

## SUPPLEMENTARY INFORMATION

### Identification of pathogenic fungi with an optoelectronic nose

Yinan Zhang,<sup>a</sup> Jon R. Askim,<sup>a</sup> Wenxuan Zhong,<sup>b</sup> Peter Orlean,<sup>c</sup> and Kenneth S. Suslick<sup>\*a</sup>

<sup>a</sup> Department of Chemistry, <sup>b</sup> Department of Statistics, and <sup>c</sup> Department of Microbiology

University of Illinois at Urbana-Champaign, 600 S. Mathews Av., Urbana, IL 61801, USA

Correspondence should be addressed to K.S.S. ([ksuslick@illinois.edu](mailto:ksuslick@illinois.edu))

**Table S1.** Summary of fungal strain information and number of replicate experiments for each strain.

Strain	Strain origin	Replicates
Control	n/a	34
<i>Candida albicans</i> (CAI-4)	Fonzi & Irwin*	12
<i>Candida albicans</i> (B311)	ATCC† 32354	11
<i>Candida albicans</i> (1-28)	Wrobel et al.‡	10
<i>Candida glabrata-1</i>	Hoyer collection§	11
<i>Candida guilliermondii</i>	Hoyer collection§	6
<i>Candida parapsilosis</i>	Hoyer collection§	6
<i>Trichosporon asahii</i> 3323	Domestic pig isolates§	6
<i>Debaryomyces hansenii</i> 3333 ( <i>Candida famata</i> )	Domestic pig isolates§	6
<i>Candida stellatoidea</i>	ATCC 36232	13
<i>Candida keyfr</i>	ATCC 46764	10
<i>Saccharomyces cerevisiae</i> 4742	Brachmann et al.**	18
<i>Kluyveromyces lactis</i>	New England Biolabs	12

\*Fonzi, W.A. & Irwin, M.Y. Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**, 717-728 (1993).

†ATCC: American Type Culture Collection number.

‡Wrobel, L. et al. Molecular phylogenetic analysis of a geographically and temporally matched set of *Candida albicans* isolates from humans and nonmigratory wildlife in central Illinois. *Eukaryotic cell* **7**, 1475-1486 (2008).

§Hoyer, Lois L <[lhoyer@illinois.edu](mailto:lhoyer@illinois.edu)>

\*\*Brachmann, C.B. et al. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115-132 (1998).

**Table S2.** List of chemically responsive dyes. Spot numbering for the 6x6 array shown in **Figure 1a** is from left to right, top to bottom.

spot #	Name
1	5,10,15,20-tetraphenylporphyrinatozinc(II)
2	5,10,15,20-tetrakis(2,4,6-trimethylphenyl)porphyrinatozinc(II)
3	5,10,15,20-tetrakis(pentafluorophenyl)porphyrinatozinc(II)
4	5,10,15,20-tetrakis(2,4,6-trimethylphenyl)porphyrinatocobalt(II)
5	5,10,15,20-tetraphenylporphyrinatocadmium(II)
6	5,10,15,20-tetraphenylporphyrinatochromium(III) chloride
7	Bromophenol Blue + TBAH
8	Methyl Red + TBAH
9	Chlorophenol Red + TBAH
10	Nitrazine Yellow + TBAH
11	Bromothymol Blue + TBAH
12	Thymol Blue + TBAH
13	m-Cresol Purple + TBAH
14	Zn(OAc) <sub>2</sub> + m-Cresol Purple + TBAH
15	HgCl <sub>2</sub> + Bromophenol Blue + TBAH
16	HgCl <sub>2</sub> + Bromocresol Green + TBAH
17	Pb(OAc) <sub>2</sub>
18	1-[4-[[4-(dimethylamino)phenyl]azo]phenyl]-2,2,2-trifluoroethanone + TsOH
19	$\alpha$ -Naphthol Red + TsOH
20	Tetraiodophenolsulfonephthalein
21	Fluorescein
22	Bromocresol Green
23	Methyl Red
24	Bromocresol Purple
25	Bromophenol Red
26	Rosolic Acid
27	Bromopyrogallol Red
28	Pyrocatechol Violet
29	Nile Red
30	Disperse Orange #25
31	4-(4-Nitrobenzyl)pyridine + N-Benzylaniline
32	Pyrylium, 4-[2-[4-(dimethylamino)phenyl]ethenyl]-2,6-dimethyl-, tetrafluoroborate
33	LiNO <sub>3</sub> + Cresol Red
34	Acridine Orange Base
35	AgNO <sub>3</sub> + Bromophenol Blue
36	AgNO <sub>3</sub> + Bromocresol Green

