

Supplemental table 1. The functions and group classification of genes involved in this study.

Gene descriptions are compiled from Saccharomyces Genome Database (SGD,

<http://www.yeastgenome.org>). Group 8 (“Other”) includes genes whose proteins cannot be

confidently located in a specific cellular component.

Gene	Description	Group	Group Description
<i>AGA1</i>	Anchorage subunit of a-agglutinin of a-cells	1	Sexual adhesin
<i>AGA2</i>	Adhesion subunit of a-agglutinin of a-cells	1	Sexual adhesin
<i>BGL2</i>	Endo-beta-1,3-glucanase involved in cell wall maintenance	4	Cell wall catalytic protein
<i>CCW12</i>	Cell wall mannoprotein	3	Cell wall structural protein
<i>CCW14</i>	Covalently linked cell wall glycoprotein	3	Cell wall structural protein
<i>CDA1</i>	Chitin deacetylase	8	Other
<i>CDA2</i>	Chitin deacetylase	8	Other
<i>CHS2</i>	Chitin synthase II	5	Transmembrane protein
<i>CIS3</i>	Mannose-containing glycoprotein constituent of the cell wall	3	Cell wall structural protein
<i>CRR1</i>	Putative glycoside hydrolase	4	Cell wall catalytic protein
<i>CWP1</i>	Cell wall mannoprotein	3	Cell wall structural protein
<i>CWP2</i>	Covalently linked cell wall mannoprotein	3	Cell wall structural protein
<i>DAN1</i>	Cell wall mannoprotein	3	Cell wall structural protein
<i>DAN2</i>	Cell wall mannoprotein	3	Cell wall structural protein
<i>DAN3</i>	Cell wall mannoprotein	3	Cell wall structural protein
<i>DCW1</i>	Putative mannosidase	4	Cell wall catalytic protein
<i>DDR2</i>	Also known as: DDRA2	4	Cell wall catalytic protein
<i>DFG5</i>	Putative mannosidase	4	Cell wall catalytic protein
<i>DSE2</i>	Daughter cell-specific secreted protein with similarity to glucanases	2	Other extracellular protein
<i>DSE4</i>	Daughter cell-specific secreted protein with similarity to glucanases	2	Other extracellular protein
<i>ECM10</i>	Heat shock protein of the Hsp70 family	6	Cytoplasmic protein
<i>EXG1</i>	Major exo-1,3-beta-glucanase of the cell wall	4	Cell wall catalytic protein
<i>FIG1</i>	Integral membrane protein required for efficient mating	5	Transmembrane protein
<i>FIG2</i>	Cell wall adhesin expressed specifically during mating	1	Sexual adhesin
<i>FIT2</i>	Mannoprotein incorporated into the cell wall via a GPI anchor	4	Cell wall catalytic protein
<i>FIT3</i>	Mannoprotein incorporated into the cell wall via a GPI anchor	4	Cell wall catalytic protein
<i>FKS3</i>	Protein involved in spore wall assembly	5	Transmembrane protein
<i>FLO10</i>	Lectin-like protein thought to be involved in flocculation	2	Other extracellular protein
<i>GAS1</i>	Beta-1,3-glucanosyltransferase required for cell wall assembly	4	Cell wall catalytic protein
<i>GAS2</i>	1,3-beta-glucanosyltransferase involved in spore wall assembly	4	Cell wall catalytic protein
<i>GAS5</i>	1,3-beta-glucanosyltransferase localized to the cell wall	4	Cell wall catalytic protein
<i>GFA1</i>	Glutamine-fructose-6-phosphate amidotransferase	6	Cytoplasmic protein
<i>GPI13</i>	ER membrane localized phosphoryltransferase	6	Cytoplasmic protein
<i>GSC2</i>	Catalytic subunit of 1,3-beta-glucan synthase	4	Cell wall catalytic protein
<i>HLR1</i>	Involved in regulation of cell wall composition and osmotic stress response	8	Other
<i>HOR7</i>	Protein of unknown function	5	Transmembrane protein

<i>KNH1</i>	Protein involved in cell wall beta 1,6-glucan synthesis	8	Other
<i>KRE9</i>	Glycoprotein involved in cell wall beta-glucan assembly	2	Other extracellular protein
<i>LRE1</i>	Protein involved in control of cell wall structure and stress response	8	Other
<i>MFA1</i>	Mating pheromone a-factor, made by a-cells	7	Pheromone and receptor
<i>MFA2</i>	Mating pheromone a-factor, made by a-cells	7	Pheromone and receptor
<i>MFAL1</i>	Mating pheromone α -factor, made by α -cells	7	Pheromone and receptor
<i>MFAL2</i>	Mating pheromone α -factor, made by α -cells	7	Pheromone and receptor
<i>MID2</i>	O-glycosylated plasma membrane protein	5	Transmembrane protein
<i>NCA3</i>	Nuclear Control of ATPase	6	Cytoplasmic protein
<i>PAU2</i>	Member of the seripauperin multigene family	3	Cell wall structural protein
<i>PEP4</i>	Vacuolar aspartyl protease (proteinase A)	6	Cytoplasmic protein
<i>PHO3</i>	Constitutively expressed acid phosphatase on the cell surface	4	Cell wall catalytic protein
<i>PHO11</i>	One of three repressible acid phosphatases	8	Other
<i>PIR1</i>	O-glycosylated protein required for cell wall stability	3	Cell wall structural protein
<i>PIR3</i>	O-glycosylated covalently-bound cell wall protein required for cell wall stability	3	Cell wall structural protein
<i>PLB1</i>	Phospholipase B (lysophospholipase) involved in lipid metabolism	8	Other
<i>PLB3</i>	Phospholipase B (lysophospholipase) involved in phospholipid metabolism	8	Other
<i>PRY2</i>	Protein of unknown function	2	Other extracellular protein
<i>ROT1</i>	Molecular chaperone involved in protein folding in the ER	6	Cytoplasmic protein
<i>SAG1</i>	Alpha-agglutinin of alpha-cells	1	Sexual adhesin
<i>SCW11</i>	Cell wall protein with similarity to glucanases	4	Cell wall catalytic protein
<i>SIM1</i>	Protein of the SUN family that may participate in DNA replication	8	Other
<i>SRL1</i>	Mannoprotein that exhibits a tight association with the cell wall	3	Cell wall structural protein
<i>SSA1</i>	ATPase involved in protein folding and nuclear transport	6	Cytoplasmic protein
<i>SSA3</i>	ATPase involved in protein folding and the response to stress	6	Cytoplasmic protein
<i>SSA4</i>	cytoplasmic protein that concentrates in nuclei upon starvation	6	Cytoplasmic protein
<i>SSE1</i>	ATPase that is a component of the heat shock protein Hsp90 chaperone complex	6	Cytoplasmic protein
<i>STE2</i>	Receptor for α -factor pheromone	7	Pheromone and receptor
<i>STE3</i>	Receptor for a-factor pheromone	7	Pheromone and receptor
<i>SUN4</i>	Cell wall protein related to glucanases	4	Cell wall catalytic protein
<i>TIP1</i>	Major cell wall mannoprotein	3	Cell wall structural protein
<i>TIR2</i>	Putative cell wall mannoprotein of the Srp1p/Tip1p family	3	Cell wall structural protein
<i>TIR3</i>	Cell wall mannoprotein of the Srp1p/Tip1p family	3	Cell wall structural protein
<i>TOS1</i>	Covalently-bound cell wall protein of unknown function	3	Cell wall structural protein
<i>UTR2</i>	Chitin transglycosylase	4	Cell wall catalytic protein
<i>YGP1</i>	Cell wall-related secretory glycoprotein	3	Cell wall structural protein
<i>YPS7</i>	Putative GPI-anchored aspartic protease	6	Cytoplasmic protein

