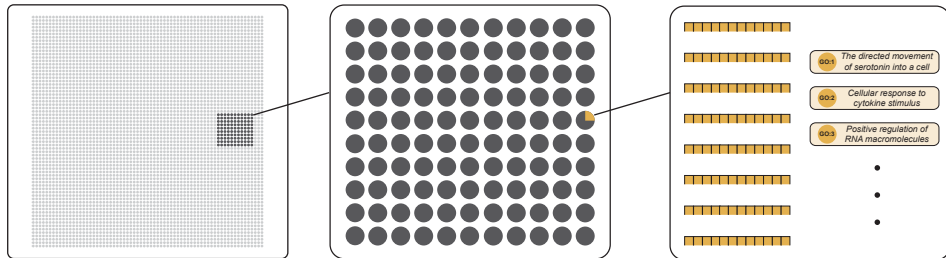


A~1B sequences, a fraction of 20^L possible sequences

~250M UniProt sequences

~600K SwissProt seqs.
with verified functions**B**

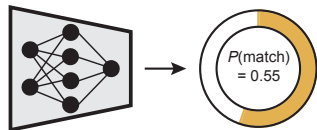
SwissProt sequences

~32K GO annotations / sequence



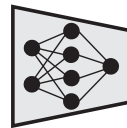
Positives: SwissProt annotations

Negatives: remaining labels

**ProtNote**Protein-function
likelihood**C**

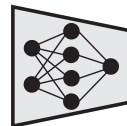
Supervised prediction: seen functions

MAGDLSIAGGL...

The directed movement
of serotonin into a cell**Supervised:**new sequence,
seen function**ProtNote****D**

Zero-shot annotation: unseen, novel functions

MAGDLSIAGGL...

Oxidoreductase, acting
on CH-OH with NAD+**Zero-shot:**new sequence;
unseen, novel function**ProtNote**