

Supplementary Online Materials

Figure S1. Independent components of the amino acid sequence around the isolated glycosylation sites, corresponding to the high existence probability of (a) Val at -3 and (b) Ala at $+2$.

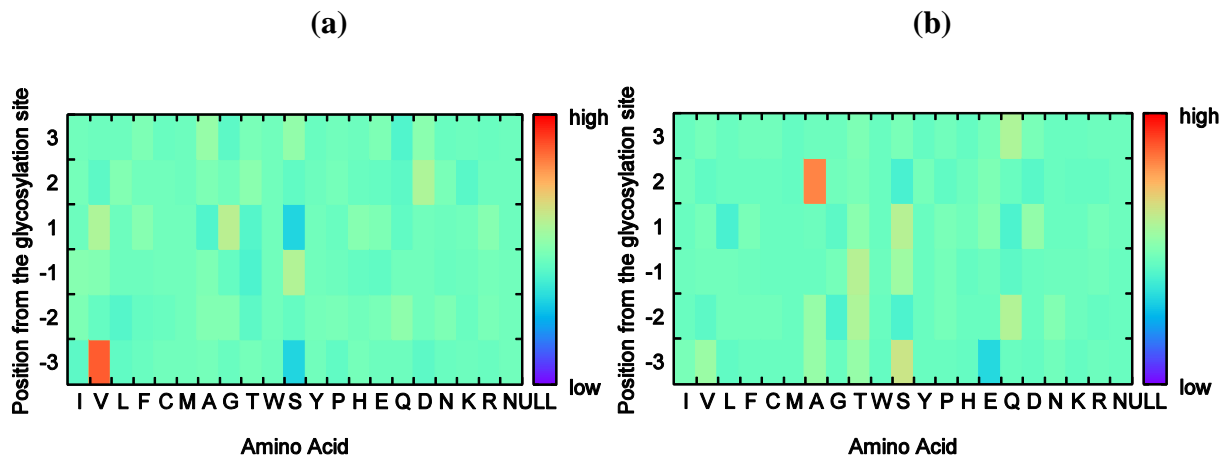


Figure S2. Independent components of the amino acid sequence around the clustered glycosylation sites, corresponding to (a) the high existence probability of Pro at $+3$ and (b) no amino acid with high probability.

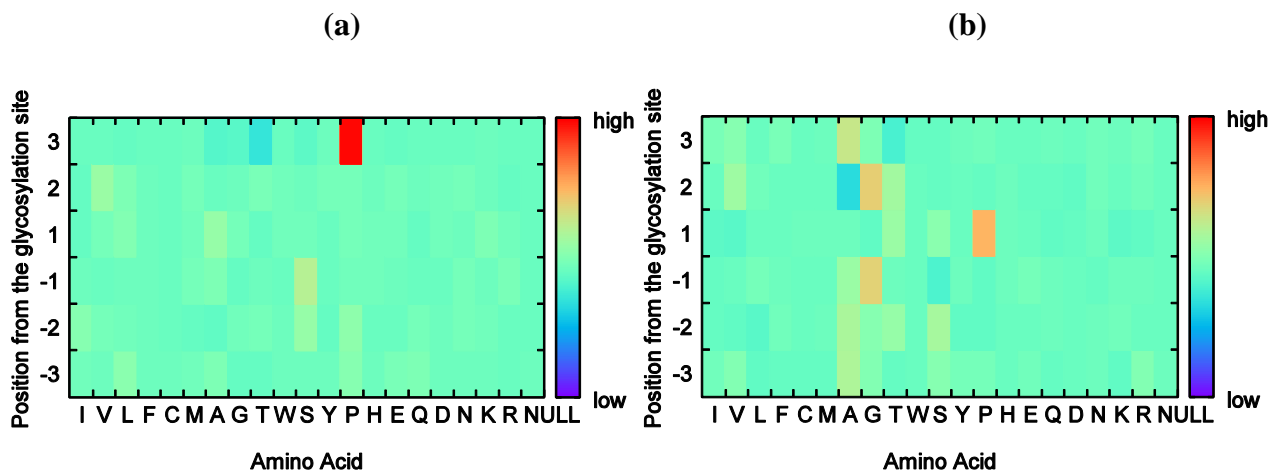


Figure S3. Glycosylation sites presented in a similar way to Figure 5, but for all the 62 human *O*-GalNAc glycoproteins examined. UniProt ID of each protein is given at the left. The light blue and red regions correspond to structural domains and ID regions obtained by DICHOT, respectively. The blue and orange dots indicate mucin-type *O*-linked (GalNAc) and *N*-linked sites, respectively.

