Supplement to: Protein markers for *Candida albicans* EVs include claudin-like Sur7 family proteins and GTPases

Charlotte Dawson[a], Donovan Garcia Ceron[a], Harinda Rajapaksha[b], Pierre Faou[b], Marilyn Anderson[a], Mark Bleackley[a]

[a]*Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, 3086, Victoria, Australia*

[b]*La Trobe Comprehensive Proteomics Platform, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, 3086, Victoria, Australia*

Corresponding author: Mark Bleackley, m.bleackley@latrobe.edu.au

1 Supplementary data

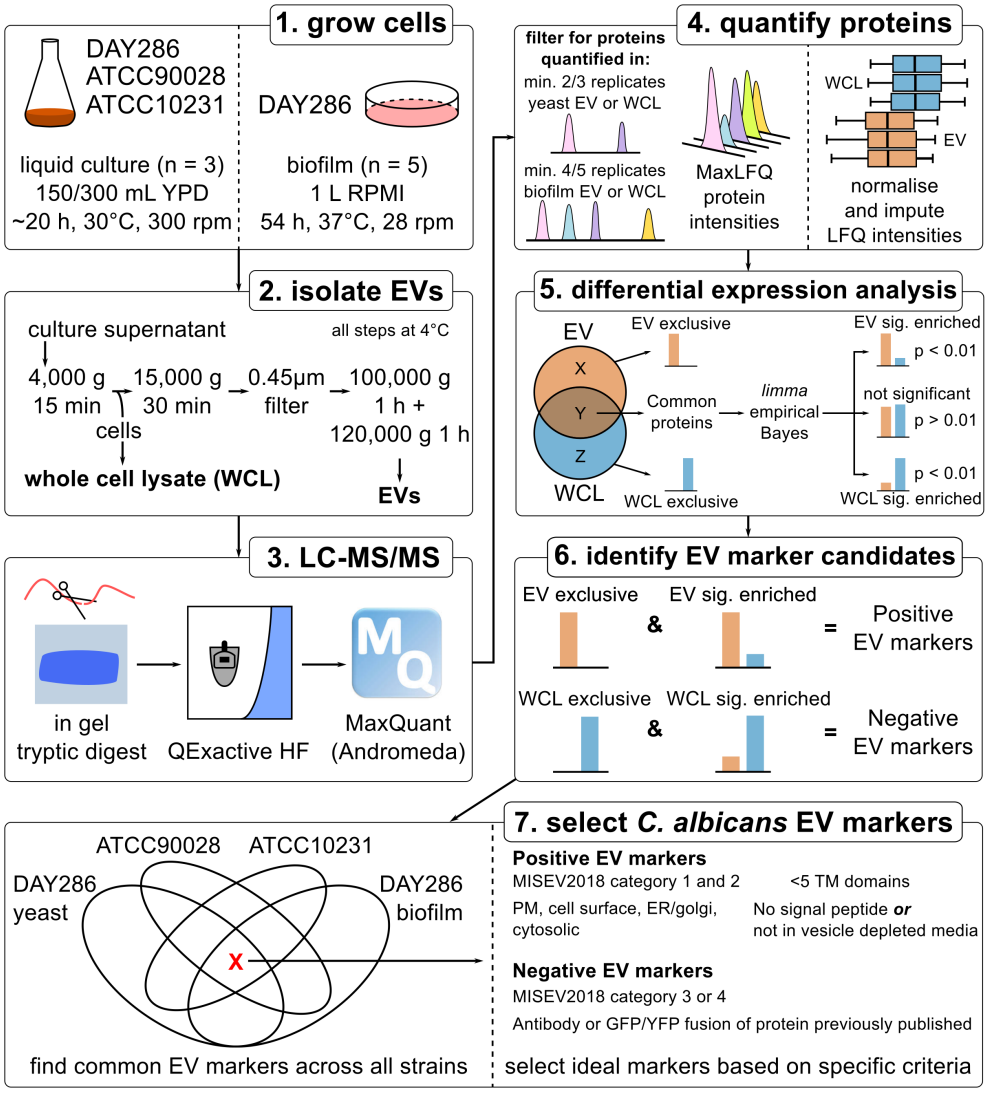

Figure S1: Schematic of the workflow used for the isolation and proteomic analysis of *C. albicans* EVs and the selection of EV protein markers. Steps 1 to 6 were performed for each of the four strains examined in this study: DAY286 yeast, ATCC90028 yeast, ATCC10231 yeast, and DAY286 biofilm. The results from the four seperate analyses were combined (step 7) to identify proteins which commonly appeared as EV marker candidates. The MISEV2018 criteria for protein content-based EV characterisation and other considerations were used to select the best *C. albicans* EV marker proteins.