Supplemental table 2. Traditional McDonald-Kreitman test for the 4 sexual adhesin genes. The test was carried out using DnaSP v5.10.01 (Librado and Rozas 2009). Columns 2-3 indicate the number of synonymous changes within and between the two yeast species, respectively. Similarly, columns 4-5 represent the number of nonsynonymous changes either within or between the two species.

Gene	S- within	S- between	N- within	N- between	<i>p</i> -value
AGA1	117	99	88	124	0.0092**
AGA2	12	16	17	19	0.8030
FIG2	193	306	257	422	0.8083
SAG1	63	117	78	151	0.9166

Supplemental Table 3. Comparison of the traditional and modified McDonald-Kreitman test results for all other genes. The traditional test was carried out using DnaSP v5.10.01 (Librado and Rozas 2009) while the modified as explained in Materials and Methods. The results from the two tests are shown in the format of “Trad/mod” denoting traditional/modified within each cell. Sp, Sd, Np, and Nd represents the numbers of segregating sites that are synonymous and polymorphic within species, synonymous and fixed differences between species, nonsynonymous and polymorphic within species, and nonsynonymous and fixed differences between species, respectively. The statistical difference is shown by the p-value, which in most cases is calculated using G-test with Williams’ correction, except of 2 cases, as indicated by ^a, in which G-value cannot be calculated and Fisher’s exact test is used. Alpha and DiS in the last column are analogous measures of the difference in standardized nonsynonymous substitution between species and that within species (but see their difference in the main text), and a positive value of either statistic indicates higher rate of nonsynonymous changes between species. And a combination of positive α/DiS and a statistically significant p-value from the McDonald-Kreitman test could suggest positive selection for amino acid changes between species. *, 0.01<p<0.05, significant; **, 0.001<p<0.01, highly significant; ***, p<0.001, very significant.

Gene	Sp	Sd	Np	Nd	p-value	α/DiS
	Trad/mod	Trad/mod	Trad/mod	Trad/mod	Trad/mod	
<i>BGL2</i>	44/18	34/58	3/1	1/1	0.4775/0.9829	-1.318/-0.0357
<i>CCW12</i>	5/3	7/8	2/1	1/2	0.4721/0.6066	-1.800/-0.0500
<i>CCW14</i>	21/6	32/60	10/4	8/10	0.2475/0.1195	-0.905/-0.2571
<i>CDA1</i>	41/17	56/100	11/4	16/17	0.8879/0.8408	0.061/-0.0452
<i>CDA2</i>	28/12	52/84	18/5	14/16	0.0417*/0.3220	-1.388/-0.1341
<i>CHS2</i>	133/52	186/315	25/7	6/9	0.00003***/0.0043**	-4.827/-0.0909
<i>CIS3</i>	24/12	36/70	12/5	16/19	0.8019/0.6806	-0.125/-0.0806
<i>CRR1</i>	38/16	75/118	26/8	21/25	0.0117*/0.1266	-1.444/-0.1585
<i>CWP1</i>	32/16	52/92	21/9	23/28	0.2979/0.2849	-0.484/-0.1267
<i>CWP2</i>	12/4	11/24	3/2	2/2	0.7614/0.3044	-0.375/-0.2564
<i>DAN1</i>	43/18	45/77	18/7	28/31	0.2844/0.8607	0.327/0.0070
<i>DAN2</i>	27/24	17/61	8/6	2/2	0.2615/0.0209	-1.519/-0.1683