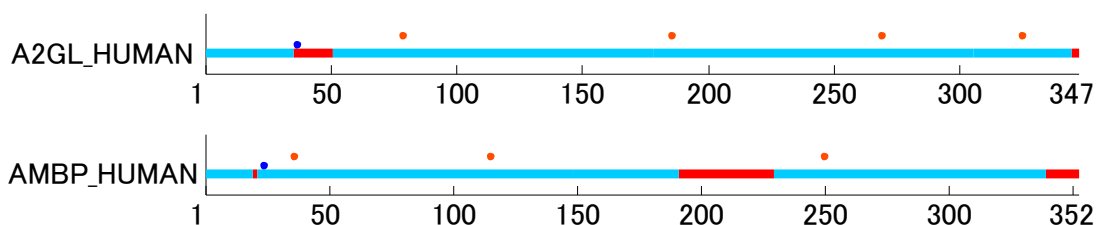


Figure S3. Cont.

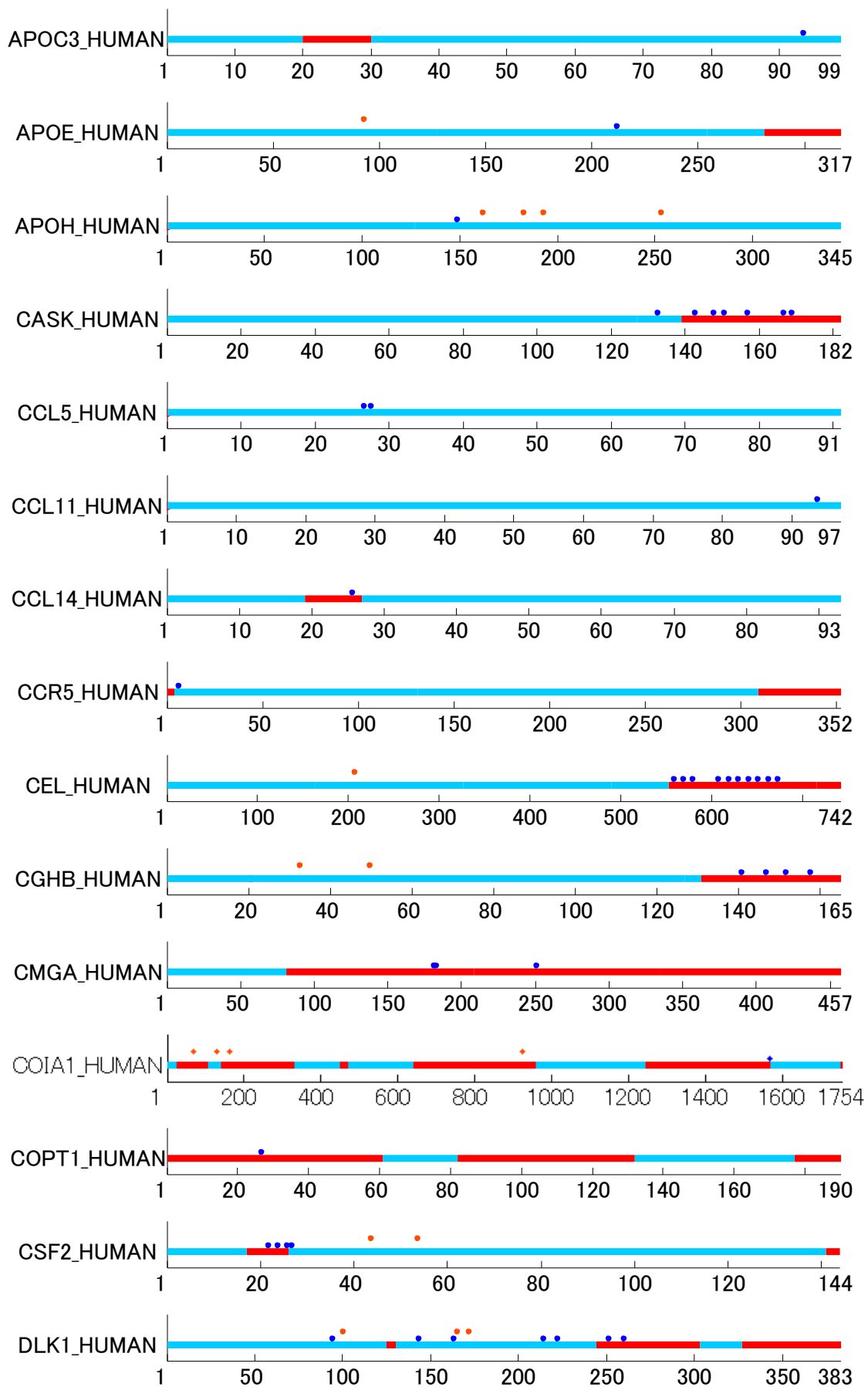