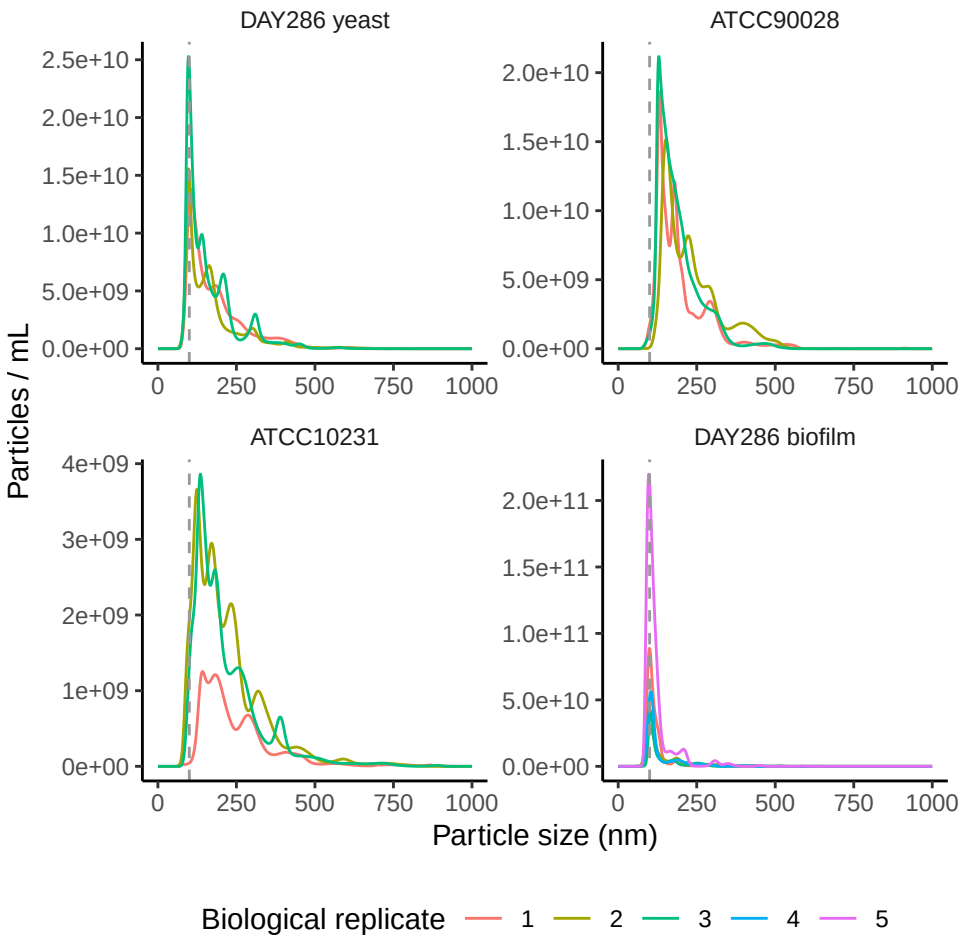

Figure S2: Size distribution of *C. albicans* EVs as determined by nanoparticle tracking analysis (NTA). A line plot for each EV biological replicate from each strain is shown. Each line plots represents the average of three technical replicates and the dashed grey lines indicate 100 nm.

Table S1: Metadata for each EV isolation analysed in this study. The optical density (at 600 nm) and volume of each culture from which the EVs were isolated is given. The protein and particle concentration of the EVs along with the NTA size distribution results for each biological replicate are also shown.