<i>DAN3</i>	25/14	6/16	8/2	3/3	0.6037/0.8354	0.360/0.0329
<i>DCW1</i>	60/22	80/135	16/6	4/6	0.0016**/0.0047**	-4.333/-0.1717
<i>DDR2</i>	13/7	9/19	2/1	2/3	0.7516/0.5987	0.308/0.0114
<i>DFG5</i>	57/24	68/110	14/5	17/20	0.9653/0.9731	0.018/-0.0186
<i>DSE2</i>	40/13	60/124	28/8	43/52	0.9412/0.5797	0.023/-0.0855
<i>DSE4</i>	124/39	155/218	96/24	68/84	0.0042**/0.1404	-0.765/-0.1028
<i>ECM10</i>	84/31	128/221	49/10	41/47	0.0182*/0.4024	-0.821/-0.0685
<i>EXG1</i>	60/27	63/113	15/4	11/13	0.4127/0.9263	-0.432/-0.0259
<i>FIG1</i>	63/29	73/159	13/3	11/12	0.483/0.9205	-0.369/-0.0236
<i>FIT2</i>	48/26	14/38	16/6	0/2	0.0348 ^a /0.1423	NA/-0.1375
<i>FIT3</i>	14/6	17/41	4/1	3/3	0.5802/0.9409	-0.619/-0.0747
<i>FLO10</i>	27/6	57/92	14/3	26/31	0.7545/0.8858	-0.137/-0.0813
<i>FKS3</i>	206/83	372/616	67/20	62/71	0.0007**/0.0115*	-0.951/-0.0908
<i>GAS1</i>	68/28	114/197	29/10	15/17	0.0007**/0.0020**	-2.241/-0.1837
<i>GAS2</i>	77/33	61/127	37/12	24/28	0.524/0.2899	-0.221/-0.0860
<i>GAS5</i>	44/18	96/147	12/4	20/21	0.5163/0.6848	-0.309/-0.0568
<i>GFA1</i>	73/29	121/176	9/2	2/2	0.0043**/0.1978	-6.459/-0.0533
<i>GPI13</i>	154/59	189/309	37/8	28/32	0.0762/0.6755	-0.622/-0.0256
<i>GSC2</i>	238/75	336/535	50/11	32/35	0.0009***/0.0429*	-1.206/-0.0665
<i>HLR1</i>	58/20	71/126	54/16	45/53	0.1524/0.1225	-0.469/-0.1484
<i>HOR7</i>	5/3	9/15	3/2	1/1	0.1863/0.2500	-4.400/-0.3375
<i>KNH1</i>	29/10	59/97	17/5	19/24	0.1419/0.3844	-0.820/-0.1350
<i>KRE9</i>	56/27	44/83	44/19	19/28	0.0765/0.0702	-0.820/-0.1608
<i>LRE1</i>	61/26	10/168	45/16	60/69	0.2663/0.3253	-0.328/-0.0898
<i>MFA1</i>	5/4	2/4	0/0	4/4	0.0606 ^a /0.2790	1.000/0.5000
<i>MFA2</i>	5/3	7/39	3/1	1/1	0.2700/0.4231	-3.200/-0.2250
<i>MFAL1</i>	14/6	27/71	3/4	11/11	0.3739/0.0900	0.474/-0.2659
<i>MFAL2</i>	21/10	17/35	14/6	11/13	0.9547/0.6356	-0.030/-0.1042
<i>MID2</i>	40/17	95/188	27/9	36/43	0.0708/0.0953	-0.781/-0.1600
<i>NCA3</i>	52/24	62/111	18/7	12/14	0.1635/0.1713	-0.788/-0.1138
<i>PAU2</i>	31/24	7/19	5/4	2/3	0.5755/0.7302	0.435/-0.0065
<i>PEP4</i>	36/14	76/123	6/2	8/8	0.4401/0.6651	-0.583/-0.0639
<i>PHO3</i>	89/41	69/134	23/7	4/6	0.0029**/0.0360*	-3.458/-0.1030
<i>PHO11</i>	150/47	22/63	43/11	1/9	0.0225*/0.4406	-5.307/-0.0647
<i>PIR1</i>	27/9	19/80	8/2	1/1	0.0716/0.0389	-4.630/-0.1695
<i>PIR3</i>	93/48	44/119	11/3	7/9	0.5762/0.9572	0.257/0.0115
<i>PLB1</i>	80/35	124/206	23/7	16/20	0.0239*/0.2053	-1.228/-0.0782
<i>PLB3</i>	93/34	133/218	22/7	26/32	0.5538/0.6191	-0.210/-0.0427
<i>PRY2</i>	50/16	43/70	34/9	18/24	0.1743/0.4308	-0.624/-0.1047
<i>ROT1</i>	30/16	52/85	10/4	7/8	0.0972/0.2709	-1.476/-0.1140
<i>SCW11</i>	85/29	107/211	38/12	48/57	0.9895/0.3451	0.003/-0.0800
<i>SIM1</i>	48/16	88/152	22/6	28/31	0.2829/0.3696	-0.440/-0.1033
<i>SRL1</i>	17/7	38/62	17/6	19/22	0.1203/0.2504	-1.000/-0.1996
<i>SSA1</i>	87/28	44/75	8/1	5/4	0.7315/0.8708	0.191/0.0162
<i>SSA3</i>	92/34	135/240	14/4	12/13	0.2001/0.3423	-0.712/-0.0539
<i>SSA4</i>	75/28	127/219	21/5	9/11	0.0008***/0.0523	-2.951/-0.1037
<i>SSE1</i>	78/32	104/178	14/3	6/7	0.0216*/0.4211	-2.111/-0.0479
<i>STE2</i>	45/18	94/149	23/7	15/16	0.0019**/0.0223*	-2.203/-0.1830
<i>STE3</i>	73/29	89/142	23/7	35/39	0.4771/0.9532	0.199/0.0210

<i>SUN4</i>	57/28	101/199	37/14	34/39	0.0237*/0.0177*	-0.928/-0.1695
<i>TIP1</i>	36/14	24/49	11/2	4/5	0.3395/0.9244	-0.833/-0.0324
<i>TIR2</i>	43/18	37/158	28/9	8/14	0.0125*/0.0005***	-2.012/-0.2519
<i>TIR3</i>	45/21	31/71	17/7	6/9	0.1987/0.1460	-0.952/-0.1375
<i>TOS1</i>	49/17	82/143	19/5	8/11	0.0018**/0.0475*	-2.974/-0.1558
<i>UTR2</i>	65/25	90/161	47/14	35/45	0.0245*/0.0934	-0.859/-0.1405
<i>YGP1</i>	45/18	61/103	23/6	18/21	0.1396/0.5172	-0.732/-0.0806
<i>YPS7</i>	80/32	115/186	51/18	52/61	0.1621/0.1404	-0.410/-0.1130

Supplemental table 4. PAML analysis of sexual adhesin genes using branch models. Shown in the second column is the log(likelihood) of branch model assuming all the branches, including the branch between the two species, have the same dN/dS ratio. The third column contains the log(likelihood) of branch model assuming the branch between the two species has a different dN/dS ratio from all the branches within the species, which are assumed to be the same.

Gene	lnL (single dN/dS ratio)	lnL (two dN/dS ratios)	$2\Delta\ln L$	p -value
<i>AGA1</i>	-4632.0013	-4630.1250	3.7526	$p > 0.05$
<i>AGA2</i>	-664.5486	-664.4730	0.1512	$p > 0.50$
<i>FIG2</i>	-11470.6848	-11469.8671	1.6354	$p > 0.05$
<i>SAG1</i>	-4599.9910	-4604.6984	-9.4148	$P < 0.01^a$

^a The result suggests the model assuming the branch between the two species has a dN/dS ratio different from all the branches within the species is significantly worse than the model assuming a single dN/dS ratio across all branches on the tree.

