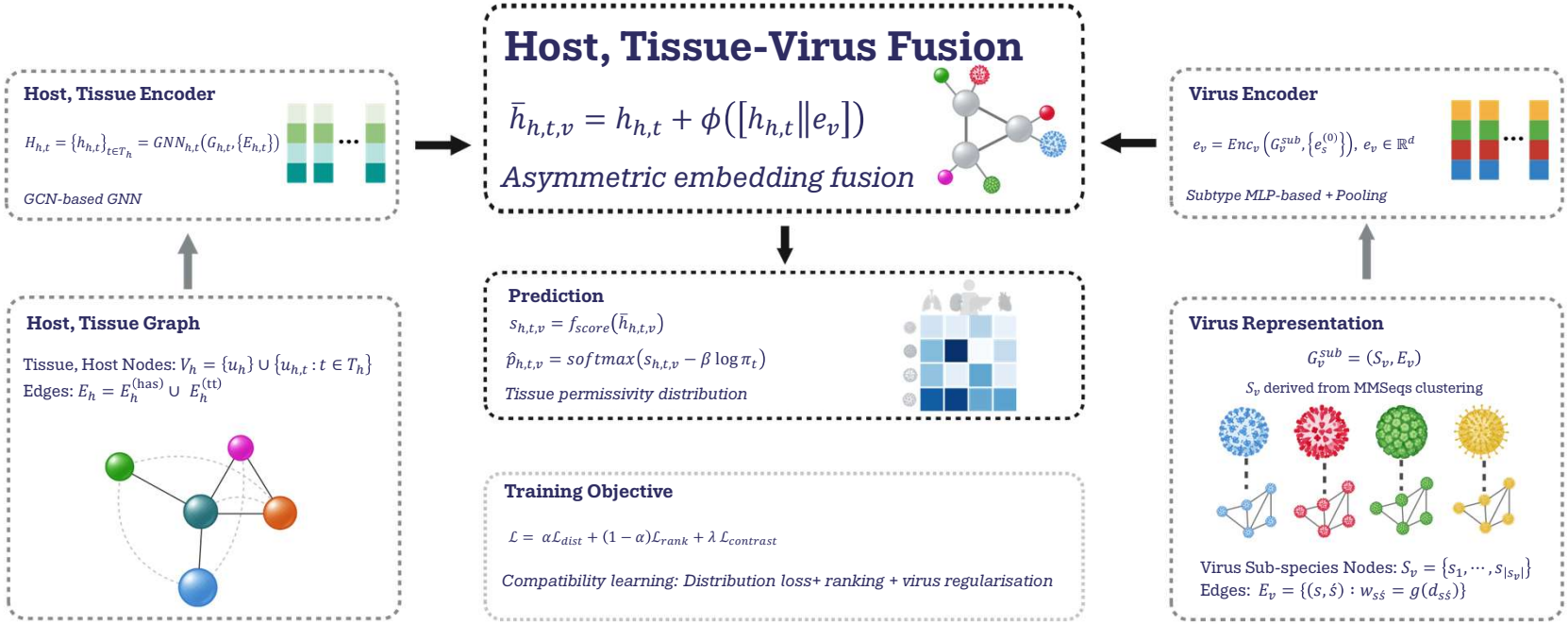
Host, Tissue Graph



Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Fusion



Medical Research Council
University of Glasgow
CVR Centre for Virus Research

vibelab.co.uk

Fusion

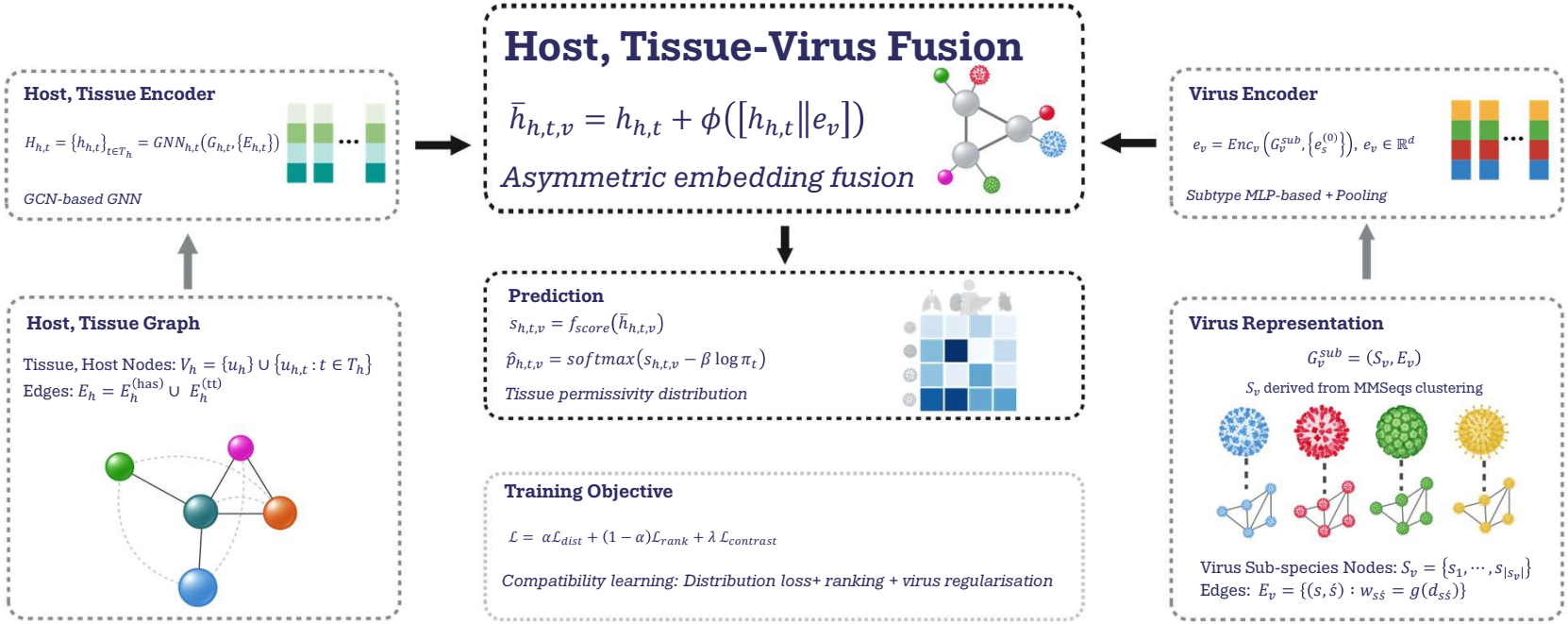
$$\bar{h}_{h,t,v} = h_{h,t} + \phi([h_{h,t} \| e_v])$$

- Tissue permissivity is virus-conditioned
- Same tissue, different virus → different outcome
  - *human-lung vs. human-lung + SARS-CoV-2*
- Virus embedding modifies each tissue representation
- Learns virus-specific tissue states, rather than fixed virus-tissue edge scores

Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

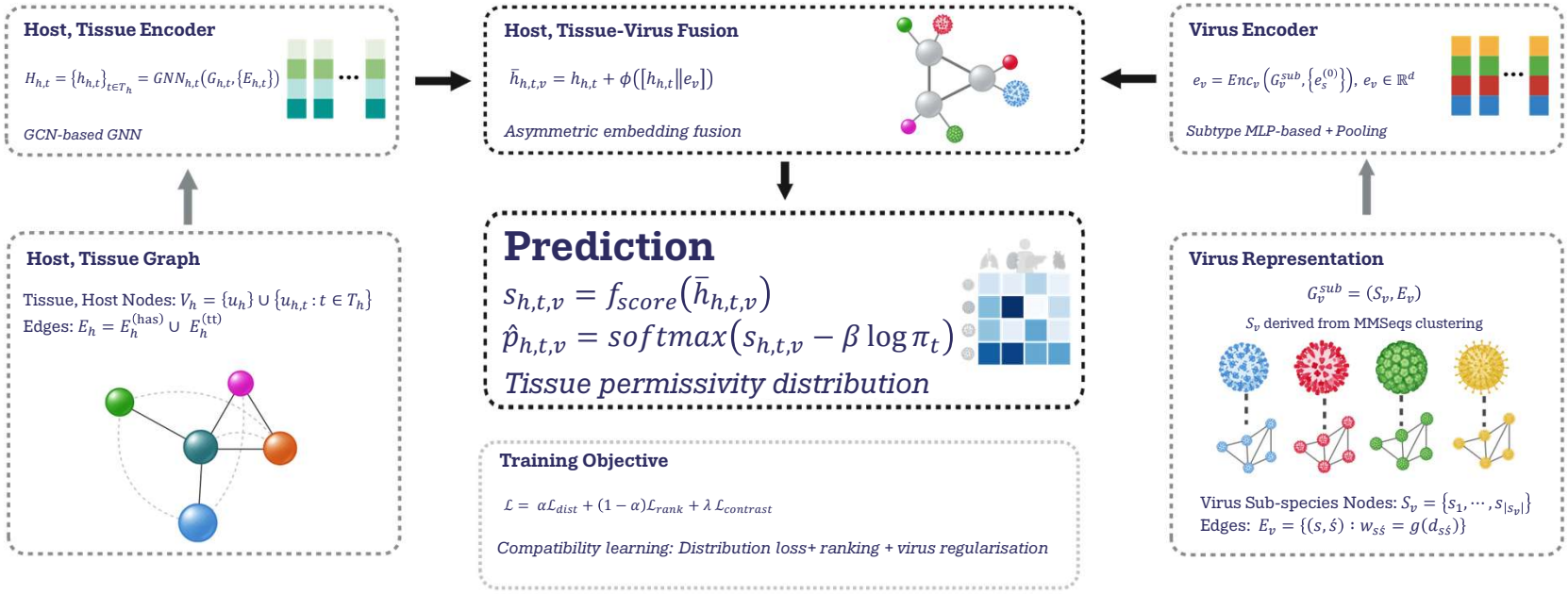
Fusion



Medical Research Council
University of Glasgow
CVR Centre for Virus Research

vibelab.co.uk

Prediction



Medical Research Council

University of Glasgow

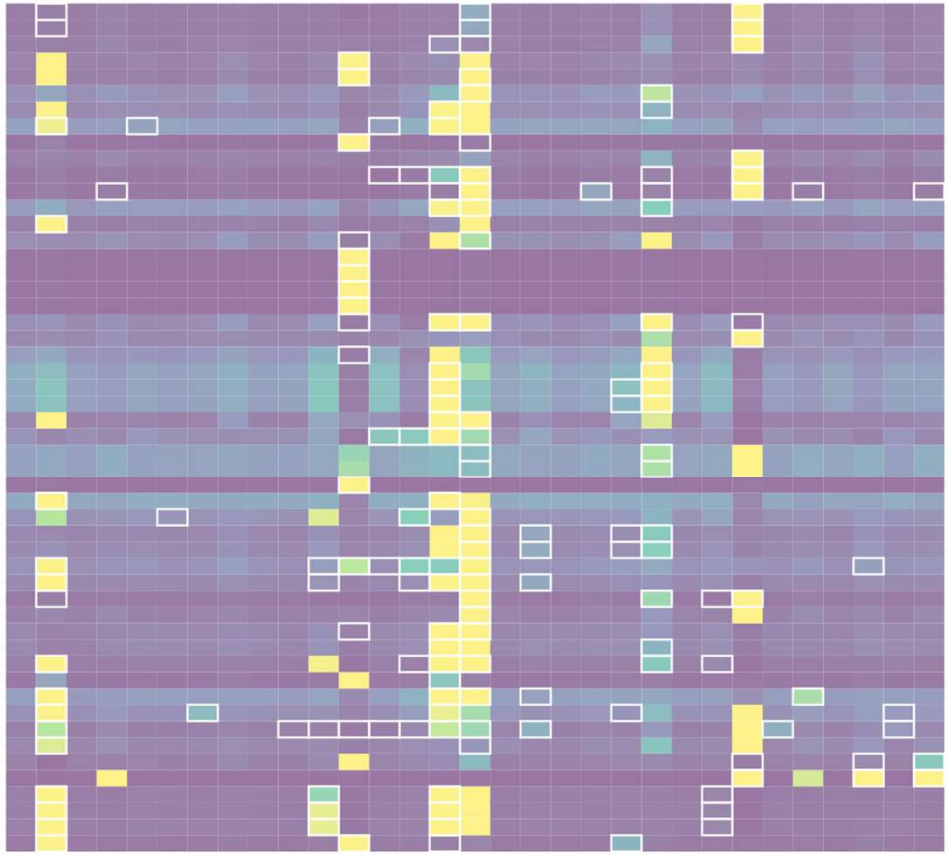
CVR Centre for Virus Research

vibelab.co.uk

Prediction: Tissue Permissivity

$$s_{h,t,v} = f_{score}(\bar{h}_{h,t,v})$$
$$\hat{p}_{h,t,v} = softmax_t(s_{h,t,v} - \beta \log \pi_t)$$

- Produces a ranked distribution over tissues
- Higher probability = higher relative permissivity
- Predictions are comparative across tissues, not absolute infection risk