						Particle diameter (nm)		
Biological replicate	OD600	Culture vol. (mL)	Protein yield (μg)	Protein conc. (μg/mL)	Particle conc. (particles/mL)	Mean	Mode	SD
DAY286 yeast								
1	36.2	300	21.85	0.437	1.30e+12	182.8	106.4	93.4
2	39.7	300	18.00	0.360	1.10e+12	174.5	99.2	93.9
3	34	300	22.35	0.447	1.57e+12	164.5	100.2	83.1
ATCC90028								
1	15.5	150	42.40	0.424	1.42e+12	197.1	135.4	87.1
2	16.85	150	71.60	0.716	1.68e+12	233.1	156.7	90.7
3	14.85	150	58.40	0.584	1.85e+12	191.0	138.0	68.4
ATCC10231								
1	9.1	150	19.20	0.192	2.04e+11	248.5	171.9	120.1
2	11.14	150	28.20	0.282	5.52e+11	218.7	155.5	112.1
3	12.3	150	28.70	0.287	4.76e+11	217.9	140.2	110.0
DAY286 biofilm								
1	NA	1000	20.40	0.204	2.97e+12	119.4	100.5	47.4
2	NA	1000	19.80	0.198	1.65e+12	142.1	94.7	64.3
3	NA	1000	29.70	0.297	1.30e+12	143.3	103.1	70.2
4	NA	1000	19.30	0.193	1.94e+12	133.3	103.1	60.7
5	NA	1000	41.60	0.416	7.94e+12	120.9	97.5	47.4