Supplemental table 5. Tests of different evolutionary rates in sexual adhesin genes using PAML site models. Shown in the second column is the log(likelihood) of site model M1a, assuming two types of sites with $\omega < 1$ and $\omega = 1$ respectively, and the third M2a, allowing one extra group of sites with $\omega > 1$.

Gene	lnL(M1a)	lnL(M2a)	$2\Delta\ln L$	p -value	Sites with $\text{Pr}(\omega > 1) > 0.99$
<i>AGA1</i>	-4584.92967	-4556.031893	57.795554	~0.00	S-190, L-192, S-224, S-226, L-381
<i>AGA2</i>	-664.305755	-664.305755	0	1	NA
<i>FIG2</i>	-11427.61331	-11422.53197	10.162694	<0.01	NA
<i>SAG1</i>	-4595.390551	-4594.597841	1.58542	>0.05	NA

Supplemental table 6. Patterns of amino acid substitutions at *AGAI* codon sites with $\text{Pr}(\omega > 1) > 99\%$. The amino acids are represented by single letters in Columns 3 and 4 and the number in parenthesis following each amino acid allele is its relative frequency.

Codon position	$\text{Pr}(\omega > 1)$	Amino acids in <i>S. cerevisiae</i>	Amino acids in <i>S. paradoxus</i>
190	1.000	S (75%), P(25%)	S (80%), P (15%), V (5%)
192	0.998	S (60%), L (40%)	S (95%), F (5%)
224	0.994	S (100%)	S (60%), P (25%), L (10%), I (5%)
226	0.998	S (100%)	S (75%), A (25%)
381	0.997	L (75%), S (25%)	L (80%), P (20%)

Supplemental table 7. Characterization of evolutionary variability and biochemical property of amino acid changes in eight functional groups. Only genes that each contains at least 5 unique sequences in both yeast species are used. All the segregating sites are classified into four categories, as shown in Columns 4-7, in which each cell contains the number and the fraction (in parenthesis) of sites in that category. The distribution of the four categories of sites is compared between sexual adhesins and each of the other functional groups and their difference statistically tested for significance using Chi-square test, the results of which are shown in the last column.

Functional Group	Number of Genes	Total Number of Segregating Sites	Sites with $U \geq 0.8$ and BLOSUM62 Score ≥ 0	Sites with $U < 0.8$ and BLOSUM62 Score ≥ 0	Sites with $U \geq 0.8$ and BLOSUM62 Score < 0	Sites with $U < 0.8$ and BLOSUM62 Score < 0	Statistical Significance vs. Sexual Adhesins
Sexual adhesins	4	884	378 (0.428)	102 (0.115)	317 (0.359)	87 (0.098)	
Other extracellular proteins	4	280	81 (0.289)	99 (0.354)	43 (0.154)	57 (0.204)	$p \approx 6.9 \times 10^{-27}$ ***
Cell wall structural proteins	13	282	103 (0.365)	127 (0.450)	18 (0.064)	34 (0.121)	$p \approx 3.1 \times 10^{-40}$ ***
Cell wall catalytic proteins	15	525	176 (0.335)	212 (0.404)	54 (0.103)	83 (0.158)	$p \approx 3.9 \times 10^{-47}$ ***
Transmembrane proteins	4	211	73 (0.346)	75 (0.355)	27 (0.128)	36 (0.171)	$p \approx 6.2 \times 10^{-21}$ ***
Cytoplasmic proteins	11	337	120 (0.356)	131 (0.389)	33 (0.098)	53 (0.157)	$p \approx 1.2 \times 10^{-34}$ ***
Pheromones and receptors	4	123	45 (0.366)	43 (0.350)	15 (0.122)	20 (0.163)	$p \approx 2.7 \times 10^{-13}$ ***
Others	9	381	134 (0.352)	147 (0.386)	49 (0.129)	51 (0.134)	$p \approx 1.1 \times 10^{-32}$ ***

Supplemental figure legends

Supplemental figure 1. Within-species nonsynonymous (dN) and synonymous (dS) substitution rates for genes in different groups. (A) *S. cerevisiae*; (B) *S. paradoxus*. All the 73 genes under study are grouped and colored in the same way as in Figure 1. The line represents $dN/dS = 1$.

Supplemental figure 2. The between-species f_N for individual genes. The variation in f_N for each gene is shown in box plot, with each circle representing an outlier pair of sequences from the two different species. The genes are shown in different colors based on their groups, with both grouping and coloring scheme being the same as those in Figure 1.

Supplemental figure 3. Evolutionary variability and functional characterization of amino acid residues in individual protein groups. Same as in Figure 5, to improve statistical reliability, the analysis included only genes that have at least five unique sequences in either species and BLOSUM62 scores were obtained for all two-state variable sites.