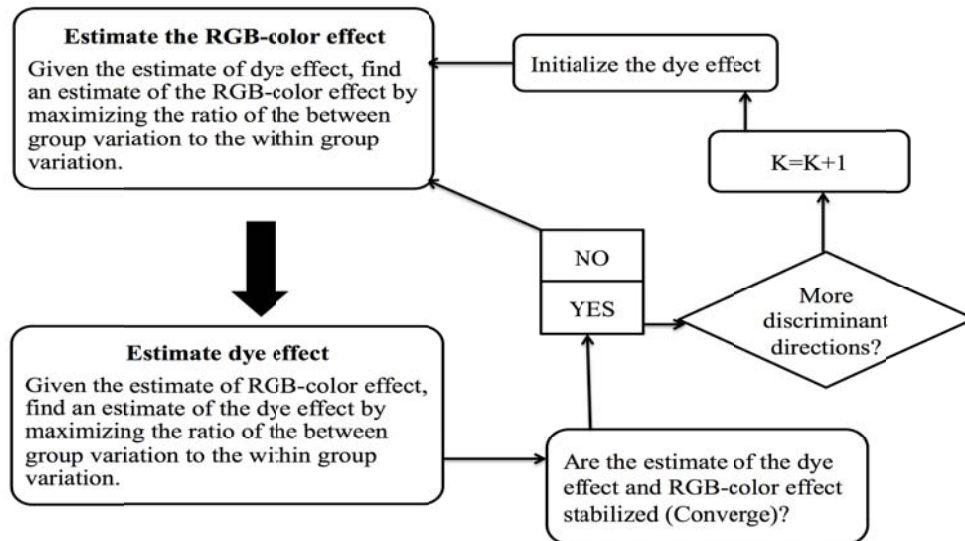
TBAH: 1.0 M tetrabutylammonium hydroxide water solution. TsOH: 1.0 M *p*-toluenesulfonic acid in 2-methoxyethanol. Red, green and blue colour channels are defined for each spot (e.g., for spot 3, G values are channel 8).