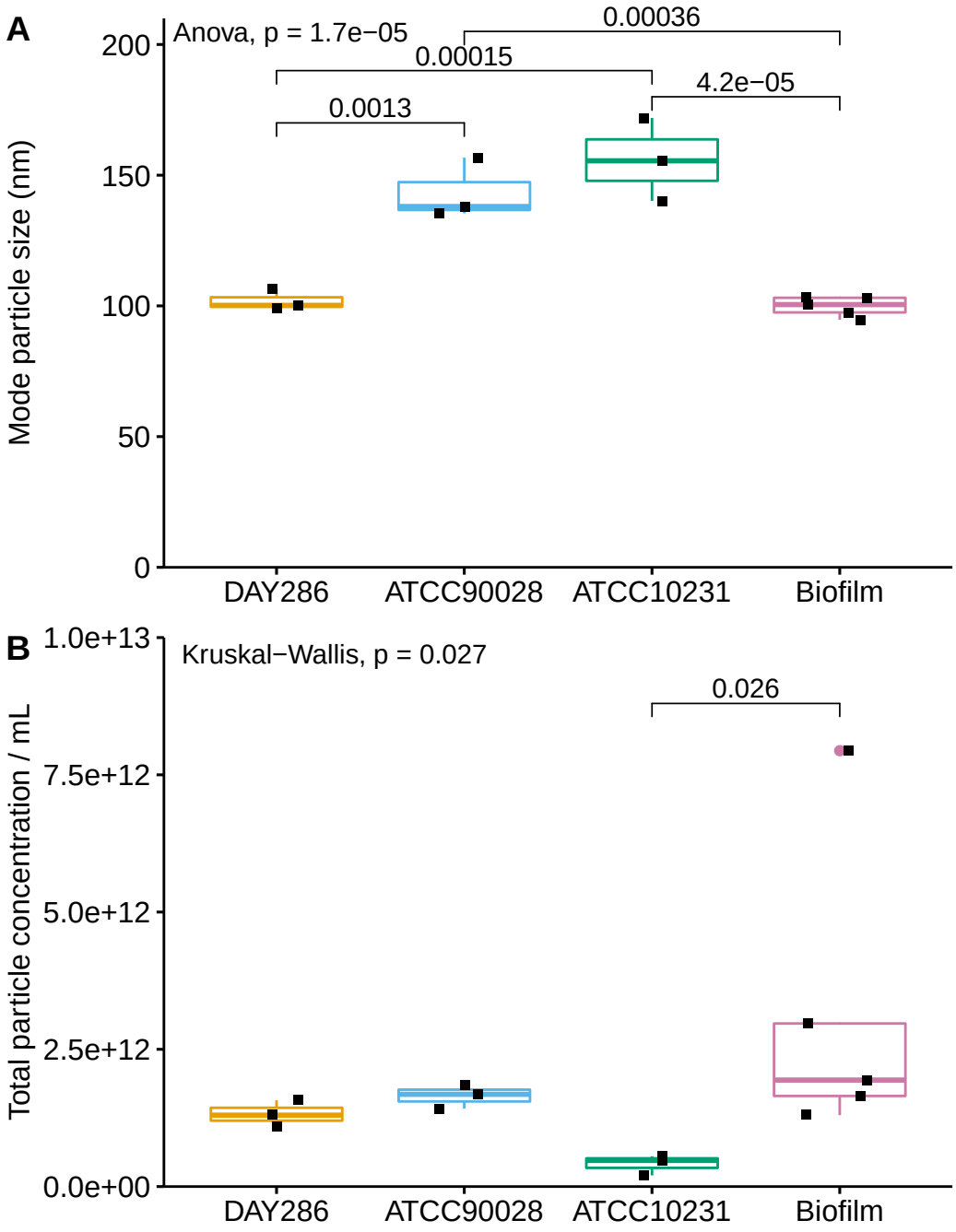

Figure S3: Statistical comparisons of *C. albicans* EV isolations. *(A)* Boxplot showing EV mode diameter measured using NTA. Differences in average EV mode diameters were compared using One-way ANOVA followed by Tukey's HSD *post-hoc* test. Adjusted p-values indicating significant differences are shown. *(B)* Boxplot showing total particle concentration for each EV preparation as measured by NTA. Differences in average total particle concentration were compared by Kruskal-Wallis test followed by Dunn's Test. Adjusted p-values indicating significant differences are shown.

Table S2: Candidate negative protein markers for *C. albicans* EVs. This list of proteins consists of those that were found to be exclusive to whole cell lysate (WCL) or significantly enriched in WCL across the four *C. albicans* strains examined in this study. Proteins are grouped according to their subcellular localisation as annotated in the Candida Genome Database (candidagenome.org) [2, 3] unless otherwise indicated. The $\log_2$ ratio of the abundance (mean MaxQuant LFQ intensity) of each protein in EVs compared to WCL for each strain is listed. A negative value indicates that a protein was enriched in WCL compared to EV. "ex" indicates where a protein was only quantified in the WCL fraction and not in EVs for that strain. The "TM" column indicates the number of transmembrane domains for each protein as annotated in UniProtKB. "SP" indicates whether a protein is annotated as having a signal peptide according to UniProt. "VDM" shows whether a protein has previously been detected in vesicle-depleted culture media (i.e. the proteins may also be in the soluble secretome) [1]. Underlined proteins are those identified as the best candidates for negative EV markers according to the criteria depicted in Supplementary Figure S1.