Supplemental Reference

Librado, P and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

FIG. S1 A

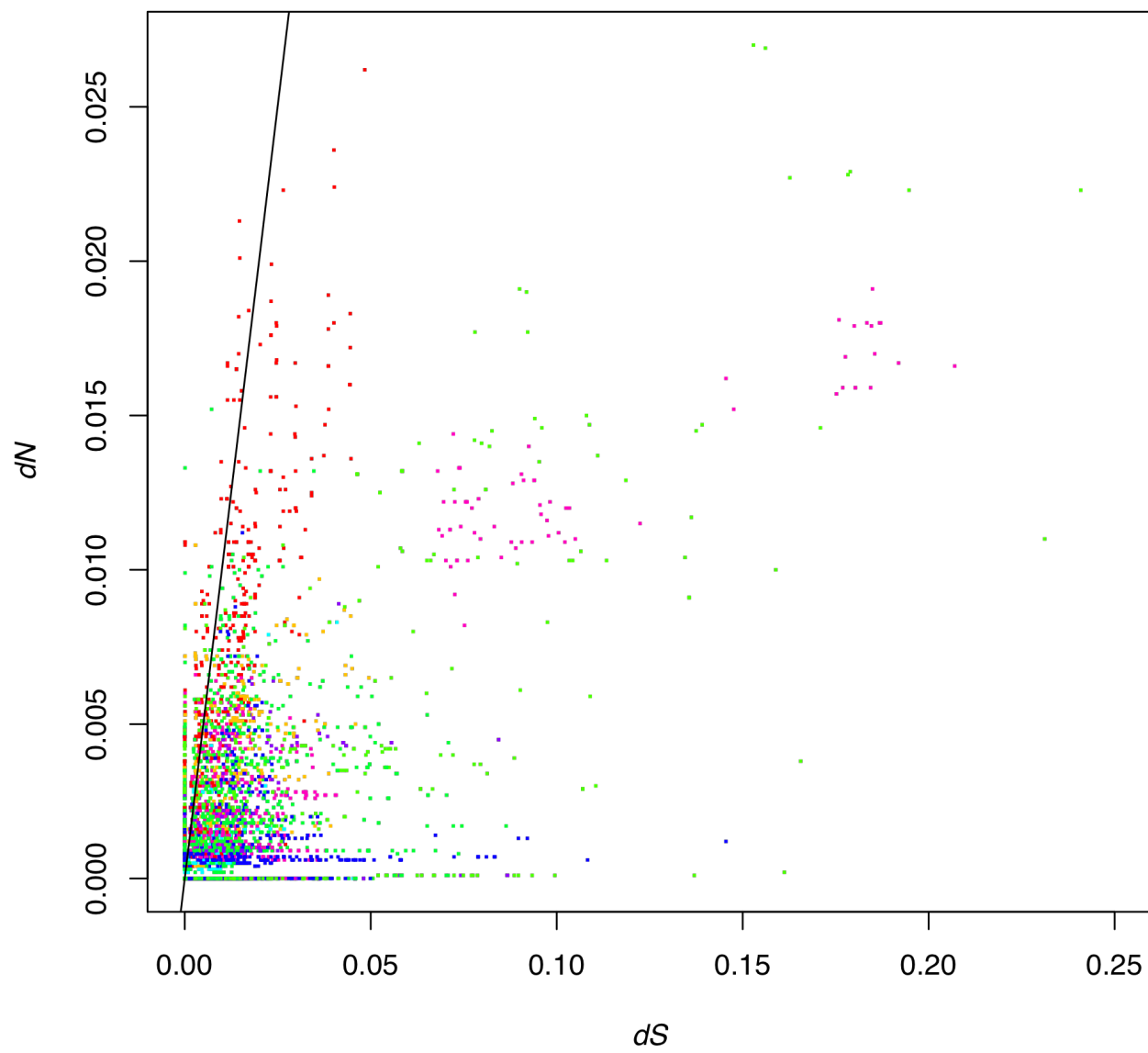