**Table S3.** Classification Matrix (Cases in row categories classified into columns).

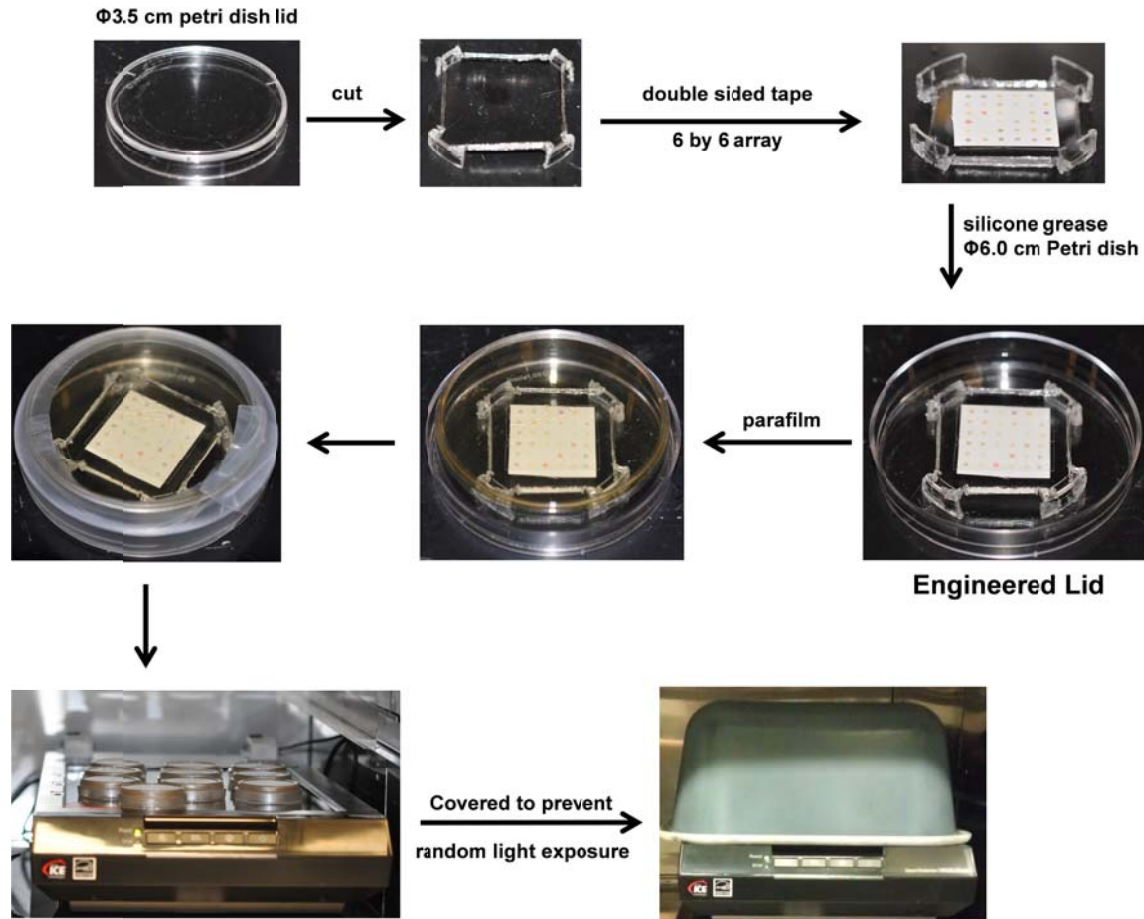
	C.stellatoidea	C.albicans 1-28	C.albicans B311	C.albicans CA14	C.glabrata	C.guilliermondii	C.keyfr	C.parapsilosis	D.hansenii	K.lactis	Medium (background)	S.cerevisiae	T.asahii	%corr.
C.stellatoidea	13	0	0	0	0	0	0	0	0	0	0	0	0	100
C.albicans 1-28	0	10	0	0	0	0	0	0	0	0	0	0	0	100
C.albicans B311	0	0	11	0	0	0	0	0	0	0	0	0	0	100
C.albicans CA14	0	0	0	12	0	0	0	0	0	0	0	0	0	100
C.glabrata	0	0	0	0	11	0	0	0	0	0	0	0	0	100
C.guilliermondii	0	0	0	0	0	6	0	0	0	0	0	0	0	100
C.keyfr	0	0	0	0	0	0	10	0	0	0	0	0	0	100
C.parapsilosis	0	0	0	0	0	0	0	6	0	0	0	0	0	100
D.Hansenii	0	0	0	0	0	0	0	0	6	0	0	0	0	100
K.lactis	0	0	0	0	0	0	0	0	0	12	0	0	0	100
Medium (background)	0	0	0	0	0	0	0	0	0	0	34	0	0	100
S.cerevisiae	0	0	0	0	0	0	0	0	0	0	0	18	0	100
T.asahii	0	0	0	0	0	0	0	0	0	0	0	0	6	100
Total	13	10	11	12	11	6	10	6	6	12	34	18	6	100