		$\log_2$(fold change) EV vs WCL						
Name	Function	DAY Y	A9	A1	DAY B	TM	SP	VDM
Cytoplasm								
ARO3	Phospho-2-dehydro-3-deoxyheptonate aldolase[b]	-1.28	-2.55	-1.51	-1.46			
ARO8	Aromatic transaminase of the Ehrlich fusel oil pathway[a]	-2.37	-1.81	ex	-2.36			
GDB1	Putative glycogen debranching enzyme	ex	ex	ex	-1.99			
HOM2	Aspartate-semialdehyde dehydrogenase[b]	ex	-3.41	ex	ex			
orf19.1889	Putative phosphoglycerate mutase family protein[b]	ex	-1.60	ex	ex			
orf19.5943.1	*S. cerevisiae* ortholog is Stm1, regulates translation[a]	ex	-3.98	-4.57	ex			
orf19.6596	Putative esterase[a]	ex	ex	ex	-2.90			
orf19.7263	Putative X-Pro aminopeptidase[a]	ex	ex	ex	ex			
SBP1	*S. cerevisiae* ortholog is Sbp1, eIF4G binding protein[a]	ex	-4.30	-3.80	-2.15			
STI1	HSP90 co-chaperone[a]	ex	-2.20	ex	ex			
URA4	Predicted succinate semialdehyde dehydrogenase[a]	ex	ex	ex	ex			
XKS1	Putative xylulokinase[b]	ex	ex	ex	ex			
YNK1	Nucleoside disphosphate kinase	-3.07	-2.57	-2.17	-1.40			
Cytosol and mitochondria								
ACH1	Acetyl-CoA hydrolase[b]	-3.24	-5.09	-2.79	-4.03			
GLR1	Glutathione reductase[b]	ex	ex	ex	-3.28			
HSP60	Heat shock protein 60[b]	-4.66	-1.59	-4.22	-4.06			
Mitochondria								
AAT1	Aspartate aminotransferase[a]	ex	-2.11	-1.48	-5.70			
BAT22	Putative branched chain amino acid aminotransferase[b]	ex	-3.57	-1.44	-3.12			
ETR1	Putative 2-enoyl thioester reductase[a]	ex	ex	ex	ex			
GCV2	Glycine decarboxylase P subunit[a]	ex	ex	ex	ex			