UKRI Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

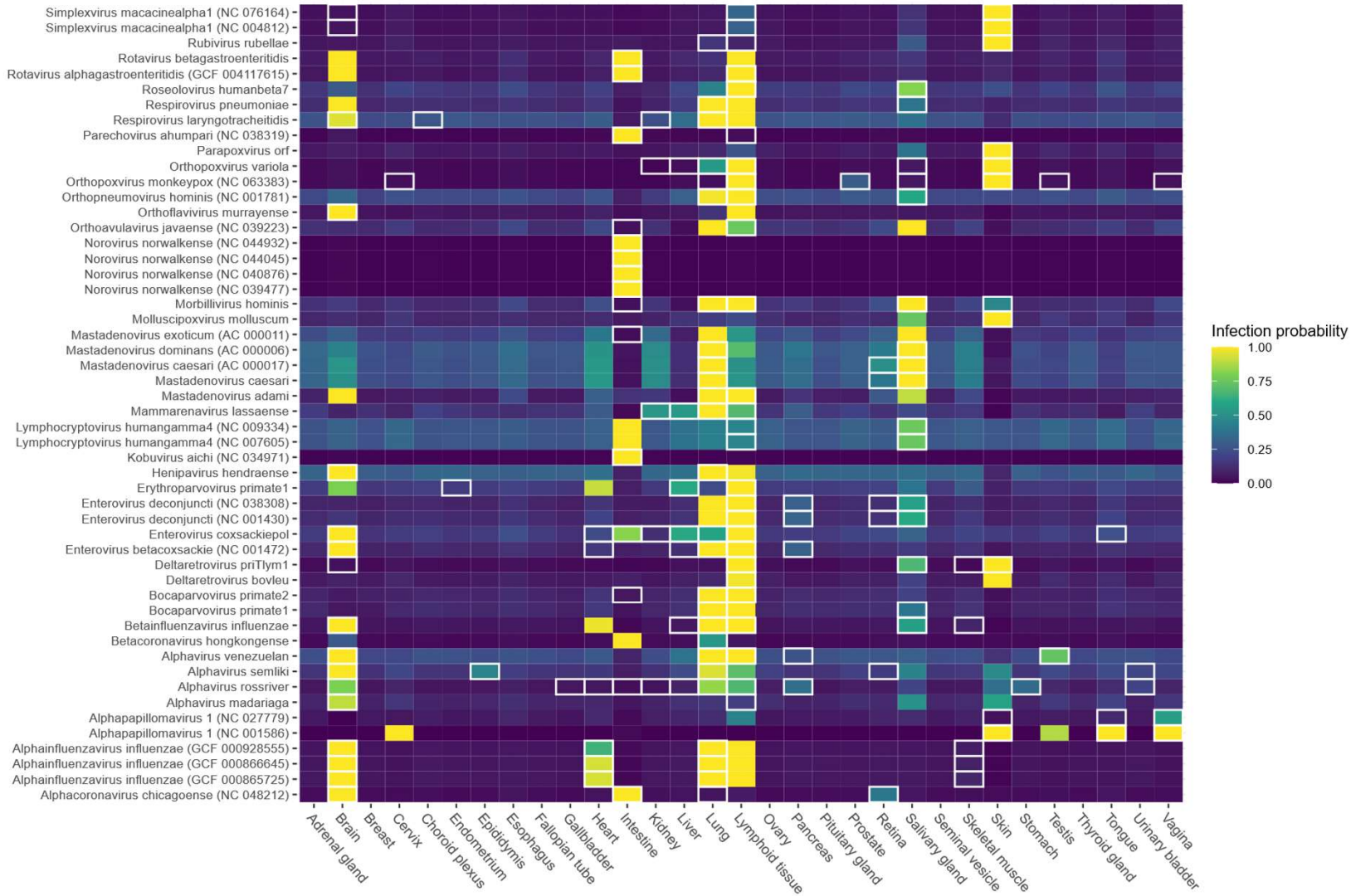
Performance

- Full model consistently performed best across ranking and classification metrics
- Virus features (sequence-based embeddings) seemed to be the most essential for discriminating tissue permissivity
- Tissue expression data improved performance beyond simple tissue labels
- Graph connectivity helped rank true permissive tissues more highly

Model setting	Top-k recall	Mean positive rank	AUC	Best pooled binary F1
Full model	**0.768**	**3.60**	**0.842**	**0.663**
No tissue expression	0.583	4.80	0.752	0.511
No tissue identity	0.620	5.22	0.761	0.562
No tissue-tissue edges	0.594	5.47	0.813	0.596
No virus features	0.367	9.82	0.688	0.392

Results

- The model recovers strong tissue-specific signals, especially for lung, brain, salivary gland, skin, lymphoid tissue, and intestine
- Lung was the best recovered tissue by held-out classification metrics; intestine was also consistently prioritised
- Predictions remain virus-specific, with distinct permissivity profiles across viral groups



Medical Research Council

University of Glasgow

CVR Centre for Virus Research

vibelab.co.uk

Virus Family Ablation