Figure S3. Cont.

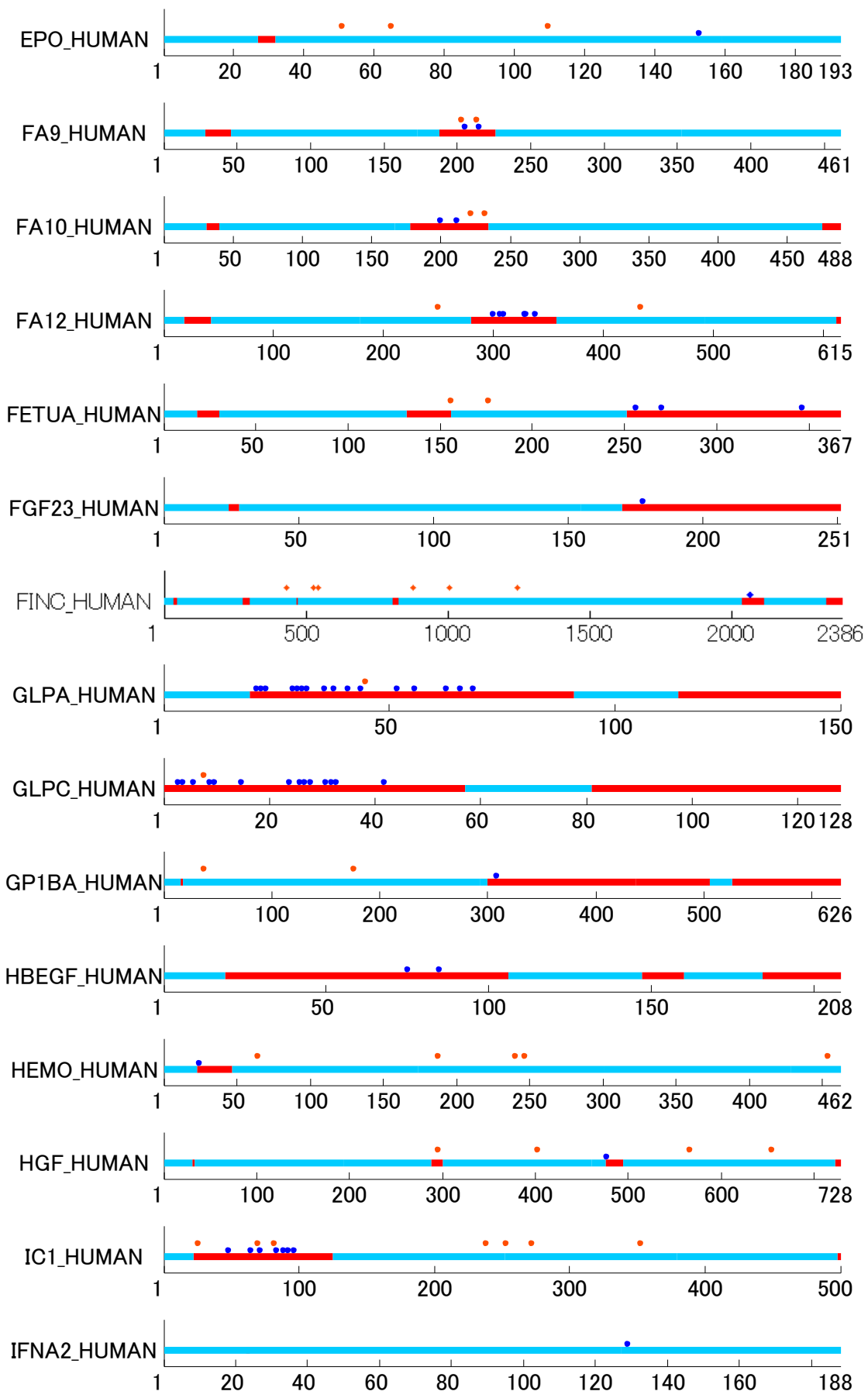


Figure S3. Cont.