**Table S4.** Jackknifed Classification Matrix of Linear Discriminant Analysis.

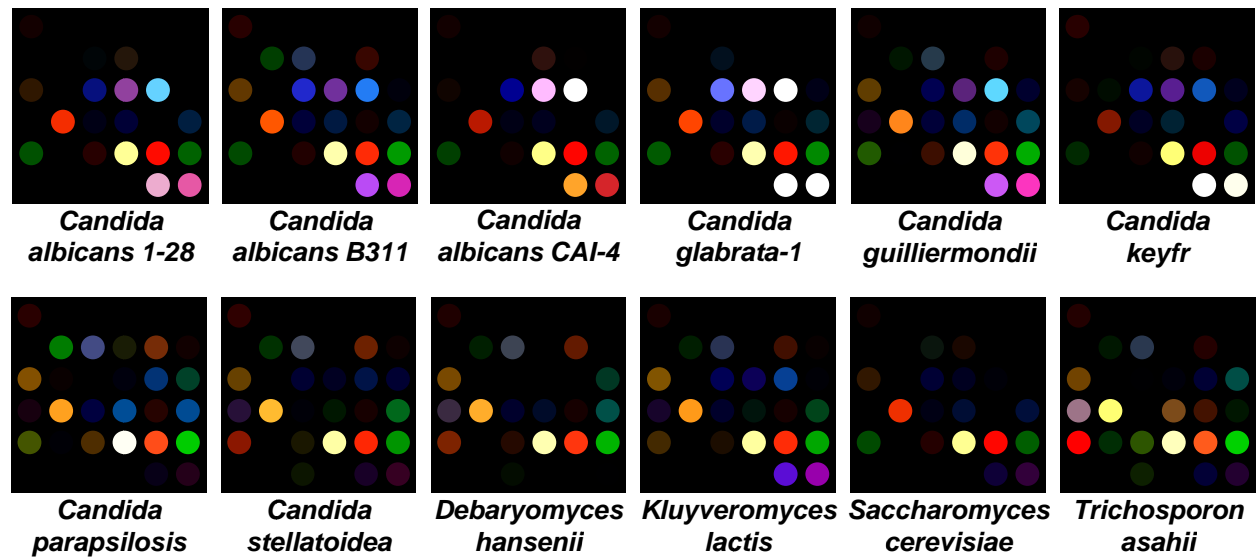
	C.Stellatoidea	C.albicans 1-28	C.albicans B311	C.albicans CA14	C.glabrata	C.guilliermondii	C.keyfr	C.parapsilosis	D.hansenii	K.lactis	Medium (background)	S.cerevisiae	T.asahii	%corr.
C.stellatoidea	12	0	0	0	0	0	0	0	0	1	0	0	0	92
C.albicans 1-28	0	10	0	0	0	0	0	0	0	0	0	0	0	100
C.albicans B311	0	1	8	0	0	0	0	0	0	0	0	2	0	73
C.albicans CA14	0	0	0	12	0	0	0	0	0	0	0	0	0	100
C.glabrata	0	0	0	0	11	0	0	0	0	0	0	0	0	100
C.guilliermondii	0	0	0	0	0	5	0	0	0	1	0	0	0	83
C.keyfr	0	0	0	0	0	0	10	0	0	0	0	0	0	100
C.parapsilosis	0	0	0	0	0	0	0	6	0	0	0	0	0	100
D.hansenii	0	0	0	0	0	0	0	0	6	0	0	0	0	100
K.lactis	0	0	0	0	0	3	0	0	0	9	0	0	0	75
Medium (background)	0	0	0	0	0	0	0	0	0	0	34	0	0	100
S.cerevisiae	0	0	1	0	0	0	0	0	0	0	0	17	0	94
T.asahii	0	0	0	0	0	0	0	0	0	0	0	0	6	100
Total	12	11	9	12	11	8	10	6	6	11	34	19	6	94