- Ablation effects were concentrated in specific virus families
- Large family-specific drops indicate that virus embeddings contribute directly to tissue ranking
- Neutral effects have been recorded in some families which suggests some reliance on shared tissued priors and sparse-label instability

Top 5 Virus Families	Recall	AUC	F1
Baseline	**0.768**	**0.842**	**0.663**
Caliciviridae	0.675	0.675	0.675
Poxviridae	0.717	0.717	0.717
Orthoherpesviridae	0.722	0.722	0.722
Sedoreoviridae	0.722	0.722	0.722
Adenoviridae	0.729	0.840	0.661

Embedding Space

UKRI Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Embedding Space

UKRI Medical Research Council
University of Glasgow
CVR Centre for Virus Research

vibelab.co.uk

Skin

Virus family / species

Adenoviridae
- Mastadenovirus adami
- Mastadenovirus caesari
- Mastadenovirus dominans
- Mastadenovirus exoticum

Arenaviridae
- Mammarenavirus lassaense

Caliciviridae
- Norovirus norwalkense

Coronaviridae
- Alphacoronavirus chicagoense
- Betacoronavirus hongkongense

Flaviviridae
- Orthoflavivirus murrayense

Matonaviridae
- Rubivirus rubellae

Orthoherpesviridae
- Lymphocryptovirus humangamma4
- Roseolovirus humanbeta7
- Simplexvirus macacinealpha1