IDH1	Isocitrate dehydrogenase subunit[a]	-2.67	-2.48	-1.85	-2.22	
IDH2	Isocitrate dehydrogenase[a]	-3.21	-2.88	-3.54	-3.37	
IDP1	Putative isocitrate dehydrogenase[a]	ex	ex	ex	ex	
ILV5	Ketol-acid reductoisomerase[b]	-3.68	-2.62	-2.27	-5.48	
KGD2	Putative dihydrolipoamide S-succinyltransferase[a]	ex	-2.20	-2.04	ex	
LPD1	Dihydrolipoamide dehydrogenase	ex	-7.49	-4.08	-7.91	
NIF3	*S. cerevisiae* ortholog is Nif3, mitochondrial protein[a]	ex	ex	ex	ex	
orf19.2966	Predicted dienelactone hydrolase domain[a]	ex	ex	ex	ex	
orf19.449	Putative phosphatidyl synthase[b]	ex	ex	ex	ex	
orf19.7215.3	*S. cerevisiae* ortholog is Hsp10, mitochondrial co-chaperonin[a]	ex	ex	ex	ex	
PDX3	Pyridoxamine-phosphate oxidase[a]	ex	-1.85	ex	ex	
SOD2	Superoxide dismutase	ex	ex	ex	ex	
Nucleus						
DOT5	Thioredoxin peroxidase[a]	ex	ex	ex	ex	
NHP6A	Non-histone chromosomal protein 6[a]	ex	ex	ex	ex	
Vacuole						
AMS1	Putative alpha-mannosidase[a]	ex	ex	ex	-2.23	
APR1	Vacuolar aspartic proteinase	-1.83	ex	ex	-1.07	Y
CPY1	Carboxypeptidase Y	ex	ex	ex	-1.59	Y
LAP41	Putative aminopeptidase yscI precursor	ex	-3.96	-1.80	-2.64	
Cell wall, cell surface, fungal biofilm matrix						
CPR3	Putative peptidyl-prolyl cis-trans isomerase	ex	-2.59	ex	ex	
GCY1	Glycerol 2-dehydrogenase	ex	ex	ex	ex	
GLX3	Glutathione-independent glyoxalase	ex	-6.21	ex	-1.49	
GPM1	Phosphoglycerate mutase	-6.84	-4.93	-3.80	-2.78	
GRP2	NAD(H)-linked methylglyoxal oxidoreductase	ex	-7.45	-2.46	-1.42	
HSP21	Small heat shock protein	ex	-4.97	ex	-2.24	
MET15	O-acetylhomoserine O-acetylserine sulfhydrylase	ex	-2.91	-2.00	-2.72	
orf19.3053	Protein of unknown function	ex	-4.71	-2.35	-3.02	
orf19.590	Putative thiamine biosynthesis enzyme	ex	ex	ex	ex	
PGI1	Glucose-6-phosphate isomerase	-6.91	-5.44	-4.14	-1.68	
PST2	Flavodoxin-like protein	ex	ex	ex	-2.07	
RIB3	3,4-Dihydroxy-2-butanone 4-phosphate synthase	ex	-1.06	-1.23	-2.11	
Proteasome						
orf19.2755	*S. cerevisiae* ortholog is Pre7, Beta 6 subunit of 20S proteasome[a]	ex	-2.34	ex	ex	
orf19.4230	20S proteasome subunit (beta7)[a]	ex	ex	ex	ex	
PUP2	Alpha5 subunit of the 20S proteasome	ex	-4.13	-2.53	ex	
SCL1	Proteasome subunit YC7alpha[a]	ex	ex	ex	ex	
Cytoplasmic stress granule						
TIF11	Translation initiation factor eIF1a[a]	ex	ex	ex	ex	
TMA19	*S. cerevisiae* ortholog is Tma19, ribosome associated protein[a]	ex	-3.05	-2.74	ex	
Peroxisome						