FIG. S1 B

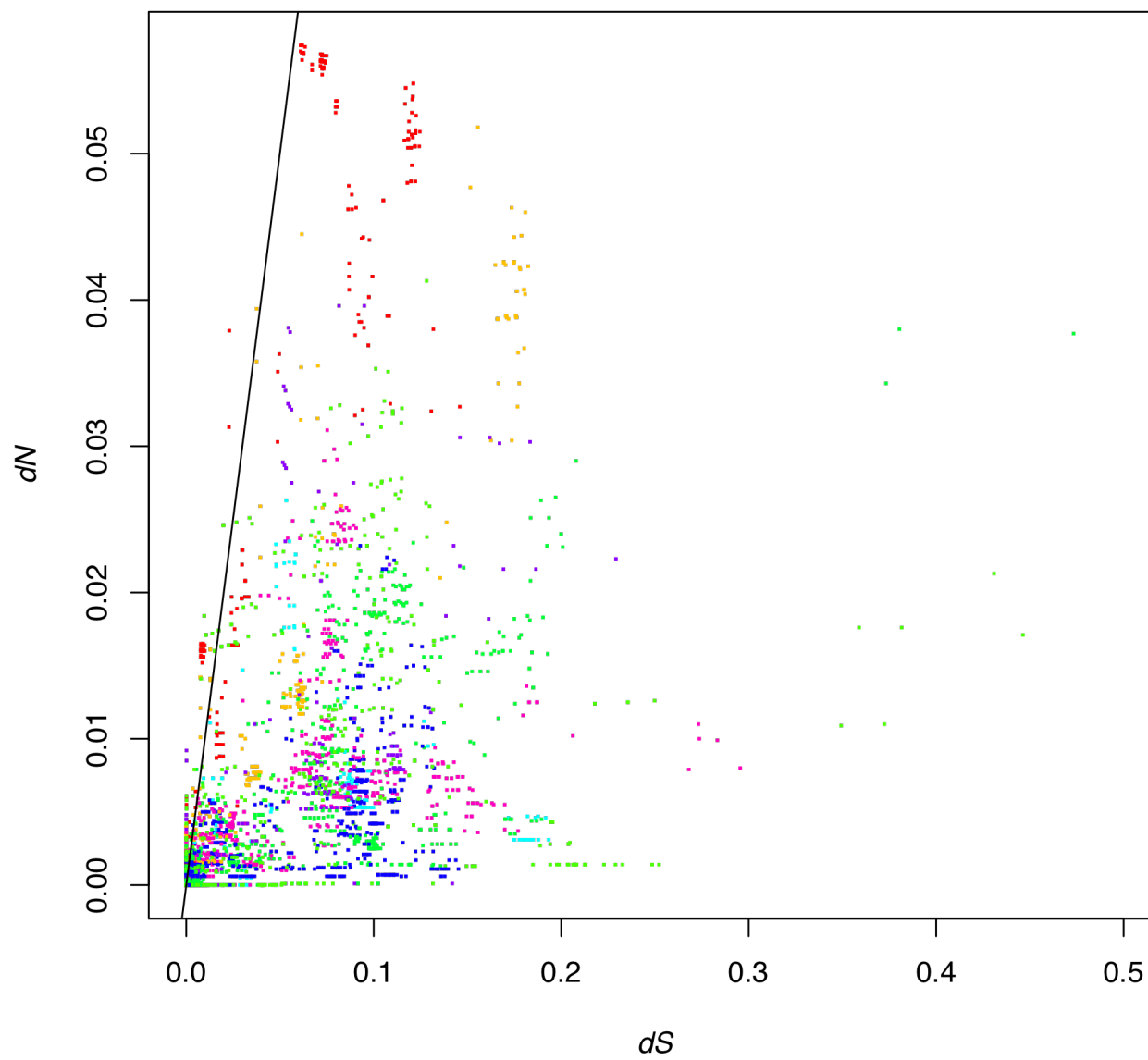


FIG. S2

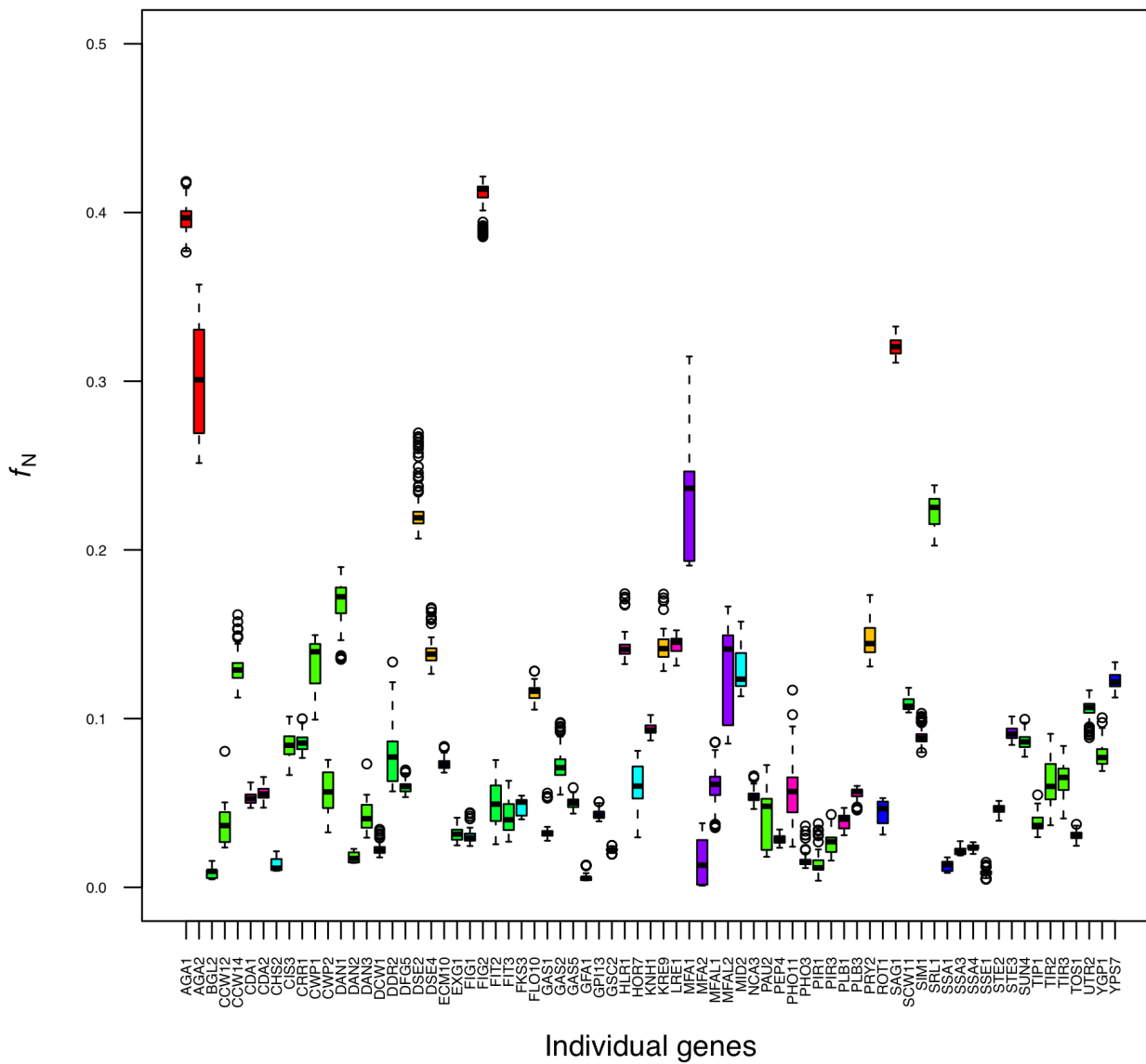
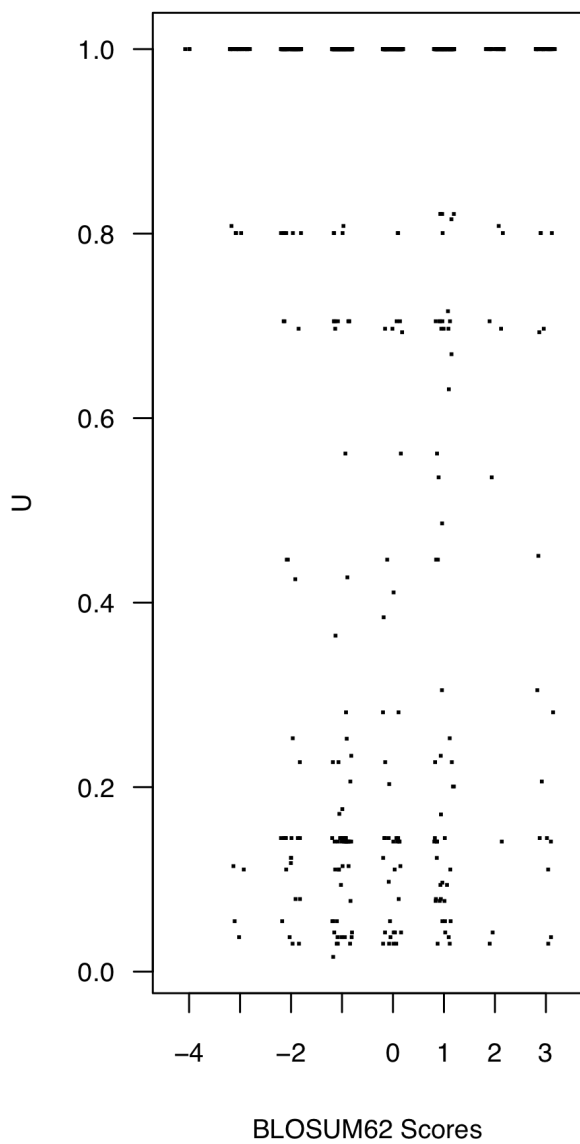