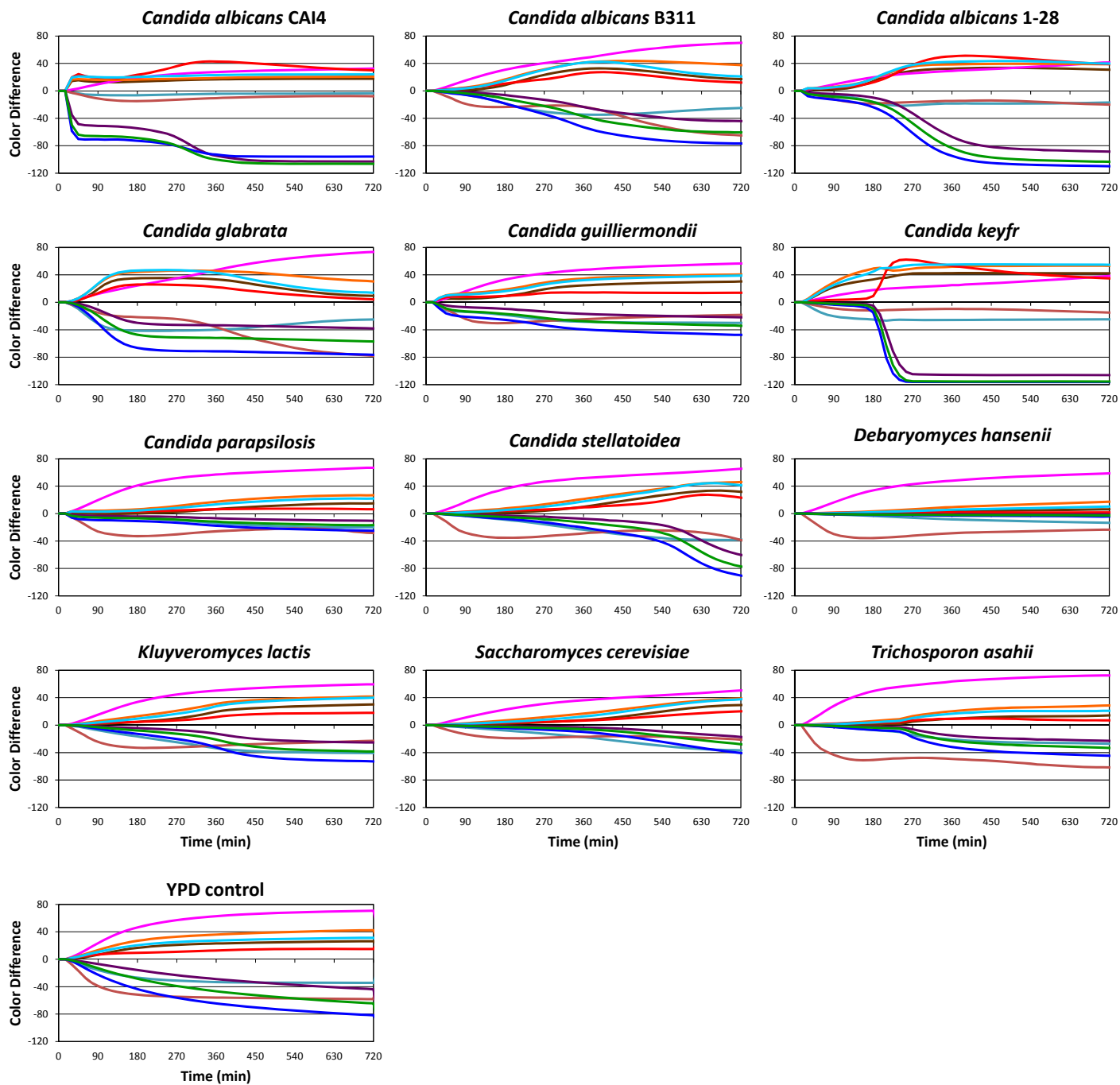
**Scheme S1.** Flow chart of tensor discriminant analysis.



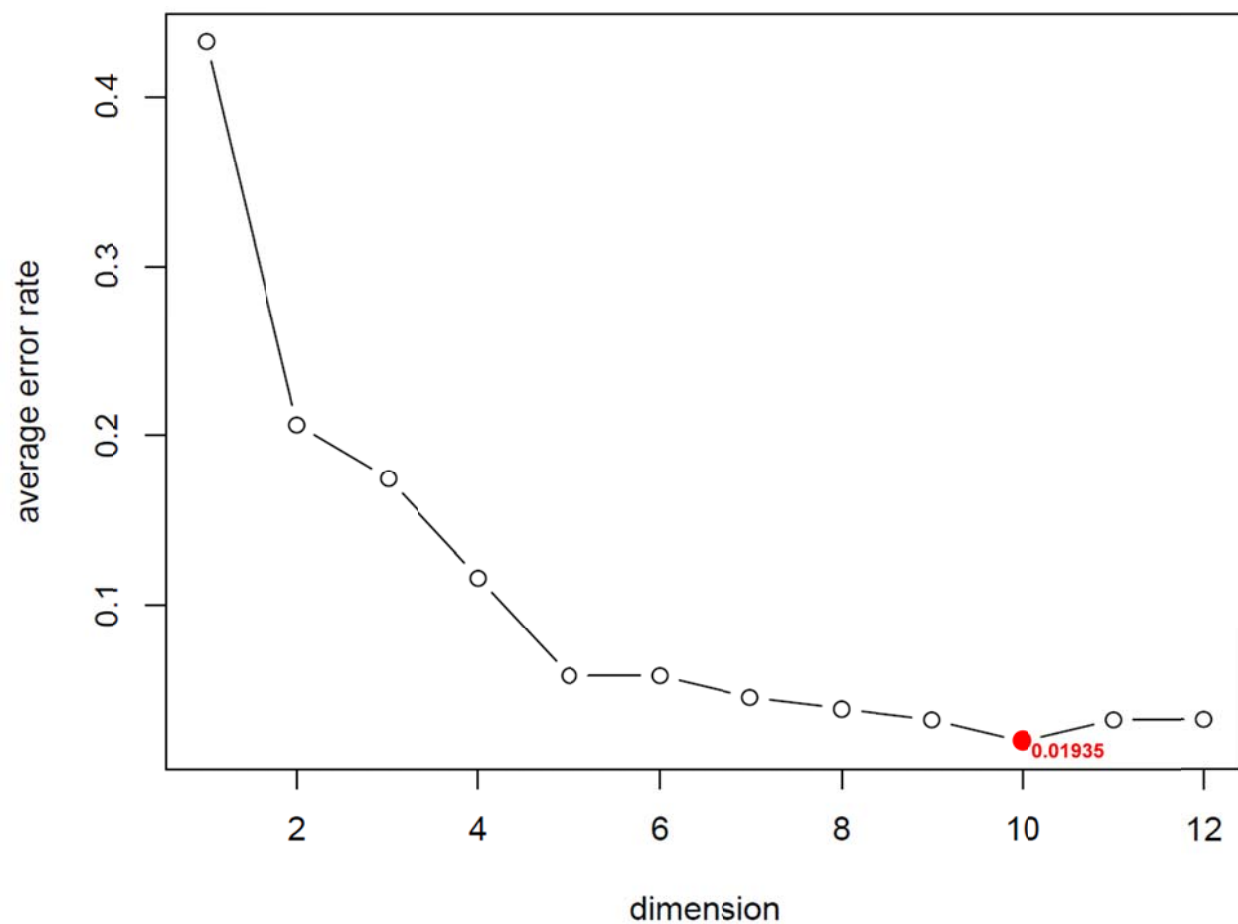
**Fig. S1** Construction of the colorimetric sensor array within Petri dishes and scanner placement within a standard incubator.



