

Table S1. Data used to draw Figs. 1 and 2 and S1 and S2.

organism <sup>1</sup>	abbrev. In figs.	genre	N- glycan Length <sup>2</sup>	QC	%AT	# sequon per protein	sequon density / 500 aa	NxT density secreted / cytosolic <sup>3</sup>	NxT secreted actual / expected <sup>4</sup>	% proteins with $\geq 1$ sequons
<i>Antonospora</i>	a	Fungi	0	N	53.7	2.0	2.9	1.0	0.9	78.5
<i>Encephalitozoon</i>	e	Fungi	0	N	52.5	1.6	2.3	1.0	0.9	73.3
<i>Cryptosporidium</i>	A	Protist	8	N	68.3	6.7	4.7	1.1	1.1*	93.1
<i>Giardia</i>	D	Protist	2	N	51.0	3.2	3.4	1.4**	1.1	71.9
<i>Plasmodium</i>	F	Protist	2	N	75.7	6.9	5.7	0.9	1.0	93.9
<i>Theileria</i>	G	Protist	0	N	64.2	5.6	4.7	1.0	1.1	92.0
<i>Toxoplasma</i>	K	Protist	10	N	41.0	1.6	1.6	1.2**	1.0	68.8
<b>AVERAGE</b>			<b>3.1</b>		<b>58.1</b>	<b>3.9</b>	<b>3.6</b>	<b>1.1</b>	<b>1.0</b>	<b>81.6</b>
<i>Aspergillus</i>	b	Fungi	14	Y	46.6	3.2	3.6	2.0**	1.6**	80.9
<i>Candida</i>	c	Fungi	14	Y	64.9	4.2	4.4	1.2**	1.1**	84.8
<i>Cryptococcus</i>	d	Fungi	11	Y	49.0	3.4	4.0	2.5**	1.8**	82.2
<i>Gibberella</i>	f	Fungi	14	Y	48.3	2.8	3.3	1.8**	1.4**	76.3
<i>Kluyveromyces</i>	g	Fungi	14	Y	59.8	3.7	3.9	1.1	1.2**	87.1
<i>Magnaporthe</i>	h	Fungi	14	Y	42.0	2.5	3.4	2.1**	1.6**	71.8
<i>Neurospora</i>	i	Fungi	14	Y	43.9	2.9	3.5	2.0**	1.5**	81.5
<i>Saccharomyces</i>	j	Fungi	14	Y	60.5	3.7	4.9	1.6**	1.5**	77.7
<i>Schizosaccharomyces</i>	k	Fungi	14	Y	60.3	5.0	5.0	2.0**	1.7**	89.8
<i>Ustilago</i>	m	Fungi	14	Y	43.8	2.9	3.2	1.7**	1.3**	71.9
<i>Yarrowia</i>	n	Fungi	14	Y	46.1	3.5	3.9	1.9**	1.5**	82.5
<i>Anopheles</i>	1	Metazoa	14	Y	43.6	1.9	3.1	1.3**	1.4**	68.4
<i>Caenorhabditis</i>	2	Metazoa	14	Y	56.7	3.0	3.8	1.6**	1.4**	77.1
<i>Canis</i>	3	Metazoa	14	Y	49.3	3.1	3.4	1.8**	1.6**	80.4
<i>Ciona</i>	4	Metazoa	14	Y	57.0	2.9	4.1	1.5**	1.3**	83.5
<i>Danio</i>	6	Metazoa	14	Y	49.7	3.3	3.6	1.7**	1.5**	83.8
<i>Drosophila</i>	5	Metazoa	14	Y	46.2	3.0	3.5	1.5**	1.4**	73.6
<i>Homo</i>	7	Metazoa	14	Y	47.9	3.8	3.6	2.0**	1.7**	81.2
<i>Mus</i>	8	Metazoa	14	Y	47.7	3.7	3.7	2.2**	1.8**	82.3
<i>Tetraodon</i>	9	Metazoa	14	Y	43.1	2.9	3.3	1.9**	1.7**	78.5
<i>Arabidopsis</i>	+	Plant	14	Y	55.8	3.4	4.3	2.0**	1.6**	82.5
<i>Dictyostelium</i>	B	Protist	14	Y	72.6	7.3	6.8	1.9**	1.4**	93.2
<i>Entamoeba</i>	C	Protist	7	Y	72.1	4.3	4.4	1.5**	1.2**	86.1
<i>Leishmania</i>	E	Protist	9	Y	37.4	2.1	2.0	1.2**	1.3**	76.7
<i>Trichomonas</i>	J	Protist	7	Y	64.4	5.6	6.3	1.8**	1.6**	87.5
<i>Trypanosoma</i>	H	Protist	11	Y	47.5	2.7	3.2	1.5**	1.4**	80.9
<b>AVERAGE</b>			<b>13</b>		<b>52.2</b>	<b>3.5</b>	<b>3.9</b>	<b>1.7</b>	<b>1.5</b>	<b>80.9</b>

<sup>1</sup>Organisms lacking N-glycan-dependent QC of protein folding are colored red, while those with QC are colored blue. <sup>2</sup>N-glycan length predicted as in ref. 2. In the case of *Trypanosoma*, both the predicted 11 sugar precursor and a truncated precursor containing 7 sugars are added to N-glycans. <sup>3</sup>Measures extent of positive selection for sequons with Thr in secreted proteins. <sup>4</sup>Measures extent of conditional selection for sequons with Thr. \*\*Statistically greater at  $\alpha = 1\%$ . \*Statistically greater at  $\alpha = 5\%$ .