orf19.1433	*S. cerevisiae* ortholog is Lpx1, peroxisomal lipase[a]	ex	ex	ex	ex	
Actin cortical patch						
ABP1	Actin-binding protein	ex	-1.47	ex	ex	
Unknown						
orf19.2125	Protein of unknown function[c]	ex	ex	ex	ex	1
orf19.2737	Carbohydrate kinase domain-containing protein[c]	ex	ex	ex	-1.48	
orf19.5620	Stationary phase enriched protein[c]	ex	ex	ex	-3.43	
OYE32	NADPH oxidoreductase family protein[c]	ex	ex	ex	-2.90	

[a] Protein localisation was inferred from sequence similarity with *S. cerevisiae* homolog as annotated in the Candida Genome Database [2, 3]

[b] Protein localisation was obtained from the GO Cellular Component annotation in the *C. albicans* UniProt reference proteome UP000000559 [4].

[c] Protein and has no Cellular Component annotation in the Candida Genome Database or UniProt reference proteome.

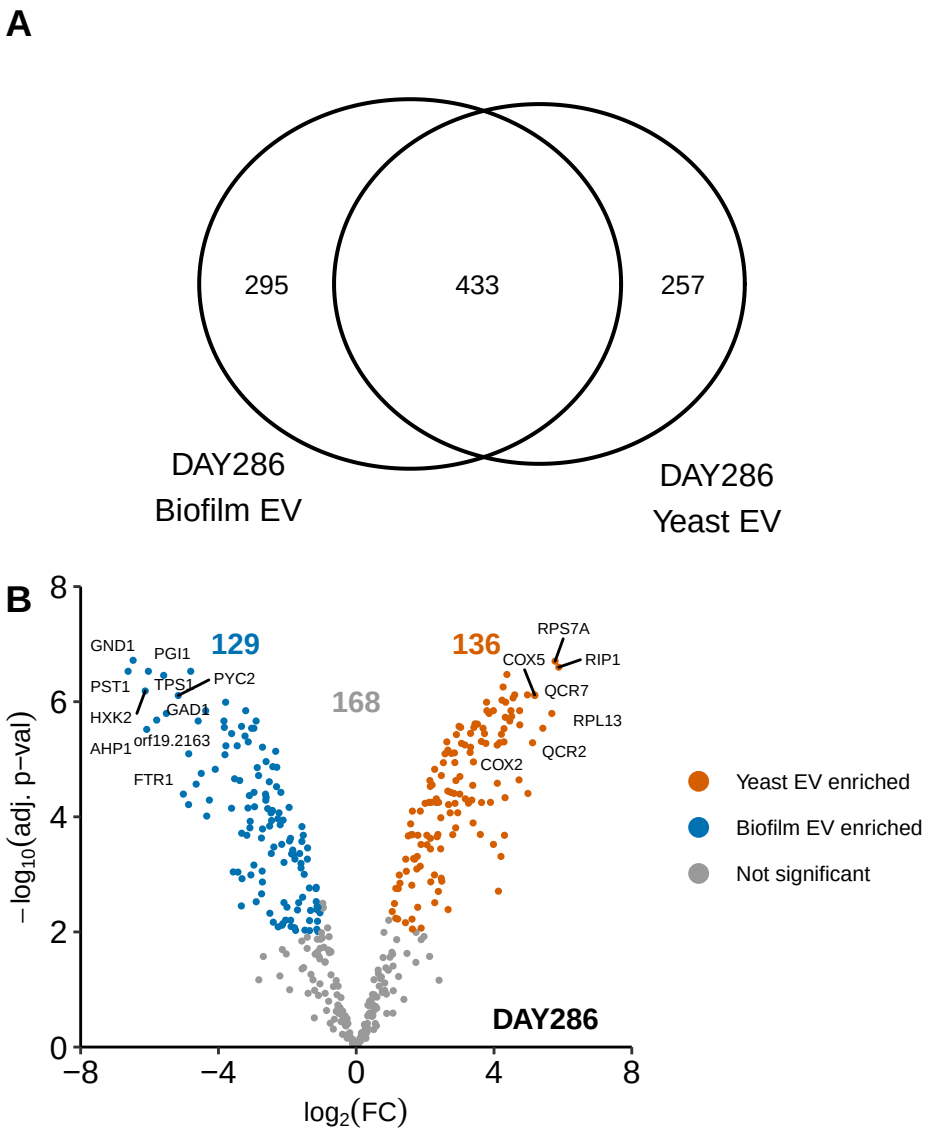

Figure S4: Differential abundance analysis of proteins identified in EVs isolated from two different morphologies of the DAY286 *C. albicans* strain. *(A)* Venn diagram comparing the biofilm EV and yeast EV proteome data sets shown in Figure 2A. *(B)* Volcano plot depicting significantly enriched biofilm EV or yeast EV proteins. Differential abundance analysis was performed by comparing the mean normalised LFQ intensities of common DAY286 EV proteins (i.e. proteins in the Venn overlap) using the package *limma* [5, 6]. Significantly enriched proteins were identified using a Benjamini-Hochberg adjusted p-value cut-off of 0.01 and a $\log_2$(FC) cut-off of 1. Counts of significant and non-significant proeins are indicated. Proteins with a $\log_2$(FC) greater than 5 or less than -5 are labelled. Data underlying this plot are provided in Supplementary Data S6.