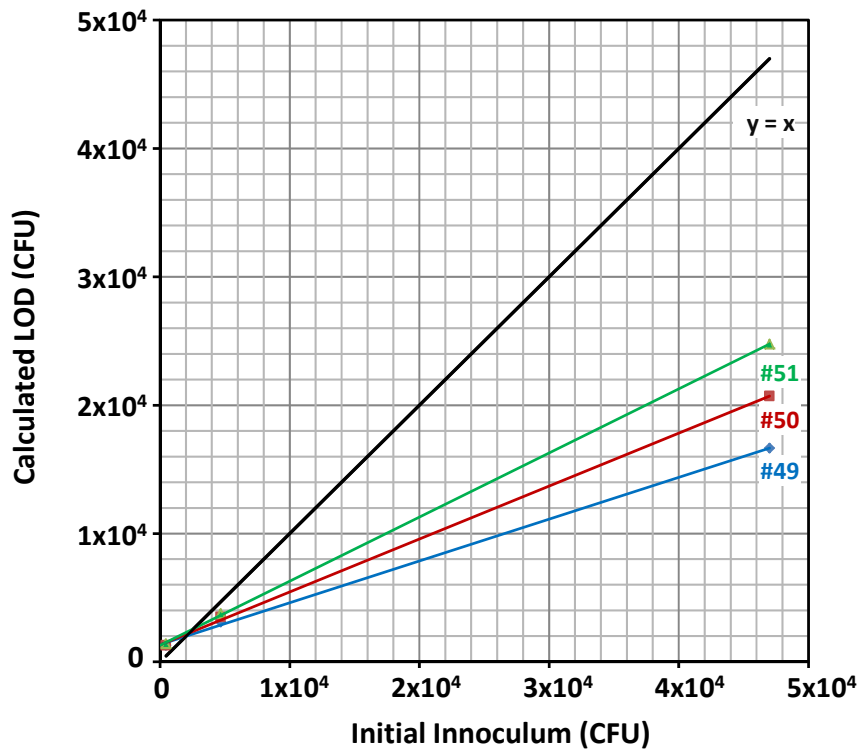
**Fig. S2** Colour difference maps for 12 different fungal strains grown on YPD medium at 180 minutes. Each fungal species has a unique pattern.