FIG. S3 A&B

A. Sexual adhesins (N=884)



B. Other extracellular proteins (N=280)

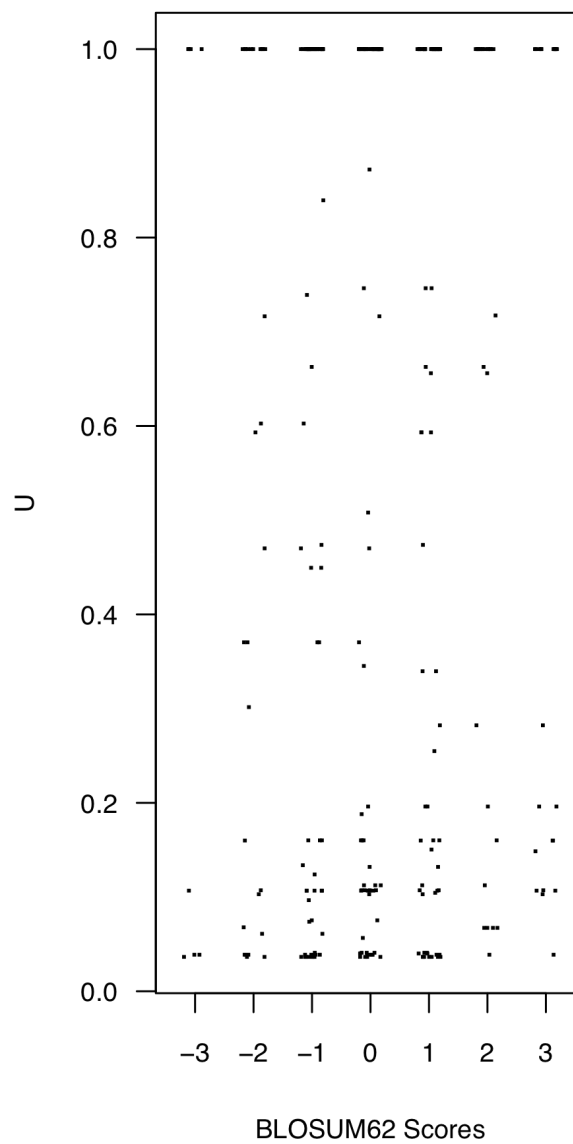
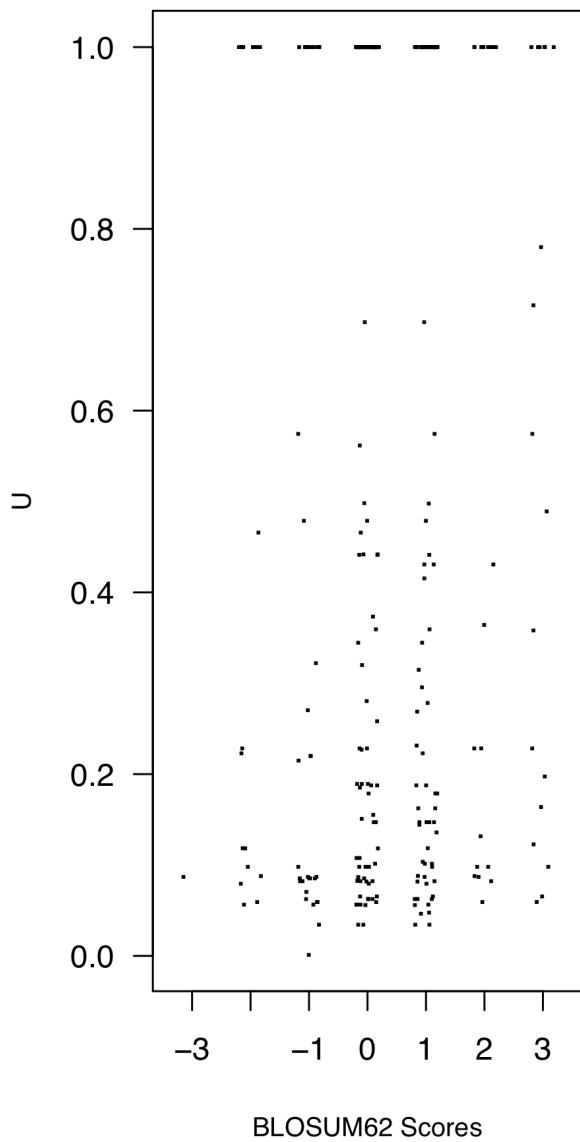


FIG. S3 C&D

C. Cell wall structural proteins (N=282)



D. Cell wall catalytic proteins (N=525)

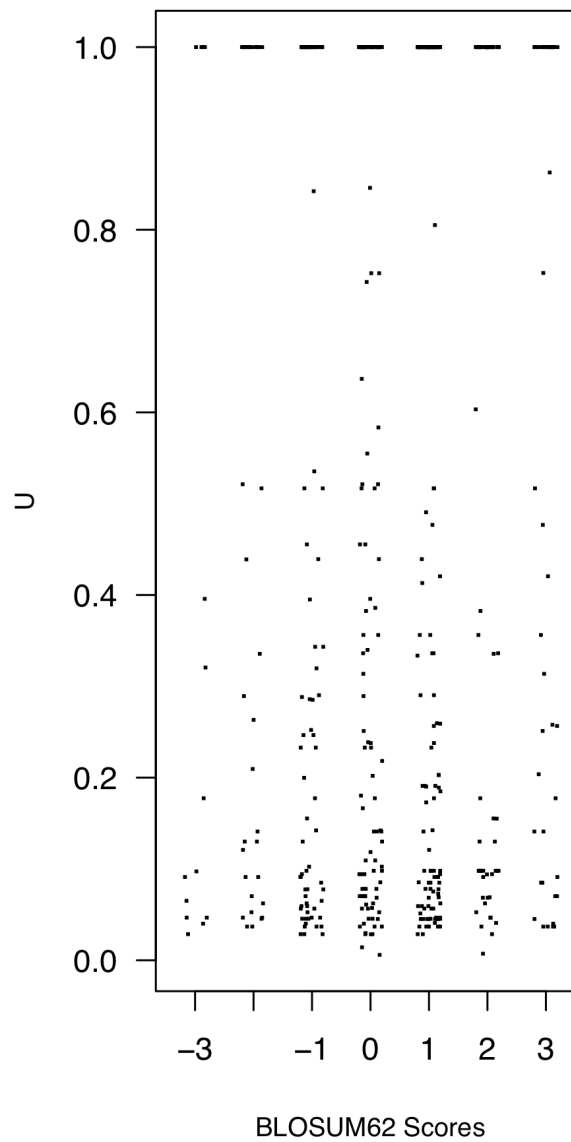
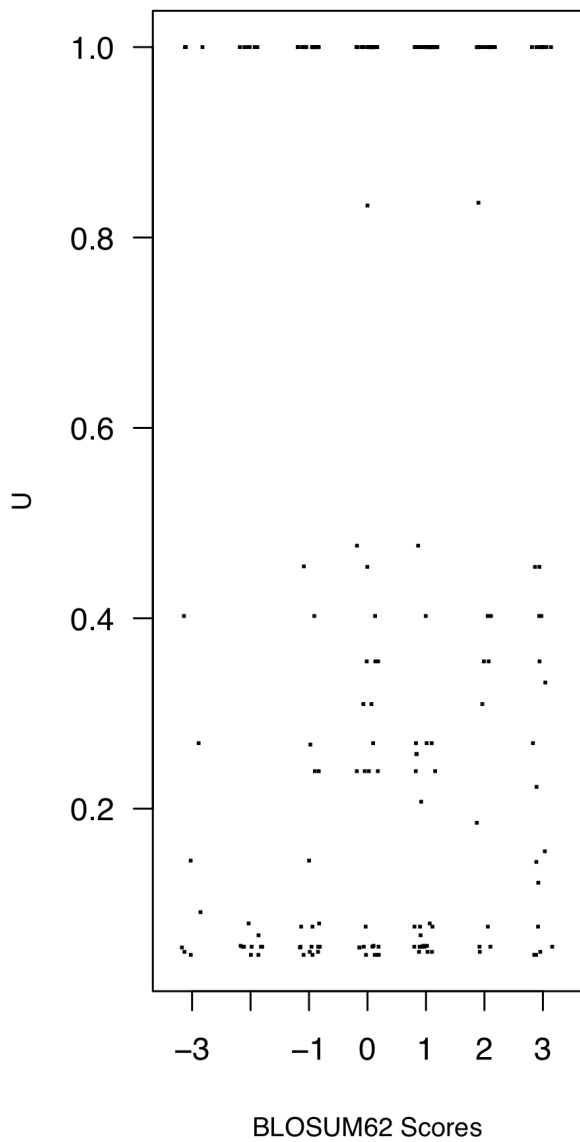