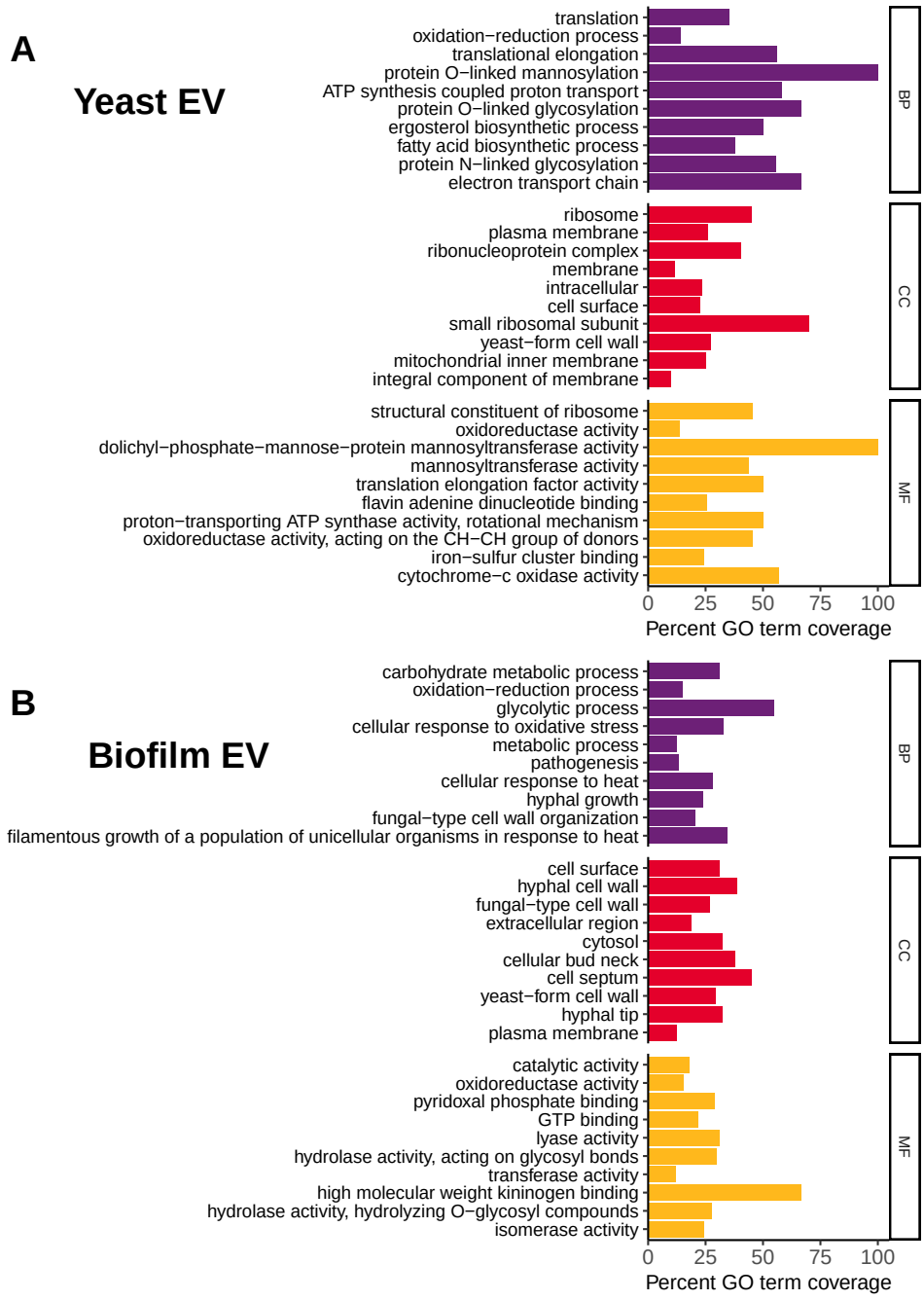

Figure S5: Functional enrichment analyses of EVs from two DAY286 morphologies. Top 10 significantly enriched biological process (BP), cellular component (CC), and molecular function (MF) GO terms for the *(A)* yeast EV enriched and exclusive proteins or *(B)* biofilm EV enriched or exclusive proteins shown in Supplementary Figure S4. GO analyses were performed using the online tool FungiFun2 (elbe.hki-jena.de/fungifun) (Hypergeometric distribution, Benjamini-Hochberg adjusted p-value < 0.01) [7]. The GO terms are presented top to bottom in order of increasing adjusted p-value. Full lists of enriched GO terms can be found in Supplementary Data S7.

2 References

[1] Gil-Bona, A.; Llama-Palacios, A.; Parra, C. M.; Vivanco, F.; Nombela, C.; Monteoliva, L.; Gil, C. Proteomics Unravels Extracellular Vesicles as Carriers of Classical Cytoplasmic Proteins in Candida Albicans. *Journal of Proteome Research*, **2015**, *14* (1), 142–153. https://doi.org/10.1021/pr5007944.

[2] Arnaud, M. B.; Costanzo, M. C.; Skrzypek, M. S.; Binkley, G.; Lane, C.; Miyasato, S. R.; Sherlock, G. The Candida Genome Database (CGD), a Community Resource for Candida Albicans Gene and Protein Information. *Nucleic Acids Research*, **2005**, *33* (suppl_1), D358–D363. https://doi.org/10.1093/nar/gki003.

[3] Skrzypek, M. S.; Binkley, J.; Binkley, G.; Miyasato, S. R.; Simison, M.; Sherlock, G. The Candida Genome Database (CGD): Incorporation of Assembly 22, Systematic Identifiers and Visualization of High Throughput Sequencing Data. *Nucleic Acids Research*, **2017**, *45* (D1), D592–D596. https://doi.org/10.1093/nar/gkw924.

[4] The UniProt Consortium. UniProt: A Worldwide Hub of Protein Knowledge. *Nucleic Acids Research*, **2019**, *47* (D1), D506–D515. https://doi.org/10.1093/nar/gky1049.

[5] Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Molecular & Cellular Proteomics*, **2014**, *13* (9), 2513–2526. https://doi.org/10.1074/mcp.M113.031591.

[6] Ritchie, M. E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C. W.; Shi, W.; Smyth, G. K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Research*, **2015**, *43* (7), e47–e47. https://doi.org/10.1093/nar/gkv007.

[7] Priebe, S.; Kreisel, C.; Horn, F.; Guthke, R.; Linde, J. FungiFun2: A Comprehensive Online Resource for Systematic Analysis of Gene Lists from Fungal Species. *Bioinformatics*, **2015**, *31* (3), 445–446. https://doi.org/10.1093/bioinformatics/btu627.