Orthomyxoviridae
- Alphainfluenzavirus influenzae
- Betainfluenzavirus influenzae

Papillomaviridae
- Alphapapillomavirus 1

Paramyxoviridae
- Henipavirus hendraense
- Morbillivirus hominis
- Orthoavulavirus javaense
- Respirovirus laryngotracheitidis
- Respirovirus pneumoniae

Parvoviridae
- Bocaparvovirus primate1
- Bocaparvovirus primate2
- Erythroparvovirus primate1

Picornaviridae
- Enterovirus betacoxsackie
- Enterovirus coxsackiepol
- Enterovirus deconjuncti
- Kobuvirus aichi
- Parechovirus ahumpari

Pneumoviridae
- Orthopneumovirus hominis

Poxviridae
- Molluscipoxvirus molluscum
- Orthopoxvirus monkeypox
- Orthopoxvirus variola
- Parapoxvirus orf

Retroviridae
- Deltaretrovirus bovleu
- Deltaretrovirus priTlym1

Sedoreoviridae
- Rotavirus alphagastroenteritidis
- Rotavirus betagastroenteritidis

Togaviridae
- Alphavirus madariaga
- Alphavirus rossriver
- Alphavirus semliki
- Alphavirus venezuelan

UKRI Medical Research Council
University of Glasgow
CVR Centre for Virus Research

Intestine

Virus family / species

Adenoviridae
- Mastadenovirus adami
- Mastadenovirus caesari
- Mastadenovirus dominans
- Mastadenovirus exoticum

Arenaviridae
- Mammarenavirus lassaense

Caliciviridae
- Norovirus norwalkense

Coronaviridae
- Alphacoronavirus chicagoense
- Betacoronavirus hongkongense

Flaviviridae
- Orthoflavivirus murrayense

Matonaviridae
- Rubivirus rubellae

Orthoherpesviridae
- Lymphocryptovirus humangamma4
- Roseolovirus humanbeta7
- Simplexvirus macacinealpha1

Orthomyxoviridae
- Alphainfluenzavirus influenzae
- Betainfluenzavirus influenzae

Papillomaviridae
- Alphapapillomavirus 1

Paramyxoviridae
- Henipavirus hendraense
- Morbillivirus hominis
- Orthoavulavirus javaense
- Respirovirus laryngotracheitidis
- Respirovirus pneumoniae

Parvoviridae
- Bocaparvovirus primate1
- Bocaparvovirus primate2
- Erythroparvovirus primate1

Picornaviridae
- Enterovirus betacoxsackie
- Enterovirus coxsackiepol
- Enterovirus deconjuncti
- Kobuvirus aichi
- Parechovirus ahumpari

Pneumoviridae
- Orthopneumovirus hominis

Poxviridae
- Molluscipoxvirus molluscum
- Orthopoxvirus monkeypox
- Orthopoxvirus variola
- Parapoxvirus orf

Retroviridae
- Deltaretrovirus bovleu
- Deltaretrovirus priTlym1

Sedoreoviridae
- Rotavirus alphagastroenteritidis
- Rotavirus betagastroenteritidis

Togaviridae
- Alphavirus madariaga
- Alphavirus rossriver
- Alphavirus semliki
- Alphavirus venezuelan

UKRI Medical Research Council
University of Glasgow
CVR Centre for Virus Research

vibelab.co.uk

Retina

Virus family / species

Adenoviridae
- Mastadenovirus adami
- Mastadenovirus caesari
- Mastadenovirus dominans
- Mastadenovirus exoticum

Arenaviridae
- Mammarenavirus lassaense

Caliciviridae
- Norovirus norwalkense

Coronaviridae
- Alphacoronavirus chicagoense
- Betacoronavirus hongkongense

Flaviviridae
- Orthoflavivirus murrayense

Matonaviridae
- Rubivirus rubellae

Orthoherpesviridae
- Lymphocryptovirus humangamma4
- Roseolovirus humanbeta7
- Simplexvirus macacinealpha1