FIG. S3 E&F

E. Transmembrane proteins (N=211)



F. Cytoplasmic proteins (N=337)

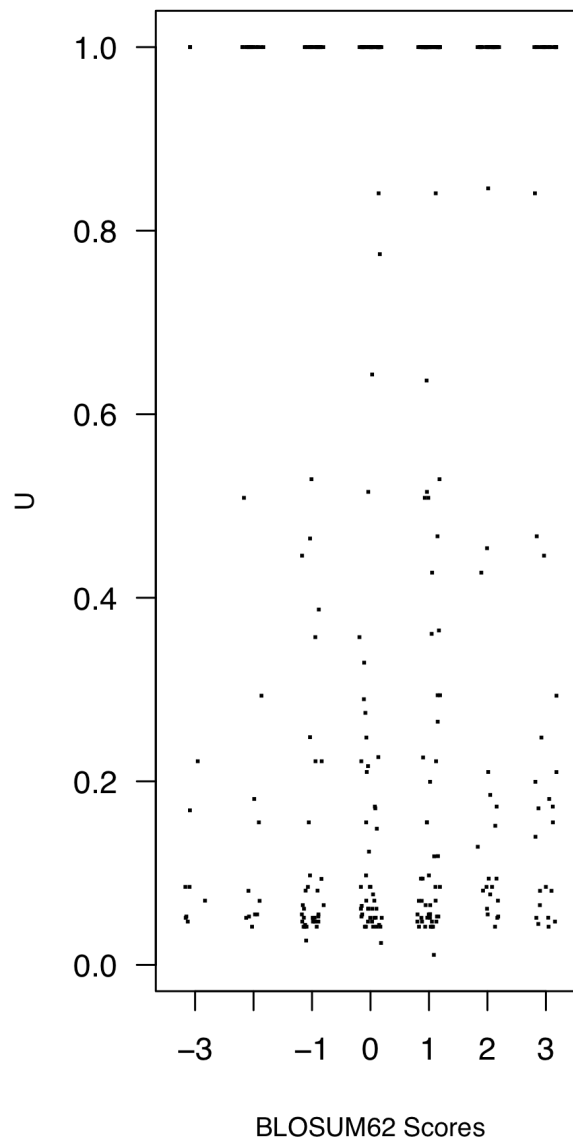
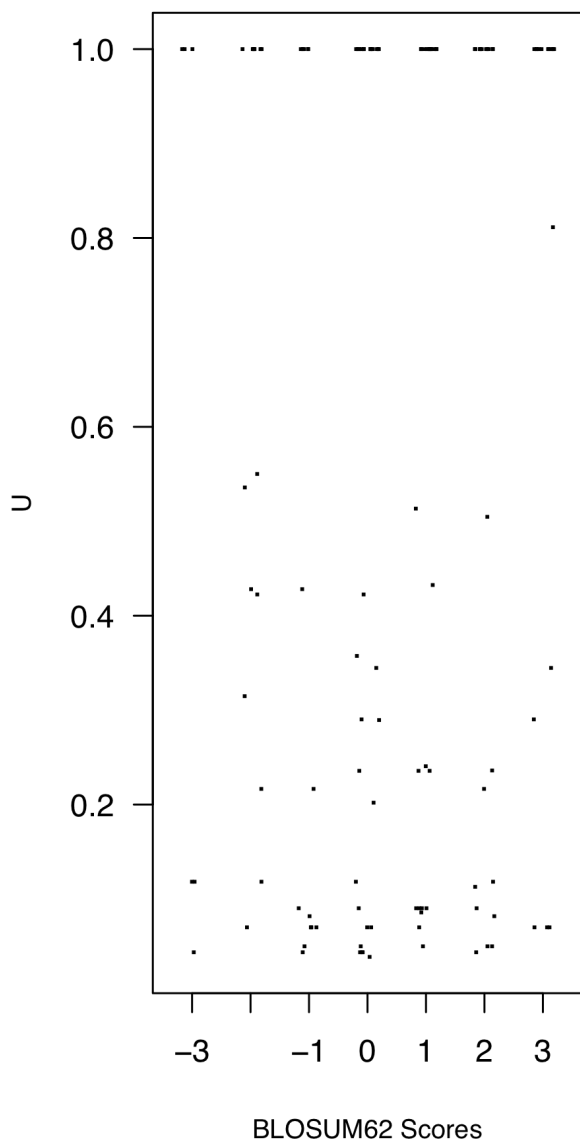


FIG. S3 G&H

G. Sex pheromones and receptors (N=123)



H. Others (N=381)

