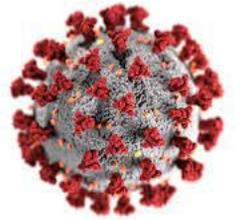


# Developing and exploring a novel software application for understanding HLA-type-associated disease risk for COVID-19

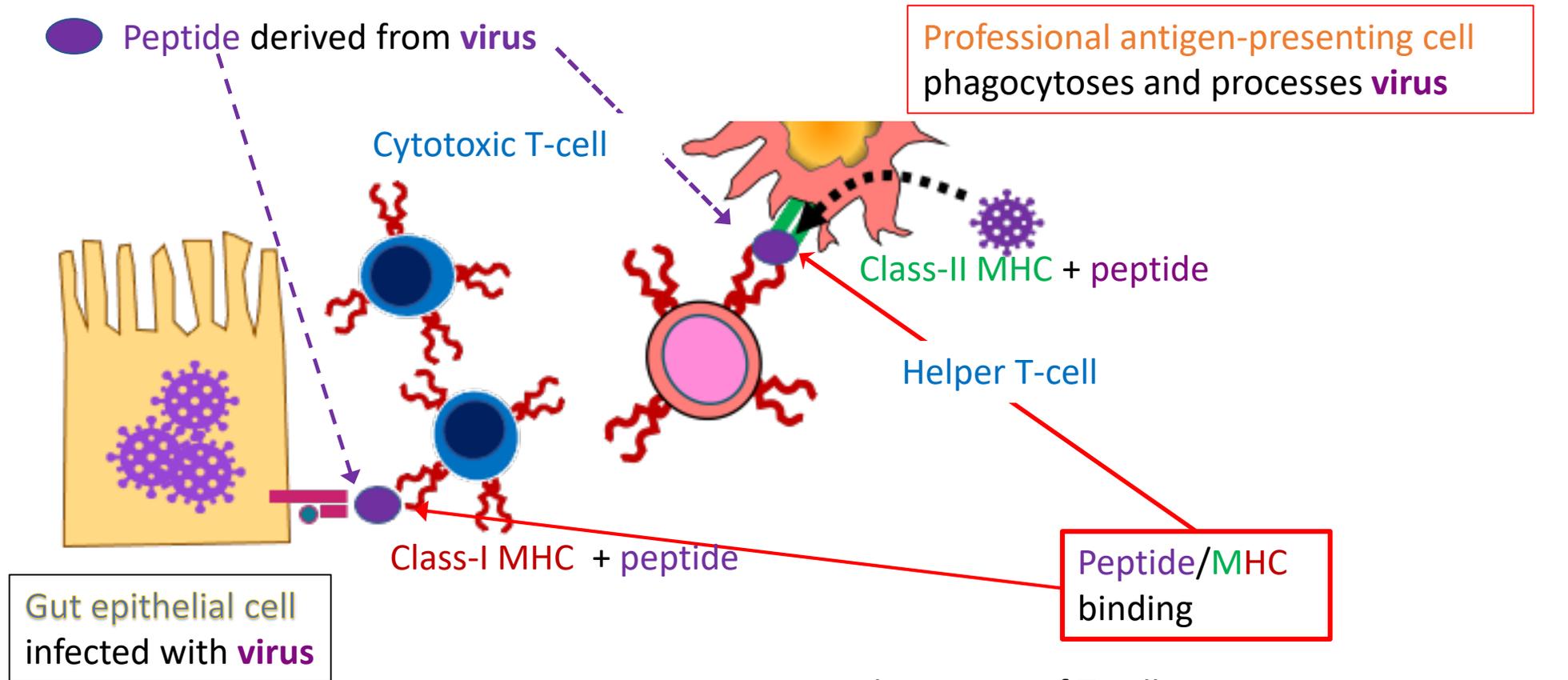
Prof Florian Kern (CEM/BSMS)  
Dr Bernhard Reus (FoSS/ENGINEF)

**Sussex COVID Research Networking Event,  
November 1<sup>st</sup>, 2021**



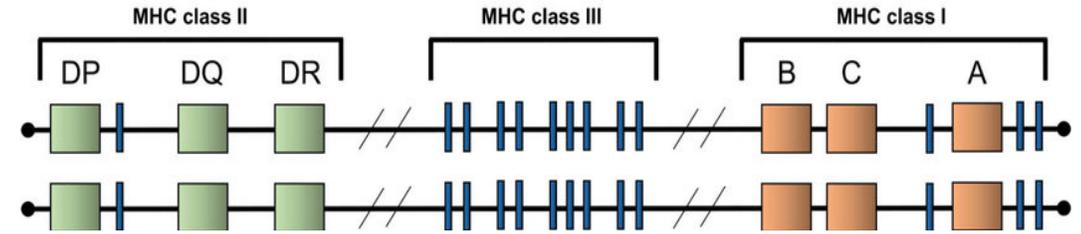
# HLA peptide presentation as a mediator of HLA/disease associations

Note: HLA = human MHC (major histocompatibility complex)



- determines if T-cell response present or not
- MHC locus polymorphism: presence of many alternative forms of the gene ( alleles ) in the population

# HLA (human MHC) polymorphism



- T-cells only recognise peptides that are bound to MHC-molecules in a **'binding groove'** whose amino acid sequence is determined by genes on chromosome 6.
- The binding of certain peptides to certain alleles will produce **'target structures'** for T-cells and may induce/trigger immunity. This may be desirable (infection) or undesirable (autoimmunity).
- MHC molecules are inherited by 'haplotypes' (one set of genes from each parent).
- Known **HLA/disease associations** include DRB1\*03:01 for diabetes, for example
- The loci (*fixed position on a chromosome where a particular gene is located*) encoding class I and class II MHC molecules are the **most polymorphic known in higher vertebrates**:  $1.2 \times 10^7$  different Class I combinations and  $1.8 \times 10^{11}$  different Class II combinations

**Goal:** Find associations of HLA-type and COVID-19 severity

**How?**

- Analyse using Biobank data (~20K patients)
- Our approach: Statistical analysis of associations between HLA binding groove amino acid sequence (rather than just 'allele name') and COVID-19 severity and outcomes
- HLA allele clusters based on biochemical properties (dimension reduction)

**Impact?**

- Such associations may help **explain the role of biogeographical ancestry** in disease severity/outcomes and **inform vaccine development/strategies**

**Goal:** Find associations of HLA-type and COVID-19 severity

## Challenges

- MHC Polymorphism + using AA **sequences** = BIG DATA (combinatorial explosion)
- Using pure sequence data or (predicted) HLA/peptide binding properties?
- Finding the right statistical models

## Deliverables

- Software tool for immunologist that (semi-)automatically analyses data for COVID-19 (and other disease) associations and visualises data