Orthomyxoviridae
- Alphainfluenzavirus influenzae
- Betainfluenzavirus influenzae

Papillomaviridae
- Alphapapillomavirus 1

Paramyxoviridae
- Henipavirus hendraense
- Morbillivirus hominis
- Orthoavulavirus javaense
- Respirovirus laryngotracheitidis
- Respirovirus pneumoniae

Parvoviridae
- Bocaparvovirus primate1
- Bocaparvovirus primate2
- Erythroparvovirus primate1

Picornaviridae
- Enterovirus betacoxsackie
- Enterovirus coxsackiepol
- Enterovirus deconjuncti
- Kobuvirus aichi
- Parechovirus ahumpari

Pneumoviridae
- Orthopneumovirus hominis

Poxviridae
- Molluscipoxvirus molluscum
- Orthopoxvirus monkeypox
- Orthopoxvirus variola
- Parapoxvirus orf

Retroviridae
- Deltaretrovirus bovleu
- Deltaretrovirus priTlym1

Sedoreoviridae
- Rotavirus alphagastroenteritidis
- Rotavirus betagastroenteritidis

Togaviridae
- Alphavirus madariaga
- Alphavirus rossriver
- Alphavirus semliki
- Alphavirus venezuelan

UKRI Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Urinary Bladder

Virus family / species

Adenoviridae
- Mastadenovirus adami
- Mastadenovirus caesari
- Mastadenovirus dominans
- Mastadenovirus exoticum

Arenaviridae
- Mammarenavirus lassaense

Caliciviridae
- Norovirus norwalkense

Coronaviridae
- Alphacoronavirus chicagoense
- Betacoronavirus hongkongense

Flaviviridae
- Orthoflavivirus murrayense

Matonaviridae
- Rubivirus rubellae

Orthoherpesviridae
- Lymphocryptovirus humangamma4
- Roseolovirus humanbeta7
- Simplexvirus macacinealpha1

Orthomyxoviridae
- Alphainfluenzavirus influenzae
- Betainfluenzavirus influenzae

Papillomaviridae
- Alphapapillomavirus 1

Paramyxoviridae
- Henipavirus hendraense
- Morbillivirus hominis
- Orthoavulavirus javaense
- Respirovirus laryngotracheitidis
- Respirovirus pneumoniae

Parvoviridae
- Bocaparvovirus primate1
- Bocaparvovirus primate2
- Erythroparvovirus primate1

Picornaviridae
- Enterovirus betacoxsackie
- Enterovirus coxsackiepol
- Enterovirus deconjuncti
- Kobuvirus aichi
- Parechovirus ahumpari

Pneumoviridae
- Orthopneumovirus hominis

Poxviridae
- Molluscipoxvirus molluscum
- Orthopoxvirus monkeypox
- Orthopoxvirus variola
- Parapoxvirus orf

Retroviridae
- Deltaretrovirus bovleu
- Deltaretrovirus priTlym1

Sedoreoviridae
- Rotavirus alphagastroenteritidis
- Rotavirus betagastroenteritidis

Togaviridae
- Alphavirus madariaga
- Alphavirus rossriver
- Alphavirus semliki
- Alphavirus venezuelan

UKRI Medical Research Council
University of Glasgow
CVR Centre for Virus Research