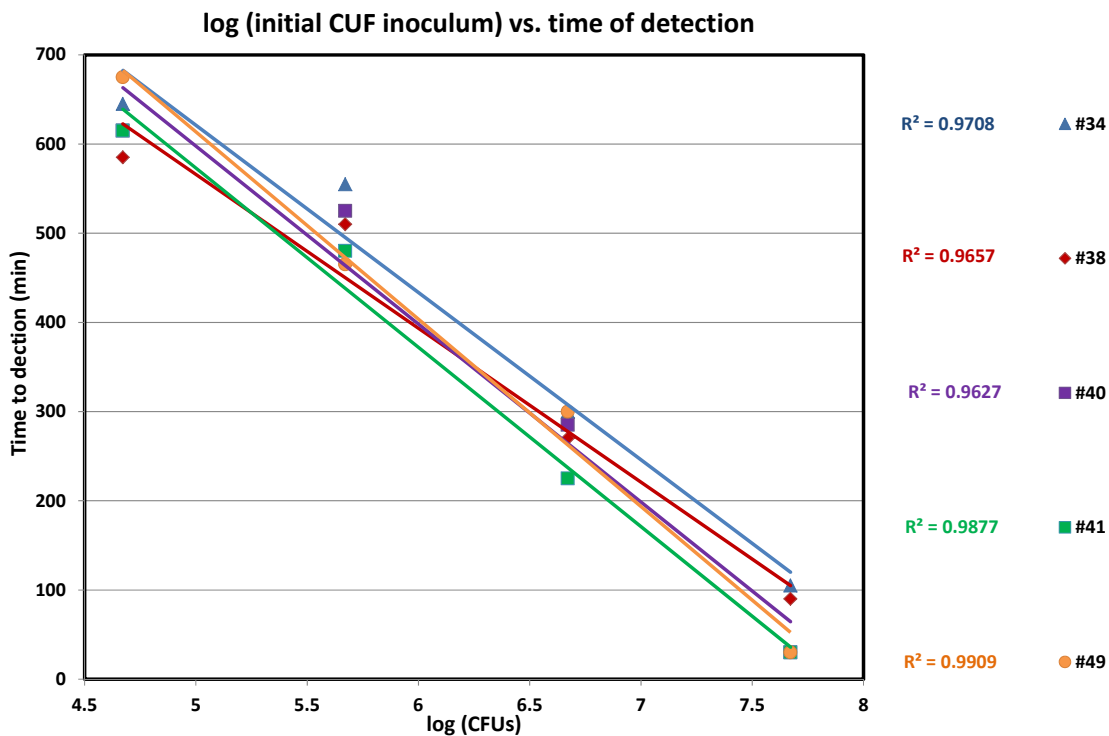
**Fig. S3** Time response profiles for all 12 different fungal strains. For each strain, the colour changes versus time are plotted for all 108 colour channels. For the clarity of the figures, 10 most responsive channels (out of 108 channels) are shown in the figure.



**Fig. S4** Misclassification rate (error rate) results from tensor discriminant analysis. Data between 120 to 360 minutes in 30-minute interval was used. Error rate for each trial was calculated by testing it with the training model built with the rest of the 154 trials. The final error rate was the average of all 155 error rates. The lowest error rate was 0.01935 using 10 dimensions.



**Fig. S5** Interpolation of the limit of detection (LOD) for *Candida albicans* (CAI-4) for three channels (#49, #50 and #51, which correspond to spot 17,  $\text{Pb}(\text{O}_2\text{CCH}_3)_2$ ). At any given time, the LOD for any given colour change (i.e.,  $\Delta\text{RGB}$  value) of any given channel of any given spot is defined as the initial inoculum that gives three times the standard deviation of the  $\Delta\text{RGB}$ , which is approximately 8 colour difference units (the range of  $\Delta\text{RGB}$  is -255 to +255). The LOD at 720 min was interpolated from the three lowest initial inocula ( $4.7 \times 10^3$ ,  $4.7 \times 10^4$  and  $4.7 \times 10^5$  CFU) and was determined to be  $2 \times 10^4$  CFU for the initial inoculum with channel 49, as indicated in the figure. The standard deviation is approximately 8 colour difference units (i.e., RGB units ranging from 0 to 255).



**Fig. S6** Correlation of the time to detection (as determined by the most responsive channel) versus log(initial inoculum in CFU), where the time to detection is defined as the time at which the total signal is greater than three times the standard deviation of the  $\Delta$ RGB value. The linear regression coefficients are above 0.95.

**Dataset S1.** The complete experimental database of colour changes for all 155 trials from 30 min to 735 min in 15 min intervals is provided as an Excel spreadsheet.