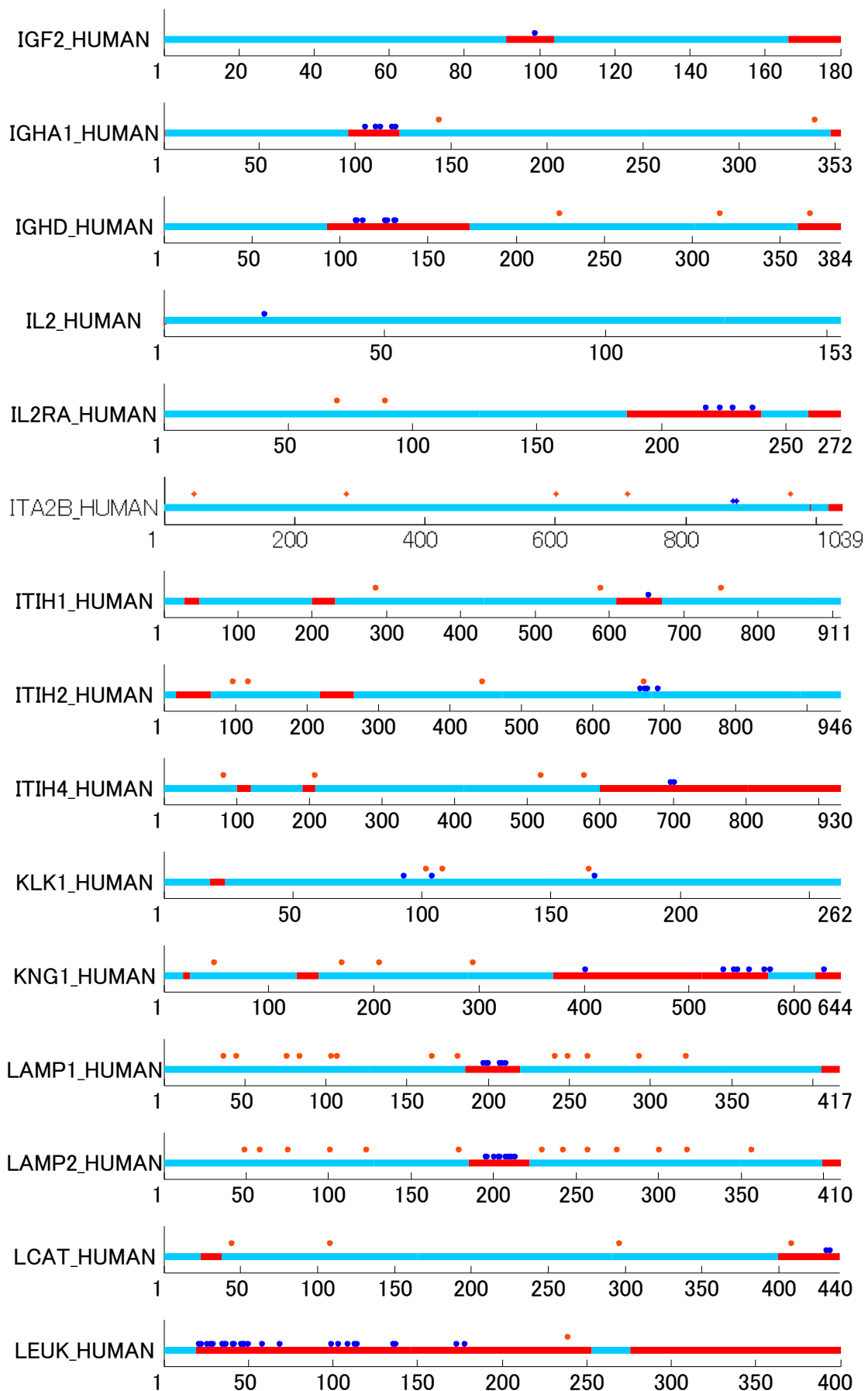


Figure S3. Cont.