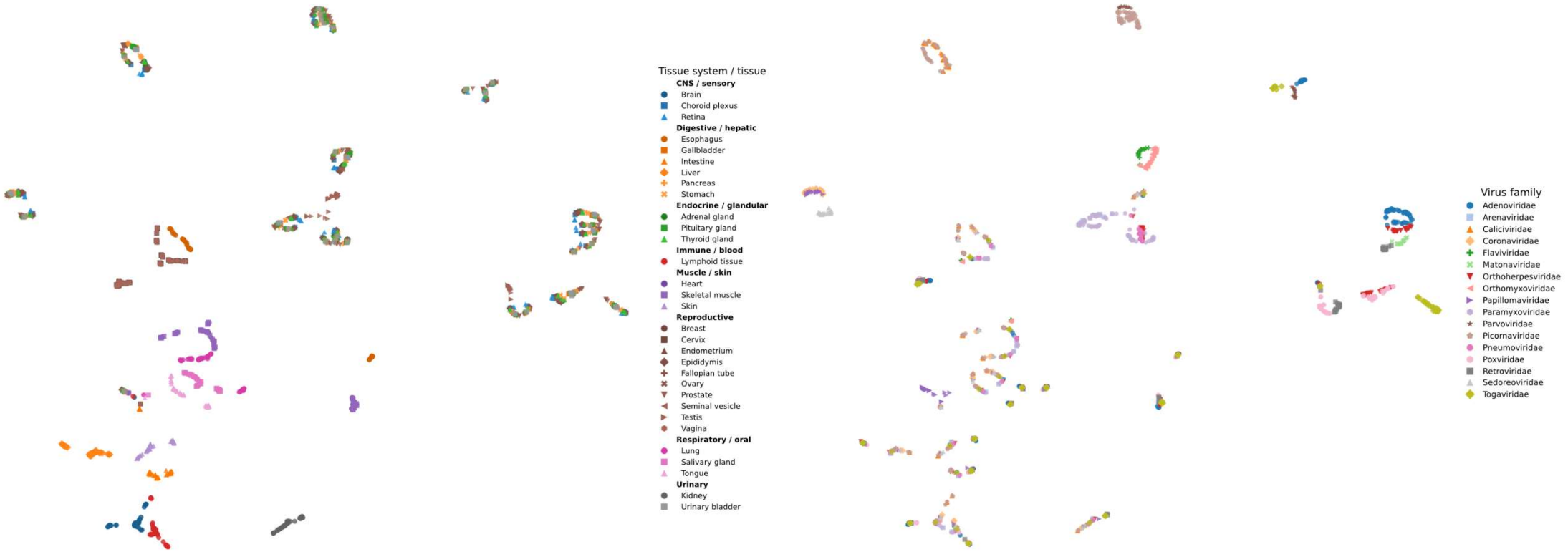
vibelab.co.uk

Training-positive Tissue Density

UKRI Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Embedding Space



UKRI Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Future Work & Challenges

- Expand to broader host ranges to evaluate host-specific tissue tropism and see the relationships across hosts
- Build an updated curated virus–tissue dataset
- Improve calibration of permissivity scores beyond ranking performance
- We would like to validate predictions against experimental and pathology-based evidence

Acknowledgements

Brierley Lab

- **Dr. Liam Brierley**
- **Dr. Jamie Herzig**
- Dr. Jordan Bone
- Bailey Atkinson
- Lukas Visockas

Robertson Lab

- Dr. Kieran Lamb

Funding

This work is supported by the MRC Career Development Fund.

Contact: @stonehc

Medical Research Council | University of Glasgow | CVR Centre for Virus Research

vibelab.co.uk

Learn from observed virus–tissue interactions

$$\mathcal{L} = \alpha\mathcal{L}_{dist} + (1-\alpha)\mathcal{L}_{rank} + \lambda\,\mathcal{L}_{contrast}$$

- Learn tissue permissivity from sparse virus–tissue observations
- Optimisation focuses on relative compatibility, not binary classification
- Loss computed only on observed labels (masked supervision)
- Shared model generalises across viruses and tissues

Tissue-Level Ablation

- Heart ablation produced the strongest overall performance drop
- Urinary bladder ablation caused the largest recall decrease
- Tissue effects were not uniform, suggesting that different tissue embeddings contribute differently to ranking, discrimination, and threshold classification

Top 5 Tissues	Recall % decrease	AUC % decrease	F1 % decrease
Heart	9.02	10.31	12.55
Skeletal muscle	7.04	4.1	3.53
Retina	6.43	2.53	1.7
Endometrium	0.88	1.8	4.11
Urinary bladder	11.68	1.78	5.21